Ergod. Th. & Dynam. Sys. (2002), **22**, 1429–1449
DOI: 10.1017/S0143385702000998
© *2002 Cambridge University Press*
Printed in the United Kingdom

Ergodic skew-systems on $\mathbb{T}^d \times SO(3, \mathbb{R})$

L. H. ELIASSON
Department of Mathematics, University of Paris 7, Paris, France
(e-mail: hakane@math.jussieu.fr)

Abstract. Analytic quasi-periodic skew-systems with close to constant coefficients and with Diophantine frequencies seem to be, in a measure theoretical sense, 'generically' reducible. In this paper we show that such systems are, in a topological sense, generically uniquely ergodic and, in particular, non-reducible.

1. *Introduction*

In this paper, we consider *skew-systems* on $\mathbb{T}^d \times SO(3, \mathbb{R})$, $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$:

$$\begin{aligned} X' &= F(x)X \\ x' &= \omega \end{aligned} \qquad (*)$$

where $F : \mathbb{T}^d \to so(3, \mathbb{R})$ is real analytic and $\omega \in \mathbb{R}^d$ satisfies a *Diophantine condition*

$$|\langle n, \omega\rangle| \geq \frac{1}{|n|^\tau}, \quad n \in \mathbb{Z}^d \setminus 0,$$

for some $\tau > d-1$. (Here $\langle \cdot, \cdot \rangle$ is the usual scalar product in $\mathbb{R}^d$ and $|\cdot|$ is the corresponding norm.) We assume that F is defined in a complex strip of width r, and we measure it by

$$|F|_r = \sup_{|\mathrm{Im}\, x| < r} |F(x)|,$$

where $|F(x)|$ is the operator norm. The flow of $(*)$:

$$Q_t(x, Q) = (x + t\omega, X(t, x)Q)$$

preserves the product Haar measure $\mu \times \nu$ on $\mathbb{T}^d \times SO(3, \mathbb{R})$. It preserves distances both in the $\mathbb{T}^d$-direction and in the $SO(3, \mathbb{R})$-direction, but it is not an isometry. Its second factor satisfies the differential equation

$$\frac{d}{dt}X(t, x) = F(x + t\omega)X(t, x) \qquad (**)_x$$

with initial value $X(0, x) = I$.

'Most' systems $(*)$, and very likely 'almost all' systems $(*)$, are *reducible* (to a constant skew-system) in the following sense. There exist $Y : \mathbb{T}^d \to SO(3, \mathbb{R})$ real analytic and $A \in so(3, \mathbb{R})$ such that

$$\partial_\omega Y(x) = F(x)Y(x) - Y(x)A,$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

where $\partial_\omega = \langle \nabla, \omega \rangle$. It follows that the transformation

$$\tilde{x} = x, \quad \tilde{X} = Y^*(x)X$$

transforms $(*)$ to the constant system

$$\tilde{X}' = A\tilde{X}$$
$$\tilde{x}' = \omega.$$

Hence, each orbit $\{Q_t(x, Q)\}_{t\in\mathbb{R}}$ is contained in an analytic torus of dimension $d+1$ and analytically conjugate to a Kronecker flow with direction (ω, α), where $\pm i\alpha, 0$ are the eigenvalues of A. In particular, $(*)$ has an uncountable number of invariant ergodic measures. Another consequence is that the system $(**)_x$ has a *Floquet representation*

$$X(t, x) = Y(x + t\omega)e^{At}Y^*(x)$$

as a quasi-periodic matrix times an exponential matrix. (That the quasi-periodic part has exactly the same frequencies as the system itself is particular for the group $SO(3, \mathbb{R})$.) We shall not try to give a precise meaning to 'most' and 'almost all' since in this paper we are mainly concerned with systems that are not reducible. The existence of such systems is made plausible by the fact that quasi-periodic linear systems do not always admit a Floquet representation. The first example of this was, we believe, [**M**]. The first analytic examples with ω Diophantine was [**FSW**]—non-perturbative—and [**E1**]—perturbative. In both these cases the group is $SL(2, \mathbb{R})$. In [**R**] a smooth example on the compact group $SU(2, \mathbb{C})$ is given. For time-discrete systems there are several examples in [**H**].

So one should not expect that all systems $(*)$, even small in the $|\cdot|_r$-norm, are reducible, and this is indeed not the case. One could think that such systems still are reducible in a weaker sense, for example by requiring that the transformation $Y(x)$ is just measurable in x and, say, analytic along the lines $\{x + t\omega : t \in \mathbb{R}\}$. There is, in fact, a dense set of such systems in a small $|\cdot|_r$-neighborhood of $F = 0$, but this is not the generic situation. We shall prove the following.

THEOREM. *There exists a constant $C = C(r, \tau)$ such that for a dense $\mathcal{G}_\delta$ in*

$$|F|_r < C$$

the skew-system $(*)$ has a unique invariant measure.

A related result has been proven in [**N**] for a Liouville vector ω using the idea of fast periodic approximations. Our proof will consist in applying the approach of [**E1**]—used to study skew-systems on $\mathbb{T}^d \times SL(2, \mathbb{R})$—to the compact group $SO(3, \mathbb{R})$. We shall, by a KAM-iteration, construct an infinite sequence of functions

$$Y_{j+1} : \mathbb{T}^d \to SO(3, \mathbb{R}), \quad j \leq 1,$$

whose infinite product 'reduces' the system to constant coefficients. If this product converges uniformly on $\mathbb{T}^d$ then we are in the reducible case. For a generic system, however, the product does not converge uniformly, but it still provides information about the system. In particular, the information needed to prove unique ergodicity. We have

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

chosen to formulate our result for systems near $F = 0$, but we could just as well have done it for systems near any $A \in so(3, \mathbb{R})$—the proof will go through almost without change. We also believe that the results and proofs of this paper can be generalized to any group $SO(n, \mathbb{R})$.

The outline of this paper is as follows. In §2 we describe some properties of real anti-symmetric matrices and prove the usual small divisor lemma. In §3 we describe the KAM-iteration. In §4 we use this iterative construction to show that 'most' small systems are reducible. In §5 we prove the unique ergodicity of $(*)$ under a certain condition on F and in §6 we prove the density of this condition. This turns out to be surprisingly complicated because of the difficulty to control parameter-dependence during the iteration—something which is often easy in KAM approaches.

1.1. *Notes.* This work appeared as a preprint in 1991 [**E2**]. The present version is the same, with minor changes, except for §6 where the proof of genericity has been simplified.

Since 1991 there has been progress on the question of reducibility which is reported in [**E3**]. In particular, the conjecture on the one-parameter family $\lambda A + F(x)$, $\det A \neq 0$, in §4 has been proven by R. Krikorian [**K**].

2. *Preliminaries*

Consider the Lie algebra $so(3, \mathbb{R})$ of real anti-symmetric 3×3-matrices. We have a basis

$$J_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

satisfying

$$[J_1, J_2] = J_3$$

plus cyclic permutations.

LEMMA 1. *For any $A \in so(3, \mathbb{R})$ there is a $M \in SO(3, \mathbb{R})$ such that*

$$M^* A M = \alpha J_1.$$

If we fix the sign of α, $\alpha > 0$ say, then M is determined up to a transformation e^{tJ_1}. If A depends C^1 on some parameter λ and if $\alpha \neq 0$, then we can choose M to be C^1 locally and

$$|\partial M| \leq \text{const.} \left| \frac{1}{\alpha} \partial A \right|,$$

where $\partial = \partial/\partial\lambda$. The constant is independent of A.

