Communications on Pure and Applied
Mathematics 44 (2001), 761-825

A New Approach to Variational Problems with Multiple Scales

GIOVANNI ALBERTI
Università di Pisa
AND
STEFAN MÜLLER
Max Planck Institute for Mathematics in the Sciences

Abstract

We introduce a new concept, the Young measure on micro-patterns, to study singularly perturbed variational problems which lead to multiple small scales depending on a small parameter ε. This allows one to extract, in the limit $\varepsilon \to 0$, the relevant information at the macroscopic scale as well as the coarsest microscopic scale (say ε^α), and to eliminate all finer scales.

To achieve this we consider rescaled functions $\mathsf{R}^\varepsilon_s x(t) := x(s + \varepsilon^\alpha t)$ viewed as maps of the macroscopic variable $s \in \Omega$ with values in a suitable function space. The limiting problem can then be formulated as a variational problem on the Young measures generated by $\mathsf{R}^\varepsilon x$. As an illustration we study a one-dimensional model that describe the competition between formation of microstructure and highest gradient regularization. We show that the unique minimizer of the limit problem is a Young measure supported on sawtooth functions with a given period.

1. Introduction

Many problems in science involve structures on several distinct length scales. Two typical examples are the hierarchy of domains, walls and (Bloch) lines in ferromagnetic materials [16], [26] and the layers-within-layers pattern often observed in fine phase mixtures induced by symmetry breaking solid-solid phase transitions [6], [30], [43], [56].

An important feature in these examples is that the relevant length scales are not known a priori, but emerge from an attempt of the system to reach its minimum energy (or maximum entropy) or at least an equilibrium state. In ferromagnetic materials, for example, the typical length scale of Bloch walls can be predicted by dimensional analysis but the size of the domains is determined by a complex interplay of specimen geometry, anisotropy and (nonlocal) magnetostatic energy.

De Giorgi's notion of Γ-convergence has proven to be very powerful to analyze variational problems with one small length scale and the passage from phase field models (with small, but finite, transition layers between different phases) to sharp interface models (the rapidly growing literature begins with [37], [38], [36], recent work includes [1], [8], where many further references can be found). More recently an alternate approach, mostly for evolution problems, based on viscosity solutions has been applied very successfully to situations where a maximum principle is available (see for instance [7], [12], [21], [20], [27] [28]).

Much less is known for problems with multiple small scales. Matched asymptotics expansion, renormalization or intermediate asymptotics are powerful methods to predict the limiting behavior but few rigorous results are known.

In this paper, we propose a new approach for a rigorous analysis of variational problems with two small scales, based on an extension of the Γ-convergence approach. As in formal asymptotics we begin by introducing a slow (i.e., order one) and a fast scale. Instead of the original quantity $v^\varepsilon(s)$, where ε represents a parameter that determines the smallness of the scales, we consider rescaled functions $\mathsf{R}^\varepsilon_s v^\varepsilon(t) = \varepsilon^{-\beta} v^\varepsilon(s+\varepsilon^\alpha t)$ of the two variables s and t, where ε^α represents the fast scale and $\varepsilon^{-\beta}$ is a suitable renormalization. We then consider $s \mapsto \mathsf{R}^\varepsilon_s v^\varepsilon$ as a map from the original domain Ω to a function space K (which can be chosen compact and metrizable). Finally we derive a variational problem for the Young measure that arises as limit of the maps $s \mapsto \mathsf{R}^\varepsilon_s v^\varepsilon$.

The Young measure (see Section 2 for precise definitions and references) is a map ν from Ω to the space of probability measures on K, and for each $s \in \Omega$ the measure $\nu(s)$, often written as ν_s, represents the probability that $\mathsf{R}^\varepsilon_s v^\varepsilon$ assumes a certain value in a small neighborhood of s in the limit $\varepsilon \to 0$. In terms of the original problem, ν_s gives the probability to find a certain pattern (i.e., an element of the function space K) on the scale ε^α near the point s. We thus refer to ν sometimes as a *Young measure on (micro-) patterns*, or a *two-scale Young measure*. A precise description is given in Section 3 below.

To illustrate our concept and its application we consider the following one dimensional problem which already shows a rather interesting two-scale behavior: minimize

$$I^\varepsilon(v) := \int_0^1 \varepsilon^2 \ddot{v}^2 + W(\dot{v}) + a(s)\, v^2 \, ds \tag{1.1}$$

among one-periodic functions $v : \mathbb{R} \to \mathbb{R}$, where $\dot{v}$ and $\ddot{v}$ denote the first and second derivative, respectively. A typical choice for the double-well potential W is

$$W(t) := (t^2-1)^2$$

but any other continuous function W that vanishes exactly at ± 1 and is bounded from below by $c|v|$ at infinity will do.

If $\varepsilon = 0$ and $a = 0$ then there exist infinitely many minimizers, indeed any sawtooth function with slope ± 1 realizes the minimum. If $\varepsilon > 0$ is small and $a = 0$ a unique (up to translation and reflection) minimizer is selected. It is very close to a sawtooth function with slope ± 1 and two corners per period. Such a result is a typical application of classical Γ-convergence; indeed for $a = 0$ the Γ-limit of $\frac{1}{\varepsilon} I^\varepsilon$ is only finite on sawtooth functions and counts the number of corners (cf. the sketch of proof after Theorem 1.2).

If $\varepsilon = 0$ but $a > 0$ then no minimizers exist and minimizing sequences are (essentially) given by highly oscillatory sawtooth functions with slope ± 1 that converge uniformly to 0 (more precisely, the Young measure generated by the derivatives of any minimizing sequence is $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ at almost every point).

If $\varepsilon > 0$ and $a > 0$ the excitation of oscillations due to $a > 0$ and their penalization due to $\varepsilon > 0$ lead to the emergence of a new structure.

THEOREM 1.1 (see [39]). *Suppose that a is constant and strictly positive. Then, for ε positive and sufficiently small, all minimizers of I^ε among one-periodic functions have minimal period*

$$P^\varepsilon = L_0 a^{-1/3} \varepsilon^{1/3} + O(\varepsilon^{2/3}),$$

where $L_0 := \left(96 \int_{-1}^1 \sqrt{W}\right)^{1/3}$.

The derivatives of minimizers exhibit indeed a structure with two fast scales: transition layers of order ε are spaced periodically with the period $P^\varepsilon \sim \varepsilon^{1/3}$, as seen in Figure 1.1.

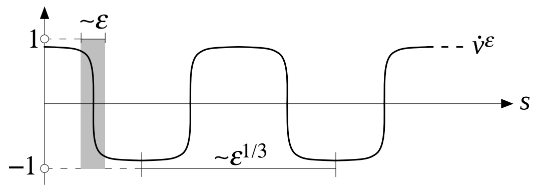

FIGURE 1.1. Two scale structure of minimizers

This behavior was predicted by Tartar [53] on the basis of matched asymptotic expansions. It can equivalently be guessed by a formal application of Γ-convergence. The purpose of our work is to create a framework in which such reasoning can be made rigorous. As corollary of our new approach we obtain the following result (see Section 3, and in particular Corollary 3.13, for precise definitions and a more detailed statement).

THEOREM 1.2. *Suppose that a belongs to L^∞ and is strictly positive a.e., let v^ε be a sequence of minimizers of I^ε, and take $\mathsf{R}^\varepsilon_s v^\varepsilon$ and the Young measure ν as above. Then for a.e. s the measure ν_s is supported on the set of all translations of the sawtooth function y_h with slope ± 1 and period $h := L_0 (a(s))^{-1/3}$; see Figure 1.2.*

Notice that when a is not constant, the minimizers v^ε are by no means periodic. Yet, Theorem 1.2 says that close to a.e. point s (of approximate continuity for a) v^ε "resembles" more and more as $\varepsilon \to 0$ the periodic sawtooth function y_h, with h now depending on s. Thus the Young measure on patterns ν provides a way to localize the result in Theorem 1.1. More importantly, it gives a precise meaning to the statement that v^ε is locally nearly periodic with a period which depends on the point s.

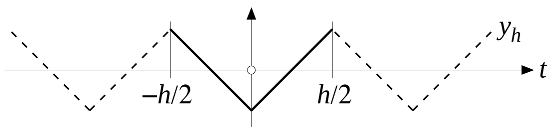

FIGURE 1.2. The sawtooth function y_h

In addition to this, the main advantage of the new object ν is, in our view, the possibility to make a formal reasoning rigorous.

Let us illustrate this in the context of Theorem 1.1. Denote by H^2_{per} the Sobolev space of functions on the interval $(0,1)$ whose periodic extension belongs to $H^2_{\mathrm{loc}}(\mathbb{R})$, and by $\mathscr{S}_{\mathrm{per}}$ the space of functions on $(0,1)$ whose periodic extension are (continuous) sawtooth functions with slope ± 1. Consider the functionals

$$J^\varepsilon(v) := \int_0^1 \varepsilon \ddot{v}^2 + \frac{1}{\varepsilon} W(\dot{v}) \quad \text{if } v \in H^2_{\mathrm{per}},$$
$$J(v) := \frac{A_0}{2} \int_0^1 |\ddot{v}| = A_0 \, \#\big(S\dot{v} \cap [0,1)\big) \quad \text{if } v \in \mathscr{S}_{\mathrm{per}},$$

where $A_0 := 2\int_{-1}^1 \sqrt{W}$, and $S\dot{v}$ denotes the set of points of discontinuity of $\dot{v}$.

We know that J^ε, extended to $+\infty$ on $W^{1,1} \setminus H^2_{\mathrm{per}}$, Γ-converges in the $W^{1,1}$ topology to J, extended to $+\infty$ on $W^{1,1} \setminus \mathscr{S}_{\mathrm{per}}$ (this easily follows from the one-dimensional version of the result in [36]). Thus it is plausible to replace

$$I^\varepsilon(v) = \varepsilon J^\varepsilon(v) + \int_0^1 a\, v^2 ds \tag{1.2}$$

by

$$\tilde{I}^\varepsilon(v) := \varepsilon J(v) + \int_0^1 a\, v^2 ds. \tag{1.3}$$

The minimization of $\tilde{I}^\varepsilon$ is a discrete problem since J is only finite on sawtooth functions with a finite even numbers of corners $0 \leq s_1 < s_2 < \ldots < s_{2N} < 1$. A short calculation yields the (sharp) bound

$$\int_{s_i}^{s_{i+1}} v^2 ds \geq \frac{1}{12}(s_{i+1} - s_i)^3$$

and a convexity argument shows that for a given number $2N$ of corners the minimum of $\tilde{I}^\varepsilon(v)$ is given by $2\varepsilon A_0 N + \frac{a}{48} N^{-2}$, and is achieved by the sawtooth function with period $1/N$ and vanishing average. Finally minimization over N yields the assertion

$$P^\varepsilon = \frac{1}{N} \sim \Big(\frac{48 A_0 \varepsilon}{a}\Big)^{1/3} = L_0 a^{-1/3} \varepsilon^{1/3},$$

while the energy of minimizers is

$$E^\varepsilon \sim E_0 a^{1/3} \varepsilon^{2/3} \quad \text{with } E_0 := \Big(\frac{3}{4} A_0\Big)^{2/3} = \Big(\frac{3}{2}\int_{-1}^1 \sqrt{W}\Big)^{2/3}. \tag{1.4}$$

The main point is to justify the passage from (1.2) to (1.3). This hinges on fact that the scale ε involved in the passage from J^ε to J (removal of ε-transition layers) is much smaller than $\varepsilon^{1/3}$. By introducing the rescaling $\mathsf{R}^\varepsilon_s v(t) := \varepsilon^{-1/3} v(s + \varepsilon^{1/3} t)$ and by replacing derivatives of v with respect to s by derivatives of $\mathsf{R}^\varepsilon_s v$ with respect to t, we may represent $I^\varepsilon(v)$ as an integral over functionals in $\mathsf{R}^\varepsilon_s v$

$$\varepsilon^{-2/3} I^\varepsilon(v) = \int_0^1 f^\varepsilon(\mathsf{R}^\varepsilon_s v)\, ds, \tag{1.5}$$

where

$$f^\varepsilon(x) := \fint_{-r}^{r} \varepsilon^{2/3} \ddot{x}^2 + \varepsilon^{-2/3} W(\dot{x}) + a\, x^2 \, dt$$

for a given positive r. Now we have that f^ε Γ-converge to f, where

$$f(x) := \frac{A_0}{2r} \#\big(S\dot{x} \cap (-r,r)\big) + a \fint_{-r}^{r} x^2 \, dt$$

if x agrees with a sawtooth function on $(-r,r)$, and is $+\infty$ otherwise. We then essentially have to show that the Γ-limit commutes with the integration in s in (1.5). More precisely we reformulate all functionals in term of Young measures and we show that the limiting (rescaled) energy $\varepsilon^{-2/3} I^\varepsilon(v^\varepsilon)$ of a sequence (v^ε) is given by

$$\int_0^1 \langle \nu_s, f \rangle \, ds, \tag{1.6}$$

where ν is the Young measure (on patterns) generated by $\mathsf{R}^\varepsilon v^\varepsilon$.

To determine the minimizing Young measure we need to know which Young measures arise as limits of $\mathsf{R}^\varepsilon v^\varepsilon$. This is not obvious since for finite ε the blowups $\mathsf{R}^\varepsilon_s v^\varepsilon$ at different points s are not independent. In the limit, however, the measures ν_s become independent and the only restriction is that ν_s be invariant under translation in the space of patterns K (see Proposition 3.1 and Remark 3.2). Thus the minimization of (1.6) can be done independently for each s and one easily arrives at the conclusions of Theorem 1.2, at least for a constant. The details of this argument are carried out in Section 3.

There are a number of other mathematical approaches to problems with small scales. For sequences v^ε converging weakly to 0 in $L^2_{\mathrm{loc}}(\mathbb{R}^N)$, Tartar [50] and Gérard [23] introduced independently a measure on $\mathbb{R}^N \times S^{N-1}$ (called the H-measure or microlocal defect measure, respectively) that estimates how much energy (in the sense of squared L^2-norm) concentrates at $x \in \mathbb{R}^N$ and high frequency oscillations with direction $\zeta \in S^{N-1}$. While this measure has no natural length scale, there are variants with characteristic scale $\delta(\varepsilon) \to 0$ ([24]). An interesting issue is to design similar objects for problems with multiple length scales; this is easy if the oscillations are additively superimposed but, as Gérard and Tartar pointed out, multiplicative interaction leads to new phenomena due to interference. The H-measure and its variants can only predict the limits of quadratic expressions in v^ε (the expression may, however, involve pseudodifferential operators) and hence have no direct applications to the study of I^ε.

The classical Young measure, by contrast, gives the limit of arbitrary (continuous) nonlinearities but contains no information on patterns. For further discussion of H-measures and their relation with Young measures see [51], [52].

Our work was inspired by the concept of two-scale convergence, although our approach is ultimately rather different. Two-scale convergence was introduced in [41] and employed by a number of researchers, cf. [2], [3], [17]. The main idea is to recover additional structure in a weakly converging sequence v^ε by using test functions of the form $\phi(s, s/\varepsilon^\alpha)$, where ϕ is periodic in the last variable.

If v^ε is of the form $v^\varepsilon(s) := v_0(s) + v_1(s, s/\varepsilon^\alpha) + o(1)$, with v_1 P-periodic in the last variable and $\int_0^P v_1(s,t)\,dt = 0$ for every s, and if one takes a test function $\phi(s,t) := \psi(s) + \eta_1(s)\,\eta_2(t)$, where η_2 is P-periodic and has vanishing average on the period, then

$$\int_0^1 \phi\Big(s, \frac{s}{\varepsilon^\alpha}\Big)\, v^\varepsilon(s)\, ds \to \int_0^1 \psi(s)\, v_0(s)\, ds + \int_0^1 \fint_0^P \eta_1(s)\,\eta_2(t)\, v_1(s,t)\, dt\, ds.$$

Thus both the weak limit v_0 and the oscillatory term on the scale ε^α related to v_1 can be retrieved.

If, however, the period or even the phase of the oscillatory part is not exactly known, then this method cannot be applied. Consider for instance

$$v^\varepsilon(s) := v_1\Big((1+\varepsilon^\beta)\frac{s}{\varepsilon^\alpha}\Big)$$

with $0 < \beta < \alpha$, v_1 continuous, one periodic, and with vanishing average on the period, and let ϕ be a test function as above. Then

$$\int_0^1 \phi\Big(s, \frac{s}{\varepsilon^\alpha}\Big)\, v^\varepsilon(s)\, ds \to 0.$$

Since we do not know the precise period of the minimizers of I^ε (and moreover we cannot expect precise periodicity if a is not constant) two-scale convergence does not suffice for our purposes.

The organization of the paper is as follows. In Section 2 we recall the notions of Young measures (associated to sequences of functions with values in a metric space) and Γ-convergence, We follow mainly [9] and [5], the omly new result concerns the convergence of functionals defined on Young measures (Theorem 2.12). Section

3 is the core of the paper, we obtain the Γ-limit of the functionals I^ε defined in (1.1) after suitable rescaling and extension to Young measures (Theorem 3.4). As a corollary we obtain Theorem 1.2 above (see Corollary 3.13). The proof of Theorem 3.4 is contained in Section 3 up to a density result to be discussed in Sections 4 and 5. More precisely, in these sections we show that every translation invariant measure on the space K of patterns can be approximated by a sequence of invariant measures, each of them being supported on the class of all translations of one-periodic function (see Theorems 4.4 and 4.15, and Corollary 5.11). In Section 6 we briefly sketch some extensions of our approach to other variational problems with multiple scales.

2. Young Measures which take values in a metric space

Young measures are maps from a measure space Ω to a space of probability measures on another space K. They arise naturally as limits of (usually rapidly oscillating) sequences of maps from Ω to K, and provide a good framework for existence of minimizers and optimal controls. Since L.C. Young's pioneering work [57], [58] there has been a large number of important contributions to this area, often in settings that are much more general than the one discussed below. We only mention here the fundamental papers of Berliocchi and Lasry [9] and Balder [4], the recent reviews of Valadier [54], [55] and the book by Roubiček [44]. A closely related but slightly different approach was pursued by Sychev [47], who emphasizes the view of Young measures as measurable maps into a suitable metric space and the use of selection theorems rather than the L^1-L^∞ duality. The theory of Young measures gained important momentum from the connections with partial differential equations and the theory of compensated compactness discovered by Tartar ([48], [49], [15], [40]), and with fine phase mixtures that arise in phase transitions modelled by nonconvex variational problems ([6], [13], [31], [32], [46], [42], [34]).

Our approach is inspired by [5] (see also the comments in [9], p. 180). The main new result concerns the convergence of functions defined on Young measures (see Theorem 2.12(iv)). Our point of view is the following: one can obtain precise information about the asymptotics of minimizers for a sequence of problems (such as the singularly perturbed problems studied in [39]) that involve maps from Ω to K by studying a limit problem defined on Young measures.

To proceed we first fix the notation. Throughout this paper, a measure on a topological space X is a σ-additive function on the σ-algebra of Borel sets. Unless stated differently, measurability always means Borel measurability.

In the rest of this section Ω is a locally compact separable and metrizable space, endowed with a finite measure λ (however, most of the results can be extended with some care to σ-finite measures). We often suppress explicit reference to λ. The case of an open set $\Omega \subset \mathbb{R}^n$ equipped with the Lebesgue measure suffices for the applications we have in mind.

We also consider a compact metric space (K,d), the class $\mathsf{Meas}(\Omega,K)$ of all measurable maps from Ω to K, the Banach space $C(K)$ of all continuous real functions on K, and the space $\mathscr{M}(K)$ of finite real Borel measures on K; $\mathscr{M}(K)$ is identified with the dual of $C(K)$ by the duality pairing $\langle\mu,g\rangle := \int_K g\,d\mu$ for $\mu\in\mathscr{M}(K)$, $g\in C(K)$, and is always endowed with the corresponding weak-star topology. For every $x\in K$, δ_x is the Dirac mass at x; $\mathscr{P}(K)$ is the set of all probability measure on K (that is, positive measures with mass equal to 1).

As far as possible we shall conform to the following notation: the letter s denotes a point in Ω, and x a point in K, μ is a measure in $\mathscr{M}(K)$, k is a positive integer, g, h, and f are real functions on K, on Ω, and on $\Omega\times K$, respectively; ϕ is a map from Ω to $C(K)$ and ν a map from Ω to $\mathscr{M}(K)$; we often use the notation f_s and ν_s to denote the function $f(s,\cdot)$ and the measure $\nu(s)$ respectively.

By $L^1\big(\Omega,C(K)\big)$ we denote the Banach space of all *measurable* maps $\phi:\Omega\to C(K)$ such that $\|\phi\|_1 := \int_\Omega |\phi(s)|_{C(K)}\,ds$ is finite. The space $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ is the Banach space of all *weak-star measurable* maps $\nu:\Omega\to\mathscr{M}(K)$ which are λ-essentially bounded, endowed with the obvious norm. More precisely, the elements of $L^1\big(\Omega,C(K)\big)$ and $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ are equivalence classes of maps which agree a.e.; we usually do not distinguish a map from its equivalence class.

Remarks. Since $C(K)$ is a separable Banach space, and Ω is endowed with a σ-finite measure λ, then the Banach space $L^1\big(\Omega,C(K)\big)$ is separable, while $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ is never separable unless λ is purely atomic and K is a finite set.

By definition, a map $\nu:\Omega\to\mathscr{M}(K)$ is weak-star measurable if the pre-image of every set in the Borel σ-algebra generated by the weak-star topology of $\mathscr{M}(K)$ is a Borel subset of Ω. Therefore the map ν is weak-star measurable if and only if the function $s\mapsto\langle\nu_s,g\rangle$ is measurable for every g in (a dense subset of) $C(K)$. Since $\mathscr{M}(K)$ is not separable, there are many weak-star measurable maps that are not strongly measurable; a typical example is the map which takes every s in an interval I into $\delta_s\in\mathscr{M}(I)$. Indeed the σ-algebra generated by the weak-star topology and the one generated by strong topology do not agree (the strong topology itself has cardinality strictly larger than the σ-algebra generated by the weak-star topology).

The space $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ is isometrically isomorphic to the dual of of the space $L^1\big(\Omega,C(K)\big)$ via the duality pairing (see [18], sec. 8.18.1)

$$\langle\nu,\phi\rangle_{L^\infty,L^1} := \int_\Omega \langle\nu_s,\phi_s\rangle_{\mathscr{M},C}\;ds,$$

with $\nu\in L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ and $\phi\in L^1\big(\Omega,C(K)\big)$. In the following we shall refer to the weak-star topology of $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ as the topology induced by this duality pairing. Since $L^1\big(\Omega,C(K)\big)$ is a separable Banach space, every closed ball in $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ endowed with the weak-star topology is compact and metrizable.

Remark 2.1. Given $g\in C(K)$ and $h\in L^1(\Omega)$, the map $h\otimes g$ which takes every $s\in\Omega$ into $h(s)\cdot g\in C(K)$ belongs to $L^1\big(\Omega,C(K)\big)$, and the class of all $h\otimes g$ with g and h ranging in dense subsets of $C(K)$ and $L^1(\Omega)$, respectively, spans a dense subspace of $L^1\big(\Omega,C(K)\big)$. Hence a bounded sequence (ν^k) in $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ weak-star converges to ν if and only if

$$\int_\Omega \langle\nu^k_s,g\rangle\,h(s)\,ds \to \int_\Omega\langle\nu_s,g\rangle\,h(s)\,ds \tag{2.1}$$

for every g, h in dense subsets of $C(K)$ and $L^1(\Omega)$, respectively. In particular this condition is immediately verified when $\nu^k_s \stackrel{*}{\rightharpoonup} \nu_s$ for almost every $s\in\Omega$.

Furthermore, on every bounded subset of $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ the weak-star topology is induced by the following norm:

$$\Phi(\nu) := \sum_{i,j}\frac{1}{2^{i+j}\alpha_{i,j}}\int_\Omega\langle\nu_s,g_i\rangle\,h_j(s)\,ds, \tag{2.2}$$

where the functions g_i, with $i=1,2,\dots$, are dense in $C(K)$, the functions h_j, with $j=1,2,\dots$, are bounded and dense in $L^1(\Omega)$, and $\alpha_{i,j} := \|g_i\|_\infty\cdot\|h_j\|_\infty$. In fact one easily checks that $\Phi(\nu^k-\nu)$ tends to 0 if and only if (2.1) holds with g and h replaced by g_i and h_j for all i,j.

Definition 2.2. *We call any map ν in $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ such that ν_s is a probability measure for a.e. $s\in\Omega$ a K-valued* Young measure *on Ω. The* elementary *Young measure associated to a measurable map $u:\Omega\to K$ is the map $\underline{\delta}_u$ given by*

$$\underline{\delta}_u(s) := \delta_{u(s)} \quad \text{for } s\in\Omega.$$

We say that sequence of measurable maps $u^k:\Omega\to K$ generates *the Young measure ν, if the corresponding elementary Young measures $\underline{\delta}_{u^k}$ converge to ν in the weak-star topology of $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$.*

We denote by $\mathsf{YM}(\Omega,K)$ (resp. $\mathsf{EYM}(\Omega,K)$) the set all Young measures (resp. elementary Young measures); $\mathsf{YM}(\Omega,K)$ is always endowed with the weak-star topology of $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$, and hence metrized by the norm Φ in (2.2).

Remarks. The map $\underline{\delta}_u$ is weak-star measurable if and only if u is measurable, and thus $\mathsf{EYM}(\Omega,K)$ is exactly the set of all $\nu\in L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$ such that ν_s is a Dirac mass for a.e. $s\in\Omega$.

Young measures are often defined as the weak-star closure of the class of elementary Young measures in $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$. By Theorem 2.3(iii) below, this definition turns out to be equivalent to ours when the measure λ is non-atomic. In [9] and [4] Young measures are endowed with the so-called narrow topology, which in the particular case we consider agrees with the weak-star topology of $L^\infty_w\big(\Omega,\mathscr{M}(K)\big)$.

The following theorem characterizes $\mathsf{YM}(\Omega,K)$ as the closure of $\mathsf{EYM}(\Omega,K)$.

Theorem 2.3. *Assume that the measure λ is non-atomic. Then*

(i) $\mathsf{YM}(\Omega, K)$ *is a weak-star compact, convex and metrizable subset of the space* $L^\infty_w\big(\Omega, \mathscr{M}(K)\big)$*;*
(ii) $\mathsf{EYM}(\Omega, K)$ *is the set of all extreme points of* $\mathsf{YM}(\Omega, K)$*;*
(iii) $\mathsf{EYM}(\Omega, K)$ *is weak-star dense in* $\mathsf{YM}(\Omega, K)$.

