Introduction to KAM theory

A.A. Glutsyuk

May 21, 2025

HSE, spring semester, 2024/2025

Contents

1 Linearization of germs of conformal maps at fixed point

Consider a germ of conformal map $f : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ at fixed point 0:

$$f(z) = \lambda z + O(z^2), \ \lambda \in \mathbb{C} \setminus \{0\}. \tag{1.1}$$

Question 1.1 Does there exist a germ of conformal map $h : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$ that conjugates f to its linear part, i.e., for which $h \circ f \circ h^{-1}(z) = \lambda z$? Or equivalently,

$$h \circ f(z) = \lambda h(z). \tag{1.2}$$

If such h exists, then the germ f is called *linearizable.*

First we give a proof of positive answer in the case, when $|\lambda| \neq 0$, which is done by showing that the conjugating map h can be found as a fixed point of a contracting map of an appropriate metric space.

The case, when $|\lambda| = 1$, i.e., $\lambda = e^{2\pi i\theta}$, $\theta \in \mathbb{R}$, is much more difficult. It is easy to show that for every fixed rational $\theta \in \mathbb{Q}$ a typical f with multiplier λ is not linearizable. Sufficient conditions on θ guaranteeing linearizability of every f with given multiplier $\lambda = e^{2\pi i\theta}$ were obtained by K.Siegel and A.Bruno. Siegel Theorem states that f is linearizable whenever θ is a Diophantine number. Bruno Theorem is the same in more general case: under a weaker Bruno Diophantine condition. A Theorem of J.-C.Yoccoz (Fields medal, 1994) states that Bruno's condition is sharp: *for every* $\theta \in \mathbb{R}$

that does not satisfy Bruno's Diophantine condition there exists an f with multiplier $\lambda = e^{2\pi i\theta}$ that is not linearizable.

Below we state and prove Siegel Theorem and state Bruno's and Yoccoz's results without proofs.

1.1 Hyperbolic case: $|\lambda| \neq 1$.

Theorem 1.2 *Every germ of conformal mapping (1.1) with $|\lambda| \neq 1$ is conformally conjugated to its linear part. More precisely, there exists a unique conformal germ $h : (\mathbb{C}, 0) \to (\mathbb{C}, 0)$, $h(0) = 0$, $h'(0) = 1$, satisfying (1.2).*

Proof Without loss of generality we consider that $0 < |\lambda| < 1$ (replacing f by f^{-1}, if this is not the case). Equation (1.2) is equivalent to the statement that h is a fixed point of the transformation

$$\mathcal{L} : h \mapsto \lambda^{-1} h \circ f.$$

We will show that $\mathcal{L}$ is a contraction in appropriate complete metric space and hence, has a unique fixed point there.

Fix a $\mu > 0$ such that

$$0 < \mu^2 < |\lambda| < \mu < 1. \tag{1.3}$$

Fix an $r > 0$ such that f is holomorphic on $\overline{D}_r$ and

$$|f(z)| \leq \mu |z| \text{ whenever } z \in \overline{D}_r. \tag{1.4}$$

In particular, (1.4) implies that $f(\overline{D}_r) \subset D_r$.

For every function $q(z)$ holomorphic on D_r and continuous on $\overline{D}_r$ such that $q(0) = q'(0) = 0$ set

$$||q|| := \sup_{|z| \leq r} \frac{|q(z)|}{|z|^2}.$$

Let M denote the space of functions h holomorphic on D_r and continuous on $\overline{D}_r$ such that

$$h(0) = 0, \ h'(0) = 1,$$

equipped with the distance $dist(h_1, h_2) = ||h_1 - h_2||$. This is a complete metric space. Indeed, a sequence fundamental in the norm converges uniformly, by definition. Hence, its limit is holomorphic, by Weierstrass Theorem, and vanishes at 0. The derivatives also converge uniformly in compact set to the derivative of the limit, by Cauchy integral formula for the derivative and convergence of the function. Therefore, the limit has unit derivative at 0. Finally, the limit of a converging sequence is an element of the space M, and hence, M is complete.

Proposition 1.3 $\mathcal{L}(M) \subset M$.

Proof If $h(0) = 0$, then $(\mathcal{L}h)(0) = 0$ and $(\mathcal{L}h)'(0) = h'(0)$. If h is holomorphic on D_r and continuous on $\overline{D}_r$, then so is the composition $h \circ f$, since f is holomorphic on $\overline{D}_r$ and $f(\overline{D}_r) \subset D_r$. This implies that $\mathcal{L}$ preserves the space M and proves the proposition. □

Proposition 1.4 $||\mathcal{L}h_1 - \mathcal{L}h_2|| \leq \nu||h_1 - h_2||$, $\nu = |\lambda|^{-1}\mu^2 < 1$.

Proof The operator $\mathcal{L}$ being linear, it suffices to show that $||\mathcal{L}q|| \leq \nu||q||$ for every q as above. One has

$$\frac{|(\mathcal{L}q)(z)|}{|z^2|} = |\lambda|^{-1}\frac{|q(f(z))|}{|f(z)|^2}\frac{|f(z)|^2}{|z|^2} \leq |\lambda|^{-1}||q||\mu^2,$$

by definition, (1.4) and since $f(z) \in D_r$ whenever $z \in \overline{D}_r$. This implies that the norm of the image $\mathcal{L}q$ is no greater than $\nu||q||$. The proposition is proved. □

The two latter propositions together imply that $\mathcal{L} : M \to M$ is a contraction. Hence, $\mathcal{L}$ has a unique fixed point $h \in M$, which obviously represents a conjugating germ we are looking for. Its uniqueness follows from the above uniqueness of fixed point and the fact that the above argument holds for every r small enough. This proves Theorem 1.2. □

1.2 Siegel Theorem

Definition 1.5 A number $\theta \in \mathbb{R}$ is *Diophantine,* if there exist $C, \gamma > 0$ such that for every rational number $\frac{m}{n}$, $(m, n) = 1$, one has

$$|\theta - \frac{m}{n}| > \frac{C}{|n|^\gamma}; \tag{1.5}$$

in this case it is called (C, γ)-Diophantine. A number is γ-Diophantine, if it is (C, γ)-Diophantine for some C.

Exercise 1.6 Prove that for $\gamma > 2$ the complement of the set of γ-Diophantine numbers has Lebesgue measure zero, i.e., for $\gamma > 2$ typical numbers are γ-Diophantine.

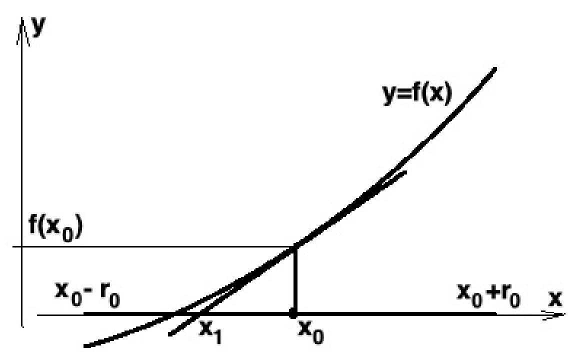
Theorem 1.7 *(Siegel). For every Diophantine number θ every conformal germ f with multiplier $e^{2\pi i\theta}$ is linearizable.*

The proof of Theorem 1.7 is based on an infinite-dimensional version of Newton method of finding root of a functional equation.

1.3 The Newton method with estimates

This is the method of finding root of a strictly monotonous function $f(x)$ on a segment $I_0 := [x_0 - r_0, x_0 + r_0]$ by Newton approximations, see Fig. 1:

$$x_1 := x_0 - \frac{f(x_0)}{f'(x_0)},\ x_{j+1} := x_j - \frac{f(x_j)}{f'(x_j)},\ j = 1, 2, \dots. \tag{1.6}$$



The next lemma not only provides sufficient conditions for convergence, but also gives useful estimates on the step of Newton method.

Lemma 1.8 *Every universal constant $c \geq 3$ satisfies the following statements. Let $r_0 > 0$, $\eta_0 \in (0, \frac{1}{4}]$, $\delta_0 \in (0, \eta_0^{4c})$. Let $f(x)$ be a C^2-smooth function on $I_0 := [x_0 - r_0, x_0 + r_0]$ satisfying the following inequalities:*

$$|f(x_0)| \leq \delta_0 r_0, \tag{1.7}$$

and for every $x \in [x_0 - r_0, x_0 + r_0]$ one has

$$|f'(x)| \geq \eta_0,\ |f''(x)| \leq \frac{1}{\eta_0 r_0}. \tag{1.8}$$

Set

$$\eta_1 = \frac{\eta_0}{2}, \quad r_1 = r_0(1 - \eta_0), \quad \delta_1 = \delta_0^2 \eta_0^{-c}. \tag{1.9}$$

Then the above x_1 given by (1.6) satisfies the following inequalities:

$$|x_1 - x_0| \leq \eta_0 r_0, \tag{1.10}$$

$$|f(x_1)| \leq \delta_1 r_1, \tag{1.11}$$

$$I_1 = [x_1 - r_1, x_1 + r_1] \subset I_0, \tag{1.12}$$

$$|f'(x)| \geq \eta_1, \ |f''(x)| \leq \frac{1}{\eta_1 r_1}. \tag{1.13}$$

Moreover,

$$\delta_1 < \eta_1^{4c}, \quad \delta_1 < \delta_0^{\frac{7}{4}}. \tag{1.14}$$

Proof One has

$$|x_1 - x_0| = \frac{|f(x_0)|}{|f'(x_0)|} \leq \frac{\delta_0 r_0}{\eta_0} \leq \eta_0^{c-1} r_0 \leq \eta_0 r_0,$$

$$f(x_1) = f(x_0) + f'(x_0)(x_1 - x_0) + R,$$

$$|R| \leq \frac{1}{2} \max_{[x_0, x_1]} |f''(x)|(x_1 - x_0)^2 \leq \frac{1}{2\eta_0 r_0}(x_1 - x_0)^2,$$

$$f(x_0) + f'(x_0)(x_1 - x_0) = 0,$$

by Taylor series remainder estimate and definition. Therefore,

$$|f(x_1)| = |R| \leq \frac{1}{2\eta_0 r_0}(x_1 - x_0)^2 = \frac{(f(x_0))^2}{2\eta_0 r_0 (f'(x_0))^2}$$

$$\leq \frac{1}{2}\delta_0^2 \eta_0^{-3} r_0 \leq \delta_0^2 \eta_0^{-3} r_1 \leq \delta_0^2 \eta_0^{-c} r_1,$$

for every $c \geq 3$. One has

$$\delta_1 = \delta_0^2 \eta_0^{-c} \leq \eta_0^{7c} < \eta_1^{4c}.$$

This proves (1.10), (1.11) and the first inequality in (1.14). Inclusion (1.12) is implied by the immediate inequality $x_1 + r_1 < x_0 + r_1 < x_0 + r_0$ and by the inequality

$$x_1 - r_1 \geq x_0 - \eta_0 r_0 - r_1 = x_0 - (\eta_0 + (1 - \eta_0)) r_0 = x_0 - r_0.$$

Inequalities (1.13) follows from (1.8) and the inclusion $I_1 \subset I_0$. The second inequality in (1.14) follows from definition and the inequality $\delta_0 < \eta_0^{4c}$. □

1.4 Background material: Cauchy bounds in one variable

In the proof of Siegel Theorem and analytic KAM theorem we use Cauchy bounds on holomorphic functions given by the next well-known theorem.

Theorem 1.9 (Cauchy bounds). *Let $U \subset \mathbb{C}$ be a domain, $z_0 \in U$ and $r > 0$ be such that the disk*

$$D_r(z_0) := \{|z - z_0| < r\}$$

is contained in U. For every bounded holomorphic function $f : U \to \mathbb{C}$ with Taylor series

$$f(z) = \sum_{k=0}^{+\infty} a_k (z - z_0)^k$$

for every $k \in \mathbb{Z}_{\geq 0}$ one has

$$|a_k| \leq \frac{\sup_U |f|}{r^k}. \tag{1.15}$$

Proof The Taylor series converges uniformly on every smaller disk $\overline{D}_\rho(z_0)$, $\rho < r$. Then each Taylor coefficient a_k can be found by well-known Cauchy integral formula

$$a_k = \frac{1}{2\pi i} \oint_{|\zeta - z_0| = \rho} \frac{f(\zeta)}{(\zeta - z_0)^{k+1}} d\zeta \tag{1.16}$$

(residue formula for integral: a_k is the residue at z_0 of the form $(z - z_0)^{-(k+1)} f(z) dz$). The module of the right-hand side in (1.16) is bounded from above by $\frac{\sup_U |f|}{\rho^k}$. This proves (1.15) with r replaced by arbitrary $\rho \in (0, r)$, and hence, implies (1.15). □

1.5 Proof of Siegel Theorem

Recall that we suppose that

$$\lambda = e^{2\pi i \theta}, \ |\theta - \frac{m}{n}| > \frac{c_0}{|n|^\gamma} \text{ for every } \frac{m}{n} \in \mathbb{Q}, \ (m, n) = 1. \tag{1.17}$$

Let $f_0(z)$ be a germ of conformal map at the origin, $f_0(0) = 0$, $f_0'(0) = \lambda$:

$$f_0(z) = \lambda z + \widehat{f}_0(z), \quad \widehat{f}_0(z) = \sum_{j=2}^{\infty} a_j z^j. \tag{1.18}$$

Applying conjugacy with rescaling, i.e., replacing $f(z)$ by

$$g_\mu(z) = \mu^{-1} f(\mu z) = \lambda z + \mu^{-1}\widehat{f}_0(\mu z) = \lambda z + \sum_{j=2}^{+\infty} \mu^{j-1} a_j z^j, \ \mu \in (0,1),$$

one can make the nonlinear part well-defined on an arbitrarily given disk $\overline{D}_r$ (μ should be small depending on r) and arbitrarily small on $\overline{D}_r$. This follows from the asymptotics and inequality

$$\widehat{f}_0(z) = O(z^2), \quad |\widehat{f}_0(\mu z)| \leq \nu\mu^2 \text{ for every } \mu \in (0,1).$$

Here ν is a constant depending only on the germ f.

Let us choose some $r_0 > 0$, $\eta_0 \in (0, \frac{1}{4})$ and $\delta_0 \in (0,1)$ (which eventually will be very small) and normalize f_0 by rescaling so that f_0 be holomorphic on D_{r_0}, continuous on $\overline{D}_{r_0}$ and satisfying the upper bound

$$|\widehat{f}_0(z)| = |f_0(z) - \lambda z| \leq \delta_0 r_0 \text{ for every } z \in \overline{D}_{r_0}. \tag{1.19}$$

Finding ψ that conjugates f_0 to its linear part is equivalent to solving the following functional equation in infinite-dimensional space:

$$G(\psi) := f_0 \circ \psi(z) - \psi(\lambda z) = 0. \tag{1.20}$$

To solve it, we apply a multidimensional version of the Newton method, starting with $\psi_0 = Id$, constructing next Newton approximation ψ_1 of ψ etc. To prove its convergence, we prove the following estimates on Newton method step, analogous to those of Lemma 1.8 but more tricky.

Lemma 1.10 *There exists a universal constant $c = c(c_0, \gamma) \geq 3$ satisfying the following statements. Let $\lambda = e^{2\pi i\theta}$ be the same, as in (1.17). Take arbitrary numbers*

$$r_0 > 0, \quad \eta_0 \in (0, \frac{1}{4}), \ \delta_0 \in (0, \eta_0^{4c}).$$

Let $f_0(z)$ be the same, as in (1.18), holomorphic on D_{r_0}, continuous on $\overline{D}_{r_0}$ and satisfying inequality (1.19). Set

$$r_1 = (1-\eta_0) r_0, \ \delta_1 = \delta_0^2 \eta_0^{-c}, \quad \eta_1 = \frac{\eta_0}{2}.$$

Then there exists an injective map $\psi_1 : \overline{D}_{r_1} \to D_{r_0}$ holomorphic on D_{r_1} and continuous on $\overline{D}_{r_1}$, such that the conjugated map

$$f_1(z) := \psi_1^{-1} \circ f_0 \circ \psi_1(z) \tag{1.21}$$

is holomorphic on D_{r_1}, continuous on $\overline{D}_{r_1}$ and satisfies the estimate

$$|\widehat{f}_1(z)| = |f_1(z) - \lambda z| \leq \delta_1 r_1 \textit{ for every } z \in \overline{D}_{r_1}, \tag{1.22}$$

and the map ψ_1 satisfies the estimate

$$|\psi_1(z) - z| \leq \delta_0 \eta_0^{-c} \ \textit{ for every } \ z \in \overline{D}_{r_1}. \tag{1.23}$$

Moreover, one has

$$\delta_1 < \eta_1^{4c}, \quad \delta_1 < \delta_0^{\frac{7}{4}}. \tag{1.24}$$

Proof of Siegel Theorem modulo Lemma 1.10. Define by induction

$$r_{k+1} = (1 - \eta_k) r_k, \ \delta_{k+1} = \delta_k^2 \eta_k^{-c}, \ \eta_{k+1} = \frac{\eta_k}{2}.$$

The sequence η_k being a decreasing geometric progression, the sequence r_k is decreasing and converges to some limit $r > 0$. The sequence δ_k decreases to zero superexponentially, by (1.24). Let ψ_k be the maps constructed successively by applying Lemma 1.10: ψ_2 satisfies (1.23) with δ_0, η_0 replaced by δ_1, η_1 etc. By construction

$$|\psi_k(z) - z| \leq \delta_{k-1} \eta_{k-1}^{-c} \ \text{ on } \overline{D}_{r_k} \Supset \overline{D}_r.$$

The sequence of the latter right-hand sides tends to zero superexponentially, by (1.24) and since η_k is a geometric progression. Therefore, the function sequence $\psi_k(z)$ converges uniformly and superexponentially to z on $\overline{D}_r$. Set now

$$\psi(z) := \lim_{k \to +\infty} H_k(z), \quad H_k(z) := \psi_0 \circ \psi_1 \circ \dots \psi_k(z).$$

For every k the above composition H_k is a well-defined map $\overline{D}_r \to D_{r_0}$, since ψ_k sends $\overline{D}_{r_k}$ to $D_{r_{k-1}}$.

Claim 1. *The sequence H_k converges uniformly with derivatives on $\overline{D}_{\frac{r}{2}}$ to a map $\psi(z)$ holomorphic on $D_{\frac{r}{2}}$ and continuous on $\overline{D}_{\frac{r}{2}}$.*

Proof Indeed, the derivatives H_k' are uniformly bounded on the latter disk, since $|H_k(z)| \leq r_0$ and by the Cauchy bound (1.15):

$$|H_k'(z)| \leq \frac{\max_{|\zeta| \leq r} |H_k(\zeta)|}{r - |z|} \leq \frac{2r_0}{r} \ \text{ whenever } \ |z| \leq \frac{r}{2}.$$

Therefore, for every $z \in \overline{D}_{\frac{r}{2}}$ one has

$$|H_{k+1}(z) - H_k(z)| \leq \frac{2r_0}{r} \max_{|z| \leq \frac{r}{2}} |\psi_{k+1}(z) - z| \leq \frac{2r_0}{r} \delta_k \eta_k^{-c}.$$

The latter right-hand side tends to zero, as $k \to \infty$. This proves Claim 1. □

Clearly ψ conjugates f_0 to the limit $\lim_{k\to\infty} f_k$ on $D_{\frac{r}{2}}$. The latter limit clearly coincides with λz: the map H_k conjugates f_0 with f_k, and $|f_k(z) - \lambda z| \leq \delta_{k-1} r_{k-1} \to 0$ on $\overline{D}_r$. Thus, ψ is the conjugacy from Siegel Theorem we are looking for. This proves Siegel Theorem modulo Lemma 1.10. □

Proof of Lemma 1.10. Proof of (1.24) repeats the proof of (1.14).

