# EFFECTIVE REDUCIBILITY OF QUASI-PERIODIC LINEAR EQUATIONS CLOSE TO CONSTANT COEFFICIENTS*

ÀNGEL JORBA†, RAFAEL RAMÍREZ-ROS†, AND JORDI VILLANUEVA†

**Abstract.** Let us consider the differential equation

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x, \quad |\varepsilon| \leq \varepsilon_0,$$

where $A$ is an elliptic constant matrix and $Q$ depends on time in a quasi-periodic (and analytic) way. It is also assumed that the eigenvalues of $A$ and the basic frequencies of $Q$ satisfy a diophantine condition. Then it is proved that this system can be reduced to

$$\dot{y} = (A^*(\varepsilon) + \varepsilon R^*(t, \varepsilon))y, \quad |\varepsilon| \leq \varepsilon_0,$$

where $R^*$ is exponentially small in $\varepsilon$, and the linear change of variables that performs such a reduction is also quasi-periodic with the same basic frequencies as $Q$. The results are illustrated and discussed in a practical example.

**Key words.** quasi-periodic Floquet theorem, quasi-periodic perturbations, reducibility of linear equations

**AMS subject classifications.** 34A30, 34C20, 34C27, 34C50, 58F30

**1. Introduction.** The well-known Floquet theorem states that any linear periodic system $\dot{x} = A(t)x$ can be reduced to constant coefficients $\dot{y} = By$ by means of a periodic change of variables. Moreover, this change of variables can be taken over $\mathbb{C}$ with the same period as $A(t)$.

A natural extension is to consider the case in which the matrix $A(t)$ depends on time in a quasi-periodic way. Before beginning the discussion of this issue, let us recall the definition and basic properties of quasi-periodic functions.

DEFINITION 1.1. *A function $f$ is a* quasi-periodic function *with a vector of basic frequencies* $\omega = (\omega_1, \ldots, \omega_r)$ *if* $f(t) = F(\theta_1, \ldots, \theta_r)$, *where $F$ is $2\pi$ periodic in all its arguments and $\theta_j = \omega_j t$ for $j = 1, \ldots, r$. Moreover, $f$ is called* analytic *on a strip of width $\rho$ if $F$ is analytical on an open set containing $|\operatorname{Im} \theta_j| \leq \rho$ for $j = 1, \ldots, r$.*

It is also known that an analytic quasi-periodic function $f(t)$ on a strip of width $\rho$ has Fourier coefficients defined by

$$f_k = \frac{1}{(2\pi)^r} \int_{\mathbb{T}^r} F(\theta_1, \ldots, \theta_r) e^{-(k,\theta)\sqrt{-1}} \, d\theta$$

such that $f$ can be expanded as

$$f(t) = \sum_{k \in \mathbb{Z}^r} f_k e^{(k,\omega)\sqrt{-1}t}$$

for all $t$ such that $|\operatorname{Im} t| \leq \rho/\|\omega\|_\infty$. We denote by $\|f\|_\rho$ the norm

$$\|f\|_\rho = \sum_{k \in \mathbb{Z}^r} |f_k| e^{|k|\rho},$$

---

† Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain (jorba@ma1.upc.es, rafael@tere.upc.es, jordi@tere.upc.es).

and it is not difficult to check that it is well defined for any analytical quasi-periodic function defined on a strip of width $\rho$. Finally, to define an analytic quasi-periodic matrix, we note that all of these definitions hold when $f$ is a matrix-valued function. In this case, to define $\|f\|_\rho$, we use the infinity norm (which will be denoted by $|\cdot|_\infty$) for the matrices $f_k$.

With these definitions and properties, let us return to the problem of the reducibility of a linear quasi-periodic equation $\dot{x} = \widehat{A}(t)x$ to constant coefficients. The approach of this paper is to assume that the system is close to constant coefficients, that is, $\widehat{A}(t) = A + \varepsilon Q(t, \varepsilon)$, where $\varepsilon$ is small. This case has already been considered in many papers (see [2], [8], and [9] among others), and the results can be summarized as follows. Let $\lambda_i$ be the eigenvalues of $A$ and let $\alpha_{ij} = \lambda_i - \lambda_j$ for $i \neq j$. Then if all the values $\operatorname{Re} \alpha_{ij}$ are different from zero, the reduction can be performed for $|\varepsilon| < \varepsilon_0$, $\varepsilon_0$ sufficiently small (see [2]). If some of the $\operatorname{Re} \alpha_{ij}$ are zero (this happens, for instance, if $A$ is elliptic, that is, if all the $\lambda_i$ are on the imaginary axis), more hypotheses are needed—usually these are (i) a diophantine condition involving the $\alpha_{ij}$ and the basic frequencies of $Q(t, \varepsilon)$ and (ii) to assume a nondegeneracy condition with respect to $\varepsilon$ on the corresponding $\alpha_{ij}(\varepsilon)$ of the matrix $A + \varepsilon \overline{Q}(\varepsilon)$ ($\overline{Q}(\varepsilon)$ denotes the average of $Q(t, \varepsilon)$). This allows to prove (see [9] for details) that there exists a Cantorian set $\mathcal{E}$ such that the reduction can be performed for all $\varepsilon \in \mathcal{E}$. Moreover, the relative measure of the set $[0, \varepsilon_0] \setminus \mathcal{E}$ in $[0, \varepsilon_0]$ is exponentially small in $\varepsilon_0$.

