AUBRY-MATHER THEORY FOR FUNCTIONS ON LATTICES

Hans Koch, Rafael de la Llave, Charles Radin

Abstract. We generalize the Aubry-Mather theorem on the existence of quasi-periodic solutions of one dimensional difference equations to situations in which the independent variable ranges over more complicated lattices. This is a natural generalization of Frenkel-Kontorova models to physical situations in a higher dimensional space. We also consider generalizations in which the interactions among the particles are not just nearest neighbor, and indeed do not have finite range.

1. Introduction

The goal of this paper is to generalize the Aubry-Mather theorem on the existence of quasi-periodic solutions of certain difference equations to cases in which the unknown depends not just on one, but on several variables by modifying a method used in [Go] in the one-dimensional case.

Such generalizations are natural from the point of view of the solid state motivations of the theorem. For example, the Frenkel-Kontorova model considered by Aubry and Mather describes configurations which are one-dimensional chains of particles in a periodic external potential and interacting with their nearest neighbors via a harmonic potential. Such models are one-dimensional caricatures of the physical situation of a layer of a material over a substratum of other material. If we denote by $V(x) = V(x+1)$ the periodic potential and by u_i the displacement of the i^{th} particle from the its equilibrium position, we seek configurations that are critical points of the energy:

$$\mathcal{S}(u) = \sum_{i \in \mathbb{Z}} \frac{1}{2}(u_i - u_{i+1} - a)^2 + V(u_i) \tag{1.1}$$

(We also point out that there are other physical interpretations. For example, we could have atoms whose state is described by an internal variable u_i. The internal energy is periodic but there is a coupling between nearest variable. For example, the model (1.1) has been used as models of spin waves. In that interpretation, $V(u_i)$ would be the magnetic energy of the i^{th} atom and the term $\frac{1}{2}(u_i - u_{i+1})^2$ — or a modification — would be the exchange interaction between neighboring spins.)

Even if the sum on the R. H. S. of (1.1) is only formal, the variational equations are quite well defined, namely:

$$u_{i+1} + u_{i-1} - 2u_i + V'(u_i) = 0 \tag{1.2}$$

1991 *Mathematics Subject Classification.* AMS Subject Classification (1991) 58F99 82B20 49J40.

which are equivalent to the well known standard-like mappings of Hamiltonian mechanics.

The Aubry-Mather theorem establishes, among other things, that for every $\omega \in \mathbb{R}$ the variational equations (1.2) admit a solution u such that $\sup_i |u_i - \omega i| < \infty$. In this paper we will be concerned with generalizing this theorem to situations in which:

- The physical space is higher dimensional.
- The interactions are not necessarily just nearest neighbor.
- The interactions are invariant under a smaller symmetry group than the full lattice.

An example to keep in mind of a model to which our results apply is the n-dimensional analogue of the Frenkel-Kontorova model, described by the action

$$\mathcal{S}(u) = \sum_{i \in \mathbb{Z}^n} \frac{1}{2n} \sum_{j : |j-i|=1} |u_i - u_j|^2 + V(u_i)$$

which leads to the variational equations:

$$(\Delta u)_i + V'(u_i) = 0$$

where Δ denotes the discrete Laplacian. Such models for $n = 2$ have been considered in [V], where one can also find an extensive discussion of properties of solutions, and physical consequences.

Another generalization which we will be able to deal with is

$$\mathcal{S}(u) = \sum_{i \in \mathbb{Z}^n} \frac{1}{2n} \sum_{j : |j-i|=1} |u_i - u_j|^2 + V(u_i, i)$$

where $V(x, i)$ is periodic in i and similar generalizations. In the case $n = 1$, this admits the dynamic interpretation of a composition of finite number of standard maps or, alternatively, as a map in a higher dimensional space. These generalizations, however do not seem to include higher dimensional standard maps.

The method of proof we will use is motivated by a recent paper of Golé [Go] which presents a new proof of the classical Aubry-Mather theorem for one dimensional Frenkel-Kontorova models. We point out that it is also possible using the methods employed here to give a simple proof of generalizations of a theorem of Moser [Mo] on partial differential equations. Since the latter results seem to require other ingredients from P.D.E., and have different motivation, we will report on them elsewhere. The use of heat flow methods in Aubry-Mather theory was introduced in the paper [An], but the methods we use here are more elementary.

We also point out that results very related to those of this paper, can be found in [Bl1], [Bl2] by using quite different methods. We also note that the papers [Bl1] and [Bl2] look for minimal solutions. Following the lead of [Mo], [An] and [Go], we look for solutions of the variational problem that are well ordered. This makes the estimates at infinity much simpler. In this respect, it is interesting to point out that if we considered functions taking values in higher dimensional spaces, the comparison methods that we use would not work and, indeed, there are examples in [Bl2] of minimal solutions which are not well ordered. These phenomena have been observed numerically in [OV].

2. Acknowledgments

We thank C. Golé for acquainting us with [Go] prior to publication, and for his encouragement. He and F. Tangerman made very useful comments on a preliminary version of this paper. We also thank L. M. Floria for making us aware of [V] and providing us with a copy and J. Moser for several discussions and making us aware of [Bl1], [Bl2]. The authors have been supported by NSF grants. R.L. also acknowledges the support of the AMS Centennial Fellowship and a URI from U. Texas during the last revision of this paper.

3. Notation and statement of results

We will be concerned with functions defined on lattices (i.e. discrete subgroups of $\mathbb{R}^d$ with its additive structure). We will find it useful to distinguish between the lattices as subsets of $\mathbb{R}^d$ and as groups of motions. (For example, we will find it useful to consider smaller groups of motions than the full lattice.) Hence, we will use the word crystals when we want to emphasize the fact that the group structure is irrelevant We will also refer to functions on lattices as configurations. This is motivated by the physical interpretation of the models in (1.1) for which the i represent labels of atoms, u_i represent characteristics of each atom. The function u_i represents the state — or configuration of the crystal — and the action in (1.1) is the energy associated to the configuration.

As it is well known, discrete subgroups of $\mathbb{R}^d$ can be identified with sets of the form $\{x \in \mathbb{R}^d \mid x = \sum_{i=1}^n k_i v_i, k_i \in \mathbb{Z}\}$ for some linearly independent vectors v_i. (Without loss of generality we will assume that $n = d$ since we can consider the space over the reals spanned by the generators of the lattice.) Hence, all crystals can be identified with $\mathbb{Z}^d$. The identification is not unique, but we will assume it is done and fixed through the proof. This allows us to define the distance between two points in the crystal using the absolute value of points in $\mathbb{Z}^d$ (The sum of the absolute value of the coordinates) Again, this is a non-canonical choice and many others would have worked, but we fix it for the sake of definiteness.

We will denote by Γ the points of a lattice and denote by G a group of translations that leaves invariant. Clearly, G is a subgroup of the lattice and indeed we will find it useful to consider the case where it is a non-trivial subgroup of finite index.

If k is an element of the lattice, we denote by T_k the translation by k on the crystal. That is $T_k x = x + k$.

For each subset B of Γ we denote by $|B|$ the number of elements of B, and by $\partial_n B$ the set:

$$\partial_n B \equiv \{x \in \Gamma - B \ : \ \mathrm{dist}(x, B) \le n\} \cup \{x \in B \ : \ \mathrm{dist}(x, \Gamma - B) \le n\}$$

We write that $\lim_n \Lambda_n = \Gamma$ if for any site $j \in \Gamma$ there is an $N(j)$ such that $n > N(j)$ implies $j \in \Lambda_n$.