Without loss of generality we can and will consider that $r_0 = 1$, applying rescaling by $\mu = r_0$: the rescaling does not change the bound (1.19), which now becomes

$$|\widehat{f}_0(z)| \leq \delta_0, \quad \widehat{f}_0(z) := f_0(z) - \lambda z = \sum_{j=2}^{\infty} a_j z^j. \tag{1.25}$$

Write

$$\psi_1(z) = z + \widehat{\psi}_1(z), \quad \widehat{\psi}_1(z) = \sum_{j=2}^{\infty} b_j z^j.$$

Our goal is to find a conjugacy ψ, i.e., a map satisfying $f_0 \circ \psi(z) = \psi(\lambda z)$. Applying a Newton method step presented below will yield the first Newton approximation ψ_1 of ψ so that the difference $f_0 \circ \psi_1(z) - \psi_1(\lambda z)$ admits a good upper bound. The terms of the latter difference are given by

$$f_0 \circ \psi_1(z) = \lambda(z + \widehat{\psi}_1(z)) + \widehat{f}_0(z) + R(z), \quad R(z) = \widehat{f}_0(z + \widehat{\psi}_1(z)) - \widehat{f}_0(z), \tag{1.26}$$

$$\psi_1(\lambda z) = \lambda z + \widehat{\psi}_1(\lambda z). \tag{1.27}$$

Equating (1.26) and (2.5) yields

$$\Lambda \widehat{\psi}_1(z) = \widehat{f}_0(z) + R(z), \quad \Lambda \widehat{\psi}_1(z) := \widehat{\psi}_1(\lambda z) - \lambda \widehat{\psi}_1(z). \tag{1.28}$$

Definition 1.11 The *homological equation* is the following equation on $\widehat{\psi}_1(z)$:

$$\Lambda \widehat{\psi}_1(z) = \widehat{f}_0(z). \tag{1.29}$$

In what follows we prove the statements of the lemma for the map $\psi_1(z) = z + \widehat{\psi}_1(z)$, where $\widehat{\psi}_1$ is a solution of the homological equation.

Proposition 1.12 *The homological equation has a unique formal solution, i.e., a unique solution of type of a formal Taylor series; it is given by*

$$\widehat{\psi}_1(z) = \sum_{j=2}^{\infty} b_j z^j, \; b_j = \frac{a_j}{\lambda(\lambda^{j-1} - 1)}. \tag{1.30}$$

The latter series converges uniformly on $\overline{D}_\nu$, $\nu = 1 - \frac{\eta_0}{4}$, and satisfies (1.23) on the latter disk with appropriate universal constant $c = c(c_0, \gamma)$.

Proof Uniqueness and formula (1.30) are obvious. Let us prove convergence and (1.23). One has $|\lambda| = 1$, $|\lambda^{j-1} - 1| = |e^{2\pi i(j-1)\theta} - 1|$. The latter right-hand side is the length of the chord connecting 1 to the point on the unit circle with argument $2\pi i(j-1)\theta(mod 2\pi\mathbb{Z})$. The chord length is bigger than half of the smallest circular arc length for an arc connecting the same points. The smallest arc length is clearly equal to 2π times the minimal distance of the number $(j-1)\theta$ to a point of the integer lattice. Thus, it is greater than $2\pi\frac{c_0}{|j-1|^{\gamma-1}} \geq \frac{2\pi c_0}{j^\gamma}$. Finally,

$$|b_j| = \frac{|a_j|}{\lambda(\lambda^{j-1} - 1)} \leq (\pi c_0)^{-1} j^\gamma |a_j|. \tag{1.31}$$

Set

$$\nu = 1 - \frac{\eta_0}{4}, \quad u := -\ln\nu \in (0, \frac{1}{4}); \quad \nu = e^{-u}.$$

On the other hand, $|a_j| \leq \max_{\overline{D}_1} |\widehat{f}_0| \leq \delta_0$, by Cauchy bound (1.15) and (1.25). Therefore, the following inequality holds on $\overline{D}_\nu$:

$$\sum_{j=2}^{\infty} |b_j z^j| \leq (\pi c_0)^{-1}\delta_0 \sum_{j=2}^{\infty} j^\gamma \nu^j = (\pi c_0)^{-1}\delta_0 \sum_{j=2}^{\infty} j^\gamma e^{-ju}. \tag{1.32}$$

Proposition 1.13 (The main power series inequality). *For every $\gamma > 0$ there exists a constant $c_2 = c_2(\gamma) > 0$ such that for $u \in (0, \frac{1}{4})$ one has*

$$\sum_{j=2}^{\infty} j^\gamma e^{-ju} \leq \frac{c_2}{u^{\gamma+1}}. \tag{1.33}$$

Proof Multiplying the above series by $u^{\gamma+1}$ transforms it to an integral sum with step u for the integral $\int_0^{+\infty} v^\gamma e^{-v}$. The latter integral sum is uniformly bounded in u varying on any finite segment in $\mathbb{R}_{\geq 0}$. Denoting by c_2 its uniform upper bound on the segment $u \in [0, \frac{1}{4}]$ we get (1.33). □

Let us now prove inequality (1.23) on $\overline{D}_\nu \Supset \overline{D}_{r_1}$. Substituting (1.33) to the right-hand side in (1.32) yields that for every $z \in \overline{D}_\nu$

$$|\widehat{\psi}_1(z)| \leq \sum_{j=2}^{\infty} |b_j z^j| \leq (\pi c_0)^{-1} c_2 \delta_0 u^{-(\gamma+1)}.$$

Taking into account that $u = -\ln(1-\frac{\eta_0}{4})$ is no less than η_0 times a universal constant, we get that the latter right-hand side is less than $\delta_0\eta_0^{-c}$ with some universal constant $c > 0$, which we can and will take no less than 3. This implies convergence of the series $\widehat{\psi}_1(z)$ and the bound (1.23) for the function $\widehat{\psi}_1(z)$ on the disk $\overline{D}_\nu$. □

Proposition 1.14 *The function $f_1(z) = \psi_1^{-1} \circ f_0 \circ \psi_1(z)$ is holomorphic on $\overline{D}_{r_1}$ and satisfies bound (1.22).*

Proof Let us introduce the following four radii:

$$r_1 < \nu_1 < \nu_2 < \nu_3 < \nu_4, \qquad r_1 = 1 - \eta_0,$$

$$\nu_1 = 1 - \frac{3\eta_0}{4}, \quad \nu_2 = 1 - \frac{\eta_0}{2}, \quad \nu_3 = 1 - \frac{3\eta_0}{8}, \quad \nu_4 = \nu = 1 - \frac{\eta_0}{4}.$$

Claim 2. *The inverse ψ_1^{-1} is holomorphic on $\overline{D}_{\nu_2}$ and sends it to D_{ν_3}.*
Proof Let us prove injectivity of the map ψ_1 on $\overline{D}_{\nu_3}$. Indeed, ψ_1 is holomorphic on $\overline{D}_{\nu_4}$ and satisfies upper bound (1.23) there, as was proved above. On the disk $\overline{D}_{\nu_3}$ the derivative $\widehat{\psi}_1' = \psi_1' - 1$ has modulus no greater than the maximum modulus of a value $|\widehat{\psi}_1(z)|$, $z \in \overline{D}_{\nu_4}$, divided by $\nu_4 - \nu_3 = \frac{\eta_0}{8}$, by Cauchy bound (1.15). That is, for every $z \in \overline{D}_{\nu_3}$, one has

$$|\psi_1'(z) - 1| \leq \delta_0\eta_0^{-c} 8\eta_0^{-1} < 8\eta_0^{3c-1} < \frac{1}{4}. \tag{1.34}$$

For every $z, w \in \overline{D}_{\nu_3}$ one has

$$\psi_1(z) - \psi_1(w) = \int_{[z,w]} \psi_1'(\zeta) d\zeta = (z - w) + \int_{[z,w]} (\psi'(\zeta) - 1) d\zeta,$$

The right-hand side is non-zero, since the latter integral has module less than $\frac{1}{4}|z - w|$, by (1.34). Injectivity on $\overline{D}_{\nu_3}$ is proved.

The minimal distance between the ψ_1-image of a point of the boundary ∂D_{ν_3} and the same boundary ∂D_{ν_3} is no greater than $\max_{\overline{D}_{\nu_3}} |\widehat{\psi}_1| < \delta_0\eta_0^{-c} < \eta_0^{3c} < \frac{\eta_0}{8} = \nu_3 - \nu_2$. This implies that $\psi_1(\partial D_{\nu_3})$ lies outside the disk $\overline{D}_{\nu_2}$, and hence, $\overline{D}_{\nu_2} \Subset \psi_1(D_{\nu_3})$. Therefore, the image of the disk D_{ν_3} under injective holomorphic map ψ_1 contains $\overline{D}_{\nu_2}$. Hence, the inverse ψ_1^{-1} is holomorphic on $\overline{D}_{\nu_2}$ and sends it to D_{ν_3}. Claim 2 is proved. □

The function ψ_1 is holomorphic on $\overline{D}_{r_1}$ and sends it to D_{ν_1}. Indeed, $|\widehat{\psi}_1(z)| = |\psi_1(z) - z| < \delta_0\eta_0^{-c}$ on $\overline{D}_{\nu_4} \Supset D_{\nu_1}$, and $\delta_0\eta_0^{-c} < \eta_0^{3c} < \nu_1 - r_1$.

Thus, $|\psi_1(z)| < |z| + (\nu_1 - r_1) \leq \nu_1$ on $\overline{D}_{r_1}$. The function f_0 sends $\overline{D}_{\nu_1}$ to D_{ν_2}, since

$$||f_0(z)| - |z|| \leq |f_0(z) - \lambda z| \leq \delta_0 < \eta_0^{4c} < \frac{\eta_0}{4} = \nu_2 - \nu_1.$$

The function ψ_1^{-1} sends $\overline{D}_{\nu_2}$ to D_{ν_3}, by Claim 2. This implies that the composition $f_1 = \psi_1^{-1} \circ f_0 \circ \psi_1$ is holomorphic on $\overline{D}_{r_1}$ and sends it to $D_{\nu_3} \subset D_1$.

Let us now prove (1.22). One has

$$f_0 \circ \psi_1(z) = \psi_1 \circ f_1(z) = \psi_1(\lambda z) + (\psi_1(f_1(z)) - \psi_1(\lambda z)),$$

$$f_0 \circ \psi_1(z) - \psi_1(\lambda z) = (\widehat{f}_0(z) - \Lambda\widehat{\psi}_1(z)) + R(z) = R(z) = \widehat{f}_0(\psi_1(z)) - \widehat{f}_0(z),$$

by (1.28) and since $\widehat{\psi}_1(z)$ is a solution of homological equation (1.29). Thus,

$$\psi_1(f_1(z)) - \psi_1(\lambda z) = R(z).$$

On the disk $\overline{D}_{r_1}$ the latter left-hand side has module no less than $\left(\min_{f_1(\overline{D}_{r_1})} |\psi_1'|\right) |\widehat{f}_1(z)|$, $\widehat{f}_1(z) = f_1(z) - \lambda z$. The latter minimum of module of derivative is no less than $\frac{3}{4}$, since $f_1(\overline{D}_{r_1}) \subset D_{\nu_3}$ and $|\psi_1' - 1| = |\widehat{\psi}'| < \frac{1}{4}$ on $\overline{D}_{\nu_3}$, by (1.34). Recall that $\psi_1(\overline{D}_{r_1}) \subset \overline{D}_{\nu_1}$. Therefore,

$$|\widehat{f}_1(z)| < \frac{4}{3}|R(z)| = \frac{4}{3}|\widehat{f}_0(\psi_1(z)) - \widehat{f}_0(z)| \leq \frac{4}{3}\max_{\overline{D}_{\nu_1}} |\widehat{f}_0'||\widehat{\psi}_1(z)|,$$

$$\max_{\overline{D}_{\nu_1}} |\widehat{f}_0'| \leq \frac{\sup_{D_1} |\widehat{f}_0|}{1 - \nu_1} < \frac{\delta_0}{1 - \nu_1} = \frac{4\delta_0}{3\eta_0},$$

by Cauchy bound (1.15) and (1.19). This together with (1.23) implies that

$$|\widehat{f}_1(z)| < \frac{4}{3}\frac{4\delta_0}{3\eta_0}\delta_0\eta_0^{-c} = \frac{16}{9}\delta_0^2\eta_0^{-c-1} < \delta_0^2\eta_0^{-(c+2)}r_1.$$

After denoting $c + 2$ by c this yields (1.22). Proposition 1.14 is proved. □

Lemma 1.10 follows from Proposition 1.14 and the above arguments. The proof of Siegel Theorem is complete. □

1.6 General results: Bruno's and Yoccoz's theorem

Consider presentation of a real number θ as a continued fraction:

$$\theta = a_1 + \cfrac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\dots}}} = \lim_{n\to+\infty} \frac{p_n}{q_n}, \quad \frac{p_n}{q_n} = a_1 + \cfrac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}.$$

Definition 1.15 An irrational number θ is a *Bruno number,* if

$$\text{the series} \quad \sum_{n=1}^{+\infty} q_n^{-1} \ln q_{n+1} \quad \text{converges.}$$

Theorem 1.16 (A.D.Bruno, 1971). *For every Bruno number θ every conformal germ $f(z) = \lambda z + O(z^2)$, $\lambda = e^{2\pi i\theta}$, is holomorphically linearizable.*

In 1988 J.-C.Yoccos proved his famous converse theorem.

Theorem 1.17 (J.-C.Yoccoz, 1988 [5, 6], Fields Medal 1994). *Let θ be a fixed irrational number, $\lambda = e^{2\pi i\theta}$. Then*

1) every conformal germ $f(z) = \lambda z + O(z^2)$ is linearizable, if and only if θ is a Bruno number;

2) the germ of quadratic polynomial $f(z) = \lambda z + z^2$ is linearizable, if and only if θ is a Bruno number.

Remark 1.18 In his papers Yoccoz proves not only his converse Theorem 1.17, but also Bruno's Linearization Theorem 1.16. His proofs are beautiful mixture of geometric and analytic arguments, based on a priori bounds for univalent functions (injective holomorphic functions).

Exercise 1.19 Prove that θ is Diophantine, if and only if it satisfies the Siegel condition

$$\sup_{n\in\mathbb{N}} \frac{\ln q_{n+1}}{\ln q_n} < +\infty.$$

Exercise 1.20 Prove that if θ is Diophantine, then it is a Bruno number.

Hint to Exercise 2.5. Use the following classical properties of the continued fractions for every $n \in \mathbb{N}$:

- θ lies between successive approximating fractions $\frac{p_n}{q_n}$ and $\frac{p_{n+1}}{q_{n+1}}$,
- $\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n q_{n+1}}$;
- one has $q_n = a_n q_{n-1} + q_{n-2}$.

2 Circle diffeomorphisms, rotation number and analytic conjugacy to rotation

2.1 Rotation number: definition, continuity, monotonicity

Here we briefly recall the material given in [1].

We will be dealing with orientation-preserving homeomorphisms of circle $\mathbb{R}_\phi/2\pi\mathbb{Z}$. Each of them takes the form

$$f(\phi) = \phi + g(\phi), \quad g(\phi + 2\pi) = g(\phi).$$

We can lift f to a homeomorphism $F : \mathbb{R}_\phi \to \mathbb{R}_\phi$ of the universal covering line, which is given by the above formula. The lifting F is uniquely determined up to translation $F \mapsto F + 2\pi m$, $m \in \mathbb{Z}$.

Definition 2.1 The *rotation number* of F is

$$\rho := \lim_{n\to+\infty} \frac{f^n(\phi) - \phi}{2\pi n}.$$

Theorem 2.2 *1) The rotation number ρ exists and is independent on ϕ.*

2) Replacing F by $F + 2\pi m$ adds m to ρ.

3) The rotation number taken modulo $\mathbb{Z}$ is a well-defined invariant of circle homeomorphism, independent of choice of lifting.

Example 2.3 1) For a rotation $f(\phi) = R_\alpha(\phi) = \phi + \alpha$ one has $\rho = \frac{\alpha}{2\pi}$.

2) One has $\rho = 0$, if and only if f has a fixed point.

3) One has $\rho = \frac{p}{q}$, if and only if f has a q-periodic point.

4) If f, w are orientation preserving circle homeomorphisms conjugated by an orientation preserving homeomorphism h, i.e., $f = h^{-1} \circ w \circ h$, then $\rho(f) = \rho(w)$.

Proposition 2.4 *1) The rotation number depends continuously on the homeomorphism in the C^0-topology.*

2) It is monotonous. Namely, let homeomorphisms f_1, f_2 have liftings $F_1 \leq F_2$. Then $\rho(F_1) \leq \rho(F_2)$.

For the proofs of statements of the above theorem and example see [1].

Exercise 2.5 Prove the proposition.

Exercise 2.6 Prove that if $F_1 < F_2$ and F_1 is a translation (i.e., f_1 is a rotation, then $\rho(F_2) > \rho(F_1)$.

2.2 Denjoy Conjugacy Theorem and example

Theorem 2.7 *(Denjoy, [1]). Every circle C^2-diffeomorphism with irrational rotation number is conjugated to rotation. Thus, its orbits are dense.*

Theorem 2.8 *(Denjoy, [1]). For every irrational number θ there exists a C^1-smooth circle diffeomorphism with $\rho = \theta$ that has an invariant Cantor set. In particular, it is not conjugated to a rotation.*

2.3 Herman Conjugacy Regularity Theorem

Theorem 2.9 *(M.Herman) Let $\mathcal{D} \subset \mathbb{R}$ denote the union of those irrational numbers θ such that for every $\gamma > 2$ there exists a $C = C(\gamma) > 0$ such that θ is (C, γ)-Diophantine. Then every C^k-smooth (analytic) circle diffeomorphism with rotation number lying in $\mathcal{D}$ is C^{k-2}-smoothly (analytically) conjugated to rotation.*

2.4 Analytic circle diffeomorphisms close to rotations. Kolmogorov–Arnold analytic conjugacy theorem

Consider a family of analytic circle diffeomorphisms

$$f(\phi) = \phi + \beta + g(\phi), \quad \beta \in \mathbb{R}, \quad g(\phi + 2\pi) = g(\phi). \tag{2.1}$$

If $g \equiv 0$, this is a family of rotations. We prove the next analytic conjugacy theorem, not covered by Herman's Theorem, stating that if $\rho(f)$ is (C, γ)-Diophantine and g is "small enough" depending on C, γ and the complex definition domain of f, then f is analytically conjugate to rotation.

Theorem 2.10 *(Kolmogorov, Arnold). For every $\sigma, C, \gamma > 0$ there exists a $\delta = \delta(C, \gamma, \sigma) > 0$ satisfying the following statements. Let f be as in (2.1), and let $\rho = \rho(f)$ be (C, γ)-Diophantine, set*

$$\alpha := 2\pi\rho, \quad U_\sigma := \{|\operatorname{Im}\phi| < \sigma\} \subset \mathbb{C}/2\pi\mathbb{Z}.$$

Let g in (2.1) be holomorphic on the cylinder U_σ and continuous on $\overline{U}_\sigma$. Let

$$\max_{\overline{U}_\sigma} |g| \leq \delta. \tag{2.2}$$

Then there exists an analytic diffeomorphism h conjugating f to rotation:

$$f \circ h(\phi) = h(\phi + \alpha). \tag{2.3}$$

The proof of Theorem 2.10 given below is similar to the above proof of Siegel Theorem. We construct successive conjugacies h_{n-1} to new functions

$$f_n(\phi) = \phi + \beta_n + g_n(\phi) = h_{n-1}^{-1} \circ f_{n-1} \circ h_{n-1},\ f_0 = f,$$

with g_n holomorphic on $\overline{U}_{\sigma_n}$, where σ_n decreases to a positive number $2\sigma_*$, so that $\max|g_n| \to 0$ and $H_n = h_0 \circ h_1 \circ \cdots \circ h_n$, converge uniformly to a map H on $\overline{U}_{\sigma_*}$. This will imply H conjugates f with the rotation R_α.