Proof. These three statements are given in [9], sec. II.2, as propositions 1 (p. 144), 3 (p. 146), and 4 (p. 148), respectively. □

Remarks. Statement (i) holds when λ is not non-atomic too. Statements (i) and (iii) show that from every sequence of measurable maps we can extract a subsequence which generates a Young measure, and conversely all Young measures are generated by sequences of measurable maps.

When the measure λ has atoms (namely, points with positive measure) it can be decomposed in a unique way as the sum of a non-atomic measure λ_n and a purely atomic measure λ_a (i.e., a countable linear combination of Dirac masses), and statements (ii) and (iii) of Theorem 2.3 should be modified as follows: the extreme points of $\mathsf{YM}(\Omega, K)$ are the Young measures ν such that ν_s is a Dirac mass for λ_n a.e. s, and the weak-star closure of $\mathsf{EYM}(\Omega, K)$ is the set of all $\nu \in \mathsf{YM}(\Omega, K)$ such that ν_s is a Dirac mass for λ_a a.e. s. The proof of this generalization is left to the interested reader.

Theorem 2.4 – Fundamental theorem of Young measures. *For every sequence of measurable maps* $u^k : \Omega \to K$ *there exists a subsequence (not relabeled) which generates a Young measure* ν*. Moreover* ν *has the following properties.*

(i) *If* $f : \Omega \times K \to \mathbb{R}$ *is measurable, continuous with respect to the second variable, and satisfies* $|f(s,x)| \le h(s)$ *for some* $h \in L^1(\Omega)$*, then*

$$\int_\Omega f\big(s, u^k(s)\big)\, ds \longrightarrow \int_\Omega \langle \nu_s, f_s \rangle\, ds \quad \text{as } k \to +\infty. \tag{2.3}$$

(ii) *The maps* u^k *converge in measure to some* $u : \Omega \to K$ *if and only if* ν *is the elementary Young measure associated to* u.
(iii) *Assume that* K *is a subset of a (separable) Banach space* E*, let* Id *denote the identity map on* E*, and define* $u : \Omega \to E$ *by*

$$u(s) := \int_K \mathrm{Id}\, d\nu_s. \tag{2.4}$$

Then $u(s)$ *is well-defined and belongs to the convex hull of* K *for a.e.* s*, and the maps* u^k *weak-star converge to* u *in* $L^\infty_w(\Omega, E)$*, that is, the functions* $s \mapsto \langle \Lambda, u^k(s)\rangle$ *weak-star converge to* $s \mapsto \langle \Lambda, u(s)\rangle$ *in* $L^\infty(\Omega)$ *for every* $\Lambda \in E^*$.

Regarding statement (iii), it is necessary to embed K in a linear structure in order to define the average (or expectation) $u(s)$. Notice moreover that the integral (2.4) is well-defined (e.g., as a Riemann integral) because K is compact and metrizable and Id is a continuous map on K. Moreover u is measurable because one has $\langle \Lambda, u(s)\rangle = \langle \nu_s, \Lambda_{|K}\rangle$, and $\Lambda_{|K}$ belongs to $C(K)$ for every $\Lambda \in E^*$.

Proof. The existence of a subsequence of (u^k) which generates a Young measure ν follows from the compactness and metrizability of $\mathsf{YM}(\Omega, K)$ (Theorem 2.3(i)).

To prove (i), notice that the map $s \mapsto f_s$ belongs to $L^1\big(\Omega, C(K)\big)$ (cf. [9], remark 5, sec. I.1, p. 135), and then (2.3) follows immediately from the definition of weak-star convergence in $L^\infty_w\big(\Omega, \mathscr{M}(K)\big)$.

We assume now that the maps u^k generate the Young measure $\underline{\delta}_u$, and we apply statement (i) with $f(s,x) := d\big(x, u(s)\big)$. Then

$$\int_\Omega d\big(u^k(s), u(s)\big)\, ds \to \int_\Omega d\big(u(s), u(s)\big)\, ds = 0,$$

and we deduce that u^k converge to u in measure.

Conversely, assume that the maps u^k converge to u in measure. Taking f as above, the integrals $\int f(s, u^k(s))ds$ converge to 0 by the Lebesgue dominated convergence theorem, and by (2.3) we obtain

$$\int_\Omega \langle \nu_s, f_s \rangle\, ds = 0. \tag{2.5}$$

Since f is non-negative, (2.5) implies that for a.e. $s \in \Omega$ the measure ν_s is supported on the set of all $x \in K$ such that $f_s(x) = 0$, that is, on the point $u(s)$. Thus $\nu_s = \delta_{u(s)}$, and statement (ii) is proved.

Finally (iii) follows by applying (i) with $f(s,x) := h(s)\,\langle \Lambda, x\rangle$ for $h \in L^1(\Omega)$, $\Lambda \in E^*$. □

Before discussing functionals on $\mathsf{YM}(\Omega, K)$, we add some elementary but useful remarks.

Remark 2.5. If the functions u^k generate a Young measure ν on Ω, then they generate the same Young measure on every Borel subset of Ω, that is, $\underline{\delta}_{u^k}$ weak-star converges to ν in $L^\infty\big(A, \mathscr{M}(K)\big)$ for every $A \subset \Omega$. Consequently, if two sequences (u_1^k) and (u_2^k) generate the Young measures ν_1 and ν_2, respectively, and the maps u_1^k and u_2^k agree on a fixed Borel set A for k sufficiently large, then ν_1 and ν_2 agree (a.e.) on A.

Remark 2.6. We say that the sequences (u_1^k) and (u_2^k) are asymptotically equivalent when the functions $s \mapsto d\big(u_1^k(s), u_2^k(s)\big)$ converge in measure to 0 as $k \to +\infty$. One easily checks, using the convergence criterion in Remark 2.1, that asymptotically equivalent sequences generate the same Young measures.

Lemma 2.7. *Let* $\Omega \subset \mathbb{R}^n$ *be endowed with Lebesgue measure. Consider a sequence of maps* u^k *which generate the Young measure* ν*, and are defined on a fixed neighborhood of* Ω*, and a sequence of vectors* τ_k *in* $\mathbb{R}^n$ *which converge to* 0*. Then the translated maps* $u^k(\cdot - \tau_k)$ *also generate the Young measure* ν.

PROOF. Take $g \in C(K)$ and $h \in L^1(\Omega)$, extended to 0 on $\mathbb{R}^n \setminus \Omega$. Then

$$\begin{aligned}\int h(s)\, g\big(u^k(s-\tau_k)\big)\, ds &\\ &= \int h(s+\tau_k)\, g\big(u^k(s)\big)\, ds \\ &= \int \big(h(s+\tau_k) - h(s)\big)\, g\big(u^k(s)\big)\, ds + \int h(s)\, g\big(u^k(s)\big)\, ds. \end{aligned} \tag{2.6}$$

Now the second integral in line (2.6) converges to $\int h(s)\langle \nu_s, g\rangle\, ds$ by assumption, while the modulus of the first one is controlled by $\big\|h(\cdot+\tau_k) - h(\cdot)\big\|_1 \cdot \big\|g(\cdot)\big\|_\infty$, and since the first term tends to 0 for every $h \in L^1(\mathbb{R}^n)$, we obtain

$$\int h(s)\, g\big(u^k(s-\tau_k)\big)\, ds \longrightarrow \int h(s)\, \langle \nu_s, g\rangle\, ds.$$

By Remark 2.1 this suffices to prove the assertion. □

In the following we introduce integral functionals on the class of measurable maps $\mathsf{Meas}(\Omega, K)$, and we show how to extend them to all Young measures. Then we discuss some semicontinuity properties of these extensions, and their behavior with respect to relaxation and Γ-convergence (Theorem 2.12). In order to do this, we briefly recall the definitions of relaxation and Γ-convergence (we refer to [14], chaps. 3-8, for more general definitions and further details).

DEFINITION 2.8 - RELAXATION. *Let X be a metric space and let $F : X \to [0, +\infty]$. The* relaxation $\overline{F}$ of F on X *is the lower semicontinuous envelope of F, that is, the supremum of all lower semicontinuous functions which lie below F. Alternatively $\overline{F}$ is characterized by the formula:*

$$\overline{F}(x) = \inf\big\{ \liminf_{k\to\infty} F(x^k) \ :\ x^k \to x \big\}. \tag{2.7}$$

DEFINITION 2.9 - Γ-CONVERGENCE. *Let X be a metric space. A sequence of functions $F^\varepsilon : X \to [0, +\infty]$* Γ-converge *to F on X, and we write $F^\varepsilon \xrightarrow{\Gamma} F$, if the following two properties are fulfilled:*

- lower bound inequality: $\forall\, x \in X$, $\forall\, x^\varepsilon \to x$, $\liminf F^\varepsilon(x^\varepsilon) \geq F(x)$;
- upper bound inequality: $\forall\, x \in X$, $\exists\, x^\varepsilon \to x$ *s.t.* $\lim F^\varepsilon(x^\varepsilon) = F(x)$.

We say that the functions F^ε converge continuously *to F on X if $F^\varepsilon(x^\varepsilon) \to F(x)$ whenever $x^\varepsilon \to x$, and that they are* equicoercive *on X if every sequence (x^ε) such that $F^\varepsilon(x^\varepsilon)$ is bounded is pre-compact in X.*

Here and in the following we use the term "sequence" also to denote families (of points of X) labelled by the continuous parameter ε, which tends to 0. A subsequence of (x^ε) is any sequence (x^{ε_n}) such that $\varepsilon_n \to 0$ as $n \to +\infty$, and we say that (x^ε) is pre-compact in X if every subsequence admits a sub-subsequence which converges in X. To simplify the notation we often omit to relabel subsequences, and we say "a countable sequence (x^ε)" to mean a sequence defined only for countably many $\varepsilon = \varepsilon_n$ such that $\varepsilon_n \to 0$.

REMARK 2.10. Given a lower semicontinuous function $F : X \to [0, +\infty]$, we say that a set $\mathscr{D}$ is F-dense in X if for every $x \in X$ with $F(x) < +\infty$ there exists a sequence $(x^k) \subset \mathscr{D}$ such that $x^k \to x$ and $F(x^k) \to F(x)$. A simple diagonal argument shows that the upper bound inequality in Definition 2.9 is verified provided that for every x in some F-dense set $\mathscr{D}$ and every $\delta > 0$ we can find a sequence (x^ε) such that

$$\limsup_{\varepsilon\to 0} d(x^\varepsilon, x) \leq \delta \quad \text{and} \quad \limsup_{\varepsilon\to 0} F^\varepsilon(x^\varepsilon) \leq F(x) + \delta. \tag{2.8}$$

PROPOSITION 2.11 (see [14], chaps. 6 and 7). *We have the following:*

(i) *every Γ-limit F is lower semicontinuous on X;*
(ii) *the constant sequence $F^\varepsilon := F$ Γ-converge on X to $\overline{F}$;*
(iii) $F^\varepsilon \xrightarrow{\Gamma} F$ *if and only if* $\overline{F^\varepsilon} \xrightarrow{\Gamma} F$;
(iv) *if $F^\varepsilon \xrightarrow{\Gamma} F$ and $G^\varepsilon \to G$ continuously, then $(F^\varepsilon + G^\varepsilon) \xrightarrow{\Gamma} (F + G)$;*
(v) *assume that the functions F^ε are equicoercive and Γ-converge to F on X, and that X is continuously embedded in X': if we extend F^ε and F to $+\infty$ on $X' \setminus X$, then $F^\varepsilon \xrightarrow{\Gamma} F$ on X';*
(vi) *if the points $\bar{x}^\varepsilon$ minimize F^ε for every ε, then every cluster point of the sequence $(\bar{x}^\varepsilon)$ minimizes F.*

We next consider integral functionals on measurable maps from Ω to K and their extension to Young measures. An *integrand* on $\Omega \times K$ is a measurable function $f : \Omega \times K \to [0, +\infty]$. Each integrand f defines a functional on $\mathsf{Meas}(\Omega, K)$ via

$$u \mapsto \int f(s, u(s))\, ds.$$

This can be viewed as a functional on elementary Young measures, and extended to $\mathsf{YM}\,(\Omega, K)$ in two natural ways: by $+\infty$ or by linearity. More precisely, we set

$$\mathscr{F}_f(\nu) := \begin{cases} \int_\Omega f\big(s, u(s)\big)\, ds & \text{if } \nu = \underline{\delta}_u \text{ for some } u, \\ +\infty & \text{otherwise,} \end{cases} \tag{2.9}$$

and

$$F_f(\nu) := \int_\Omega \langle \nu_s, f_s\rangle\, ds \qquad \text{for every } \nu \in \mathsf{YM}\,(\Omega, K). \tag{2.10}$$

Clearly for every elementary Young measure $\underline{\delta}_u$ we have

$$F_f(\underline{\delta}_u) = \mathscr{F}_f(\underline{\delta}_u) = \int_\Omega f\big(s, u(s)\big)\, ds. \tag{2.11}$$

Theorem 2.12 below shows that the relaxation or the Γ-convergence of functionals of type (2.9) always lead to functionals of type (2.10). We recall that the set $\mathsf{YM}(\Omega, K)$ is always endowed with the weak-star topology of $L^\infty_w\big(\Omega, \mathscr{M}(K)\big)$, which makes it compact and metrizable.

THEOREM 2.12. *If the measure λ is non-atomic the following holds.*

(i) *If the integrand f satisfies $f(s,x) \le h(s)$ for some $h \in L^1(\Omega)$ and f_s is continuous on K for a.e. $s \in \Omega$, then F_f is continuous and finite on $\mathsf{YM}(\Omega, K)$.*

(ii) *If f_s is lower semicontinuous on K for a.e. $s \in \Omega$ then F_f is lower semicontinuous on $\mathsf{YM}(\Omega, K)$.*

(iii) *The relaxation of $\mathscr{F}_f$ and F_f on $\mathsf{YM}(\Omega, K)$ is the functional $F_{\hat f}$ where $\hat f$ is any integrand such that $\hat f_s$ agrees with the relaxation of f_s on K for a.e. $s \in \Omega$.*

(iv) *Assume that the integrands f^ε satisfy $f^\varepsilon_s \xrightarrow{\Gamma} f_s$ on K for a.e. $s \in \Omega$, and that the envelope functions $\mathrm{E}f^\varepsilon$ defined by*

$$\mathrm{E}f^\varepsilon(s) := \inf_{x\in K} f^\varepsilon(s,x) \quad \text{for } s \in \Omega, \tag{2.12}$$

are equi-integrable on Ω. Then $\mathscr{F}_{f^\varepsilon} \xrightarrow{\Gamma} F_f$ and $F_{f^\varepsilon} \xrightarrow{\Gamma} F_f$ on $\mathsf{YM}(\Omega, K)$.

REMARKS. Concerning statement (iii), we remark that such an integrand $\hat f$ exists in view of Lemma 2.14 below (this is a subtle point: the map $(s,x) \mapsto \bar f_s(x)$ may be not Borel measurable on $\Omega \times K$).

In statement (iv), we notice that the assumption $f^\varepsilon_s \xrightarrow{\Gamma} f_s$ for almost every $s \in \Omega$ is quite strong, and far from necessary. Indeed the Γ-convergence of the functionals may occur even with a more complicate asymptotic behavior of the integrands (e.g., some kind of homogenization), but the analysis of such situations is beyond the purposes of this paper.

If the functions $\mathrm{E}f^\varepsilon$ in (2.12) are not equi-integrable on Ω, some concentration effect occurs, and the Γ-convergence results may not hold. In particular, if $\|\mathrm{E}f^\varepsilon\|_1 \to +\infty$ then $\mathscr{F}_{f^\varepsilon}$ and F_{f^ε} Γ-converge to the constant functional $+\infty$. On the other hand, if there exist sets $B^\varepsilon \subset \Omega$ such that $|B^\varepsilon| \to 0$, the restrictions of $\mathrm{E}f^\varepsilon$ to the complements of B^ε are equi-integrable on Ω, and $\int_{B^\varepsilon} \mathrm{E}f^\varepsilon ds$ converge to some constant c, then both $\mathscr{F}_{f^\varepsilon}$ and F_{f^ε} Γ-converge to $F_f + c$ (this generalization of statement (iv) can be proved by suitably modifying the proof below). However, both $\mathscr{F}_{f^\varepsilon}$ and F_{f^ε} verify the lower bound inequality without any assumption on $\mathrm{E}f^\varepsilon$.

Finally we notice that the functions $\mathrm{E}f^\varepsilon$ are λ-measurable (see for instance [11], lemma III.39) and therefore they agree a.e. with Borel functions.

PROOF OF STATEMENTS (i) AND (ii) OF THEOREM 2.12. Regarding (i), one can easily verify that the map $s \mapsto f_s$ belongs to $L^1\big(\Omega, C(K)\big)$ (cf. [9], remark 5, sec. I.1, p. 135). Hence F_f belongs to the pre-dual of $L^\infty_w\big(\Omega, \mathscr{M}(K)\big)$, and is therefore (weak-star) continuous on $\mathsf{YM}(\Omega, K)$.

Assertion (ii) is contained in [9], proposition 3, sec. II.1, p. 152, and theorem 2, sec. I.3, p. 138. Alternatively one can use (i) and the approximation from below established in Lemmas 2.13 and 2.14 below. □

To prove assertions (iii) and (iv) of Theorem 2.12 we need two lemmas on approximation by continuous integrands and a density result for Young measures ν with finite energy $F_f(\nu)$.

LEMMA 2.13. *Consider an integrand f and for every integer k set*

$$f^k(x,s) := k \wedge \inf_{x'\in K} \big[f(x') + k \cdot d(x,x')\big] \quad \text{for } s \in \Omega,\ x \in K \tag{2.13}$$

(here $a \wedge b$ denotes, as usual, the minimum of a and b). Then

(i) *for every s, f^k_s is k-Lipschitz on K and $0 \le f^k_s \le k$;*

(ii) *for every s, f^k_s increases to the relaxation of f_s as $k \nearrow +\infty$;*

(iii) *there exists a negligible set $N \subset \Omega$ such that each f^k is measurable on $(\Omega \setminus N) \times K$.*

PROOF. Statements (i) and (ii) follow by straightforward computations. Statement (iii) is slightly more subtle, and indeed f^k may be not Borel measurable on $\Omega \times K$. Let k be fixed. For every Borel function g on $\Omega \times K$ the map $s \mapsto \inf\big\{g(s,x) : x \in K\big\}$ is λ-measurable (cf. [11], lemma III.39) and thus it agrees a.e. in Ω with a Borel function. Hence for every $x \in K$ we can find a negligible Borel set $N^k_x \subset \Omega$ such that the map $s \mapsto f^k(s,x)$ is Borel measurable on $\Omega \setminus N^k_x$ (cf. (2.13)). Now we take N^k as the union of all N^k_x as x ranges in a countable dense subset $\mathscr{D}$ of K, thus N^k is a negligible Borel set, f^k is Borel measurable in $(\Omega \setminus N^k) \times \mathscr{D}$, and then also on $(\Omega \setminus N^k) \times K$ because f^k is continuous in the second variable and $\mathscr{D}$ is dense in K. Finally we take $N := \cup_k N^k$. □

LEMMA 2.14. *Consider an integrand f and let $\bar f_s$ be the relaxation of f_s for every $s \in \Omega$. Then there exists a negligible Borel set $N \subset \Omega$ such that $\bar f$ is Borel-measurable on $(\Omega \setminus N) \times K$. In particular there exists an integrand $\hat f$ such that $\hat f_s$ is the relaxation of f_s for a.e. $s \in \Omega$.*

PROOF. Take N as in statement (iii) of Lemma 2.13: all the functions f^k are Borel-measurable on $(\Omega \setminus N) \times K$, and then the same holds for $\bar f$ by statement (ii) of Lemma 2.13. □

PROPOSITION 2.15 ([9], proposition 1, sec. II.2, p. 144). *Assume that λ is non-atomic. Consider an integrand f such that f_s is lower semicontinuous on K for a.e. $s \in \Omega$, and the set*

$$M_f := \big\{\nu \in \mathsf{YM}(\Omega, K) : F_f(\nu) \le 1\big\}. \tag{2.14}$$

Then $\mathsf{EYM}(\Omega, K) \cap M_f$ is dense in M_f.

THEOREM 2.16. *Take f as in Proposition 2.15. Then $\mathsf{EYM}(\Omega, K)$ is F_f-dense in $\mathsf{YM}(\Omega, K)$, that is, for every $\nu \in \mathsf{YM}(\Omega, K)$ such that $F_f(\nu) < +\infty$ there exist* *a sequence of elementary Young measures ν^k such that $\nu^k \xrightarrow{*} \nu$ and $F_f(\nu^k) \to F_f(\nu)$.*

PROOF. We may assume without loss of generality that $F_f(\nu) = 1$. Then $\nu \in M_f$ and by Proposition 2.15 we can find a sequence of elementary Young measures $(\nu^k) \subset M_f$ which converge to ν. Then $F_f(\nu^k) \le F_f(\nu)$ for every k, and since F_f is lower semicontinuous, we deduce that $F_f(\nu^k)$ converge to $F_f(\nu)$. □

PROOF OF STATEMENTS (iii) AND (iv) OF THEOREM 2.12. Statement (iii) of Theorem 2.12 follows from statement (iv) and Proposition 2.11(ii).

To prove statement (iv), it suffices to prove the lower bound inequality for the functionals F_{f^ε}, and then the upper bound inequality for the functionals $\mathscr{F}_{f^\varepsilon}$ (recall that $F_{f^\varepsilon} \le \mathscr{F}_{f^\varepsilon}$).

For the lower bound inequality, we begin with a simple remark: if $g^\varepsilon \xrightarrow{\Gamma} g$ on K, then for every continuous function g' such that $g > g'$ on K there holds $g^\varepsilon \ge g'$ on K for every ε sufficiently small (this can be easily proved by contradiction).

We fix now an integer k and we take f^k as in (2.13) (setting it equal to 0 in the set N given in Lemma 2.13(iii) to make it Borel measurable). Since $f_s \ge f^k_s$ on K (see Lemma 2.13(ii)), there holds $f_s > f^k_s - 1/k$, and since f^k_s is continuous on K (Lemma 2.13(i)), by the previous remark

$$f^\varepsilon_s \ge f^k_s - \frac{1}{k} \tag{2.15}$$

for ε sufficiently small. Consider a sequence of Young measures ν^ε which converge to ν in $\mathsf{YM}(\Omega, K)$. Then (2.15) yields

$$\begin{aligned} \liminf_{\varepsilon\to 0} F_{f^\varepsilon}(\nu^\varepsilon) = \liminf_{\varepsilon\to 0} \int_\Omega \langle \nu^\varepsilon_s, f^\varepsilon_s \rangle \, ds &\ge \liminf_{\varepsilon\to 0} \int_\Omega \Big\langle \nu^\varepsilon_s . f^k_s - \frac{1}{k} \Big\rangle \, ds \\ &= \int_\Omega \langle \nu_s, f^k_s \rangle \, ds - \frac{1}{k}\lambda(\Omega), \end{aligned} \tag{2.16}$$

where the last equality follows from statement (i) of Theorem 2.12. Now we pass to the limit in (2.16) as $k \to +\infty$, and by Lemma 2.13(ii) and the monotone convergence theorem we deduce

$$\liminf_{\varepsilon\to 0} F_{f^\varepsilon}(\nu^\varepsilon) \ge \int_\Omega \lim_{k\to\infty} \langle \nu_s, f^k_s \rangle \, ds = \int_\Omega \langle \nu_s, f_s \rangle \, ds = F_f(\nu).$$

We consider now the upper bound inequality. Since $\mathsf{EYM}(\Omega, K)$ is F_f-dense in $\mathsf{YM}(\Omega, K)$ (Theorem 2.16) and each $\mathscr{F}_{f^\varepsilon}$ is finite only on $\mathsf{EYM}(\Omega, K)$, by Remark 2.10 it suffices to show that every elementary Young measure can be approximated in energy by a sequence of elementary Young measures; more precisely, for every $u \in \mathsf{Meas}(\Omega, K)$ we shall exhibit a sequence of maps u^ε which converge to u a.e. in Ω and satisfy

$$\lim_{\varepsilon\to 0} \int_\Omega f^\varepsilon\big(s, u^\varepsilon(s)\big) \, ds = \int_\Omega f\big(s, u(s)\big) \, ds. \tag{2.17}$$

Since $f^\varepsilon_s \xrightarrow{\Gamma} f_s$ for a.e. $s \in \Omega$, for every $\varepsilon > 0$ and a.e. $s \in \Omega$ we can choose $x^\varepsilon_s \in K$ so that

$$\lim_{\varepsilon\to 0} x^\varepsilon_s = u(s) \quad \text{and} \quad \lim_{\varepsilon\to 0} f^\varepsilon(s, x^\varepsilon_s) = f\big(s, u(s)\big). \tag{2.18}$$

By (2.12), for every $\varepsilon > 0$ and a.e. $s \in \Omega$ we can also choose y^ε_s so that

$$f^\varepsilon(s, y^\varepsilon_s) \le \mathrm{E}f^\varepsilon(s) + 1. \tag{2.19}$$

We thus define the approximating maps $u^\varepsilon : \Omega \to K$ by

$$u^\varepsilon(s) := \begin{cases} x^\varepsilon_s & \text{if } f^\varepsilon(s, x^\varepsilon_s) \le f\big(s, u(s)\big) + 1, \\ y^\varepsilon_s & \text{otherwise.} \end{cases} \tag{2.20}$$

From (2.18) we deduce that for a.e. $s \in \Omega$ there holds $u^\varepsilon(s) = x^\varepsilon_s$ for ε small enough, and thus $u^\varepsilon(s) \to u(s)$ and $f^\varepsilon\big(s, u^\varepsilon(s)\big) \to f\big(s, u(s)\big)$. We claim that the functions $s \mapsto f^\varepsilon\big(s, u^\varepsilon(s)\big)$ are equi-integrable, henceforth (2.17) follows from Lebesgue's dominated convergence theorem. To prove the claim, notice that by (2.20) and (2.19)

$$f^\varepsilon\big(s, u^\varepsilon(s)\big) \le \mathrm{E}f^\varepsilon(s) + f\big(s, u(s)\big) + 1,$$

and that the functions $\mathrm{E}f^\varepsilon$ are equi-integrable by assumption, while $f\big(s, u(s)\big)$ is summable.

