Chapter 8

Technical details

We state and prove the properties of the weighted norm $[\cdot]_{r,s,w}$ in Section 8.1, and those of the functions $\Delta$, $\Gamma$ and $\Psi$ in Section 8.2. In the third section we collect a number of useful lemmas. In the final section we calculate weighted norms of several perturbation terms corresponding to a row of weakly coupled pendulums.

8.1 Properties of the weighted norm

In this section the properties of the norm defined in Section 4.2 are studied. This norm is an adaptation of the weighted norm of [P1] such that it is invariant under unitary transformations. The other properties of the weighted norm of [P1] are inherited without much change.

We start with checking that the weighted norm does indeed satisfy the norm properties. Then we compare the weighted norm with the supremum norm, and show that it is invariant under unitary coordinate transformations. Next we show that the space of functions with finite weighted norm is complete. For the submultiplicative property we can follow [P1] quite literally, but in the Cauchy inequalities an extra factor depending on $m$ (the number of normal degrees of freedom) turns up. In the estimates involving the Poisson bracket we assume a certain parameter to be sufficiently small such that the extra dependence on $m$ disappears. The estimate of coordinate transformations close to the identity then is similar to the one in [P1].

8.1.1 Checking the norm properties

Lemma 8.1 (The weighted norm is a norm indeed) For real-analytic functions $F, G$ on $\mathcal{D}_{r,s} \times W_w$ for which $[F]_{r,s,w}$ and $[G]_{r,s,w}$ are finite, and for $\mu \in \mathbb{C}$ we have

1. $[F]_{r,s,w} \geq 0$, where equality holds for and only for $F \equiv 0$. (positivity)
2. \([\mu F]_{r,s,w} = |\mu| [F]_{r,s,w}\)  
   \(\text{(homogeneity)}\)

3. \([F + G]_{r,s,w} \leq [F]_{r,s,w} + [G]_{r,s,w}\)  
   \(\text{(triangle inequality)}\)

**Proof** The proof follows from well-known properties of Fourier and Taylor coefficients:

1. From Definition 4.3 we see that the weighted norm is nonnegative, if it exists, and strictly positive if at least one of the Fourier-Taylor coefficients \(F_{ab}^k\) is nonzero for some value of \((I; \omega)\). Only for the null function all Fourier-Taylor coefficients are identical to zero.

2. Both Fourier and Taylor coefficients are homogeneous in the function they represent.

3. For the triangle inequality, look back at equation (4.15):

\[
[(F + G)_{k}^{(N,M)}] = \left( \sum_{a \in \{1,\ldots,m\}^N, b \in \{1,\ldots,m\}^M} |(F + G)_{ab}^k|^2 \right)^{1/2}
\leq \left( \sum_{a,b} |F_{ab}^k|^2 \right)^{1/2} + \left( \sum_{a,b} |G_{ab}^k|^2 \right)^{1/2}
= [F_{k}^{(N,M)}] + [G_{k}^{(N,M)}],
\]

etcetera.

\[\square\]

### 8.1.2 Comparisons with the supremum norm

On the one hand the weighted norm \([\cdot]_{r,s,w}\) majorizes the supremum norm \(|\cdot|_{r,s,w}\) on \(D_{r,s} \times W_w\). On the other hand it is smaller than a constant times the supremum norm on a sufficiently larger region.

**Lemma 8.2 (The supremum norm is majorized by the weighted norm)** Suppose \(F\) is analytic on \(D_{r,s} \times W_w\). If \([F]_{r,s,w}\) is finite, it bounds \(|F|\) on \(D_{r,s} \times W_w\):

\[
[F]_{r,s,w} := \sup_{D_{r,s} \times W_w} |F| \leq [F]_{r,s,w}.
\]

**Proof** We first prove that

\[
\sup_{|z| \leq s} \left| \sum_{a \in \{1,\ldots,m\}^N, b \in \{1,\ldots,m\}^M} F_{ab}^k z_a \xi_b \right| \leq [F_{k}^{(N,M)}]_{s^{N+M}} \quad (8.1)
\]
pointwise in $((I; \omega)/)$. The following version of the Cauchy-Schwarz inequality

$$\left| \sum_{i=1}^{m} u_i v_i \right|^2 \leq \sum_{i=1}^{m} |u_i|^2 \sum_{j=1}^{m} |v_j|^2, \ u, v \in \mathbb{C}^m.$$ 

is repeatedly used for proving this:

$$\sup_{z,\xi} \left| \sum_{a,b} F_{ab}^k z_a \xi_b \right|^2 = \sup_{z,\xi} \left( \sum_{a_1=1}^{m} \sum_{b_1=1}^{m} \sum_{a_{N-1}=1}^{m} \sum_{b_{M-1}=1}^{m} F_{ab}^k z_{a_1} \ldots z_{a_{N-1}} \xi_{b_1} \ldots \xi_{b_{M-1}} \right)^2 \leq s^2 \sum_{a_1=1}^{m} \sup_{z,\xi} \left| \sum_{b_1=1}^{m} \sum_{a_{N-1}=1}^{m} \sum_{b_{M-1}=1}^{m} F_{ab}^k z_{a_1} \ldots z_{a_{N-1}} \xi_{b_1} \ldots \xi_{b_{M-1}} \right|^2 \leq \ldots \leq s^{2(N+M)} \sum_{a,b} |F_{ab}^k|^2 = \left[ F_{k}^{(N,M)} \right]_{2,2(N+M)}^2.$$ 

Together with an obvious estimate for the supremum norm, inequality (8.1) finishes the proof:

$$|F|_{r,s,w} \leq \sum_{k} e^{k\rho r} \sup_{I,\omega} \sum_{N,M} \sup_{z,\xi} \left| \sum_{a \in \{1, \ldots, m\}^N} \sum_{b \in \{1, \ldots, m\}^M} F_{ab}^k z_a \xi_b \right| \leq \sum_{k} e^{k\rho r} \sup_{I,\omega} \sum_{N,M} \left[ F_{k}^{(N,M)} \right]_{2,N+M} \leq \left[ F \right]_{r,s,w}.$$ 

Conversely, we have the following bound:

**Lemma 8.3 (The weighted norm is bounded by a constant times the supremum on a sufficiently larger region)** Suppose that $s < S/\sqrt{m}$ and $r < R$. Then

$$[F]_{r,s,w} \leq \bullet |F|_{r,s,w}$$

where $\bullet$ indicates a factor depending only on $n, m, R-r, S/s$. 


Proof We start with proving Lemma 8.3 for functions $F$ depending only on $z \in \mathbb{C}^m$ and analytic on the ball $\|z\| < S$. Dropping unnecessary sub-and superscripts the definition of the weighted norm reduces to

$$[F]_s = \sum_{N=0}^{\infty} [F^{(N)}] s^N,$$

where $[F^{(N)}] = \left( \sum_{a \in \{1, \ldots, m\}^N} |F_a|^2 \right)^{1/2}$.

We also use multi-index coefficients in which $F$ reads

$$F(z) = \sum_{p \in \mathbb{N}_0^m} F_p z^p = \sum_{p \in \mathbb{N}_0^m} F_{p_1 \ldots p_m} z_1^{p_1} \ldots z_m^{p_m},$$

where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$.

By repeated application of the Cauchy inequality the multi-index coefficients are bounded from above by

$$|F_p| \leq \frac{\sup_{|z| < S/\sqrt{m}} |F(z)|}{(S/\sqrt{m}) |p|} \leq \frac{|F|_s}{(S/\sqrt{m}) |p|},$$

where $[F]_s$ is an abbreviation of $\sup_{|z| < S} |F(z)|$. Now we crudely estimate $[F^{(N)}]$:

$$[F^{(N)}] = \left( \sum_{a \in \{1, \ldots, m\}^N} |F_a|^2 \right)^{1/2} \leq \left( \sum_{p \in \mathbb{N}_0^m, |p| = N} |F_p|^2 \right)^{1/2} \leq \sum_{p \in \mathbb{N}_0^m, |p| = N} |F_p| \leq \frac{|F|_s}{(S/\sqrt{m})^N} \left( \frac{N + m - 1}{N} \right).$$

where

$$\left( \frac{N + m - 1}{N} \right) = \frac{(N + m - 1)!}{N!(m - 1)!}$$

is the number of vectors $p \in \mathbb{N}_0^m$ with $|p| = N$.

