Seminar on Hamilton PDE [1]

R. E. Lee DeVille[2]
Professor: C. Eugene Wayne, Boston Univerity

Last modified: June 6, 2000

[1]Notes for MA 876, Spring 2000
[2]based exclusively on notes given by C. E. Wayne and others.

Contents

Introduction

In Spring 2000, Professor Gene Wayne gave a class at Boston University titled “MA 876 - Partial Differential Equations Seminar”. Most of the lectures were given by Prof. Wayne, but others were given by several of the graduate students, including Doug Wright, Stephanie Jones, Tony Harkin, Slava Krigman, Bill Basener, and myself. The following pages are a (hopefully accurate) copy of the notes I took in this class.

This document was produced using $\mathcal{AMS}$-LaTeX macros on top of Version 3.1415 of the LaTeX2ϵ compiler. I used the amsbook documentclass. The only packages I had to import were epsfig and graphicx for the pictures, and theorem to define the various theorem environments. All of the pictures were made either with `xfig` or Mathematica, and then dumped into PostScript.

For further information (including how to get a copy of this PostScript file), please go to

`http://math.bu.edu/people/deville/Notes/`. The latest versions of this file will be maintained there when corrections are made. If you find any errors, or have any general comments on the notes, please send email to me at `deville@math.bu.edu`. I will maintain an errata page for these notes (and all future releases), and any and all comments are greatly appreciated.

Chapter 1

Hamiltonian ODE

1.1 An introduction

Most of the following introduction will be taken from [Arn], Chapters 8 and 9.

An example of a Hamiltonian system is as follows: Let

$$H(p,q) = \frac{1}{2}p^2 + V(q),$$

where p and q are scalar quantities. Then **Hamilton's Equations** define an ODE on $\mathbb{R}^2$:

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial q} = -\frac{\partial V}{\partial q} \\ \dot{q} &= \frac{\partial H}{\partial p} = p. \end{aligned}$$

We will need some differential geometry to get everything working smoothly, so we proceed:

1.2 Symplectic Geometry

Let M^{2n} be a smooth manifold of dimension $2n$. Let $\omega^{(2)}$ be a closed, non-degenerate differential 2-form on M.

To say that $\omega^{(2)}$ is a **differential 2-form** means that $\omega^{(2)}$ is a bilinear map

$$\omega^{(2)} \colon T_x M^{2n} \times T_x M^{2n} \to \mathbb{R},$$

with

$$\omega^{(2)}(\xi, \eta) = -\omega^{(2)}(\eta, \xi).$$

To say that $\omega^{(2)}$ is **closed** means that $d\omega^{(2)} = 0$.

To say that $\omega^{(2)}$ is **non-degenerate** means that given $\xi \in T_x M^{2n} \neq 0$, there is an η with

$$\omega^{(2)}(\xi, \eta) \neq 0.$$

The pair $(M^{2n}, \omega^{(2)})$ is called a **symplectic manifold**.

Example: Let $M^{2n} = \mathbb{R}^2$, and $\omega^{(2)} = dp \wedge dq$. If

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix},$$

then

$$\omega^{(2)}(\xi,\eta) = \begin{vmatrix} \xi_1 & \xi_2 \\ \eta_1 & \eta_2 \end{vmatrix} = \xi_1\eta_2 - \xi_2\eta_1.$$

We can extend this example by choosing $M^{2n} = \mathbb{R}^{2n}$ and

$$\omega^{(2)} = \sum_{j=1}^{n} dp_j \wedge dq_j.$$

We will not prove the following theorem, but we should note that it essentially tells us that all symplectic manifolds look like $\mathbb{R}^{2n}$, at least locally.

Theorem 1.1 (Darboux' Theorem)
Let $(M^{2n}, \omega^{(2)})$ be a symplectic manifold. Pick $\underline{x} \in M^{2n}$. In a neighborhood of $\underline{x}$, there are coordinates $(\underline{p}, \underline{q})$ with

$$\omega^{(2)} = \sum_{j=1}^{n} dp_j \wedge dq_j.$$

♠

Proposition 1.2
The symplectic form defines a vector space isomorphism between vector fields and 1-forms. ♠

Proof: Let ξ be a vector field. Define a 1 form

$$\omega_\xi^{(1)}(\eta) = \omega(\eta, \xi).$$

The map $\xi \mapsto \omega_\xi^{(1)}$ is obviously linear. To see that it is injective, suppose that $\omega_\xi^{(1)} = \omega_{\tilde{\xi}}^{(1)}$. Then

$$\begin{aligned} \omega_\xi^{(1)}(\eta) - \omega_{\tilde{\xi}}^{(1)}(\eta) &= 0, \quad \text{for all}\eta \\ \omega(\xi,\eta) - \omega(\tilde{\xi},\eta) &= 0 \quad \text{for all}\eta \\ \xi &= \tilde{\xi}. \end{aligned}$$

■

Example: For this example, we will again choose the symplectic manifold $(\mathbb{R}^2, dp \wedge dq)$, and let I be the map $\omega_\xi^{(1)} \mapsto \xi$. We will compute the matrix of I.

1. By the Riesz Representation Theorem, we know that

$$\omega_\xi^{(1)} = \langle \Omega, \eta \rangle.$$

2. $\omega_\xi^{(1)}(\eta) = \omega^{(2)}(\eta,\xi) = \eta_1\xi_2 - \eta_2\xi_1$.

3. One can write any 2-form as

$$\begin{aligned} \omega^{(2)}(\eta,\xi) &= \langle \eta, J\xi \rangle \\ &= \eta_1(J_{11}\xi_1 + J_{12}\xi_2) + \eta_2(J_{21}\xi_1 + J_{22}\xi_2), \end{aligned}$$

   for an anti-symmetric matrix J.

   Comparing Steps 2 and 3 gives us that

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

4.

$$\langle \eta, J\xi \rangle = \omega^{(2)}(\eta, \xi) = \omega^{(1)}_{\xi}(\eta) = \langle \Omega, \eta \rangle = \langle \eta, \Omega \rangle \,,$$

so that

$$J\xi = \Omega, \qquad \xi = J^{-1}\Omega = -J\Omega,$$

so that

$$I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

For $(\mathbb{R}^{2n}, \sum dp_j \wedge dq_j)$, the idea is the same, but the notation makes it a little more complicated. Doing the same calculation gives us

$$I = \begin{pmatrix} 0 & -\mathbf{1}_n \\ \mathbf{1}_n & 0 \end{pmatrix},$$

where $\mathbf{1}_n$ is the $n \times n$ identity matrix.

Definition 1.1
*Given a smooth function H on $(M^{2n}, \omega^{(2)})$, we defined the **Hamiltonian vector field associated with H** by*

$$X_H = I(dH),$$

i.e. that X_H is the vector field which satisfies the equation

$$dH(\eta) = \omega^{(2)}(\eta, X_H).$$

♣

Example: Pick $H = \dfrac{p^2}{2} + V(q)$ on $(\mathbb{R}^2, dq \wedge dq)$.
Then

$$dH = \frac{\partial H}{\partial p}\, dp + \frac{\partial H}{\partial q}\, dq = p\, dp + \frac{\partial V}{\partial q} dq,$$

which makes $\Omega = \begin{pmatrix} p \\ \dfrac{\partial V}{\partial q} \end{pmatrix}$.

Then

$$\begin{aligned} X_H = I(dH) &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Omega = \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p \\ \dfrac{\partial V}{\partial q} \end{pmatrix} \\ &= \begin{pmatrix} -\dfrac{\partial V}{\partial q} \\ p \end{pmatrix}. \end{aligned}$$

Definition 1.2
*For a given Hamiltonian H, **Hamilton's Equations** are just the differential equations on $(M^{2n}, \omega^{(2)})$ given by*

$$\dot{x} = X_H.$$

♣

In the above example, this gives us that

$$\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} -\frac{\partial V}{\partial q} \\ p \end{pmatrix}.$$

In the next Proposition, we will prove that H is preserved by the flow:

Proposition 1.3
Let g^t be the flow associated with H. Then

$$H(x_0) = H(g^t(x_0)).$$

♠

Proof: By definition,

$$\frac{d}{dt}g^t(x_0) = X_H(g^t(x_0)).$$

Calculating:

$$\begin{aligned} \frac{d}{dt}(H(g^t(x_0))) &= \sum_{j=1}^{2n} \frac{\partial H}{\partial x_j}\frac{d}{dt}(g_j^t(x_0)) \\ &= \sum \frac{\partial H}{\partial x_j}\left(X_{H,j}(g^t(x_0))\right), \end{aligned}$$

or

$$\frac{d}{dt}\left(H(g^t(x_0))\right) = d_{g^t(x_0)}H\left(X_H(g^t(x_0))\right).$$

Recalling that

$$dH(\eta) = \omega^{(2)}(\eta, X_H),$$

we have that

$$\frac{d}{dt}\left(H(g^t(x_0))\right) = \omega^{(2)}(X_H(g^t(x_0)), X_H(g^t(x_0))) = 0.$$

■

The following theorem will say that the Hamiltonian flow preserves the symplectic form:

Theorem 1.4
Let g^t be the flow associated with some Hamiltonian H. Then

$$\left(g^t\right)^* \omega^{(2)} = \omega^{(2)},$$

i.e.

$$\left(g^t\right)^* \omega^{(2)}(\xi,\eta) := \omega^{(2)}(g_*^t\xi, g_*^t\eta) = \omega^{(2)}(\xi,\eta).$$

♠

Proof: We will prove this result for the manifold $(\mathbb{R}^2, dp\wedge dq)$, and postpone the proof of the general result for later. Recall that if we have a set $A \subset \mathbb{R}^2$, then

$$\int_A \omega^{(2)} - \int_A dp\wedge dq = \text{ area of } A.$$

To prove the theorem for $\mathbb{R}^2$, we will prove the equivalent statement for $\mathbb{R}^2$ that, for any region $A\subset\mathbb{R}^2$, the area of A is the same as the area of $g^t(A)$, i.e. the Hamiltonian flow preserves area.

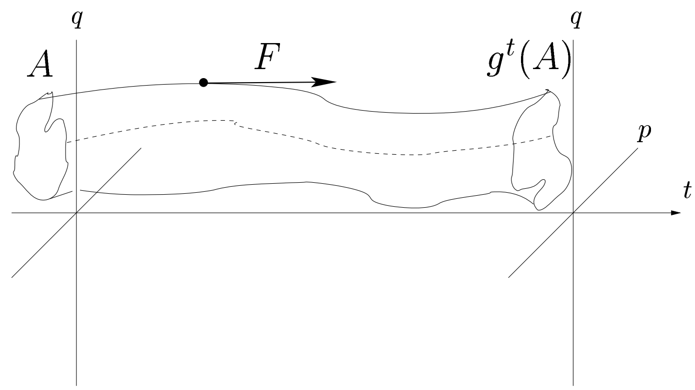

Figure 1.1: The region A flowing under the Hamiltonian

Let us consider a region A in the (p, q) plane, and let it flow under the Hamiltonian for some time (see Figure 1.1).

Let $S = A \cup C \cup g^t(A)$, the "total boundary" of the above cylinder. Consider the vector field

$$F = \begin{pmatrix} X_H \\ 1 \end{pmatrix},$$

where

$$X_H = \left(-\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}\right).$$

Apply the Divergence Theorem:

$$\iiint_{\text{int } S} \text{div} F \, dV = \iint_S F \cdot \hat{n} \, dS.$$

We can calculate that

$$\begin{aligned} \text{div} F &= \frac{\partial F_1}{\partial p} + \frac{\partial F_2}{\partial q} + \frac{\partial F_3}{\partial t} \\ &= \frac{-\partial^2 H}{\partial p \partial q} + \frac{\partial^2 H}{\partial p \partial q} + 0 = 0. \end{aligned}$$

Thus we have that

$$\iint_S F \cdot \hat{n} \, dS = 0.$$

Since $F \cdot \hat{n} = 0$ on C (the sides of the cylinder are of course tangent to the flow), we have

$$\iint_A F \cdot \hat{n} \, dS + \iint_{g^t(A)} F \cdot \hat{n} \, dS = 0.$$

Now, on A, $\hat{n} = (0, 0, -1)$, and on $g^t(A)$, $\hat{n} = (0, 0, 1)$. Thus

$$F \cdot \hat{n} = \begin{cases} -1 & \text{on } A \\ 1 & \text{on } g^t(A) \end{cases},$$

giving

$$\begin{aligned} \iint_A -1 + \iint_{g^t(A)} 1 &= 0 \\ \iint_A 1 &= \iint_{g^t(A)} 1 \\ \text{area of } A &= \text{ area of } g^t(A). \end{aligned}$$

■

Example: Let's consider the Hamiltonian $H(p,q) = \frac{1}{2}p^2 - \frac{1}{2}q^2$, so that

$$\begin{aligned}\dot{p} &= q,\\ \dot{q} &= p.\end{aligned}$$

Then

$$g^t \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix},$$

which means that

$$\begin{aligned}\begin{pmatrix} P \\ Q \end{pmatrix} &= \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} \begin{pmatrix} P \\ Q \end{pmatrix}\\ \begin{pmatrix} p \\ q \end{pmatrix} &= \begin{pmatrix} \cosh(t) & -\sinh(t) \\ -\sinh(t) & \cosh(t) \end{pmatrix} \begin{pmatrix} P \\ Q \end{pmatrix}.\end{aligned}$$

Then we calculate

$$\begin{aligned}\omega &= dp \wedge dq\\ &= (\cosh(t)\,dP - \sinh(t)\,dQ) \wedge (-\sinh(t)\,dP + \cosh(t)\,dQ)\\ &= \cosh^2(t)(dP \wedge dQ) - \sinh^2(t)(dP \wedge dQ)\\ &= dP \wedge dQ.\end{aligned}$$

1.3 Canonical Transformations

Definition 1.3
*A **canonical transformation** is a map $g\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that g preserves the symplectic form, i.e.*

$$g^*\omega^{(2)} = \omega^{(2)}.$$

♣

So, if we have the set of equations

$$\begin{pmatrix} \dot{p} \\ q \end{pmatrix} = \begin{pmatrix} -\frac{\partial H}{\partial q} \\ \frac{\partial H}{\partial p} \end{pmatrix},$$

and we make the change of variables $(P,Q) = g(p,q)$ where g is canonical, then we have

$$\begin{pmatrix} \dot{P} \\ \dot{Q} \end{pmatrix} = \begin{pmatrix} -\frac{\partial K}{\partial Q} \\ \frac{\partial H}{\partial p} \end{pmatrix},$$

with $K(P,Q) = H \circ g^{-1}(P,Q)$.

Recall that any 2-form on $\mathbb{R}^n$ can be written as

$$\omega^{(2)}(\xi,\eta) = \langle A\xi, \eta \rangle,$$

for some anti-symmetric matrix A.

Corollary 1.5
Given any 2-form $\omega^{(2)}$ on $\mathbb{R}^{2n+1}$, there is a $\xi \neq 0$ (called the **null-vector** or **vortex vector**) such that

$$\omega^{(2)}(\xi, \eta) = 0 \text{ for all } \eta \in \mathbb{R}^{2n+1}.$$

♠

Definition 1.4
*We say that $\omega^{(2)}$ is **non-degenerate** if the dimension of the set of null-vectors is as small as possible, i.e.*

$$\begin{array}{l} 0, \textit{ even dimension} \\ 1, \textit{ odd dimension.} \end{array}$$

♣

Proof: Write $\omega^{(2)}(\cdot,\cdot)$ as $\langle A\cdot,\cdot\rangle$. Since A is anti-symmetric,

$$\det(A) = \det(A^{\mathsf{T}}) = \det(-A) = -\det(A),$$

(the last equality being true since we are in an odd-dimensional space). Since $\det(A) = -\det(A)$, we have $\det(A) = 0$.

This means that 0 is an eigenvalue for A, so let ξ be its eigenvector. Then

$$\omega^{(2)}(\xi, \eta) = \langle A\xi, \eta\rangle = 0 \text{ for all } \eta.$$

■

Exercise 1.1
Let ω^1 be any 1-form on $\mathbb{R}^{2n+1}$, where we use the variables $(\underline{p}, \underline{q}, t)$, and p and q are n-dimensional. Define

$$\omega^{(2)}(\xi, \eta) = \sum_{j=1}^{n} dp_j \wedge dq_j - \omega^1 dt.$$

Show that $\omega^{(2)}$ is non-degenerate (has only a 1-dimensional null space).

Remark. Given any 1-form μ^1 on M^{2n+1}, we see from above that for every $x \in M$, $d\mu^1$ has at least one null-vector in $T_x M^{2n+1}$. If μ^1 is non-degenerate, then for every $x \in M$, $d\mu^1$ has exactly one null-vector in $T_x M^{2n+1}$.

Definition 1.5
*If μ^1 is non-degenerate, then the **vortex field** of μ^1 is the set of vortex directions of $d\mu^1$. An integral curve of the vortex field is called a **vortex line** of μ^1. Given a closed curve γ, the set of vortex lines through γ is called a **vortex tube**.*

♣

Lemma 1.6
Let σ be a vortex tube for μ^1 and let γ_1, γ_2 be closed curves encircling σ so that $\gamma_1 - \gamma_2 = \partial\tilde{\sigma}$ for $\tilde{\sigma}$ some piece of σ. Then

$$\oint_{\gamma_1} \mu^1 = \oint_{\gamma_2} \mu^1.$$

♠

Proof: We have that

$$\begin{aligned} &\oint_{\gamma_1} \mu^1 - \oint_{\gamma_2} \mu^1 = \\ &= \oint_{\gamma_1 - \gamma_2} \mu^1 \\ &= \iint_{\tilde{\sigma}} d\mu^1 = 0, \end{aligned}$$

since the tangent space to $\tilde{\sigma}$ contains a null-vector for $d\mu^1$ at every point (see Figure 1.2).

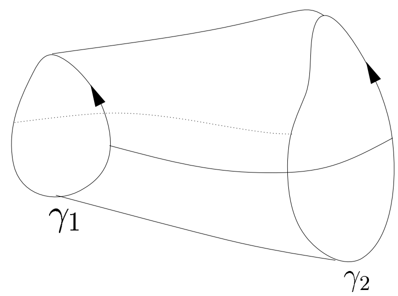

Figure 1.2: A picture for Stokes' Theorem

■

If we have a Hamiltonian $H = H(\underline{p}, \underline{q}, t)$, we define

$$\omega^1 = \sum_{j=1}^{n} p_j dq_j - H dt.$$

By the above exercise, this is non-degenerate.

Theorem 1.7
The vortex lines of ω^1 as defined above can be written as graphs $(\underline{p}(t), \underline{q}(t), t)$ where $\underline{p}, \underline{q}$ satisfy

$$\dot{p}_j = -\frac{\partial H}{\partial q_j}, \qquad \dot{q}_j = \frac{\partial H}{\partial p_j}. \tag{1.1}$$

♠

Proof: By calculation,

$$d\omega^1 = \sum dp_j \wedge dq_j - dH \wedge dt,$$

and we know that there is an A such that

$$d\omega^1(\xi, \eta) = \langle A\xi, \eta \rangle .$$

We will prove the case $\boxed{n = 1}$ explicitly. If $n = 1$, then we are on $\mathbb{R}^3$, so that

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}, \qquad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}.$$

Then

$$\begin{aligned} d\omega^1(\xi, \eta) &= dq \wedge dq(\xi, \eta) - dH \wedge dt(\xi, \eta) \\ &= (dq \wedge dq)(\xi, \eta) - \left\{ \left(\frac{\partial H}{\partial p} dp + \frac{\partial H}{\partial q} dq + \frac{\partial H}{\partial t} dt \right) \wedge dt \right\} (\xi, \eta) \\ &= (dp \wedge dq)(\xi, \eta) - \frac{\partial H}{\partial p} (dp \wedge dt)(\xi, \eta) - \frac{\partial H}{\partial q} (dq \wedge dt)(\xi, \eta) \\ &= \xi_1 \eta_2 - \xi_2 \eta_1 - \frac{\partial H}{\partial p} (\xi_1 \eta_3 - \xi_3 \eta_1) - \frac{\partial H}{\partial q} (\xi_2 \eta_3 - \xi_3 \eta_2). \end{aligned}$$

If we also write

$$d\omega^1(\xi,\eta) = \langle A\xi,\eta\rangle\,,$$

then equating coefficients gives us that

$$A = \begin{pmatrix} 0 & -1 & \dfrac{\partial H}{\partial p} \\ 1 & 0 & \dfrac{\partial H}{\partial q} \\ -\dfrac{\partial H}{\partial p} & -\dfrac{\partial H}{\partial q} & 0 \end{pmatrix}.$$

Remark: If we do the same proof for general n, we have that

$$A = \begin{pmatrix} \mathbf{0}_n & -\mathbf{1}_n & \left(\dfrac{\partial H}{\partial p_j}\right)_{j=1}^n \\ \mathbf{1}_n & \mathbf{0}_n & \left(\dfrac{\partial H}{\partial q_j}\right)_{j=1}^n \\ -\left(\dfrac{\partial H}{\partial p_j}\right)_{j=1}^n & -\left(\dfrac{\partial H}{\partial q_j}\right)_{j=1}^n & 0 \end{pmatrix}.$$

Now, we need to find the eigenvector of A with eigenvalue 0. Let

$$\xi^n = \begin{pmatrix} -\dfrac{\partial H}{\partial q} \\ \dfrac{\partial H}{\partial p} \\ 1 \end{pmatrix},$$

and we can check that for any n, $A\xi^n = 0$.

The vortex tube is parallel to ξ^n at every point. But if $(p(t), q(t))$ solve (1.1), they are also tangent to ξ^n. By Existence-Uniqueness, we are done.

■

Corollary 1.8
Let γ_1 be a simple closed curve in $\mathbb{R}^{2n+1}$. Consider the tube of solutions to Hamilton's Equations passing through γ_1. Let γ_2 be any other simple closed curve encircling the same tube of trajectories. Then

$$\oint_{\gamma_1} (p\,dq - H\,dt) = \oint_{\gamma_2} (p\,dq - H\,dt),$$

and this is called **Poincaré's Integral Invariant**. ♠

Corollary 1.9
Theorem 1.4 is true for all n. ♠

Proof: Suppose that t is a constant on both γ_1, γ_2 (see Figure 1.3). By the above Corollary,

$$\oint_{\gamma_1} p\,dq = \oint_{\gamma_2} p\,dq.$$

Then by Stokes',

$$\iint_{\sigma_1} d(p\,dq) = \iint_{\sigma_2} d(p\,dq),$$

or

$$\iint_{\sigma_1} dp\wedge dq = \iint_{g^{(t_2-t_1)}(\sigma_1)} dq\wedge dq.$$

■

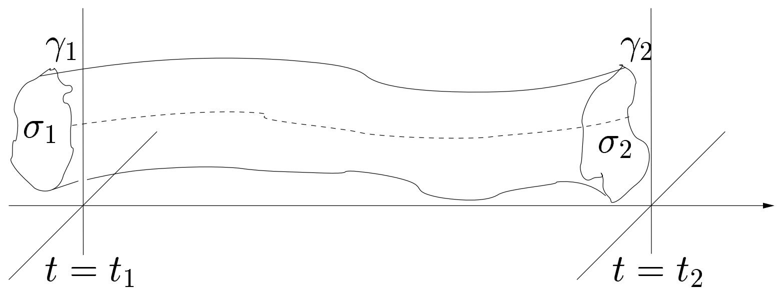

Figure 1.3:

1.4 Changing Variables

Consider a Hamiltonian system on $\mathbb{R}^{2n}$

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial q} \\ \dot{q} &= \frac{\partial H}{\partial p}, \end{aligned} \qquad \text{(eq:Hamiltons)}$$

and let $g\colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a canonical transformation

$$(P, Q) = g(p, q).$$

Theorem 1.10
Let $K(P, Q) = H \circ g^{-1}(p, q)$. Then the solutions of (1.1) expressed with respect to (P, Q) satisfy

$$\begin{aligned} \dot{P} &= -\frac{\partial K}{\partial Q}, \\ \dot{Q} &= \frac{\partial K}{\partial P}. \end{aligned}$$

♠

We will use the following proposition in the proof of the theorem, then prove the proposition.

Proposition 1.11
Let (P, Q, T) be coordinates on $\mathbb{R}^{2n+1}$ and let $K(P, Q, T)$ and $S(P, Q, T)$ satisfy

$$p\,dq - H\,dt = P\,dQ - K\,dT + dS.$$

Then the solutions of (1.1) expressed in the (P, Q, T) coordinates are integral curves of

$$\begin{aligned} \frac{\partial P}{\partial T} &= -\frac{\partial K}{\partial Q}, \\ \frac{\partial Q}{\partial T} &= \frac{\partial K}{\partial P}. \end{aligned}$$

♠

Proof: (Proof of the Theorem) Consider the 1-form $\mu^1 - p\,dq - P\,dQ$. (It seems strange to mix variables like this, but there's nothing illegal about it.)

By Poincaré's Integral Invariant, we know that

$$\oint_\gamma p\,dq = \oint_\gamma P\,dQ,$$

so that

$$\oint_\gamma \mu^1 = 0.$$

Let us define

$$S(p,q) = \int_{(p_0,q_0)}^{(p,q)} \mu^1.$$

(This is well-defined since $\oint \mu^1 = 0$ for every closed path.) Also, since S is independent of path,

$$dS = \mu^1 = p\,dq - P\,dQ,$$

or

$$p\,dq = P\,dQ + dS.$$

Thus

$$\begin{aligned} p\,dq - H\,dt &= P\,dQ - H\,dt + dS \\ &= P\,dQ - K\,dT + dS. \end{aligned}$$

Apply the Proposition, and we're done.

■

Proof: (Proof of the Proposition) By Theorem 1.7, the solutions of (1.1) are the vortex lines of $\omega^1 = p\,dq - H\,dt$. Since the vortex lines are independent of coordinates, they will also be vortex lines of

$$\tilde{\omega}^1 = P\,dQ - K\,dT + dS.$$

To find the equations of the vortex lines, consider

$$d\tilde{\omega}^1 = dP \wedge dQ - dK \wedge dT.$$

By the same calculation as before, this says that the vortex lines satisfy

$$\begin{aligned} \dot{P} &= -\frac{\partial K}{\partial Q}, \\ \dot{Q} &= \frac{\partial K}{\partial P}. \end{aligned}$$

■

1.5 Generating Functions and the Hamilton-Jacobi Equation

Recall that we showed before that if

$$(P(p,q), Q(p,q)) = g(p,q)$$

is a canonical transformation, then

$$p\,dq - P\,dQ = dS. \tag{1.2}$$

Let's suppose that we can choose q, Q as independent variables, i.e. that we can solve $Q = Q(p,q)$ to express $p = p(Q,q)$. By the Implicit Function Theorem, this is possible if

$$\det\left|\frac{\partial Q}{\partial p}\right| \neq 0.$$

Definition 1.6
Such canonical transformations are called ***free canonical transformations****.* ♣

We define the **generating function** by

$$S_1(Q, q) = S(p, (Q, q), q).$$

Note that

$$dS = \frac{\partial S_1}{\partial Q}\, dQ + \frac{\partial S_1}{\partial q}\, dq = p\, dq - P\, dQ,$$

or

$$p = \frac{\partial S_1}{\partial q}, \qquad P = -\frac{\partial S_1}{\partial Q}. \tag{1.3}$$

In summary, for every free canonical transformation, there is some S_1 so that (1.3) is satisfied.

Theorem 1.12
Given any smooth function $S_1(Q, q)$ that satisfies

$$\det\left(\frac{\partial^2 S_1}{\partial Q \partial q}\right),$$

then S_1 generates a canonical transformation via (1.3). ♠

Proof: Start with $p = \dfrac{\partial S_1}{\partial q}(Q, q)$. Since

$$\det\left(\frac{\partial p}{\partial Q}\right) = \det\left(\frac{\partial^2 S_1}{\partial Q \partial q}\right) \neq 0,$$

we can solve to find $Q = Q(p, q)$. The second equation in (1.3) gives

$$P = -\frac{\partial S_1}{\partial Q}\left(Q(p, q), q\right).$$

We need to check that the transformation is, in fact, canonical:

$$p\, dq - P\, dQ = \frac{\partial S_1}{\partial q}\, dq - \frac{\partial S_1}{\partial Q}\, dQ = dS_1.$$

■

Question: Are all canonical transformations free? By considering the identity transformation $(P, Q) = (p, q)$, we see that all are not.

Remark. To specify a general transformation $g \colon \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ requires $2n$ functions, each of which has $2n$ variables. On the other hand, a free canonical transformation requires only 1 function of $2n$ variables, which is considerably less information.

Jacobi tried to transform the Hamiltonian so that in the new variables,

$$K(P, Q) = K(Q).$$

This would be a good transformation since we would have

$$\dot{P} = -\frac{\partial K}{\partial Q}$$
$$\dot{Q} = \frac{\partial K}{\partial P} = 0.$$

So we would have $Q(t) = Q(0)$ for all t. Since Q would not change, neither would $\dfrac{\partial K}{\partial Q}$, so that $\dot{P} = \Omega$, where Ω is independent of time. Then

$$\begin{aligned} P(t) &= \Omega t + P(0) \\ Q(t) &= Q(0), \end{aligned}$$

and we would have essentially solved the system. Then we have

$$K(Q) = K(P,Q) = H(p,q) = H\left(\frac{\partial S_1}{\partial q}(Q,q), q\right).$$

So Jacobi said, try to solve

$$H\left(\frac{\partial S_1}{\partial q}(Q,q), q\right) = K(Q), \tag{\textbf{HJ}}$$

and find K and S_1. Equation (**HJ**) is known as the **Hamilton-Jacobi Equation** for H.