Here we deal with functions holomorphic on an annulus in $\mathbb{C}_z$, $z = e^{i\phi}$. We use the following Cauchy bound for their Laurent series coefficients.

Proposition 2.11 *Let $f(z)$ be a function holomorphic on an annulus $\mathcal{A}_{\nu,r} := \{\nu < |z| < r\}$, $0 < \nu < r$, and continuous on its closure. Let $f(z) = \sum_{k\in\mathbb{Z}} a_k z^k$ be its Laurent series. Then*

$$|a_k| \leq \frac{\max_{\mathcal{A}_r}|f|}{r^k} \tag{2.4}$$

Proof One has

$$a_k = \frac{1}{2\pi i}\oint_{|\zeta|=r}\frac{f(\zeta)}{\zeta^{k+1}},$$

by the Residue Formula applied to $f(z)z^{-(k+1)}$. This implies (2.4). □

The inductive construction of h_n and f_n with bounds implying convergence is based on the following lemma.

Lemma 2.12 (Main Lemma). *There exists a universal constant $c_* = c_*(C,\gamma) > 0$ satisfying the following statements. Let $f_0 : S^1 \to S^1$ be a circle diffeomorphism*

$$f_0(\phi) = \phi + \beta_0 + g_0(\phi) \tag{2.5}$$

with a (C,γ)-Diophantine rotation number ρ, set $\alpha = 2\pi\rho$. Let $\sigma_0 > 0$, and let g_0 be holomorphic on U_{σ_0} and continuous on its closure. Let

$$\eta_0 > 0, \quad \eta_0 < \min\left\{\frac{1}{4}, \frac{\sigma_0}{4}\right\},\ \delta_0 > 0,\ \delta_0 < \eta_0^{4c_*},$$

$$|g_0| \leq \delta_0 \text{ on } \overline{U}_{\sigma_0}. \tag{2.6}$$

Set

$$\sigma_1 = \sigma_0 - \eta_0,\ \delta_1 = \delta_0^2\eta_0^{-c_*},\ \eta_1 = \frac{\eta_0}{2}. \tag{2.7}$$

Then there exists a holomorphic map

$$h_0 : \overline{U}_{\sigma_0-\frac{\eta_0}{4}} \to \mathbb{C}/2\pi\mathbb{Z},\ h_0(\phi) = \phi + q_0(\phi),\quad |q_0| \leq \delta_0\eta_0^{-c^*} \text{ on } \overline{U}_{\sigma_0-\frac{\eta_0}{4}}, \tag{2.8}$$

whose restriction to $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ is a diffeomorphism $S^1 \to S^1$, such that

$$f_1(\phi) := h_0^{-1} \circ f_0 \circ h_0(\phi)$$

is holomorphic on $\overline{U}_{\sigma_1}$ and can be written as

$$f_1(\phi) = \phi + \beta_1 + g_1(\phi), \quad \max_{\overline{U}_{\sigma_1}} |g_1| \le \delta_1. \tag{2.9}$$

One has $\delta_1 < \eta_1^{4c^*}$, $\delta_1 < \delta_0^{\frac{7}{4}}$.

Proof of Theorem 2.10 modulo Lemma 2.12. For every $n \ge 0$ we construct inductively h_n and f_n,

$$h_n(\phi) = \phi + q_n(\phi), \quad f_{n+1} = h_n^{-1} \circ f_n \circ h_n,$$

by successively applying the lemma to f_0 replaced by f_n. Set

$$H_n := h_0 \circ h_1 \circ \cdots \circ h_n : \quad f_n = H_{n-1}^{-1} \circ f_0 \circ H_{n-1},$$

$$\sigma_* := \frac{1}{2}(\sigma_0 - 2\eta_0) = \frac{1}{2}(\sigma_0 - \sum_j \eta_j) > 0.$$

Claim 3. *The maps H_n are holomorphic and they converge uniformly to some H on $\overline{U}_{\sigma_*}$.*

Proof Each map h_{n-1} is holomorphic on $\overline{U}_{\sigma_n}$ and sends it to $U_{\sigma_{n-1}}$, since

$$|q_{n-1}| \le \delta_{n-1}\eta_{n-1}^{-c_*} < \eta_{n-1}^{3c_*} < \eta_{n-1} = \sigma_{n-1} - \sigma_n \text{ on } \overline{U}_{\sigma_n},$$

by (2.8). Therefore, H_{n-1} sends $\overline{U}_{\sigma_n}$, and hence, $\overline{U}_{2\sigma_*} \subset U_{\sigma_n}$, holomorphically to U_{σ_0}. On the set $\overline{U}_{\sigma_*}$ one has $h_n(\phi) \in U_{\sigma_n}$, and hence,

$$|H_n(\phi) - H_{n-1}(\phi)| = |H_{n-1} \circ h_n(\phi) - H_{n-1}(\phi)| \le \max_{\overline{U}_{\sigma_n}} |H_{n-1}'| \max_{\overline{U}_{\sigma_{n+1}}} |q_n|. \tag{2.10}$$

To estimate the latter derivative, let us deal with H_{n-1} as a mapping written in the coordinate

$$z = e^{i\phi}, \quad \ln |z| \le \sigma_n$$

Set

$$r_n = e^{\sigma_n}, \; r_* = e^{\sigma_*}, \; \mathcal{A}_n = \mathcal{A}(r_n^{-1}, r_n), \; \mathcal{A}_* = \mathcal{A}(r_*^{-1}, r_*).$$

Then $H_{n-1}(\overline{\mathcal{A}}_n) \subset \mathcal{A}_0$. The derivative module $|H_{n-1}'|$ on $\overline{\mathcal{A}}_*$ is no greater than $\frac{r_0}{r_0 - r_*}$, by Cauchy bound. Therefore, on the same set in the coordinate

ϕ one has $|H'_{n-1}| \leq K := \frac{r_0 r_*^2}{r_0 - r_*}$, since r_* is the maximum of derivative of the exponential map $\overline{U}_{\sigma_*} \to \overline{\mathcal{A}}_*$, $\phi \mapsto e^{i\phi}$, and of holomorphic branches of its inverse. Thus, the left-hand side in (2.10) is bounded by $K|\max_{\overline{U}_{\sigma_{n+1}}} |q_n| < K\delta_n \eta_n^{-c_*} < K\eta_n^{3c_*}$, where η_n is a decreasing geometric progression. This implies uniform convergence of H_n. □

Finally, the maps H_n are holomorphic, they fix the circle S^1 and their restrictions there are analytic diffeomorphisms. They converge to a map H uniformly on $\overline{U}_{\sigma_*}$. They satisfy the relation $f_0 \circ H_n = H_n \circ f_n$. The map f_n converge to the rotation $R_\alpha : \phi \mapsto \phi + \alpha$. Indeed, the difference $g_n = f_n - \beta_n$ have module less than $\delta_n \to 0$ on $\overline{U}_{\sigma_*}$, and hence, converge uniformly to 0. This implies that each subsequence of f_n contains a uniformly converging subsequence, and its limit is a rotation. The latter rotation coincides with R_α since the rotation numbers of all of f_n, and hence, of the limit, are equal to $\rho = \frac{\alpha}{2\pi}$. Thus, $f_n \to R_\alpha$, and passing to limit, we get that

$$f_0 \circ H(\phi) = H(\phi + \alpha). \tag{2.11}$$

The map H is analytic on U_{σ_*}, and hence, on S^1, by Weierstrass Theorem, as a limit of uniformly converging sequence of holomorphic functions.

It remains to show that $H : S^1 \to S^1$ is an analytic diffeomorphism, i.e., its inverse is analytic. One can prove this by showing that H_n are injective on $U_{\sigma_{n+1}}$ and their derivatives are bounded from below there by using Cauchy bound for the derivative of the difference $H_n(\phi) - \phi$. But we present a different proof. Suppose the contrary: $H : S^1 \to S^1$ is not a diffeomorphism. Then either H is a non-constant non-injective analytic map, or H is constant, or H is injective and there is a $\phi_0 \in \mathbb{R}$ such that $H'(\phi_0) = 0$. The first case is impossible, since then the restrictions $H_n : S^1 \to S^1$ are non-injective for n large enough, which contradicts to the statement that h_n, and hence, H_n, being restricted to S^1, are diffeomorphisms. If $H \equiv const$, then H sends S^1 to a point, hence H_n with big n send S^1 to a small circle arc, which is impossible for a circle diffeomorphism. If H is injective and $H'(\phi_0) = 0$ for some ϕ_0, then $H'(\phi_0 + n\alpha) = 0$ for all n, by (2.11). Hence, $H' \equiv 0$ on S^1, by density of the sequence $n\alpha (\bmod 2\pi\mathbb{Z})$ (density of orbits of the rotation R_α with irrational rotation number). Therefore, $H \equiv const$ on S^1, – a contradiction. Thus, H is an analytic diffeomorphism $S^1 \to S^1$ conjugating f_0 to R_α. Theorem 2.10 is proved modulo Main Lemma 2.12. □

2.5 Proof of Main Lemma 2.12

Note that the last inequalities on δ_1, η_1 of the lemma were already proved in the proof of the Main Lemma for Siegel Theorem. Let us prove the main part of Lemma 2.12.

Step 1. A priori bounds on Fourier coefficients of g_0. Let us write $f_0(\phi) = \phi + \beta_0 + g_0(\phi)$ as

$$f_0(\phi) = \phi + \alpha + \zeta_0 + \widetilde{g}_0(\phi), \quad \widetilde{g}_0(\phi) = \sum_{k \in \mathbb{Z}_{\neq 0}} a_k e^{ik\phi}, \quad a_{-k} = \overline{a}_k, \tag{2.12}$$

$$\widetilde{g}_0(\phi) = g_0(\phi) - a_0, \quad a_0 = \frac{1}{2\pi} \int_0^{2\pi} g_0(\phi) d\phi, \quad \zeta_0 = a_0 + \beta_0 - \alpha.$$

Claim 4. *One has*

$$|\zeta_0|, \ \max_{\overline{U}_{\sigma_0}} |\widetilde{g}_0| \leq 2\delta_0. \tag{2.13}$$

Proof The number a_0 has module no greater than $\max_{\overline{U}_{\sigma_0}} |g_0| \leq \delta_0$. This implies the bound on $\widetilde{g}_0$ in (2.13). Suppose the contrary to the first bound in (2.13): say, $\zeta_0 > 2\delta_0$; the case $\zeta < -2\delta_0$ is treated analogously. Then

$$f_0(\phi) = \phi + \alpha + (\zeta_0 + \widetilde{g}_0(\phi)) > R_\alpha(\phi) = \phi + \alpha.$$

Therefore, $\rho(f_0) > \rho(R_\alpha) = \frac{\alpha}{2\pi}$, by Exercise 2.6, – a contradiction. □

Step 2. Homological equation and bound on its solution. We are looking for a conjugating diffeomorphism h_0 killing a big part of $\widetilde{g}_0$:

$$h_0 : S^1 \to S^1, \quad h_0(\phi) = \phi + q_0(\phi), \quad q_0(\phi) = \sum_{k \in \mathbb{Z}_{\neq 0}} b_k e^{ik\phi}, \quad b_{-k} = \overline{b}_k.$$

If h_0 conjugates f_0 to the rotation R_α, then $f_0 \circ h_0(\phi) - h_0(\phi + \alpha) = 0$. Below we construct an h_0 for which the latter difference is very small. To do this, we use the following formula for the difference:

$$f_0 \circ h_0(\phi) - h_0(\phi + \alpha) = \phi + q_0(\phi) + \alpha + \zeta_0 + \widetilde{g}_0(\phi + q_0(\phi)) - (\phi + \alpha + q_0(\phi + \alpha))$$

$$= (\widetilde{g}_0(\phi) + q_0(\phi) - q_0(\phi + \alpha)) + \zeta_0 + R(\phi), \quad R(\phi) = \widetilde{g}_0(\phi + q_0(\phi)) - \widetilde{g}_0(\phi). \tag{2.14}$$

The *homological equation* is vanishing of the expression in the first brackets:

$$q_0(\phi + \alpha) - q_0(\phi) = \widetilde{g}_0(\phi). \tag{2.15}$$

Proposition 2.13 *The homological equation (2.15) has a unique holomorphic solution on a neighborhood of the circle S^1 with zero average along S^1:*

$$q_0(\phi) = \sum_{k\in\mathbb{Z}_{\neq 0}} b_k e^{ik\phi}, \quad b_{-k} = \overline{b}_k, \quad b_k = \frac{a_k}{e^{ik\alpha}-1}. \tag{2.16}$$

The function $q_0(\phi)$ is holomorphic on U_{σ_0}. There exists a universal constant $c_1 = c_1(C,\gamma)$ such that for every $\eta_0 > 0$, $\eta_0 < \min\{\frac{1}{4}, \frac{\sigma_0}{4}\}$,

$$\max_{\overline{U}_{\sigma_0-\frac{\eta_0}{4}}} |q_0(\phi)| \leq \delta_0 \eta_0^{-c_1}. \tag{2.17}$$

Proof The Fourier series (2.16), if converges, obvious solves (2.15), and is obviously unique in the above sense. Let us prove its holomorphicity and bound (2.17). The sum of series (2.16) admits the bound on $\overline{U}_{\sigma_0-\frac{\eta_0}{4}}$:

$$|\sum_{k\in\mathbb{Z}_{\neq 0}} b_k e^{ik\phi}| \leq 2\sum_{k\in\mathbb{N}} |b_k| e^{k(\sigma_0-\frac{\eta_0}{4})} = 2\sum_{k\in\mathbb{N}} \frac{|a_k|}{|e^{ik\alpha}-1|} e^{k(\sigma_0-\frac{\eta_0}{4})}, \tag{2.18}$$

since $b_{-k} = \overline{b}_k$ and by (2.16). For every $k\in\mathbb{N}$ one has

$$|a_k| e^{k\sigma_0} \leq \max_{\overline{U}_{\sigma_0}} |\widetilde{g}_0| \leq 2\delta_0, \tag{2.19}$$

by Cauchy bound (Proposition 2.11), writing $\widetilde{g}_0$ as a Laurent series in the variable $z = e^{i\phi}$,

$$|e^{ik\alpha}-1| > \frac{C'}{|k|^\gamma}, \tag{2.20}$$

by Diophantine inequality on α, as in the proof of Siegel's Theorem. Here $C' = C'(C,\gamma) > 0$. Substituting (2.19) and (2.20) to (2.18) yields

$$|\sum_{k\in\mathbb{Z}_{\neq 0}} b_k e^{ik\phi}| \leq 2(C')^{-1}\delta_0 \sum_{k\in\mathbb{N}} |k|^\gamma e^{-\frac{k\eta_0}{4}}. \tag{2.21}$$

The series in (2.21) is no greater than some universal constant times $\eta_0^{-(\gamma+1)}$, see Proposition 1.13. Therefore, the right-hand side in (2.21) is no greater than $\delta_0\eta_0^{-c_1}$ with some universal constant c_1, since $\eta_0 < \frac{1}{4}$. This implies that the series for the function q_0 converges uniformly on every annulus strictly smaller than U_{σ_0} and satisfies (2.17). Proposition 2.13 is proved. □

Thus, the bound (2.8) is proved with a universal constant c_1, and hence, with every constant $c_* \geq c_1$. We will choose $c_* \geq 4$. Recall that we choose δ_0 so that $\delta_0 < \eta_0^{4c_*}$.

Step 3. Holomorphicity of the composition $f_1 = h_0^{-1} \circ f_0 \circ h_0$ **and its bound (2.9).** To prove its holomorphicity, we have to prove injectivity of the map h_0 on appropriate annulus. To this end, let us introduce the numbers $\nu_0 = \sigma_1 < \nu_1 < \cdots < \nu_5 = \sigma_0$:

$$\nu_1 = \sigma_0 - \frac{3\eta_0}{4}, \ \nu_2 = \sigma_0 - \frac{\eta_0}{2}, \ \nu_3 = \sigma_0 - \frac{3\eta_0}{8}, \ \nu_4 = \sigma_0 - \frac{\eta_0}{4}.$$

Proposition 2.14 *The map* $h_0(\phi) = \phi + q_0(\phi)$ *sends each* $\overline{U}_{\nu_j}$*,* $j \leq 4$*, to* $U_{\nu_{j+1}}$*. Its derivative satisfies the upper bound*

$$|q_0'| < \delta_0 \eta_0^{-c_1'} < \frac{1}{2} \quad \text{on } \overline{U}_{\nu_3}; \quad c_1' = c_* + 3 \text{ is a universal constant.} \tag{2.22}$$

Proof The first statement follows from (2.8) and the elementary inequality $\delta_0 \eta_0^{-c_*} < \eta^{3c_*} < \nu_{j+1} - \nu_j$. To prove (2.22), let us consider the function $q_0(\phi)$ restricted to the strips $\widetilde{U}_{\nu_j} = \{|\operatorname{Im}\phi| < \nu_j\} \subset \mathbb{C}$. On $\widetilde{U}_{\nu_3}$ one has

$$|\frac{dq_0}{dz}| \leq \frac{\sup_{\widetilde{U}_{\nu_4}} |q_0|}{\nu_4 - \nu_3} \leq 8\delta_0 \eta_0^{-(c_*+1)} < \delta_0 \eta_0^{-(c_*+3)} < \eta_0^{3(c_*-3)} < \frac{1}{2},$$

by Cauchy bound and elementary inequalities. □

Proposition 2.15 *The map* $h_0 : \overline{U}_{\nu_3} \to U_{\nu_4}$ *injective.*

Proof The lifting of the map h_0 to the closure of the strip $\widetilde{U}_{\nu_3}$ is injective. Indeed, for every two distinct ϕ_1, ϕ_2 lying there one has

$$h_0(\phi_1) - h_0(\phi_2) = (\phi_1 - \phi_2) + (q_0(\phi_1) - q_0(\phi_2)), \tag{2.23}$$

$|q_0(\phi_1) - q_0(\phi_2)| \leq \frac{1}{2}|\phi_1 - \phi_2|$, by (2.22). Hence, the right-hand side in (2.23) is non-zero. Let us prove injectivity of the map h_0 as an annulus map. Suppose the contrary: h_0 is not injective. Then there exist ϕ_1, ϕ_2 in the closure of the strip $\widetilde{U}_{\nu_3}$ such that $\phi_1 - \phi_2 \notin 2\pi\mathbb{Z}$ while the difference (2.23) is equal to $2\pi n$ for some $n \in \mathbb{Z}$. Shifting ϕ_2 by $2\pi n$ one can achieve that the difference (2.23) is equal to zero. But this is impossible, since $\phi_1 \neq \phi_2$ and hence, the first term in the right-hand side in (2.23) dominates the second term, as above. The proposition is proved. □

Proposition 2.16 *The map* h_0 *sends* $\partial\overline{U}_{\nu_3}$ *outside* U_{ν_2}*. The map* h_0^{-1} *is holomorphic on* $\overline{U}_{\nu_2}$*, and it sends is to* U_{ν_3}*.*

Proof The first statement follows from (2.8), as in the proof of Proposition 2.14. The second statement follows from the first one and injectivity of the map h_0 on $\overline{U}_{\nu_3}$, see Proposition 2.15. $\square$

Proposition 2.17 *The map* $f_1 = h_0^{-1} \circ f_0 \circ h_0$ *is holomorphic on* $\overline{U}_{\sigma_1}$ *and*

$$f_1(\phi) = \phi + \alpha + \zeta_0 + g_1(\phi), \quad |g_1| \leq \delta_0^2 \eta_0^{-c_3}, \quad c_3 \text{ is a universal constant.} \tag{2.24}$$

Proof The map h_0 is holomorphic on $\overline{U}_{\sigma_1}$ and sends it to U_{ν_1}, by Proposition 2.14. The map f_0 sends $\overline{U}_{\nu_1}$ to U_{ν_2}, which is proved similarly to the analogous statement for the map h_0 in Proposition 2.14. The map h_0^{-1} is holomorphic on $\overline{U}_{\nu_2}$. This implies holomorphicity of the composition f_1 on $\overline{U}_{\sigma_1}$. Let us write f_1 as in (2.24) and prove the upper bound (2.24) on g_1. First, let us estimate the difference between the images of $f_1(\phi)$ and $\phi + \alpha + \zeta_0$ under the map h_0. One has

$$h_0 \circ f_1(\phi) - h_0(\phi + \alpha + \zeta_0) = f_0 \circ h_0(\phi) - h_0(\phi + \alpha + \zeta_0) \tag{2.25}$$
$$= (\widetilde{g}_0(\phi) + q_0(\phi) - q_0(\phi + \alpha)) + B(\phi) + R(\phi) = B(\phi) + R(\phi),$$
$$B(\phi) = q_0(\phi + \alpha) - q_0(\phi + \alpha + \zeta_0), \; R(\phi) = \widetilde{g}_0(\phi + q_0(\phi)) - \widetilde{g}_0(\phi),$$

by (2.14) and since q_0 is the solution of the homological equation (2.15).

