Hamiltonian Mechanics and Integrable Systems

Summer School, Lake Arrowhead *

September 12th - 17th 2004

Organizers:

Rowan Killip, University of California, Los Angeles

Wilhelm Schlag, California Institute of Technology, Pasadena

Christoph Thiele, University of California, Los Angeles

*supported by NSF grant DMS 9985572 and NSF grant DMS 0300081

Contents

1 Gaps and bands of one dimensional periodic Schrödinger operators **4**
Kaihua Cai, California Institute of Technology 4
1.1 Introduction 4
1.2 Preliminaries 5
1.3 Lengths and harmonic measures 6
1.4 The Marcenko-Ostrovshii Theorem 7
1.5 Proof of Theorem 1, 2 and 3 8

2 A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation **10**
Tonči Crmarić, California Institute of Technology 10

3 A Lagrangian Proof of the Invariant Curve Theorem for Twist Mappings **16**
Silvius Klein, University of California, Los Angeles 16

4 Instability of dynamical systems with several degrees of freedom **23**
Irina Nenciu, California Institute of Technology, Pasadena 23

5 Compatible brackets in Hamiltonian mechanics **26**
Irina Nenciu, California Institute of Technology, Pasadena 26

6 Construction of Quasi-Periodic Solutions for Hamiltonian Perturbations of Linear Equations **29**
Dmitry Pavlov, California Institute of Technology, Pasadena 29

7 Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations **33**
Stefanie Petermichl, Brown University, Providence 33

8 The Hofer-Zehnder capacity **39**
8.1 Symplectic geometry preliminaries 39
8.1.1 From Hamiltonian mechanics to symplectic geometry. . 39
8.1.2 The problem of symplectic embeddings and symplectic invariants 40

8.1.3 Periodic orbits of Hamiltonian vector fields 41
8.2 Capacities 42
8.2.1 Definition and general properties 42
8.2.2 The linear case 42
8.3 Rigidity Results and Symplectic homeomorphisms 44
8.4 The Hofer-Zehnder capacity 46
8.4.1 Definition and verification of axioms. 46
8.4.2 Capacity of strictly convex domains 54

9 A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the mKdV equation 57
Svetlana Roudenko, Duke University, Durham 57
9.1 Introduction 57
9.2 Outline of the proof steps 59
9.3 Splitting of the two crosses (Section 3) 60
9.4 Reduction to the real line (Section 4) 61

10 On Nekhoroshev's Estimate at an Elliptic Equilibrium 64
Eric Ryckman, University of California, Los Angeles 64
10.1 Overview 64
10.2 Normal Form 65
10.3 Local Stability Estimates 67
10.4 Global Stability Estimates 68

11 Approximation of Solutions of the Cubic Nonlinear Schrodinger Equations by Finite-Dimensional Equations and Nonsqueezing Properties 70
Monica Visan, University of California, Los Angeles 70

12 Integrable systems and reductions of the self-dual Yang-Mills equations 74
Brett Wick, Brown University, Providence 74

1 Gaps and bands of one dimensional periodic Schrödinger operators

after J. Garnett and E. Trubowitz
A summary written by Kaihua Cai

Abstract

We give a simple necessary and sufficient condition on the length of the gaps corresponding to one dimensional periodic Schrödinger operators.

1.1 Introduction

Let $q(x)$ be the periodic extension to the whole line of a real valued function in $L^2_R[0,1]$. The spectrum of $-(d^2/dx^2)+q(x)$ acting on $L^2(R)$ is the union of purely absolutely continuous bands $B_n(q), n \geq 1$. It is well known that the bands may touch but never overlap. We assume that the bottom of the first band is at 0. The complement of the spectral band in $(0,+\infty)$ is a sequence of open intervals, called the gaps. To each set of bands $B_n(q), n \geq 1$, we associate the sequence of nonnegative numbers

$$\gamma_1(q), \gamma_2(q), \cdots$$

where $\gamma_n(q)$ is the distance between the top of the n-th band and the bottom of the next. The main theorems of this paper is to describe the set of all possible configurations of bands by understanding the distribution of gaps.

Theorem 1 *Let $\gamma_n, n \geq 1$, be any nonnegative sequence satisfying $\Sigma_{n\geq 1}\gamma_n^2 < \infty$. Then, there is a way of placing the sequence of open tiles of lengths γ_n, in order on the positive axis $(0,\infty)$ so that the compliment is the set of bands for a function q in $L^2_R[0,1]$. In other words, the map*

$$q \to \gamma(q) = (\gamma_n(q), n \geq 1),$$

from $L^2_R[0,1]$ to $(l^2)^+$, is onto.

It is natural to ask how many different ways a sequence of open tiles, whose lengths are $\gamma_n, n \geq 1$, can be placed so that the complement is a set of bands. This is answered by the following theorem:

Theorem 2 *There is just one way to place a sequence of open tiles, satisfying the hypothesis of Theorem 1, on the positive real axis so that they are genuine gaps.*

To prove these two theorems, we use a characterization of bands due to Marcenko and Ostrovskii. They identify band configurations with slit quarter planes.

Let $\mu_n(q), n \geq 1$, and $\nu_n(q), n \geq 0$, be the Dirichlet and Neumann spectrum of q in $L^2_R[0,1]$, that is, the spectra of

$$-y'' + q(x)y = \lambda y \tag{1}$$

for the boundary condition $y(0) = 0, y(1) = 0$ and $y'(0) = 0, y'(1) = 0$ respectively. If q is an even function, then $\gamma_n(q) = |\mu_n(q) - \nu_n(q)|$. We define the signed gap lengths of q in E_0, the subspace of even functions in $L^2_R[0,1]$ with mean 0, to be the sequence $(\mu_n(q) - \nu_n(q)), n \geq 1$. We have the following:

Theorem 3 *The map from q to its signed gap lengths is a real analytic isomorphism between E_0 and l^2.*

In fact, we obtain three real analytic isomorphisms between the three spaces E_0, l^2 and l^2_1, the space of real sequences $\{h_n\}$, satisfying $\Sigma n^2 h_n^2 < \infty$.

1.2 Preliminaries

Let $y_1(x,\lambda,q)$ and $y_2(x,\lambda,q)$ be the solutions of 1 satisfying

$$y_1(0,\lambda) = y_2'(0,\lambda) = 1, \qquad y_1'(0,\lambda) = y_2(0,\lambda) = 0$$

and set $\Delta(\lambda) = \Delta(\lambda,q) = y_1(1,\lambda) + y_2'(1,\lambda)$. The sequence of roots

$$\lambda_0 < \lambda_1 \leq \lambda_2 < \lambda_3 \leq \lambda_4 < \cdots$$

of $\Delta^2(\lambda) - 4 = 0$ is the spectrum of equation 1 with periodic conditions of period 2. We have $\Delta(\lambda_{2n-1}) = \Delta(\lambda_{2n}) = 2(-1)^n$ and $\lambda_{2n-1}, \lambda_{2n} = n^2\pi^2 + \int_0^1 q(x)dx + l^2(n)$. The problem of describing band configurations is equivalent to the characterization of all periodic spectra. We assume $\lambda_0(q) = 0$.

Lemma 4 *Let*

$$\delta(\lambda) = \delta(\lambda, q) = \int_0^\lambda \frac{-\dot{\Delta}(\mu)}{\sqrt{4 - \Delta^2(\mu)}} d\mu,$$

then $\delta(\lambda)$ is a conformal mapping of the upper half plane to the slit quarter plane

$$\Omega(h) = \{Rez > 0, Imz > 0\} \setminus \bigcup_{n=1}^{\infty} T_n$$

where $T_n = \{n\pi + iy : 0 < y \leq h_n\}$ and $\Sigma_{n \geq 1} n^2 h_n^2 < \infty$. Moreover,

$$\Delta(\lambda) = 2 \cos \delta(\lambda).$$

1.3 Lengths and harmonic measures

Write $(l^2)^+$ for the space of square-summable nonnegative sequences and $(l_1^2)^+$ for nonnegative sequences h_n, such that $\Sigma n^2 h_n^2 < \infty$. Say $h \in (l_1^2)^+$ is finite, if $h_n = 0$ for n sufficiently large. Let $h \in (l_1^2)^+$ and φ_h be the unique conformal map from the upper half plane to the slit quarter plane $\Omega(h)$, such that $\varphi_h(0) = 0, \varphi_h(\infty) = \infty$ and $\varphi_h^{-1}(z) = z^2 + b + O(1/|z|^2)$, for $|z|$ large. Set $\lambda_0 = 0$; $\lambda_{2n-1} = \varphi_h^{-1}(n\pi -)$ and $\lambda_{2n} = \varphi_h^{-1}(n\pi +)$. We also denote

$$\alpha_n = \lambda_{2n-1} - \lambda_{2n-2}, \qquad \gamma_n = \lambda_{2n} - \lambda_{2n-1}.$$

Lemma 5 *Assume h is finite. Let $u(z)$ be a bounded harmonic function on $\Omega(h)$ such that $u(z) = 0$ if $z \in \partial\Omega(h)$, $|z|$ large and let $U(\lambda) = u(\varphi_h(\lambda))$. Then for Lebesgue almost all $t \in R$, the limit $U(t) = \lim_{\eta \to 0} U(t + i\eta)$ exists and is integrable, and*

$$\int_{-\infty}^{\infty} U(t) dt = \lim_{x \to \infty} 2\pi x^2 u(x + ix). \tag{2}$$

In particular, the limit in (2) is finite and it is strictly positive if $u(z)$ is nonnegative but not identically zero.

Notice that if $u(z)$ is the harmonic measure of a bounded Borel set $E \subset \partial\Omega(h)$, then $U(t)$ agrees almost everywhere with the characteristic function of $\varphi^{-1}(E)$ and the limit in (2) evaluates the length of $\varphi^{-1}(E)$. The following two theorems are simple consequences of the maximum principle:

Theorem 6 *For all $h \in (l_1^2)^+$, and all $n \geq 1$, $\alpha_n(h) \leq (2n-1)\pi^2$. Equality holds for a single n if and only if $h = 0$.*

Theorem 7 *If $h \in (l_1^2)^+$, then $\gamma_n(h) \leq 4\max(2\pi n h_n, h_n^2)$. Furthermore,*

$$\|\gamma_n(h)\|_2^2 \leq const.(\sigma_n n^2 h_n^2 + (\sigma_n n^2 h_n^2)^2).$$

Note that if $\gamma_n^*(h) = \sup_k \gamma_n(h^{(k)})$, where $h^{(k)}$ is the truncation of h, then by the above theorem, $\gamma_n^*(h) \leq 4\max(2\pi n h_n, h_n^2)$, which implies that $\Sigma(\gamma_n^*)^2 < \infty$.

For the Marcenko-Ostrovskii characterization of spectra we need two further estimates.

Theorem 8 *Let $h \in (l_1^2)^+$. Then there is a constant $c = c(h)$ such that*

$$\gamma_{2n-1}(h) = n^2\pi^2 + c + l^2(h), \qquad \gamma_{2n}(h) = n^2\pi^2 + c + l^2(h)$$

and $\lim_{k\to\infty} \Sigma_n(\gamma_{2n}(h^{(k)}) - \gamma_{2n}(h))^2 = 0$ where $h^{(k)}$ is the truncation of h.

1.4 The Marcenko-Ostrovshii Theorem

Let E denote the subspace of even functions in $L_R^2[0,1]$.

Theorem 9 *Let $h \in (l_1^2)^+$ and let φ_h be the conformal mapping from the upper half plane to the slit quarter plane $\Omega(h)$. Then there exists $q \in E$ such that $\varphi_h(\lambda) = \delta(\lambda, q)$.*

Except for the fact that the potential is even, i.e. $q(x) = q(1-x)$, this theorem was first proved by [1]. We will give a different proof, using the estimate of the previous section and some ideas from [2]. First, we consider the roots $\mu_1(q) < \mu_2(q) < \cdots$ of $y_2(1,\lambda,q) = 0$. The sequences μ_n is called the Dirichlet spectrum of q. It is well known that $\lambda_{2n-1} \leq \mu_n \leq \lambda_{2n}$. Let S be the Hilbert manifold of all increasing sequences $\sigma_n = n^2\pi^2 + l^2(n)$, $n \geq 1$, and $E_0 \subset E$ be the subspace of even function with mean 0.

Theorem 10 *The map from* E_0 *to* S *defined by*

$$q \in E_0 \to (\mu_1(q), \mu_2(q), \cdots) \in S$$

is one-to-one, onto and bianalytic.

Refer [2] for the proof. We next make a list of all possible functions $\Delta(\lambda, q)$.

Lemma 11 *Let* $\sigma \in S$*, then the series*

$$\Delta_\sigma(\lambda) = 2\cos\sqrt{\lambda} + \Sigma_{n\geq 1} 2[(-1)^n - \cos\sqrt{\sigma_n}] \prod_{m\neq n} \frac{\sigma_m - \lambda}{\sigma_m - \sigma_n}$$

converges, uniformly on bounded subsets of the complex plane $\mathcal{C}$*, to an entire function. Moreover, there is an even function* $q \in E_0$ *such that* $\Delta(\lambda, q) = \Delta_\sigma(\lambda)$ *and* $\sigma_n = \mu_n(q)$*,* $n \geq 1$*. Conversely, if* $q \in E_0$*, then* $\Delta(\lambda, q) = \Delta_\mu(\lambda)$*, where* $\mu = (\mu_1, \mu_2, \cdots)$ *is the Dirichlet spectrum of* q*.*

1.5 Proof of Theorem 1, 2 and 3

Write $\mathbb{R}_+^N = \{x \in \mathbb{R}^N : x_n > 0, 1 \leq n \leq N\}$ and regard $\mathbb{R}_+^N$ both as the subspace $\{h_n = 0, n > N; h_n > 0, n \leq N\}$ of $(l_1^2)^+$ and as the subspace $\{\gamma_n = 0, n > N; \gamma_n > 0, n \leq N\}$ of $(l^2)^+$. Then we have defined a mapping $h \to \gamma_n(h)$ from $\mathbb{R}_+^N$ into $\mathbb{R}_+^N$ because $\gamma_n = 0$ if and only if $h_n = 0$.

Theorem 1 and 2 are consequences of the following lemmas:

Lemma 12 *From* $\mathbb{R}_+^N$ *to* $\mathbb{R}_+^N$ *the map* $h \to \gamma_n(h)$ *is real analytic. It satisfies:*

$$\frac{\partial\gamma_n}{\partial h_n} > 0 \qquad \frac{\partial\gamma_k}{\partial h_n} < 0 \qquad k \neq n$$
$$\frac{\partial\gamma_n}{\partial h_n} + \Sigma_{k;k\neq n}\frac{\partial\gamma_k}{\partial h_n} > Const.n \exp^{-(M+h_n)/2},$$

where $M = \max\{h_n : n = 1, 2, \cdots\}$*.*

Lemma 13 *If* $h \in (l_1^2)^+$ *and* N *is large, then*

$$\{\Sigma_{n=1}^N(\gamma_n(h) - \gamma_n(h^{(N)}))^2\}^{1/2} \leq \frac{C}{\sqrt{N}}\{\Sigma_{k=N}^\infty k^2 h_k^2\}^{1/2}$$

The main use of Lemma 12 is the observation that the Jacobian of $\gamma_n(h)$ is never zero.

References

[1] Marcenko, V. A. and Ostrovshii, I. V., *A Characterization of the spectrum of Hill's operator.* Math. USSR-Sbornik, 97 (1975), pp. 493-554.

[2] Poschel, J. and Trubowitz, E., *Lectures on Inverse Spectral Theory.*

[3] Mckean, H.P. and Trubowitz, E., Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points. Comm. Pure Appl. Math. 29 (1976) pp. 143-226

KAIHUA CAI: DIVISION OF ASTRONOMY, MATHEMATICS, AND PHYSICS, 253-37 CALTECH, PASADENA, C.A. 91125, U.S.A.
email: **kaihua@its.caltech.edu**

2 A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation

after P Deift and X. Zhou
A summary written by Tonči Crmarić

Abstract

We give a summary of the results in the first part of [2].

Introduction

In [2], a new and general method to analyzing the asymptotics of oscillatory Riemann-Hilbert problems is presented. The method is applied to finding the long-time asymptotics of the modified Korteweg-de Vries (MKdV) equation,

$$y_t - 6y^2 y_x + y_{xxx} = 0, \quad -\infty < x < \infty,\ t \geq 0, \tag{3}$$
$$y(x, t = 0) = y_0(x). \tag{4}$$

The associated Riemann-Hilbert problem (see [1] for a description of the inverse scattering method for MKdV) is

$$m_+(z) = m_-(z) v_{x,t}(z), \quad z \in \mathbf{R}, \tag{5}$$
$$m(z) \to I \quad \text{as } z \to \infty, \tag{6}$$

where $m(z; x, t)$ maps to 2×2 matrices and is analytic in z for $\mathrm{Im} z \neq 0$,

$$m_\pm(z) = \lim_{\epsilon \downarrow 0} m(z \pm i\epsilon; x, t)$$

and

$$v_{x,t}(z) = e^{-i(4tz^3 + xz)\sigma_3} v(z) e^{i(4tz^3 + xz)\sigma_3},$$

where

$$v(z) = \begin{pmatrix} 1 - |r(z)|^2 & -\overline{r(z)} \\ r(z) & 1 \end{pmatrix}.$$

Here $r(z)$ lies in Schwartz space and satisfies

$$r(z) = -\overline{r(-z)}, \quad \sup_{z\in\mathbf{R}}|r(z)| < 1.$$

The matrix $v(z)$ has unique upper/lower factorization with 1's on the diagonal,

$$v(z) = b_-^{-1} b_+ = \begin{pmatrix} 1 & -\overline{r} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix}. \tag{7}$$

The solution of the inverse problem is given by

$$y(x,t) = \left(\left[\sigma_3, \int_{\mathbf{R}} \mu(z;x,t) w_{x,t}(z) \frac{\mathrm{d}z}{2\pi i}\right]\right)_{21}, \tag{8}$$

where

$$w_{x,t} = (w_+)_{x,t} + (w_-)_{x,t} \tag{9}$$

and

$$w_\pm = \pm(b_\pm - I) \tag{10}$$

and $\mu(z;x,t)$ is the solution of a Fredholm integral equation for an explicit singular integral operator on the contour associated with problem (5), i.e. on $\mathbf{R}$. Although μ can be presented via Neumann series, it is not possible to obtain a sensible estimate on the long-term asymptotics of $y(x,t)$ by estimating the sizes of the addends, because of the large amount of cancellation. See [1].

The long-time behavior of the MKdV equationand the other nonlinear wave equations solvable by the inverse scattering method is a well-studied problem. The method presented here has an advantage that it is general and algorithmic and does not require a hypothesis for the form of the solution of the asymptotic problem. For the statement of the results we refer the reader to the abstract of the second part of the article.

Altering the contour

First, we present an example which contains the basic ideas used in the part of the [2] discussed here. Let $y(x,t)$ be a solution obtained from (8). We will show that, for fixed t, $y(x,t)$ decays rapidly as $x \to \pm\infty$. For the simplicity of the notation, assume $t=0$. We write

$$e^{2izx} r(z) = \int_x^\infty \hat{r} e^{iz(2x-s)} \mathrm{d}s + \int_{-\infty}^x \hat{r}(s) e^{iz(2x-s)} \mathrm{d}s. \tag{11}$$

Note that the first term, to be denoted by $h_1(x,z)$, decays rapidly as $x \to +\infty$ for $z \in \mathbf{R}$ and that the second term, $h_2(x,z)$, has an analytic continuation to $\{z \in \mathbf{C} : \mathrm{Re}(izx) < 0\} = \{z \in \mathbf{C} : \mathrm{Im} z > 0\}$, which is bounded and converges to 0 as $z \to +\infty$ and decays exponentially as $x \to +\infty$. If we set

$$v^+_{x,t=0} = \begin{pmatrix} 1 & -\overline{h}_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ h_1 & 1 \end{pmatrix},$$

$$m^+ = \begin{cases} m \begin{pmatrix} 1 & 0 \\ -h_2(x,z) & 1 \end{pmatrix} & \mathrm{Im} z > 0 \\ m \begin{pmatrix} 1 & -\overline{h_2(x,\overline{z})} \\ 0 & 1 \end{pmatrix} & \mathrm{Im} z < 0 \end{cases},$$

then (5) is transformed into the Riemann-Hilbert problem

$$m^+_+(z) = m^+_-(z) v^+_{x,t=0}(z), \quad z \in \mathbf{R}, \tag{12}$$

$$m^+(z) \to I \quad \text{as } z \to +\infty. \tag{13}$$

The rapid decay of $y(x,0)$ is now immediate from the rapid decay of $v^+_{x,t=0} - I$ and (8), (9), (10).