We carry out the summation over $N$ to obtain $[F]_s$:

$$[F]_s = \sum_{N=0}^{\infty} [F^{(N)}] s^N \leq [F]_s \sum_{N} \left( \frac{\sqrt{m} s}{S} \right)^N \frac{(N + m - 1)!}{N!(m - 1)!}.$$

The summation converges for $s < S/\sqrt{m}$ since the first factor in the last summation then decreases exponentially while the second grows only polynomially. So for $s < S/\sqrt{m}$ we have

$$[F]_s \leq \bullet [F]_s.$$
8.1. Properties of the weighted norm

With minor adaptations the previous argument also applies to functions $F(z, \xi)$ analytic on

$$\{(z, \xi) \in \mathbb{C}^m : \|z\|, \|\xi\| < S\},$$

and this can be used directly in the estimate of $[F]_{r,s,w}$:

$$[F]_{r,s,w} = \sum_k e^{[k]} \sup_{I,\omega} \sum_{N,M} |F_k^{(N,M)}|^2 s^{N+M}$$

$$\leq \sum_k |k| \cdot \sup_{|I| < k, |\xi| \leq S} |F_k(I, z, \xi; \omega)|$$

$$\leq \sum_k e^{-|I|} \sup_{\mathcal{D}_r \times \mathcal{W}_w} |F(I, \varphi, z, \xi; \omega)|$$

$$\leq \cdot |F|_{r,s,w}.$$

We have used the Paley-Wiener theorem (Lemma 8.19 in Subsection 8.3.5) for the last but one inequality. \hfill \Box

8.1.3 Invariance under unitary transformations

The weighted norm is invariant under unitary coordinate transformations. We state this result in a way that is convenient for the application.

Families of unitary coordinate transformations and induced mappings Suppose that $U : \mathcal{O} \to U(m, \mathbb{C})$ is a family of complex unitary $m \times m$ matrices, so

$$U^*_\omega U_\omega = U_\omega U^*_\omega = 1 \text{ for all } \omega \in \mathcal{O}.$$ 

We do not assume the family to be continuous in $\omega$. This family induces the following transformation on functions on $\mathcal{D}_{r,s} \times \mathcal{W}_w$:

$$F \mapsto (uF) \text{ defined by } (uF)(\varphi, I, z, \xi; \omega) = F(\varphi, I, U^*_\omega z, U_\omega \xi; \omega).$$

The weighted norm is insensitive to such coordinate transformations:

Lemma 8.4 (Invariance under unitary transformations) Suppose that $F$ is analytic on $\mathcal{D}_{r,s} \times \mathcal{W}_w$ and has finite weighted norm. Let $U$ be a family of unitary transformations. Then we have

$$[(uF)]_{r,s,w} = [F]_{r,s,w},$$

and, more particularly,

$$[(uF)_k^{(N,M)}] = [F_k^{(N,M)}], \quad N, M \in \mathbb{N}_0,$$

where $(uF)_k^{(N,M)}$ is the collection of Fourier-Taylor coefficients of $(uF)$ corresponding to wave vector $k$ and of degree $N$ in $z$ and $M$ in $\xi$. 
**Remark** Note that the transformed function \((uF)\) is generally not real-analytic in \(\omega\). But, as we have remarked before, the weighted norm remains well-defined.

**Proof** We first prove the above lemma for functions depending only on \(z \in \mathbb{C}^n\) and analytic on \(|z| < s\), and then argue why it holds for analytic functions on \(D_{r,s} \times W_\omega\).

We calculate the tensor coefficients of degree \(N\) in \(z\) of

\[(uF)(z) = F(U^*z).\]

Using the abbreviations

\[z_a = \prod_{i=1}^{N} z_{a_i},\]
\[U^*_a = \prod_{i=1}^{N} U^*_{a_i},\]

where \(a, b \in \{1, \ldots, m\}^N\), it is easily verified that

\[(uF)^a = \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b.\]

The collection \((uF)^{(N)}\) of tensor coefficients is symmetric (see Definition 4.2) if the original collection \(F^{(N)}\) is symmetric: let \(\sigma\) be an arbitrary permutation on \(\{1, \ldots, N\}\), then

\[(uF)^{\sigma a} = \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b \sigma a = \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b \sigma a = \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b \sigma a = (uF)^a,\]

where the fourth equality holds since the original tensor coefficients are symmetric, the second and third because sum and product are commutative, and the others by definition of \((uF)\).

The next step is to prove that \([ (uF)^{(N)} ] = [F^{(N)}] \):

\[\sum_{a \in \{1, \ldots, m\}^N} \left( (uF)^a \right)^2 = \sum_{a \in \{1, \ldots, m\}^N} \left( (uF)^a \right)^2 = \sum_{a \in \{1, \ldots, m\}^N} \left( \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b \right)^2 = \sum_{a \in \{1, \ldots, m\}^N} \left( \sum_{b \in \{1, \ldots, m\}^N} F^b U^*_b \right)^2\]
8.1. Properties of the weighted norm

\[
\sum_a \left( \sum_b F^b U^c_{_{ba}} \right) \left( \sum_c F^c U^c_{_{ca}} \right) = \sum_b \sum_c F^b F^c \left( \sum_a U^c_{_{ba}} (\bar{U})_{ac} \right) = \sum_b \sum_c F^b F^c \delta_{bc} = \sum_b |F^b|^2 = |F(N)|^2.
\]

Since this holds for all \( N \), the proof is complete for functions depending only on \( z \). Again, the proof for analytic functions on \( \{(z, \xi) \in \mathbb{C}^n \times \mathbb{C}^m : \|z\|, \|\xi\| < s\} \) goes the same way, and since the other variables \( (I, \varphi; \omega) \) are not influenced by the unitary transformations, this implies validity of the lemma for functions depending on all variables.

8.1.4 The Banach property

We show here that the space of analytic functions on \( \mathcal{D}_{r,s} \times \mathcal{W}_w \) with finite \( \Gamma_{r,s,w} \) norm becomes a Banach space when it is equipped with the \( \Gamma_{r,s,w} \) topology.

Remarks

1. The Banach property makes it possible to use the generalized Cauchy inequality, see Subsection 8.3.1.

2. It is well-known that the space of analytic functions on a certain region is a Banach space when the supremum norm topology is used.

Proposition 8.5 (Banach property) The space \( (\mathcal{C}^\omega, \Gamma) \) of analytic functions on \( \mathcal{D}_{r,s} \times \mathcal{W}_w \) with finite weighted norm \( \Gamma \) is a Banach space.

In the proof we use a generalization of the concept of continuity:

Definition 8.6 (Upper semicontinuity) A real-valued function \( f \) is called upper semi-continuous (USC) if for all \( \lambda \in \mathbb{R} \) the set \( f^{-1}(] \lambda, \infty[) \) is open.

Lemma 8.7 (Basic properties of USC functions) Suppose \( A \) is any set of USC functions with common domain.

1. If \( f, g \in A \) then \( f + g \) is USC.

2. The pointwise supremum function \( \bar{f}(x) := \sup_{f \in A} f(x) \) is also USC.
We skip the proofs since they are easy and standard.

**Proof of the proposition**  By definition a Banach space is a **closed** normed space. In other words, if $(C^ω, [\cdot])$ is a Banach space, any Cauchy sequence $(F_n)_{n \in \mathbb{N}}$ in $(C^ω, [\cdot])$ has a limit, say $F_\infty$, in this space.

Since the weighted norm majorizes the supremum norm (see Lemma 8.2), a Cauchy sequence $(F_n)_{n \in \mathbb{N}}$ in $(C^ω, [\cdot])$ is automatically Cauchy in $(C^ω, [\cdot])$, which is the space of bounded analytic functions on $\mathcal{D}_r \times \mathcal{W}_w$ equipped with the supremum norm topology. Since the latter space is Banach, there indeed exists a limit $F_\infty$ with finite supremum. The question is if $F_\infty$ also has finite weighted norm.