Proposition 2.18 *The remainders* B, R *are holomorphic on* $\overline{U}_{\nu_3}$ *and satisfy the inequalities*

$$|B(\phi)|, |R(\phi)| \leq \delta_0^2 \eta_0^{-c_2} \quad \text{on } \overline{U}_{\nu_3}, \quad c_2 \text{ is an universal constant.} \tag{2.26}$$

Proof Holomorphicity follows from holomorphicity of the function $\widetilde{g}_0$ on U_{σ_0} and the inclusion $h_0(\overline{U}_{\nu_4}) \subset U_{\nu_5} = U_{\sigma_0}$, see Proposition 2.14. One has

$$\max_{\overline{U}_{\nu_2}} |R| \leq \max_{\overline{U}_{\nu_3}} |\widetilde{g}_0'| \max_{\overline{U}_{\nu_2}} |q_0| \leq \frac{2\delta_0^2 \eta_0^{-c_*}}{\nu_4 - \nu_3} < \delta_0^2 \eta_0^{-c_2},$$

where $c_2 > 0$ is a universal constant, by the inclusion $h_0(\overline{U}_{\nu_2}) \subset U_{\nu_3}$, Cauchy bound, (2.13), (2.8) and elementary inequalities,

$$\max_{\overline{U}_{\nu_2}} |B| \leq \max_{\overline{U}_{\nu_2}} |q_0'||\zeta_0| \leq 2\delta_0^2 \eta_0^{-c_1'},$$

by (2.22) and (2.13). The latter right-hand side is no greater than $\delta_0^2 \eta_0^{-c_2}$ with some universal constant c_2. The proposition is proved. $\square$

Let us now prove inequality (2.9). On $\overline{U}_{\sigma_1}$ one has

$$f_1(\phi) - (\phi + \alpha + \zeta_0) = h_0^{-1}(f_0 \circ h_0(\phi)) - h_0^{-1}(h_0(\phi + \alpha + \zeta_0))$$
$$= h_0^{-1}(h_0(\phi + \alpha + \zeta_0) + B(\phi) + R(\phi)) - h_0^{-1}(h_0(\phi + \alpha + \zeta_0)). \qquad (2.27)$$

Both maps $h_0(\phi + \alpha + \zeta_0)$ and $h_0(\phi + \alpha + \zeta_0) + B(\phi) + R(\phi)$ send $\overline{U}_{\sigma_1}$ to U_{ν_2}, since the former map sends it to U_{ν_1} and the difference $B(\phi) + R(\phi)$ between the maps is much less than the gap $\nu_2 - \nu_1$, see (2.26). One has

$$\max_{\overline{U}_{\nu_2}} |h_0^{-1})'(\phi)| \leq \max_{\overline{U}_{\nu_3}} |(h_0'(\phi))^{-1}| < 2, \qquad (2.28)$$

by (2.22), elementary inequalities and since $h_0^{-1}(\overline{U}_{\nu_2}) \subset U_{\nu_3}$. Therefore, on $\overline{U}_{\nu_2}$ the module of the right-hand side in (2.27) is less than $2(|B(\phi)| + |R(\phi)|) \leq 4\delta_0^2\eta_0^{-c_2} < \delta_0^2\eta_0^{-c_3}$, by (2.28) and (2.26), with some universal constant $c_3 \geq 4$. Proposition 2.17 is proved. □

Now it remains to correct c_* to be greater than all the above-mentioned universal constants $c_* + 3$, $c_1, \dots, c_3$ and $\delta_0 < \eta_0^{4c_*}$ and set $\delta_1 = \delta_0^2\eta_0^{-c_*}$. Then (2.9) follows from Proposition 2.17. The Main Lemma 2.12 is proved.

3 Introduction to symplectic geometry and dynamics

3.1 Symplectic manifolds. Basic examples

Definition 3.1 Consider an antisymmetric bilinear form ω on vector space $\mathbb{R}^m$. For every vector $u \in \mathbb{R}^m$ let us introduce the linear functional

$$i_u\omega : \mathbb{R}^m \to \mathbb{R} : \quad (i_u\omega)(v) := \omega(u, v).$$

The *kernel* of the form ω is kernel of the linear operator $\mathbb{R}^m \to \mathbb{R}^{m*}$: $u \mapsto i_u\omega$. That is, the set of vectors u such that $i_u\omega(v) = 0$ for all $v \in \mathbb{R}^m$. A form ω is called *non-degenerate,* if its kernel is zero.

Remark 3.2 A non-degenerate antisymmetric bilinear form ω exists only on even-dimensional vector spaces $\mathbb{R}^{2n}$. It always generates a non-zero volume form $\omega^{\wedge n}$.

Definition 3.3 A *symplectic manifold* is a manifold M equipped with a closed non-degenerate 2-form ω. Non-degenerate means that at each point $x \in M$ the corresponding antisymmetric bilinear form on T_xM is non-degenerate: has zero kernel. Due to the above remark, each symplectic manifold is even-dimensional and has a natural volume form $\omega^{\wedge n}$, $n = \frac{1}{2}\dim M$.

Example 3.4 The *standard symplectic space* equipped with standard (canonical) symplectic form is

$$(\mathbb{R}^{2n}_{q_1,\dots,q_n,p_1,\dots,p_n}, \omega_{st}), \quad \omega_{st} := dq \wedge dp = \sum_{j=1}^{n} dq_j \wedge dp_j.$$

Another similar example is the torus $(\mathbb{T}^{2n}, \omega_{st})$.

Example 3.5 The cotangent bundle T^*M of a smooth n-dimensional manifold M carries a natural symplectic form ω. Indeed, consider the *Liouville 1-form* α on T^*M defined as follows. Take an arbitrary $q \in M$, a $p \in T^*_q M$ and a vector $v \in T_{(q,p)}(T * M)$. Let $\pi : T^*M \to M$ denote the standard bundle projection. Set

$$\alpha(v) := p(d\pi(v)), \quad \omega := -d\alpha.$$

Let now U be a local chart on M identified with a domain in $\mathbb{R}^n_{q_1,\dots,q_n}$. There is a canonical isomorphism $\pi^{-1}(U) = U \times \mathbb{R}^n_{p_1,\dots,p_n}$ such that for every points $\widetilde{q} \in U$ and $p = (p_1, \dots, p_n) \in \mathbb{R}^n$ the corresponding 1-form is on $T_{\widetilde{q}}U$ is $pdq = \sum_{j=1}^{n} p_j dq_j$. Then for every v as above one has

$$\alpha(v) = \sum_j p_j dq_j(d\pi(v)) = (\sum_j p_j dq_j)(v).$$

This means that

$$\alpha = \sum_j p_j dq_j, \quad \omega = -d\alpha = \omega_{st} = dq \wedge dp.$$

3.2 Hamiltonian vector fields. Basic conservation laws

Definition 3.6 Let (M, ω) be a symplectic manifold, $H : M \to \mathbb{R}$ be a smooth function. The *Hamiltonian vector field* with the Hamiltonian function H is the unique vector field X_H on M such that

$$i_{X_H}\omega = dH. \tag{3.1}$$

Proposition 3.7 *The flow of the Hamiltonian vector field X_H preserves H and the symplectic form ω.*

Proof One has $\frac{dH}{dX_H} = \omega(X_H, X_H) = -\omega(X_H, X_H)$, by definition and since ω is anti-symmetric. Therefore, the latter quantity vanishes, thus, H has zero derivative along the field X_H. Hence, it is invariant under the flow.

The proof of invariance of the form ω is based on Cartan Formula for Lie derivative. Namely, recall that for a given k-form ω and a vector field X on a manifold M the Lie derivative $L_X\omega$ of the form ω along the field X is defined as follows. Let g_X^t denote the time t flow map of the field X. Then

$$L_X\omega := \frac{d}{dt}\left((g_X^t)^*\omega\right)|_{t=0}.$$

Here $(g_X^t)^*\omega$ is the pullback of the form ω under the time t flow map. The *Cartan Formula* is

$$L_X\omega = i_X(d\omega) + d(i_X\omega). \tag{3.2}$$

Applying it to the symplectic form ω and the field X_H we get

$$L_{X_H}\omega = i_{X_H}(d\omega) + d(i_{X_H}\omega) = 0 + d(dH) = 0,$$

since $d\omega = 0$ and $i_{X_H}\omega = dH$, by definition. Finally, the Lie derivative of the form ω along the field X_H vanishes. Hence, ω is invariant under its flow. The proposition is proved. □

Corollary 3.8 *Each Hamiltonian vector field (and in general, each vector field preserving the symplectic form) is always volume-preserving for the volume form generated by the symplectic form.*

Exercise 3.9 Prove Cartan Formula.

Hint. It suffices to prove it in the case, when M is a domain in $\mathbb{R}^m$ and X is a constant vector field, say $X = \frac{\partial}{\partial x_1}$. In this case

$$L_X\Big(\sum_{i_1<\dots<i_k} a_{i_1\dots i_k}(x)dx^{i_1}\wedge\dots\wedge dx^{i_k}\Big) = \sum \frac{\partial}{\partial x_1}a_{i_1\dots i_k}(x)dx^{i_1}\wedge\dots\wedge dx^{i_k}.$$

Prove the latter equality.

3.3 Lie algebra structures. Poisson brackets of functions and Lie bracket of vector fields

Definition 3.10 Let (M,ω) be a symplectic manifold. The *Poisson bracket* of smooth functions $F, G : M \to \mathbb{R}$ is

$$\{F,G\} := \omega(X_F, X_G) = dF(X_G) = -dG(X_F). \tag{3.3}$$

Let us recall the following definition from differential geometry.

Definition 3.11 The *Lie bracket* (or commutator) of two vector fields X and Y on a manifold M is the field $Z = [X, Y]$ defined in any of the two following equivalent ways.

1) Set $[X, Y]$ to be the *Lie derivative* $L_X Y$ of the field Y along X, that is

$$[X, Y] = L_X Y := \frac{d}{dt}\left((g_X^{-t})_* Y\right).$$

Here $Y_t = (g_X^{-t})_* Y$ is the pushforward of the field Y under the flow map g_X^{-t}, that is,

$$Y_t(q) := (dg_X^{-t})(g_X^t(q))Y(g_X^t(q)).$$

2) For every vector field X on M consider the corresponding first order linear differential operator $L_X : f \mapsto L_X f := df(X)$ acting on smooth functions on M. For every two vector fields X, Y on M one has

$$L_X L_Y - L_Y L_X = L_Z, \quad Z \text{ is a vector field on } M. \tag{3.4}$$

Indeed, the latter commutator is a first order differential operator: the terms with second derivatives cancel out due to relation $\frac{\partial^2 f}{dx_i dx_j} = \frac{\partial^2 f}{dx_j dx_i}$. The field Z is called the Lie bracket $[X, Y]$ of the fields X and Y.

Theorem 3.12 *The C^∞-smooth vector fields on M form a Lie algebra under the Lie bracket.*

Proof Recall that an abstract associative algebra always carries a standard Lie algebra structure, $[a, b] := ab - ba$, and the Jacobi identity follows immediately by definition. Let us embed the vector fields to the associative algebra of differential operators acting on functions, with multiplication given by composition via the operator $X \mapsto L_X$. The restriction of the associative algebra bracket to the vector fields coincides with the Lie bracket, by (3.4). This together with the Jacobi identity in the associative algebra implies the Jacobi identity for the Lie bracket. □

Theorem 3.13 *1) The Poisson bracket is a Lie algebra structure on the space of C^∞-smooth functions, i.e., it satisfies the Jacobi identity*

$$\{\{A, B\}, C\} + \{\{C, A\}, B\} + \{\{B, C\}, A\} = 0. \tag{3.5}$$

2) One has

$$[X_F, X_G] = -X_{\{F,G\}} \quad \textit{for every two functions } F, G : M \to \mathbb{R}. \tag{3.6}$$

In other terms, the map $F \mapsto -X_F$ sending a function to minus the corresponding Hamiltonian vector field is a Lie algebra homomorphism between the function space equipped with Poisson bracket and the vector field space equipped with the Lie brackets. The Hamiltonian vector fields thus form a Lie subalgebra in the Lie algebra of vector fields.

Proof We use the following Leibnitz rule for the Lie derivative.

Proposition 3.14 *For every C^1-smooth k-form ω on a manifold M and every collection of C^1-smooth vector fields Y, $X_1, \dots, X_k$ one has*

$$L_Y(\omega(X_1, \dots, X_k)) = (L_Y\omega)(X_1, \dots, X_k)$$

$$+\sum_{j=1}^{k} \omega(X_1, \dots, X_{j-1}, L_Y X_j, X_{j+1}, \dots, X_k). \tag{3.7}$$

Proof Let $\omega_t = (g_Y^t)^*\omega$ denote the pullback of the form ω under the time t flow map g_Y^t of the field Y. Let $X_{j,t} := (g_Y^{-t})_* X_j$ denote the pushforwards of the fields X_j under the inverse map. The value at a point $p \in M$ of the Lie derivatives in (3.7) are equal to the derivatives in t at $t = 0$ of the families ω_t and $X_{j,t}$ respectively. Now let us fix a point p and consider the latter families as families of forms and vectors in $T_pM = \mathbb{R}^n$. The above statement together with the usual Leibnitz rule applied to the family of functions $\omega_t(X_{1,t}, \dots, X_{k,t})$ in t as sums of products of the components of the vectors $X_{j,t}$ and coefficients of the form ω_t imply (3.7). $\square$

Let us now prove (3.5). One has

$$\{\{A, B\}, C\} = \{\omega(X_A, X_B), C\} = L_{X_C}(\omega(X_A, X_B))$$

$$= \omega([X_C, X_A], X_B) + \omega(X_A, [X_C, X_B]), \tag{3.8}$$

by (3.7) and since $L_{X_C}\omega = 0$ (invariance of the form ω under a Hamiltonian flow) and since the Lie derivative of a vector field is a commutator. The latter right-hand side is equal to

$$-dB([X_C, X_A]) + dA([X_C, X_B]) = (L_{X_A}L_{X_C} - L_{X_C}L_{X_A})B + (L_{X_C}L_{X_B} - L_{X_B}L_{X_C})A$$

$$= \{\{B, C\}, A\} - \{\{B, A\}, C\} + \{\{A, B\}, C\} - \{\{A, C\}, B\},$$

since $L_{X_C}B = \{B, C\}$ etc. This together with (3.8) yields

$$\{\{B, C\}, A\} - \{\{B, A\}, C\} - \{\{A, C\}, B\} = 0,$$

which is equivalent to (3.5).

Let us prove Statement 2). The sum of the two first terms in (3.5) is equal to

$$(L_{X_C}L_{X_B} - L_{X_B}L_{X_C})A = L_{[X_C,X_B]}A.$$

The third term is equal to minus the derivative of the function A along the field $X_{\{B,C\}}$. The above statements hold for every function A. This together with (3.5) yields $[X_C, X_B] = X_{\{B,C\}} = -X_{\{C,B\}}$. □

3.4 Canonical coordinates. Darboux Theorem

Definition 3.15 Local coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ on a symplectic manifold are called *canonical,* if in these coordinates the symplectic form is *standard:*

$$\omega = \omega_{st} := dq \wedge dp = \sum_{j=1}^{n} dq_j \wedge dp_j.$$

Theorem 3.16 (Darboux). *On every symplectic manifold each point has a neighborhood on which canonical coordinates exists. That is, each symplectic form is standard in some local chart.*

Proof We prove Theorem 3.16 by induction in dimension. To do this, let us first make the following preliminary construction. Let (M, ω) be a symplectic manifold. Fix a point $x \in M$ and its neighborhood $U = U(x) \subset M$ on which some smooth local chart exists. Thus, we can and will consider that U is a domain in $\mathbb{R}^{2n}$ equipped with some symplectic form ω. Take an arbitrary smooth function $H : U \to \mathbb{R}$ without critical points, say, a coordinate function, $H(x) = 0$. Fix a hypersurface $\Gamma \subset U$ through x transversal to the Hamiltonian vector field X_H. For every $y \in U$ let $t(y)$ denote the time needed to go to y from the hypersurface Γ along an orbit of the field X_H. The above Γ exists and the function $t : U \to \mathbb{R}$ is well-defined if U is chosen small enough. Then

$$\{t, H\} = dt(X_H) = 1, \tag{3.9}$$

by construction, and $t(x) = H(x) = 0$.

Induction base: $n = 1$. Then $q = t$, $p = H$ are canonical coordinates.

Induction step. Let Theorem 3.16 be proved for $n = k$. Let us prove it for $n = k + 1$. Set $q_1 = t$, $p_1 = H$. Let us construct additional functions $q_2, \dots, q_n, p_2, \dots, p_n$ forming, together with q_1, p_1, a system of canonical coordinates. To do this, consider the projection $\pi : U \to \mathbb{R}^2$, $y \mapsto (q_1, p_1)$ and for every $(a, b) \in \mathbb{R}^2$ set

$$S_{a,b} := \pi^{-1}(a, b) \subset \Gamma_a = \{H = a\}, \quad S := S_{0,0} \subset \Gamma = \Gamma_0.$$

Each $S_{a,b}$ is a hypersurface in Γ_a transversal to the field X_H. The restriction to $S_{a,b}$ of the form ω is symplectic due to the next proposition.

Proposition 3.17 *Consider the vector space $\mathbb{R}^{2n}$ equipped with a non-degenerate skew-symmetric bilinear form ω. For every codimension one vector subspace $\Gamma \subset \mathbb{R}^{2n}$ the restriction to Γ of the form ω has one-dimensional kernel. For every codimension one vector subspace $S \subset \Gamma$ transversal to the kernel the restriction $\omega|_S$ is non-degenerate.*

Proof Fix a vector $v \in \mathbb{R}^{2n}$ transversal to Γ. Suppose the contrary: there exist two linearly independent vectors $u_1, u_2 \in \Gamma$ lying in the kernel of the form $\omega|_\Gamma$. There exists their non-trivial linear combination $\eta = a_1u_2 + a_2u_2$ such that $\omega(\eta, v) = 0$. Then $\omega(\eta, z) = 0$ for every $z \in \mathbb{R}^{2n}$, by construction. Hence, the form ω is degenerate on the ambient space $\mathbb{R}^{n2}$, – a contradiction.