A particularly important role in the argument will be played by collections of sets that cover the whole space but are hierarchically ordered in sizes of different scales. The sets of one scale can be decomposed into pieces which are sets of the smaller scale. (This is quite similar to the usual Calderón-Zygmund decompositions of harmonic analysis)

For example, using the representation of the crystal as $\mathbb{Z}^d$, we can consider the collection $\mathcal{C}$ of finite subsets Λ which are of the form:

$$\Lambda = [k_1 - l_m, k_1 + l_m] \times [k_2 - l_m, k_2 + l_m] \times \cdots [k_n - l_m, k_n + l_m]$$

for some $k_1, \dots, k_n \in \mathbb{Z}$, $l_m = (3^{m+1} - 1)/2$. In other words, cubes of sizes l_m growing exponentially and centered at any point.

It is important to note that since $[k - l_m, k + l_m] = [(k - l_m + l_{m-1}) - l_{m-1}, (k - l_m + l_{m-1}) + l_{m-1}] \cup [(k - l_m + 3l_{m-1} + 1) - l_{m-1}, (k - l_m + 3l_{m-1} + 1) + l_{m-1}] \cup [(k - l_m + 5l_{m-1} + 2) - l_{m-1}, (k - l_m + 5l_{m-1} + 2) + l_{m-1}]$ (Recall that $-l_m + 6l_{m-1} + 2 = l_m$.) we conclude that the cubes of size l_m can be divided into 3^n disjoint cubes of size l_{m-1}.

We also note that as the sizes of the cubes grow, their boundaries are negligible with respect to their volume. This property will also play an important role.

We will record these in a definition to make clear what properties of the collection $\mathcal{C}$ will be needed later. This may lead to developing an Aubry-Mather theory in a more abstract setting. We also note that very similar properties play a role in the theory of the thermodynamic limit (See e.g., [Ru]).

Definition 3.1. *We say that a collection $\mathcal{C}$ of finite subsets of a lattice Γ admits an inflation rule if:*

i) *If $\{\Lambda_n\}_{n \in \mathbb{N}} \subset \mathcal{C}$ and $|\Lambda_n| \to \infty$, then there are translations $k_n \in G$ such that $\lim_n T_{k_n}(\Lambda_n) = \Gamma$.*
ii) *There exist $N, L \in \mathbb{N}$ and $M \in \mathbb{R}$ such that for any set $\Lambda \in \mathcal{C}$ with $|\Lambda| > L$, we can find N mutually disjoint sets $\Lambda_1, \cdots, \Lambda_N \in \mathcal{C}$, $|\Lambda_i| > M|\Lambda|$, such that $\Lambda = \Lambda_1 \cup \Lambda_2, \cup \cdots, \cup \Lambda_N$.*
iii) *For every $n \in \mathbb{N}$, if we define $\phi_n(m) \equiv \max\{|\partial_n \Lambda| \ : \ |\Lambda| \le m, \ \Lambda \in \mathcal{C}\}$, we have $\lim_{m \to \infty} \phi_n(m)/m^\alpha = 0$, for any $\alpha > \log(M)/\log(N)$ for the N, M in* ii).

It is obvious that for the M, N in $ii)$ we should have $M \ge N$. The difference between M and N is a measure of the difficulty of fitting sets of the same size in $\mathcal{C}$ together. For example, if we consider the sets of cubes with all possible translations, we can take $M = N$. For the set consisting of up and down triangles of size at least S on a two dimensional hexagonal lattice we can take $N = 4$ and $M = 4 + O(1/S)$.

For the above two cases it is easy to verify condition $ii)$ with $\alpha = 1$, but nevertheless for fixed n and sufficiently large sets, $|\partial_n \Lambda| \approx n|\delta_n \Lambda| \approx n|\Lambda|^{1-1/\nu}|$ where ν is the dimension of the lattice.

Aubry-Mather theory will assert the existence of real valued functions on the lattice satisfying certain properties. Again, we point out that we will use the physically motivated name "configurations" often.

Definition 3.2. *A configuration is a map $u : \Gamma \mapsto \mathbb{R}$. We define sums of configurations and their multiplication by numbers in the usual way for maps.*

Our next goal is to find a definition, analogous to that of Birkhoff orbits, that generalizes to higher dimensional situations. We observe that the essential part of the definition of Birkhoff orbits is that translations preserve order properties.

In what follows, we note that the group of translations that we consider may be a subgroup of finite index of the full group of translations of the lattice. To emphasize that, we will denote it by $\tilde{G}$.

Definition 3.3. *If we have a crystal Γ, a subgroup $\tilde{G}$ — of finite index — of the group G of translations of the lattice, we will denote the action of $\tilde{G}$ acting on configurations by $(\mathcal{T}_k u)_i = u_{i+k}$. Similarly, for $\ell \in \mathbb{Z}$ we denote by $\mathcal{R}_\ell$ the mapping defined by $(\mathcal{R}_\ell u)_k = u_k + \ell$.*

Note that $\mathcal{T}_k$ are linear operators and the $\mathcal{R}_\ell$ are affine.

An immediate consequence of these definitions is:

Proposition 3.4. *$\mathcal{T}_k \mathcal{T}_{k'} = \mathcal{T}_{k'} \mathcal{T}_k = \mathcal{T}_{k+k'}$, $\mathcal{R}_\ell \mathcal{R}_{\ell'} = \mathcal{R}_{\ell'} \mathcal{R}_\ell = \mathcal{R}_{\ell+\ell'}$, and, for k in $\tilde{G}$, $\mathcal{T}_k \mathcal{R}_\ell = \mathcal{R}_\ell \mathcal{T}_k$.*

Definition 3.5. *Given two configurations u, u', we write $u \le u'$ if $u_i \le u'_i$ for all $i \in \Gamma$, with analogous definitions for any of the other comparison symbols $\ge$, $<$ or $>$.*

Notice that these comparisons among configurations are not total orders (except in the trivial case where Γ consists of one element). That is, there are pairs of configurations which do not satisfy any of the comparisons.

This order works very nicely with all the other elements of structure that we have introduced so far (actions of translations, addition of configurations)

Proposition 3.6. *If $u \le u'$, then for every $k \in \tilde{G}$ and every $\ell \in \mathbb{Z}$ and for any configuration u'' we have: $\mathcal{T}_k u \le \mathcal{T}_k u'$, $\mathcal{R}_\ell u \le \mathcal{R}_\ell u'$, and $u + u'' \le u' + u''$.*

Definition 3.7. *Given a vector $\omega \in \mathbb{R}^n$, we say that a configuration u is of type ω if*

$$\sup_{j\in\Gamma} ||u_j - \omega \cdot j|| < \infty \tag{3.1}$$

where we denote by $\omega \cdot j$ the usual scalar product. We denote by $\mathcal{O}_\omega$ the set of configurations of type ω.

We refer to the configuration u^ω defined by $u^\omega_k = \omega \cdot k$ as the plane wave of frequency vector ω.

We observe that $\mathcal{T}_k \mathcal{R}_\ell(\mathcal{O}_\omega) = \mathcal{O}_\omega$. Notice also that $\mathcal{O}_\omega$ is an affine space modeled on $\ell^\infty(\Gamma, \mathbb{R}) \equiv \{u : \Gamma \to \mathbb{R} \ : \ ||u|| \equiv \sup_{j\in\mathbb{Z}^n} |u_j| < \infty.\}$ Part of the difficulties of the theory arise from the fact that ℓ^∞ is a space ill suited for many constructions in the calculus of variations.

Nevertheless, there is a natural way to restore some control at ∞. The following is a definition that has played an important role in dynamical systems.

Definition 3.8. *Given a discrete subgroup $\tilde{G}$ of translations leaving invariant a crystal Γ, we say that a configuration $u \in \mathcal{O}_\omega$ is a Birkhoff configuration if for every $k \in \tilde{G}$, every $\ell \in \mathbb{Z}$ we have either $\mathcal{T}_k \mathcal{R}_\ell u \ge u$ or $\mathcal{T}_k \mathcal{R}_\ell u \le u$*

We denote by $\mathcal{B}_\omega$ the set of Birkhoff configurations in $\mathcal{O}_\omega$. We note that this set is not empty (e.g., the plane wave u^ω belongs to it).