To prove the rapid decay as $x \to -\infty$, we must replace the upper/lower factorization by a lower/upper factorization. We find that

$$v(z) = \begin{pmatrix} 1 & 0 \\ r(1-|r|^2)^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1-|r|^2 & 0 \\ 0 & (1-|r|^2)^{-1} \end{pmatrix} \begin{pmatrix} 1 & -\overline{r}(1-|r|^2)^{-1} \\ 0 & 1 \end{pmatrix}.$$

One can remove the diagonal term in the following way. Consider the scalar Riemann-Hilbert problem

$$\delta_+(z) = \delta_-(z)(1-|r(z)|^2), \tag{14}$$

$$\delta(z) \to 1 \quad \text{as } z \to \infty \tag{15}$$

for $\delta(z)$ analytic in $\mathbf{C} \setminus \mathbf{R}$. After conjugating v by δ^{σ_3}, problem (14) is transformed into

$$m^-_+(z) = m^-_-(z) v^-_{x,t=0}(z), \quad z \in \mathbf{R}, \tag{16}$$

$$m^-(z) \to I \quad \text{as } z \to \infty, \tag{17}$$

where

$$v^-_{x,t=0} = e^{-izx\sigma_3} \begin{pmatrix} 1 & 0 \\ \delta_-^{-2} r(1-|r|^2)^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & -\bar{r}\delta_+^2(1-|r|^2)^{-1} \\ 0 & 1 \end{pmatrix} e^{izx\sigma_3}$$

and

$$m^-(z) = m(z)\delta^{-\sigma_3}(z).$$

The decay of $y(x,0)$ as $x \to -\infty$ can now be inferred from the problem (16) in the same way as in the case $x \to \infty$.

In the problem of long-time asymptotics, t is a variable and it is no longer possible to decompose

$$e^{2i(4z^3t+xz)} r(z) = h_1 + h_2$$

as in formula (11) into parts h_1 and h_2, which are rapidly decaying as $t \to \infty$ and are uniformly bounded in $\text{Im}z > 0$, respectively. The crucial observation is that the complex plane must be decomposed according to the sign of $\text{Re}(iF), F = 4tz^3 + xz$.

We will now outline the first steps of the procedure for finding long-time asymptotics of MKdV, in the case $x > 0$. The remaining steps of the procedure are presented in the abstract of the second part of [2]. We will transform (5) to an equivalent Riemann-Hilbert problem on a contour which will reflect the sign of $\text{Re}(iF)$. Let $z_0 = \sqrt{-x/12t}$. Note that$\text{Re}(iF)$ changes sign at $\pm z_0$ and we have $\text{Re}(iF(z_0+\epsilon(1+i)) < 0,\ \text{Re}(iF(z_0+\epsilon(-1+i)) > 0$ etc. The new contour, Σ, is defined as

$$\Sigma = L \cup \overline{L} \cup \mathbf{R},$$

where

$$\begin{aligned} L = & \{z = z_0 + z_0 u e^{3i\pi/4} : -\infty < u \le \sqrt{2}\} \\ & \cup \{z = -z_0 + z_0 u e^{i\pi/4} : -\infty < u \le \sqrt{2}\}. \end{aligned}$$

Set $L_\epsilon = \{z = z_0 + z_0 u e^{3i\pi/4} : \epsilon < u \le \sqrt{2}\} \cup \{z = -z_0 + z_0 u e^{i\pi/4} : \epsilon < u \le \sqrt{2}\}$ and $\Sigma' = \Sigma \setminus (L_\epsilon \cup \overline{L_\epsilon})$. We orient Σ in an appropriate way.

As in the example above, we consider scalar Riemann-Hilbert problem

$$\delta_+(z) = \delta_-(z)(1-|r(z)|^2), \quad |z| < z_0, \tag{18}$$

$$= \delta_-(z), \quad |z| > z_0, \tag{19}$$

$$\delta(z) \to 1 \quad \text{as } z \to \infty \tag{20}$$

which has solution

$$\delta(z) = \left(\frac{z - z_0}{z + z_0}\right)^{i\nu} g^{\chi(z)},$$

where

$$\nu = -\frac{1}{2\pi} \log 1 - |r(z_0)|^2$$

and

$$\chi(z) = \frac{1}{2\pi i} \int_{-z_0}^{z_0} \log\left(\frac{1 - |r(\zeta)|^2}{1 - |r(z_0)|^2}\right) \frac{\mathrm{d}\zeta}{\zeta - z}.$$

The solution δ has the property $\delta(z) = \overline{\delta(-\overline{z})} = \left(\overline{\delta(\overline{z})}\right)^{-1}$ and, consequently, $|\delta(z)| < c < \infty$ and $|\delta(z)|^{-1} < c < \infty$ for every z.

Conjugating the Riemann-Hilbert problem (5) by δ^{σ_3} leads to the Riemann-Hilbert problem for $m\delta^{-\sigma_3}$,

$$\left(m\delta^{-\sigma_3}\right)_+(z) = \left(m\delta^{-\sigma_3}\right)_-(z)\delta_-^{\sigma_3} v_{x,t} \delta_+^{-\sigma_3}(z), \quad z \in \mathbf{R}, \tag{21}$$

$$m\delta^{-\sigma_3} \to I \quad \text{as } z \to \infty, \tag{22}$$

where $\delta_-^{\sigma_3} v_{x,t} \delta_+^{-\sigma_3}$ has factorization

$$e^{-iF\sigma_3} \begin{pmatrix} 1 & 0 \\ r\delta_-^{-2}(1 - |r|^2)^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & -\overline{r}\delta_+^2(1 - |r|^2)^{-1} \\ 0 & 1 \end{pmatrix} e^{iF\sigma_3},$$

for $|z| < z_0$, and

$$e^{-iF\sigma_3} \begin{pmatrix} 1 & -\overline{r}\delta^2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r\delta^{-2} & 1 \end{pmatrix} e^{iF\sigma_3},$$

for $|z| > z_0$.

Let $\Phi(z) = 4(z^3 - 3z_0^2 z)$. Then $F = t\Phi(z)$, in the analogy with xz from the above example. Let $|z| < z_0$, k any positive integer and let

$$\rho(z) = -\overline{r(z)}(1 - |r|^2)^{-1} = h(z) + R(z),$$

where R is a polynomial of order $2k + 1$ such that h vanishes to order k at z_0 and$-z_0$. By Plancharel, it follows that h can be split, using the Fourier transform (in variable Φ) as in the example above, into a part, h_1, which decays to a high order as $t \to \infty$ and a part, h_2, which has a continuation into upper halfplane. Splitting for $r(z)(1 - |r|^2)^{-1}$ is then given by $-\overline{h_1(\overline{z})}$ –

$\overline{(h_2(\overline{z}) + \overline{R(\overline{z})})}$ and the splitings for the entries of $\delta_-^{\sigma_3} v_{x,t} \delta_+^{-\sigma_3}$ are easy to obtain.

We are led to a Riemann-Hilbert problem

$$m_+(z) = m_- - (z) v_{x,t}(z), \quad z \in \Sigma, \tag{23}$$
$$m\ (z) \to I \quad \text{as } z \to \infty. \tag{24}$$

On $\mathbf{R}$, the coefficients of $v_{x,t}(z)$ depend only on h_1 and on $\Sigma \setminus \mathbf{R}$ they depend also on h_2 and on R.

As $t \to \infty$, we have $h_1 \to 0$ on Σ. On $\Sigma \setminus \mathbf{R}$, the terms of the type $h_2 e^{-2it\Phi(z)}$ converge to 0. On $L_\epsilon \cup \overline{L_\epsilon}$, the terms of the type $Re^{-2it\Phi(z)}$ converge to 0. It follows that it suffices to analyze the Riemann-Hilbert problem on Σ', the reduced contour, i.e.

$$m'_+(z) = m'_-(z) v_{x,t}(z), \quad z \in \Sigma', \tag{25}$$
$$m'(z) \to I \quad \text{as } z \to \infty. \tag{26}$$

References

[1] Beals R. and Coifman R. *Scattering and inverse scattering for first order systems* Comm. Pure Appl. Math **37** [1984], pp. 39-90.

[2] Deift P. and Zhou X. *A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation* Annals of Math. **137** [1993], pp. 295-368.

3 A Lagrangian Proof of the Invariant Curve Theorem for Twist Mappings

after M. Levi and J. Moser
A summary written by Silvius Klein

Abstract

The paper summarized here gives a new proof (of a version) of Moser's classical invariant curve theorem. The main idea of the proof is to reformulate the condition for a curve to be invariant as a second order difference equation for the $x-$coordinate of the curve. This difference equation is solved using a Newton's iteration scheme. In the appendix the authors give a proof of the small-twist theorem and apply it to prove a stability theorem for elliptic fixed points.

Background on twist maps

Let $C := S^1 \times \mathbb{R}$ be the (unbounded) cylinder (where $S^1 = \mathbb{R}/\mathbb{Z}$). The universal cover of the cylinder C is the plane $\mathbb{R}\times\mathbb{R}$. A lift of a map $\Phi : C \to C$ is a map $\phi : \mathbb{R}\times\mathbb{R} \to \mathbb{R}\times\mathbb{R}$ such that if $\pi : \mathbb{R}\times\mathbb{R} \to S^1\times\mathbb{R}$, $\pi(x, y) := ([x], y)$ is the projection to the quotient C, then $\pi \circ \phi = \Phi \circ \pi$. If continuous, such a lift is unique up to an additive integer constant.

Definition 14 *A surjective diffeomorphism $\Phi : C \to C$ of the cylinder is called an area-preserving twist map if Φ preserves area, Φ preserves orientation, Φ preserves boundary components and if*

$$\partial_y \phi_1(x, y) > 0$$

for any lift $\phi = (\phi_1, \phi_2)$ of Φ to the universal cover $\mathbb{R} \times \mathbb{R}$ of C.

Example 15 A twist map is called integrable if it is of the form

$$\Phi(x, y) = (x + g(y), y).$$

Integrable twist maps leave all circles $S^1 \times \{y\}$ invariant and rotate them by the function (assumed monotone) g.

The map Φ on the cylinder C induced by

$$\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}, (x, y) \mapsto (x + y, y + V(x + y)),$$

with V periodic, is called the standard map.It is immediate that the standard map is a twist map. For $V = 0$ we get a very simple integrable twist.

Definition 16 *Let Φ be a(n area preserving twist) function and ϕ a lift to its universal cover. A generating function for Φ is a map $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ so that:*

$$\begin{array}{c} h_1(x, \phi_1(x, y)) = -y, \\ h_2(x, \phi_1(x, y)) = \phi_2(x, y) \end{array} \tag{27}$$

We impose moreover that

$$h(x_1 + 1, x_2 + 1) = h(x_1, x_2)$$

and

$$h_{12} < 0,$$

where h_1, h_2 are the partial derivatives of h with respect the its first and second variables and h_{12} is its second derivative with respect to both variables.

It is not hard to show that any area-preserving twist map has a generating function.

We will define the notion of invariant curve for a (twist) map. Proving the existence of such a curve is the purpose of this paper.

Definition 17 *Consider a closed curve $w(\theta)$ wrapping around the cylinder. It's lift to the covering $(x, y)-$plane of the cylinder has parametric equations $x = u(\theta)$, $y = v(\theta)$ where $u(\theta) - \theta$ and $v(\theta)$ are periodic functions of period 1.*

An invariant curve for the (twist) map Φ is any such a curve $w(\theta) = (u(\theta), v(\theta))$ which satisfies the invariance condition:

$$\phi(w(\theta)) = w(\theta + \omega) \tag{28}$$

with a prescribed rotation number ω, and so that $u(\theta)$ is strictly monotone.

Reduction to a difference equation

The Hamiltonian problem of finding such a curve $w(\theta)$ can be reduced, via elementary manipulations of the formulas (27) to a Lagrangian equation for a single function $u(\theta)$. More precisely, we have the following:

Proposition 18 *The curve $w(\theta) = (u(\theta), v(\theta))$ satisfies the invariance condition (28) with ϕ given by (27) iff the horizontal function $u(\theta)$ satisfies the second order difference equation*

$$E(u(\theta)) := h_1(u(\theta), u)\theta + \omega)) + h_2(u(\theta - \omega), u(\theta)) = 0 \tag{29}$$

Example 19 For the integrable twist map $\phi(x, y) = (x + y, y)$, the generating function is $h(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$, while the difference equation (29) takes the form

$$E(u(\theta)) := u(\theta + \omega) - 2u(\theta) + u(\theta - \omega) = 0$$

It will be convenient for a function $u(\theta)$ and a rotation number ω to use the following notations: $u^+(\theta) := u(\theta + \omega)$, $u^-(\theta) := u(\theta - \omega)$.

Main Theorem

The invariant curve theorem in this paper says that in the vicinity of a near-solution to (29): $E(u_0) \approx 0$ there exists an exact solution: $E(u) = 0$ provided that ω is Diophantine.

Let W_r be the space of 1-periodic real analytic functions of θ, bounded on the strip $|\Im\Theta| \le r$ and equipped with the maximum norm denoted $|f|_r$.

We will make the following assumptions on the data (h, the generating function of the twist map and u_0, the near-solution to (29)).

The function h is analytic for $(x_1, x_2) \in D \subset \mathbb{C}^2$ and it is real for $(x_1, x_2) \in \mathbb{R}^2$. Denote by D_R the largest set whose $R-$neighborhood lies in D. Moreover, assume that

$$\min_D |h_{12}| > \kappa > 0, \tag{30}$$

and that

$$|h|_{C^3(D)} < M. \tag{31}$$

Let $u_0(\theta) - \theta \in W_r$ for some $0 < r < 1$. Moreover, assume that

$$(u_0, u_0^+) \in D_R \text{ for } |\Im\theta| \le r \tag{32}$$

and that

$$|\partial_\theta u_0|, |(\partial_\theta u_0)^{-1}| < N_0 \tag{33}$$

Finally assume ω satisfies a Diophantine condition (DC). Then we have:

Theorem 20 *There exists $\delta = \delta(r, h, M, N_0, \kappa, DC) > 0$ such that if $|E(u_0)|_r < \delta$, then there exists a unique solution u near u_0 of $E(u) = 0$ with $u(\theta) - \theta \in W_{r/2}$ and mean value of $u(\theta) - \theta$ equal to zero.*

This theorem can be applied in particular to small perturbations of the standard map. Indeed, we have:

Corollary 21 *Consider a small perturbation of the standard map*

$$\phi(x, y) = (x + y + O(\epsilon), y + O(\epsilon))$$

It's generating function has the form

$$h(x_1, x_2) = \frac{1}{2}(x_2 - x_1)^2 + O(\epsilon)$$

which we assume to be defined on

$$D = \{(x_1, x_2) : a < \Re(x_2 - x_1) < b, |\Im x_1|, |\Im x_2| < 1\}.$$

Choose $u_0(\theta) = \theta$, and ω Diophantine. Then it is easy to see that $|E(u_0)|_r = O(\epsilon)$.

Then Theorem 20 yields an invariant curve in the annulus $a < y < b$ provided ϵ is small enough.

Idea of the proof

The equation $E(u) = 0$ is solved using a modification of Newton's method. Starting with u_0 such that $E(u_0)$ is sufficiently small, we seek an improvement $u_1 = u_0 + v$ (so that u_1 is even smaller). Continuing inductively, we obtain a sequence u_n which is shown to converge to an exact solution u.

Here is how we seek the improvement: having u, find v so that

$$E(u+v) = E(u) + E'(u)v + ...$$

becomes quadratically small compared with $E(u)$.

The first idea would be of course to eliminate the linear term completely, therefore to solve for v the linear equation

$$E(u) + E'(u)v = 0. \tag{34}$$

This equation is difficult to solve, because $E'(u)$ is not easy to invert. The idea is to modify (34) so that v could be easily determined (the new equation will not be equivalent to (34) though) and so that $E(u+v)$ would be quadratically small. Introducing the finite-difference operators $\nabla^+ u(\theta) = u(\theta+\omega) - u(\theta)$ and $\nabla^- u(\theta) = u(\theta) - u(\theta - \omega)$ we transform (34) into the following homological equation:

$$\nabla^-(h_{12}^+ \, \partial_\theta u \, \partial_\theta u^+ \, \nabla^+ w) + \partial_\theta u \, E(u) = 0 \tag{35}$$

Given u, we will solve the homological equation (35) and find w. The improvement sought for u will be

$$\tilde{u} = u + v, v = \partial_\theta w$$

The following lemma says that for appropriate u the solution w to (35) exists:

Lemma 22 *Assume that $u(\theta)$ satisfies*

$$(u_0, u_0^+) \in D_R \ \ \textit{for} \ \ |\Im\theta| \le r$$

$$|\partial_\theta u_0|, |(\partial_\theta u_0)^{-1}| < N \ \ \textit{for} \ \ |\Im\theta| \le r$$

Then(35) has a unique solution $w \in W_\rho$, with mean zero, for any $0 < \rho < r$ and the correction $v = \partial_\theta w$ satisfies the estimates

$$|v|_\rho \lesssim \frac{1}{(r-\rho)^\sigma} |E(u)|_r, \ \ |\partial_\theta v|_\rho \lesssim \frac{1}{(r-\rho)^{\sigma+1}} |E(u)|_r,$$

where $\sigma = \sigma(DC)$.

The following lemma shows that the updated $\tilde{u} = u + v, (v = \partial_\theta w)$ gives a quadratic improvement of the error.

Lemma 23 *Let u satisfy the assumptions of Lemma 22 and let $\tilde{u} = u + v$ satisfy $(u, u^+) \in D_R$ for $|\Im\theta| < \rho$, for some $0 < \rho < r$. Then*

$$|E(\tilde{u})|_\rho \lesssim \frac{1}{(r-\rho)^{2\sigma}} |E(u)|_r^2$$

Other results and applications

Using the method described above, the following "small twist" case is treated.

Assume that the area-preserving analytic map ϕ depends on a parameter $\gamma \in (0,1)$ and has the form

$$\phi(x,y) = (x + a(\gamma) + \gamma y + f(x,y,\gamma),\ y + g(x,y,\gamma))$$

where the functions f, g are analytic in the complex disc $|y| \le 1$ and in $|\Im x| \le r$. Denote by $W_{r,s}$ the space of real analytic functions bounded in $|y| \le s$, $|\Im x| \le r$, of period 1 in x and with the supremum norm $|f|_{r,s}$ on this domain. Then we have the following:

Theorem 24 *There exists $\delta > 0$, independent of γ such that if*

$$|f|_{r,1} + |g|_{r,1} < \gamma\delta$$

then there exists an invariant curve $x = u(\theta)$, $y = v(\theta)$ in $-1 < v(\theta) < 1$ with both $u(\theta) - \theta$ and $v(\theta)$ periodic of period 1, real analytic and with $u(\theta)$ monotone increasing.

This theorem differs from the previous one as the rotation ω is not prescribed but has to be constructed in an interval of length γ.

This result can be used to prove stability of elliptic fixed points for certain area-preserving analytic maps.

Let ϕ be such a map. In the neighborhood of a fixed point (which we assume to be the origin) ϕ can be written in complex coordinates as

$$w \to we^{i\alpha} + O(|w|^2)$$

If $\frac{q\alpha}{2\pi}$ is not an integer for $q = 1, 2, ..., k$, then by changing coordinates, ϕ can be written in Birkhoff normal form up to order $k-1$, i.e.

$$w \to we^{i\psi} + O(|w|^k) \tag{36}$$

where ψ is a real polynomial of degree $\le \frac{k}{2} - 1$ in $|w|^2$.

We obtain the following

Theorem 25 *If ψ in (36) is not a constant, then the origin is a stable fixed point under ϕ.*

Final remarks

The original invariant curve theorem was proved under different assumptions. The map there was only sufficiently smooth, satisfying a certain curve intersection property. The present statement is more restrictive, but still sufficient for many applications.

Theorem 20 in this paper was extended in [Z] to the case of quasi-periodic twist mappings. The approach is very similar. The coordinates of the invariant curve obtained are quasi-periodic.

References

[KH] Anatole Katok and Boris Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge University Press, Cambridge, 1995, With a supplementary chapter by Katok and Leonardo Mendoza. MR 96c:58055

[LM] M. Levi and J. Moser, *A Lagrangian proof of the invariant curve theorem for twist mappings*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 733–746. MR 2002h:37066

[M] J. Moser, *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II **1962** (1962), 1–20. MR 26 #5255

[Z] Vadim Zharnitsky, *Invariant curve theorem for quasiperiodic twist mappings and stability of motion in the Fermi-Ulam problem*, Nonlinearity **13** (2000), no. 4, 1123–1136. MR 2001h:37087

4 Instability of dynamical systems with several degrees of freedom

after V. I. Arnol'd
A summary written by Irina Nenciu

Abstract

We give a summary of the results in Arnold [1]. Following Wayne [3], one may informally state the KAM theorem as saying that most trajectories of nearly integrable Hamiltonian systems remain quasi-periodic. Arnol'd's avowed purpose in [1] was to give an example of a system containing unstable trajectories.

Introduction and main result

Recall that a Hamiltonian system is a system of $2d$ equations $\dot{p}_j = -\frac{\partial H}{\partial q_j}$ and $\dot{q}_j = \frac{\partial H}{\partial p_j}$ for $1 \leq j \leq d$. If the system is completely integrable, then there exists a special set of variables $(I, \phi) \in \mathbf{R}^d \times \mathbf{T}^d$, called action-angle variables, in which $H(I, \phi) = H_0(I)$. Therefore the equations of motion reduce to $\dot{I}_j = 0$, and hence the I's are constant, and $\dot{\phi}_j = \frac{\partial H}{\partial I_j} \equiv \omega_j(I)$. So the motion of the system is quite simple: $I_j(t) = I_j(0)$ are constants of motion, and $\phi_j(t) = t\omega_j + \phi_j(0)$ describes the evolution of the trajectory along a d-dimensional torus.