Proof. Let $\langle \cdot, \cdot \rangle$ be the usual sesqui-linear product on $\mathbb{C}^3$—anti-linear in the first and linear in the second variable. Suppose

$$Au = \lambda u, \quad |u|^2 = \langle u, u \rangle = 2.$$

Then

$$\overline{\lambda} \langle u, u \rangle = \langle Au, u \rangle = -\langle u, Au \rangle = -\lambda \langle u, u \rangle$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

so $\lambda = i\alpha$ is purely imaginary. Since A is real, both $\pm i\alpha$ are eigenvalues and

$$i\alpha\langle\overline{u}, u\rangle = \langle A\overline{u}, u\rangle = -\langle\overline{u}, Au\rangle = -i\alpha\langle\overline{u}, u\rangle,$$

so $\langle\overline{u}, u\rangle = 0$. If $u = u_1 + iu_2$, then

$$0 = \langle\overline{u}, u\rangle = |u_1|^2 - |u_2|^2 + i2\langle u_1, u_2\rangle,$$

which implies that

$$|u_1|^2 = |u_2|^2 = 1, \quad \langle u_1, u_2\rangle = 0.$$

Now choose v real so that

$$\langle v, u_1\rangle = \langle v, u_2\rangle = 0, \quad |v| = 1.$$

Then $Av = 0$ and we can take $M = (u_1\ u_2\ v)$, where we determine the sign of v so that $\det M = 1$. The second statement of the lemma is an easy calculation. Let $A = xJ_1 + yJ_2 + zJ_3$. Then $\alpha = \sqrt{x^2 + y^2 + z^2}$ and the characteristic equation of A is

$$t(t^2 + \alpha^2) = 0.$$

Now let $w = A(A + i\alpha I)e$. If $e = e_1 = (1\ 0\ 0)^{\mathrm{t}}$, then

$$w = \begin{pmatrix} -x^2 - z^2 \\ -yz \\ xy \end{pmatrix} + i\alpha \begin{pmatrix} 0 \\ -x \\ -z \end{pmatrix}.$$

This implies that

$$|w| \geq c_1|\alpha||A|$$

unless $x^2 + z^2 < y^2$, in which case we get the same inequality by taking e to be e_2 or e_3. Now $u = 2w/|w|$ from which we get

$$|\partial u| \leq c_2\frac{|\partial w|}{|w|} \leq c_3\frac{|\partial A|}{|\alpha|}.$$

If we let $w = (A^2 + \alpha^2 I)e$ with $e = e_1$ (or e_2 or e_3), then we get $v = \pm w/|w|$ and we get the estimate

$$|\partial v| \leq c\frac{|\partial A|}{|\alpha|}$$

in the same way. □

From the real normal form we easily get a complex diagonal form. In fact

$$C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \Longrightarrow C^*\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}C = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

We shall use this for the next lemma. Let $\mathcal{B}_r$ be the space of all analytic functions $F : \mathbb{T}^d \to gl(3, \mathbb{R})$ for which $|F|_r < \infty$ and let $\mathcal{B} = \cup\mathcal{B}_r$. Let $F \in \mathcal{B}_r$ and suppose that F depends on a parameter λ in some open set $\Delta \subset \mathbb{R}^3$. We say that F is C^1 *in* λ if $\lambda \mapsto F_\lambda \in \mathcal{B}_r$ is C^1. We say that F is *piecewise* C^1 *in* λ if $\Delta = \cup\Delta_i$ and, for each i, there is an open set $\Omega_i \supset \Delta_i$ such that F/Δ_i extends to a C^1-function on Ω_i. Let

$$\hat{\sigma}(F) = \{n \in \mathbb{Z}^d : \hat{F}(n) \neq 0\}$$

where $\hat{F}(n)$ is the nth Fourier coefficient of F.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

LEMMA 2. *Let $A \in so(3,\mathbb{R})$ and $F \in \mathcal{B}_r \cap so(3,\mathbb{R})$. Assume that $0 \notin \hat{\sigma}(F)$ and*

$$|\langle n, \omega\rangle \pm \alpha| \geq \frac{1}{K|n|^\tau}, \quad n \in \hat{\sigma}(F),$$

where $\pm i\alpha, 0$ are the eigenvalues of A. Then there is a unique $Y \in \mathcal{B} \cap so(3,\mathbb{R})$ such that

$$\partial_\omega Y = [A, Y] + F, \quad \hat{\sigma}(Y) \subset \hat{\sigma}(F).$$

Y satisfies the estimate

$$|Y|_s \leq \text{const.}\, \frac{K}{(r-s)^\tau}|F|_r, \quad s < r.$$

Moreover, if A and F are C^1 in $\lambda \in \Delta$, then $Y \in \mathcal{B}_s$ also and

$$|\partial Y|_s \leq \text{const.}\left[\frac{K^2}{(r-s)^{2\tau}}|\partial A||F|_r + \frac{K}{(r-s)^\tau}|\partial F|_r\right], \quad s < r,$$

where $\partial = \partial/\partial\lambda$. The constant only depends on τ.

Proof. We can assume that $A = \alpha J_1$. If we write

$$Y = \sum Y_i J_i \quad \text{and} \quad F = \sum F_i J_i,$$

then we get the equations

$$\partial_\omega Y_1 = F_1, \quad \partial_\omega Y_2 + \alpha Y_3 = F_2, \quad \partial_\omega Y_3 - \alpha Y_2 = F_3.$$

Now consider

$$C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$$

and write

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} Y \quad \text{and} \quad G = \begin{pmatrix} 1 & 0 \\ 0 & C^* \end{pmatrix} F.$$

Then

$$\partial_\omega Z_1 = G_1, \quad \partial_\omega Z_2 + i\alpha Z_2 = G_2, \quad \partial_\omega Z_3 - i\alpha Z_3 = G_3,$$

with $\hat{\sigma}(Z_i) \subset \hat{\sigma}(F)$. These equations can now be solved in Fourier series and the estimates are standard. From the estimate of $|Y|_s$ we get the C^1-dependence in λ and the estimate of ∂Y follows from

$$\partial_\omega(\partial Y) = [A, \partial Y] + [\partial A, Y] + \partial F. \qquad \square$$

3. *The inductive lemma*

Let $F_1(\cdot,\lambda) \in \mathcal{B}_{r_1} \cap so(3,\mathbb{R})$ and $A_1(\lambda) \in so(3,\mathbb{R})$. Assume that F_1 and A_1 depend piecewise C^1 on a parameter $\lambda \in \Delta$. Assume also that

$$|\partial^\nu A_1| < 1, \quad |\partial^\nu F_1|_{r_1} < \varepsilon_1$$

for $\nu = 0, 1$, where $\partial = \partial/\partial\lambda$. Let $0 < \sigma < 1/36$ and let $\{r_j\}$ be a decreasing sequence of positive numbers such that

$$r_j - r_{j+1} \geq \frac{r_1}{2^{j+1}}, \quad j \geq 1.$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

Let

$$\varepsilon_{j+1} = \varepsilon_j^{1+\sigma}, \quad N_j = \frac{2\sigma}{r_j - r_{j+1}} \log\left(\frac{1}{\varepsilon_j}\right)$$

for each $j \geq 1$. It follows that

$$\varepsilon_j^{\sigma} \leq N_j^{-4\tau}$$

for all j, if just ε_1 is small enough. For example, it suffices that

$$\varepsilon_1 \leq \text{const.}\, r_1^{((4\tau/\sigma)+1)}$$

for some small constant that only depends on τ and σ.

LEMMA 3. *There exists a constant $C = C(\tau, \sigma)$ such that if*

$$\varepsilon_1 < C r_1^{((4\tau/\sigma)+1)}$$

then, for all $j \geq 1$, there exist $F_{j+1} \in \mathcal{B} \cap so(3, \mathbb{R})$, $Y_{j+1} \in \mathcal{B} \cap SO(3, \mathbb{R})$ and $A_{j+1} \in so(3, \mathbb{R})$, all piecewise C^1 in λ, satisfying

$$\partial_\omega Y_{j+1}(x) = (A_j + F_j(x))Y_{j+1}(x) - Y_{j+1}(x)(A_{j+1} + F_{j+1}(x)).$$

Moreover, if $\pm i\alpha_j, 0$ are the eigenvalues of A_j, then we have the following estimates for $\nu = 0, 1$:

$$|\partial^\nu (Y_{j+1} - e^{(\langle m_j, \cdot\rangle/\alpha_j)A_j})|_{r_{j+1}} < \varepsilon_j^{1/2}, \tag{1}_{j+1}$$

where either m_j is the unique integer valued vector restricted to $0 < |m| \leq N_j$ such that $|\alpha_j - \langle m, \omega\rangle| < \varepsilon_j^{\sigma}$ and $r_j - r_{j+1} = \frac{1}{2} r_j$, or, if no such vector exists, $m_j = 0$ and $r_j - r_{j+1} = r_1/2^{j+1}$,

$$|\partial^\nu F_{j+1}|_{r_{j+1}} < \varepsilon_{j+1}, \tag{2}_{j+1}$$

$$\left|\partial^\nu \left(A_{j+1} - \left(1 - \frac{\langle m_j, \omega\rangle}{\alpha_j}\right) A_j\right)\right| < \varepsilon_j^{2/3}, \tag{3}_{j+1}$$

$$|\partial A_{j+1}| < \varepsilon_{j+1}^{-\sigma}. \tag{4}_{j+1}$$

Proof. Let us show how to construct Y_{j+1}, F_{j+1} and A_{j+1} assuming the previous steps are done. We must handle two different cases.