Note that the concept of Birkhoff depends on the subgroup $\tilde{G}$ that we are considering. We will omit this from the notation when it does not lead to confusion. Usually in Aubry-Mather theory, one just takes $\tilde{G} = G$.

Proposition 3.9. *Let $u \in \mathcal{B}_\omega$. If for some $k \in \tilde{G}$, $\ell \in \mathbb{Z}$, $\omega \cdot k + \ell < 0$. Then $\mathcal{T}_k \mathcal{R}_\ell u \leq u$. If $\omega \cdot k + \ell > 0$ then $\mathcal{T}_k \mathcal{R}_\ell u \geq u$.*

(In other words, the choice of sign in the Birkhoff definition is the same as for the plane wave of frequency ω. In particular, it is the same for all the functions in $\mathcal{B}_\omega$)

Proof. We will only prove the first case. Assume the hypothesis was true and the conclusion false. By the Birkhoff property, if $\mathcal{T}_k \mathcal{R}_\ell u \geq u$ is false then, we should have $\mathcal{T}_k \mathcal{R}_\ell u \leq u$. If $\mathcal{T}_k \mathcal{R}_\ell u \leq u$ then, for any natural number n we also have $(\mathcal{T}_k \mathcal{R}_\ell)^n u \leq u$.

If we compare with the plane wave solution, using that $(\mathcal{T}_k \mathcal{R}_\ell)^n u^\omega = u^\omega + (k \cdot \omega + \ell) n$, we have using that $\mathcal{T}_k$ is linear and $\mathcal{R}_\ell$ affine

$$(\mathcal{T}_k \mathcal{R}_\ell)^n u - (\mathcal{T}_k \mathcal{R}_\ell)^n u^\omega = (\mathcal{T}_k \mathcal{R}_\ell)^n u - u + (u - u^\omega) - n(k \cdot \omega + \ell) \leq (u - u^\omega) - n(k \cdot \omega + \ell).$$

On the other hand $\inf_j [(\mathcal{T}_k \mathcal{R}_\ell)^n u]_j - [(\mathcal{T}_k \mathcal{R}_\ell)^n u^\omega]_j = \inf_j u_j - u_j^\omega$

This is the desired contradiction. □

As a corollary, we obtain a priori bounds for the oscillation of the Birkhoff configurations uniform on the function.

Proposition 3.10. *For every $k \in \tilde{G}$, there exist ℓ^-, ℓ^+ such that for every $u \in \mathcal{B}_\omega$, $\mathcal{R}_{\ell^-} u \leq \mathcal{T}_k u \leq \mathcal{R}_{\ell^+} u$.*

Actually, this is the main property of the set of Birkhoff orbits that we will use. Notice that its strength comes from the fact that ℓ^-, ℓ^+ depend only on k and ω and not on the function $u \in \mathcal{B}_\omega$.

If $\omega \cdot k$ is not an integer, we can get ℓ^+ and ℓ^- to differ by 1. If $\omega \cdot k$ happens to be an integer, we can get them to differ by 2.

Another property of Birkhoff sets we will use is:

Proposition 3.11. *$\mathcal{T}_k \mathcal{R}_\ell \mathcal{B}_\omega = \mathcal{B}_\omega$, for $k \in \tilde{G}$.*

From this we deduce that if we define an equivalence relation among configurations by $u \mathcal{R} u'$ if and only if $u = \mathcal{R}_\ell u'$ for some ℓ, it is possible to restrict this relation to $\mathcal{B}_\omega$ and hence, consider the set $\mathcal{B}_\omega / \mathcal{R}$. This set will play an important role. Note that we can identify an element in $\mathcal{B}_\omega / \mathcal{R}$ with a Birkhoff configuration u such that $u_0 \in [0, 1)$. Then, Proposition 3.10 gives us a priori bounds on the values at all points.

The previous definitions involve only the geometry of configurations. To obtain meaningful physical models it is necessary to define interactions among the particles. The following definition is standard in statistical mechanics. See e.g., [Ru].

Definition 3.12. *An interaction is a collection of maps $\{H_B \ : \ B \subset \Gamma,\ B \text{ finite}\}$ that to each configuration u associate a number $H_B(u)$ which depends only on the restriction of u to B.*

Definition 3.13. *We will say that an interaction is finite range when we can find $L \in \mathbb{N}$ such that $H_B \equiv 0$ if $|B| \geq L$. We will say that an interaction is invariant under the group $\tilde{G}$ of translations if for all the translations $k \in \tilde{G}$, all B, and all* *configurations u, $H_{\mathcal{T}_k B}(\mathcal{T}_k u) = H_B(u)$. We will say that an interaction has periodic phase space if for all $\ell \in \mathbb{Z}$, all B, and all configurations u,*

$$H_B(\mathcal{R}_\ell u) = H_B(u). \tag{3.2}$$

One consequence of (3.2) is that the maps H_B can be considered not just as maps $\mathcal{B}_\omega \to \mathbb{R}$ but rather as $\mathcal{B}_\omega/\mathcal{R} \to \mathbb{R}$.

Notice that the Frenkel-Kontorova models correspond to taking $H_{\{i\}}(u) = V(u_i)$, $H_{\{i,j\}}(u) = \frac{1}{2n}(u_i - u_j)^2$ when $|i-j| = 1$, and all other $H_B(u)$ equal to zero. They are translation invariant and have periodic phase space.

The variational principle we will consider comes from the formal action:

$$\mathcal{S}(u) = \sum_{\substack{B\subset\Gamma \\ B\,\text{finite}}} H_B(u) \tag{3.3}$$

The variational equations corresponding to the formal variational principle above are:

$$\sum_{B\ni i} \frac{\partial}{\partial u_i} H_B(u) = 0, \quad \text{for all } i \in \Gamma \tag{3.4}$$

The equations (3.4) are well defined whenever the sums converge in a sufficiently strong sense. This motivates the following definitions.

Definition 3.14. *We say that an interaction is r-bounded on the configurations in the convex set $\mathcal{O}$ of configurations if:*

$$||H||_r \equiv \sup_{u\in\mathcal{O}} \sup_{j_1} \sum_{j_2,\cdots,j_r} \Big| \sum_{B\ni j_1,\cdots j_r} \frac{\partial}{\partial u_{j_1}} \cdots \frac{\partial}{\partial u_{j_r}} H_B(u) \Big| < \infty \tag{3.5}$$

Remark. Note that we have chosen the definition in such a way that $||H||_r$ depends only on the derivatives of H of order exactly r. In particular, assuming that a function is r-bounded does not afford any control on derivatives of order less than r.

Remark. Note that the equations (3.4) make perfectly good sense when the interaction is 1 bounded.

Remark. Notice that the property of being 1 bounded depends on the set of configurations we are considering. For example, if we consider the Frenkel-Kontorova model, it is 1 bounded on sets of the form $\mathcal{O}_\omega^K = \{u \ : \ \sup_i |u_i - \omega \cdot i| \le K < \infty\}$ (Hence, it is 1 bounded in $\mathcal{B}_\omega$) Nevertheless, it is not 1 bounded on $\mathcal{O}_\omega$. We will still omit the dependence on the set when it is obvious to which set we are referring.

Notice that by Proposition 3.10 the set $\mathcal{B}_\omega$ is contained in $\mathcal{O}_\omega^2 + \mathbb{N}$. Hence the Frenkel-Kontorova-model will be 1-bounded on $\mathcal{B}_\omega$.

The reason for introducing these semi-norms is that the variational equations can be written as $\mathcal{F}_i(u) = 0$ where $\mathcal{F}_i(u) = \sum_{B\in i} \frac{\partial}{\partial u_i} H_B(u)$. In the cases that we will be interested in, $\mathcal{O}$ will be an affine space over ℓ^∞ and the conditions (3.5) are conditions that ensure that the $r-1$ derivative of $\mathcal{F}$ exists and is uniformly bounded in the sense of derivatives in the Banach space ℓ^∞.