We are interested in the case where a perturbation is added to H_0:

$$H(I, \phi) = H_0(I) + \epsilon H_1(I, \phi).$$

Then, for small enough ϵ and for almost all nonresonant frequency vectors ω (i.e. so that $\langle \omega, n \rangle \neq 0$ for all $n \in \mathbf{Z}^d \backslash 0$), there exists an invariant torus $T(\omega)$ of the perturbed system which is close to the corresponding invariant torus $T_0(\omega)$ of the completely integrable system (see, for example, [2, Theorem 21.7]).

Let us introduce Arnol'd's example. We work on the 5-dimensional phase space $\Omega = \mathbf{R}^2 \times \mathbf{T}^3$, with $(I_1, I_2) \in \mathbf{R}^2$ and $(\phi_1, \phi_2, t) \in \mathbf{T}^3$. The Hamiltonian depends on 2 parameters, ϵ and μ, and is given by

$$H = \frac{1}{2}(I_1^2 + I_2^2) + \epsilon(\cos\phi_1 - 1)(1 + \mu B), \quad B = \sin\phi_2 + \cos t.$$

In other words, we consider the system of differential equations

$$\dot{\phi}_j = I_j\,,\ \dot{I}_1 = \epsilon \sin\phi_1(1+\mu B)\,,\ \dot{I}_2 = \epsilon(1-\cos\phi_1)\mu\cos\phi_2 \tag{37}$$

with $j = 1, 2$ and $0 < \mu \ll \epsilon \ll 1$.

Theorem 26 *Assume* $0 < A < B$. *For every* $\epsilon > 0$ *there exists a* $\mu_0 = \mu_0(A, B, \epsilon) > 0$ *such that for* $0 < \mu < \mu_0$ *the system (37) has a solution satisfying* $I_2(0) < A$, $I_2(t) > B$ *for a certain* t.

What makes Arnol'd's example work

The main mechanism behind the construction of such unstable trajectories is the *transition chain.* Let us clarify the notions. Assume that in the phase space of the dynamical system there is an invariant torus T and on it a quasi-periodic motion with dense orbits.

Definition 27 *We shall call* T *a whiskered torus if* T *is a component of the intersection of two invariant, open manifolds* Y^- *and* Y^+*, where all the trajectories of the arriving whisker* Y^- *approach* T *as* $t \to \infty$*, and on the departing whisker* Y^+ *all the trajectories approach* T *as* $t \to -\infty$.

If, for a whiskered torus T, neighborhoods of points of Y^- stay very close to Y^+ as they evolve (in a sense that will be made rigorous during the talk), then we call T a transition torus. To construct a transition chain one needs a sequence $T_1, \ldots, T_s$ of transition tori with $Y_j^+ \bigcap Y_{j+1}^- \neq \emptyset$ for $1 \leq j \leq s-1$. The following Lemma ends the story in the abstract case (see [2, Lemma 23.8]):

Lemma 28 *Let* $T_1, \ldots, T_s$ *be a transition chain. Then an arbitrary neighborhood* U *of an arbitrary point* $\xi \in Y_1^-$ *is connected with an orbit* $\zeta(t)$ *to an arbitrary neighborhood* V *of an arbitrary point* $\eta \in Y_s^+$.

In view of Lemma 28, the proof of Theorem 26 reduces to the identification of a transition chain $T_1, \ldots, T_s$ for the system (37) with $I_2 < A$ on T_1 and $I_2 > B$ on T_s.

The transition tori T_ω that make the transition chain are defined by the equations

$$I_1 = \phi_1 = I_2 - \omega = 0$$

for ω irrational. That these tori are invariant is clear, and their whiskers can be explicitly written for $\mu = 0$. That they are still whiskered tori of the full perturbed system (37) requires an additional argument using contracting maps (see [2, Ch. 3,§15]). The following Lemma ends the proof of Theorem 26:

Lemma 29 *Assume $\tilde{A} < \omega < \tilde{B}$. Then there exists a constant*

$$\kappa = \kappa(\tilde{A}, \tilde{B}, \epsilon, \mu) > 0$$

so that for all $|\omega' - \omega| < \kappa$ we have that $Y_\omega^+ \bigcap Y_{\omega'}^- \neq \emptyset$.

The proof of this Lemma involves first order perturbative calculations in μ. Its statement allows us to choose a transition chain of tori $T_{\omega_1}, \ldots, T_{\omega_s}$ with $I_2(T_{\omega_1}) = \omega_1 < A$ and $I_2(T_{\omega_s}) = \omega_s > B$. But in this case, by Lemma 28, there exists a trajectory $\zeta(t)$ that passes arbitrarily close to both T_{ω_1} and T_{ω_s}; on this trajectory, I_2 evolves from being less than A to being greater than B in finite time, which is exactly the result that we were seeking.

References

[1] Arnol'd, V.I. *Instability of dynamical systems with several degrees of freedom.* Soviet Math **5** [1964], no.3, pp. 581-585.

[2] Arnol'd, V. I., Avez, A. *Ergodic Problems in Classical Mechanics.* Benjamin, New York, 1967.

[3] Wayne, C. E. *An introduction to KAM theory.* Dynamical systems and probabilistic methods in partial differential equations, pp. 3-29, Lectures in Applied Mathematics **31**, Amer. Math. Soc., Providence, RI, 1996.

5 Compatible brackets in Hamiltonian mechanics

after H. P. McKean
A summary written by Irina Nenciu

Abstract

We give a summary of the results in [3]. In the first part we present the general bi-Hamiltonian approach to complete integrability for Hamiltonian systems, due to [1]. The second part is dedicated to explaining Magri's [2] result on how to construct such a bi-Hamiltonian structure.

Introduction and general results

Our main purpose in this section is to understand how the existence and compatibility of two brackets for a Hamiltonian system allows one to find a sequence of commuting Hamiltonians for that system.

Let d be fixed and let $C^\infty(\mathbf{R}^{2d})$ be a Lie algebra with the bracket

$$[H_1, H_2]_J = (\nabla H_1)^t J \nabla H_2 \tag{38}$$

where ∇ is the gradient and $J : R^{2d} \to Gl(2d, \mathbf{R})$ is smooth and skew. Assume also that there exists a second such bracket

$$[H_1, H_2]_K = (\nabla H_1)^t K \nabla H_2.$$

Definition 30 *Given J and K as above, we say that a Hamiltonian H_0 can be raised if there exists another Hamiltonian H_1 so that*

$$[f, H_0]_K = [f, H_1]_J \tag{39}$$

for all $f \in C^\infty(\mathbf{R}^{2d})$. In this case we say that H_0 raises to H_1 or that H_1 lowers to H_0, and denote these relations by $H_0 \uparrow H_1$ and $H_1 \downarrow H_0$, respectively.

We immediately obtain the following

Lemma 31 *Consider J and K as above and assume that there exists a sequence of Hamiltonians so that $H_0 \uparrow H_1 \uparrow H_2 \uparrow \dots$. Then*

$$[H_i, H_j]_J = [H_i, H_j]_K = 0 \tag{40}$$

for all $i, j \geq 0$.

The next natural question is when is it possible to have such infinite sequences of raising Hamiltonians $H_0 \uparrow H_1 \uparrow H_2 \uparrow \dots$. In other words, we ask for a condition on J and K so that, if a Hamiltonian H_0 can be raised to H_1, then H_1 itself can be raised. Clearly this is equivalent to saying that a Hamiltonian can be raised if and only if it can be lowered. So the next Lemma gives the natural condition to answer our question.

Lemma 32 *If J, K and $J+K$ all produce honest brackets (i.e. the brackets they define obey the Jacobi identity), then the class of Hamiltonians that can be raised coincides with the class of those that can be lowered.*

In this case we say that J and K are compatible and write that $\uparrow = \downarrow$.

Magri's result

In this section we concern ourselves with the question of finding a second, compatible, bracket for a given completely integrable Hamiltonian system. To achieve this, let us first assume the existence of such a bracket, and find a necessary form for it. An essential role in this study is played by the spectrum of the matrix $J^{-1}K$ and, in particular, by the traces of its powers.

Let $\pi_k = \mathrm{Tr}(J^{-1}K)^k$ for all $k \geq 1$. Then one can prove that

$$\pi_1 \uparrow \frac{\pi_2}{2} \uparrow \frac{\pi_3}{3} \uparrow \dots \tag{41}$$

and, consequently, the π_k's commute. Let us moreover assume that the spectrum of $J^{-1}K$ is *ample*, in the sense that the gradients $\nabla\pi_1, \cdots, \nabla\pi_d$ are independent. In this case, fixing the values of the π_k's defines a smooth d-dimensional manifold $M \subset \mathbf{R}^{2d}$ on which the $\nabla\pi_k$'s are normal and the $J\nabla\pi_k$'s are tangent for all $1 \leq k \leq d$.

So if we consider J and K to be compatible and the spectrum of $J^{-1}K$ to be ample, the functions $H_k = \frac{\pi_k}{k}$ $(1 \leq k \leq d)$ commute and can be completed by functions G_k $(1 \leq k \leq d)$, canonically conjugate relative to J. One can then, using the fact that $H_k \uparrow H_{k+1}$, find the brackets of the H's and G's relative to K. But ∇H_k and ∇G_k $(1 \le k \le d)$ span $\mathbf{R}^{2d}$, and hence K is uniquely determined by Magri's rule:

$$\begin{aligned} K = & \textstyle\sum_{1\le i,j\le d} c_{i,j}(J\nabla H_i \otimes J\nabla H_j - J\nabla H_j \otimes J\nabla H_i) \\ & + \textstyle\sum_{j=1}^{d}(J\nabla H_j^+ \otimes J\nabla G_j - J\nabla G_j \otimes J\nabla H_j^+) \end{aligned}$$

where the coefficients $c_{ij} = c_{ij}(H_1, \dots, H_d) = [G_i, G_j]_K$ are constant on M. Moreover, both 2-forms

$$c_- = \sum_{i<j} c_{ij} dH_i \wedge dH_j$$

and

$$c_+ = \sum_{i<j} (dH_i^+ \wedge dH_j + dH_i \wedge dH_j^+),$$

are closed.

By following essentially the backwards road, one can finally prove the main theorem:

Theorem 33 *If H_i, $i \ge 1$, is any family of commuting functions, if the supplementary family G_j, $1 \le j \le d$, is canonically paired, if H_j, $j > d$, stands in the same relation to H_i, $1 \le i \le d$, as $\frac{\pi_j}{j}$, $j > d$, does to $\frac{\pi_i}{i}$, $1 \le i \le d$, and if $[c_{ij}(H_1, \dots, H_d)]_{1\le i,j\le d}$ is such that both forms c_- and c_+ defined above are closed, then, with $H_n^+ = H_{n+1}$, $1 \le n \le d$, the skew form K produced by Magri's rule is compatible with J. Moreover, $H_1 \uparrow H_2 \uparrow \dots \uparrow H_d$.*

References

[1] Magri, F. *A simple model of an integrable Hamiltonian equation.* J. Math. Phys. **19** [1978], pp. 1156-1162.

[2] Magri, F. Private communication to H. P. McKean, [1989]

[3] McKean, H. P. *Compatible brackets in Hamiltonian mechanics.* Important Developments in Soliton Theory, pp. 344-354, Springer Ser. Nonlinear Dynamics, Springer, Berlin, 1993.

6 Construction of Quasi-Periodic Solutions for Hamiltonian Perturbations of Linear Equations

after Jean Bourgain
A summary written by Dmitry Pavlov

Abstract

We give a summary of the results in [1]. Detailed explanation of the technique can be also found in [3]. Further development of this approach can be found in [2].

Linear case

We will consider the linear system

$$\begin{cases} \dot{p}_\alpha = -\lambda_\alpha q_\alpha, \\ \dot{q}_\alpha = \lambda_\alpha p_\alpha, \\ \dot{p}_\beta = -\mu_\beta q_\beta, \\ \dot{q}_\beta = \mu_\beta p_\beta, \end{cases} \tag{42}$$

where $\alpha = 1, 2$ and $\beta > 2$. This system has a quasi-periodic solution $p_\alpha = a_\alpha \cos \lambda t$, $q_\alpha = a_\alpha \sin \lambda_a t$, $(\alpha = 1, 2)$, $p_\beta = q_\beta = 0$ $(\beta > 2)$. Our aim is to prove persistency of this solution under small Hamiltonian perturbation of the linear system.

Let's denote $u_\alpha = p_\alpha + iq_\alpha$, $v_\alpha = p_\alpha - iq_\alpha$. Then we can write the perturbed system as

$$\begin{cases} -i\dot{u}_\alpha = \lambda_\alpha u_\alpha + \xi_\alpha u_\alpha + \varepsilon \partial_{v_\alpha} H, \\ i\dot{v}_\alpha = \lambda_\alpha v_\alpha + \xi_\alpha v_\alpha + \varepsilon \partial_{u_\alpha} H, \\ -i\dot{u}_\beta = \mu_\beta u_\beta + \varepsilon \partial_{v_\beta} H, \\ i\dot{v}_\beta = \mu_\beta v_\beta + \varepsilon \partial_{u_\beta} H. \end{cases} \tag{43}$$

Here

$$\xi_\alpha = \xi_\alpha(\varepsilon, \lambda), \quad \xi_\alpha(0, \lambda) = 0$$

is the frequency correction. The main result is given by the following theorem.

Theorem 34 *Consider the system (43) with H polynomial in u, $\bar{u}$. Let $\lambda = (\lambda_1, \lambda_2)$ be a parameter and μ_β depends smoothly on λ. Assume the following nonresonance condition:*

$$|\mu_\beta(\lambda) + m_1\lambda_1 + m_2\lambda_2| > N_0^{-C}, \tag{44}$$

for $\beta > 2$, $|m_1|, |m_2| < N_0$ and N_0 sufficiently large.

Then there is persistency of the unperturbed solution given above for λ taken outside an exceptional set, which has measure tending to 0 when $\varepsilon \to 0$. This set also depends on the nonlinearity and a_α. The perturbed solution $u_\alpha = \sum_m \hat{u}_\alpha(m_1, m_2) e^{i(m_1\lambda_1 + m_2\lambda_2)t}$ satisfies an estimate

$$\sum_m |\hat{u}_\alpha(m)| e^{|m|^c} < \infty \tag{45}$$

for some $c > 0$.

Basic Construction

To prove the persistency of the quasi-periodic solution we will solve by Newton's method the following system (P-equation):

$$\begin{cases} (-\lambda_1 m_1 - \lambda_2 m_2 + (\lambda_\alpha + \xi_\alpha))\hat{u}_\alpha(m) + \varepsilon \widehat{\partial_{v_\alpha} H}(m) = 0 & (\alpha = 1, 2;\ m \neq e_\alpha), \\ (\lambda_1 m_1 + \lambda_2 m_2 + (\lambda_\alpha + \xi_\alpha))\hat{v}_\alpha(m) + \varepsilon \widehat{\partial_{u_\alpha} H}(m) = 0 & (\alpha = 1, 2;\ m \neq -e_\alpha), \\ (-\lambda_1 m_1 - \lambda_2 m_2 + \mu_\beta)\hat{u}_\beta(m) + \varepsilon \widehat{\partial_{v_\beta} H}(m) = 0 & (\beta > 2), \\ (\lambda_1 m_1 + \lambda_2 m_2 + \mu_\beta)\hat{v}_\beta(m) + \varepsilon \widehat{\partial_{u_\beta} H}(m) = 0 & (\beta > 2), \end{cases} \tag{46}$$

where $e_1 = (1, 0)$, $e_2 = (0, 1)$. The frequency correction is determined by the equation

$$\xi_\alpha \hat{u}_\alpha(e_\alpha) = -\varepsilon \widehat{\partial_{v_\alpha} H}(e_\alpha), \quad \alpha = 1, 2.$$

Recall the algorithm. We are solving the system

$$F(y) = 0.$$

Define the approximations by the formula

$$y_{i+1} = y_i - (F'(y_i))^{-1} F(y_i).$$

In our case $y_0 = 0$ and the linearized operator S has the form $S = D + \varepsilon T$, where T depends on $y_i = (\hat{u}_i, \hat{v}_i)$ and D is the diagonal

$$D_{m,\alpha,\pm 1} = \begin{cases} \mp < \lambda, m > +\lambda_\alpha \text{ for } \alpha = 1,2 \\ \mp < \lambda, m > +\mu_\alpha \text{ for } \alpha > 2. \end{cases}$$

The problem now is to estimate the norms of the inverses of operators S arising in this process. The main difficulty here is that the diagonal elements may be arbitrarily small. Using certain multiscale analysis we can obtain estimates for S^{-1} at each step of the Newton's method. Also at each step we get an exceptional λ-parameter set on which these estimates don't hold. Estimating measure of these sets we can ensure that the total measure is small provided ε-parameter is small.

The NLS case

Consider 1D nonlinear Schrodinger equation (NLS)

$$iu_t - u_{xx} + g(x)u + \varepsilon \partial_{\bar{u}} H(u, \bar{u}) = 0, \tag{47}$$

where g is a real analytic function on $\mathbf{T}$. Denote by $\o_n$ the spectrum of the operator $-u_{xx} + g(x)u$ and let ψ_n be the corresponding eigenvalues. We will find quasi-periodic solutions in the form

$$u = \sum_{m_1, m_2, n} \hat{u}(m_1, m_2, n) e^{i(m_1\lambda_1 + m_2\lambda_2)t} \psi_n(x). \tag{48}$$

Using this form of the solution we can rewrite the equation as

$$(-(\lambda_1 m_1 + \lambda_2 m_2) + \o_n)\hat{u}(m_1, m_2, n) + \varepsilon \widehat{\partial_{\bar{u}} H}(m_1, m_2, n) = 0. \tag{49}$$

The main result of this section is the following theorem analogous to Theorem 1.

Theorem 35 *Let H be a polynomial and consider $\lambda = (\lambda_1, \lambda_2)$ as a parameter. Assume the condition 44 holds.*

Then there is persistency of the unperturbed solution $u = a_1 e^{i\lambda_1 t}\psi_{n_1}(x) + a_2 e^{i\lambda_2 t}\psi_{n_2}(x)$ for the perturbed system for λ taken outside a set of measure tending to 0 when $\varepsilon \to 0$. The perturbed solution satisfies the condition

$$\sum_{m,n} |\hat{u}(m,n)| e^{(|m|+|n|)^c} < \infty$$

for some $c > 0$.

Basic Construction

Denoting $v = \bar{u}$ we can rewrite the equation (49) as

$$\begin{cases} (-(\lambda_1 m_1 + \lambda_2 m_2) + \o_n)\hat{u}(m_1, m_2, n) + \varepsilon\widehat{\partial_v H}(m_1, m_2, n) = 0, \\ (-(\lambda_1 m_1 + \lambda_2 m_2) + \o_n)\hat{v}(m_1, m_2, n) + \varepsilon\widehat{\partial_u H}(m_1, m_2, n) = 0. \end{cases}$$

Then the linearized operator has the form $S = D + T$, where

$$D = \begin{pmatrix} (-(\lambda_1 m_1 + \lambda_2 m_2) + \o_n)\mathbb{I} & 0 \\ 0 & (\lambda_1 m_1 + \lambda_2 m_2 + \o_n)\mathbb{I} \end{pmatrix}$$

and

$$T = \varepsilon \begin{pmatrix} -T_{\partial^2_{uv}H} & T_{\partial^2_{vv}H} \\ T_{\partial^2_{uu}H} & T_{\partial^2_{uv}H} \end{pmatrix}.$$

Here T_f is the operator of multiplication by f if Fourier series.

Like in the previous case Newton's method is used to solve this system. Again the diagonal elements are small and multiscale analysis can be used to get estimates for the inverse operators. The technique used in both parts is essentially the same.

References

[1] J.Bourgain, *Construction of Quasi-Periodic Solutions for Hamiltonian Perturbations of Linear Equations and Applications to Nonlinear PDE,* International Math Research Notices 1994 No. **11** pp. 475-497.

[2] J.Bourgain, *Quasi-Periodic Solutions of Hamiltonian Perturbations of 2D Linear Shrodinger Equations*, The Annals of Math, vol. **148**, no. 2, (1998), 363–439.

[3] W.Craig and C.Wayne, *Newton's Method and Periodic Solutions of Nonlinear Wave Equations,* Comm. Pure Appl. Math. **46** (1993), 1409–1501.

7 Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations

after Michael Christ, James Colliander and Terrence Tao
A summary written by Stefanie Petermichl

Abstract

We give a summary of the results in [1] concerning the defocusing cubic non-linear Schrödinger equation. We construct classes of solutions and use them to demonstrate lack of local well-posedness in Sobolev spaces H^s for $s < 0$. Here we mean well-posedness in the sense that the dependence of solutions upon initial data fails to be uniformly continuous.