We argue that this is certainly true if for any $\lambda \in \mathbb{R}$ the solid ball

$$B_\lambda = \{ F \in C^ω(\mathcal{D}_r \times \mathcal{W}_w) : [F]_{r,s,w} \leq \lambda \}$$

is closed in the supremum norm topology: For $\lambda$ sufficiently large each element of the sequence $(F_n)_{n \in \mathbb{N}}$ is contained in the ball $B_\lambda$, and provided this ball is closed in the supremum norm topology, it also contains the limit $F_\infty$.

It remains to check that the solid ball $B_\lambda$ is closed. We do this by showing that the weighted norm is USC in the supremum norm topology:

The tensor coefficients $F_k^{ab}$ are continuous in $(C^ω, [\cdot])$. Subsequently, the expression $[F_k]_{r,s,w}$ is evidently continuous in these tensor coefficients. The next step in the definition of the $[\cdot]_{r,s,w}$ norm is an infinite summation of $[F_k]_{(N,M)}$ over $N,M$. Realizing that this summation is over nonnegative elements only, it is equal to the supremum over the finite partial sums. Application of Lemma 8.7 yields upper semicontinuity of the summation. The final summation and supremum are dealt with in the same manner. Hence, the weighted norm is USC on $(C^ω, [\cdot])$. By Definition 8.6 the ball $B_\lambda$ is closed in the supremum norm topology, for all real $\lambda$. In the third paragraph of this proof we argued that the limit $F_\infty$ of the Cauchy sequence then has finite weighted norm. This finishes the proof. □

### 8.1.5 Submultiplicativity

The Banach space $(C^ω, [\cdot])$ is even a Banach algebra: the weighted norm is submultiplicative. We sketch the proof.

**Lemma 8.8 (Submultiplicativity)**  If $F$ and $G$ are in $(C^ω, [\cdot])$ then their product is that as well. Moreover, the weighted norm of the product $FG$ is at most the product of the weighted norms:

$$[FG]_{r,s,w} \leq [F]_{r,s,w} [G]_{r,s,w}.$$

**Sketch of the proof**  We prove submultiplicativity of the weighted norm for analytic functions depending only on $z \in \mathbb{C}^n$, and for analytic functions in $\varphi \in (\mathbb{C}/2\pi \mathbb{Z})^p$. In this way the essential ingredients for the proof for functions depending on all variables $(\varphi, I, z, \xi; \omega)$ are shown. We skip the actual proof since it involves very long formulas.
8.1. Properties of the weighted norm

**Proof for functions depending only on $z$** We consider functions $F, G$ that are analytic on $\|z\| \leq s$. The product of the degree $N$ part of $F$ with the degree $M$ part of $G$ can be written as

$$
\left( \sum_{a \in \{1, \ldots, m\}^N} F^{a_1, \ldots, a_N} z_{a_1} \cdots z_{a_N} \right) \left( \sum_{b \in \{1, \ldots, m\}^M} G^{b_1, \ldots, b_M} z_{b_1} \cdots z_{b_M} \right) = 
\sum_{c \in \{1, \ldots, m\}^{N+M}} F^{c_1, \ldots, c_N} G^{c_{N+1}, \ldots, c_{N+M}} z_{c_1} \cdots z_{c_{N+M}}.
$$

The collection of tensor coefficients of this product is denoted by $F^{[N]} \times G^{[M]}$, and one easily verifies that

$$
[F^{[N]} \times G^{[M]}] = [F^{[N]}] [G^{[M]}].
$$

However, the above coefficients are not symmetric, in general. According to the fourth remark following Definition 4.3 the previous equality becomes an inequality after symmetrization:

$$
[S(F^{[N]} \times G^{[M]})] \leq [F^{[N]}] [G^{[M]}].
$$

The degree $N$ symmetric coefficients of $FG$ are obtained by convolution followed by symmetrization:

$$
(FG)^{(N)} = \sum_{M=0}^{N} S(F^{[M]} \times G^{[N-M]}).
$$

Now the proof follows quickly:

$$
[FG]_s = \sum_{N=0}^{\infty} [F^{[N]}] [G^{[N]}] s^N
= \sum_{N=0}^{\infty} \left( \sum_{M=0}^{N} S(F^{[M]} \times G^{[N-M]} \right) s^N
\leq \sum_{N=0}^{\infty} \sum_{M=0}^{N} [F^{[M]}] s^M [G^{[N-M]}] s^{N-M}
= [F]_s [G]_s,
$$

where we have dropped the unnecessary subscripts $r$ and $w$. This finishes the proof for functions only depending on $z \in \mathbb{C}^n$. The proof for functions of $(z, \xi) \in \mathbb{C}^n \times \mathbb{C}^m$ goes the same, but it is notationally longer.
Proof for functions depending only on $\varphi$ Next we show submultiplicativity of the weighted norm for functions $F(\varphi)$ that are analytic on a complex neighborhood of the $n$-torus. In this case the norm reduces to

$$[F]_r = \sum_{k \in \mathbb{Z}^n} e^{\langle d \rangle |F_k|},$$

where the $F_k$ are Fourier coefficients. The unnecessary subscripts $s, w$ are dropped. Submultiplicativity is quickly proved:

$$[FG]_r = \sum_{k \in \mathbb{Z}^n} e^{\langle d \rangle \sum_{l \in \mathbb{Z}^n} |F_k||G_l|}$$

$$\leq \sum_{k \in \mathbb{Z}^n} e^{\langle d \rangle \sum_{l \in \mathbb{Z}^n} |F_k||G_l|}$$

$$\leq \sum_{l \in \mathbb{Z}^n} \sum_{k \in \mathbb{Z}^n} e^{\langle d \rangle \sum_{l \in \mathbb{Z}^n} |F_k||G_l|}$$

$$= [F]_r [G]_r.$$

As we have said before, the proof for functions of all variables $(\varphi, \nu, z, \xi; \omega)$ needs no new ideas but involves rather long formulas, and is therefore omitted.

8.1.6 Cauchy inequalities

We estimate the weighted norm of various partial derivatives of functions in $(C^\omega, [\cdot]_{r, s, w})$.

Lemma 8.9 (Cauchy inequalities)

$$\sum_{i=1}^n \left[ \frac{\partial F}{\partial \varphi_i} \right]_{r-\rho, s, w} \leq \frac{1}{\epsilon \rho} [F]_{r, s, w}$$

$$\sup_{i=1, \ldots, n} \left[ \frac{\partial F}{\partial I_i} \right]_{r, s-\sigma, w} \leq \frac{1}{s^2 - (s - \sigma)^2} [F]_{r, s, w}$$

$$\left( \sum_{i=1}^m \left[ \frac{\partial F}{\partial z_j} \right]_{r, s-\sigma, w}^2 \right)^{1/2} \leq \sqrt{\frac{m}{s \delta}} [F]_{r, s, w}$$

$$\left( \sum_{i=1}^m \left[ \frac{\partial F}{\partial \xi_j} \right]_{r, s-\sigma, w}^2 \right)^{1/2} \leq \sqrt{\frac{m}{s \delta}} [F]_{r, s, w}.$$  

Proof For the first two inequalities we can follow [P1] almost verbatim: We have

$$\sum_{i=1}^n \left[ \frac{\partial F}{\partial \varphi_i} \right]_{r-\rho, s, w} = \sum_{k} e^{\langle d \rangle \sum_{l} \sup_{s, M} \sum_{N} \left[ f_k(N, M) \right]_{s, N+M}$$

$$= \sum_{k} |k| \left[ d \right]_{r-\rho, s, w} \sup_{s, M} \sum_{N} \left[ f_k(N, M) \right]_{s, N+M}$$

$$\leq \sup_{t \geq 0} t e^{-t \rho} [F]_{r, s, w}.$$
and the supremum equals $1/\pi$. This proves the first inequality.

