Proceedings of the Edinburgh Mathematical Society (2002) **45**, 731–759 ©
DOI:10.1017/S001309150000064X Printed in the United Kingdom

PERTURBATIONS OF VECTOR FIELDS ON TORI: RESONANT NORMAL FORMS AND DIOPHANTINE PHENOMENA

DETTA DICKINSON[1], TODOR GRAMCHEV[2] AND MASAFUMI YOSHINO[3]

[1]*Department of Mathematics, Logic House, NUI Maynooth, Co. Kildare, Republic of Ireland* (ddickinson@maths.may.ie)
[2]*Dipartimento di Matematica, Università di Cagliari, 09124 Cagliari, Italia* (todor@mat33.sc.unica.it)
[3]*Department of Mathematics, Chuo University, Tokyo 192-0393, Japan* (yoshinom@tamacc.chuo-u.ac.jp)

(Received 31 May 2000)

Abstract This paper concerns perturbations of smooth vector fields on $\mathbb{T}^n$ (constant if $n \geqslant 3$) with zeroth-order C^∞ and Gevrey G^σ, $\sigma \geqslant 1$, pseudodifferential operators. Simultaneous resonance is introduced and simultaneous resonant normal forms are exhibited (via conjugation with an elliptic pseudodifferential operator) under optimal simultaneous Diophantine conditions outside the resonances. In the C^∞ category the results are complete, while in the Gevrey category the effect of the loss of the Gevrey regularity of the conjugating operators due to Diophantine conditions is encountered. The normal forms are used to study global hypoellipticity in C^∞ and Gevrey G^σ. Finally, the exceptional sets associated with the simultaneous Diophantine conditions are studied. A generalized Hausdorff dimension is used to give precise estimates of the 'size' of different exceptional sets, including some inhomogeneous examples.

Keywords: vector fields on tori; simultaneous resonances; Diophantine conditions; resonant normal forms; global hypoellipticity; generalized Hausdorff dimension

AMS 2000 *Mathematics subject classification:* Primary 37C15; 11J13
Secondary 58J40; 11J20; 35H05

1. Introduction

The purpose of this paper is to study resonant normal forms and the Diophantine phenomena of perturbations of commuting resonant vector fields X_k on the n-dimensional torus $\mathbb{T}^n = \mathbb{R}^n/(2\pi\mathbb{Z})^n$. Here

$$X_k = \langle \omega^k, \partial_x \rangle = \sum_{j=1}^{n} \omega_j^k \partial_{x_j}, \quad 1 \leqslant k \leqslant d,$$

where $x = (x_1, \ldots, x_n) \in \mathbb{T}^n$ and $\omega^k = (\omega_1^k, \ldots, \omega_n^k) \in \mathbb{R}^n$, and $\partial_x = (\partial_{x_1}, \ldots, \partial_{x_n})$, $\partial_{x_j} = \partial/\partial x_j$. If $n = 2$ we also allow the X_k to be non-constant vector fields.

The X_k are perturbed with zeroth-order classical pseudodifferential operators (denoted PDOs hereafter) so that the operators

$$b^k(x,D) := -\mathrm{i}X_k + a^k(x,D), \quad 1 \leqslant k \leqslant d, \tag{1.1}$$

are commuting, where $D = (D_{x_1}, \dots, D_{x_n})$, $D_{x_j} = -\mathrm{i}\partial_{x_j}$, $j = 1, \dots, n$, and $a^k(x,D)$ is a zeroth-order PDO defined by

$$a^k(x,D)f = \sum_{\xi \in \mathbb{Z}^n} \mathrm{e}^{\mathrm{i}x\xi} a^k(x,\xi)\hat{f}(\xi),$$

where

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) = \int_{\mathbb{T}^n} \mathrm{e}^{-\mathrm{i}x\xi} f(x)\,\mathrm{d}x$$

is the Fourier transform on the n-dimensional torus $\mathbb{T}^n$ and $a^k(x,\xi) \in S^0(\mathbb{T}^n \times \mathbb{R}^n)$ is a classical symbol of PDOs. As a special case of (1.1), we also consider the following first-order commuting differential operators:

$$\mathcal{L}_j = \langle \omega^j, D_x \rangle + a_j(x), \quad a_j \in C^\infty(\mathbb{T}^n), \quad j = 1, \dots, d. \tag{1.2}$$

We will actually consider perturbations with functions in the Gevrey class $G^\sigma(\mathbb{T}^n)$. Recall that $G^\sigma(\mathbb{T}^n)$ is the set of all $f \in C^\infty(\mathbb{T}^n)$ such that there exists $C > 0$ satisfying

$$\max_{x \in \mathbb{T}^n} |\partial^\beta f(x)| \leqslant C^{|\beta|+1}(\beta!)^\sigma, \quad \beta = (\beta_1, \dots, \beta_n) \in \mathbb{Z}^n,$$

where $\beta! = \beta_1! \cdots \beta_n!$ and $|\beta| = \beta_1 + \cdots + \beta_n$. Clearly, $G^1(\mathbb{T}^n)$ coincides with the set of all analytic functions on $\mathbb{T}^n$.

For a single vector field (respectively, map) with an isolated singular (respectively, fixed) point the formal reduction to its linear part requires a non-resonance condition on the eigenvalues of the linear part, while in order to show the convergence of formal transformations, arithmetic conditions are imposed (see the surveys [**26**, **40**] and the references therein). If commuting vector fields or maps are considered, it is not necessary for each vector field or map to satisfy these conditions. More precisely, simultaneous arithmetic conditions are required (see [**11**, **14**, **24**, **34**, **39**, **40**]). Certain exceptional sets will arise in connection with some of the parameters in these simultaneous arithmetic conditions (cf. [**18**]) and these will be discussed later.

This paper deals with four closely related problems. First, after introducing the concept of simultaneous (non)resonance for $\omega^1, \dots, \omega^d$, we are interested in the simultaneous resonant normal forms of $\{b^k(x,D)\}_1^d$ and/or $\{\mathcal{L}_k\}_1^d$ via conjugation with an elliptic PDO (which in the case of differential operators reduces to a change of variables on $\mathbb{T}^n$) in the framework of the C^∞ and the Gevrey spaces. Roughly speaking, resonant normal forms mean that lower-order terms depend only on resonant variables. For this purpose

the geometrical properties of a simultaneously resonant set associated with $\omega := \{\omega^k\}_1^d$ and defined by

$$\Gamma_{\mathbb{Z}}^{\omega} := \{z \in \mathbb{Z}^n;\ \langle \omega^1, z\rangle = \cdots = \langle \omega^d, z\rangle = 0\}$$

will be considered. We say that $\omega^1, \ldots, \omega^d$ are *simultaneously non-resonant* (respectively, *resonant*) if $\Gamma_{\mathbb{Z}}^{\omega} = \{0\}$ (respectively, $\Gamma_{\mathbb{Z}}^{\omega} \neq \{0\}$). Using theorems on finitely generated free abelian groups and exact sequences the existence of suitable bases for $\Gamma_{\mathbb{Z}}^{\omega}$ will be obtained and will lead naturally to the introduction of resonant variables on the torus. These constructions are invariant under the action of the group of linear automorphisms of the lattice $\mathbb{Z}^n$. A simultaneous resonant normal form will be exhibited for every perturbed system under simultaneous Diophantine conditions for $\omega^1, \ldots, \omega^k$ on the set $\mathbb{Z}^n \setminus \Gamma_{\mathbb{Z}}^{\omega}$ in the class C^∞ or Gevrey G^σ. In particular, if $\Gamma_{\mathbb{Z}}^{\omega} = \{0\}$, it is possible to transform $\{b^k\}_1^d$ (respectively, $\{\mathcal{L}_k\}_1^d$) simultaneously into a constant PDO (respectively, differential operator). It will also be shown that in general the Diophantine condition is necessary to transform the perturbed systems to their normal forms: constant pseudodifferential operators on $\mathbb{T}^n$. For results on normal forms in dynamical systems by means of KAM-type methods and Diophantine phenomena we refer, for example, to [**1**, **7**, **9**, **13**, **19**, **29**, **35**, **40**] and the references therein.

We then consider the simultaneous reduction of non-constant vector fields to constant ones for the case $n = 2$. There are very few results regarding normal forms of vector fields and differential operators on $\mathbb{T}^n$ except for those vector fields on $\mathbb{T}^2$ of the form $\partial_{x_1} + \lambda(x)\partial_{x_2}$, where the rotation number of the Poincaré map plays an essential role (cf. [**1**]). For $n \geqslant 3$, there is a recent result of Wenyi and Chi [**12**] for the reduction of a smooth vector field $\mathcal{L}$ on $\mathbb{T}^n$ to a constant one provided that the adjoint operator $\mathcal{L}^*$ is globally C^∞ hypoelliptic. We stress that the classical results of Arnold and Moser in KAM theory are not applied to the original vector field, but to a small perturbation of it (cf. [**40**] and the references therein).

In the second part of this paper we estimate the loss of the Gevrey regularity of the conjugating PDO if the lower-order perturbations are Gevrey G^σ. Although we are not able to prove that our Gevrey estimates are best possible, there is some evidence to suggest this from the sharp loss of Sobolev regularity in each inductive step. This phenomenon resembles a similar one in the effective stability (Nekhoroshev estimates) of normal forms in dynamical systems and their applications (cf. [**2**, **21**, **22**, **30**, **33**]; see also [**23**] for Nekhoroshev estimates for billiard ball maps in $\mathbb{R}^n$, $n \geqslant 3$, by means of Gevrey techniques).

In the third part of the paper an application of the resonant normal forms to the study of the global hypoellipticity of commuting systems of pseudodifferential operators is presented. Indeed, assuming $\Gamma_{\mathbb{Z}}^{\omega} = \{0\}$, we show that certain inhomogeneous Diophantine conditions completely characterize the global properties in C^∞ and in G^θ for large values of θ by virtue of the Nekhoroshev-type estimates. The case $\Gamma_{\mathbb{Z}}^{\omega} \neq \{0\}$ is more difficult. A natural extension of completely resonant systems is introduced and it will be shown that the discrete condition is necessary and sufficient for the global hypoellipticity of the perturbations of such systems. We will also consider systems where no discrete phenomena appear.

The final question considered in this paper is that of the 'size' of the exceptional sets of ω associated with G^σ Siegel conditions, for which perturbations of the vector fields X_k are not reducible to normal forms. The inhomogeneous Diophantine conditions also play an essential role when studying the global properties of the reduced operators on the torus. Here the problem is twofold: firstly, ordinary Hausdorff dimension gives little information about the exceptional sets in the Gevrey category and so we use logarithmic Hausdorff dimension; secondly, the inhomogeneity of the Diophantine conditions causes some difficulties.

The paper is organized in the following way. Section 2 deals with the geometry of the resonant set and § 3 is devoted to simultaneous normal forms. The Nekhoroshev-type estimates are shown in § 4 and global hypoellipticity and solvability are discussed in § 5. Finally, estimates for the exceptional sets using a generalized concept of Hausdorff dimension are proved in § 6.

2. Simultaneous resonance

The resonant set $\Gamma^\omega_{\mathbb{Z}}$ will now be investigated. We observe that if $P \in \mathrm{GL}(n;\mathbb{Z})$, where $\mathrm{GL}(n;\mathbb{Z})$ is the set of all $n \times n$ integer matrices with determinant ± 1 (i.e. the group of linear automorphisms of the lattice $\mathbb{Z}^n$), and $y = {}^{\mathrm{tr}}Px$, then $\langle \omega^j, \partial_x\rangle \to \langle P\omega^j, \partial_y\rangle$, which implies the invariance property $P(\Gamma^\omega_{\mathbb{Z}}) = \Gamma^{P\omega}_{\mathbb{Z}}$, $P\omega = \{P\omega^1, \dots, P\omega^k\}$. Also note that if the $\mathbb{R}$-span of $\tilde\omega^1, \dots, \tilde\omega^\ell \in \mathbb{R}^n$ coincides with the $\mathbb{R}$-span of $\omega^1, \dots, \omega^k \in \mathbb{R}^n$, then $\Gamma^\omega_{\mathbb{Z}} = \Gamma^{\tilde\omega}_{\mathbb{Z}}$.

Theorem 2.1. *Assume $\Gamma^\omega_{\mathbb{Z}} \neq \{0\}$. Then the following conditions hold.*

(i) *There exist $\kappa^1, \dots, \kappa^r \in \mathbb{Z}^n \setminus \{0\}$ linearly independent over $\mathbb{R}$ such that the $\{\kappa^\nu\}_{\nu=1}^r$ form a basis of $\Gamma^\omega_{\mathbb{Z}}$, i.e.*

$$\Gamma^\omega_{\mathbb{Z}} = \Big\{ z \in \mathbb{Z}^n;\ z = \sum_{\nu=1}^r t_j \kappa^j,\ \text{with } t_j \in \mathbb{Z} \text{ uniquely determined} \Big\}.$$

The number $r = r(\omega) = \dim_{\mathbb{Z}} \Gamma^\omega_{\mathbb{Z}}$ is invariant under the action of $\mathrm{GL}(n;\mathbb{Z})$, i.e. $r(\omega) = r(P\omega)$ for every $P \in \mathrm{GL}(n;\mathbb{Z})$. It is always possible to find κ^j (after changes of indices $x_\mu \to x_\nu$) to satisfy

$$\kappa^j = (\kappa^j_1, \kappa^j_2, \dots, \kappa^j_n), \quad \kappa^j_j \in \mathbb{N}, \quad \kappa^j_\nu = 0, \quad 1 \leqslant \nu \leqslant j-1. \tag{2.1}$$

Every other basis $\{\tilde\kappa^\nu\}_{\nu=1}^r$ is represented uniquely in the form, $(\tilde\kappa^1, \dots, \tilde\kappa^r) = (\kappa^1, \dots, \kappa^r)Q$, for some $Q \in \mathrm{GL}(r;\mathbb{Z})$.

(ii) *Let $\kappa^1, \dots, \kappa^r$ be a basis of $\Gamma^\omega_{\mathbb{Z}}$. Then there exists $P \in \mathrm{GL}(n;\mathbb{Z})$ such that its first r columns coincide with $\kappa^1, \dots, \kappa^r$, i.e. the basis of $\Gamma^\omega_{\mathbb{Z}}$ is extended into a basis of $\mathbb{Z}^n$. In particular, if $y = (y', y'') = {}^{\mathrm{tr}}Px$ with $y' = (y_1, \dots, y_r)$ and $y'' = (y_{r+1}, \dots, y_n)$, then*

$$P\omega^k = (0', \theta^k), \quad \theta^k \in \mathbb{R}^{n-r}, \quad 1 \leqslant k \leqslant d, \tag{2.2}$$

with $\Gamma_{\mathbb{Z}}^{\theta} = \{0\}$, i.e. $\theta^1, \dots, \theta^k$ are non-resonant in $\mathbb{R}^{n-r}$. In the y-variables the resonant set $\Gamma_{\mathbb{Z}}^{P\omega}$ coincides with $\mathbb{Z}^r_{y'}$, where $\mathbb{Z}^r_{y'}$ is the lattice in $\mathbb{Z}^r$ with basis y'. The set of all matrices P from $\mathrm{GL}(n;\mathbb{Z})$ with such a property is isomorphic to $\mathcal{M}_{(n-r)\times r}(\mathbb{Z}) \oplus \mathrm{GL}(n-r;\mathbb{Z})$, where $\mathcal{M}_{(n-r)\times r}(\mathbb{Z})$ denotes the group of $(n-r)\times r$ integer matrices.

In order to prove the theorem we recall the following well-known assertion (see [**6**] for more general statements).

Lemma 2.2. *Let $A \subset B := \mathbb{Z}^n$ be a subgroup. Set $C = A^{\perp}_{\mathbb{Z}} = \{\xi \in \mathbb{Z}^n : \xi \perp A\}$. Then*

$$B = A \oplus C \tag{2.3}$$

holds if and only if the sequence

$$0 \to A \xrightarrow{\mathrm{id}} B \xrightarrow{\pi} C \to 0 \tag{2.4}$$

is exact (i.e. $\operatorname{Ker}\pi = \mathrm{id}(A)$) where π (respectively, id) stands for the orthogonal projection on C (respectively, the identity map).