So if we could solve (**HJ**), then we'd be in good shape. But we've replace our ODE system with a nonlinear PDE, which is not usually an improvement.

Question: How do we deal with canonical transformations which are not free, e.g. the identity?

Suppose that we are given a canonical transformation and we can choose (P, q) as independent coordinates.

$$\begin{aligned} &p\,dq - P\,dQ = dS \\ &p\,dq + Q\,dP = dS + P\,dQ + Q\,dP = d(PQ + S). \end{aligned}$$

We define the **generating function of second type** to be

$$S_2(P,q) = PQ + S(p,q),$$

where everything on the right-hand side is thought of as functions of P and q. In a fashion very similar to above,

$$p\,dq + Q\,dP = dS_2 = \frac{\partial S_2}{\partial P}\,dP + \frac{\partial S_2}{\partial q}\,dq,$$

or

$$p = \frac{\partial S_2}{\partial q} \qquad Q = \frac{\partial S_2}{\partial P}. \tag{1.4}$$

As before, any smooth function S_2 satisfying

$$\det\left(\frac{\partial^2 S_2}{\partial P \partial q}\right) \neq 0,$$

Equation (1.4) defines a canonical transformation.

Question: What is the generating function for the identity?

$$S_2(P,q) = \langle P, q\rangle = \sum_{j=1}^{n} P_j q_j.$$

1.6 Near Identity Canonical Transformations

Let

$$S_\epsilon(P,q,\epsilon) = Pq + \epsilon\Sigma(P,q,\epsilon).$$

Then

$$\begin{aligned} p &= \frac{\partial S_\epsilon}{\partial q} = P + \epsilon\frac{\partial \Sigma}{\partial q}(P,q,\epsilon) \\ Q &= \frac{\partial S_\epsilon}{\partial P} = q + \epsilon\frac{\partial \Sigma}{\partial P}(P,q,\epsilon), \end{aligned}$$

where P, Q depend on ϵ. So exactly how do P and Q depend on ϵ?

$$\begin{aligned} \frac{\partial P}{\partial \epsilon} &= -\frac{\partial \Sigma}{\partial q}(P,q,\epsilon) - \epsilon\frac{\partial^2 \Sigma}{\partial \epsilon\partial q}(P,q,\epsilon), \\ \frac{\partial Q}{\partial \epsilon} &= -\frac{\partial \Sigma}{\partial P}(P,q,\epsilon) - \epsilon\frac{\partial^2 \Sigma}{\partial \epsilon\partial P}(P,q,\epsilon). \end{aligned}$$

If we let $\epsilon \to 0$, then

$$\begin{aligned} \frac{\partial P}{\partial \epsilon} &= -\frac{\partial \Sigma}{\partial q}(p,q,0) \\ \frac{\partial Q}{\partial \epsilon} &= \frac{\partial \Sigma}{\partial p}(p,q,0) \end{aligned}$$

since $P \to p$, $Q \to q$ as $\epsilon \to 0$. This infinitesimal canonical transformations satisfy Hamilton's Equations.

1.7 Canonical Perturbation Theory

Definition 1.7
A Hamiltonian system if called ***integrable*** *if there is a canonical change of variables*

$$g(p,q) = (I,\phi)$$

where $I \in \mathbb{R}^n$, $\phi \in \mathbb{T}^n$, such that

$$K(I,\phi) = H \circ g^{-1}(I,\phi) = h(I).$$

(I,ϕ) are called the ***action-angle variables****.* ♣

Recall as before, if we have an integrable system, then

$$\begin{aligned} \dot{I} &= -\frac{\partial K}{\partial \phi} = -\frac{\partial h}{\partial \phi} = 0 \implies I(t) = I(0) \\ \dot{\phi} &= \frac{\partial K}{\partial I} = \frac{\partial h}{\partial I} = \omega, \end{aligned}$$

where ω is independent of time, and thus

$$\phi(t) = \omega t + \phi(0).$$

So ϕ just precesses linearly around a torus.

Example: Let $H = \frac{1}{2}(p^2+q^2)$. Then

$$\begin{aligned} \dot{p} &= -q \\ \dot{q} &= p. \end{aligned}$$

We define

$$\psi = \tan^{-1}(q/p)$$
$$I = \frac{1}{2}(p^2+q^2) = H.$$

Then

$$dI = p\,dp + q\,dq$$
$$d\phi = \frac{1}{1+(q/p)^2}\left(\frac{dq}{p} - \frac{q\,dp}{p^2}\right).$$

We can calculate that $dI \wedge d\phi = dq \wedge dq$, and $K(I,\phi) = I$. So this system is integrable, and the action is just the Hamiltonian.

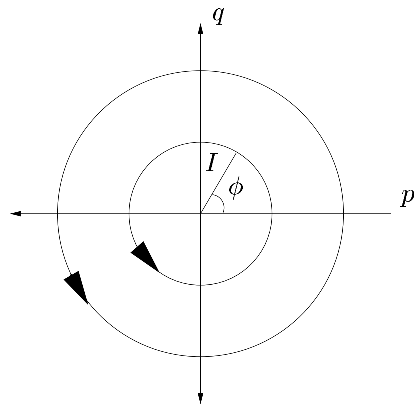

Figure 1.4: The flow for $H = (1/2)(p^2+q^2)$

Example: Let $H = \frac{1}{2}p^2 - \cos q$. Let's look for a generating function $S(I,q)$ such that

$$p = \frac{\partial S}{\partial q}$$
$$\phi = \frac{\partial S}{\partial I},$$

and if we write

$$H\left(\frac{\partial S}{\partial q}, q\right) = h(I)$$
$$\frac{1}{2}\left(\frac{\partial S}{\partial q}\right)^2 - \cos q = h(I).$$

Call $M_{h(I)}$ a level curve of $h(I)$. Then

$$dS = \frac{\partial S}{\partial I}\,dI + \frac{\partial S}{\partial q}\,dq.$$

Restricting to one trajectory, we have that

$$dS|_{M_{h(I)}} = p\,dq,$$

so that

$$S(I,q) = \int_{q_0}^{q} p\,dq$$

where the integral is done on $M_{h(I)}$.

Remark. If we integrate around $M_{h(I)}$, we don't get 0, so $S(I,q)$ is really a multi-valued function.

Calculating,

$$\begin{aligned}\oint_{M_h(I)} p\,dq &= \text{ area enclosed by } M_{h(I)} \\ &= \Pi_{M_{h(I)}}.\end{aligned}$$

Since $\phi = \dfrac{\partial S}{\partial I}$, and

$$2\pi = \frac{\partial}{\partial I}\left(\Pi_{M_{h(I)}}\right),$$

we have that

$$\Pi_{M_{h(I)}} = 2\pi I,$$

or

$$I = \frac{1}{2\pi}\Pi_{M_{h(I)}}.$$

Thus the generating function for the action-angle variables is

$$S(I,q) = \left.\left(\int_{q_0}^{q} p\,dq\right)\right|_{M_{h(I)}},$$

and

$$I = \frac{1}{2\pi}\Pi_{M_{h(I)}}.$$

Exercise 1.2
Work out the action-angle variables for the Kepler problem.

Exercise 1.3
Consider $H = \frac{1}{2}\langle p,p\rangle + \frac{1}{2}\langle q,Vq\rangle$ with $(p,q) \in \mathbb{R}^{2n}$, and V is a symmetric $n\times n$ matrix with positive eigenvalues. Show that H is integrable, and construct action-angle variables.

1.8 Nearly Integrable Hamiltonian Systems and KAM Theory

This section is based on [Way96].

Consider the harmonic oscillator

$$\ddot{x} + m^2 x = f(t).$$

If $F(t)$ oscillates with the same frequency of the "natural frequency" of the system, the solutions grow without bound (even if f is small). This phenomenon is known as resonace. In nonlinear systems, the problem is even worse, since we can have resonance even when the frequencies are rationally related, instead of the same.

We should note that the 3-body problem, when we ignore the interaction of the two small masses, is an integrable system. We could try to understand it perturbatively, but we could end up with the same problem.

1.8.1 The formal calculation

So, let's assume that we have a Hamiltonian as:

$$H(I,\phi) = h(I) + \epsilon f(I,\phi),$$

with $\epsilon \ll 1$. If we consider $\epsilon = 0$, the solutions of Hamilton's equations lie on n-dimensional tori.

Question: Suppose $\epsilon > 0$. Do these tori survive?

Our first try will be to make a canonical change of variables $(\tilde{I},\tilde{\phi}) = g(I,\phi)$ so that

$$\tilde{H}(\tilde{I},\tilde{\phi}) = \tilde{h}(\tilde{I}).$$

Since our system is nearly integrable, we'll look for a change of variables that is close to the identity, i.e. we'll assume that the generating function is

$$S(\tilde{I},\phi) = \left\langle \tilde{I},\phi \right\rangle + \epsilon\Sigma(\tilde{I},\phi),$$

and then

$$I = \frac{\partial S}{\partial \phi} = \tilde{I} + \epsilon\frac{\partial \Sigma}{\partial \phi},$$
$$\tilde{\phi} = \frac{\partial S}{\partial I} = \phi + \epsilon\frac{\partial \Sigma}{\partial I}.$$

In essence, we're looking for $S(\tilde{I},\phi)$ such that

$$I = \tilde{I} + \frac{\partial S}{\partial \phi}$$
$$\tilde{\phi} = \phi + \frac{\partial S}{\partial I},$$

and this transforms H to $\tilde{H}(\tilde{I},\tilde{\phi}) = \tilde{h}(\tilde{I})$.

Thus the Hamiltonian-Jacobi equation we want to solve is

$$h\left(\tilde{I} + \frac{\partial S}{\partial \phi}\right) + \epsilon f\left(\tilde{I} + \frac{\partial S}{\partial \phi},\phi\right) = \tilde{H}(\tilde{I}).$$

Linearizing,

$$h\left(\tilde{I} + \frac{\partial S}{\partial \phi}\right) = h(\tilde{I}) + \left\langle \frac{\partial h}{\partial I}(\tilde{I}), \frac{\partial S}{\partial \phi} \right\rangle + \mathcal{O}\left(\left(\frac{\partial S}{\partial \phi}\right)^2\right),$$

and

$$\epsilon f\left(\tilde{I} + \frac{\partial S}{\partial \phi},\phi\right) = \epsilon f(\tilde{I},\phi) + \mathcal{O}\left(\epsilon\frac{\partial S}{\partial \phi}\right)$$

Let's assume that S (and its derivatives) are $\mathcal{O}(\epsilon)$. Then the Hamilton-Jacobi Equation simplifies to

$$\tilde{h}(\tilde{I}) = h(\tilde{I}) + \left\langle \frac{\partial h}{\partial I}(\tilde{I}), \frac{\partial S}{\partial \phi} \right\rangle + \epsilon f(\tilde{I},\phi) + \mathcal{O}(\epsilon^2),$$

and we're looking for S, and we're free to choose $\tilde{h}$. This is a big improvement, since the above equation is linear. Let's ignore the $\mathcal{O}(\epsilon^2)$ term and try to find an approximate solution.

Let us denote

$$\omega(\tilde{I}) = \frac{\partial h}{\partial I}(\tilde{I}).$$

Then our linear equation is

$$\left\langle \omega(\tilde{I}), \frac{\partial S}{\partial \phi} \right\rangle + \epsilon f(\tilde{I}, \phi) = \tilde{h}(\tilde{I}) - h(\tilde{I}) \tag{\textbf{L}}$$

Since f is periodic in ϕ, we can expand

$$f(I, \phi) = \sum_{k \in \mathbb{Z}^n} \hat{f}(I, k) e^{i\langle k, \phi \rangle}.$$

Let's look for

$$S(I, \phi) = \sum_{k \in \mathbb{Z}^n} \hat{S}(I, k) e^{i\langle k, \phi \rangle}.$$

Then if everything is right with the world,

$$\frac{\partial S}{\partial \phi}(\tilde{I}, \phi) = \sum_{k \in \mathbb{Z}^n} ik\hat{S}(\tilde{I}, k) e^{i\langle k, \phi \rangle}.$$

Plugging this into (**L**), we have that

$$\sum_{k \in \mathbb{Z}^n} \left\{ i \left\langle \omega(\tilde{I}), k \right\rangle \hat{S}(\tilde{I}, k) + \epsilon \hat{f}(\tilde{I}, k) \right\} e^{i\langle k, \phi \rangle} = \tilde{h}(\tilde{I}) - h(\tilde{I}).$$

If we consider the case $k = 0$, then we have to satisfy

$$\epsilon \hat{f}(\tilde{I}, 0) = \tilde{h}(\tilde{I}) - h(\tilde{I}),$$

but this can be satisfied by choosing

$$\tilde{h}(\tilde{I}) = h(\tilde{I}) + \epsilon \hat{f}(\tilde{I}, 0).$$

Now $\tilde{h}$ has been chosen for us. If we try to solve for the other choices of k, we have

$$i \left\langle \omega(\tilde{I}), k \right\rangle \hat{S}(\tilde{I}, k) + \epsilon \hat{f}(\tilde{I}, k) = 0,$$

so that

$$\hat{S}(\tilde{I}, k) = \frac{i\epsilon \hat{f}(\tilde{I}, k)}{\left\langle \omega(\tilde{I}), k \right\rangle},$$

or

$$S(\tilde{I}, \phi) = i\epsilon \sum_{k \in \mathbb{Z}^n - \{0\}} \frac{\hat{f}(\tilde{I}, k)}{\left\langle \omega(\tilde{I}), k \right\rangle} e^{i\langle k, \phi \rangle}. \tag{1.5}$$

Equation (1.5) presents some problems. Let's look at them in dimension 2. In this case, the denominator is $\omega_1(\tilde{I})k_1 + \omega_2(\tilde{I})k_2$.

1. What if $\left\langle \omega(\tilde{I}), k \right\rangle = 0$?

   This will happen if

$$\frac{\omega_1(\tilde{I})}{\omega_2(\tilde{I})} = -\frac{k_2}{k_1},$$

   i.e. is ω_1, ω_2 are rationally related. This is the resonance problem.

2. Suppose that $\omega_1/\omega_2 \notin \mathbb{Q}$. Then Dirichlet's Theorem says that there is a constant C and an infinite number of pairs of integers (p_j, q_j) such that

$$\left|\frac{\omega_1}{\omega_2} - \frac{p_j}{q_j}\right| \leq \frac{C}{q_j^2},$$

or

$$|q_j\omega_1 - p_j\omega_2| \leq \frac{C\omega_2}{q_j} \to 0.$$

This means that even if $\omega_1/\omega_2 \notin \mathbb{Q}$ and we don't have resonance, then there are pairs (k_1, k_2) such that the denominator becomes arbitrarily small. This is the **Small Denominator Problem**.

Remark. The denominators get arbitrarily small for any choice of ω, but we'll show that for "most" choices of ω, the denominators get small in a controlled way.

1.8.2 Some analysis

Proposition 1.13
Fix $\gamma > 1$. There is a $C_0 > 0$ such that if $0 < C < C_0$, there is a subset of $\{(\omega_1, \omega_2) | \, |\omega_1| + |\omega_2| < 1\}$ of positive measure such that

$$|k_1\omega_1 + k_2\omega_2| \geq \frac{C}{|K|^\gamma}, \text{ for all } k \neq 0, \tag{\textbf{D}}$$

where $|k| = |k_1| + |k_2|$. ♠

Definition 1.8
*A pair (ω_1, ω_2) satisfying $(\mathbf{D})$ is called a **number of type** (C, γ). The measure of the complement of this set goes to 0 as $C \to 0$.* ♣

Proof: Fix some $k \neq 0$. We will compute the measure of the set of points (ω_1, ω_2) such that

$$|k_1\omega_1 + k_2\omega_2| < \frac{C}{|k|^\gamma}. \tag{1.6}$$

Note that

$$\begin{aligned} \langle k, \omega\rangle &= |k|\,|\omega| \cos\theta, \\ &= |k| \cdot \text{length of the projection of } \omega \text{ onto } k. \end{aligned}$$

Rephrasing our inequality geometrically (see Figure 1.5), we're looking for the set of ω such that

$$|k|\,|\text{proj of } \omega \text{ onto } k| < \frac{C}{|k|^\gamma},$$

or

$$|\text{proj of } \omega \text{ onto } k| < \frac{\tilde{C}}{|k|^{\gamma+1}},$$

where $\tilde{C} < \sqrt{2}C$, which we have to keep track of since we mean two different things by $|k|$ above.
Geometrically, we see that

$$m\left(\left\{\omega \,\middle|\, \langle k, \omega\rangle < \frac{C}{|k|^\gamma}\right\}\right) \leq \frac{4\tilde{C}}{|k|^{\gamma+1}}.$$

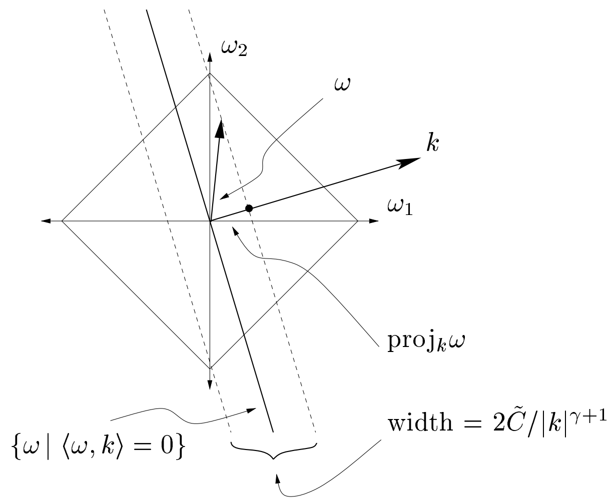

Figure 1.5: The set of "bad" numbers

The set of all ω which do not satisfy (**D**) is the same as the set of ω which satisfy (1.6) for some k. Thus its measure is bounded above by

$$\sum_{k\in\mathbb{Z}^n-\{0\}} \frac{4\tilde{C}}{|k|^{\gamma+1}}.$$

We will estimate the sum by approximating it with an integral:

$$\begin{aligned}
\sum_{k\in\mathbb{Z}^n-\{0\}} \frac{4\tilde{C}}{|k|^{\gamma+1}} &\approx \int_{|k|\geq 1} \frac{4\tilde{C}}{|k|^{\gamma+1}} d^2k \\
&= 4\tilde{C} \int_{r=1}^{\infty} \int_{\theta=0}^{2\pi} \frac{r\,dr}{r^{\gamma+1}}\, d\theta \\
&= 4\tilde{C}2\pi \int_{r=1}^{\infty} \frac{dr}{r^{\gamma}} \\
&= \frac{8\pi\tilde{C}}{\gamma-1}.
\end{aligned}$$

■

Remark: If $\omega \in \mathbb{R}^n$, the set of numbers in a bounded region which satisfy

$$|\langle k, \omega\rangle| \geq \frac{C}{|k|^{\gamma}}, \text{ for all } k \in \mathbb{Z}^n - \{0\},$$

has positive measure if $\gamma > n-1$ and C is sufficiently small.

To summarize, the denominators get small, but not too small.

So what about the numerators? Since we've controlled the denominators somewhat, if we can have the numerators shrinking sufficiently quickly, the sum will converge.

Suppose that $f(\tilde{I}, \phi)$ is analytic on a domain

$$\mathcal{A}_{\sigma,\rho}(\tilde{I}) := \{\tilde{I} \in \mathbb{C}^n, \phi \in \mathbb{C}^n \,\Big|\, \left|\tilde{I} - I^*\right| < \rho, |\text{Im}\,\phi_j| < \sigma, j = 1, 2, \ldots n\}.$$

Definition 1.9
We define

$$\|f\|_{\sigma,\rho} = \sup_{(\tilde{I},\phi)\in\mathcal{A}_{\sigma,\rho}(I^*)} \left| f(\tilde{I},\phi) \right| .$$

$\clubsuit$

Proposition 1.14

$$\left| \hat{f}(\tilde{I},k) \right| \leq \|f\|_{\sigma,\rho}\, e^{-\sigma|k|}.$$

$\spadesuit$

Proof: The proof will use Cauchy's Theorem, and the version of the theorem which says that integration is independent of path. Recall that by definition

$$\hat{f}(\tilde{I},k) = \left(\frac{1}{2\pi}\right)^n \int_0^{2\pi} \cdots \int_0^{2\pi} \hat{f}(\tilde{I},\phi) e^{i\langle k,\phi\rangle} d\phi_1 \ldots d\phi_n .$$

By shifting the ϕ_1 contour (see Figure 1.6), we have that

$$\begin{aligned}
&\int_0^{2\pi} f(\tilde{I},\phi_1,\ldots) e^{ik_1\phi_1} d\phi_1 \\
&= \int_0^{2\pi} f(\tilde{I},\phi_1+i\mu,\ldots) e^{ik_1(\phi_1+i\mu)}\, d\phi_1 \\
&= \int_0^{2\pi} f(\tilde{I},\phi_1+i\mu,\ldots) e^{ik_1\phi_1} e^{-k_1\mu}\, d\phi_1 ,
\end{aligned}$$

for any $\mu < \sigma$. By shifting each contour in a like fashion, we have

$$\begin{aligned}
\left| \hat{f}(\tilde{I},k) \right| &= \left(\frac{1}{2\pi}\right)^n \left| \int \cdots \int \hat{f}(\tilde{I},\phi+i\mu) e^{i\langle k,\phi\rangle} e^{-|k|\mu}\, d\phi \right| \\
&\leq e^{-|k|\mu} \left(\frac{1}{2\pi}\right)^n \int \cdots \int d\phi\, \|f\|_{\sigma,\rho} \\
&\leq \|f\|_{\sigma,\rho}\, e^{-|k|\mu}.
\end{aligned}$$

Since this is true for all $\mu < \sigma$, it is true for σ by continuity.

$\blacksquare$

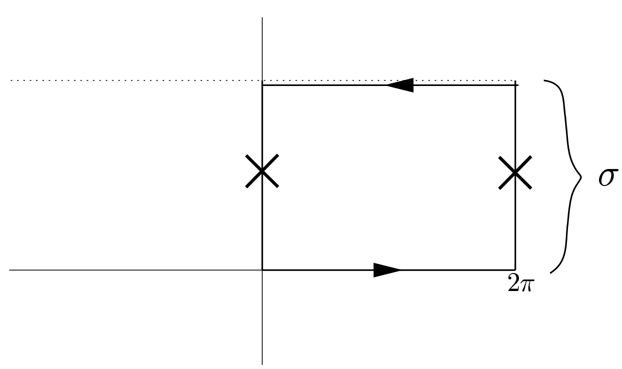

Figure 1.6: Shifting the contour upward

Exercise 1.4
Suppose that $f(x)$ is a real-valued, $\mathcal{C}^1$ function, with $f(x+2\pi) = f(x)$. Then we know that

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(k) e^{ikx}.$$

Suppose that $\hat{f}(k) \leq M e^{-\mu|k|}$ for some M, μ. Show that $f(x)$ can be extended to an analytic function $f(z)$ on the set

$$\{z \in \mathbb{C} | \ |\mathrm{Im}\, z| < \sigma\} \text{ for all } \sigma < \mu.$$

1.8.3 The ground step

$\boxed{\text{Kolmogorov's Idea:}}$ Let's "solve" Equation (**HJ**) by "Newton's Method".

1. Take $H(I, \phi) = h(I) + \epsilon f(\tilde{I}, \phi)$ as an approximate solution of (**HJ**).
2. Linearize to obtain (**L**) and solve this.
3. Make the canonical transformation whose generating function is the solution of (**L**).
4. We hope that

$$\tilde{H}(\tilde{I}, \tilde{\phi}) = H \circ \Phi(\tilde{I}, \phi) = \tilde{h}(\tilde{I}) + \tilde{f}(\tilde{I}, \tilde{\phi})$$

   with $\tilde{f} = \mathcal{O}(\epsilon^2)$.
5. Lather, rinse, repeat.

Let's make some standing hypotheses, which we will denote **H1,H2,H3**:

1. **(H1)** Suppose that $H(I\phi) = h(I) + f(I, \phi)$ is analytic on $\mathcal{A}_{\sigma_0,\rho_0}(I^*)$ for some $\sigma_0, \rho_0 > 0$. Assume without loss of generality that $\rho_0 < 1$ and $\sigma_0 < 1/2\pi$.
2. **(H2)** I^* is chosen that that

$$|\langle \omega(I^*), k\rangle| \geq \frac{C^*}{|k|^\gamma} \text{ for all } k \in \mathbb{Z}^n - \{0\}.$$

3. **(H3)** Suppose that $\frac{\partial^2 h}{\partial I_j \partial I_k}(I^*)$ is invertible.

We define

$$\Omega = \max\left(\sup_{|I-I^*|<\rho_0} \left\| \left(\frac{\partial^2 h}{\partial I \partial I} \right) \right\|, \sup_{|I-I^*|<\rho_0} \left\| \left(\frac{\partial^2 h}{\partial I \partial I} \right)^{-1} \right\| \right).$$

Also, we have

Definition 1.10
*$f(x)$ is **periodic with frequency** ω if $f(x) = \sum \hat{f}(k) e^{i\omega kx}$, or if there is a function $g\colon \mathbb{T}^1 \to \mathbb{R}$ with*

$$f(t) = g(\omega t).$$

*We say that f is **quasi-periodic** with frequency $\underline{\omega} = (\omega_1, \ldots, \omega_n)$ if there is a function $g\colon \mathbb{T}^n \to \mathbb{R}$ with*

$$f(t) = g(\omega_1 t, \ldots, \omega_n t).$$

♣

Theorem 1.15 (KAM)
Under hypotheses **H1,H2,H3**, there is an $\epsilon_0 > 0$ such that if $\|f\|_{\sigma_0,\rho_0} < \epsilon_0$, then the Hamiltonian system

$$H(I,\phi) = h(I) + f(I,\phi)$$

has a quasi-periodic orbit with frequency $\omega(I^*)$. ♠

Example: Let $H(I,\phi) = h(I) + f(I,\phi)$. If $f = 0$, then

$$\dot{I} = 0,$$
$$\dot{\phi} = \omega(I).$$

In particular, for initial conditions I^*, ϕ_0, we have that

$$I(t) = I^*,$$
$$\phi(t) = \phi_0 + \omega(I^*)t.$$

Now, if the components of $\omega(I^*)$ are all rationally related, then we will have a periodic orbit. If they are not, the orbit will be quasi-periodic. The KAM theorem says that under the right perturbations, these quasi-periodic orbits survive.

Recall that under our formal calculation, we had that

$$S(\tilde{I},\phi) = i \sum_{k\in\mathbb{Z}^n-\{0\}} \frac{\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}}{\left\langle \omega(\tilde{I}),k\right\rangle}.$$

We want the transformed Hamiltonian to be analytic, and this would require S to be analytic. Unfortunately, there is a dense set of $\omega \in \mathbb{T}^n$ for which $\langle\omega,k\rangle = 0$. Our nondegeneracy condition **H2** says, essentially, that when we move $\tilde{I}$, this moves $\omega(\tilde{I})$.

In other words: Given an $\tilde{I}$, it may not be true that $\left\langle \omega(\tilde{I}),k\right\rangle = 0$ for some k. But for some $\tilde{I}'$ close to $\tilde{I}$, there is a k such that $\left\langle \omega(\tilde{I}'),k\right\rangle = 0$ for some k, and thus our formal calculation blows up. So as the situation stands, our calculation is defined only on a set whose complement is dense.

The idea to fix this problem is to *truncate* the sum for S in such a way that the omitted terms are of the same order of the stuff we already threw away. Then we're not doing any worse, and S will make sense in a much better way.

So we define

$$S^{\leq}(\tilde{I},\phi) = i \sum_{0<|k|\leq N_0} \frac{\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}}{\left\langle \omega(\tilde{I}),k\right\rangle},$$

where $|k| = \sum_{j=1}^n |k_j|$.