To complete the proof of the upper bound inequality, we have to show that for every fixed $\varepsilon > 0$ the maps $s \mapsto y^\varepsilon_s$ and $s \mapsto x^\varepsilon_s$ can be chosen Borel measurable.

In the first case, we modify $\mathrm{E}f^\varepsilon$ in a negligible set in order to make it Borel measurable (cf. the remarks after Theorem 2.12); hence the set of all $(s, y) \in \Omega \times K$ which satisfy $f^\varepsilon(s, y) \le \mathrm{E}f^\varepsilon(s) + 1$ is Borel measurable and the projection on Ω is equal to Ω, and we can apply the Von Neumann-Aumann measurable selection theorem (see [11], theorem III.22) to find a λ-measurable selection $s \mapsto y^\varepsilon_s$ (which henceforth fulfills (2.19)); finally we modify such a map in a negligible set to make it Borel measurable.

In the second case we need to refine the previous argument. First we set

$$h^\varepsilon(s) := \inf_{x\in K} \Big[\big| f^\varepsilon(s, x) - f(s, u(s)) \big| + d(x, u(s)) \Big];$$

the function h^ε is λ-measurable (see [11], lemma III.39), and thus we can modify it in a negligible set to make it Borel measurable. Hence the set of all $(s, x) \in \Omega \times K$ which satisfy

$$\big| f^\varepsilon(s, x) - f(s, u(s)) \big| + d(x, u(s)) \le h^\varepsilon(s) + \varepsilon \tag{2.21}$$

is Borel measurable, its projection on Ω is equal to Ω, and can we proceed as before to find a Borel measurable selection map $s \mapsto x^\varepsilon_s$ which satisfies (2.21) for a.e. $s \in \Omega$. One readily checks that $h^\varepsilon(s) \to 0$ for a.e. $s \in \Omega$, and thus (2.18) holds. □

3. Application to a two-scale problem

In this section we apply the notion of Young measure developed in Section 2 to the two-scale problem presented in the introduction.

We first introduce some additional notation. As in Section 2, measurability always means Borel measurability; for sequences we follow the convention introduced after Definition 2.9. Throughout this section Ω is a bounded open interval endowed with Lebesgue measure, the letter s denotes the (slow) variable in Ω and v is a real function on Ω, periodically extended out of Ω; The letter x denotes functions of the (fast) variable $t \in \mathbb{R}$; the space of patterns K is the set of all measurable functions $x : \mathbb{R} \to [-\infty, +\infty]$ modulo equivalence almost everywhere, and G is the group of functional translations on K. We represent G by $\mathbb{R}$: for every $\tau \in \mathbb{R}$ and every function $x \in K$, $T_\tau x$ is the translated function $x(t - \tau)$. Thus a function x in K is h-periodic if $T_h x = x$.

By identifying the extended real line $[-\infty, +\infty]$ with the closed interval $[-1, 1]$ via the function $x \mapsto \frac{2}{\pi}\arctan(x)$, we can identify K with the closed unit ball of $L^\infty(\mathbb{R})$ and endow it with the weak-star topology. Thus K is compact and metrizable (a distance is given in (5.1) taking $n = 1$) and G acts continuously on K (cf. Proposition 5.3). If the functions x^k converge to some x pointwise a.e., or even in measure, then they converge to x also in the topology of K; in particular the Fréchet space $L^p_{\mathrm{loc}}(\mathbb{R})$ embeds continuously in K for $1 \le p \le \infty$. See Section 5 for more details and precise statements.

For every measure μ on K and every $\tau \in \mathbb{R}$, $T_\tau^\# \mu$ is the push-forward of the measure μ according to the map $T_\tau : K \to K$, that is, $T_\tau^\# \mu(B) := \mu(T_\tau^{-1} B)$ for every measurable $B \subset K$. We say that a probability measure μ on K is *invariant* if it is invariant under the action of the group G, namely if $\mu(T_\tau B) = \mu(B)$ for every $B \subset K$ and every $\tau \in \mathbb{R}$; $\mathscr{I}(K)$ is the class of all invariant probability measures on K. The *orbit* of $x \in K$ is the set $\mathscr{O}(x)$ of all translations of x; this set is compact in K whenever x is periodic. In this case ϵ_x is the measure defined by

$$\langle \epsilon_x, g \rangle = ⨍_0^h g(T_\tau x)\, d\tau \tag{3.1}$$

for every positive Borel function g on K (here h is the period of x); ϵ_x is the unique invariant probability measure supported on $\mathscr{O}(x)$, and we call it the *elementary invariant measure associated to* x (see Section 4, and in particular Lemma 4.10).

For every bounded open interval I, we write $H^2_{\mathrm{per}}(I)$ (resp., $W^{k,p}_{\mathrm{per}}(I)$) to denote the Sobolev space of all real functions on I, extended to $\mathbb{R}$ by periodicity, which belong to $H^2_{\mathrm{loc}}(\mathbb{R})$ (resp., to $W^{k,p}_{\mathrm{loc}}(\mathbb{R})$), and by $\mathscr{S}(I)$ the class of all functions $x \in K$ which are continuous and piecewise affine on the interval I with slope ± 1 only (sawtooth functions); we denote by $S\dot{x}$ the set of all points in $\mathbb{R}$ where x is not differentiable, and thus the points in $S \cap I$ are "corners" of x; $\mathscr{S}_{\mathrm{per}}(I)$ is the class of all real functions on I extended to $\mathbb{R}$ by periodicity and of class $\mathscr{S}$ on every bounded interval. The space $\mathscr{S}(I)$ can be characterized as the class of all functions $x \in K$ which are continuous on I and whose distributional derivative $\dot{x}$ is a BV function on I and takes values ± 1 only; if $S\dot{x} \cap I$ consists of the points $t_1 < t_2 < \ldots < t_N$, then $\ddot{x} = \pm \sum (-1)^i 2\delta_{t_i}$, and in particular the total variation of the measure $\ddot{x}$ on I is twice the number of points of $S\dot{x} \cap I$. In short $\|\ddot{x}\| = 2\,\#(S\dot{x} \cap I)$.

For every function v and every $s \in \Omega$ the ε-*blowup* of v at s is defined by

$$\mathsf{R}^\varepsilon_s v(t) := \varepsilon^{-1/3}\, v(s + \varepsilon^{1/3} t) \quad \text{for } t \in \mathbb{R}. \tag{3.2}$$

The ε-*blowup* of v is the map $\mathsf{R}^\varepsilon v$ which takes every $s \in \Omega$ into $\mathsf{R}^\varepsilon_s v \in K$.

As we explained in the introduction, our goal is to identify the Young measures $\nu \in \mathsf{YM}(\Omega, K)$ generated as $\varepsilon \to 0$ by ε-blowups of minimizers v^ε of the functionals

$$I^\varepsilon(v) := \int_\Omega \varepsilon^2 \ddot{v}^2 + W(\dot{v}) + a\, v^2\, ds, \tag{3.3}$$

where $v \in H^2_{\mathrm{per}}(\Omega)$, $a \in L^\infty(\Omega)$ is strictly positive a.e., and W is a continuous non-negative function on $\mathbb{R}$ which vanishes at ± 1 only and has growth at least linear at infinity. Notice that the assumption $a \in L^\infty(\Omega)$ (see Theorem 1.2) can be probably replaced by $a \in L^1(\Omega)$; this would only require a modification of the final part of the proof of the upper bound inequality in Theorem 3.4, and precisely the proof of estimates (3.29 – 32).

The goal is achieved in several steps:

Step 1. Identify the class of all Young measures $\nu \in \mathsf{YM}(\Omega, K)$ which are generated by sequences of ε-blowups of functions v^ε (Proposition 3.1 and Remark 3.2).

Step 2. Write the rescaled functionals $\varepsilon^{-2/3} I^\varepsilon(v)$ as $\int_\Omega f^\varepsilon_s(\mathsf{R}^\varepsilon_s v)\, ds$ for suitable functionals f^ε_s on K (cf. (3.6) and (3.7)).

Step 3. Identify the Γ-limit f_s of f^ε_s as $\varepsilon \to 0$ for a.e. $s \in \Omega$ (Proposition 3.3).

Step 4. Prove that the Γ-limit of the rescaled functionals $\varepsilon^{-2/3} I^\varepsilon$, viewed as functionals of the elementary Young measures associated with ε-blowups of functions, is given by $F(\nu) := \int_\Omega \langle \nu_s, f_s \rangle\, ds$ for all Young measures ν described in Step 1 (Theorem 3.4).

Step 5. If ν is a Young measure generated by ε-blowups of minimizers v^ε of I^ε, use Step 4 to show that ν minimizes F (Corollary 3.11); then use this fact to identify ν (Theorem 3.12 and Corollary 3.13).

Step 1. Admissible Young measures

The first step of our program consists in understanding which $\nu \in \mathsf{YM}(\Omega, K)$ are generated by the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$ of functions v^ε. We have the following result:

Proposition 3.1. *Let $\nu \in \mathsf{YM}(\Omega, K)$ be a Young measure generated by the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$ of a countable sequence of measurable functions v^ε. Then ν_s is an invariant measure on K of a.e. $s \in \Omega$.*

Proof. Set $u^\varepsilon := \mathsf{R}^\varepsilon v^\varepsilon$ for every $\varepsilon > 0$ and fix $\tau \in \mathbb{R}$. By (3.2) we have

$$T_\tau\big(u^\varepsilon(s)\big) = u^\varepsilon(s - \varepsilon^{1/3}\tau). \tag{3.4}$$

Since the functions u^ε generate the Young measure ν, the functions $T_\tau u^\varepsilon$ generate the Young measure $T_\tau^\# \nu$; on the other hand Lemma 2.7 shows that the functions $u^\varepsilon(\cdot + \varepsilon^{1/3}\tau)$ generate the same Young measure as the functions u^ε, and thus identity (3.4) yields $T^\# \nu = \nu$. Therefore we can find a negligible set $N_\tau \subset \mathbb{R}$ such that

$$T_\tau^\# \nu_s = \nu_s \qquad \text{for every } s \in \mathbb{R} \setminus N_\tau .$$

Let now N be the union of N_τ over all rational τ. Then N is negligible, and for every $s \in \Omega \setminus N$ there holds $T_\tau^\# \nu_s = \nu_s$ for every rational τ, and by approximation also for every real τ (the map $\tau \mapsto T_\tau^\# \mu$ is weak-star continuous for every $\mu \in \mathscr{P}(K)$). Hence ν_s is an invariant measure. □

Remark 3.2. The converse of Proposition 3.1 is also true: for every $\nu \in \mathsf{YM}(\Omega, K)$ such that $\nu_s \in \mathscr{I}(K)$ for a.e. $s \in \Omega$ we can find functions v^ε such that $\mathsf{R}^\varepsilon v^\varepsilon$ generate ν. The proof of this fact is more difficult, and is essentially included in the proof of Theorem 3.4 below.

Step 2. Rewriting $I^\varepsilon(v)$ in term of $\mathsf{R}^\varepsilon_s v$

We extend a by periodicity out of Ω and set $a_s^\varepsilon(t) := a(s + \varepsilon^{1/3} t)$ for every s and t. We fix a function $v \in H^2_{\mathrm{per}}(\Omega)$ and set $x_s := \mathsf{R}^\varepsilon_s v$ for every $s \in \Omega$. Thus

$$\begin{aligned} x_s(t) &= \varepsilon^{-1/3} v(s + \varepsilon^{1/3} t), \\ \dot x_s(t) &= \dot v(s + \varepsilon^{1/3} t), \\ \ddot x_s(t) &= \varepsilon^{1/3} \ddot v(s + \varepsilon^{1/3} t). \end{aligned}$$

Hence

$$\varepsilon^{4/3}\ddot v^2 + \varepsilon^{-2/3} W(\dot v) + \varepsilon^{-2/3} a v^2 = \varepsilon^{2/3} \ddot x_s^2 + \varepsilon^{-2/3} W(\dot x_s) + a_s^\varepsilon x_s^2, \tag{3.5}$$

where all functions on the left-hand side are computed at $s + \varepsilon^{1/3} t$, and those on the right-hand side are computed at t.

Now we fix $r > 0$ and for every x of class H^2 on $(-r, r)$ we set

$$f_s^\varepsilon(x) := \fint_{-r}^{r} \varepsilon^{2/3} \ddot x^2 + \varepsilon^{-2/3} W(\dot x) + a_s^\varepsilon x^2 \, dt. \tag{3.6}$$

Therefore, taking the average of the right-hand side of (3.5) over all $t \in (-r, r)$ and then integrating over all $s \in \Omega$ we get $\int_\Omega f_s^\varepsilon(x_s)\, ds$. On the other hand, if we integrate the left-hand side of (3.5) over all $s \in \Omega$ we get $\varepsilon^{-2/3} I^\varepsilon(v)$ for every t, and nothing changes if we take the average over all $t \in (-r, r)$. Therefore

$$\varepsilon^{-2/3} I^\varepsilon(v) = \int_\Omega f_s^\varepsilon(x_s)\, ds. \tag{3.7}$$

Step 3. Asymptotic behavior of f_s^ε

We fix now $s \in \Omega$ and consider the Γ-limit on K of the functionals f_s^ε defined in (3.6).

Proposition 3.3. *Let s be a point in Ω such that the function a is L^1-approximately continuous at s. Then the functionals f_s^ε, extended to $+\infty$ on all functions $x \in K$ which are not of class H^2 on $(-r, r)$, Γ-converge on K to*

$$f_s(x) := \begin{cases} \dfrac{A_0}{2r} \#\big(S\dot x \cap (-r, r)\big) + a(s) \displaystyle\fint_{-r}^{r} x^2 \, dt & \text{if } x \in \mathscr{S}(-r, r) \\ +\infty & \text{otherwise,} \end{cases} \tag{3.8}$$

where $A_0 := 2 \int_{-1}^{1} \sqrt{W}$.

Proof. This proposition is an immediate consequence of the following theorem by L. Modica and S. Mortola (see [37], [38], [36]): for every bounded open set $\Omega \subset \mathbb{R}^n$ the functionals given by $\int_\Omega \varepsilon |\nabla y|^2 + \frac{1}{\varepsilon} W(y)$ for all $y \in H^1(\Omega)$, such that $|y| \le 1$ – and extended to $+\infty$ elsewhere – are equicoercive on $L^1(\Omega)$ and Γ-converge to the functional given by $A_0 \, \|Dy\|$ when y is a function of bounded variation on Ω which takes only the values ± 1 a.e., and $+\infty$ otherwise. We immediately deduce that the functionals

$$\fint_{-r}^{r} \varepsilon^{2/3} \ddot x^2 + \varepsilon^{-2/3} W(\dot x)$$

Γ-converge on on $W^{1,1}(-r, r)$ to the functional given by $\frac{A_0}{2r} \#\big(S\dot x \cap (-r, r)\big)$, if $x \in \mathscr{S}(-r, r)$, and by $+\infty$ otherwise.

The assumption that a is L^1-approximately continuous at s implies that the rescaled functions $a_s^\varepsilon(t) := a(s + \varepsilon^{1/3} t)$ converge in $L^1_{\mathrm{loc}}(\mathbb{R})$ to the constant value $a(s)$. Thus the functionals $\fint_{-r}^{r} a_s^\varepsilon x^2$ converge to $a(s) \fint_{-r}^{r} x^2$ continuously on $W^{1,1}(-, r, r)$.

Hence the functionals f_s^ε are equicoercive on $W^{1,1}(-r, r)$ and Γ-converge to f_s. Now it suffices to take into account that $W^{1,1}(-r, r)$ embeds continuously in K and apply Proposition 2.11(v). □

Step 4. The main Γ-convergence result

Using identity (3.7), we can view the rescaled functionals $\varepsilon^{-2/3} I^\varepsilon(v)$ as functionals on $\mathsf{YM}(\Omega, K)$. More precisely we set

$$F^\varepsilon(\nu) := \begin{cases} \int_\Omega \langle \nu_s, f_s^\varepsilon \rangle \, ds & \text{if } \nu \text{ is the elementary Young measure} \\ & \text{associated to } \mathsf{R}^\varepsilon v \text{ for some } v \in H^2_{\mathrm{per}}(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.9}$$

Hence $F^\varepsilon(\nu)$ is finite if and only if ν is the elementary Young measure associated with the ε-blowup $\mathsf{R}^\varepsilon v$ of some $v \in H^2(\Omega)$, and (cf. (3.7))

$$F^\varepsilon(\nu) = \varepsilon^{-2/3} I^\varepsilon(v). \tag{3.10}$$

Propositions 3.1 and 3.3 clearly suggest the Γ-limit of F^ε, and indeed we have:

THEOREM 3.4. *The functionals F^ε in (3.9) Γ-converge on $\mathsf{YM}(\Omega, K)$ to*

$$F(\nu) := \begin{cases} \int_\Omega \langle \nu_s, f_s \rangle \, ds & \text{if } \nu_s \in \mathscr{I}(K) \text{ for a.e. } s \in \Omega, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.11}$$

REMARK. If x belongs to $\mathscr{S}_{\mathrm{per}}(0,h)$, and ϵ_x is the associated elementary invariant measure, a simple computation yields (cf. (5.8))

$$\langle \epsilon_x, f_s \rangle = \frac{A_0}{h} \# \big(S\dot{x} \cap [0,h) \big) + a(s) \fint_0^h x^2 \, dt. \tag{3.12}$$

Hence the value of $\langle \mu, f_s \rangle$ does not depend on the the constant r which appears in (3.6) and (3.8) when μ is an elementary invariant measure, and the same conclusion holds by density for every invariant measure (cf. Corollary 5.11). Therefore also F does not depend on r.

Proof of Theorem 3.4

In view of future applications we will try to present a proof of Theorem 3.4 as much independent as possible of the particular example we have considered so far. In fact, one could be tempted to view Theorem 3.4 as a particular case of the following general result: if the functionals F^ε are defined as in (3.9) for some integrands f_s^ε which Γ-converge on K to f_s, then they Γ-converge on $\mathsf{YM}(\Omega, K)$ to the functional F defined in (3.11). Unfortunately the convergence of the integrands alone seems not sufficient to guarantee the convergence of the functionals (cf., for instance, Theorem 2.12(iv)).

REMARK 3.5 – ESSENTIAL INGREDIENTS OF THE PROOF. The proof of Theorem 3.4 below can be adapted to a large class of problems with few modifications (even though this is not always case, see Section 6). In order to make its structure clear we have gathered here all the relevant properties of f_s^ε and f_s. Indeed the whole

proof will be derived by these properties, with the only exception of estimates $(3.30-32)$, where we use more specific arguments based on the definition of f_s^ε. In what follows, $B(s,\rho)$ denotes the open ball of center s and radius ρ, that is, the open interval $(s-\rho, s+\rho)$.

(1) *Pointwise convergence of the integrands*: for a.e. $s \in \Omega$, $f_s^\varepsilon \xrightarrow{\Gamma} f_s$ on K. This condition is verified in Proposition 3.3, and is one of the basis upon which we propose Theorem 3.4, the other being Proposition 3.1).
(2) *Existence of a "nice" dense subset of $\mathscr{I}(K)$*: for a.e. $s \in \Omega$, every invariant measure $\mu \in \mathscr{I}(K)$ can be approximated in the weak-star topology of $\mathscr{P}(K)$ with elementary invariant measures ϵ_x associated with functions $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ for some $h > 0$, so that $\langle \epsilon_x, f_s \rangle$ approximates $\langle \mu, f_s \rangle$. Both Sections 4 and 5 are devoted to the approximation of invariant measures by elementary invariant measures, and in Corollary 5.11 we prove that condition (2) is verified by every f_s of the form (3.8).
(3) *Uniformity in s of f_s*: there exists a negligible set $N \subset \Omega$ such that, for every $h > 0$ and $x \in \mathscr{S}_{\mathrm{per}}(0,h)$, the function $s \mapsto \langle \epsilon_x, f_s \rangle$ is L^1-approximately upper semicontinuous at every point of $\Omega \setminus N$. More precisely, formula (3.12) shows that $s \mapsto \langle \epsilon_x, f_s \rangle$ is L^1-approximately continuous at every point where a is L^1-approximately continuous. We expect that this condition is easily verified in many cases (cf., however, the situation described in Section 6.3).
(4) *Uniformity in s of the Γ-convergence of f_s^ε*: for every $h > 0$, $x \in \mathscr{S}_{\mathrm{per}}(0,h)$, and a.e. $\bar{s} \in \Omega$ there exist functions $x^\varepsilon \in H^2_{\mathrm{per}}(0,h)$ which converge to x in K and satisfy

$$\limsup_{\varepsilon \to 0} \fint_{\substack{\tau \in [0,h] \\ s \in B(\bar{s},\rho)}} f_s^\varepsilon(T_\tau x^\varepsilon) \, d\tau \, ds \le \fint_{\substack{\tau \in [0,h] \\ s \in B(\bar{s},\rho)}} f_s(T_\tau x) \, d\tau \, ds + \eta(\rho), \tag{3.13}$$

where the error $\eta(\rho)$ tends to 0 as $\rho \to 0$. Moreover one can assume $|\dot{x}^\varepsilon| \le 1$.

PROPOSITION 3.6. *The integrands f_s^ε defined in* (3.6) *satisfy condition* (4) *above.*

PROOF. We prove a stronger assertion: for every $\bar{s} \in \Omega$ and $\rho > 0$, the functional given by the average on the left-hand side of (3.13) for all 1-Lipschitz functions x in $H^2_{\mathrm{per}}(0,h)$, and extended to $+\infty$ elsewhere, Γ-converges on $W^{1,1}_{\mathrm{per}}(0,h)$ to the functional equal to the average on the right-hand side of (3.13) for $x \in \mathscr{S}_{\mathrm{per}}(0,h)$, and to $+\infty$ elsewhere.

Hence, for every $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ we could find 1-Lipschitz functions x^ε which converge to x in $W^{1,1}_{\mathrm{per}}(0,h)$, and thus in K, and satisfy (3.13) with $\eta(\rho) \equiv 0$.

To prove the claim, we first notice that for every $x \in H^2_{\mathrm{per}}(0,h)$ the average on the left-hand side of (3.13) can be written as

$$\fint_0^h \left[\varepsilon^{2/3} \ddot{x}^2 + \varepsilon^{-2/3} W(\dot{x}) \right] + \fint_{\substack{\tau \in [0,h] \\ s \in B(\bar{s},\rho)}} \left[\fint_{-r}^r a_s^\varepsilon \, (T_\tau x)^2 \right] d\tau \, ds, \tag{3.14}$$

and for every function $x \in \mathscr{S}_{\rm per}(0,h)$ the integral on the right-hand side of (3.13) can be written as

$$\frac{A_0}{h}\,\#\big(S\dot x \cap [0,h)\big) + \fint_{\substack{\tau\in[0,h]\\ s\in B(\bar s,\rho)}} \Big[a(s)\fint_{-r}^{r} (T_\tau x)^2\Big]\, d\tau\, ds. \tag{3.15}$$

Now we proceed as in the proof of Proposition 3.3: the first integral in (3.14) Γ-converge on $W^{1,1}_{\rm per}(0,h)$ to the first integral in (3.15), while the second integral in (3.14) converge continuously on $W^{1,1}_{\rm per}(0,h)$ to the second integral in (3.15) for every $\bar s, \rho$. □

Remark. Given positive functions f^ε_s on a metric space X which Γ-converge to f_s for every parameter s, it may be not true that the average of the functions f^ε_s (with respect to a fixed probability distribution on the space of parameters s) Γ-converges to the average of f_s (consider for instance the situation described in Section 6.3: any non-trivial average of f_s there is identically equal to $+\infty$, because the functionals f_s have pairwise disjoint supports).

In particular, condition (1) above does not yield condition (4). In fact, condition (1) implies that for every $x \in K$, $\tau \in \mathbb{R}$, and a.e. $s \in \Omega$ there exists a sequence (x^ε), depending on x, s and τ, such that $x^\varepsilon \to x$ in K and $f^\varepsilon_s(T_\tau x^\varepsilon) \to f_s(T_\tau x)$, while in (4) we essentially require that such a sequence can be chosen independent of $\tau \in [0,h]$ and of s in a neighborhood of a given $\bar s$.

We now give the proof of Theorem 3.4, starting with the lower bound inequality, namely that $\liminf F^\varepsilon(\nu^\varepsilon) \geq F(\nu)$ whenever $\nu^\varepsilon \to \nu$ in $\mathsf{YM}(\Omega,K)$. We may assume that the left-hand side of this inequality is finite (otherwise there is nothing to prove), and, possibly passing to a subsequence, that the liminf is actually a limit. By the definition of F^ε, each ν^ε has to be the elementary Young measures associated to some ε-blowup, which in view of Proposition 3.1 implies that ν_s is an invariant measure for a.e. s. By the definition of F we are left to show that

$$\liminf_{\varepsilon\to 0} \int_\Omega \langle \nu^\varepsilon_s, f^\varepsilon_s\rangle\, ds \geq \int_\Omega \langle \nu_s, f_s\rangle\, ds. \tag{3.16}$$

Since $f^\varepsilon_s \xrightarrow{\Gamma} f_s$ on K for a.e. $s \in \Omega$ (condition (1) of Remark 3.5), then (3.16) follows from Theorem 2.12(iv). We remark that since we only use the lower bound part of the convergence result stated in Theorem 2.12, as remarked after that theorem we do no need to verify the equi-integrability of the envelope functions in (2.12).

While the proof of the lower bound inequality follows from a quite general and relatively simple convergence result for functionals on Young measures, the proof of the upper bound inequality is definitely more delicate. The first step is to find a set $\mathscr{D}$ of Young measures with relatively simple structure which is F-dense in $\mathsf{YM}(\Omega,K)$ (cf. Remark 2.10).

Definition 3.7. *Let $\mathscr{D}$ be the class of all Young measures $\nu \in \mathsf{YM}(\Omega,K)$ which satisfy the following condition: there exist countably many disjoint intervals which cover almost all of Ω, and on every such interval ν agrees a.e. with an elementary invariant measure ϵ_x, with $x \in \mathscr{S}_{\rm per}(0,h)$ and $h>0$ (depending on the interval).*

Lemma 3.8. *The set $\mathscr{D}$ is F-dense in $\mathsf{YM}(\Omega,K)$, that is, for every $\nu \in \mathsf{YM}(\Omega,K)$ such that $F(\nu)$ is finite there exist $\nu^k \in \mathscr{D}$ such that $\nu^k \to \nu$ in $\mathsf{YM}(\Omega,K)$, and $\limsup F(\nu^k) \leq F(\nu)$.*

Proof. We first recall that there exists a norm ϕ on the space of all measures $\mathscr{M}(K)$ which induces the weak-star topology on every bounded subset, and in particular on $\mathscr{P}(K)$ (cf. Proposition 4.8).