Proof of Theorem 2.1. The existence of a basis follows from the well-known fact that any subgroup A (in our case $A = \Gamma^{\omega}_{\mathbb{Z}}$) of a finitely generated abelian group B (in our case $B = \mathbb{Z}^n$) is also a finitely generated abelian group [**6**]. Define $\pi_j : \mathbb{R}^n \to \mathbb{R}^n$ by $\pi_j(\xi) = \langle \xi, \vec{e}_j\rangle \vec{e}_j = \xi_j \vec{e}_j$, where $\vec{e}_j$ is the jth unit vector. Without loss of generality we may assume that for each $j \in \{1, \dots, n\}$ we have $\pi_j(\omega^k) \neq 0$ for some k, for otherwise we are reduced to a space of dimension $n-1$. Because $\Gamma^{\omega}_{\mathbb{Z}}$ is a linear $\mathbb{Z}$-submodule of $\mathbb{Z}^n$, $\pi_1(\Gamma^{\omega}_{\mathbb{Z}})$ is closed under summation, and therefore there exists $p_1 \in \mathbb{N}$ such that $\pi_1(\Gamma^{\omega}_{\mathbb{Z}}) = p_1\mathbb{Z}\vec{e}_1$. Let $\kappa^1 \in \Gamma^{\omega}_{\mathbb{Z}}$ be such that $\pi_1(\kappa^1) = p_1\vec{e}_1$. Let $\Lambda_2 \subset \Gamma^{\omega}_{\mathbb{Z}}$ be the $\mathbb{Z}$-submodule such that $\pi_1(\Lambda_2) = \{0\}$. In fact, Λ_2 is uniquely determined: $\Lambda_2 = \{(\xi_1, \dots, \xi_n) \in \Gamma^{\omega}_{\mathbb{Z}} : \xi_1 = 0\}$. If $\Lambda_2 = \{0\}$, the set $\Gamma^{\omega}_{\mathbb{Z}}$ is generated by κ^1 and $r = 1$. If $\Lambda_2 \neq \{0\}$, this procedure is continued and sequences $\{\kappa^k\} \subset \Gamma^{\omega}_{\mathbb{Z}}$, $\{p_k\} \subset \mathbb{N}$ are constructed together with a nested family of $\mathbb{Z}$-submodules $\{\Lambda_k\}$, $\Lambda_k \subset \Lambda_{k-1}$ such that $\pi_k(\Lambda_k) = p_k\mathbb{Z}\vec{e}_k$, $\pi_k(\kappa^k) = p_k\vec{e}_k$. Clearly, this procedure ends after a finite number of steps. Let $r \leqslant n-1$ be the smallest integer such that $\Lambda_{r+1} = \{0\}$. Then the vectors κ^j form a basis of $\Gamma^{\omega}_{\mathbb{Z}}$ and satisfy (2.1).

Now let $\{\tilde{\kappa}^j\}_1^r \subset \Gamma^{\omega}_{\mathbb{Z}}$ be another basis of $\Gamma^{\omega}_{\mathbb{Z}}$. Then there exist two $r \times r$ integer matrices Q and $\tilde{Q}$ such that $(\kappa^1 \dots \kappa^r)Q = (\tilde{\kappa}^1 \dots \tilde{\kappa}^r)$ and $(\tilde{\kappa}^1 \cdots \tilde{\kappa}^r)\tilde{Q} = (\kappa^1 \cdots \kappa^r)$. Hence $\tilde{Q} = Q^{-1}$ and $\det Q \det \tilde{Q} = 1$. Because $\det Q$ and $\det \tilde{Q}$ are integers it follows that $\det Q = \det \tilde{Q} = \pm 1$, i.e. $Q, \tilde{Q} \in \mathrm{GL}(r;\mathbb{Z})$. Conversely, for any $Q \in \mathrm{GL}(r;\mathbb{Z})$, $(\kappa^1 \dots \kappa^r)Q$ gives a basis of $\Gamma^{\omega}_{\mathbb{Z}}$. This proves (i).

Set $A = \Gamma^{\omega}_{\mathbb{Z}}$ and $C = A^{\perp}_{\mathbb{Z}} = \{\xi \in \mathbb{Z}^n : \xi \perp A\}$. Evidently, $A \subset C^{\perp}_{\mathbb{Z}}$. If we show the exactness of (2.4), namely $\rho \in \mathbb{Z}^n$ and $\rho \perp C$ imply $\rho \in A$, then Lemma 2.2 yields the first assertion of (ii), since by (2.3) we can always extend a basis of A with a basis of C and obtain a basis of $\mathbb{Z}^n$.

Let $\rho \in \mathbb{Z}^n$, $\rho \perp C$. Set $C_{\mathbb{R}}^{\perp} = \{\xi \in \mathbb{R}^n;\ \xi \perp C\}$. Fix a basis as in (2.1). Since $\kappa^j \perp C$ for $j = 1, \dots, r$, and $\dim_{\mathbb{R}} C_{\mathbb{R}}^{\perp} = r$, we obtain that $\kappa^1, \dots, \kappa^r$ is a basis of $C_{\mathbb{R}}^{\perp}$ and ρ is a linear combination of $\kappa^1, \dots, \kappa^r$, implying that $\langle \omega^k, \rho \rangle = 0$ for all $1 \leqslant k \leqslant d$. Since $\rho \in \mathbb{Z}^n$, this implies that $\rho \in A$.

Next, note that $\kappa^1, \dots \kappa^r$ become the first r unit vectors in $\mathbb{R}^n_y$. Thus the invariance of the resonant set implies that $\Gamma_{\mathbb{Z}}^{P\omega} = \{\zeta \in \mathbb{Z}^n : \zeta_{r+1} = \cdots = \zeta_n = 0\}$ and (2.2) is therefore true. If $\theta^1, \dots \theta^{n-r}$ are resonant in $\mathbb{R}^{n-r}_{y''}$, there will exist a vector $\zeta^0 \in \mathbb{Z}^{n-d} \setminus \{0\}$ orthogonal to $\theta^1, \dots, \theta^{n-r}$, so that $(0', \zeta^0) \in \Gamma_{\mathbb{Z}}^{P\omega}$. This is impossible as $\zeta^0 \neq 0''$.

The proof is concluded by observing that $P \in \mathrm{GL}(n; \mathbb{Z})$ preserves the basis $\vec{e}_1, \dots, \vec{e}_r$ of $\Gamma_{\mathbb{Z}}^{P\omega} = \mathbb{Z}^r \times 0$ if and only if

$$P = \begin{pmatrix} I_r & 0_{(n-r)\times r} \\ M & P'' \end{pmatrix}$$

with $P'' \in \mathrm{GL}(n-r; \mathbb{Z})$ and $M \in \mathcal{M}_{(n-r)\times r}(\mathbb{Z})$. The basis $\kappa^1, \dots, \kappa^r$ is called a *canonical basis* of $\Gamma_{\mathbb{Z}}^{\omega}$, and r is a ($\mathbb{Z}$-)dimension of $\Gamma_{\mathbb{Z}}^{\omega}$ over $\mathbb{Z}$, so that we write $\dim_{\mathbb{Z}} \Gamma_{\mathbb{Z}}^{\omega} = r$. We stress that $\mathbb{Z}^n \setminus \Gamma_{\mathbb{Z}}^{\omega}$ is not a linear $\mathbb{Z}$-submodule of $\mathbb{Z}^n$. □

Remark 2.3. Take $A = \{(2t, -2t);\ t \in \mathbb{Z}\}$, $C = A_{\mathbb{Z}}^{\perp} = \{(t, t);\ t \in \mathbb{Z}\}$. Then although A, $B = \mathbb{Z}^2$ and C are finitely generated abelian groups, the decomposition (2.3) does not hold (the sequence (2.4) is not exact because $C_{\mathbb{Z}}^{\perp} \neq A$).

The vectors $\omega^k \in \mathbb{R}^n$ ($1 \leqslant k \leqslant d$) are said to satisfy the *simultaneous resonant σ-Siegel condition* if, for every $\varepsilon > 0$,

$$\liminf_{|\xi| \to \infty,\ \xi \in \mathbb{Z}^n \setminus \Gamma_{\mathbb{Z}}^{\omega}} (\mathrm{e}^{\varepsilon |\xi|^{1/\sigma}} \max_{1 \leqslant k \leqslant d} |\langle \omega^k, \xi \rangle|) > 0, \tag{Si}_{\sigma}$$

for $1 \leqslant \sigma < \infty$, and

$$\liminf_{|\xi| \to \infty,\ \xi \in \mathbb{Z}^n \setminus \Gamma_{\mathbb{Z}}^{\omega}} (|\xi|^{\tau} \max_{1 \leqslant k \leqslant d} |\langle \omega^k, \xi \rangle|) > 0, \tag{Si}_{\infty}$$

when $\sigma = \infty$ for some $\tau \in \mathbb{R}$.

Remark 2.4. Observe that $(\mathrm{Si})_{\sigma}$ and $(\mathrm{Si})_{\infty}$ are invariant under the action of $\mathrm{GL}(n; \mathbb{Z})$. In particular, extending to a basis of $\mathbb{Z}^n$ and in view of (2.2), we get that $(\mathrm{Si})_{\sigma}$ (respectively, $(\mathrm{Si})_{\infty}$) becomes, for each $\varepsilon > 0$,

$$\liminf_{|\xi''| \to \infty,\ \xi'' \in \mathbb{Z}^{n-r}} (\mathrm{e}^{\varepsilon |\xi''|^{1/\sigma}} \max_{1 \leqslant k \leqslant d} |\langle \theta^k, \xi'' \rangle|) > 0 \tag{Si}'_{\sigma}$$

(respectively,

$$\liminf_{|\xi''| \to \infty,\ \xi'' \in \mathbb{Z}^{n-r}} (|\xi''|^{\tau} \max_{1 \leqslant k \leqslant d} |\langle \theta^k, \xi'' \rangle|) > 0, \tag{Si}'_{\infty}$$

for the non-resonant vectors $\theta^1, \dots, \theta^k$ in $\mathbb{R}^{n-k}_{y''}$).

If $d = 1$ and $\omega^1 \in \mathbb{R}^n$ is non-resonant, i.e. $\Gamma^\omega_{\mathbb{Z}} = \{0\}$, then $(\mathrm{Si})_\infty$ (respectively, $(\mathrm{Si})_\sigma$, $1 \leqslant \sigma < \infty$) is the well-known small divisor condition for ω^1 and the number τ in $(\mathrm{Si})_\infty$ is necessarily not less than $n-1$ [**38**] (respectively, coincides with the Gevrey type Siegel condition used in [**25**]; see also [**4**, **37**] for $\sigma = 1$). Furthermore, if $n \geqslant 3$ and $d = n-1$, $(\mathrm{Si})_\infty$ in the non-resonance case is the same as the simultaneous Diophantine condition used in [**5**].

We exhibit examples related to the above notions. In the completely resonant case $r = n-1$, according to Theorem 2.1, after a linear automorphism of the torus $\mathbb{T}^n$ we can reduce the original system to the one consisting of vector fields proportional to ∂_{y_n}. Next, consider two vector fields

$$L_1 = \partial_{x_1} + \alpha_1 \partial_{x_2}, \qquad L_2 = \partial_{x_1} + \alpha_2 \partial_{x_3}$$

on $\mathbb{T}^3$, where α_1 and α_2 are irrational. Then the resonant set of L_j has dimension 1, $j = 1, 2$, and L_j satisfies $(\mathrm{Si})_\sigma$ if and only if α_j is not Liouville, for $\sigma = \infty$, and α_j is not σ-Liouville. Thus for every $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that

$$\left|\alpha_j - \frac{p}{q}\right| \geqslant C_\varepsilon \exp(-\varepsilon q^{1/\sigma}), \quad p \in \mathbb{Z}, \quad q \in \mathbb{N}.$$

Furthermore, the system $\chi = \{L_1, L_2\}$ is simultaneously non-resonant if and only if α_1/α_2 is irrational. Finally, if α_1 and α_2 are rationally independent σ-Liouville numbers, then the simultaneous σ-Siegel condition $(\mathrm{Si})_\sigma$ is satisfied if and only if for $\sigma = \infty$, α_1 and α_2 satisfy the simultaneous Diophantine condition in [**34**] (see also [**5**, **27**]), while for $1 \leqslant \sigma < \infty$ and for every $\varepsilon > 0$ we can find $C_\varepsilon > 0$ such that

$$\max_{j=1,2}\left\{\left|\alpha_j - \frac{p}{q}\right|\right\} \geqslant C_\varepsilon \exp(-\varepsilon q^{1/\sigma}), \quad p \in \mathbb{Z}, \quad q \in \mathbb{N}$$

(cf. [**4**] for $\sigma = 1$, see also [**24**] for other simultaneous Gevrey arithmetic conditions).

Let $C^\infty_{2\pi}(\mathbb{R}^n_x)$ be the space of all 2π-periodic C^∞ functions (i.e. the lift on $\mathbb{R}^n$ of all functions from $C^\infty(\mathbb{T}^n)$), where $\mathbb{R}^n_x$ is the lattice in $\mathbb{R}^n$ with basis x. It is readily verified that this space is invariant under the action of $\mathrm{GL}(n;\mathbb{Z})$, namely $f(y) \in C^\infty_{2\pi}(\mathbb{R}^n_y)$ if and only if $P^* f(x) := f(Px) \in C^\infty_{2\pi}(\mathbb{R}^n_x)$ for some $P \in \mathrm{GL}(n;\mathbb{Z})$. We will identify $C^\infty(\mathbb{T}^n)$ and $G^\sigma(\mathbb{T}^n)$ with $C^\infty_{2\pi}(\mathbb{R}^n)$ and $G^\sigma_{2\pi}(\mathbb{T}^n)$, respectively.