Proposition 1.16
If $\rho_0 < \frac{L}{2\Omega n^2} N_0^{-(\gamma+1)}$, and f is analytic on $\mathcal{A}_{\sigma_0,\rho_0}$, then $S^{\leq}$ is analytic on $\mathcal{A}_{\sigma_o-\delta,\rho_0}$ and

$$\left\|S^{\leq}\right\|_{\sigma_0-\delta,\rho_0} \leq \frac{C(\gamma,n)}{L\delta^{(\gamma_1)n}} \|f\|_{\sigma_o,\rho_0}.$$ ♠

Proof: Recall from Proposition 1.14 that

$$\left|\hat{f}(\tilde{I},k)\right| \leq e^{-\sigma_0|k|} \|f\|_{\sigma_0,\rho_0}.$$

On the set $\mathcal{A}_{\sigma_0-\delta,\rho_0}$, we have that $|\operatorname{Im}\phi_j| < \sigma_0 - \delta$, so that

$$\left|e^{i\langle k,\phi\rangle}\right| \leq e^{(\sigma_o-\delta)|k|},$$

and

$$\left|\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}\right| \leq e^{-\delta|k|}\,\|f\|_{\sigma_0,\rho_0}\,. \tag{1.7}$$

So we've estimated the numerator, and the time have come for the denominator. From the Mean Value Theorem, we know

$$\begin{aligned}
\left|\omega(\tilde{I})-\omega(I^*)\right| &\leq \left|\int_{I^*}^{\tilde{I}} \frac{\partial\omega}{\partial I}\,dI\right| \\
&= \left|\int_{I^*}^{\tilde{I}} \left(\frac{\partial^2 h}{\partial I\partial I}\right)\,dI\right| \\
&\leq n^2\Omega(\tilde{I}-I^*) \\
&\leq n^2\Omega\rho_0.
\end{aligned}$$

Thus we can say that

$$\begin{aligned}
\left|\left\langle\omega(\tilde{I}),k\right\rangle\right| &= \left|\langle\omega(I^*),k\rangle+\left\langle\omega(\tilde{I}),k\right\rangle-\langle\omega(I^*),k\rangle\right| \\
&\geq \frac{L}{|k|^\gamma}-\left|\left\langle\omega(\tilde{I})-\omega(I^*),k\right\rangle\right| \\
&\geq \frac{L}{|k|^\gamma}-n^2\Omega\rho_o\,|k| \\
&\geq \frac{L}{2\,|k|^\gamma}.
\end{aligned}$$

Then we can say that

$$\begin{aligned}
\left|S^{\leq}(\tilde{I},\phi)\right| &\leq \sum_{0<|k|\leq N_0}\frac{e^{-\delta|k|}\,\|f\|_{\sigma_0,\rho_0}}{\frac{L}{2|k|^\gamma}} \\
&\leq \left(\frac{2\,\|f\|_{\sigma_0,\rho_0}}{L}\right)\sum_{0<|k|\leq N_0}|k|^\gamma\,e^{-\delta|k|} \\
&\leq \left(\frac{2\,\|f\|_{\sigma_0,\rho_0}}{L}\right)\left(1+2\sum_{j=1}^{\infty}j^\gamma e^{-\delta j}\right)^n.
\end{aligned}$$

Using the fact that

$$\begin{aligned}
\sum_{j=0}^{\infty}j^\gamma e^{-\delta j} &= \left(\frac{d}{d\delta}\right)^\gamma\sum_{j=0}^{\infty}e^{-\delta j} \\
&= \left(\frac{d}{d\delta}\right)^\gamma\frac{1}{1-e^{-\delta}} \\
&\approx \left(\frac{d}{d\delta}\right)^\gamma\frac{1}{\delta} \\
&= \frac{C}{\delta^{\gamma+1}},
\end{aligned}$$

we have

$$\begin{aligned}
\left|S^{\leq}(\tilde{I},\phi)\right| &\leq \left(\frac{2\,\|f\|_{\sigma_0,\rho_0}}{L}\right)\left(1+\frac{2C}{\delta^{\gamma+1}}\right)^n \\
&\leq \frac{C(\gamma,n)}{L}\frac{1}{\delta^{n(\gamma+1)}}\,\|f\|_{\sigma_0,\rho_0}\,.
\end{aligned}$$

■

So let's do the canonical transformation and check to see if everything works out. Let's look at

$$I = \tilde{I} + \frac{\partial S^{\leq}}{\partial \phi}, \qquad \tilde{\phi} = \phi + \frac{\partial S^{\leq}}{\partial I}. \tag{1.8}$$

We want to solve the first equation for $\tilde{I} = \tilde{I}(I, \phi)$. By the Inverse Function Theorem, we can solve for $\tilde{I}$ if

$$\frac{\partial}{\partial \tilde{I}}\left(\tilde{I} + \frac{\partial S^{\leq}}{\partial \phi}\right) = \mathbf{1} + \frac{\partial^2 S^{\leq}}{\partial \tilde{I} \partial \phi}$$

is invertible. We know that a matrix of the form $(\mathbf{1} + M)$ is invertible if $\|M\| < 1$, which case

$$(\mathbf{1} + M)^{-1} = \mathbf{1} - M + M^2 + \cdots$$

Theorem 1.17 (Cauchy)
If $f(z)$ is analytic on B_r, then

$$\sup_{z \in B_{r-\delta}} |f'(z)| \leq \frac{1}{\delta} \sup_{z \in B_r} |f(z)|.$$

♠

Using Cauchy's Theorem, we have that

$$\begin{aligned}
\left\| \left(\frac{\partial^2 S^{\leq}}{\partial \tilde{I} \partial \phi} \right) \right\|_{\sigma_0 - 2\delta, \rho_0/2} &\leq n^2 \max_{j,k} \left\| \frac{\partial^2 S^{\leq}}{\partial \tilde{I}_j \partial \phi_k} \right\|_{\sigma_0 - 2\delta, \rho_0/2} \\
&\leq \frac{n^2}{\delta \rho_0 / 2} \left\| S^{\leq} \right\|_{\sigma_o - \delta, \rho_0} \\
&\leq \frac{2C \|f\|_{\sigma_0, \rho_0}}{\rho_0 L \delta^{n(\gamma+1)+1}}.
\end{aligned}$$

We can state this as a

Proposition 1.18
If

$$\frac{2C \|f\|_{\sigma_0, \rho_0}}{\rho_0 L \delta^{n(\gamma+1)+1}} < 1,$$

then the first equation in (1.8) can be inverted to find $\tilde{I}(I, \phi)$.

♠

We can show that under exactly the same condition, the second equation can bve inverted to find $\phi = \phi(\tilde{I}, \tilde{\phi})$. Summarizing, we have

Proposition 1.19
Suppose that

$$\frac{2C \|f\|_{\sigma_0, \rho_0}}{\rho_0 L \delta^{n(\gamma+1)+1}} < 1, \tag{1.9}$$

and

$$\rho_0 < \frac{L}{2 \Omega n^2} N^{-(\gamma+1)}. \tag{1.10}$$

Then the system

$$I = \tilde{I} + \frac{\partial S^{\leq}}{\partial \phi}, \qquad \tilde{\phi} = \phi + \frac{\partial S^{\leq}}{\partial I}$$

defines an analytic and invertible transformation on $\mathcal{A}_{\sigma_0 - 3\delta, \rho_0/4}$.

♠

Now we want to do the canonical transformation and show that we really have a reduction in order. Writing $(I,\phi)=\Phi(\tilde{I},\tilde{\phi})$, we have

$$
\begin{aligned}
H^1(\tilde{I},\tilde{\phi}) &:= H\circ\Phi(\tilde{I},\tilde{\phi})\\
&= h(I(\tilde{I},\tilde{\phi}))+f(I(\tilde{I},\tilde{\phi}),\phi(\tilde{I},\tilde{\phi}))\\
&= h(\tilde{I}+\frac{\partial S^{\le}}{\partial\phi}(\tilde{I},\phi(\tilde{I},\tilde{\phi}))+f(\tilde{I}+\frac{\partial S^{\le}}{\partial\phi}(\tilde{I},\phi(\tilde{I},\tilde{\phi})),\phi(\tilde{I},\tilde{\phi})).
\end{aligned}
$$

We can write

$$
\begin{aligned}
h(\tilde{I}+\frac{\partial S^{\le}}{\partial\phi})-h(\tilde{I}) &= h\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi}\right)\Big|_{t=0}^{t=1}\\
&= \int_0^1\frac{d}{dt}\left(h\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi}\right)\right)\,dt\\
&= \int_0^1\left\langle\frac{\partial h}{\partial I}\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi}\right),\frac{\partial S^{\le}}{\partial\phi}\right\rangle\,dt\\
&= \left\langle\frac{\partial h}{\partial I}(\tilde{I}),\frac{\partial S^{\le}}{\partial\phi}\right\rangle+\int_0^1\left\langle\frac{\partial h}{\partial I}\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi}\right)-\frac{\partial h}{\partial I}(\tilde{I}),\frac{\partial S^{\le}}{\partial\phi}\right\rangle\,dt\\
&= \left\langle\frac{\partial h}{\partial I}(\tilde{I}),\frac{\partial S^{\le}}{\partial\phi}\right\rangle+\int_{t=0}^1\int_{s=0}^t\left\langle\left\langle\frac{\partial^2 h}{\partial^2 I}\left(\tilde{I}+s\frac{\partial S^{\le}}{\partial\phi}\right),\frac{\partial S^{\le}}{\partial\phi}\right\rangle,\frac{\partial S^{\le}}{\partial\phi}\right\rangle\,ds\,dt.
\end{aligned}
$$

Similarly, writing

$$
f\left(\tilde{I}+\frac{\partial S^{\le}}{\partial\phi},\phi\right)-f(\tilde{I},\phi)=\int_0^1\frac{\partial f}{\partial I}\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi},\phi\right)\,dt
$$

gives us that

$$
\begin{aligned}
H^1(\tilde{I},\tilde{\phi}) &= h(\tilde{I})+\left\langle\frac{\partial h}{\partial I}(\tilde{I}),\frac{\partial S^{\le}}{\partial\phi}\right\rangle+f(\tilde{I},\phi)\\
&\quad+\int_0^1\int_0^t\left\langle\left\langle\frac{\partial^2 h}{\partial^2 I}\left(\tilde{I}+s\frac{\partial S^{\le}}{\partial\phi}\right),\frac{\partial S^{\le}}{\partial\phi}\right\rangle,\frac{\partial S^{\le}}{\partial\phi}\right\rangle\,ds\,dt+\int_0^1\frac{\partial f}{\partial I}\left(\tilde{I}+t\frac{\partial S^{\le}}{\partial\phi},\phi\right)\,dt.
\end{aligned}
$$

Above, we see that the two integrals are formally $\mathcal{O}(\|f\|^2)$. The second and third terms are formally $\mathcal{O}(\|f\|)$, but this is bad and we don't want them there. Fortunately, they almost cancel. For if we not truncated S, these terms would exactly cancel, since we picked S to do exactly that.

Recall that

$$
S^{\le}(\tilde{I},\phi)=i\sum_{0<|k|\le N_0}\frac{\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}}{\left\langle\omega(\tilde{I}),k\right\rangle},
$$

and so

$$
\begin{aligned}
\left\langle\omega(\tilde{I}),\frac{\partial S^{\le}}{\partial\phi}\right\rangle &= \left\langle\omega(\tilde{I}),\left(-\sum_{0<|k|\le N_0}\frac{k\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}}{\left\langle\omega(\tilde{I}),k\right\rangle}\right)\right\rangle\\
&= -\sum_{0<|k|\le N_0}\hat{f}(\tilde{I},k)e^{i\langle k,\phi\rangle}.
\end{aligned}
$$

Then we have that

$$\left\langle \omega(\tilde{I}), \frac{\partial S^{\leq}}{\partial \phi} \right\rangle + f(\tilde{I}, \phi) = -\sum_{0<|k|\leq N_0} \hat{f}(\tilde{I}, k) e^{i\langle k, \phi \rangle} + \sum_k \hat{f}(\tilde{I}, k) e^{i\langle k, \phi \rangle}$$
$$= f(\tilde{I}, 0) + \sum_{|k|>N_0} \hat{f}(\tilde{I}, k) e^{i\langle k, \phi \rangle}.$$

So we can write

$$H^1 = h(\tilde{I}) + \hat{f}(\tilde{I}, 0) + \sum_{|k|>N_0} \hat{f}(\tilde{I}, k) e^{i\langle k, \phi \rangle}$$
$$+ \int_0^1 \int_0^t \left\langle \left\langle \frac{\partial^2 h}{\partial^2 I} \left(\tilde{I} + s \frac{\partial S^{\leq}}{\partial \phi} \right), \frac{\partial S^{\leq}}{\partial \phi} \right\rangle, \frac{\partial S^{\leq}}{\partial \phi} \right\rangle ds\, dt + \int_0^1 \frac{\partial f}{\partial I} \left(\tilde{I} + t \frac{\partial S^{\leq}}{\partial \phi}, \phi \right) dt.$$

Looking at this, we see that $\hat{f}(\tilde{I}, 0)$ doesn't depend on ϕ, so we can throw that into h. So as long as the integrals are $\mathcal{O}(\|f\|^2)$ (which is certainly formally true) and the tail of the sum is $\mathcal{O}(\|f\|^2)$ (which we will make true by a choice of N_0, then we have significantly improved our situation.

We should remark that H^1 is guaranteed to be analytic on $\mathcal{A}_{\sigma_0 - 4\delta, \rho/4}(I^*)$

Lemma 1.20

$$\sup_{\left|\tilde{I} - I^*\right| < \rho_0/4} \left| \tilde{h}(\tilde{I}) - h(\tilde{I}) \right| \leq \|f\|_{\sigma_o, \rho_0}.$$

♠

Let's write the terms we want to bound as:

$$\mathbf{I} = \int_0^1 \left\langle \frac{\partial f}{\partial I} \left(\tilde{I} + t \frac{\partial S}{\partial \phi} \right), \frac{\partial S^{\leq}}{\partial \phi} \right\rangle dt,$$
$$\mathbf{II} = \int_0^1 \int_0^t \left\langle \left\langle \frac{\partial^2 h}{\partial^2 I} \left(\tilde{I} + s \frac{\partial S^{\leq}}{\partial \phi} \right), \frac{\partial S^{\leq}}{\partial \phi} \right\rangle, \frac{\partial S^{\leq}}{\partial \phi} \right\rangle ds\, dt,$$
$$\mathbf{III} = \sum_{|k|>N_0} \hat{f}(\tilde{I}, k) e^{i\langle k, \phi \rangle}.$$

Lemma 1.21 (I)

$$\|\mathbf{I}\|_{\sigma_0 - 4\delta, \rho/4} \leq n \frac{\|f\|_{\sigma_0, \rho_0}}{\rho_0/2} \left(C(\gamma, n) \frac{\|f\|_{\sigma_0, \rho_0}}{\delta^{(\gamma_1)n} L} \right) \frac{1}{\delta}$$
$$\leq \frac{C(\gamma, n)}{\rho_0 \delta^{(\gamma+1)n+1} L} \|f\|^2_{\sigma_0, \rho_0}.$$

♠

Lemma 1.22 (II)

$$\|\mathbf{II}\|_{\sigma_0 - 4\delta, \rho_0/4} \leq n^2 \Omega \left(\frac{C(\gamma, n) \|f\|_{\sigma_o, \rho_0}}{\delta^{(\gamma+1)n} L} \frac{1}{\delta} \right)^2$$
$$\leq \frac{C(\gamma, n)\Omega}{\delta^{2(\gamma+1)n+2} L^2} \|f\|^2_{\sigma_0, \rho_0}.$$

♠

And if $(\tilde{I},\tilde{\phi}) \in \mathcal{A}_{\sigma_0-4\delta,\rho_0/4}(I^*)$, then we can say that

$$\begin{aligned}
\mathbf{III} &\leq \sum_{|k|>N_0} \|f\|_{\sigma_0,\rho_0} e^{-|k|\sigma_0} e^{|k|(\sigma_o-3\delta)} \\
&= \left(\sum_{|k|>N_0} e^{-3\delta|k|} \right) \|f\|_{\sigma_0,\rho_0} \\
&\leq \left(\sum_{|k|>N_0} e^{-2\delta|k|} \right) e^{-\delta N_0} \|f\|_{\sigma_0,\rho_0} \\
&\leq \left(1 + 2\sum_{j=0}^{\infty} e^{-2\delta j} \right)^n e^{-\delta N_0} \|f\|_{\sigma_0,\rho_0} \\
&\leq \left(1 + \frac{2}{1-e^{-2\delta}} \right)^n e^{-\delta N_0} \|f\|_{\sigma_0,\rho_0} .
\end{aligned}$$

And so we get

Lemma 1.23 (III)

$$\|\mathbf{III}\|_{\sigma_0-4\delta,\rho_0/4} \leq \frac{C(\gamma,n)}{\delta^n} e^{-\delta N_0} \|f\|_{\sigma_0,\rho_0} .$$ ♠

Let's summarize all of this in a

Proposition 1.24
If Equations (1.9) and (1.10) hold, then

$$\tilde{H}(\tilde{I},\tilde{\phi}) = \tilde{h}(\tilde{I}) + \tilde{f}(\tilde{I},\tilde{\phi})$$

is analytic on $\mathcal{A}_{\sigma_0-4\delta,\rho_0/4}(I^*)$ with

1. $$\left|\tilde{h}(\tilde{I}) - h(\tilde{I})\right| \leq \|f\|_{\sigma_0,\rho_0} \text{ for } \left|\tilde{I} - I^*\right| < \rho_0/4,$$

2. and
$$\left\|\tilde{f}\right\|_{\sigma_0-4\delta,\rho/4} \leq C(\gamma,n) \left\{ \frac{\Omega n^2}{\delta^{2+2(\gamma+1)+n} L} + \frac{1}{\rho_0 \delta^{(\gamma+1)n+1} L} \right\} \|f\|^2_{\sigma_0,\rho_0} + \frac{C(\gamma,n)}{\delta^n} e^{-\delta N_0} \|f\|_{\sigma_0,\rho_0} .$$ ♠

We can make this term $\mathcal{O}(\|f\|^2)$ if we choose N_0 so that

$$e^{-\delta N_0} \approx \|f\|_{\sigma_0,\rho_0} ,$$

or

$$N_0 = \frac{\left|\log \|f\|_{\sigma_0,\rho_0}\right|}{\delta}.$$

So all is well if we choose N_0 large enough. But recall then that that makes

$$\rho_0 < \frac{L}{2n^2\Omega \left(\frac{\left|\log \|f\|_{\sigma_0,\rho_0}\right|}{\delta} \right)^{\gamma+1}},$$

so making N_0 large makes ρ_0, our domain of analyticity, much smaller. So at the end of the day, we *must* check that all of the inequalities are consistent.

Remark. Another problem is that if we set $\tilde{\omega}(\tilde{I}) = \frac{\partial \tilde{h}}{\partial \tilde{I}}(\tilde{I})$, then in general $\tilde{o}(I^*) \neq \omega^*$. In particular, we can no longer be sure that $\tilde{o}(I^*)$ is of type (L, γ).

Kolmogorov's solution to this was to notice that we're not concerned with I^* in particular, just ω^*. So we change the action variables from I^* to $I^{(1)}$ where $\tilde{\omega}(I^{(1)}) = \omega^*$. Then we know that

$$\left| \left\langle \tilde{\omega}(I^{(1)}), k \right\rangle \right| \geq \frac{L}{|k|^\gamma}.$$

So what should we choose $I^{(1)}$ to be? Let $I^{(1)} = I^* + \delta I$. Then

$$\begin{aligned}
\tilde{o}(I^{(1)}) &= \frac{\partial \tilde{h}}{\partial \tilde{I}}(I^{(1)}) = \frac{\partial h}{\partial \tilde{I}}(I^{(1)}) + \frac{\partial \hat{f}(\tilde{I}, 0)}{\partial \tilde{I}}(I^{(1)}, 0) \\
&= \omega(I^* + \delta I) + \frac{\partial \hat{f}}{\partial \tilde{I}}(I^* + \delta I, 0) \\
&= \underbrace{\omega(I^*)}_{\omega^*} + \left\langle \frac{\partial \omega}{\partial I}(I^*), \delta I \right\rangle + \frac{\partial \hat{f}(\tilde{i}, 0)}{\partial \tilde{I}}(I^*, 0) + \mathcal{O}\left((\delta I)^2, \left\langle \frac{\partial \hat{f}}{\partial I}, \delta I \right\rangle \right).
\end{aligned}$$

Up to first order, we simply choose δI so that

$$\left\langle \frac{\partial \omega}{\partial I}(I^*), \delta I \right\rangle + \frac{\partial \hat{f}(\tilde{I}, 0)}{\partial \tilde{I}}(I^*, 0) = 0,$$

or

$$\begin{aligned}
\delta I &= \left[\frac{\partial \omega}{\partial I}(I^*) \right]^{-1} \left[\frac{\partial \hat{f}}{\partial \tilde{I}}(I^*, 0) \right] \\
&= \left[\frac{\partial^2 h}{\partial I \partial I}(I^*) \right]^{-1} \left[\frac{\partial \hat{f}}{\partial \tilde{I}}(I^*, 0) \right].
\end{aligned}$$

Since the linear piece is invertible, the rigorous justification follows from the Implicit Function Theorem.

Exercise 1.5
Complete details of the rigorous argument, and show that there is a $C > 0$ such that

$$\left| I^{(1)} - I^* \right| \leq \frac{C\Omega}{\rho_0} \left\| f \right\|_{\sigma_o, \rho_0}.$$

We'll choose ρ_1 so that $\tilde{H}$ is analytic on $\mathcal{A}_{\sigma_1, \rho_1}(I^{(1)})$, where $\sigma_1 = \sigma_0 - 4\delta$, and we must always check (see Figure 1.7) that

$$\frac{C\Omega}{\rho} \left\| f \right\|_{\sigma_0, \rho_0} < \frac{\rho_0}{\delta}.$$

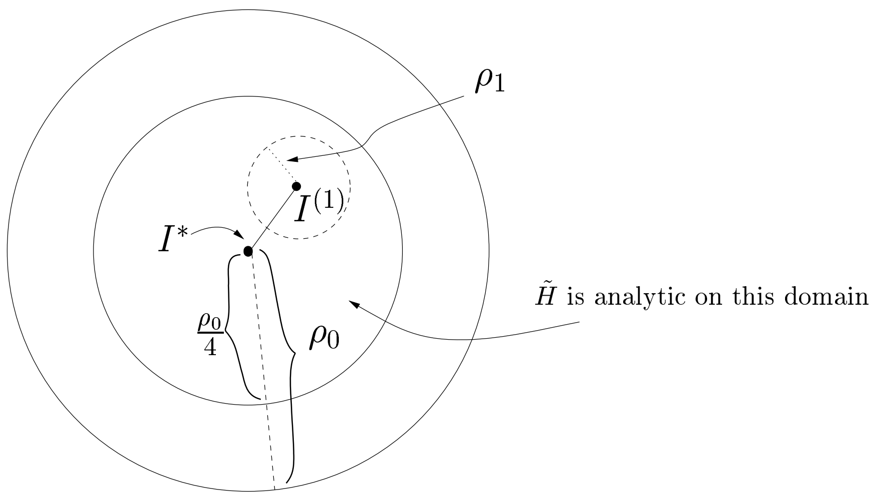

Figure 1.7: Shrinking the domain and moving I^*

1.8.4 The inductive argument

Let use define the following inductive constants:

$$\delta_n = \frac{\sigma_0}{36(1+n^2)}, \qquad n \geq 0$$
$$\sigma_{n+1} = \sigma_n - 4\delta_n,$$
$$\rho_{n+1} = \rho_n/8,$$
$$\epsilon_0 = \|f\|_{\sigma_0,\rho_0}, \qquad \epsilon_n = \epsilon_0^{(3/2)^{(n/\gamma)}},$$
$$N_n = \frac{|\log \epsilon_n|}{\delta_n}.$$

Note that $\sigma_n > \sigma_0/2$ for all n, and that $\rho_n \to 0$.

Theorem 1.25 (The Inductive Lemma)
There is a $C(\gamma, n) > 0$ such that if

1.
$$\rho_0 < \frac{C(\gamma,n)L}{\Omega N_0^{\gamma+1}},$$

2.
$$\epsilon_0 < \frac{C(\gamma,n)}{\Omega^8}\sigma_0^{8n(4\gamma+1)}\rho_0^8 L^{16},$$

then

1. there is a sequence of canonical transformations

$$\Phi_k \colon \mathcal{A}_{\sigma_k - 3\delta_k, \rho_k/4}(I_k^*) \to \mathcal{A}_{\sigma_k - 2\delta_k, \rho_k/2}(I_k^*)$$

whose generating function satisfies

$$\left\| S_k^{\leq} \right\|_{\sigma_k - \delta_k, \rho_k} \leq \frac{C(\gamma, n)}{\delta_k^{n(\gamma+1)}} \frac{\epsilon_k}{L}$$

2. If we set

$$H^{k+1} = H^k \circ \Phi_k = h^{k+1}(I) + f^{k+1}(I, \phi),$$

then

(a)

$$\left\| f^{k+1} \right\|_{\sigma_{k+1}, \rho_{k+1}} \leq \epsilon_{k+1},$$

and

(b)

$$\left| h^{k+1}(I) - h^k(I) \right| \leq \epsilon_k \text{ for all } \left| I - I_{k+1}^* \right| < \rho_{k+1}.$$

3. If we set

$$\omega_k(I) := \frac{\partial h^k}{\partial I}(I),$$

then

$$\omega(I_k^*) = \omega^*(:= \omega(I^*)),$$

and

$$\left| I_k^* - I_{k+1}^* \right| < \rho_k/8.$$

♠ Whew!

Now, if we look at the induction lemma carefully, we see that at each step, the new Hamiltonian that pops out is closer to integrable. We'd be done, but note that the domain of analyticity actually shrinks all the way to 0. We'd be hard pressed to take a derivative of a function defined at a point, so we've a bit more work to do.

Define

$$\Psi_k = \Phi_0 \circ \cdots \circ \Phi_k,$$

so that

$$\Psi_k : \mathcal{A}_{\sigma_k - 3\delta_k, \rho_k/4}(I_k^*) \to \mathcal{A}_{\sigma_0, \rho_0}(I^*),$$

and

$$H^k = H^0 \circ \Psi_k.$$

Now, we know that Ψ_k must be canonical, so that if $(I^{(k)}(t), \phi^{(k)}(t))$ is a solution of the system with Hamiltonian H^k, then

$$\Psi_k \left(I^{(k)}(t), \phi^{(k)}(t) \right)$$

is a solution of the system with Hamiltonian H^0.

The equations of motion of H^k are

$$\dot{I}^{(k)} = -\frac{\partial f^k}{\partial \phi},$$
$$\dot{\phi}^{(k)} = \omega^k(I) + \frac{\partial f^k}{\partial I}.$$

By Cauchy's Theorem, we know that

$$\left\|\frac{\partial f^k}{\partial \phi}\right\|_{\sigma_k - \delta_k, \rho_k} \leq \frac{n\epsilon_k}{\rho_k},$$
$$\left\|\frac{\partial f^k}{\partial I}\right\|_{\sigma_k, \rho_k/2} \leq \frac{2n\epsilon_k}{\rho_k}.$$

(We do need to check that $\epsilon_k/\rho_k \to 0$ as $k \to \infty$, but it's true.) We also need to check that the above tranjectory doesn't leave the domain of analyticity.

For example, suppose that $I^{(k)}(0) = I_k^*$. Then what is the shortest time it can take to leave the domain? We know that

$$\left|I^{(k)}(t) - I_k^*\right| \leq \frac{tn\epsilon_k}{\delta_k} < \rho_k.$$

So we're ok if

$$t < \frac{\rho_k \delta_k}{n\epsilon_k} \to \infty.$$

Lemma 1.26
If $(I^{(k)}(t), \phi^{(k)}(t))$ is a solution of the Hamiltonian system with Hamiltonian H^k and initial condition (I_k^*, ϕ_0) (with ϕ_0 arbitrary), then

1. $(I^{(k)}(t), \phi^{(k)}(t))$ remains in $\mathcal{A}_{\sigma_k - 3\delta_k, \rho_k/4}(I_k^*)$ for $|t| \leq T_k$,
2. $$\max\left(\sup_{|t|<T_k} |I_k(t) - I_k^*|, \sup_{|t|<T_k} \left|\phi^{(k)}(t) - (\omega^* t + \phi_0)\right|\right) \leq \frac{C(n,\gamma)\Omega\epsilon_k}{\rho_k \delta_k},$$

where $T_k = 2^k$. ♠

(Note that, by the choice of the ρ_k, we have that $I_k^* \to I_\infty^*$. If we pick any compact set $M \subset \mathbb{R}$, then for all $x \in M$, we have

$$(I^{(k)}(t), \phi^{(k)}(t)) \to (I_\infty^*, \omega^* t + \phi_o) \qquad (k \to \infty).$$

Now consider

$$\begin{aligned}\Psi_{k+1} - \Psi_k &= \Psi_k \circ \Phi_{k+1} - \Psi_k \\ &= \Psi_k \circ (\Phi_{k+1} - \mathbf{1})\end{aligned}$$

Recall that we define Φ_{k+1} so that

$$I = \tilde{I} + \frac{\partial S^{\leq}{}_{k+1}}{\partial \phi}$$
$$\phi = \tilde{\phi} + \frac{\partial Sl_{k+1}}{\partial I},$$

or, roughly, $\Phi_{k+1} = \mathbf{1} + \mathcal{O}(\epsilon_{k+1})$.

More rigorously, we have

$$\|\Psi_{k+1}(I,\phi) - \Psi_k(I,\phi)\|_{\sigma_{k+1},\rho_{k+1}} \leq \frac{C(n,\gamma)}{\delta_{k+1}^{n(\gamma+1)+2}\rho_{k+1}^2}\frac{\epsilon_{k+1}}{L} \to 0 \qquad (k \to \infty).$$

e checked!

This gives

1. $$\lim_{k\to\infty} \Psi_k(I_\infty^*, \omega^* t + \phi_0) = (I^*(t), \phi^*(t)),$$

2. $$\lim_{k\to\infty} \left|\Psi_k(I^{(k)}(t), \phi^{(k)}(t)) - \Psi_k(I_\infty^*, \omega^* t + \phi_0)\right| = 0.$$

Also note that $\Psi_k(I^{(k)}(t), \phi^{(k)}(t))$ is a solution of the system with Hamiltonian H^0 for all k, so the above equations show that $(I^*(t), \phi^*(t))$ is a quasi-periodic solution of the system with Hamiltonian H^0.