Take $\nu \in \mathsf{YM}(\Omega,K)$ such that $\int \langle \nu_s, f_s\rangle\, ds$ is finite, and fix $\eta>0$. By condition (2) of Remark 3.5, for a.e. $\bar s \in \Omega$ we can find $h(\bar s)>0$ and $x_{\bar s} \in \mathscr{S}_{\rm per}(0,h(\bar s))$ so that

$$\phi(\epsilon_{x_{\bar s}} - \nu_{\bar s}) \leq \eta \quad\text{and}\quad \langle \epsilon_{x_{\bar s}}, f_{\bar s}\rangle \leq \langle \nu_{\bar s}, f_{\bar s}\rangle + \eta. \tag{3.17}$$

For a.e. $\bar s \in \Omega$ we can also take $\rho(\bar s) > 0$ such that, for every $\rho \leq \rho(\bar s)$ there holds

$$\fint_{B(\bar s,\rho)} \phi(\nu_{\bar s} - \nu_s)\, ds \leq \eta \quad\text{and}\quad \langle \nu_{\bar s}, f_{\bar s}\rangle \leq \fint_{B(\bar s,\rho)} \langle \nu_s, f_s\rangle\, ds + \eta, \tag{3.18}$$

and (cf. condition (3) of Remark 3.5)

$$\fint_{B(\bar s,\rho)} \langle \epsilon_{x_{\bar s}}, f_s\rangle\, ds \leq \langle \epsilon_{x_{\bar s}}, f_{\bar s}\rangle + \eta. \tag{3.19}$$

Putting together (3.17 − 19) we get

$$\begin{aligned} &\fint_{B(\bar s,\rho)} \phi(\epsilon_{x_{\bar s}} - \nu_s)\, ds \leq 2\eta, \\ &\fint_{B(\bar s,\rho)} \langle \epsilon_{x_{\bar s}}, f_s\rangle\, ds \leq \fint_{B(\bar s,\rho)} \langle \nu_s, f_s\rangle\, ds + 3\eta. \end{aligned} \tag{3.20}$$

By the Besicovitch covering theorem (see [19], chap. 2), we can find countably many disjoint intervals $B_i = B(\bar s_i, \rho_i)$ with $\rho_i \leq \rho(\bar s_i)$ which cover almost all of Ω. For every i we set $x_i := x_{\bar s_i}$, $f_i := f_{\bar s_i}$, and finally we define $\nu^\eta \in \mathscr{D}$ by

$$\nu^\eta_s := \epsilon_{x_i} \quad \text{if } s \in B_i \text{ for some } i.$$

Then ν^η belongs to $\mathscr{D}$, and (3.20) yields

$$\int_\Omega \phi(\nu^\eta_s - \nu_s)\, ds \leq \sum_i \int_{B_i} \phi(\epsilon_{x_i} - \nu_s)\, ds \leq \sum_i 2\eta|B_i| = 2\eta|\Omega|, \tag{3.21}$$

and

$$\int_\Omega \langle \nu_s^\eta, f_s\rangle\, ds = \sum_i \int_{B_i} \langle \epsilon_{x_i}, f_s\rangle\, ds \le \sum_i \Big[\int_{B_i} \langle \nu_s, f_s\rangle\, ds + 3\eta |B_i| \Big]$$
$$= \int_\Omega \langle \nu_s, f_s\rangle\, ds + 3\eta|\Omega|. \tag{3.22}$$

Inequality (3.21) shows that $\phi(\nu_s - \nu_s^\eta)$ converge in measure to 0 as $\eta \to 0$, and then pointwise a.e. provided that we pass to a suitable subsequence. Hence ν_s^η weak-star converge to ν_s for a.e. $s \in \Omega$, and ν^η converge to ν in $\mathsf{YM}\,(\Omega, K)$ (cf. Remark 2.1). Inequality (3.22) yields $\limsup F(\nu^\eta) \le F(\nu)$, and the proof is complete. □

According to Remark 2.10, to prove the upper bound inequality for the functionals F^ε – thus completing the proof of Theorem 3.4 – it suffices to show that every $\nu \in \mathscr{D}$ can be approximated (in energy) by ε-blowups of functions on Ω. We first construct the approximating sequence for a constant Young measure ν, and then we show how to localize such a construction to adapt to a general Young measure in $\mathscr{D}$.

Let be given a bounded interval I, a function $x \in \mathscr{S}_{\rm per}(0,h)$ with $h > 0$, and a sequence of functions $x^\varepsilon \in H^2_{\rm per}(0,h)$ which converge to x in K and satisfy

$$\limsup_{\varepsilon\to 0} \fint_{\substack{\tau\in[0,h]\\ s\in I}} f_s^\varepsilon(T_\tau x^\varepsilon)\, d\tau\, ds \le \fint_{s\in I} \langle \epsilon_x, f_s\rangle\, ds + \eta. \tag{3.23}$$

For every $\varepsilon > 0$ we choose $\tau^\varepsilon \in [0,h]$ and we set

$$v^\varepsilon(s) := \varepsilon^{1/3} x^\varepsilon(\varepsilon^{-1/3}s - \tau^\varepsilon) \quad \text{for every } s \in \mathbb{R}. \tag{3.24}$$

Lemma 3.9. *The functions v^ε in (3.24) belong to $H^2_{\rm per}(0, h\varepsilon^{1/3})$, and the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$ generate on I the constant Young measure ϵ_x. Moreover the numbers τ^ε in (3.24) can be chosen so that*

$$\limsup_{\varepsilon\to 0} \fint_I f_s^\varepsilon(\mathsf{R}^\varepsilon v^\varepsilon)\, ds \le \fint_I \langle \epsilon_x, f_s\rangle\, ds + \eta. \tag{3.25}$$

Proof. Let ν be a Young measure on I generated by a subsequence of $\mathsf{R}^\varepsilon v^\varepsilon$. For every $s \in \mathbb{R}$ we have (cf. (3.2))

$$\mathsf{R}_s^\varepsilon v^\varepsilon = T_{(\tau^\varepsilon - \varepsilon^{-1/3}s)} x^\varepsilon. \tag{3.26}$$

Since x^ε tends to x in K, $\mathsf{R}_s^\varepsilon v^\varepsilon$ tends to the orbit $\mathscr{O}(x)$ for every $s \in \Omega$, and then $\bar\nu_s$ is supported on $\mathscr{O}(x)$ for a.e. s. Thus $\bar\nu_s = \epsilon_x$ because $\bar\nu_s$ is invariant (Proposition 3.1), and the only invariant probability measure supported on $\mathscr{O}(x)$ is ϵ_x.

Let us consider the second part of the assertion. By identity (3.26) we get

$$\int_I f_s^\varepsilon(\mathsf{R}^\varepsilon v^\varepsilon)\, ds = \int_I f_s^\varepsilon(T_{(\tau^\varepsilon - \varepsilon^{-1/3}s)} x^\varepsilon)\, ds.$$

Now we choose τ^ε so that the integral on the right-hand side is not larger than the average of $\int_I f_s^\varepsilon(T_{(\tau - \varepsilon^{-1/3}s)} x^\varepsilon)\, ds$ over all $\tau \in [0,h]$, and taking into account that x^ε is h-periodic we get

$$\int_I f_s^\varepsilon(\mathsf{R}^\varepsilon v^\varepsilon)\, ds \le \fint_0^h \Big[\int_I f_s^\varepsilon(T_{(\tau-\varepsilon^{-1/3}s)} x^\varepsilon)\, ds \Big] d\tau = \fint_0^h \Big[\int_I f_s^\varepsilon(T_\tau x^\varepsilon)\, ds \Big] d\tau.$$

Finally we pass to the limit as $\varepsilon \to 0$ and apply inequality (3.23). □

We have thus shown that the ε-blowups of the functions v^ε defined in (3.24) converge in energy to the constant Young measure $\nu_s = \epsilon_x$, provided that the functions x^ε fulfill (3.23). Using condition (4) in Remark 3.5 we can show that such approximating sequence exist "locally" for every $\nu \in \mathscr{D}$.

Lemma 3.10. *Let be given $\nu \in \mathscr{D}$ and $\eta > 0$. Then there exist finitely many intervals I_i with pairwise disjoint closures which cover Ω up to an exceptional set with measure less than η, so that the following statements hold for every i:*

(i) *there exist $h_i > 0$ and $x_i \in \mathscr{S}_{\rm per}(0,h_i)$ such that $\nu_s = \epsilon_{x_i}$ for a.e. $s \in I_i$;*

(ii) *for every $\varepsilon > 0$ there exist 1-Lipschitz function $x_i^\varepsilon \in H^2_{\rm per}(0,h_i)$ which converge to x_i in K and satisfy (3.23) (with I, h, x, x^ε replaced by $I_i, h_i, x_i, x_i^\varepsilon$).*

Proof. Since ν belongs to $\mathscr{D}$, for almost every point $\bar s \in \Omega$ we can find a function $x \in \mathscr{S}_{\rm per}(0,h)$ with $h > 0$ and an interval I of the form $I = B(\bar s, \rho) \subset \Omega$ so that $\nu_s = \epsilon_x$ for a.e. $s \in I$. Moreover, by the uniformity assumption (4) in Remark 3.5, for almost every such $\bar s \in \Omega$ and for ρ sufficiently small we can find functions $x^\varepsilon \in H^2_{\rm per}(0,h)$ which converge to x in K and satisfy inequality (3.13) or, equivalently, (3.23). (Notice that the right-hand sides of (3.13) and (3.23) agree because $I = B(\bar s, \rho)$ and $\langle \epsilon_x, f_s\rangle$ is the average of $f_s(T_\tau x)$ over all $\tau \in [0,h]$).

We apply now Besicovitch covering theorem to find finitely many intervals of the type above whose closures are pairwise disjoint and cover Ω up to an exceptional set with measure less than η. □

We can now complete the proof of the upper bound inequality.

Since $\mathscr{D}$ is F-dense in $\mathsf{YM}\,(\Omega, K)$, by Remark 2.10 it suffices to construct, for every $\delta > 0$ and $\nu \in \mathscr{D}$, functions $v^\varepsilon \in H^2_{\rm per}(\Omega)$ so that the elementary Young measures ν^ε associated with the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$ satisfy (cf. (2.8))

$$\begin{aligned} &\limsup_{\varepsilon\to 0} \Phi(\nu^\varepsilon - \nu) \le \delta, \\ &\limsup_{\varepsilon\to 0} \int_\Omega f_s(\mathsf{R}_s^\varepsilon v^\varepsilon)\, ds \le \int_\Omega \langle \nu_s, f_s\rangle\, ds + \delta, \end{aligned} \tag{3.27}$$

where Φ is the norm which metrizes $\mathsf{YM}(\Omega, K)$ defined in (2.2).

We fix $\nu \in \mathscr{D}$, $\delta > 0$ and $\eta > 0$ (which will be later chosen in order to get (3.27)). We take I_i, x_i, h_i and x_i^ε as in Lemma 3.10, and define v_i^ε as in (3.24), namely

$$v_i^\varepsilon(s) := \varepsilon^{1/3} x_i^\varepsilon(\varepsilon^{-1/3} s - \tau_i^\varepsilon) \quad \text{for every } s \in \mathbb{R},$$

where τ_i^ε are chosen as in Lemma 3.9. We denote the intervals I_i by (a_i, b_i), ordered so that $a_i < b_i < a_{i+1} < b_{i+1}$, and set

$$v^\varepsilon(s) := v_i^\varepsilon(s) \quad \text{if } s \in (a_i + r\varepsilon^{1/3}, b_i - r\varepsilon^{1/3}) \text{ for some } i, \tag{3.28}$$

where r is the constant which appears in the definition of f_s (see (3.6)). It remains to extend the function v^ε out of the union of the intervals $(a_i + r\varepsilon^{1/3}, b_i - r\varepsilon^{1/3})$.

Take a positive number M (larger than 1 and r) such that $|x_i(t)| + 1 \le M$ for every i and every $t \in \mathbb{R}$. Since the functions x_i^ε converge to x_i in K and are 1-Lipschitz, then they also converge uniformly; in particular, for ε sufficiently small, $|x_i^\varepsilon(t)| \le M$ for every i, t, and thus $|v^\varepsilon(s)| \le M\varepsilon^{1/3}$ for every s where it is defined. Notice that M depends on the choice of x_i, and ultimately on η; therefore the dependence on M cannot be neglected in the estimates below.

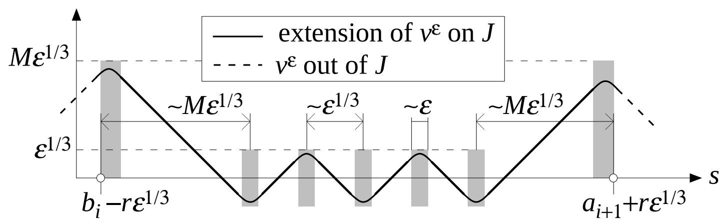

FIGURE 3.1. Construction of v^ε in $J := [b_i - r\varepsilon^{1/3}, a_{i+1} + r\varepsilon^{1/3}]$

For ε sufficiently small, we extend v^ε to the interval $[b_i - r\varepsilon^{1/3}, a_{i+1} + r\varepsilon^{1/3}]$ as shown in Figure 3.1. More precisely, $\dot v^\varepsilon$ takes alternately the values $+1$ and -1 in a sequence of intervals with length of order $\varepsilon^{1/3}$ (except the first and the last one, which have length of order $M\varepsilon^{1/3}$); two consecutive intervals are separated by a transition layer (marked in grey in the figure above) with length of order ε where $\ddot v^\varepsilon$ is of order ε^{-1}. The value of v^ε is of order $\varepsilon^{1/3}$ in each interval except the first and the last one where it is of order $M\varepsilon^{1/3}$.

Let us prove the first inequality in (3.27).

Let $\bar\nu$ be a Young measure generated by any subsequence (not relabeled) of the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$. Since v^ε and v_i^ε agree on $(a_i + r\varepsilon^{1/3}, b_i - r\varepsilon^{1/3})$ (cf. (3.28)), given a point $s \in I_i = (a_i . b_i)$, the ε-blowups $\mathsf{R}_s^\varepsilon v_\varepsilon$ and $\mathsf{R}_s^\varepsilon v_i^\varepsilon$ agree on the larger and larger intervals

$$\big(-(s - a_i)\varepsilon^{-1/3} + r, (b_i - s)\varepsilon^{-1/3} - r \big),$$

and therefore their distance in K vanishes as $\varepsilon \to 0$ (Proposition 5.1).

Hence $\mathsf{R}^\varepsilon v_\varepsilon$ and $\mathsf{R}^\varepsilon v_i^\varepsilon$ generate on I_i the same Young measure (see Remark 2.6), that is, $\bar\nu_s = \epsilon_{x_i}$ for a.e. $s \in I_i$ (see Lemma 3.9). On the other hand $\epsilon_{x_i} = \nu_s$ for a.e. $s \in I_i$ by construction (cf. Lemma 3.10), and then ν and $\bar\nu$ agree on the union of the intervals I_i; taking into account that the complement in Ω of this union has measure less than η, by the definition of Φ in (2.2) we get $\Phi(\bar\nu - \nu) \le \eta$. This gives the first inequality in (3.27), provided we choose η smaller than δ.

Let us consider now the second inequality in (3.27). If s belongs to the interval $(a_i + 2r\varepsilon^{1/3}, b_i - 2r\varepsilon^{1/3})$ for some i, the function $\mathsf{R}_s^\varepsilon v^\varepsilon$ agrees with $\mathsf{R}_s^\varepsilon v_i^\varepsilon$ on the interval $(-r, r)$, and then (cf. (3.28), (3.6))

$$f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon) = f_s^\varepsilon\big(\mathsf{R}_s^\varepsilon v_i^\varepsilon\big). \tag{3.29}$$

If s belongs to $(b_i + M\varepsilon^{1/3}, a_{i+1} - M\varepsilon^{1/3})$, $\mathsf{R}_s^\varepsilon v^\varepsilon$ agrees on $(-r, r)$ with the ε-blowup of the extension described in figure 2, and then it is of order 1, while its derivative is always $+1$ or -1 apart a number – not exceeding $2r + 1$ – of transition layers with size $\varepsilon^{2/3}$, where the second derivative is of order $\varepsilon^{-2/3}$. A direct computation gives the estimate

$$f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon) = O(1). \tag{3.30}$$

If s belongs to $(a_i - M\varepsilon^{1/3}, a_i)$ or $(b_i, b_i + M\varepsilon^{1/3})$, then $\mathsf{R}_s^\varepsilon v^\varepsilon$ agrees on $(-r, r)$ with the ε-blowup of the extension described in figure 2, but it is now of order M, and reasoning as before we get

$$f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon) = O(M^2). \tag{3.31}$$

Finally, if s belongs to $(a_i, a_i + 2r\varepsilon^{1/3})$ or $(b_i - 2r\varepsilon^{1/3}, b_i)$, then $\mathsf{R}_s^\varepsilon v^\varepsilon$ agrees partly with the ε-blowup of v_i^ε and partly with the ε-blowup of the extension described in Figure 3.1. By coupling estimates (3.29) and (3.30), we get

$$f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon) \le f_s^\varepsilon\big(\mathsf{R}_s^\varepsilon v_i^\varepsilon\big) + O(M^2). \tag{3.32}$$

Now we put together $(3.29 - 32)$, and since the measure of the complement of the union of all I_i is less than η, we obtain

$$\int_\Omega f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon)\, ds \le \sum_i \int_{I_i} f_s^\varepsilon\big(\mathsf{R}_s^\varepsilon v_i^\varepsilon\big)\, ds + O(1)\cdot\eta + O(M^3)\cdot\varepsilon^{1/3}.$$

Passing to the limit as $\varepsilon \to 0$, and recalling inequality (3.25), we get

$$\begin{aligned}\limsup_{\varepsilon\to 0} \int_\Omega f_s^\varepsilon(\mathsf{R}_s^\varepsilon v^\varepsilon)\, ds &\le \sum_i \Big[\int_{I_i} \langle \epsilon_{x_i}, f_s\rangle\, ds + \eta |I_i| \Big] + O(1)\cdot\eta \\ &\le \int_\Omega \langle \nu_s, f_s\rangle\, ds + O(1)\cdot\eta,\end{aligned}$$

which gives the second inequality in (3.27) if we choose η small enough. □

Step 5. Minimizers of F

An immediate consequence of Theorem 3.4 is the following:

COROLLARY 3.11. *For every $\varepsilon > 0$, let v^ε be a minimizer of I^ε on $H^2_{\mathrm{per}}(\Omega)$, and let ν be a Young measure in $\mathsf{YM}(\Omega, K)$ generated by a subsequence of the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$. Then ν minimizes the functional F in (3.11), which means that for a.e. $s \in \Omega$ the measure ν_s minimizes $\langle \mu, f_s \rangle$ among all invariant probability measures μ on K.*

PROOF. Apply Proposition 2.11(vi) and Theorem 3.4, taking into account (3.10) and (3.11). □

Now we want to show that every Young measure generated by the ε-blowups of the minimizers of I^ε is uniquely determined by the minimality property established in the previous corollary. For every $h > 0$, let y_h be the h-periodic function on $\mathbb{R}$ given by

$$y_h(t) := |t| - h/4 \quad \text{for } t \in (-h/2, h/2] \tag{3.33}$$

(cf. Figure 1.2). We have the following.

THEOREM 3.12. *Fix $s \in \Omega$ and let f_s be given in (3.8). If $\bar{\mu}$ minimizes $\langle \mu, f_s \rangle$ among all $\mu \in \mathscr{I}(K)$, then $\bar{\mu}$ is the elementary invariant measure associated with the function $y_{h(s)}$ where*

$$h(s) := L_0 (a(s))^{-1/3}, \tag{3.34}$$

and $L_0 := (48 A_0)^{1/3} = (96 \int_{-1}^1 \sqrt{W})^{1/3}$.

Taking Corollary 3.11 into account, we immediately deduce the following, which concludes our analysis of the asymptotic behavior of the minimizers of I^ε.

COROLLARY 3.13. *For every $\varepsilon > 0$, let v^ε be a minimizer of I^ε on $H^2_{\mathrm{per}}(\Omega)$. Then the ε-blowups $\mathsf{R}^\varepsilon v^\varepsilon$ generate a unique Young measure $\nu \in \mathsf{YM}(\Omega, K)$, and, for a.e. $s \in \Omega$, ν_s is the elementary invariant measure associated with the sawtooth function $y_{h(s)}$.*

Proof of Theorem 3.12

Throughout this subsection $s \in \Omega$ is fixed, and for simplicity we write $\bar{h}$, $\bar{y}$ instead of $h(s)$, $y_{h(s)}$. We begin with a computation which shows the optimality of $\bar{y}$.

LEMMA 3.14. *The measure $\epsilon_{\bar{y}}$ minimizes $\langle \mu, f_s \rangle$ among all invariant measures μ.*

PROOF. Fix $x \in \mathscr{S}_{\mathrm{per}}(0, h)$ with $h > 0$. Up to a suitable translation, we may assume that $S\dot{x} \cap [0, h]$ consists of the points $t_0 = 0 < t_1 < t_2 < \ldots < t_n = h$ and $n = \#\big(S\dot{x} \cap [0, h)\big)$ is an even number. For every $i = 1, \ldots, n$, let I_i be the interval (t_{i-1}, t_i), $h_i := |I_i| = t_i - t_{i-1}$, and p_i be the average of x on I_i. Thus, recalling (3.12) and taking into account that $\dot{x}$ is constant ± 1 on each I_i, we get

$$\begin{aligned}
\langle \epsilon_x, f_s \rangle = \frac{A_0}{h} n + a(s) \fint_0^h x^2 \, dt &= \sum_{i=1}^n \frac{1}{h} \Big[A_0 + a(s) \int_{I_i} x^2 \, dt \Big] \\
&= \sum_{i=1}^n \frac{h_i}{h} \Big[\frac{A_0}{h_i} + a(s) \fint_{-h_i/2}^{h_i/2} (t + p_i)^2 dt \Big] \\
&= \sum_{i=1}^n \frac{h_i}{h} \Big[\frac{A_0}{h_i} + \frac{a(s)}{12} h_i^2 + a(s)\, p_i^2 \Big].
\end{aligned}$$

We rewrite the last identity as

$$\langle \epsilon_x, f_s \rangle = \sum_{i=1}^n \frac{h_i}{h} \, g(h_i, p_i), \tag{3.35}$$

where we have set

$$g(h, p) := \frac{A_0}{h} + \frac{a(s)}{12} h^2 + a(s)\, p^2. \tag{3.36}$$

A simple computation shows that $(\bar{h}/2, 0)$ is the unique minimum point of g. Furthermore, for $x := \bar{y}$ we have $n = 2$, $h_1 = h_2 = \bar{h}/2$, $p_1 = p_2 = 0$, and (3.35) becomes

$$\langle \epsilon_{\bar{y}}, f_s \rangle = g(\bar{h}/2, 0) = \min g. \tag{3.37}$$

On the other hand, (3.35) and the fact that $\sum h_i/h = 1$ yield, for a general x,

$$\langle \epsilon_x, f_s \rangle = \sum_{i=1}^n \frac{h_i}{h} g(h_i, p_i) \geq \min g = \langle \epsilon_{\bar{y}}, f_s \rangle.$$

We have thus proved that $\epsilon_{\bar{y}}$ minimizes $\langle \mu, f_s \rangle$ among all elementary invariant measures μ. We conclude by a density argument based on Corollary 5.11. □

A careful examination of the previous proof leads to the conclusion that no other *elementary* invariant measure minimizes $\langle \mu, f_s \rangle$ among all $\mu \in \mathscr{I}(K)$. But proving Theorem 3.12 means showing that no other invariant measure minimizes $\langle \mu, f_s \rangle$, and this requires a more refined argument.

Since we know that every invariant measure can be approximated by elementary invariant measures, we first look for general criteria which ensure that a sequence of elementary invariant measures converges to a given elementary invariant measure.

LEMMA 3.15. *Let $\tilde{x} \in \mathscr{S}_{\mathrm{per}}(0, \tilde{h})$ be given with $\tilde{h} > 0$, and, for $k = 1, 2, \ldots$, $x^k \in \mathscr{S}_{\mathrm{per}}(0, h^k)$ with $h^k > 0$. Then the elementary invariant measures ϵ_{x^k} weak-star converge to $\epsilon_{\tilde{x}}$ if (and only if) the probability that $\tau \in (0, h^k)$ satisfies*

$d\big(T_\tau x^k, \mathscr{O}(\tilde x)\big) > \varepsilon$ vanishes as $k \to +\infty$ for every $\varepsilon > 0$ (here d is the distance in K and $\mathscr{O}(\tilde x)$ is the orbit of $\tilde x$).

Proof. Let μ be an invariant measure on K. Since $\epsilon_{\tilde x}$ is the only invariant measure supported on the orbit of $\tilde x$, μ is equal to $\epsilon_{\tilde x}$ if (and only if) μ is supported on the compact set $\mathscr{O}(\tilde x)$, that is to say, $\mu(A_\varepsilon) = 0$ for every $\varepsilon > 0$, where A_ε is the open set of all $x \in K$ such that $d\big(x, \mathscr{O}(\tilde x)\big) > \varepsilon$.