In particular we have:

Corollary 3.15. *If we define $\mathcal{F}$ as above, we have $||\mathcal{F}(u) - \mathcal{F}(\tilde{u})||_{\ell^\infty} \leq ||H||_2||u - \tilde{u}||_{\ell^\infty}$.*

Finally, we will need an extra hypothesis which is analogous to the twist condition in Hamiltonian mechanics and to ferromagnetism in statistical mechanics.

Definition 3.16. *We will say that an interaction which is 2 bounded on $\mathcal{O}$ satisfies the twist condition — or is ferromagnetic — if for all configurations $u \in \mathcal{O}$, and all $j \neq j' \in \Gamma$,*

$$\sum_{B \ni j} \frac{\partial^2}{\partial u_j \partial u_{j'}} H_B(u) \geq 0 \tag{3.6}$$

The property (3.6) is obviously implied by the stronger one:

$$\frac{\partial^2}{\partial u_j \partial u_{j'}} H_B(u) \geq 0 \tag{3.7}$$

for all $j \neq j' \in \Gamma$, $B \subset \Gamma$. We refer to (3.7) as the strong twist condition or the strong ferromagnetic condition.

(Notice that the Frenkel-Kontorova models satisfy (3.7).)

Theorem 3.17. *Let Γ be a crystal with group of translations G. Let H be an interaction invariant under the group $\tilde{G}$ of finite index in G, 2 bounded on $\mathcal{O}^3_\omega$. Assume furthermore either:*

a) *H satisfies the twist condition and is finite range.*
b) *H satisfies the strong twist condition.*

Then, there is a solution of (3.4) which lies in $\mathcal{B}_\omega$.

The crux of the proof will be to give a proof using a). We also note that the hypothesis of boundedness of the flow can be weakened considerably. See the remarks at the end of the proof for the details.

4. Proof of Theorem 3.17

Our first task will be to prove Theorem 3.17 under the extra assumption that the interaction is of finite range and that it is 3-bounded. Later, we will use this result to prove the full result by approximating our original problem by 3-bounded finite range models.

The proof of this weak version of Theorem 3.17 will follow, roughly, the scheme in [Go]. We will derive a contradiction to the assumption that there are no critical points of (3.4) of the desired type; namely, we will show that if there is no critical point of (3.4) in $\mathcal{B}_\omega$, then we can find a map $T : \mathcal{B}_\omega \to \mathcal{B}_\omega$, and a continuous function $G : \mathcal{B}_\omega \to \mathbb{R}$, such that $G(T(u)) \geq G(u) + \delta$ for some $\delta > 0$, and this will lead to a contradiction. The argument in [Go] uses the fact that if a configuration u_0 is not a critical point of (1.2) we can find a G such that $G(T(u)) \geq G(u)$ with the inequality being strict for u in a neighborhood of u_0. Unfortunately, some steps of the argument of [Go] do not generalize when the dimension of "time" is greater than one and we have to use a more indirect argument.

Lemma 4.1. *$\mathcal{B}_\omega/\mathcal{R} \subset \mathcal{O}_\omega/\mathcal{R}$ is compact when $\mathcal{O}_\omega/\mathcal{R}$ is given the topology generated by the following basis: $\{O_{B,A} : B$ a finite subset of Γ, A an open subset of $\mathbb{R}^d\}$, where $\{O_{B,A}\} \equiv \{\hat{u} \ : \ (u_j - \omega \cdot j) \subset A$ for $j \in B$, for some $u \in \hat{u}\}$. (We will refer to this topology as the topology of component-wise convergence since convergence in it is just convergence component by component.)*

Proof. If we pick p representatives $\{i_1, \cdots, i_p\}$ of the classes $\Gamma/\tilde{G}$, — here we use that $\tilde{G}$ is of finite index in G —, given a configuration u we can apply one and only one $\mathcal{R}_\ell$ in such a way that $0 < u_{i_\alpha} <= 1$.

This means that we can identify $\mathcal{B}_\omega/\mathcal{R}$ as the set of configurations u in $\mathcal{B}_\omega$ which satisfy $u_{i_\alpha} \in (0, 1]$ if we identify the ends of $(0, 1]$.

If a configuration $u \in \mathcal{O}_\omega$ is Birkhoff by Proposition 3.10

$$[\omega \cdot k] + u_{i_\alpha} \le u_{i_\alpha + k} \le [\omega \cdot k] + 1 + u_{i_\alpha} \tag{4.1}$$

for every $\alpha = 1, \cdots, p$ and every $k \in \tilde{G}$, where $[\quad]$ denotes integer part. Since, by assumption, the $i_1, \cdots, i_p$ cover all the conjugacy classes of $\Gamma/\tilde{G}$, we see that if an orbit belongs to $\mathcal{B}_\omega$ and the u_{i_α} remain bounded, then the values at each site have a bounded range. Applying Tychonov's theorem, we can conclude that $\mathcal{B}_\omega/\mathcal{R}$ is precompact when we give it the topology of point-wise convergence. By the definition of Birkhoff orbit, it is clear that $\mathcal{B}_\omega$ is closed in this topology since it is expressed as the intersection of conditions on each coordinate that are preserved under point-wise limits. $\square$

Lemma 4.2. *Consider the map $\mathcal{F}$ defined on $\mathcal{O}^3_\omega$ by:*

$$[\mathcal{F}(u)]_j = -\sum_{B \ni j} \frac{\partial}{\partial u_j} H_B(u); \tag{4.2}$$

If H is 1-bounded on bounded sets of $\mathcal{O}_\omega$, then $\mathcal{F}$ maps $\mathcal{O}_\omega$ into ℓ^∞. Moreover if H is $r-$differentiably bounded, then $\mathcal{F}^i$ is a $r - 1$ map from $\mathcal{O}_\omega$ to ℓ^∞, and the norm of the derivatives of order $r - 1$ is uniformly bounded.

Proof. The proof just consists of restating the definition of 1 bounded and noting that if an interaction is $1, 2$ bounded the derivative of the right hand side of (4.2) can be obtained by taking derivatives term by term in the sum. $\square$

Remark. Notice that $\mathcal{F}$ is formally the derivative of the variational principle (3.3). Notice also that in the Frenkel-Kontorova case it reduces to an analogue of the heat equation in which the Laplacian is discrete.

If we remember that $\mathcal{O}_\omega$ is an affine space modeled on ℓ^∞, the usual existence theorem of differential equations implies that we can define a local flow $\Phi_t : \mathcal{O}_\omega \to \mathcal{O}_\omega$ satisfying

$$\frac{d}{dt}\Phi_t(u) = \mathcal{F}(\Phi_t(u)). \tag{4.3}$$

From the fact that the right hand side of (4.2), is uniformly Lipschitz, we conclude that the solutions are defined for all time.

Using the hypothesis about the symmetry of the coefficients of the interaction we obtain that for $k \in \tilde{G}$, $q \in \mathbb{Z}$.

$$\mathcal{T}_k R_q \mathcal{F} = \mathcal{F} \mathcal{T}_k. \tag{4.4}$$

Hence, using the uniqueness of solutions of the differential equations defining the flow, we also have for $k \in \tilde{G}$ and $q \in \mathbb{Z}$:

$$\mathcal{T}_k \mathcal{R}_q \Phi_t = \Phi_t \mathcal{T}_k \mathcal{R}_q \tag{4.5}$$

We show next that the partial orders of Definition 3.5 are also preserved by the flow when the interaction is ferromagnetic.