Introduction

The Cauchy problem for the cubic one-dimensional defocusing non-linear Schrödinger equation (NLS) is

$$\begin{cases} -iu_t + u_{xx} = |u|^2 u; \quad u : (0, T) \times \mathbb{R}_x \mapsto \mathbb{C} \\ u(0, x) = u_0(x) \end{cases} \tag{50}$$

where u_0 is an element of a Sobolev space H^s_x for some $s \in \mathbb{R}$.

If the initial datum is in the Schwartz space $\mathcal{S}$, then there exists a unique global smooth solution u. For fixed t, this gives rise to a non-linear evolution operator

$$S(t) : \mathcal{S} \to \mathcal{S}, u_0 \mapsto u(t)$$

and a well-defined solution map

$$S : \mathcal{S} \to C^\infty_t(\mathbb{R}, \mathcal{S}), u_0 \mapsto u.$$

We ask if S can be extended to the space H^s_x for some $s \in \mathbb{R}$.

Definition 36 *If for every radius $R > 0$ there exists a time $T > 0$ such that S can be uniformly continuously and uniquely extended to a map from the $R-$ball in H^s_x to $C^0_t((-T, T); H^s_x)$, then we say that (50) is locally well-posed in H^s_x.*

If one can make T arbitrarily large and independent of R, then we say that (50) is globally well-posed in H^s_x.

The following is due to Tsutsumi [2]

Theorem 37 *If $s \geq 0$, then the equation (50) is globally well-posed in H^s_x.*

We raise the question what happens for $s < 0$ and revisit the case $s = 0$.

Proofs

It will be helpful to transform the problem to a 'backward Cauchy problem' via the pseudo-conformal transformation or *pc*-transform. Introduce the change of variables

$$(t, x) = (y/s, 1/s - 1) \quad (s, y) = (1/(t+1), x/(t+1))$$

and $v = pc(u), u = pc^{-1}(v)$ by

$$u(t, x) := (1+t)^{-1/2} \exp{(-ix^2/4(t+1))} v(s, y)$$

$$v(s, y) := s^{-1/2} \exp{(iy^2/4s)} u(t, x).$$

The map *pc* transforms the Cauchy problem (50) into a backwards Cauchy problem

$$\begin{cases} iv_s + v_{yy} = s^{-1}|v|^2 v; \\ v(1, y) = v_1(y); 0 < s \leq 1 \end{cases} \tag{51}$$

where $v_1(y) := \exp{(iy^2)} u_0(y)$.

For each fixed time t, the map $u(t) \mapsto v(s)$ is a linear isometry in $L^2_x = H^0_x$ but is not well-behaved in other Sobolev spaces, so we work in weighted Sobolev spaces. For $k \geq 0$ define $H^{k,k}_x$ to be the closure of the Schwartz class under the norm

$$\|u\|_{H^{k,k}_x} := \sum_{i,j \geq 0 : i+j \leq k} \|x^i \partial^j_x u\|_{L^2_x}.$$

Now we can control the Sobolev norm of u by the weighted Sobolev norm of its *pc*-transform. We can also control the behavior of the $H^{k,k}_x$ spaces under pointwise multiplication, *pc*-transform at the endpoint $t = 0$ and free Schrödinger flow $u \mapsto \exp{(it\partial_{xx})} u$.

We drop the dispersive term v_{yy} and end up with the associated ODE

$$iv_s = s^{-1}|v|^2 v \tag{52}$$

for which there are explicit solutions

$$v^{[w]}(s,y) = w(y)\exp\left(-i|w(y)|^2 \log s\right)$$

for any function $w(y)$ which can be thought of as initial datum. We denote by $v^{[w]}$ an approximate solution and by $v^{\langle w\rangle}$ an associated exact solution to (51).

The following asymptotic completeness result is crucial to our proofs of ill-posedness:

Lemma 38 *Let $K \geq 5$ be an integer and let $w \in H_y^{K+2,K+2}$ have norm of order ε. Then, if ε is sufficiently small, there exists an initial datum $v_1(y) \in H_y^{K,K}$ such that the unique solution $v = v^{\langle w\rangle}$ to the backwards Cauchy problem (51) with initial datum v_1 satisfies*

$$\|v^{\langle w\rangle}(s) - v^{[w](s)}\|_{H_y^{K,K}} \lesssim \varepsilon s(1+\log s)^C \text{ for all } 0 < s \leq 1. \tag{53}$$

Furthermore, the map $w \mapsto v$ is Lipschitz continuous from the ε-ball in the $H^{K+2,K+2}$-norm to the space $L_s^\infty((0,1]; H_y^{K,K})$, meaning

$$\sup_{0<s\leq 1} \|v^{\langle w\rangle}(s) - v^{\langle \tilde{w}\rangle}(s)\|_{H_y^{K,K}} \lesssim \|w-\tilde{w}\|_{H_y^{K+2,K+2}} \tag{54}$$

Proof.

We summarize the proof of lemma (38). Using perturbation analysis we make the Ansatz: $v^{\langle w\rangle} = v^{[w]} + \Phi$. Then Φ will solve the DE

$$i\Phi_s + \Phi_{yy} = -v_{yy}^{[w]} + s^{-1}F(\Phi)$$

where $F(\Phi) = 2|v^{[w]}|^2\Phi + 2v^{[w]}|\Phi|^2 + (v^{[w]})^2\overline{\Phi} + \overline{v^{[w]}}\Phi^2$. We proceed by solving the forward Cauchy problem

$$\begin{cases} i\Phi_s + \Phi_{yy} = -v_{yy}^{[w]} + s^{-1}F(\Phi); \\ \Phi(0,y) = 0; 0 < s \leq 1 \end{cases} \tag{55}$$

in the sense that $\|\Phi(s)\|_{H_y^{K,K}} \to 0$ as $s \to 0^+$. We then define the intitial datum for (51) to be $v_1(y) = v^{\langle w\rangle}(1,y)$.

We write equation (55) in integral form:

$$\Phi(s) = \int_0^s \exp\left(i(s-\tilde{s})\partial_{yy}\right)(-v_{yy}^{[w]}(\tilde{s}) + \tilde{s}^{-1}F(\Phi(\tilde{s})))d\tilde{s}$$

and set up an iteration

$$\Phi^{(0)} = 0$$

$$\Phi^{(k+1)}(s) = \int_0^s \exp\left(i(s-\tilde{s})\partial_{yy}\right)(-v_{yy}^{[w]}(\tilde{s}) + \tilde{s}^{-1}F(\Phi^{(k)}(\tilde{s})))d\tilde{s}.$$

One can show inductively that

$$\|\Phi^{(k)}(s)\|_{H^{j,j}} \lesssim \varepsilon s(1+\log s)^C$$

for all $0 < s \leq 1, 0 \leq j \leq K$ and all k.

The iterates $\Phi^{(k)}$ converge in $L_s^\infty((0,1];H^{K,K})$ to a solution Φ of the forward Cauchy problem (55) with $\|\Phi(s)\|_{H_y^{K,K}} \to 0$ as $t \to 0^+$. We see that $v_1(y) = v^{\langle w\rangle}(1,y) = v^{[w]}(1,y) + \Phi(1,y)$ belongs to the space $H_y^{K,K}$ and we can obtain the Lipschitz bound in the statement of the lemma.

The case $s = 0$.

Now let $K = 5$ and consider $w(y) = \varepsilon \exp(-y^2)$. Its $H^{7,7}$ norm is clearly of order ε. For any $a \in [1/2, 2]$ apply lemma (38) to aw. The map

$$a \mapsto v^{\langle aw\rangle}(1,y)$$

is continuous in the $H_y^{5,5}$-norm, but there is a decoherence property as $s \to 0^+$ for $a \neq \tilde{a} \in [1/2, 2]$:

$$\limsup_{s\to 0^+} \|v^{[aw]} - v^{[\tilde{a}w]}\| \gtrsim (a+\tilde{a})\|w\|_{L_y^2}.$$

Since exact solutions and approximate solutions are close even in the $H_y^{K,K}$-norm, we have that

$$\limsup_{s\to 0^+} \|v^{\langle aw\rangle} - v^{\langle \tilde{a}w\rangle}\|_{L_y^2} \gtrsim (a+\tilde{a})\varepsilon$$

and cannot be arbitrarily small. This shows that the backwards Cauchy problem is not uniformly well-posed in L^2 on the backward interval $(0,1]$, even for initial data with arbitrarily small $H^{5,5}$-norm.

Now let $u^{\langle aw \rangle} := pc^{-1}(v^{\langle aw \rangle})$. This is a solution to (50) (but not with initial datum $u(0) = aw$!).

We can control the pc-transform through the weighted Sobolev norms, so we can conclude that $\|u^{\langle aw \rangle}(t)\|_{H^5_x}$ is uniformly small and $u^{\langle aw \rangle}$ inherits a decoherence property at $t \to +\infty$. We conclude that even in L^2 the solution map fails to be uniformly continuous, so the global well-posedness in L^2 is not uniform in time.

Ill-posedness for $s < 0$

The solutions $u^{\langle w \rangle}$ to (50) constructed in this section will be used to disprove uniform continuity of the solution operator $\mathcal{S}$ in H^s for $s < 0$. Let $0 < \delta << \epsilon << 1$ and let $T > 0$ be arbitrary. We will find two solutions $\Phi^{\langle a \rangle}, \Phi^{\langle \tilde{a} \rangle}$ to (50) with

$$\|\Phi^{\langle a \rangle}(0)\|_{H^s_x}, \|\Phi^{\langle \tilde{a} \rangle}(0)\|_{H^s_x} \lesssim \epsilon \tag{56}$$

$$\|\Phi^{\langle a \rangle}(0) - \Phi^{\langle \tilde{a} \rangle}(0)\|_{H^s_x} \lesssim \delta \tag{57}$$

$$\sup_{0 \le t < T} \|\Phi^{\langle a \rangle}(t) - \Phi^{\langle \tilde{a} \rangle}(t)\|_{H^s_x} \gtrsim \epsilon \tag{58}$$

Since $\dot{H}^{-1/2}$ is invariant under scaling, the scale invariance

$$u(t,x) \mapsto \lambda^{-1} u(1/\lambda^2, x/\lambda)$$

suggests that scaling should help us for the case $s < -1/2$. Also, L^2 is invariant under the Gallilean transformation $u(t,x) \mapsto \exp{(i\alpha x/2 + i\alpha^2 t/4)} u(t, x + \alpha t)$, suggesting we try this for other $s < 0$. Indeed, we have to distinguish between the cases $s < -1/2$, $s = -1/2$ and $s \in (-1/2, 0)$.

Subcritical case, $s \in (-1/2, 0)$

We use the Gallilean transformations and scaling of $u^{\langle aw \rangle}$ for this range of s, with w as before. For parameters $a, \tilde{a} \in [1/2, 2]$, $N >> 0$ and $\lambda > 0$, consider

$$\Phi^{\langle a \rangle}(t,x) = \lambda \exp{(iNx + iN^2 t)} u^{\langle aw \rangle}(\lambda^2 t, \lambda(x + 2tN))$$

and similarly for $\Phi^{\langle \tilde{a} \rangle}$. If one chooses the parameters correctly, these solutions will satisfy (56) through (58).

Critical case, $s = -1/2$

We want to use scaling alone and consider

$$\Phi^{\langle a \rangle}(t, x) = \lambda u^{\langle aw \rangle}(\lambda^2 t, \lambda x).$$

There is a slight issue of making the $\dot{H}^{-1/2}$-norm converge. This can be overcome by choosing w odd and noticing that the oddness reproduces itself in $v^{\langle aw \rangle}$ and hence in the intital datum $u^{\langle aw \rangle}(0)$. Picking λ correctly, these scaled solutions will satisfy (56) through (58).

Supercritical case, $s < -1/2$

The construction for this case does not rely on the solutions $u^{\langle aw \rangle}$ constructed above but rather ones arising from the small dispersion NLS:

$$\begin{cases} -iv_t + \delta v_{xx} = |v|^2 v; \\ v(0, x) = aw(x) \end{cases} \tag{59}$$

with $0 < \delta \leq 1$, a in the unit ball of $\mathbb{C}$ and $w \in \mathcal{S}$. We have good control over the closeness of the exact and approximate solution for times before $T(\delta)$ where $T(\delta) \to \infty$ as $\delta \to 0$. Careful considerations lead to a decoherence property for certain times τ for the solution v of (59) . We now consider δ-dilates of v, which will be exact solutions of (50) with intital datum $aw(\delta x)$. The correct λ-scaling will produce solutions satisfying (56) through (58).

References

[1] M. Christ, J. Colliander, T. Tao, *Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations,* Amer. J. of math. **125** [2003], pp. 1235-1293.

[2] Y. Tsutsumi, L^2 *solutions for non-linear Schrödinger equations and non-linear groups,* Funk. Ekva.**30** [1987], pp. 115-125.

8 The Hofer-Zehnder capacity

Olga Radko

Abstract

We give an exposition of a special class of invariants of symplectic manifolds, called *symplectic capacities.* These invariants provide the link between the field of symplectic topology (the study of global phenomenon on symplectic manifolds) and Hamiltonian dynamics. In particular, the Hofer-Zehnder capacity is related to the behavior of periodic solutions of Hamiltonian equations.

8.1 Symplectic geometry preliminaries

8.1.1 From Hamiltonian mechanics to symplectic geometry.

The equations of motion in classical mechanics arise as solutions of variational problems. Let $\mathbb{R}^{2n}$ be the phase space of a system with *Hamiltonian coordinates* $(q,p) \in \mathbb{R}^{2n}$, where $q = (q_1, \ldots, q_n)$ are space coordinates, and $p = (p_1, \ldots, p_n)$ are momenta of the system. Let H be a smooth Hamiltonian function. The associated *action functional* of a path $z(t) = (q(t), p(t)) : [0,1] \to \mathbb{R}^{2n}$ is defined by

$$\Phi_H(z(t)) = \int_0^1 \left(\langle p, \dot{q} \rangle - H(q,p,t)\right) dt. \tag{60}$$

The corresponding equations of motions are the well-known Hamiltonian equations

$$\dot{z}(t) = -J\nabla H(z(t)), \tag{61}$$

where ∇H denotes the gradient of H and J is the $2n \times 2n$-matrix $J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$. The vector field $X_H = -J\nabla H$ is called the *Hamiltonian vector field* of H, generating the Hamiltonian flow φ_H^t.

This setting for Hamiltonian dynamics can be translated into the language of symplectic geometry. The differential 2-form $\omega_0 = \sum_{i=1}^n dq_i \wedge dp_i$ on $M = \mathbb{R}^{2n}$ is called the (standard) *symplectic form.* In terms of such a form, the Hamiltonian vector field of H is determined by the identity

$$\omega_0(X_H, Y)(z) = dH(Y)(z) \quad \forall Y \in T_zM,\ z \in M.$$

More generally, a *symplectic structure* on a $2n$-dimensional manifold is defined as a 2-form ω which is closed ($d\omega = 0$) and non-degenerate ($\forall v \in T_x M$ the condition $\omega(v, w)(x) = 0$ $\forall w \in T_x M$ implies $v(x) = 0$).

A *symplectomorphism*, i.e., diffeomorphism $\phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ preserving the symplectic form, $\phi^* \omega_0 = \omega_0$, corresponds to a canonical transformation in classical mechanics.

Locally all symplectic manifolds of a fixed dimension look the same: Darboux's theorem states that for any point x on a $2n$-dimensional symplectic manifold (M, ω), there is a local coordinate system $(q_i, p_i)_{i=1}^n$ in a neighborhood U of x in which $\omega_{|U} = \sum_{i=1}^n dq_i \wedge dp_i$. The change of coordinates in this neighborhood is, of course, a (local) symplectomorphism. Thus, the only non-trivial *local* invariant of a symplectic structure is the dimension of the manifold.

8.1.2 The problem of symplectic embeddings and symplectic invariants

Question. Given two open subsets U and V of $\mathbb{R}^{2n}$, under what conditions does there exist a symplectic embedding $(U, \omega_0) \to (V, \omega_0)$?

It is clear that one must require $\mathrm{vol}(U) \leq \mathrm{vol}\ (V)$, where $\mathrm{vol}(U) = \int_U \omega^{\wedge n}$ is the Liouville volume, since any symplectic embedding must preserve the Liouville volume form $\omega^{\wedge n}$. However, the following example shows that there should be other invariants. Let (x, y) be the Darboux coordinates on $\mathbb{R}^{2n}$ such that $\omega = \sum_{i=1}^n dx_i \wedge dy_i$. Let

$$B(r) = \{(x, y) \in \mathbb{R}^{2n} : |x|^2 + |y|^2 < r^2\} \tag{62}$$

$$Z(R) = \{(x, y) \in \mathbb{R}^{2n} : x_1^2 + y_1^2 < R^2\} \tag{63}$$

be a symplectic ball of radius r and a cylinder based over a symplectic disc of radius R. What is the relation of r and R which guarantees that a symplectic embedding $B(r) \to Z(R)$ exists?[1] It is not a problem to construct (a linear) *volume*-preserving embedding. Take, e.g.,

$$\psi(x, y) = \left(\varepsilon x_1, \frac{x_2}{\varepsilon}, x_3, \ldots, x_n; \varepsilon y_1, \frac{y_2}{\varepsilon}, y_3, \ldots, y_n\right)$$

[1] It is crucial that the base disc of the cylinder lies in a symplectic subspace. E.g., if the base lies in a lagrangian subspace, a symplectic embedding $B(r) \to Z(R)$ exists independently of the relation between $r > 0$ and $R > 0$.

for ε sufficiently small. However, this map is not symplectic unless $\varepsilon = 1$, which implies $r \leq R$. For $r > R$, it is easy to see that there are no other *linear* map giving a symplectic embedding. More surprisingly, as Gromov [Gro85] proved using the technique of J-holomorphic curves, the same is true even for non-linear maps:

Theorem 39 *(Gromov's (non)-squeezing theorem). A symplectic embedding $B(r) \to Z(R)$ exists iff $r \leq R$.*

This discovery showed that the symplectic maps are much more rigid then the volume preserving maps, and motivated the definition of symplectic capacities in the work of Ekeland and Hofer [EH89].

The first symplectic capacity, motivated by the example above, was the *Gromov width*, defined by

$$w_G((M,\omega)) = \sup\{\pi r^2 : \; B^{2n}(r) \text{ embeds symplectically into } M\}. \quad (64)$$

8.1.3 Periodic orbits of Hamiltonian vector fields

Another problem which lead to the development of symplectic capacities is the existence of periodic orbits of Hamiltonian systems. Let, as before, $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian, and X_H be the corresponding Hamiltonian vector field. The flow φ_H^t of this vector field leaves H invariant. A *regular energy surface* is a compact surface $S = \{x \in \mathbb{R}^{2n} \; : \; H(x) = \text{const}\}$ such that $dH(x) \neq 0$ for $x \in S$. Recall the following *Poincare's recurrence theorem*:

Theorem 40 *(Poincare) Let S be a compact regular energy surface of H. Then for almost any $x \in S$ there is an increasing sequence $t_i \to \infty$ satisfying $\lim_{i\to\infty} \varphi^{t_i}(x) = x$.*

This theorem motivates the search for periodic phenomena on S, thus leading to the following

QUESTION: Let S be a compact regular energy surface of the Hamiltonian vector field X_H in $\mathbb{R}^{2n}$. Is there a periodic orbit of X_H lying on S?

This question is still open. The breakthrough results of Weinstein [Wei78] and Rabinowitz [Rab78] established the existence of periodic orbits in the case of convex and star-shaped hypersurfaces, respectively. The proofs used variational techniques, which were later applied in the Viterbo's proof [Vit87] of Weinstein's conjecture [Wei79] that every hypersurface of contact type carries a periodic orbit. Similar variational methods are used in the proof of existence of a capacity function (Section 8.4).

8.2 Capacities

8.2.1 Definition and general properties

Consider symplectic manifolds (M, ω) of fixed dimension $2n$.

Definition 41 *A* symplectic capacity *is a map c associating to each symplectic manifold a non-negative number or infinity, such that the following axioms are satisfied:*

A1. Monotonicity: $c(M, \omega) \leq c(N, \tau)$ *if there is a symplectic embedding* $\varphi : (M, \omega) \to (N, \tau)$.

A2. Conformality: $c(M, \alpha\omega) = |\alpha| c(M, \omega)$ *for any* $\alpha \neq 0$.

A3. Normalization : $c(B(1), \omega_0) = \pi = c(Z(1), \omega_0)$,
for the open unit ball $B(1)$ and open symplectic cylinder $Z(1)$ in $(\mathbb{R}^{2n}, \omega_0)$.

To see that $c(M, \omega)$ is indeed a symplectic invariant, apply A1 to a symplectomorphism φ and its inverse φ^{-1} to see that $c(M, \omega) = c(\varphi(M, \omega))$. Moreover, A1 implies that capacities are obstructions to symplectic embeddings. One can show that the existence of a capacity function (which is highly nontrivial) turns out to be equivalent to the Gromov's (non-)squeezing theorem.