Now a few words about the proof of the inductive lemma:

So, we start with

$$H^k(I,\phi) = h^k(I) + f^k(I,\phi).$$

We want to construct a Φ_k so that

$$H^{k+1}(\tilde{I},\tilde{\phi}) = H^k \circ \Phi_k(\tilde{I},\tilde{\phi}) = h^{k+1}(\tilde{I}) + f^{k+1}(\tilde{I},\tilde{\phi}),$$

where

$$\left\|f^{k+1}\right\| \ll \left\|f^k\right\|.$$

Take a generating function for Φ_k to be of the form

$$\Sigma_k(\tilde{I},\phi) = \left\langle \tilde{I}, \phi \right\rangle + S_{k+1}^{\leq}(\tilde{I},\phi),$$

and

$$I = \tilde{I} + \frac{\partial S_{k+1}^{\leq}}{\partial \phi},$$
$$\phi = \tilde{\phi} + \frac{\partial S_{k+1}^{\leq}}{\partial I}.$$

So we want to have

$$\begin{aligned} H^k \circ \Phi^k(\tilde{I},\tilde{\phi}) &= h^k\left(\tilde{I} + \frac{\partial S_{k+1}^{\leq}}{\partial \phi}\right) + f^k\left(\tilde{I} + \frac{\partial S_{k+1}^{\leq}}{\partial \phi}, \phi\right) \\ &= h^{k+1}(\tilde{I}) + f^{k+1}(\tilde{I},\phi). \end{aligned}$$

Formally expanding:

$$h^k(\tilde{I}) + \left\langle \frac{\partial h^k}{\partial I}, \frac{\partial S_{k+1}^{\leq}}{\partial \phi} \right\rangle + \mathcal{O}(\|S\|^2) + f^k(\tilde{I},\phi) + \mathcal{O}(\left\|f^k\right\| \|S\|) = h^{k+1}(\tilde{I}) + \mathcal{O}(\left\|f^k\right\|^2).$$

So, if we can get that $\mathcal{O}(\|S\|) = \mathcal{O}(\|f\|)$, then we can throw away all the higher order terms and get

$$\left\langle \omega^k(\tilde{I}), \frac{\partial S_{k+1}^{\leq}}{\partial \phi} \right\rangle + f^k(\tilde{I}, \phi) = h^{k+1}(\tilde{I}) - h^k(\tilde{I}).$$

This is the linearized Hamilton-Jacobi equation. We can solve this by Fourier series

$$f^k(\tilde{I}, \phi) = \sum_m \hat{f}^k(\tilde{I}, m) e^{i\langle m, \phi \rangle},$$
$$S_{k+1}^{\leq}(\tilde{I}, \phi) = \sum_m \hat{S}_{k+1}^{\leq}(\tilde{I}, m) e^{i\langle m, \phi \rangle},$$

so that

$$\sum_m \left\{ i \left\langle m, \omega^k(\tilde{I}) \right\rangle \hat{S}_{k+1}^{\leq}(\tilde{I}, m) + \hat{f}(\tilde{I}, m) \right\} e^{i\langle m, \phi \rangle} = \hat{h}^{k+1}(\tilde{I}) - h^k(\tilde{I}).$$

Looking at the $m = 0$ term,

$$\hat{f}^k(\tilde{I}, 0) = h^{k+1}(\tilde{I}) - h^k(\tilde{I}),$$

or

$$h^{k+1}(\tilde{I}) := h^k(\tilde{I}) + \hat{f}^k(\tilde{I}, 0).$$

For all the $m \neq 0$ terms,

$$\hat{S}_{k+1}^{\leq}(\tilde{I}, m) = \frac{-i \hat{f}^k(\tilde{I}, m)}{\left\langle m, \omega^k(\tilde{I}) \right\rangle},$$

or

$$S_{k+1}^{\leq}(\tilde{I}, \phi) = -i \sum_{m \neq 0} \frac{\hat{f}^k(\tilde{I}, m)}{\left\langle \omega^k(\tilde{I}), m \right\rangle} e^{i\langle m, \phi \rangle}.$$

This means that $S_{k+1}^{\leq}$ no longer solves the linearized Hamilton-Jacobi equation, because of the error term

$$\sum_{|m| > N_k} \hat{f}(\tilde{I}, m) e^{i\langle m, \phi \rangle}.$$

How big is this error term? Recall that f^k is analytic on $\mathcal{A}_{\sigma_k, \rho_k}(I^*)$, which implies that

$$\left| e^{i\langle m, \phi \rangle} \right| \leq e^{|m|(\sigma_k - \delta_k)}.$$

Thus

$$\begin{aligned} \left\| \sum_{|m| > N_k} \hat{f}^k(\tilde{I}, m) e^{i\langle m, \phi \rangle} \right\|_{\sigma_k - \delta_k, \rho_k} &\leq \sum_{|m| > N_k} \left\| f^k \right\|_{\sigma_k, \rho_k} e^{-|m|\sigma_k} e^{|m|(\sigma_k - \delta_k)} \\ &\leq \sum_{|m| > N_k} \left\| f^k \right\|_{\sigma_k, \rho_k} e^{-|m|\delta_k} \\ &\leq C \|f\|_{\sigma_k, \rho_k} e^{-\delta_k N_k}. \end{aligned}$$

If we plug in our assumed value for N_k, we get that this term is $\mathcal{O}(\left\| f^k \right\|^2)$, and we are done.

Chapter 2

Hamiltonian PDE

Recall that we had to develop the following concepts for ODEs:

1. Symplectic geometry
2. Hamilton's Equations
3. Canonical Transformations
4. Analytic Maps
5. Complex Analysis, Cauchy's Theorem

We'll develop the theory of Hamiltonian PDE by thinking of them as infinite dimensional analogs of Hamiltonian ODE. In the following two examples, we see that PDE can be thought of as infinite-dimensional ODE:

Example: Consider the PDE

$$u_t = u_{xx}, \qquad u(0,t) = u(\pi,t) = 0.$$

In light of the boundary conditions, we can write u as its Fourier series:

$$u(x,t) = \sum_k a_k(t) \sin(kx).$$

If we plug this into the PDE, we get the infinite collection of ODEs

$$a_k'(t) = -k^2 a_k(t), \qquad k = 0, 1, 2, \ldots$$

Example: In ODEs, we consider linear systems of the form

$$\dot{x} = F(x), \qquad x \in M^n.$$

The solutions are paths in M, which we can write as $\phi_t(x)$. Analogously, if we think of X as a function space, we can write a PDE as

$$\frac{\partial u}{\partial t} = Ju,$$

for J some linear operator on X.

2.1 Differentiation in Banach space

Definition 2.1
*A **Hilbert space** is a complete inner-product space.* ♣

Some examples are

$$L^2[0,2\pi] = \{f\colon [0,2\pi] \to \mathbb{R} | \int_0^{2\pi} |f(x)|^2 \, dx < \infty\},$$

or

$$\ell^2 = \left\{ \{a_n\}_{n=0}^{\infty} | \; a_n \in \mathbb{C}, \sum |a_n|^2 < \infty \right\}.$$

Definition 2.2
*Let $O \subset X$ be a domain, and X, Y be Hilbert spaces. If $f\colon O \to Y$ is continuous, then f is **continuously differentiable** if there is a bounded linear map (which we call the derivative)*

$$f_*(x)\colon X \to Y$$

which depends continuously on x, such that if $x, x + x_1 \in O$, then

$$f(x + x_1) - f(x) = f_*(x)x_1 + o(\|x_1\|_X).$$ ♣

We can also define the (Hilbert) **adjoint** of the derivative as

$$f^*(x) = (f_*(x))^* \colon Y \to X$$

so that

$$\langle u, f_*(x)v \rangle_Y = \langle f^*(x)u, v \rangle_X .$$

A special case of the above is if $Y = \mathbb{R}$, i.e. we have a map

$$f\colon O \subset X \to \mathbb{R}.$$

Then we have that $f_*(x) \in X^*$ (the dual of X).

Theorem 2.1 (Riesz)
If $L \in X^*$, then there is a unique $y_L \in X$ so that

$$Lv = \langle y_L, v \rangle_X .$$

Consequently, $X \approx X^*$. ♠

Since $f_*(x) \in X^*$, there is an element in X, call it $\nabla f(x)$, with

$$f_*(x)v = \langle \nabla f(x), v \rangle .$$

Thus we have a map, called the **gradient map**

$$\begin{aligned} \nabla f\colon O &\to X \\ x &\mapsto \nabla f(x) \end{aligned}$$

Example: Let $X = \ell^2$, with

$$\begin{aligned} f\colon \ell^2 &\to \mathbb{R} \\ \{x_n\} &\mapsto \|\{x_n\}\|^2 , \end{aligned}$$

where we assume that the x_n are real. **Finish this example**

Recall that for $f\colon O \to \mathbb{R}$, we have the gradient map $\nabla f\colon O \to X$.

Proposition 2.2

$$\langle \nabla f(x)_* \xi, \eta \rangle = \langle \nabla f(x)_* \eta, \xi \rangle \,. \tag{2.1}$$

♣

Proof: Define a map

$$\begin{aligned} \mathbb{R}^2 &\to \mathbb{R} \\ (\alpha, \beta) &\mapsto f(x + \alpha\eta + \beta\xi) \end{aligned}$$

with x, η, ξ fixed in X. Then

$$\begin{aligned} \frac{\partial}{\partial \beta} f(x + \alpha\eta + \beta\xi) &= \lim_{h \to 0} \frac{f(x + \alpha\eta + (\beta + h)\xi) - f(x + \alpha\eta + \beta\xi)}{h} \\ &= \lim_{h \to 0} \frac{f_*(x + \alpha\eta + \beta\xi)(h\xi)}{h} \\ &= f_*(x + \alpha\eta + \beta\xi)(\xi) \\ &= \langle \nabla f(x + \alpha\eta + \beta\xi), \xi \rangle \,. \end{aligned}$$

Then we have

$$\begin{aligned} \frac{\partial}{\partial \alpha} \langle \nabla f(x + \alpha\eta + \beta\xi), \xi \rangle &= \lim_{h \to 0} \frac{\langle \nabla f(x + (\alpha + h)\eta + \beta\xi) - \nabla f(x + \alpha\eta + \beta\xi), \xi \rangle}{h} \\ &= \langle \nabla f_*(x + \alpha\eta + \beta\xi)\eta, \xi \rangle|_{\alpha=\beta=0} \\ &= \langle \nabla f_*(x)\eta, \xi \rangle \,. \end{aligned}$$

If we calculate the partial derivatives in the opposite order, we get the other term of (2.1), and by equality of mixed partials, we are done.

■

2.2 Analyticity

Definition 2.3
*Let X be a real Hilbert space. The we define the **complexification** of X as*

$$X^{\mathbb{C}} = X \oplus X,$$

where we have

$$i(x_1, x_2) = (-x_2, x_1).$$

♣

Now, consider $O^{\mathbb{C}} \subset X^{\mathbb{C}}$, $O^{\mathbb{C}}$ a domain, and

$$f \colon O^{\mathbb{C}} \to Y^{\mathbb{C}}.$$

Definition 2.4
*We say that f is **analytic** if it is differentiable when we consider $X^{\mathbb{C}}$ and $Y^{\mathbb{C}}$ as a real vector space, and, moreover, $f_*(x)$ is complex linear.*

♣

Now, consider $O \subset X$, X and Y (perhaps real) Hilbert spaces, and $f \colon O \to Y$.

Definition 2.5
*Then we say f is **analytic** if it can be extended to a map*

$$F \colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$$

which is complex analytic.

♣

Recall that we defined $f_*(x)$ so that

$$f_*\colon O^{\mathbb{C}} \to L(X^{\mathbb{C}}, Y^{\mathbb{C}}),$$

and note that $L(X^{\mathbb{C}}, Y^{\mathbb{C}})$ is a Banach space.

Definition 2.6

$$f_*^{(2)}\colon O^{\mathbb{C}} \to L(X^{\mathbb{C}}, L(X^{\mathbb{C}}, Y^{\mathbb{C}})) \approx L^2(X^{\mathbb{C}}, Y^{\mathbb{C}}),$$

where we define

$$L^2(X^{\mathbb{C}}, Y^{\mathbb{C}}) := \{\mathcal{L}\colon X^{\mathbb{C}} \times X^{\mathbb{C}} \to Y^{\mathbb{C}} |\ \mathcal{L} \text{ is bounded and bilinear}\}.$$

Analogously, we can define

$$f_*^{(p)}\colon O^{\mathbb{C}} \to L^p(X^{\mathbb{C}}, Y^{\mathbb{C}}).$$

♣

Theorem 2.3 (Taylor)
If $f\colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$ is $\mathcal{C}^p$ for $p \geq 1$ and $x + th \in O^{\mathbb{C}}$ for all $t \in [0,1]$, then

$$f(x+h) = f(x) + f_*(x)(h) + \frac{1}{2} f_*^{(2)}(x)(h,h) + \cdots + \frac{1}{(p-1)!} f_*^{(p-1)}(x)(h,h,h,\ldots,h) + R,$$

where

$$\|R\|_Y \leq \frac{\|h\|_{X^{\mathbb{C}}}^p}{p!} \sup_{0\leq t\leq 1} \left\| f_*^{(p)}(x+th) \right\|.$$

If $f \in \mathcal{C}^\infty$ and $R \to 0$ as $p \to \infty$ for small h, then f has a Taylor series

$$f(x+h) = \sum_{k\geq 0} \frac{1}{k!} f_*^{(k)}(x)(h,\ldots,h).$$

♠

Definition 2.7
$f\colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$ *is* ***weakly analytic*** *on* $O^{\mathbb{C}}$ *if for all* $x \in O^{\mathbb{C}}, h \in X^{\mathbb{C}}$ *and* $\mathcal{L} \in Y^*$, *the map*

$$\begin{aligned} \phi\colon \mathbb{C} &\to \mathbb{C} \\ x &\mapsto \mathcal{L}(f(x+zh)) \end{aligned}$$

is analytic in a neighborhood of $0 \in \mathbb{C}$.

♣

Definition 2.8
We say that $f\colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$ *is* ***locally bounded*** *if for all* $x \in O^{\mathbb{C}}$, *there is a neighborhood* $N \ni x$, $N \subset O^{\mathbb{C}}$, *such that* f *is bounded on* N.

♣

Theorem 2.4 (Big Ol' Theorem)
Let $f\colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$, then the following are equivalent

1. f is analytic on $O^{\mathbb{C}}$,
2. f is locally bounded on $O^{\mathbb{C}}$ and f is weakly analytic on $O^{\mathbb{C}}$,
3. $f \in \mathcal{C}^\infty$ and can be represented by a Taylor series in a neighborhood of any $x \in O^{\mathbb{C}}$.

Moreover, if $X^{\mathbb{C}}, Y^{\mathbb{C}}$ are Hilbert, and suppose $\{y_i\}$ is a countable subset of $Y^{\mathbb{C}}$ such that $\{y_i\}$ spans $Y^{\mathbb{C}}$. Let $\mathcal{L}_i = \langle \cdot, y_i \rangle_{Y^{\mathbb{C}}}$, then the following is also equivalent to the above:

4. f is locally bounded and the maps

$$\begin{aligned} \phi_i \colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \mathcal{L}_i(f(x+zh)) \end{aligned}$$

are analytic in a neighborhood of 0. ♠

Definition 2.9
We define the ***operator norm*** *of a linear map as*

$$\|F\|_{X^{\mathbb{C}},Y^{\mathbb{C}}} = \sup_{\substack{x \in X^{\mathbb{C}} \\ \|x\|_{X^{\mathbb{C}}}=1}} \|F(x)\|_{Y^{\mathbb{C}}} .$$

♣

Theorem 2.5 (Cauchy)
If $F \colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$ admits a bounded analytic extension to $O^{\mathbb{C}} + \delta$, then for all $u \in O^{\mathbb{C}}$, we have

$$\|F_*(u)\|_{X^{\mathbb{C}},Y^{\mathbb{C}}} \leq \frac{1}{\delta} \sup_{x \in O^{\mathbb{C}}+\delta} \|F(x)\|_{Y^{\mathbb{C}}} .$$

♠

Proof: Consider the map

$$\begin{aligned} \phi \colon B_\delta(0) &\to \mathbb{C} \\ \lambda &\mapsto \langle F(u+\lambda x), y \rangle_Y , \end{aligned}$$

where $O^{\mathbb{C}}$ is fixed and $\|x\| = \|y\| = 1$. Note that ϕ is analytic on $B_\delta(0)$. If we let γ_δ be the contour around 0 with radius δ, then Cauchy's Theorem for one complex variable tells us that

$$\begin{aligned} |\phi'(0)| &= \frac{1}{2\pi} \left| \int_{\gamma_\delta} \frac{\phi(\zeta)}{(z-\zeta)^2} \, d\zeta \right| \\ &\leq \frac{1}{2\pi} 2\pi\delta \frac{\sup_{z \in B_\delta(0)} |\phi(z)|}{\delta^2} \\ &= \frac{1}{\delta} \sup_{z \in B_\delta(0)} \phi(z). \end{aligned}$$

Looking at the right hand side, we have

$$\begin{aligned} \frac{1}{\delta} \sup_{\lambda \in B_\delta(0)} |\phi(\lambda)| &= \frac{1}{\delta} \sup_{\lambda \in B_\delta(0)} |\langle F(u+\lambda x), y \rangle_Y| \\ &= \frac{1}{\delta} \sup_{u' \in O^{\mathbb{C}}+\delta} |\langle F(u'), y \rangle_Y| \\ &\leq \frac{1}{\delta} \sup_{u' \in O^{\mathbb{C}}+\delta} \|F(u')\|_Y \, \|y\|_Y \\ &= \frac{1}{\delta} \sup_{u' \in O^{\mathbb{C}}+\delta} \|F(u')\|_Y . \end{aligned}$$

One the other hand, by definition

$$\begin{aligned} \phi'(0) &= \lim_{h \to 0} \frac{\langle F(u+hx), y \rangle_Y - \langle F(u), y \rangle_Y}{h} \\ &= \lim_{h \to 0} \left\langle \frac{F(u+hx) - F(u)}{h}, y \right\rangle_Y \\ &= \langle F_*(u)(x), y \rangle_Y . \end{aligned}$$

Choosing

$$y = \frac{F_*(u')(x)}{\|F_*(u')(x)\|},$$

we get

$$\|F_*(u)(x)\|_Y \leq \frac{1}{\delta} \sup_{u' \in O^{\mathbb{C}}+\delta} \|F(u')\|_Y .$$

Taking the supremum over all x with norm 1 gives us the conclusion.

■

Theorem 2.6 (Inverse Function Theorem)
Let $F\colon O^{\mathbb{C}} \to Y^{\mathbb{C}}$ be an analytic map and assume there is an $x \in O^{\mathbb{C}}$ such that $F_*(x)$ is an isomorphism. Then F can be analytically inverted in a neighborhood of $F(x)$. ♠

2.3 Hilbert scales

Let X_0 be a Hilbert space with orthonormal basis $\{\phi_k\}_{k\in\tilde{Z}}$, where $\tilde{Z} \subset \mathbb{Z}^n$ has the property that $-\tilde{\mathbb{Z}} = \tilde{\mathbb{Z}}$.

Let $\{\theta_j\}_{j\in\tilde{Z}}$ be a positive sequence such that $\theta_j \to \infty$ and $\theta_{-j} = \theta_j$ for all j.

Definition 2.10
For $s \in \mathbb{R}$, let X_s be a Hilbert space with basis $\{\phi_k\theta_k^{-s}\}$. If $u, v \in X_s$, then we can write

$$u = u_k\phi_k\theta_k^{-s},$$
$$v = v_k\phi_k\theta_k^{-s},$$

and we define

$$\langle u, v\rangle_s := \sum_{k\in\mathbb{Z}} \overline{u}_k v_k.$$

♣

Now, let $u \in X_0$. We know that $u = \sum u_k\phi_k = \sum u_k\theta_k^s \left(\theta_k^{-s}\phi_k\right)$, so that

$$\|u\|_s := \langle u, u\rangle_s = \sum \left(u_k\theta_k^s\right)\overline{\left(u_k\theta_k^s\right)} = \sum |u_k|^2 \theta_k^{2s}.$$

Definition 2.11
The sequence of spaces $\{X_s\}_s$ is called the ***Hilbert scale*** *of X_0, and $\{\phi_k\}$ is called the* ***basis of the scale.*** ♣

For example, suppose that $u \in X_0$. Then we can write $u = \sum u_k\phi_k$. We can also write this as

$$u = \sum u_k\theta_k^s \left(\theta_k^{-s}\phi_k\right)$$

which is in X_s provided that

$$\|u\|_s = \sum_k |u_k|^2 \theta_k^{2s} < \infty.$$

Example: Let $X_0 = L^2[0, 2\pi]$, with $\phi_k e^{ikx}$ for $k \in \mathbb{Z}$. Let's choose $\theta_k = |k|$. Then X_m has basis $\{|k|^{-m} e^{ikx}\}$. Let f be m-times differentiable, i.e. $f^{(m)} \in L^2$. Then we can calculate that

$$f^{(m)} = \sum i^m k^m u_k e^{ikx}.$$

But since $f^{(m)} \in L^2$, we have that

$$\sum_k |k|^{2m} |u_k|^2 < \infty,$$

and thus $f \in X_m$. In this case, up to some more proof, we have that $X_s = H^s$, the s-th Sobolev space.

Proposition 2.7
For a Hilbert scale $\{X_s\}$, if $s > r$, then $X_s \subset\subset X_r$, and X_s is a dense subset of X_r. ♠

Proof: (We'll wait on the proof of compactly contained, but we'll prove that X_s is a dense subset of X_r.

1. $\boxed{X_s \subset X_r}$ Let's choose $u \in X_s$. Then writing u in this basis, we have

$$\begin{aligned} u &= \sum u_k(\theta_k^{-s}\phi_k) \\ &= \sum u_k\theta_k^{-s}\theta_k^{r}(\theta_k^{-r}\phi_k). \end{aligned}$$

So we have that

$$u \in X_r \Leftrightarrow \sum |u_k|^2 \theta_k^{2(r-s)} < \infty.$$

Now, we know that $\sum |u_k|^2 < \infty$, since $u \in X_s$, and that

$$\theta_k^{2(r-s)} \to 0 \text{ as } k \to \infty.$$

2. $\boxed{X_s \text{ is dense in } X_r}$
Choose $v \in X_r$. For any $\epsilon > 0$, we'll find $u \in X_s$ with

$$\|u - v\|_r < \epsilon.$$

We can write $v = \sum_k v_k\theta_k^{-r}\phi_k$, and thus $\|v\|_r^2 = \sum |v_k|^2 < \infty$. Since this is convergent series, there is a K so that

$$\sum_{k>K} |v_k|^2 < \epsilon^2.$$

Choose

$$u = \sum_{k\leq K} v_k\theta_k^{s-r}\theta_k^{-s}\phi_k.$$

We clearly have that $u \in X_s$, since this is a finite sum. Then we have that

$$v - u = \sum_{k>K} v_k\theta_k^{-r}\phi_k,$$

and

$$\|v - u\|_r^2 = \sum_{k<K} |v_k|^2 < \epsilon^2.$$

3. $\boxed{\text{compact}}$
We won't prove it here, but for the record, if we were to show compact containment, we would have to show one of the following (which are equivalent):

   (a) Every bounded sequence in X_s is precompact when considered as a sequence in X_r.

   (b) The identity embedding $i_{s,r}\colon X_s \to X_r$ (which is bounded by part 1 above) is a compact operator, i.e. sends closed and bounded sets of X_s to compact sets of X_r.

■

Definition 2.12
We define $X_\infty = \cap_s X_s$ *and* $X_{-\infty} = \cup_s X_s$. ♣

Proposition 2.8
The spaces X_s and X_{-s} are conjugate with respect to the X_0 inner product. ♠

Proof: Forgetaboutit

■

Proposition 2.9
If $-\infty < a < b < \infty$, with $0 \le \theta \le 1$, and $c = (1-\theta)a + \theta b$, then X_c interpolates X_a and X_b in the sense that for all $u \in X_b \subset X_c \subset X_a$,

$$\|u\|_c \le \|u\|_a^{1-\theta} \|u\|_b^{\theta}.$$ ♠

Proof: This will follow from a rarely seen formulation of the Hölder inequality, i.e.

$$\sum_k a_k^p b_k^q \le \left(\sum a_k\right)^p \left(\sum b_k\right)^q.$$

So let's write $u = \sum u_k \phi_k$. Then

$$\begin{aligned}
\|u\|_c &= \sum_k |u_k|^2 \theta_k^{2c} \\
&= \sum_k |u_k|^2 \theta_k^{2(1-\theta)a} \theta_k^{2\theta b} \\
&= \sum_k |u_k|^{2(1-\theta)} \theta_k^{2(1-\theta)a} |u_k|^{2\theta} \theta_k^{2\theta b} \\
&\le \left(\sum |u_k|^2 \theta_k^{2a}\right)^{1-\theta} \left(\sum_k |u_k|^2 \theta_k^{2b}\right)^{\theta} \\
&= \|u\|_a^{1-\theta} \|u\|_b^{\theta}.
\end{aligned}$$

■

Now we'll prove the last part of the above theorem, that X_s is compactly contained in X_r for $s > r$.

Proof: What we will do is start with a bounded sequence in X_s and show that it is precompact in X_r. So assume that we have $\{a^j\}$ with

$$\left\|a^j\right\|_s \le B.$$

Let A be the set which is the sequence $\{a^j\}$. Let's assume that $\overline{A}$ is not compact in the X_r, which means that there is a sequence $\{b^j\} \subset \overline{A}$ which has no convergent subsequence. Let's choose $\{c^j\} \subset \{b^j\}$ so that

$$\left\|c^{j'} - c^j\right\|_r \ge 4\epsilon$$

for some $\epsilon > 0$, and for all j, j'. We know that we can do this, or $\{c^j\}$ would have a convergent subsequence.

Since the c^j are limit points of A, we know that for all j, we can find $d^j \in A$ with $\left\|c^j - d^j\right\|_r < \epsilon$. Then we have that

$$\left\|d^{j'} - d^j\right\| \ge 2\epsilon.$$

Since we can write

$$a^j = \sum a_k^j \theta_k^{-s} \phi_k = \sum a_k^j \theta_k^{r-s} \theta_k^{-r} \phi_k,$$

we know that

$$\|a^j\|_r^2 = \sum_k \left|a_k^j\right|^2 \theta_k^{2(r-s)}.$$

Since $r < s$, we know that $\theta_k^{2(r-s)} \to 0$ as $|k| \to \infty$. Then we can conclude that there is a K, so that for $|k| > K$, we have

$$\theta_k^{2(r-s)} \leq \frac{\epsilon^2}{4B^2}.$$

Thus we have

$$\begin{aligned}
\|a^j\|_r &\leq \sum_{|k|<K} \left|a_k^j\right|^2 \theta_k^{2(r-s)} + \sum_{|k|\geq K} \left|a_k^j\right|^2 \theta_k^{2(r-s)} \\
&\leq \sum_{|k|<K} \left|a_k^j\right|^2 \theta_k^{2(r-s)} + \sum_{|k|\geq K} \left|a_k^j\right|^2 \frac{\epsilon^2}{4B^2} \\
&\leq \sum_{|k|<K} \left|a_k^j\right|^2 \theta_k^{2(r-s)} + \frac{\epsilon^2}{4}.
\end{aligned}$$

Writing each d^j as $d^j_{\leq} + d^j_{>}$ in the obvious way, we can conclude that

$$\left\|d^j_{>}\right\|_r \leq \frac{\epsilon}{2}.$$

Putting it all together, we have

$$2\epsilon \leq \left\|d^{j'} - d^j\right\|_r \leq \left\|d^{j'}_{\leq} - d^j_{\leq}\right\|_r + \left\|d^{j'}_{>} - d^j_{>}\right\|_r,$$

and thus

$$\left\|d^{j'}_{\leq} - d^j_{\leq}\right\|_r \geq \epsilon.$$

What we have now is an infinite sequence which is contained in a bounded set, yet it's elements are separated by a constant amount. Although this can happen in infinite dimensions, it can't in finite dimensions. More specifically, we can consider the d^j's lying in $\mathbb{C}^K$, using the obvious map. Since all norms are equivalent in $\mathbb{C}^K$, we have a bounded infinite sequence, no two elements of which are close. A little measure theory or a picture will show that this can't happen.

■

The last result of this section is a theorem which tells us how to compare two differently constructed Hilbert scales.

Let's consider a Hilbert scale X_0 with basis $\{\phi_k\}$. Choose

$$\begin{aligned}
\theta_k &= c\,|k|^m + o(|k|^m), \\
\tilde{\theta}_k &= |k| + 1.
\end{aligned}$$

Let $\{X_s\}$ be the Hilbert scale formed from using the weights θ_k, and $\{\tilde{X}_s\}$ be the one formed from $\tilde{\theta}_k$.

Theorem 2.10 (Christmas theorem)
For all $s \in \mathbb{R}$, $X_s \approx \tilde{X}_{ms}$, in the sense that the identity map from one space to the other is one-to-one, onto, and bounded.

♠

Proof: We choose $u \in X_s$. Then $u = \sum u_k \theta_k^{-s} \phi_k$. We want to show that $u \in \tilde{X}_{ms}$. We can write

$$u = \sum u_k \theta_k^{-s} \tilde{\theta}_k^{ms} \tilde{\theta}_k^{-ms} \phi_k,$$

giving

$$\|u\|_{\tilde{X}_{ms}}^2 = \sum_k |u_k|^2 \left(\theta_k^{-s} \tilde{\theta}_k^{ms}\right)^2.$$

But we know that the sequence

$$\theta_k^{-1} \tilde{\theta}_k^m = \frac{(|k|+1)^m}{c\,|k|^m + o(|k|^m)} \to \frac{1}{c},$$

and we can conclude that $u \in \tilde{X}_{ms}$.

