

Max-Planck-Institut  
für Mathematik  
in den Naturwissenschaften  
Leipzig

A new approach to variational problems  
with multiple scales

by

*Giovanni Alberti and Stefan Müller*

Preprint no.: 60

1999





# *A New Approach to Variational Problems with Multiple Scales\**

GIOVANNI ALBERTI <sup>a</sup>, STEFAN MÜLLER <sup>b</sup>

**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  $\varepsilon$ . This allows one to extract, in the limit  $\varepsilon \rightarrow 0$ , the relevant information at the macroscopic scale as well as the coarsest microscopic scale (say  $\varepsilon^\alpha$ ), and to eliminate all finer scales. To achieve this we consider rescaled functions  $R_\varepsilon^\alpha x(t) := x(s + \varepsilon^\alpha t)$  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  $R^\varepsilon x$ . As an illustration we study a one-dimensional model that describes 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 ([Do], [HS]) and the layers-within-layers pattern often observed in fine phase mixtures induced by symmetry breaking solid-solid phase transitions ([BJ], [Kh], [PZ], [WLR]).

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  $\Gamma$ -convergence has proved 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 [MM1], [MM2], [Mo], recent work includes [AB], [BF], where many further references can be found). More recently an alternate approach, mostly for evolution

---

\* 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.

<sup>a</sup> Dipartimento di Matematica, via Buonarroti 2, 56127 Pisa, Italy (e-mail: alberti@dm.unipi.it)

<sup>b</sup> Max Planck Institute for Mathematics in the Sciences, Inselstr. 22-26, 04103 Leipzig, Germany (e-mail: sm@mis.mpg.de)

problems, based on viscosity solutions has been applied very successfully to situations where a maximum principle is available (see for instance [BES], [CGG], [ES], [ESS], [KS1-2]).

Much less is known for problems with multiple small scales. Matched asymptotics expansion, renormalization or intermediate asymptotics are powerful methods to predict the limiting behaviour 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  $\Gamma$ -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  $\varepsilon$  represents a parameter that determines the smallness of the scales, we consider rescaled functions  $\mathbf{R}_s^\varepsilon v^\varepsilon(t) = \varepsilon^{-\beta} v^\varepsilon(s + \varepsilon^\alpha t)$  of the two variables  $s$  and  $t$ , where  $\varepsilon^\alpha$  represents the fast scale and  $\varepsilon^{-\beta}$  is a suitable renormalization. We then consider  $s \mapsto \mathbf{R}_s^\varepsilon v^\varepsilon$  as a map from the original domain  $\Omega$  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  $s \mapsto \mathbf{R}_s^\varepsilon v^\varepsilon$ .

The Young measure (see section 2 for precise definitions and references) is a map  $\nu$  from  $\Omega$  to the space of probability measures on  $K$ , and for each  $s \in \Omega$  the measure  $\nu(s)$ , often written as  $\nu_s$ , represents the probability that  $\mathbf{R}_s^\varepsilon v^\varepsilon$  assumes a certain value in a small neighbourhood of  $s$  in the limit  $\varepsilon \rightarrow 0$ . In terms of the original problem,  $\nu_s$  gives the probability to find a certain pattern (i.e., an element of the function space  $K$ ) on the scale  $\varepsilon^\alpha$  near the point  $s$ . We thus refer to  $\nu$  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 behaviour: minimize

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

among one-periodic functions  $v : \mathbb{R} \rightarrow \mathbb{R}$ , 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  $\pm 1$  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  $\pm 1$  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  $\pm 1$  and two corners per period. Such a result is a typical application of classical  $\Gamma$ -convergence; indeed for  $a = 0$  the  $\Gamma$ -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 conversely  $\varepsilon = 0$  but  $a > 0$  then no minimizers exist and minimizing sequences are (essentially) given by highly oscillatory sawtooth functions with slope  $\pm 1$  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).

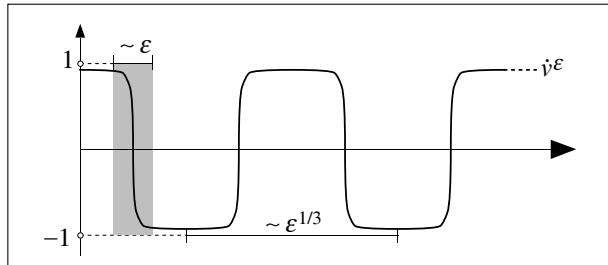
If  $\varepsilon > 0$  and  $a > 0$  the excitation of oscillations due to  $a > 0$  and their penalization due to  $\varepsilon > 0$  lead to the emergence of a new structure.

**Theorem 1.1.** ([Mu]) *Suppose that  $a$  is constant and strictly positive. Then, for  $\varepsilon$  positive and sufficiently small, all minimizers of  $I^\varepsilon$  among one-periodic functions have minimal period*

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

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

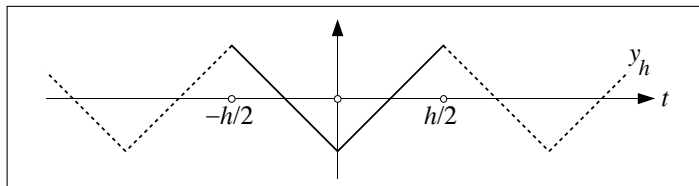
The derivatives of minimizers exhibit indeed a structure with two fast scales: transition layers of order  $\varepsilon$  are spaced periodically with the period  $P^\varepsilon \sim \varepsilon^{1/3}$  (see Fig. 1).



**Fig.1.** Two scale structure of minimizers

This behaviour was predicted by Tartar [Ta3] on the basis of matched asymptotic expansions. It can equivalently be guessed by a formal application of  $\Gamma$ -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  $a \in L^1$  and  $a > 0$  a.e., let  $v^\varepsilon$  be a sequence of minimizers of  $I^\varepsilon$  and let  $R_s^\varepsilon v^\varepsilon$  and the Young measure  $\nu$  be as above. Then for a.e.  $s$  the measure  $\nu_s$  is supported on the set of all translations of the sawtooth function  $y_h$  with slope  $\pm 1$  and period  $h = L_0(a(s))^{-1/3}$  (see Fig. 2).*



**Fig.2.** The sawtooth function  $y_h$

Thus the Young measure on patterns  $\nu$  provides a useful tool to localize the result in Theorem 1.1. More importantly, it gives a precise meaning to the statement that  $v^\varepsilon$  is locally nearly periodic with a period  $L_0(\varepsilon/a(s))^{-1/3}$  which depends on the point  $s$ . In addition to this, the main advantage of the new object  $\nu$  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  $H_{\text{per}}^2$  the Sobolev space of functions on the interval  $(0, 1)$  whose periodic extension belongs to  $H_{\text{loc}}^2(\mathbb{R})$ , and by  $\mathcal{S}_{\text{per}}$  the space of functions on  $(0, 1)$  whose periodic extension are (continuous) sawtooth functions with slope  $\pm 1$ . Consider the functionals

$$J^\varepsilon(v) := \int_0^1 \varepsilon \dot{v}^2 + \frac{1}{\varepsilon} W(\dot{v}) \quad \text{if } v \in H_{\text{per}}^2,$$

and

$$J(v) := \frac{A_0}{2} \int_0^1 |\dot{v}| = A_0 \#(S\dot{v} \cap [0, 1)) \quad \text{if } v \in \mathcal{S}_{\text{per}},$$

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

We know that ([MM1], [MM2])  $J^\varepsilon$ , extended to  $+\infty$  on  $W^{1,1} \setminus H_{\text{per}}^2$ ,  $\Gamma$ -converges in the  $W^{1,1}$  topology to  $J$ , extended to  $+\infty$  on  $W^{1,1} \setminus \mathcal{S}_{\text{per}}$ . Thus it is plausible to replace

$$I^\varepsilon(v) = \varepsilon J^\varepsilon(v) + \int_0^1 a v^2 ds \quad (1.2)$$

by

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

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  $0 \leq s_1 < s_2 < \dots < s_{2N} < 1$ . 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

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

while the energy of minimizers is

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

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

$$\varepsilon^{-2/3} I^\varepsilon(v) = \int_0^1 f^\varepsilon(R_s^\varepsilon v) ds , \quad (1.5)$$

where

$$f^\varepsilon(x) := \int_{-r}^r [\varepsilon^{2/3} \dot{x}^2 + \varepsilon^{-2/3} W(\dot{x}) + a x^2] dt$$

for a given positive  $r$ . Now we have that  $f^\varepsilon$   $\Gamma$ -converge to  $f$ , where

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

if  $x$  agrees with a sawtooth function on  $(-r, r)$ , and is  $+\infty$  otherwise. We then essentially have to show that the  $\Gamma$ -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  $(v^\varepsilon)$  is given by

$$\int_0^1 \langle \nu_s, f \rangle ds , \quad (1.6)$$

where  $\nu$  is the Young measure (on patterns) generated by  $R^\varepsilon v^\varepsilon$ .

To determine the minimizing Young measure we need to know which Young measures arise as limits of  $R^\varepsilon v^\varepsilon$ . This is not obvious since for finite  $\varepsilon$  the blowups  $R_s^\varepsilon v^\varepsilon$  at different points  $s$

are not independent. In the limit, the measures  $\nu_s$ , however, become independent and the only restriction is that  $\nu_s$  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 in the case  $a$  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  $v^\varepsilon$  converging weakly to 0 in  $L^2_{\text{loc}}(\mathbb{R}^N)$ , Tartar [Ta4] and Gérard [Ge1] introduced independently a measure on  $\mathbb{R}^N \times S^{N-1}$  (called the  $H$ -measure or microlocal defect measure, respectively) that measures 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  $\delta(\varepsilon) \rightarrow 0$  ([Ge2]). 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  $v^\varepsilon$  (the expression may, however, involve pseudodifferential operators) and hence have no direct applications to the study of  $I^\varepsilon$ .

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 [Ta5], [Ta6].

Our work was inspired by the concept of two-scale convergence, although our approach is ultimately rather different. Two-scale convergence was introduced by [Ng] and employed by a number of researchers, in particular Allaire ([A1], [A1B], see also [E]). The main idea is to recover additional structure in a weakly converging sequence  $v^\varepsilon$  by using test functions of the form  $\phi(s, s/\varepsilon^\alpha)$ , where  $\phi$  is periodic in the last variable.

If  $v^\varepsilon$  is of the form  $v^\varepsilon(s) = v_0(s) + v_1(s, s/\varepsilon^\alpha) + o(1)$ , where  $v_1$  is  $P$ -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  $\eta_2$  is  $P$ -periodic and has vanishing average on the period, then

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

Thus both the weak limit  $v_0$  and the oscillatory term on the scale  $\varepsilon^\alpha$  related to  $v_1$  can be retrieved.

If, however, the period or even the phase of the oscillatory part is not exactly known, then it can usually not be retrieved. Consider for instance

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

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

$$\int_0^1 \phi\left(s, \frac{s}{\varepsilon^\alpha}\right) v^\varepsilon(s) ds \rightarrow 0 .$$

Since we do not know the precise period of the minimizers of  $I^\varepsilon$  (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  $\Gamma$ -convergence. We follow mainly [BL] and [Ba], the main 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  $\Gamma$ -limit of the functionals  $I^\varepsilon$  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).

**Acknowledgements:** 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.

## 2. Young Measures which take values in a metric space

Young measures are maps from a measure space  $\Omega$  to probability measures on another space  $K$ . They arise naturally as limits of (usually rapidly oscillating) sequences of maps from  $\Omega$  to  $K$ , and provide a good framework for existence of minimizers and optimal controls. Since the pioneering work of L.C. Young [Yo1-2] 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 [BL] and Balder [Bal], the recent reviews of Valadier [Va1-2] and the book by Roubiřek [Ro]. A closely related but slightly different approach was pursued by Sychev [Sy], who emphasizes the view of Young measures as (strongly) measurable maps into a suitable metric space and the use of selection theorems rather than the  $L^1$ - $L^\infty$  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 ([Ta1-2], [DP], [MT]), and with fine phase mixtures that arise in phase transitions modelled by nonconvex variational problems ([BJ], [CK], [KP1-2], [Sv], [Pe], [Kr]).

Our approach is inspired by [Ba] (see also the comments in [BL], 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 [Mu]) that involve maps from  $\Omega$  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  $\sigma$ -additive function on the  $\sigma$ -algebra of Borel sets. Unless stated differently, measurability always means Borel measurability.

In the rest of this section  $\Omega$  is a locally compact separable and metrizable space, endowed with a finite measure  $\lambda$  (however, most of the results can be extended with some care to  $\sigma$ -finite measures). We often suppress explicit reference to  $\lambda$ . 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  $\text{Meas}(\Omega, K)$  of all measurable maps from  $\Omega$  to  $K$ , the Banach space  $C(K)$  of all continuous real functions on  $K$ , and the space  $\mathcal{M}(K)$  of finite real Borel measures on  $K$ ;  $\mathcal{M}(K)$  is identified with the dual of  $C(K)$  by



the duality pairing  $\langle \mu, g \rangle := \int_K g d\mu$  for  $\mu \in \mathcal{M}(K)$ ,  $g \in C(K)$ , and is always endowed with the corresponding weak\* topology. For every  $x \in K$ ,  $\delta_x$  is the Dirac mass at  $x$ ;  $\mathcal{P}(K)$  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  $\Omega$ , and  $x$  a point in  $K$ ,  $\mu$  is a measure in  $\mathcal{M}(K)$ ,  $k$  is a positive integer,  $g$ ,  $h$ , and  $f$  are real functions on  $K$ , on  $\Omega$ , and on  $\Omega \times K$ , respectively;  $\phi$  is a map from  $\Omega$  to  $C(K)$  and  $\nu$  a map from  $\Omega$  to  $\mathcal{M}(K)$ ; we often use the notation  $f_s$  and  $\nu_s$  to denote the function  $f(s, \cdot)$  and the measure  $\nu(s)$  respectively.