Our purpose here is somewhat different. Instead of looking for a total reduction to constant coefficients (this seems to lead us to eliminate a dense set of values of $\varepsilon$; see [8] or [9]), we try to minimize the quasi-periodic part without taking out any value of $\varepsilon$. The result obtained is that the quasi-periodic part can be made exponentially small. Since all of the proof is constructive (and can be carried out with a finite number of steps), it can be applied to practical examples in order to perform an "effective" reduction: if $\varepsilon$ is small enough, the remainder will be so small that, for practical purposes, it can be taken equal to zero. The error produced with this dropping can be easily bounded by means of the Gronwall lemma. Finally, we want to stress that we have also eliminated the nondegeneracy hypothesis of previous papers [8], [9].

Before finishing this introduction, we want to mention a similar result obtained when the dynamics of the system is slow: $\dot{x} = \varepsilon(A + \varepsilon Q(t, \varepsilon))x$. This case is contained in [14], which is an extension of [12]. The result obtained is also that the quasi-periodic part can be made exponentially small in $\varepsilon$. Total reducibility has been also considered in this case: in [15], it is stated that the reduction can be performed except for a set of values of $\varepsilon$ of measure exponentially small.

There are many other results for the reducibility problem. For instance, in the case of the Schrödinger equation with quasi-periodic potential, we mention [3], [4], [5], [10], [11], and [13]. Another classical and remarkable paper is [7], where the general case (that is, without asking to be close to constant coefficients) is considered. Finally, classical results for quasi-periodic systems can be found in [6].

In order to simplify reading, the paper has been divided as follows. Section 2 contains the exposition (without technical details) of the main ideas and methodology, section 3 contains the main theorem, sections 4 and 5 are devoted to the proofs and, finally, section 6 contains an example to show how these results can be applied to a concrete problem.

**2. The method.** The method used is based on the same inductive scheme as [8]. Let us write our equation as

$$\dot{x} = (A + \varepsilon Q(t, \varepsilon))x, \tag{1}$$

where $A$ is an elliptic $d \times d$ matrix and $Q(t, \varepsilon)$ is quasi-periodic with $\omega = (\omega_1, \ldots, \omega_r)$ as vector of basic frequencies and analytic on a strip of width $\rho$. First of all, let us rewrite this equation as

$$\dot{x} = (A_0(\varepsilon) + \varepsilon \widetilde{Q}(t, \varepsilon))x,$$

where $A_0(\varepsilon) = A + \overline{Q}(\varepsilon)$ and $\widetilde{Q}(t, \varepsilon) = Q(t, \varepsilon) - \overline{Q}(\varepsilon)$. Now let us assume that we are able to find a quasi-periodic $d \times d$ matrix $P$ (with the same basic frequencies as $Q$) verifying

$$(2) \qquad \dot{P} = A_0(\varepsilon)P - PA_0(\varepsilon) + \widetilde{Q}(t, \varepsilon)$$

such that $\|\varepsilon P(t, \varepsilon)\|_\sigma < 1$ for some $\sigma > 0$. In this case, it is not difficult to check that the change of variables $x = (I + \varepsilon P(t, \varepsilon))y$ transforms equation (1) into

$$(3) \qquad \dot{y} = (A_0(\varepsilon) + \varepsilon^2(I + \varepsilon P(t, \varepsilon))^{-1}\widetilde{Q}(t, \varepsilon)P(t, \varepsilon))y.$$

Since this equation is similar to (1) but with $\varepsilon^2$ instead of $\varepsilon$, the inductive scheme seems clear: average the quasi-periodic part of (3) and restart this process. The main difficulty that appears in this process comes from equation (2) because the solution contains the denominators $\lambda_i(\varepsilon) - \lambda_j(\varepsilon) + \sqrt{-1}(k, \omega)$, $1 \le i, j \le d$, where $\lambda_i(\varepsilon)$ are the eigenvalues of $A_0(\varepsilon)$. (This is shown in the proof of Lemma 4.2.) This divisor appears in the $k$th Fourier coefficient of $P$. Note that if the values $\lambda_i(\varepsilon) - \lambda_j(\varepsilon)$ are outside the imaginary axis, the (modulus of the) divisor can be bounded from below, making it easy to prove the convergence. On the other hand, the value $\lambda_i(\varepsilon) - \lambda_j(\varepsilon) + \sqrt{-1}(k, \omega)$ can be arbitrarily small, giving rise to convergence problems.