Now, if μ is the limit of (a subsequence of) the measures ϵ_{x^k}, which in turn are the averages of the Dirac masses centered at $T_\tau x^k$ over all $\tau \in (0, h^k)$ (see (4.6)), then

$$\mu(A_\varepsilon) \le \liminf_{k\to\infty} \epsilon_{x^k}(A_\varepsilon) \le \liminf_{k\to\infty} \frac{1}{h^k}\big|\{\tau \in (0,h^k) : T_\tau x^k \in A_\varepsilon\}\big|.$$

Since the last term in the previous line vanishes by assumption, it follows that $\mu = \epsilon_{\tilde x}$, and the assertion is proved (the converse is immediate). □

The criterion in the previous lemma can be consistently improved when $\tilde x$ is of the form $\tilde x = y_{\tilde h}$ for some $\tilde h > 0$ (cf. (3.33)). For every k, we define n^k, I_i^k, h_i^k, p_i^k as in the proof of Lemma 3.14 (replacing x and h by x^k and h^k), and consider the probability measures λ^k on $(0,+\infty) \times \mathbb{R}$ given by

$$\lambda^k := \sum_i \frac{h_i^k}{h^k}\, \delta_{(h_i^k, p_i^k)}. \tag{3.38}$$

We would expect that ϵ_{x^k} converge to $\epsilon_{\tilde x}$ if the numbers h_i^k and p_i^k converge, in the limit $k \to +\infty$, to $\tilde h/2$ and 0, respectively, at least "in probability". That is, if the measures λ^k converge to the Dirac mass centered at $(\tilde h/2, 0)$. In fact, we need a slightly stronger requirement:

Lemma 3.16. *Assume that there exists $\tilde h$ such that*

$$\lim_{k\to\infty} \int_C \Big(1 + \frac{1}{h}\Big)\, d\lambda^k(h,p) = 0 \tag{3.39}$$

for every closed set $C \subset (0,+\infty) \times \mathbb{R}$ which does not contain the point $(\tilde h/2, 0)$. Then ϵ_{x^k} weak-star converge to $\epsilon_{\tilde x}$ with $\tilde x := y_{\tilde h}$.

Proof. In view of Lemma 3.15, it suffices to show that for every $\varepsilon > 0$ the probability that $\tau \in (0, h^k)$ satisfies $d\big(T_\tau x^k, \mathscr{O}(\tilde x)\big) > \varepsilon$ vanishes as $k \to +\infty$.

Let $\varepsilon > 0$ be fixed. We can assume with no loss in generality that $h^k \to +\infty$ as $k \to +\infty$ (if x is h-periodic then it is also nh-periodic for every positive integer n). We also use the fact that, since the distance on K is the one in (5.1) for $n = 1$, by Remark 5.2 there exists m such that

$$d(x_1, x_2) \le \varepsilon/2 + \|x_1 - x_2\|_{L^\infty(-m,m)} \quad \text{for } x_1, x_2 \in K. \tag{3.40}$$

The proof is now divided in two steps.

Step 1.

Consider $\delta > 0$ and $\tau \in (0, h^k)$ such that
(a) τ belongs to $(m, h^k - m)$;
(b) for every index i such that I_i^k and $(\tau - m, \tau + m)$ intersect, there holds $|h_i^k - \tilde h/2| \le \delta$;
(c) there exists an index j such that I_j^k and $(\tau - m, \tau + m)$ intersect, and $|p_j^k| \le \delta$.

We claim that for a suitable choice of the parameter δ (depending on ε, but not on τ and k), there holds

$$d\big(T_{-\tau} x^k, \mathscr{O}(\tilde x)\big) \le \varepsilon. \tag{3.41}$$

More precisely, in case that x^k has slope -1 in I_j^k, we prove that x^k is close to $T_{t_j}\tilde x$ (the case when x^k has slope $+1$ in I_j^k can be treated in a similar way). We set $\bar x := T_{t_j}\tilde x$, and notice that $x^k(t_j) = p_j^k - h_j^k/2$ and $\bar x(t_j) = \tilde x(0) = -\tilde h/4$; by assumptions (b) and (c) we infer

$$\big|x^k(t_j) - \bar x(t_j)\big| \le |p_j^k| + \frac{1}{2}|h_j^k - \tilde h/2| \le 2\delta. \tag{3.42}$$

We label the points of $S\dot{\bar x}$ as $\bar t_i$, so that $\bar t_{i-1} < \bar t_i$ for every i and $\bar t_j = t_j$ (t_j belongs to $S\dot{\bar x}$ because 0 belongs to $S\dot{\tilde x}$), and we let $\bar I_i$ denote the interval $(\bar t_{i-1}, \bar t_i)$.

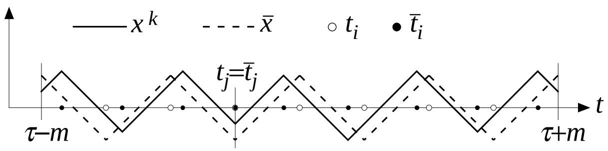

Figure 3.2. The functions x^k and $\bar x := T_{t_j}\tilde x$ in $(\tau - m, \tau + m)$

Thus x^k and $\bar x$ have the same derivative in $\bar I_i \cap I_i^k$ for every i (cf. Figure 3.2); since $t_j = \bar t_j$ by construction, assumption (b) implies that the measure of $\bar I_i \setminus I_i^k$ is less than δ when $i = j, j+1$, less than 2δ when $i = j-1, j+2$, less than 3δ when $i = j-2, j+3$, and so on.

Taking into account that the total number of indexes i such that $\bar I_i$ and $(\tau - m, \tau + m)$ intersect does not exceed $N := 1 + 4m/\tilde h$, we obtain that $|\bar I_i \setminus I_i^k| \le N\delta$ for all such i, and then the derivatives of x^k and $\bar x$ agree in $(\tau - m, \tau + m)$ minus a set with measure less than $N^2\delta$. Using (3.42) we deduce that for every $t \in (\tau - m, \tau + m)$

$$\big|x^k(t) - \bar x(t)\big| \le \big|x^k(t_j) - \bar x(t_j)\big| + \int_{t_j}^t \big|\dot x^k - \dot{\bar x}\big| \le 2(1 + N^2)\,\delta.$$

Therefore, if we choose δ so that $2(1+N^2)\,\delta \le \varepsilon/2$, by (3.40) we get

$$d\big(T_{-\tau}x^k, T_{t_j-\tau}\tilde{x}\big) = d\big(T_{-\tau}x^k, T_{-\tau}\bar{x}\big) \le \varepsilon/2 + \|x^k - \bar{x}\|_{L^\infty(\tau-m,\tau+m)} \le \varepsilon,$$

which implies (3.41).

Step 2.

We show that the probability that $\tau \in (0, h^k)$ does not satisfy assumption (a), (b) or (c) above vanishes as $k \to +\infty$.

The probability that (a) fails amounts to $2m/h^k$, which vanishes as $k \to +\infty$ because $h^k \to +\infty$.

The points $\tau \in (m, h^k - m)$ which do not satisfy (b) belong to the union of all interval $(t^k_{i-1} - m, t^k_i + m)$ over all indexes i such that $|h^k_i - \tilde{h}/2| \ge \delta$; therefore they occur with probability not exceeding

$$\sum_{\substack{i \text{ such that}\\ |h^k_i - \tilde{h}/2| \ge \delta}} \frac{h^k_i + 2m}{h^k} = \int_C \Big(1 + \frac{2m}{h}\Big)\, d\lambda^k(h,p),$$

where the measures λ^k are defined in (3.38) and C is the set of all (h,p) such that $|h - \tilde{h}/2| \ge \delta$. The integral on the right-hand side vanishes as $k \to +\infty$ by assumption (3.39).

The points $\tau \in (0, h^k)$ which do not satisfy (c) certainly belong to the union of the intervals (t_{i-1}, t_i) over all indexes i such that $|p^k_i| \ge \delta$; therefore they occur with probability not exceeding

$$\sum_{\substack{i \text{ such that}\\ |p^k_i| \ge \delta}} \frac{h^k_i}{h^k} = \int_C d\lambda^k(h,p),$$

where C is the set of all (h,p) such that $|p| \ge \delta$, and again the integral on the right-hand side vanishes as $k \to +\infty$ by (3.39). □

We can now conclude the proof of Theorem 3.12.

Let $\bar{\mu}$ minimize $\langle \mu, f_s \rangle$ among all $\mu \in \mathscr{I}(K)$. By Lemma 3.14 and equality (3.37) we deduce that $\langle \bar{\mu}, f_s \rangle = \langle \epsilon_{\bar{y}}, f_s \rangle = \min g$, with g given in (3.36). By applying Corollary 5.11 we find elementary invariant measures ϵ_{x^k}, with $x^k \in \mathscr{S}_{\rm per}(0, h^k)$ for some $h^k > 0$, which converge weak-star to $\bar{\mu}$ and satisfy

$$\langle \epsilon_{x^k}, f_s \rangle \to \langle \bar{\mu}, f_s \rangle = \min g. \tag{3.43}$$

Hence, to prove the assertion of Theorem 3.12, namely that $\bar{\mu} = \epsilon_{\bar{y}}$, it suffices to show that assumption (3.39) of Lemma 3.16 is verified when $\tilde{h}$ is equal to $\bar{h}$.

Possibly passing to a subsequence, we may assume that the measures λ^k weak-star converge on $[0,+\infty]\times[-\infty,+\infty]$ to a probability measure λ. Since g, extended to $+\infty$ at the boundary of $(0,+\infty)\times\mathbb{R}$, is a positive lower semicontinuous function, (3.35) and (3.43) yield

$$\min g \le \langle \lambda, g \rangle \le \liminf_{k\to\infty} \langle \lambda^k, g \rangle = \lim_{k\to\infty} \langle \epsilon_{x^k}, f_s \rangle = \min g.$$

Hence $\langle \lambda, g \rangle = \min g$, which implies that λ is supported on the set of all minimum points of g, that is, λ is the Dirac mass centered at $(\bar{h}/2, 0)$.

Moreover $\langle \lambda^k, g \rangle \to \langle \lambda, g \rangle$, which implies that the measures $g\cdot\lambda^k$ converge weak star and in variation to $g\cdot\lambda$, which is supported at the point $(\bar{h}/2, 0)$. Therefore, for every closed set C which does not contain $(\bar{h}/2, 0)$ there holds

$$\lim_{k\to\infty} \int_C g\, d\lambda^k = 0.$$

This implies (3.39) because, up to a suitable multiplicative constant, the function $g(h,p)$ is larger than the function $1 + 1/h$.

4. Approximation of invariant measures on abstract spaces

In this section we will focus on the approximation properties of probability measures on a compact metric space K which are invariant under the action of a certain group G of transformations of K. In the applications we have in mind K is a space of functions on $\mathbb{R}^n$ and G is the group of translations (cf. Section 3); this specific case is discussed in detail in Section 5. Since the case of a non-commutative group G presents some additional difficulties which would make the exposition of the results more technical, we restrict our attention to the commutative case; the non-commutative case is briefly discussed at the end of this section.

We first fix some notation. Throughout this section (K,d) is a compact metric space, $\mathscr{M}(K)$ is the Banach space of finite real Borel measures on K and $\mathscr{P}(K)$ is the subset of all probability measures; we usually denote by the letter x a point of K, and by the letter μ a measure on K. If K' is a locally compact topological space, μ is a measure on K, and f is a Borel map from K to K', then the *push-forward* of μ on K' via f is the measure $f^\#\mu$ given by $(f^\#\mu)(B) := \mu(f^{-1}(B))$ for every Borel set $B \subset K'$.

A topological group G is also given, that is first countable and locally compact, and acts on K via the continuous left action $(T,x) \mapsto Tx$; every element of G is regarded as an homeomorphism of K onto itself, and is usually denoted by the capital letter T. Given a map g and a measure μ defined on K, gT and $T^\#\mu$ denote the composed function $g\circ T$ and the push-forward of μ according to T, respectively. Notice that $T^\#\delta_x = \delta_{Tx}$ for every $x \in K$, and $\int_K g\, d(T^\#\mu) = \int_K gT\, d\mu$ for every μ, g. A measure μ on K is called *invariant* if it is invariant under the action of

G, that is, if $T^{\#}\mu = \mu$ for every $T \in G$; $\mathscr{I}(K)$ denotes the class of all invariant probability measures on K.

If H is a subgroup of G, G/H is the left quotient of G, and $[T]$ is the equivalence class in G/H which contains T. If H is closed then G/H is a Hausdorff locally compact space; if in addition G/H is compact we say that H is co-compact. The *orbit* of a point $x \in K$ is the set $\mathscr{O}(x) := \{T(x) : T \in G\}$ (notice that G is not assumed to act transitively on K). The point x has a *period* T if $Tx = x$; the set of all periods of x is denoted by $P(x)$. Thus $P(x)$ is always a closed subgroup of G, and $P(x) = P(x')$ whenever x and x' belong to the same orbit. We distinguish some cases:

- when $P(x)$ is not co-compact we say that x is *non-periodic*;
- when $P(x)$ is co-compact we say that x is *periodic*;
- when $P(x)$ includes a co-compact subgroup H we say that x is *H-periodic*.

Notice that the map $[T] \mapsto Tx$ is continuous and one-to-one from $G/P(x)$ to $\mathscr{O}(x)$. If $P(x)$ is co-compact, then $\mathscr{O}(x)$ is compact and homeomorphic to $G/P(x)$.

We assume now that G is commutative. Thus the quotient $G/P(x)$ is also a group, and if in addition x is periodic, $G/P(x)$ is a compact group which acts continuously and transitively on the orbit of x. Therefore there exists a unique probability measure ϵ_x, called the *elementary invariant measure* associated to x, which is supported on $\mathscr{O}(x)$ and is invariant under the action of $G/P(x)$ (see for instance [45], theorem 5.14, or [22], sec. 2.7; cf. also Lemma 4.10 below, and the remarks on the non-commutative case at the end of this section). It may be easily verified that ϵ_x is also invariant under the action of G, and that $\epsilon_x = \epsilon_{x'}$ when x and x' belong to the same orbit.

The elementary invariant measures are the simplest invariant probability measures we can construct on K, and within the class of invariant probability measures, they play a rôle similar to Dirac masses within the class of all probability measures (cf. Remark 4.7). So the following question naturally arises.

Problem. *Under which hypotheses is it possible to approximate (in the weak-star topology of $\mathscr{M}(K)$) every invariant probability measure by convex combinations of elementary invariant measures?*

When G is a compact group, such an approximation is easily obtained by exploiting the existence of a finite Haar measure on G (see Remark 4.6). When G is not compact we can obtain this approximation under some additional hypotheses on G and K, to state which we need some more definitions.

Let H be a co-compact subgroup of G and let $\pi : G \to G/H$ be the canonical projection of G onto G/H. Since G/H is a compact group, there exists a unique (left) Haar probability measure Φ on G/H, that is, a probability measure which is invariant under the left action of G/H on itself (see [45], theorem 5.14, or [22], sec. 2.7).

Definition 4.1. *Let H be a co-compact subgroup of G, and let $\Phi_{G/H}$ denote the unique Haar probability measure on G/H. We say that a Borel set $A \subset G$ is a representation of the quotient G/H if A is pre-compact in G and π is one-to-one from A to G/H. We denote by Φ_A the push-forward of the measure $\Phi_{G/H}$ onto A according to the inverse of π restricted to A.*

Notice that such an inverse is a Borel measurable map, and then Φ_A is well-defined; in fact $\pi^{\#}\Phi_A = \Phi_{G/H}$. In the following G/H and A are always endowed with the measures $\Phi_{G/H}$ and Φ_A given above. When no confusion may arise, we omit to write explicitly the measure Φ_A (resp. $\Phi_{G/H}$) in integrals on A (resp. on G/H).

The existence of a representation is guaranteed by the following result.

Proposition 4.2. *A representation A of G/H exists for every co-compact subgroup H.*

Proof. Since the topology of G is first countable, it can be metrized by a distance d_G which satisfies $d_G(T_1, T_2) = d_G(ST_1, ST_2)$ for every $T_1, T_2, S \in G$ (cf. [29], chap. 6, exercise O, or [10]). Thus G/H can be metrized by the quotient distance

$$d_{G/H}\big([T_1], [T_2]\big) := \inf\big\{d_G(ST_1, T_2) : S \in H\big\} \quad \text{for } [T_1], [T_2] \in G/H.$$

The first step is to construct a compact set $K \subset G$ such that $\pi(K) = G/H$.

Since G/H is compact, then it is totally bounded with respect to the quotient distance, and for every integer $k \geq 0$ we can find finitely many points y_i^k in G/H (the total number of which depends on k) so that the balls with radius $2^{-(k+2)}$ centered at these points cover G/H. We choose a representative T_i^k in every equivalence class y_i^k by the following inductive procedure: if $k = 0$, we just take T_i^0 in $\pi^{-1}(y_i^0)$; if $k > 0$, for every y_i^k there exists y_j^{k-1} such that $d_{G/H}(y_i^k, y_j^{k-1}) \leq 2^{-(k+1)}$, and by the definition of $d_{G/H}$ we can choose T_i^k in $\pi^{-1}(y_i^k)$ such that $d_G(T_i^k, T_j^{k-1}) \leq 2^{-k}$. According to this procedure, for every T_i^k and every integer $h < k$, there exists T_j^h such that

$$d_G(T_i^k, T_j^h) \leq 2^{-k} + 2^{-k+1} + \ldots + 2^{-(h+1)} \leq 2^{-h}.$$

Let K be the closure of the collection of T_i^k for all k, i. Thus K is closed and totally bounded (because for every $h > 0$ it is covered by the closed balls with radius 2^{-h} centered at the points $T_i^{\bar h}$ with $\bar h \leq h$, which are finitely many), and therefore compact. Hence $\pi(K)$ is compact too, and contains all points y_i^k, and since these are dense in G/H, $\pi(K) = G/H$.

Finally we consider the multifunction which takes every $y \in G/H$ into the non-empty closed set $\pi^{-1}(y) \cap K$. Since the graph of this multifunction is closed in $(G/H) \times K$, by theorem III.6 in [11] we can find a Borel selection, namely, a Borel map $\sigma : G/H \to K$ such that $\pi(\sigma(y)) = y$ for every $y \in G/H$. We conclude by taking A equal to the image of σ (which is Borel measurable because G/H is compact and σ is one-to-one, cf. [22], sec. 2.2.10). $\square$

Definition 4.3. *A set $X \subset K$ is called* uniformly approximable *if for every $\varepsilon > 0$ there exists a co-compact subgroup H and a representation A of G/H such that for every point $x \in X$ we may find an H-periodic point $\bar{x} \in K$ which satisfies*

$$\int_A d(Tx, T\bar{x})\, d\Phi_A(T) \leq \varepsilon. \tag{4.1}$$

Roughly speaking this definition means that we can approximate every point $x \in X$ by a periodic point $\bar{x}$ so that not only $\bar{x}$ is close to x, but also $T\bar{x}$ is close to Tx for "most" T. Moreover we ask that this approximation is in some sense uniform in x. Using the compactness of K it may be proved that the notion of uniform approximability depends only on the topology of K (and on the action of G) but not on the specific choice of the distance d.

We can now state the main result of this section.

Theorem 4.4. *If K is uniformly approximable in the sense of Definition 4.3, then every invariant probability measure μ on K can be approximated (in the weak-star topology of $\mathscr{M}(K)$) by a sequence of convex combinations μ_k of elementary invariant measures. More precisely, each μ_k can be taken the form $\sum_i \sigma_i \epsilon_{\bar{x}_i}$ where all points $\bar{x}_i$ are H-periodic for some co-compact group H that depends only on k.*

Comments and remarks on Theorem 4.4

We do not know if the uniform approximability assumption in Theorem 4.4 is necessary or not. In particular we do not know if it suffices to assume that periodic points are dense in K (which would already give a large class of elementary invariant measures).

Remark 4.5. When G is the additive group $\mathbb{R}^n$ and H is a subgroup of the form $(a\mathbb{Z})^n$ with $a > 0$, a representation of G/H is given by the cube $A := (0, a)^n$ endowed with Lebesgue measure $\mathscr{L}_n$ suitably renormalized. In particular K is uniformly approximable when the following condition holds: for every $\varepsilon > 0$ there exists $a > 0$ such that for every $x \in K$ we may find an $(a\mathbb{Z})^n$-periodic point $\bar{x}$ which satisfies

$$\fint_{T \in (0,a)^n} d(Tx, T\bar{x})\, d\mathscr{L}_n(T) \leq \varepsilon.$$

Remark 4.6. If G is compact it is always possible to approximate an invariant probability measure by convex combinations of elementary invariant measures. A simple direct proof of this fact can be obtained by considering a (left) Haar probability measure Φ on G. To every $\mu \in \mathscr{P}(K)$ we can associate an invariant probability measure $P\mu$ by taking the average of all $T^{\#}\mu$ with respect to the measure Φ, that is

$$\langle P\mu, g\rangle := \int_G \langle T^{\#}\mu, g\rangle\, d\Phi(T) = \Big\langle \mu, \int_G gT\, d\Phi(T)\Big\rangle \quad \forall\, g \in C(K)\,. \tag{4.2}$$

Thus P is a projection of $\mathscr{P}(K)$ onto $\mathscr{I}(K)$ that is continuous with respect to the weak-star topology, and takes every Dirac mass δ_x into the elementary invariant measure ϵ_x (recall that every point of K is periodic because G is compact). Let now μ be an invariant measure on K, and let μ_k be convex combinations of Dirac masses which converge to μ. Then the measures $P\mu_k$ are convex combinations of elementary invariant measures, and converge to $P\mu = \mu$.

Remark 4.7. The set $\mathscr{I}(K)$ of all invariant probability measures on K is weak-star compact and convex, thus it is natural to look for its extreme points: indeed every point in a compact convex subset C of a separable locally convex space (in our case, $\mathscr{M}(K)$ endowed with the weak-star topology) can be approximated by convex combinations of extreme points of C by the Krein-Millman theorem (cf. [45], theorem 3.21). It may be proved that μ is an extreme point of $\mathscr{I}(K)$ if and only if every Borel set invariant under the action of G has either full measure or zero measure (see [35], chap. II, proposition 2.5, when G is the group generated by one transformation). Clearly every elementary invariant measure ϵ_x is an extreme point of $\mathscr{I}(K)$, but in general the converse is not true, even if periodic points are dense in K (consider for instance the product $K := (\mathbb{R}/\mathbb{Z}) \times (\mathbb{N} \cup \{\infty\})$ and the group G generated by the transformation $T(x, k) := (x + a_k, k)$ where all a_k with finite k are rational numbers and converge to a_∞ irrational).

The situation simplifies when G is compact. In this case the quotient K/G is a compact metrizable space, and for every $\mu \in \mathscr{M}(K)$ we may define the push forward $\pi^{\#}\mu \in \mathscr{M}(K/G)$, where π is the canonical projection of K into K/G. Then $\pi^{\#}$ is a weak-star continuous operator which maps $\mathscr{I}(K)$ into $\mathscr{P}(K/G)$ bijectively, and takes elementary invariant measures into Dirac masses. Hence the extreme points of $\mathscr{I}(K)$ are the elementary invariant measures only. If G is not compact, K/G may be neither metric nor even Hausdorff, that is, the quotient topology may not separate points (cf. the remark after Proposition 5.3).

Proof of Theorem 4.4

It is convenient to introduce the following norm on $\mathscr{M}(K)$: we take a sequence (g_k) of Lipschitz functions which is dense in $C(K)$, we let $\alpha_k := \|g_k\|_\infty + \mathrm{Lip}(g_k)$, and set

$$\phi(\mu) := \sum_{k=1}^{\infty} \frac{\big|\langle \mu, g_k\rangle\big|}{2^k \alpha_k}. \tag{4.3}$$

It can be easily shown (cf. Proposition 4.8 below) that ϕ induces the weak-star topology on every bounded subset of $\mathscr{M}(K)$. For the rest of this section we only consider measures in the class $\mathscr{P}(K)$, that is always endowed with the weak-star topology of $\mathscr{M}(K)$. Therefore, in the following the notions "approximation" or "distance" always refer to ϕ.

Proposition 4.8. *The function ϕ given in* (4.3) *has the following properties:*

(i) *ϕ is a norm on $\mathscr{M}(K)$, and $\phi(\mu) \leq \|\mu\|$ for every μ;*

(ii) *ϕ induces on every bounded subset of $\mathscr{M}(K)$ the weak-star topology;*

(iii) *for every $x, y \in K$ one has $\phi\big(\delta_x - \delta_y\big) \le d(x,y)$.*

Proof. The function ϕ is clearly a norm, and for every $\mu \in \mathscr{M}(K)$ there holds

$$\phi(\mu) = \sum_1^\infty \frac{\big|\langle \mu, g_k\rangle\big|}{2^k \alpha_k} \le \sum_1^\infty \frac{\|\mu\| \cdot \|g_k\|_\infty}{2^k \alpha_k} \le \sum_1^\infty \frac{\|\mu\|}{2^k} \le \|\mu\|.$$

Regarding statement (ii), it may be easily verified that $\phi(\mu^i - \mu) \to 0$ if and only if $\langle \mu^i, g_k\rangle$ converge to $\langle \mu, g_k\rangle$ for every k. Since the functions g_k are dense in $C(K)$, and the sequence (μ^i) is bounded, this implies weak-star convergence.

We finally prove (iii):

$$\begin{aligned}\phi(\delta_x - \delta_y) &= \sum_1^\infty \frac{\big|g_k(x) - g_k(y)\big|}{2^k \alpha_k} \\ &\le \sum_1^\infty \frac{\operatorname{Lip}(g_k) \cdot d(x,y)}{2^k \alpha_k} \le \sum_1^\infty \frac{d(x,y)}{2^k} \le d(x,y). \qquad \square\end{aligned}$$

The idea of the proof of Theorem 4.4 is roughly the following. We first define the notion of average for a family of measures, and show that for an H-periodic point x the average of δ_{Tx} over all T in a representation A of the quotient G/H is the elementary invariant measure ϵ_x. Then we notice that the operator P which associates to every $\mu \in \mathscr{P}(K)$ the average of the translated measures $T^\# \mu$ over all $T \in A$ is continuous. Finally we approximate an invariant probability measure μ by convex combinations μ_k of Dirac masses at H-periodic points, and then apply the averaging operator P: the measures $P\mu_k$ are then convex combination of elementary invariant measures, and approximate $P\mu$, which agrees with μ because μ is invariant.

Definition 4.9. *Let B be a bounded Borel set of a locally compact space and let λ be a probability measure supported on B. Let $\{\mu_t : t \in B\}$ be a family of measures in $\mathscr{P}(K)$ parametrized by $t \in B$ and assume that this parametrization is measurable, that is, $t \mapsto \langle \mu_t, g\rangle$ is a Borel real function for every $g \in C(K)$. The* average *of the measures μ_t over all $t \in B$ (weighted by λ) is the measure $\mu \in \mathscr{P}(K)$ defined by*

$$\langle \mu, g\rangle := \int_B \langle \mu_t, g\rangle \, d\lambda(t) \quad \forall\, g \in C(K)\,, \tag{4.4}$$

and is denoted by $\int \mu_t \, d\lambda(t)$.