3. Global canonical form of a resonant system

First we recall the basic properties of G^σ pseudodifferential operators (see [**17**, **36**]). In the C^∞ case, the constants in the estimates below may depend on the multi-indices $\alpha, \beta \in \mathbb{Z}^n_+$ and the index of homogeneity j. Let $\mathrm{FS}^m_\sigma = \mathrm{FS}^m_\sigma(\mathbb{T}^n \times \mathbb{R}^n)$ be the set of all formal sums $\sum_{j=0}^\infty a_{m-j}(x,\xi)$ such that $a_{m-j} \in G^\sigma(\mathbb{T}^n \times (\mathbb{R}^n \setminus \{0\}))$, $\operatorname{ord}_\xi a_{m-j} = m-j$ (i.e. a_{m-j} is (positively) homogeneous of order $m-j$) and there exists $A > 0$ satisfying

$$\sup_{x \in \mathbb{T}^n} |\partial^\beta_x \partial^\gamma_\xi a_{m-j}(x,\xi)| \leqslant A^{j+|\gamma|+|\beta|+1} (\beta! j!)^\sigma \gamma! |\xi|^{m-j-|\gamma|} \tag{3.1}$$

for all $\beta \in \mathbb{Z}_+^n$, $\gamma \in \mathbb{Z}_+^n$, $j \in \mathbb{Z}_+$, $\xi \in \mathbb{R}^n$. By Cauchy's integral formula and the homogeneity of a_{m-j}, (3.1) is equivalent to the existence of a neighbourhood $\tilde{S}$ in $\mathbb{C}^n$ of S^{n-1} such that

$$\sup_{x\in\mathbb{T}^n,\ \xi\in\tilde{S}} |\partial_x^\beta a_{m-j}(x,\xi)| \leqslant A^{j+|\beta|+1}(\beta!j!)^\sigma, \quad \beta \in \mathbb{Z}_+^n, \quad j \in \mathbb{Z}_+.$$

A function $a(x,\xi) \in G^\sigma(\mathbb{T}^n \times \mathbb{R}^n)$ is said to be a G^σ *realization* of $\sum_{j=0}^\infty a_{m-j}(x,\xi) \in \mathrm{FS}^m_\sigma$, written $a(x,\xi) \sim \sum_{j=0}^\infty a_{m-j}(x,\xi)$, if for every $\gamma \in \mathbb{Z}_+^n$ one can find $C = C_\gamma > 0$ such that

$$\left|\partial_x^\beta \partial_\xi^\gamma \left(a(x,\xi) - \sum_{j=0}^N a_{m-j}(x,\xi)\right)\right| \leqslant C^{N+|\gamma|+|\beta|+1}(\beta!N!)^\sigma |\xi|^{m-N-|\gamma|-1}$$

for all $\beta \in \mathbb{Z}_+^n$, $N \in \mathbb{Z}_+$, $x \in \mathbb{T}^n$ and $\xi \in \mathbb{R}^n$ with $|\xi| \geqslant 1$. The set of G^σ realizations of order m is denoted by $S^m_\sigma(\mathbb{T}^n \times \mathbb{R}^n)$ and the set of pseudodifferential operators with symbol in $S^m_\sigma(\mathbb{T}^n \times \mathbb{R}^n)$ by $\mathrm{FS}^m_\sigma(\mathbb{T}^n)$. We also recall that for each $\tilde{a} \in \mathrm{FS}^m_\sigma$ there exists a realization $a \in S^m_\sigma$ such that $a \sim \tilde{a}$. If $b(x,\xi)$ is another G^σ realization of $\tilde{a}$, then $R(x,\xi) = a(x,\xi) - b(x,\xi) \in S^{-\infty}_\sigma(\mathbb{T}^n \times \mathbb{R}^n)$, i.e. R is a G^σ smoothing symbol, so that for some $c > 0$

$$\sup_{x\in\mathbb{T}^n} |\partial_x^\beta R(x,\xi)| \leqslant c^{-|\beta|-1}(\beta!)^\sigma \mathrm{e}^{-c|\xi|^{1/\sigma}}, \quad \xi \in \mathbb{R}^n, \quad \beta \in \mathbb{Z}_+^n.$$

Note that the composition rule for PDOs on $\mathbb{T}^n$ is valid as in the local case, that is if

$$a(x,\xi) \sim \sum_{j=0}^\infty a_{\mu-j} \in S^\mu_\sigma(\mathbb{T}^n \times \mathbb{R}^n) \quad \text{and} \quad b(x,\xi) \sim \sum_{j=0}^\infty b_{\nu-j}(x,\xi) \in S^\nu_\sigma(\mathbb{T}^n \times \mathbb{R}^n),$$

then $c(x,D) := a(x,D) \circ b(x,D)$ is in $\mathrm{FS}^{\nu+\mu}_\sigma(\mathbb{T}^n)$, with symbol

$$c(x,\xi) = a(x,\xi) \circ b(x,\xi) \sim \sum_{j=0}^\infty c_{\mu+\nu-j}(x,\xi)$$

given by (see [**17**])

$$c_{\mu+\nu-j}(x,\xi) = \sum_{p+q+|\beta|=j} \frac{1}{\beta!} D_\xi^\beta a_{\mu-p}(x,\xi) \partial_x^\beta b_{\nu-q}(x,\xi), \quad j \in \mathbb{Z}_+.$$

Furthermore, if $a(x,\xi)$ is elliptic, i.e. $a_\mu(x,\xi) \neq 0$ for $(x,\xi) \in \mathbb{T}^n \times (\mathbb{R}^n \setminus 0)$, then $a(x,D)$ admits a (formal) inverse $a^{-1}(x,D) \in \mathrm{FS}^{-\mu}$ satisfying $a \circ a^{-1} - \mathrm{id} \in \mathrm{FS}^{-\infty}_\sigma$ and $a^{-1} \circ a - \mathrm{id} \in \mathrm{FS}^{-\infty}_\sigma$. Recall that for a given $P \in \mathrm{GL}(n;\mathbb{Z})$, the change of variables $y = Px$ induces the linear (symplectic) change $(x,\xi) \to (y,\eta)$, $x = P^{-1}y$ and $\xi = {}^{\mathrm{tr}}P\eta$. We have $(P^*a)(y,\eta) = a(P^{-1}y, {}^{\mathrm{tr}}P\eta)$ for $y \in \mathbb{T}^n$, $\eta \in \mathbb{R}^n$. Thus the composition is invariant, i.e. for two PDOs $a = a(x,D)$ and $b = b(x,D)$ we have

$$P^*(a \circ b) = P^*a \circ P^*b. \tag{3.2}$$

Moreover, if a is a PDO having a symbol depending in x only on resonant variables, the same is true for P^*a.

We denote by $C^\infty(\mathbb{R}^n\backslash\{0\}; G^\sigma(\mathbb{T}^n_x))$ the set of functions $a(x,\xi)$ smooth in $\xi\in\mathbb{R}^n\backslash\{0\}$ and valued in $G^\sigma(\mathbb{T}^n_x)$. Now we state the first main result on simultaneous normal forms.

Theorem 3.1. *Let $1\leqslant\sigma\leqslant\infty$ and suppose that (Si)$_\sigma$ holds for ω. If $r=\dim_{\mathbb{Z}}\Gamma^\omega_{\mathbb{Z}}\geqslant 1$, fix a resonant (canonical) basis $\{\kappa^j\}_1^r$. Assume that if $\sigma<\infty$, then $a^k_{-j}(x,\xi)$ is in $C^\infty(\mathbb{R}^n\setminus\{0\}; G^\sigma(\mathbb{T}^n_x))$ for $k=1,\dots,d$, $j\in\mathbb{Z}_+$. Then there exists an elliptic symbol $q(x,\xi)\sim\sum_{j=0}^\infty q_{-j}(x,\xi)\in \mathrm{FS}^0(\mathbb{T}^n\times\mathbb{R}^n)$ and symbols $\alpha^k(y,\xi)\sim\sum_{j=0}^\infty\alpha^k_{-j}(y,\xi)\in \mathrm{FS}^0(\mathbb{T}^r\times\mathbb{R}^n)$ such that for $1\leqslant k\leqslant d$*

$$(-\mathrm{i}X_k+a^k(x;D))\circ q(x;D)=q(x;D)\circ(-\mathrm{i}X_k+\alpha^k(\kappa x;D)+R_k(x,D)),\tag{3.3}$$

where $R_k(x,\xi)\in S^{-\infty}(\mathbb{T}^n\times\mathbb{R}^n)$ and $\kappa x:=(\langle\kappa^1,x\rangle,\dots,\langle\kappa^r,x\rangle)$.

Proof. Let $P\in\mathrm{GL}(n;\mathbb{Z})$ be as in Theorem 2.1, with the first r rows being $\kappa^1,\dots,\kappa^r$, and define $y=Px$. In view of (3.2), we will consider the new variables (y,η), with $\kappa^j=\vec{e}_j$. The resonant set is the lattice $\mathbb{Z}^r_{y'}$ in $\mathbb{R}^r_{y'}$, where $y=(y',y'')$ and $y'=(y_1,\dots,y_r)$. For brevity, $a(y,\eta)$ is used instead of $P^*a(y,\eta)=a(P^{-1}y,{}^{\mathrm{tr}}P\eta)$ for the symbol $a(x,\xi)$. We write terms of the same order of homogeneity on both sides of (3.3) in the new variables y. The first system of d equations involves both $q_0(y,\eta)$ and $\alpha^k_0(y',\eta)$ $(1\leqslant k\leqslant d)$

$$\langle\theta^k,D_{y''}\rangle q_0+a^k_0(y,\eta)q_0=\alpha^k_0(y',\eta)q_0,\quad \eta\in\mathbb{R}^n\setminus 0.\tag{3.4}$$

The advantage of the expression (3.4) is that we can regard the resonant variables y' as parameters if $r\geqslant 1$. A non-zero solution of (3.4) can be found if for every $k\in\{1,\dots,d\}$, α^k_0 is chosen to satisfy

$$\int_{\mathbb{T}^{n-r}}a^k_0(y',y'',\eta)\,\mathrm{d}y''=\alpha^k_0(y',\eta),\quad y'\in\mathbb{T}^r,\quad \eta\in\mathbb{R}^n\setminus\{0\},\tag{3.5}$$

i.e. the Fourier coefficients $\mathcal{F}_{y\to\zeta}(a^k(y,\eta)-\alpha^k_0(y',\eta))$ vanish on the resonant set, which in the canonical variables is defined by $\zeta_{r+1}=\dots=\zeta_n=0$. By the commutativity we have $\langle\theta^j,D_{y''}\rangle a^k_0(y,\eta)=\langle\theta^k,D_{y''}\rangle a^j_0$ for $1\leqslant j,\,k\leqslant d$. As $\langle\theta^j,D_{y''}\rangle\alpha^k_0\equiv 0$, it follows from the Frobenius Theorem and some simple calculations that the general solution of (3.4) is given by

$$q_0(y,\eta)=\ell(y',\eta)\mathrm{e}^{\mathrm{i}\psi(y,\eta)},\quad \ell(y',\eta)\in G^\sigma(\mathbb{T}^r\times(\mathbb{R}^n\setminus 0)),\quad \operatorname{ord}_\eta\ell=0,$$

where $\psi(y,\eta)$ is the unique solution of the following system of equations with Fourier coefficients vanishing on the resonant set $\Gamma^{P\omega}_{\mathbb{Z}}$:

$$\langle\theta^k,\partial_{y''}\rangle\psi=f^k_0(y,\eta):=-a^k_0(y,\eta)+\alpha^k_0(y',\eta),\quad 1\leqslant k\leqslant d.\tag{3.6}$$

In fact, ψ is defined by

$$\mathcal{F}_{y''\to\zeta''}(\psi(y',y'',\eta))=\frac{\mathcal{F}_{y''\to\zeta''}(f^k_0(y',y'',\eta))}{\mathrm{i}\langle\theta^k,\zeta''\rangle}$$

if $\langle \theta^k, \zeta'' \rangle \neq 0$ for some $1 \leqslant k \leqslant d$, and is 0 otherwise. In view of $(\mathrm{Si})_\sigma$ and the results in [**25**], $\psi(y,\eta)$ (and therefore $q_0(y,\eta)$) is G^σ in x. The composition rule for PDOs yields

$$(q_0^{-1} \circ b^k \circ q_0)(y,\eta) = \langle \theta^k, \eta'' \rangle + \alpha_0^k(y',\eta) + (O(|\eta|^{-1})), \quad |\eta| \to \infty$$

for all $1 \leqslant k \leqslant d$, where b^k stands for $P^*(\mathrm{i}X_k + a^k)$. We proceed by induction: suppose that there exist $\alpha_{-j}^k(y',\eta) \in C^\infty(\mathbb{R}^n \setminus \{0\}; G^\sigma(\mathbb{T}^r_{y'}))$, $\operatorname{ord}_\eta \alpha_{-j}^k(y',\eta) = -j$, $0 \leqslant j \leqslant N-1$, $1 \leqslant k \leqslant d$, such that

$$a^k(y,\eta) \sim \sum_{j=0}^{N-1} \alpha_{-j}^k(y',\eta) + \sum_{j=N}^{\infty} a_{-j}^k(y,\eta), \quad 1 \leqslant k \leqslant d.$$

Now we look for an elliptic symbol $q(y,\eta) \sim 1 + q_{-N}(y,\eta)$, $\operatorname{ord}_\eta q_{-N} = -N$, and $\alpha_{-N}^k(y',\eta) \in C^\infty(\mathbb{R}^n \setminus \{0\}; G^\sigma(\mathbb{T}^r))$ such that

$$(q^{-1} \circ b^k \circ q)(y,\eta) = \langle \theta^k, \eta'' \rangle + \sum_{j=0}^{N} \alpha_{-j}^k(y',\eta) + O(|\eta|^{-N-1}),$$

when $|\eta| \to \infty$ for all $1 \leqslant k \leqslant d$. For $k = 1, \dots, d$ this gives the system of equations

$$-\mathrm{i}\langle \theta^k, \partial_{y''} \rangle q_{-N}(y,\eta) = f_{-N}^k(y,\eta) := \alpha_{-N}^k(y',\eta) - a_{-N}^k(y,\eta), \tag{3.7}$$

for $q_{-N}(y,\eta)$ and $\alpha_{-N}^k(y',\eta)$. In view of (3.5) and (3.6) define α_{-N}^k by

$$\mathcal{F}_{y' \to \zeta'}(\alpha_{-N}^k(y',\eta)) = \mathcal{F}_{y \to \zeta}(a_{-N}^k(y',y'',\eta))|_{\zeta=(\zeta',0)}. \tag{3.8}$$

By the Frobenius Theorem, (3.8) and $(\mathrm{Si})_\sigma$, (3.7) admits a unique solution $q_{-N}(y,\xi)$, if $\mathcal{F}_{y' \to \zeta'}(q_{-N}(y',y'',\eta)) = 0$ for $\zeta' \in \mathbb{Z}^r$. Moreover, in view of $(\mathrm{Si})_\sigma$, $q_{-N}(y,\xi)$ is in $C^\infty(\mathbb{R}^n \setminus \{0\}; G^\sigma(\mathbb{T}^n))$.

Finally, construct an elliptic PDO q by

$$q(y,\eta) = \cdots \circ (1 + q_{-N}(y,\eta)) \circ \cdots \circ (1 + q_{-1}(y,\eta)) \circ q_0(y,\eta). \tag{3.9}$$

The right-hand side of (3.9) defines a formal symbol in $\mathrm{FS}^0(\mathbb{T}^n \times (\mathbb{R}^n \setminus 0))$. Taking any realization of this formal symbol, the above calculations yield (3.3) and the estimates on the remainders. □

Remark 3.2. Let $\omega \in \mathbb{R}^n$ be non-resonant and assume that ω does not satisfy $(\mathrm{Si})_\sigma$ for some $1 \leqslant \sigma \leqslant \infty$. Then we can find $f(x) \in G^\sigma(\mathbb{T}^n)$ such that the pseudodifferential operator with symbol $\langle \omega, \xi \rangle + f(x)\xi|\xi|^{-1}$ cannot be transformed via conjugation (3.3) into a constant PDO.

Indeed, in view of the assumption on ω there exists a positive constant $\delta > 0$ and a sequence $\rho^k \in \mathbb{Z}^n \setminus 0$ $(k = 1, 2, \dots)$ such that

$$\lim_{k \to \infty} \mathrm{e}^{-2\delta|\rho^k|^{1/\sigma}} |\langle \omega, \rho^k \rangle|^{-1} = \infty.$$

We choose a function $f(x)=\sum_\eta \hat f(\eta)\mathrm{e}^{\mathrm{i}\eta x}\in G^\sigma(\mathbb{T}^n)$ such that $\hat f(\eta)=\mathrm{e}^{-\delta|\rho^k|^{1/\sigma}}$ if $\eta=\rho^k$, $\hat f(\eta)=0$ for $\eta\neq\rho^k$. The solution $q_0(x,\xi)$ of the equation $\langle\omega,D_x\rangle q_0=f(x)\xi/|\xi|$ satisfies $|\mathcal{F}_{x\to\eta}(q_0(x,\xi))|\geqslant C\mathrm{e}^{\delta|\eta|^{1/\sigma}}$ for $\eta=\rho^k$, $k\in\mathbb{N}$, which contradicts the fact that $q_0\in G^\sigma$ in x. Thus, our Diophantine conditions are optimal.