Case 1. Assume that

$$|\alpha_j - \langle n, \omega\rangle| \geq \varepsilon_j^{\sigma}, \quad 0 < |n| \leq N_j$$

and let $G(x)$ be the truncated Fourier series

$$\sum_{0<|n|\leq N_j} \hat{F}_j(n) e^{i\langle n, x\rangle}.$$

Let Y to be the solution of

$$\partial_\omega Y = [A_j, Y] + G, \quad \hat{\sigma}(Y) \subset \hat{\sigma}(G),$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

which exists uniquely by Lemma 2 and define

$$\begin{aligned} Y_{j+1}(x) &= e^{Y(x)} \\ A_{j+1} &= A_j + \hat{F}_j(0) \\ F_{j+1} &= e^{-Y(x)}\{F_j(x)(e^{Y(x)} - I) + [A_j, e^{Y(x)} - I - Y(x)] \\ &\quad - (e^{Y(x)} - I)\hat{F}_j(0) - \partial_\omega(e^{Y(x)} - Y(x)) + (F_j(x) - G(x) - \hat{F}_j(0))\}. \end{aligned}$$

With these definitions, the functions F_{j+1} etc., have all the required properties so we only need to consider the estimates.

Estimates. We have

$$|\partial^\nu G|_s < c\varepsilon_j N_j^{\tau+1}, \quad s = r_{j+1} + \frac{r_j - r_{j+1}}{2}.$$

(In the following, c will denote an arbitrary constant that only depends on σ and τ.) Then by Lemma 2 and $(4)_j$ we get

$$|\partial^\nu Y|_{r_{j+1}} < \text{const.}\, \varepsilon_j^{1-(1+2\nu)\sigma} N_j^{\tau(2+\nu)+1}.$$

Also $e^{Y(x)} - I$ and $e^{-Y(x)} - I$ satisfy these estimates which proves $(1)_{j+1}$. In order to prove $(2)_{j+1}$ we first observe that

$$|\partial^\nu(F_j - G - \hat{F}_j(0))|_{r_{j+1}} < c\varepsilon_j^{1+2\sigma} N_j^{\tau+1}.$$

Moreover,

$$\begin{aligned} |\partial_\omega(e^Y - Y)|_{r_{j+1}} &< c(|A_j||Y|^2_{r_{j+1}} + |G|_{r_j}|Y|_{r_{j+1}}) \\ |\partial\partial_\omega(e^Y - Y)|_{r_{j+1}} &< c(|A_j||\partial Y|_{r_{j+1}}|Y|_{r_{j+1}} + |\partial A||Y|^2_{r_{j+1}} + |\partial G|_{r_j}|Y|_{r_{j+1}} + |G|_{r_j}|\partial Y|_{r_{j+1}}). \end{aligned}$$

From this we easily get

$$|F_{j+1}|_{r_{j+1}} < c(\varepsilon_j^{1+2\sigma} N_j^{\tau+1} + \varepsilon_j^{2-2\sigma} N_j^{4\tau+2}) < \varepsilon_j^{1+\sigma}.$$

If we differentiate the expression for F_{j+1}, we can estimate the derivative in the same way and we get

$$|\partial F_{j+1}|_{r_{j+1}} < c(\varepsilon_j^{1+2\sigma} N_j^{\tau+1} + \varepsilon_j^{2-4\sigma} N_j^{5\tau+2}) < \varepsilon_j^{1+\sigma}.$$

This proves $(2)_{j+1}$. The estimate $(3)_{j+1}$ is trivial, as is $(4)_{j+1}$.

Case 2. Suppose now that there is an integer vector m_j in $0 < |m| \le N_j$—necessarily unique—such that

$$|\alpha_j - \langle m_j, \omega\rangle| < \varepsilon_j^\sigma,$$

and let

$$Z(x) = e^{(\langle m_j, x\rangle/\alpha_j)A_j}.$$

Then

$$\partial_\omega Z(x) = (A_j + F_j(x))Z(x) - Z(x)(B_j + G_j(x))$$

where

$$B_j = \left(1 - \frac{\langle m_j, \omega\rangle}{\alpha_j}\right) A_j, \quad G_j(x) = Z^*(x)F_j(x)Z(x).$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

The eigenvalues of B_j are $\pm i\beta_j, 0$ with $\beta_j = \alpha_j - \langle m_j, \omega\rangle$ and they satisfy

$$|\beta_j - \langle n, \omega\rangle| \geq (6N_j)^{-\tau} - \varepsilon_j^\sigma \geq 2\varepsilon_j^\sigma, \quad 0 < |n| \leq 6N_j.$$

Now let now G be the truncated Fourier series

$$\sum_{0<|n|\leq 6N_j} \hat{G}_j(n) e^{i\langle n,x\rangle}$$

and let Y be the unique solution of

$$\partial_\omega Y = [B_j, Y] + G, \quad \hat{\sigma}(Y) \subset \hat{\sigma}(G).$$

We define

$$\begin{aligned}
Y_{j+1}(x) &= Z(x)e^{Y(x)} \\
A_{j+1} &= B_j + \hat{G}_j(0) \\
F_{j+1}(x) &= e^{-Y(x)}\{G_j(x)(e^{Y(x)} - I) + [B_j, e^{Y(x)} - I - Y(x)] \\
&\quad - (e^{Y(x)} - I)\hat{G}_j(0) - \partial_\omega(e^{Y(x)} - Y(x)) + (G_j(x) - G(x) - \hat{G}_j(0))\}.
\end{aligned}$$

As in Case 1, all requirements are fulfilled and we only need to consider the estimates.

Estimates. By Lemma 1, we can write

$$Z(x) = M^* e^{\langle m_j, x\rangle J_1} M,$$

where $M \in SO(3, \mathbb{R})$ and depends C^1 on λ (locally) and

$$|\partial M| \leq c\frac{|\partial A_j|}{|\alpha_j|}.$$

It follows that

$$|\partial^\nu Z|_{r_j} < cN_j^{\nu\tau}\varepsilon_j^{-(4+\nu)\sigma},$$

since $r_{j+1} = \frac{1}{2}r_j$. The same estimate holds for Z^*. This implies that

$$|\partial^\nu G_j|_{r_j} < cN_j^{\nu\tau}\varepsilon_j^{1-(8+\nu)\sigma}$$

and

$$|\partial^\nu G|_s < cN_j^{(1+\nu)\tau+1}\varepsilon_j^{1-(8+\nu)\sigma}, \quad s = r_{j+1} + \frac{r_j - r_{j+1}}{2}.$$

Moreover,

$$|\partial B_j| < c\varepsilon_j^{-\sigma}.$$

We conclude from Lemma 2 that

$$|\partial^\nu Y|_{r_{j+1}} < cN_j^{(2+2\nu)\tau+1}\varepsilon_j^{1-(9+3\nu)\sigma}$$

and

$$|\partial^\nu (Y_{j+1} - Z)|_{r_{j+1}} < cN_j^{(2+2\nu)\tau+1}\varepsilon_j^{1-(13+3\nu)\sigma} < \varepsilon_j^{1/2}.$$

This proves $(1)_{j+1}$. In order to prove $(2)_{j+1}$, we observe that

$$|\partial^\nu (G_j - G - \hat{G}_j(0))|_{r_{j+1}} < cN_j^{(1+\nu)\tau+1}\varepsilon_j^{1+(4-\nu)\sigma},$$

from which we deduce the estimate of F_{j+1} as in the previous case. The estimate $(3)_{j+1}$ is obvious and $(4)_{j+1}$ follows from the estimate of $|\partial B_j|$ given above. This completes the proof of Lemma 3. □

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

4. *Reducibility*

Each function $Y_{j+1}(x)$ is close to an exponential $e^{B_j(x)}$. If all but finitely many B_j are equal to zero, then the product

$$\prod_{j=1}^{\infty} Y_{j+1}(x)$$

converges uniformly on $\mathbb{T}^d$ to an analytic function $Y(x)$ which solves

$$\partial_\omega Y(x) = (A_1 + F_1(x))Y(x) - Y(x)A, \quad A = \lim_{j\to\infty} A_j,$$

so the skew-system is reducible. If this is not so, then the matrices A_j must satisfy

$$|\alpha_j - \langle m_j, \omega\rangle| < \varepsilon_j^\sigma, \quad \text{some } 0 < |m_j| \le N_j,$$

for infinitely many j's. Since

$$|\{\alpha : |\alpha - \langle n, \omega\rangle| < \varepsilon_j^\sigma, \text{some } 0 < |n| \le N_j\}| < 2\varepsilon_j^\sigma N_j^{\tau+1}$$

converges to zero very fast, we like to think of this as the exceptional case. In fact, it is quite easy to show that in any one-parameter family

$$\lambda A + F, \quad \det A \neq 0,$$

the system is reducible except for a small set of λ-values if $|F|_r$ is small enough. We believe that this family is reducible for a.e. λ, but this may be more delicate to prove.