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

*Remarks.* Since  $C(K)$  is a separable Banach space, and  $\Omega$  is endowed with a  $\sigma$ -finite measure  $\lambda$ , then the Banach space  $L^1(\Omega, C(K))$  is separable, while  $L_w^\infty(\Omega, \mathcal{M}(K))$  is never separable unless  $\lambda$  is purely atomic and  $K$  is a finite set.

By definition, a map  $\nu : \Omega \rightarrow \mathcal{M}(K)$  is weak\* measurable if the pre-image of every set in the Borel  $\sigma$ -algebra generated by the weak\* topology of  $\mathcal{M}(K)$  is a Borel subset of  $\Omega$ . Therefore the map  $\nu$  is weak\* 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  $\mathcal{M}(K)$  is not separable, there are many weak\* measurable maps that are not strongly measurable; a typical example is the map which takes every  $s$  in an interval  $I$  into  $\delta_s \in \mathcal{M}(I)$ . Indeed the  $\sigma$ -algebra generated by the weak\* topology and the one generated by strong topology do not agree (the strong topology itself has cardinality strictly larger than the  $\sigma$ -algebra generated by the weak\* topology).

The space  $L_w^\infty(\Omega, \mathcal{M}(K))$  is isometrically isomorphic to the dual of  $L^1(\Omega, C(K))$  via the duality pairing (see [Ed], section 8.18.1)

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

where  $\nu \in L_w^\infty(\Omega, \mathcal{M}(K))$  and  $\phi \in L^1(\Omega, C(K))$ . In the following we shall refer to the weak\* topology of  $L_w^\infty(\Omega, \mathcal{M}(K))$  as the topology induced by this duality pairing. Since  $L^1(\Omega, C(K))$  is a separable Banach space, every closed ball in  $L_w^\infty(\Omega, \mathcal{M}(K))$  endowed with the weak\* 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  $s \in \Omega$  in  $h(s) \cdot g \in C(K)$  belongs to  $L^1(\Omega, C(K))$ , 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  $L^1(\Omega, C(K))$ . Hence a bounded sequence  $(\nu^k)$  in  $L_w^\infty(\Omega, \mathcal{M}(K))$  weak\* converges to  $\nu$  if and only if

$$\int_\Omega \langle \nu_s^k, g \rangle h(s) ds \rightarrow \int_\Omega \langle \nu_s, g \rangle h(s) ds \quad (2.1)$$

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

Furthermore, on every bounded subset of  $L_w^\infty(\Omega, \mathcal{M}(K))$  the weak\* topology is induced by the following norm:

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

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.** (Young measures)

A  $K$ -valued Young measure on  $\Omega$  is a map  $\nu \in L_w^\infty(\Omega, \mathcal{M}(K))$  such that  $\nu_s$  is a probability measure for a.e.  $s \in \Omega$ . The set of all Young measures is denoted by  $\text{YM}(\Omega, K)$ ; it is always endowed with the weak\* topology of  $L_w^\infty(\Omega, \mathcal{M}(K))$ , and hence metrized by the norm  $\Phi$  in (2.2).

The elementary Young measure associated to a measurable map  $u : \Omega \rightarrow K$  is the map  $\underline{\delta}_u$  given by

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

The set of all elementary Young measures is denoted by  $\text{EYM}(\Omega, K)$ .

We say that a sequence of measurable maps  $u^k : \Omega \rightarrow K$  generates the Young measure  $\nu$ , if the corresponding elementary Young measures  $\underline{\delta}_{u^k}$  converge to  $\nu$  in the weak\* topology of  $L_w^\infty(\Omega, \mathcal{M}(K))$ .

*Remarks.* The map  $\underline{\delta}_u$  is weak\* measurable if and only if  $u$  is measurable, and thus  $\text{EYM}(\Omega, K)$  is exactly the set of all  $\nu \in L_w^\infty(\Omega, \mathcal{M}(K))$  such that  $\nu_s$  is a Dirac mass for a.e.  $s \in \Omega$ .

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

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

**Theorem 2.3.** Assume that the measure  $\lambda$  is non-atomic. Then

- (i)  $\text{YM}(\Omega, K)$  is a weak\* compact, convex and metrizable set in  $L_w^\infty(\Omega, \mathcal{M}(K))$ ;
- (ii)  $\text{EYM}(\Omega, K)$  is the set of all extreme points of  $\text{YM}(\Omega, K)$ ;
- (iii)  $\text{EYM}(\Omega, K)$  is weak\* dense in  $\text{YM}(\Omega, K)$ .

*Proof.* These three statements are given in [BL] as Proposition 1, §II.2, p. 144, Proposition 3, §II.2, p. 146, Proposition 4, §II.2, p. 148, respectively.  $\square$

*Remarks.* Statement (i) holds when  $\lambda$  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  $\lambda$  has atoms (namely, points with positive measure) it can be decomposed in a unique way as the sum of a non-atomic measure  $\lambda_n$  and a purely atomic measure  $\lambda_a$  (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  $\text{YM}(\Omega, K)$  are the Young measures  $\nu$  such that  $\nu_s$  is a Dirac mass for  $\lambda_n$  a.e.  $s$ , and the weak\* closure of  $\text{EYM}(\Omega, K)$  is the set of all  $\nu \in \text{YM}(\Omega, K)$  such that  $\nu_s$  is a Dirac mass for  $\lambda_a$  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  $u^k : \Omega \rightarrow K$  there exists a subsequence (not relabelled) which generates a Young measure  $\nu$ . Moreover  $\nu$  has the following properties.

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

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

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

$$u(s) := \int_K \text{Id} d\nu_s . \quad (2.4)$$

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

Regarding statement (iv), it is necessary to embed  $K$  in a linear structure in order to define the average (or the 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  $\text{Id}$  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  $\nu$  follows from the compactness of  $\text{YM}(\Omega, K)$  (Theorem 2.3(i)).

To prove (ii), notice that the map  $s \mapsto f_s$  belongs to  $L^1(\Omega, C(K))$  (cf. [BL], Remark 5, §I.1, p. 135), and then (2.3) follows immediately from the definition of weak\* convergence in  $L_w^\infty(\Omega, \mathcal{M}(K))$ .

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

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

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

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

$$\int_\Omega \langle \nu_s, f_s \rangle ds = 0 . \quad (2.5)$$

Since  $f$  is non-negative, (2.5) implies that for a.e.  $s \in \Omega$  the measure  $\nu_s$  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  $h \in L^1(\Omega)$ ,  $\Lambda \in E^*$ .  $\square$

Before discussing functionals on  $\text{YM}(\Omega, K)$  we add some elementary remarks which will be useful in the following.

*Remark 2.5.* If the maps  $u^k$  generate the Young measure  $\nu$  on  $\Omega$ , then they generate the same Young measure on every Borel subset of  $\Omega$ , that is,  $\underline{\nu}_{u^k}$  weak\* converges to  $\nu$  in  $L^\infty(A, \mathcal{M}(K))$  for every  $A \subset \Omega$ . We deduce the following locality property of Young measures: if two sequences  $(u_1^k)$  and  $(u_2^k)$  generate the Young measures  $\nu_1$  and  $\nu_2$ , respectively, and  $u_1^k$  and  $u_2^k$  agree on a Borel set  $A \subset \Omega$  for  $k$  sufficiently large, then  $\nu_1$  and  $\nu_2$  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  $s \mapsto d(u_1^k(s), u_2^k(s))$  converge in measure to 0 as  $k \rightarrow +\infty$ . One easily checks, using

the convergence criterion in Remark 2.1, that asymptotically equivalent sequences generate the same Young measure.

**Lemma 2.7.** *Let  $\Omega \subset \mathbb{R}^n$  be endowed with the Lebesgue measure. Consider a sequence of maps  $u^k$ , defined on a given neighbourhood of  $\Omega$ , which generate a Young measure  $\nu$  and a sequence of vectors  $\tau_k \in \mathbb{R}^n$  such that  $\tau_k \rightarrow 0$ . Then the translated maps  $u^k(\cdot - \tau_k)$  also generate the Young measure  $\nu$ .*

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

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

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  $\|h(\cdot + \tau_k) - h(\cdot)\|_1 \cdot \|g(\cdot)\|_\infty$ , and since the first term tends to 0 for every  $h \in L^1(\mathbb{R}^n)$ , we obtain

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

By Remark 2.1 this suffices to prove the assertion.  $\square$

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

**Definition 2.8.** (Relaxation)

*Let  $X$  be a metric space and let  $F : X \rightarrow [0, +\infty]$ . 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 following formula:*

$$\overline{F}(x) = \inf \left\{ \liminf_{k \rightarrow \infty} F(x^k) : x^k \rightarrow x \right\} . \quad (2.7)$$

**Definition 2.9.** ( $\Gamma$ -convergence and continuous convergence)

*Let  $X$  be a metric space. A sequence of functions  $F^\varepsilon : X \rightarrow [0, +\infty]$  are equicoercive on  $X$  if every sequence  $(x^\varepsilon)$  such that  $F^\varepsilon(x^\varepsilon)$  is bounded is pre-compact in  $X$ .*

*We say that the functions  $F^\varepsilon$   $\Gamma$ -converge to  $F$  on  $X$ , and we write  $F^\varepsilon \xrightarrow{\Gamma} F$ , if the following properties are fulfilled:*

Lower bound inequality:  $\forall x \in X, \forall (x^\varepsilon) \subset X$  s.t.  $x^\varepsilon \rightarrow x$ ,  $\liminf F^\varepsilon(x^\varepsilon) \geq F(x)$ ;

Upper bound inequality:  $\forall x \in X, \exists (x^\varepsilon) \subset X$  s.t.  $x^\varepsilon \rightarrow x$  and  $\limsup F^\varepsilon(x^\varepsilon) \leq F(x)$ .

*The functions  $F^\varepsilon$  converge continuously to  $F$  on  $X$  if  $F^\varepsilon(x^\varepsilon) \rightarrow F(x)$  whenever  $x^\varepsilon \rightarrow x$ .*

Here and in the following we use the term ‘‘sequence’’ also to denote families (of points of  $X$ ) labelled by the continuous parameter  $\varepsilon$ , which tends to 0. A subsequence of  $(x^\varepsilon)$  is any sequence  $(x^{\varepsilon_n})$  such that  $\varepsilon_n \rightarrow 0$  as  $n \rightarrow +\infty$ , and we say that  $(x^\varepsilon)$  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  $(x^\varepsilon)$ ” to mean a sequence defined only for countably many  $\varepsilon = \varepsilon_n$  such that  $\varepsilon_n \rightarrow 0$ .

*Remark 2.10.* Given a lower semicontinuous function  $F : X \rightarrow [0, +\infty]$ , we say that a set  $\mathcal{D}$  is  $F$ -dense in  $X$  if for every  $x \in X$  with  $F(x) < +\infty$  there exists a sequence of points  $x^k \in \mathcal{D}$  such that  $x^k \rightarrow x$  and  $F(x^k) \rightarrow F(x)$ . 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  $\mathcal{D}$  and every  $\delta > 0$  we can find a sequence  $(x^\varepsilon)$  such that

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

**Proposition 2.11.** (see [DM], chapters 6 and 7)

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

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

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

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

$$\mathcal{F}_f(\nu) := \begin{cases} \int_{\Omega} f(s, u(s)) ds & \text{if } \nu \in \text{EYM}(\Omega, K) \text{ and } \nu = \underline{\delta}_u, \\ +\infty & \text{elsewhere in } \text{YM}(\Omega, K), \end{cases} \quad (2.9)$$

and

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

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

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

Theorem 2.12 below shows that the relaxation or the  $\Gamma$ -convergence of functionals of the first type always lead to functionals of the second type. We recall that the set  $\text{YM}(\Omega, K)$  is always endowed with the weak\* topology of  $L_w^\infty(\Omega, \mathcal{M}(K))$ , which makes it compact and metrizable.

**Theorem 2.12.** *If the measure  $\lambda$  is non-atomic the following statements hold.*

- (i) *If the integrand  $f$  satisfies  $f(s, x) \leq h(s)$  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  $\text{YM}(\Omega, K)$ .*

- (ii) If  $f_s$  is lower semicontinuous on  $K$  for a.e.  $s \in \Omega$  then  $F_f$  is lower semicontinuous on  $\text{YM}(\Omega, K)$ .
- (iii) The relaxation of  $\mathcal{F}_f$  and  $F_f$  on  $\text{YM}(\Omega, K)$  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  $f^\varepsilon$  satisfy  $f_s^\varepsilon \xrightarrow{\Gamma} f_s$  on  $K$  for a.e.  $s \in \Omega$ , and that the envelope functions  $\text{E}f^\varepsilon$  defined by

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

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

*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  $f_s^\varepsilon \xrightarrow{\Gamma} f_s$  for almost every  $s \in \Omega$  is quite strong, and far from necessary. Indeed the  $\Gamma$ -convergence of the functionals may occur even with a more complicate asymptotic behaviour 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  $\text{E}f^\varepsilon$  in (2.12) are not equi-integrable on  $\Omega$ , some concentration effect occurs, and the  $\Gamma$ -convergence results may not hold. In particular, if  $\|\text{E}f^\varepsilon\|_1 \rightarrow +\infty$  then  $\mathcal{F}_{f^\varepsilon}$  and  $F_{f^\varepsilon}$   $\Gamma$ -converge to the constant functional  $+\infty$ . On the other hand, if there exist sets  $B^\varepsilon \subset \Omega$  such that  $|B^\varepsilon| \rightarrow 0$ , the restrictions of  $\text{E}f^\varepsilon$  to the complements of  $B^\varepsilon$  are equi-integrable on  $\Omega$ , and  $\int_{B^\varepsilon} \text{E}f^\varepsilon ds$  converge to some constant  $c$ , then both  $\mathcal{F}_{f^\varepsilon}$  and  $F_{f^\varepsilon}$   $\Gamma$ -converge to  $F_f + c$  (this generalization of statement (iv) can be proved by suitably modifying the proof below). However, both  $\mathcal{F}_{f^\varepsilon}$  and  $F_{f^\varepsilon}$  verify the lower bound inequality without any assumption on  $\text{E}f^\varepsilon$ .

