Computer Assisted Proofs
for the FPU Model

Gianni Arioli

Miami, December 17-20, 2003
The Fermi-Pasta-Ulam model consists of $P$ particles whose dynamics is described by the equations

$$\ddot{q}_m = \phi'(q_{m+1} - q_m) - \phi'(q_m - q_{m-1}), \quad m = 1, 2, \ldots, P,$$

where $\phi(x) = \frac{x^2}{2} + \alpha \frac{x^3}{3} + \beta \frac{x^4}{4}$.

We investigate time-periodic solutions with $\alpha = 0$.

By homogeneity, it suffices to consider the case $\beta = 1$. Similarly, the fundamental period $T$ of a non-constant solution can be normalized to $2\pi$ by a rescaling of time.

This leads us to consider the equation

$$\omega^2 \ddot{q} = -\nabla^* \left[ \nabla q + (\nabla q)^3 \right],$$

where $\omega = 2\pi / T$, and where $\nabla$ and $\nabla^*$ are defined by

$$(\nabla q)_m(t) = q_{m+1}(t) - q_m(t), \quad (\nabla^* q)_m(t) = q_{m-1}(t) - q_m(t).$$
The best known periodic solutions of the $\beta$-model are near $q = 0$. In this regime, the cubic term is small compared to $q$, and we have $L_\omega q \approx 0$, where

$$L_\omega q = -\omega^2 \ddot{q} - \nabla^* \nabla q.$$ 

The values of $\omega > 0$ for which $L_\omega$ has an eigenvalue zero, and the corresponding nonzero solutions of $L_\omega q = 0$, will be referred to as resonant frequencies and normal modes, respectively.

**Proposition 1.** A frequency $\omega_n > 0$ is resonant, that is, $L_\omega$ has an eigenvalue zero, if and only if

$$\omega_n k = \pm 2 \sin(h\theta/2),$$

for some nonzero $k \in \mathbb{Z}$ and $h \in \mathcal{P} := \mathbb{Z}/(P\mathbb{Z})$. 
For every $h \in \mathcal{P}$, let $\mathbb{P}_h$ be the orthogonal projection on the $h$-th normal mode.

As a measure for the size of the $h$-th spatial mode of $q \in H^1_0$, we consider its “harmonic energy”

$$E_h(q) = E(\mathbb{P}_h'q), \quad E(q) = \frac{1}{2} \langle \dot{q}, \dot{q} \rangle + \frac{1}{2} \langle \nabla q, \nabla q \rangle.$$

These energies are not directly related to the FPU Hamiltonian (unless $\alpha = \beta = 0$), but they have the advantage of being additive, that is, the sum of $E_h(q)$ over all $h \in \mathbb{Z}/(2P\mathbb{Z})$ is equal to $E(q)$.
Theorem 1. For $P = 32$ and $\omega = 0.1989$, the equation has a set of 11 real analytic solutions, $\{f_A, f_B, \ldots, f_K\}$, with the properties listed in the following table, where $\Phi_\omega$ denotes the value of the functional, $E$ is the harmonic energy of the given solution, and $E_h = E_h / E$. The symbol $\epsilon$ stands for a real number of modulus less than 0.002, which may vary from one instance to the next.
<table>
<thead>
<tr>
<th>solution</th>
<th>( E )</th>
<th>( \mathcal{E}_1 )</th>
<th>( \mathcal{E}_2 )</th>
<th>( \mathcal{E}_{11} )</th>
<th>( \mathcal{E}_{14} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_A )</td>
<td>5.71…</td>
<td>0.248…</td>
<td>0.109…</td>
<td>0.195…</td>
<td>0.442…</td>
</tr>
<tr>
<td>( f_B )</td>
<td>5.67…</td>
<td>0.185…</td>
<td>0.123…</td>
<td>0.215…</td>
<td>0.470…</td>
</tr>
<tr>
<td>( f_C )</td>
<td>5.48…</td>
<td>( \epsilon )</td>
<td>0.243…</td>
<td>( \epsilon )</td>
<td>0.755…</td>
</tr>
<tr>
<td>( f_D )</td>
<td>5.38…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>0.375…</td>
<td>0.622…</td>
</tr>
<tr>
<td>( f_E )</td>
<td>5.21…</td>
<td>0.537…</td>
<td>( \epsilon )</td>
<td>0.458…</td>
<td>( \epsilon )</td>
</tr>
<tr>
<td>( f_F )</td>
<td>5.16…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>0.999…</td>
</tr>
<tr>
<td>( f_G )</td>
<td>5.02…</td>
<td>0.814…</td>
<td>0.137…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
</tr>
<tr>
<td>( f_H )</td>
<td>4.97…</td>
<td>0.974…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
</tr>
<tr>
<td>( f_I )</td>
<td>4.95…</td>
<td>0.883…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>0.075…</td>
</tr>
<tr>
<td>( f_J )</td>
<td>4.81…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
<td>0.999…</td>
<td>( \epsilon )</td>
</tr>
<tr>
<td>( f_K )</td>
<td>3.65…</td>
<td>( \epsilon )</td>
<td>0.996…</td>
<td>( \epsilon )</td>
<td>( \epsilon )</td>
</tr>
</tbody>
</table>
Proof of the theorem