Let us now turn to the case when there are infinitely many j's for which $m_j \neq 0$. The important facts to keep in mind are:

(i)
$$|Y_{j+1} - e^{B_j}|_{r_{j+1}} < \varepsilon_j^{1/2}, \quad j \ge 1; \tag{5}$$

(ii) the matrices B_j are of the form

$$B_j(x) = \frac{\langle m_j, x\rangle}{\alpha_j} A_j,$$

where $\pm i\alpha_j, 0$ are the eigenvalues of A_j and $0 < |m_j| \le N_j$;

(iii) the matrices B_j are all equal to zero except for some subsequence

$$1 \le j_1 < j_2 < \dots;$$

(iv)
$$|B_{j_{n+1}}(\omega)| < 2\varepsilon_{j_n}^\sigma, \quad n \ge 1; \tag{6}$$

(v)
$$|B_{j_{n+1}}(\omega)| > \tfrac{1}{2} N_{j_{n+1}}^{-\tau}, \quad n \ge 0. \tag{7}$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

Equation (5) is just a reformulation of (1), (7) follows from the form of B_j and from the Diophantine condition, and (6) follows because

$$\begin{aligned}|\langle m_{j_{n+1}}, \omega\rangle| &< \varepsilon_{j_{n+1}}^{\sigma} + |\alpha_{j_{n+1}}| < \varepsilon_{j_{n+1}}^{\sigma} + |A_{j_{n+1}}| \\ &< \varepsilon_{j_{n+1}}^{\sigma} + 2\varepsilon_{j_n}^{2/3} + \left|\left(1 - \frac{\langle m_{j_n}, \omega\rangle}{\alpha_{j_n}}\right) A_{j_n}\right| < 2\varepsilon_{j_n}^{\sigma}.\end{aligned}$$

A consequence of (6) and (7) is that the sequence $\{j_n\}$ increases extremely fast. In fact, the right-hand side of (7) decreases in the same way as $(2+2\sigma)^{-\tau j_{n+1}}$ and the right-hand side of (6) in the same way as $\varepsilon_1^{(1+\sigma)^{j_n}}$. Hence

$$j_{n+1} \geq \text{const.} \log\left(\frac{1}{\varepsilon_1}\right)(1+\sigma)^{j_n}.$$

We now have

$$r_j \searrow 0, \quad |F_j|_{r_j} \to 0, \quad A_j \to 0$$

where the last limit follows from (6) because $|A_{j_n}| \sim |B_{j_n}(\omega)|$. It also follows that the product

$$\prod_{j=1}^{\infty} Y_{j+1}(t\omega)$$

converges, uniformly on compact t-intervals, to a solution $X(t)$ of $(**)_0$, because for t fixed we have

$$|e^{B_{j_n}(t\omega)} - I| < \text{const.}\, \varepsilon_{j_{n-1}}^{\sigma}|t|$$

if n is large enough. So we have a representation of a solution $X(t)$ of $(**)_0$ as a product $\prod Y_{j+1}(t\omega)$, where $\{Y_{j+1}\}$ satisfy (1)–(7). Such a representation is of course not unique, but the proof of Lemma 3 provides a unique—though arbitrary—construction of such a representation. Let us finally notice that

$$|X(0) - I| < 2\varepsilon_1^{1/2}$$

and, for simplicity, we shall assume in this section that the Y_j are normalized so that $X(0) = I$. The solution $X(t)$ of $(**)_0$ may of course still be quasi-periodic, but it must then have extra frequencies. Indeed, if $X(t)$ is quasi-periodic then its limit values 'above' the point $0 \in \mathbb{T}^d$ are an Abelian subgroup of $SO(3,\mathbb{R})$. By this we mean that the set of all $B \in SO(3,\mathbb{R})$ such that there exists a sequence $\{t_j\}$ with

$$t_j\omega \to 0 \text{ in } \mathbb{T}^d, \quad X(t_j) \to B$$

is an Abelian subgroup (see Theorem 1.17 in [**F**]). Moreover, if $X(t)$ has the same frequencies as $F(t\omega)$, then this subgroup must be the identity (see Theorem 4.5 in [**F**]). Now, the existence of limit values other than I follows from the following lemma.

LEMMA 4. *Given t and $n \geq n_0(\tau)$, there exists a $\tilde{t}$ such that, for all x,*

$$|X(\tilde{t}, x - \tilde{t}\omega) - e^{B_{j_n}(t\omega)}| < \varepsilon_{j_{n-2}}^{\sigma}.$$

In particular, if $\{t_n\}$ is a sequence such that $t_n\omega \to 0$ in $\mathbb{T}^d$ and e^B is an accumulation point of $\{e^{B_{j_n}(t_n\omega)}\}$, then there is a sequence $\{\tilde{t}_n\}$ such that $\tilde{t}_n\omega \to 0$ in $\mathbb{T}^d$ and $X(\tilde{t}_n)$ has e^B as an accumulation point.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

Proof. For given n, choose $\delta = \delta_n > 0$ such that

$$\left(C\frac{n}{\delta}\right)^{(\tau+1)(d+n-2)(d+n-1)/2} = \varepsilon_{j_{n-1}}^{-\sigma/4}.$$

Since $t \mapsto e^{B_{j_n}}(t\omega)$ is periodic with period $\leq 2\pi N_{j_n}^{\tau}$ we can assume that t is bounded by this number. By Corollary A.1 (see the Appendix) there exists, in any interval of length greater than

$$T = N_{j_{n-1}}^{\tau}\left(C\frac{n}{\delta}\right)^{(\tau+1)(d+n-2)(d+n-1)/2},$$

a point $\tilde{t}$ such that

$$\|\tilde{t}\omega_1\|, \ldots, \|\tilde{t}\omega_d\| \quad \text{and} \quad \|\tilde{t}\langle m_{j_1}, \omega\rangle\|, \ldots, \|\tilde{t}\langle m_{j_{n-1}}, \omega\rangle\|$$

are all less than δ/n. We take such a $\tilde{t}$ in the interval $[t, t+T]$. We now have

$$\begin{aligned}
|e^{B_{j_l}(\tilde{t}\omega)} - I| &< 2\frac{\delta}{n}, \quad \text{for } l \leq n-1,\\
|e^{B_{j_l}(\tilde{t}\omega)} - I| &< 2|\tilde{t}||\langle m_{j_l}, \omega\rangle| < 2(2\pi N_{j_n}^{\tau} + N_{j_{n-1}}^{\tau}\varepsilon_{j_{n-1}}^{-\sigma/4})\varepsilon_{j_{l-1}}^{\sigma}\\
&< \varepsilon_{j_{l-1}}^{\sigma/2}, \quad \text{for } l \geq n+1,\\
|e^{B_{j_n}(\tilde{t}\omega)} - e^{B_{j_n}(t\omega)}| &< 2\varepsilon_{j_{n-1}}^{\sigma}|T| < \varepsilon_{j_{n-1}}^{\sigma/2},
\end{aligned}$$

and

$$|Y_{j+1}(\tilde{t}\omega) - e^{B_j(\tilde{t}\omega)}| < \text{const.}\,\varepsilon_j^{1/2}\frac{1}{r_{j+1}}\|\tilde{t}\omega\| < \text{const.}\,\varepsilon_j^{1/2}\frac{1}{r_{j+1}}\delta < \varepsilon_j^{\sigma}\delta.$$

From this we deduce that

$$|X(\tilde{t}) - e^{B_{j_n}(t\omega)}| < 2\delta + 2\varepsilon_{j_{n-1}}^{\sigma/2},$$

which is less than $\varepsilon_{j_{n-2}}^{\sigma}$. In order to see this we observe that

$$\delta = \delta_n = Cn\varepsilon_{j_{n-1}}^{(\sigma/2(\tau+1))[(d+n-2)(d+n-1)]^{-1}} < \tfrac{1}{4}\varepsilon_{j_{n-2}}^{\sigma},$$

holds if

$$(1+\sigma)^{j_{n-1}-j_{n-2}} \gg 4(\tau+1)(d+n-2)(d+n-1).$$

However, this is certainly true for n greater than some n_0—depending only on τ. The rest of the lemma is an obvious consequence of this result. □

Remark. We have not yet proven that there are any F at all in $|F|_r < C$ for which the sequence $j_1 < j_2 < \cdots$ is actually infinite. This will follow from §6 where we prove the density of a much stronger statement.