On the other hand, the axioms above do not characterize the capacity function uniquely. We will concentrate on Hofer-Zehnder capacity in Section 8.4.

Note that the symplectic invariant $(\text{vol})^{1/n}$ on $\mathbb{R}^{2n}$ is not a capacity (for $n \neq 1$), since it does not satisfy the normalization axiom: the symplectic volume of a cylinder $Z(R)$ is infinite.

The situation is special in dimension $2n = 2$, where all the capacities coincide:

Proposition 42 *(Siburg [Sib93]) The only capacity function on $(\mathbb{R}^2, \omega_0)$ satisfying axioms A1-A3 is the total volume (or, equivalently, the Lebesgue measure): for a compact connected domain (D, ω_0) we have $c(D, \omega_0) = vol(D, \omega_0)$.*

8.2.2 The linear case

The special case of the symplectic embedding problem is that of embeddings by *linear* symplectic maps $\varphi \in \text{Sp}(\mathbb{R}^{2n})$. The linear non-squeezing theorem states that a ball $B(r)$ can be embedded into a cylinder $Z(R)$ over a symplectic disc by a linear symplectomorphism if and only if $r \leq R$. Moreover, this non-squeezing property allows one to distinguish the symplectic maps among all linear ones in the following sense. A linear (orientation-preserving) map φ is a symplectomorphism iff the existence of a symplectic embedding $\varphi : B(r) \to Z(R)$ implies $r \leq R$.

Accordingly, one can define the *linear capacity* as a function on open sets in $\mathbb{R}^{2n}$ satisfying the same axioms A1-A3 above, where the allowed maps φ are *linear symplectomorphisms.* In particular, an example of a linear symplectic capacity is the *linear symplectic width* of a subset U, which can be defined exactly as in (64), with the extra condition that all the maps are linear.

It turns out that for ellipsoids the usual (non-linear) capacity coincides with their linear symplectic width. To see this, recall first a well-known linear algebra result stating that a positive definite quadratic form $q(z)$ on $(\mathbb{R}^{2n}, \omega)$ can be diagonalized by a linear symplectic map $\varphi \in \mathrm{Sp}(\mathbb{R}^{2n})$. Geometrically this means the φ puts the ellipsoid $E(q) = \{z \in \mathbb{R}^{2n} : q(z) < 1\}$ corresponding to the quadratic form $q(z)$ into the standard form $\varphi(E(q)) = \{z \in \mathbb{C}^n \simeq \mathbb{R}^{2n} : \sum_{i=1}^n \frac{|z_i|^2}{r_i}\}$, where $r = (r_1, \ldots, r_n)$, $0 < r_1 \leq \ldots, \leq r_n$, is the linear symplectic invariant of $E(q)$ called the *linear spectrum.* One can show that the linear width of an ellipsoid is determined by the smallest element of the spectrum, $w_L(E) = \pi r_1^2(E)$. The monotonicity of the usual (non-linear) symplectic capacity applied to the symplectic embeddings

$$B^{2n}(r_1) \subset \varphi(E) \subset Z^{2n}(r_1).$$

implies that $c(E) = \pi r_1^2 = w_L(E)$.

The linear symplectic invariant $r(E)$ can be interpreted from the dynamical point of view: each $r_i(E)$ is the action of certain periodic orbit on the boundary of E. To be more precise, let $E = E(q)$ be an ellipsoid in $\mathbb{R}^{2n}$ corresponding to the standard positive definite quadratic form q with the linear spectrum $r = (r_1, \ldots, r_n)$. Let $\partial E = \{z : q(z) = 1\}$ be the boundary of E, which is a compact hypersurface. Let $X_q(x)$ be the (linear) Hamiltonian vector field of q. Let $\mathbb{R}^{2n} = V_1 \oplus V_2 \oplus \cdots \oplus V_n$, where V_i has a symplectic basis (x_i, y_i), be a decomposition of $\mathbb{R}^{2n}$ into a direct sum of symplectic planes. In each intersection $\partial E \cap V_i$ there is the distinguished periodic orbit of X_q given by $w_i(t) = (0, \ldots, z_i(t), \ldots, 0)$, where $z_i(t) = e^{\frac{2}{r_i^2}tJ} r_i$. The period of this solution is $T_i = \pi r_i^2$. The action is given by

$$A(w_j) = \frac{1}{2} \int_0^T \langle -J\dot{w}_j, w_j \rangle = \pi r_j^2.$$

In particular, the invariant $c(E) = w_L(E) = w_G(E) = \pi r_1^2$ is represented by (the action of) a periodic orbit on ∂E. Moreover,

$$\pi r_1(E)^2 = \min\{|A(z)| \ : \ z \text{ is a periodic orbit on } \partial E\}.$$

Thus, the linear symplectic capacity is the action of a minimizing periodic orbit on the boundary of an ellipsoid. As we will see in Section 8.4.2, a similar statement can be made in the case of Hofer-Zehnder capacity of a compact connected convex domain.

8.3 Rigidity Results and Symplectic homeomorphisms

In this section we will assume that capacities exists, and will consider the rigidity properties of symplectic maps. First, in the linear case, the invariance of capacities of ellipsoids with respect to a linear operator is equivalent to the invariance of the symplectic form:

Lemma 43 *Let $L \in \mathrm{End}(\mathbb{R}^2)$ be a linear orientation-preserving map on a symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$. Then L preserves the capacities of ellipsoids iff L is symplectic, i.e., $L^*\omega_0 = \omega_0$.*

The proof is by contradiction: if we assume that L preserves the capacities of ellipsoids but not the symplectic form, by composing L with specific linear symplectic maps on the left and on the right, one can construct an embedding of the unit ball $B(1)$ into an arbitrarily thin cylinder $Z(\varepsilon)$, which is a contradiction.

To use this lemma in establishing the corresponding nonlinear result (Theorem 45), we need to show that the capacity of convex sets is continuous with respect to the Hausdorff topology on sets. This is the content of the following

Lemma 44 *Let U be a compact convex subset of $\mathbb{R}^{2n}$. Then*

$$\lim_{\substack{V \to U \\ V \text{ is convex}}} c(V) = c(U),$$

where the limit is taken over compact convex sets.

For the proof, recall that the Hausdorff distance $d(U, V)$ between two closed subsets in $\mathbb{R}^{2n}$ is defined by

$$d(U, V) = \max_{x \in U}(\min_{y \in V} ||x - y||) + \max_{y \in V}(\min_{x \in U} ||x - y||).$$

Let U be a compact convex set containing the origin. Then one can show that for any $\varepsilon > 0$ there is a $\delta > 0$ such that for any compact convex subset V satisfying $d(U, V) < \delta$, we have

$$(1-\varepsilon)U \subset V \subset (1+\varepsilon)U.$$

By the monotonicity, this implies that for $d(U, V) < \delta$, the capacity satisfies

$$(1-\varepsilon)c(U) < c(V) < (1+\varepsilon)U^2,$$

which establishes the continuity of the restriction of c to compact convex sets with respect to Hausdorff metric.

Combining the last two lemmas, one can prove the following theorem, which characterizes the symplectomorphisms as capacity-preserving orientation-preserving diffeomorphisms:

Theorem 45 *(Ekeland-Hofer [EH89]) Let $\varphi \in \mathrm{Diff}^+(\mathbb{R}^{2n})$ be a orientation-preserving diffeomorphism. Then φ is symplectic iff it preserves the capacity of all open subsets of $\mathbb{R}^{2n}$.*

The following corollary is immediate:

Corollary 46 *(Eliashberg [Eli87], Gromov [Gro85], Ekeland-Hofer [EH89]). The group of symplectomorphisms $\mathrm{Symp}(M)$ is C^0-closed in the group of all diffeomorphisms.*

The last theorem and corollary were originally proved by very difficult methods, using the non-embedding results and pseudoholomorphic curves. On the other hand, the proof of these results using the existence of capacities is almost immediate.

One of the application of the last corollary is the notion of *symplectic homeomorphisms*. Recall that it is quite easy to define volume-preserving homeomorphism: a diffeomorphism preserves the volume iff it preserves the corresponding measure; thus, one can define volume-preserving homeomorphisms as simply the homeomorphisms preserving the measure. However, the definition of a symplectomorphism involves the differential of φ and can not be straightforwardly generalized to the continuous category. However, Eliashberg and Ekeland-Hofer showed, one can use capacities to give a definition of symplectic homeomorphism as the one that preserves the capacities of open sets. Thus, each capacity function gives rise to a subgroup of homeomorphisms which preserve this capacity for all open sets, and which is closed in the group of all homeomorphisms with respect to the C^0-topology. The elements of this subgroup which happen to be smooth and orientation-preserving are automatically symplectic. This group can, therefore, be viewed as a topological version of the group of symplectomorphisms.

8.4 The Hofer-Zehnder capacity

8.4.1 Definition and verification of axioms.

In this section, we will discuss the Hofer-Zehnder capacity function c_0, related to the properties of the periodic orbits of Hamiltonian flows on a symplectic manifold.

Definition 47 *Let $\mathcal{H}(M,\omega)$ be a special class of compactly supported Hamiltonians such that*

1. There is a compact set $K \subset M$ such that $K \subset M \setminus \partial M$ and
$$H(M \setminus K) = m(H) = \text{const}.$$
2. There is an open set $U \subset M$ such that $H(U) = 0$.
3. $0 \leq H(x) \leq m(H)$ for all $x \in M$.

Note that the sets $K$and U depend on H. The number $m(H) = \max(H) - \min(H)$ is called the oscillation of H.

Definition 48 *A function $H \in \mathcal{H}(M,\omega)$ will be called* admissible *if all the periodic solutions of $\dot{z} = X_H(z)$ are either constant, or have a period $T > 1$. Denote by $\mathcal{H}_a(M,\omega)$ the set of all admissible Hamiltonians.*

Note that by rescaling one can make any function in $\mathcal{H}(M,\omega)$ to be an admissible one: for any $H \in \mathcal{H}(M,\omega)$, there is an $\varepsilon > 0$ such that $\varepsilon \cdot H \in \mathcal{H}_a(M,\omega)$. (Roughly speaking, if we make the oscillation of H sufficiently small, the corresponding vector field X_H will become sufficiently slow, so that the period of any periodic orbit will be longer than 1).

Definition 49 *([HZ90]) The Hofer-Zehnder capacity is defined by*

$$c_0(M,\omega) = \sup\{m(H) \mid H \in \mathcal{H}_a(M,\omega)\}.$$

In other words, if $c_0(M,\omega)$ is finite, it is characterized by the property that the Hamiltonian vector field X_H of any function H with an oscillation $m(H) > c_0(M,\omega)$ has a non-constant periodic orbit which is "fast" (the period $T \in (0,1]$), and $c_0(M,\omega)$ is the infimum of the numbers with this property.

Theorem 50 *(Hofer-Zehnder [HZ90]) The function c_0 is a symplectic capacity (i.e., satisfies axioms A1-A3).*

The monotonicity follows easily from the definition. Let $\varphi : (M_1,\omega_1) \to (M_2,\omega_2)$ be a symplectic embedding. Then for any Hamiltonian $H_1 \in C^\infty(M_1)$ one can define $H_2 \in C^\infty(M_2)$ by putting $H_2(\varphi(x)) = H_1(x)$ on $\varphi(M_1) \subset M_2$ and $H_2(y) = 0$ on $M_2 \setminus \varphi(M_1)$. Then φ intertwines the flows of X_{H_1} and X_{H_2}, thus giving a correspondence of periodic orbits of these vector fields on M_1 and M_2. This easily implies A1.

The conformality follows from the simple observation that $\mathcal{H}_a(M,\lambda\omega) = \{\lambda H \ : \ H \in \mathcal{H}_a(M,\omega)\}$.

To prove the normalization axiom, we will first shows that $c_0(B(1)) \geq \pi$. To do so, for each $\varepsilon > 0$ we will construct a function $H \in \mathcal{H}_a(B(1))$ with oscillation $m(H) = \pi - \varepsilon$. Taking the limit $\varepsilon \to 0$, we will conclude that $c_0(B(1)) \geq \lim_{\varepsilon\to 0} m(H) = \pi$.

Define $H(z) = f(|z|^2)$ for $z = x + iy \in B(1)$, where $f : [0,1] \to \mathbb{R}$ is a smooth non-decreasing function, which is 0 near 0, has a value of $\pi - \varepsilon$ near 1, and whose derivative satisfies $0 < f'(r) < \pi$. Then $H \in \mathcal{H}(B(1))$ and $m(H) = \pi - \varepsilon$. The solutions of the corresponding Hamiltonian equations are all periodic, and the condition on the derivative of f implies that the period the period is greater than 1, $T = \pi/|f'(r)| > 1$, thus proving that $H \in \mathcal{H}_a(M,\omega)$.

Next, the monotonicity axiom applied to the standard embedding $B(1) \to Z(1)$ together with the estimate $c_0(B(1)) \geq \pi$ implies that $c_0(Z(1)) \geq \pi$. Thus, it only remains to prove that $c_0(Z(1)) \leq \pi$. This is by far the hardest part of the proof. By definition of c_0, the inequality $c_0(Z(1)) \leq \pi$ is equivalent to the following

Theorem 51 *Let $H \in \mathcal{H}(Z(1))$ with $\sup(H) > \pi$. Then the hamiltonian flow of H has a nonconstant periodic orbit of period 1. (In particular, $H \notin \mathcal{H}_a(Z(1))$).*

Outline of the Proof. First, by replacing H with $H \circ \psi$ (where $\psi \in \mathrm{Symp}(Z(1))$ is a compactly supported symplectomorphism of the cylinder), one can assume that H vanishes in a neighborhood of the origin.

The key in the proof is the following

VARIATIONAL PRINCIPLE: *the 1-periodic solutions of the Hamiltonian equation $\dot{z} = X_{\bar{H}}(z)$ corresponding to $\bar{H} \in C^\infty(M)$ are the critical points of the functional*

$$\Phi(z) = \int_0^1 \left(\frac{1}{2}\langle -J\dot{z},\ z\rangle - \bar{H}(z(t)) \right) dt, \tag{65}$$

which is defined on the loop-space $C^\infty(S^1, \mathbb{R}^{2n})$.

STEP 1. QUADRATIC EXTENSION OF A HAMILTONIAN FUNCTION $H \in \mathcal{H}(Z(1))$.

Let $H \in \mathcal{H}(Z(1))$. Let K be the compact set such that $H(M \backslash K) = m(H)$ and $U \subset K$ be an open set such that $H(U) = 0$, as in the definition of $\mathcal{H}(Z(1))$. To apply the above principle, we first need to extend $H \in \mathcal{H}(Z(1))$ to a Hamiltonian $\bar{H}$ on $\mathbb{R}^{2n}$.

Since H is constant near the boundary of $Z(1)$, such an extension can be easily found. However, we will need the extension of a special form: namely, we will find $\bar{H}$ which has *quadratic growth,* i.e., $||d^2H(z)|| \leq c\ \forall z \in \mathbb{R}^{2n}$ for some constant $c > 0$. Such an extension is convenient because in the region of quadratic growth of the Hamiltonian the corresponding Hamiltonian vector field has linear growth, the Hamiltonian equation splits, and can be explicitly solved. More precisely, one can construct $\bar{H} : \mathbb{R}^{2n} \to \mathbb{R}$ such that

1. $\exists$ a ball $B(R)$ of radius R such that $K \subset B(R)$ and $\bar{H}(z) = H(z)$ for $z \in Z(1) \cap B(R)$.

2. $\bar{H}$ has quadratic growth.

First, let $\varepsilon \in (0, \pi/2)$ be such that $m(H) > \pi + \varepsilon$. Choose a function $f : [0, \infty) \to \mathbb{R}$ such that

$$\begin{aligned} f(s) &= m(H), & s &\in [0,1] \\ f(s) &= (\pi+\varepsilon)s, & s &\in [m(H), \infty) \\ f(s) &\geq (\pi+\varepsilon)s, & & \\ 0 \leq f'(s) &\leq (\pi+\varepsilon). & & \end{aligned}$$

Let $R > 1$ be such that $K \subset B(R)$. Choose a function $g : [0, \infty) \to \mathbb{R}$ such that

$$\begin{gathered} g(s) = 0, \quad s \in [0, R^2] \\ g(s) = \pi s/2, \quad s \geq 3R^2 \\ 0 \leq g'(s) < \pi. \end{gathered}$$

Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ and introduce complex coordinate $z = (z_1, w) \in \mathbb{C}^n$, where $z_1 \in \mathbb{C}$ and $w = (z_2, \dots, z_n) \in \mathbb{C}^{n-1}$. The extension now is given by

$$\bar{H}(z) = \begin{cases} H(z), & z \in Z(1) \cap B(R), \\ f(|z_1|^2) + g(|w|^2), & z \notin Z(1) \cap B(R). \end{cases}$$

STEP 2. 1-PERIODIC SOLUTIONS OF $\dot{z} = X_H(z)$ WITH $\Phi(z) > 0$.

Lemma 52 *Let $z(t)$ be a 1-periodic solution of the Hamiltonian equation $\dot{z} = X_{\bar{H}}(z)$, satisfying $\Phi(z) > 0$. Then $z(t)$ is non-constant, and lies in K.*

For a constant solution $z(t) = z_0 = \text{const}$, the action is $\Phi(z(t)) = -\bar{H}(z_0) < 0$, since $\bar{H} > 0$. Thus, a solution satisfying $\Phi(z) > 0$ must be nonconstant.

Now we will prove that for every 1-periodic solution (starting) outside of K we have $\Phi(z) \leq 0$. Assume that $z(t)$ is a 1-periodic solution with $\Phi(z) > 0$, and $z(t_0) \in \mathbb{R}^{2n} \setminus K$ for some time $t = t_0$. Then $H(z(t_0)) > m(H)$, and, since H is invariant under the Hamiltonian flow, the same is true for all t. Thus $z(t)$ lies in the region where $\bar{H}(z) = f(|z_1|^2) + g(|w|^2)$. Hence, the solution $z(t) = (z_1(t), w(t))$ is given by

$$z_1(t) = e^{-2if'(|z_1|^2)} z_1(0), \quad w(t) = e^{-2ig'(|w|^2)} w(0).$$

Since $2g' \in [0, 2\pi)$, the function $w(t)$ can not have period $T = 1$ unless either $w(t) = 0$ or $g' = 0$. In either of these cases $w = \text{const}$. This implies that $|z_1| > 1$ and

$$0 < 2f'(|z_1|^2) \leq 2\pi + 2\varepsilon.$$

Since $z_1(t) = z_1(t+1)$, it follows that $2f'(|z_1|^2) = 2\pi$. Thus, we can estimate $\Phi(z(t))$ as follows:

$$\begin{aligned} \Phi(z) &= \int_0^1 \frac{1}{2}(y_1 \dot{x}_1 - x_1 \dot{y}_1) - f(|z_1^2|) \\ &= f(|z_1|^2)|z_1|^2 - f(|z_1|^2) = -\pi |z_1|^2 - f(|z_1|^2) \leq -\varepsilon |z_1|^2. \end{aligned}$$

Hence, if there is a 1-periodic solution of $\dot{z} = X_{\bar{H}}(z)$ with positive action, it must lie inside of K and thus, be a solution of the original Hamiltonian system $\dot{z} = X_H(z)$. It only remains to prove the existence of a 1-periodic solution $z(t)$ with $\Phi(z(t)) > 0$.

STEP 3. THE ANALYTICAL SETTING

In order to apply the minimax techniques to prove the existence of a periodic orbit with prescribed properties, we first extend the functional $\Phi : C^\infty(S^1, \mathbb{R}^{2n}) \to \mathbb{R}$ to an appropriate Hilbert space. First, represent the periodic loops in $C^\infty(S^1, \mathbb{R}^{2n})$ by their Fourier series

$$z(t) = \sum_{k \in \mathbb{Z}} e^{k \cdot 2\pi J t} z_k, \quad z_k \in \mathbb{R}^{2n}, \tag{66}$$

which converge together with all the derivatives in the sup norm. The first term (which corresponds to the symplectic structure) of the action functional (60) can be written as $a(z, z)$, where

$$\begin{aligned} a(z, w) &= \int_0^1 \frac{1}{2} \langle -J\dot{z}, w \rangle dt = \pi \sum_{k \in \mathbb{Z}} k \langle z_k, w_k \rangle = \\ &= \pi \sum_{k>0} |k| \langle z_k, w_k \rangle - \pi \sum_{k<0} |k| \langle z_k, w_k \rangle. \end{aligned}$$

Thus, $a(z, w)$ can be extended to a continuous bilinear form on the Sobolev space $H^{1/2}(S^1)$, defined by

$$H^{1/2}(S^1) = \left\{ z \in L^2(S^1) \mid \sum_{k \in \mathbb{Z}} |k|^{2 \cdot \frac{1}{2}} \cdot |z_k|^2 < \infty \right\},$$

where $z(t)$ is given by the Fourier expansion as in (66). The inner product on the Hilbert space $H^{1/2}(S^1)$ is defined by

$$\langle z, z' \rangle = \langle z_0, z'_0 \rangle + 2\pi \sum_{k \in \mathbb{Z}} |k| \cdot \langle z_k, z'_k \rangle.$$

The Hilbert space $E = H^{1/2}(S^1)$ has an orthogonal splitting $E = E^- \oplus E^0 \oplus E^+$ into the spaces of elements having only Fourier coefficients with $k < 0$, $k = 0$ and $k > 0$. The corresponding orthogonal projections are P^-, P^0 and P^+, and $z = z^- + z^0 + z^+$ is the unique decomposition of an element $z \in E$.