■

As an example as to why we would want the previous theorem, consider the PDE

$$iu_t = \frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right),$$

with $a(x) \geq a_0 > 0$, defined for $0 < x < 2\pi$ and with periodic boundary conditions (we enforce $u(x,t) = u(x+2\pi,t)$). Then we can define the Hamiltonian for the system for be

$$H := \int_0^{2\pi} a(x) \left|\frac{\partial u}{\partial x}\right|^2 dx,$$

but enough of that for now.

Consider the eigenvalue problem

$$\frac{\partial}{\partial x}\left(a(x)\frac{\partial \phi_k}{\partial x}\right) = \theta_k \phi_k.$$

Then Sturm-Liouville theory gives us that $\theta_k = c\,|k|^2 + o(|k|^2)$. If we expand u as

$$u(x,t) = \sum a_k(t)\phi_k(x),$$

then

$$H = \sum_k \theta_k^2 \,|a_k|^2 \approx \|\cdot\|_{X_1} \approx \|\cdot\|_{\tilde{X}_2},$$

so that H is equivalent to the standard Sobolev H^2 norm.

2.4 Linear maps between scales

Consider two Hilbert scales $\{X_s\}, \{Y_s\}$ and a linear map

$$L \colon X_\infty \to Y_{-\infty}.$$

We will have a string of definitions following:

Definition 2.13
*We define the **operator norm** of L from X_{s_1} to Y_{s_2} as*

$$\|L\|_{s_1,s_2} = \sup_{\substack{x \in X_{s_1} \\ x \neq 0}} \frac{\|L(x)\|_{Y_{s_2}}}{\|x\|_{X_{s_1}}}.$$

♣

Definition 2.14
*L is a **morphism of order** d between $\{X_s\}$ and $\{Y_s\}$ for $s \in [s_0, s_1]$ if*

$$\|L\|_{s,s-d} < \infty \text{ for } s \in [s_0, s_1].$$

*If L^{-1} exists and L^{-1} defines a morphism of order $-d$ from $\{Y_s\}$ to $\{X_s\}$ for $s \in [s_0 - d, s_1 - d]$, then L is an **isomorphism of order** d. If $\{X_s\} = \{Y_s\}$, then L is an **automorphism of order** d. If the order of L is $-\Delta < 0$, then we say that L is a **Δ-smoothing morphism**.*

♣

Definition 2.15
Let $L\colon X_s \to Y_{s-d}$ be of order d for $s \in [s_0, s_1]$. Then

$$\begin{aligned} L^*: Y^*_{s-d} &\to X^*_s \\ Y_{d-s} &\to X_{-s} \\ \phi &\mapsto \phi \circ L, \end{aligned}$$

*is of order d for $s \in [-s_1 + d, -s_0 + d]$. We call L^* the **adjoint** of L.*

♣

Definition 2.16
If $X_s = Y_s$, and we have that

$$\langle Lu, v\rangle = \langle u, Lv\rangle \text{ for all } s \in [s_0, s_1],$$

*then we say that L is **symmetric**. If $s \in [s_0, d - s_0]$, then we say that L is **self-adjoint**.*

♣

Theorem 2.11 (Interpolation)
If $\{X_s\}$ and $\{Y_s\}$ are real Hilbert scales, and $L\colon X_\infty \to Y_{-\infty}$ is a linear map such that

$$\begin{aligned} \|L\|_{a_1,b_1} &= c_1, \\ \|L\|_{a_2,b_2} &= c_2, \end{aligned}$$

then for all $\theta \in [0, 1]$, we have that

$$\|L\|_{a,b} \le c_\theta = c_1^\theta c_2^{1-\theta},$$

where

$$\begin{aligned} a &= \theta a_1 + (1-\theta)a_2, \\ b &= \theta b_1 + (1-\theta)b_2. \end{aligned}$$

♠

Also, more generally, if we have $L = L_u$, where L_u depends analytically on $u \in H$, H A complex Hilbert space, and the maps

$$\begin{aligned} L_u\colon X_{a_1} &\to X_{b_1} \\ L_u\colon X_{a_2} &\to X_{b_2} \end{aligned}$$

are bounded uniformly in norm, then for all $\theta \in [0, 1]$, L_u depends analytically on u as an operator $X_{a_\theta} \to X_{b_\theta}$.

2.5 Differential Forms

Let $d > 0$ and X_d be a Hilbert space. For O and open subset of X_d, the tangent space to O at $\mathfrak{x}$, $T_{\mathfrak{x}}O$, is X_d itself.

Definition 2.17
*A **differential** k**-form** on O is a continuous function*

$$O \times X_d \times \cdots \times X_d \to \mathbb{R}$$

which is polylinear and antisymmetric in the last k entries. ♣

Definition 2.18
A 1-form is denoted $a(\mathfrak{x})\, d\mathfrak{x}$, $a\colon O \to X_{-d}$, and we define

$$a(\mathfrak{x})\, d\mathfrak{x}(\xi) = \langle a(\mathfrak{x}), \xi \rangle_0 \,, \text{ for } \xi \in X_{-d}.$$ ♣

Definition 2.19
A 2-form is denoted $A(\mathfrak{x})\, d\mathfrak{x} \wedge d\mathfrak{x}$, where

$$A(\mathfrak{x})\, d\mathfrak{x} \wedge d\mathfrak{x}[\xi, \eta] = \langle A(\mathfrak{x})\xi, \eta \rangle_0 \,, \text{ for } \xi, \eta \in X_d,$$

as long as $A(\mathfrak{x})\colon X_d \to X_{-d}$ is bounded and anti-self-adjoint. ♣

For example, Let $X_0 = \mathbb{R}^n$, and let A be an $n \times n$ anti-symmetric matrix, then $A(x)\, dx \wedge dx = -\sum_{i<j} A_{ij}\, dx_i \wedge dx_j$.

Definition 2.20
*A k-form ω_k on $O \subset X_d$ is **analytic** if the corresponding map from O to the linear space of skew-symmetric polylinear functions is analytic, i.e. depends analytically on the base point.* ♣

Theorem 2.12 (Cartan's Formula)

$$d\omega(\mathfrak{x})[\xi_1, \ldots, \xi_{k+1}] = \sum_{i=1}^{k+1} \frac{\partial}{\partial \xi_i} \omega_k(\mathfrak{x})[\xi_1, \ldots, \hat{\xi}_i, \ldots, \xi_{k+1}].$$ ♠

Example:

1. Let f be a $\mathcal{C}^1$ function on O. Then $df = \nabla f(\mathfrak{x})\, d\mathfrak{x}$.
2. Let ω be a 1-form on O, $\omega = a(\mathfrak{x})\, d\mathfrak{x}$. Then $d(a(\mathfrak{x})\, d\mathfrak{x}) = (a_* - a^*)\, d\mathfrak{x} \wedge d\mathfrak{x}$.

2.6 Flows

Let ω_t be any $\mathcal{C}^1$-smooth closed k-form on $O \subset X_d$ which depends $\mathcal{C}^2$-smoothly on $t \in [0,1]$. Let $V(\mathfrak{x}, t)$ be a nonautonomous $\mathcal{C}^1$-smooth Lipschitz vector field on O. Consider

$$\dot{\mathfrak{x}} = V(\mathfrak{x}, t), \mathfrak{x} \in O,$$

and denote the flow maps $S^t_{t_0}$, where $S^t_{t_0}(\mathfrak{x}(t_o)) = \mathfrak{x}(t)$. Let $Q \subset O$ such that $S^t_0 Q \subset O$ for $t \in [0,1]$ and write ϕ^t for S^t_0.

Theorem 2.13 (Cartan's Identity)

$$\begin{aligned}\frac{d}{dt}\left(\phi^t\right)^*\omega_t &= \phi^{t*}\frac{\partial\omega_t}{\partial t} + d\phi^{t*}\left(V\rfloor_{\omega_t}\right)\\ &= \phi^{t*}\left(\frac{\partial\omega_t}{\partial t} + D\left(V\rfloor_{\omega_t}\right)\right)\end{aligned}$$

everywhere in Q, where

$$V\rfloor_{\omega_t}(\xi_2,\ldots,\xi_k) = \omega_t(V,\xi_2,\ldots,\xi_k).$$

♠

Proof: We'll prove it in the case where $k=0$. In that case, we have $\omega_t = \omega(x,t)$, so that $\phi^{t*}\omega_t = \omega(\phi^t(x),t)$ and

$$\begin{aligned}\frac{d}{dt}\phi^{t*}\omega_t &= \frac{d}{dt}\omega(\phi^t(x),t)\\ &= \frac{\partial\omega}{\partial t}\left(\phi^t(x),t\right) + \frac{\partial\omega}{\partial x}\left(\phi^t(x),t\right)\cdot\frac{\partial\phi^t}{\partial t}\\ &= \left(\phi^t\right)^*\frac{\partial\omega}{\partial t} + \frac{\partial\omega}{\partial x}(\phi^t,t)\cdot V(\phi^t(x),t)\\ &= \left(\phi^t\right)^*\frac{\partial\omega}{\partial t} + \left(\phi^t\right)^*\left(\frac{\partial\omega}{\partial t}(x,t)V(\phi^t(x),t)\right).\end{aligned}$$

■

2.7 Symplectic Structures and Hamiltonian Equations

Let $O\subset X_d$ for $d\geq 0$ and $\alpha_2 = \overline{J}\,d\mathfrak{x}\wedge d\mathfrak{x}$ be a closed 2-form. We define (O,α_2) to be a **symplectic manifold** and $(\{X_s\},\alpha_2)$ to be a **symplectic Hilbert scale**. Recall that symplectic forms define an isomorphism between vector fields and 1-forms. If we let h be a $\mathcal{C}^1$ function $h\colon O_d\to\mathbb{R}$, then we can define a Hamiltonian vector field by

$$\alpha_2(V_h,\xi) = -dh(\xi)$$

for all $\xi\in TO_d$.

There is a Riesz-like theorem for 2-forms which gives us that

$$\left\langle\overline{J}(\mathfrak{x})V_h,\xi\right\rangle = -\left\langle\nabla h(\mathfrak{x}),\xi\right\rangle$$

implies

$$V_h = J(\mathfrak{x})\nabla h(\mathfrak{x}),$$

where $J = -\overline{J}^{-1}$. We call J the **operator of the Poisson structure** and $\overline{J}$ the **operator of the symplectic structure**. Also, we usually will assume that V_h is at least $\mathcal{C}^1$.

If h is $\mathcal{C}^1$-smooth on $O_d\times\mathbb{R}$, the Hamiltonian equation is

$$\dot{\mathfrak{x}}(t) = J(\mathfrak{x})\nabla h(\mathfrak{x},t). \tag{2.2}$$

If the order of V_h is 0, i.e. $V_h : X_d\to X_d$, as V_h is $\mathcal{C}^1$-smooth and Lipschitz, then (2.2) is **well-posed**. This will mean that for any $\mathfrak{x}(0)\in O_d$, there is a unique solution which depends smoothly on initial conditions.

Definition 2.21
A PDE with boundary conditions is called a ***Hamiltonian PDE*** *if for a suitable symplectic Hilbert scale, and a suitable Hamiltonian, the PDE can be written in the form (2.2).*

♣

Note that if V_h is unbounded, or $\operatorname{ord} V_h = d_1 > 0$, then

$$V_h \colon O_d \times \mathbb{R} \to X_{d-d_1}.$$

Definition 2.22
*A continuous curve $\mathfrak{x}\colon [0,T] \to O_d$ is a **solution of the Hamiltonian equation** (2.2) in X_d if it defines a $\mathcal{C}^1$ map $: [0,T] \to X_{d-d_1}$ and both sides of (2.2) coincide as curves in X_{d-d_1}. If $\mathfrak{x}(t), t \geq \tau$ exists and is unique, then we write*

$$\mathfrak{x}(t) = S_\tau^t \mathfrak{x}(\tau),$$

or

$$\mathfrak{x}(t) = S^{t-\tau} r(\tau) \text{ if equation is autonomous,}$$

*and S is called the **flow map**.* ♣

We should be aware that for Hamiltonian equations with $\operatorname{ord} V_h > 0$, there are **no** general existence theorems for flow maps.

Example: [Semi-linear Equation] Let

$$\dot{\mathfrak{x}}(t) = V(\mathfrak{x}),$$

where $V = B + V^0$, with B linear (possible with positive order) and V^0 nonlinear, but $\operatorname{ord} V^0 = 0$, and V^0 Lipschitz bounded on bounded sets.

For a specific example, consider the nonlinear Schrödinger equation

$$\dot{u}(x,t) = i(\nabla u + f(|u|^2)u), \qquad x \in \mathbb{T}^n.$$

Proposition 2.14
If (2.2) is semi-linear, and $\operatorname{ord}(V^0) = 0$, then for any c, the flow maps $S^t\colon O_c \to X_d$ are well defined for $|t| < T$, where $T = T(c) > 0$. If V^0 is $\mathcal{C}^1$-smooth, then S^t is $\mathcal{C}^1$. ♠

Example: [Nonlinear string equation] We consider the equation

$$u_{tt} = u_{xx} + \frac{\partial}{\partial x} f(u_x),$$

where

$$f(v) = C + av^2 + \dots, \qquad a \neq 0,$$

and with boundary conditions

$$u(x,t) = u(x + 2\pi, t), \qquad \int_0^\pi u(x,t)\, dx = 0.$$

We rewrite it as a system

$$\begin{aligned} \dot{u} &= v \\ \dot{v} &= u_{xx} + \frac{\partial}{\partial x}\left(f\left(\frac{\partial u}{\partial x}\right)\right). \end{aligned}$$

Let $w = \begin{pmatrix} u \\ v \end{pmatrix}$. Then we can rewrite the equation as the system

$$\dot{w} = Aw + F(w),$$

where

$$A = \begin{pmatrix} 0 & 1 \\ \partial_{xx} & 0 \end{pmatrix},$$

$$F(w) = \begin{pmatrix} 0 \\ \dfrac{\partial}{\partial x} f(u_x) \end{pmatrix}.$$

We want to write the equation as

$$\dot{w} = J \nabla h.$$

Our Hilbert scale will be $Z_s = H_0^s \times H_0^s$, where H_0^s is the Sobolev space of functions which have s derivatives in L^2, and mean value 0. We let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Then we let

$$h = \int_0^{2\pi} \left\{ \frac{1}{2} |v|^2 + \frac{1}{2} |u_x|^2 + f(u_x) \right\} dx.$$

So what is dh? We check:

$$\begin{aligned} dh(u,v)\xi &= \int_0^{2\pi} \{ v\xi_2 + u_x \xi_1' + f'(u_x)\xi_1' \} \, dx \\ &= u_x \xi_1 |_0^{2\pi} + f'(u_x)\xi_1 |_0^{2\pi} + \int_0^{2\pi} \left\{ v\xi_2 - u_{xx}\xi_1 - \frac{\partial}{\partial x} f(u_x)\xi_1 \right\} dx \\ &= \langle (-u_{xx} - f'(u_x), v), \xi \rangle_{L^2}, \end{aligned}$$

since the boundary terms drop out because we're in H_0^s.

Definition 2.23
Quasi-linear Hamiltonian PDE *are equations of the form*

$$\dot{u} = Au + F(u),$$

and $\operatorname{ord} A > \operatorname{ord} F$.

♣

Example: [Equations of KdV Type] Let $X_s = H_0^s$, and $J = \dfrac{\partial}{\partial x}$, so that $\operatorname{ord} J = 1$. We let $\overline{J} = -J^{-1}$. Then our symplectic scale is $(X_s, \overline{J}\, du \wedge du) = (H_0^s, -\left(\dfrac{\partial}{\partial x}\right)^{-1} du \wedge du)$.

Let us consider

$$h(u) = \int_0^{2\pi} \left(-\frac{1}{8}(u')^2 + f(u) \right) dx.$$

Then

$$\begin{aligned} dh(u)v &= \int_0^{2\pi} \left(-\frac{1}{4} u'v' + f'(u)v \right) dx \\ &= -\frac{1}{4} u'v' |_0^{2\pi} + \int_0^{2\pi} \left\{ \frac{1}{4} u''v + f'(u)v \right\} dx \\ &= \left\langle \frac{1}{4} u'' + f'(u), v \right\rangle_{L^2}, \end{aligned}$$

since the boundary term again drops out.

Then we have $\nabla h(u) = \frac{1}{4}u'' + f'(u)$, which gives

$$\begin{aligned}\dot{u} = J\nabla h &= \frac{\partial}{\partial x}\left(\frac{1}{4}u'' + f'(u)\right)\\ &= \frac{1}{4}u''' + \frac{\partial}{\partial x}f'(u).\end{aligned}$$

The above is an equation of **KdV** type. If we choose $f = \frac{1}{4}u^3$, then we get the classical KdV. A non-trivial result is that if we choose initial conditions in H_0^s with $s \geq 3$, then there exists a unique solution for all time.

2.8 Symplectic Transformations

Let $\{X_s\}, \{Y_s\}$ be Hilbert scales, with $O \subset X_d, Q \subset Y_{\tilde{d}}$, with $d, \tilde{d} > 0$. Let us consider $\alpha_2 = \overline{J}\, d\mathfrak{x} \wedge d\mathfrak{x}$, symplectic on O, and $\beta_2 = \overline{\Upsilon}\, d\mathfrak{y} \wedge d\mathfrak{y}$, symplectic on Q.

Definition 2.24
*A $\mathcal{C}^1$-smooth map $\Phi\colon Q \to O$ is called a **symplectic map** if*

$$\Phi^*\alpha_2 = \beta_2,$$

i.e. for any $\mathfrak{y} \in Q$ with $\Phi(\mathfrak{y}) = \mathfrak{x}$,

$$\left\langle \overline{J}\Phi_*(\mathfrak{y})\xi, \Phi_*(\mathfrak{y})\eta \right\rangle_{X_0} = \left\langle \overline{\Upsilon}(\mathfrak{y})\xi, \eta \right\rangle_{Y_0},$$

for all $\xi, \eta \in Y_{\tilde{d}}$, or

$$\Phi^*(\mathfrak{y}) \circ \overline{J}(\mathfrak{x}) \circ \Phi_*(\mathfrak{y}) = \overline{\Upsilon}(\mathfrak{y}).$$

*If $\Phi_*(\mathfrak{y})$ defines isomorphisms of $Y_{\tilde{d}}, X_d$, then we say that Φ is a **symplectomorphism**.* ♣

For right now, we will restrict to the case where $d = \tilde{d} \geq 0$, such that $\operatorname{ord}\overline{J} = \operatorname{ord}\overline{\Upsilon} = -d_J$.

Consider the following hypothesis on Φ:

(C1): Let $\Phi\colon Q \to O$ be a $\mathcal{C}^1$-smooth symplectomorphism such that $\Phi_*(\mathfrak{y})\colon Y_s \to X_s$ is a linear map continuous in $\mathfrak{y} \in Q$, for all $|s| < d$.

Incidentally, it can be shown that if $\overline{J}, \overline{\Upsilon}$ are constant isomorphisms on order 0, then **(C1)** is automatically satisfied.

Theorem 2.15
Let $O \subset X_d, Q \subset Y_d$, $d \geq 0$, be given symplectic structures by α_2, β_2 and assume that $\operatorname{ord}\overline{J} = \operatorname{ord}\overline{\Upsilon} = -d_J$. Define the vector field $V_h = J\nabla h$, and assume that it defines a $\mathcal{C}^1$-smooth map

$$V_h\colon O \times \mathbb{R} \to X_{d-d_1},$$

of order $d_1 \leq 2d$. Consider the differential equation

$$\dot{\mathfrak{x}}(t) = V_h(\mathfrak{x}, t).$$

Let $\Phi\colon Q \to O$ be a symplectic map satisfying **(C1)** such that V_h in O is tangent to $\Phi(Q)$ in the following sense:

$$V_h(\Phi(\mathfrak{y})) = \Phi_*(\mathfrak{y})\xi, \text{ for all } \mathfrak{y} \in Q, \text{ for some } \xi = \xi(\mathfrak{y}) \in Y_{d-d_1}. \qquad (\textbf{(C2)})$$

Then the map Φ transforms solutions of the equation

$$\dot{\mathfrak{y}} = \overline{\Upsilon}(\mathfrak{y})\nabla_{\mathfrak{y}}H(\mathfrak{y}, t), \qquad H = h \circ \Phi, \tag{2.3}$$

to solutions of

$$\dot{\mathfrak{x}}(t) = J\nabla h = V_h(\mathfrak{x}, t).$$ ♠

Note: Let's consider this **(C2)**. We know that $V_h\colon O\times\mathbb{R}\to X_{d-d_1}$, be definition. Since we know that $\Phi(\mathfrak{y})\in O$, we have that $V_h(\Phi(\mathfrak{y}))\in X_{d-d_1}$.

On the other hand, since $\Phi\colon Q\to O$, we know that $\Phi_*\colon Y_d\to X_d$. To make **(C2)** true, we would want that Φ_* maps from Y_{d-d_1} to X_{d-d_1}. This is not going to be true in general. But if we assume **(C1)**, we will have it. Since $d_1\le 2d$, this means that $|d-d_1|<d$, which, with **(C1)**, implies that $\Phi_*(\mathfrak{y})\colon Y_{d-d_1}\to X_{d-d_1}$, and, furthermore, since $\Phi_*(\mathfrak{y})$ is an invertible map on these spaces (being a symplectomorphism), we can solve the equation for ξ.

Proof: Let $\mathfrak{y}(t)$ be a solution of (2.3). By **(C1)** the curve $\mathfrak{x}(t)=\Phi(\mathfrak{y}(t))$ is $\mathcal{C}^1$-smooth in Y_{d-d_1}, and continuous in Y_d. Since

$$\dot{\mathfrak{x}}(t)=\Phi_*(\mathfrak{y})\dot{\mathfrak{y}},$$

and

$$\nabla_{\mathfrak{y}}H=\Phi^*(\mathfrak{y})\nabla_{\mathfrak{x}}h,$$

we calculate that

$$\begin{aligned}\dot{\mathfrak{x}}&=\Phi_*(\mathfrak{y})\Upsilon(\mathfrak{y})\Phi^*(\mathfrak{y})\nabla_{\mathfrak{x}}h\\&=-\Phi_*(\mathfrak{y})\Upsilon(\mathfrak{y})\Phi^*(\mathfrak{y})\overline{J}\Phi_X(\mathfrak{y})\xi\\&=-\Phi_*(\mathfrak{y})\Upsilon(\mathfrak{y})\overline{\Upsilon}(\mathfrak{y})\xi\\&=\Phi_*(\mathfrak{y})\xi\\&=V_h(\mathfrak{x}).\end{aligned}$$

■

Corollary 2.16
Let $O\subset(X_d,\alpha_2)$ and the Hamiltonian vector field V_h be as in the theorem, and let $\Phi\colon O\to X_d$ be $\mathcal{C}^1$-smooth satisfying **(C1)** such that

$$\Phi^*\alpha_2=K\alpha_2,$$

for some $K\neq 0$. Then Φ transforms solutions of

$$\dot{\mathfrak{x}}(t)=K^{-1}J(\mathfrak{x})\nabla H(\mathfrak{x},t)$$

to

$$\dot{\mathfrak{x}}(t)=V_h(\mathfrak{x},t).$$

♠

Recall that a useful way to construct symplectic maps of domains $O\subset(X_d,\alpha_2)$ is to get them as flow maps S^t_τ of additional nonautonomous Hamiltonian equations

$$\dot{\mathfrak{x}}(t)=J\nabla f(\mathfrak{x},t)=V_f(\mathfrak{x},t),$$

where f is such that V_f is Lipschitz.

Theorem 2.17
Let f be a $\mathcal{C}^1$-smooth function on $\mathbb{R}\times O$ such that the map

$$V_f\colon\mathbb{R}\times O\to X_d$$

is Lipschitz in $\mathfrak{x},t$ and $\mathcal{C}^1$-smooth in $\mathfrak{x}$. Let $O^1\subset O$. Then the flow maps

$$S^\tau_t\colon(O^1,\alpha_2)\to(O,\alpha_2)$$

are symplectomorphisms provided that they map O^1 to O. Moreover,

$$\|(S_t^\tau)_*(\mathfrak{x})\|_{d,d} \leq \exp(|t-\tau|\, C_*),$$

where

$$C_* = \sup_{t,\mathfrak{x}} \left\|(V_f)_*(\mathfrak{x},t)\right\|_{d,d}.$$

If V_f are analytic, then so are the S_t^τ. ♠

Consider the constant coefficient form $\alpha_2 = \overline{J}\, d\mathfrak{x} \wedge d\mathfrak{x}$ where $\operatorname{ord} J = d_J$. (Constant coefficient means $\overline{J}$ doesn't depend on $\mathfrak{x}$!) For any $\mathfrak{y}$, the maps $(S_t^\tau)_*(\mathfrak{y})$ define zero-order morphisms of the scale for $s \in [-d-d_J, d]$. Let's further assume that V_f is Δ-smoothing, and

$$\|(V_f)_*(\mathfrak{x},t)\| \leq C_*.$$

Since

$$S_t^\tau(\mathfrak{x}) = \mathfrak{x} + \int_t^\tau V_f(S_t^\theta(\mathfrak{x}),\theta)\, d\theta,$$

then

$$(S_t^\tau)_* = \mathbf{1} + \int_t^\tau (V_f)_*\,(S_t^\theta(\mathfrak{x}),\theta)\cdot(S_t^\theta)_*(\mathfrak{x})\, d\theta.$$

(We can take the derivative this way since the symplectic form is constant coefficient.)

Since the $(V_f)_*$ are Δ-smoothing, and $(S_t^\tau)_*$ satisfy the estimate above, the flow maps are symplectomorphisms, and are close to the identity up to Δ-smoothing maps, i.e.

$$\|(S_t^\tau)_* - \mathbf{1}\|_{s,s+\Delta} \leq C_*\,|t-\tau|\, e^{C_*|t-\tau|},$$

for $s \in [d-\Delta-d_J, d+\Delta]$.

2.9 Linearizing

Definition 2.25 (Linearized Equation)
Let $\mathfrak{x}(t), t \in \mathbb{R}$ be a solution for $\dot{\mathfrak{x}} = V_h$. If for each $\zeta \in X_d$ and each θ the linearized equation

$$\dot{\zeta}(t) = V_{h*}(\mathfrak{x}(t),t)\zeta(t), \qquad \zeta(\theta) = \zeta$$

has a unique solution $\zeta(t) \in X_d$ defined for all t, and such that $\|\zeta(t)\|_d \leq C\,\|\zeta\|$, uniformly in θ, t (for θ, t lying in a compact region), then we write

$$\zeta(t) = S_{\theta**}^t(\mathfrak{x})\zeta$$

and say the flow $\{S_{\theta**}^t(\mathfrak{x})\}$ of the linearized equation is well-defined on X_d. ♣

Note that when the S_τ^t are $\mathcal{C}^1$-smooth, we have that $S_{t\,*}^\tau = S_{t\,**}^\tau$.

Theorem 2.18
Assume that the Hamiltonian vector field V_f defines a $\mathcal{C}^1$-smooth map on $\mathbb{R}\times O \to X_{d-d_1}$ and $d_1 - d_J \leq 2d$. Let a point $\mathfrak{x}_0$ be in O such that the solution

$$\mathfrak{x}(t) = S_{t_0}^t(\mathfrak{x}_0)$$

exists for all $t \in [t_o, T]$ and, for these t's, the flow maps $S_{t_0**}^t(\mathfrak{x}_0)$ for the linearized equation are well-defined in X_d. Then the maps $S_{t_0}^t$ are symplectic. ♠

NOTE: Does the first condition in the above theorem refer to the whole right hand side of the equation, or just the nonlinearity?

The idea of the proof is this: We check that $\alpha_2(\mathfrak{x}(t))[S^t_{t_0**}\xi, S^t_{t_0**}\eta] = \alpha_2(\mathfrak{x}(t))[\xi,\eta]$, or, equivalently, that

$$l(t) = \alpha_2(\mathfrak{x}(t))[\xi(t),\eta(t)]$$

should be independent of t, where ξ,η are solutions of the linearized equation. Thus we have to check that

$$\frac{d}{dt}\left\langle \overline{J}(\mathfrak{x}(t))\xi(t),\eta(t)\right\rangle = 0.$$

Corollary 2.19
If $\dot{\mathfrak{x}}(t) = V_f(\mathfrak{x},t)$ is semi-linear (i.e. $V_f = B + V^0$ and V^0 defines a $\mathcal{C}^1$-smooth map $\mathbb{R}\times O \to X_d$), then the flow maps $S^t_{t_0}$ are $\mathcal{C}^1$-smooth symplectomorphisms. ♠

Note: The previous note asks the question on whether or not there is any restriction on the order of the linear piece. We know that the nonlinear piece has order 0, is this enough?

Definition 2.26
*Consider $\mathcal{C}^1$ functions H_1,H_2 on O which have continuous gradient maps of orders $d_1,d_2 \le 2d$ such that $d_1+d_2+d_j \le 2d$. Then the **Poisson bracket** $\{H_1,H_2\}$ is defined to be the continuous function given as*

$$\left\langle J(\mathfrak{x})\nabla H_1(\mathfrak{x}),\nabla H_2(\mathfrak{x})\right\rangle.$$
♣

We should note that

- The Poisson bracket is skew-symmetric (since J is),
- $\{H,H\}=0$, as long as $\operatorname{ord}\nabla H \le d - d_J/2$.

Now, let $H_1,H_2,\overline{J}$ be γ-analytic on $O\subset X_d$ (i.e., extend to bounded analytic maps on $O+\gamma\in X_d^{\mathbb{C}}$. Let $\operatorname{ord}\nabla H_1 \le -d_J$. Consider $\dot{\mathfrak{x}} = J\nabla H_1(\mathfrak{x})$ in O, and let S^τ be its flow map.