Remark. The whole group of accumulation points of

$$\{e^{B_j(t_j\omega)} : t_j\omega \to 0 \text{ in } \mathbb{T}^d\},$$

is attained as limit values of $X(t)$ 'above' $0 \in \mathbb{T}^d$. We do not claim that this group contains all such limit values. In particular, the Abelian character of this group is a necessary, but not sufficient, condition for $X(t)$ to be quasi-periodic.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

5. *Ergodicity*

In the previous section we showed that the product

$$\prod_{j=1}^{\infty} Y_{j+1}(x + t\omega)$$

is uniformly convergent on compact t-intervals when the base point x is equal to 0. In order to also treat other values of x, we must modify the product. Let

$$\tilde{Y}_{j+1}(y, x) = Y_2(x) \cdots Y_j(x) Y_{j+1}(y + x) Y_{j+1}^*(x) Y_j^*(x) \cdots Y_2^*(x).$$

Then one checks easily that

$$X(t, x) = \prod \tilde{Y}_{j+1}(t\omega, x)$$

is uniformly convergent on compact t-intervals, for x fixed, and that $X(t, x)$ solves $(**)_x$ with $X(0, x) = I$. We need the following version of Lemma 4.

LEMMA 5. *Given t and $n \geq n_0(\tau)$, there exists a $\tilde{t}$ such that, for all x,*

$$|X(\tilde{t}, x - \tilde{t}\omega) - Y_2(x) \cdots Y_{j_n}(x) e^{B_{j_n}(t\omega)} Y_{j_n}^*(x) \cdots Y_2^*(x)| < \varepsilon_{j_n-2}^{\sigma}.$$

Proof. Let n be given and define δ, t and $\tilde{t}$ as in Lemma 4. Let

$$Z_j(x) = Y_2(x) \cdots Y_j(x).$$

Then for all j,

$$\begin{aligned} |\tilde{Y}_{j+1}(\tilde{t}\omega, x - \tilde{t}\omega) - I| &= |Z_j(x - \tilde{t}\omega) Y_{j+1}(x) Y_{j+1}^*(x - \tilde{t}\omega) Z_j^*(x - \tilde{t}\omega) - I| \\ &\leq |Y_{j+1}(x) - Y_{j+1}(x - \tilde{t}\omega)| \\ &\leq |(Y_{j+1}(x) - e^{B_j(x)}) - (Y_{j+1}(x - \tilde{t}\omega) - e^{B_j(x - \tilde{t}\omega)})| \\ &\quad + |e^{B_j(\tilde{t}\omega)} - I|. \end{aligned}$$

The first term is estimated by $(1)_{j+1}$ and the second term can be estimated as in the proof of Lemma 4, for all j except $j = j_n$. For $j = j_n$, we have

$$\begin{aligned} &|\tilde{Y}_{j+1}(\tilde{t}\omega, x - \tilde{t}\omega) - Z_j(x) e^{B_j(t\omega)} Z_j(x)^*| \\ &\quad \leq 2 \sum_{k=1}^{j-1} |(Y_{k+1}(x - \tilde{t}\omega) - e^{B_k(x - \tilde{t}\omega)}) - (Y_{k+1}(x) - e^{B_k(x)})| + |e^{B_k(\tilde{t}\omega)} - I| \\ &\quad\quad + |e^{B_j(t\omega)} - e^{B_j(\tilde{t}\omega)}| + |Y_{j+1}(x) Y_{j+1}^*(x - \tilde{t}\omega) - e^{B_j(\tilde{t}\omega)}|. \end{aligned}$$

The first term is estimated as above, the second term can be estimated as in the proof of Lemma 4, and we must only consider the last term. We can write it as

$$\begin{aligned} &|Y_{j+1}(x) - e^{B_j(\tilde{t}\omega)} Y_{j+1}(x - \tilde{t}\omega)| \\ &\quad = |(e^{-B_j(x)} Y_{j+1}(x) - I) - (e^{-B_j(x - \tilde{t}\omega)} Y_{j+1}(x - \tilde{t}\omega) - I)| \\ &\quad \leq \text{const.} \, \|\tilde{t}\omega\| \frac{1}{r_{j+1}} |e^{-B_j} Y_{j+1} - I|_{r_{j+1}} \\ &\quad \leq \text{const.} \, \|\tilde{t}\omega\| \frac{1}{r_{j+1}} \varepsilon_j^{1/2} |e^{B_j}|_{r_{j+1}}. \end{aligned}$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

However,

$$|e^{B_j}|_{r_{j+1}} \leq e^{N_j r_{j+1}} = e^{-4\sigma \log \varepsilon_j} = e_j^{-4\sigma}.$$

Hence, we get the estimate $\delta\varepsilon_j^{\sigma}$ and now everything goes through as in Lemma 4. □

Consider the scalar product on $so(3, \mathbb{R})$:

$$\langle A, B\rangle = \operatorname{tr}(A^* B)$$

with corresponding norm $|A| = \sqrt{\langle A, A\rangle}$. Let us now assume that F satisfies the following condition:

$$\liminf_{n\to\infty} \frac{|\langle B_{j_n}(\omega), B_{j_{n+1}}(\omega)\rangle|}{|B_{j_n}(\omega)||B_{j_{n+1}}(\omega)|} < 1. \tag{8}$$

LEMMA 6. *Under the assumption* (8)*, there exist two sequences*

$$\tilde{t}_n\omega \to 0, \quad \tilde{s}_n\omega \to 0 \quad \text{in } \mathbb{T}^d$$

such that, for all $x \in \mathbb{T}^d$*,*

$$X(\tilde{t}_n\omega, x - \tilde{t}_n\omega) \quad \text{and} \quad X(\tilde{s}_n\omega, x - \tilde{s}_n\omega)$$

have accumulation points e^{B_x} *and* e^{C_x}*, respectively, with*

$$\frac{|\langle B_x, C_x\rangle|}{|B_x||C_x|} < 1.$$

The eigenvalues $0, \pm i\beta$ *and* $0, \pm i\gamma$ *of* B_x *and* C_x*, respectively, are independent of* x*, and* β *and* γ *are both irrational. Moreover, each subsequence of* $\{n\}$ *has the same property.*

Proof. Consider two sequences

$$t_n\omega \to 0, \quad s_n\omega \to 0 \quad \text{in } \mathbb{T}^d$$

such that $e^{B_{j_n}(t_n\omega)}$ has accumulation point e^B, $B = \beta M^* J_1 M$, with β irrational and $e^{B_{j_{n+1}}(s_n\omega)}$ has accumulation point e^C, $C = \gamma N^* J_1 N$, with γ irrational. By condition (8) there exist sequences such that

$$\frac{|\langle B, C\rangle|}{|B||C|} < 1.$$

For each n, consider $\tilde{t}_n$ and $\tilde{s}_n$ as given by Lemma 5. It follows that for each $x \in \mathbb{T}^d$, $X(\tilde{t}_n\omega, x - \tilde{t}_n\omega)$ has an accumulation point of the form

$$e^{B_x} = Z_x e^B Z_x^*,$$

where Z_x is an accumulation point of $\{Y_2(x)\cdots Y_{j_n}(x)\}_{n=1}^{\infty}$, and $X(\tilde{s}_n\omega, x - \tilde{s}_n\omega)$ has an accumulation point of the form

$$e^{C_x} = Z_x W_x e^C W_x^* Z_x^*,$$

where W_x is an accumulation point of $\{Y_{j_n+1}(x)\cdots Y_{j_{n+1}}(x)\}_{n=1}^{\infty}$. It follows from the form of $Y_j(x)$ that W_x must be an accumulation point of $\{e^{B_{j_n}(x)}\}_{n=1}^{\infty}$, and therefore of the form $M^* e^{\beta_x J_1} M$. Hence, W_x commutes with B. (Notice that we do not claim any measurability

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

of Z_x and B_x in x—only that for each x there are accumulation points of this form.) We can assume without restriction that $M = I$ so $B = \beta J_1$. Then

$$C = a_1 J_1 + a_2 J_2 + a_3 J_3, \quad (a_2, a_3) \neq 0.$$

An easy calculation shows that, for all x,

$$\langle B_x, C_x \rangle = 2a_1\beta, \quad \langle B_x, B_x \rangle = 2\beta^2, \quad \langle C_x, C_x \rangle = 2(a_1^2 + a_2^2 + a_3^2)$$

from which we get the result. The last statement of the lemma is clear from the construction. □

Proof of unique ergodicity under assumption (8). Let $f(x, Q)$ be a continuous function on $\mathbb{T}^d \times SO(3, \mathbb{R})$ and let

$$g(x, Q) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(x + t\omega, X(t, x)Q)\, dt.$$