**2.1. Avoiding the small divisors.** Let us begin by assuming that the eigenvalues $\lambda_i$ of the original unperturbed matrix $A$ (see equation (1)) and the basic frequencies of $Q$ satisfy the diophantine condition

$$(4) \qquad |\lambda_i - \lambda_j + \sqrt{-1}(k, \omega)| \ge \frac{c}{|k|^\gamma} \quad \forall k \in \mathbb{Z}^r \setminus \{0\},$$

where $|k| = |k_1| + \cdots + |k_r|$. Note that, in principle, we cannot guarantee that this condition holds in equation (2) because the eigenvalues of $A_0(\varepsilon)$ have been changed with respect to the ones of $A$ (by an amount of $\mathcal{O}(\varepsilon)$) and some of the divisors can be very small or even zero.

The key point is to realize that as the eigenvalues of $A$ move by an amount of $\mathcal{O}(\varepsilon)$ at most, the quantities $\lambda_i(\varepsilon) - \lambda_j(\varepsilon)$ are contained in a (complex) ball $B_{i,j}(\varepsilon)$ centered in $\lambda_i - \lambda_j$ and with radius $\mathcal{O}(\varepsilon)$. Since the center of the ball satisfies condition (4), the values $(k, \omega)$ cannot be inside that ball if $|k|$ is less than some value $M(\varepsilon)$. This implies that it is possible to cancel all of the harmonics such that $0 < |k| < M(\varepsilon)$ because they do not produce small divisors. (Note that we can have resonances only when $(k, \omega)$ is inside $B_{i,j}(\varepsilon)$.) The harmonics with $|k| \ge M(\varepsilon)$ are exponentially small in $M(\varepsilon)$ (when $M(\varepsilon) \to \infty$), that is, exponentially small in $\varepsilon$ (when $\varepsilon \to 0$), so we do not need to eliminate them.

The idea of considering only frequencies less than some threshold $M$ has already been applied in other contexts (see, for instance, [1]).

**2.2. The iterative scheme.** To apply the considerations above, we define, as before, $A_0(\varepsilon) = A + \varepsilon \overline{Q}(\varepsilon)$, $\widetilde{Q}(t, \varepsilon) = Q(t, \varepsilon) - \overline{Q}(\varepsilon)$ and we split $\widetilde{Q}(t, \varepsilon)$ into the sum of two matrices $Q_0(t, \varepsilon)$ and $R_0(t, \varepsilon)$: $Q_0(t, \varepsilon)$ contains the harmonics $Q_k e^{(k, \omega)\sqrt{-1}t}$

with $|k| < M(\varepsilon)$ and $R_0(t, \varepsilon)$ contains those with $|k| \geq M(\varepsilon)$. Therefore, (1) can be rewritten as

$$\dot{x} = (A_0(\varepsilon) + \varepsilon Q_0(t, \varepsilon) + \varepsilon R_0(t, \varepsilon))x. \tag{5}$$

Now the idea is to cancel $Q_0(t, \varepsilon)$ and to leave $R_0(t, \varepsilon)$. (It is already exponentially small with $\varepsilon$.) Therefore, we compute $P_0$ such that

$$\dot{P}_0 = A_0(\varepsilon)P_0 - P_0 A_0(\varepsilon) + Q_0(t, \varepsilon).$$

Then the change $x = (I + \varepsilon P_0(t, \varepsilon))y$ gives

$$\dot{y} = \left[A_0 + \varepsilon^2 (I + \varepsilon P_0)^{-1} Q_0 P_0 + \varepsilon (I + \varepsilon P_0)^{-1} R_0 (I + \varepsilon P_0)\right] y.$$

This equation can be rewritten to be like (5) to repeat the process. Note that the size of the harmonics with $0 < |k| < M(\varepsilon)$ has been squared. As we will see in the proofs, this is enough to guarantee convergence of those terms to zero. Thus the final equation has a purely quasi-periodic part that is exponentially small with $\varepsilon$.

**2.3. Remarks.** It is interesting to note that it is enough to apply a finite number of steps of the inductive process. We do not need to completely cancel the harmonics with $0 < |k| < M(\varepsilon)$, but we can stop the process when they are of the same size as those of $R$. (From the proof, it can be seen that the number of steps needed to achieve this is of order $|\ln|\varepsilon||$.) This allows us to (with the help of a computer) apply this procedure on a practical example.

Another remarkable point concerns the diophantine condition. Note that we need the condition only up to a finite order ($M(\varepsilon)$, which is of order $(1/|\varepsilon|)^{1/\gamma}$, as we shall see in the proofs). This means that in a practical example when the perturbing frequencies are known with finite precision, the diophantine condition can be easily checked.

**3. The theorem.** In what follows, $\mathcal{Q}_d(\rho, \omega)$ stands for the set of analytic quasi-periodic $d \times d$ matrices on a strip of width $\rho$ and that have $\omega$ as their vector of basic frequencies. Moreover, $i$ will denote $\sqrt{-1}$.