The previous definition is well-posed because the right-hand side of (4.4) is a well-defined bounded linear functional on $C(K)$. Notice moreover that the class $\mathcal{F}$ of all bounded function $g : K \to \mathbb{R}$ such that the map $t \mapsto \langle \mu_t, g\rangle$ is Borel measurable contains $C(K)$ by definition, and is closed with respect to pointwise convergence; thus $\mathcal{F}$ contains all bounded Borel functions, and identity (4.4) can be extended to every bounded Borel function $g : K \to \mathbb{R}$.

Fix now $\mu \in \mathscr{P}(K)$ and consider the push-forward measures $T^\# \mu$ with $T \in G$. The identity $\langle T^\# \mu, g\rangle = \langle \mu, gT\rangle$ immediately shows that the parametrization $T \mapsto \langle T^\# \mu, g\rangle$ is measurable in T, and for every probability measure Φ on G and every $g \in C(K)$ one has

$$\Big\langle \int_G T^\# \mu \; d\Phi(T)\,,\, g \Big\rangle = \int_G \langle \mu, gT\rangle \, d\Phi(T) = \langle \mu, \tilde g\rangle, \tag{4.5}$$

where $\tilde g(x) := \int_G g(Tx) \, d\Phi(T)$ for every $x \in K$.

Lemma 4.10. *Let H be a co-compact subgroup of G, and let A be a representation of G/H. Then the elementary invariant measure ϵ_x associated with an H-periodic point x is given by*

$$\epsilon_x = \int_A T^\# \delta_x \, dT = \int_A \delta_{Tx} \, dT. \tag{4.6}$$

Proof. Obviously the two integrals in (4.6) define the same probability measure μ on K, which is supported on $\mathscr{O}(x)$, and since ϵ_x is the only invariant measure supported on $\mathscr{O}(x)$, it suffices to verify that μ is invariant. To this end we recall that $[T] \mapsto [T]x := Tx$ is a well-defined continuous map from G/H to $\mathscr{O}(x)$, and that the push-forward of the canonical measure on A by the canonical projection of G onto G/H is (by definition) the Haar probability measure on G/H (see Definition 4.1). Hence for every function $g \in C(K)$ and every $S \in G$ we have

$$\begin{aligned}\langle S^\# \mu, g\rangle = \langle \mu, gS\rangle = \Big\langle \int_A \delta_{Tx} \, dT\,,\, gS \Big\rangle &= \int_{T \in A} g(STx) \, dT \\ &= \int_{[T] \in G/H} g\big([S][T]x\big) \, d[T] = \int_{[T] \in G/H} g\big([T]x\big) \, d[T].\end{aligned}$$

This shows that for every $g \in C(K)$ the value of $\langle S^\# \mu, g\rangle$ is independent of S, and thus μ is invariant. $\square$

Lemma 4.11. *Let Φ be any probability measure on G. Then every $\mu \in \mathscr{P}(K)$ can be approximated by convex combination μ_k of Dirac masses so that*

$$\lim_{k \to \infty} \phi\Big(\int_G T^\# \mu \, d\Phi(T) - \int_G T^\# \mu_k \, d\Phi(T) \Big) = 0. \tag{4.7}$$

Proof. For every $\mu \in \mathscr{P}(K)$, let $P\mu$ be the average of $T^\# \mu$ over all $T \in G$ weighted by the measure Φ, that is, $P\mu := \int_G T^\# \mu \, d\Phi(T)$. Thus P is a continuous operator from $\mathscr{P}(K)$ into $\mathscr{P}(K)$ (use for instance identity (4.5)). Now we take *any* sequence of convex combinations μ_k of Dirac masses which converge to μ; thus

$P\mu_k$ converge to $P\mu$, and by Proposition 4.8(ii) we get $\phi(P\mu_k - P\mu) \to 0$, which is (4.7). □

LEMMA 4.12. *Assume that K is uniformly approximable, consider $\varepsilon > 0$ and a co-compact subgroup H as in Definition 4.3, and let A be a representation of G/H. Then for every $\mu \in \mathscr{P}(K)$ we may find a convex combination of elementary invariant measures $\bar\mu = \sum \sigma_i \epsilon_{\bar x_i}$ so that all $\bar x_i$ are H-periodic and*

$$\phi\Big(\int_A T^\# \mu \, dT \; - \bar\mu\Big) \le 2\varepsilon. \tag{4.8}$$

PROOF. By applying Lemma 4.11 with Φ replaced by Φ_A we may find a convex combination of Dirac masses $\hat\mu = \sum_i \sigma_i \, \delta_{x_i}$ so that

$$\phi\Big(\int_A T^\# \mu \, dT \; - \int_A T^\# \hat\mu \, dT\Big) \le \varepsilon. \tag{4.9}$$

Now we exploit the fact that the subgroup H was chosen according to Definition 4.3, and we approximate every x_i with an H-periodic point $\bar x_i$ so that (4.1) holds. Therefore, recalling statement (iii) of Proposition 4.8, we obtain

$$\begin{aligned}\phi\Big(\int_A T^\# \delta_{x_i} \, dT - \int_A T^\# \delta_{\bar x_i} \, dT\Big) &\le \int_A \phi\big(T^\# \delta_{x_i} - T^\# \delta_{\bar x_i}\big) \, dT \\ &\le \int_A d(Tx_i, T\bar x_i) \, dT \le \varepsilon. \end{aligned}\tag{4.10}$$

By Lemma 4.10 the average of the measures $T^\# \delta_{\bar x_i}$ over all $T \in A$ is the elementary invariant measure $\epsilon_{\bar x_i}$ (recall that $\bar x_i$ is H-periodic). Hence we set

$$\bar\mu := \sum_i \sigma_i \, \epsilon_{\bar x_i},$$

and by (4.10) we get

$$\begin{aligned}\phi\Big(\int_A T^\# \hat\mu \, dT \; - \; \bar\mu\Big) &\le \sum_i \sigma_i \phi\Big(\int_A T^\# \delta_{x_i} \, dT - \int_A T^\# \delta_{\bar x_i} \, dT\Big) \\ &\le \sum_i \sigma_i \varepsilon = \varepsilon. \end{aligned}\tag{4.11}$$

Inequalities (4.9) and (4.11) yield (4.8). □

We can now prove Theorem 4.4. Let μ be an invariant probability measure and fix $\varepsilon > 0$. Apply Lemma 4.12 to find a convex combination $\bar\mu$ of elementary

invariant measures such that (4.8) holds. Since $\mu = T^\# \mu$ for every $T \in G$, (4.8) becomes

$$\phi(\mu - \bar\mu) \le 2\varepsilon.$$

Approximation in energy

In the applications we have in mind, K is a function space endowed with some "natural" lower semicontinuous functional $f : K \to [0, +\infty]$. In this situation we may need to approximate an invariant probability measure μ on K by convex combinations μ_k of elementary invariant measures which verify the additional constraint

$$\lim_{k\to\infty} \langle \mu_k, f\rangle = \langle \mu, f\rangle. \tag{4.12}$$

In the following we modify Definition 4.3 and Theorem 4.4 in order to incorporate such constraint.

REMARK 4.13. Notice that the map $\mu \mapsto \langle \mu, f\rangle$ is well-defined and weak-star lower semicontinuous on $\mathscr{P}(K)$ because f is non-negative and lower semicontinuous. Therefore (4.12) holds whenever $\limsup \langle \mu_k, f\rangle \le \langle \mu, f\rangle$.

DEFINITION 4.14. *A set $X \subset K$ is called* f-uniformly approximable *if for every $\varepsilon > 0$ there exists a co-compact subgroup H and a representation A of G/H such that for every point $x \in X$ we may find an H-periodic point $\bar x \in K$ which satisfies*

$$\int_A d(Tx, T\bar x) \, dT \le \varepsilon, \tag{4.13}$$

$$\int_A f(T\bar x) \, dT \le \int_A f(Tx) \, dT \; + \varepsilon. \tag{4.14}$$

THEOREM 4.15. *If K is f-uniformly approximable, then every invariant probability measure μ on K can be approximated by convex combinations μ_k of elementary invariant measures so that* (4.12) *holds.*

The proof of this theorem is obtained by adapting the proof of Theorem 4.4. To this end we have to modify Lemmas 4.11 and 4.12.

LEMMA 4.16. *Let Φ be a probability measure on G. Then every $\mu \in \mathscr{P}(K)$ can be approximated by convex combinations μ_k of Dirac masses which satisfy* (4.7) *and*

$$\Big\langle \int_G T^\# \mu_k \, d\Phi(T), f \Big\rangle \le \Big\langle \int_G T^\# \mu \, d\Phi(T), f \Big\rangle \quad \textit{for every } k. \tag{4.15}$$

PROOF. For every $\mu \in \mathscr{P}(K)$ we consider $P\mu := \int_G T^\# \mu \, d\Phi(T)$ as in the proof of Lemma 4.11. We claim that every $\mu \in \mathscr{P}(K)$ may be approximated by a sequence (μ_k) of convex combinations of Dirac masses so that (4.15) holds, that is, $\langle P\mu_k, f\rangle \le \langle P\mu, f\rangle$ for every k. Once this claim is proved, the rest of the proof of Lemma 4.16 follows that of Lemma 4.11.

Fix now $\mu \in \mathscr{P}(K)$ and set $a := \langle P\mu, f\rangle$. With no loss of generality we may assume that a is finite, and then set

$$C := \big\{\lambda \in \mathscr{P}(K) : \langle P\lambda, f\rangle \le a\big\}. \tag{4.16}$$

By (4.5) we have that $\langle P\lambda, f\rangle = \langle \lambda, \tilde f\rangle$ where $\tilde f(x) := \int f(Tx)\, d\Phi(T)$ for every $x \in K$, and since $\tilde f$ is lower semicontinuous and positive, the set C is convex and compact. Moreover the extreme points of C are convex combinations of two Dirac masses (see [9], proposition 2, sec. II.2, p. 145). Since μ belongs to C, we can apply the Krein-Milman theorem to approximate μ with convex combinations μ_k of extreme points of C, and thus (4.15) follows from (4.16). □

LEMMA 4.17. *Assume that K is f-uniformly approximable, consider $\varepsilon > 0$ and a co-compact subgroup H as in Definition 4.14, and let A be a representation of G/H. Then for every $\mu \in \mathscr{P}(K)$ we may find a convex combination of elementary invariant measures $\bar\mu = \sum_i \sigma_i \epsilon_{\bar x_i}$ so that each $\bar x_i$ is H-periodic,* (4.8) *holds and*

$$\langle \bar\mu, f\rangle \le \Big\langle \int_A T^\# \mu\, dT, f\Big\rangle + \varepsilon. \tag{4.17}$$

PROOF. We proceed as in the proof of Lemma 4.12: we apply Lemma 4.16 to find a convex combination of Dirac masses $\hat\mu = \sum_i \sigma_i\, \delta_{x_i}$ so that (4.9) holds and

$$\Big\langle \int_A T^\# \hat\mu\, dT, f\Big\rangle \le \Big\langle \int_A T^\# \mu\, dT, f\Big\rangle. \tag{4.18}$$

Now we can exploit the choice of H and approximate every x_i with an H-periodic point $\bar x_i$ so that (4.13) and (4.14) hold. We define $\bar\mu := \sum \sigma_i \epsilon_{\bar x_i}$, and hence (4.8) follows as in the proof of Lemma 4.12. On the other hand by identity (4.6) and inequality (4.14) we get

$$\begin{aligned}\langle \bar\mu, f\rangle = \sum_i \sigma_i \int_A f(T\bar x_i)\, dT &\le \sum_i \sigma_i \int_A f(Tx_i)\, dT \ + \varepsilon \\ &= \Big\langle \int_A T^\# \hat\mu\, dT, f\Big\rangle + \varepsilon,\end{aligned}$$

which, together with inequality (4.18), implies (4.17). □

We can now prove Theorem 4.15. As in the proof of Theorem 4.4 we fix a real number $\varepsilon > 0$ and an invariant probability measure μ on K such that $\langle \mu, f\rangle$ is finite. Then we apply Lemma 4.17 to get a convex combinations of elementary

invariant measures $\bar\mu$ so that both (4.8) and (4.17) hold. Since μ is invariant (4.8) and (4.17) become respectively

$$\phi(\mu - \bar\mu) \le 2\varepsilon \quad \text{and} \quad \langle \bar\mu, f\rangle \le \langle \mu, f\rangle + \varepsilon.$$

By Remark 4.13 this concludes the proof of Theorem 4.15.

Extension to the non-commutative case

Theorems 4.4 and 4.15 hold also when the group G is a non-commutative. In this case, however, some of the previous definitions need to be modified. We first remark that if x is a periodic point but $P(x)$ is not a normal subgroup, then the quotient $G/P(x)$ is not a group.

Therefore our construction of the elementary invariant measure ϵ_x fails, and in fact the orbit of x, although compact, may support no invariant probability measure. Consider for instance the following example: K is the projective line $\mathbb{R} \cup \{\infty\}$ and G the group of all projective transformations of K, that is, transformations of the form $x \mapsto (ax+b)\big/(cx+d)$ with $ad - bc \ne 0$. Then the orbit of any point x is K, $G/P(x)$ is homeomorphic to K and then $P(x)$ is co-compact, but K supports no invariant measures (since translations $x \mapsto x + b$ are projective transformations, any invariant measure should be supported at ∞, but this is impossible, too, because G acts transitively on K).

The previous example motivates the following definition: we say that a co-compact subgroup H of G is a *W-subgroup* if there exists a probability measure on G/H which is invariant under the left action of G.

This probability measure is unique (see [22], theorem 2.7.11(2)), and is denoted by $\Phi_{G/H}$. A co-compact subgroup H is a W-subgroup if and only if it satisfies the so-called Weil's condition, namely that the modular functions of G and H agree on H; in particular Weil's condition is verified when H is normal, or when G is compact (see [22], theorem 2.7.11 and sec. 2.7.12, or [25], sec. 15). Notice that if H is a W-subgroup, then also every co-compact subgroup H' which includes H is a W-subgroup. When x is a periodic point, the map $[T] \mapsto Tx$ is a homeomorphism of G/H into $\mathscr{O}(x)$, and then $\mathscr{O}(x)$ supports an invariant probability measure if and only if $P(x)$ is a W-subgroup.

Therefore the following modifications should be introduced to adapt the results of this section to the non-commutative case: the elementary invariant measures can be defined only for periodic points x such that $P(x)$ is a W-subgroup, and in Definitions 4.1, 4.3, 4.14, and Proposition 4.2, it must be required that H is a W-subgroup.

5. Approximation of invariant measures on function spaces

In this section we present in detail the case where K is a space of functions on $\mathbb{R}^n$, and show that the assumptions of Theorem 4.4 are verified. Then we restrict our attention to the particular situation considered in Section 3, we show that the assumptions of Theorem 4.15 are satisfied, and obtain the approximation in energy used in the proof of Theorems 3.4 and 3.12.

We conform to the notation of Section 4, with the only difference that K is now the set of all Borel functions $x : \mathbb{R}^n \to [-\infty, +\infty]$ modulo equivalence almost everywhere, and G is the group of functional translations, and is represented by $\mathbb{R}^n$: for every $\tau \in \mathbb{R}^n$ and every $x \in K$, $T_\tau x$ is the translated function $x(t-\tau)$.

By identifying the extended real line $[-\infty, +\infty]$ with the closed interval $[-1, 1]$ via the function $x \mapsto \frac{2}{\pi}\arctan(x)$, we can identify K with the closed unit ball of $L^\infty(\mathbb{R}^n)$ and endow it with the weak-star topology of $L^\infty(\mathbb{R}^n)$. Thus K is compact and metrizable. In particular we can consider the following distance: let (y_k) be a sequence of bounded functions which are dense in $L^1(\mathbb{R}^n)$, and such that each y_k has support included in the cube $(-k, k)^n$; for every $x_1, x_2 \in K$ set

$$d(x_1, x_2) := \sum_{k=1}^{\infty} \frac{1}{2^k \alpha_k} \Big| \int_{\mathbb{R}^n} y_k \Big(\frac{2}{\pi}\arctan x_1 - \frac{2}{\pi}\arctan x_2 \Big)\, d\mathscr{L}_n \Big|, \tag{5.1}$$

where $\alpha_k := \|y_k\|_1 + \|y_k\|_\infty$.

It follows immediately from (5.1) that when the functions x_k converge to x locally in measure, then they converge to x also in the distance d. Hence $L^p_{\mathrm{loc}}(\mathbb{R}^n)$ embeds continuously in K for $1 \le p \le \infty$. Moreover, (5.1) yields, for every $p \in [1, \infty]$,

$$\begin{aligned} d(x_1, x_2) &\le \sum_{k=1}^{\infty} \frac{1}{2^k \alpha_k} \int_{\mathbb{R}^n} |y_k|\, |x_1 - x_2| \\ &\le \sum_{k=1}^{\infty} \frac{\|y_k\|_q \|x_1 - x_2\|_p}{2^k\big(\|y_k\|_1 + \|y_k\|_\infty\big)} \le \|x_1 - x_2\|_p. \end{aligned} \tag{5.2}$$

(The first inequality follows from the fact that $\frac{2}{\pi}\arctan$ is 1-Lipschitz, the second one is Hölder's, and the last one follows from the interpolation $\|y_k\|_q \le \|y_k\|_1 + \|y_k\|_\infty$).

Remark. Embedding into K may be no longer continuous if we consider weaker forms of convergence. For instance, if the functions $x_k : \mathbb{R}^n \to \{a, b\}$ weak-star converge to the constant function $\frac{1}{2}(a+b)$ in $L^\infty(\mathbb{R}^n)$, then they converge on K to the constant function $\tan\big(\frac{1}{2}(\arctan a + \arctan b)\big)$.

The main feature of the distance d is the following locality property, which in fact is shared by every distance which metrizes K.

Proposition 5.1. *For every $\varepsilon > 0$ there exists $m > 0$ such that the following implication holds for $x_1, x_2 \in K$:*

$$(x_1 \wedge m) \vee -m = (x_2 \wedge m) \vee -m \ \text{a.e. in } (-m, m)^n \ \Rightarrow\ d(x_1, x_2) \le \varepsilon. \tag{5.3}$$

(Here $a \wedge b$ and $a \vee b$ denote respectively the minimum and the maximum of a and b.)

Proof. Fix a positive real number m and take x_1, x_2 such that the hypothesis of (5.3) holds. Then $\big|\arctan x_1(t) - \arctan x_2(t)\big| \le \pi/2 - \arctan m$ for a.e. $t \in (-m, m)^n$, and since $\operatorname{spt} y_k \subset (-k, k)^n$, for $k \le m$ we have

$$\Big| \int_{\mathbb{R}^n} y_k \big(\arctan x_1 - \arctan x_2\big)\, d\mathscr{L}_n \Big| \le \|y_k\|_1 \Big(\frac{\pi}{2} - \arctan m\Big),$$

while for $k > m$ the integral on the left-hand side is controlled by $\pi\|y_k\|_1$. Hence, recalling formula (5.1) and that $\alpha_k \ge \|y_k\|_1$,

$$d(x_1, x_2) \le \sum_{k=1}^{m} \frac{1}{2^k}\Big(\frac{\pi}{2} - \arctan m\Big) + \sum_{k=m+1}^{\infty} \frac{\pi}{2^k} \le \frac{\pi}{2} - \arctan m + \frac{\pi}{2^m}.$$

To finish the proof it suffices to choose m large enough. □

Remark 5.2. Given $x_1, x_2 \in K$, let x be the function which agrees with x_1 in the cube $(-m, m)^n$, and with x_2 elsewhere. Hence $d(x_1, x_2) \le d(x_1, x) + d(x, x_2)$, and if we estimate $d(x_1, x)$ by (5.3), and $d(x, x_2)$ by (5.2), we obtain the following useful inequality:

$$d(x_1, x_2) \le \varepsilon + \|x_1 - x_2\|_{L^p((-m,m)^n)} \quad \text{for } x_1, x_2 \in K, \tag{5.4}$$

where m and ε are taken as in Proposition 5.1, and p is any number in $[1, +\infty]$.

Proposition 5.3. *The group of functional translations G acts continuously on K, and K is uniformly approximable. Thus Theorem 4.4 applies, and every invariant probability measure on K can be approximated by convex combinations of elementary invariant measures.*

Proof. We prove that G acts continuously on K by showing that the group of translations act (sequentially) continuously on $L^\infty(\mathbb{R}^n)$ endowed with the weak-star topology. Consider $\tau_k \to \tau$ in $\mathbb{R}^n$, $x_k \stackrel{*}{\rightharpoonup} x$ in $L^\infty(\mathbb{R}^n)$, and $y \in L^1(\mathbb{R}^n)$. Then $T_{-\tau_k} y \to T_{-\tau} y$ in $L^1(\mathbb{R}^n)$ and thus

$$\langle T_{\tau_k} x_k - T_\tau x, y \rangle = \langle x_k, T_{-\tau_k} y \rangle - \langle x, T_{-\tau} y \rangle \longrightarrow 0.$$

Since this holds for every $y \in L^1(\mathbb{R}^n)$ we deduce that $T_{\tau_k} x_k \stackrel{*}{\rightharpoonup} T_\tau x$ in $L^\infty(\mathbb{R}^n)$.

Let us show that K is uniformly approximable. Fix $\varepsilon > 0$, take m so that implication (5.3) holds, and then choose a so that $a\varepsilon \ge m$. For every $x \in K$, let $\bar{x}$ be the function on $\mathbb{R}^n$ which agrees with x on the cube $(0, a)^n$ and is extended periodically to the whole of $\mathbb{R}^n$. Then $\bar{x}$ is $(a\mathbb{Z}^n)$-periodic, and $\bar{x}(t-\tau) = x(t-\tau)$ whenever $t \in (-m, m)^n$ and $\tau \in (m, a-m)^n$. Hence (5.3) yields $d(T_\tau \bar{x}, T_\tau x) \le \varepsilon$

for every $\tau \in (m, a-m)^n$; on the other hand the distance d is never larger than one, and recalling that $a\varepsilon \geq m$ we obtain

$$\fint_{(0,a)^n} d(T_\tau \bar{x}, T_\tau x)\, d\mathscr{L}_n(\tau) \leq \frac{\varepsilon\, a^n + 2nm\, a^{n-1}}{a^n} \leq (1+2n)\varepsilon. \qquad \square$$

Remark. Notice that there exist points $x \in K$ whose orbits are dense in K. In other words $\mathscr{O}(x)$ is an element of the quotient space K/G which is dense in K/G, and then the topology of K/G is not Hausdorff, and not even T_0. To construct such a function x, we take a sequence (x_k) which is dense in K, and for every k we choose the positive real number m_k corresponding to $\varepsilon = 1/k$ in Proposition 5.1; then we take pairwise disjoint open cubes $C_k = -\tau_k + (-m_k, m_k)^n$, and choose as x any function which agrees with $T_{-\tau_k} x_k$ on each cube C_k. Hence $T_{\tau_k} x = x_k$ in $(-m_k, m_k)^n$ for every k, and (5.3) yields $d(T_{\tau_k} x, x_k) \leq 1/k$ for every k. Hence the orbit of x is dense in K.

A similar argument can be used to prove that every convex combination of elementary invariant measures can be approximated by elementary invariant measures. Together with Proposition 5.3, this would yield that every invariant probability measure on K is in fact the limit of a sequence of elementary invariant measures. In Lemma 5.10 we prove this fact, and something more, for $n = 1$.

A one-dimensional example

We apply now Theorem 4.15 to the choice of K and f considered in Section 3. Thus G and K are given as before with $n = 1$, and in particular G is represented by $\mathbb{R}$. Every proper co-compact subgroup of $\mathbb{R}$ is of the form $h\mathbb{Z}$ for some $h > 0$, and a representation is given by the interval $(0, h)$, endowed with Lebesgue measure, suitably renormalized.

For the rest of this section the letter h will be mainly used to denote periods of elements of X. The spaces $\mathscr{S}(I)$ and $\mathscr{S}_{\mathrm{per}}(0,h)$ are defined at the beginning of Section 3, while $\mathscr{S}_{\mathrm{per},0}(0,h)$ denotes the space of all $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ such that $x(0) = x(h) = 0$; r is a fixed positive real number and we set (cf. (3.8))

$$f(x) := \begin{cases} \dfrac{1}{2r} \#\big(S\dot{x} \cap (-r,r)\big) + \displaystyle\fint_{-r}^{r} x^2(t)\, dt & \text{if } x \in \mathscr{S}(-r,r), \\ +\infty & \text{otherwise.} \end{cases} \tag{5.5}$$

Proposition 5.4. *The function f is lower semicontinuous on K.*

Proof. Let be given functions x_k that converge to x in K such that the values $f(x_k)$ are uniformly bounded. Then the functions x_k belong to $\mathscr{S}(-r,r)$ for every k, they are 1-Lipschitz on $(-r,r)$ and uniformly bounded in $L^2(-r,r)$. Hence they converge to x uniformly in $(-r,r)$. This implies that the distributional derivatives $\dot{x}_k$ converge to $\dot{x}$ weak-star in $BV(-r,r)$, hence x belongs to $\mathscr{S}(-r,r)$, and $\liminf f(x_k) \geq f(x)$ (recall that $2\#\big(S\dot{x} \cap (-r,r)\big)$ is the total variation of $\ddot{x}$ on $(-r,r)$). $\square$

Now we want to prove that K is f-uniformly approximable (recall Definition 4.14). To this end we need some preliminary lemmas and definitions. In what follows, $*$ denotes the usual convolution products, 1_B is the characteristic function of the set B, and we set

$$\rho(t) := \frac{1}{2r}\, 1_{[-r,r]}(t) \quad \text{for } t \in \mathbb{R}. \tag{5.6}$$

Lemma 5.5. *Let $x \in K$ satisfy $\int_0^h f(T_\tau x)\, d\tau < +\infty$. Then $x \in \mathscr{S}(I)$ for every I relatively compact in $(-r, h+r)$ and*

$$\fint_0^h f(T_\tau x)\, d\tau = \frac{1}{h}\Big[\sum_{t \in S\dot{x}} (\rho * 1_{[0,h]})(t) \;+ \int_{\mathbb{R}} (\rho * 1_{[0,h]})\, x^2\, dt\Big] \tag{5.7}$$

(notice that the convolution product $\rho * 1_{[0,h]}$ vanishes outside $(-r, h-r)$). Moreover, if x is h-periodic, then $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ and

$$\langle \epsilon_x, f\rangle = \fint_0^h f(T_\tau x)\, d\tau = \frac{1}{h}\Big[\#\big(S\dot{x} \cap [0,h)\big) + \int_0^h x^2 dt.\Big] \tag{5.8}$$

Proof. As $\int_0^h f(T_\tau x)\, d\tau < +\infty$, $f(T_\tau x)$ is finite for a.e. $\tau \in (0,h)$, which implies $x \in \mathscr{S}(\tau - r, \tau + r)$, and since every interval I relatively compact in $(-r, h+r)$ can be covered by finitely many such intervals $(\tau - r, \tau + r)$, then $x \in \mathscr{S}(I)$.