Since the product Haar measure $\mu \times \nu$ is invariant under the flow $(x, Q) \to Q_t(x, Q)$, it follows that $g(x, Q)$ exists for a.e. (x, Q) and is measurable. Since the flow preserves distances in the $SO(3, \mathbb{R})$-direction it even holds that, for a.e. $x \in \mathbb{T}^d$, $g(x, Q)$ exists for all Q and $Q \mapsto g(x, Q)$ is equicontinuous. Moreover, for a.e. $x \in \mathbb{T}^d$,

$$g(x, Q) = g(x + t\omega, X(t, x)Q), \quad \text{for all } t, Q,$$

or, equivalently,

$$g(x - t\omega, Q) = g(x, X(t, x - t\omega)Q), \quad \text{for all } t, Q.$$

Finally, if $t_j\omega \to 0$ in $\mathbb{T}^d$ then, for all Q,

$$g(x - t_j\omega, Q) \to g(x, Q) \quad \text{in } L^2(\mathbb{T}^d)$$

since as one easily verifies, one has weak convergence and convergence in L^2-norm. In particular we have for each Q, pointwise convergence almost everywhere for some subsequence $\{j_i\}$. Because of equicontinuity, it follows that there exists a set $E_1 \subset \mathbb{T}^d$ of full μ-measure such that

$$g(x - t_{j_i}\omega, Q) \to g(x, Q), \quad (x, Q) \in E_1 \times SO(3, \mathbb{R})$$

where the set E_1 depends, of course, on the subsequence $\{t_{j_i}\}$. Now let

$$W(x) = \{U \in SO(3, \mathbb{R}) : g(x, UQ) = g(x, Q) \text{ for all } Q\}.$$

Clearly $W(x)$ is a closed subgroup for all x in a set E_2 of full μ-measure. Let us now choose sequences

$$t_n\omega \to 0, \quad s_n\omega \to 0 \quad \text{in } \mathbb{T}^d,$$

and let e^{B_x} and e^{C_x} be accumulation points as in Lemma 6. Since we are allowed to take subsequences we may assume that

$$g(x - \tilde{t}_n\omega, Q) \to g(x, Q), \quad g(x - \tilde{s}_n\omega, Q) \to g(x, Q)$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

as $n \to \infty$ for a.e. x and all Q. Then, by Lemma 6, for a.e. x there exists a subsequence $\{n_i\}$ such that

$$\begin{aligned} g(x, Q) &= \lim_{i\to\infty} g(x - \tilde{t}_{n_i}\omega, Q) \\ &= \lim_{i\to\infty} g(x, X(\tilde{t}_{n_i}\omega, x - \tilde{t}_{n_i}\omega)Q) \\ &= g(x, e^{B_x}Q), \quad \text{for all } Q. \end{aligned}$$

In the same way

$$g(x, e^{C_x}Q) = g(x, Q), \quad \text{for all } Q.$$

Hence, $W(x)$ contains e^{B_x} and e^{C_x} with

$$\frac{|\langle B_x, C_x\rangle|}{|B_x||C_x|} < 1,$$

for a.e. x. However, this implies that $W(x) = SO(3, \mathbb{R})$, so g is independent of Q for a.e. x. Now the flow-invariance of g implies the invariance under $x \mapsto x + t\omega$ and

$$\int_{\mathbb{T}^d \times SO(3,\mathbb{R})} f(x, Q)\, d\mu\, d\nu = g(x, Q), \quad \text{a.e. } x, \text{ for all } Q.$$

Now let m be another flow-invariant Borel measure on $\mathbb{T}^d \times SO(3, \mathbb{R})$. Then

$$\tilde{m}(A) = m(A \times SO(3, \mathbb{R})), \quad A \subset \mathbb{T}^d,$$

is invariant under $x \mapsto x + t\omega$. Hence $\tilde{m} = \mu$. It now follows that

$$\begin{aligned} \int_{\mathbb{T}^d \times SO(3,\mathbb{R})} f(x, Q)\, dm &= \int_{\mathbb{T}^d \times SO(3,\mathbb{R})} g(x, Q)\, dm = \int_{\mathbb{T}^d} g(x, Q)\, d\tilde{m} \\ &= \int_{\mathbb{T}^d} g(x, Q)\, d\mu = \int_{\mathbb{T}^d \times SO(3,\mathbb{R})} g(x, Q)\, d\mu\, d\nu \\ &= \int_{\mathbb{T}^d \times SO(3,\mathbb{R})} f(x, Q)\, d\mu\, d\nu. \end{aligned}$$

Since this holds for any continuous function f it follows that $m = \mu \times \nu$. Hence $(*)$ has a unique invariant Borel-measure. □

6. *Genericity*

We must now prove the density of condition (8). This turns out to be surprisingly difficult for the following reason. It is hard to control parameter dependence during the iteration because, if $B, D \in so(3, \mathbb{R})$, then

$$\frac{1}{T}\int_0^T e^{-Bt} D e^{Bt}\, dt = c_D B,$$

where T is the period of $t \mapsto e^{Bt}$. This means that if we put F in an $so(3, \mathbb{R})$-parameter family $D + F$, then the dependence on D of the mean values $A_j(D)$ will be very poor as soon as $j > j_1$—the first index j for which the transformation Y_j is not close to I. This will cause problems in proving the density of condition (8).

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

LEMMA 7. *There exists a constant $C = C(\sigma, \tau, r_1)$ such that in*

$$|F|_{r_1} < C,$$

the set of F which has a representation (1)–(7) *with*

$$B_j(\omega) = \frac{\langle m_j, \omega\rangle}{\alpha_j} A_j = 0, \quad \textit{for all } j > j_0(F)$$

and

$$B_j(\omega) = \frac{\langle m_j, \omega\rangle}{\alpha_j} A_j \neq 0, \quad \textit{for } j = j_0(F)$$

for some $j_0(F)$ is dense. In particular, the set of reducible systems is dense in a neighborhood of zero. Moreover, for each N, the set of such F for which $j_0(F) \geq N$ is also dense.

Proof. Let F be given and consider a representation (1)–(7)—this always exists by Lemma 3. Let $j_1 < j_2 < \cdots$ be all indices j for which $B_j \neq 0$ and let

$$Z_{j+1}(x) = \prod_{k=1}^{j} Y_{k+1}(x), \quad C_j = \prod_{k=1}^{j} e^{B_k}.$$

Fix a j_{n-1} arbitrarily large and consider the perturbation

$$\tilde{F}(x) = F(x) + \delta C_{j_{n-1}} D C_{j_{n-1}}^*$$

with $D \in so(3, \mathbb{R})$, $|D| \leq 1$ and δ small. Then

$$|\tilde{F} - F|_{r_1} < \text{const.}\, \delta |D| \prod_{j=j_1,\ldots,j_{n-1}} e^{2N_j r_1} < \text{const.}\, \delta |D| \varepsilon_1^{-a\sigma}$$

and

$$|Z_{j+1}^*(\tilde{F} - F) Z_{j+1}|_{r_{j+1}} < \text{const.}\, \delta |D| \varepsilon_1^{-a\sigma}$$
$$|\langle Z_{j+1}^*(\tilde{F} - F) Z_{j+1}\rangle - \delta D| < 2\varepsilon\delta |D|,$$

for $j_{n-1} \leq j < j_n$, where $a = 16(2 + 2\sigma)^{j_{n-1}}$. (Here $\langle\cdot\rangle$ stands for the mean value of a function over the torus $\mathbb{T}^d$.) The estimates of derivatives with respect to D are trivial because the expressions are linear in D. Let $j_{n-1} < k < j_n$ be such that

$$\varepsilon_k \geq 2\varepsilon_{j_n}^{\sigma} \varepsilon_1^{-a\sigma} > \varepsilon_{k+1},$$

and define $\delta = \varepsilon_k \varepsilon_1^{a\sigma}$. From this we get

$$\delta \varepsilon_1^{-a\sigma} = \varepsilon_k, \quad \delta \geq 2\varepsilon_{j_n}^{\sigma}, \quad \varepsilon_1 \delta \gg \varepsilon_{k+1}.$$

The last inequality follows because

$$\varepsilon_{k+1} = \varepsilon_k^{1+\sigma} = \delta(\varepsilon_{k+1} \varepsilon_1^{-a(1+\sigma)})^{\sigma/(1+\sigma)} \leq \delta(2\varepsilon_{j_n}^{\sigma} \varepsilon_1^{-a(1+2\sigma)})^{\sigma/(1+\sigma)} \ll \varepsilon_1 \delta$$

for n large enough because of the rapid increase of $\{j_n\}_1^{\infty}$. We can now transform $\tilde{F}$ by the transformation Z_k,

$$\partial_\omega Z_k(x) = \tilde{F}(x) Z_k(x) - Z_k(x)(A_k + F_k(x) + Z_k^*(x)(\tilde{F}(x) - F(x)) Z_k(x)).$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