THEOREM 3.1. *Consider the equation* $\dot{x} = (A + \varepsilon Q(t, \varepsilon))x$, $|\varepsilon| \leq \varepsilon_0$, $x \in \mathbb{R}^d$, *where we have the following hypotheses:*

1. *$A$ is a constant $d \times d$ matrix with different eigenvalues $\lambda_1, \ldots, \lambda_d$.*
2. *$Q(\cdot, \varepsilon) \in \mathcal{Q}_d(\rho, \omega)$ with $\|Q(\cdot, \varepsilon)\|_\rho \leq q \ \forall |\varepsilon| \leq \varepsilon_0$, for some $\omega \in \mathbb{R}^r$, and where $q, \rho > 0$.*
3. *The vector $\omega$ satisfies the diophantine conditions*

$$|\lambda_j - \lambda_\ell + i(k, \omega)| \geq \frac{c}{|k|^\gamma} \quad \forall k \in \mathbb{Z}^r \setminus \{0\} \quad \forall j, \ell \in \{1, \ldots, d\} \tag{6}$$

*for some constants $c > 0$ and $\gamma > r - 1$. As usual, $|k| = |k_1| + \cdots + |k_r|$.*

*Then there exist positive constants $\varepsilon^*$, $a^*$, $r^*$, and $m$ such that for all $\varepsilon$, $|\varepsilon| \leq \varepsilon^*$, the initial equation can be transformed into*

$$\dot{y} = (A^*(\varepsilon) + \varepsilon R^*(t, \varepsilon))y, \tag{7}$$

*where*

1. *$A^*$ is a constant matrix with $|A^*(\varepsilon) - A|_\infty \leq a^* |\varepsilon|$ and*
2. *$R^*(\cdot, \varepsilon) \in \mathcal{Q}_d(\rho, \omega)$ and $\|R^*(\cdot, \varepsilon)\|_{\rho - \delta} \leq r^* \exp(-(m/|\varepsilon|)^{1/\gamma} \delta) \ \forall \delta \in {]0, \rho]}$.*

*Furthermore, the quasi-periodic change of variables that performs this transformation is also an element of $\mathcal{Q}_d(\rho,\omega)$. Finally, a general explicit computation of $\varepsilon^*$, $a^*$, $r^*$, and $m$ is possible:*

$$\varepsilon^* = \min\left(\varepsilon_0, \frac{\alpha}{eq\beta(3d-1)}\right), \qquad a^* = \frac{eq\beta^2}{e-1}, \qquad r^* = ea^*, \qquad m = \frac{c}{10eq\beta}$$

*where $e = \exp(1)$, $\alpha = \min_{j\neq\ell}(|\lambda_j - \lambda_\ell|)$, and $\beta$ is the condition number of a regular matrix $S$ such that $S^{-1}AS$ is diagonal, that is, $\beta = C(S) = |S^{-1}|_\infty |S|_\infty$.*

*Remark* 3.1. For fixed values of $\lambda_1, \ldots, \lambda_d$ and $\gamma$, hypothesis 3 is not satisfied for any $c > 0$ only for a set of values of $\omega$ of zero measure if $\gamma > r - 1$.

*Remark* 3.2. In case the eigenvalues of the perturbed matrices move on balls of radius $\mathcal{O}(\varepsilon^p)$ (that is, if the nondegeneracy hypothesis needed in [8] or [9] is not satisfied), it is not difficult to show that the bound of the exponential can be improved: $\|R^*(\cdot,\varepsilon)\|_{\rho-\delta} \leq r^* \exp(-(m/|\varepsilon|)^{p/\gamma}\delta)$. The proof is very similar but uses $M(\varepsilon) = (m/|\varepsilon|)^{p/\gamma}$ instead of $(m/|\varepsilon|)^{1/\gamma}$.

Remark 3.2 seems to show that this nondegeneracy hypothesis is not necessary, and it is only used for technical reasons. In fact, the results seem to be better when this hypothesis is not satisfied.

*Remark* 3.3. If the unperturbed matrix $A$ has multiple eigenvalues (that is, if hypothesis 1 is not satisfied), the theorem is still true, but the exponent of $\varepsilon$ in the exponential of the remainder is slightly worse. This happens because the (small) divisors are now raised to a power that increases with the multiplicity of the eigenvalues. The proof is not included since it does not introduce new ideas and the technical details are rather tedious.

*Remark* 3.4. The values of $\varepsilon^*$, $a^*$, $r^*$, and $m$ given in the theorem are rather pessimistic. In the proof, we have used simple (but rough) bounds instead of cumbersome but more accurate ones. If one is interested in realistic bounds for a given problem, the best thing to do is to rewrite the proof for that particular case. We have done this in section 6 where, with the help of a computer program, we have applied some steps of the method to an example. This allows us to obtain not only better bounds but also (numerically) the reduced matrix as well as the corresponding change of variables.