Theorem 2.20
Take V_1 as above, and further assume that $\|V_1\|_d < K$. Then the maps $S^\tau\colon Q\to O$, for $\tau\in[0,\delta/K]$ are well-defined analytic symplectomorphisms, and

$$H_2(S^\tau(\mathfrak{x})) = H_2 + \tau\{H_1,H_2\} + \mathcal{O}((\tau K)^2), \qquad x\in Q,$$

or

$$\left.\frac{d}{dt}H_2(S^\tau(\mathfrak{x}))\right|_{t=0} = \{H_1,H_2\}(\mathfrak{x}).$$

In words, let's consider the flow under H_1, and then follow a trajectory of that system (a level curve of H_1). If we evaluate H_1 along this curve, it of course doesn't change. However, if we evaluate H_2 along this curve, the theorem tells us that, to first order, the change is $\{H_1,H_2\}$. ♠

Proof: The flow maps S^τ are well-defined for sufficiently small τ since V_1 is uniformly Lipschitz by the Cauchy estimate. In other words, since we know that V_1 is γ-analytic, and $\|V_1\|_d < K$, then Cauchy's Theorem tells us that $\sup_{\mathfrak{x}\in O} V_{1*} \le K/\gamma$. Let's choose a subregion Q of O, such that the minimum distance from Q to ∂O is δ. If $\mathfrak{x}\in Q$, then $\|S^\tau(\mathfrak{x})-\mathfrak{x}\|_d \le \tau K$, and thus S^τ stays in O for $\tau\in[0,\delta/K]$. The previous theorem says that since V_1 is uniformly Lipschitz, then S^τ is a symplectomorphism. Since

$$V_1(S^\tau(\mathfrak{x})) = V_1(\mathfrak{x}) + \mathcal{O}(\tau K^2),$$

we have that

$$\begin{aligned} S^\tau(\mathfrak{x}) &= \mathfrak{x} + \int_0^\tau V_1(S^\theta(\mathfrak{x}))\,d\theta \\ &= \mathfrak{x} + \tau V_1(\mathfrak{x}) + \mathcal{O}(\tau^2 K^2). \end{aligned}$$

Hence,

$$\begin{aligned} H_2(S^\tau(\mathfrak{x})) - H_2(\mathfrak{x}) &= \langle \nabla H_2(\mathfrak{x}), S^\tau(\mathfrak{x}) - \mathfrak{x}\rangle + \mathcal{O}(\|S^\tau(\mathfrak{x}) - \mathfrak{x}\|_d^2) \\ &= \tau\,\langle \nabla H_2, J\nabla H_1\rangle + \mathcal{O}(\tau^2 K^2). \end{aligned}$$

■

Example: [nonlinear Schrödinger, revisited] Consider the equation

$$\dot{u}(x,t) = i(-u_{xx} + P(|u|^2)u), \qquad x \in S^1,$$

where P is a real polynomial. Consider the Hilbert scale

$$Z_d = H^d(S^1;\mathbb{C}),$$

with the scalar product in Z_0 defined as

$$\langle u, v\rangle = \int_0^{2\pi} u\overline{v}\,dx.$$

Choose $\omega_2 = J\,du \wedge du$, where $Ju(x) = iu(x)$. Of course, the order of J is 0. Let our Hamiltonian be

$$h = \frac{1}{2}\int_0^{2\pi} \left(|u_x|^2 + Q(|u|^2)u\right)\,dx, \qquad Q' = P,$$

and we see that $\nabla h\colon Z_d \to Z_{d-2}$ (essentially because the right hand side of the equation has two x derivatives). There is a result that says that for all $d > 1/2$, the Sobolev space H^d is an algebra, i.e. the product of any two functions in H^d is again in H^d. (Note that this is not true for all d, most notably in the case where $d = 0$, for the product of two L^2 functions is not in L^2.) Since H^d is an algebra for $d > 1/2$, the operator

$$u \mapsto P(|u|^2)u$$

is of order 0 as long as $d > 1/2$.

Thus the NLS equation is semi-linear and well-defined in Z_d. Since $d_J = 0$, and the order of ∇h is 2, and h is continuous in Z_1, we can conclude that $h(S^t(u))$ is constant for $u \in Z_1$.

Note: The next comment will harken back to the question of whether the corollary about semi-linear equations stated above was true for a linear part of any order.

For $d \in (1/2, 1)$, the flow maps are continuous, but the Hamiltonian is not. It turns out, though, that if we define $h(u) = \infty$ for $u \in Z_d \setminus Z_1$, then we still have $H(S^t(u))$ is constant, even for the "orbits with infinite energy".

Chapter 3

Integrable Hamiltonian PDE

3.1 Integrable Subsystems of Hamiltonian Equations

3.1.1 Some examples

Consider a Hilbert scale $\{Z_s\}$ defined by $\{\theta_k |\ k \in \tilde{\mathbb{Z}} \subset \mathbb{Z}\}$, with $\theta_k > 0$, $\theta_k = c\,|k|^m + o(|k|^m)$. By Theorem 2.10, we know that $Z_s \equiv \tilde{Z}_{ms}$, where $\tilde{\theta}_k = (|k| + 1)$.

Let $\alpha_2 = \overline{J} dz \wedge dz$ supply the symplectic structure, where $\overline{J}$ is an anti-selfadjoint morphism of the Hilbert scale of the order $-d_J < 0$. Let $\mathcal{H} = 1/2\,\langle Az, z\rangle + H(z)$, where A is a self-adjoint morphism with order d_A. The symplectic structure corresponds to the Hamiltonian equation

$$\dot{u} = J\nabla\mathcal{H} = J(Au + \nabla H(u)) = V_{\mathcal{H}}(u).$$

Definition 3.1
*We say that $\mathcal{H}$ is **quasi-linear** if H is analytic on a domain $O_d \subset Z_d$, for $d > d_A/2$, and ∇H is analytic of order d_H, with $d_H < d_A$.* ♣

We state the following corollary which follows from interpolation:

Corollary 3.1
Let L be a bounded linear operator $L\colon X_a \to X_b$ that is symmetric (resp. anti-symmetric) in a real or complex Hilbert scale $\{X_s\}$. Then L extends to a selfadjoint (resp. anti-selfadjoint) morphism of the scale of order $a - b$, for $s \in [-b, a]$. ♠

So for any $u \in O_d \subset Z_d$, since $(\nabla H)_*\colon Z_d \to Z_{d-d_H}$ is a bounded linear operator, then $(\nabla H)_*$ extends to a selfadjoint morphism of the scale $\{Z_s\}$, of order d_H, for $s \in [-d - d_H, d]$, or

$$\begin{aligned}\nabla H_*\colon Z_s &\to Z_{s-d_H}\\ \|\nabla H_*\|_{s,s-d_H} &< \infty.\end{aligned}$$

We denote $d_1 = \operatorname{ord} V_{\mathcal{H}} = d_A + d_J \leq 2d + d_J$. We do not assume that the flow maps of the the equation are defined on the whole domain O_d.

Example: [KdV] Let $\{Z_s |\ s \in \mathbb{Z}\} = \{H_0^s(S^1, \mathbb{R})\}$. This the Sobolev scale of 2π-periodic functions with 0 mean value.

Let

$$J = \frac{\partial}{\partial x}, A = \frac{1}{4}\frac{\partial^2}{\partial x^2}, H(u) = \frac{1}{4}\int u^3\,dx.$$

Then we have that

$$(Z_s, \overline{J} du \wedge du) = \left(H_0^s, -\left(\frac{\partial}{\partial x}\right)^{-1} du \wedge du\right).$$

The Hamiltonian is

$$\begin{aligned}\mathcal{H}(u) &= \frac{1}{2}\langle Au, u\rangle + H(u)\\ &= \frac{1}{2}\int_0^{2\pi}\frac{1}{4}u''u\,dx + \frac{1}{4}\int_0^{2\pi}u^3\,dx\\ &= \frac{1}{8}uu'|_0^{2\pi} - \frac{1}{8}\int_0^{2\pi}(u')^2\,dx + \frac{1}{4}\int_0^{2\pi}u^3\,dx\\ &\int_0^{2\pi}\left[-\frac{1}{8}(u')^2 + \frac{1}{4}u^3\right]dx.\end{aligned}$$

Then (3.1.1) becomes

$$\begin{aligned}\dot{u} &= J\nabla\mathcal{H} = J(Au + \nabla H)\\ &= \frac{\partial}{\partial x}\left(\frac{1}{4}u'' + \frac{3}{4}u^2\right)\\ &= \frac{1}{4}\frac{\partial}{\partial x}(u'' + 3u^2),\end{aligned}$$

under the boundary conditions

$$\begin{aligned}u(x,t) &= u(x+2\pi, t),\\ \int_0^{2\pi}u\,dx &= 0,\end{aligned}$$

and we have that $\operatorname{ord} V_{\mathcal{H}} = d_1 = d_A + d_J = 3$.

Example: [Higher-order KdV] The KdV equation is an equation from an infinite hierarchy of Hamiltonian PDE's. The lth equation from the hierarchy can be written as

$$\dot{u} = J\nabla\mathcal{H}_l,$$

with the Hamiltonian

$$\mathcal{H}_l(u) = K_l\int_0^{2\pi}\left[(-1)^l\left(u^{(l)}\right)^2 + \text{ h.o.t.}\right]dx,$$

where the higher-order terms have no derivatives with order more than $l-2$, and $K_l \neq 0$. The first Hamiltonian in the hierarchy, $\mathcal{H}_1$, is the Hamiltonian in the previous example. It turns out that

$$\mathcal{H}_2(u) = \frac{1}{8}\int_0^{2\pi}\left[(u'')^2 - 5u^2u'' - 5u^4\right]dx,$$

and we can calculate that the actual second KdV equation is

$$\dot{u} = \frac{1}{4}u^{(5)} - \frac{1}{4}\frac{\partial}{\partial x}\left(5\left(u'\right)^2 + 10uu' + 10u^3\right).$$

Example: [Sine-Gordon Equation] **Fill this in later.**

3.1.2 Integrable Subsystems

We assume that (3.1.1) has an invariant submanifold $\mathcal{T}^{2n} \subset O_d \cap Z_\infty$ such that the restriction of (3.1.1) to $\mathcal{T}^{2n}$ is integrable.

Recall that a finite-dimensional Hamiltonian system is integrable is there is a canonical change of variables

$$g\colon (p,q) \to (I,\phi)$$

such that

$$K(I,\phi) = H \circ g^{-1}(I,\phi) = h(I),$$

and the system then becomes

$$\dot{I} = -\frac{\partial K}{\partial \phi} = 0, \qquad I(t) = I(0),$$
$$\dot{\phi} = \frac{\partial K}{\partial I} = \omega(I(t)) = \omega(I(0)),$$

so

$$\phi(t) = \omega t + \phi_0.$$

We need to make some assumptions on $\mathcal{T}^{2n}$, most of which have to do with whether or not it has singularities, and then α_2 restricted to $\mathcal{T}^{2n}$ would be degenerate. We'll fix this potential problem by working on a submanifold of $\mathcal{T}^{2n}$, where the singularities have been removed.

First, let's write

$$\mathcal{T}^{2n} = \Phi_0(R \times \mathbb{T}^n), R \subset \mathbb{R}^n, \text{ and } \mathbb{T}^n \text{ is a standard } n\text{-torus.}$$

We assume that R is the real part of a complex analytic subset R^c, which is itself the zeros of an analytic map : $\Pi^c \to \mathbb{C}^{N-n}$. Choose R_s^c to be a proper subset of R^c which contains all of the singularities, and denote R_s to be the real part of R_s^c. Choose R_{s1}^c to be a proper subset of R^c which contains all of the points where $\Phi_0^*(\alpha_2)$ is non-degenerate, and denote R_{s1} to be the real part of R_{s1}^c. In summary,

$$R_0 := R \setminus (R_s \cup R_{s1})$$

is a set on which

1. the manifold $\mathcal{T}^{2n}$ is non-degenerate, and
2. $\Phi_0^*\alpha_2$ restricted to R_0 is non-degenerate.

We also denote

$$\mathcal{T}_0^{2n} := \Phi_0(R_0 \times \mathbb{T}^n).$$

We will make four assumptions on $\mathcal{T}^{2n}$.

1. $\Phi_0 \colon R \times \mathbb{T}^n \to Z_l$ is analytic for all l. From this we can also deduce that for some $\rho > 0$, Φ_0 extends to an analytic map : $\Pi^c \times \{|\mathrm{Im}\, z| \le \rho\} \to Z_l^c$.
2. We can choose R_s and R_{s1} above in a way so that R_0, R_0^c are smooth analytic manifolds, and
$$\Phi_0(R_0 \times \mathbb{T}^n) \to Z_l$$
is an analytic immersion.
3. The set $\mathcal{T}_0^{2n}$ is a smooth analytic submanifold of each space Z_l, invariant for the equation (3.1.1), as well as in each torus
$$T^n(x) = \Phi_0(\{x\} \times \mathbb{T}^n), \qquad x \in \mathbb{R}^n.$$
Also, the equation restricted to $T^n(x)$ takes the form
$$\dot{z} = \omega(x),$$
where ω extends to an analytic map $\omega \colon \Pi^c \to \mathbb{C}^n$.
Note: Assumptions (2) and (3) assure than that manifold $\mathcal{T}_0^{2n}$ is filled with smooth time-quasiperiodic solutions of (3.1.1).

4. The frequency map $\omega(x)$ is assumed to be non-degenerate, i.e.

$$\omega_*(x)\colon T_x R_0 \to \mathbb{R}^n$$

is an isomorphism.

This can be viewed as amplitude-frequency modulation, i.e. by changing the amplitude vector x, we can change the frequency vector ω in a prescribed direction.

This implies that, for almost all x, the components of the vector $\omega(x)$ rationally independent. Thus the flow $\dot{z} = \omega(x)$ will be ergodic. Then we have that $\mathcal{T}_0^{2n}$ is filled with smooth time-QP solutions of (3.1.1).

Using Theorem 2.15, Equation (3.1.1) restricted to $\mathcal{T}_0^{2n}$ is Hamiltonian. To check this, we need that

1. Φ_* is linear and continuous (follows from assumption #2),
2. $V_{\mathcal{H}}$ restricted to Z_l is tangent to $\Phi_0(R_0 \times \mathbb{T}^n)$ (follows from assumption #3)

We want to claim that our system here is, in some sense, integrable. First we need some definitions.

Definition 3.2
Two functions f_1, f_2 are ***in involution*** *if $\{f_1, f_2\} = 0$. We also say that f_1 and f_2 are* ***independent****. A function f is a* ***constant of motion*** *(or* ***first integral****) of a Hamiltonian system with Hamiltonian H iff $\{f, H\} = 0$.* ♣

Definition 3.3
If, in a system with n degrees of freedom, we know n constants of motion that are in involution, then then system is ***Liouville-Arnold integrable*** *(or, classically,* ***integrable in quadrature****).* ♣

Lemma 3.2
The Hamiltonian equation (3.1.1) which satisfies assumptions (1)–(4) is Liouville-Arnold integrable in $\mathcal{T}_0^{2n}$ ♣

We need one more definition before we prove the lemma:

Definition 3.4
A manifold M is ***Lagrangian*** *if for any two vectors tangent to the manifold, $\xi, \eta \in T_x M$, we have that $\alpha_2(\xi, \eta) = 0$.* ♣

Proof: Our goal is to show that we can find n independent constants of motion.
From assumption (4) we know that the flow

$$\dot{z} = \omega(x), \qquad \text{on } T^n(x)$$

is ergodic, for almost all x.
Claim: A torus with ergodic flow is Lagrangian.
Since the flow maps of equation (3.1.1) are symplectic, their restriction to the torus preserve the form

$$\Omega_2 := \left.\alpha_2\right|_{T^n(x)} .$$

Since the flow on the torus is ergodic, then

$$\Omega_2 = \sum a_{ij}\, dz_i \wedge dz_j,$$

with some constant coefficients a_{ij}.
Of course, if we wanted to calculate a_{ij}, we could simply average Ω_2 along the 2-torus $T_{ij} = \{z |\ z_l = 0 \text{ for } l \neq i, j\}$. Noting that Ω_2 is exact, and using Stokes' Theorem,

$$a_{ij} = \int_{T_{ij}} \Omega_2 = \int_{T_{ij}} d\xi = \int_{\partial T_{ij}} \xi = 0.$$

Thus $\Omega_2 = 0$, and the torus is Lagrangian, and the claim is proven.

This is true for almost all tori, and by continuity (Φ_0 is analytic), we have that all the tori $T^n(x)$ are Lagrangian.

Choose $x \in R_0$. We choose coordinates $x_1, \dots, x_n$ in the vicinity of the torus $T^n(x) \subset \mathcal{T}_0^{2n}$, and consider the functions

$$f_j : \mathcal{T}_0^{2n} = \Phi_0(x, y) \to x_j$$

As the f_j's are constant on each torus $T^n(x)$, then for any $z \in T^n(x)$, and every tangent vector $\xi \in T_z T^n(x)$, we have

$$\langle df_j(z), \xi \rangle = -\Omega_2(V_{f_j}(z), \xi) = 0.$$

Thus the vectors $V_{f_j}(z)$ lie in the skew-orthogonal complement to $T_z T^n(x)$ (which is equal to $T_z T^n(x)$, since the torus is Lagrangian).

Here, the functions f_j are in involution as

$$\begin{aligned} \{f_j, f_k\} &= \Omega_2(V_{f_j}, V_{f_k}) \\ &= -\langle df_j, V_{f_k} \rangle \\ &= -\langle df_j, J\nabla f_k \rangle = -\{f_j, f_k\}. \end{aligned}$$

Similarly, each f_j commutes with the Hamiltonian of the equation, and we thus have n independent constants of motion.

■

By the last lemma, we know that R_0 can be covered by a countable system of domains $R_{0,j}$, with $r_0 = \cup R_{0,j}$, such that (3.1.1) restricted to each manifold

$$\mathcal{T}_{0,j}^{2n} = \Phi_0(R_{0,j} \times \mathbb{T}^n)$$

admits action-angle variables (p, q) with actions $p \in P_j \subset\subset \mathbb{R}^n$, and angles $q \in \mathbb{T}^n$, so that we have $\omega_2 = dp \wedge dq$ and (3.1.1) restricted to $\mathcal{T}_{0,j}^{2n}$ takes the form

$$\begin{aligned} \dot{p} &= 0 \\ \dot{q} &= \nabla h(p), \end{aligned}$$

and the actions p depend only on x.

If we want to construct the action-angle variables (p, q): choose the cycles $Q_1, \dots, Q_n$,

$$Q_l = \{(q_1, \dots, q_n) |\, q_i \in S^1, q_j = 0, j \neq l\}$$

to be homotopic to any n-cycles forming a basis of $H_1(\mathbb{T}^n, \mathbb{Z})$. One choice is

$$Q_l \sim \mathcal{Z}_l := \{\text{pt.}\} \times \cdots \times S^1 \times \cdots \times \{\text{pt.}\} \subset \mathbb{T}^n,$$

with the circle in the lth place.

Lemma 3.3
Under this choice of the Q_l, and assumptions (1)–(3), the gradient $\nabla h(p(x))$ equals $\omega(x)$. If, in addition, assumption (4) holds, then $q = z + q^0(x)$.

♠

Proof: We know that

$$\dot{q}_j = \frac{\partial h}{\partial p_j},$$

so

$$q(t) = \frac{\partial h}{\partial p_j} t + q^0,$$

and thus

$$\frac{q(t)}{t} \to \frac{\partial h}{\partial p}, \qquad t \to \infty.$$

But this number can be interpreted geometrically as the number of intersections of any trajectory on $T^n(x)$ with the cycle Q_j, i.e. the number of times is wraps around the torus.

Similarly, we can write

$$\frac{\omega_j t + \omega^0}{t} \to \omega_j,$$

which can be interpreted as the number of intersections of any trajectory with the cycle $\mathcal{Z}_l$.

Since $Q_l \sim \mathcal{Z}_l$, we have that

$$\frac{\partial h}{\partial p} = \omega.$$

By (3), we have that

$$\dot{z} = \omega(x),$$

so that

$$\frac{d}{dt}(q_i - z_i) = \nabla h - \omega = 0,$$

which gives us that

$$q = z + \text{ const, along each trajectory.}$$

If we further assume (4), we get that the trajectories are dense in the torus $T^n(x)$, so the assertion follows.

■

3.2 Lax-Integrable Equations

3.2.1 General Lax-Integrable Theory

Consider a Hamiltonian PDE

$$\dot{u} = J\nabla H(u). \tag{3.1}$$

Definition 3.5
This equation is called ***Lax-integrable*** *(or of* ***Lax type****) if there exist linear operators $\mathcal{L}_u, \mathcal{A}_u$, $u \in Z_\infty$, such that a curve $u(x,t)$ is a smooth solution of (3.1) iff*

$$\frac{d}{dt}\mathcal{L}_{u(t)} = \left[\mathcal{A}_{u(t)}, \mathcal{L}_{u(t)}\right] = \mathcal{A}\mathcal{L} - \mathcal{L}\mathcal{A}. \tag{3.2}$$

♣

We should note that KdV, Sine-Gordon equations are equations of Lax type. We further assume that the maps

$$u \mapsto \mathcal{L}_{u(t)}, \qquad u \mapsto \mathcal{A}_{u(t)}$$

are analytic, and that $\frac{d}{dt}\mathcal{L}_{u(t)}$ is well-defined for any $u(t) \in Z_s$, which is $\mathcal{C}^1$. We will also write $\mathcal{L}_t = \mathcal{L}_{u(t)}, \mathcal{A}_t = \mathcal{A}_{u(t)}$.

Lemma 3.4
Let $\chi_0 \in Z_\infty$ be a smooth eigenvector of $\mathcal{L}_0$, i.e. $\mathcal{L}_0\chi_0 = \lambda\chi_0$. Also, assume that the initial value problem

$$\dot{\chi} = \mathcal{A}_t\chi, \qquad \chi(0) = \chi_0 \tag{3.3}$$

has a unique smooth solution $\chi(t) \in \mathbb{Z}_\infty$. Then

$$\mathcal{L}_t\chi(t) = \lambda\chi(t), \qquad \text{for all } t.$$

♠

Proof: We denote

$$\begin{aligned} \xi(t) &= \mathcal{L}_t\chi(t), \\ \eta(t) &= \lambda\chi(t). \end{aligned}$$

Calculating,

$$\frac{d}{dt}\xi = \frac{d}{dt}\mathcal{L}\chi = [\mathcal{A}, \mathcal{L}]\,\chi + \mathcal{L}\mathcal{A}\chi = \mathcal{A}\mathcal{L}\chi = \mathcal{A}\xi,$$

whereas

$$\frac{d}{dt}\eta = \frac{d}{dt}\lambda\chi = \lambda\mathcal{A}\chi = \mathcal{A}\eta.$$

Thus both $\xi(t)$ and $\eta(t)$ both solve the initial-value problem, and by the uniqueness assumption, they must be equal.

■

Note: The L-periodic spectrum of $\mathcal{L}_u$ is made up of first integrals for (3.1.1) whenever (3.3) defines a flow in the space of L-periodic vector functions.

3.2.2 The KdV Equation(s)

For KdV, we have an infinite sequence of first integrals,

$$I_s = \int p_s(u, \ldots, u^{(s)})\, dx,$$

with p_s a polynomial. For example, we have

$$\begin{aligned} I_{-1} &= \int u\, dx, \\ I_0 &= \int u^2\, dx, \\ I_1 &= \int \left(\frac{(u')^2}{2} + u^3 \right) dx. \end{aligned}$$

Consider the KdV equation

$$\begin{aligned} \dot{u} &= \frac{1}{4}\frac{\partial}{\partial x}(u_{xx} + 3u^2), \\ u(x,t) &= u(x + 2\pi, t), \\ \int_0^{2\pi} u(x,t)\, dx &= 0. \end{aligned}$$

The Lax pair will be

$$\begin{aligned} \mathcal{L}_u &= -\frac{\partial^2}{\partial x^2} - u, \\ \mathcal{A}_u &= \frac{\partial^3}{\partial x^3} + \frac{3}{2}u\frac{\partial}{\partial x} + \frac{3}{4}u_x. \end{aligned}$$

From [Mc77],[MT76], we know that the spectrum of the problem $\mathcal{L}_u v = \lambda v$ acting on $\mathcal{C}^2$ functions of period 4π will be single or double eigenvalues

$$\lambda_0 < \lambda_1 \leq \lambda_2 < \lambda_3 \leq \lambda_4 \to \infty.$$

The corresponding eigenfunctions will be smooth if $u(x)$ is smooth.

Example: Let $u = 0$, then we have the eigenvalue problem

$$-\frac{\partial^2 v}{\partial x^2} = \lambda v,$$

and the eigenvalues are

$$\lambda_{2k} = \frac{k^2}{4},$$
$$\lambda_{2l-1} = \frac{l^2}{4},$$

and the corresponding eigenfunctions are

$$\phi_{2k} = (2\pi)^{-1/2} \cos\left(\frac{kx}{2}\right),$$
$$\phi_{2l-1} = (2\pi)^{-1/2} \sin\left(\frac{lx}{2}\right).$$

We also know that if $u(x,t)$ is a smooth function periodic in x, then the equation

$$\dot{v} = \mathcal{A}_{u(x,t)} v, \qquad v(x,0) = v_0(x),$$

has a unique, smooth, periodic (in x) solution $v(x,t)$ for smooth, periodic initial data $v_0(x)$.

If we let our Hilbert scale be $Z_s = H^s(\mathbb{R}/4\pi\mathbb{Z})$, then the above lemma says that $\lambda(u(t,\cdot))$ is time-independent if $u(x,t)$ solves KdV.

Definition 3.6
*The segment $\Delta_j = [\lambda_{2j-1}, \lambda_{2j}]$ is called the jth **spectral gap**. The gap is **open** if $\lambda_{2j} > \lambda_{2j-1}$, otherwise it is **closed**.* ♣

We define an integer n-vector ν as

$$\nu = (\nu_1, \ldots, \nu_n), \qquad \nu_1 < \cdots < \nu_n$$

and consider the set $\mathcal{T}_\nu^{2n}$ defined as

$$\mathcal{T}_\nu^{2n} = \{u(x) | \text{ the gap } \Delta_j(u) \text{ is open, iff } j = \nu_i\}$$

Furthermore, if $u \in |{}^{2n}_{\nu}$, i.e. the operator $\mathcal{L}_u$ has n open spectral gaps, we say that u is an **n-gap potential**.

Now, given a $r = (r_1, \ldots, r_j)$, we define

$$\mathcal{T}_\nu^n(r) = \{u \in \mathcal{T}_\nu^{2n} | \ \left|\Delta_{\nu_j}\right| = r_j\}.$$

Note that, by definition,

$$\mathcal{T}_\nu^{2n} = \bigcup_{r \in \mathbb{R}_+^n} \mathcal{T}_\nu^n(r).$$

Using, again, results from elsewhere: In [GT87], the spectrum $\lambda(u)$ is actually defined by the n-vector r, and depends analytically on it. This is nontrivial, since we see that we can specify the spectrum simply by specifying the gaps. Furthermore, each set $\mathcal{T}_\nu^n(r)$ is non-empty, and is an analytic n-torus for $H_0^s(S^1)$, for all s. From [Mc77],[Mos78],[GT87],[MT76], we can also say that the $\mathcal{T}_\nu^n(r)$ are analytically glued together, so that $\mathcal{T}_\nu^{2n}$ is an analytic $2n$-submanifold on each H_0^s.

In summary: Since we've established that $\lambda(u(t,\cdot))$ is independent of time, then each set $\mathcal{T}_\nu^n(r)$ is invariant for the KdV flow. Thus we have the existence of KdV-invariant $2n$-manifolds $\mathcal{T}_\nu^{2n}$ that are foliated by invariant n-tori $\mathcal{T}_\nu^n(r)$.