Lemma 4.3. *Let $\mathcal{F}$ be defined as before and let H be a ferromagnetic interaction. Then if $u \le \tilde{u}$ (resp. $u \ge \tilde{u}$) we have $\Phi_t(u) \le \Phi_t(\tilde{u})$ (resp. $\Phi_t(u) \ge \Phi_t(\tilde{u})$).*

Proof. It suffices to prove the result in the case $u \le \tilde{u}$. Define $u_\lambda = \lambda u + (1-\lambda)\tilde{u}$. By the theorem on smooth dependence on parameters of solutions of O.D.E.'s in Banach space, we have that $\Phi_t(u_\lambda)$ is C^{r-2} with respect to parameters if H is $1,\dots,r$-bounded. (In finite dimensions, one does not need an extra derivative since, in that case, continuous functions on closed bounded sets are uniformly continuous.) Hence, if $r \ge 3$ we have that $\Phi_t(u_\lambda)$ is C^1 and, hence, to prove the conclusion it suffices to show $\frac{d}{d\lambda}\Phi_t(u_\lambda) \le 0$.

Recall that if we consider $\frac{d}{d\lambda}\Phi_t(u_\lambda)$ as a function of t it satisfies the variational equations:

$$\begin{aligned} &\frac{d}{dt}\left[\frac{d}{d\lambda}\Phi_t(u_\lambda)\right] = D\mathcal{F}(\Phi_t(u_\lambda)\left[\frac{d}{d\lambda}\Phi_t(u_\lambda)\right] \\ &\frac{d}{d\lambda}\Phi_t(u_\lambda)\big|_{t=0} = u - \tilde{u} \end{aligned} \tag{4.6}$$

Observe that, for fixed t and λ, $D\mathcal{F}(\Phi_t(u))$ is a bounded linear operator on $\ell^\infty(\Gamma, \mathbb{R})$. And note that in this representation the matrix elements of $D\mathcal{F}(\Phi_t(u_\lambda))$ are positive if they are off the diagonal, and are bounded on the diagonal.

Let $M(t) = D\mathcal{F}(\Phi_t(u_\lambda))$, D_t be the diagonal part of $M(t)$ and $N(t)$ be the non-diagonal part. If H is C^3 it follows that the mappings $t \mapsto M(t), D(t), N(t)$ are C^1 and the derivatives are uniformly bounded. Following the method of variation of constants, we try to write the solution of (4.6) as $\exp\left[\int_0^t D(s)ds\right] C(t)$. Since all the diagonal elements commute among themselves, we have:

$$\frac{d}{dt}\exp\left[\int_0^t D(s)ds\right] = \exp\left[\int_0^t D(s)ds\right] D(t) = D(t)\exp\left[\int_0^t D(s)ds\right].$$

Hence equation (4.6) becomes:

$$\begin{aligned} &\frac{d}{dt}C(t) = \left(\exp\left[-\int_0^t D(s)ds\right] N(t) \exp\left[-\int_0^t D(s)ds\right]\right) C(t) \\ &C(0) = u - \tilde{u} \end{aligned} \tag{4.7}$$

Noting that $\exp\left[-\int_0^t D(s)ds\right]$, $N(t)$, $\exp\left[-\int_0^t D(s)ds\right]$ all have non-negative matrix elements, it follows that the initial value problem (4.7) is equivalent to:

$$C(t) = (u - \tilde{u}) + \int_0^t R(s)C(s)ds \tag{4.8}$$

where $R(s)$ is a matrix that has non-negative entries. As is standard, we obtain that the R.H.S. of (4.8) can be considered an operator acting on continuous functions defined on an interval $[0, T]$. It is a contraction if this space of continuous functions has the metric $d(C, \tilde{C}) = \sup_{t\in[0,T]} ||C(t) - \tilde{C}(t)||e^{At}$, where A is sufficiently large.

From the fact that $R(s)$ has positive entries and that $u - \tilde{u} \le 0$ it follows that the operator defined by the R.H.S. of (4.8) preserves the set of C's such that $C(t) \ge 0$ for all t. And finally, since the solution can be obtained by iterating the R.H.S. of (4.8) with starting point $u - \tilde{u}$, it follows that the solution is negative and the proof of Lemma 4.3 is complete. □

Since the fact that a configuration is Birkhoff can be expressed in terms of the translations that commute with the heat flow, and of the order (which we have shown is preserved), we obtain:

Corollary 4.4. *Under the above conditions,*

$$\Phi_t(\mathcal{B}_\omega) \subset \mathcal{B}_\omega.$$

Moreover, the flow Φ_t can be defined on $\mathcal{B}_\omega/\mathcal{R}$.

Given any finite $\Lambda \subset \Gamma$ we define:

$$\mathcal{S}_\Lambda(u) = \sum_{B\cap\Lambda\neq\emptyset} H_B(u).$$

We want to compute the derivative of S_Λ along the flow Φ_t. Note that Φ_t is the gradient flow for S, and that outside of the boundary of Λ the gradient flows of S_Λ and S are the same. Therefore, it is reasonable to we expect that the derivative is the square of the gradient up to terms that are related to the boundary. This can be shown by a simple calculation.

$$\begin{aligned}
\frac{d}{dt} S_\Lambda\left(\Phi_t(u)\right)\Big|_{t=0} &= \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B} \frac{\partial H_B(u)}{\partial u_j}\mathcal{F}(u)_j(u) \\
&= \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B\cap\Lambda} \frac{\partial H_B(u)}{\partial u_j}\mathcal{F}(u)_j(u) + \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B/\Lambda} \frac{\partial H_B(u)}{\partial u_j}\mathcal{F}(u)_j(u) \\
&= \sum_{j\in\Lambda}\mathcal{F}(u)_j \sum_{B\ni j}\frac{\partial H_B(u)}{\partial u_j} + \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B/\Lambda} \frac{\partial H_B(u)}{\partial u_j}\mathcal{F}(u)_j(u) = \\
&= -\sum_{j\in\Lambda}|\mathcal{F}(u)_j(u)|^2 + \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B/\Lambda} \frac{\partial H_B(u)}{\partial u_j}\mathcal{F}(u)_j(u)\ .
\end{aligned}$$

Notice that on the last line the first term is obviously non positive and that the second one only involves boundary terms. Intuitively, the first term should dominate since it is a term that depends on the bulk while the other depends only on the boundary. In the case $m = 1$ and nearest neighborhood interactions this was proved in [Go]. Nevertheless, that proof does not generalize to higher dimensions since it relies on the fact that there are only two boundary terms independently of the size of the cube. Fortunately, for our purposes considerably less is needed.

Lemma 4.5. *Let $\mathcal{C}$ be a collection of subsets of Γ as in Definition 3.1. Assume that there is no critical point in $\mathcal{B}_\omega$. Then we can find $\epsilon_0 > 0$ and $L_0 \in \mathbb{N}$ such that for all configurations $u \in \mathcal{B}_\omega$, all sets $\Lambda \in \mathcal{C}$, and $|\Lambda| > L_0$, we have:*

$$\sum_{j\in\Lambda} |\mathcal{F}(u)_j|^2 \geq \epsilon_0.$$

Proof. Let $E_\Lambda(u) \equiv \sum_{j\in\Lambda} |\mathcal{F}(u)_j|^2$. Notice that $E_{\Lambda'}(u) \leq E_\Lambda(u)$ whenever $\Lambda' \subseteq \Lambda$. We now proceed by contradiction and assume, contrary to the conclusion, that we can find a sequence $u^{(n)} \in \mathcal{B}_\omega$ and a sequence Λ_n of sets in $\mathcal{C}$ with $|\Lambda_n| \to \infty$ such that

$$E_{\Lambda_n}(u^{(n)}) \longrightarrow 0 \qquad \text{as } n \to \infty.$$

By assumption $i)$ in Definition 3.1, we can find $k_n \in \tilde{G}$ such that $T_{k_n}\Lambda_n$ converges to Γ. Since E_Λ is invariant under translations in $\tilde{G}$, (i.e. $E_\Lambda(u) = E_{T_k\Lambda}(\mathcal{T}_k u)$ whenever $k \in \tilde{G}$), if we consider $\Lambda'_n = T_{k_n}\Lambda_n$ and $u'^{(n)} = \mathcal{T}_{k_n} u^{(n)}$ we have obtained a sequence of sets ${\Lambda'}_n$ converging to Γ and a sequence of configurations $u'^{(n)} \in \mathcal{B}_\omega$ such that $E(\Lambda'_n)(u'^{(n)})$ converges to zero.