Suppose, by contradiction, that the form $\omega|_S$ has a non-zero kernel, let u be its element. Then $\omega(u, X_H) = 0$, and hence u lies in the kernel of the restriction $\omega|_\Gamma$, as does X_H. Therefore, the vectors u and X_H are proportional, which contradicts transversality of the vector X_H and S in Γ. The proposition is proved. □

Proposition 3.18 *There is a smooth foliation $\mathcal{F}$ of the domain U by surfaces tangent to the commuting fields X_t and X_H. Its leaves are transversal to the fibers $S_{a,b}$, or equivalently, they are locally diffeomorphically projected to $\mathbb{R}^2$ by the projection π to the coordinates (q_1, p_1).*

The proposition follows from commutativity and linear independence of the fields X_t, X_H at each point and the equality $dt(X_H) = 1 = -dH(X_t) = 1$.

There exist canonical coordinates $(q_2, \dots, q_n, p_2, \dots, p_n)$ on the symplectic manifold $S \cap U$ after shrinking U, by the induction hypothesis. Let us extend them to the whole U by requiring that they be constant along the leaves of the foliation $\mathcal{F}$.

Proposition 3.19 *The coordinates $(q_1, \dots, q_n, p_1, \dots, p_n)$ on U thus constructed are canonical.*

Proof For every coordinate $\zeta = q_j, p_j$ with $j \geq 2$ one has

$$\{\zeta, q_1\} = d\zeta(X_t) = \{\zeta, p_1\} = d\zeta(X_H) = 0,$$

since ζ is constant along surfaces tangent to X_t and X_H. It remains to check that for every $j, \ell = 2, \dots, n$ one has

$$\{q_j, q_\ell\} = \{p_j, p_\ell\} = 0, \quad \{q_j, p_\ell\} = \delta_{j\ell}. \tag{3.10}$$

Proposition 3.20 *Each function $\zeta = q_j, p_j$, $j \geq 2$, satisfies the following statements.*

1) The vector field X_ζ is tangent to the fibers $S_{a,b}$, i.e., has zero projection $d\pi$ to the coordinates (q_1, p_1).

2) The field X_ζ restricted to each fiber $S_{a,b}$ coincides with the Hamiltonian vector field $X_{\widetilde{\zeta}}$ of the restriction $\widetilde{\zeta} := \zeta|_{S_{a,b}}$.

Proof One has $dq_1(X_\zeta) = \{q_1, X_\zeta\} = 0$. This imples Statement 1). Thus, at each point $y \in S_{a,b}$ both vectors $u := X_\zeta(y)$ and $v := X_{\widetilde{\zeta}}(y)$ lie in $T_y S_{a,b}$ and satisfy the relation $i_u\omega = i_v\omega = dq_j$ on $T_y S_{a,b}$. This together with non-degeneracy of the form ω on $T_y S_{a,b}$ implies that $u = v$ and proves the proposition. □

The Poisson bracket of any two coordinate functions $q_2, \dots, q_n, p_2, \dots, p_\ell$ is equal to the value of the form ω on the corresponding Hamiltonian vector fields. The latter coincide with the Hamiltonian fields of the restriction to $S_{a,b}$ of the functions in question, by Statement 2) of Proposition 3.20. Therefore, the bracket in question is equal to the bracket of the restrictions of the functions to $S_{a,b}$. But the above coordinates restricted to each $S_{a,b}$ are canonical, since this holds on $S = S_{0,0}$, by assumption, and by invariance of the form and the fibration under flow maps of the fields X_{q_1} and X_{p_1}. This implies (3.10) and proves Proposition 3.19. □

Proposition 3.19 immediately implies the Darboux Theorem. □

3.5 Symplectomorphisms

Definition 3.21 A diffeomorphism between two symplectic manifolds is called a *symplectomorphism,* if its pullback map sends one symplectic form into the other. A self-diffeomorphism F of a symplectic manifold that preserves the symplectic form is called a symplectomorphism or a symplectic automorphism. If the underlying invariant symplectic form is standard, i.e., $dq \wedge dp$ in some coordinates (q, p) on a source domain and on its image, then F is called a *canonical transformation.*

Example 3.22 1) Translations of the space $\mathbb{R}^{2n}_{q_1,\dots,q_n,p_1,\dots,p_n}$ equipped with the standard form $dq \wedge dp$ are symplectomorphism.

2) Consider the product $M = \mathbb{R}^n_q \times \mathbb{R}^{n*}_p$ of a vector space and the space of linear functionals on it. Let α be the 1-form on it that is defined as follows: for every $(q,p) \in M$ and every $v \in T_{(q,p)}M$ one has $\alpha(v) = p(d\pi(v))$, where π is the projection to the q-space. The differential $d\alpha$ is the standard symplectic form on M. In the coordinates $(q_1,\dots,q_n,p_1,\dots,p_n)$, where p_j are the coordinates on $\mathbb{R}^{n*}$ dual to q_ℓ, one has $\omega = \omega_{st} = dq \wedge dp$. Each linear automorphism $A : \mathbb{R}^n \to \mathbb{R}^n$ induces the conjugate $A^* : \mathbb{R}^{n*} \to \mathbb{R}^{n*}$ acting by taking pullback: $(A^*p)(v) = p(Av)$. In the above coordinates A^* is given by the matrix A^t transposed to A. The linear automorphism

$$F : M \to M, \quad F(q,p) := (A^{-1}q, A^*p)$$

preserves the form α, and hence ω. See a more general statement below.

Example 3.23 3) Let N be an arbitrary smooth manifold, set $M = T^*N$. Recall that M is equipped with the standard symplectic form $\omega = d\alpha$, where α is the Liouville form. We denote points in M by (q,p), $q \in N$, $p \in T^*_qN$. Let β be a smooth 1-form on N. Its restriction to the fiber T_qN will be denoted by β_q. *The map*

$$F_\beta : (q,p) \mapsto (q, p + \beta_q) \tag{3.11}$$

is a symplectomorphism, if and only if the form β is closed. Indeed,

$$F^*_\beta \alpha = \alpha + \pi^*\beta, \tag{3.12}$$

which follows by definition, since for every $(q,p) \in M$ and $v \in T_{(q,p)}M$ one has $\alpha(v) = p(d\pi(v))$. Therefore, $F^*_\beta\omega = \omega + \pi^*(d\beta)$. This implies the above symplectomorphicity criterium. In standard coordinates (q,p) on M over a local chart $(q_1,\dots,q_n)$ on N the form β is represented by a vector function

$$\beta = (\beta_1(q), \dots, \beta_n(q)).$$

Symplectomorphicity of the map F_β, i.e., closeness of the form β, is equivalent to the system of equations

$$\frac{\partial \beta_i}{\partial q_j} = \frac{\partial \beta_j}{\partial q_i} \quad \text{for every } i,j = 1,\dots,n. \tag{3.13}$$

4) In the above conditions let $f : N \to N$ be a smooth diffeomorphism, and let $F = f^* : T^*N \to T^*N$ be the corresponding pullback map acting

on 1-forms on fibers. *Then $F : M \to M$ preserves the Liouville form α, and hence, is a symplectomorphism.* Indeed, it sends each point (q, p) to $(f^{-1}(q), (df(f^{-1}(q)))^* p)$. Its differential sends a vector $v \in T_{(q,p)}M$ to the vector in $T_{f(q,p)}M$ whose projection to the base is $(df(f^{-1}(q)))^{-1}(d\pi(v)) \in T_{f^{-1}(q)}N$. Therefore, the value on v of the pullback of the form α under the map F is equal to

$$(df(f^{-1}(q)))^* p)[(df(f^{-1}(q)))^{-1}(d\pi(v))]$$
$$= p(df(f^{-1}(q)(df(f^{-1}(q)))^{-1}d\pi(v)) = p(d\pi(v)) = \alpha(v).$$

Thus, F is a symplectomorphism. Let $q_1, \dots, q_n$ be local coordinates near a point $y \in N$ and let us denote local coordinates near $f(y)$ by the same symbol $q_1, \dots, q_n$. Let $p_1, \dots, p_n$ denote the corresponding dual linear coordinates on the fibers in T^*N: the coordinates (q, p) are canonical for the form ω. Then in the coordinates (q, p) the symplectomorphism F is a canonical map given by the formula

$$(q, p) \mapsto (f^{-1}(q), (df(f^{-1}(q)))^* p). \tag{3.14}$$

3.6 Generating functions. Local space of symplectomorphisms

Definition 3.24 Consider an arbitrary diffeomorphism between domains,

$$F : U \to V, \quad U \subset \mathbb{R}^{2n}_{q,p}, \quad V \subset \mathbb{R}^{2n}_{Q,P},$$
$$q = (q_1, \dots, q_n), \ p = (p_1, \dots, p_n), \ Q = (Q_1, \dots, Q_n), \ P = (P_1, \dots, P_n).$$

We say that the coordinates (q, Q) are *free,* if the map $(q, p) \mapsto (q, Q \circ F)$ is a diffeomorphism on U. Note that a necessary condition to be free is that the Jacobian matrix $\frac{\partial Q}{\partial p}$ is non-degenerate.

In what follows we consider that U is simply connected. We equip the ambient space $\mathbb{R}^{2n}$ with the standard symplectic form $\omega_{st} = dq \wedge dp$, respectively $dQ \wedge dP$.

Theorem 3.25 *In the above assumptions the diffeomorphism F is a canonical transformation, i.e., a symplectomorphism, if and only if*

$$PdQ - pdq = dg(q, Q), \quad g \textit{ is a function on } U,$$

the function g is written as a function in free coordinates (q, Q) parametrizing U. Here PdQ denotes the 1-form on U that is the pullback of the form PdQ under the map F. The above function g, if exists, is unique up to additive constant.

Proof The form $-pdq$ is a primitive of the symplectic form: $d(-pdq) = dq \wedge dp$. Therefore, the map F is a symplectomorphism, if and only if the pullback of the form PdQ and the form pdq differ by a closed 1-form. Each closed 1-form on a simply connected domain U is exact, the differential of a function uniquely determined by the form up to additive constant. This implies the statement of the theorem. □

Definition 3.26 If in the conditions of Theorem 3.25 the map F is a symplectomorphism, then the corresponding function g is called its *generating function.*

Remark 3.27 Theorem 3.25 implies that locally the space of symplectomorphisms looks like the space of functions of $2n$ variables.

4 Integrable systems. Arnold–Liouville Theorem

The *standard* integrable Hamiltonian system is a system in canonical variables (q, p) with a Hamiltonian function on $U \subset \mathbb{R}^{2n}$ that depends only on n variables, say, $H = H(p)$. In this case the system takes the form

$$\begin{cases} \dot{q} = w(p) = \frac{dH}{dp} \\ \dot{p} = 0, \end{cases} \tag{4.1}$$

and the system has n first integrals $p_1, \ldots, p_n$. They are *in involution,* which means that $\{p_i, p_j\} = 0$. The next classical theorem states the converse.

Theorem 4.1 (Arnold – Liouville). *Let a Hamiltonian system on a $2n$-dimensional symplectic manifold (M, ω) with a Hamiltonian function H have n independent first integrals in involution*

$$H_1 = H,\ H_2, \ldots, H_n : \quad \text{"in involution" means that } \{H_i, H_j\} = 0.$$

Independence means that $dH_j(y)$ are linearly independent at every point $y \in M$. Consider the vector function $\mathcal{H} := (H_1, \ldots, H_n)$.

1) The level surfaces $\mathcal{H}^{-1}(a)$, $a = (a_1, \ldots, a_n) \in \mathcal{H}(M) \subset \mathbb{R}^n$, are submanifolds tangent to all the vector fields X_{H_j}. They are **Lagrangian,** *which means that the restriction to each of them of the symplectic form is zero.*

2) If S_α is a compact connected component of the level set $\mathcal{H}^{-1}(\alpha)$, then the following statements hold:

a) There exists a diffeomorphism $S_\alpha \to \mathbb{T}^n = \mathbb{R}^n_{\phi_1,\ldots,\phi_n}/2\pi\mathbb{Z}^n$.

b) The above diffeomorphism can be chosen to transform X_{H_j} to constant vector fields on $\mathbb{T}^n$ depending only on the parameter α. The restriction to S_α of the system thus takes the form

$$\dot{\phi} = w(\alpha), \quad w(\alpha) = (w_1(\alpha), \dots, w_n(\alpha)).$$

c) There exists a neighborhood $U = U(\alpha) \subset \mathbb{R}^n$ such that for every $a \in U$ the level set $\mathcal{H}^{-1}(a)$ contains a compact connected component S_a close to S_α and the open subset

$$N := \cup_{a \in U} S_a \subset M$$

admits a symplectomorphism

$$F : (N, \omega) \to (\mathbb{T}^n \times V, \omega_{st}), \;\; \mathbb{T}^n = \mathbb{R}^n_{\phi_1,\dots,\phi_n}/2\pi\mathbb{Z}^n, \;\; V \subset \mathbb{R}^n_{I_1,\dots,I_n}, \; \omega_{st} = d\phi \wedge dI,$$

satisfying the following statements:

(i) one has $I \circ F = g \circ \mathcal{H}$, where g is a diffeomorphism $U \to V$, i.e., F sends each fiber S_a to a toric fiber;

(ii) in the new coordinates (ϕ, I) the Hamiltonian functions H_j are functions just of I, thus so are the fields X_{H_j}, and the system takes the form

$$\begin{cases} \dot{\phi} = w(I) := \frac{\partial H}{\partial I} \\ \dot{I} = 0. \end{cases} \tag{4.2}$$

In particular, the latter system is integrable in quadratures.

Definition 4.2 The above coordinates (ϕ, I) are called the *action-angle coordinates: I* is the action; ϕ is the angle.

Proof of Theorem 4.1. Let us first prove Statements 1), 2a),b). The level subsets $\{\mathcal{H} = a\}$ are smooth n-dimensional submanifolds, by independence of the functions H_i. One has $dH_j(X_{H_i}) = \{H_j, H_i\} = \omega(X_{H_j}, X_{H_i}) = 0$ for every i and j. Thus, $H_j = const$ along orbits of each field X_{H_i}. Hence, X_{H_i} is tangent to each above level submanifold, and hence, to S_α. The above statement also implies that the fields $X_{H_i}(y)$ generate the whole tangent space to the level submanifold through each y, and ω vanishes on each pair of them. Hence, it is a Lagrangian submanifold. Statement 1) is proved.

The vector fields X_{H_i} are linearly independent at each point $y \in M$, as are the differentials $dH_i(y)$. They commute, since $\{H_i, H_j\} = 0$. For every compact component S_a of the submanifold $\{\mathcal{H} = a\}$ the restrictions to S_a of the fields X_{H_i} have well-defined flow maps $g^t_{X_{H_i}} : S_a \to S_a$ for every $t \in \mathbb{R}$. Thus, for every $x_0 \in S_a$ there is a smooth map

$$\Pi_a : \mathbb{R}^n_{t_1,\dots,t_n} \to S_a, \; (t_1, \dots, t_n) \mapsto g^{t_1}_{X_{H_1}} \circ \dots \circ g^{t_n}_{X_{H_n}}(x_0). \tag{4.3}$$

Claim 5. *The map $\Pi_a : \mathbb{R}^n \to S_a$ is a universal covering.*
Proof The image $\Pi_a(\mathbb{R}^n) \subset S_a$ is open. Moreover, there exist $\delta, \varepsilon > 0$ such that the map Π_a is injective on the ball of radius δ around each point $z \in \mathbb{R}^n$, and its Π_a-image contains the ε-ball in S_a centered at $\Pi_a(z)$. By definition, the latter ε-ball is the intersection with S_a of the ε-neighborhood of the point x_0 in some fixed Riemannian metric on M. The above statement follows by linear independence of the fields X_{H_i} and compactness. It implies that the image $\Pi_a(\mathbb{R}^n)$ has no boundary points and hence, coincides with all of S_a, and the map Π_a is a covering, hence universal. □

The above coordinates $(t_1, \dots, t_n)$ on the universal coverings are uniquely defined up to change of the point x_0, which results in translation of the coordinates t by a vector depending on the transversal parameter a. The liftings to $\mathbb{R}^n$ of the vector fields X_{H_j} coincide with the fields $\frac{\partial}{\partial t_j}$, by construction. The covering desk transformations $\mathbb{R}^n \to \mathbb{R}^n$ are translations, since they preserve the above fields. Thus, S_a is isomorphic to the quotient of the space $\mathbb{R}^n$ by a discrete additive subgroup Λ with compact quotient. The subgroup Λ has a finite number, say k, of generators. They are linearly independent over $\mathbb{R}$. Indeed, if, to the contrary, there were a subsystem of generators that are linearly dependent over $\mathbb{R}$ but not over $\mathbb{Q}$, then they would generate a non-discrete subgroup, – a contradiction. Their number k is equal to n: otherwise their linear combinations with real coefficients do not cover all of $\mathbb{R}^n$, then the quotient is not compact. Finally, Λ has n generators, linearly independent over $\mathbb{R}$. Hence, there is a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ sending Λ to the lattice $2\pi\mathbb{Z}^n$. Together with the covering projection $\Pi_a : \mathbb{R}^n \to S_a$, it induces a diffeomorphism $S_a \to \mathbb{T}^n$ sending each X_{H_j} to a constant field. Statements 2a),b) are proved.

Let us prove Statement 2c). The first statement, existence of a compact component S_a close to S_α in $\mathcal{H}^{-1}(a)$ for a close to α follows immediately by transversality. It implies, again by transversality, that the union N is a n-dimensional submanifold. It remains to prove existence of a rectifying symplectomorphism F. This will be done in the following steps. They deal with the universal covering $\widetilde{N}$ of the manifold N, which is the union of the universal coverings $\widetilde{S}_a$ of the topological tori S_a.

Step 1). Construction of preliminary canonical coordinates $(t_1, \dots, t_n, \mathcal{H}_1, \dots, \mathcal{H}_n)$ on the universal cover $\widetilde{N}$. We have already constructed some preliminary coordinates $t_1, \dots, t_n$ on universal coverings $\widetilde{S}_a$, see (4.3). They were defined by a base point $x_0 \in S_a$, but for different choises of x_0 the corresponding t-charts on $\widetilde{S}_a$ differ by translation. The above coordinates $t_1, \dots, t_n$ identify each $\widetilde{S}_a$ with $\mathbb{R}^n$, and $X_{H_i} = \frac{d}{dt_i}$. Let us correct the cordinates t by fiber-depending translations in order make the coordinates $(t, \mathcal{H})$ canonical:

$$\{t_i, t_j\} = 0, \{t_i, H_j\} = \delta_{ij}. \tag{4.4}$$

1a) Construction of the coordinate t_1. Recall that the vector fields X_{H_i} commute. Therefore, there exists a foliation, denoted by $\mathcal{F}_1$, by $(n-1)$-dimensional leaves (local submanifolds), each leaf being saturated by orbits of the fields $X_{H_2}, \dots, X_{H_n}$. Fix an integral hypersurface Σ_1 of the foliation $\mathcal{F}_1$, i.e., a hypersurface that is a union of leaves, that is transversal to the vector field X_{H_1}. It cuts each universal cover $\widetilde{S}_a$ by an affine subspace parallel to the subspace generated by the vectors $X_{H_i} = \frac{d}{dt_i}$. For every $y \in \widetilde{N}$ close to $\widetilde{S}_\alpha$ let $t_1(y)$ denote the time needed to reach y along a trajectory of the field X_{H_1} starting at Σ_1. On each universal covering it coincides with the previously constructed coordinate t_1 up to translation. One has

$$\{t_1, H_j\} = dt_1(X_{H_j}) = \delta_{1j}, \tag{4.5}$$

Indeed, the above equalities holds on Σ_1, by construction. They remain valid along orbits of the field X_{H_1}: its time τ flow translates t_1 by τ and preserves each field X_{H_j}, since $[X_{H_1}, X_{H_j}] = 0$.