Remark 3.3. In Theorem 3.1 perturbations of constant vector fields X_k are considered. If X_k is a variable coefficient vector field, the problem is essentially a global nonlinear one. There follow two examples of a system of overdetermined vector fields being simultaneously transformed into constant vector fields. First, any non-singular hypoelliptic real vector field on $\mathbb{T}^2$ is diffeomorphic to a non-zero multiple of $\partial_t+\rho\partial_x$ with ρ being an irrational non-Liouville number (cf. [**28**]). The second example is an overdetermined system of vector fields, $L=d_t+\omega(t)\wedge\partial_x$, $x\in\mathbb{T}^1$, $t=(t_1,\dots,t_n)\in\mathbb{T}^n$, $\omega(t)=\sum_{j=1}^n\omega_j(t)\,\mathrm{d}t_j$ being a real-valued smooth closed one-form on $\mathbb{T}^n$. The corresponding family of n commuting vector fields associated with L (cf. [**5**]) is given by $L_j=\partial_{t_j}+\omega_j(t)\partial_x$, $1\leqslant j\leqslant n$. Straightforward calculations show the existence of a real-valued function $h(t)$ on $\mathbb{T}^n$ such that $\partial_{t_j}h(t)=\omega_j(t)-c_j$, with c_j being the mean value $[\omega_j(\cdot)]_n$ of ω_j, $1\leqslant j\leqslant n$. Then the family $\{L_j\}_1^n$ is transformed into $\{\partial_{s_j}+c_j\partial_y\}_1^n$ via the diffeomorphism of $\mathbb{T}^{n+1}$: $y=x-h(t)$, $s=t$.

We now study the simultaneous reduction of a family of commuting vector fields with variable coefficients on $\mathbb{T}^2$

$$X_j=\partial_t+h_j(t,x)\partial_x,\quad j=1,\dots,d.$$

The commutativity relations are given by

$$\partial_t(h_k-h_j)=h_k\partial_xh_j-h_j\partial_xh_k,\quad j,k=1,\dots,d. \tag{3.10}$$

Let $\psi_j^t:\mathbb{T}^1\to\mathbb{T}^1$ be the time map of X^j, defined by $\psi_j^t(\xi)=x_j(t;\xi)$, where $x_j(t;\xi)$ is a solution of $\dot x_j=h_j(t,x_j)$, $x_j(0)=\xi$. Clearly, (3.10) is equivalent to $\psi_j^t\circ\psi_k^s=\psi_k^s\circ\psi_j^t$, $s,t\in\mathbb{R}$, $j,k=1,\dots,d$.

Recall that the Poincaré map of X_j, given by $P_j:=\psi_j^{2\pi}:\mathbb{T}^1\to\mathbb{T}^1$, is a diffeomorphism of $\mathbb{T}^1$ and assume the orientation-preserving property of P_j. The rotation number of P_j is denoted by $\rho_j\in\mathbb{R}$ (ρ_j is also called the rotation number of the vector field X_j). Moreover, suppose that there exists a smooth diffeomorphism u simultaneously conjugating the maps $P_1,\dots,P_d$ to the rotations $R_{\rho_1},\dots,R_{\rho_d}$, where $R_{\rho_j}(z)=z+\rho_j$, $z\in\mathbb{T}$ ($j=1,\dots,d$), so that

$$u^{-1}\circ P_j\circ u=R_{\rho_j},\quad j=1,\dots,d. \tag{3.11}$$

Theorem 3.4. *Let $1\leqslant\sigma\leqslant+\infty$. If $d\geqslant2$, assume that the Poincaré maps $P_1,\dots,P_d$ are orientation preserving and that there exists an index $j\in\{1,\dots,d\}$ such that $(2\pi)^{-1}\rho_j$ is irrational. Then if a G^σ diffeomorphism u on $\mathbb{T}^1$ satisfying (3.11) can be found, the map $\Phi:t=s$, $x=\phi(s,y)$, where $\phi(s,y)=x_1(s,u(y-s(2\pi)^{-1}\rho_1))$ if $d=1$ and $\phi(s,y)=x_j(s,u(y-s(2\pi)^{-1}\rho_j))$ if $d\geqslant2$, is a G^σ diffeomorphism of $\mathbb{T}^2$ transforming $\partial_t+h_k(t,x)\partial_x$ to $\partial_s+(2\pi)^{-1}\rho_k\partial_y$ for all $1\leqslant k\leqslant d$.*

Proof. We want to reduce the vector fields X_j to $\partial_t + \rho_j \partial_x$ by the change of variables on $\mathbb{T}^2$, $(t, x) = \Phi(s, y)$ with $t = s$ and $x = \phi(s, y)$ simultaneously. By simple computation it can be seen that $\partial_s + (2\pi)^{-1}\rho_j \partial_y = \partial_t + (\phi_s + \rho_j \phi_y)\partial_x$. Hence it is necessary to solve the equation

$$\phi_s + \rho_j \phi_y = h_j(s, \phi(s, y)), \quad j = 1, \dots, d.$$

By the definition of $x_j(t; \xi)$, we obtain for $\xi = u(y - s(2\pi)^{-1}\rho_j)$, $t = s$,

$$\begin{aligned}\partial_s \phi + (2\pi)^{-1}\rho_j \partial_y \phi &= \dot{x}_j(t, \xi) - (2\pi)^{-1}\rho_j \partial_\xi x_j(t; \xi) u'(y - s(2\pi)^{-1}\rho_j) \\ &\quad + (2\pi)^{-1}\rho_j \partial_\xi x_j(t; \xi) u'(y - s(2\pi)^{-1}\rho_j) \\ &= \dot{x}_j(t; \xi) \\ &= h_j(t, x_j(t; \xi)) = h_j(s, \phi(s, y)).\end{aligned}$$

Hence ϕ transforms $\partial_t + h_j(t, x)\partial_x$ into $\partial_s + (2\pi)^{-1}\rho_j \partial_y$. If $d \geqslant 2$ for $i \neq j$, $\partial_t + h_i(t, x)\partial_x$ is transformed to $\partial_s + g_i(s, y)\partial_y$ for some $g_i \in G^\sigma(\mathbb{T}^2; \mathbb{R})$. The commutativity implies that

$$(\partial_s + (2\pi)^{-1}\rho_j \partial_y) g_i(s, y) = 0, \quad i = 1, \dots, d, \quad i \neq j.$$

Since $(2\pi)^{-1}\rho_j$ is irrational, it is readily verified from the Fourier expansion that the g_i are constants. Because the rotation number is invariantly defined for vector fields on $\mathbb{T}^2$, we have $g_i(s, y) \equiv (2\pi)^{-1}\rho_i$.

In order to prove that the map Φ is a G^σ diffeomorphism on $\mathbb{T}^2$, note that the Jacobian of Φ does not vanish, as it is equal to $\partial_y \phi$. For $\xi = u(y - s(2\pi)^{-1}\rho_j)$ we have $\partial_y \phi(s, y) = \partial_\xi x_j(t; \xi) u'(y - s(2\pi)^{-1}\rho_j)$, and since u is a G^σ diffeomorphism, $u'(y) \neq 0$ for every $y \in \mathbb{R}$. On the other hand, $\theta(t; z) := \partial_z x_j(t; z)$ satisfies

$$\dot{\theta} = \partial_x h_j(t, x_j(t; z))\theta, \quad \theta|_{t=0} = 1,$$

which implies that $\theta(t; \xi) = \partial_\xi x_j(t; \xi) \neq 0$ for all $t \in \mathbb{R}$, $\xi \in \mathbb{R}$. Thus $\partial_y \phi(s, y) \neq 0$ for all $s \in \mathbb{R}$, $y \in \mathbb{R}$, and the map Φ is a G^σ diffeomorphism. In order to complete the proof, we will show that $\phi(s + 2\pi, y) = \phi(s, y)$ and $\phi(s, y + 2\pi) = \phi(s, y)$. The latter relation is easy; for the former one, recall that $x_j(2\pi, u(y)) = u(y + \rho_j)$ by the definition of u. The function $x_j(t + 2\pi, u(y - \rho_j - s\rho_j/2\pi))$ solves $\dot{x}_j = h_j(t, x_j)$ with initial value $\xi = x_j(2\pi, u(y - \rho_j - s\rho_j/2\pi)) = u(y - s\rho_j/2\pi)$. By uniqueness it is therefore equal to $x_j(t, u(y - s\rho_j/2\pi))$. □

Remark 3.5.

(a) Let $d = 1$ and suppose either that $\sigma = 1$ and $\rho/(2\pi) = \rho_1/(2\pi)$ satisfies the Bruno condition [**10**] or that $\sigma = \infty$ and $\rho/(2\pi)$ satisfies the usual Siegel condition. Then by the global reduction theorems due to Herman and Bruno, the circle map $P = P_1$ is conjugate to a rotation by a G^1 or C^∞ diffeomorphism.

(b) Let $d = 1$ and $1 < \sigma < +\infty$. We are not aware of any result on global reduction in G^σ Gevrey classes.

(c) Let $d \geqslant 2$. If $\sigma = +\infty$ or $\sigma = 1$, there is a local simultaneous reduction theorem due to Moser [**34**], under a simultaneous Siegel condition. A global C^∞ simultaneous reduction theorem was given by Katznelson, Kra and Ornstein [**31**].

(d) For $d \geqslant 1$ and $1 \leqslant \sigma < +\infty$, Gramchev and Yoshino [**24**] proved local simultaneous reductions of commuting G^σ circle maps under weak exponential arithmetic conditions.

4. Nekhoroshev-type estimates for G^σ PDO

The next result shows that perturbations with G^σ PDOs are conjugated with G^θ elliptic PDOs for some $\theta > \sigma$ if simultaneous small divisor conditions in the non-resonant case are satisfied. This shows a sharp contrast with the case of commuting diffeomorphisms, where even weaker arithmetic conditions are sufficient.

Theorem 4.1. *Let $\Gamma^\omega_{\mathbb{Z}} = \{0\}$ and let $(Si)_\infty$ in § 2 hold for some $\tau \leqslant -1$. Let $\theta = \max\{1, 1+\tau\}$. Assume that $a^k(x, D)$ $(1 \leqslant k \leqslant d)$ is a G^σ PDO. Then there exist symbols $q_{-j}(x, \xi)$ of the conjugating PDO with the following properties: there exists a neighbourhood $\tilde{S} \subset \mathbb{C}^n$ of the unit sphere S^{n-1} such that for every $s > n/2$, $\varepsilon > 0$ we can find $M > 0$, $N > 0$ satisfying*

$$\sup_{\xi \in \tilde{S}} \|\partial_x^\gamma q_{-j}(\cdot, \xi)\|_{H^s} \leqslant \varepsilon M^j N^{|\gamma|}((\theta j + |\gamma|)!)^\sigma, \quad \gamma \in \mathbb{Z}^n_+, \quad j \in \mathbb{N}, \tag{4.1}$$

$$\sup_{\xi \in \tilde{S}} |\alpha^k_{-j}(\xi)| \leqslant \varepsilon M^j ((\theta j)!)^\sigma, \quad k = 1, \dots, d, \quad j \in \mathbb{N}, \tag{4.2}$$

where $\|\cdot\|_{H^s}$ denotes the usual Sobolev norm. It follows that

$$\sum_{j=0}^\infty q_{-j}(x, \xi) \in \mathrm{FS}^0_{\theta\sigma} \quad \text{and} \quad \sum_{j=0}^\infty \alpha_{-j}(\xi) \in \mathrm{FS}^0(\mathbb{R}^n \setminus \{0\}).$$

Set $\theta' = \theta$ if $\sigma > 1$ or $\tau > 0$, and $\theta' = 1 + \rho$, $0 < \rho \ll 1$ if $\theta = 1$ (i.e. $\sigma = 1$ and $\tau \leqslant 0$). Then there exist $G^{\theta\sigma}$ realizations $q(x, \xi)$, $\alpha^k(\xi)$, $1 \leqslant k \leqslant d$, and positive constants C, δ such that

$$\sup_{x \in \mathbb{T}^n} |\partial_x^\gamma \partial_\xi^\beta R_k(x, \xi)| \leqslant C_2^{|\beta|+|\gamma|} (\gamma!)^\sigma (\beta!)^{\sigma\theta'} |\xi|^{-|\gamma|} \exp(-\delta |\xi|^{1/(\sigma\theta)}),$$

for all $\beta, \gamma \in \mathbb{Z}^n_+$, $\xi \in \mathbb{R}^n$ and $1 \leqslant k \leqslant d$. Thus $R_k(x, D) \in \mathrm{FS}^{-\infty}_{\sigma\theta}(\mathbb{T}^n)$, $1 \leqslant k \leqslant d$. Finally, if the symbols $a_j(x, \xi)$ depend on one base variable, the system can be transformed into a simultaneous normal form without any restrictions on τ in the Diophantine condition for $\omega^1, \dots, \omega^d$. In this case there is no loss of regularity in the remainders; more precisely, if $a_k(x, D)$, $1 \leqslant k \leqslant d$, are G^σ PDO, then $R_k(x, D)$ are G^σ smoothing PDOs.

First we need a technical assertion, in which $C^\infty_0(\mathbb{T}^n)$ denotes the set of smooth functions with mean value zero.

Lemma 4.2. *Suppose that $(Si)_\infty$ holds. Then for every $f_k(x) \in C_0^\infty(\mathbb{T}^n)$ such that $X_k f_r = X_r f_k$, $1 \leqslant r, k \leqslant d$, there exists a unique solution $q(x) \in C_0^\infty(\mathbb{T}^n)$ of the system $\langle \omega^k, \partial_x \rangle q(x) = f_k(x)$, $1 \leqslant k \leqslant d$, satisfying $\hat{q}(\xi) = 0$, $\xi \in \Gamma_{\mathbb{Z}}^\omega$. Furthermore, for some $C_\tau > 0$ this unique solution satisfies the estimate*

$$\|q\|_{H^s} \leqslant C_\tau \max_{1 \leqslant k \leqslant d} \|f_k\|_{H^{s+\tau}}, \quad s \geqslant 0.$$

Proof. Using the Fourier transform we obtain $\langle \omega^k, \eta \rangle \hat{q}(\eta) = \hat{f}_k(\eta)$. It follows that $\hat{f}_k(\eta) = 0$ when $\langle \omega^k, \eta \rangle = 0$. The compatibility condition and the definition of $\Gamma_{\mathbb{Z}}^\omega$ implies that we can decompose $\mathbb{Z}^n \setminus \Gamma_{\mathbb{Z}}^\omega = \bigcup_{m=1}^d Z'(m)$, $Z'(m) \bigcap Z'(j) = \emptyset$, $1 \leqslant j < m \leqslant d$, with the property that $\langle \omega^k, \eta \rangle \neq 0$ for $\eta \in Z'(k)$, $1 \leqslant k \leqslant d$. Hence for $1 \leqslant k \leqslant d$ the inverse Fourier transform gives

$$\begin{aligned} \|q\|_{H^s} &\leqslant (2\pi)^{-n/2} \sum_{k=1}^d \Big(\sum_{\eta \in Z'(k)} \frac{|\eta|^{2s}}{|\langle \omega^k, \eta \rangle|^2} |\hat{f}_k(\eta)|^2 \Big)^{1/2} \\ &\leqslant C_\tau \max_{1 \leqslant k \leqslant d} \Big(\sum_{\eta \in \mathbb{Z}^n \setminus 0} \langle \eta \rangle^{2(s+\tau)} |\hat{f}_k(\eta)|^2 \Big)^{1/2} = C_\tau \max_{1 \leqslant k \leqslant d} \|f_k\|_{H^{s+\tau}}, \end{aligned}$$

where $\langle \eta \rangle = (1 + |\eta|^2)^{1/2}$. □

Proof of Theorem 4.1. Suppose that τ is an integer (the proof is the same for $\tau \notin \mathbb{Z}$, but there are some technical complications). Clearly, after conjugation with the symbol $\mathrm{e}^{\mathrm{i}\psi(x,\xi)}$, we may suppose that $a_0^k(x,\xi) = \alpha_0(\xi)$ for $1 \leqslant k \leqslant d$, and hence that $\psi = 0$, $q_0 = 1$ and the conjugating PDO is of the form $q(x,\xi) \sim 1 + \sum_{j=1}^\infty q_{-j}(x,\xi)$. Then, by the composition rule, q_{-j} satisfies the following equations

$$\langle \omega^k, D_x \rangle q_{-j} = F_{-j}^k(x,\xi) + G_{-j}^k(x,\xi) - a_{-j}^k(x,\xi) + \alpha_{-j}^k(\xi) \tag{4.3}$$

for $1 \leqslant k \leqslant d$, where

$$\begin{aligned} F_{-j}^k(x,\xi) &= -\sum_{r=1}^{j-1} \sum_{|\beta| \leqslant j-r} \frac{1}{\beta!} D_\xi^\beta a_{-(j-r-|\beta|)}^k(x,\xi) \partial_x^\beta q_{-r}(x,\xi); \\ G_{-j}^k(x,\xi) &= \sum_{r=1}^{j-1} q_{-r}(x,\xi) \alpha_{-(j-r)}^k(\xi). \end{aligned}$$