By the compactness of $\mathcal{B}_\omega/\mathcal{R}$ we can find a subsequence $u'^{(n_k)}$ that converges (in $\mathcal{B}_\omega/\mathcal{R}$) to $u^{(\infty)}$. We want to conclude that $u^{(\infty)}$ is a critical point. (This would contradict the assumption that there is no critical point and hence prove Lemma 4.5.

In effect, if we fix $j \in \Gamma$ we have

$$|\mathcal{F}(u^{(\infty)})_j|^2 = \lim_k |\mathcal{F}(u'^{n_k})_j|^2 \leq \lim_k E(\Lambda'_{n_k})(u'^{(n_k)})$$

The first equality is true because $\mathcal{F}_j$ depends only on the value of the configuration on a finite number of sites and we have point-wise convergence. The second holds because E is a sum of non-negative terms one of which is eventually $|\mathcal{F}_j|^2$; since the limit of the R.H.S. is zero, we conclude that $\mathcal{F}_j = 0$ for any i, j. Hence, $u^{(\infty)}$ is a critical point and our contradiction is established. □

Corollary 4.6. *Assume there are no critical points in $\mathcal{B}_\omega$. Then we can find $\epsilon^* > 0$ and $L^* \in \mathbb{N}$ such that for all configurations $u \in \mathcal{B}_\omega$, and all sets in $\mathcal{C}$ of size longer than L^*, we have*

$$\sum_{j\in\Lambda; i} |\mathcal{F}(u)_j|^2 \geq \epsilon^* |\Lambda|^\alpha$$

where α is as in Definition 3.1.

Proof. Let L^* be as in Lemma 4.5. Applying repeatedly part $iii)$ of Definition 3.1, we obtain that all sets Λ of sufficiently large size can be broken into N^k disjoint pieces contained in $\mathcal{C}$, each of which has size larger than $|\Lambda|/M^k$. If we choose $k = [\log(|\Lambda|/N*)/\log(N)]$ we conclude that we can divide the set into at least $K|\Lambda|^\alpha$ disjoint pieces each of which is of size at least N^*. Since the sum can be divided into the sum of each of these pieces, and Lemma 4.5 tells us that they are bounded uniformly away from zero, the claim is established. $\square$

Continuing the argument to prove Theorem 3.17 for finite range systems, we have on the other hand that if the range of the interaction is r and Λ is a set in $\mathcal{C}$:

$$\left|\sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B/\Lambda;i}\frac{\partial H_B(u)}{\partial u_j^i}\mathcal{F}(u)_j\right| \leq \sum_{B\cap\Lambda\neq\emptyset}\sum_{j\in B/\Lambda;i}\left|\frac{\partial H_B(u)}{\partial u_j}\right| M_1 \leq K\phi(r,|\Lambda|)$$

where ϕ has the same meaning as in Definition 3.1. Since by assumption $iii)$ of Definition 3.1 this is negligible compared with the lower bounds in Corollary 4.6, we obtain:

Corollary 4.7. *If there are no critical points in $\mathcal{B}_\omega$, then for all large enough sets Λ of $\mathcal{C}$ we have:*

$$\frac{d}{dt}S_\Lambda(u)\big|_{t=0} \geq K(|\Lambda|) > 0$$

where $K(|\Lambda|)$ does not depend on u.

Furthermore, since $\frac{d^2}{dt^2}\Phi_t(u)$ is a continuous function on $\mathcal{B}_\omega/\mathcal{R}$, which is a compact set, we can bound it uniformly in u. It follows that if there is no critical point in $\mathcal{B}_\omega$ we can find $t_0 > 0$ and a cube $\Lambda \subset \mathbb{Z}^m$ such that

$$S_\Lambda(\Phi_{t_0}(u)) \geq S_\Lambda(u) + \delta$$

with $\delta > 0$. This is impossible since $\mathcal{S}_\Lambda$ is a continuous function on the compact set.

From this contradiction we conclude there is a critical point in $\mathcal{B}_\omega$ and Theorem 3.17 is proved under the extra assumption that H is of finite range and 3 bounded.

To conclude the proof of Theorem 3.17 as stated, we just have to show that a $1, 2$ bounded interaction can be approximated by a sequence of interactions for which the theorem as proved so far applies in such a way that the critical points thus produced converge to a critical point of the interaction.

Lemma 4.8. *Let $\mathcal{F}^n$ be a sequence of vector fields derived as in (4.2) from an interaction H^n satisfying condition (3.2) (periodic phase space). Assume that each of the $\mathcal{F}$'s admits a critical point in $\mathcal{B}_\omega$. Assume furthermore that there is another vector field $\mathcal{F}^\infty$ coming from an interaction H^∞ satisfying condition (3.2). Finally, assume:*

i) *For every $i \in \Gamma$, every $u \in \mathcal{B}_\omega$, $\mathcal{F}^n(u)_i \to \mathcal{F}^\infty(u)_i$*

ii) $\sum_l \sup_n \sum_{|B|=l} \sup_{u\in\mathcal{O}_\omega} \sum_{i\in\Gamma} |\frac{\partial^2}{\partial u_i \partial u_j} H_B^n(u)|$

Then $\mathcal{F}^\infty$ has a critical point in $\mathcal{B}_\omega$.

Remark. The interpretation of condition $ii)$ is that if we truncate any of the interactions H^n to those interaction terms which correspond to sets of size L, the error incurred can be made arbitrarily small uniformly in n, in the sense of the $|| \quad ||_2$ semi-norm. This is a technical condition that will be easy to verify for the cases that we have in mind. The main one is when we consider a 2 bounded interaction and approximate it by finite range ones which are obtained by ignoring the interaction terms that have size larger than a certain number.

Proof. Denote by $\mathcal{F}_n$ the vector field in $\mathcal{O}_\omega$ corresponding to the interaction H^n. Denote by $\mathcal{F}_\infty$ the vector field corresponding to the limiting interaction. If u^n are the critical points corresponding to H^n, we have $\mathcal{F}_n(u^n) = 0$. Given the invariance under $\mathcal{R}$ of the interaction, by substituting $\mathcal{R}_{q_n} u^n$ for our original configurations and using (4.1), we can assume that for all $n, n' \in \mathbb{N}$ we have $||u^n - u^{n'}||_{\ell^\infty} \le 2p$, so

$$||u^n - u^\infty||_{\ell^\infty} \le 2p. \tag{4.9}$$

By the compactness of $\mathcal{B}_\omega/\mathcal{R}$ given a sequence as above, we can obtain a subsequence that converges to a limiting configuration in the sense of point-wise convergence. So that, by passing to a subsequence, we can assume without loss of generality that $u^n \to u^\infty$ in the point-wise sense and $\mathcal{F}^n(u^n) = 0$.