1b) Construction of the coordinate t_2. The vector fields $X_{t_1}, X_{H_1}, \dots, X_{H_n}$ commute, by assumption and (4.5). They are linearly independent at each point, by construction. Let $\mathcal{F}_2$ denote the foliation by n-dimensional leaves saturated by orbits of the fields $X_{t_1}, X_{H_1}, X_{H_3}, \dots, X_{H_n}$. Fix its integral hypersurface Σ_2 that is transversal to the field X_{H_2}. Let us define the time function t_2 as above with Σ_1, X_{H_1} replaced by Σ_2, X_{H_2}. One has

$$\{t_2, H_j\} = \delta_{2j}, \ \{t_1, t_2\} = dt_2(X_{t_1}) = 0, \tag{4.6}$$

as in the above construction.

1c) Construction of $t_3, \dots, t_n$. The fields $X_{t_1}, X_{t_2}, X_{H_1}, \dots, X_{H_n}$ commute, by (4.6). Let $\mathcal{F}_3$ denote the foliation by $(n+1)$-dimensional leaves saturated by orbits of the fields $X_{t_1}, X_{t_2}, X_{H_1}, X_{H_2}, X_{H_4}, \dots, X_{H_n}$. Fix its integral surface Σ_3 transversal to X_{H_3} and construct t_3 as above. Continuing this procedure we construct coordinates $t_1, \dots, t_n$ satisfying (4.4) and hence, canonical.

The vector fields X_{H_j} lifted to $\widetilde{N}$ are tangent to the fibers $\mathcal{H} = const$ and coincide with the fields $\frac{\partial}{\partial t_j}$. The desk transformations of the covering $\widetilde{N}$ preserve the lifted fields and the vector function $\mathcal{H}$. Hence, they are translations by vectors depending on the transversal parameter $a = \mathcal{H}$. For every a the translation vectors form a cocompact lattice, i.e., an additive subgroup in $\mathbb{R}^n$ with compact quotient. Hence the above lattice is generated over $\mathbb{Z}$ by n vectors, which we will denote $v_1, \dots, v_n$. Each $v_d = v_d(\mathcal{H})$ is a function of the coordinates $\mathcal{H} = (H_1, \dots, H_n)$.

Step 2) Simultaneous symplectic rectification of the lattices. Here we construct a symplectomorphism

$$F : (\widetilde{N}, \omega) \to (\mathbb{R}^n_\phi \times V, \omega_{st}) \;\; V \subset \mathbb{R}^n_I, \quad \phi = (\phi_1, \dots, \phi_n), \quad I = (I_1, \dots, I_n).$$

Here by ω we denote the lifting to $\widetilde{N}$ of the form ω, and $\omega_{st} = d\phi \wedge dI$. We construct F so that I is a vector function of the coordinates $\mathcal{H}$, and F sends fibers to fibers by affine maps depending on the parameter $\mathcal{H}$. We construct F so that the corresponding linear maps send the basis lattice vectors $v_d(\mathcal{H})$ to the standard basis vectors $2\pi e_j$. This will imply that F passes to a quotient symplectomorphism $N \to \mathbb{T}^n \times V$. Then the corresponding coordinates (ϕ, I) are action-angle coordinates we are looking for.

In the above coordinates $(t, \mathcal{H})$ each covering desk transformation

$$T_d : \widetilde{N} \to \widetilde{N}, \; T_d : (t, \mathcal{H}) \to (t + v_d(\mathcal{H}), \mathcal{H})$$

is a symplectomorphism for the standard form ω_{st}. Let $v_d = (v_{1d}, \dots, v_{nd})$ be the coordinate representation of the vector function v_d. The 1-form $\sum_j v_{jd}(\mathcal{H}) dH_j$ is closed, by symplectomorphicity, see Example 3.23. We consider that it is exact, choosing U simply connected. Thus,

$$v_d = \left(\frac{\partial g_d}{\partial H_1}, \dots, \frac{\partial g_d}{\partial H_n}\right) \;\text{ for some function }\; g_d(\mathcal{H}) \text{ on } U.$$

Consider the map

$$g : U \to \mathbb{R}^n, (H_1, \dots, H_n) \mapsto (g_1(\mathcal{H}), \dots, g_n(\mathcal{H})).$$

It is a local diffeomorphism, since the gradients $v_d(\mathcal{H})$ of the functions $g_d(\mathcal{H})$ are linearly independent at every $\mathcal{H} \in U$. We can and will consider that it is a diffeomorphism $U \to V \subset \mathbb{R}^n$, shrinking U. Consider the map

$$F : \widetilde{N} \to \mathbb{R}^n_\phi \times \mathbb{R}^n_I, \quad (t, \mathcal{H}) \mapsto (2\pi t (dg(\mathcal{H}))^{-1}, (2\pi)^{-1} g(\mathcal{H})). \tag{4.7}$$

Here $t(dg(\mathcal{H}))^{-1}$ is the product of a horizontal vector $t = (t_1, \dots, t_n)$ and the matrix $(dg(\mathcal{H}))^{-1}$. The above map F can be considered as a map between the cotangent bundles of the domains U and V, where the bases are equipped with coordinates $\mathcal{H}$ and I respectively and the fibers with coordinates t and ϕ respectively. Recall that $(t, \mathcal{H})$ and (ϕ, I) are canonical coordinates.

The map F is the lifting to the cotangent bundle of the diffeomorphism $(2\pi)^{-1}g(\mathcal{H})$, which follows from (4.7). Hence, it is a symplectomorphism, see Example 3.23, 4). Its restriction to each fiber is a linear operator $\mathbb{R}^n \to \mathbb{R}^n$ sending each vector $v_d(\mathcal{H})$ to $2\pi e_d$; here e_d is the d-th vector in the standard basis. Thus, F induces a quotient symplectomorphism $N \to \mathbb{T}^n_\phi \times V_I$. The action-angle coordinates (ϕ, I) are constructed. Theorem 4.1 is proved. $\square$

Remark 4.3 In the proof of Arnold–Liouville Theorem given in [1] the construction of action-angle coordinates is done by constructing the generating function of a symplectomorphism F we are looking for.

5 The KAM invariant tori theorem: analytic case

5.1 Analytic theorem on persistence of invariant tori

This theorem deals with perturbations of a standard integrable system on a domain $\mathbb{T}^n \times U \subset \mathbb{T}^n_\phi \times \mathbb{R}^n_p$, given by an *analytic* Hamiltonian function $K = K(p)$ depending only on the p-variables:

$$\mathbb{T}^n = \mathbb{R}^n_{\phi_1,\dots,\phi_n}/2\pi\mathbb{Z}^n, \quad p = (p_1, \dots, p_n).$$

We impose the following

non-degeneracy condition: the Hessian matrix $\frac{\partial^2 K}{\partial p^2}$ is non-degenerate for every $p \in U$.

The corresponding Hamiltonian system takes the form

$$\begin{cases} \dot\phi = w(p) \\ \dot p = 0, \end{cases} \qquad w(p) = (w_1(p), \dots, w_n(p)), \quad w_j(p) = \frac{\partial K(p)}{\partial p_j}. \tag{5.1}$$

Remark 5.1 The non-degeneracy condition is equivalent to the condition that the frequency map $p \mapsto w(p)$ is a local diffeomorphism.

Recall that a vector $w = (w_1, \dots, w_n) \in \mathbb{R}^n$ is called *Diophantine,* if there exist constants $C, \gamma > 0$ such that

$$| < w, m > | \geq \frac{C}{|m|^\gamma} \quad \text{for every} \;\; m = (m_1, \dots, m_n) \in \mathbb{Z}^n \setminus \{0\}. \tag{5.2}$$

In this case it is called (C, γ)-Diophantine. If for a given number γ a vector is (C, γ)-Diophantine for appropriate $C > 0$, then it is called γ-Diophantine.

Exercise 5.2 For every $\gamma > n-1$ the set of γ-Diophantine vectors has full measure.

In what follows for every $s > 0$ we denote

$$\Delta_s := \{(r_1, \dots, r_n) \in \mathbb{C}^n \mid \max_j |r_j| < s\},$$

$$\mathbb{T}^n_s := \{\phi = (\phi_1, \dots, \phi_n) \in \mathbb{C}^n / 2\pi\mathbb{Z}^n \mid |\operatorname{Im} \phi_j| < s \text{ for all } j = 1, \dots, n\},$$

For a domain $U \subset \mathbb{R}^n$ we set

$$U_s := \{r \in \mathbb{C}^n \mid |\operatorname{Im} r_j| < s, \text{ and } \operatorname{Re} r \text{ lies in the } s-\text{neighborhood of } U\}.$$

Theorem 5.3 (Kolmogorov – Arnold). *Fix $C, \gamma, s > 0$. Let $K(p)$ be a function analytic on $\overline{U}_s$ and satisfying the non-degeneracy condition on $\overline{U}_s$. Then there exists an $\varepsilon_0 = \varepsilon_0(C, \gamma, K) > 0$ satisfying the following statement. Consider a new, perturbed Hamiltonian function $H(\phi, p)$ analytic on $\mathbb{T}^n_s \times U_s$ and continuous on its closure. Let for some $\varepsilon \in (0, \varepsilon_0)$ one have*

$$|H(\phi, p) - K(p))| \leq \varepsilon \quad \textit{for every } (\phi, p) \in \mathbb{T}^n_s \times U_s. \tag{5.3}$$

Then for every $\zeta \in U$ corresponding to a (C, γ)-Diophantine frequency $w(\zeta)$ the Hamiltonian system with the Hamiltonian function H has an invariant n-dimensional torus

$$\mathcal{T}^n_\zeta = \{p = \zeta + g_\zeta(\phi)\}, \ \ g_\zeta \textit{ is an analytic function on } \mathbb{T}^n, \ \max |g_\zeta| < \delta(K, \varepsilon),$$

$$\delta(K, \varepsilon) \to 0, \quad \textit{as} \ \ \varepsilon \to 0.$$

The Hamiltonian flow on the torus $\mathcal{T}^n_\zeta$ is analytically conjugated to the flow $\dot{\phi} = w(\zeta)$ of the non-perturbed system on the torus $\mathbb{T}^n \times \{\zeta\}$.

5.2 Reduction to KAM theorem for one torus

We deduce Theorem 5.3 from a similar theorem concerning just one Diophantine torus. We prove the latter theorem using a method analogous to proofs of Siegel Theorem and its analogue for circle diffeomorphisms close to rotations.

Consider a Hamiltonian system on $\mathbb{T}^n \times U \subset \mathbb{T}^n \times \mathbb{R}^n$, $0 \in U$, with a Hamiltonian function of type

$$K_0(\phi, r) = c_0 + < \alpha, r > + O(r^2), \quad \text{as} \ \ r \to 0. \tag{5.4}$$

Then the torus

$$\mathbb{T}_0^n := \mathbb{T}^n \times \{0\} \subset \mathbb{T}^n \times \mathbb{R}^n$$

is invariant for the Hamiltonian vector field X_{K_0}, and the corresponding differential equation on $\mathbb{T}_0^n$ takes the form

$$\begin{cases} \dot{\phi} = \alpha \\ \dot{r} = 0. \end{cases} \tag{5.5}$$

In what follows we denote

$$Q(\phi) := \frac{\partial^2 K_0}{\partial r^2}(\phi, 0), \quad Q^0 := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} Q(\phi) d\phi_1 \dots d\phi_n : \tag{5.6}$$

$Q(\phi)$ is the Hessian matrix function, and Q^0 is its average over ϕ, a constant matrix. The latter matrices are symmetric, and we also deal with them as with quadratic forms. Set

$$\mathcal{K}^\alpha := \text{ the set of functions } K_0(\phi, r) \text{ of type (5.4)},$$

$$M_s := \mathbb{T}_s^n \times \Delta_s,$$

$$\mathcal{A}(U) := \{\text{functions holomorphic on } U \text{ and continuous on } \overline{U}\},$$

$$\mathcal{A}_s := \mathcal{A}(M_s), \ \mathcal{K}_s^\alpha := \mathcal{K}^\alpha \cap \mathcal{A}_s.$$

The space $\mathcal{A}_s$ is a Banach space with the norm

$$||f||_s := \max_{\overline{M_s}} |f|.$$

We show that the dynamics on the invariant torus of the perturbed system is conjugated to that of the non-perturbed system by a symplectomorphism of the following group:

$$\Xi := \text{ the group of analytic symplectomorphisms}$$

$$G : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n \times \mathbb{R}^n, \quad G(\phi, r) = (\theta(\phi), (r + \rho(\phi))(\theta'(\phi))^{-1}), \tag{5.7}$$

$$\rho(\phi) d\phi \text{ is a closed 1-form, i.e., } \rho = (\rho_1(\phi), \dots, \rho_n(\phi)), \ \frac{\partial \rho_i}{\partial \phi_j} = \frac{\partial \rho_j}{\partial \phi_i}, \tag{5.8}$$

$$\phi \mapsto \theta(\phi) \text{ is an analytic diffeomorphism } \mathbb{T}^n \to \mathbb{T}^n.$$

Note that closeness condition (5.8) is equivalent to the condition that

$$\rho(\phi) = R + S'(\phi), \quad S \text{ is a function on } \mathbb{T}^n, \ S' = (\frac{\partial S}{\partial \phi_1}, \dots, \frac{\partial S}{\partial \phi_n}). \tag{5.9}$$

Theorem 5.4 (Kolmogorov – Arnold). *For every collection of positive constants $C, \gamma, s, d > 0$ there exist $\varepsilon_0 > 0$ and a family of constants $\delta(\varepsilon) > 0$ depending on $\varepsilon \in (0, \varepsilon_0]$ such that $\delta(\varepsilon) \to 0$, as $\varepsilon \to 0$, that satisfy the following statements. Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ be a (C, γ)-Diophantine vector. Let a function $K_0(\phi, r) \in \mathcal{K}_s^\alpha$ be such that*

$$\max_{\overline{M}_s} |K_0| \le d, \quad ||(Q^0)^{-1}|| \le d. \tag{5.10}$$

Let a function $H(\phi, r) \in \mathcal{A}_s$ be such that

$$||H - K_0||_s \le \varepsilon, \quad \varepsilon \le \varepsilon_0. \tag{5.11}$$

Then there exists an analytic symplectomorphism $G \in \Xi$ such that

$$K_* := H \circ G(\phi, r) \in \mathcal{K}^\alpha, \tag{5.12}$$

$$\max_{\mathbb{T}^n} |G(\phi, 0) - (\phi, 0)| \le \delta(\varepsilon). \tag{5.13}$$

In particular, the Hamiltonian field X_H on $\mathbb{T}^n \times \mathbb{R}^n$ has an invariant torus

$$\mathcal{T}^n = G(\mathbb{T}^n_0), \ \mathcal{T}^n = \{r = g(\phi)\}, \quad g(\phi) \text{ is analytic on } \mathbb{T}^n, \ \max |g| \le \delta(\varepsilon)$$

and the restriction $X_H|_{\mathcal{T}^n}$ is conjugated by G to the standard flow $\dot{\phi} = \alpha$.

Proof of Theorem 5.3. For every $\zeta \in U$ with (C, γ)-Diophantine frequency $w(\zeta)$ let us consider the auxiliary Hamiltonian function

$$K_\zeta(\phi, r) := K(\zeta + r) = K(\zeta) + < \alpha, r > + O(r^2) \in \mathcal{K}_s^\alpha, \quad \alpha = w(r).$$

Its Hessian matrix Q in r is equal to that of the function K. It is nondegenerate and depends only on r. The norm $||Q^{-1}||$ is uniformly bounded on M_s by the same constant d, independent on ζ, that bounds the norm $|| \left(\frac{\partial^2 K}{\partial p^2}\right)^{-1} ||$ on $\overline{U}_s$. We choose the above d greater than $\max_{\overline{U}_s} |K|$, then $||K_\zeta(\phi, r)||_s \le d$ for all $\zeta \in U$. Let now $\varepsilon_0 = \varepsilon_0(C, \gamma, s, d)$ be the same, as in Theorem 5.4, and let $\varepsilon \in (0, \varepsilon_0)$. Then the function $H_\zeta(\phi, r) := H(\phi, \zeta + r)$ is ε-close to K_ζ on M_s. Hence, it satisfies the statements of Theorem 5.4: the Hamiltonian vector field X_{H_ζ} has an invariant torus $\delta(\varepsilon)$-close to the invariant torus $\{r = 0\} = \{p = \zeta\}$ of the nonperturbed system X_{K_ζ}. In the coordinates (ϕ, p) it is an invariant torus of the Hamiltonian vector field X_H. This proves Theorem 5.3 modulo Theorem 5.4. □

Theorem 5.4 will be proved below following a proof due to J.Féjoz [2, 3]. Its proof will be somewhat similar to the proof of Siegel Theorem and to the proof of Kolmogorov-Arnold theorem on analytic circle diffeomorphisms close to rotations. We find the symplectomorphism G as an infinite composition of symplectomorphisms of type (5.7), each obtained as a solution of the KAM analogue of homological equation.

To prove convergence of compositions, we prove an upper bound on solution of KAM homological equation using multidimensional Cauchy bounds. The corresponding background material is presented below. Then we prove a preparatory upper bound on solution of a simplified homological equation. Afterwards we prove an upper bound of solution of the KAM homological equation. Then we prove convergence of compositions to a limit symplectomorphism G. Apriori bounds will imply that G is $\delta(\varepsilon)$-close to the identity.

5.3 Background material 1. Several complex variables. Multidimensional Cauchy formula and Taylor series

For every $R = (R_1, \dots, R_n) \in \mathbb{R}_+^n$ and $a = (a_1, \dots, a_n) \in \mathbb{C}^n$ by $\Delta_R(a)$ we denote the *polydisk centered at a of multiradius R:*

$$\Delta_R(a) := \prod_{j=1}^{n} D_{R_j}(a_j) = \{(z_1, \dots, z_n) \mid |z_j - a_j| < R_j\}.$$

$$\Delta_R := \Delta_R(0), \quad \Delta_s := \Delta_{(s,s,\dots,s)} \text{ for every } s > 0.$$

Recall that a function $f(z_1, \dots, z_n)$ of several complex variables, defined in a domain $U \subset \mathbb{C}^n$, is *holomorphic,* if it is differentiable as a function of $2n$ real variables $\operatorname{Re} z_j$, $\operatorname{Im} z_j$, and its differential at every point z, which is a $\mathbb{R}$-linear functional $\mathbb{R}^{2n} = T_z\mathbb{C}^n \to \mathbb{C}$ is in fact a $\mathbb{C}$-linear functional $\mathbb{C}^n \to \mathbb{C}$.

Theorem 5.5 *(Hartogs). A function $f(z_1, \dots, z_n)$ is holomorphic on a domain $\omega = \omega_1 \times \dots \times \omega_n \subset \mathbb{C}^n$, if and only if it is* **separately holomorphic:** *for every $j = 1, \dots, n$ and every given collection of points $z_s \in \omega_s$, $s \neq j$, the function $g(z) = f(z_1, \dots, z_{j-1}, z, z_{j+1}, \dots, z_n)$ is holomorphic on ω_j.*

Remark 5.6 Separate holomorphicity obviously follows from holomorphicity. The nontrivial part of the theorem says that if a function is separately holomorphic, then it is holomorphic as a function of several variables. We will not prove Theorem 5.5 in full generality. We will prove its weaker version under continuity assumption (Osgood Lemma).