**4. Lemmas.** We will use some lemmas to simplify the proof of Theorem 3.1.

**4.1. Basic lemmas.**
LEMMA 4.1. *Let $Q(t) = \sum_{k\in\mathbb{Z}^r} Q_k e^{i(k,\omega)t}$ be an element of $\mathcal{Q}_d(\rho,\omega)$ and $M > 0$. Let us define $\overline{Q} = Q_0$, $\widetilde{Q}(t) = Q(t) - Q_0$,*

$$Q_{\geq M}(t) = \sum_{\substack{k\in\mathbb{Z}^r \\ |k|\geq M}} Q_k e^{i(k,\omega)t},$$

*and $\widetilde{Q}_{<M} = \widetilde{Q} - Q_{\geq M}$. Then we have the bounds*
  1. $|\overline{Q}|_\infty$, $\|\widetilde{Q}\|_\rho$, $\|\widetilde{Q}_{<M}\|_\rho \leq \|Q\|_\rho$ *and*
  2. $\|Q_{\geq M}\|_{\rho-\delta} \leq \|Q\|_\rho e^{-M\delta} \ \forall\delta \in \,]0,\rho]$.
*Proof.* The proof follows immediately.    ☐

The next lemma is used to control the variation of the eigenvalues of a perturbed diagonal matrix.

LEMMA 4.2. *Let $D$ be a $d\times d$ diagonal matrix with different eigenvalues $\lambda_1, \ldots, \lambda_d$ and $\alpha = \min_{j\neq\ell}(|\lambda_j - \lambda_\ell|)$. Then if $A$ verifies $|A-D|_\infty \leq b \leq \alpha/(3d-1)$, the following conditions hold:*

1. *A has different eigenvalues $\mu_1, \ldots, \mu_d$ and $|\lambda_j - \mu_j| \leq b$ if $j = 1, \ldots, d$.*
2. *There exists a regular matrix $S$ such that $S^{-1}AS = D^* = \mathrm{diag}(\mu_1, \ldots, \mu_d)$ satisfying $C(S) \leq 2$.*

*Proof.* The proof is contained in [8].  ☐

LEMMA 4.3. *Let $(q_n)_n$, $(a_n)_n$, and $(r_n)_n$ be sequences defined by*

$$q_{n+1} = q_n^2, \qquad a_{n+1} = a_n + q_{n+1}, \qquad r_{n+1} = \frac{2 + q_n}{2 - q_n} r_n + q_{n+1}$$

*with initial values $q_0 = a_0 = r_0 = e^{-1}$. Then $(q_n)_n$ is decreasing to zero and $(a_n)_n$ and $(r_n)_n$ are increasing and convergent to some values $a_\infty$ and $r_\infty$, respectively, with $a_\infty < 1/(e-1)$ and $r_\infty < e/(e-1)$.*

*Proof.* It is immediate that $q_n$ goes to zero quadratically, and this implies that $a_n$ is convergent to the value $a_\infty$:

$$a_\infty = \sum_{j=0}^{\infty} q_j < \sum_{j=1}^{\infty} e^{-j} = \frac{1}{e-1}.$$

Then

$$r_n \leq p \left( r_0 + \sum_{j=1}^{n} q_j \right) \leq p a_\infty,$$

where $p = \prod_{j=0}^{\infty}(2 + q_j)/(2 - q_j)$. This product is convergent. In fact,

$$\ln p = \sum_{j=0}^{\infty} \left[ \ln \left( 1 + \frac{q_j}{2} \right) - \ln \left( 1 - \frac{q_j}{2} \right) \right] \leq \frac{3}{2} a_\infty \leq \frac{3}{2(e-1)} < 1,$$

and so $p < e$, where we have used the fact that $\ln(1 + x) \leq x$ and $-\ln(1 - x) \leq 2x$ for $x \in (0, 1/2)$.  ☐

**4.2. The inductive lemma.** The next lemma is used to perform a step of the inductive procedure.

Before stating the result, let us introduce some notation. Let $D$ and $\alpha$ be as in Lemma 4.2 and let $\varepsilon^*$, $q^*$, $L$, and $M(\varepsilon)$ be positive constants. We consider the equation at the step $n$ of the iterative process:

$$(8) \qquad \dot{x}_n = (A_n(\varepsilon) + \varepsilon Q_n(t, \varepsilon) + \varepsilon R_n(t, \varepsilon)) x_n, \quad |\varepsilon| \leq \varepsilon^*,$$

where $Q_n(\cdot, \varepsilon)$, $R_n(\cdot, \varepsilon) \in \mathcal{Q}_d(\rho, \omega)$ and $\overline{Q}_n(\varepsilon) = Q_n(\cdot, \varepsilon)_{\geq M(\varepsilon)} = 0$. We assume that for some $a_n$, $q_n$, $r_n \geq 0$ and $|\varepsilon| < \varepsilon^*$, the following bounds hold:

$$|A_n(\varepsilon) - D| \leq q^* a_n |\varepsilon|, \qquad \|Q_n(\cdot, \varepsilon)\|_\rho \leq q^* q_n, \qquad \|R_n(\cdot, \varepsilon)\|_{\rho - \delta} \leq q^* r_n e^{-M(\varepsilon)\delta},$$