One would expect, if we take an n-gap potential, and shrink one of the gaps, we should get an $n-1$-gap potential. More specifically, let $\nu = (\nu_1, \ldots, \nu_n)$, and take $u \in \mathcal{T}_n^{2n}$. If we were to change u so that the ν_kth gap shrank to a point, we would expect to get an $n-1$-gap potential from the set $\mathcal{T}^{2n-2}_{(\nu_1,\ldots,\hat{\nu_k},\ldots,\nu_n)}$. The following theorem says that this occurs in an analytic way:

Theorem 3.5
The closure of $\mathcal{T}_\nu^{2n}$ in H_0^s (for $s \geq 1$), is a $2n$-dimensional analytic submanifold of H_0^s, diffeomorphic to $\mathbb{R}^{2n} = \{z\}$. This manifold contains all finite-gap manifolds $\mathcal{T}_{\nu^m}^{2m}$ where $\nu^m \subset \nu (m < n)$. It passes through the origin, and its tangent space at the origin is spanned by $e_l^\pm \in H_0^s$, with

$$e_l^+ = \frac{1}{\sqrt{\pi}}\cos(\nu_l x),$$
$$e_l^- = \frac{1}{\sqrt{\pi}}\sin(\nu_l x).$$

Furthermore, we can choose coordinates z_k, $k = 1, \ldots, 2n$, so that

1. $$\frac{\partial}{\partial z_{2l-1}} = e_l^+, \qquad \frac{\partial}{\partial z_{2l}} = e_l^-.$$

2. If we take u with $\|u\|_s$ small, and write it as

$$u = \sqrt{\pi}\sum_{j=1}^{\infty}\left(u_j^+\cos jx - u_j^-\sin jx\right),$$

then

$$z_{2k-1} = u_{\nu_k}^+ + \mathcal{O}(\|u\|_s^2),$$
$$z_{2k} = u_{\nu_k}^- + \mathcal{O}(\|u\|_s^2).$$

3. We also have that

$$z_{2j-1}^2 + z_{2j}^2 = r_j^2, \text{ for all } j.$$

♠

Using Vey's version of the integrability theorems (see [Vey78],[Ito89]), we can reword the above theorem as below:

Theorem 3.6
If δ is sufficiently small, and $s \geq 1$, then there is a $\delta_1 > 0$ and an analytic map

$$U\colon O_{\delta_1}(\mathbb{R}^{2n}) \to \mathcal{T}_\nu^{\leq 2n} \subset H_0^s, \qquad y \mapsto U(\cdot, y),$$

such that its image is contained in $\mathcal{T}^{\leq 2n}$. The transformation $y \mapsto z = z(U(\cdot,y))$ is a diffeomorphism of the form $z = y + \mathcal{O}(|y|^2)$ (so that $y(0) = 0$). Furthermore,

1. $U^*\alpha_2 = \sum_{l=1}^n \nu_l^{-1} dy_{2l-1} \wedge dy_{2l}$,

2. The pullback of the Hamiltonian of the KdV equation is an analytic function, h^n, of the arguments $y_1^1 + y_2^2, \ldots, y_{2n-1}^2 + y_{2n}^2$,

3. for any $l \leq n$, the submanifold formed by the potentials $u(x)$ such that $|\Delta_{\nu_i}| = 0$ corresponds to the subspace $\{y|\ y_{2l-1} = y_{2l} = 0\}$,
4. the finite-gap torus $\mathcal{T}^{2n}(r)$ can be written, in y coordinates, in the form $\{y_{2l-1}^2 + y_{2l}^2 = C_l(r)\}$. ♠

These coordinates will give us action-angle variables on $\mathcal{T}_\nu^{2n}$. We choose

$$I_j = \frac{1}{2\nu_j}(y_{2j-1}^2 + y_{2j}^2), \tag{3.4}$$

$$q_j = \operatorname{Arg}(y_{2j-1} + iy_{2j}). \tag{3.5}$$

These coordinates are symplectic, since we can see by calculation that $U^*\alpha_2 = dI \wedge dq$. The KdV Equation restricted to $\mathcal{T}_\nu^{\leq 2n}$ takes the form

$$\begin{aligned} \dot{I} &= 0, \\ \dot{q} &= \nabla h^n(I). \end{aligned}$$

Then we can write the finite-gap solutions which fill the manifold $\mathcal{T}^{2n}$ as

$$u(t,x) = U(x; I, q + t\nabla h^n(I)). \tag{3.6}$$

Note that (see [Kuk00]) we can also write

$$R_j = \sqrt{y_{2j-1}^2 + y_{2j}^2} = \sqrt{2\nu_j I_j},$$

so that

$$U^*\alpha_2 = \frac{1}{2}\sum_j dR_j^2 \wedge dq_j,$$

and since we know

$$I_j = \frac{r_j^2}{2\nu_j}(1 + \mathcal{O}(|r|^2)),$$

then we can say that

$$R_j = r_j(1 + \mathcal{O}(|r|^2)). \tag{3.7}$$

3.2.3 An integrable example

Most of this section is from [AS]. Recall the KdV equation:

$$u_t = \frac{1}{4}\left(u_{xx} + 3u^2\right)_x. \tag{3.8}$$

We have the Lax pair

$$\begin{aligned} \mathcal{L}_u &= -\frac{d^2}{dx^2} - u, \\ \mathcal{A}_u &= \frac{d^3}{dx^3} + \frac{3}{2}u\frac{d}{dx} + \frac{3}{4}u_x. \end{aligned}$$

We know that $u(x,t)$ solves the KdV iff $\dot{\mathcal{L}}_u = [\mathcal{A}_u, \mathcal{L}_u]$.
Recall that the eigenvalues of $\mathcal{L}_u$ are constants of motion, i.e.

$$\mathcal{L}_{u(t)}\chi_t = \lambda\chi_t,$$

where

$$\chi_t = \dot{\chi} = \mathcal{A}_u \chi.$$

$$\chi'' - u\chi = \lambda\chi \Leftrightarrow \chi'' + (u+\lambda)\chi = 0. \tag{3.9}$$

Let's fix some point $x_0 \in [0, 2\pi]$. (Recall that we're looking for solutions of KdV with periodic boundary conditions, i.e. $u(x,t) = u(x+2\pi, t)$.)

Define $\phi(x; x_0, \sqrt{\lambda}), \phi^*(x; x_0, \sqrt{\lambda})$ to be solutions of (3.9) that satisfy initial conditions

$$\begin{aligned} \phi(x_0; x_0, \sqrt{\lambda}) &= 1 \\ \phi^*(x_0; x_0, \sqrt{\lambda}) &= 1 \\ \phi_x(x_0; x_0, \sqrt{\lambda}) &= i\sqrt{\lambda} \\ \phi^*_x(x_0; x_0, \sqrt{\lambda}) &= -i\sqrt{\lambda} \end{aligned} \tag{3.10}$$

Define

$$\Phi(x; x_0, \sqrt{\lambda}) = \begin{pmatrix} \phi & \phi_x \\ \phi^* & \phi^*_x \end{pmatrix} \tag{3.11}$$

Since we know that ϕ, ϕ^* are two linearly independent solutions of a second-order initial value problem (and thus span the set of solutions), we know that we can write

$$\Phi(x+2\pi; x_0, \sqrt{\lambda}) = T(x_0, \sqrt{\lambda})\Phi(x; x_0, \sqrt{\lambda}). \tag{3.12}$$

Definition 3.7
The matrix T above is known as the ***monodromy matrix****.* ♣

By calculation, we see that T must have the form

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}.$$

Note that $\det \Phi$, the **Wronskian** of (3.9), is independent of x. Thus

$$\det \Phi(x+2\pi; x_0, \sqrt{\lambda}) = \det \Phi(x, ; x_0, \sqrt{\lambda}),$$

and therefore

$$\det T = 1.$$

Putting these two together, we have that

$$|a|^2 - |b|^2 = 1. \tag{3.13}$$

We define $E_1, E_2, \ldots$ to be the eigenvalues of $\mathcal{L}_u$ acting on functions periodic of period 4π. (See Figure 3.1)

Consider a new boundary eigenvalue problem for (3.9):

$$\mathcal{L}_u r = \gamma r, \qquad r(x_0) = r(x_0 + 2\pi) = 0. \tag{3.14}$$

Since (r, r_x) solve (3.9), we have that

$$\begin{aligned} r &= A\phi + B\phi^*, \\ r_x &= A\phi_x + B\phi^*_x, \end{aligned}$$

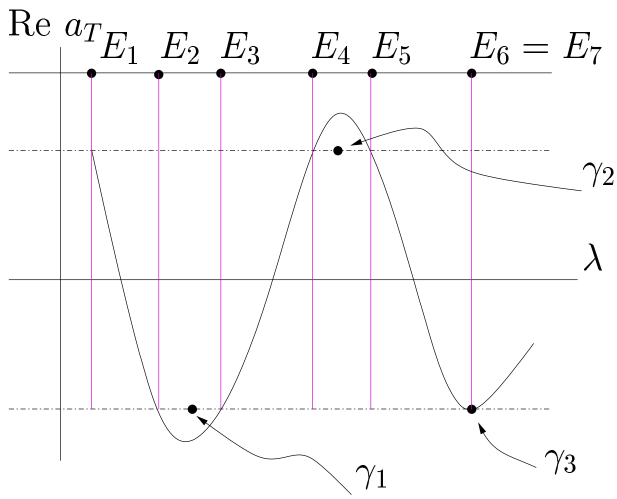

Figure 3.1: The eigenvalues E_j and γ_j

so that

$$(r, r_x) = (A, B)\Phi.$$

Since

$$r(x_0) = 0 = A + B,$$

then

$$-A = B.$$

Also,

$$\begin{aligned} r(x_0 + 2\pi) = 0 &= (A, -A)\Phi(x_0 + 2\pi) \\ &= (A, -A)T\Phi(x_0) \\ &= (A, -A)\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}\begin{pmatrix} 1 & i\sqrt{\lambda} \\ 1 & -i\sqrt{\lambda} \end{pmatrix}. \end{aligned}$$

The eigenvalues γ that satisfy (3.14) have the property that $\operatorname{Im} a + \operatorname{Im} b = 0$. From (3.13), we have that

$$\begin{aligned} a_R^2 + a_I^2 - (b_R^2 + b_I^2) &= 1 \\ a_R^2 - b_R^2 + a_I^2 - b_I^2 &= 1. \end{aligned}$$

Since $a_I + b_I = 0$, we know that $a_I^2 - b_I^2 = 0$, and thus, for eigenvalues γ, we have

$$a_R^2 = 1 + b_R^2 \geq 1.$$

This says that the γ's must be inside the gaps (see figure).

Theorem 3.7
Suppose that we have N open gaps. Then the solution of KdV satisfies

$$u(x_0) = \sum_{j=1}^{2N+1} E_j - 2\sum_{j=1}^{N} \gamma_j. \tag{3.15}$$

♠

Proof: We will consider yet more solutions of equation (3.9). We define the **Bloch eigenfunctions**, which are solutions $\psi_\pm(x; x_0, \sqrt{\lambda})$ that satisfy

$$\psi_\pm(x_0; x_0, \sqrt{\lambda}) = 1,$$
$$\psi_\pm(x_0 + 2\pi; x_0, \sqrt{\lambda}) = \Lambda_\pm.$$

A calculation similar to above gives us that

$$\Lambda_\pm^2 - 2a_R\Lambda_\pm + 1 = 0. \tag{3.16}$$

For λ large, we expect that

$$\psi_\pm(x; x_0, \sqrt{\lambda}) \approx \exp(\pm i\sqrt{\lambda}(x - x_0)).$$

Let's consider

$$\chi_\pm = \frac{-i\psi'_\pm}{\psi_\pm} = -i\frac{d}{dx}(\log \psi_\pm).$$

Thus, for λ large, we have that

$$\chi_\pm \approx \pm\sqrt{\lambda}.$$

Furthermore, if we use asymptotic methods and plug into Equation (3.9), we can calculate that

$$\chi_\pm \approx \left(\sqrt{\lambda} - \frac{u(x_0)}{2\sqrt{\lambda}} - \frac{2iu_x(x_0)}{4\lambda} + \ldots\right)$$

We also know that

$$\psi_+ = A\phi + B\phi^*.$$

So we have

$$\psi_+(x_0) = 1 = A + B$$
$$\psi'_+(x_0) = i\chi_+^0 := i\chi_+(x_0) = i\sqrt{\lambda}(A - B).$$

This and similar calculations for ψ_- gives us that

$$\psi_\pm(x; x_0, \sqrt{\lambda}) = \frac{1}{2}\left(1 + \frac{\chi_\pm^0}{\sqrt{\lambda}}\right)\phi + \frac{1}{2}\left(1 - \frac{\chi_\pm^0}{\sqrt{\lambda}}\right)\phi^* \tag{3.17}$$

We also know that

$$\psi_\pm(x + 2\pi; x_0, \sqrt{\lambda}) = \Lambda_\pm\psi_\pm(x; x_0, \sqrt{\lambda})$$

Using Equations (3.16) and (3.17), we can equate coefficients to get

$$\chi_\pm^0 = \left(\frac{\pm\sqrt{1 - a_R^2} + ib_R}{a_I + b_I}\right)\sqrt{\lambda}. \tag{3.18}$$

Remarks: It can be checked (but won't be here) that $(1 - a_R^2)$ and $(a_I + b_I)/\sqrt{\lambda}$ are entire functions of λ. We know that a general entire function X can be written

$$X(z) = \left[\prod_{j=1}^{\infty}(z - r_j)\right] \cdot \Xi(z),$$

where r_j are the roots of X, and $\Xi(z)$ is entire and has no zeros.

Recall that the E_j's are the roots of $(1-a_R^2)$, and the γ_j's are the roots of a_I+b_I. Renumbering the E_j's and γ_j's, to get the open gaps first, we can write

$$1-a_R^2 = g_1(\lambda)\prod_{j=1}^{2N+1}(\lambda-E_j)\prod_{k=1}^{\infty}\left(\lambda-E_{2(N+k)}\right)^2,$$
$$\frac{a_I+b_I}{\sqrt{\lambda}} = g_2(\lambda)\prod_{j=1}^{N}(\lambda-\gamma_j)\prod_{k=1}^{\infty}(\lambda-\gamma_{N+k}).$$

We should note that since all the gaps past N are closed, we have double roots in the first expression above, and also, that $\gamma_{N+k}=E_{2(N+k)}$.

Now, we have

$$\begin{aligned}\chi_{0R}^{+} &= \frac{\sqrt{1-A_R^2}\sqrt{\lambda}}{(a_I+b_I)}\\ &= \frac{\sqrt{g_1(\lambda)}\left(\pm\sqrt{\prod_{j=1}^{N}(\lambda-E_j)}\right)\prod_{k=1}^{\infty}(\lambda-E_{2(N+k)})}{g_2(\lambda)\prod_{j=1}^{N}(\lambda-\gamma_j)\prod_{k=1}^{\infty}(\lambda-\gamma_{N+k})}\\ &= \frac{\sqrt{g_1(\lambda)}\left(\pm\sqrt{\prod_{j=1}^{N}(\lambda-E_j)}\right)}{g_2(\lambda)\prod_{j=1}^{N}(\lambda-\gamma_j)}.\end{aligned}$$

So as $\lambda\to\infty$,

$$\chi_{0R}^{+}\overset{\text{inspection}}{\approx}\lambda\left(\frac{g_1(\lambda)}{(g_2(\lambda))^2}\right)\overset{(3.18)}{\approx}\lambda.$$

Therefore

$$\frac{g_1(\lambda)}{(g_2(\lambda))^2}\approx 1$$

as $\lambda\to\infty$. Since g_1 and g_2 are both entire functions without zeros, the quotient is entire, so by Liouville's Theorem,

$$\frac{g_1(\lambda)}{(g_2(\lambda))^2}=1.$$

So we can deduce that

$$\begin{aligned}\left(\chi_{0R}^{+}\right)^2 &= \frac{\prod_{j=1}^{2N+1}(\lambda-E_j)}{\prod_{k=1}^{N}(\lambda-\gamma_j)^2}\\ &\approx \lambda+\sum_{j=1}^{2N+1}-E_j+2\sum_{j=1}^{N}\gamma_j+\mathcal{O}\left(\frac{1}{\sqrt{\lambda}}\right).\end{aligned}$$

Equating coefficients gives

$$-u=\sum_{j=1}^{2N+1}-E_j+2\sum_{j=1}^{N}\gamma_j,$$

or

$$u=\sum_{j=1}^{2N+1}E_j-2\sum_{j=1}^{N}\gamma_j.$$

■

Now, we can calculate that

$$\frac{d\gamma_j}{dx_0} = \frac{2i\left(\pm\sqrt{\prod_{j=1}^{2N+1}(\gamma - E_j)}\right)}{\prod_{\substack{k=1\\k\neq j}}^{N}(\gamma_j - \gamma_k)}$$

$$\frac{d\gamma_j}{dt} = \frac{8i\left(\prod_{k=1}^{N}\gamma_k - \frac{1}{2}\sum_{k=1}^{2N+1}E_k\right)\left(\pm\sqrt{\prod_{j=1}^{2N+1}(\gamma - E_j)}\right)}{\prod_{\substack{k=1\\k\neq j}}^{N}(\gamma_j - \gamma_k)}.$$

Define the Riemann surface

$$R(E) = \prod_{j=1}^{2N+1}(\lambda - E_j),$$

putting the branch cuts in the open gaps. (see Figure 3.2)

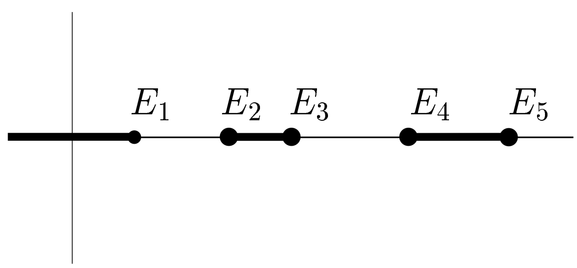

Figure 3.2: Da branch cutz

Define a coordinate $P_j = (\gamma_j, \sigma_j)$ on the Riemann surface, with $\sigma = \pm 1$, which tells us which branch that we're on.

Define

$$\Omega_m(E) = \sum_{k=0}^{N-1} C_{km}\frac{E^k\, dE}{(R(E))^{1/2}}.$$

We introduce new coordinates

$$\eta_m(P_1, \ldots, P_N) = \sum_{j=1}^{N}\int_{E_{2j}}^{P_j}\Omega_m(E),$$

making the normalization

$$\oint_{a_j}\Omega_m(E) = 2\pi i\delta_{m,j},$$

where a_j is the loop encircling the jth branch cut.

It is shown in [AS] that the change of coordinates

$$(P_1, \ldots, P_N) \mapsto (\eta_1, \ldots, \eta_N)$$

is invertible, so we can write

$$P_j = f_j(\eta_1, \ldots, \eta_N).$$

We calculate

$$\begin{aligned}
\frac{d\eta_m}{dx_0} &= \sum_{j=1}^{N} \Omega_m(P) \frac{d\gamma_j}{dx_0} \\
&= 2i \sum_{j=1}^{N} \left(\sum_{k=0}^{N-1} C_{km} \frac{\gamma^k}{R^{1/2}(\gamma)} \right) \frac{(R(\gamma))^{1/2}}{\prod\limits_{\substack{l=1 \\ l \neq j}}^{N} (\gamma_j - \gamma_l)} \\
&= 2i \sum_{k=0}^{N-1} C_{km} \sum_{j=1}^{N} \frac{\gamma_j^k}{\prod\limits_{\substack{l=1 \\ l \neq j}}^{N} (\gamma_j - \gamma_l)}.
\end{aligned}$$

Note: By Cauchy's Theorem,

$$\sum_{j=1}^{N} \frac{\gamma_j^k}{\prod\limits_{\substack{l=1 \\ l \neq j}}^{N} (\gamma_j - \gamma_l)} = \delta_{k,N-1},$$

and so

$$\frac{d\eta_m}{dx_0} = 2iC_{N-1,m} =: iV_m.$$

Similarly,

$$\frac{d\eta_m}{dt} = -8iC_{N-2,m} - 4iC_{N-1,m} \left(\sum_{j=1}^{2N+1} E_j \right) =: -iW_m.$$

So we can write

$$\eta_m(x_0, t) = iV_m x - iW_m t + \eta_m(0,0),$$

or

$$P_j = f_j(i\mathbf{V}x - i\mathbf{W}t + \eta(0)), \tag{3.19}$$

which implies that the γ_j's are quasi-periodic functions.

Thus when we write

$$u(x,t) = \sum_{j=1}^{2N+1} E_j - \sum_{j=1}^{N} \gamma_j,$$

we see that the E_j's are independent of t, and the γ_j's are quasi-periodic with respect to t, thus u is quasi-periodic.

3.2.4 A theorem on the frequencies

As we saw in the last section, when we have finite-gap solutions to KdV, they have space-quasiperiodic profiles (with frequency $\mathbf{V}$), which precess around tori (with time frequencies $\mathbf{W}$).

It it proven in [Kuk00] that $\mathbf{V} = \nu$, the numbering of the open gaps. The properties of $\mathbf{W}$ are going to be important, since, as we recall, in KAM, the way we precess around tori is important to whether or not the tori will survive under perturbation.

What we will show is

Theorem 3.8

$$W_j(r) = \frac{-V_j^3}{4} + \frac{3}{8\pi V_j} r_j^2 + \cdots \tag{3.20}$$

♠

We should note that this theorem says quite a lot. First, it says that as $r \to 0$, W_j looks like $-V_j^3/4$. Secondly, we claim that W never depends on r linearly, but quadratically, and, further, that the jth frequency component depends quadratically only on the jth radial component. A priori, W_j could depend on r_k, but this must happen at a very high order.

We will prove the theorem in the 1-gap and 2-gap cases, and see a technique to push the theorem for all finite-gap solutions.

It is proven in [Kuk00] that as $u \to 0$, for fixed $\mathbf{V}$, we can say that

$$W_j \to \frac{-V_j^3}{4}.$$

What does it mean to fix $\mathbf{V}$ and send $u \to 0$? Fixing $\mathbf{V}$ tells us which gaps are open, but not how far open they are. As we send $u \to 0$, the gaps must necessarily close, but the open ones are staying open, so essentially we are sending each $r_j \to 0$. In other words,

$$W_j \to \frac{-V_j^3}{4}, \qquad r \to 0, \tag{3.21}$$

proving the first term in the theorem.

By statement #2 in Theorem 3.6, we know that $\mathbf{W}$ is an analytic function of $R_1^2, \ldots, R_n^2$, with $R_j = r_j(1 + \mathcal{O}(|r|^2))$ (see Equation (3.7).

Let's consider a small-gap solution, i.e. a solution u with $|r| \ll 1$. Using (3.6) and the fact that $\nabla h^n = \mathbf{W}$, we have

$$\begin{aligned} u &= U(x; R, t\mathbf{W}(R) + q_0) \\ &= U(0; R, t\mathbf{W}(R) + \mathbf{V}x + q_0). \end{aligned}$$

The second equality follows from the last section where we showed that these finite-gap solutions must have a quasi-periodic profile. For simplicity, we denote

$$G(q, R) = U(0; R, q),$$

so that we have

$$u(t, x; R, q) = G(\mathbf{W}(R)t + \mathbf{V}x + q_0, R) \tag{3.22}$$

(compare this to Equation 3.19).

We can deduce from Theorem 3.5 that

$$G(q, R) = \frac{1}{\sqrt{\pi}} \sum R_j \cos q_j + \mathcal{O}(|R|^2). \tag{3.23}$$

Example: [1-gap potentials] Let $n = 1$, so that $\mathbf{V} = V_1 =: k$. We have an invariant 2-torus $\mathcal{T}_k^2$ filled with time-periodic solutions $w(x,t)$. We know that

$$w = G(kx + Wt + q_0, R),$$

where we know that $G(Y,R)$ is analytic, 2π-periodic in Y and W, and analytic in R^2. Since

$$\int w \, dx = 0 \text{ for all } t,$$

then

$$\int G \, dY = 0.$$

By (3.23), we have that

$$G(Y,R) = R\frac{1}{\sqrt{\pi}}\cos Y + R^2 g_2(Y) + R^3 g_3(Y) + \cdots,$$

where g_2 and g_3 also satisfy the periodic boundary conditions, and by (3.21),

$$W = \frac{-1}{4}k^3 + R^2 W_2 + \mathcal{O}(R^4).$$

We really want to find W_2, and we'll do so by a compatibility condition. Let's plug the formula for w into KdV to get

$$WG' = \frac{1}{4}k^3 G''' + \frac{3}{4}k(G^2)',$$

or

$$k^3 G'' - 4WG + 3kG^2 = \text{const}.$$

Plugging Equation (3.23) into the above, and equating the coefficients for R^2 and R^3, we get:

$$R^2: \qquad k^3 g_2'' + k^3 g_2 + \frac{3k}{\pi}\cos^2(Y) = \text{const.}, \tag{3.24}$$

$$R^3: \qquad k^3 g_3'' + k^3 g_3 - 4W_2\cos(Y) + 6kg_2\cos(Y) = \text{const.}. \tag{3.25}$$

Solving equation (3.24) gives us that

$$g_2 = -\frac{1}{4k^2\pi}\cos(2x).$$

If we plug this into (3.25), we get

$$g_3'' + g_3 = \frac{-1}{k^3}\left(\frac{3}{2k\pi} - 4W_2\right)\cos(Y) + \frac{3}{k^4\pi}\cos(Y)\sin^2(Y).$$

This is an inhomogeneous equation, but note that $\cos(Y)$ solves the homogeneous equation. So forcing the equation with $\cos(Y)$ will lead to resonance (and we couldn't get a solution with periodic boundary conditions). Therefore

$$\frac{3}{2k\pi} = 4W_2,$$

or

$$W_2 = \frac{3}{8k\pi},$$

proving the theorem for the 1-gap case with the formula

$$W(R) = -\frac{k^3}{4} + \frac{3R^2}{8k\pi} + \mathcal{O}(R^4) \tag{3.26}$$

For more gaps, we will use a "method of descent"-like method. First, some notation:
If $U = (U_1, \ldots, U_n)$, and $m \leq n$, we denote

$$U^{\hat{m}} = (U_1, \ldots, \widehat{U_m}, \ldots, U_n).$$

Recalling part 3 of Theorem 3.6, consider a n-gap manifold $\mathcal{T}_{\mathbf{V}}^{\leq 2n}$. When we close a gap, we get the manifold $\mathcal{T}_{\mathbf{V}^{\hat{m}}}^{2n-2}$.

We should also note that

$$\mathbf{W}^{\hat{m}}(R)\big|_{R_m=0} = \mathbf{W}(R^{\hat{m}}), \tag{3.27}$$

which is simply a restatement of part 3 of Theorem 3.6.

Theorem 3.9

1. Let $m \leq n$, $R \in \mathbb{R}^n$ with $|R| \ll 1$. Assume that $R_m = 0$,and $r_l \neq 0$ for $l \neq m$. Then

$$u_{n-1}(t, x; R^{\hat{m}}, q) = G(\mathbf{V}x + \mathbf{W}t + q; R^{\hat{m}})$$

is an $(n-1)$-gap solution in $T_{\mathbf{V}^{\hat{m}}}^{n-1}(R^{\hat{m}})$ with frequency $\mathbf{W}^{\hat{m}}$.

2. Let r^ϵ be the vector $(R_1, \ldots, \epsilon, \ldots, R_n)$, with the ϵ in the mth spot. If we choose $q_m \in S^1$, then

$$v = \frac{\partial}{\partial \epsilon} G(\mathbf{V}x + \mathbf{W}t + q_m, R^\epsilon)\bigg|_{\epsilon=0}$$

solves

$$\dot{v} - \frac{1}{4} v_{xxx} = \frac{3}{2} \frac{\partial}{\partial x}(u_{n-1} v). \tag{3.28}$$

♠

Proof: Clear after Equation (3.27).

■

Also, since U is analytic in y, then so is G. If we change variables to

$$\begin{aligned} w_j &= y_{2j-1} + i y_{2j} =: R_j e^{i q_j}, \\ w_{-j} &= \overline{w_j}, \end{aligned}$$

then G must also be analytic in the w variables, so we can write

$$G(q, R) = \sum_{s \in \mathbb{R}^{2n}_{>0}} C_s w^s = \sum_{s \in \mathbb{R}^{2n}_{>0}} C_s \prod_{p=1}^{n} R_p^{s_p + s_{-p}} \exp(i q_p (s_p - s_{-p})). \tag{3.29}$$

Example: [2-gap potential] Let's assume that we have a 2-gap potential with the mth and kth gaps open, $k \neq m$. Assume that we have a

$$u \in T^2(R, R') \subset \mathcal{T}_{k,m}^{\leq 4}, \qquad 0 < R' \ll R \ll 1.$$

So we know we have

$$u(t, x; R, R', q) = G(\mathbf{V}x + \mathbf{W}t + q, R, R') \tag{3.30}$$

We know that $\mathbf{V} = (k, m)$, so denote $\mathbf{W} = (W_k, W_m), q = (q_k, q_m)$.

If we set R' to 0, then we get the 1-gap potential

$$w(t, x; R) = u(t, x; R, 0),$$

and we know that

$$v := \frac{\partial}{\partial R'} u(t, x, R, 0)\bigg|_{R'=0}$$

solves equation (3.28) with $u_{n-1} = w$. By earlier arguments, $W_k(R, 0) = W(R)$, so that $W_k(R, 0)$ satisfies (3.26) with $R = R_k$.

Since $\mathbf{W}(R, R')$ is an analytic function of $R^2, (R')^2$, we have that

$$W'_{R'}(R, 0) = 0.$$

Differentiating Equation (3.30) with respect to R', we have

$$u'_{R'}|_{R'=0} = G'_{R'}(\mathbf{V}x + \mathbf{W}t + q, R, 0).$$

Considering Equation (3.29) gives us

$$\frac{\partial G}{\partial R'} = \sum_{s \in \mathbb{R}_+^2} C_s(R^{s_1+s_{-1}} \exp(iq_1(s_1 - s_{-1}))(s_2 + s_{-2})(R')^{s_2+s_{-2}+1} \exp(iq_2(s_2 - s_{-2}))).$$

If we plug in $R' = 0$, and write

$$Z = mx + W_m t + q_m,$$
$$Y = kx + W_k t + q_k,$$

we get

$$C_1(1 + f(Y, R))e^{iZ} + C_2(1 + g(Y, R))e^{-iZ},$$

with $f(Y, 0) = g(Y, 0) = 0$, and $|C_1| + |C_2| \neq 0$.

Taking linear combinations, we can write

$$v = e^{iZ} H(Y, R)$$
$$H = 1 + Rh_1(Y) + R^2 h_2(Y) + \cdots,$$

where we've simply expanded H in a Fourier series with respect to R.