Finally we notice that the functions  $\text{E}f^\varepsilon$  are  $\lambda$ -measurable (see for instance [CV], 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  $L^1(\Omega, C(K))$  (cf. [BL], Remark 5, §I.1, p. 135). Hence  $F_f$  belongs to the pre-dual of  $L_w^\infty(\Omega, \mathcal{M}(K))$ , and is therefore weak\* continuous on  $\text{YM}(\Omega, K)$ .

Assertion (ii) is contained in [BL], Proposition 3, §II.1, p. 152, and Theorem 2, §I.3, p. 138. Alternatively one can use (i) and the approximation from below established in Lemmas 2.13 and 2.14 below.  $\square$

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  $\nu$  with finite energy  $F_f(\nu)$ .

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

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

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

- (i) for every  $s$ ,  $f_s^k$  is  $k$ -Lipschitz on  $K$  and  $0 \leq f_s^k \leq k$ ;
- (ii) for every  $s$ ,  $f_s^k$  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  $s \mapsto \inf \{g(s, x) : x \in K\}$  is  $\lambda$ -measurable (cf. [CV], Lemma III.39) and thus it agrees a.e. in  $\Omega$  with a Borel function. Hence for every  $x \in K$  we can find a negligible Borel set  $N_x^k \subset \Omega$  such that the map  $s \mapsto f^k(s, x)$  is Borel measurable on  $\Omega \setminus N_x^k$  (cf. (2.13)). Now we take  $N^k$  as the union of all  $N_x^k$  as  $x$  ranges in a countable dense subset  $\mathcal{D}$  of  $K$ , thus  $N^k$  is a negligible Borel set,  $f^k$  is Borel measurable in  $(\Omega \setminus N^k) \times \mathcal{D}$ , and then also on  $(\Omega \setminus N^k) \times K$  because  $f^k$  is continuous in the second variable and  $\mathcal{D}$  is dense in  $K$ . Finally we take  $N := \cup_k N^k$ .  $\square$

**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 measurable set  $N \subset \Omega$  such that  $\bar{f}$  is 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 measurable on  $(\Omega \setminus N) \times K$ , and then the same holds for  $\bar{f}$  by statement (ii) of Lemma 2.13.  $\square$

**Proposition 2.15.** ([BL], Proposition 1, §II.2, p. 144)

*Assume that  $\lambda$  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*

$$M_f := \{\nu \in \text{YM}(\Omega, K) : F_f(\nu) \leq 1\} . \quad (2.14)$$

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

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

*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  $\nu$ . Then  $F_f(\nu^k) \leq F_f(\nu)$  for every  $k$ , and since  $F_f$  is lower semicontinuous, we deduce that  $F_f(\nu^k)$  converge to  $F_f(\nu)$ .  $\square$

*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  $F_{f^\varepsilon}$ , and then the upper bound inequality for the functionals  $\mathcal{F}_{f^\varepsilon}$  (recall that  $F_{f^\varepsilon} \leq \mathcal{F}_{f^\varepsilon}$ ).

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

We fix now an integer  $k$  and we take  $f^k$  as in (2.13) (we set it equal to 0 in the set  $N$  given in Lemma 2.13(iii) to make it Borel measurable). Since  $f_s \geq f_s^k$  on  $K$  (see Lemma 2.13(ii)), there holds  $f_s > f_s^k - \frac{1}{k}h(s)$ , and since  $f_s^k$  is continuous on  $K$  (Lemma 2.13(i)), by the previous remark for  $\varepsilon$  sufficiently small there holds

$$f_s^\varepsilon \geq f_s^k - \frac{1}{k} . \quad (2.15)$$

Consider a maps  $\nu^\varepsilon$  which converge to  $\nu$  in  $\text{YM}(\Omega, K)$ . Then (2.15) yields

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} F_{f^\varepsilon}(\nu^\varepsilon) &= \liminf_{\varepsilon \rightarrow 0} \int_{\Omega} \langle \nu_s^\varepsilon, f_s^\varepsilon \rangle ds \\ &\geq \liminf_{\varepsilon \rightarrow 0} \int_{\Omega} \langle \nu_s^\varepsilon, f_s^k - \frac{1}{k} \rangle ds = \int_{\Omega} \langle \nu_s, f_s^k \rangle ds - \frac{1}{k} \lambda(\Omega) , \end{aligned} \quad (2.16)$$

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

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

We consider now the upper bound inequality. Since  $\text{EYM}(\Omega, K)$  is  $F_f$ -dense in  $\text{YM}(\Omega, K)$  (Theorem 2.16) and each  $\mathcal{F}_{f^\varepsilon}$  is finite only on  $\text{EYM}(\Omega, K)$ , 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  $u \in \text{Meas}(\Omega, K)$  we will exhibit a sequence of maps  $u^\varepsilon$  which converge to  $u$  a.e. in  $\Omega$  and satisfy

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

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

$$\lim_{\varepsilon \rightarrow 0} x_s^\varepsilon = u(s) \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} f^\varepsilon(s, x_s^\varepsilon) = f(s, u(s)) . \quad (2.18)$$

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

$$f^\varepsilon(s, y_s^\varepsilon) \leq \text{E}f^\varepsilon(s) + 1 . \quad (2.19)$$

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

$$u^\varepsilon(s) := \begin{cases} x_s^\varepsilon & \text{if } f^\varepsilon(s, x_s^\varepsilon) \leq f(s, u(s)) + 1, \\ y_s^\varepsilon & \text{otherwise.} \end{cases} \quad (2.20)$$

From (2.18) we deduce that for a.e.  $s \in \Omega$  there holds  $u^\varepsilon(s) = x_s^\varepsilon$  for  $\varepsilon$  small enough, and thus  $u^\varepsilon(s) \rightarrow u(s)$  and  $f^\varepsilon(s, u^\varepsilon(s)) \rightarrow f(s, u(s))$ . We claim that the functions  $s \mapsto f^\varepsilon(s, u^\varepsilon(s))$  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)

$$f^\varepsilon(s, u^\varepsilon(s)) \leq \text{E}f^\varepsilon(s) + f(s, u(s)) + 1 ,$$

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

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

In the first case, we modify  $\text{E}f^\varepsilon$  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  $f^\varepsilon(s, y) \leq \text{E}f^\varepsilon(s) + 1$  is Borel measurable and the projection on  $\Omega$  is equal to  $\Omega$ , and we can apply the Von Neumann-Aumann measurable selection theorem (see [CV], Theorem III.22) to find a  $\lambda$ -measurable selection  $s \mapsto y_s^\varepsilon$  (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

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



the function  $h^\varepsilon$  is  $\lambda$ -measurable (see [CV], 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

$$|f^\varepsilon(s, x) - f(s, u(s))| + d(x, u(s)) \leq h^\varepsilon(s) + \varepsilon \quad (2.21)$$

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

### 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  $\Omega$  is a bounded open interval endowed with Lebesgue measure, the letter  $s$  denotes the (slow) variable in  $\Omega$  and  $v$  is a real function on  $\Omega$ , periodically extended out of  $\Omega$ ; 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  $x : \mathbb{R} \rightarrow [-\infty, +\infty]$  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\* 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  $L_{\text{loc}}^p(\mathbb{R})$  embeds continuously in  $K$  for  $1 \leq p \leq \infty$ . See section 5 For more details and precise statements.

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

$$\langle \epsilon_x, g \rangle = \int_0^h g(T_\tau x) d\tau \quad (3.1)$$

for every positive Borel function  $g$  on  $K$  (here  $h$  is the period of  $x$ );  $\epsilon_x$  is the unique invariant probability measure supported on  $\mathcal{O}(x)$ , 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 denote by  $H_{\text{per}}^2(I)$  (resp.,  $W_{\text{per}}^{k,p}(I)$ ) the Sobolev space of all real functions on  $I$ , extended to  $\mathbb{R}$  by periodicity, which belong to  $H_{\text{loc}}^2(\mathbb{R})$  (resp., to  $W_{\text{loc}}^{k,p}(\mathbb{R})$ ), and by  $\mathcal{S}(I)$  the class of all functions  $x \in K$  which are continuous and piecewise affine on the interval  $I$  with slope  $\pm 1$  only (sawtooth functions);  $Sx$  is the set of all points in where  $x$  is not differentiable, and thus the points in  $S \cap I$  are “corners” of  $x$ ;  $\mathcal{S}_{\text{per}}(I)$  is the

class of all real functions on  $I$  extended to  $\mathbb{R}$  by periodicity and of class  $\mathcal{S}$  on every bounded interval. The space  $\mathcal{S}(I)$  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  $\pm 1$  only; if  $S\dot{x} \cap I$  consists of the points  $t_1 < t_2 < \dots < t_N$ , then  $\pm \ddot{x} = \sum (-1)^i 2\delta_{t_i}$ , 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  $\varepsilon$ -blowup of  $v$  at  $s$  is defined by

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

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

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

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

where  $v \in H_{\text{per}}^2(\Omega)$ ,  $a \in L^1(\Omega)$  is strictly positive a.e., and  $W$  is a continuous non-negative function on  $\mathbb{R}$  which vanishes at  $\pm 1$  only and has growth at least linear at infinity. This goal is achieved in several steps:

- Step 1.** Identify the class of all Young measures  $\nu \in \mathbf{YM}(\Omega, K)$  which are generated by sequences of  $\varepsilon$ -blowups of functions  $v^\varepsilon$  (Proposition 3.1 and Remark 3.2).
- Step 2.** Write the rescaled functionals  $\varepsilon^{-2/3} I^\varepsilon(v)$  as  $\int_{\Omega} f_s^\varepsilon(\mathbf{R}_s^\varepsilon v) ds$  for suitable functionals  $f_s^\varepsilon$  on  $K$  (cf. (3.6) and (3.7)).
- Step 3.** Identify the  $\Gamma$ -limit  $f_s$  of  $f_s^\varepsilon$  as  $\varepsilon \rightarrow 0$  for a.e.  $s \in \Omega$  (Proposition 3.3).
- Step 4.** Prove that the  $\Gamma$ -limit of the rescaled functionals  $\varepsilon^{-2/3} I^\varepsilon$ , viewed as functionals of the elementary Young measures associated with  $\varepsilon$ -blowups of functions, is given by  $F(\nu) := \int_{\Omega} \langle \nu_s, f_s \rangle ds$  for all Young measures  $\nu$  described in Step 1 (Theorem 3.4).
- Step 5.** Let  $\nu$  be a Young measure generated by  $\varepsilon$ -blowups of minimizers  $v^\varepsilon$  of  $I^\varepsilon$ , and use Step 4 to show that  $\nu$  minimizes  $F$  (Corollary 3.11); use this fact to identify  $\nu$  (Theorem 3.12 and Corollary 3.13).

#### *Step 1. Admissible Young measures*

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

**Proposition 3.1.** *Let  $\nu \in \mathbf{YM}(\Omega, K)$  be a Young measure generated by the  $\varepsilon$ -blowups  $\mathbf{R}^\varepsilon v^\varepsilon$  of a countable sequence  $(v^\varepsilon) \subset H^2(\Omega)$ . Then  $\nu_s$  is an invariant measure on  $K$  of a.e.  $s \in \Omega$ .*

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

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

Since the functions  $u^\varepsilon$  generate the Young measure  $\nu$ , 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  $u^\varepsilon$ , and thus identity (3.4) yields  $T^\# \nu = \nu$ . Therefore we can find a negligible set  $N_\tau \subset \mathbb{R}$  such that

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

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

*Remark 3.2.* The converse of Proposition 3.1 is also true: for every  $\nu \in \text{YM}(\Omega, K)$  such that  $\nu_s \in \mathcal{A}(K)$  for a.e.  $s \in \Omega$  we can find a sequence  $(v^\varepsilon)$  such that  $\mathbb{R}^\varepsilon v^\varepsilon$  generate  $\nu$ . 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  $\mathbb{R}^\varepsilon v$*

We extend  $a$  by periodicity out of  $\Omega$  and set  $a_s^\varepsilon(t) := a(s + \varepsilon^{1/3}t)$  for every  $s$  and  $t$ . We fix a function  $v \in H_{\text{per}}^2(\Omega)$  and set  $x_s := \mathbb{R}_s^\varepsilon v$  for every  $s \in \Omega$ . Thus

$$x_s(t) = \varepsilon^{-1/3}v(s + \varepsilon^{1/3}t), \quad \dot{x}_s(t) = \dot{v}(s + \varepsilon^{1/3}t), \quad \ddot{x}_s(t) = \varepsilon^{1/3}\ddot{v}(s + \varepsilon^{1/3}t).$$

Hence

$$\varepsilon^{4/3}\ddot{v}^2 + \varepsilon^{-2/3}W(\dot{v}) + \varepsilon^{-2/3}av^2 = \varepsilon^{2/3}\ddot{x}_s^2 + \varepsilon^{-2/3}W(\dot{x}_s) + a_s^\varepsilon x_s^2, \quad (3.5)$$

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

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

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

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

*Step 3. Asymptotic behaviour of  $f_s^\varepsilon$*

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

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

$$f_s(x) := \begin{cases} \frac{A_0}{2r} \#(S\dot{x} \cap (-r, r)) + a(s) \int_{-r}^r x^2 dt & \text{if } x \in \mathcal{S}(-r, r) \\ +\infty & \text{otherwise,} \end{cases} \quad (3.8)$$

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 [MM1-2], [Mo]): 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  $|y| \leq 1$  – and extended to  $+\infty$  elsewhere – are equicoercive on  $L^1(\Omega)$  and  $\Gamma$ -converge to the functional given by  $A_0 \|Dy\|$  when  $y$  is a function of bounded variation on  $\Omega$  which takes only the values  $\pm 1$  a.e., and  $+\infty$  otherwise. We immediately deduce that the functionals

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

$\Gamma$ -converge on  $W^{1,1}(-r, r)$  to the functional given by  $\frac{A_0}{2r} \#(S\dot{x} \cap (-r, r))$ , if  $x \in \mathcal{S}(-r, r)$ , 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  $L^1_{\text{loc}}(\mathbb{R})$  to the constant value  $a(s)$ . Thus the functionals  $\int_{-r}^r a_s^\varepsilon x^2$  converge to  $a(s) \int_{-r}^r x^2$  continuously on  $W^{1,1}(-r, r)$ .