The quadratic growth of $H(z)$ guarantees that the second part of the functional $\Phi(z)$, given by

$$b(z) = \int_0^1 H(z(t))dt,$$

can be extended to $z \in L^2(S^1)$, and, in particular, to $H^{1/2}(S^1) \subset L^2(S^1)$. Moreover, this function is differentiable, and $b \in C^\infty(H^{1/2}, \mathbb{R})$ provided $H \in C^\infty(\mathbb{R}^{2n}, \mathbb{R})$. (This follows from Sobolev's embedding theorem, which in our case states the existence of a compact embedding $H^{1/2}(S^1) \to L^2(S^1)$, and gives the estimates on the norms, $||z||_{L^2} \leq C \cdot ||z||_{H^{1/2}}$ for some constant C). Moreover,

$$\nabla b(x) = j^* \nabla H(x),$$

where $j : E \to L^2$ is the natural inclusion map.

Thus, the extended action functional $\Phi : E \to \mathbb{R}$ given by $\Phi(z) = a(z) - b(z)$ is differentiable, and one can show that its gradient is given by $\nabla\Phi(z) = z^+ - z^- - \nabla b(z)$.

STEP 4. REGULARITY.

The following regularity statement shows that all the critical points of the extended functional are actually smooth periodic solutions of the Hamiltonian equation.

Lemma 53 *Let $z \in E$ be a critical point $\nabla\Phi(z) = 0$. Then $z \in C^\infty(S^1)$ and, moreover, is a 1-periodic solution of the Hamiltonian equation $\dot{z}(t) = X_H(z(t))$.*

To prove this lemma, one first represents $z(t)$ and $H(z(t))$ by their Fourier expansions in L^2 and writes down the condition $\nabla\Phi(z) = 0$ in terms of the Fourier coefficients, which leads to

$$\langle (P^+ - P^-)z,\ v\rangle = \int_0^1 \langle \nabla H(x), v\rangle dt, \qquad \forall v \in E.$$

Choosing the test function $v(t) = e^{k\cdot 2\pi Jt}$, one finds $2\pi k z_k = (\nabla H(z))_k$ for all k, which allows to conclude that $z \in H^1(S^1)$, another Sobolev space. In particular, $z(t) \in C^1(S^1, \mathbb{R}^{2n})$. Hence, $\nabla H(z(t)) \in C^1(S^1, \mathbb{R}^{2n})$. Then the

Hamiltonian equation implies that $z \in C^2(S^1, \mathbb{R}^{2n})$. Iterating this argument, one concludes that $z \in C^\infty(S^1)$.

STEP 5. THE PALAIS-SMALE CONDITION AND THE MINIMAX LEMMA.

When searching for critical points of a differentiable function $\Phi : E \to \mathbb{R}$ on a Hilbert space E, it is sometimes convenient to take the dynamical point of view and interpret the critical points of Φ as the equilibrium points of the gradient equation $\dot{z} = -\nabla\Phi(z)$. Assume that this differential equation has a global flow $\varphi^t(z)$, i.e., the corresponding Cauchy problem can be solved for all times $t \in \mathbb{R}$. The crucial property of the gradient flow is that $\Phi(\varphi^t(z))$ decreases along non-constant solutions. More precisely,

$$\Phi(\varphi^t(z)) - \Phi(z) = -\int_0^t ||\nabla\Phi(\varphi^s(z))||^2 ds. \qquad (67)$$

Since E is not compact, there is no guarantee that Φ will have any critical points unless we impose some conditions on Φ. Palais and Smale [PS64] gave the following strong compactness condition which guarantees a critical point in many variational problems:

Definition 54 *A C^1-function $\Phi : E \to \mathbb{R}$ satisfies the* Palais-Smale condition *(PS condition) if every sequence $x_i \in E$ satisfying*

1. *$\nabla\Phi(z_i) \to 0$ in E;*

2. *$|\Phi(z_i)| \le c < \infty$ for some $c \in \mathbb{R}$;*

possesses a convergent subsequence. The limit of this subsequence is then a critical point of Φ.

In particular, if $\mathcal{F}$ is a family of subsets $F \subset E$ one defines the *minimax* $c(\Phi, \mathcal{F})$ of Φ on this family by

$$c(\Phi, \mathcal{F}) = \inf_{F\in\mathcal{F}} \sup_{x\in F} \Phi(x) \quad \in \mathbb{R} \cup \{\pm\infty\}. \qquad (68)$$

Lemma 55 (Minimax Lemma) *Let $\Phi \in C^1(E)$ and $\mathcal{F}$ is a family of subsets such that*

1. *Φ satisfies the PS condition.*

2. *$\dot{z} = -\nabla\Phi(z)$ defines a global flow $\varphi^t(z)$.*

3. *The family $\mathcal{F}$ is positively invariant under the flow, i.e., if $F \in \mathcal{F}$, then $\varphi^t(F) \in \mathcal{F}$ for every $t \geq 0$.*

4. $c(\Phi, \mathcal{F}) \in \mathbb{R}$ *(i.e., is finite).*

Then $c(\Phi, \mathcal{F})$ is a critical value of Φ.

We will apply this lemma in the proof of existence of 1-periodic orbits of X_H with positive action. First, we have to make sure that all the conditions of the Minimax Lemma are satisfied.

The fact that Φ satisfies the PS condition follows from the assumption that $\bar{H}$ growth quadratically with such a speed that there are no nontrivial 1-periodic solutions in the region where $\bar{H}$ is quadratic. (See [HZ87] for details of the proof). The theory of ordinary differential equations implies that the gradient equation $\dot{z}(t) = -\nabla\Phi(z(t))$ defines a unique global flow. Thus, the first two conditions in the Minimax Lemma are satisfied.

STEP 6. THE LINKING ARGUMENT AND THE EXISTENCE OF A CRITICAL POINT.

Now we have to find a family of subsets $\mathcal{F}$ which is positively invariant, and such that the corresponding minimax $c(\Phi, \mathcal{F})$ is finite. In order to do it, the following linking argument is used.

For $\tau > 0$, let

$$\Sigma_\tau = \{z \mid z = z^- + z^0 + sz^+, \ ||z^- + z^0| \leq \tau \quad \text{and} \quad 0 \leq s \leq \tau\}. \tag{69}$$

Using the assumptions that $H \in [0, m(H)]$ as well as the asymptotic behavior of $\bar{H}$, one can show that there exists τ_0 such that for all $\tau \geq \tau_0$ we have

$$\Phi|_{\partial\Sigma} \leq 0.$$

On the other hand, for $\alpha > 0$ define $\Gamma_\alpha = \{z \in E^+ \mid ||z|| = \alpha\}$. Using the assumption that H is identically zero near the origin, one can show that there exists $\alpha > 0$ such that $\Phi(\Gamma) \geq \beta > 0$.

Let $\varphi^t(z)$ be the gradient flow of the equation $\dot{z} = -\nabla\Phi(z)$. Since $\Phi(\varphi^t(z))$ decreases with t, it follows that $\Phi(\varphi^t(\partial\Sigma) \leq 0$ for all $t \geq 0$. Since $\Phi(\Gamma) \geq \beta > 0$, it follows that $\varphi^t(\partial\Sigma) \cap \Gamma = 0$. Using the generalization of the Brouwer mapping degree to infinite dimensional spaces (called the Leray-Schauder degree), one can show that this implies that

$$\varphi^t(\Sigma) \cap \Gamma \neq 0, \quad \forall t \geq 0. \tag{70}$$

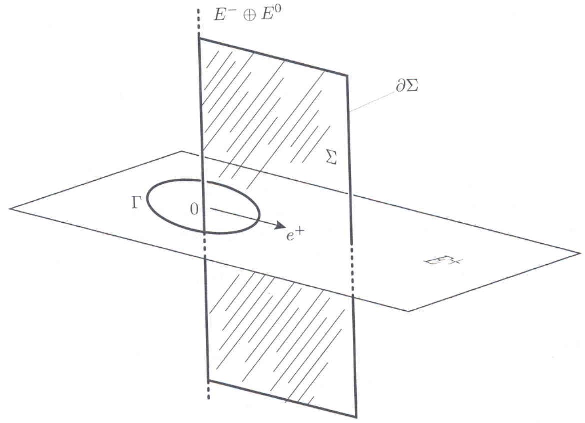

Figure 1: Linking argument [HZ94, p. 95]

The following figure helps to visualize the subsets in question:

Define now the family of subsets $\mathcal{F} = \{\Phi = \varphi^t(\Sigma) \mid t \geq 0\}$, and let

$$c(\Phi, \mathcal{F}) = \inf_{t \geq 0} \sup_{z \in \varphi^t(\Sigma)} \Phi(z)$$

be the corresponding minimax of Φ on this family. The family $\mathcal{F}$ is clearly positive-invariant. The property (70) implies that $\inf_{z \in \Gamma} \Phi(z) \leq \sup_{z \in \varphi^t(\Sigma)} \Phi(z)$. Since $\Phi(\Gamma) \geq \beta > 0$, it follows that $c(\Phi, \mathcal{F}) > -\infty$. Since Φ maps bounded sets to bounded sets, it follows that $\sup_{z \in \varphi^t(\Sigma)} \Phi(z) < \infty$, and thus, $c(\Phi, \mathcal{F}) < \infty$. Thus, the minimax lemma implies that $c(\Phi, \mathcal{F})$ is the required critical value of $\Phi(z)$, and there exists z_0 such that $\Phi(z_0) = c(\Phi, \mathcal{F}) \geq \beta > 0$. By regularity, $z_0 \in C^\infty(S^1, \mathbb{R}^{2n})$, and geometrically represents a 1-periodic solution of the Hamiltonian equation $\dot{z} = -\nabla H(z)$ with $\Phi(z(t)) > 0$. This completes the proof of the inequality $c_0(Z(1)) \leq \pi$, and thus establishes that the Hofer-Zehnder function is indeed a symplectic capacity.

8.4.2 Capacity of strictly convex domains

As we have seen above, the capacity of ellipsoids agrees with the action of a periodic solution of its boundary. More generally, in this section we outline the proof of the fact that the Hofer-Zehnder capacity function of a convex bounded domain can be represented by a distinguished periodic orbit on its boundary. First, recall that in this setting a periodic orbit always exists:

Theorem 56 *(Rabinowitz [Rab78], Weinstein [Wei78]) Let $S \in (\mathbb{R}^{2n}, \omega)$ be a smooth (of class C^2) boundary of a compact strictly convex region in $\mathbb{R}^{2n}$. Then there exists a periodic orbit on S.*

Hofer and Zehnder proved the following

Theorem 57 *Let $S = \partial C$ be the smooth compact boundary of a convex region $C \subset \mathbb{R}^{2n}$. Let $z_0(t)$ be the periodic Hamiltonian trajectory such that*

$$A(z_0(t)) = \inf\{|A(z(t))| \ : \ z(t) \in S \ \ \textit{is a period orbit}\},$$

where $A(z(t) = \frac{1}{2}\int_0^T \langle -J\dot{z}, z\rangle dt$ is the reduced action. Then $c_0(C) = A(z_0(t))$.

Outline of the proof. Assume that C is a strictly convex containing the origin. Let $r(x) \geq 0$ be a function which is smooth away from 0 and differentiable at least once at 0, and such that $C = \{z \in \mathbb{R}^{2n} \mid |r(z) \leq 1\}$. Let $\alpha = \inf\{|A(z)|, \ z$ is a periodic orbit on $S\}$. To prove that $c_0(C) \geq \alpha$ one constructs an admissible Hamiltonian $H \in \mathcal{H}_a(C)$ with the oscillation $m(H) = \alpha - \varepsilon$. (This part of the proof is very similar to the proof of the inequality $c_0(B(1)) \geq \pi$, as in Section 8.4).

To prove that $c_0(C) \leq \alpha$, it is enough to take $H \in \mathcal{H}(C_0)$ with $m(H) > \alpha$ and prove the existence of a 1-periodic non-constant solution of the Hamiltonian equation in C. This construction is a modification of the variational method used in the proof of $c_0(Z(1)) \leq \pi$.

To go from a strictly convex region to a convex region C, one approximates C with strictly convex regions and uses natural estimates on the action.

References

[EH89] I. Ekeland and H. Hofer, *Symplectic topology and Hamiltonian dynamics*, Math. Z. **200** (1989), no. 3, 355–378.

[Eli87] Ya. M. Eliashberg, *A theorem on the structure of wave fronts and its application in symplectic topology*, Funktsional. Anal. i Prilozhen. **21** (1987), no. 3, 65–72, 96.

[Gro85] M. Gromov, *Pseudoholomorphic curves in symplectic manifolds*, Invent. Math. **82** (1985), no. 2, 307–347.

[HZ87] H. Hofer and E. Zehnder, *Periodic solutions on hypersurfaces and a result by C. Viterbo*, Invent. Math. **90** (1987), no. 1, 1–9.

[HZ90] H. Hofer and E. Zehnder, *A new capacity for symplectic manifolds*, Analysis, et cetera, Academic Press, Boston, MA, 1990, pp. 405–427.

[HZ94] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1994.

[PS64] R. S. Palais and S. Smale, *A generalized Morse theory*, Bull. Amer. Math. Soc. **70** (1964), 165–172.

[Rab78] P. H. Rabinowitz, *Periodic solutions of Hamiltonian systems*, Comm. Pure Appl. Math. **31** (1978), no. 2, 157–184.

[Sib93] K. Siburg, *Symplectic capacities in two dimensions*, Manuscripta Math. **78** (1993), no. 2, 149–163.

[Vit87] C. Viterbo, *A proof of Weinstein's conjecture in* $\mathbf{R}^{2n}$, Ann. Inst. H. Poincaré Anal. Non Linéaire **4** (1987), no. 4, 337–356.

[Wei78] A. Weinstein, *Periodic orbits for convex Hamiltonian systems*, Ann. of Math. (2) **108** (1978), no. 3, 507–518.

[Wei79] A. Weinstein, *On the hypotheses of Rabinowitz' periodic orbit theorems*, J. Differential Equations **33** (1979), no. 3, 353–358.

9 A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the mKdV equation

after P. Deift and X. Zhou
A summary written by Svetlana Roudenko

Abstract

We give a summary of the results in [1], in particular, we discuss the deformation techniques for the RHP developed in Sections 3 and 4.

9.1 Introduction

The authors in [1] are analyzing the long time behavior of the mKdV equation

$$y_t - 6y^2 y_x + y_{xxx} = 0, \qquad -\infty < x < \infty, \quad t \geq 0$$

$$y(x,0) = y_0(x)$$

by the methods of the inverse scattering (see [2] and the summary in [4]) which leads to studying the asymptotics of the associated Riemann-Hilbert problem (RHP), which we denote by $(m, v_{x,t}, \mathbb{R})$. For each x and t, find a 2×2 matrix-valued function $m(z; x, t)$ such that

- $m(z)$ is analytic for $z \in \mathbb{C} \setminus \mathbb{R}$; (analyticity condition)
- $m_+(z) = m_-(z)\, v_{x,t}(z), \quad z \in \mathbb{R}$; (jump condition)
- $m(z) \to I$ as $z \to \infty$, (normalization condition)

here

$$m_\pm(z) = \lim_{\epsilon \searrow 0} m(z \pm i\epsilon; x, t)$$

and

$$v_{x,t}(z) = e^{-i(4tz^3 + xz)\sigma}\, v(z)\, e^{i(4tz^3+xz)\sigma}, \quad \sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with

$$v(z) = \begin{pmatrix} 1 - |r(z)|^2 & -\overline{r(z)} \\ r(z) & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\bar{r} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix} = (b_-^{-1})\, b_+.$$

Since $y_0 \in \mathcal{S}$, we have $r \in \mathcal{S}$, and $r(z) = -\overline{r(-z)}$, $\sup_{z\in\mathbb{R}} |r(z)| < 1$.

Inversely, given the jump matrix v, solve the RHP $(m, v_{x,t}, \mathbb{R})$ for each x and t, then consider the associated quantity $\mu(z; x, t) = m_+(z; x, t)\ (b_+^{-1})_{x,t} = m_-(z; x, t)\,(b_-^{-1})_{x,t}$ and the integral equation

$$\mu = I + C\,\mu, \qquad \text{or} \quad \mu = (1_{\mathbb{R}} - C)^{-1} I, \tag{71}$$

where C is the Cauchy operator on $\mathbb{R}$ (see next section). Then (see [2]) the solution of the inverse problem is given by

$$y(x, t) = \left(\left[\sigma, \int_{\mathbb{R}} \mu(z; x, t)\, w_{x,t}(z)\, \frac{dz}{2\pi i} \right] \right)_{21}, \tag{72}$$

where $w_{x,t} = (w_+)_{x,t} + (w_-)_{x,t}$ and $w_\pm = \pm(b_\pm - I)$.

Therefore, to compute the long-time asymptotics of $y(x, t)$, (72) is analyzed with μ given implicitly by (71).

We consider the case when $x < 0$ and decompose the complex plane into the regions of the sign of $\operatorname{Re} i(4tz^3 + xz)$. In particular, we concentrate on the physically interesting region, called region II in [1],

$$z_0 \le M \quad \text{for a constant} \quad M > 1, \quad \text{and} \quad \frac{-x}{t^{1/3}} \to \infty.$$

Here, $z_0 = \sqrt{\frac{-x}{12t}}$ is the stationary phase point of $4tz^3 + xz$ for $x < 0$.

The main result is the following:

Theorem 58 *Let $y_0(x) \in \mathcal{S}$ (and so the reflection coefficient $r(z) \in \mathcal{S}$). The solution $y(x, t)$ of mKdV with the initial data $y_0(x)$ has uniform leading asymptotics as $t \to \infty$ in the region II:*

$$y(x, t) = y_a + \frac{c(z_0)}{(tz_0)^{1/2}}\, O\left(\frac{1}{(tz_0^3)^{1/2}} + \frac{\log(tz_0)}{(tz_0)^{1/2}} \right) \tag{73}$$

where

$$y_a = (\nu/(3tz_0))^{1/2} \cos(16\tau - \nu \log(192\tau) + \varphi(z_0)),$$

$$\varphi(z_0) = \arg \Gamma(i\nu) - \frac{\pi}{4} - \arg r(z_0) + \frac{1}{\pi} \int_{-z_0}^{z_0} \log|s - z_0|\, d(\log(1 - |r(s)|^2)),$$

$\nu = -\frac{1}{2\pi} \log(1 - |r(z_0)|^2) > 0$, *$\tau = tz_0^3$ and Γ is the standard gamma function.*

9.2 Outline of the proof steps

The above RHP is being studied by deforming contours, extracting the leading terms from such deformations, and therefore, producing a "simplified" RHP on a deformed contour. Each step peels off asymptotic terms from the solution $y(x,t)$ in (72).

In what follows, $C_\pm$ is the Cauchy operator defined on $L^2(\Gamma)$ with Γ being an oriented contour in $\mathbb{C}$:

$$(C_\pm f)(z) = \int_\Gamma \frac{f(\xi)}{\xi - z_\pm} \frac{d\xi}{2\pi i}, \qquad z \in \Gamma,$$

where "+"("-") is to the left (right) of Γ following its orientation. The operators $C_\pm$ are bounded on $L^2(\Gamma)$ and $C_+ - C_- = 1$. Set

$$C_w f = C_+(f w_-) + C_-(f w_+),$$

and note that $w \in L^1(\Gamma) \cap L^\infty(\Gamma)$ implies $C_w : L^2(\Gamma) + L^\infty(\Gamma) \to L^2(\Gamma)$. If the contour is important, we use the superscript Γ: C_w^Γ.

First, we change the RHP $(m, v_{x,t}, \mathbb{R})$ from the real line to the contour $\Sigma = L \cup \bar{L} \cup \mathbb{R}$, where

$$L = \{z = z_0 + z_0 u e^{i3\pi/4} : -\infty < u \le \sqrt{2}\} \cup$$
$$\{z = -z_0 + z_0 u e^{i\pi/4} : -\infty < u \le \sqrt{2}\}.$$

Let δ be the solution of the scalar RHP $(\delta, v = (1 - |r(z)|^2)\chi_{\{|z|<z_0\}} + \chi_{\{|z|>z_0\}}, \mathbb{R})$. The jump matrix $v_{x,t}$ after conjugation with δ^σ, i.e. $\delta_-^\sigma v_{x,t} \delta_+^{-\sigma}$, decomposes into factors, off-diagonal terms of which have the following property: each term can be split into an analytic part and a small non-analytic remainder decaying to high order as $t \to \infty$. For instance, when $|z| < z_0$ the off-diagonal term (in the upper factorization) $-\overline{r(z)}(1 - |r(z)|^2)^{-1}$ being split into $h_1 + h_2 + R$ where R is a polynomial in z, h_1 decays to high order as $t \to \infty$ and h_2 has analytic continuation into the upper half plane. Thus, the RHP $(m, v_{x,t}, \mathbb{R})$ is being extended to the $(m^\#, v_{x,t}^\#, \Sigma)$. The coefficients of the new jump matrix $v^\#$ depend on h_1 for $z \in \mathbb{R}$ and on both R and h_2 for $z \in \Sigma \setminus \mathbb{R}$.