Substituting v into

$$\dot{v} - \frac{1}{4} v_{xxx} = \frac{3}{2}(u_{n-1} v)_x,$$

and multiplying both sides by e^{-iZ}, we get

$$e^{-iZ}\left(\frac{\partial}{\partial t} - \frac{1}{4}\frac{\partial^3}{\partial x^3}\right) e^{iZ} H = \frac{3}{2} e^{-iZ} \frac{\partial}{\partial x}(we^{iZ} H). \tag{3.31}$$

According to Equation (3.21), we know that the function $W_m(R, 0)$ has the form

$$W_m(R, 0) = -\frac{m^3}{4} + \omega_2 R^2 + \mathcal{O}(R^4),$$

and thus

$$e^{-iZ}\left(\frac{\partial}{\partial t} - \frac{1}{4}\frac{\partial^3}{\partial x^3}\right) e^{iZ} = i\omega_2 R^2 + \mathcal{O}(R^4).$$

The left hand side of (3.31) is, after some serious calculation,

$$-R\frac{k}{4}\frac{\partial}{\partial Y}Mh_1 + R^2\left(i\omega_2 - \frac{k}{4}\frac{\partial}{\partial Y}Mh_2\right) + \cdots$$

where

$$M = k^2\frac{\partial^3}{\partial Y^3} + 3imk\frac{\partial^2}{\partial Y^2} + (k^2 - 3m^2)\frac{\partial}{\partial Y}.$$

After some more calculation, we can get that the right hand side of (3.31) is

$$\frac{3}{2}R\left(imw_1 + k\frac{\partial}{\partial Y}w_1\right) + \frac{3}{2}R^2\left(imw_2 + imw_1h_1 + k\frac{\partial}{\partial Y}(w_1h_1 + w_2)\right) + \cdots$$

Equating coefficients for the R and R^2 terms gives us

$$-\frac{k}{4}Mh_1 = \frac{3}{2}im\left(\frac{\partial}{\partial Y}\right)^{-1}w_1 + \frac{3}{2}kw_1 = \frac{3}{2\sqrt{\pi}}(im\sin Y + k\cos Y)$$

$$-\frac{k}{4}Mh_2 = i\left(\frac{\partial}{\partial Y}\right)^{-1}\left[\frac{3}{2}m(w_2 + w_1h_1) - \omega_2\right] + \frac{3}{2}k(w_1h_1 + w_2)$$

Since the average

$$\left\langle\frac{3}{2}m(w_2 + w_1h_1) - \omega_2\right\rangle = \left\langle\frac{3}{2}m\left(\frac{\cos 2Y}{2k^2\pi} - \frac{i}{m\pi}\sin Y\cos Y\right) - \omega_2\right\rangle = -\omega_2$$

must vanish, we know that $\omega_2 = 0$. Therefore we know v has the form

$$v = e^{i(mx + W_mt + q_m)}\left(1 - \frac{iR}{m\sqrt{\pi}}\sin(kx + W_kt + q_k) + \mathcal{O}(R^2)\right),$$

and since $\omega_2 = 0$, we know that $W_m = -m^3/4 + \mathcal{O}(r^4)$. So for $R' = 0$, and $R \to 0$, we have

$$W_k = -\frac{k^3}{4} + \frac{3R^2}{8k\pi} + \mathcal{O}(R^4), \tag{3.32}$$

$$W_m = -\frac{m^3}{4} + \mathcal{O}(R^4). \tag{3.33}$$

Lemma 3.10
For any finite-gap manifold $\mathcal{T}_{\mathbf{V}}^{\leq 2n}$ the corresponding frequency vector $\mathbf{W}(R)$ has the following asymptotics as $R = (R_1, \ldots, R_n) \to 0$:

$$W_j(R) = -\frac{1}{4}V_j^3 + \frac{3}{8\pi V_j}R_j^2 + \mathcal{O}(|R|^4),$$

and the result remains true with the R variables replaced with the r variables. ♠

Proof: The zeroth-order terms follows from (3.21). For small R, each W_j is an analytic function of the $R_l^2 = \mu_l$'s. If we apply (3.27) iteratively, we get that $W_j(0, \ldots, R_j, \ldots, 0)$ is the frequency of the one-gap solution w above, and thus (3.26) implies that

$$\frac{\partial W_j}{\partial \mu_j}(0) = \frac{3}{8\pi V_j}.$$

if we apply (3.27) again, we find that the function $W_j(0, \ldots, R_j, \ldots, R_l, \ldots, 0)$ is the first component of the frequency vector of a two-gap solution. Equation (3.32) gives us that

$$\frac{\partial W_j}{\partial \mu_l} = 0,$$

and the result follows.

■

3.3 Linearized Equations and their Floquet Solutions

3.3.1 Motivation

Most of this motivation comes straight from [Arn83].

Consider the ordinary differential equation

$$\dot{x} = A(t)x,$$

where $A(t)\colon \mathbb{C} \to \mathbb{C}$ is 2π-periodic in time. Since we have a 2π-periodic vector field, the solutions must also be 2π-periodic, and thus to solve the system, it suffices to know the **monodromy map**

$$\begin{aligned} M\colon \mathbb{C}^n &\to \mathbb{C}^n \\ x(0) &\mapsto x(2\pi). \end{aligned}$$

This is the analytic version of the Poincaré map idea in dynamical systems. There is a theorem

Theorem 3.11 (Floquet's Theorem)
If M is non-degenerate (in the sense of linear algebra), then there is a 2π-periodic change of variables $x = B(t)y$ so that

$$\dot{y} = \Lambda y.$$

Also, if the eigenvalues of Λ are λ_s, then the eigenvalues of M are $\mu_s = e^{2\pi\lambda_s}$. ♠

The situation is a little more complicated in the case where we have a quasi-periodic vector field. For example, assume we have the differential equation

$$\dot{x} = \Lambda x + v(t)x,$$

where $v(t)$ is small and quasi-periodic. Let's look for a change of variables

$$x = y + h(t)y,$$

with h small, to put our system in the form

$$\dot{y} = \Lambda y.$$

Then we calculate:

$$\begin{aligned} \dot{x} &= \dot{y} + \dot{h(t)}y + h(t)\dot{y} \\ &= \Lambda y + h\Lambda y + \dot{h}y. \end{aligned}$$

On the other hand,

$$\begin{aligned} \dot{x} &= \Lambda x + hx \\ &= \Lambda y + \Lambda hy + vy + vhy. \end{aligned}$$

If both h and v are small, we can neglect their product and get

$$\Lambda y + h\Lambda y + \dot{h}y = \Lambda y + \Lambda hy + vy,$$

or

$$h\Lambda y - \Lambda hy + \dot{h}y = vy,$$

or

$$[h, \Lambda] + \dot{h} = v. \tag{3.34}$$

This is known as the **homological equation**.

Recall that Λ and v are given, and we are solving for h. Of course, we may not be able to solve this equation for every choice of Λ, v.

We assume that Λ has a compete eigenbasis, i.e. that there are n vectors e_n such that

$$\Lambda e_n = \lambda_n e_n.$$

If we write h and v in Taylor series as

$$v = \sum_{k,n} v_{k,n} e^{ikt} e_n,$$
$$h = \sum_{k,n} h_{k,n} e^{ikt} e_n,$$

plug these into (3.34), and equate coefficients, we have

$$h_{k,n} = \frac{v_{k,n}}{ik - \lambda_n}.$$

Thus we have a resonance condition. If the frequencies of v are the same as the eigenvalues of Λ, we're in a bit of trouble. One way to look at this is that if we have a vector field v and a linear constant-coefficient piece Λ such that there are resonances, Equation (3.34) is not solvable. Another way to look at this is, given a linear, constant-coefficient piece Λ, we are restricted in the class of time-dependent perturbations v we can allow. We will see an analogy in the PDE's later.

In summary, we take a quasi-periodic vector field, and make it autonomous by a change of variables. We should note that the change of variables is essentially to use our quasi-periodic solutions to parameterize our new vector field. These solutions are known as **Floquet solutions**.

3.3.2 Floquet theory for Hamiltonian PDEs

We assume that we have a symplectic space (Z, α_2) with $Z = Z_d$ for some fixed d, and $\alpha_2 = \overline{J} dz \wedge dz$. We assume that we have a quasi-linear PDE

$$\dot{u} = J\nabla\mathcal{H}(u) = J(Au + \nabla H(u)) =: V_{\mathcal{H}}(u), \tag{3.35}$$

where $\operatorname{ord} A = d_A, \operatorname{ord} \nabla H = d_H < d_A, \operatorname{ord} J = d_J$, and we assume that $d \geq d_A/2$.

We also assume that we have an invariant $2n$-torus $\mathcal{T}^{2n} \subset Z$, where we parameterize

$$\mathcal{T}^{2n} = \Phi_0(R \times \mathbb{T}^n).$$

Recall that we also excise all of the singularities in this torus by restricting our action variables and get

$$\mathcal{T}_0^{2n} = \Phi_0(R_0 \times \mathbb{T}^n).$$

See Theorem 3.6 for a review of the properties of Φ_0 and R_0.

By part 3 of Theorem 3.6, we know we have a quasi-periodic solution

$$u_0(t) = u_0(t; r_0, z_0) = \Phi_0(w(t)),$$

where

$$w_0(t) = (r_0, z_0 + t\omega(r_0)).$$

We will linearize (3.35) about u_0 to get

$$\dot{v} = J(Av + (\nabla H)_*(u_0)v) =: JA_t v. \tag{3.36}$$

We denoted the flow maps of this linearized equation as $S^t_{\tau**}(u_0(\tau))$ if they exist. We make the further assumption that for all $u_0 \in \mathcal{T}_0^{2n}$, the maps $S^t_{\tau**}(u_0(\tau))$ are well-defined in Z_d for $-\infty < t, \tau < \infty$.

This is in analogy to the previous section in that, we now have a linear equation, but it nonautonomous, and, furthermore, the vector field depends quasi-periodically on time.

What we will show in this section is, if we assume that we have a "complete" system of Floquet solutions (in a sense to be made precise later), which is essentially a nondegeneracy condition on our PDE, then we can write down the change of variables that makes our PDE constant-coefficient.

It should be noted that this assumption is not restrictive. For example, Kuksin shows in [Kuk00] that any Lax-integrable equation has such a system.

Definition 3.8
*We say that $v(t)$ is a **Floquet solution** if there is a Ψ, a section of the complexified tangent bundle on Z, based at $\mathcal{T}_0^{2n}$, i.e. $\Psi\colon R_0 \times \mathbb{T}^n \to T^c Z|_{\mathcal{T}_0^{2n}}$, and a function $\nu(r)$ such that*

$$v(t) = v(t; r_0, z_0) = e^{i\nu(r_0)t}\Psi(w_0(t)),$$

where

$$w_0(t) = (r_0, z_0 + t\omega(r_0)).$$

*We also say that v is a **skew-orthogonal Floquet solution** is Ψ is a section of $T^{\perp c}\mathcal{T}_0^{2n}$, i.e. Ψ is a normal vector to our torus, instead of a tangent vector.* ♣

We should note that, given v, the function ν is not uniquely defined. For example, if we write

$$\Psi = e^{i\langle s,z\rangle}\Psi_1(r,z), \qquad s \in \mathbb{Z}^n,$$

then we can write

$$v = e^{i(\nu(r_0)+\langle\omega(r_0),s\rangle)t}e^{i\langle s,z_0\rangle}\Psi_1(w_0(t)),$$

so that $\nu(r)$ is really only defined up to the $\mathbb{Z}$-module $\omega(r)\cdot\mathbb{Z}^n$.

We will assume

1. We have a family $v_j(t)$ of Floquet solutions, such that different solutions have different exponents. Note that if v_j is a solution, then so is $\overline{v_j}$ with the exponent $-\overline{\nu_j}(r)$. In other words, we know our family of Floquet exponents is invariant with respect to the involution

$$\nu \mapsto -\overline{\nu}.$$

2. Our family of Floquet exponents is invariant with respect to the involution $\nu \mapsto \overline{\nu}$, which will imply that it is also invariant with respect to $\nu \mapsto -\nu$.
3. For any $j \neq k$, there is an $r \in R_0$ with $\nu_j \neq \nu_k$.
4. The ν_k are real-analytic on R, and Ψ_k extends to an analytic map $\Pi^c \times \{|\mathrm{Im}\, z| < \delta\} \to Z^c$.

Just to make our life easy, we will enumerate our solutions to that $\nu_{-j} = -\nu_j$.

Note that Ψ_j (restricted to an n-torus $T^n(r)$) are eigenvectors of S^1_{0*} acting on $T^{\perp c}\mathcal{T}_0^{2n}\big|_{T^n(r)}$, because

$$S^1_{0*}\Psi_j = e^{i\nu_j(r)}\Psi_j.$$

Assumption #4 above is crucial, and a bit of a cheat. In general, this will not be true. What can be shown in a general setting (e.g. Lax-integrable equations) is that by excising even more of our action variables, we can have, for all "large" indices ($|j| > j_1$), that the ν_j are analytic, by playing with some algebraic geometry (see [Kuk00]). This would cause us a lot of problem in the proofs later, but the results will be essentially the same. But we're making our lives easy here.

As simple example of what can go wrong is the 1-parameter family

$$B_a = \begin{pmatrix} 1 & -a \\ 1 & -1 \end{pmatrix}.$$

Although the eigenvalues of B_a are continuous functions of a, they aren't even differentiable. Of course, the differentiability fails only at one point, so by excising some of the a values, we can get the eigenvalues to be analytic.

3.3.3 Complete Floquet systems

Let $\{\phi_j\}$ be a symplectic Hilbert basis of the real Hilbert space Z_s, in that it respects the form, i.e. for $j > 0$,

$$\alpha_2[\phi_j, \phi_{-k}] = \langle \overline{J}\phi_j, \phi_{-k} \rangle = \mu_j \delta_{j,k},$$

with $\mu_j > 0$. We set $\mu_{-j} = -\mu_j$. Since the ϕ_k form a Hilbert space basis, we know that

$$\overline{J}\phi_k = \mu_k \phi_{-k}.$$

This means that with respect to the basis ϕ_k, the operator $\overline{J}$ is almost diagonal, in that the indices almost match up. Remember that $\operatorname{ord} \overline{J} = -d_J \leq 0$, and thus $\operatorname{ord} J = d_J \geq 0$.

Let's denote $\nu_j^J = \mu_j^{-1}$, and thus we have, by Theorem 2.10, that

$$C_1^{-1} k^{d_J} \leq \nu_k^J \leq C_1 k^{d_J}.$$

Now we define a new basis on Z^c as

$$\psi_j = \frac{1}{\sqrt{2}}(\phi_j - i\phi_{-j})$$
$$\psi_{-j} = \frac{1}{\sqrt{2}}(\phi_j + i\phi_{-j}) = \overline{\psi_j}.$$

We can easily check that

$$\alpha_2[\psi_j, \psi_{-k}] = i\delta_{j,k}\mu_j.$$

Using the definition of inner product in a complex Hilbert space, this means that

$$\overline{J}\psi_j = i\mu_j\psi_j,$$

or

$$J\psi_j = i\nu_j^J\psi_j.$$

Note that we have a basis which still respects the symplectic form, but for which J is now diagonal.

Assume that we have a system of Floquet solutions $\{v_j\}$, with the corresponding sections Ψ_j. For each $(r, z) \in R_0 \times \mathbb{T}^n$, we define the map Φ_1 as

$$\begin{aligned} \Phi_1(r,z)\colon Y_d^c &\to Z^c \\ \psi_j &\mapsto \Psi_j(r,z). \end{aligned}$$

We would like to define a complete family of Floquet solutions in the most natural sense, in that we can fill up our space with Floquet solutions. Our definition would be

Definition 3.9 (But not really the definition)
*A family of Floquet solutions $\{v_j\}$ is called **complete** if the exponents satisfy assumptions (1)–(4) above, and the sections form a skew-orthogonal basis of $T^{\perp}\mathcal{T}^{2n,c}$.* ♣

What we'll really define is

Definition 3.10
*A family of Floquet solutions $\{v_j\}$ is called **complete** if the exponents satisfy assumptions (1)–(4) above, and*

1. *They are skew-orthogonal Floquet solutions,*
2. (a) *$\beta_j = -i\alpha_2[\Psi_j, \Psi_{-j}]$ is independent of z,*

   (b) *there is a nonempty subdomain of* R_0 *where* $\beta_j \not\equiv 0$,

   (c) $\alpha_2[\Psi_j, \Psi_{-k}] = i\beta_j(r)\delta_{j,k}$,

3. *The* $\Psi_j(w)$ *are analytic in* w *and asymptotically close to* Ψ_j, *and the* $\nu_j(r)$ *are asymptotically close to constants, i.e.*

   (a) *The map* $\Phi_1(w)$ *is analytic in* w *and equals the identity up to a* Δ*-smoothing transformation, i.e.*

$$\|\Phi_1(w) - \mathbf{1}\|_{d,d+\Delta} \leq C, \qquad \text{for all } w \in Y^c.$$

   (b) *The* $\beta_j(r)$ *are analytic in* R_1^c, *and there exists* μ_j *such that*

$$|\beta_j(r) - \mu_j| \leq C_2 \, |j|^{-d_J - \Delta},$$

   (c) *We have that*

$$|\nu_j(r)| \leq C_3 \, |j|^{d_A + d_J}$$
$$|\nabla \nu_j(r)| \leq C_4 \, |j|^{\tilde{\Delta}},$$

   with $\tilde{\Delta} \leq d_A + d_J$.

♣

In words, assumption #2 essentially means that the Ψ_j act like a symplectic basis, and when they show up in the inner product, they are independent of z.

We have a lemma (proven in [Kuk00]) that says that these two definitions are the same, i.e.

Lemma 3.12
Under the assumptions of our definition, the Ψ_j form a complete (in the sense of span) skew-orthogonal basis of $T^\perp \mathcal{T}^{2n,c}$.

♠

3.3.4 Non-resonance conditions

Definition 3.11
Assume that we have $\{v_j\}$ *a family of Floquet solutions with exponents* $\nu_j(r)$, *which satisfy assumptions (1)–(4). Let us further assume that there is an open set* $O \subset R_0$ *such that for all* $s \in \mathbb{Z}^n$, *and for* $j, k \in \mathbb{Z}$ *with* $j \neq -k$, *then*

$$\langle \omega(r), s \rangle + \nu_j(r) \not\equiv 0, \text{ in } O,$$
$$\langle \omega(r), s \rangle + \nu_j(r) + \nu_k(r) \not\equiv 0, \text{ in } O,$$

where (recall that) $\omega(r)$ *comes from the pullback of the equation to* $R_0 \times \mathbb{T}^n$,

$$\dot{r} = 0,$$
$$\dot{z} = \omega(r).$$

Note that we can also restate these conditions as

$$\nu_j(r) \not\equiv 0,$$
$$\nu_j \not\equiv \nu_k,$$

if we think of these as lying in $\mathbb{R}^n/(\omega(r) \cdot \mathbb{Z}^n)$. *If this is true, we say the exponents are* ***non-resonant***.

♣

Note that the definition has to only hold on an open set, but we get a lot for free, simply by analytic continuation. If an analytic function is equal to 0 on an open set, then it must be 0 almost everywhere, which we will state in the following

Lemma 3.13
If a system of Floquet exponents is non-resonant, then each resonance function is not zero almost everywhere. ♠

Definition 3.12
An Floquet family which satisfies assumptions (1)–(4) on the exponents, satisfies (1)–(3) of the completeness definition, and satisfies the non-resonance condition is called ***complete non-resonant****.* ♣

It is shown in [Kuk00] that if the non-resonance condition is satisfied, then we also have (1), (2a), and (2c) of the completeness definition. We can summarize this in the following:

Corollary 3.14
If we have a Floquet family $\{v_j\}$ which satisfies assumptions (1)–(4) on the exponents, (2b) and (3) of the completeness definition, and the non-resonance condition, then it is a complete non-resonant family, i.e. gives a family of sections which are skew-orthogonal to $\mathcal{T}^{2n}$, span the orthogonal complement, and are non-resonant. ♠

This corollary is actually a significant improvement. The reason is that it is typically very difficult to check assumption (2) in the completeness definition, since it is a condition on the Ψ_j. On the other hand, non-resonance is relatively easy to check, since it's a restriction on the exponents.

The next theorem is the culmination of this whole section, in that we state and prove the analogous result to Floquet's Theorem for ODEs.

Theorem 3.15 (The Badness!)
Given a complete non-resonant Floquet family, we can normalize Ψ_j, ψ_j such that

$$\alpha_2[\phi_j, \phi_{-k}] = i\mu_k\delta_{j,k}$$

still holds, and $J\psi_j = i\nu_j^J\psi_j$. Furthermore,

1. for all (r, z), the map $\Phi_1(r, z)$ defines a symplectic isomorphism

$$\Phi_1(r, z)\colon Y^c \to T^{\perp}_{\Phi_0(r,z)}\mathcal{T}^{2n,c}$$

   which is analytic in r and z.

2. The nonautonomous map Φ_1 sends solutions $y(t)$ of the autonomous Hamiltonian equation

$$\dot{y} = JB_r y, \qquad y \in Y^c \tag{3.37}$$

   to solutions of (3.36). Also, B_r is self-adjoint, $\operatorname{ord} B_r = \operatorname{ord} A$, and

$$\operatorname{ord} \nabla_r B_r \leq \tilde{\Delta} - d_J.$$

3. JB_r is diagonal in the basis formed by the ψ_j, and, for all j,

$$JB_r\psi_j = i\nu_j(r)\psi_j.$$

♠

Before we prove this, we should again remark that assumption #4 on the exponents, the assumption of analyticity, will make this proof much, much easier. The general case is dealt with in [Kuk00], and is much more difficult.

Proof: The numbers $\beta_j(r)$ are real, nonzero, and odd in j. Assumption (3b) of completeness gives us that

$$|\beta_j(r)| \geq \frac{1}{2}\mu_j, \qquad \text{as } j \to \infty,$$

and thus is true for $|j| > j_1$ for some j_1. For $|j| > j_1$, the function $\operatorname{sgn} j \cdot \beta_j > 0$. If, for some j with $|j| \leq j_1$ the function $\operatorname{sgn} j \cdot \beta_j < 0$, then simply interchange v_j and v_{-j} (because we didn't want 'em like that anyway).

We know that $\beta_j(r)\nu_j^J$ is positive (since the map $j \mapsto \nu_j^J$ is odd in j). All we have to do now is normalize: Replace Ψ_j by $(\nu_j\beta_j(r))^{-1/2}\,\Psi_j$, and we can check that

$$\alpha_2[\Psi_j, \Psi_{-j}] = \alpha[\psi_j, \psi_{-j}].$$

Define the linear operator B_r on this basis as

$$B_r\psi_j = \frac{\nu_j(r)}{\nu_j^J}\psi_j,$$

and extend linearly,so that

$$JB_r\psi_j = i\nu_j(r)\psi_j.$$

Clearly, we can define B_j in a self-adjoint way, and we are done.

■

3.4 A normal form theorem

To summarize, what we did in the last section was put new coordinates on the invariant torus $\mathcal{T}^{2n}$ to put the equation in constant coefficient form. What we will do here is put a skew-orthogonal coordinate neat the invariant torus, and show that we can put the Hamiltonian into a nice normal form near the invariant torus.

Recall that we have action-angle variables (p, q) on the torus, and the previous section says that we have coordinates to give us

$$\dot{p} = 0 \tag{3.38}$$
$$\dot{q} = \nabla h = \omega \tag{3.39}$$
$$\dot{y} = JB_p y. \tag{3.40}$$

Let's assume that we have a complete Floquet family, so that all of Z is filled by either p, q, or y coordinates. We define

$$\alpha_2^Y := \alpha_2|_Y\ ,$$

and we know in these coordinates that

$$\alpha_2^Y = \overline{J}dy \wedge dy.$$

Recall that both Φ_0 and Φ_1 are functions of p, q, so we define a function

$$\Phi(p, q, y) = \Phi_0(p, q) + \Phi_1(p, q)y.$$

Since

$$\begin{aligned} \Phi|_{W\times\{0\}} &= \Phi_0 \\ \Phi|_{\{w\}\times\mathcal{O}_\delta(y)} &= \Phi_1(w) \text{ (up to translation),} \end{aligned}$$

we know that Φ is a symplectomorphism, and if we define $\omega_2\Phi^*\alpha_2$, we know that $\omega_2 = dp \wedge dq$ on the set $\{y = 0\} = \mathcal{T}_0^{2n}$.

We would like to say that

$$\omega_2 = dp \wedge dq \oplus \alpha_2^Y,$$

but this isn't exactly true. We can say that they differ by an exact piece, as in the following lemma from [Kuk00]:

Lemma 3.16
If $\omega_2 = \Phi^* \alpha_2$, then

$$\omega_2 = dp \wedge dq \oplus \alpha_2^Y + d(L(w,y)\, dw),$$

where

$$L = \frac{1}{2}\alpha_2 \left[\Phi_1(w)y, \nabla_w \Phi_1(w)y\right]. \qquad \spadesuit$$

We want to see what happens to the Hamiltonian under this change of variables. We will denote $w = (p,q) \in W$. Let

$$\Phi^t = \Phi_1(p, q + t\nabla h).$$

Note that $\Phi^0 = \Phi_1(w)$, and that the Floquet-like theorem from the last section says that

$$v(t) = \Phi^t y(t),$$

where y solves (3.38) and v solves (3.36).

Recall that the Hamiltonian

$$\mathcal{H}(v) = \frac{1}{2}\langle Av, v\rangle + H(v).$$

If we change variables, we have

$$\mathcal{H} \circ \Phi = \frac{1}{2}\langle A\Phi_0, \Phi_0\rangle + \langle A\Phi_0, \Phi_1 y\rangle + \frac{1}{2}\langle A\Phi_1 y, \Phi_1 y\rangle + H(\Phi).$$

If we expand this in y, then we have

$$\begin{aligned}
&\left(\frac{1}{2}\langle A\Phi_0, \Phi_0\rangle + H(\Phi_0)\right) + \\
&+ \left(\langle A\Phi_0, \Phi_1 y\rangle + \langle \nabla H(\Phi_0), \Phi_1 y\rangle\right) + \\
&+ \left(\frac{1}{2}\langle A\Phi_1 y, \Phi_1 y\rangle + H(\Phi) - H(\Phi_0) - \langle \nabla H(\Phi_0), \Phi_1 y\rangle\right).
\end{aligned}$$

The constant term is just the Hamiltonian on the torus (which we know is integrable). The linear term is

$$\nabla_y (\mathcal{H} \circ \Phi)\big|_{y=0}.$$

Since the torus is invariant, this must be 0. Thus we have

$$\mathcal{H} \circ \Phi = h(p) + \frac{1}{2}\langle B_p y, y\rangle + h_2(p,q,y),$$

where $h_2 = \mathcal{O}(\|y\|^2)$.

If we expand h_2 in y, we get

$$h_2(p,q,y) = \frac{1}{2}\langle \mathcal{B}(p,q)y, y\rangle + h_3(p,q,y),$$

where $h_3 = \mathcal{O}(\|y\|^3)$.

It is proven in [Kuk00] that

Theorem 3.17
There is a symplectic transformation ϕ that gives

$$H \circ \Phi \circ \phi = h(p) + \frac{1}{2}\langle B_p y, y\rangle + \frac{1}{2}\langle \mathcal{B}(p,q)y, y\rangle + h_3(p,q,y), \tag{3.41}$$

with $h_3 = \mathcal{O}(\|y\|^3)$, and, furthermore, in these coordinates,

$$\alpha_2 = dp \wedge dq \oplus \alpha_2^Y. \qquad \spadesuit$$

We see that ϕ must be $\mathcal{O}(y^2)$ close to $\mathbf{1}$. Thus $\phi_*(w,0)|_{\{0\}\times Y} = \mathbf{1}$. Then we know that

$$(\Phi \circ \phi)_* \, (w(t),0)y(t) = (\Phi_*(w(t),0))\, y(t)$$

A priori, the map

$$y \mapsto (\Phi_*(w(t),0))y$$

sends solutions of (3.38) to solutions of (3.36). On the other hand $\Phi \circ \phi$ sends solutions of the system with Hamiltonian (3.41) to solutions of (3.35). Thus $(\Phi \circ \phi)_*(w)$ sends solutions of

$$\dot{y} = J(B_p + \mathcal{B}(p,q))y \tag{3.42}$$

to solutions of (3.35), and $\phi_*(w)$ sends solutions of (3.42) to solutions of (3.38). Therefore

$$JB_p y = J(B_p + \mathcal{B}(w(t)))y,$$

or

$$\mathcal{B}(p,q) \equiv 0.$$

Index

Bibliography

[Arn] V. I. Arnold. *Mathematical Method in Classical Mechanics.* ?, ?

[Arn83] V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations.* Springer, 1983.

[AS] Ablowitz and Siegen. *Solitons and the Inverse Scattering Transform.*

[GT87] J. Garnet and E. Trubowitz. Gaps and bands on one-dimensional schrödinger operators ii. *Comment. Math Helvetici*, 1987.

[Ito89] H. Ito. Conbergence of birkhoff normal forms for integrable systems. *Comment. Math. Helvetici*, 1989.

[Kuk00] Sergei Kuksin. *Analysis of Hamiltonian PDE.* 2000.

[Mc77] V. A. Marˇ cenko. *Sturn-Liouville operators and applications.* Naukova Dumka, 1977.

[Mos78] J. Moser. *Various aspects of integrable hamiltonian systems.* Bressansone, 1978.

[MT76] H. P. McKean and E. Trubowitz. Hill's operator and hyperelliptic function theory in the presence of infinitely many branch points. *Comm. Pure Appl. Math.*, 1976.

[Vey78] J. Vey. Sur certain systèmes dynamiques séperables. *Amer. J. Math.*, 1978.

[Way96] C. Eugene Wayne. An introduction to kam theory. *Lectures in Applied Mathematics*, 1996.