If we fix $i \in \Gamma$, we have for any $n, L \in \mathbb{N}$:

$$\begin{aligned} |\mathcal{F}^\infty(u^\infty)_i| =& |\mathcal{F}^\infty(u^\infty)_i - \mathcal{F}^n(u^n)_i| \\ &\le |\mathcal{F}^\infty(u^\infty)_i - \mathcal{F}^n(u^\infty)_i| + |\mathcal{F}^n(u^\infty)_i - \mathcal{F}^n(u^n)_i| \\ &\le |\mathcal{F}^\infty(u^\infty)_i - \mathcal{F}^n(u^\infty)_i| \\ &\quad + \Big| \sum_{\substack{B \ni i \\ |B| \le L}} \frac{\partial}{\partial u_i} H_B^n(u^\infty) - \sum_{\substack{B \ni i \\ |B| \le L}} \frac{\partial}{\partial u_i} H_B^n(u^n) \Big| \\ &\quad + \sum_{\substack{B \ni i \\ |B| > L}} \left| \frac{\partial}{\partial u_i} H_B^n(u^\infty) - \frac{\partial}{\partial u_i} H_B^n(u^n) \right| \end{aligned} \tag{4.10}$$

Given any $\epsilon > 0$, by condition $ii)$ $\sup_{u \in \mathcal{O}_\omega} \sum_{|B|>L} \sum_{i'} |\frac{\partial^2}{\partial u_i \partial u_{i'}} H_B^n(u^n)| \le \epsilon/6p$. Using (4.9) and Corollary 3.15, this implies that $|\frac{\partial}{\partial u_i} H_B^n(u^\infty) - \frac{\partial}{\partial u_i} H_B^n(u^n)| \le \epsilon/3$ independently of n. Once L is chosen, we can find n large enough so that the first term in the sum in (4.10) is smaller that $\epsilon/3$, and given the uniform convergence assumption in Lemma 4.8 we can make arbitrarily small the last term in (4.10). □

To conclude the proof of Theorem 3.17, we observe that a 1, 2 bounded function of finite range can be approximated, in the sense required Lemma 4.8, by C^3 functions of finite range by using just the usual smoothing of finite dimensional variables for each of the H_B's. The twist condition is preserved since it only depends on the second derivatives of sums of of finitely many functions of finitely many variables, which we are approximating. Hence, we can use the results we have already proved to prove $a)$ of Theorem 3.17.

In the case where we have an interaction satisfying $b)$ of Theorem 3.17, we can just approximate the interaction by the cut-off interactions which agree with the original interaction on sets of size up to n and are zero for larger sets. It is easy to check that under the hypothesis of the strong twist condition all these interactions will also satisfy the strong twist condition and also converge in a sense strong enough for Lemma 4.8 to apply. This concludes the proof of Theorem 3.17. □

Remark. We note that the regularity hypothesis in Theorem 3.17 can be slightly weakened. The most delicate analytical point of the proof is that the heat flow can be defined on $\mathcal{B}_\omega$ for all time. Then, one can use the argument starting in Lemma 4.5. We concluded the existence of the flow for all time by making assumptions that implied that $\mathcal{F}$ was uniformly Lipschitz and then showed that the comparison principle and symmetries lead to the preservation of the Birkhoff character. Note that since we know that $\mathcal{B}_\omega \subset \mathcal{O}^2_\omega$, it would have sufficed to show that there is a flow defined for a positive time for conditions starting in $\mathcal{O}^3_\omega$ and that comparison held. Then, the same argument we have detailed shows that the $\mathcal{B}_\omega$ is preserved, hence, that the flow exits for all times when the conditions are Birkhoff. There are indeed some models for which this generality leads to new results, but we do not know of any physically interesting one. Nevertheless, analogues of these arguments are needed when generalizing this argument to PDE's.

5. Applications

In this section, we discuss some models that can be reduced to situations in which $G \neq \tilde{G}$. That is, the group that leaves invariant the interaction — and with respect to which we define Birkhoff orbits — is smaller than the group that leaves invariant the lattice. Besides the conceptual simplification of not identifying objects that play different roles, having these models was an important motivation.

As a first example we show how finite compositions of twist mappings can be considered in this framework. (This application had been considered by other methods in [Ma3]) This addresses the fact that even if the conclusions in the dynamical version of Aubry-Mather theory are invariant under changes of coordinates, the twist hypothesis is not. Relatedly, even if the conclusions of Aubry-Mather theory for a map f imply the same conclusions for f^n, it is not true in general that f^n satisfies the twist hypothesis. The class of maps that are finite compositions of twist maps, is clearly closed under compositions. It is a deeper result (See [Ma3]) that this class is also invariant under changes of variables since it can be given a topological characterizaton as "positive tilt" maps.

We recall given a diffeomorphisms M of $\mathbb{T}^n \times \mathbb{R}^n$ that preserves the symplectic form $\omega = \sum_i dp_i \wedge dq_i$ and is exact $(M^*(\sum_i p_i dq_i) = \sum_i p_i dq_i + dS)$ where $S : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$. If we denote by (P', Q') the image of (P, Q) under the map and assume that (Q, Q') is a good system of coordinates. (This happens if $|\det(\partial Q' \partial P)| \geq a > 0$.)

Then, we can find a generating function h such that the statement $M(P, Q) = (P', Q')$ is equivalent to $P' = \partial_2 h(Q, Q')$, $P = -\partial_1 h(Q, Q')$. In that case, a sequence Q_n is the projection of an orbit if and only if it is a critical point of the formal action $\sum_{n\in\mathbb{Z}} h(Q_n, Q_{n+1})$

If h^α, $\alpha = 1, \cdots, p$, is a finite set of these generating functions, corresponding to twist maps M_α, we see that if we consider the formal action principle:

$$\mathcal{S}(Q) = \sum_n h^\alpha(Q_{np+\alpha}, Q_{np+\alpha+1}) \tag{5.1}$$

the variational equations are:

$$\partial_1 h^\alpha(Q_{np+\alpha}, Q_{np+\alpha+1}) + \partial_2 h^{\alpha-1}(Q_{np+\alpha-\alpha}, Q_{np+\alpha}). \tag{5.2}$$

That is, if we set

$$P_{np+\alpha} = -\partial_1 h^\alpha(Q_{np+\alpha}, Q_{np+\alpha+1}), \tag{5.3}$$

the equations (5.2) imply that:

$$P_{np+\alpha+1} = \partial_2 h^\alpha(Q_{np+\alpha}, Q_{np+\alpha+1}). \tag{5.4}$$

Taking together (5.3) and (5.4), we obtain:

$$(Q_{np+\alpha+1}, P_{np+\alpha+1}) = M^\alpha(Q_{np+\alpha}, P_{np+\alpha}).$$

Hence $\{(Q_{np}, P_{np})\}$ is an orbit for $M^p \circ M^{p-1} \circ M^1$.

The variational principle for the action (5.1) can be fitted into those considered in Theorem 3.17 if we consider a one-dimensional lattice $\Gamma = \mathbb{Z}$ with group of translations $\mathbb{Z}$ and reduced group $p\mathbb{Z}$. Clearly, the action is invariant under these reduced groups. The only partial derivatives that are not zero are those corresponding to neighboring points, and in this case they reduce to the twist condition for the individual maps.

Notice that this problem is equivalent to the problem of considering the map obtained by composing the p twist maps. Nevertheless, in general the map will not satisfy the twist condition.

This construction has an analogue in higher dimensions if we associate to a site in a lattice $\mathbb{Z}^d$ a value in $\mathbb{R}^{l^{(d-1)2d}}$. Notice that there is an identification σ between $[1, l^{(d-1)2d}]$ and the boundary of Λ_l, the cube of size l. Given a state u we can associate to it the energy $V(u) = \min\{E(\tilde{u}) \ : \ \tilde{u} : \Lambda_l \to \mathbb{R}, \tilde{u}|_{\partial\Lambda_l} = u \circ \sigma\}$, where $E(\tilde{u})$ satisfies the twist conditions. The Frenkel-Kontorova model for this interaction can be identified with a one-component Frenkel-Kontorova model in which each of the sites is blown up to a cube Λ_l and the interaction between the new sites is given by the expression E if they are in the same block, and by the original Frenkel-Kontorova if they have adjacent boundaries.

In general this process leads to models in which the potential cannot be expressed as a sum of functions of the one-dimensional variables. Hence the multidimensional analogue of the twist condition is violated. This is very similar to the process that is called *"conjunction"* in [Ma3]. It can also be considered an analogue of the block spin renormalization of statistical mechanics, and indeed the renormalization group picture of [McK] can also be formulated in this language. Notice that if the dimension of the *"time"* is larger than one the renormalization process increases the number of variables. In the one-dimensional time case — the one most interesting for dynamics — the number of variables does not increase.