The next theorem generalizes the classical Cauchy formula for holomorphic functions in one variable.

Theorem 5.7 *(Multidimensional Cauchy formula). Let $f : \overline{\Delta}_r \to \mathbb{C}$ be a continuous function that is separately holomorphic on Δ_r. Then for every $z = (z_1, \dots, z_n) \in \Delta_r$ one has*

$$f(z) = \frac{1}{(2\pi i)^n} \oint_{|\zeta_1|=r_1} \cdots \oint_{|\zeta_n|=r_n} \frac{f(\zeta)}{\prod_{j=1}^n (\zeta_j - z_j)} d\zeta_1 \dots d\zeta_n. \tag{5.14}$$

Remark 5.8 Let $g(\zeta)$ denote the sub-integral function in the latter right-hand side. The multiple integral in (5.14) is independent of integration order (Fubini's theorem and continuity of the function $g(\zeta)$). It is equal to the integral of the complex-valued differential n-form $g(\zeta) d\zeta_1 \wedge \dots \wedge d\zeta_n$ on the n-torus $\mathbb{T}^n = \prod_{j=1}^n S_j^1$, $S_j^1 = \{|\zeta_j| = r_j\}$, oriented as a product of positively (i.e., counterclockwise) oriented circles. That is, an orienting basis $v_1, \dots, v_n \in T_\zeta \mathbb{T}^n$ is formed by vectors $v_j \in T_{\zeta_j} S_j^1$ oriented counterclockwise.

Proof It suffices to prove the statement of the theorem in the case, when f is holomorphic in each variable on a domain containing the closed polydisk $\overline{\Delta}_r$: the general case is reduced to it via scaling the function f to $f_\varepsilon(z) = f(\varepsilon z)$, $0 < \varepsilon < 1$ (which is holomorphic in each variable on $\overline{\Delta}_r$) and passing to the limit under the integral, as $\varepsilon \to 1$. We prove formula (5.14) by induction in n.

Induction base: for $n = 1$ this is the classical Cauchy formula for one variable.

Induction step. Let formula (5.14) be proved for the given $n = k$. Let us prove it for $n = k + 1$. For every $w = (w_1, \dots, w_k) \in \mathbb{C}^k$ set

$$f_w(t) = f(w_1, \dots, w_k, t).$$

For every fixed $z_{k+1} \in D_{r_{k+1}}$ the function $g(w_1, \dots, w_k) = f_w(z_{k+1})$ is holomorphic on $\overline{\Delta}_{(r_1,\dots,r_k)}$. Hence,

$$f(z_1, \dots, z_{k+1}) = \frac{1}{(2\pi i)^k} \oint_{|\zeta_1|=r_1} \cdots \oint_{|\zeta_k|=r_k} \frac{f_\zeta(z_{k+1})}{\prod_{j=1}^k (\zeta_j - z_j)} d\zeta_1 \dots d\zeta_k, \tag{5.15}$$

by the induction hypothesis. The function $f_\zeta(t)$ being holomorphic in $t \in \overline{D}_{r_{k+1}}$ for every $\zeta = (\zeta_1, \dots, \zeta_k)$, it is expressed by Cauchy Formula

$$f_\zeta(t) = \frac{1}{2\pi i} \oint_{|\zeta_{k+1}|=r_{k+1}} \frac{f_\zeta(\zeta_{k+1})}{\zeta_{k+1} - t} d\zeta_{k+1} \text{ for every } t \in D_{r_{k+1}}.$$

Substituting the latter formula with $t = z_{k+1}$ to (5.15) yields (5.14), by continuity and Fubini Theorem. □

Lemma 5.9 *(Osgood). Every continuous function on a domain in $\mathbb{C}^n$ that is holomorphic in each individual variable is holomorphic.*

Proof It sufficed to prove the statement of the lemma for a function continuous on a closed polydisk $\overline{\Delta}_r$. Then Multidimensional Cauchy Formula (5.14) holds, and its subintegral expression is a continuous family of rational functions in $z \in \Delta_r$. Therefore, the subintegral expressions are holomorphic on Δ_r. They are uniformly bounded and continuous together with derivatives on compact subsets in Δ_r. Therefore, the integral is C^1-smooth and its partial derivatives are equal to the integrals of partial derivatives in z of the subintegral expression (here one can differentiate the integral by the above boundedness and continuity statements). This imply holomorphicity of the Cauchy integral. □

Theorem 5.10 *Let a sequence of holomorphic functions on a domain $\Omega \subset \mathbb{C}^n$ converge uniformly on compact subsets. Then its limit is holomorphic on Ω.*

Proof The Cauchy formula passes to limit and thus, holds for the limit function. This together with the above argument implies its holomorphicity. □

Set

$$\mathbb{Z}_{\geq 0} = \mathbb{N} \cup \{0\}.$$

Theorem 5.11 *Every function f holomorphic at $0 \in \mathbb{C}^n$ is a sum of power series converging to f uniformly on a neighborhood of 0:*

$$f(z) = \sum_{k \in \mathbb{Z}^n_{\geq 0}} c_k z^k; \ c_k \in \mathbb{C}, \ z^k = z_1^{k_1} \dots z_n^{k_n}, \ c_0 = f(0), \tag{5.16}$$

$$c_k = \frac{1}{(2\pi i)^n} \oint_{|\zeta_n|=\delta} \cdots \oint_{|\zeta_1|=\delta} \frac{f(\zeta)}{\zeta_1^{-k_1-1} \dots \zeta_n^{-k_n-1}} d\zeta_1 \dots d\zeta_n. \tag{5.17}$$

Proof Fix a $\delta > 0$ such that f is holomorphic on the closed polydisk $\overline{\Delta}_\delta = \overline{\Delta}_{(\delta,\dots,\delta)}$. Let us show that the right-hand side of the Cauchy formula written in the same polydisk is a sum of power series converging on Δ_δ. For every ζ_j and z_j with $|z_j| < \delta = |\zeta_j|$ one has

$$\frac{1}{\zeta_j - z_j} = \zeta_j^{-1} \frac{1}{1 - \frac{z_j}{\zeta_j}} = \sum_{l=0}^{+\infty} \zeta_j^{-l-1} z_j^l. \tag{5.18}$$

This series converges absolutely uniformly on every disk $|z_j| \leq \delta'$ with $\delta' < \delta$. Hence, the product of the latter series for all $j = 1, \dots, n$ also absolutely uniformly converges to $\frac{1}{\prod_j (\zeta_j - z_j)}$ on $\Delta_{\delta'}$. Substituting formulas (5.18) for all j to (5.14) together with permutability of integration and series summation (ensured by absolute uniform convergence of subintegral series and uniform boundedness of the function on $\partial\Delta$) yields (5.16) with c_k given by (5.17). Substituting $k = 0$ yields $c_0 = f(0)$, by (5.14). □

5.4 Background material 2. Cauchy bounds

Theorem 5.12 *Let $f(z)$ be a bounded holomorphic function on a domain $U \subset \mathbb{C}^n$, $M := \sup_U |f|$. Let $a \in U$, and let*

$$f(z) = \sum_{k \in \mathbb{Z}^n_{\geq 0}} f_k (z - a)^k$$

be its Taylor series at a. Let $\eta > 0$ be such that $\Delta_\eta(a) \subset U$. Then for every $k = (k_1, \dots, k_n)$ one has

$$|f_k| \leq \frac{M}{\eta^{|k|}}, \quad |k| := \sum_j |k_j|. \tag{5.19}$$

Proof Let us fix an arbitrary $\nu \in (0, \eta)$ and write the multidimensional Cauchy Formula for the coefficient f_k in the closed polydisk $\overline{\Delta_\nu(a)} \subset U$. The module of the subintegral expression is no greater than $\frac{M}{\nu^{|k|+n}}$, which implies that $|f_k| \leq \frac{M}{\nu^{|k|}}$. Since $\nu > 0$ is arbitrary less than η, this yields (5.19). □

Corollary 5.13 *Let $U \subset \mathbb{C}^n$ be a domain, and let $V \subset U$ be a compact subset. Let $\eta > 0$ be such that $\Delta_\eta(a) \subset U$ for every $a \in V$. Then for f and M as above one has*

$$\max_V |\frac{\partial^k f}{\partial z^k}| \leq k_1! \dots k_n! \frac{M}{\eta^{|k|}}. \tag{5.20}$$

Theorem 5.14 *Let f be a bounded holomorphic function on $\mathbb{T}^n_s$. Let us denote $M := \sup_{\mathbb{T}^n_s} |f|$. Let us write f as a Fourier series:*

$$f(\phi) = \sum_{k \in \mathbb{Z}^n} f_k e^{i<k,\phi>}. \tag{5.21}$$

Then

$$|f_k| \leq M e^{-|k|s}. \tag{5.22}$$

Proof One has

$$f_k = \frac{1}{(2\pi)^n} \int \cdots \int f(\phi) e^{-i<k,\phi>} d\phi_1 \dots d\phi_n. \tag{5.23}$$

The subintegral expression is holomorphic and 2π-periodic in each variable ϕ_j lying in the strip $\{|\operatorname{Im}\phi_j| < s\}$, i.e., holomorphic in $z_j = e^{i\phi_j}$ lying in the annulus bounded by circles of radii $e^{\pm s}$. Therefore, the above multiple integral can be replaced by multiple integrals with integration in ϕ_j taken along any counterclockwise oriented circle $\{\operatorname{Im}\phi_j = \nu_j\}$, $|\nu_j| < s$. Fix an arbitrary $\nu \in (0, s)$ and set

$$\nu_j := -\nu \operatorname{sign} k_j \quad \text{for every} \quad j = 1, \dots, n.$$

Then along the product of the above integration circles the module of the subintegral expression is equal to $|f(\phi)|e^{-|k|\nu}$. Therefore, the right-hand side in (5.23) is no greater than $Me^{-|k|\nu}$ for arbitrary positive $\nu < s$. This implies (5.22). □

5.5 Main differential equation and bound of its solution

Here we prove the following proposition, which is one of the key statements in proofs of upper bounds of a solution of homological equation.

In what follows by $\mathcal{A}_s(\mathbb{T}^n)$ we denote the space of functions holomorphic on $\mathbb{T}^n_s$ and continuous on its closure. It is a Banach space with the norm

$$||f||_s := \max_{\mathbb{T}^n_s} |f|.$$

Proposition 5.15 *For every $C, \gamma > 0$ there exists a $\chi = \chi(C, \gamma) > 0$ satisfying the following statement. Let $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ be a (C, γ)-Diophantine vector. Then for every function $u \in \mathcal{A}_s(\mathbb{T}^n)$ with zero average, i.e., zero free term of the Fourier series, there exists a unique analytic function $h(\phi)$ on $\mathbb{T}^n$ with zero average that is a solution of the differential equation*

$$\frac{dh}{d\alpha} = u. \tag{5.24}$$

Moreover, it is holomorphic on $\mathbb{T}^n_s$, and for every $\eta \in (0, s)$ one has

$$||h||_{s-\eta} \leq \chi \eta^{-(\gamma+1)} ||u||_s. \tag{5.25}$$

Proof Let $u(\phi) = \sum_{k\neq 0} u_k e^{i<k,\phi>}$. The formal Fourier series solution of (5.24) is

$$h(\phi) = \sum_{k\neq 0} h_k e^{i<k,\phi>}, \quad h_k = \frac{u_k}{i<k,\alpha>},$$

$$|h_k| \leq C^{-1}|k|^\gamma |u_k| \leq C^{-1}|k|^\gamma ||u||_s e^{-|k|s},$$

by Diophantine property and (5.22). Therefore, for every $\phi(mod 2\pi\mathbb{Z}^n) \in \mathbb{T}^n_{s-\eta}$ one has

$$\sum_{k\neq 0} |h_k e^{i<k,\phi>}| \leq C^{-1}||u||_s \sum_{k\in\mathbb{Z}^n\setminus\{0\}} |k|^\gamma e^{-|k|s} e^{|k|(s-\eta)}$$

$$= C^{-1}\eta^{-(\gamma+n)} \sum_{k\in\mathbb{Z}^n\setminus\{0\}} |k\eta|^\gamma e^{-|k|\eta}\eta^n.$$

The latter sum is an integral sum with step η for the function $|x|^\gamma e^{-|x|}$, $x \in \mathbb{R}^n$, whose integral converges. Therefore, the sums are uniformly bounded, as is η. This implies holomorphicity of the function h on $\mathbb{T}^n_s$ and (5.25). □

5.6 The Main Lemma

In what follows we denote

$\Xi_s :=$ {symplectomorphisms (5.7) analytic on M_s and continuous on $\overline{M}_s$}.

The distance between two symplectomorphisms $G_1, G_2 \in \Xi_s$ is defined as the norm $||G_1 - G_2||_s$, which is the sum of the s-norms of the $2n$ components of the difference $G_1(\phi, r) - G_2(\phi, r)$.

Lemma 5.16 (Main Lemma). *For every $C, \gamma > 0$ and $d > 4$ there exists a*

$$c = c(C, \gamma, d) > 4d + 4 \tag{5.26}$$

satisfying the following statements. Let

$$s_0 \in (0, \frac{1}{2}), \quad 0 < \eta_0 \leq \frac{s_0}{4}, \quad 0 < \delta_0 \leq \eta_0^{4c}, \quad s_1 := s_0 - \eta_0 < s_{10} := s_0 - \frac{\eta_0}{4}.$$

Let $\alpha \in \mathbb{R}^n$ be a (C,γ)-Diophantine vector. Let $K_0 \in \mathcal{K}^\alpha_{s_0}$, and let $Q(\phi)$ be its Hessian matrix in r at $r = 0$, see (5.6), Q^0 be its average over $\mathbb{T}^n$. Let

$$||K_0||_{s_0}, \ ||(Q^0)^{-1}|| \leq d, \tag{5.27}$$

$$H_0 \in \mathcal{A}_{s_0}, \quad ||H_0 - K_0||_{s_0} \le \delta_0. \tag{5.28}$$

Then there exist a $K_1 \in \mathcal{K}^{\alpha}_{s_1}$ and a symplectomorphism

$$G_0 \in \Xi_{s_{10}}, \; ||G_0 - Id||_{s_{10}} \le \delta_0 \eta_0^{-c}, \tag{5.29}$$

such that the function $H_1 := H_0 \circ G_0$ lies in $\mathcal{A}_{s_1}$ and

$$||H_1 - K_1||_{s_1} \le \delta_1 := \delta_0^2 \eta_0^{-c}. \tag{5.30}$$

One has

$$\delta_1 < \eta_1^{4c}, \quad \eta_1 := \frac{\eta_0}{2}. \tag{5.31}$$

Proof of Theorem 5.4 modulo the Main Lemma. Let K_0, d be the same, as in Theorem 5.4. Let $D = 4d$, $c(C, \gamma, D)$ be the same, as in the Main Lemma. Set $H_0 = H$. Let us construct s_1, δ_1, η_1, s_2, δ_2, η_2, K_1, G_0, H_1, K_2, G_2, etc. successively applying the Main Lemma. The sequence s_j decreases and converges to a number $s_* \ge \frac{s_0}{2}$. One has

$$H_j = H_0 \widehat{G}_{j-1}, \quad \widehat{G}_j := G_0 \circ \dots \circ G_j. \tag{5.32}$$

Let us show that the above compositions are well-defined,

$$H_j, K_j \text{ converge in } \mathcal{A}_{s_*} \text{ to the same function } K_*, \tag{5.33}$$

$$\widehat{G}_j \text{ converge in the } \frac{s_*}{2}\text{−norm to an analytic symplectomorphism } G_* \in \Xi_{\frac{s_*}{2}}. \tag{5.34}$$

The limit function K_* lies in $\mathcal{K}^{\alpha}_{s_*}$, as do K_j. Then $H \circ G_* = K_*$, by (5.30) applied to H_j, K_j, since $\delta_j \to 0$. This will prove Theorem 5.4.

We have not only to prove the above statements, but also to show that the Main Lemma indeed can be applied infinitely many times. To this end, we use the following claim for each application of the Main Lemma.

Claim 6. *The map G_j satisfying the statement of the Main Lemma is injective on $\overline{M}_{s_{j+1}}$ and*

$$G_j(\overline{M}_{s_{j+1}}) \Subset M_{s_j - \frac{\eta_j}{2}} \Subset M_{s_j}. \tag{5.35}$$

Proof Inclusion (5.35) follows from (5.29) applied to G_j, since its right-hand side is

$$\delta_j \eta_j^{-c} \le \eta_j^{3c} < \frac{\eta_j}{2} = \frac{s_j - s_{j+1}}{2}.$$

Recall that upper bound (5.29) holds on $M_{s_{j+1}+\frac{3\eta_j}{4}}$. On the set $M_{s_{j+1}}$ one has

$$||dG_j - Id|| < \delta_j \eta_j^{-c} \left(\frac{3\eta_j}{4}\right)^{-1} < \eta_j^{2c} < \frac{1}{16}, \tag{5.36}$$

by (5.29) and Cauchy bound. This together with Newton–Leibniz Integral Formula for the difference $G_j(x) - G_j(y)$ implies injectivity on $M_{s_{j+1}}$, as in the proof of Claim 2 in the proof of Siegel Theorem in Subsection 1.5. Claim 6 is proved. □

Claim 7. *The functions K_j constructed on each step and their averaged Hessian matrices Q_j^0 satisfy the inequalities*

$$||K_{j-1} - K_j||_{s_j} \leq (4d+3)\delta_{j-1}, \tag{5.37}$$

$$||Q_j^0 - Q_{j-1}^0|| \leq (4d+3)\delta_{j-1}s_{j-1}^{-2}, \tag{5.38}$$

$$||(Q_j^0)^{-1} - (Q_{j-1}^0)^{-1}|| \leq 2^8 d^3 \delta_{j-1} s_{j-1}^{-2}. \tag{5.39}$$

Proof One has $H_j = H_{j-1} \circ G_{j-1}$. Therefore,

$$\begin{aligned}||K_{j-1} - K_j||_{s_j} &\leq ||H_j - H_{j-1}||_{s_j} + ||H_j - K_j||_{s_j} + ||H_{j-1} - K_{j-1}||_{s_j} \\ &\leq ||H_{j-1} \circ G_{j-1} - H_{j-1}||_{s_j} + \delta_j + \delta_{j-1}.\end{aligned} \tag{5.40}$$

One has

$$||H_{j-1} \circ G_{j-1} - H_{j-1}||_{s_j} \leq \delta_j \eta_j^{-1} \max_{V_j} ||H'_{j-1}||, \quad V_j = M_{s_j + \delta_j \eta_j^{-1}}, \tag{5.41}$$

by Lagrange Increment Theorem and (5.29) applied to G_{j-1},

$$\max_{V_j} ||H'_{j-1}|| \leq ||H_{j-1}||_{s_{j-1}} (\frac{\eta_{j-1}}{2})^{-1} \leq (4d + \delta_{j-1})\eta_{j-1}^{-c-2}, \tag{5.42}$$

by Cauchy bound and the inequality

$$||H_{j-1}||_{s_{j-1}} \leq 4d + \delta_{j-1}. \tag{5.43}$$

Indeed, the Main Lemma with $c = c(C, \gamma, 4d)$ being applicable for K_{j-1}, H_{j-1}, this means that

$$||K_{j-1}||_{s_{j-1}} \leq 4d, \quad ||(Q_{j-1}^0)^{-1}|| \leq 4d, \tag{5.44}$$

and hence, (5.43) holds by (5.30). Substituting (5.42) to (5.41) and then substituting everything to (5.40) by elementary inequalities yields

$$||K_{j-1} - K_j||_{s_j} \leq \delta_j + \delta_{j-1} + (4d + \delta_{j-1})\eta_{j-1}^{-c-3}\delta_j \leq (4d+3)\delta_{j-1}.$$

Inequality (5.37) is proved. Inequality (5.38) then follows by Cauchy bound. To prove (5.39), let us note that $(||(Q_j^0)^{-1}||)^{-1}$ is the minimum $m_j := \min_{|v|=1} ||Q_j^0 v||$. One has $m_{j-1} \geq (4d)^{-1}$, by (5.44),

$$|m_j - m_{j-1}| \leq (4d+3)\delta_{j-1} s_{j-1}^{-2}, \tag{5.45}$$

by (5.38). Therefore,

$$\big| ||(Q_j^0)^{-1}|| - ||(Q_{j-1}^0)^{-1}|| \big| = |m_j^{-1} - m_{j-1}^{-1}|$$

$$= \frac{|m_j - m_{j-1}|}{m_j m_{j-1}} \leq \frac{(4d+3)\delta_{j-1} s_{j-1}^{-1}}{(4d)^{-1}((4d)^{-1} - (4d+3)\delta_{j-1} s_{j-1}^{-2})} < 2^8 d^3 \delta_{j-1} s_{j-1}^{-1},$$

since $\delta_{j-1} \leq s_{j-1}^{4c}$, $s_{j-1} < \frac{1}{2}$, $c > 4d+4$, see (5.26), by elementary inequalities. □

Claim 7 implies that on each step of the Main Lemma the bounds (5.27) for the norms of the functions K_j and their averaged Hessian matrices with inverses remain valid with d replaced by $4d$, since the sum of right-hand sides of inequalities (5.37), (5.39) though all j is less than 4, since $\delta_j \leq \eta_j^{4c}$, $c > 4d+4$, see (5.26), by elementary inequalities. Hence, the Main Lemma is applicable infinitely many times and yields a sequence of well-defined functions K_j, H_j and maps G_j.