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

#### Step 4. The main $\Gamma$ -convergence result

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

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

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

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

Propositions 3.1 and 3.3 clearly suggest the  $\Gamma$ -limit of  $F^\varepsilon$ , and indeed we have:

**Theorem 3.4.** *The functionals  $F^\varepsilon$  in (3.9)  $\Gamma$ -converge on  $\text{YM}(\Omega, K)$  to*

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

*Remark.* If  $x$  belongs to  $\mathcal{S}_{\text{per}}(0, h)$ , and  $\epsilon_x$  is the associated elementary invariant measure, a simple computation yields (cf. (5.8))

$$\langle \epsilon_x, f_s \rangle = \frac{A_0}{h} \#(S\dot{x} \cap [0, h]) + a(s) \int_0^h x^2 dt . \quad (3.12)$$

Hence the value of  $\langle \mu, f_s \rangle$  does not depend on the constant  $r$  which appears in (3.6) and (3.8) when  $\mu$  is an elementary invariant measure, and the same conclusion holds for every invariant measure by 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  $F^\varepsilon$  are defined as in (3.9) for some integrands  $f_s^\varepsilon$  which  $\Gamma$ -converge on  $K$  to  $f_s$ , then they  $\Gamma$ -converge on  $\text{YM}(\Omega, K)$  to the functional  $F$  defined as in (3.11). Unfortunately no such abstract result holds: the convergence of the integrands alone is not sufficient to guarantee the convergence of the functionals.

However, the proof of Theorem 3.4 below can be adapted to a large class of problems with few straightforward modifications, and in order to make its structure clear we have gathered in the paragraph below the relevant properties of  $f_s^\varepsilon$  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  $f_s^\varepsilon$ .

In the following,  $B(s, \rho)$  denotes the open ball of center  $s$  and radius  $\rho$ , that is, the open interval  $(s - \rho, s + \rho)$ .

### 3.5. Essential ingredients of the proof

- (a) 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).

- (b) Existence of a “nice” dense subset of  $\mathcal{A}(K)$ : *for a.e.  $s \in \Omega$ , every invariant measure  $\mu \in \mathcal{A}(K)$  can be approximated in the weak\* topology of  $\mathcal{P}(K)$  with elementary invariant measures  $\epsilon_x$  associated with functions  $x \in \mathcal{L}_{\text{per}}(0, h)$  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 (b) is verified by every  $f_s$  of the form (3.8).

- (c) Uniformity in  $s$  of  $f_s$ : *there exists a negligible set  $N \subset \Omega$  such that, for every  $h > 0$  and  $x \in \mathcal{L}_{\text{per}}(0, h)$ , 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 condition (c) is easily verified by any “reasonable” integrand.

- (d) Uniformity in  $s$  of the  $\Gamma$ -convergence of  $f_s^\varepsilon$ : *for every  $h > 0$ ,  $x \in \mathcal{L}_{\text{per}}(0, h)$ , and a.e.  $\bar{s} \in \Omega$  there exist functions  $x^\varepsilon \in H_{\text{per}}^2(0, h)$  which converge to  $x$  in  $K$  and satisfy*

$$\limsup_{\varepsilon \rightarrow 0} \int_{\substack{\tau \in [0, h] \\ s \in B(\bar{s}, \rho)}} f_s^\varepsilon(T_\tau x^\varepsilon) d\tau ds \leq \int_{\substack{\tau \in [0, h] \\ s \in B(\bar{s}, \rho)}} f_s(T_\tau x) d\tau ds + \eta(\rho), \quad (3.13)$$

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

**Proposition 3.6.** *The integrands  $f_s^\varepsilon$  defined in (3.6) satisfy condition (d) above.*

*Proof.* We prove a stronger assertion: for every  $\bar{s} \in \Omega$  and  $\rho > 0$ , the functional given by the average at the left-hand side of (3.13) for all 1-Lipschitz functions  $x$  in  $H_{\text{per}}^2(0, h)$ , and

extended to  $+\infty$  elsewhere,  $\Gamma$ -converges on  $W_{\text{per}}^{1,1}(0, h)$  to the functional equal to the average at the right-hand side of (3.13) for  $x \in \mathcal{S}_{\text{per}}(0, h)$ , and to  $+\infty$  elsewhere.

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

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

$$\int_0^h [\varepsilon^{2/3} \ddot{x}^2 + \varepsilon^{-2/3} W(\dot{x})] + \int_{\substack{\tau \in [0, h] \\ s \in B(\bar{s}, \rho)}} \left[ \int_{-r}^r a_s^\varepsilon (T_\tau x)^2 \right] d\tau ds, \quad (3.14)$$

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

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

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

*Remark.* Given positive functions  $f_s^\varepsilon$  on a metric space  $X$  which  $\Gamma$ -converge to  $f_s$  for every parameter  $s$ , it may be not true that the average of the functions  $f_s^\varepsilon$  (with respect to a fixed probability distribution on the space of parameters  $s$ )  $\Gamma$ -converge to the average of  $f_s$ . In particular, condition (a) above does not immediately imply condition (d). In fact, condition (a) implies that for every  $x \in K$ ,  $\tau \in \mathbb{R}$ , and a.e.  $s \in \Omega$  there exists a sequence  $(x^\varepsilon)$ , depending on  $x$ ,  $s$  and  $\tau$ , such that  $x^\varepsilon \rightarrow x$  in  $K$  and  $f_s^\varepsilon(T_\tau x^\varepsilon) \rightarrow f_s(T_\tau x)$ , while in (d) we essentially require that such a sequence can be chosen independent of  $\tau \in [0, h]$  and of  $s$  in a neighbourhood of a given  $\bar{s}$ .

We now start with the proof of Theorem 3.4.

We begin with the lower bound inequality. In view of Proposition 3.1 and the definitions of  $F^\varepsilon$  and  $F$ , it suffices to show that

$$\liminf_{\varepsilon \rightarrow 0} \int_{\Omega} \langle \nu_s^\varepsilon, f_s^\varepsilon \rangle ds \geq \int_{\Omega} \langle \nu_s, f_s \rangle ds \quad (3.16)$$

where  $\nu^\varepsilon$  are the elementary Young measures associated with the  $\varepsilon$ -blowup  $\mathbb{R}^\varepsilon v^\varepsilon$  of a countable sequence of functions  $v^\varepsilon$ , and  $\nu$  is the limit of  $\nu^\varepsilon$ . Since  $f_s^\varepsilon \xrightarrow{\Gamma} f_s$  on  $K$  for a.e.  $s \in \Omega$  (condition (a) of §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 not 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  $\mathcal{D}$  of Young measures with relatively simple structure which is  $F$ -dense in  $\text{YM}(\Omega, K)$  (cf. Remark 2.10).

**Definition 3.7.** *Let  $\mathcal{D}$  be the class of all Young measures  $\nu \in \text{YM}(\Omega, K)$  which satisfy the following condition: there exist countably many disjoint intervals which cover almost all of*

$\Omega$ , and on every such interval  $\nu$  agrees a.e. with an elementary invariant measure  $\epsilon_x$ , with  $x \in \mathcal{L}_{\text{per}}(0, h)$  and  $h > 0$  (depending on the interval).

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

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

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

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

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

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

and (cf. condition (c) of §3.5)

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

Putting together (3.17 – 19) we get

$$\int_{B(\bar{s}, \rho)} \phi(\epsilon_{x_{\bar{s}}} - \nu_s) ds \leq 2\eta \quad \text{and} \quad \int_{B(\bar{s}, \rho)} \langle \epsilon_{x_{\bar{s}}}, f_s \rangle ds \leq \int_{B(\bar{s}, \rho)} \langle \nu_s, f_s \rangle ds + 3\eta. \quad (3.20)$$

Using Besicovitch covering theorem (see [EG], chapter 2) we cover almost all of  $\Omega$  by countably many disjoint intervals  $B_i = B(\bar{s}_i, \rho_i)$  with  $\rho_i \leq \rho(\bar{s}_i)$ . For every  $i$  we set  $x_i := x_{\bar{s}_i}$ ,  $f_i := f_{\bar{s}_i}$ , and finally we define  $\nu^\eta \in \mathcal{D}$  by

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

Then  $\nu^\eta$  belongs to  $\mathcal{D}$ , and (3.20) yields

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

and

$$\begin{aligned} \int_{\Omega} \langle \nu_s^\eta, f_s \rangle ds &= \sum_i \int_{B_i} \langle \epsilon_{x_i}, f_s \rangle ds \leq \sum_i \left[ \int_{B_i} \langle \nu_s, f_s \rangle ds + 3\eta |B_i| \right] \\ &= \int_{\Omega} \langle \nu_s, f_s \rangle ds + 3\eta |\Omega|. \end{aligned} \quad (3.22)$$

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

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

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

$$\limsup_{\varepsilon \rightarrow 0} \int_{\substack{\tau \in [0, h] \\ s \in I}} f_s^\varepsilon(T_\tau x^\varepsilon) d\tau ds \leq \int_{s \in I} \langle \epsilon_x, f_s \rangle ds + \eta. \quad (3.23)$$

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

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

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

$$\limsup_{\varepsilon \rightarrow 0} \int_I f_s^\varepsilon(R^\varepsilon v^\varepsilon) ds \leq \int_I \langle \epsilon_x, f_s \rangle ds + \eta. \quad (3.25)$$

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

$$R_s^\varepsilon v^\varepsilon = T_{(\tau^\varepsilon - \varepsilon^{-1/3} s)} x^\varepsilon. \quad (3.26)$$

Since  $x^\varepsilon$  tends to  $x$  in  $K$ ,  $R_s^\varepsilon v^\varepsilon$  tends to the orbit  $\mathcal{O}(x)$  for every  $s \in \Omega$ , and then  $\bar{\nu}_s$  is supported on  $\mathcal{O}(x)$  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  $\mathcal{O}(x)$  is  $\epsilon_x$ .

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

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

Now we choose  $\tau^\varepsilon$  so that the integral at 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  $x^\varepsilon$  is  $h$ -periodic we get

$$\int_I f_s^\varepsilon(R^\varepsilon v^\varepsilon) ds \leq \int_0^h \left[ \int_I f_s^\varepsilon(T_{(\tau - \varepsilon^{-1/3} s)} x^\varepsilon) ds \right] d\tau = \int_0^h \left[ \int_I f_s^\varepsilon(T_\tau x^\varepsilon) ds \right] d\tau.$$

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

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

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

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



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

*Proof.* Using the fact that  $\nu$  belongs to  $\mathcal{D}$  and the uniformity assumption (d) in §3.5, for almost every point  $\bar{s} \in \Omega$  we can find an interval of the form  $I = B(\bar{s}, \rho) \subset \Omega$  with arbitrary small  $\rho$ , and a function  $x \in \mathcal{L}_{\text{per}}(0, h)$  for some  $h > 0$ , so that  $\nu_s = \epsilon_x$  for a.e.  $s \in I$ , and there exist functions  $x^\varepsilon \in H_{\text{per}}^2(0, h)$  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  $\Omega$  up to an exceptional set with measure less than  $\eta$ .  $\square$

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

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

$$\limsup_{\varepsilon \rightarrow 0} \Phi(\nu^\varepsilon - \nu) \leq \delta \quad \text{and} \quad \limsup_{\varepsilon \rightarrow 0} \int_{\Omega} f_s(\mathbb{R}_s^\varepsilon v^\varepsilon) ds \leq \int_{\Omega} \langle \nu_s, f_s \rangle ds + \delta, \quad (3.27)$$

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

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

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

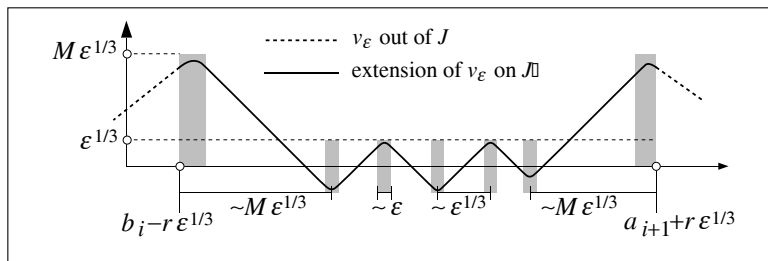
where  $\tau_i^\varepsilon$  are chosen as in Lemma 3.9. We denote the intervals  $I_i$  by  $(a_i, b_i)$ , ordered so that  $a_i < b_i < a_{i+1} < b_{i+1}$ , and set

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

It remains to extend the function  $v^\varepsilon$  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  $|x_i(t)| + 1 \leq M$  for every  $i$  and every  $t \in \mathbb{R}$ . Since the functions  $x_i^\varepsilon$  converge to  $x_i$  in  $K$  and are 1-Lipschitz, then they also converge uniformly; in particular, for  $\varepsilon$  sufficiently small,  $|x_i^\varepsilon(t)| \leq M$  for every  $i, t$ , and thus  $|v^\varepsilon(s)| \leq M\varepsilon^{1/3}$  for every  $s$  where it is defined. Notice that  $M$  depends on the choice of  $x_i$ , and ultimately on  $\eta$ ; therefore the dependence on  $M$  cannot be neglected in the estimates below.