Clearly, in order to solve (4.3) we must choose

$$\alpha_{-j}^k(\xi) = -[F_{-j}^k(\cdot,\xi)]_n - [G_{-j}^k(\cdot,\xi)]_n + [a_{-j}^k(\cdot,\xi)]_n, \quad 1 \leqslant k \leqslant d. \tag{4.4}$$

Because $a^k(x,\xi) \in S_\sigma^0(\mathbb{T}^n)$, $1 \leqslant k \leqslant d$, there exists a complex neighbourhood $\tilde{S}$ in $\mathbb{C}^n$ of the unit sphere S^{n-1} in $\mathbb{R}^n$ so that for fixed $s > n/2$ and τ there exists a positive constant A such that

$$\sup_{\xi \in \tilde{S}} \|\partial_x^\beta a_{-j}^k(\cdot,\xi)\|_{H^{s+\tau}} \leqslant A^{j+|\beta|+1} (j!\beta!)^\sigma, \quad j \in \mathbb{Z}_+, \quad \beta \in \mathbb{Z}_+^n. \tag{4.5}$$

By the Schauder Lemma for multiplication in Sobolev spaces, if $s+\tau > n/2$ there exists $\omega > 0$ such that

$$\|fg\|_{H^{s+\tau}} \leqslant \omega \|f\|_{H^{s+\tau}} \|g\|_{H^{s+\tau}}, \quad f, g \in H_0^{s+\tau}(\mathbb{T}^n). \tag{4.6}$$

Estimates (4.1) and (4.2) will be proved by induction. They both hold for $j = 0$. Suppose that they are true for $0 \leqslant j \leqslant i-1$. By (4.6) we obtain for $\xi \in \tilde{S}$ that

$$\begin{aligned}
&\|\partial_x^\gamma F_{-i}^k(\cdot,\xi)\|_{H^{s+\tau}} \\
&\quad \leqslant \omega \sum_{r=1}^{i-1} \sum_{|\beta|\leqslant i-r} \sum_{\gamma'\leqslant\gamma} \binom{\gamma}{\gamma'} \frac{1}{\beta!} \|\partial_x^{\gamma-\gamma'} D_\xi^\beta a^k_{-(i-r-|\beta|)}(\cdot,\xi)\|_{H^{s+\tau}} \|\partial_x^{\beta+\gamma'} q_{-r}(\cdot,\xi)\|_{H^{s+\tau}} \\
&\quad \leqslant \varepsilon\omega A \sum_{r=1}^{i-1} \sum_{|\beta|\leqslant i-r} \sum_{\gamma'\leqslant\gamma} \binom{\gamma}{\gamma'} A^{i-r+|\gamma|-|\gamma'|} ((i-r-|\beta|)!(\gamma-\gamma')!)^\sigma M^r \\
&\qquad \times N^{|\beta|+\tau+|\gamma'|} ((\theta r + |\beta| + |\gamma'| + \tau)!)^\sigma \leqslant \varepsilon M^i N^{|\gamma|} ((\theta i + |\gamma|)!)^\sigma \omega A N^\tau K,
\end{aligned} \tag{4.7}$$

with

$$\begin{aligned}
K &= \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r} \sum_{|\beta|\leqslant i-r} \sum_{\gamma'\leqslant\gamma} \binom{\gamma}{\gamma'} \\
&\qquad \times \left(\frac{A}{N}\right)^{|\gamma|-|\gamma'|} \left(\frac{(i-r-|\beta|)!(\gamma-\gamma')!(\theta r+|\beta|+\tau+|\gamma'|)!}{(\theta i+|\gamma|)!}\right)^\sigma \\
&\leqslant \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r} \sum_{|\beta|\leqslant i-r} \sum_{\gamma'\leqslant\gamma} \frac{(|\gamma'|+1)\dots(|\gamma'|+|\gamma|-|\gamma'|)}{(\gamma-\gamma')!} \\
&\qquad \times \left(\frac{A}{N}\right)^{|\gamma|-|\gamma'|} \left(\frac{(\gamma-\gamma')!}{(|\gamma'|+1)\dots(|\gamma'|+|\gamma|-|\gamma'|)}\right)^\sigma \\
&\qquad \times \left(\frac{(i-r-|\beta|)!(|\gamma'|+1)\dots(|\gamma'|+|\gamma|-|\gamma'|)(\theta r+|\beta|+|\gamma'|+\tau)!}{(\theta i+|\gamma|)!}\right)^\sigma \\
&\leqslant \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r} \sum_{|\beta|\leqslant i-r} \sum_{\gamma'\leqslant\gamma} \left(\frac{A}{N}\right)^{|\gamma|-|\gamma'|} C_1^{\sigma-1} C_2^\sigma,
\end{aligned} \tag{4.8}$$

where

$$C_1 = \frac{(\gamma-\gamma')!}{(|\gamma'|+1)\dots(|\gamma'|+|\gamma|-|\gamma'|)} \leqslant 1,$$

while

$$\begin{aligned}
C_2 &= \frac{(i-r-|\beta|)!(|\gamma'|+1)\dots(|\gamma'|+|\gamma|-|\gamma'|)(\theta r+|\beta|+|\gamma'|+\tau)!}{(\theta i+|\gamma|)!} \\
&\leqslant \frac{(i-r-|\beta|)!(\theta r+|\beta|+|\gamma|+\tau)!}{(\theta i+|\gamma|)!} \leqslant \frac{(\theta i+|\gamma|-(\theta-1)(i-r)+\tau)!}{(\theta i+|\gamma|)!} \leqslant 1
\end{aligned}$$

for all $\gamma, \gamma', \beta \in \mathbb{Z}_+^n$, $\gamma' \leqslant \gamma$, $r, i \in \mathbb{N}$, $i - r \geqslant 1$. We used the inequality $(\theta - 1)(i - r) \geqslant \theta - 1 \geqslant \tau$ if $i - r \geqslant 1$. Choosing first $N > A$ and then $M > 2AN$ large enough so that $AN/M(1 - A/N)^{-n} \sum_{\nu=0}^{\infty} 2^{-\nu}(\nu + 2)^n \leqslant (6C_\tau \omega A N^\tau)^{-1}$, and using the estimates on C_1 and C_2, we get from (4.8) that

$$\begin{aligned} K &\leqslant \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r} \sum_{|\beta| \leqslant i-r} \sum_{\gamma' \leqslant \gamma} \left(\frac{A}{N}\right)^{|\gamma| - |\gamma'|} \\ &\leqslant \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r} \sum_{|\beta| \leqslant i-r} \left(1 - \frac{A}{N}\right)^{-n} \\ &\leqslant \left(\frac{AN}{M}\right) \left(1 - \frac{A}{N}\right)^{-n} \sum_{r=1}^{i-1} \left(\frac{AN}{M}\right)^{i-r-1} (i - r + 1)^n \\ &\leqslant \left(\frac{AN}{M}\right) \left(1 - \frac{A}{N}\right)^{-n} \sum_{\nu=0}^{\infty} 2^{-\nu}(\nu + 2)^n \leqslant (6C_\tau \omega A N^\tau)^{-1}. \end{aligned}$$

Hence, by the choice of M and N,

$$\sup_{\xi \in \tilde{S}} \|\partial_x^\gamma F_{-i}^k(\cdot, \xi)\|_{H^{s+\tau}} \leqslant (3C_\tau)^{-1} \varepsilon M^i N^{|\gamma|} ((\theta i + |\gamma|)!)^\sigma, \quad \gamma \in \mathbb{Z}_+^n. \tag{4.9}$$

Next, for any $\xi \in \tilde{S}$,

$$\begin{aligned} \|\partial_x^\gamma G_{-i}^k(\cdot, \xi)\|_{H^{s+\tau}} &\leqslant \sum_{r=1}^{i-1} \|\partial^\gamma q_{-r}(\cdot, \xi)\|_{H^{s+\tau}} |\alpha_{-(i-r)}^k(\xi)| \\ &\leqslant \varepsilon^2 M^i N^{|\gamma|+\tau} ((\theta i + |\gamma|)!)^\sigma L, \end{aligned}$$

$$\begin{aligned} L &= \sum_{r=1}^{i-1} \left(\frac{(\theta r + |\gamma| + \tau)!(\theta(i - r))!}{(\theta i + |\gamma|)!}\right)^\sigma \\ &= \left(\frac{(\tau + 1)!}{\theta i + |\gamma|}\right)^\sigma \sum_{r=1}^{i-1} \left(\frac{(\tau + 2) \dots (\tau + 1 + \theta r + |\gamma| - 1)}{(\theta(i - r) + 1) \dots (\theta(i - r) + \theta r + |\gamma| - 1)}\right)^\sigma. \end{aligned} \tag{4.10}$$

Since $i - r \geqslant 1$, $r \geqslant 1$ and $\theta \geqslant (1 + \tau)$,

$$\begin{aligned} \tilde{c} &:= \frac{(\tau + 2) \dots (\tau + 1 + \theta r + |\gamma| - 1)}{(\theta(i - r) + 1) \dots (\theta(i - r) + \theta r + |\gamma| - 1)} \\ &\leqslant \frac{\theta \dots (\theta + \theta r + |\gamma| - 1)}{(\theta(i - r) + 1) \dots (\theta(i - r) + \theta r + |\gamma| - 1)} \leqslant 1. \end{aligned} \tag{4.11}$$

It is readily verified from (4.10) that, for all $\gamma \in \mathbb{Z}_+^n$,

$$\begin{aligned} L &\leqslant \left(\frac{(\tau + 1)!}{\theta i + |\gamma|}\right)^\sigma \sum_{r=1}^{i-1} 1 \\ &= \left(\frac{(\tau + 1)!}{\theta i + |\gamma|}\right)^\sigma (i - 1) \leqslant \frac{((\tau + 1)!)^\sigma}{(\theta i + |\gamma|)^{\sigma - 1}} \leqslant ((\tau + 1)!)^\sigma. \end{aligned}$$

Hence, by choosing $\varepsilon > 0$ and $N > 0$ so that $\varepsilon N^\tau((\tau+1)!)^\sigma \leqslant \frac{1}{3}$, we obtain, from (4.1) and (4.2) with $1 \leqslant j \leqslant i-1$,

$$\sup_{\xi\in\tilde{S}} \|\partial_x^\gamma G^k_{-i}(\cdot,\xi)\|_{H^{s+\tau}} \leqslant 3^{-1}\varepsilon M^i N^{|\gamma|}((\theta i+|\gamma|)!)^\sigma, \quad \gamma \in \mathbb{Z}^n_+. \tag{4.12}$$

Finally, observe that (4.5) leads to

$$\sup_{\xi\in\tilde{S}} \|\partial_x^\gamma a^k_{-j}(\cdot,\xi)\|_{H^{s+\tau}} \leqslant 3^{-1}\varepsilon M^j N^{|\gamma|}((\theta j+|\gamma|)!)^\sigma, \quad j \in \mathbb{N}, \quad \gamma \in \mathbb{Z}^n_+, \tag{4.13}$$

provided that $M \geqslant 3A/\varepsilon$ and $N \geqslant A$. By Lemma 4.2, (4.4), (4.12) and (4.13) we have, for any $\xi \in \tilde{S}$,

$$\begin{aligned}
|\alpha^k_{-i}(\xi)| &\leqslant |[F^k_{-i}(\cdot,\xi)]_n| + |[G^k_{-i}(\cdot,\xi)]_n| + |[a^k_{-i}(\cdot,\xi)]_n| \\
&\leqslant \|F^k_{-i}(\cdot,\xi)\|_{L^2} + \|G^k_{-i}(\cdot,\xi)\|_{L^2} + \|a^k_{-i}(\cdot,\xi)\|_{L^2} \\
&\leqslant \|F^k_{-i}(\cdot,\xi)\|_{H^{s+\tau}} + \|G^k_{-i}(\cdot,\xi)\|_{H^{s+\tau}} + \|a^k_{-i}(\cdot,\xi)\|_{H^{s+\tau}} \\
&\leqslant \varepsilon(C_\tau)^{-1}M^i((\theta i)!)^\sigma \leqslant \varepsilon M^i((\theta i)!)^\sigma.
\end{aligned}$$

Thus (4.1) and (4.2) for $1 \leqslant j \leqslant i-1$ imply (4.2) for $j = i$, while (4.1) for $j = i$ follows from Lemma 4.2 applied to $f_k = F^k_{-i}(x,\xi) + G^k_{-i}(x,\xi) - a^k_{-i}(x,\xi) + \alpha^k_{-i}(\xi)$, $1 \leqslant k \leqslant d$, and the fact that the $H^{s+\tau}(\mathbb{T}^n)$ norm of f_k does not depend on α^k_{-i}. □

Remark 4.3. As the proof of Theorem 4.1 shows, our approach resembles methods used for obtaining effective stability (Nekhoroshev-type) estimates (cf. [**2**,**21**–**23**,**29**,**33**]). However, we stress that the study of the Gevrey regularity of the conjugating PDO $q(x,D)$ presents new features and difficulties in comparison with the aforementioned results in dynamical systems. This is due to the presence of compositions of PDOs and the global regularity in $\xi \in \mathbb{R}^n$. In fact, the composition rule of PDOs is a major technical obstacle for getting Gevrey estimates in the presence of resonances. Finally, we point out that our iterative approach shows that the dominating term in $q_j(x,\xi)$, $|\xi| = 1$, is given by $C^{j+1}|D^{j\theta}q_{-1}(x,\xi)| \sim (j!)^\theta$, which suggests that the Gevrey index σ is sharp.

5. Global properties of perturbations of resonant vector fields

In this section we study the global hypoellipticity of first-order overdetermined systems by use of global normal forms.

Let $\mathcal{L}_j$ be defined by (1.2) with $a_j \in G^\sigma(\mathbb{T}^n)$. We say that $\mathcal{L}_j$ is G^σ *globally hypoelliptic* ($1 \leqslant \sigma \leqslant \infty$) if every distribution u such that $\mathcal{L}_j u \in G^\sigma$ ($j = 1,\dots,d$) is in G^σ. Set $r = \dim_{\mathbb{Z}} \Gamma^\omega_{\mathbb{Z}}$ and, for $r \geqslant 1$, take a canonical basis $\eta^1,\dots,\eta^r$ of $\Gamma^\omega_{\mathbb{Z}}$. Suppose that it is possible to find $\alpha_j(y) \in G^\sigma(\mathbb{T}^r)$ ($j = 1,\dots,d$) for which the system

$$\langle\omega^j,\partial_x\rangle\psi = a_j(x) - \alpha_j(\langle\eta^1,x\rangle,\dots,\langle\eta^r,x\rangle), \quad 1 \leqslant j \leqslant d, \tag{5.1}$$

admits a unique solution ψ with $\hat{\psi}(\xi) = 0$ for $\xi \in \Gamma^\omega_{\mathbb{Z}}$. Note that (5.1) is weaker than the simultaneous σ-Siegel-type condition on $\mathbb{Z}^n \setminus \Gamma^\omega_{\mathbb{Z}}$, since if that holds we can solve (5.1) by Fourier series expansion.