where $\delta$ is such that $0 < \delta \leq \rho$. (The constant $q^*$ has been introduced to simplify the proof of the theorem later on.) We want to see if it is possible to apply a step of the iterative process to equation (8) to obtain

$$(9) \qquad \dot{x}_{n+1} = (A_{n+1}(\varepsilon) + \varepsilon Q_{n+1}(t, \varepsilon) + \varepsilon R_{n+1}(t, \varepsilon)) x_{n+1}, \quad |\varepsilon| \leq \varepsilon^*,$$

such that $Q_{n+1}(\cdot, \varepsilon)$, $R_{n+1}(\cdot, \varepsilon) \in \mathcal{Q}_d(\rho, \omega)$ and $\overline{Q}_{n+1}(\varepsilon) = Q_{n+1}(\cdot, \varepsilon)_{\geq M(\varepsilon)} = 0$. We also want to relate the bounds $a_{n+1}$, $q_{n+1}$, and $r_{n+1}$ of the terms of this equation with the corresponding bounds of equation (8).

LEMMA 4.4. *Let $\lambda_1^{(n)}(\varepsilon), \ldots, \lambda_d^{(n)}(\varepsilon)$ be the eigenvalues of $A_n(\varepsilon)$. Under the previous notation, if*

1. $L \geq 8q^*$, $\varepsilon^* \leq \alpha/q^*(3d-1)$,
2. $a_n \leq 1$, $q_n \leq e^{-1}$, and
3. *the condition*

$$|\lambda_j^{(n)}(\varepsilon) - \lambda_\ell^{(n)}(\varepsilon) + i(k,\omega)| \geq L|\varepsilon|, \quad |\varepsilon| \leq \varepsilon^*,$$

*is satisfied for all $j$ and $\ell$ and for all $k \in \mathbb{Z}^r$ such that $0 < |k| < M(\varepsilon)$,*
*then equation* (8) *can be transformed into* (9) *and*

$$q_{n+1} = q_n^2, \qquad a_{n+1} = a_n + q_{n+1}, \qquad r_{n+1} = \frac{2+q_n}{2-q_n}r_n + q_{n+1}.$$

*The quasi-periodic change of variables that performs this transformation is*

$$(10) \qquad\qquad x_n = (I + \varepsilon P_n(t,\varepsilon))x_{n+1},$$

*where $P_n(\cdot,\varepsilon)$ is the (only) solution of*

$$(11) \qquad\qquad \dot{P}_n = A_n(\varepsilon)P_n - P_nA_n(\varepsilon) + Q_n(t,\varepsilon), \quad \overline{P}_n = 0,$$

*that belongs to $\mathcal{Q}_d(\rho,\omega)$. Moreover, $\|\varepsilon P_n(\cdot,\varepsilon)\|_\rho \leq q_n/2 < 1/2$.*

*Remark* 4.1. $A_n$, $Q_n$, $R_n$, $P_n$, $M$, and $\lambda_j^{(n)}$ depend on $\varepsilon$ but, for simplicity, we will not write this explicitly.

*Proof of Lemma* 4.4. Let us begin by studying the solutions of (11). Let $S_n$ be the matrix found in Lemma 4.2 with $S_n^{-1}A_nS_n = D_n = \mathrm{diag}(\lambda_1^{(n)},\ldots,\lambda_d^{(n)})$ and $C(S_n) \leq 2$. This lemma can be applied because

$$|A_n - D|_\infty \leq q^* a_n|\varepsilon| \leq q^*\varepsilon^* \leq \frac{\alpha}{3d-1} \quad \forall\, |\varepsilon| \leq \varepsilon^*.$$

Making the change of variables $P_n = S_nX_nS_n^{-1}$ and defining $Y_n = S_n^{-1}Q_nS_n$, equation (11) becomes

$$\dot{X}_n = D_nX_n - X_nD_n + Y_n, \quad \overline{Y}_n = 0.$$

Since $D_n$ is a diagonal matrix, we can handle this equation as $d^2$ unidimensional equations, which can be easily solved by expanding in Fourier series. If $X_n = (x_{\ell j,n})$ and $Y_n = (y_{\ell j,n})$ with

$$x_{\ell j,n}(t) = \sum_{\substack{k \in \mathbb{Z}^r \\ 0 < |k| < M}} x_{\ell j,n}^k e^{i(k,\omega)t}, \qquad y_{\ell j,n}(t) = \sum_{\substack{k \in \mathbb{Z}^r \\ 0 < |k| < M}} y_{\ell j,n}^k e^{i(k,\omega)t},$$

the coefficients must be

$$x_{\ell j,n}^k = \frac{y_{\ell j,n}^k}{\lambda_j^{(n)} - \lambda_\ell^{(n)} + i(k,\omega)},$$

and by hypothesis 3, they can be bounded by $|x_{\ell j,n}^k| \leq (L|\varepsilon|)^{-1}|y_{\ell j,n}^k|$, which implies

$$\|P_n\|_\rho \leq C(S_n)\|X_n\|_\rho \leq C(S_n)(L|\varepsilon|)^{-1}\|Y_n\|_\rho \leq C(S_n)^2(L|\varepsilon|)^{-1}\|Q_n\|_\rho$$
$$\leq 4(L|\varepsilon|)^{-1}q^*q_n \leq |\varepsilon|^{-1}\frac{q_n}{2}.$$