For  $\varepsilon$  sufficiently small, we extend  $v^\varepsilon$  to the interval  $[b_i - r\varepsilon^{1/3}, a_{i+1} + r\varepsilon^{1/3}]$  as shown in the figure below:



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

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  $\varepsilon$  where  $\ddot{v}^\varepsilon$  is of order  $\varepsilon^{-1}$ . The value of  $v^\varepsilon$  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 relabelled) of the  $\varepsilon$ -blowups  $\mathbb{R}^\varepsilon v^\varepsilon$ . Since  $v^\varepsilon$  and  $v_i^\varepsilon$  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  $\varepsilon$ -blowups  $\mathbb{R}_s^\varepsilon v_\varepsilon$  and  $\mathbb{R}_s^\varepsilon v_i^\varepsilon$  agree on the larger and larger intervals

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

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

Hence  $\mathbb{R}^\varepsilon v_\varepsilon$  and  $\mathbb{R}^\varepsilon v_i^\varepsilon$  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  $\nu$  and  $\bar{\nu}$  agree on the union of the intervals  $I_i$ ; taking into account that the complement in  $\Omega$  of this union has measure lower than  $\eta$ , by the definition of  $\Phi$  in (2.2) we get  $\Phi(\bar{\nu} - \nu) \leq \eta$ . This gives the first inequality in (3.27), provided we choose  $\eta$  smaller than  $\delta$ .

Let us consider now the second inequality in (3.27). For simplicity we assume that the function  $a$  in the definition of the energy  $I^\varepsilon$  is bounded. The proof in the general case requires some additional care when dealing with the estimates (3.29 – 32) below.

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

$$f_s^\varepsilon(\mathbb{R}_s^\varepsilon v^\varepsilon) = f_s^\varepsilon(\mathbb{R}_s^\varepsilon v_i^\varepsilon) . \quad (3.29)$$

If  $s$  belongs to  $(b_i + M\varepsilon^{1/3}, a_{i+1} - M\varepsilon^{1/3})$ ,  $\mathbb{R}_s^\varepsilon v^\varepsilon$  agrees on  $(-r, r)$  with the  $\varepsilon$ -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

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

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

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

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

$$f_s^\varepsilon(\mathbb{R}_s^\varepsilon v^\varepsilon) \leq f_s^\varepsilon(\mathbb{R}_s^\varepsilon v_i^\varepsilon) + O(M^2) . \quad (3.32)$$

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

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

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

$$\limsup_{\varepsilon \rightarrow 0} \int_{\Omega} f_s^\varepsilon(R_s^\varepsilon v^\varepsilon) ds \leq \sum_i \left[ \int_{I_i} \langle \varepsilon_{x_i}, f_s \rangle ds + \eta |I_i| \right] + O(1) \eta \leq \int_{\Omega} \langle \nu_s, f_s \rangle ds + O(1) \eta,$$

which gives the second inequality in (3.27) provided that we choose  $\eta$  small enough.  $\square$

### Step 5. Minimizers of $F$

An immediate consequence of Theorem 3.4 is the following:

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

*Proof.* Apply Proposition 2.11(vi) and Theorem 3.4, taking into account (3.9), (3.11), and (3.10).  $\square$

Now we want to show that every Young measure generated by the  $\varepsilon$ -blowups of the minimizers of  $I^\varepsilon$  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

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

(cf. Fig. 2 in section 1). 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  $\mu \in \mathcal{S}(K)$ , then  $\bar{\mu}$  is the elementary invariant measure associated with the function  $y_{h(s)}$  where*

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

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

Taking Corollary 3.11 into account, we immediately deduce the following, which concludes our analysis of the asymptotic behaviour of the minimizers of  $I^\varepsilon$ .

**Corollary 3.13.** *For every  $\varepsilon > 0$ , let  $v^\varepsilon$  be a minimizer of  $I^\varepsilon$  on  $H_{\text{per}}^2(\Omega)$ . Then the  $\varepsilon$ -blowups  $R^\varepsilon v^\varepsilon$  generate a unique Young measure  $\nu \in \text{YM}(\Omega, K)$ , and, for a.e.  $s \in \Omega$ ,  $\nu_s$  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 determines the optimal periodic function  $\bar{y}$ .

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

*Proof.* Fix  $x \in \mathcal{S}_{\text{per}}(0, h)$  with  $h > 0$ . Up to a suitable translation, we may assume that  $S\dot{x} \cap [0, h]$  consists of the points  $t_0 = 0 < t_1 < t_2 < \dots < t_n = h$  and  $n = \#(S\dot{x} \cap [0, h])$  is an even number. For every  $i = 1, \dots, n$ , 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  $\pm 1$  on each  $I_i$ , we get

$$\begin{aligned} \langle \epsilon_x, f_s \rangle &= \frac{A_0}{h} n + a(s) \int_0^h x^2 dt = \sum_{i=1}^n \frac{1}{h} \left[ A_0 + a(s) \int_{I_i} x^2 dt \right] \\ &= \sum_{i=1}^n \frac{h_i}{h} \left[ \frac{A_0}{h_i} + a(s) \int_{-h_i/2}^{h_i/2} (t + p_i)^2 dt \right] \\ &= \sum_{i=1}^n \frac{h_i}{h} \left[ \frac{A_0}{h_i} + \frac{a(s)}{12} h_i^2 + a(s) p_i^2 \right]. \end{aligned}$$

We rewrite the last identity as

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

where we have set

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

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

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

Hence, (3.35), (3.37), and the fact that  $\sum h_i/h = 1$ , yield

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

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

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  $\mu \in \mathcal{S}(K)$ . However, 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 converge to a given elementary invariant measure.

**Lemma 3.15.** *Let be given  $\tilde{x} \in \mathcal{S}_{\text{per}}(0, \tilde{h})$  with  $\tilde{h} > 0$ , and, for  $k = 1, 2, \dots$ ,  $x^k \in \mathcal{S}_{\text{per}}(0, h^k)$  with  $h^k > 0$ . Then the elementary invariant measures  $\epsilon_{x^k}$  weak\* converge to  $\epsilon_{\tilde{x}}$  if (and only if) the probability that  $\tau \in (0, h^k)$  satisfies  $d(T_\tau x^k, \mathcal{O}(\tilde{x})) > \varepsilon$  vanishes as  $k \rightarrow +\infty$  for every  $\varepsilon > 0$  (here  $d$  is the distance in  $K$  and  $\mathcal{O}(\tilde{x})$  is the orbit of  $\tilde{x}$ ).*

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

Now, if  $\mu$  is the limit of (a subsequence of) the measures  $\epsilon_{x^k}$ , 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

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

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).  $\square$

The criterion in the previous lemma can be consistently improved when  $\tilde{x}$  has the special form described in (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 on  $(0, +\infty) \times \mathbb{R}$  given by

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

**Lemma 3.16.** *Let be given  $\tilde{h} > 0$  such that*

$$\lim_{k \rightarrow \infty} \int_C \left(1 + \frac{1}{h}\right) d\lambda^k(h, p) = 0 \quad (3.39)$$

for every closed set  $C \subset (0, +\infty) \times \mathbb{R}$  which does not contain the point  $(\tilde{h}/2, 0)$ . Then  $\epsilon_{x^k}$  weak\* 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  $d(T_\tau x^k, \mathcal{O}(\tilde{x})) > \varepsilon$  vanishes as  $k \rightarrow +\infty$ .

Let  $\varepsilon > 0$  be fixed. We can assume with no loss in generality that  $h^k \rightarrow +\infty$  as  $k \rightarrow +\infty$  (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 Proposition 5.1 and Remark 5.2 there exists  $m > 0$  such that

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

The proof is now divided in two steps.

**Step 1.** *Consider  $\delta > 0$  and  $\tau \in (0, h^k)$  such that*

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

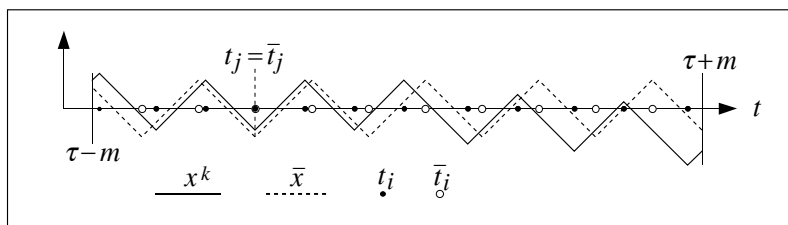
Then, for a suitable choice of the parameter  $\delta$  (not depending on  $\tau$  and  $k$ ), there holds

$$d(T_{-\tau} x^k, \mathcal{O}(\tilde{x})) \leq \varepsilon. \quad (3.41)$$

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

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

We label the points of  $S\tilde{x}$  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  $S\tilde{x}$  because  $0$  belongs to  $S\tilde{x}$ ), and we let  $\bar{I}_i$  denote the interval  $(\bar{t}_{i-1}, \bar{t}_i)$ .



**Fig.4.** 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. Fig. 4); 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  $\delta$  when  $i = j, j+1$ , less than  $2\delta$  when  $i = j-1, j+2$ , less than  $3\delta$  when  $i = j-2, j+3$ , and so on.

Taking into account that the total number of indices  $i$  such that  $\bar{I}_i$  and  $(\tau - m, \tau + m)$  intersect does not exceed  $N := 1 + 4m/\tilde{h}$ , we obtain that  $|\bar{I}_i \setminus I_i^k| \leq N\delta$  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  $t \in (\tau - m, \tau + m)$

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

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

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

which implies (3.41).

**Step 2.** *The probability that  $\tau \in (0, h^k)$  does not satisfy either assumption (a) or (b) or (c) above vanishes as  $k \rightarrow +\infty$ .*

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

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

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

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

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

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

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

We can now conclude the proof of Theorem 3.12.

Let  $\bar{\mu}$  minimize  $\langle \mu, f_s \rangle$  among all  $\mu \in \mathcal{S}(K)$ . 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  $\epsilon_{x^k}$ , with  $x^k \in \mathcal{L}_{\text{per}}(0, h^k)$  for some  $h^k > 0$ , which converge weakly\* to  $\bar{\mu}$  and satisfy

$$\langle \epsilon_{x^k}, f_s \rangle \rightarrow \langle \bar{\mu}, f_s \rangle = \min g. \quad (3.43)$$

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  $\lambda^k$  weak\* converge on  $[0, +\infty] \times [-\infty, +\infty]$  to a probability measure  $\lambda$ . 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

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

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

Moreover  $\langle \lambda^k, g \rangle \rightarrow \langle \lambda, g \rangle$ , which implies that the measures  $g \cdot \lambda^k$  converge weakly\* 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

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

This implies (3.39) because, up to a suitable multiplicative constant, the function  $g$  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,  $\mathcal{M}(K)$  is the Banach space of finite real Borel measures on  $K$  and  $\mathcal{P}(K)$  is the subset of all probability measures; we usually denote by the letter  $x$  a point of  $K$ , and by the letter  $\mu$  a measure on  $K$ . If  $K'$  is a locally compact topological space,  $\mu$  is a measure on  $K$ , and  $f$  is a Borel map from  $K$  to  $K'$ , then the *push-forward* of  $\mu$  on  $K'$  via  $f$  is the measure  $f\#\mu$  given by  $(f\#\mu)(B) := \mu(f^{-1}(B))$  for every Borel set  $B \subset K'$ .

It is also given a topological group  $G$  which 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  $\mu$  defined on  $K$ ,  $gT$  and  $T\#\mu$  denote the composed function  $g \circ T$  and the push-forward of  $\mu$  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  $\mu, g$ . A measure  $\mu$  on  $K$  is called *invariant* if it is invariant under the action of  $G$ , that is, if  $T\#\mu = \mu$  for every  $T \in G$ ;  $\mathcal{I}(K)$  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  $\mathcal{O}(x) := \{T(x) : T \in G\}$  (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  $\mathcal{O}(x)$ . If  $P(x)$  is co-compact, then  $\mathcal{O}(x)$  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  $\epsilon_x$ , called the *elementary invariant measure associated to  $x$* , which is supported on  $\mathcal{O}(x)$  and is invariant under the action of  $G/P(x)$  (see for instance [Ru], Theorem 5.14, or [Fe], section 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  $\epsilon_x$  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 hypothesis is it possible to approximate (in the weak\* topology of  $\mathcal{M}(K)$ ) 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 hypothesis on  $G$  and  $K$ , to state which we need some more definitions.

Let  $H$  be a co-compact subgroup of  $G$  and let  $\pi : G \rightarrow G/H$  be the canonical projection of  $G$  onto  $G/H$ . Since we assumed that  $G$  is commutative,  $G/H$  is a compact group and then there exists a unique (left) Haar probability measure  $\Phi$  on  $G/H$ , that is, a probability measure which is invariant under the left action of  $G/H$  on itself (see [Ru], Theorem 5.14, or [Fe], section 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  $\pi$  is one-to-one from  $A$  to  $G/H$ . We denote by  $\Phi_A$  the push-forward of the measure  $\Phi_{G/H}$  onto  $A$  according to the inverse of  $\pi$  restricted to  $A$ .*

Notice that such an inverse is a Borel measurable map, and then  $\Phi_A$  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  $\Phi_A$  given above. When no confusion may arise, we omit write explicitly the measure  $\Phi_A$  (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. [Ke], chapter 6, exercise O, or [Bi]). Thus  $G/H$  can be metrized by the quotient distance

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



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 representant  $T_i^k$  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

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

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  $y_i^k$ , which are dense in  $G/H$ , that is,  $\pi(K) = G/H$ .