Claim 6 implies that each composition $\widehat{G}_j$ is a well-defined injective holomorphic map $\overline{M}_{s_{j+1}} \to M_{s_0}$, and H_0 is holomorphic on M_{s_0}. Therefore, the compositions $H \circ \widehat{G}_j$ are well-defined and holomorphic on $\overline{M}_{s_{j+1}}$. Afterwards uniform convergence of the maps $\widehat{G}_j$ on $\overline{M_{s_*}}$ to a limit symplectomorphism G satisfying $H \circ G = K_* := \lim_{j\to\infty} K_j \in \mathcal{K}^\alpha$ follows as in the proofs of Siegel Theorem and Kolmogorov–Arnold Theorem on circle diffeomorphisms. □

5.7 Proof of the Main Lemma

The Lie algebra ξ of the group Ξ is the space of vector fields on $\mathbb{T}^n \times \mathbb{R}^n$ of the type

$$(\dot\theta(\phi),\ \dot\rho(\phi) - r\dot\theta'(\phi)), \quad \dot\rho = \dot R + \dot S'(\phi), \quad \dot R = const. \tag{5.46}$$

Here $\dot\theta(\phi)$ is a vector field on $\mathbb{T}^n$, $S(\phi)$ is a function on $\mathbb{T}^n$. The above Lie algebra description follows by differentiating a one-parametric symplectomorphism family

$$(\theta_t(\phi), (r + \rho_t(\phi))(\theta_t'(\phi))^{-1}), \quad \theta_0(\phi) = \phi, \quad \rho_0(\phi) = 0,$$

in t at $t = 0$, denoting

$$\dot{\theta} = \frac{d\theta_t(\phi)}{dt}|_{t=0}, \quad \dot{\rho} = \frac{d\rho_t(\phi)}{dt}|_{t=0},$$

and taking into account that $(\theta'_t(\phi))^{-1} = 1 - t\dot{\theta}'(\phi) + o(t)$, as $t \to 0$. In these notations the derivative of the above family is given by (5.46). The representation of the vector function ρ given by (5.46) follows from (5.9). Thus the Lie algebra ξ is identified, as a vector space, with the space

$$(\dot{\theta}(\phi), \dot{R}, \dot{S}(\phi)) \mid \dot{\theta}(\phi) \text{ is a vector field on } \mathbb{T}^n, \ \dot{S}(\phi) \ \text{ is a function on } \mathbb{T}^n.$$

For a vector field v by $\exp(v) = g_v^1$ we denote its unit time flow map. In what follows we construct appropriate vector field

$$v = (\dot{\theta}, \dot{R}, \dot{S}) \in \xi$$

and then prove the lemma for $G_0 = \exp(v)$.

Every function $H(\phi, r)$ on $\mathbb{T}^n \times \mathbb{R}^n$ can be written as

$$H(\phi, r) = H^0 + H_{0,1}(\phi, r) + O(r^2), \quad H^0 = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} H(\phi, 0) d\phi_1 \dots d\phi_n,$$

$$H_{0,1}(\phi, r) = h_0(\phi) + < h_1(\phi), r >, \qquad \int_{\mathbb{T}^n} h_0(\phi) d\phi_1 \dots d\phi_n = 0.$$

Remark 5.17 One has $H \in \mathcal{K}^\alpha$, if and only if $H_{0,1}(\phi, r) =< \alpha, r >$. For every H we set

$$\pi_\alpha(H)(\phi, r) = H(\phi, r) - H_{0,1}(\phi, r) + < \alpha, r > . \tag{5.47}$$

The function $\pi_\alpha(H)$ lies in $\mathcal{K}^\alpha$, by construction.

In what follows we denote $(H_0)_{0,1}$ by $H_{0,1}$ for simplicity.

First let us motivate the construction of the above field v. In order to conjugate the perturbed Hamiltonian function H_0 to a function of class $\mathcal{K}^\alpha$, we have to construct a conjugating symplectomorphism that kills the terms $H_{0,1}(\phi, r) - < \alpha, r >$. *On heuristic level,* we treate the small differences $H_0 - K_0$, $H_{0,1}(\phi, r) - < \alpha, r >$ and the vector field v as *"quantities of first order"*. We first construct a field v so that the conjugacy by $\exp(v)$ kills the difference $H_{0,1}(\phi, r) - < \alpha, r >$ on the first order level. The field v will be the unique solution of homological equation introduced below. We deduce an a priori upper bound on its solution directly from results of Subsection 5.5 and then prove that $G = \exp(v)$ satisfies (5.29). Afterwards we show that the error term has a quadratic bound and prove (5.30).

5.7.1 Construction of v. Homological equation. Proof of (5.29)

In the subsection by h.o.t. we mean "terms of order greater than one" in the above, heuristic sense. Let

$$H_{0,1}(\phi,r) - <\alpha,r> = h_0(\phi) + <b(\phi),r>, \quad b(\phi) = h_1(\phi) - \alpha.$$

Up to terms of order greater than one, we have

$$\exp(v)(\phi,r) = (\phi + \dot{\theta}(\phi), r(Id - \dot{\theta}'(\phi)) + \dot{\rho}(\phi)),$$

$$H_0 \circ \exp(v) = H_0 + \frac{dH_0}{dv} + h.o.t = H_0 + \frac{dK_0}{dv} + h.o.t.$$

$$= \pi_\alpha(H_0 + \frac{dK_0}{dv}) + h_0(\phi) + <b(\phi),r> + (\frac{dK_0}{dv})_{0,1} + h.o.t.$$

Proposition 5.18 *1) Let α be a (C,γ)-Diophantine vector,*

$$K_0(\phi,r) = c_0 + <\alpha,r> + \frac{1}{2} <Q(\phi)r,r> + O(r^2) \in \mathcal{K}^\alpha.$$

Let the averaged Hessian matrix $Q^0 = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} Q(\phi) d\phi_1,\dots,d\phi_n$ be non-degenerate. Then for every pair of analytic functions h_0, b on $\mathbb{T}^n$, h_0 being with zero average, the **homological equation**

$$h_0(\phi) + <b(\phi),r> + (\frac{dK_0}{dv})_{0,1} = 0 \qquad (5.48)$$

has a solution $v = (\dot{\theta}(\phi), \dot{\rho}(\phi))$ analytic on a neighborhood of the torus $\mathbb{T}^n$ in $\mathbb{C}^n/2\pi\mathbb{Z}^n$, unique up to adding a constant vector to $\dot{\theta}$: it is unique after its normalization with $\dot{\theta}$ having zero average.

2) Let in addition $K_0 \in \mathcal{A}_{s_0}$ and $h_0(\phi)$, $b(\phi)$ be analytic on $\mathbb{T}^n_{s_0}$ and continuous on its closure. Let $d > 0$ be such that

$$||K_0||_{s_0} \leq d, \quad ||(Q^0)^{-1}|| \leq d. \qquad (5.49)$$

Then for every $\eta \in (0, s_0)$ the above solution v admits the upper bound

$$||v||_{s_0-\eta} \leq \sigma(C,\gamma,d)(||h_0||_{s_0} + ||b||_{s_0})\eta^{-4(\gamma+1)}. \qquad (5.50)$$

Here $\sigma(C,\gamma,d) > 0$ is a constant depending only on (C,γ,d).

Proof Recall that in the coordinates on $\mathbb{T}^n \times \mathbb{R}^n$ the field v takes the form $(\dot{\theta}(\phi), \dot{R} + \dot{S}'(\phi) - r\dot{\theta}'(\phi))$. The (0,1)-part of the derivative $\frac{dK_0}{dv}$ is equal to

$$\left(\frac{dK_0}{dv}\right)_{0,1}(\phi, r) = \Big(< \alpha + Q(\phi)r, \dot{R} + \dot{S}'(\phi) - r\dot{\theta}'(\phi) >\Big)_{0,1}$$

$$= \frac{d\dot{S}}{d\alpha}(\phi) + < r, -\frac{d\dot{\theta}}{d\alpha}(\phi) + Q(\phi)(\dot{R} + \dot{S}'(\phi)) >, \tag{5.51}$$

since $\dot{R} = const$ and the derivative $\frac{d\dot{S}(\phi)}{d\alpha}$ has zero average, as does any partial derivative of a function on $\mathbb{T}^n$. The second vector function in the scalar product in (5.51) is the sum of its average, equal to $Q^0\dot{R} + (Q\dot{S}')^0$, and the following expression with zero average:

$$-\frac{\dot{\theta}(\phi)}{d\alpha} + Q(\phi)(\dot{R} + \dot{S}'(\phi)) - Q^0\dot{R} - (Q\dot{S}')^0.$$

Therefore, (5.48) is equivalent to the system of equations

$$\begin{cases} \frac{d\dot{S}}{d\alpha}(\phi) = -h_0(\phi) \\ Q^0\dot{R} = -b^0 - (Q\dot{S}')^0 \\ \frac{d\dot{\theta}}{d\alpha}(\phi) = b_1(\phi) + Q(\phi)(\dot{R} + \dot{S}'(\phi)) - Q^0\dot{R} - (Q\dot{S}')^0, \end{cases} \tag{5.52}$$

$$b^0 = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} b(\phi)d\phi_1 \dots d\phi_n, \quad b_1(\phi) = b(\phi) - b^0.$$

This system is clearly triangular and has a unique solution $(\dot{\theta}, \dot{R}, \dot{S})$ with $\dot{\theta}$ being of zero average: we find S from the first equation as in Subsection 5.5, then we set

$$\dot{R} = -((Q^0)^{-1}(b^0 + (Q\dot{S}')^0),$$

and find $\dot{\theta}$ from the third equation as in Subsection 5.5. The upper bound of $(\dot{\theta}, \dot{R}, \dot{S})$ given by (5.50) follows from the upper bound (5.25) of solution of equation $\frac{dh}{d\alpha} = u$, inequalities (5.49), the upper bound on $Q(\phi)$ given by Cauchy bounds and the elementary inequalities

$$|b^0| \le ||b||_{s_0}, \quad ||b_1||_{s_0} = ||b - b^0||_{s_0} \le ||b||_{s_0} + |b^0| \le 2||b||_{s_0}.$$

□

Corollary 5.19 *In the conditions of the Main Lemma one has*

$$||v||_{s_0 - \eta} \le 3\sigma(C, \gamma, d)\delta_0\eta^{-5(\gamma+1)}. \tag{5.53}$$

Proof Inequality (5.53) follows from (5.50) and the inequalities

$$||h_0||_{s_0} \le ||H_0 - K_0 - (H_0 - K_0)^0)||_{s_0}$$

$$\le ||H_0 - K_0||_{s_0} + |(H_0 - K_0)^0| \le 2||H_0 - K_0||_{s_0} \le 2\delta_0,$$

$$||b||_{s_0} \le ||H_0 - K_0||_{s_0} s_0^{-1} \le \delta_0 \eta^{-1}.$$

The latter inequality follows from (5.28) and Cauchy bounds. □

Proof of (5.29). Applying (5.53) to $\eta = \frac{\eta_0}{8}$ yields

$$||v||_{s_0 - \frac{\eta_0}{8}} \le 3\sigma(C, \gamma, d)\delta_0 \eta_0^{-20(\gamma+1)} \le \delta_0 \eta_0^{-c_1}, \ c_1 = c_1(C, \gamma, d) > 4. \quad (5.54)$$

We choose $\delta_0 < \eta_0^{4c_1}$. Then

$$||v||_{s_0 - \frac{\eta_0}{8}} \le \delta_0 \eta_0^{-c_1} \le \eta_0^{3c_1} < \frac{\eta_0}{8}. \quad (5.55)$$

Thus if the initial conditions for the differential equation given by v lie in $\overline{M}_{s_0 - \frac{\eta_0}{4}}$, then in times $t \in [0, 1]$ the corresponding solution will move with velocity less than $\frac{\eta_0}{8}$, and hence, will not escape from $M_{s_0 - \frac{\eta_0}{8}}$. Therefore, the symplectomorphisms

$$G^t = g_v^t, \ t \in [0, 1], \quad G_0 = G^1 = \exp(v)$$

satisfy the inequality

$$||G^t - Id||_{s_0 - \frac{\eta_0}{4}} \le ||v||_{s_0 - \frac{\eta_0}{8}} \le \delta_0 \eta_0^{-c_1}. \quad (5.56)$$

This proves (5.29). □

5.7.2 Quadratic error bounds. End of proof of the Main Lemma

Here we prove (5.30) for the function

$$K_1 := \pi_\alpha H_1 \in \mathcal{K}^\alpha, \quad H_1 = H_0 \circ G_0 = H_0 \circ \exp(v).$$

To do this, we use the second Taylor remainder integral formula for functions in one variable:

$$f(\tau) - f(0) = \tau f'(0) + \int_0^\tau (\tau - t) f''(t) dt. \quad (5.57)$$

One has

$$H_1 - K_1 = (H_0 \circ G^1 - K_1)_{0,1} = (H_0 \circ G^1 - H_0)_{0,1} + (H_0 - K_1)_{0,1}$$

$$= (H_0 \circ G^1 - H_0)_{0,1} + h_0(\phi) + < b(\phi), r >, \tag{5.58}$$

$$H_0 \circ G^1 - H_0 = \frac{dH_0}{dv} + \int_0^1 (1-t) \frac{d^2(H_0 \circ G^t(\phi, r))}{dt^2} dt,$$

$$\frac{d^2(H_0 \circ G^t(\phi, r))}{dt^2} = \frac{d}{dt}\left((dH_0(G^t(\phi, r)))v(G^t(\phi, r))\right)$$

$$=< (\partial^2 H_0(G^t(\phi, r)))v(G^t(\phi, r)), v(G^t(\phi, r)) >$$

$$+(dH_0(G^t(\phi, r)))dv(G^t(\phi, r))v(G^t(\phi, r)).$$

Note that $G^t(\phi, r) \in M_{s_0 - \frac{\eta_0}{4}}$ for every $(\phi, r) \in \overline{M}_{s_1 + \frac{\eta_0}{2}}$, by (5.55) and since $s_1 + \frac{\eta_0}{2} = s_0 - \frac{\eta_0}{2}$. Thus, the composition $H_0 \circ G^t(\phi, r)$ is well-defined and lies in $\mathcal{A}_{s_1 + \frac{\eta_0}{2}}$ for every $t \in [0, 1]$. This together with (5.54), (5.56) and Cauchy bounds implies that

$$||H_0 \circ G_0 - H_0 - \frac{dH_0}{dv}||_{s_1 + \frac{\eta_0}{4}} \leq \delta_0^2 \eta_0^{-c_2}, \quad c_2 = c_2(C, \gamma, d). \tag{5.59}$$

Hence, the same inequality with maybe a bigger constant $c_2' = c_2'(C, \gamma, d)$ holds also for the norm $||(H_0 \circ G_0 - H_0 - \frac{dH_0}{dv})_{0,1}||_{s_1}$, by Cauchy bounds and elementary inequalities, as in the proof of Corollary 5.19. Substituting the latter inequality to (5.58) together with triangle inequality yields

$$H_1 - K_1 = \left(\frac{dH_0}{dv}\right)_{0,1} + h_0(\phi) + < b(\phi), r > + \mathcal{R}_1(\phi, r), \ \ ||\mathcal{R}_1(r, \phi)||_{s_1} \leq \delta_0^2 \eta_0^{-c_2'}. \tag{5.60}$$

On the other hand, the sum of the first three terms in the latter right-hand side is equal to

$$\left(\frac{dK_0}{dv}\right)_{0,1} + h_0(\phi) + < b(\phi), r > + \mathcal{R}_2(\phi, r), \ \ \mathcal{R}_2(\phi, r) = \left(\frac{d(H_0 - K_0)}{dv}\right)_{0,1}. \tag{5.61}$$

The sum of the first three terms in (5.61) vanishes, since v is a solution of the homological equation, by construction. One has

$$||\mathcal{R}_2(r, \phi)||_{s_1} \leq \delta_0^2 \eta_0^{-c_3}, \quad c_3 = c_3(C, \gamma, d) > c_2', \tag{5.62}$$

by (5.28), (5.54) and Cauchy bounds. Finally,

$$H_1 - K_1 = \mathcal{R}_1 + \mathcal{R}_2,$$

where $\mathcal{R}_1$, $\mathcal{R}_2$ satisfy (5.60) and (5.62) respectively. This implies (5.30) with appropriate constant $c = c(C, \gamma, d) > 0$. Inequality (5.31) was already deduced from the inequality $\delta_0 \leq \eta_0^{4c}$ in Subsection 1.3. The Main Lemma is proved.

References

[1] Arnold, V.I. *Geometric methods in the theory of ordinary differential equations.* Second edition. Grundlehren der mathematischen Wissenschaften **250**, Springer, 2012. [Russian version of previous edition: Dopolnitelnyie glavy teorii obyknovennyh differentsialnyh uravnenii, Moskva, Nauka, 1978.]

[2] Féjoz, J. *A proof of invariant torus theorem of Kolmogorov.* Preprint.

[3] Féjoz, J. *Introduction to KAM theory.* A short course at Università degli Studi di Milano–Bicocca.

[4] Herman, M. R. *Sur les courbes invariantes par les difféomorphismes de l'anneau.* Astérisque, **144** (1986), 1–248.

[5] Yoccoz, J.-C. *Linéarisation des germes de difféomorphismes holomorphes de* $(\mathbb{C}, 0)$. C. R. Acad. Sci. Paris, **306** (1988), Série I, 55–58.

[6] Yoccoz, J.-C. *Théorème de Siegel, nombres de Bruno et polynômes quadratiques.* Astérisque, **231** (1995), 1–88.