Hence $\|\varepsilon P_n\|_\rho \leq q_n/2 < 1/2$. Thus $I + \varepsilon P_n$ is invertible and

$$\|(I + \varepsilon P_n)^{-1}\|_\rho \leq \frac{1}{1 - \|\varepsilon P_n\|_\rho} < 2.$$

Now applying the change of (10) to (8) and defining $Q_n^* = \varepsilon(I+\varepsilon P_n)^{-1}Q_n P_n$, $A_{n+1} = A_n + \varepsilon\overline{Q_n^*}$, $Q_{n+1} = (\widetilde{Q_n^*})_{<M}$, and $R_{n+1} = (I + \varepsilon P_n)^{-1}R_n(I + \varepsilon P_n) + (Q_n^*)_{\geq M}$, it is easy to derive equation (9). Finally, we use Lemma 4.1 to bound the terms of this equation:

$$\|Q_n^*\|_\rho \leq \|(I + \varepsilon P_n)^{-1}\|_\rho\|Q_n\|_\rho\|\varepsilon P_n\|_\rho \leq \|Q_n\|_\rho q_n \leq q^*q_n^2 = q^*q_{n+1},$$
$$\|Q_{n+1}\|_\rho \leq \|Q_n^*\|_\rho \leq q^*q_{n+1},$$
$$|A_{n+1} - D|_\infty \leq |A_n - D|_\infty + |\varepsilon\overline{Q_n^*}|_\infty \leq q^*(a_n + q_{n+1})|\varepsilon| = q^*a_{n+1}|\varepsilon|,$$
$$\|R_{n+1}\|_{\rho-\delta} \leq \frac{1 + \|\varepsilon P_n\|_\rho}{1 - \|\varepsilon P_n\|_\rho}\|R_n\|_{\rho-\delta} + \|(Q_n^*)_{\geq M}\|_{\rho-\delta}$$
$$\leq \left(\frac{1 + q_n/2}{1 - q_n/2}r_n + q_{n+1}\right)q^*e^{-M\delta} = q^*r_{n+1}e^{-M\delta} \quad \forall \delta \in ]0, \rho],$$

and the proof is complete. $\quad\square$

**5. Proof of Theorem 3.1.** Let $S$ be a regular matrix such that $S^{-1}AS = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_d)$. We define $\varepsilon^*$, $\alpha$, $\beta$, and $m$ as in Theorem 3.1. We also define $q^* = e\beta q$, $M = M(\varepsilon) = (m/|\varepsilon|)^{1/\gamma}$, and $L = 8q^*$.

The (constant) change $x = Sx_0$ transforms the initial equation into

$$(12) \qquad\qquad \dot{x}_0 = (D + \varepsilon Q^*(t, \varepsilon))x_0,$$

where $Q^* = S^{-1}QS$ and so $\|Q^*\|_\rho \leq e^{-1}q^*$ for $|\varepsilon| \leq \varepsilon^*$. We split equation (12) as follows:

$$\dot{x} = (A_0 + \varepsilon Q_0(t) + \varepsilon R_0(t))x_0,$$

where $A_0 = D + \varepsilon\overline{Q^*}$, $Q_0 = \widetilde{Q^*}_{<M}$, and $R_0 = Q_{\geq M}^*$. Using Lemma 4.1, it is easy to see that

$$|A_0 - D|_\infty \leq q^*a_0|\varepsilon|, \qquad \|Q_0\|_\rho \leq q^*q_0, \qquad \|R_0\|_{\rho-\delta} \leq q^*r_0e^{-M\delta}$$

$\forall \delta \in ]0, \rho]$ and $|\varepsilon| \leq \varepsilon^*$ if $a_0 = q_0 = r_0 = e^{-1}$.

We will show that in all of the steps, the hypotheses of Lemma 4.4 are satisfied. Since hypothesis 1 and 2 are easy to check, we focus on hypothesis 3.

Now since $a_n \leq 1$, $|\varepsilon| \leq \varepsilon^*$, and $|A_n - D|_\infty \leq q^*|\varepsilon| \leq \alpha/(3d - 1)$, Lemma 4.2 gives that

$$|\alpha_{j\ell}^{(n)} - \alpha_{j\ell}| < 2q^*|\varepsilon| \quad \forall j, \ell, \ |\varepsilon| \leq \varepsilon^*,$$

where $\alpha_{j\ell} = \lambda_j - \lambda_\ell$ and $\alpha_{j\ell}^{(n)} = \lambda_j^{(n)} - \lambda_\ell^{(n)}$, where $\lambda_1^{(n)}, \ldots, \lambda_d^{(n)}$ are the eigenvalues of $A_n(\varepsilon)$.