As to the second inequality,

$$
\left( \frac{\partial F}{\partial I_{i}} \right)_{k}^{ab} (I, \omega) = \frac{1}{2\pi i} \oint_{I} \frac{F_{k}^{ab}(J; \omega)}{(J - I)^{2}} dJ
$$

by Cauchy’s formula with a suitable path $\gamma$ in a plane containing $I$ and parallel to the $I_{i}$-plane. Hence

$$
\sup_{|1| < (s - \omega)^{2}} \left| \left( \frac{\partial F}{\partial I_{i}} \right)_{k}^{ab} \right| \leq \sup_{|1| < (s - \omega)^{2}} \frac{|F_{k}^{ab}|}{s^{2} - (s - \omega)^{2}},
$$

and consequently

$$
\left[ \frac{\partial F}{\partial I_{i}} \right]_{r,s}^{\alpha,\omega} = \sum_{k} e^{i|\chi|} \sup_{|1| < (s - \omega)^{2}} \sum_{N,M} [(\partial F/\partial I_{i})_{N}^{(N,M)}]_{r,s}^{\alpha,\omega} (s - \omega)^{N+M}
$$

$$
\leq \frac{1}{s^{2} - (s - \omega)^{2}} \sum_{k} e^{i|\chi|} \sup_{|1| < (s - \omega)^{2}} \sum_{N,M} [F_{k}^{(N,M)}]_{r,s}^{\alpha,\omega} (s - \omega)^{N+M}
$$

$$
\leq \frac{|F|_{r,s}^{\alpha,\omega}}{s^{2} - (s - \omega)^{2}}.
$$

This holds for all $i \in \{1, \ldots, n\}$, proving the second inequality.

For the last inequality, we first compute the symmetric tensor coefficients of $\partial F/\partial z_{i}$ in terms of those of $F$. For $a \in \{1, \ldots, m\}^{N}$, we have

$$
\left( \frac{\partial F}{\partial z_{i}} \right)_{k}^{ab} = \sum_{j=1}^{N+1} F_{k}^{ab} (\delta_{j}, \ldots, \delta_{j-1}, \ldots, 0)_{r,s}^{\alpha,\omega},
$$

where we have supposed the latter coefficients symmetric. It follows that

$$
\left[ \frac{\partial F}{\partial z_{i}} \right]_{r,s}^{\alpha,\omega}
$$

$$
= \sum_{k} e^{i|\chi|} \sup_{I,\omega} \sum_{N,M} \left( \sum_{a,b} \left| \sum_{j=1}^{N+1} F_{k}^{ab} (\delta_{j}, \ldots, \delta_{j-1}, \ldots, 0)_{r,s}^{\alpha,\omega} \right|^{2} \right)^{1/2} (s - \omega)^{N+M}
$$

$$
\leq \sum_{k} e^{i|\chi|} \sup_{I,\omega} \sum_{N,M} \left( \sum_{a,b} \left( N + 1 \right) \sum_{j=1}^{N+1} \left| F_{k}^{ab} (\delta_{j}, \ldots, \delta_{j-1}, \ldots, 0)_{r,s}^{\alpha,\omega} \right|^{2} \right)^{1/2} (s - \omega)^{N+M}
$$

$$
\leq \sum_{k} e^{i|\chi|} \sup_{I,\omega} \sum_{N,M} \sqrt{N + 1} \left( \sum_{a,b} \left| F_{k}^{ab} \right|^{2} \right)^{1/2} (s - \omega)^{N+M}
$$

$$
= \sum_{k} e^{i|\chi|} \sup_{I,\omega} \sum_{N,M} \sqrt{N + 1} \left| F_{k}^{(N+1,M)} \right|_{r,s}^{\alpha,\omega} (s - \omega)^{N+M}
$$

$$
= \sum_{k} e^{i|\chi|} \sup_{I,\omega} \sum_{N,M} \left| F_{k}^{(N+1,M)} \right|_{s}^{N+M+1} \left( \frac{\sqrt{N + 1}(s - \omega)^{N+M}}{s^{N+M+1}} \right).
$$
The supremum over \( N, M \) of the fraction in the last line is estimated as follows
\[
\sup_{N,M} \frac{\sqrt{N + 1}(s - \sigma)^{N+M}}{N^{N+M+1}} = \sup_N \frac{\sqrt{N + 1}(s - \sigma)^N}{N^{N+1}} \leq \sqrt{\frac{1}{s\sigma}}.
\]
Taking the square root of the sum of the squares of \( [\partial F/\partial z]_{r,\sigma,w} \) over \( i \) finishes the proof of third inequality.

The fourth Cauchy inequality follows from the third by symmetry. \( \Box \)

**An extra dependence on** \( m \) **in the last two Cauchy inequalities there is a dependence on** \( m \) **not present in** [P1]. **In the following estimates we eliminate this extra factor by presuming a certain parameter to be sufficiently small.**

### 8.1.7 Estimates for the Poisson bracket

The estimate for the Poisson bracket is stated in a form that is convenient for our applications.

**Lemma 8.10 (Poisson bracket)** Suppose \( \rho_0 \leq 1/m \) and
\[
\frac{1}{\rho_0} [F]_{\gamma_0-\rho_0,\rho_0,w} \sum_i [\frac{\partial F}{\partial \varphi_i}]_{\gamma_0-\rho_0,\rho_0,w} \leq C.
\]
Then
\[
[[F,G]]_{r,\rho,s-\sigma,w} \leq \left( \frac{1}{s\sigma} + \frac{\rho_0}{\rho} \frac{\rho_0}{e(s_0^2 - (s - \sigma)^2)} + \frac{2}{\sqrt{s_0(s_0 - (s - \sigma)s\sigma)}} \right) C[G]_{r,s,w},
\]
for \( 0 < \rho < r \) and \( 0 < \sigma < s \) with \( r - \rho \leq r_0 - \rho_0 \) and \( s - \sigma \leq s_0 \).

**Proof** Recalling Definition 4.1 of the complex variables, one easily verifies that
\[
\{F,G\} = \{F_{\varphi}, G_{\xi}\} + i\{F_{\xi}, G_{\varphi}\} + i\{F_{\varphi}, G_{\xi}\} - i\{F_{\xi}, G_{\varphi}\},
\]
where we have abbreviated \( \partial F/\partial \varphi \) to \( F_{\varphi} \), etcetera.

Let \([\cdot] = [\cdot]_{r,\rho,s-\sigma,w}. \) By Lemmas 8.8 and 8.9 we have
\[
[\{F_{\varphi}, G_{\xi}\}]_\gamma \leq \sum_i [F_{\varphi_i}, G_{\xi_i}]_\gamma \leq \sum_i [F_{\varphi_i}]_\gamma [G_{\xi_i}]_\gamma \leq \frac{C}{s\sigma}[G]_{r,s,w},
\]
since \( 2s - \sigma \geq s \). Similarly,
\[
[\{F_{\xi}, G_{\varphi}\}]_\gamma \leq \sup_i [F_{\xi_i}]_\gamma \sum_i [G_{\varphi_i}]_\gamma \leq \frac{1}{s_0^2 - (s - \sigma)^2} [F]_{r,\rho_0,\rho_0,w} \frac{1}{e}\rho \rho_0 [G]_{r,\rho,s,w} \leq \frac{\rho_0}{e(s_0^2 - (s - \sigma)^2)} C[G]_{r,s,w}.
\]
The estimate for the third term is quite rough and yields a factor \( m \) that is not present in \([P1]\):

\[
\sum_i [F_i]_\sigma \leq \sum_i [G_i]_\sigma 
\]

\[
\leq \left( \sum_i [F_i]_\sigma^2 \right)^{1/2} \left( \sum_i [G_i]_\sigma^2 \right)^{1/2} 
\]

\[
\leq \sqrt{\frac{m}{s_0(s_0 - (s - \sigma))}} [F]_{\rho_0, \rho, s, w} \sqrt{\frac{m}{s\sigma}} [G]_{\tau, s, w} 
\]

\[
= \sqrt{\frac{s_0(s_0 - (s - \sigma))s\sigma}{m\rho_0}} C[G]_{\tau, s, w}.
\]

The estimate for the fourth term is identical. Adding up the estimates for the four terms of \( \{F, G\} \) and using the assumption \( \rho_0 < 1/m \) finishes the proof. \( \square \)

### 8.1.8 Estimate for small symplectic transformations

Unlike the familiar supremum-norm, the \([\cdot]_{r, s, w}\)-norm of a function is sensitive to coordinate transformations. Fortunately, we only need to consider canonical transformations that are close to the identity. The following lemma is therefore stated with our specific application in mind.