The transformed system can be written as $A_k(D) + \tilde{F}_k(x, D)$, where

$$A_k(D) = A_k + \langle Z_k^*(\tilde{F} - F)Z_k \rangle.$$

These matrices satisfy

$$|A_k(D)| < \text{const.}, \quad |\partial(A_k(D) - \delta D)| < \text{const.}\, \varepsilon_1 \delta,$$
$$|\partial^\nu \tilde{F}_k|_{r_k} < \text{const.}\, \varepsilon_k, \quad \nu = 0, 1.$$

(The partial derivatives are with respect to the parameters D.) Notice that the mean value of $\tilde{F}_k$ is equal to the mean value of F_k, so it is independent of D. The inequality

$$|\alpha_k - \langle n, \omega \rangle| \geq \varepsilon_k^\sigma, \quad 0 < |n| \leq N_k,$$

implies that

$$|\alpha_k(D) - \langle n, \omega \rangle| \geq \tfrac{1}{2}\varepsilon_k^\sigma, \quad 0 < |n| \leq N_k,$$

since $A_k(0) = A_k$ and

$$|\alpha_k(D) - \alpha_k(0)| \leq \text{const.}\, \delta \leq \tfrac{1}{2}\varepsilon_k^\sigma.$$

We can therefore transform the system by a transformation $\tilde{Y}_{k+1}$—close to I and depending on the parameters D—to $A_{k+1}(D) + \tilde{F}_{k+1}(x, D)$ with

$$|A_{k+1}(D)| < \text{const.}, \quad |\partial(A_{k+1}(D) - \delta D)| < \text{const.}\, \varepsilon_1 \delta,$$
$$|\partial^\nu \tilde{F}_{k+1}|_{r_{k+1}} < \text{const.}\, \varepsilon_{k+1}, \quad \nu = 0, 1.$$

(The second estimate follows because the mean value of $\tilde{F}_k$ is independent of D.) Since $\varepsilon_{k+1} \ll \varepsilon_1 \delta$ we will have good control on the D-derivatives in the following iteration steps as long as we transform the system by a transformation close to I. Since $\delta \geq 2\varepsilon_{j_n}^\sigma$, we can determine D to make sure that for all $j \geq j_n$,

$$|\alpha_j(D) - \langle m, \omega \rangle| \geq \varepsilon_j^\sigma, \quad 0 < |m| \leq N_j.$$

This gives a perturbation $\tilde{F}$ of F of size $\delta\varepsilon_1^{-a\sigma} = \varepsilon_k \leq \varepsilon_{j_{n-1}}$ with the properties stated in the lemma. In order to prove the second part, we fix an N and we assume that $j_{n-1} = j_0(F) < N$. But then, in the same D-parameter family $\tilde{F}$ as above, we will find, for arbitrarily small D, elements with $j_0(\tilde{F}) > j_0(F)$. This proves the second statement by a finite induction. □

LEMMA 8. *There exists a constant $C = C(\sigma, \tau, r_1)$ such that in*

$$|F|_{r_1} < C,$$

for any N, the set of F which has a representation (1)–(7) *with*

$$\langle B_{j_n}(\omega), B_{j_{n-1}}(\omega) \rangle = 0$$

for some $j_n \geq N$ is dense.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

Proof. By Lemma 7 we can suppose that $j_0(F) = j_{n-1} \geq N$. We can suppose without restriction that $B_{j_{n-1}}(\omega) = \gamma J_1$. Now consider the perturbation

$$\tilde{F}(x) = F(x) + \delta\lambda C_{j_{n-2}} J_1 C^*_{j_{n-2}}$$

with $|\lambda| \leq 1$ and δ small. Then

$$|\tilde{F} - F|_{r_1} < \text{const.}\, \delta|\lambda| \prod_{j=j_1,\dots,j_{n-2}} e^{2N_j r_1} < \text{const.}\, \delta|\lambda|\varepsilon_1^{-a\sigma}$$

and

$$|Z^*_{j+1}(\tilde{F} - F)Z_{j+1}|_{r_{j+1}} < \text{const.}\, \delta|\lambda|\varepsilon_1^{-a\sigma}$$
$$|\langle Z^*_{j+1}(\tilde{F} - F)Z_{j+1}\rangle - \delta\lambda J_1| < 2\varepsilon_1\delta|\lambda|,$$

for $j_{n-1} \leq j$, where $a = 16(2 + 2\sigma)^{j_{n-2}}$. The estimates of derivatives with respect to λ are trivial because the expressions are linear in λ. In order for $\tilde{F}$ to be a perturbation of F it is enough that $\delta\varepsilon_1^{-a\sigma}$ is small, which is achieved for, say,

$$\delta = 2\varepsilon^{\sigma}_{j_{n-1}},$$

due to the fast increase of the sequence $\{j_i\}_{i=1}^{n-1}$. Let $k = j_{n-1}$. We can now transform $\tilde{F}$ by the transformation Z_k,

$$\partial_\omega Z_k(x) = \tilde{F}(x)Z_k(x) - Z_k(x)(A_k + F_k(x) + Z^*_k(x)(\tilde{F}(x) - F(x))Z_k(x)).$$

The transformed system can be written as $A_k(\lambda) + \tilde{F}_k(x, \lambda)$, where

$$A_k(\lambda) = A_k + \langle Z^*_k(\tilde{F} - F)Z_k\rangle.$$

These matrices satisfy

$$|A_k(\lambda)| < \text{const.}, \quad |\partial(A_k(\lambda) - \delta\lambda J_1)| < \text{const.}\, \varepsilon_1\delta,$$
$$|\partial^\nu \tilde{F}_k|_{r_k} < \text{const.}\, \tilde{\varepsilon}_k = \text{const.}\, \delta\varepsilon_1^{-a\sigma}, \quad \nu = 0, 1.$$

Since

$$|\alpha_k - \langle m, \omega\rangle| < \varepsilon^{\sigma}_k, \quad \text{some } 0 < |m| \leq N_k,$$

and $\delta = 2\varepsilon^{\sigma}_k$, we can determine $\lambda = \lambda_k$ so that

$$|\alpha_k(\lambda_k) - \langle m, \omega\rangle| = 0$$

and restrict λ to a neighborhood of λ_k—of size ~ 1. We can then transform the system by a transformation $\tilde{Y}_{k+1}$—close to $e^{\langle m,x\rangle J_1}$ and depending on the parameter λ—to $A_{k+1}(\lambda) + \tilde{F}_{k+1}(x, \lambda)$ with

$$|A_{k+1}(\lambda_k)| < \text{const.}, \quad |\partial(A_{k+1}(\lambda) - \delta\lambda J_1)| < \text{const.}\, \varepsilon_1\delta,$$
$$|\partial^\nu \tilde{F}_{k+1}|_{r_{k+1}} < \tilde{\varepsilon}_{k+1} = \tilde{\varepsilon}^{1+\sigma}_k, \quad \nu = 0, 1.$$

(The second estimate follows because the mean value of $\tilde{F}_k$ is independent of λ.) Since $|\langle A_{k+1}(\lambda_k), J_1\rangle| \leq \varepsilon_k$ we can determine $\lambda = \lambda_{k+1}$ so that

$$\langle A_{k+1}(\lambda_{k+1}), J_1\rangle = 0,$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

and restrict λ to a neighborhood of λ_{k+1}—of size $\sim\tilde{\varepsilon}_{k+1}/\delta \ll 1$. Then we repeat the same process again. Since $\tilde{\varepsilon}_{k+1} \ll \varepsilon_1\delta$, we will have good control on the λ-derivatives in the following iteration steps as long as we transform the system by a transformation close to I. This may stop at some step $k = l$, because we must make a transformation which is not close to I, after which we will lose control on the parameter dependence in the J_1-direction. In that case, we determine λ so that the J_1-part of $A_l(\lambda)$ is $= 0$. Then $B_l \neq 0$ and

$$\langle J_1, B_l\rangle = 0,$$

and we are done. Otherwise it does not stop and we can control the J_1-part of $A_k(\lambda)$ for all $k \geq j_{n-1} + 1$. However, this means that we can determine $\lambda = \lambda_\infty$, so that

$$|\langle A_k(\lambda_\infty), J_1\rangle| \leq 2\tilde{\varepsilon}_k, \quad B_k = 0$$

for all $k \geq j_{n-1} + 1$. For an F with this property, however, we can choose a perturbation

$$\tilde{F}(x) = F(x) + C_{j_{n-1}} D C^*_{j_{n-1}}$$

as in Lemma 7 and determine D—arbitrarily small—so that

$$\langle A_k, J_1\rangle = 0, \quad B_k \neq 0$$

for some k sufficiently large. Then in this case we are also done. □

Proof of the Theorem. It follows from Lemma 8 that the set of F, in some small neighborhood

$$|F|_r < C = C(\tau, r),$$

which has a representation (1)–(7) with

$$\liminf_{n\to\infty} \frac{|\langle B_{j_n}(\omega), B_{j_{n-1}}(\omega)\rangle|}{|B_{j_n}(\omega)||B_{j_{n+1}}(\omega)|} = 0$$

is dense in the $|\cdot|_r$-topology. In §5 we have proven that the skew-system $(*)$ defined by such an F is uniquely ergodic. Hence, the set of uniquely ergodic systems is dense. However, it is known [**FH**] that the set of uniquely ergodic systems is always a $\mathcal{G}_\delta$ in the $\mathcal{C}^1$-topology, hence also in the stronger $|\cdot|_r$-topology. This completes the proof of the theorem. □