We also point out that an statistical mechanics interpretation of these models is molecules laid in a linear substratum. The atoms in the molecule interact with their nearest neighbors and with the substratum in a way that depends on the chemical element. Since the molecules are arranged periodically, the interactions are periodic with a period equal to the number of atoms in the molecule.

We also point out that this formalisms also allows to deal with some restricted classes of several component models. In the statistical mechanics interpretation, this would correspond to situations where the internal state of an atom is describe by several parameters.

Frenkel-Kontorova models with p components have configurations which are maps $u : \mathbb{Z}^n \to \mathbb{R}^p$. We will denote the components by superscripts and, as before, the site of the lattice by a subindex. So for a vector $i \in \mathbb{Z}^n$, u_i denotes the vector in $\mathbb{R}^p$ which is the value taken by the configuration at the site i. For an integer $j \in \{1, \cdots, p\}$, u_i^j will denote component j of the $\mathbb{R}^p$ vector u_i. Similarly, we will denote by u^j the mapping $\mathbb{Z}^n \mapsto \mathbb{R}$ that to each site associates the j^{th} component of the configuration at the site. The interaction of such a multicomponent Frenkel-Kontorova model is given by $H_{\{i\}} = V(u_i)$, where $V(x+e) = V(x)$ if e is a vector with integer components of length 1, $H_{\{i,j\}}(u) = 1/2n\ |u_i - u_j|^2$ if $|i-j| = 1$, and $H_B(u) = 0$ for any other set B.

To reduce such models to the situation discussed in Theorem 3.17, we introduce an auxiliary one-component model in which $\hat{\Gamma} = \mathbb{Z}^n \times \mathbb{Z}$. The group G of the new model will be $\mathbb{Z}^n \times \mathbb{Z}$ acting in the obvious way, and $\tilde{G} = \mathbb{Z}^n \times p\mathbb{Z}$. The configurations $\hat{u}$ in the auxiliary model can be obtained from those in the old one by setting $\hat{u}_{(i_1,\cdots,i_p,i_{p+1})} = u^{i_{p+1} \bmod p}_{(i_1,\cdots,i_p)}$. The interactions are given by:

$$\hat{H}_{\{(i_1,\cdots,i_p,pn+1),\cdots,(i_1,\cdots,i_p,pn+p)\}} = V(\hat{u}_{(i_1,\cdots,i_p,pn+1)}, \cdots, \hat{u}_{(i_1,\cdots,i_p,pn+p)})$$

If $\sum_{\alpha=1,\dots,p} |i_\alpha - j_\alpha| = 1$, we define:

$$\hat{H}_{\{(i_1,\dots,i_p,\ell),(j_1,\dots,j_p,\ell)\}} = \frac{1}{2}\left(\hat{u}_{(i_1,\cdots,i_p,\ell)} - \hat{u}_{(j_1,\cdots,j_p,\ell)}\right)^2$$

This identification can be interpreted as saying that we lump together segments of p elements in the vertical direction and consider them as a site in the original problem.

If we have a configuration $\hat{u}$ of the extended system satisfying

$$\hat{u}_{(i_1,\dots,i_p,i_{p+1})} = \hat{u}_{(i_1,\dots,i_p,i_{p+1}+p)} \tag{5.5}$$

and it satisfies the equations (3.4) the p component configuration u satisfies the equation (3.4) for the reduced model.

Given a frequency $\omega = (\omega_1, \dots, \omega_p)$ we can consider an extended frequency $\hat{\omega} = (\omega_1, \dots, \omega_p, 1/p)$. The Birkhoff configurations of this extended frequency have to satisfy the periodicity condition (5.5).

Unfortunately, these models, even if they have a natural interpretation in Statistical Mechanics, they do not have such a nice interpretation as twist maps in the annulus. (They are dynamical systems in $\mathbb{R}^{2d}$ not in $\mathbb{T}^d \times \mathbb{R}^d$.)

This can be illustrated by trying to get the Frenkel Kontorova model for functions taking values in $\mathbb{R}^d$. (when $n = 1$, this has the interpretation of a twist map in $\mathbb{T}^d \times \mathbb{R}^d$, sometimes called Froeschlé maps). To verify the twist condition we have to assume that $\frac{\partial^2}{\partial u_i \partial u_{i'}} V(u) \geq 0$ when $i' \neq i$. Unfortunately, if we want that they are functions in a torus $V(u_1+1, u_2, \cdots, u_p) = V(u_1, u_2+1, \cdots, u_p) = \cdots = V(u_1, u_2, \cdots, u_p+1) = V(u_1, u_2, \cdots, u_p)$ we have considered this implies that $V(u) = V_1(u_1) + \cdots + V_p(u_p)$, so the model reduces to uncoupled one-dimensional cases. Therefore, we conclude that this construction does not extend to being able to apply Theorem 3.17 to the Froeschlé maps.

References

[AD] S. Aubry, P. Y. Le Daeron, *The discrete Frenkel Kontorova model and its extensions: Exact results for the ground state*, Physica **8D** (1983), 240–258.

[An] S. Angenent, *Monotone recurrence relations, their Birkhoff orbits and topological entropy*, Ergod. Th. and Dyn. Syst. **10** (1990), 15–41.

[BK] D. Bernstein, A. Katok, *Birkhoff periodic orbits for small perturbations of completely integrable systems with convex Hamiltonians*, Inv. Mat. **88** (1987), 225–241.

[Ba] V. Bangert, *Mather sets for twist maps and geodesics on tori*, Dynamics Reported **1** (1988), 1–54.

[B1] M.L. Blank, *Metric properties of minimal solutions of discrete periodical variational problems*, Nonlinearity **2** (1989), 1–22.

[B2] M.L. Blank, *Chaos and order in the multidimensional Frenkel-Kontorova model*, Sov. Phy. JEPT **85** (1989), 1255–1268.

[He] M. Herman, *Inégalités " a priori" pour des tores Lagrangiens invariantes par des diff'eomorphismes symplectiques*, Pub. Mat. I.H.E.S.

[Go] C. Golé, *A new proof of Aubry-Mather's theorem*, E.T.H. preprint.

[Gr] R. Griffiths, *Frenkel-Kontorova models of commensurate-incommensurate transitions*, preprint (1991).

[McK] R.S. McKay, *Scaling exponents at the transition by breaking of analyticity for incommensurate structures*, Physica **50D** (1991), 71–79.

[Ma1] J. Mather, *A criterion for non existence of invariant circles*, Pub. Mat. I.H.E.S. **63** (1987), 153–204.

[Ma2] J. Mather, *Existence of quasiperiodic orbits for twist homeomorphisms of the annulus*, Topology **21** (1982), 457–467.

[Ma3] J. Mather, *Variational construction of orbits of twist diffeomorphisms*, Jour. Am. Math. Soc. **4** (1991), 207–263.

[Mo] J. Moser, *Minimal solutions of variational problems on a torus*, Ann. Inst. H. Poinc. An. non lin. **3** (1986), 229–272.

[Mo2] J. Moser, *A stability theorem for minimal foliations on a torus*, Ergod. Theo. Dyn. Syst. **8** (1988), 251–281.

[OV] A. Olvera, C. Vargas, *A continuation method to study periodic orbits in the Froeschlé map*, Physica **72 D** (1992,1994), 351–371.

[Ru] D. Ruelle, *Statistical Mechanics*, Benjamin, Reading MA, 1969.

[V] F. Vallet, *Thermodynamique unidimensionelle, et structures bidimensionelles de quelques modèles pour des systèmes incommensurables*, U. Paris VI thesis (1986).

Mathematics Department, The University of Texas at Austin, Austin TX 78712
E-mail address: koch@math.utexas.edu; llave@math.utexas.edu; radin@math.utexas.edu