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 [CV], we can find a Borel selection, namely, a Borel map  $\sigma : G/H \rightarrow K$  such that  $\pi(\sigma(y)) = y$  for every  $y \in G/H$ . We conclude by taking  $A$  equal to the image of  $\sigma$  (which is Borel measurable because  $G/H$  is compact and  $\sigma$  is one-to-one, cf. [Fe], §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*

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

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  $\mu$  on  $K$  can be approximated (in the weak\* topology of  $\mathcal{M}(K)$ ) by a sequence  $(\mu_k)$  of convex combinations of elementary invariant measures. More precisely, each  $\mu_k$  can be taken the form  $\sum_i \sigma_i \delta_{\bar{x}_i}$  where all points  $\bar{x}_i$  are  $H$ -periodic for some co-compact group  $H$  (which 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

$\mathcal{L}_n$  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

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

*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  $\Phi$  on  $G$ . To every  $\mu \in \mathcal{P}(K)$  we can associate an invariant probability measure  $P\mu$  by taking the average of all  $T^\# \mu$  with respect to the measure  $\Phi$ , that is

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

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

*Remark 4.7.* The set  $\mathcal{I}(K)$  of all invariant probability measures on  $K$  is weakly\* 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,  $\mathcal{M}(K)$  endowed with the weak\* topology) can be approximated by convex combinations of extreme points of  $C$  by the Krein-Millman theorem (cf. [Ru], Theorem 3.21).

It may be proved that  $\mu$  is an extreme point of  $\mathcal{I}(K)$  if and only if every Borel set invariant under the action of  $G$  has either full measure or zero measure (see [Ma], chapter II, Proposition 2.5, when  $G$  is the group generated by one transformation). Clearly every elementary invariant measure  $\epsilon_x$  is an extreme point of  $\mathcal{I}(K)$ , 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  $a_\infty$  irrational).

The situation simplifies when  $G$  is compact. In this case the quotient  $K/G$  is a compact metrizable space, and for every  $\mu \in \mathcal{M}(K)$  we may define the push forward  $\pi^\# \mu \in \mathcal{M}(K/G)$ , where  $\pi$  is the canonical projection of  $K$  into  $K/G$ . Then  $\pi^\#$  is a weak\* continuous operator which maps  $\mathcal{I}(K)$  into  $\mathcal{P}(K/G)$  bijectively, and takes elementary invariant measures into Dirac masses. Hence the extreme points of  $\mathcal{I}(K)$  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  $\mathcal{M}(K)$ : we take a sequence  $(g_k)$  of Lipschitz functions which is dense in  $C(K)$ , we let  $\alpha_k := \|g_k\|_\infty + \text{Lip}(g_k)$ , and set

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

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

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

- (i)  $\phi$  is a norm on  $\mathcal{M}(K)$ , and  $\phi(\mu) \leq \|\mu\|$  for every  $\mu$ ;
- (ii)  $\phi$  induces on every bounded subset of  $\mathcal{M}(K)$  the weak\* topology;
- (iii) for every  $x, y \in K$  one has  $\phi(\delta_x - \delta_y) \leq d(x, y)$ .

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

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

Regarding statement (ii), it may be easily verified that  $\phi(\mu^i - \mu) \rightarrow 0$  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  $(\mu^i)$  is bounded, this implies weak\* convergence.

We finally prove (iii):

$$\phi(\delta_x - \delta_y) = \sum_1^\infty \frac{|g_k(x) - g_k(y)|}{2^k \alpha_k} \leq \sum_1^\infty \frac{\text{Lip}(g_k) \cdot d(x, y)}{2^k \alpha_k} \leq \sum_1^\infty \frac{d(x, y)}{2^k} \leq d(x, y) . \quad \square$$

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  $\delta_{T^x}$  over all  $T$  in a representation  $A$  of the quotient  $G/H$  is the elementary invariant measure  $\epsilon_x$ . Then we notice that the operator  $P$  which associates to every  $\mu \in \mathcal{P}(K)$  the average of the translated measures  $T^\# \mu$  over all  $T \in A$  is continuous. Finally we approximate an invariant probability measure  $\mu$  by convex combinations  $\mu_k$  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  $\mu$  because  $\mu$  is invariant.

**Definition 4.9.** *Let  $B$  be a bounded Borel set of a locally compact space and let  $\lambda$  be a probability measure supported on  $B$ . Let  $\{\mu_t : t \in B\}$  be a family of measures in  $\mathcal{P}(K)$  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  $\mu_t$  over all  $t \in B$  (weighed by  $\lambda$ ) is the measure  $\mu \in \mathcal{P}(K)$  defined by*

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

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  $g : K \rightarrow \mathbb{R}$  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  $g : K \rightarrow \mathbb{R}$ .

Fix now  $\mu \in \mathcal{P}(K)$  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  $\Phi$  on  $G$  and every  $g \in C(K)$  one has

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

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  $\epsilon_x$  associated with an  $H$ -periodic point  $x$  is given by*

$$\epsilon_x = \int_A T^\# \delta_x dT = \int_A \delta_{Tx} dT . \quad (4.6)$$

*Proof.* Obviously the two integrals in (4.6) define the same probability measure  $\mu$  on  $K$ , which is supported on  $\mathcal{O}(x)$ , and since  $\epsilon_x$  is the only invariant measure supported on  $\mathcal{O}(x)$ , it suffices to verify that  $\mu$  is invariant. To this end we recall that  $[T] \mapsto [T]x := Tx$  is a well-defined continuous map from  $G/H$  to  $\mathcal{O}(x)$ , 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

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

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

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

$$\lim_{k \rightarrow \infty} \phi \left( \int_G T^\# \mu d\Phi(T) - \int_G T^\# \mu_k d\Phi(T) \right) = 0 . \quad (4.7)$$

*Proof.* For every  $\mu \in \mathcal{P}(K)$ , let  $P\mu$  be the average of  $T^\# \mu$  over all  $T \in G$  weighed by the measure  $\Phi$ , that is,  $P\mu := \int_G T^\# \mu d\Phi(T)$ . Thus  $P$  is a continuous operator from  $\mathcal{P}(K)$  into  $\mathcal{P}(K)$  (use for instance identity (4.5)). Now we take *any* sequence of convex combinations  $\mu_k$  of Dirac masses which converge to  $\mu$ ; thus  $P\mu_k$  converge to  $P\mu$ , and by Proposition 4.8(ii) we get  $\phi(P\mu_k - P\mu) \rightarrow 0$ , which is (4.7).  $\square$

**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  $\mu \in \mathcal{P}(K)$  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*

$$\phi \left( \int_A T^\# \mu dT - \bar{\mu} \right) \leq 2\varepsilon . \quad (4.8)$$

*Proof.* By applying Lemma 4.11 with  $\Phi$  replaced by  $\Phi_A$  we may find a convex combination of Dirac masses  $\hat{\mu} = \sum_i \sigma_i \delta_{x_i}$  so that

$$\phi \left( \int_A T^\# \mu dT - \int_A T^\# \hat{\mu} dT \right) \leq \varepsilon . \quad (4.9)$$

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

$$\begin{aligned} \phi\left(\int_A T^\# \delta_{x_i} dT - \int_A T^\# \delta_{\bar{x}_i} dT\right) &\leq \int_A \phi(T^\# \delta_{x_i} - T^\# \delta_{\bar{x}_i}) dT \\ &\leq \int_A d(Tx_i, T\bar{x}_i) dT \leq \varepsilon . \end{aligned} \quad (4.10)$$

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

$$\phi\left(\int_A T^\# \hat{\mu} dT - \bar{\mu}\right) \leq \sum_i \sigma_i \phi\left(\int_A T^\# \delta_{x_i} dT - \int_A T^\# \delta_{\bar{x}_i} dT\right) \leq \sum_i \sigma_i \varepsilon = \varepsilon . \quad (4.11)$$

Inequalities (4.9) and (4.11) yield (4.8).  $\square$

We can now prove Theorem 4.4. Let  $\mu$  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

$$\phi(\mu - \bar{\mu}) \leq 2\varepsilon .$$

#### *Approximation in energy*

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

$$\lim_{k \rightarrow \infty} \langle \mu_k, f \rangle = \langle \mu, f \rangle . \quad (4.12)$$

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\* lower semicontinuous on  $\mathcal{P}(K)$  because  $f$  is non-negative and lower semicontinuous. Therefore (4.12) holds whenever  $\limsup \langle \mu_k, f \rangle \leq \langle \mu, f \rangle$ .

**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*

$$\int_A d(Tx, T\bar{x}) dT \leq \varepsilon . \quad (4.13)$$

and

$$\int_A f(T\bar{x}) dT \leq \int_A f(Tx) dT + \varepsilon . \quad (4.14)$$

**Theorem 4.15.** *If  $K$  is  $f$ -uniformly approximable, then every invariant probability measure  $\mu$  on  $K$  can be approximated by convex combinations  $\mu_k$  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  $\Phi$  be a probability measure on  $G$ . Then every  $\mu \in \mathcal{P}(K)$  can be approximated by convex combinations  $\mu_k$  of Dirac masses which satisfy (4.7) and*

$$\langle \int_G T^\# \mu_k d\Phi(T); f \rangle \leq \langle \int_G T^\# \mu d\Phi(T); f \rangle \quad \text{for every } k. \quad (4.15)$$

*Proof.* For every  $\mu \in \mathcal{P}(K)$  we consider  $P\mu = \int_G T^\# \mu d\Phi(T)$  as in the proof of Lemma 4.11. Then we claim that every  $\mu \in \mathcal{P}(K)$  may be approximated by a sequence  $(\mu_k)$  of convex combinations of Dirac masses so that (4.15) holds, that is,  $\langle P\mu_k, f \rangle \leq \langle P\mu, f \rangle$  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  $\mu \in \mathcal{P}(K)$  and set  $a := \langle P\mu, f \rangle$ . With no loss of generality we may assume that  $a$  is finite, and then set

$$C := \{ \lambda \in \mathcal{P}(K) : \langle P\lambda, f \rangle \leq a \} . \quad (4.16)$$

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 weak\* compact. Moreover the extreme points of  $C$  are convex combinations of two Dirac masses (see [BL], Proposition 2, §II.2, p. 145). Since  $\mu$  belongs to  $C$ , we can apply the Krein-Milman theorem to approximate  $\mu$  with convex combinations  $\mu_k$  of extreme points of  $C$ , and thus (4.15) follows from (4.16).  $\square$

**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  $\mu \in \mathcal{P}(K)$  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*

$$\langle \bar{\mu}, f \rangle \leq \langle \int_A T^\# \mu dT; f \rangle + \varepsilon . \quad (4.17)$$

*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

$$\langle \int_A T^\# \hat{\mu} dT; f \rangle \leq \langle \int_A T^\# \mu dT; f \rangle . \quad (4.18)$$

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

$$\langle \bar{\mu}, f \rangle = \sum_i \sigma_i \int_A f(T\bar{x}_i) dT \leq \sum_i \sigma_i \int_A f(Tx_i) dT + \varepsilon = \langle \int_A T^\# \hat{\mu} dT; f \rangle + \varepsilon ,$$

which together with inequality (4.18) implies (4.17).  $\square$

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  $\mu$  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  $\mu$  is invariant (4.8) and (4.17) become respectively

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

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  $\epsilon_x$  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  $x \mapsto (ax+b)/(cx+d)$  with  $ad-bc \neq 0$ . 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  $\infty$ , but on this is also impossible, 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 [Fe], 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 [Fe], Theorem 2.7.11 and §2.7.12, or [HR], section 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  $G/H$  to  $\mathcal{O}(x)$ , and then  $\mathcal{O}(x)$  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  $x : \mathbb{R}^n \rightarrow [-\infty, +\infty]$  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\* 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

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

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  $L_{\text{loc}}^p(\mathbb{R}^n)$  embeds continuously in  $K$  for  $1 \leq p \leq \infty$ . Moreover, (5.1) yields, for every  $p \in [1, \infty]$ ,

$$d(x_1, x_2) \leq \sum_{k=1}^{\infty} \frac{1}{2^k \alpha_k} \int_{\mathbb{R}^n} |y_k| |x_1 - x_2| \leq \sum_{k=1}^{\infty} \frac{\|y_k\|_q \|x_1 - x_2\|_p}{2^k (\|y_k\|_1 + \|y_k\|_\infty)} \leq \|x_1 - x_2\|_p. \quad (5.2)$$

(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  $\|y_k\|_q \leq \|y_k\|_1 + \|y_k\|_\infty$ ).

*Remark.* Embedding into  $K$  may be no longer continuous if we consider weaker forms of convergence: if the functions  $x_k : \mathbb{R}^n \rightarrow \{a, b\}$  weak\* 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  $\tan(\frac{1}{2}(\arctan a + \arctan b))$ .