We rewrite equation (main) in the form \( F(q) = q \), where

\[
F(q) = \omega^{-2} \partial^{-2} \nabla^* \left( \nabla q + (\nabla q)^3 \right),
\]

where \( \partial^{-1} \) denotes the antiderivative operator on the space of continuous

\( 2\pi \)-periodic functions with average zero and we look for fixed points of \( F \) in the

space of functions \( q : \mathbb{R} \to \mathbb{R}^P \) which extend analytically to a strip

\[ D_\rho = \{ t \in \mathbb{C} : |\text{Im}(t)| < \rho \} \].
To be more precise, given $\rho > 0$, denote by $\mathcal{F}_\rho$ the vector space of all $2\pi$-periodic analytic functions $f : D_\rho \rightarrow \mathbb{C}$,

$$f(t) = \sum_{k=1}^{\infty} f_k \sin(kt) + \sum_{k=0}^{\infty} f'_k \cos(kt), \quad t \in D_\rho,$$

which take real values for real arguments and for which the norm

$$\|f\|_\rho = \sum_{k=1}^{\infty} e^{\rho k} |f_k| + \sum_{k=0}^{\infty} e^{\rho k} |f'_k|$$

is finite. When equipped with this norm, $\mathcal{F}_\rho$ is a Banach space.

On the direct sum $\mathcal{F}_\rho^P$, we define the norm

$$\|q\|_\rho = \max_{1 \leq i \leq P} \|q_i\|.$$
We note that $\mathcal{F}_\rho$ is a Banach algebra, that is, $\|fg\|_\rho \leq \|f\|_\rho \|g\|_\rho$, for all $f$ and $g$ in $\mathcal{F}_\rho$. Furthermore, $\partial^{-2}$ acts as a compact linear operator on $\mathcal{A}_\rho$, as well as on $\mathcal{A}_\rho^P$. This shows that the equation above defines a differentiable map $F$ on $\mathcal{A}_\rho^P$ with compact derivatives $DF(q)$. Thus, $F$ can be well approximated locally by its restriction to a suitable finite dimensional subspace of $\mathcal{A}_\rho$. This property makes it ideal for a computer-assisted analysis.
Our goal is to find fixed points for $F$ by using a Newton like iteration, starting with a numerical approximation $q_0$ for the desired fixed point. The Newton map $\mathcal{N}$ associated with $F$ is given by

$$\mathcal{N}(q) = F(q) - \mathcal{M}(q)[F(q) - q],$$

with

$$\mathcal{M}(q) = [DF(q) - I]^{-1} + I.$$

If the spectrum of $DF(q)$ is bounded away from 1, and $q_0$ is sufficiently close to a fixed point of $F$, then $\mathcal{N}$ is a contraction in some neighborhood of $q_0$. 
Due to the compactness of $DF(q)$, this contraction property is preserved if we replace $\mathcal{M}(q)$ by a fixed linear operator $M$ close to $\mathcal{M}(q_0)$. This leads us to consider the new map $C$, defined by

$$C(q) = F(q) - M[F(q) - q], \quad q \in A^P_\rho.$$  

$M$ will be chosen to be a “matrix”, in the sense that $M = P_\ell M P_\ell$ for some $\ell > 0$, where $P_\ell$ denotes the canonical projection in $A^P_\rho$ onto Fourier polynomials of degree $k \leq \ell$. We also verify that $M - I$ is invertible, so that $C$ and $F$ have the same set of fixed points. In order to prove that $C$ is a contraction on some ball $B(q_0, r)$ in $A^P_\rho$ of radius $r > 0$, centered at $q_0$, it suffices to verify the inequalities

$$\|C(q_0) - q_0\|_\rho < \varepsilon, \quad \|DC(q)\| < K, \quad \varepsilon + Kr < r,$$

for some real numbers $r, \varepsilon, K > 0$, and for arbitrary $q$ in the ball $B(q_0, r)$. These bounds imply that $C$, and thus $F$, has a unique fixed point in $B(q_0, r)$. 
Theorem 2. In each of the 11 cases described in Theorem 1, there exists a Fourier polynomial \( q_0 \), and real numbers \( \rho, \varepsilon, r, K > 0 \), such that the inequalities above hold. Furthermore, the numerical bounds given in these theorems are satisfied for all function in the corresponding ball \( B(q_0, r) \).