To obtain (5.7), we consider the measure μ given by $\mu(B) := \#\big(B \cap S\dot{x} \cap (-r, h+r)\big) + \int_B x^2\, dt$, and thus we write $f(T_\tau x)$ as

$$f(T_\tau x) = \int_{\mathbb{R}} \rho(t-\tau)\, d\mu(t).$$

Integration over $\tau \in (0,h)$ yields (5.7). The second part of the assertion follows from the fact that on $\mathbb{R}$ modulo h there holds $\rho * 1_{[0,h]} = \rho * 1 = 1$. $\square$

Definition 5.6. *For every $h > 2r$ and every $x \in K$, the h-periodic function $R_h x$ is defined as follows (see Figure 5.1):*

- *for $0 \leq t < r$, $R_h x(t) := t \wedge (-t + r)$;*
- *for $r \leq t < h/2)$, $R_h x(t)$ is set equal to $x(t)$ if $|x(t)| \leq t - r$, to $t - r$ if $x(t) > t - r$, and to $-(t-r)$ if $x(t) < -(t-r)$;*
- *for $h/2 \leq t < h - r$, $R_h x(t)$ is set equal to $x(t)$ if $|x(t)| \leq h - r - t$, to $h - r - t$ if $x(t) > h - r - t$, and to $-(h - r - t)$ if $x(t) < -(h - r - t)$;*
- *for $h - r \leq t < h)$, $R_h x(t) := (t - h + r) \vee (-t + h)$.*

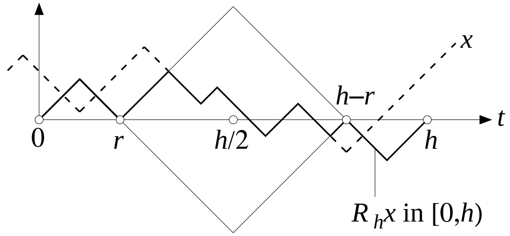

FIGURE 5.1. The function $R_h x$

LEMMA 5.7. *Let h, x and $R_h x$ as in Definition 5.6. Thus $R_h x$ is h-periodic and $R_h x(0) = R_h x(h) = 0$ by construction. Moreover*

(i) *for every $m > 0$ and $t \in (m + r, h - m - r)$, either $x(t) = R_h x(t)$, or $x(t) \geq R_h x(t) \geq m$, or $x(t) \leq R_h x(t) \leq -m$;*

(ii) *if $x \in \mathscr{S}(r, h - r)$ then $R_h x \in \mathscr{S}_{\mathrm{per},0}(0, h)$ and $S\dot{x} \cap (r, h - r)$ contains $S(R_h x)' \cap [0, h)$ except at most six points.*

PROOF. Straightforward (see Figure 5.1). □

PROPOSITION 5.8. *For every $\varepsilon > 0$ there exists $h > 0$ such that for every $x \in K$*

$$\begin{aligned} &\fint_0^h d(T_\tau x, T_\tau R_h x)\, d\tau \leq 2\varepsilon, \\ &\fint_0^h f(T_\tau R_h x)\, d\tau \leq \fint_0^h f(T_\tau x)\, d\tau + \varepsilon. \end{aligned} \tag{5.9}$$

In particular, K is f-uniformly approximable (recall Definition 4.14).

PROOF. Fix $m > 0$ such that implication (5.3) holds, and take $h > 2(m + r)$. Then statement (i) of Lemma 5.7 and (5.3) imply that $d(T_\tau x, T_\tau R_h x) \leq \varepsilon$ for every $\tau \in (m + r, h - m - r)$. Hence, taking into account that $d \leq 1$,

$$\fint_0^h d(T_\tau x, T_\tau R_h x)\, d\tau \leq \varepsilon + \frac{2(m + r)}{h},$$

and the first inequality in (5.9) is recovered by choosing $h \geq \frac{2(m+r)}{\varepsilon}$.

Consider now the second inequality in (5.9). We can assume that the integral $\fint_0^h f(T_\tau x)\, d\tau$ is finite (otherwise there is nothing to prove). Therefore $R_h x \in \mathscr{S}_{\mathrm{per},0}(0, h)$, and

- $\#\big(S(R_h x)' \cap [0, h)\big) \leq \#\big(S\dot{x} \cap (r, h - r)\big) + 6$ (see Lemma 5.7(ii)),
- $|R_h x| \leq |x|$ in $(r, h - r)$ (see Lemma 5.7(i)),
- $|R_h x| \leq r/2$ in $(0, r)$ and $(h - r, h)$ (by construction).

Hence, by (5.8),

$$\begin{aligned} \fint_0^h f(T_\tau R_h x)\, d\tau &= \frac{1}{h}\Big[\#\big(S(R_h x)' \cap [0, h)\big) + \int_0^h \big(R_h x\big)^2\, dt\Big] \\ &\leq \frac{1}{h}\Big[\#\big(S\dot{x} \cap (r, h - r)\big) + 6 + \int_r^{h-r} x^2\, dt + \frac{r^2}{2}\Big] \end{aligned}$$

and since $\rho * 1_{[0,h]} = 1$ in $(r, h - r)$,

$$\leq \frac{1}{h}\Big[\sum_{t \in S\dot{x}} (\rho * 1_{[0,h]})(t) + \int_{\mathbb{R}} (\rho * 1_{[0,h]})\, x^2\, dt + \frac{12 + r^2}{2}\Big]$$

and by (5.7),

$$= \fint_0^h f(T_\tau x)\, d\tau + \frac{12 + r^2}{2h}.$$

The second inequality in (5.9) is thus recovered by choosing $h \geq \frac{12+r^2}{2\varepsilon}$. □

COROLLARY 5.9. *Every invariant probability measure μ on K can be approximated by a sequence of convex combinations μ_k of elementary invariant measures so that (4.12) holds. Moreover, if $\langle \mu, f \rangle$ is finite, each μ_k can be taken of the form $\mu_k = \sum \sigma_i \epsilon_{\bar{x}_{ki}}$ with $\bar{x}_{ki} \in \mathscr{S}_{\mathrm{per},0}(0, h_k)$ for every i and a suitable $h_k > 0$.*

PROOF. By Proposition 5.8 the space K is f-uniformly approximable, and then the first part of Corollary 5.9 follows from Theorem 4.15. Furthermore Proposition 5.8 shows that for every $x \in K$ the approximating point $\bar{x}$ in Definition 4.14 can be taken equal to $R_h x$, and if we examine the construction of the measure $\bar{\mu}$ described in the proof of Lemma 4.17, we see that $\bar{\mu}$ can be taken of the form $\sum \sigma_i \epsilon_{\bar{x}_i}$ with $\bar{x}_i = R_h x_i$, and then $\bar{x}_i$ is h-periodic and $\bar{x}_i(0) = \bar{x}_i(h) = 0$.

Thus the same holds for the approximating measures μ_k given in Theorem 4.15. Moreover $\langle \mu, f \rangle < +\infty$ implies that $\langle \mu_k, f \rangle$ is finite (for k large enough). Hence $\langle \epsilon_{\bar{x}_{ki}}, f \rangle$ is also finite for every i, and Lemma 5.5(ii) yields $\bar{x}_{ik} \in \mathscr{S}_{\mathrm{per},0}(0, h_k)$. □

We can refine the statement of Corollary 5.9 by showing that μ can be directly approximated by elementary invariant measures.

LEMMA 5.10. *Given $\varepsilon > 0$, $h > 0$ and $\mu = \sum_1^N \sigma_i \epsilon_{x_i}$ such that $x_i \in \mathscr{S}_{\mathrm{per},0}(0, h)$ for every $i = 1, \ldots, N$, we can find $\bar{h} > 0$ and $x \in \mathscr{S}_{\mathrm{per},0}(0, \bar{h})$ such that*

$$\phi(\mu - \epsilon_x) \leq 2\varepsilon \quad \textit{and} \quad \big|\langle \mu, f \rangle - \langle \epsilon_x, f \rangle\big| \leq \varepsilon. \tag{5.10}$$

PROOF. First of all, notice that all σ_i can be assumed rational (by a standard density argument). We fix $m > 0$ such that implication (5.3) holds, and we write every σ_i as $\sigma_i = p_i/q$ with positive integers q and p_i. Notice that q can be taken arbitrarily large.

We set $q_0 := 0$, $q_i := q_{i-1} + p_i$ for $i = 1, \dots, N$ (in particular $q_N = q$), and we take $x \in \mathscr{S}_{\mathrm{per},0}(0, qh)$ defined by

$$x(t) := x_i(t) \text{ for every } t \in [q_{i-1}h, q_i h) \text{ and } i = 1, \dots, N. \tag{5.11}$$

In other words x is equal to x_1 in the first p_1 periods of length h, it is equal to x_2 in the following p_2 periods, and so on for a total of q periods; cf. Figure 5.2.

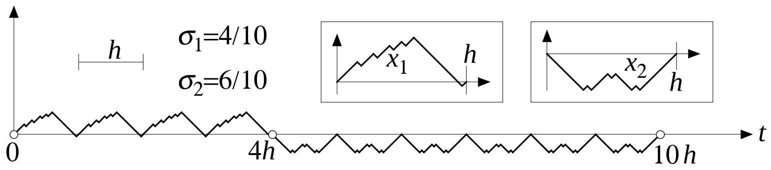

FIGURE 5.2. Construction of x for $N = 2$, $q = 10$, $p_1 = 4$, $p_2 = 6$

By formula (4.6) we get

$$\epsilon_x = \frac{1}{qh}\int_0^{qh} \delta_{T_\tau x} d\tau, \qquad \epsilon_{x_i} = \frac{1}{h}\int_0^h \delta_{T_\tau x_i} d\tau = \frac{1}{p_i h}\int_{q_{i-1}h}^{q_i h} \delta_{T_\tau x_i} d\tau.$$

Hence

$$\mu - \epsilon_x = \Big(\sum_i \sigma_i \epsilon_{x_i}\Big) - \epsilon_x = \sum_{i=1}^N \frac{1}{qh}\int_{q_{i-1}h}^{q_i h} \big(\delta_{T_\tau x_i} - \delta_{T_\tau x}\big)\, d\tau,$$

and by Proposition 4.8(iii)

$$\phi(\mu - \epsilon_x) \le \sum_{i=1}^N \frac{1}{qh}\int_{q_{i-1}h}^{q_i h} d(T_\tau x, T_\tau x_i)\, d\tau. \tag{5.12}$$

Thus we need to estimate the distance $d(T_\tau x, T_\tau x_i)$. From (5.11) we deduce that for every $\tau \in (q_{i-1}h + m, q_i h - m)$ and every i there holds $x = x_i$ in $(\tau - m, \tau + m)$, and then $d(T_\tau x, T_\tau x_i) \le \varepsilon$ by (5.3). Hence inequality (5.12) becomes (recall that $d \le 1$)

$$\phi(\mu - \epsilon_x) \le \sum_{i=1}^N \frac{1}{qh}(p_i h\varepsilon + 2m) = \sum_{i=1}^N \frac{p_i \varepsilon}{q} + \frac{2mN}{qh} = \varepsilon + \frac{2mN}{qh},$$

and the first inequality in (5.10) is recovered by choosing $q \ge \frac{2mN}{q\varepsilon}$.

Let us prove the second inequality in (5.10). From (5.11) we get $S\dot{x} \cap (q_{i-1}h, q_i h) = S\dot{x}_i \cap (q_{i-1}h, q_i h)$ for every i, and then $\#\big(S\dot{x} \cap [0, qh)\big) \le N + \sum_i \#\big(S\dot{x}_i \cap [q_{i-1}h, q_i h)\big)$. Hence (5.8) yields

$$\begin{aligned}
\langle \epsilon_x, f\rangle &= \frac{1}{qh}\Big[\#\big(S\dot{x} \cap [0, qh)\big) + \int_0^{qh} x^2(t)\, dt\Big] \\
&\le \frac{N}{qh} + \sum_i \Big[\frac{1}{qh}\#\big(S\dot{x}_i \cap [q_{i-1}h, q_i h)\big) + \frac{1}{qh}\int_{q_{i-1}h}^{q_i h} x_i^2(t)\, dt\Big] \\
&= \frac{N}{qh} + \sum_i \Big[\frac{p_i}{qh}\#\big(S\dot{x}_i \cap [0, h)\big) + \frac{p_i}{qh}\int_0^h x_i^2(t)\, dt\Big] \\
&= \frac{N}{qh} + \sum_i \sigma_i \langle \epsilon_{x_i}, f\rangle = \frac{N}{qh} + \langle \mu, f\rangle,
\end{aligned}$$

and the second inequality in (5.10) is recovered by choosing $q \ge \frac{N}{h\varepsilon}$. □

COROLLARY 5.11. *Every invariant probability measure μ on K which satisfies $\langle \mu, f\rangle < +\infty$ can be approximated by a sequence of elementary invariant measures (ϵ_{x_k}) so that $x_k \in \mathscr{S}_{\mathrm{per},0}(0, h_k)$ for some $h_k > 0$ and*

$$\lim_{k\to\infty} \langle \epsilon_{x_k}, f\rangle = \langle \mu, f\rangle.$$

PROOF. Apply Corollary 5.9 and Lemma 5.10. □

6. Overview of further applications

In this section we briefly sketch some extensions of our approach to other variational problems with multiple scales. We begin with some variations of the one-dimensional problem studied in Section 3.

6.1 Boundary conditions

The periodicity constraint imposed in the study of the functional I^ε (see (3.3)) can be replaced by any reasonable boundary condition (Dirichlet, natural, or even none at all) without changing the limit problem. In other words, the Γ-limit F defined in (3.11) is independent of boundary conditions.

Indeed, the presence of different boundary conditions only affects the formulation of f_s^ε for all s whose distance from the boundary is less than $r\varepsilon^{1/3}$ (cf. the paragraph "Rewriting $I^\varepsilon(v)$ in term of $\mathsf{R}_s^\varepsilon v$" in Section 3, and particularly formula (3.6)). Thus, given $s \in \Omega$, for every ε sufficiently small f_s^ε is the same as before, and so is the Γ-limit f_s; moreover one can easily adapt the proof of Theorem 3.4 to include this "vanishing" perturbation.

This is not surprising, since we already know from the form of F that in the limit $\varepsilon \to 0$ there are no correlations between different values of the slow variable s

(and indeed the minimization of F reduces to a family of independent minimization problems parametrized in s – cf. Corollary 3.11), which implies in particular that the effect of the boundary condition on the behaviour of minimizers near any point s vanishes as $\varepsilon \to 0$.

6.2 Additional externally imposed scales

The only property of the lower order term $\int av^2$ in I^ε (cf. (3.3)) used in the proof is that the rescaled functionals $\fint_{-r}^{r} a_s^\varepsilon x^2$ converge continuously for a.e. $s \in \Omega$ as $\varepsilon \to 0$ (see Definition 2.9). More precisely we need that the integrals

$$\fint_{-r}^{r} a(s+\varepsilon^{1/3}t)(x^\varepsilon(t))^2 dt$$

converge for any sequence x^ε which converges strongly in $W^{1,1}(-r,r)$.

Thus the proof of Theorem 3.4. can be extended (with almost no modifications) to more complex lower order terms. In particular we can consider highly oscillatory coefficients. For example, we can take (cf. (3.3))

$$I^\varepsilon := \int_\Omega \varepsilon^2 \ddot{v}^2 + W(\dot{v}) + a(\varepsilon^{-\beta} s)\, v^2\, ds,$$

where the function a is a bounded, strictly positive and periodic, and has average $\bar{a}$.

If $\beta > \frac{1}{3}$, i.e., if the externally imposed scale ε^β is shorter than the fast scale $\varepsilon^{1/3}$ used in our blow-up procedure, then Theorem 3.4 holds true, provided that we replace $a(s)\fint_{-r}^{r} x^2$ with $\bar{a}\fint_{-r}^{r} x^2$ in (3.8). This requires no modifications in the proof, since the rescaled functions $a_s^\varepsilon(t) := a(\varepsilon^{-\beta}s + \varepsilon^{1/3-\beta}t)$ converge weakly to the constant function $\bar{a}$, and then the functionals $\fint_{-r}^{r} a_s^\varepsilon(t)x^2$ converge continuously to $\bar{a}\fint_{-r}^{r} x^2$.

If $\beta < \frac{1}{3}$ then this convergence no longer holds. We expect that minimizers of the ε-problem are locally well approximated by periodic sawtooth functions with period $L_0(a(\varepsilon^{-\beta}s))^{-1/3}\varepsilon^{1/3}$ and generate the homogeneous two-scale Young measure

$$\nu_s = \fint \epsilon_{x_q}\, dq \quad \text{for a.e. } s \in (0,1),$$

where x_q is the sawtooth function with period $h_q := L_0 a^{-1/3}(q)$ defined in (3.33), and the average is taken over a period of the function a.

The (rescaled) limiting energy is thus given by $E_0\big(\fint a^{1/3}(q)dq\big)$ (cf. (1.4)). In this case the Γ-limit F (if it exists) cannot have the simple form (3.11), in fact it cannot be affine on the affine set defined by the condition $\nu_s \in \mathscr{I}(K)$ a.e. This follows from the fact that for the homogeneous two-scale Young measure $\nu_s = \epsilon_x$, where x is an h-periodic sawtooth function, one has (cf. (3.12))

$$F(\nu) = \frac{A_0}{h}\#(S\dot{x} \cap [0,h)) + \bar{a}\fint_0^h x^2\, dt.$$

Hence if F was affine the minimal energy would involve $\bar{a}^{1/3} = (\fint a)^{1/3}$ rather than the smaller value $\fint a^{1/3}$. In this case a more natural representation of the limit might be achievable by performing a (hierarchical) blow-up with two small scales ε^β and $\varepsilon^{1/3}$ and looking at the corresponding Young measures and limit functionals. A detailed implementation of this idea (and the verification of the statements above) is left to the courageous reader.

The case $\beta = \frac{1}{3}$ is particularly interesting since in this case the externally imposed scale and the internally created scale are of the same order and relative phases may play an important role. Note that the formula for the (rescaled) limiting energy changes discontinuously at $\beta = \frac{1}{3}$, since it is given by $E_0(\fint a)^{1/3}$ for $\beta > \frac{1}{3}$ (cf. (1.4)), and by $E_0(\fint a^{1/3})$ for $\beta < \frac{1}{3}$.

6.3 Additional penalizing term

A natural modification of the energy in (1.1) is obtained by replacing the term $a(s)\, v^2$ with the penalization $a(v-u)^2$, where u is a given function and a a positive constant, that is,

$$I^\varepsilon(v) := \int_\Omega \varepsilon^2 \ddot{v}^2 + W(\dot{v}) + a(v-u)^2\, ds. \tag{6.1}$$

By analogy with the case $u = 0$, we expect that when $|\dot{u}| \le 1$, the minimizers v^ε of I^ε get closer and closer to u as $\varepsilon \to 0$, while the derivatives $\dot{v}^\varepsilon$ take values closer and closer to ± 1 (because of the term $W(\dot{v})$ and present a "microstructure" at the scale $\varepsilon^{1/3}$ which is once again induced by the term $\varepsilon^2\ddot{v}^2$. Both the local volume fractions of the + and the − "phase" and the typical length-scale of the microstructure we expect to depend on $\dot{u}(s)$, with the length-scale approaching $+\infty$ (measured in units of $\varepsilon^{1/3}$) as $|\dot{u}(s)|$ approaches 1. A totally different behavior is expected where $|\dot{u}| > 1$, because a function with derivative larger than 1 cannot be approximated by sawtooth functions with derivative ± 1.

To proceed, we assume that u is of class C^1 and $|\dot{u}| < 1$. The effect we want to analyze is captured by the following ε-blowups (cf. (3.2)):

$$\mathsf{R}_s^\varepsilon v(t) := \varepsilon^{-1/3}\big(v(s+\varepsilon^{1/3}t) - u(s) - \dot{u}(s)\varepsilon^{1/3}t\big). \tag{6.2}$$

The program outlined in Section 3 may be carried out, with some modifications, in this case too.

The first step consists in proving that if ν is the Young measure generated by the ε-blowups associated to any sequence of functions v^ε, then ν_s is an invariant measure on K for a.e. s. With the usual notation $x_s^\varepsilon := \mathsf{R}_s^\varepsilon v^\varepsilon$, for a given $h \in \mathbb{R}$ we get

$$\begin{aligned}\tau_h x_s^\varepsilon - x^\varepsilon_{s+\varepsilon^{1/3}h} &= \\ &= \varepsilon^{-1/3}\big(u(s+\varepsilon^{1/3}h) - u(s) - \dot{u}(s)\varepsilon^{1/3}h\big) + \big(\dot{u}(s+\varepsilon^{1/3}h) - \dot{u}(s)\big)\, t.\end{aligned}$$

Using the fact that u is of class C^1 it is not difficult to verify that the distance between $\tau_h x_s^\varepsilon$ and $x_{s+\varepsilon^{1/3}h}^\varepsilon$ tends to 0 as $\varepsilon \to 0$ for (almost) every s. Therefore the two sequences generate the same Young measure. On the other hand, the first sequence generates (by definition) the Young measure $\tau_h\nu_s$, while the second one generates ν_s, and this suffices to prove that ν_s is invariant for a.e. s (cf. the proof of Proposition 3.1).

In order to rewrite the functionals $I^\varepsilon(v)$ in (6.1) in term of the ε-blowups $\mathsf{R}_s^\varepsilon v$, we define

$$b_s := \dot u(s) \quad \text{and} \quad a_s^\varepsilon(t) := \varepsilon^{-1/3}\big(u(s+\varepsilon^{1/3}t) - u(s) - \dot u(s)\varepsilon^{1/3}t\big).$$

Then

$$v - u = \varepsilon^{1/3}(x_s + a_s^\varepsilon), \quad \dot v = \dot x_s - b_s, \quad \ddot v = \varepsilon^{-1/3}\ddot x_s,$$

where all the left-hand sides are computed at $s+\varepsilon^{1/3}t$, and all the right-hand sides are computed at t. Proceeding as in Section 3, we get

$$\varepsilon^{-2/3} I^\varepsilon(v) = \int_\Omega f_s^\varepsilon(x_s)\,ds,$$

where the functionals f_s^ε are defined by (cf. (3.6))

$$f_s^\varepsilon(x) := \fint_{-r}^{r} \varepsilon^{2/3}\ddot x^2 + \varepsilon^{-2/3}W(\dot x - b_s) + a(x + a_s^\varepsilon)^2\,dt,$$

and since $a_s^\varepsilon \to 0$ in $L^2_{\mathrm{loc}}(\mathbb{R})$ for every s, they Γ-converge on K to

$$f_s(x) := \begin{cases} \dfrac{A_0}{2r}\#\big(S\dot x \cap (-r,r)\big) + a\fint_{-r}^{r} x^2\,dt & \text{if } x \in \mathscr{S}^s(-r,r) \\ +\infty & \text{otherwise,} \end{cases} \tag{6.3}$$

where $A_0 := 2\int_{-1}^{1}\sqrt{W}$ (as usual) and $\mathscr{S}^s(-r,r)$ is the class of all functions $x \in K$ which are continuous and piecewise affine on the interval $(-r,r)$ with slope b_s+1 or b_s-1 (cf. Proposition 3.3).

Now the Γ-convergence result in Theorem 3.4 can be restated without modifications.

However, a modification is needed in the proof of the upper bound inequality, due to the fact that the required uniformity of f_s in s (cf. Remark 3.5) cannot be achieved unless u is piecewise affine (notice for instance that the domain of f_s is $\mathscr{S}^s$, and $\mathscr{S}^s \cap \mathscr{S}^{s'} = \emptyset$ when $\dot u(s) \neq \dot u(s')$). The F-dense class $\mathscr{D}$ of Young measures for which we actually construct an approximation in energy by ε-blowups (cf. Definition 3.7 and following lemmas) has to be defined in a different way. To begin with, we denote by $\mathscr{S}^s_{\mathrm{per}}$ the class of all periodic and piecewise affine functions in K with slope $b_s \pm 1$, and we say that a family of functions x_s in $\mathscr{S}^s_{\mathrm{per}}$ are invariant in s when

- the functions have the same period h,
- the singular sets $S\dot x_s \cap [0,h)$ have the same cardinality n and consist of points $t_0^s = 0 < t_1^s < t_2^s < \ldots < t_n^s = h$ so that each x^s has negative slope in $[0, t_1^s]$,
- t_i^s and $x_s(t_i^s)$ do not depend on s for even index i (notice that every function in $\mathscr{S}^s_{\mathrm{per}}$ is uniquely determined by the position of even singular points and the corresponding value of the function – cf. the Figure 6.1.)

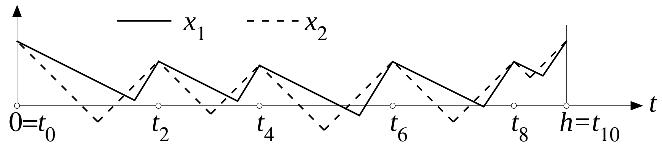

FIGURE 6.1. Example of functions x_1 and x_2 in $\mathscr{S}^{s_1}_{\mathrm{per}}$ and $\mathscr{S}^{s_2}_{\mathrm{per}}$, with $b_{s_1} = 1/2$ and $b_{s_2} = 0$, which are invariant in s.

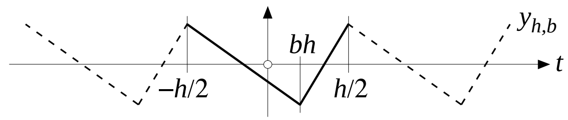

FIGURE 6.2. The asymmetric sawtooth function $y_{h,b}$.

Then $\mathscr{D}$ is defined as the class of all young measures $\nu \in \mathsf{YM}(\Omega, K)$ such that ν_s is the elementary invariant measure associated to a function $x_s \in \mathscr{S}^s_{\mathrm{per}}$ for a.e. s and there exists countably many pairwise disjoint intervals which cover almost all of Ω so that the functions x_s are invariant for s running through each such interval (cf. Definition 3.7). It is not difficult to check that $\mathscr{D}$ is F-dense in $\mathsf{YM}(\Omega, K)$ (cf. Lemma 3.8). Then one can construct an approximation in energy by ε-blowups for each $\nu \in \mathscr{D}$ following the procedure described for the proof of Theorem 3.4 in Section 3 (from Lemma 3.9 to the end of that subsection).