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  $x_1, x_2 \in K$ :*

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

(Here  $a \wedge b$  and  $a \vee b$  denote respectively the infimum 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  $|\arctan x_1(t) - \arctan x_2(t)| \leq \pi/2 - \arctan m$  for a.e.  $t \in (-m, m)^n$ , and since  $\text{spt } y_k \subset (-k, k)^n$ , for  $k \leq m$  we have

$$\left| \int_{\mathbb{R}^n} y_k (\arctan x_1 - \arctan x_2) d\mathcal{L}_n \right| \leq \|y_k\|_1 (\pi/2 - \arctan m) \leq \pi/2 - \arctan m.$$

Hence

$$d(x_1, x_2) \leq \sum_{k=1}^m \frac{\pi/2 - \arctan m}{2^k} + \sum_{k=m+1}^{\infty} \frac{\pi \|y_k\|_1}{2^k} \leq \frac{\pi}{2} - \arctan m + \frac{1}{2^m}.$$

To finish the proof it suffices to choose  $m$  large enough.  $\square$

*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  $d(x_1, x_2) \leq d(x_1, x) + d(x, x_2)$ , 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:

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



where  $m$  and  $\varepsilon$  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. Hence Theorem 4.4 applies, and thus 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\* topology. Consider  $\tau_k \rightarrow \tau$  in  $\mathbb{R}^n$ ,  $x_k \xrightarrow{*} x$  in  $L^\infty(\mathbb{R}^n)$ , and  $y \in L^1(\mathbb{R}^n)$ . Then  $T_{-\tau_k}y \rightarrow T_{-\tau}y$  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  $T_{\tau_k}x_k \xrightarrow{*} T_\tau x$  in  $L^\infty(\mathbb{R}^n)$ .

Let us show that  $K$  is uniformly approximable. Fix  $\varepsilon > 0$  and take  $m$  so that implication (5.3) holds, and then choose  $a$  so that  $a\varepsilon \geq m$ . 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  $d(T_\tau \bar{x}, T_\tau x) \leq \varepsilon$  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

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

*Remark.* Notice that there exist points  $x \in K$  whose orbit is dense in  $K$ . In other words  $\mathcal{O}(x)$  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 we 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  $\mathcal{S}(I)$  and  $\mathcal{S}_{\text{per}}(0, h)$  are defined at the beginning of section 3, while  $\mathcal{S}_{\text{per}, 0}(0, h)$  denotes the space of all  $x \in \mathcal{S}_{\text{per}}(0, h)$  such that  $x(0) = x(h) = 0$ ;  $r$  is a fixed positive real number and we set (cf. (3.8))

$$f(x) := \begin{cases} \frac{1}{2r} \#(Sx \cap (-r, r)) + \int_{-r}^r x^2(t) dt & \text{if } x \in \mathcal{S}(-r, r), \\ +\infty & \text{otherwise.} \end{cases} \quad (5.5)$$

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

*Proof.* Let  $(x_k)$  be a sequence such that  $x_k \rightarrow x$  in  $K$  and the values  $f(x_k)$  are uniformly bounded. Then the functions  $x_k$  belong to  $\mathcal{S}(-r, r)$  for every  $k$ , and in particular 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  $\dot{x}$  weakly\* in  $BV(-r, r)$ , hence  $x$  belongs to  $\mathcal{S}(-r, r)$ , and  $\liminf f(x_k) \geq f(x)$ .  $\square$

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

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

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

$$\int_0^h f(T_\tau x) d\tau = \frac{1}{h} \left[ \sum_{t \in S\dot{x}} (\rho * 1_{[0, h]})(t) + \int_{\mathbb{R}} (\rho * 1_{[0, h]}) x^2 dt \right] \quad (5.7)$$

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

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

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

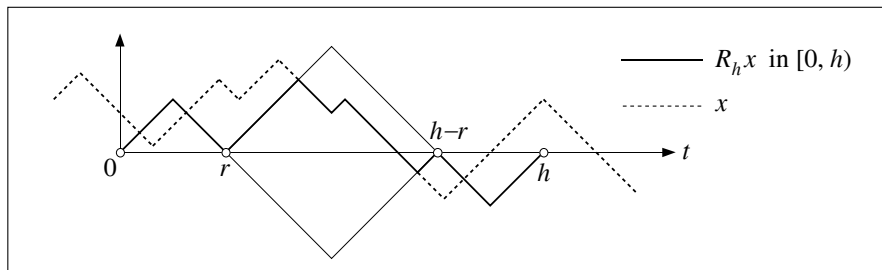
To obtain (5.7), we consider the measure  $\lambda$  given by  $\mu(B) := \#(B \cap S\dot{x} \cap (-r, h+r)) + \int_B x^2 dt$ , and thus we write  $f(T_\tau x)$  as

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

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 Fig. 5):*

- in the interval  $[0, r)$ ,  $R_h x(t)$  is equal to  $t \wedge (-t+r)$ ;
- in the interval  $[r, h/2)$ ,  $R_h x(t)$  is equal to  $x(t)$  if  $|x(t)| \leq t-r$ , to  $t-r$  if  $x(t) > t-r$ , and to  $-(t-r)$  if  $x(t) < -(t-r)$ ;
- in the interval  $[h/2, h-r)$ ,  $R_h x(t)$  is 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  $x(t) < -(h-r-t)$ ;
- in the interval  $[h-r, h)$ ,  $R_h x(t)$  is equal to  $(t-h+r) \vee (-t+h)$ .



**Fig.5.** 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  $x \in \mathcal{S}(r, h-r)$  then  $R_h x \in \mathcal{S}_{\text{per},0}(0, h)$  and  $S\dot{x} \cap (r, h-r)$  contains  $S(R_h x) \cap [0, h)$  except at most six points.

*Proof.* Straightforward (see Fig. 5). □

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

$$\int_0^h d(T_\tau x, T_\tau R_h x) d\tau \leq 2\varepsilon \quad \text{and} \quad \int_0^h f(T_\tau R_h x) d\tau \leq \int_0^h f(T_\tau x) d\tau + \varepsilon. \quad (5.9)$$

In particular,  $K$  is  $f$ -uniformly approximable (see 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  $d \leq 1$ ,

$$\int_0^h d(T_\tau x, T_\tau R_h x) d\tau \leq \varepsilon + \frac{2(m+r)}{h},$$

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

Let us consider now the second inequality in (5.9). We can assume that the integral  $\int_0^h f(T_\tau x) d\tau$  is finite (otherwise there is nothing to prove). Therefore  $R_h x \in \mathcal{S}_{\text{per},0}(0, h)$ , and

- $\#(S(R_h x) \cap [0, h)) \leq \#(S\dot{x} \cap (r, h-r)) + 6$  (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),

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

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

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

and by (5.7),

$$= \int_0^h f(T_\tau x) d\tau + \frac{12 + r^2}{2h}.$$

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

**Corollary 5.9.** *Every invariant probability measure  $\mu$  on  $K$  can be approximated by a sequence of convex combinations  $\mu_k$  of elementary invariant measures so that (4.12) holds. Moreover, if  $\langle \mu, f \rangle$  is finite, each  $\mu_k$  can be taken of the form  $\mu_k = \sum \sigma_i \epsilon_{\bar{x}_{ki}}$  with  $\bar{x}_{ki} \in \mathcal{L}_{\text{per},0}(0, h_k)$  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 keeping this fact in mind, 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  $\mu_k$  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  $\bar{x}_{ik} \in \mathcal{L}_{\text{per},0}(0, h_k)$ .  $\square$

We can refine the approximation result of Corollary 5.9 by showing that  $\mu$  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  $x_i \in \mathcal{L}_{\text{per},0}(0, h)$  for every  $i = 1, \dots, N$ , we can find  $\bar{h} > 0$  and  $x \in \mathcal{L}_{\text{per},0}(0, \bar{h})$  such that*

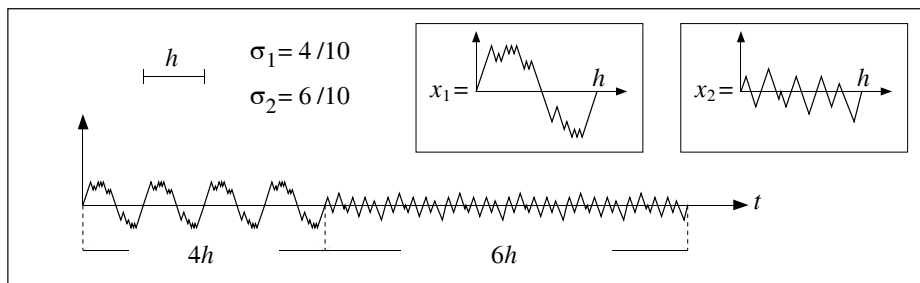
$$\phi(\mu - \epsilon_x) \leq 2\varepsilon \quad \text{and} \quad |\langle \mu, f \rangle - \langle \epsilon_x, f \rangle| \leq \varepsilon. \quad (5.10)$$

*Proof.* First of all, notice that all  $\sigma_i$  can be assumed rational (by a standard density argument). We fix  $m > 0$  such that implication (5.3) holds, and we write every  $\sigma_i$  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  $x \in \mathcal{L}_{\text{per},0}(0, qh)$  defined by

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

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. Fig. 6 below).



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

Taking into account (4.6) we get

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

Hence

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

and by Proposition 4.8(iii)

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

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  $d(T_\tau x, T_\tau x_i) \leq \varepsilon$  by (5.3). Hence inequality (5.12) becomes (recall that  $d \leq 1$ )

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

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

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  $\#(S\dot{x} \cap [0, qh]) \leq N + \sum_i \#(S\dot{x}_i \cap [q_{i-1}h, q_i h])$ . Hence (5.8) yields

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

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

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

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

*Proof.* Apply Corollary 5.9 and Lemma 5.10.  $\square$

## 6. Overview of further applications

In this section we briefly sketch in different detail 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.

### Boundary conditions

The periodic boundary conditions imposed in the study of the functional  $I^\varepsilon$  (see (3.3)) can be replaced by other boundary conditions (Dirichlet, natural, ...) without changing the limit problem. In other words, the  $\Gamma$ -limit  $F$  defined in (3.11) is independent of the boundary conditions. This is not surprising since the limit functional contains no correlations between values of the slow variable  $s$ .

### Additional externally imposed scales

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

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

converge for any sequence  $x^\varepsilon$  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))

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

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

If  $\beta > \frac{1}{3}$ , i.e., if the externally imposed scale  $\varepsilon^\beta$  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  $a(s)\int_{-r}^r x^2$  with  $\bar{a}\int_{-r}^r x^2$  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  $\int_{-r}^r a_s^\varepsilon(t)x^2$  converge continuously to  $\bar{a}\int_{-r}^r x^2$ .

If  $\beta < \frac{1}{3}$  then this convergence no longer holds. We expect that minimizers of the  $\varepsilon$ -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

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

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

The (rescaled) limiting energy is thus given by  $E_0(\int a^{1/3}(q)dq)$  (cf. (1.4)). In this case the  $\Gamma$ -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  $\nu_s \in \mathcal{S}(K)$  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))

$$F(\nu) = \frac{A_0}{h} \#(S\dot{x} \cap [0, h)) + \bar{a} \int_0^h x^2 dt .$$

Hence if  $F$  was affine the minimal energy would involve  $\bar{a}^{1/3} = (\int a)^{1/3}$  rather than the smaller value  $\int a^{1/3}$ . In this case a more natural representation of the limit might be achievable by

performing a (hierarchical) blow-up with two small scales  $\varepsilon^\beta$  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  $E_0(f a)^{1/3}$  for  $\beta > \frac{1}{3}$  (cf. (1.4)), and by  $E_0(f a^{1/3})$  for  $\beta < \frac{1}{3}$ .

### 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 [KM]):

$$I^\varepsilon(v) := \int_{\Omega} \varepsilon^2 \dot{v}^2 + W(\dot{v}) ds + \|v\|_{\dot{H}^{1/2}}^2. \quad (6.1)$$

The minimization is taken over functions in  $v \in H_{\text{per}}^2(\Omega)$  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

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

and can be written as

$$\|v\|_{\dot{H}^{1/2}}^2 = \int_{\Omega \times \Omega} g(s-s') (v(s) - v(s'))^2 ds' ds = \int_{\Omega \times \Omega} h(s-s') \dot{v}(s) \dot{v}(s') ds' ds, \quad (6.2)$$

where we have set

$$g(t) := \frac{\pi}{4(1 - \cos \pi t)} \quad \text{and} \quad h(t) := -\frac{1}{2\pi} [\ln 2 + \ln(1 - \cos \pi t)]. \quad (6.3)$$

(Notice that the second identity in (6.2) holds for functions in  $W^{1,1}$  only).