**Lemma 8.11 (Transformation)** Suppose that \( \rho_0 \leq 1/m \) and

\[
\frac{1}{\rho_0} [F]_{\rho_0, \rho, s, w} \sum_{i=1}^n \left[ \frac{\partial F}{\partial \varphi_i} \right]_{\rho_0, \rho, s, w} \leq C \leq \frac{s^2}{35}.
\]

Then, for \( 0 < \rho < \rho_0 \leq r = \rho_0 \) and \( 0 < s \leq s_0/2 \), one has

\[
\left[ G \circ \Phi \right]_{\rho, s/2, w} \leq 2 [G]_{\tau, s, w},
\]

where \( \Phi \) denotes the time-1-map of the Hamiltonian vector field \( X_F \).

**Remarks**

1. The hypotheses of the lemma imply that

\[
X_t^F : D_{r, \rho, s/2} \rightarrow D_{r, \sigma}, \quad 0 \leq t \leq 1,
\]

which is easily verified using the arguments of Subsection 7.1.4.

2. The lemma holds when \( \Phi \) is replaced by \( X_t^F \), \( 0 \leq t \leq 1 \) since the time-\( t \)-flow according to the Hamiltonian \( F \) is identical to the time-1-flow of \( tF \):

\[
X_t^F = X_{tF}^1.
\]

Moreover, if the conditions in the lemma hold for \( F \), they also hold for \( tF \) for \( 0 \leq t \leq 1 \) by homogeneity.
Proof of Lemma 8.11  Consider the Lie series expansion
\[ G \circ \Phi = \sum_{h=0}^{\infty} \frac{1}{h!} \text{ad}_F^h G, \]
where
\[ \text{ad}^0_F G = G, \quad \text{ad}^h_F G = \{ \text{ad}^{h-1}_F G, F \}, \quad h > 0. \]

Using the notation \([\cdot]_h\) for \([\cdot]_{r,p,s-rh,sw}\) we have for arbitrary \(\rho, \sigma\) and positive integers \(h\) with \(0 < h\rho \leq \rho < r\) and \(0 < h\sigma \leq s/2\) that
\[
\begin{align*}
[\text{ad}^h_F G]_h & = [\{\text{ad}^{h-1}_F G, F\}]_h \\
& \leq \left( \frac{1}{(s - (h - 1)\sigma)\sigma} + \frac{\rho_0 - \rho}{s(\rho_0 - (s - h\sigma)^2)} + \frac{2m\rho_0}{\sqrt{s_0(s_0 - (s - h\sigma))(s - (h - 1)\sigma)\sigma}} \right) C[\text{ad}^{h-1}_F G]_{h-1} \\
& \leq \left( \frac{2}{s\sigma} + \frac{4\rho_0 - \rho}{15s^2} + \frac{2\sqrt{2}}{\sqrt{3s^3(\sigma)}} \right) C[\text{ad}^{h-1}_F G]_{h-1}, \quad (8.2)
\end{align*}
\]
by the preceding lemma and the assumption \(s_0 \geq 2s\). Iterating this \(h\) times gives
\[
[\text{ad}^h_F G]_h \leq \left( \frac{2}{s\sigma} + \frac{4\rho_0 - \rho}{15s^2} + \frac{2\sqrt{2}}{\sqrt{3s^3(\sigma)}} \right)^h C[G]_{r,s,w}.
\]
Replacing \(\rho, \sigma\) by \(\rho/h, s/2h\), respectively, and using the assumption \(\rho_0 \leq \rho < r\) this yields
\[
[\text{ad}^h_F G]_{r-p,s/2w} \leq \left( \frac{6.41hC}{s^2} \right)^h [G]_{r,s,w}.
\]
By Stirling’s formula, \(h^p / h! \leq e^h\) for \(h \geq 1\). Taylor expanding \(G \circ \Phi\), we obtain
\[
[\text{ad}^h_F G]_{r-p,s/2w} \leq \sum_{h=0}^{\infty} \frac{1}{h!} \text{ad}^h_F G \leq \sum_{h=0}^{\infty} \frac{1}{h!} [\text{ad}^{h-1}_F G]_{r-p,s/2w} \\
\leq \sum_{h=0}^{\infty} \left( \frac{35C}{2s^2} \right)^h [G]_{r,s,w}.
\]
Supposing that \(35C \leq s^2\) we have
\[
[\text{ad}^h_F G]_{r-p,s/2w} \leq 2 [G]_{r,s,w}, \quad (8.3)
\]
which concludes the proof. \(\square\)
8.2 The functions $\Delta$, $\Gamma$ and $\Psi$

We define the functions $\Delta$, $\Gamma$ and $\Psi$, which turn up in the proof of the KAM theorem, and discuss some of their properties.

The small divisor function $\Delta$  The first of these three functions characterizes the small divisor conditions (7.2). We define it as follows:

$$\Delta(t) = (1 + t)^\tau. \quad (8.4)$$

For the measure estimate it is necessary that $\tau > n$. We even assume

$$\tau \geq n + 1.$$

The above definition of $\Delta$ makes the small divisor conditions look like the usual Diophantine conditions. In [R1] the class of approximation functions is introduced to characterize a large class of small divisors for which the KAM procedure is applicable.

The function $\Gamma$  The function $\Gamma$ arises in Proposition 7.2, which estimates the size of the transformation Hamiltonian $F$ in terms of the size of the truncated perturbation $R$. It is defined on the positive real axis by

$$\Gamma(\rho) = \sup_{t \geq 0} (1 + t) \Delta(t)e^{-\rho t}. \quad (8.5)$$

Obviously, $\Gamma(\rho)$ is decreasing in $\rho$. For $\Delta(t) = (1 + t)^\tau$ with positive $\tau$ and $0 < \rho \leq 1$ we find

$$\Gamma(\rho) = \left(\frac{\tau + 1}{e\rho}\right)^{\tau + 1} e^\rho. \quad (8.6)$$

For $\tau \geq 1$ it is easily checked that

$$\sup_{t \geq 0} \Delta(t)e^{-\rho t} \leq \rho \Gamma(\rho). \quad (8.7)$$

Summing $\Delta(|k|)$ over $\mathbb{Z}^n$  With the measure estimate in mind we present the following estimate:

Lemma 8.12 For $\Delta(t) = (1 + t)^\tau$, with $\tau \geq n + 1$ we have

$$\sum_{k \in \mathbb{Z}^n} \Delta(|k|)^{-1} \leq 1 + 2^n.$$
Proof The estimate is given by a number of inequalities, which are explained afterwards:

\[
\sum_{k \in \mathbb{Z}^n} \frac{1}{(1 + |k|)^\tau} \leq 1 + 2^n \sum_{k \in (\mathbb{N}_0)^n \setminus \{0\}} \frac{1}{(1 + |k|)^\tau}
\]

\[
\leq 1 + 2^n \sum_{t=1}^{\infty} \frac{1}{(1 + t)^\tau} \left( \binom{n-1}{t} \right)
\]

\[
\leq 1 + 2^n \sum_{t=1}^{\infty} \frac{1}{(1 + t)^{\tau+n}}
\]

\[
\leq 1 + 2^{2n-\tau+1} \leq 1 + 2^n.
\]

1. Since the summand depends only on $|k|$, the summation over $\mathbb{Z}^n$ is bounded from above by $2^k$ times the summation over $(\mathbb{N}_0)^n$, where $2^n$ is the number of “quadrants”. In this way vectors $k$ with components equal to zero are counted too often. We improve on this coarse estimate by considering $k = 0$ separately.

2. The number of different $n$-vectors with nonnegative integer components with sum $t$ is equal to

\[
\binom{n-1+t}{t}.
\]

Such vectors can be uniquely represented by a row of $t$ blue balls and $n$ red balls. The number of different rows is then obtained using elementary combinatorics.