In the last step of our program we show (cf. Theorem 3.12) that for every s the *only* invariant measure which minimizes $\langle \mu, f_s\rangle$ is the elementary invariant measure associated with the periodic function $y_{h,b}$ given in Figure 6.2 with

$$b = b_s := \dot u(s) \quad \text{and} \quad h = h_s := (48A_0/a)^{1/3}(1-b^2)^{-2/3} \tag{6.4}$$

In particular, the Young measure ν generated by the ε-blowups of a sequence of minimizers v^ε of I^ε is given by this elementary invariant measure for a.e. s (cf. Corollary 3.13).

To prove this minimality result one can argue as in the proof of Theorem 3.12, with a slight modification in the first part, and more precisely in the proof of Lemma 3.14. From (6.3) we derive an explicit formula for the integral $\langle \mu, f_s\rangle$ when μ is the elementary invariant measure ϵ_x associated with an h-periodic function

$x \in \mathscr{S}^s_{\mathrm{per}}$ (cf. (3.12))

$$\langle \epsilon_x, f_s \rangle = \frac{A_0}{h} \#\big(S\dot{x} \cap [0,h)\big) + a \fint_0^h x^2 \, dt. \tag{6.5}$$

We may assume that $S\dot{x} \cap [0,h)$ consists of the points $t_0 = 0 < t_1 < t_2 < \ldots < t_{2n} = h$, and that the slope of x in the interval $[t_0, t_1]$ is $b_s - 1$. Now one can see that within each interval $I_i := [t_{2i}, t_{2i+2}]$ the function f is uniquely determined by the length h_i of I_i, the average m_i of x on I_i, and the difference between the average of x on $[t_{2i+1}, t_{2i+2}]$ and the average on $[t_{2i}, t_{2i+1}]$, which we denote by $d_i h_i$. Consequently the integral of x^2 on I_i can be expressed in terms of such quantities. Indeed a long but straightforward computation yields

$$\fint_{I_i} x^2 dt = h_i^2 \Big[\frac{(1-b^2)^2}{48} + \frac{1-b^2}{2} d_i^2 + \frac{4b}{3} d_i^3 - d_i^4 \Big] + m_i^2.$$

Thus we can rewrite the right-hand side of (6.5) as an average (cf. (3.35)):

$$\langle \epsilon_x, f_s \rangle = \sum_{i=1}^n \frac{h_i}{h} g(h_i, d_i, m_i)$$

with

$$g(h,m,d) := \frac{2A_0}{h} + a h^2 \Big[\frac{(1-b^2)^2}{48} + \frac{1-b^2}{2} d^2 + \frac{4b}{3} d^3 - d^4 \Big] + a m^2.$$

Finally, a direct computation shows that the minimum of g over all admissible (h,m,d), and precisely $h > 0$, $m \geq 0$, and $\frac{b+1}{2} \geq d \geq \frac{b-1}{2}$, is achieved only for $m = 0$, $d = 0$, and $h = (48A_0/a)^{1/3}(1-b^2)^{-2/3}$ (cf. (6.4)). Now we proceed as in the proof of Lemma 3.14 and show that the ϵ_x minimizes $\langle \mu, f_s \rangle$ among all (elementary) invariant measure, if and only if $m_i = d_i = 0$ and $h_i = (48A_0/a)^{1/3}(1-b^2)^{-2/3}$ for all i, that is, if and only if x is the function $y_{h,b}$ described above. The rest of the proof, namely that there are no other minimizers among all invariant measures, can be obtained as in Section 3.

6.4 Nonlocal terms, $H^{1/2}$-norm

A one-dimensional ansatz for a two-dimensional model of an austenite finely-twinned martensite phase boundary leads to a functional which involves the homogeneous $H^{1/2}$-norm rather than the L^2-norm (see [33]):

$$I^\varepsilon(v) := \int_\Omega \varepsilon^2 \ddot{v}^2 + W(\dot{v}) ds + \|v\|^2_{\mathring{H}^{1/2}}. \tag{6.6}$$

The minimization is taken over functions in $v \in H^2_{\mathrm{per}}(\Omega)$ with zero average, and $\Omega = (-1,1)$.

In the Fourier expansion $v = \sum \hat{v}(k) e^{\pi i k s}$, the homogeneous $H^{1/2}$-norm is given by

$$\|v\|^2_{\mathring{H}^{1/2}} := 2\pi \sum_{k=-\infty}^{+\infty} |k| |\hat{v}|^2 = \frac{2}{\pi} \sum_{k=-\infty}^{+\infty} \frac{1}{|k|} |\hat{\dot{v}}|^2,$$

and can be written as

$$\begin{aligned} \|v\|^2_{\mathring{H}^{1/2}} &= \int_{\Omega\times\Omega} g(s-s') \, (v(s) - v(s'))^2 ds' ds \\ &= \int_{\Omega\times\Omega} h(s-s') \, \dot{v}(s) \, \dot{v}(s') \, ds' ds, \end{aligned} \tag{6.7}$$

where we have set

$$g(t) := \frac{\pi}{4(1-\cos(\pi t))}, \quad h(t) := -\frac{1}{2\pi} \big[\ln 2 + \ln(1 - \cos(\pi t)) \big]. \tag{6.8}$$

(Notice that the second identity in (6.7) makes sense for functions of class $W^{1,1+\varepsilon}$ only).

The scheme developed in Section 3 applies to this functional, too, but again some essential modifications are required. First one easily checks that the fast scale is now $\varepsilon^{1/2}$ rather than $\varepsilon^{1/3}$ (see, e.g., [33]). Second the functional is invariant under the addition of constants and therefore it is more natural to look for the Young measures generated by the blow-up of the derivative rather than the function itself (for the latter choice it is easy to construct minimizing sequences whose Young measure on micropatterns is concentrated both at the function that is identically $+\infty$ and at the function that is identically $-\infty$). Let therefore consider the blowup

$$\mathsf{R}^\varepsilon_s \dot{v} := \dot{v}(s + \varepsilon^{1/2} t). \tag{6.9}$$

The competitors for the limit problem will be all Young measures $\nu \in \mathsf{YM}(\Omega, K)$ generated by sequences $\mathsf{R}^\varepsilon_s \dot{v}^\varepsilon$. As in Section 3, these Young measures are characterized as those ν such that ν_s is an invariant measure on K for a.e. $s \in \Omega$ (cf. Proposition 3.1).

The second step in the program developed in Section 3 consists in rewriting $I^\varepsilon(v)$ in terms of the ε-blowups. To avoid problems with integrals over unbounded domains we choose a smooth positive function ρ on $\mathbb{R}$ such that $\int \rho(t) \, dt = 1$. For $v \in H^2_{\mathrm{per}}(\Omega)$ we set

$$x_s := \varepsilon^{-1/2} v(s + \varepsilon^{1/2} t).$$

Thus

$$\dot{x}_s = \dot{v}(s + \varepsilon^{1/2} t) = \mathsf{R}^\varepsilon_s \dot{v} \quad \text{and} \quad \ddot{x}_s = \varepsilon^{1/2} \ddot{v}(s + \varepsilon^{1/2} t).$$

Setting $s'' := s + \varepsilon^{1/2} t$, $s' := s + \varepsilon^{1/2}(t + \tau)$ we get

$$\begin{aligned} &\varepsilon^2 \ddot{v}^2(s'') + W(\dot{v}(s'')) + \int_\Omega g(s'' - s') \, (v(s'') - v(s'))^2 ds' \\ &\qquad = \varepsilon \ddot{x}_s^2(t) + W(\dot{x}_s(t)) + \int_{-\varepsilon^{-1/2}}^{\varepsilon^{-1/2}} \varepsilon g(\varepsilon^{1/2}\tau) \, (x_s(t+\tau) - x_s(t))^2 d\tau. \end{aligned}$$

Integrating in $s \in \Omega$ and then taking the average over all $t \in \mathbb{R}$ with respect to the weight ρ we obtain

$$\varepsilon^{-1/2} I^\varepsilon(v) = \int_\Omega f^\varepsilon(\dot{x}_s), \tag{6.10}$$

where

$$\begin{aligned} f^\varepsilon(\dot{x}) := &\int_{\mathbb{R}} \big(\varepsilon^{1/2}\ddot{x}^2 + \varepsilon^{-1/2} W(\dot{x})\big)\rho(t)\,dt \\ &+ \int_{\mathbb{R}\times\mathbb{R}} g^\varepsilon(\tau)(x(t+\tau)-x(t))^2 \rho(t)\,dt\,d\tau, \end{aligned} \tag{6.11}$$

and

$$g^\varepsilon(\tau) := \begin{cases} \varepsilon g(\varepsilon^{1/2}\tau) & \text{if } |\tau| \le \varepsilon^{-1/2}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.12}$$

Note that f^ε is invariant under addition of constants, and then only depends on x through $\dot{x}$.

From (6.8) we have that $g(\tau) \simeq 1/2\pi\tau^2$, then the functions $g^\varepsilon(\tau)$ converge to $1/2\pi\tau^2$, and we claim that f^ε Γ-converge on K to the functional f given by

$$f(\dot{x}) := A_0 \sum_{t \in S\dot{x}} \rho(t) + \int_{\mathbb{R}\times\mathbb{R}} \frac{1}{2\pi\tau^2} (x(t+\tau)-x(t))^2 \rho(t)\,dt\,d\tau \tag{6.13}$$

for $x \in \mathscr{S}_{\mathrm{loc}}(\mathbb{R})$, and $+\infty$ elsewhere (here we view f^ε and f as functionals of $\dot{x} \in K$). To prove the claim, we proceed as for Proposition 3.3: the functionals

$$\int \big(\varepsilon^{1/2}\ddot{x}^2 + \varepsilon^{-1/2} W(\dot{x})\big)\rho \tag{6.14}$$

are equicoercive and Γ-converge on $W^{1,1}_{\mathrm{loc}}(\mathbb{R})$ – and therefore also in K – to the first term on the right-hand side of (6.13), while the double integrals on the right-hand side of (6.11) converge to the double integral on the right-hand side of (6.13) for all sequences x^ε which converge to x uniformly on $\mathbb{R}$, and are uniformly Lipschitz. Unfortunately such a convergence is not implied by convergence in $W^{1,1}_{\mathrm{loc}}(\mathbb{R})$, and one has to be more careful: given functions $\dot{x}^\varepsilon \to \dot{x}$ such that $f^\varepsilon(\dot{x}^\varepsilon)$ is bounded, we have $\dot{x}^\varepsilon \to \dot{x}$ in $L^1_{\mathrm{loc}}(\mathbb{R})$, and, modulo addition of suitable constants, $x^\varepsilon \to x$ in $W^{1,1}_{\mathrm{loc}}(\mathbb{R})$. Then a careful application of Fatou's lemma, and the fact that g and ρ are positive functions, give the lower-bound inequality. To prove the upper-bound inequality for x, it suffices to construct functions x_ε which converge uniformly to x, are uniformly Lipschitz, and satisfy the upper-bound inequality for the functionals in (6.14).

Now we can proceed as in Section 3, and prove a suitable version of Theorem 3.4. which leads to the following equivalent of Corollary 3.11: *Suppose that the functions v^ε minimize I^ε and the ε-blowups $\mathsf{R}^\varepsilon \dot{v}^\varepsilon$ generate a Young measure ν. Then, for a.e. $s \in \Omega$, the measure ν_s minimizes $\langle \mu, f\rangle$ among all invariant measures $\mu \in \mathscr{I}(K)$.*

We have not been able to carry out the last step of our program, the characterization of the minimizing measures μ. We conjecture that minimality implies that the measure is supported on the orbit of the derivative of a single periodic sawtooth function like in Figure 1.2.

Indeed, for every $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ one has (cf. (3.12))

$$\langle \epsilon_{\dot{x}}, f\rangle = \frac{A_0}{h} \#\big(S\dot{x} \cap [0,h)\big) + \int_0^h \Big[\int_{-\infty}^{\infty} \frac{1}{2\pi\tau^2}(x(t+\tau)-x(t))^2 d\tau\Big]\,dt.$$

It can be verified that the minimum of $\langle \epsilon_{\dot{x}}, f\rangle$ over all $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ and $h > 0$ is strictly positive, and hence the minimum of $\langle \mu, f\rangle$ over all $\mu \in \mathscr{I}(K)$ is also strictly positive. This shows in particular that the minima of energies I^ε in (6.6) are exactly of order $\varepsilon^{1/2}$.

As a first step in the characterization of minimizing measures μ, one should prove that the minimum of $\langle \epsilon_{\dot{x}}, f\rangle$ over all $x \in \mathscr{S}_{\mathrm{per}}(0,h)$ with h and $2n := \#(S\dot{x} \cap [0,h))$ fixed is given by the sawtooth function $y_{h/n}$ (see (3.33)). Then one could determine the optimal one by minimization over all $h > 0$. As discussed in Section 3 this is, however, only the first step in the proof of the conjecture stated above.

6.5 Concentration effects

A suitable modification of the Young measure on micropatterns which uses the energy density rather than the Lebesgue measure as background measure can also capture certain concentration effects that occur, for example, in the passage from diffuse interface models to sharp interface models. The simplest possible example is the minimization of the one-dimensional functional (already introduced in the proof of Proposition 3.3)

$$J^\varepsilon(v) = \int_0^1 \varepsilon \dot{v}^2 + \frac{1}{\varepsilon} W(v)\,ds,$$

subject to periodic boundary conditions and volume constraint $\int v = 0$. As $\varepsilon \to 0$ minimizers v^ε converge to a piecewise constant function v with two equidistantly spaced jumps. The corresponding energy density

$$e^\varepsilon = \varepsilon \dot{v}^2 + \frac{1}{\varepsilon} W(v)$$

converges (in the weak-star sense) to a measure $\mu = A_0\delta_a + A_0\delta_b$, where a and b are the positions of the jumps and $A_0 := 2\int_{-1}^1 \sqrt{W}$.

We consider now the ε-blowups $\mathsf{R}^\varepsilon_s v^\varepsilon(t) := v^\varepsilon(s+\varepsilon t)$, and define the associated measures ν^ε on $\Omega \times K$ by

$$\nu^\varepsilon = \int_\Omega (\delta_s \times \nu^\varepsilon_s)\, e^\varepsilon(s)\,ds,$$

where ν_s^ε is the Dirac mass concentrated at $\mathsf{R}_s^\varepsilon v^\varepsilon$ for every s. Then the measures ν^ε converge, up to a subsequence, to a limit measure ν on $\Omega \times K$. Since the projection of each ν_ε on Ω is the measure associated to the energy density e^ε, the projection of ν is the limiting energy measure μ, and we can thus write ν as

$$\nu = A_0 \delta_a \times \mu_a + A_0 \delta_b \times \mu_b,$$

where μ_a and μ_b care probability measures on K which capture the asymptotic behavior of minimizers v^ε near the jumps a and b resp. If we assume that the limit v of the minimizers v^ε jumps from -1 to 1 at a, and denote by x the optimal profile for the transition between the two minima of W, namely the solution of

$$2\ddot{x} = W'(x), \qquad \lim_{t \to \pm\infty} x(t) = \pm 1,$$

(which is unique up to translations), and by $e = \dot{x}^2 + W(x)$ the associated energy density, then one can prove that

$$\mu_a = \frac{1}{A_0} \int_{\mathbb{R}} \delta_{T_\tau x}\, e(t)\, dt,$$

and a similar result holds for μ_b.

Acknowledgements. The research of G.A. was supported by the Max Planck Institute for Mathematics in the Sciences in Leipzig, and by Max Planck prize and the University of Freiburg. Part of this work was done during a joint stay of both authors at the Institute for Mathematics and its Applications in Minneapolis.

In the preparation of this work we benefitted from many very helpful suggestions, and we would like to thank in particular Bernd Kirchheim for sharing his knowledge in measure theory with us, and Luc Tartar for his encouragement and inspiration. Of course responsibility for the final outcome rests solely with the authors. G.A. would like to thank people at the Max Planck Institute in Leipzig for making his stay there such a pleasant and productive one.

Bibliography

[1] Alberti, G.; Bellettini, G.: A non-local anisotropic model for phase transitions. Asymptotic behavior of rescaled energies. *European J. Appl. Math.* **9** (1998), no.3, 261–284.

[2] Allaire, G.: Homogenization and two-scale convergence. *SIAM J. Math. Anal.* **23** (1992), no. 6, 1482–1518.

[3] Allaire, G.; Briane, M.: Multiscale convergence and reiterated homogenisation. *Proc. Roy. Soc. Edinburgh Sect. A* **126** (1996), no. 2, 297–342.

[4] Balder, E.J.: A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.* **22** (1984), no. 4, 570–598.

[5] Ball, J.M.: A version of the fundamental theorem for Young measures. *PDE's and continuum models of phase transitions (Nice, 1988)*, 207–215, edited by M. Rascle et al., Lecture Notes in Phys., 344. Springer, Berlin-New York, 1989.

[6] Ball, J.M.; James, R.D.: Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.* **100** (1987), no. 1, 13–52.

[7] Barles, G.; Evans, L.C.; Souganidis, P.E.: Wavefront propagation for reaction-diffusion systems of PDE. *Duke Math. J.* **61** (1990), no. 3, 835–858.

[8] Barroso, A.C.; Fonseca, I.: Anisotropic singular perturbations. The vectorial case. *Proc. Roy. Soc. Edinburgh Sect. A* **124** (1994), no. 3, 527–571.

[9] Berliocchi, H.; Lasry, J.-M.: Intégrandes normales et measures paramétrées en calcul des variations. *Bull. Soc. Math. France* **101** (1973), 129–184.

[10] Birkhoff, G.: A note on topological groups. *Compositio Math.* **3** (1936), 427-430.

[11] Castaing, C.; Valadier, M.: *Convex analysis and measurable multifunctions.* Lecture Notes in Mathematics, Vol. 580. Springer, Berlin-New York, 1977.

[12] Chen, Y.G.; Giga, Y.; Goto S.: Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differential Geom.* **33** (1991), no. 3, 749–786.

[13] Chipot, M.; Kinderlehrer, D.: Equilibrium configurations of crystals. *Arch. Rational Mech. Anal.* **103** (1988), no. 3, 237–277.

[14] Dal Maso, G.: *An introduction to Γ-convergence.* Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston, Boston, MA, 1993.

[15] DiPerna, R.J.: Compensated compactness and general systems of conservation laws. *Trans. Amer. Math. Soc.* **292** (1985), no. 2, 383–420.

[16] Döring, W.: Mikromagnetismus. *Handbuch der Physik 18/2*, 341-347, edited by S. Flügge. Springer, Berlin-Heidelberg, 1966.

[17] E, W.: Homogenization of linear and nonlinear transport equations. *Comm. Pure Appl. Math.* **45** (1992), no. 3, 301–326.

[18] Edwards, R.E.: *Functional analysis. Theory and applications.* Holt, Rinehart and Winston, New York-Toronto-London, 1965. Reprinted by Dover Publications, New York, 1995.

[19] Evans, L.C.; Gariepy, R.F.: *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[20] Evans, L.C.; Soner, H.M.; Souganidis, P.E.: Phase transitions and generalized motion by mean curvature. *Comm. Pure Appl. Math.* **45** (1992), no. 9, 1097–1123.

[21] Evans, L.C.; Spruck, J.: Motion of level sets by mean curvature, I-IV. *J. Differential Geom.* **33** (1991), no. 3, 635–681.; *Trans. Amer. Math. Soc.* **330** (1992), no. 1, 321–332; *J. Geom. Anal.* **2** (1992), no. 2, 121–150; *J. Geom. Anal.* **5** (1995), no. 1, 77–114.

[22] Federer, H.: *Geometric measure theory.* Grundlehren der mathematischen Wissenschaften, 153. Springer, Berlin-New York, 1969. Reprinted in the series Classics in Mathematics. Springer, Berlin-Heidelberg, 1996.

[23] Gérard, P.: Microlocal defect measures. *Comm. Partial Differential Equations* **16** (1991), no. 11, 1761–1794.

[24] Gérard, P.: Mesures semi-classiques et ondes de Bloch. *Séminaire sur les Équations aux Dérivées Partielles, 1990–1991*, Exp. No. XVI, 19 pp., École Polytech., Palaiseau, 1991.

[25] Hewitt, E.; Ross, K.A.: *Abstract harmonic analysis. Vol. I. Structure of topological groups, integration theory, group representations.* Second edition. Grundlehren der Mathematischen Wissenschaften, 115. Springer, Berlin-New York, 1979.

[26] Hubert, A.; Schäfer, R.: *Magnetic domains - The analysis of magnetic microstructures.* Springer, Berlin-Heidelberg, 1998.

[27] Katsoulakis, M.A.; Souganidis, P.E.: Interacting particle systems and generalized evolution of fronts. *Arch. Rational Mech. Anal.* **127** (1994), no. 2, 133–157.

[28] Katsoulakis, M.A.; Souganidis, P.E.: Stochastic Ising models and anisotropic front propagation. *J. Statist. Phys.* **87** (1997), no. 1-2, 63–89.

[29] Kelley, J.L.: *General topology.* D. Van Nostrand Company, Toronto-New York-London, 1955. Reprinted in the series Graduate Texts in Mathematics, 27. Springer, Berlin-New York, 1975.

[30] Khachaturyan, A.: *Theory of structural transformations in solids.* J. Wiley and Sons, New York, 1983.

[31] Kinderlehrer, D.; Pedregal, P.: Characterization of Young measures generated by gradients. *Arch. Rational Mech. Anal.* **115** (1991), no. 4, 329–365.

[32] Kinderlehrer, D.; Pedregal, P.: Gradient Young measure generated by sequences in Sobolev spaces. *J. Geom. Analysis* **4** (1994), no. 1, 59–90.

[33] Kohn, R.V.; Müller, S.: Branching of twins near an austenite-twinned-martensite interface. *Philosophical Magazine A* **66** (1992), 697-715.

[34] Kristensen, J.: Lower semicontinuity in spaces of weakly differentiable functions. *Math. Ann.* **313** (1999), no. 4, 653–710.

[35] Mañé, R.: *Ergodic theory and differentiable dynamics.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), 8. Springer, Berlin-New York, 1987

[36] Modica, L.: The gradient theory of phase transitions and the minimal interface criterion. *Arch. Rational Mech. Anal.* **98** (1987), no. 2, 123–142.

[37] Modica, L.; Mortola, S.: Un esempio di Γ^--convergenza. *Boll. Un. Mat. Ital. B (5)*, **14** (1977), no. 1, 285–299.

[38] Modica, L.; Mortola, S.: Il limite nella Γ-convergenza di una famiglia di funzionali ellittici. *Boll. Un. Mat. Ital. A (5)* **14** (1977), no. 3, 526–529.

[39] Müller, S.: Singular perturbations as a selection criterion for periodic minimizing sequences. *Calc. Var. Partial Differential Equations* **1** (1993), no. 2, 169–204.

[40] Murat, F.; Tartar, L.: On the relation between Young measures and H-measures. In preparation.

[41] Nguetseng, G.: A general convergence result for a functional related to the theory of homogenization. *SIAM J. Math. Anal.* **20** (1989), no. 3, 608–623.

[42] Pedregal, P.: *Parametrized measures and variational principles.* Progress in Nonlinear Differential Equations and their Applications, 30. Birkhäuser, Basel, 1997.

[43] Pitteri, M.; Zanzotto, G.: *Continuum models for phase transitions and twinning in crystals.* Chapman and Hall, to appear.

[44] Roubiček, T.: *Relaxation in optimization theory and variational calculus.* De Gruyter Series in Nonlinear Analysis and Applications, 4. Walter de Gruyter & Co., Berlin, 1997.

[45] Rudin, W.: *Functional analysis.* Second edition. International Series in Pure and Applied Mathematics. McGraw-Hill, New York, 1991.

[46] Šverák, V.: Lower-semicontinuity of variational integrals and compensated compactness. *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, 1153–1158, Birkhäuser, Basel, 1995.

[47] Sychev, M.A.: A new approach to Young measure theory, relaxation and convergence in energy. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **16** (1999), no. 6, 773–812.

[48] Tartar, L.: Compensated compactness and applications to partial differential equations. *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, 136–212, edited by R.J. Knops. Res. Notes in Math., 39. Pitman, Boston-London, 1979.

[49] Tartar, L.: The compensated compactness method applied to systems of conservation laws. *Systems of nonlinear partial differential equations (Oxford, 1982)*, 263–285, edited by J.M. Ball. NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., 111. Reidel, Dordrecht-Boston, 1983.

[50] Tartar, L.: H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations. *Proc. Roy. Soc. Edinburgh Sect. A* **115** (1990), no. 3-4, 193–230.

[51] Tartar, L.: Beyond Young measures. *Microstructure and phase transitions in solids (Udine, 1994). Meccanica* **30** (1995), no. 5, 505–526.

[52] Tartar, L.: Homogenization, compensated compactness and H-measures. *CBMS-NSF conference (Santa Cruz 1993).* Lecture notes in preparation.

[53] Tartar, L.: Unpublished notes.

[54] Valadier, M.: Young measures. *Methods of nonconvex analysis (Varenna, 1989)*, 152–188, edited by A. Cellina. Lecture Notes in Math., 1446. Springer, Berlin, 1990.

[55] Valadier, M.: A course on Young measures. *Workshop on Measure Theory and Real Analysis (Grado, 1993)*, edited by Aljoša Volčič. *Rend. Istit. Mat. Univ. Trieste* **26** (1994), suppl., 349–394

[56] Wechsler, M.S.; Liebermann, D.S.; Read, T.A.: On the the theory of the formation of martensite. *Trans. AIME J. Metals* **197** (1953), 1503-1515.

[57] Young, L.C.: Generalized curves and the existence of an attained absolute minimum in the calculus of variations. *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie (classe III)* **30** (1937), 212-234.

[58] Young, L.C.: *Lectures on the calculus of variations and optimal control theory.* Saunders, Philadelphia-London-Toronto, 1969. Reprinted with corrections by American Mathematical Society, Providence. R.I., 2000

Giovanni Alberti
Dipartimento di Matematica
Università di Pisa
via Buonarroti 2
56127 Pisa
ITALY
E-mail: alberti@dm.unipi.it

Stefan Müller
Max Planck Institute for
Mathematics in the Sciences
Inselstr. 22-26
04103 Leipzig
GERMANY
E-mail: sm@mis.mpg.de