The scheme developed in section 3 applies to this functional, too, even though 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., [KM]). 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

$$R_s^\varepsilon \dot{v} := \dot{v}(s + \varepsilon^{1/2} t). \quad (6.4)$$

The competitors for the limit problem will be the class of all Young measures  $\nu \in \text{YM}(\Omega, K)$  generated by sequences  $R_s^\varepsilon \dot{v}^\varepsilon$ . As in section 3, these Young measures are characterized as those  $\nu$  such that  $\nu_s$  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  $\varepsilon$ -blowups. To avoid problems with integrals over unbounded domains we choose a smooth positive function  $\rho$  on  $\mathbb{R}$  such that  $\int \rho(t) dt = 1$ . For  $v \in H_{\text{per}}^2(\Omega)$  we set

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

Thus

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

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

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

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

$$\varepsilon^{-1/2} I^\varepsilon(v) = \int_{\Omega} f^\varepsilon(\dot{x}_s) , \quad (6.5)$$

where

$$f^\varepsilon(\dot{x}) := \int_{\mathbb{R}} (\varepsilon^{1/2} \ddot{x}^2 + \varepsilon^{-1/2} W(\dot{x})) \rho(t) dt + \int_{\mathbb{R} \times \mathbb{R}} g^\varepsilon(\tau) (x(t + \tau) - x(t))^2 \rho(t) dt d\tau , \quad (6.6)$$

and

$$g^\varepsilon(\tau) := \begin{cases} \varepsilon g(\varepsilon^{1/2}\tau) & \text{if } |\tau| \leq \varepsilon^{-1/2}, \\ 0 & \text{otherwise.} \end{cases} \quad (6.7)$$

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

From (6.3) 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  $f^\varepsilon$   $\Gamma$ -converge on  $K$  to the functional  $f$  given by

$$f(\dot{x}) := A_0 \sum_{t \in S\dot{x}} \rho(t) + \int_{\mathbb{R} \times \mathbb{R}} \frac{1}{2\pi\tau^2} (x(t + \tau) - x(t))^2 \rho(t) dt d\tau \quad (6.8)$$

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

$$\int (\varepsilon^{1/2} \ddot{x}^2 + \varepsilon^{-1/2} W(\dot{x})) \rho \quad (6.9)$$

are equicoercive and  $\Gamma$ -converge on  $W_{\text{loc}}^{1,1}(\mathbb{R})$  – and therefore also in  $K$  – to the sum at the right-hand side of (6.8), while the double integrals at the right-hand side of (6.6) converge to the double integral at the right-hand side of (6.8) for all sequences  $x^\varepsilon$  which converge to  $x$  uniformly on  $\mathbb{R}$ , and are uniformly Lipschitz. Unfortunately such a convergence is not implied by convergence in  $W_{\text{loc}}^{1,1}(\mathbb{R})$ , and one has to be more careful: given functions  $\dot{x}^\varepsilon \rightarrow \dot{x}$  such that  $f^\varepsilon(\dot{x}^\varepsilon)$  is bounded, we have  $\dot{x}^\varepsilon \rightarrow \dot{x}$  in  $L_{\text{loc}}^1(\mathbb{R})$ , and, modulo addition of suitable constants,  $x^\varepsilon \rightarrow x$  in  $W_{\text{loc}}^{1,1}(\mathbb{R})$ . Then a careful application of Fatou's lemma, and the fact that  $g$  and  $\rho$  are positive functions, give the lower-bound inequality. To prove the upper-bound inequality for  $x$ , it suffices to construct functions  $x_\varepsilon$  which converge uniformly to  $x$ , are uniformly Lipschitz, and satisfy the upper-bound inequality for the functionals in (6.9).

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  $v^\varepsilon$  minimize  $I^\varepsilon$  and the  $\varepsilon$ -blowups  $\mathbf{R}^\varepsilon \dot{v}^\varepsilon$  generate a Young measure  $\nu$ . Then, for a.e.  $s \in \Omega$ , the measure  $\nu_s$  minimizes  $\langle \mu, f \rangle$  among all invariant measures  $\mu \in \mathcal{S}(K)$ .*



We have not been able to carry out the last step of our program, the characterization of the minimizing measures  $\mu$ . We conjecture that minimality implies that the measure is supported on the orbit of the derivative of a single periodic sawtooth function like in Fig. 2 of section 1.

For every  $x \in \mathcal{S}_{\text{per}}(0, h)$  one has (cf. (3.12))

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

It can be verified that the infimum of  $\langle \epsilon_{\dot{x}}, f \rangle$  over all  $x \in \mathcal{S}_{\text{per}}(0, h)$  and  $h > 0$  is strictly positive, and hence the minimum of  $\langle \mu, f \rangle$  over all  $\mu \in \mathcal{S}(K)$  is also strictly positive. This shows in particular that the minima of energies  $I^\varepsilon$  in (6.1) are exactly of order  $\varepsilon^{1/2}$ .

As a first step in the characterization of minimizing measures  $\mu$ , one should prove that the infimum of  $\langle \epsilon_{\dot{x}}, f \rangle$  over all  $x \in \mathcal{S}_{\text{per}}(0, h)$  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.

### *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  $\varepsilon \rightarrow 0$  minimizers  $v^\varepsilon$  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\* 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  $\varepsilon$ -blowups  $\mathbb{R}_s^\varepsilon v^\varepsilon(t) := v^\varepsilon(s + \varepsilon t)$ , and define the associated measures  $\nu^\varepsilon$  on  $\Omega \times K$  by

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

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

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

where  $\mu_a$  and  $\mu_b$  are probability measures on  $K$  which capture the asymptotic behaviour of minimizers  $v^\varepsilon$  near the jumps  $a$  and  $b$  resp. If we assume that the limit  $v$  of the minimizers  $v^\varepsilon$

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

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

(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  $\mu_b$ .

## References

- [AB] G. Alberti, G. Bellettini: A non-local anisotropic model for phase transitions - Asymptotic behaviour of rescaled energies, *European J. Appl. Math.* **9** (1998), 261-284.
- [Al] G. Allaire: Homogenization and two-scale convergence, *SIAM J. Math. Anal.* **23** (1992), 1482-1518.
- [AIB] G. Allaire, M. Briane: Multi-scale convergence and reiterated homogenization, *Proc. Roy. Soc. Edinburgh-Sect. A*-**126** (1996), 297-342.
- [Bal] E.J. Balder: A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM J. Control and Optimization* **22** (1984), 570-598.
- [Ba] J.M. Ball: A version of the fundamental theorem for Young measures, in *PDE's and Continuum Models of Phase Transitions* (M. Rascle et al., eds.), *Lect. Notes in Physics* **344**, Springer, Berlin 1989.
- [BJ] J.M. Ball, R.D. James: Fine phase mixtures as minimizers of energy, *Arch. Rational Mech. Anal.* **100** (1987), 13-52.
- [BES] G. Barles, L.C. Evans, P.E. Souganidis: Wavefront propagation for reaction-diffusion systems of PDE, *Duke Math. J.* **61** (1990), 835-858.
- [BF] A.C. Barroso, I. Fonseca: Anisotropic singular perturbations - The vectorial case, *Proc. Roy. Soc. Edinburgh* **A-124** (1994), 527-571.
- [BL] H. Berliocchi, J.-M. Lasry: Intégrales normales et mesures paramétrées en calcul des variations, *Bull. Soc. Math. France* **101** (1973), 129-184.
- [Bi] G. Birkhoff: A note on topological groups, *Compositio Math.* **3** (1936), 427-430.
- [CGG] Y.G. Chen, Y. Giga, S. Goto: Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations, *J. Differential Geom.* **33** (1991), 749-786.
- [CK] M. Chipot, D. Kinderlehrer: Equilibrium configurations of crystals, *Arch. Rational Mech. Anal.* **103** (1988), 237-277.
- [DM] G. Dal Maso: *An Introduction to  $\Gamma$ -Convergence*, Progress in Nonlinear Diff. Eq. Appl. **8**, Birkhäuser, Boston 1993.
- [DP] R.J. DiPerna: Compensated compactness and general systems of conservation laws, *Trans. Amer. Math. Soc.* **292** (1985), 383-420.
- [Do] W. Döring: Mikromagnetismus, in *Handbuch der Physik* **18/2** (S. Flügge ed.), Springer, Heidelberg 1966, 341-347.
- [E] Weinan E: Homogenization of linear and nonlinear transport equations, *Comm. Pure Appl. Math.* **45** (1992), 301-326.
- [Ed] R.E. Edwards: *Functional Analysis, Theory and Applications*, Holt, Rinehart and Winston, New York 1965 (reprinted by Dover, New York 1995).

- [EG] L.C. Evans, R. Gariepy: *Measure Theory and Fine Properties of Functions*, Studies in Advanced Math., CRC Press, Boca Raton 1992.
- [ES] L.C. Evans, J. Spruck: Motion of level sets by mean curvature I-IV, resp. in *J. Differential Geom.* **33** (1991), 635-681; *Trans. Amer. Math. Soc.* **330** (1992), 321-332; *J. Geom. Anal.* **2** (1992), 121-150; *J. Geom. Anal.* **5** (1995), 77-114.
- [ESS] L.C. Evans, H.M. Soner, P.E. Souganidis: Phase transitions and generalized motion by mean curvature, *Comm. Pure Appl. Math.* **45** (1992), 1097-1123.
- [Fe] H. Federer: *Geometric Measure Theory*, Grundlehren der math. Wissensch. **153**, Springer, Berlin 1969 (reprinted in the series Classics in Mathematics, Springer, Berlin 1996).
- [Ge1] P. Gérard: Microlocal defect measures, *Comm. PDE* **16** (1991), 1761-1794.
- [Ge2] P. Gérard: Mesures semi-classiques et ondes de Bloch, in *Seminaire sur les Equations aux Derivees Partielles, 1990-1991*, Ecole Polytechnique, Palaiseau 1991.
- [HR] E. Hewitt, K.A. Ross: *Abstract Harmonic Analysis, vol. I*, second edition, Grundlehren der math. Wissensch. **115**, Springer, New York 1979.
- [HS] A. Hubert, R. Schäfer: *Magnetic Domains - The Analysis of Magnetic Microstructures*, Springer, Heidelberg 1998.
- [KS1] M.A. Katsoulakis, P.E. Souganidis: Interacting particle systems and generalized evolution of fronts, *Arch. Rational Mech. Anal.* **127** (1994), 133-157.
- [KS2] M.A. Katsoulakis, P.E. Souganidis: Stochastic Ising models and anisotropic front propagation, *J. Statist. Phys.* **87** (1997), 63-89.
- [Ke] J.L. Kelley: *General Topology*, D. Van Nostrand Co., Toronto 1955, (reprinted in the series Graduate Texts in Mathematics, Springer, New York 1975).
- [Kh] A. Khachaturyan: *Theory of Structural Transformations in Solids*, J. Wiley and Sons, New York 1983.
- [KP1] D. Kinderlehrer, P. Pedregal: Characterization of Young measures generated by gradients, *Arch. Rational Mech. Anal.* **115** (1991), 329-365.
- [KP2] D. Kinderlehrer, P. Pedregal: Gradient Young measure generated by sequences in Sobolev spaces, *J. Geom. Analysis* **4** (1994), 59-90.
- [KM] R.V. Kohn, S. Müller: Branching of twins near an austenite-twinned-martensite interface, *Philosophical Magazine A* **66** (1992), 697-715.
- [Kr] J. Kristensen: Lower semicontinuity in spaces of weakly differentiable functions, *Math. Ann.* **313** (1999), 653-710.
- [Ma] R. Mañé: *Ergodic Theory and Differentiable Dynamics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **8**, Springer, Berlin 1987.
- [Mo] L. Modica: The gradient theory of phase transitions and the minimal interface criterion, *Arch. Rational Mech. Anal.* **98** (1987), 123-142.
- [MM1] L. Modica, S. Mortola: Un esempio di  $\Gamma$ -convergenza, *Boll. Un. Mat. Ital. (5)*, **14-B** (1977), 285-299.
- [MM2] L. Modica, S. Mortola: Il limite nella  $\Gamma$ -convergenza di una famiglia di funzionali ellittici, *Boll. Un. Mat. Ital. (5)*, **14-A** (1977), 526-529.
- [Mu] S. Müller: Singular perturbations as a selection criterion for periodic minimizing sequences, *Calc. Var.* **1** (1993), 169-204.
- [MT] F. Murat, L. Tartar: On the relation between Young measures and  $H$ -measures, paper in preparation.
- [Ng] G. Nguetseng: A general convergence result for a functional related to the theory of homogenization, *SIAM J. Math. Anal.* **20** (1989), 608-623.
- [Pe] P. Pedregal: *Parametrized Measures and Variational Principles*, Progress in Nonlinear Differential Equations and their Applications **30**, Birkhäuser, Basel 1997.

- [PZ] M. Pitteri, G. Zanzotto: *Continuum Models for Phase Transitions and Twinning in Crystals*, Chapman and Hall, book in preparation.
- [Ro] T. Roubiřek: *Relaxation in Optimization Theory and Variational Calculus*, de Gruyter Series in Nonlinear Analysis and Applications **4**, Walter de Gruyter & Co., Berlin 1997.
- [Ru] W. Rudin: *Functional Analysis*, second edition, McGraw-Hill, New York 1991.
- [Sv] V. řverák: Lower semicontinuity of variational integrands and compensated compactness, in *Proceedings of the ICM 1994* (S.D. Chatterji ed.), Birkhauser, Basel 1995, 1153-1158.
- [Sy] M.A. Sychev: A new approach to Young measure theory, relaxation and convergence in energy, to appear on *Ann. Inst. H. Poincaré, Anal. non linéaire*.
- [Ta1] L. Tartar: Compensated compactness and application to partial differential equations, in *Nonlinear Analysis and Mechanics: Heriot-Watt Symposium IV* (R. Knops, ed.), Pitman Research Notes in Math. **39**, Boston 1979, 136-212.
- [Ta2] L. Tartar: The compensated compactness method applied to systems of conservation laws, in *Systems of Nonlinear Partial Differential Equations (Oxford, 1982)* (J.M. Ball, ed.), NATO-ASI Series C **111**, Reidel, Dordrecht 1983.
- [Ta3] L. Tartar: unpublished notes.
- [Ta4] L. Tartar: *H*-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations, *Proc. Roy. Soc. Edinburgh* **A-115** (1990), 193-230.
- [Ta5] L. Tartar: Beyond Young measures, *Meccanica* **30** (1995), 505-526.
- [Ta6] L. Tartar: Homogenization, compensated compactness and H-measures, in the CBMS-NSF conference (Santa Cruz 1993), lecture notes in preparation.
- [Va1] M. Valadier: Young measures, in *Methods of Non-convex Analysis* (A. Cellina ed.), Lect. Notes in Math. **1446**, Springer, Berlin 1990.
- [Va2] M. Valadier: A course on Young measures, *Rend. Istit. Mat. Univ. Trieste* **26** (1994) suppl., 349-394.
- [WLR] M.S. Wechsler, D.S. Liebermann, T.A. Read: On the the theory of the formation of martensite, *Trans. AIME J. Metals* **197** (1953), 1503-1515.
- [Yo1] L.C. Young: 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.
- [Yo2] L.C. Young: *Lectures on the Calculus of Variations and Optimal Control Theory*, W.B. Saunders, Philadelphia 1969, (reprinted by Chelsea, 1980).