3. Using the inequality $t + s \leq s(t + 1)$ for $s, t \geq 1$, the binomial coefficient is estimated by

\[
\binom{n-1+t}{t} = \frac{(n-1+t)!}{(n-1)! t!}
\]

\[
= \frac{(n-1+t)(n-2+t)\ldots(1+t)}{(n-1)(n-2)\ldots1}
\]

\[
\leq (1 + t)^{n-1}.
\]

4. The summation in the last but one line is bounded from above by:

\[
\sum_{t=1}^{\infty} \frac{1}{(1 + t)^{\tau-n+1}} \leq 2^{-(\tau-n+1)} + \int_{t=1}^{\infty} \frac{dt}{(1 + t)^{\tau-n+1}}
\]

\[
= 2^{-\tau+n-1} + 2^{-\tau+n}/(\tau - n).
\]

For $\tau \geq n + 1$ this is smaller than $2^{-\tau+n+1}$

This finishes the proof of Lemma 8.12. \qed
8.3. Auxiliary lemmas

The sequence $\{\kappa_v\}$ Before defining $\Psi(\rho)$, we recall the sequence $\{\kappa_v\}$, which was defined in (7.35) by

$$
\kappa_v = \frac{\kappa - 1}{\kappa^v}, \quad v = 0, 1, 2, \ldots,
$$

where

$$
\kappa = \frac{4}{3}.
$$

This sequence has the following properties

$$
\sum_{\nu=0}^{\infty} \kappa_{v} = 1, \quad \text{and} \quad \sum_{\nu=0}^{\infty} \nu \kappa_{v} = \frac{1}{\kappa - 1}.
$$

The function $\Psi(\rho)$ The function $\Psi(\rho)$ is defined as

$$
\Psi(\rho) = \inf \prod_{\nu=0}^{\infty} \Gamma_{\nu}(\rho_{\nu}), \quad (8.8)
$$

where the infimum is over positive decreasing sequences $\{\rho_{\nu}\}$ with

$$
\sum_{\nu=0}^{\infty} \rho_{\nu} = \rho.
$$

Calculating the argument of the infinite product for the sequence $\{\rho_{\nu}\}$ with

$$
\rho_{\nu} = \kappa_{\nu} \rho,
$$

gives the following bound:

$$
\Psi(\rho) \leq \overline{\Psi}(\rho) := \left(\frac{10(\tau + 1)}{\rho e} \right) e^{\frac{\tau + 1}{\rho}}. \quad (8.9)
$$

8.3 Auxiliary lemmas

8.3.1 Cauchy inequalities and related estimates

The usual Cauchy inequalities are generalized to inequalities concerning functions that are analytic on an open subset of a complex Banach space. We also give estimates on how well such functions are approximated by their constant and affine parts on smaller domains.

Definition of norms Let $A$ and $B$ be two complex Banach spaces with norms $|\cdot|_A$ and $|\cdot|_B$, and let $F$ be an analytic map from an open subset of $A$ into $B$. The first derivative $d_v F$ of $F$ at $v$ is a linear map from $A$ into $B$, whose induced operator norm is

$$
|d_v F|_{B,A} = \max_{u \neq 0} \frac{|d_v F(u)|_B}{|u|_A}.
$$
Analogously, the second derivative \( d_v^2F \) is a quadratic map whose norm is defined as
\[
|d_v^2F|_{B,A} = \max_{u \neq \emptyset} \frac{|d_v^2F(u)|_B}{|u|_A^2}.
\]

The Cauchy inequalities estimating the norms of \( d_vF \) and \( d_v^2F \) can be stated as follows:

**Lemma 8.13 (Generalized Cauchy inequalities)** Let \( F \) be an analytic map from the open ball of radius \( r \) around \( v \) in \( A \) into \( B \) such that \( |F|_B \leq M \) on this ball. Then the inequalities
\[
|d_vF|_{B,A} \leq \frac{M}{r} \quad (8.10)
\]
\[
|d_v^2F|_{B,A} \leq \frac{2M}{r^2} \quad (8.11)
\]
hold.

**Proof** For the first inequality we repeat the proof from [P1]. Let \( u \neq 0 \) be in \( A \). Then \( f(z) = F(v + zu) \) is an analytic map from the complex disc \( |z| < r/|u|_A \) in \( C \) into \( B \) that is uniformly bounded by \( M \). Hence
\[
|d_0f|_B = |d_vF(u)|_B \leq \frac{M}{r} \cdot |u|_A
\]
\[
|d_0^2f|_B = |d_v^2F(u)|_B \leq \frac{2M}{r^2} \cdot |u|^2_A
\]
by the corresponding usual Cauchy inequalities. The statements in the lemma follow, since \( u \neq 0 \) was arbitrary.

**Lemma 8.14 (Approximating analytic functions by constant and affine parts)**

Again, let \( F \) be an analytic map from the open ball of radius \( r \) around \( v \) in \( A \) into \( B \) such that \( |F|_B \leq M \) on this ball. Then for \( 0 < \gamma < 1 \) the inequalities
\[
\max_{|u-v|_A \leq \gamma r} |F(u) - F(v)|_B \leq \frac{\gamma M}{1 - \gamma}
\]
\[
\max_{|u-v|_A \leq \gamma r} |F(u) - d_vF(u) - F(v)|_{B} \leq \frac{\gamma^2 M}{1 - \gamma}
\]
hold.

**Proof** The difference between \( F(u) \) and \( F(v) \) equals the line-integral of the derivative of \( F \) along the straight line from \( v \) to \( u \):
\[
F(u) - F(v) = \int_{s=0}^{1} (d_{v+s|u-v|}F)(u - v) \, ds.
\]
Using the first generalized Cauchy inequality, this difference is estimated by

\[ |F(u) - F(v)|_B \leq \int_{s=0}^{1} |d_{v+s(u-v)}F|_{B,A}|u - v|_A \, ds \]

\[ \leq \int_{s=0}^{1} \frac{M}{s - s|u - v|_A} \cdot |u - v|_A \, ds \]

\[ = \int_{s=0}^{1} \frac{M}{1 - s\gamma} \cdot \gamma \, ds \]

\[ = -M\log(1 - \gamma). \]

The inequality

\[ -\log(1 - \gamma) \leq \gamma/(1 - \gamma), \] which holds for \( \gamma < 1, \]

finishes the proof of the first estimate in Lemma 8.14.

The difference between \( F(u) \) and its affine approximation \( F(v) + (d_v F)(u - v) \) is

\[ F(u) - F(v) - (d_v F)(u - v) = \int_{s=0}^{1} \int_{r=0}^{s} (d^2_{v+r(u-v)}F)(u - v) \, dr \, ds \]

\[ = \int_{s=0}^{1} (1 - s) (d^2_{v+s(u-v)}F)(u - v) \, ds - \]

\[ \left( (1 - s) \int_{r=0}^{s} (d^2_{v+r(u-v)}F)(u - v) \, dr \right) \left|_{s=0}^{1} \right. , \]

where the last equality has been obtained by partial integration. The last term is zero. Using the generalized Cauchy inequality for the second derivative of \( F \), we obtain

\[ |F(u) - F(v) - (d_v F)(u - v)|_B \leq \int_{s=0}^{1} (1 - s) |d^2_{v+s(u-v)}F|_{B,A}|u - v|_A^2 \, ds \]

\[ \leq \int_{s=0}^{1} (1 - s) \frac{2M}{(s - s|u - v|_A)^2} \cdot |u - v|_A^2 \, ds \]

\[ \leq \int_{s=0}^{1} (1 - s) \frac{2M}{(1 - s\gamma)^2} \cdot \gamma^2 \, ds \]

\[ = -2M \gamma (\gamma + \log(1 - \gamma)). \]

Application of inequality (8.12) finishes the proof.

\[ \square \]

8.3.2 The operator and Frobenius norm on matrices

**Definition 8.15** For matrices \( A \in \mathfrak{gl}(m, \mathbb{C}) \) we define its operator norm \( \|A\| \) and Frobenius norm \( \|A\|_F \) as follows:

\[ \|A\| = \sup_{z \in \mathbb{C}^m \setminus \{0\}} \frac{\|Az\|}{\|z\|}, \] (8.13)
\[ \|A\|_F = \left( \sum_{i,j=1}^{m} |A_{ij}|^2 \right)^{1/2}, \]  
(8.14)

where \( \| \cdot \| \) in the fraction in (8.13) denotes the Euclidean norm on \( \mathbb{C}^m \).

We state, without proof, a number of useful (in)equalities:

**Lemma 8.16** For \( A, B \in \text{gl}(\mathbb{C}, m) \) we have

\[
\begin{align*}
\|A^T\| &= \|A\|, \\
\|A^T\|_F &= \|A\|_F, \\
\|A\| &\leq \|A\|_F, \\
\|AB\|_F &\leq \|A\|\|B\|_F. 
\end{align*}
\]

### 8.3.3 The difference between the eigenvalues of different Hermitian matrices

We copy the following useful estimate from [Pa] (Fact 1-11 on page 14, and Section 10-3). It is a special case of a theorem first written down by Weyl [We], which was already known in the nineteenth century.