Second, the RHP $(m^\#, v^\#, \Sigma)$ is reduced to a RHP (m', v', Σ') on a pair of crosses $\Sigma' = \Sigma_{A'} \cup \Sigma_{B'}$ localized around $\pm z_0$ (i.e. $\Sigma' = L|_{\{-\infty<u<\epsilon\}} \cup \overline{L|_{\{-\infty<u<\epsilon\}}}$), since the jump matrix $v^\#$ converges rapidly to the identity as

$t \to \infty$ on $\mathbb{R}$ and on the compact part of $\Sigma \backslash \mathbb{R}$ away from $\pm z_0$ (on $\Sigma \backslash (\mathbb{R} \cup \Sigma')$).

The main result in this step is the following proposition:

Lemma 59 *The solution y(x,t) in (72) takes the form*

$$y(x,t) = \left(\int_{\Sigma'} \left[\sigma, \left((1_{\Sigma'} - C_{w'}^{\Sigma'})^{-1} I \right) (\xi)\, w'(\xi) \right] \frac{d\xi}{2\pi i} \right)_{21} + E(x,t), \tag{74}$$

where the controlled error term $E(x,t) = O(z_0 \tau^{-l}; |x|^{-l})$, $l \in \mathbb{N}$.

(Here, w' is the corresponding quantity to $w_{x,t}$ in (72) changed thru the extension to Σ and restriction to Σ'.)

Third, the interaction between $\Sigma_{A'}$ and $\Sigma_{B'}$ is observed to vanish to high order as $t \to \infty$ and so the contribution from Σ' in (74) can be approximated by the sum of contributions from $\Sigma_{A'}$ and $\Sigma_{B'}$. See next section for details.

The last step is to extend and rescale each cross $\Sigma_{A'}$ and $\Sigma_{B'}$ to the RHP on the simple cross at the origin with jump matrices independent of time and solve the factorization problems in terms of parabolic cylindric functions (see [3]). Refer to the last section for details.

9.3 Splitting of the two crosses (Section 3)

Write $\Sigma' = \Sigma_{A'} \cup \Sigma_{B'}$, where $\Sigma_{A'}$ is the cross centered at $-z_0$ and $\Sigma_{B'}$ centered at z_0. Define $w^{A'} = w'$ for $z \in \Sigma_{A'}$ and 0 on $\Sigma_{B'}$, similarly define $w^{B'}$. Then $w' = w^{A'} + w^{B'}$. Set $A' = C_{w^{A'}}^{\Sigma'}$ and similarly B'. Observe that $C_{w'}^{\Sigma'} = A' + B'$.

The main tool in this section is the existence and uniform boundedness of the operators $(1_{\Sigma'} - A')^{-1}$ and $(1_{\Sigma'} - B')^{-1}$ on $L^2(\Sigma')$ which is obtained in step-by-step deformation of contour Σ' to $\mathbb{R}$ and successive estimates from the associated RHPs.

Using the boundedness of the above operators, we write

$$\int_{\Sigma'} \left[(1_{\Sigma'} - C_{w'}^{\Sigma'})^{-1} I \right] (\xi)\, w'(\xi)\, d\xi = \int_{\Sigma'} \left[(1_{\Sigma'} + A'(1_{\Sigma'} - A')^{-1} \right.$$

$$\left. + B'(1_{\Sigma'} - B')^{-1}) I \right] (\xi)\, w'(\xi)\, d\xi + O\left(\frac{c(z_0)}{\tau^{1/2} (t z_0)^{1/2}} \right)$$

as $t \to \infty$. Separating w' into $w^{A'}$ and $w^{B'}$ and using their L^1 and L^2 bounds, we obtain

$$\left| \int_{\Sigma'} \left[A'(1_{\Sigma'} - A')^{-1} \right] (\xi) w^{B'}(\xi)\, d\xi \right| \leq \frac{c(z_0)}{\tau^{1/2}(tz_0)^{1/2}}$$

and the similar estimate with B' and $w^{A'}$.

Now restrict the operator $C^{\Sigma'}_{w^{A'}}$ defined on Σ' to $C^{A'}_{w^{A'}} = C^{\Sigma_{A'}}_{w^{A'}}$ defined only on the left cross $\Sigma_{A'}$ (with $w^{A'}|_{\Sigma_{A'}}$ instead of $w^{A'}$ on all Σ').

Using the restriction identity $(1_{\Sigma_{A'}} - C^{A'}_{w^{A'}})^{-1} I = [(1_{\Sigma'} - A')^{-1} I]\,|_{\Sigma_{A'}}$ (and analogous one for B'), we decompose the solution $y(x,t)$ from the formula (74) into the following terms:

Lemma 60

$$y(x,t) = \left(\int_{\Sigma_{A'}} \left[\sigma, \left((1_{\Sigma_{A'}} - C^{A'}_{w^{A'}})^{-1} I \right) (\xi)\, w^{A'}(\xi) \right] \frac{d\xi}{2\pi i} \right)_{21}$$

$$+ \left(\int_{\Sigma_{B'}} \left[\sigma, \left((1_{\Sigma_{B'}} - C^{B'}_{w^{B'}})^{-1} I \right) (\xi)\, w^{B'}(\xi) \right] \frac{d\xi}{2\pi i} \right)_{21} \qquad (75)$$

$$+ O(z_0 \tau^{-l}; |x|^{-l}) + O\left(\frac{c(z_0)}{\tau^{1/2}(tz_0)^{1/2}} \right),$$

as $\tau \to \infty$.

9.4 Reduction to the real line (Section 4)

We further analyze the solution $y(x,t)$ from (75), in particular, we are interested in estimating

$$\int_{\Sigma_{A'}} \left[(1_{\Sigma_{A'}} - C^{A'}_{w^{A'}})^{-1} I \right] (\xi)\, w^{A'}(\xi)\, d\xi.$$

Extending Σ'_A to the full cross $\hat{\Sigma}_{A'}$ and then shifting it to the cross at the origin Σ_A, we rewrite the above integral as

$$(48tz_0)^{-1/2} U_A \left\{ \int_{\Sigma_A} \left((1_{\Sigma_A} - A)^{-1} I \right) (\xi)\, w^A(\xi)\, d\xi \right\} U_A^{-1},$$

where U_A is a unitary operator on $L^2(\Sigma_A)$ and $A = C^{\Sigma_A}_{w^A}$. Let w^{A^0} be the leading term in w^A for large t, then denoting by A^0 the Cauchy operator associated to w^{A^0}, we rewrite the integral in curly brackets as

$$\int_{\Sigma_A} [(1_A - A^0)^{-1} I](\xi) w^{A^0}(\xi)\, d\xi + O\left(c(z_0) \left[\frac{1}{\tau^{1/2}} + \frac{\log(tz_0)}{(tz_0)^{1/2}} \right] \right), \qquad (76)$$

for $tz_0 \geq 2$.

Now we again turn back to the RHP principle: if $m^{A^0} = I + C_{w^{A^0}}(1_A - A^0)^{-1} I$, then m^{A^0} solves the RHP $(m^{A^0}, v^{A^0}, \Sigma_A)$, where $v^{A^0} = (I - w_-^{A^0})^{-1}(I + w_+^{A^0})$. The series representation of $m^{A^0}(z)$ gives

$$m^{A^0}(z) = I - \frac{m_1^{A^0}}{z} + O(z^{-2}), \quad \text{as} \quad z \to \infty,$$

and thus, $m_1^{A^0}$ is nothing else but the integral in (76). Similar steps can be repeated for Σ_B and the RHP is $(m^{B^0}, v^{B^0}, \Sigma_B)$ in that case. Using the definitions of w^{A^0} and w^{B^0}, we connect the jump matrices v^{A^0} and v^{B^0}, as well as m^{A^0} and m^{B^0}: $m^{A^0}(z) = \sigma \overline{m^{B^0}(-\bar{z})} \sigma$; and expanding both sides, we obtain $m_1^{A^0} = -\sigma \overline{m_1^{B^0}(-\bar{z})} \sigma$, so the commutators relate as $[\sigma, m_1^{A^0}] = \overline{[\sigma, m_1^{B^0}]}$. Therefore, the solution (75) simplifies to

$$y(x,t) = (48tz_0)^{-1/2} \left((\delta_B^0)^{-2} (-2(m_1^{B^0})_{21} + \overline{(\delta_B^0)^{-2}(-2(m_1^{B^0})_{21}} \right)$$

$$+ O(z_0 \tau^{-l}; |x|^{-l}) + O\left(\frac{c(z_0)\, \log(tz_0)}{tz_0} \right), \quad \text{for} \quad tz_0 \geq 2, \quad \text{as} \quad \tau \to \infty. \qquad (77)$$

So what left to compute is $(m_1^{B^0})_{21}$, which is done by considering the RHP with the jump across the real line and reducing it to the parabolic-cylindric equation (ref. [3])

$$g''(\xi) + (1/2 - \xi^2/4 + a) g(\xi) = 0,$$

solution of which is the linear combination of the standard parabolic-cylindric function $D_a(\xi)$. Putting everything back together into (77), we obtain the leading asymptotic solution $y(x,t)$ in (73).

References

[1] Deift, P. and Zhou, X. *A steepest descent method for oscillatory Riemann Hilbert problems. Asymptotics for the mkdv equation,* Annals of Math. 137 (1993), no. 2, 295-368.

[2] Beals, R. and Coifman, R. *Scattering and inverse scattering for first order systems,* Comm. Pure Appl. Math. 37 (1984), 39-90.

[3] Whittaker, E.T. and Watson, G.N. *A course of Modern Analysis,* Cambridge University Press, Cambridge, 1963.

[4] Crmaric, T. Summary of *"A steepest descent method for oscillatory Riemann Hilbert problems. Asymptotics for the mKdV equation",* part 1.

10 On Nekhoroshev's Estimate at an Elliptic Equilibrium

after Jürgen Pöschel
A summary written by Eric Ryckman

10.1 Overview

In his 1977 paper, Nekhoroshev proves that under a perturbation of order ϵ of an integrable Hamiltonian H_0, the actions of an arbitrary orbit vary only of the order ϵ^b over a time interval of the order $\exp(\epsilon^{-a})$, where a and b are positive *stability exponents* that depend only on the number of degrees of freedom of the Hamiltonian and the steepness of H_0 (for convex Hamiltonians like we will consider, the optimal values of a and b seem to be $a = b = \frac{1}{2n}$ where n is the dimension).

While Nekhoroshev proved his result only for Hamiltonians in action-angle coordinates, he conjectured that it should also hold in neighborhoods of elliptic equilibria. Guzzo, Fassò & Benettin and Niederman simultaneously and independently showed that the conjecture was correct in the case of convex Hamiltonians. Unfortunately their proofs were rather technical and so not easily accessible.

In this paper, Pöschel presents an easier proof that is essentially an extension of Lochak's method [2] for determining the stability exponents. Basically, the analysis is done in neighborhoods of periodic orbits of the unperturbed Hamiltonian, then all other initial positions are approximated by periodic ones.

Consider a real analytic Hamiltonian near an elliptic equilibrium whose characteristic frequencies are nonresonant up to order $l \geq 4$ (i.e., $k_1\omega_1 + \cdots + k_n\omega_n \neq 0$ for all $k_1, \ldots, k_n \in \mathbf{Z}$ with $0 < |k_1| + \cdots + |k_n| \leq l$). It is a general fact [1] that in this case there are symplectic coordinates such that

$$H = \langle \alpha, I \rangle + \tfrac{1}{2}\langle AI, I \rangle + G(I) + \epsilon F(z, \epsilon) = H_0 + \epsilon F(z, \epsilon).$$

Here $z = (x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n)$, $I = I(z) = \big(\ldots, \frac{1}{2}(z_j^2 + z_{n+j}^2), \ldots\big)$, $F(\cdot, \epsilon)$ is real analytic (by which we mean analytic in each complex variable and real for real arguments) of order $l+1$ in z and uniformly bounded on a fixed ball around the origin, G is at least order 3 in I and absent if $l = 4$ or 5. For this Hamiltonian we prove the main result:

Theorem 61 *Suppose A is positive definite. If ϵ is sufficiently small, then for every orbit of the Hamiltonian H with $|I(0)| < 1$ one has*

$$|I(t) - I(0)| < c\epsilon^a \quad \textit{for} \quad |t| < \frac{1}{|\alpha|}\exp(d\epsilon^{-a})$$

with $a = \frac{1}{2n}$. The constants c and d depend only on A and the dimension, n.

So the functions I (which one can think of as action coordinates, even though they were not introduced as such) vary only of the order of ϵ^a over a time interval of the order $\exp(\epsilon^{-a})$. In other words, Nekhoroshev's conjecture holds in this case. Note also that Theorem 61 yields the same stability exponents as the usual Nekhoroshev estimates for convex Hamiltonians in action-angle coordinates.

The proof of Theorem 61 proceeds in three steps: normal form, local stability estimates, and global stability estimates.

10.2 Normal Form

We begin by finding symplectic coordinates in a neighborhood of the origin so that the perturbed Hamiltonian is the sum of functions that we can estimate. We begin with some notation. Let z_0 be an initial position of a real solution of H_0 and define $I_0 = I(z_0)$. The motion is quasi-periodic with frequencies

$$\omega_0 = \omega(I_0) = \frac{\partial H_0}{\partial I}(I_0).$$

Say ω_0 is T-periodic and gives rise to a nontrivial T-periodic solution, where we say a frequency vector is *T-periodic* if

$$T\omega_0 \in 2\pi\mathbf{Z}^n, \quad T\omega_0 \neq 0,$$

for some real T. The set of all real initial positions z with these frequencies ω_0 forms a torus

$$\mathcal{T}(I_0) = \{z : I(z) = I_0\} \subseteq \mathbf{R}^n \times \mathbf{R}^n.$$

We want to obtain a normal form for H in a neighborhood of $\mathcal{T}(I_0)$.

Let $J = I - I_0$. Up to an irrelevant additive constant we can write

$$H_0 = \langle \omega_0, I \rangle + g(I)$$

where

$$g(I) = \frac{1}{2}\langle AJ, J\rangle + G(I) - G(I_0) - dG_{I_0}(J).$$

(That this is correct is easily checked using the relation $\omega_0 = \frac{\partial H_0}{\partial I}(I_0)$ and the original form of H_0.) The total Hamiltonian then becomes

$$H = h(I) + g(I) + f(z)$$

where $h(I) = \langle \omega_0, I\rangle$ and $f(z) = \epsilon F(z, \epsilon)$.

We'll study this Hamiltonian on the complex domains

$$D_{r,s} = \{z : |I(z) - I_0| < r, \quad \|z\| < s\} \subseteq \mathbf{R}^n \times \mathbf{R}^n.$$

Let $|\cdot|_{r,s}$ denote the sup-norm on $D_{r,s}$ and X_h the vector field of the Hamiltonian h.

Finally, we'll write $a \ll b$ if $ca < b$ for some constant $c \geq 1$, and similarly $a \lesssim b$ if $a \leq cb$.

The result we'll use is

Lemma 62 (Normal Form Lemma) *Consider $H = h+g+f$ as described above. If*

$$mT\epsilon \ll r, \quad mTr \ll 1$$

with an integer $m \geq 1$ and $0 < r \ll 1$, then there exists a real analytic, symplectic transformation $\Psi : D_{2r,2} \to D_{3r,3}$ with $|\Psi - \mathrm{id}|_{2r,2} \lesssim T\epsilon$, such that

$$H \circ \Psi = h + g + \hat{g} + \hat{f}$$

with $\{h, \hat{g}\} = 0$ and

$$|\hat{g}|_{2r,2} \lesssim \epsilon, \quad |\hat{f}|_{2r,2} \lesssim 2^{-m}\epsilon, \quad |X_{\hat{f}}|_{2r,2} < 2^{-m}\epsilon.$$

Here, the implicit constants are in fact independent of m and n.

We won't actually prove Lemma 62, but will instead prove the following (which also applies and yields the same results).

Lemma 63 *Consider a real analytic Hamiltonian $H = h + f + g$ on $D_{3r,3}$, where h and g are integrable, the flow of $h = \langle \omega, I\rangle$ is T-periodic, and*

$$|X_g|_{3r,3} < \delta, \quad |X_f|_{3r,3} < \epsilon.$$

If

$$\delta \lesssim r, \quad mT\epsilon \ll r, \quad mTr \ll 1$$

with some integer $m \geq 1$ *and* $0 < r \ll 1$, *then there exists a real analytic, symplectic transformation* $\Psi : D_{2r,2} \to D_{3r,3}$ *with* $|\Psi - \mathrm{id}|_{2r,2} \lesssim T\epsilon$, *such that* $H \circ \Psi = h + \tilde{g} + \hat{f}$ *with*

$$|X_{\tilde{g}} - X_g|_{2r,2} < 2\epsilon, \quad |X_{\hat{f}}| < 2^{-m}\epsilon,$$

and $\{h, \tilde{g}\} = 0$.

To prove Lemma 63 we iterate an averaging transformation m times. The general step is given by

Lemma 64 (Iterative Lemma) *Suppose the Hamiltonian* $H = h + g + f$ *is real analytic on* $D_{r,s}$, *where the flow of* $h = \langle \omega, I \rangle$ *is T-periodic and*

$$|X_g - X_{g_0}|_{r,s} < \gamma, \quad |X_{g_0}|_{r,s} < \delta, \quad |X_f|_{r,s} < \epsilon$$

with an integrable Hamiltonian g_0. *If*

$$T\epsilon \ll \rho \ll \sigma$$

with $0 < \rho < r$ *and* $0 < \sigma < s$, *then there exists a real analytic, symplectic transformation* $\Phi : D_{r-\rho,s-\sigma} \to D_{r,s}$, *such that* $H \circ \Phi = h + g_+ + f_+$ *with*

$$|X_{g_+} - X_g|_{r,s} < \epsilon, \quad |X_{f_+}|_{r-\rho,s-\sigma} \lesssim T\Big(\frac{\delta}{\sigma} + \frac{\gamma + \epsilon}{\rho}\Big)\epsilon,$$

and $\{h, g_+ - g\} = 0$. *Moreover,* $|\Phi - \mathrm{id}|_{r-\rho,s-\sigma} < T\epsilon$.

The desired map Φ will be the time-1 map of the Hamiltonian vector field X_ϕ where ϕ solves an appropriate homological equation involving h and f. The various inequalities follow from integral representations of the Hamiltonian vector fields and Cauchy's estimates.

10.3 Local Stability Estimates

Once the Hamiltonian is in normal form, we can prove a stability result near periodic solutions.

Lemma 65 (Local Stability Estimate) *Consider the Hamiltonian in normal form, and let z_0 be an initial position in $\|z_0\| < 1$ with a T-periodic vector $\omega_0 = \omega(I_0)$. If*

$$\epsilon \ll r^2, \quad mTr \ll 1$$

with an integer $m \geq 1$ and $0 < r \ll 1$, then for every initial position with amplitudes $I(0)$ satisfying $|I(0) - I_0| \ll r$ one has

$$|I(t) - I_0| < r \quad \textit{for} \quad |t| < \frac{2^m}{|\omega_0|}.$$

The proof of Lemma 65 follows from the estimates obtained in Lemma 62.

10.4 Global Stability Estimates

Following [2], we use the local stability estimate to determine global stability. As mentioned earlier, we do this by approximating an arbitrary initial position by periodic orbits of the unperturbed Hamiltonian, then applying the local stability estimate around these periodic orbits. We will use Dirichlet's theorem on simultaneous approximations:

Theorem 66 *For every $\omega \in \mathbf{R}^n$ and every integer $Q \geq 1$,*

$$\min_{\substack{1 \leq q \leq Q \\ q \in \mathbf{Z}}} \min_{p \in \mathbf{Z}^n} |q\omega - p|_\infty \leq \frac{1}{Q^{1/n}}$$

where $|\cdot|_\infty$ denotes the sup-norm.

Now, let z be an arbitrary initial position with amplitudes $|I| < 1$, and let $\omega = \omega(I)$ be its frequencies with respect to H_0. Let $\epsilon > 0$. We can scale down one component of ω until it is an integer, apply Dirichlet's theorem to the remaining $n-1$ components to approximate them with a rational frequency vector, then scale back. This yields a $2\pi T$-periodic frequency vector ω_0 such that $|\omega - \omega_0|_\infty \leq \epsilon^a/T$ where $\frac{1}{2} \leq T \leq Q = \epsilon^{-a(n-1)}$ with a to be chosen. From the convexity of H_0 there is a unique amplitude I_0 corresponding to ω_0. If we set $r \simeq \epsilon^a/T$ for an appropriately chosen implicit constant, then $|I - I_0| \ll r$.