Acknowledgements. The author would like to thank the Forschungsinstitut für Mathematik at ETH-Zürich for its hospitality during the preparation of this work.

A. *Appendix*

Let $\omega \in \mathbb{R}^d$ and assume that

$$\|\langle n, \omega\rangle\| = \inf_{p\in\mathbb{Z}} |\langle n, \omega\rangle - 2\pi p| \geq \frac{K}{|n|^\tau}, \quad \text{for all } n \in \mathbb{Z}^d \setminus 0,$$

for some $K, \tau > 0$—here we do not require that $\tau > d - 1$.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

LEMMA A.1. *There exists a constant C—independent of τ and d—such that, for all $x \in \mathbb{T}^d$ and all $\delta > 0$, there is a positive integer*

$$k \le \mathcal{N}_d(K,\delta) = \frac{1}{K}\left(\frac{C}{\delta}\right)^{(\tau+1)d(d+1)/2}$$

such that $\|k\omega_i - x_i\| < \delta$ for all $i = 1,\dots,d$.

Proof. By Dirichlet's Theorem, there is a positive integer $k \le (2\pi/\delta)^d$ such that

$$\|k\omega_i\| < \delta, \quad \text{for all } i.$$

Now let $\omega' = k\omega$, mod $(2\pi\mathbb{Z})^d$, such that

$$2\pi\alpha_d = |\omega_1'| \ge |\omega_2'| \ge \cdots \ge |\omega_d'|.$$

Then

$$\|\langle n,\omega'\rangle\| = \|\langle k\cdot n,\omega\rangle\| \ge \frac{K}{k^\tau}\frac{1}{|n|^\tau}, \quad \text{for } n \in \mathbb{Z}^d \setminus 0.$$

Now consider

$$\omega'' = 2\pi\left(\frac{\omega_2'}{\omega_1'},\dots,\frac{\omega_d'}{\omega_1'}\right).$$

Then

$$\begin{aligned}
\|\langle n,\omega''\rangle\| &= \inf_{p\in\mathbb{Z}} |\langle n,\omega''\rangle + p| \\
&= \frac{1}{\alpha_d}\inf_{p\in\mathbb{Z}} |n_2\omega_2' + \cdots + n_d\omega_d' + p\omega_1'| \\
&= \frac{1}{\alpha_d}\inf\{|n_2\omega_2' + \cdots + n_d\omega_d' + 2\pi p\omega_1'| : |p| \le \sqrt{d}|n|\} \\
&\ge \frac{1}{\alpha_d}\inf\left\{\frac{K}{k^\tau}(|n|^2+|p|^2)^{-\tau/2} : |p| \le \sqrt{d}|n|\right\} \\
&\ge \frac{K}{\alpha_d}\frac{1}{(d+1)^{\tau/2}}\left(\frac{\delta}{2\pi}\right)^{d\tau}\frac{1}{|n|^\tau} = \frac{K}{\alpha_d}f_d\frac{1}{|n|^\tau}.
\end{aligned}$$

Moreover,

$$\alpha_d = \|k\omega_1\| \ge K\left(\frac{\delta}{2\pi}\right)^{d\tau} \ge Kf_d.$$

Since $\omega'\mathbb{Z}$ is α_d-dense on the line $\omega\mathbb{R}$ and since the projection of this line onto $\mathbb{T}^d$ intersects $\{0\times\mathbb{T}^{d-1}\}$ in the points $\omega''\mathbb{Z}$, it follows that

$$\mathcal{N}_d(K_d,\delta) \le \left(\frac{2\pi}{\delta}\right)^d \frac{1}{\alpha_d}\mathcal{N}_{d-1}(K_{d-1},\delta), \quad K_{d-1} = K_d\frac{f_d}{\alpha_d}.$$

If $d = 1$, it is easily shown that

$$\mathcal{N}_1(K_1,\delta) \le \frac{1}{K_1}\left(\frac{2\pi}{\delta}\right)^{\tau+1}.$$

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press

Then we get by a finite induction that

$$\mathcal{N}_d(K_d, \delta) \leq \left(\frac{2\pi}{\delta}\right)^{d+(d-1)+\cdots+2+\tau+1} \frac{(2\pi)^{d-1}}{\alpha_d \cdots \alpha_2} \frac{1}{K_1},$$

and

$$K_1 = \frac{f_d \cdots f_2}{\alpha_d \cdots \alpha_2} K_1.$$

Hence,

$$\mathcal{N}_d(K_d, \delta) \leq \left(\frac{2\pi}{\delta}\right)^{d+(d-1)+\cdots+2+\tau+1} \frac{(2\pi)^{d-1}}{f_d \ldots f_2} \frac{1}{K} \leq \left(\frac{C}{\delta}\right)^{(\tau+1)d(d+1)/2} \frac{1}{K}. \qquad \square$$

Now let $\omega \in \mathbb{R}^d$ and assume that

$$|\langle n, \omega \rangle| \geq \frac{K}{|n|^\tau}, \quad \text{for all } n \in \mathbb{Z}^d \setminus 0,$$

for some $K, \tau > 0$. We have the following immediate consequence of Lemma A.1.

COROLLARY A.1. *There exists a constant C—independent of τ and d—such that, for all $x \in \mathbb{T}^d$ and all $\delta > 0$, there is a positive number*

$$T \leq \frac{1}{K} \left(\frac{C}{\delta}\right)^{(\tau+1)(d-1)d/2}$$

such that $\|T\omega_i - x_i\| < \delta$ for all $i = 1, \ldots, d$.

REFERENCES

[E1] L. H. Eliasson. Floquet solutions for the one-dimensional quasi-periodic Schrödinger equation. *Commun. Math. Phys.* **146** (1992), 447–482.

[E2] L. H. Eliasson. Ergodic skew-systems on $SO(3, \mathbb{R})$. *Preprint*, ETH, Zürich, 1991.

[E3] L. H. Eliasson. Reducibility and point spectrum for linear quasi-periodic skew-products. *Proc. ICM (Berlin 1998). Documenta Mathematica* (Extra Volume **2**) (1998), 779–788. http://www.matematik.uni-bielefeld.de/documenta.

[F] A. Fink. *Almost-periodic Differential Equations (Lecture Notes in Mathematics, 377)*. Springer, 1977.

[FH] A. Fahti and M. Herman. Existence de diffeomorphismes minimaux. *Asterisque* **49** (1977), 37–59.

[FSW] J. Fröhlich, T. Spencer and P. Wittver. Localization for a class of one-dimensional quasi-periodic Schrödinger operators. *Commun. Math. Phys.* **132** (1990), 5–25.

[H] M. Herman. Non topological conjugacy of skew products on $SU(2)$. Manuscript, 1989.

[K] R. Krikorian. Reductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts. *Ann. Sci. École Norm. Sup.* **32** (1999), 187–240.

[M] V. M. Millionščikov. Proof of the existence of irregular systems of linear differential equations with almost-periodic coefficients. *Diff. Eqns* **4** (1968), 203–205.

[N] M. G. Nerurkar. On the construction of smooth ergodic skew-products. *Ergod. Th. & Dynam. Sys.* **8** (1988), 311–326.

[R] M. Rychlik. Renormalization of cocycles and linear ODE with almost periodic coefficients. *Invent. Math.* **110** (1992), 173–206.

https://doi.org/10.1017/S0143385702000998 Published online by Cambridge University Press