**Lemma 8.17** Suppose that \( H \) and \( H' \) are Hermitian \( m \times m \) matrices with eigenvalues \( (\lambda_1, \ldots, \lambda_m) \) and \( (\lambda'_1, \ldots, \lambda'_m) \), respectively, which are ordered from small to large. In terms of the norm

\[
\|H\| = \sup_{z \in \mathbb{C}^m \setminus \{0\}} \frac{\|Hz\|}{\|z\|}
\]

we have the following estimate of the difference between the eigenvalues:

\[ |\lambda' - \lambda|_\infty \leq \|H' - H\|. \]

### 8.3.4 Gershgorin’s theorem

For easy reference, we copy Gershgorin’s theorem from [C]:

**Theorem 8.18 (Gershgorin)** Each eigenvalue of \( A = (a_{ij}) \) lies in at least one of the Gershgorin disks

\[ \left\{ z : |z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}, \text{ for } i = 1, 2, \ldots, n. \]
8.3.5 The Paley-Wiener theorem

**Lemma 8.19 (Theorem of Paley-Wiener, [BHS])** Suppose that

\[ F(\varphi) = \sum_{k \in \mathbb{Z}^n} F_k e^{i(k,\varphi)} \]

is analytic and bounded on

\[ \mathcal{T}_r = \{ \varphi \in (\mathbb{C}/2\pi\mathbb{Z})^n : |\text{Im } \varphi|_\infty < r \}. \]

Then the Fourier coefficients are bounded as follows:

\[ |F_k| \leq \sup_{\varphi \in \mathcal{T}_r} |F(\varphi)| e^{-|k|} \text{ for all } k \in \mathbb{Z}^n. \]

8.3.6 An inverse function theorem

From [P1], we copy the inverse function theorem that is used for the reparametrization of \( \omega \) during the KAM-step (Subsection 7.1.6). Recall that \( \mathcal{W}_w \) is the open complex neighborhood of radius \( w \) (in the maximum norm) around some closed subset \( \mathcal{O} \) of \( \mathbb{R}^n \):

\[ \mathcal{W}_w := \{ \omega \in \mathbb{C}^n : \exists \omega' \in \mathcal{O} : |\omega' - \omega|_\infty < w \}. \]

**Lemma 8.20 (Inverse function theorem, [P1])** Suppose \( f \) is a real-analytic map from \( \mathcal{W}_w \) into \( \mathbb{C}^n \). If

\[ |f - \text{id}|_\infty \leq \delta \leq w/4 \]

on \( \mathcal{W}_w \), then \( f \) has a real-analytic inverse \( \varphi \) on \( \mathcal{W}_w/4 \). Moreover,

\[ |\varphi - \text{id}|_\infty \leq \delta, \quad w|d\varphi - 1|_\infty \leq \frac{2\delta/w}{1 - 2\delta/w} \]

on this domain.

8.3.7 A Lipschitz estimate for analytic functions

**Lemma 8.21** Suppose that \( f \) is an analytic map from \( \mathcal{W}_w \) to a Banach space \( (B, |\cdot|) \), which is bounded in the following way

\[ \sup_{\omega \in \mathcal{W}_w} |f(\omega)| \leq E. \]

Then the following estimate for the Lipschitz seminorm of \( f \) on \( \mathcal{O} \) holds:

\[ |f|_{Lip, \mathcal{O}} := \sup_{\omega, \omega' \in \Omega, \omega \neq \omega'} \frac{|f(\omega) - f(\omega')|_B}{|\omega - \omega'|_\infty} \leq \frac{2E}{w}. \]
Proof Consider for arbitrary \( \omega, \overline{\omega} \in \mathcal{O} \) the closed interval
\[
[\omega, \overline{\omega}] := \{(1 - \lambda)\omega + \lambda\overline{\omega}, \lambda \in [0, 1]\}.
\]
The intersection \([\omega, \overline{\omega}] \cap \mathcal{O}\) is closed in the standard topology of the real line. Hence its complement \([\omega, \overline{\omega}] \setminus \mathcal{O}\) is open and consists of a countable union of disjoint open intervals:
\[
|[\omega, \overline{\omega}] \setminus \mathcal{O}| = \bigcup_{n \in \mathbb{N}} [\omega_n, \omega'_n].
\]
Now consider the union \( \mathcal{U} \) of \([\omega, \overline{\omega}] \cap \mathcal{O}\) with those open intervals \([\omega_n, \omega'_n]\) for which \(|\omega'_n - \omega_n|_{\infty} < w\). In other words, \( \mathcal{U} \) is obtained from \([\omega, \overline{\omega}] \cap \mathcal{O}\) by filling all gaps shorter than \(w\) in the \(| \cdot |_{\infty}\)-norm. The remaining gaps are at least of length \(w\), and therefore finite in number. Hence \( \mathcal{U} \) is a finite union of disjoint closed intervals:
\[
\mathcal{U} = [\omega_1, \overline{\omega}_1] \cup [\omega_2, \overline{\omega}_2] \cup \cdots \cup [\omega_N, \overline{\omega}_N],
\]
where we suppose \(\omega = \omega_1 \leq \overline{\omega}_1 < \omega_2 \leq \cdots \leq \omega_N = \overline{\omega}\). Obviously, \( \mathcal{U} \) is contained in \( \mathcal{W}_w \), moreover, the distance from \( \mathcal{U} \) to the boundary \( \partial \mathcal{W}_w \) is at least \(w/2\), since \( \mathcal{U} \) is completely within distance \(w/2\) of \( \mathcal{O} \). The generalized Cauchy inequality (Lemma 8.3.1) gives
\[
|d_{\omega}f|_{L, \infty} \leq \frac{2E}{w}, \; \omega \in \mathcal{U}.
\]
and this implies that
\[
|f(\overline{\omega}) - f(\omega)|_B \leq \frac{2E}{w} |\overline{\omega}_i - \omega_i|_{\infty}, \; i = 1, \ldots, N.
\]
From \(|f|_B \leq E\) it follows that across the gaps in \( \mathcal{U} \) we have
\[
|f(\omega_{i+1}) - f(\omega_i)|_B \leq 2E \leq \frac{2E}{w} |\omega_{i+1} - \omega_i|_{\infty}, \; i = 1, \ldots, N - 1.
\]
Summation of these estimates implies the statement of the lemma. \(\square\)

8.3.8 The volume of a projection of a hypercube

Lemma 8.22 The \((n - 1)\)-dimensional volume of any orthogonal projection of the \(n\)-dimensional unit cube is bounded from above by \(\sqrt{n}\).

8.4 Estimating the size of the perturbation

In a number of lemmas we estimate the weighted norms of the perturbation components \(P_1, P_2, P_3, P_4\) given in (5.2, ..., 5.5). Together, these lemmas imply Proposition 5.2. At the end of the section we outline how the improved estimate of Proposition 5.3 is obtained.
Lemma 8.23 (Estimating the size of \( P^1 \)) Under Conditions 5.1 on \((r, s, w)\) we have

\[
[P^1]_{r,s,w} = \left[ \sum_{i=1}^{n} f(J(\omega_i) + I_i) - \omega_i I_i \right]_{r,s,w} \leq (4/s^4) s^4.
\]

Proof Since \( P^1 \) does not depend on \((\varphi, z, \xi)\), its weighted norm reduces to

\[
[P^1]_{r,s,w} = \sup_{\omega \in \mathbb{W}_s} \left| \sum_{i=1}^{n} f(I_i) - f(J(\omega_i)) - \omega_i I_i \right| 
\leq \sup_{\omega} \sup_{\mu \in \mathbb{R}^{*}} \sum_{i=1}^{n} \sup_{|I| < \mu^2 s^4} \left| f(I_i) - f(J(\omega_i)) - \omega_i I_i \right|.
\]

Now we apply Lemma 8.14 to find

\[
[P^1]_{r,s,w} \leq \sup_{\mu} \sum_{l} \frac{2(\mu_4 s^2 / s^2)^2}{1 - \mu_4 s^2 / s^2} \leq (4/s^4) s^4 \sup_{\mu} \sum_{l} \mu_4^l \leq (4/s^4) s^4,
\]

where we have used \( s^2 \leq \bar{s}^2 / 2 \).