If the multiplication operator q is defined by $qv(x) = \mathrm{e}^{\mathrm{i}\psi(x)}v(x)$, the proof of Theorem 3.1 and (5.1) imply that for every $j \in \{1, \dots, d\}$ we have

$$q^{-1} \circ \mathcal{L}_j \circ q = L_j := \langle \omega^j, D_x \rangle + \alpha_j(\langle \eta^1, x\rangle, \dots, \langle \eta^r, x\rangle).$$

Indeed, in (3.7) $a^k_{-N}(x,\xi) = 0$ for $N \geqslant 1$. Therefore, the system $\mathcal{L}_j u = f_j$ $(1 \leqslant j \leqslant d)$ is equivalent to the system

$$L_j v = g_j, \quad 1 \leqslant j \leqslant d, \tag{5.2}$$

where $g_j = \mathrm{e}^{\mathrm{i}\psi(x)} f_j(x)$, $1 \leqslant j \leqslant d$.

We define the change of variables $x = T\tilde{x}$, $\tilde{x} = (y, z)$ with $y = (y_1, \dots, y_r)$ and $z = (z_1, \dots, z_{n-r})$ by $y_j = \langle \eta^j, x\rangle$ $(1 \leqslant j \leqslant r)$, and $z_j = \langle \kappa^j, x\rangle$ $(1 \leqslant j \leqslant n-r)$. Then the system (5.2) becomes

$$\langle \theta^j, D_z\rangle w + \alpha_j(y) w = h_j(y,z), \quad j = 1, \dots, d, \tag{5.3}$$

where the $\theta^j \in \mathbb{R}^{n-r}$ are given by $(0, \theta^j) = {}^{\mathrm{tr}}T\omega^j$, $1 \leqslant j \leqslant d$. The partial Fourier transform $z \to \zeta$ reduces (5.3) to

$$(\langle \theta^j, \zeta\rangle + \alpha_j(y))\hat{w}(y,\zeta) = \hat{h}_j(y,\zeta), \quad \zeta \in \mathbb{Z}^{n-r}, \quad j = 1, \dots, d. \tag{5.4}$$

Define $\Phi(y,\zeta) = \sum_{j=1}^d |\langle \theta^j, \zeta\rangle + \alpha_j(y)|$ and let $d_0 \leqslant n-r$ be the rank of $\theta^1, \dots, \theta^d$. Then we have the following theorem.

Theorem 5.1. *Suppose $\Gamma^\omega_{\mathbb{Z}} \neq \{0\}$ and let $\eta^1, \dots, \eta^r$ be a canonical basis of $\Gamma^\omega_{\mathbb{Z}}$. Assume that (5.1) has a unique solution and that $a^k_{-j}(x,\xi) \in C^\infty(\mathbb{R}^n_\xi \setminus \{0\}; G^\sigma(\mathbb{T}^n_x))$ if $\sigma < \infty$. Then the following conditions hold.*

(i) *The system (5.3) is G^σ globally hypoelliptic if there exists $c > 0$ such that $\Phi(y,\zeta) \geqslant c$ for all $\zeta \in \mathbb{Z}^{n-r}$ and $y \in \mathbb{T}^r$. If $d_0 = n-r$, the preceding condition is replaced by the discrete condition $\Phi(y,\zeta) \neq 0$ for all $\zeta \in \mathbb{Z}^{n-r}$ and all $y \in \mathbb{T}^n$. Note that the completely resonant case $r = n-1$ satisfies $d_0 = n-r$.*

(ii) *Suppose $d_0 = n-r$. Then the system (5.3) is G^σ globally hypoelliptic for $1 < \sigma \leqslant \infty$ if and only if the discrete condition is satisfied.*

Proof. By the representation (5.4), w satisfies $\hat{w}(y;\zeta) = \hat{h}_j(y;\zeta)(\langle \theta^j, \zeta\rangle + \alpha_j(y))^{-1}$ for some $1 \leqslant j \leqslant d$. Hence $w \in G^\sigma$. If $d_0 = n-r$, then $\sum_{j=1}^d |\langle \theta^j, \zeta\rangle| \to \infty$ as $|\zeta| \to \infty$, $\zeta \in \mathbb{Z}^{n-r}$. Hence the condition $\Phi(y,\zeta) \geqslant c$ can be replaced by the discrete condition, proving (i).

(ii) We define the distribution $w(y,z)$ by $\hat{w}(y,\zeta) \equiv 0$ if $\zeta \neq \zeta^0$, and $\hat{w}(y,\zeta^0) = \delta(y-y^0)$, with $y^0 \in \mathbb{T}^r$ and $\zeta^0 \in \mathbb{Z}^{n-r}$ satisfying $\zeta^0 + \alpha_k(y^0) = 0$ for $1 \leqslant k \leqslant d$. Here $\delta(y-y^0)$ stands for the Dirac measure massed at $y = y^0$. The distribution $w(y,z)$ gives a non-smooth zero solution of (5.3). □

Remark 5.2. If $d_0 < n-r$, the condition $\Phi(y,\zeta) \geqslant c$ is not a discrete condition, because the vectors $(\langle \theta^1, \zeta\rangle, \dots, \langle \theta^d, \zeta\rangle)$ $(\zeta \in \mathbb{Z}^{n-r})$ fill a dense subset of a d_0-dimensional

subspace of $\mathbb{R}^d$. Indeed if $d_0 < d$, this actually occurs. Next, if the simultaneous G^σ Siegel condition holds, the solvability of (5.1) is superfluous; that is, instead of the Diophantine conditions on ω^j, the solvability of (5.1) for special $a_j(x)$ is enough. This allows us to characterize larger classes of first-order operators on $\mathbb{T}^2$ with non-resonant $\omega = \omega^1$ than those in Proposition 3.2 of [**3**] ($\sigma = \infty$) and to generalize some of the results for a class of overdetermined systems of vector fields on $\mathbb{T}^n$ in Theorem 3.3 of [**4**] ($\sigma = 1$). Note that the Gevrey spaces $G^\sigma(\mathbb{T}^n)$ are not metrizable if $1 \leqslant \sigma < \infty$, and therefore Baire's Theorem is not applicable.

Now we study the global properties of $\{b^k(x,D)\}$ in the non-resonant case $\Gamma^\omega_{\mathbb{Z}} = \{0\}$. Let a corresponding normal form $\{\omega^k \cdot \xi + \alpha^k(\xi)\}_1^d$ be as in Theorem 3.1 and $1 \leqslant \theta \leqslant \infty$. We say that $\{\omega^k\xi + \alpha^k(\xi)\}_1^d$ satisfies the *simultaneous inhomogeneous θ-Siegel condition* if

$$\liminf_{|\xi|\to\infty,\ \xi\in\mathbb{Z}^n_+\setminus 0} \exp(\varepsilon|\xi|^{1/\theta}) \max_{1\leqslant k\leqslant d} |\omega^k \cdot \xi + \alpha^k(\xi)| > 0, \quad \text{for every } \varepsilon > 0 \tag{TSi}_\theta$$

for $1 \leqslant \theta < \infty$, and

$$\liminf_{|\xi|\to\infty,\ \xi\in\mathbb{Z}^n_+\setminus 0} |\xi|^N \max_{1\leqslant k\leqslant d} |\omega^k \cdot \xi + \alpha^k(\xi)| > 0, \tag{TSi}_\infty$$

for some $N \in \mathbb{R}$ when $\theta = \infty$.

If $\theta = \infty$, $(\mathrm{TSi})_\infty$ is well defined in the sense that it does not depend on realizations of $\sum_{r=0}^\infty \alpha_{-r}(\xi)$. However, if $1 \leqslant \theta < \infty$, the condition $(\mathrm{TSi})_\theta$ is independent of realizations if q is a G^θ PDO. In this case, $\alpha^j(D)$ and the remainders $R^k(x,D)$ are G^θ PDOs as well.

Theorem 5.3. *Assume that ω^j, $j = 1,\dots,d$, are simultaneously non-resonant and satisfy $(\mathrm{Si})_\sigma$ for some $1 \leqslant \sigma \leqslant \infty$. Let the hypotheses of Theorem 3.1 be true. Then $(\mathrm{TSi})_\infty$ holds for one simultaneous canonical form of $\{b_j\}_1^d$ if and only if it holds for every simultaneous canonical form of $\{b^j\}_1^d$. Moreover, $\{b^j\}_1^d$ is globally hypoelliptic on $\mathbb{T}^n$ if and only if a normal form of $\{b^j\}_1^d$ satisfies $(\mathrm{TSi})_\infty$. If $(\mathrm{Si})_\infty$ holds and $a^j(x,D)$, $j = 1,\dots,d$, are G^σ PDOs, then the operator $\{b^j\}_1^d$ is G^θ globally hypoelliptic if and only if a normal form of $\{b^j\}_1^d$ satisfies $(\mathrm{TSi})_\theta$ for $\theta \geqslant \rho\sigma$, where $\rho = \max\{1, 1+\tau\}$.*

Proof. The proof follows from Theorem 3.1 (respectively, the Nekhoroshev-type estimates) if $\theta = \infty$ (respectively, $\theta \geqslant \rho\sigma$) and the following properties: since $q(x,D)$ is elliptic in $G^{\rho\sigma}$, either both or neither systems $\{\omega^j D_x + a^j(x,D)\}$ and $\{\omega^j D_x + \alpha^j(D) + R^j(x,D)\}$ are globally hypoelliptic. The estimates for $q(x,D)$ and the remainders $R^j(x,\xi)$ imply the global G^θ hypoellipticity for $\theta \geqslant \rho\sigma$ if $(\mathrm{TSi})_\theta$ is true. □

6. Logarithmic Hausdorff dimension

For $\boldsymbol{\beta} = (\beta_1,\dots,\beta_d) \in R^d$ and $1 \leqslant \sigma \leqslant \infty$ we define $NG^\sigma(\boldsymbol{\beta};d)$ to be the set of all $\omega = (\omega^1,\dots,\omega^d)$ in $\mathbb{R}^{nd}$ such that $\{\langle\omega^j,\xi\rangle + \beta_j\}_1^d$ does not satisfy $(\mathrm{Si})_\sigma$. Also put $NG^\sigma(d) = NG^\sigma(0;d)$. Clearly, $NG^\sigma(d)$ with $\Gamma^\omega_{\mathbb{Z}} = \{0\}$ is the exceptional set for which $(\mathrm{Si})_\sigma$ fails, and therefore no reduction to normal forms is possible. Similarly, if $\boldsymbol{\beta} \neq 0$, $NG^\sigma(\boldsymbol{\beta};d)$ is related to the exceptional sets of $\{\omega^j\}_1^d$ for which the system $\{L_k + \beta_k\}_1^d$

is not globally G^σ hypoelliptic. The sizes of $NG^\sigma(d)$ and $NG^\sigma(\beta;1)$ will be estimated in terms of the Hausdorff dimension and logarithmic Hausdorff dimension associated with the Gevrey index σ.

For $n > d$ let $E_{\boldsymbol{\beta}}(\sigma) = E_{\boldsymbol{\beta}}(\sigma; n, d)$ be the set

$$E_{\boldsymbol{\beta}}(\sigma) = \{X \in \mathbb{R}^{nd} : |\boldsymbol{q}X - \boldsymbol{\beta}| < \mathrm{e}^{-|\boldsymbol{q}|^{1/\sigma}} \text{ for infinitely many } \boldsymbol{q} \in \mathbb{Z}^n\}.$$

Note that if X is identified with $\omega = (\omega^1, \dots, \omega^d)$, we have

$$E_0(\sigma) \subset NG^\sigma(d) \quad \text{and} \quad E_{\boldsymbol{\beta}}(\sigma) \subset NG^\sigma(\boldsymbol{\beta}; d).$$

Recall that a *δ-cover of a set A* is a union of sets C with diameter less than or equal to δ such that $A \subseteq \cup C$.

Definition 6.1. Let $J \subseteq \mathbb{R}^{nd}$ and define the function $f : \mathbb{R}^+ \to \mathbb{R}$ by $f(r) = r^{(n-1)d}(\log 1/r)^{-s}$. The *logarithmic Hausdorff measure* of J is

$$\mathfrak{L}^s(J) := \lim_{\delta \to 0} \inf_{\mathfrak{C}_\delta} \sum_{C \in \mathfrak{C}_\delta : \cup C \supset J} f(L(C)),$$

where the infimum is over all δ-covers, $\mathfrak{C}_\delta$ of J and $L(C) \leqslant \delta$ is the diameter of C. The *logarithmic Hausdorff dimension* is

$$\mathrm{Ldim}(J) := \sup\{s : \mathfrak{L}^s(J) = \infty\} = \inf\{s : \mathfrak{L}^s(J) = 0\}.$$

Hence $\mathrm{Ldim}(J)$ is the unique value for which $\mathcal{L}^s(J)$ changes from being ∞ to 0. The definition of Hausdorff dimension is the same as above with $f(L(C)) = L(C)^s$; its properties can be found in [**20**]. To denote the information given by the two dimensions, ordinary and logarithmic, the dimension of J will be written as the ordered pair $\vec{\dim} J = (\dim J, \mathrm{Ldim}\, J)$.

For the homogeneous case the following theorem will be proved.

Theorem 6.2. *Let $\sigma \in \mathbb{R}^+$. Then $\vec{\dim} E_0(\sigma; n, d) = ((n-1)d, n\sigma)$.*

By modifying the proof of Theorem 6.2 an inhomogeneous version will also be proved for the case $d = 1$. We will only deal with $d = 1$, as in this case the set can be changed so that the homogeneous result can be used with an appropriate change in the number of variables. For $d > 1$, the set is not a Cartesian product of the homogeneous sets we have already dealt with and so this method cannot be used. As yet we have been unable to obtain the Hausdorff dimension when $d > 1$.

Theorem 6.3. *Let $\sigma \in \mathbb{R}^+$ and $\beta \in \mathbb{R}$. Then $\vec{\dim} E_{\boldsymbol{\beta}}(\sigma; n, 1) = (n-1, n\sigma)$.*

Remark 6.4. Identify $\mathbb{R}^{nd}$ with $(\mathbb{R}^n)^d$ and note that the sets $NG^\sigma(\boldsymbol{\beta}; d)$ are invariant under the action of $\mathrm{GL}(n;\mathbb{Z})$ ($P(NG^\sigma(\boldsymbol{\beta}; d)) = NG^\sigma(\boldsymbol{\beta}; d)$ for all $P \in \mathrm{GL}(n;\mathbb{Z})$). Next, let $\sim$ be an equivalence relation in $NG^\sigma(\boldsymbol{\beta}; d)$ such that $\omega \sim \omega'$ if and only if $\omega = P\omega'$ for some $P \in \mathrm{GL}(n;\mathbb{Z})$ (i.e. they lie in the same orbit), then the space of $\mathrm{GL}(n;\mathbb{Z})$ orbits of $NG^\sigma(\boldsymbol{\beta}; d)/\sim$ has the same $\vec{\dim}$.

Without loss of generality, as the two sets are invariant under translations of integer vectors, we can restrict ourselves to the sets $E_{\boldsymbol{\beta}}(\sigma;n,d)\cap I^{nd}$, where $I=(-\frac{1}{2},\frac{1}{2}]$. The proofs are done in three parts: firstly the Hausdorff dimensions are obtained; then an upper bound for the logarithmic Hausdorff dimension of $E_{\boldsymbol{\beta}}(\sigma;n,d)$ is determined for all $\boldsymbol{\beta}$, n and d; and finally lower bounds for the logarithmic Hausdorff dimensions of the two sets are found.