We want to apply the local stability estimate around $\mathcal{T}(I_0)$. Since $T \leq Q$ we see that the hypothesis $\epsilon \ll \epsilon^{2a}/T^2$ is satisfied if $\epsilon \lesssim \frac{\epsilon^{2a}}{Q^2} = \epsilon^{2an}$, which is possible if we choose $a = 1/2n$. If we choose $m \simeq \epsilon^{-a}$ then the hypothesis $mTr \simeq m\epsilon^a \ll 1$ is satisfied as well. So if ϵ is small enough, we can use the local stability estimate to conclude

$$|I(t) - I_0| < r \quad \text{for} \quad |t| < \frac{1}{|\omega|} \exp(d\epsilon^{-a}).$$

We then also see $|I(t) - I(0)| < r \lesssim \epsilon^a$ when $|t| < \frac{1}{|\omega|} \exp(d\epsilon^{-a})$, proving Theorem 61.

References

[1] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser, Basel, 1994.

[2] P. Lochak, Hamiltonian perturbation theory: periodic orbits, resonances and intermittency, *Nonlinearity* **6** (1993), 885–904.

[3] J. Pöschel, On Nekhoroshev's estimate at an elliptic equilibrium, *Int. Math. Res. Not.* **4** (1998), 203–215.

11 Approximation of Solutions of the Cubic Nonlinear Schrodinger Equations by Finite-Dimensional Equations and Nonsqueezing Properties

after J. Bourgain
A summary written by Monica Visan

Abstract

We prove a nonsqueezing result for certain nonlinear, periodic Schrodinger equations by reducing the problem to a finite-dimensional model where Gromov's nonsqueezing theorem holds.

We will consider the cubic periodic, one dimensional NLSE

$$iu_t + u_{xx} + u|u|^2 = 0,$$

with initial data

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

Assuming local existence of solutions established, there are many other questions to be addressed concerning their behavior as $t \to \infty$ and the properties of the flow S_t in the phase space. Some of these questions deal with global existence, blowup behavior, asymptotic stability, spreading of energy to higher modes, behavior of higher Sobolev norms of smooth solutions as $t \to \infty$.

This paper deals with nonsqueezing properties of the flow S_t, extending a result obtained previously by Kuksin. Adjusting the finite-dimensional theory to an infinite dimensional phase space setting, Kuksin proved that

$$S_t(B_r) \subset \mathbb{T}_R^{(k)} \;\Leftrightarrow\; r \leq R,$$

provided that $S_t =$ linear operator+ compact smooth operator. Here, B_r is an r-ball in the symplectic Hilbert space $L^2(\mathbb{T}^d)$ and $\mathbb{T}_R^{(k)}$ stands for a translate of the cylinder $\{\sum(p_j\phi_j^+ + q_j\phi_j^-)/p_k^2 + q_k^2 < R^2\}$, with $\{\phi_j^\pm\}$ a Darboux basis of $L^2(\mathbb{T}^d)$.

The example we will deal with does not fall under the category of examples treated by Kuksin in the sense that the flow map S_t is not of the special form linear operator + compact smooth operator. The nonsqueezing result we'll prove is going to be a consequence of the following:

Proposition 67 *Consider the solutions u, v to the Cauchy problems*

$$iu_t + u_{xx} + u|u|^2 = 0,\ \ u(x,0) = \phi(x)$$

and

$$iv_t + v_{xx} + P_N(v|v|^2) = 0,\ \ v(x,0) = \phi(x),$$

where $\phi = P_N\phi$ (P_N is the Dirichlet projection with respect to the x-variable, i.e. $P_N\phi(x) = \sum_{|n|\leq N} \hat{\phi}(n)e^{inx}$).

Fix a positive integer N' and a time t. Then one has the approximation

$$\|P_{N'}(u(t) - v(t))\|_2 < \epsilon,$$

provided $N > N(N', |t|, \epsilon, \|\phi\|_2)$.

Recalling the finite-dimensional result, $S_N(t)(B_r) \subset \mathbb{T}_R^{(k)} \Leftrightarrow r \leq R$, one gets the nonsqueezing statement

Proposition 68

$$S(t)(B_r) \subset \mathbb{T}_R^{(k)} \ \ \Leftrightarrow\ \ r \leq R.$$

Now, Proposition 1 will be obtained as a consequence of the following two lemmas

Lemma 69 *Consider the solutions u, v to the Cauchy problems*

$$iu_t + u_{xx} + u|u|^2 = 0,\ \ u(x,0) = \phi(x)$$

and

$$iv_t + v_{xx} + v|v|^2 = 0,\ \ v(x,0) = \psi(x),$$

with $\|\phi\|_2 = \|\psi\|_2$. Then for $|t| < T(\|\phi\|_2)$ one has

$$\|P_{N_0}(u(t) - v(t))\|_2 < \|P_{N_1}(\phi - \psi)\|_2 + \epsilon,$$

provided $N_1 - N_0 > c\epsilon^{-c}$.

Lemma 70 *Consider the solutions u, v to the Cauchy problems*

$$iu_t + u_{xx} + u|u|^2 = 0,\ \ u(x,0) = \phi(x)$$

and

$$iv_t + v_{xx} + P_N(v|v|^2) = 0,\ \ v(x,0) = \phi(x),$$

where $\phi = P_N\phi$. Then, for $|t| < T(\|\phi\|_2)$ one has

$$\|P_{N_0}(u(t) - v(t))\|_2 < \epsilon,$$

provided $N - N_0 > c\epsilon^{-c}$.

The first hour I will present the general layout of the paper and a sketch of the proof of the wellposedness result. The second hour I will discuss the proof of the two lemmas and, if time allows, the main ingredient in controlling the nonlinearity, i.e. the estimate

$$\|f\|_{L^4(\mathbb{T}^2)} \leq c[\sum_{m,n\in\mathbb{Z}} (1+|n-m^2|)^{\frac{3}{4}} |\hat{f}(m,n)|^2]^{\frac{1}{2}}.$$

In the nonperiodic case, the nonlinearity is controlled in the iteration process by Strichartz's inequality:

$$\|U(\cdot)\psi\|_{L^q(dxdt)} \leq C\|\psi\|_{L^2(dx)},$$

for $q = \frac{2(n+2)}{n}$. In the periodic case, this type of inequalities may hold only locally in time (global solutions are not dispersive). In fact, the analogue of Strichartz's inequality for $L^q(dxdt)$ replaced by $L^q(\mathbb{T}^{n+1})$ fails. For example, when $n = 1$ and $q = 6$, one has

$$\|\sum_{|n|\leq N} a_n e^{2\pi i(nx+n^2t)}\|_{L^6(\mathbb{T}^2)} << N^\epsilon (\sum |a_n|^2)^{\frac{1}{2}},$$

for any $\epsilon > 0$, the presence of the N^ϵ-factor being necessary.

References

[1] J. Bourgain, Approximation of Solutions to the Cubic Nonlinear Schrodinger Equations by Finite-Dimensional Equations and Non-squeezing Properties, IMRN, 1994, No.2

[2] J. Bourgain, Fourier Transform Restriction Phenomena for Certain Lattice Subsets and Applications to Nonlinear Evolution Equations, Part I: Schrodinger Equations, Geom. Funct. Anal. 3(1993), 107-156

[3] J. Bourgain, Fourier Transform Restriction Phenomena for Certain Lattice Subsets and Applications to Nonlinear Evolution Equations, Part II: KdV Equation, Geom. Funct. Anal. 3(1993), 209-262

[4] J. Bourgain, Aspects of Long Time Behavior of Solutions of Nonlinear Hamiltonian Evolution Equations, Geom. Funct. Anal. 2(1995), 105-140

[5] J. Bourgain, Global Solutions of Nonlinear Schrodinger Equations, American Mathematical Society colloquium publications, ISSN 0065-9258, v. 46

[6] S. Kuksin, Infinite-Dimensional Symplectic Capacities and a Squeezing Theorem for Hamiltonian PDE's, ETH, August 1993

12 Integrable systems and reductions of the self-dual Yang-Mills equations

after Ablowitz, Chakravarty, and Halburd
A summary written by Brett D. Wick

Abstract

We give a summary of the results in [1]. In particular, we give a brief explanation of the self-dual Yang-Mills (SDYM) equations. It is then indicated how certain reductions of the SDYM equations yield the NLS, KdV, and Euler-Arnold-Manakov top equations. We also explain the significance of the Painlevé equations and indicate how they can be obtained from the SDYM equations through a different type of reduction.

Introduction

The Yang-Mills equations and the self-dual Yang-Mills equations are an important topic in mathematical physics and the field of integrable systems. From the point of view of integrable systems the study of SDYM equations became intriguing when R. S. Ward conjectured,

> *... many (and perhaps all?) of the ordinary and partial differential equations that are regarded as being integrable or solvable may be obtained from the self-dual gauge field equations (or its generalizations) by reduction.*

One of the main goals of the article [1] is to demonstrate how one can derive well-known examples of integrable equations by using the SDYM equations. Because of this, the SDYM equations are often referred to as the "master integrable system".

We first explain the Yang-Mills equations and the SDYM equations. Then we demonstrate how, by using certain symmetries in the variables, it is possible to reduce the SDYM equations to some well-known equations. Finally, we indicate how, using certain conformal symmetries, the SDYM equations reduce to the Painlevé equations.

The self-dual Yang-Mills equations

We will be working with a four-dimensional manifold M and the coordinates will be given by $\{x^\mu\}$, with $\mu = 0, \dots, 3$.

Let G be a Lie group acting on the manifold M, and let LG denote the associated Lie algebra. Then we take $A_\mu(x) \in \text{LG}$ and introduce covariant derivatives

$$D_\mu := \partial_\mu - A_\mu,$$

and their commutators

$$F_{\mu\nu} := -[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - [A_\mu, A_\nu],$$

with the first equality being the definition of $F_{\mu\nu}$ and the second equality following from a computation using the product rule.

To the functions A_μ and $F_{\mu\nu}$ we associate differential forms. The one-form $A := A_\mu dx^\mu$ is called the *connection one-form*. The two-form $F := \frac{1}{2}F_{\mu\nu}dx^\mu dx^\nu$ is called the *curvature two-form of the connection*. Here we have used the Einstein summation convention of summing over a repeated upper and lower index.

The differential forms F and A are related through an exterior covariant derivative. If we define the operator

$$D_A X := dX - A \wedge X,$$

then computation using the properties of differential forms will show that

$$F = D_A A = dA - A \wedge A.$$

The Hodge star operator, $*$, acts on differential forms. On a four-manifold $*$ acts on two-forms in the following manner: it takes $T = \frac{1}{2}T_{\mu\nu}dx^\mu dx^\nu$ to $*T = \frac{1}{2}\varepsilon_{\mu\nu}^{\gamma\delta}T_{\gamma\delta}dx^\mu dx^\nu$. Here $\varepsilon_{\mu\nu\gamma\delta}$ is the totally anti-symmetric tensor with $\varepsilon_{0123} = 1$, and the standard metric on M is used to raise and lower the index. The Yang-Mills equations are:

$$D_A(*F) = 0. \tag{YM}$$

We also have the Bianchi identity on the manifold M. This states

$$D_A F = 0. \tag{BI}$$

Now we introduce the SDYM equations. In general the YM equations are a set of coupled, second-order PDE's and are very difficult to solve, but by making a clever reduction we can look for a special class of solutions. First, observe that any F which satisfies $*F = \lambda F$ (λ a constant) also satisfies the YM equation by virtue of the Bianchi identity. Indeed,

$$D_A(*F) = D_A(\lambda F) = \lambda D_A F = 0.$$

The equations $*F = (-)F$ are called the (anti-)self-dual Yang-Mills equations. Simple computation gives that the SDYM equations are

$$F_{01} = F_{23}, \quad F_{02} = F_{31}, \quad F_{03} = F_{12}. \qquad \text{(SDYM)}$$

Introducing the null (or complex) coordinates

$$\begin{aligned} \sigma &:= \tfrac{1}{\sqrt{2}}(x^1 + ix^2),\ \tau := \tfrac{1}{\sqrt{2}}(x^0 - ix^3), \\ \tilde{\sigma} &:= \tfrac{1}{\sqrt{2}}(x^1 - ix^2),\ \tilde{\tau} := \tfrac{1}{\sqrt{2}}(x^0 + ix^3), \end{aligned}$$

we can write the connection one-form $A = A_\mu dx^\mu$ in terms of these null coordinates. It follows that

$$\begin{aligned} A_0 &= \tfrac{1}{\sqrt{2}}(A_\tau + A_{\tilde{\tau}}),\ A_1 = \tfrac{1}{\sqrt{2}}(A_\sigma + A_{\tilde{\sigma}}), \\ A_2 &= \tfrac{i}{\sqrt{2}}(A_\sigma - A_{\tilde{\sigma}}),\ A_3 = -\tfrac{i}{\sqrt{2}}(A_\tau - A_{\tilde{\tau}}). \end{aligned}$$

In these new coordinates the SDYM equations become

$$F_{\sigma\tau} = 0, \quad F_{\tilde{\sigma}\tilde{\tau}} = 0, \quad F_{\sigma\tilde{\sigma}} + F_{\tau\tilde{\tau}} = 0. \qquad \text{(SDYM)}$$

This will be the form of the SDYM equations we will use. These equations are also the compatibility conditions of the following isospectral linear problem:

$$\begin{aligned} (\partial_\sigma + \zeta\partial_{\tilde{\tau}})\Psi &= (A_\sigma + \zeta A_{\tilde{\tau}})\Psi, \\ (\partial_\tau - \zeta\partial_{\tilde{\sigma}})\Psi &= (A_\tau - \zeta A_{\tilde{\sigma}})\Psi. \end{aligned}$$

Reduction of SDYM equations to KdV, NLS, and Euler-Arnold-Manakov top

Some of the simplest reductions of the SDYM equations are found when the A_μ's are taken to be independent of certain coordinates. To show how the SDYM equations reduce to each of these above equations, reductions will be used which take advantage of certain translational symmetries.

KdV and NLS Equations

To obtain this reduction we take LG $= \mathfrak{sl}(2;\mathbf{C})$ and the A_μ to be functions of $x = \sigma + \tilde{\sigma}$ and $t = \tilde{\tau}$ only. By using the gauge freedom it is possible to take $A_\sigma = 0$.

If we consider the matrix-valued functions $P := A_\tau$, $Q := A_{\tilde{\sigma}}$ and $R := A_{\tilde{\tau}}$, then the SDYM equations become:

$$P_x = 0, \quad Q_x - P_t - [P,R] = 0, \quad R_x - Q_t - [Q,R] = 0.$$

Notice these equations are invariant if P, Q, R all undergo the same constant similarity transformation. Therefore, if P is independent of t it can be put into canonical form. The case of P being diagonalizable leads to the NLS and the case of P being non-diagonalizable leads to the KdV.

Assuming P is diagonalizable we have

$$P = \begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix},$$

and using the second SDYM equation from above we see that we can take Q to have the form

$$Q = \begin{pmatrix} 0 & q \\ -r & 0 \end{pmatrix}.$$

Finally, the last equation gives that R must have the form

$$R = \frac{1}{2k}\begin{pmatrix} qr & q_x \\ r_x & -qr \end{pmatrix},$$

and this then gives the equations

$$\begin{aligned} 2kq_t &= q_{xx} + 2q^2 r, \\ 2kr_t &= -r_{xx} - 2qr^2. \end{aligned}$$

Choosing $k = \frac{i}{2}$ and $r = \pm\overline{q}$ gives the nonlinear Schrödinger equation

$$iq_t = q_{xx} \pm 2|q|^2 q. \tag{NLS}$$

In the case when P is non-diagonalizable, we take it of the form

$$P = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We then find that Q and R have the form

$$Q = \begin{pmatrix} v & 1 \\ w & -v \end{pmatrix}, \quad R = \frac{1}{8}\begin{pmatrix} 4w_x & -8v_x \\ v_{xxx} - 4vv_{xx} - 2v_x^2 + 4v^2 v_x & -4w_x \end{pmatrix}$$

where $w = v_x - v^2$, $u = -v_x$ and u satisfies the Korteweg-de Vries equation

$$u_t = \frac{1}{4}u_{xxx} + 3uu_x. \tag{KdV}$$

Euler-Arnold-Manakov Top Equations

Here we take the functions A_μ to depend only on the null coordinate $t = \sigma$ and we choose G $= \mathrm{SL}(n; \mathbf{C})$ and LG $= \mathfrak{sl}(n; \mathbf{C})$. Then the vanishing of $F_{\tilde{\sigma}\tilde{\tau}}$ will force $A_{\tilde{\sigma}}$ and $A_{\tilde{\tau}}$ to commute. We choose these matrices to be constant diagonal matrices,

$$A_{\tilde{\sigma}} = \mathrm{diag}(a_1, \ldots, a_n), \quad A_{\tilde{\tau}} = \mathrm{diag}(b_1, \ldots, b_n).$$

The equation $F_{\sigma\tau} + F_{\tilde{\sigma}\tilde{\tau}} = 0$ then becomes the algebraic equation $[A_\sigma, A_{\tilde{\sigma}}] + [A_\tau, A_{\tilde{\tau}}] = 0$. If we call $A_\sigma = (A_{ij})$ and $A_\tau = (B_{ij})$, then computation gives (when $a_i \neq a_j$ and $b_i \neq b_j$)

$$A_{ij} = -\frac{b_j - b_i}{a_j - a_i} B_{ij}, \quad i \neq j.$$

Choosing $a_i = b_i^2$ and taking A_σ and A_τ to be skew-symmetric and using the last of the SDYM equations, $F_{\sigma\tau} = 0$, we find

$$\frac{dB_{ij}}{dt} = \sum_{k=1}^{n} \left(\frac{1}{b_j + b_k} - \frac{1}{b_k + b_i} \right) B_{ik} B_{kj}. \tag{EAM}$$

Reduction of SDYM to Painlevé equations

The Painlevé equations are important for several reasons. They arise from similarity reductions of classical soliton equations and they are also the monodromy-preserving deformation equations associated to a linear system of ODE's with rational coefficients. They were first discovered as the class of second order ODE's with the property that the only singularities which depend on the initial conditions are poles. See [2]. The Painlevé equations are the following six ODE's:

$$P_I: \; u'' = 6u^2 + t,$$
$$P_{II}: \; u'' = 2u^3 + tu + \alpha,$$
$$P_{III}: \; u'' = \frac{1}{u}u'^2 - \frac{1}{t}u' + \frac{1}{t}(\alpha u^2 + \beta) + \gamma u^3 + \frac{\delta}{u},$$

$$P_{IV}: \; u'' = \frac{1}{2u}u'^2 + \frac{3}{2}u^3 + 4tu^2 + 2(t^2 - \alpha)u + \frac{\beta}{u},$$
$$P_V: \; u'' = \left(\frac{1}{2u} + \frac{1}{u-1}\right)u'^2 - \frac{1}{t}u' + \frac{(u-1)^2}{t^2}\left(\alpha + \frac{\beta}{u}\right) + \frac{\gamma u}{t} + \frac{\delta u(u+1)}{u-1},$$
$$P_{VI}: \; u'' = \frac{1}{2}\left(\frac{1}{u} + \frac{1}{u-1} + \frac{1}{u-t}\right)u'^2 - \left(\frac{1}{t} + \frac{1}{u-1} + \frac{1}{u-t}\right)u' + \frac{u(u-1)(u-t)}{t^2(t-1)^2}\left(\alpha + \frac{\beta t}{u^2} + \frac{\gamma(t-1)}{(u-1)^2} + \frac{\delta t(t-1)}{(u-t)^2}\right).$$

The first five equations can be derived from the sixth by a process of coalescence (making changes of variables that depend on certain parameters and letting these parameters tend to zero).

To derive the Painlevé equations from the SDYM equations we will use conformal symmetries. We take LG $= \mathfrak{sl}(2;\mathbf{C})$ as the Lie algebra. The SDYM equations are invariant under the group of conformal transformations (those which preserve the metric up to a factor). Studying these conformal symmetries and looking at the generators of these conformal transformations (called conformal Killing vectors) gives rise to six four-parameter subgroups of GL$(4;\mathbf{C})$ called the Painlevé groups.

It can be shown that the SDYM equations associated to a Painlevé group gives rise to the associated Painlevé equation. In particular, certain linear combinations of the conformal Killing vectors (found from the structure of the Painlevé group) give rise to new vectors $\tilde{X}_j$ $(j = 1, 2, 3)$. One then finds new variables w^j such that $\tilde{X}_j(w^i) = \delta^i_j$. Writing the connection one-form A in terms of the new variables w^j it is then possible to discover the values of A_σ, A_τ, $A_{\tilde{\sigma}}$ and $A_{\tilde{\tau}}$, which in turn gives the SDYM equations. Solutions of the Painlevé equations are then found by taking the solution of a certain gauge-invariant equation arising from the SDYM equations.

References

[1] Ablowitz, M. J., Chakravarty, S. and Halburd, R. G. *Integrable systems and reductions of the self-dual Yang-Mills equations* J. Math. Phys. **44** [2003], pp. 3147-3173.

[2] Ince, E. L. "Ordinary Differential Equations", Dover Publications, 1926.