Lemma 8.24 (Estimating the size of \( P^2 \)) Under Conditions 5.1 we have

\[
[P^2]_{r,s,w} = \left[ \sum_{j=n+1}^{n+m} (1 - \cos q_j) - \frac{1}{2} q_j^2 \right]_{r,s,w} \leq (\cosh(\bar{s}) / 6) \sqrt{m} s^8.
\]

Proof We first do the estimate for \( m = 1 \). Taylor expanding, replacing \( q \) by \((z + \xi) / \sqrt{2} \) and rearranging gives

\[
1 - \cos q - q^2 / 2 = - \sum_{k+l=4,6,8,\ldots} (-1)^{(k+l)/2} (\sqrt{2})^k (\xi / \sqrt{2})^l / k! l!.
\]

Calculating the weighted norm now consists of taking moduli of the Taylor coefficients, and then replacing \( z \) and \( \xi \) by \( s \) as follows:

\[
\left[ 1 - q^2 - \cos q \right]_{r,s,w} = \sum_{k+l=4,6,8,\ldots} \frac{(s / \sqrt{2})^k (s / \sqrt{2})^l}{k! l!} = \cosh(\sqrt{2}s) - 1 - s^2 \leq (\cosh(\bar{s}) / 6) s^4,
\]

where the last line follows from \( s \leq \bar{s} / \sqrt{2} \) and Taylor’s formula with remainder (see for instance [Ap1], Sections 7.5, 7.6). For \( m > 1 \) the perturbation \( P^2 \) is a sum of functions depending only on one of the \( q_j \) at a time, with \( j = n + 1, \ldots, n + m \). For that reason each term in the Taylor expansion will depend only on \((z_j, \xi_j)\) for one \( j \). This facilitates computation of \([P^2]_{r,s,w}\) since its coefficients are already symmetric. In fact, the weighted norm gets only a factor \( \sqrt{m} \) bigger. \( \square \)
Lemma 8.25 (Estimating the size of $P^3_c$) Under Conditions 5.1 we have

$$[P^3_c]_{r,s,w} \leq \left( \frac{8}{1 - \exp(-\pi/2)} + \pi \right)^2 \sum_{k=1}^{n} c_k.$$  

Proof  We use submultiplicativity of the $\| \cdot \|_{r,s,w}$-norm to obtain

$$\left[ \sum_{k=1}^{n} (c_k/2)(q_k+1 - q_k)^2 \right]_{r,s,w} \leq \sum_{k=1}^{n} [q_k/2]( [q_k+1]_{r,s,w} + [q_k]_{r,s,w})^2. \quad (8.15)$$

We first estimate the size of $q_i$ for $i = 1, \ldots, n$:

$$[q_i]_{r,s,w} = \sup_{\omega \in \mathbb{W}_o} \sum_{k \leq n} d[k] \sup_{\omega \leq n} |q_k|$$

$$\leq \sum_{k \leq n} d[k] 2 e^{-\pi/2}$$

$$\leq 2 \sum_{k \leq n} e^{-\pi/2}$$

$$\leq \frac{4}{1 - e^{-\pi/2}}.$$  

The first step is by definition, the second because $q_i$ depends only on $(I_i; \omega_i)$. Then we use the Paley-Wiener Lemma and the conditions on $r, s, w$.

Now the size of $q_{n+1}$:

$$[q_{n+1}]_{r,s,w} = \left[ \frac{z_1 + \xi_1}{2} \right]_{r,s,w} = \sqrt{2} \leq \pi.$$  

Combining these results finishes the proof. \hfill \qed

Lemma 8.26 (Estimating the size of $P^4_c$) Under the conditions on $(r, s, w)$ we have

$$[P^4_c]_{r,s,w} = 6 \left( \sum_{l=m+1}^{n+m-1} |c_l|^2 \right)^{1/2} s^2.$$  

Proof  We first write $P^4_c$ in a symmetric form:

$$P^4_c(\varphi, I, z, \xi; \omega) = \sum_{k=n+1}^{n+m-1} (c_k/2)(q_k+1 - q_k)^2 \leq \sum_{k,l=n+1}^{n+m-1} q_k c_l q_l / 2, \quad (8.16)$$

where $C$ is the real symmetric tridiagonal matrix

$$C = \begin{pmatrix}
  c_{n+1} & -c_{n+1} & & \\
  -c_{n+1} & c_{n+1} + c_{n+2} & -c_{n+2} & \\
  & \ddots & \ddots & \ddots \\
  & & -c_{n+m-2} & c_{n+m-2} + c_{n+m-1} & -c_{n+m-1} \\
  & & & -c_{n+m-1} & c_{n+m}
\end{pmatrix}. \quad (8.17)$$
8.4. Estimating the size of the perturbation

The expressions in (8.16) do not depend on $(\varphi, I, \omega)$, so there is only $(z, \xi)$ to expand in:

\[
\sum_{k,l=n+1}^{n+m-1} q_k C_{kl} q_l / 2 = \sum_{k,l=n+1}^{n+m-1} (z_k + \xi_k) C_{kl} (z_l + \xi_l) / 4
\]

\[
= \sum_{k,l=n+1}^{n+m-1} (z_k C_{kl} z_l + 2 \xi_k C_{kl} z_l + \xi_k C_{kl} \xi_l) / 4.
\]

The $[\cdot]_{r,s,w}$-norm of the last expression is $|C|^2$, with

\[
|C|^2 = \sum_{k,l=n+1}^{n+m} |C_{kl}|^2 \leq 6 \sum_{k=n+1}^{n+m-1} |\xi_k|^2.
\]  

This yields the bound for the size of $P_c^4$.

Sketch of proof of Proposition 5.3  We show for positive spring constants $c = (c_{n+1}, \ldots, c_{n+m-1})$ that

\[
[P_c^4]_{r,s,w} = [(p^1 + p^2 + p^3) \circ \Phi_c]_{r,s,w}
\]

\[
\leq [p^1 + p^2 + p^3]_{r,s,w}.
\]

Using the preceding results of this section the proposition then follows.

Note that all of the functions $p^1, p^2$ and $p^3$ depend only on $(\varphi, I, q; \omega)$, i.e., not on $p$, and for such functions $F$ we prove that

\[
[F \circ \Phi_c]_{r,s,w} \leq [F]_{r,s,w}.
\]

The transformation $\Phi_c$ is first made explicit:

\[
(\Phi_c)(\varphi, I, q; \omega) = F(\varphi, I, (1 + C_c)^{-1/4} q; \omega).
\]

Since $(1 + C_c)$ is real symmetric it can be written as

\[
(1 + C_c) = T \Lambda T^{-1},
\]

with $T$ an $m \times m$ orthogonal and diagonal matrix. Then $(1 + C_c)^{-1/4}$ can be rewritten to

\[
(1 + C_c)^{-1/4} = T \Lambda^{-1/4} T^{-1}.
\]

Recall from Section 6.3 that the orthogonal and symplectic transformation

\[
\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} T & 0 \\ 0 & T^{-1} \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}
\]

is unitary, and therefore keeps invariant the $[\cdot]_{r,s,w}$ norm (Subsection 8.1.3).

For functions $F$ only depending on $(\varphi, I, q; \omega)$ we therefore only need to find out what the transformation $q \mapsto \Lambda^{-1/4} q$ does to the $[\cdot]_{r,s,w}$ norm. Recall from
Subsection 6.4.2 that for positive spring constants $c = (c_{n+1}, \ldots, c_{n+m-1})$ the eigenvalues $\lambda_1, \ldots, \lambda_m$ are at least equal to one. Consequently, the elements of $\Lambda^{-1/4}$ are positive and smaller than or equal to one. Hence the absolute values of the Fourier-Taylor coefficients of $F \circ \Phi_c$ are also smaller than or equal to those of $F$. From Definition 4.3 it follows that the same holds for the $\| \cdot \|_{r,s,w}$ norm. \hfill $\square$