In [**15**], the Hausdorff dimension of the set

$$W_0(\tau)=W_0(n,d;\tau)=\{X\in\mathbb{R}^{nd}:|\boldsymbol{q}X|<|\boldsymbol{q}|^{-\tau}\text{ for infinitely many }\boldsymbol{q}\in\mathbb{Z}^n\}$$

was shown to be $(n-1)d+n/(\tau+1)$. Plainly, the $(n-1)d$-dimensional hyperplanes $R_{\boldsymbol{q}}=\{X\in I^{nd}:\boldsymbol{q}X=0\}$ are contained in $E_0(\sigma)$ for all $\boldsymbol{q}\in\mathbb{Z}^n$ so that $\dim E_0(\sigma)\geqslant(n-1)d$. Also $E_0(\sigma)\subset W_0(\tau)$, which implies that $\dim E_0(\sigma)\leqslant\dim W_0(\tau)$ for all τ and all σ. Thus $\dim E_0(\sigma)=(n-1)d$. Similarly, in the inhomogeneous case the $(n-1)$-dimensional hyperplanes $R_{\beta,\boldsymbol{q}}=\{\boldsymbol{x}\in I^n:\boldsymbol{q}\cdot\boldsymbol{x}+\beta=0\}$ are contained in $E_{\boldsymbol{\beta}}(\sigma;n,1)$, which implies that $\dim E_{\boldsymbol{\beta}}(\sigma;n,1)\geqslant n-1$. Using the upper-bound argument in the next section with $f(L(C))$ replaced by $L(C)^s$, it is readily verified that for any $s>n-1$ the Hausdorff s-measure of $E_{\boldsymbol{\beta}}(\sigma;n,1)$ is zero. This implies that $\dim E_{\boldsymbol{\beta}}(\sigma;n,1)\leqslant n-1$, completing the result.

Now we deal with the logarithmic component of $\vec{\dim}$. In what follows $a\ll b$ (respectively, $a\gg b$) means that there exists a constant $c>0$ such that $a\leqslant cb$ ($a\geqslant cb$). If $a\ll b$ and $a\gg b$, then $a\asymp b$.

First the upper bounds are obtained for both theorems.

Lemma 6.5. *Let $n,d\in\mathbb{N}$, $n\geqslant 2$, $\sigma\in\mathbb{R}^+$ and $\boldsymbol{\beta}\in\mathbb{R}$. Then*

$$\operatorname{Ldim}E_{\boldsymbol{\beta}}(\sigma)\leqslant n\sigma.$$

Proof. For future reference note that

$$f(\mathrm{e}^{-r^{1/\sigma}}/r)=\frac{(\mathrm{e}^{-r^{1/\sigma}})^{(n-1)d}}{r^{(n-1)d}(r^{s/\sigma}+\log r)}\asymp\frac{(\mathrm{e}^{-r^{1/\sigma}})^{(n-1)d}}{r^{(n-1)d+s/\sigma}}.\tag{6.1}$$

Define $R_{\boldsymbol{\beta},\boldsymbol{q}}$ as $R_{\boldsymbol{\beta},\boldsymbol{q}}=\{X\in I^{nd}:\boldsymbol{q}X+\boldsymbol{\beta}=0\}$ and let $C(\boldsymbol{q})$ denote a collection of hypercubes C with sides of length $L(C)=2n^{1/2}\mathrm{e}^{-|\boldsymbol{q}|^{1/\sigma}}/|\boldsymbol{q}|$, centred on $R_{\boldsymbol{\beta},\boldsymbol{q}}$ with centres at a distance $\mathrm{e}^{-|\boldsymbol{q}|^{1/\sigma}}/|\boldsymbol{q}|$ apart. There are $\ll(|\boldsymbol{q}|e^{|\boldsymbol{q}|^{1/\sigma}})^{(n-1)d}$ such hypercubes. Let C_N denote the set $C_N=\{C(\boldsymbol{q}):|\boldsymbol{q}|>N\}$. For each N this set is a cover of $E_{\boldsymbol{\beta}}(\sigma)$. Hence, using (6.1),

$$\begin{aligned}\mathfrak{L}^s(E_{\boldsymbol{\beta}}(\sigma))&\leqslant\sum_{r>N}\sum_{|\boldsymbol{q}|=r}\sum_{C\in C(\boldsymbol{q})}f(L(C))\\&\ll\sum_{r>N}\sum_{|\boldsymbol{q}|=r}(|\boldsymbol{q}|\mathrm{e}^{|\boldsymbol{q}|^{1/\sigma}})^{(n-1)d}\frac{(\mathrm{e}^{-|\boldsymbol{q}|^{1/\sigma}})^{(n-1)d}}{|\boldsymbol{q}|^{(n-1)d+s/\sigma}}\\&\ll\sum_{r>N}r^{(n-1)-s/\sigma}<\varepsilon\end{aligned}$$

for $s > n\sigma$ and for N sufficiently large. Therefore, $\mathfrak{L}^s(E_{\boldsymbol{\beta}}(\sigma)) = 0$ for $s > n\sigma$ and $\operatorname{Ldim} E_{\boldsymbol{\beta}}(\sigma) \leqslant n\sigma$. □

Proof of Theorem 6.2. We now obtain a lower bound for $\operatorname{Ldim} E_0(\sigma; n, d)$. First, a Cantor subset $\mathcal{K}$ of $E_0(\sigma; n, d)$ and a probability measure supported on $\mathcal{K}$ are constructed. Then the Mass Distribution Principle is used to show that the logarithmic Hausdorff measure of $\mathcal{K}$ is infinity for $s < n\sigma$. This implies that the logarithmic Hausdorff dimension of $\mathcal{K}$ and hence of $E_0(\sigma; n, d)$ is at least $n\sigma$.

For the rest of the proof it is assumed that $s < n\sigma$. Also note that the following inequality will continually be used:

$$\frac{N^{n/d}}{16n\log N} \leqslant \left[\frac{N^{n/d}}{8n\log N}\right] \leqslant \frac{N^{n/d}}{8n\log N}, \tag{6.2}$$

where $[\cdot]$ denotes integer part.

Similar Cantor constructions to the following can be found in [**8**, p. 347] and [**16**] (for general dimension functions). Choose N_1 large enough so that

$$\mathrm{e}^{N_1^{1/\sigma}} > \frac{5N_1^{n/d-1}}{2\log N_1}, \tag{6.3}$$

$$\frac{15N_1^{n-s/\sigma}}{5n^d 2^{4nd+4}\log^d N_1} > 1. \tag{6.4}$$

Starting in one corner, divide I^{nd} into $T(N_1) = [N_1^{n/d}/8n\log N_1]^{nd}$ hypercubes C with

$$L(C) = L_{N_1} = [N_1^{n/d}/8n\log(N_1)]^{-1}.$$

Define C' to be a hypercube with the same centre as C and $L(C') = \frac{1}{2}L(C)$. Let $R_N = \{R_{\boldsymbol{q}} : N \leqslant |\boldsymbol{q}| \leqslant 2N\}$ and let $G(N_1)$ denote the set of cubes C for which there exists an $R_{\boldsymbol{q}} \in R_{N_1}$ such that $C' \cap R_{\boldsymbol{q}} \geqslant (\frac{1}{10}L(C))^{(n-1)d}$; i.e. the intersection $C \cap R_{\boldsymbol{q}}$ is relatively large and $R_{\boldsymbol{q}}$ does not just 'clip' one corner. From Lemma 2 in [**15**] it can be verified that for any $\varepsilon > 0$ and N_1 sufficiently large

$$\#\mathrm{G}(N_1) \geqslant \left[\frac{N_1^{n/d}}{8n\log N_1}\right]^{nd}(1-\varepsilon). \tag{6.5}$$

For each $C \in \mathrm{G}(N_1)$ pick one such intersection $C' \cap R_{\boldsymbol{q}}$ and call it $I(C)$. Now draw smaller hypercubes centred on $I(C)$ and contained in C' of sidelength $\mathrm{e}^{-N_1^{1/\sigma}}/N_1$ with centres at a distance $2\mathrm{e}^{-N_1^{1/\sigma}}/N_1$ apart, so that they are disjoint (see Figure 1 for $n = 2$).

Let $\#I(C)$ denote the number of these small cubes lying on the segment $I(C)$. Then, using (6.3) it can be shown that

$$\begin{aligned}(\tfrac{1}{40}N_1L(C)\mathrm{e}^{N_1^{1/\sigma}})^{(n-1)d} &\leqslant (\tfrac{1}{20}L(C)N_1\mathrm{e}^{N_1^{1/\sigma}} - 1)^{(n-1)d}\\ &\leqslant \#I(C) \leqslant (N_1\mathrm{e}^{N_1^{1/\sigma}}L(C))^{(n-1)d}.\end{aligned}$$

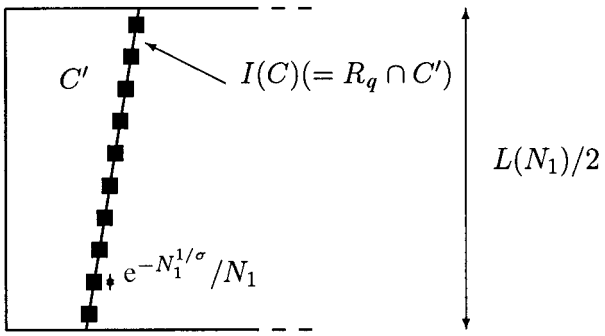

Figure 1.

The union of these cubes over all $C \in G(N_1)$ forms the first level K_1 of the Cantor set and $\#K_1$ denotes the number of them. From (6.2), (6.5) (with $\varepsilon = \frac{1}{16}$) and above

$$\begin{aligned}\#K_1 &\geqslant \frac{15}{16}\left[\frac{N_1^{n/d}}{8n\log N_1}\right]^{nd}(\tfrac{1}{40}N_1 L(C)\mathrm{e}^{N_1^{1/\sigma}})^{(n-1)d}\\ &\geqslant \frac{15N_1^{nd-d+n}(\mathrm{e}^{N_1^{1/\sigma}})^{(n-1)d}}{2^{nd+4}(8n)^d 40^{(n-1)d}(\log N_1)^d}.\end{aligned}$$

Let E_1 denote a hypercube in K_1, so that $L(E_1) = \mathrm{e}^{-N_1^{1/\sigma}}/N_1$. Then (6.4) yields

$$\#K_1 f(L(E_1)) \geqslant \frac{15N_1^{n-s/\sigma}}{5n^d 2^{4nd+4}\log^d N_1} > 1.$$

This will be needed later to apply the Mass Distribution Principle.

To obtain the second level K_2, almost the same construction is made, but instead of I^{nd}, the cubes coming from the previous level K_1 are used. Choose N_2 large enough so that

$$\frac{N_2^{n/d}}{16n\log N_2} > N_1\mathrm{e}^{N_1^{1/\sigma}}, \tag{6.6}$$

$$\frac{3N_2^{n-s/\sigma}}{5n^d\cdot 2^{5nd+2}\log^d N_2} > \frac{\mathrm{e}^{ndN_1^{1/\sigma}}N_1^{nd}}{\#K_1}, \tag{6.7}$$

and $\varepsilon > 0$ such that

$$\varepsilon \leqslant 2^{-1}e^{-2N_1^{1/\sigma}}N_1^{-nd}. \tag{6.8}$$

Divide I^{nd} into $T(N_2)$ hypercubes with sides of length $L_{N_2} = [N_2^{n/d}/8n\log N_2]^{-1}$ as before. Since $N_2 > N_1$ from (6.2),

$$\#\mathrm{G}(N_2) \geqslant \left[\frac{N_2^{n/d}}{8n\log N_2}\right]^{nd}(1-\varepsilon) > \frac{N_2^{n^2}}{(16n)^{nd}\log^{nd}N_2}(1-\varepsilon).$$

Let $T_{E_1}(N_2)$ denote the number of hypercubes contained in E_1, and $\#\mathrm{G}_{E_1}(N_2)$ be the number of those which have large intersection with a resonant set $R_{\boldsymbol{q}} \in R_{N_2}$. It is readily

verified that

$$\left(\frac{\mathrm{e}^{-N_1^{1/\sigma}}}{N_1 L_{N_2}} - 1\right)^{nd} \leqslant T_{E_1}(N_2) \leqslant \left(\frac{\mathrm{e}^{-N_1^{1/\sigma}}}{N_1 L_{N_2}} + 2\right)^{nd},$$

and (6.6) implies that

$$\left(\frac{\mathrm{e}^{-N_1^{1/\sigma}}}{2N_1 L_{N_2}}\right)^{nd} \leqslant T_{E_1}(N_2) \leqslant \left(\frac{2\mathrm{e}^{-N_1^{1/\sigma}}}{N_1 L_{N_2}}\right)^{nd}.$$

Equations (6.8) and (6.5) can be used to show that $\#\mathrm{G}(N_2) \geqslant T(N_2) - T_{E_1}(N_2)/4$, whence

$$\#\mathrm{G}_{E_1}(N_2) > \frac{3T_{E_1}(N_2)}{4} > \frac{3\mathrm{e}^{-ndN_1^{1/\sigma}}}{4 \cdot 2^{nd}(16n)^{nd} N_1^{nd} N_2^{-d^2} \log^{nd} N_2}.$$

Do exactly the same as before with the segments $I(C)$ to construct the second level of the Cantor set K_2. Let $\#H_2$ be the number of hypercubes of K_2 in one E_1, so that $\#K_2 = \#K_1 \#H_2$. Then

$$\#H_2 \geqslant \#\mathrm{G}_{E_1}(N_2)(L_{N_2} N_2 \mathrm{e}^{N_2^{1/\sigma}}/40)^{(n-1)d}$$

and

$$\#H_2 \leqslant \#\mathrm{G}_{E_1}(N_2)(L_{N_2} N_2 \mathrm{e}^{N_2^{1/\sigma}})^{(n-1)d}.$$

Thus

$$\#K_2 \geqslant \#K_1 \frac{3N_2^{nd-d+n}\mathrm{e}^{(n-1)dN_2^{1/\sigma}}}{4 \cdot 2^{2nd}(8n)^n \cdot 40^{(n-1)d} \log^d N_2} \frac{\mathrm{e}^{-ndN_1^{1/\sigma}}}{N_1^{nd}}$$

and, from (6.7),

$$\#K_2 f(L(E_2)) = \frac{3N_2^{n-s/\sigma}}{5n^d \cdot 2^{5nd+2} \log^d N_2} \frac{\mathrm{e}^{-ndN_1^{1/\sigma}}}{N_1^{nd}} \#K_1 > 1.$$

A similar process is used for the rth level. Assume that K_{r-1} has been constructed with hypercubes E_{r-1} of sidelength $\mathrm{e}^{-N_{r-1}^{1/\sigma}}/N_{r-1}$. Choose N_r sufficiently large so that

$$\frac{N_r^{n/d}}{16n \log N_r} > N_r \mathrm{e}^{N_{r-1}^{1/\sigma}}, \tag{6.9}$$

$$\frac{3N_r^{n-s/\sigma}}{5n^d \cdot 2^{5nd+2} \log^d N_r} > \frac{\mathrm{e}^{ndN_{r-1}^{1/\sigma}} N_{r-1}^{nd}}{\#K_{r-1}}, \tag{6.10}$$

and $\varepsilon > 0$ so that

$$\varepsilon \leqslant 2^{-1}\mathrm{e}^{-2N_{r-1}^{1/\sigma}} N_{r-1}^{-nd}. \tag{6.11}$$

Let $T_{E_{r-1}}(N_r)$ denote the number of hypercubes of the rth level contained in E_{r-1} (a cube in K_{r-1}) and $\#\mathrm{G}_{E_{r-1}}(N_r)$ the number which have large intersection with a resonant

set $R_{\boldsymbol{q}} \in R_{N_{r-1}}$. Following the argument as for the second level but using equations (6.9) and (6.11) rather than equations (6.6) and (6.8) we get

$$\#K_r \geqslant \#K_{r-1} \frac{3N_r^{nd-d+n} \mathrm{e}^{(n-1)dN_r^{1/\sigma}}}{4 \cdot 2^{2nd}(8n)^d 40^{(n-1)d} \log^d N_r} \frac{\mathrm{e}^{-ndN_{r-1}^{1/\sigma}}}{N_{r-1}^{nd}}, \tag{6.12}$$

and

$$\#K_r f(L(E_r)) = \frac{3N_2^{n-s/\sigma}}{5n^d \cdot 2^{5nd+2} \log^d N_r} \frac{\mathrm{e}^{-ndN_{r-1}^{1/\sigma}}}{N_{r-1}^{nd}} \#K_{r-1} > 1, \tag{6.13}$$

from (6.10).