The proof of this theorem is based on a discretization of the problem, carried out and controlled with the aid of a computer.

At the trivial level of real numbers, the discretization is implemented by using interval arithmetics. In particular, a number \( s \in \mathbb{R} \) is “represented” by an interval \( S = [S^-, S^+] \) containing \( s \), whose endpoints belong to some finite set of real numbers that are representable on the computer. Such an interval will be called a “standard sets” for \( \mathbb{R} \).
The goal now is to combine these elementary bounds to obtain e.g. a bound $G_1$ on the norm function on $\mathcal{A}_\rho^P$, and a bound $G_2$ on the map $\mathcal{C}$. Then, in order to prove the first inequality, it suffices to verify that $G_1(G_2(S)) \subset U$, where $S$ is a set in $\text{std}(\mathcal{A}_\rho^P)$ containing $g_0$, and $U$ is an interval in $\text{std}(\mathbb{R})$ with $U^+ < \varepsilon$.

We define the standard sets for $\mathcal{A}_\rho$. Let $n \geq \ell$ be a fixed integer. Given $U = (U_1, \ldots, U_n)$ in $\text{std}(\mathbb{R}^n)$, and $V = (V_0, \ldots, V_{2n})$ in $\text{std}(\mathbb{R}^{2n+1})$, denote by $S(U, V)$ the set of all functions $f$ that can be represented as

$$f(t) = \sum_{k=1}^{n} u_k \sin(kt) + \sum_{m=0}^{2n} v_m(t), \quad v_m(t) = \sum_{k=m}^{\infty} v_{m,k} \sin(kt),$$

with $u_k \in U_k$, and $v_m \in \mathcal{A}_\rho$ with $\|v_m\|_\rho \in V_m$, for all $k$ and $m$. We now define $\text{std}(\mathcal{A}_\rho)$ to be the collection of all such sets $S(U, V)$, subject to the condition that $V_m^- = 0$ for all $m$. 
It is now straightforward to implement a bound on the norm function on $\mathcal{A}_\rho^P$, or the operator $\partial^{-2}$, or the sum of two functions in $\mathcal{A}_\rho^P$. In order to obtain a bound on the product of two functions in $\mathcal{A}_\rho$, we simply multiply the representations of the two factors term by term, and write the result again as an explicit Fourier polynomial of order $n$, plus a sum of “error terms” of orders greater than $m$, for $m = 0, 1, \ldots, 2n$.

The guiding principle here is to keep as much information as possible about the order of each term in the product, since the operator $\partial^{-2}$, which is applied last in the definition of $F$, contracts higher order terms more than lower order ones. This principle also motivated our choice of standard sets for $\mathcal{A}_\rho$. 
For a bound on the linear operator $M$, we can compute explicitly its restriction to standard sets whose components $S(U_i, V_i)$ have $V_{i,m} = [0, 0]$ whenever $m \leq \ell$. The remaining terms are estimated by using that $\|Mq\|_{\rho} \leq \|M\| \|q\|_{\rho}$. The operator norm $\|L\|$ of a continuous linear operator $L$ on $A^P_{p}$ is given by the following formula. Denote by $h^{j,m}$ the function $(i, k) \mapsto \delta_{ij} e^{-k\rho} \sin(kt)$. Then

$$\|L\| = \max_{1 \leq i \leq P} \sum_{j=1}^{P} \sup_{m \geq 1} \| (Lh^{j,m})_i \|_{\rho}.$$ 

In the case where $L$ is the “matrix” $M$, the right hand side of this equation is trivial to estimate. The bounds discussed so far can be combined to yield a bound on the map $C$, suitable for proving the first and the last inequality.
In order to prove the second inequality

\[ \| DC(q) \| < K, \]

we also need a bound on the map \( q \mapsto \| D(q) \| \). Its domain only needs to include balls \( B(\rho_0, r) \) with positive representable radii, and these balls are in fact standard sets of \( A^P_\rho \). The technique used for this estimate is similar, up to some technical details.