Let $\mathcal{K} = \bigcap_{r=1}^{\infty} K_r$. It can be readily verified that $\mathcal{K} \subset E_0(\sigma; n, d)$. Now we proceed to construct a probability measure supported on $\mathcal{K}$. In order to do this the following lemma from [**20**, p. 55] is needed.

Lemma 6.6 (the Mass Distribution Principle). *Let μ be a probability measure supported on a subset F of $\mathbb{R}^k$. Suppose there are positive constants λ and δ such that $\mu(C) \leqslant \lambda f(L(C))$, for any cube C with $L(C) \leqslant \delta$. Then $\mathfrak{L}^s(F) \geqslant \lambda^{-1}$.*

Recall that E_i represents a cube in K_i that is in the ith level of $\mathcal{K}$ and define a probability measure on $\mathcal{K}$ in the following way: $\mu(E_i) = (\#K_i)^{-1} < f(L(E_i))$ from (6.13), for $i \in \mathbb{N}$. Thus the sum of the measures of all hypercubes in the rth level in one fixed cube on the $(r-1)$th level is $\mu(E_{r-1})$, i.e. $\mu(E_r) = \mu(E_{r-1})/(\#H_r)$. To use Lemma 6.6, we need to show that the measure of an arbitrary cube B is also $\ll f(L(B))$. Without loss of generality we can choose an arbitrary cube B contained in some hypercube E_{r-1} and containing at least two cubes of the rth level so that $L(B) \leqslant \mathrm{e}^{-N_{r-1}^{1/\sigma}}/N_{r-1}$. (If B never contains two hypercubes of any level, then its measure is zero.) There are two cases to consider. The first is when B intersects hypercubes lying on at least two different segments $I(C)$. In this case $L(B) \geqslant L_{N_r} = [N_r^{n/d}/8n \log N_r]^{-1}$. Let S be the number of segments $I(C)$ that B intersects. Then $S \leqslant (L(B)/L_{N_r})^{nd}$. The maximum number of hypercubes E_r of K_r that B can intersect on each segment is $(L_{N_r} N_r e^{(N_r^{1/\sigma})})^{(n-1)d}$. Evidently,

$$\mu(B) \leqslant \frac{S(L_{N_r} N_r \mathrm{e}^{N_r^{1/\sigma}})^{(n-1)d}}{\#K_r} \leqslant \frac{L(B)^{nd} N_r^{nd-d+n} \mathrm{e}^{(n-1)dN_r^{1/\sigma}}}{\log^d N_r \#K_r}.$$

From (6.12)

$$\mu(B) \ll \frac{L(B)^{nd} \mathrm{e}^{ndN_{r-1}^{1/\sigma}} N_{r-1}^{nd}}{\#K_{r-1}}.$$

However, $\#K_{r-1} \geqslant 1/f(L(E_{r-1}))$, which implies that

$$\mu(B) \ll L(B)^{nd} \mathrm{e}^{ndN_{r-1}^{1/\sigma}} N_{r-1}^{nd} f(L(E_{r-1})).$$

As the function $f(r)/r^{nd}$ decreases for $r < 1$ and since

$$L(B) \leqslant \mathrm{e}^{-N_{r-1}^{1/\sigma}}/N_{r-1} < 1,$$

we obtain $\mu(B) \ll f(L(B))$.

The second case is when B intersects cubes in K_r from only one segment $I(C)$. As before $L(B) \leqslant \mathrm{e}^{-N_{r-1}^{1/\sigma}}/N_{r-1}$, and $L(B) \geqslant L(E_r) = \mathrm{e}^{-N_r^{1/\sigma}}/N_r$. The number of cubes that B can intersect is $\ll (L(B)/L(E_r))^{(n-1)d}$. Therefore,

$$\mu(B) \ll \frac{L(B)^{(n-1)d}}{L(E_r)^{(n-1)d}\#K_r} \leqslant \frac{L(B)^{(n-1)d} f(L(E_r))}{L(E_r)^{(n-1)d}}.$$

As $f(r)/r^{(n-1)d} = 1/\log^s(1/r)$ is an increasing function and $L(B) \geqslant L(E_r)$, it is easy to verify that $\mu(B) \ll f(L(B))$.

Thus from Lemma 6.6 (the Mass Distribution Principle) for $s < n\sigma$ we have $\mathfrak{L}^s(\mathcal{K}) \geqslant 1$, which implies that $\mathrm{Ldim}(E_0(\sigma; n, d)) \geqslant n\sigma$ and completes the proof of Theorem 6.2. □

Proof of Theorem 6.3. Only the lower bound is left to prove. Let $I_\delta^m = \{\boldsymbol{x} \in I^n : |\boldsymbol{x}| > \delta\}$ and let $\|\cdot\|$ denote the distance from the nearest integer, i.e. $\|\boldsymbol{x}\| = \min_{\boldsymbol{k}\in\mathbb{Z}^2} |\boldsymbol{x} - \boldsymbol{k}|$. The following lemma is needed.

Lemma 6.7. *For every $N > 2$ and almost all $\boldsymbol{x} \in I_\delta^n$ there exists $\boldsymbol{q} \in \mathbb{Z}^n \backslash \{\boldsymbol{0}\}$ with $N/\log^4 N \leqslant |\boldsymbol{q}| \leqslant N$ such that* $\mathrm{dist}(\boldsymbol{x}, R_{\beta,\boldsymbol{q}}) < N^{-2}\log^7 N$.

Proof. From [**32**, pp. 74, 75] it can be verified that for every $N > 2$ and almost all $\boldsymbol{x} \in I^n$ there exists $\boldsymbol{q} \in \mathbb{Z}^n \backslash \{\boldsymbol{0}\}$ with $N/\log^4 N \leqslant |\boldsymbol{q}| \leqslant N$ such that

$$\|\boldsymbol{q} \cdot \boldsymbol{x} - \beta\| \leqslant N^{-n} \log^3 N. \tag{6.14}$$

The following argument is done in detail for $n = 2$ and any decreasing function $\phi(N)$ and is not difficult to extend to $n \geqslant 3$. Consider the line $x_2 = 1$. Then $|\boldsymbol{q} \cdot \boldsymbol{x} + \beta| < \phi(N)$ implies that $|q_1 x_1 + q_2 + \beta| < \phi(N)$. Thus we can consider the inequality $\|q_1 x + \beta\| < \phi(N)$. Similarly, consider the line $x_2 = \alpha$ for any $1 \geqslant \alpha > 0$. Then the inequality $|q_1 x_1 + q_2\alpha + \beta| < \phi(N)$ is the same as the inequality $\|qx + \beta'\| < \alpha^{-1}\phi(N)$, where $x = x_1/\alpha$ and $\beta' = \beta/\alpha$. Thus, if there exists a function ϕ for which almost all points on the lines $x_2 = \alpha$ are covered by intervals centred at points $(p+\beta')/q$ of length $\alpha^{-1}\phi(N)/q$ with $N/\log^4 N < q < N$, then, for some fixed $\delta > 0$ with $\alpha > \delta$ almost all points in I_δ^2 will be covered by the set

$$\bigcup_{\boldsymbol{q}: N > |\boldsymbol{q}| > N/\log^4 N} \{\boldsymbol{x} : |\boldsymbol{q} \cdot \boldsymbol{x} + \beta| < \phi(|\boldsymbol{q}|)\}$$

for N sufficiently large. On letting $\phi(N) = N^{-1}\log^3 N$ and using (6.14) and the fact that $N > |\boldsymbol{q}| > N/\log^4 N$ the lemma is proved. Exactly the same argument can be used for $n \geqslant 3$. □

Using the ubiquity argument from [**32**] it can be shown that as in the homogeneous case with $L_N = [N^2/8n\log^7 N]^{-1}$ we have

$$\#G(N) \geqslant \left[\frac{N^2}{8m\log^7 N}\right]^n (1 - \varepsilon).$$

Theorem 6.3 can now be proved by using the above and following the proof of Theorem 6.2 exactly. □

Theorems 6.2 and 6.3 show that the sets $E_0(\sigma; m, n)$ and $E_{\boldsymbol{\beta}}(\sigma; m, 1)$ consist of more than just the resonant hyperplanes $R_{\boldsymbol{q},\beta}$.

Acknowledgements. D.D. was partly supported by the EPSRC. T.G. was partly supported by Volkswagen-Stiftung (RiP-program at Oberwolfach), by a Coordinated Research Project of the University of Cagliari, and by INDAM–GNAMPA, Italy. M.Y. was partly supported by Volkswagen-Stiftung (RiP-program at Oberwolfach), Grant-in-Aid for Scientific Research (no. 11640183), Ministry of Education, Science and Culture, Japan, and by Chuo University, Tokyo, Japan. This paper was initiated when T.G. and M.Y. were staying in MFO, Oberwolfach, in the framework of an RiP program. They thank Matthias Kreck for his hospitality. The authors are also grateful to Maurice Dodson and Vasilii Bernik for some enlightening discussions on Diophantine approximation and applications of Hausdorff dimension, and to Tim Steger for suggesting the use of theorems on finitely generated abelian groups in the study of resonant sets. T.G. thanks Rafael De La Llave for bringing [**12**] to his attention and for useful discussions on the appearance of Gevrey spaces in homological equations. Finally we thank the referee and Chris Smyth for making some helpful comments.

References

1. V. I. Arnol'd, *Geometrical methods in the theory of ordinary differential equations* (Springer, 1983).
2. D. Bambusi and N. N. Nekhoroshev, A property of exponential stability in nonlinear wave equations near the fundamental linear mode, *Physica* D **122** (1998), 73–104.
3. A. Bergamasco, Perturbations of globally hypoelliptic operators, *J. Diff. Eqns* **141** (1994), 513–526.
4. A. Bergamasco, Remarks about global analytic hypoellipticity, *Trans. Am. Math. Soc.* **351** (1999), 4113–4126.
5. A. Bergamasco, P. Cordaro and P. Malagutti, Globally hypoelliptic systems of vector fields, *J. Funct. Analysis* **114** (1993), 263–285.
6. G. Birkhoff and S. MacLane, *Algebra* (Macmillan, New York, 1979).
7. F. Bonetto, G. Gallavotti, G. Gentile and V. Mastropietro, Lindstedt series, ultraviolet divergences and Moser's theorem, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **26** (1998), 545–593.
8. J. D. Bovey and M. M. Dodson, The Hausdorff dimension of systems of linear forms, *Acta Arith.* **XLV** (1986), 337–358.
9. H. W. Broer, G. B. Huitema and M. B. Sevryuk, *Quasi-periodic motions in families of dynamical systems. Order amidst chaos*, Lecture Notes in Mathematics, vol. 1645 (Springer, 1996).
10. A. D. Bruno, The analytic form of differential equations, *Tr. Mosk. Mat. O. va* **25** (1971), 119–262 and **26** (1972), 199–239 (in Russian).
11. A. D. Bruno and S. Walcher, Symmetries and convergence of normalizing transformations, *J. Math. Analysis Applic.* **183**, (1994), 571–576.
12. W. Chen and M. Y. Chi, Hypoelliptic vector fields and almost periodic motions on the torus $\boldsymbol{T}^n$, *Commun. PDEs* **25** (2000), 337–354.
13. R. De La Llave, A tutorial on KAM theory, *Smooth ergodic theory and its applications* (Seattle, WA, 1999), pp. 175–292, *Proc. Symp. Pure Math.*, vol. 69 (American Mathematical Society, Providence, RI, 2001).

14. D. DeLatte, Diophantine conditions for the linearization of commuting holomorphic functions, *Discrete Contin. Dynam. Syst.* **3**, (1997), 317–332.
15. H. Dickinson, The Hausdorff dimension of systems of simultaneously small linear forms, *Mathematika* **40** (1993), 367–374.
16. D. Dickinson and S. L. Velani, Hausdorff measure and linear forms, *J. Reine Angew. Math.* **490** (1997), 1–36.
17. D. Dickinson, T. Gramchev and M. Yoshino, First order pseudodifferential operators on the torus: normal forms, Diophantine phenomena and global hypoellipticity, *Proc. Differential Equations Conf., Ferrara, June 1996*, *Ann. Univ. Ferrara Sez. VII (NS)* **41** (Suppl.) (1997), 51–64.
18. M. M. Dodson and J. A. G. Vickers, Exceptional sets in Kolmogorov–Arnol'd–Moser theory, *J. Phys.* A **19** (1986), 349–374.
19. L. H. Eliasson, Absolutely convergent series expansions for quasi periodic motions, *Math. Phys. Electron. J.* **2** (1996), Paper 4 (electronic).
20. K. Falconer, *Fractal geometry: mathematical foundations and applications* (Wiley, 1990).
21. F. Fass, M. Guzzo and G. Benettin, Nekhoroshev-stability of elliptic equilibria of Hamiltonian systems, *Commun. Math. Phys.* **197** (1998), 347–360.
22. A. Giorgilli, A. Delshams, E. Fontich, L. Galgani and C. Simó, Effective stability for a Hamiltonian system near an elliptic equilibrium point, with an application to the restricted three body problem, *J. Diff. Eqns* **77** (1989), 167–198.
23. T. Gramchev and G. Popov, Nekhoroshev type estimates for billiard ball maps, *Annls Inst. Fourier* **45** (1995), 859–895.
24. T. Gramchev and M. Yoshino, Rapidly convergent iterative method for simultaneous normal forms of commuting maps, *Math. Z.* **231** (1999), 745–770.
25. T. Gramchev, P. Popivanov and M. Yoshino, Global properties in spaces of generalized functions on the torus for second order differential operators with variable coefficients, *Rend. Sem. Mat. Univ. Pol. Torino* **51** (1993), 144–174.
26. M. Herman, Recent results and some open questions on Siegel's linearization theorem of germs of complex analytic diffeomorphisms of $\mathbb{C}^n$ near a fixed point, in *Proc. VIIIth Int. Congr. on Mathematical Physics, Marseille, 1986* (ed. M. Mebkhout and R. Sénéor), pp. 138–184 (World Scientific, 1987).
27. A. Himonas, and G. Petronilho, Global hypoellipticity and simultaneous approximability, *J. Funct. Analysis* **170** (2000), 356–365.
28. J. Hounie, Globally hypoelliptic vector fields on compact surfaces, *Commun. PDEs* **7** (1982), 343–370.
29. A. Jorba and J. Villanueva, On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems, *Nonlinearity* **10** (1997), 783–822.
30. A. Jorba, R. Ramirez-Ros and J. Villanueva, Effective reducibility of quasi-periodic linear equations close to constant coefficients, *SIAM J. Math. Analysis* **28** (1997), 178–188.
31. Y. Katznelson, B. Kra and D. Ornstein, A global version of a theorem of J. Moser, *Israel J. Math.*, in press.
32. J. Levesley, A general inhomogeneous Jarník–Besicovitch theorem, *J. Number Theory* **71** (1998), 65–80.
33. P. Loshak, Canonical perturbation theory: an approach based on joint approximations, *Russ. Math. Surv.* **47** (1992), 57–133.
34. J. Moser, On commuting circle mappings and simultaneous Diophantine approximations, *Math. Z.* **205** (1990), 105–121.
35. J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Commun. Pure Appl. Math.* **35** (1982), 653–696.
36. L. Rodino, *Linear partial differential operators in Gevrey spaces* (World Scientific, 1993).

37. H. Rüssmann, *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, Lecture Notes in Physics, vol. 38, pp. 598–624 (Springer, 1975).
38. W. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785 (Springer, 1980).
39. L. Stolovitch, Singular complete integrability, *Inst. Hautes Etudes Sci. Publ. Math.*, no. 91 (2000), pp. 133–210.
40. J.-C. Yoccoz, Recent developments in dynamics, in *Proc. Int. Congr. of Mathematicians, Zürich, Switzerland, August 1994*, pp. 246–265 (Birkhäuser, 1995).