Using hypothesis 3 of Theorem 3.1, we obtain that if $k \in \mathbb{Z}^r$ and $0 < |k| < M(\varepsilon)$,

$$|\alpha_{j\ell}^{(n)} + i(k, \omega)| \geq |\alpha_{j\ell} + i(k, \omega)| - |\alpha_{j\ell}^{(n)} - \alpha_{j\ell}| > \frac{c}{|k|^\gamma} - 2q^*|\varepsilon|$$
$$> \left(\frac{c}{m} - 2q^*\right)|\varepsilon| = L|\varepsilon|,$$

and hypothesis 3 of Lemma 4.4 is verified.

Consequently, the iterative process can be carried out and Lemma 4.3 ensures the convergence of the process. The composition of all of the changes $I + \varepsilon P_n$ is convergent because $\|I + \varepsilon P_n\|_\rho \leq 1 + q_n/2$. Then the final equation is

$$(13) \qquad \dot{x}_\infty = (A_\infty(\varepsilon) + \varepsilon R_\infty(t, \varepsilon))x_\infty, \quad |\varepsilon| \leq \varepsilon^*,$$

where $|A_\infty(\varepsilon) - D|_\infty \leq q^* a_\infty |\varepsilon| \leq (e\beta/(e-1))q|\varepsilon|$ and

$$\|R_\infty(\cdot, \varepsilon)\|_{\rho-\delta} \leq q^* r_\infty e^{-M(\varepsilon)\delta} \leq \frac{e^2\beta}{e-1} q \exp\left\{-\left(\frac{m}{|\varepsilon|}\right)^{1/\gamma} \delta\right\} \quad \forall\, \delta \in\, ]0, \rho].$$

To complete the proof, the change $x_\infty = S^{-1}y$ transforms equation (13) into equation (7) with the bounds that we were looking for.

**6. An example.** The results of this paper can be applied in many ways according to the kind of problem we are interested in. Let us illustrate this with the help of an example.

Let us consider the equation

$$(14) \qquad \ddot{x} + (1 + \varepsilon q(t))x = 0,$$

where $q(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$ with $\omega_1 = \sqrt{2}$ and $\omega_2 = \sqrt{3}$. Defining $y$ as $\dot{x}$, we can rewrite (14) as

$$(15) \qquad \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ -q(t) & 0 \end{pmatrix}\right] \begin{pmatrix} x \\ y \end{pmatrix}.$$

Since $\lambda_{1,2} = \pm i$, the diophantine condition (6) is satisfied for $\gamma = 1$ (because the frequencies are quadratic irrationals). The value of $c$ will be discussed later. For the sake of simplicity, let us take $\rho = 2$ and $\delta = 1$. This implies that $q = \|Q\|_\rho = 2e^2$. It is not difficult to derive $\beta = 2$ and, finally, $\varepsilon^* = 4.9787\ldots \times 10^{-3}$ and $r^* = 2.5419\ldots \times 10^2$.

The value of $c$ might be calculated for all $k = (k_1,\, k_2)$, but better (larger) values can be used since we need to consider $|k|$ only up to a finite order. For instance, an easy computation shows that for $|k| \leq 125$, $c$ is 0.149. If $|k| = 126$, then $c$ must be at most 0.013 due to the quasi resonance produced by $k = (70, -56)$. In the range $126 \leq |k| \leq 10^5$, there are no more relevant resonances, so the value $c = 0.013$ suffices.

To begin our discussion, let us suppose that the value of $\varepsilon$ in (15) is $\varepsilon = 2 \times 10^{-6}$. If we take $c = 0.149$, we obtain that $m = 1.8545\ldots \times 10^{-4}$ and $M = 93$. (Recall that the process cancels frequencies such that $|k| < M(\varepsilon)$.) If the value of $M$ had been larger than 125, we would have used the value $c = 0.013$ instead. Therefore, we can reduce the system to constant coefficients with a remainder $R^*$ such that $\|R^*\|_{\rho-1} < 10^{-37}$.

If the given value of $\varepsilon$ is smaller—for instance, $\varepsilon = 10^{-7}$—the computed value of $M$ if $c = 0.149$ is 1855, so $c = 0.013$ must be used. This produces $M = 162$ and $\|R^*\|_{\rho-1} < 10^{-67}$. A value of $\varepsilon = 5 \times 10^{-8}$ implies that $M = 324$ and $\|R^*\|_{\rho-1} < 10^{-138}$. The computation of the reduced matrix as well as the quasi-periodic change of variables will be discussed below.

Another interesting problem is the study of reducibility for a value of $\varepsilon$ larger than the $\varepsilon^*$ given above. Let us continue working with the same equation but with $\varepsilon = 0.1$ as our example.

To increase the value of $\varepsilon^*$, one may try to rewrite the proof using optimal bounds at each step. This has not been done here in order to get an easy, clean, and short proof. Instead of doing this, we think that it is much better to rewrite the proof for our example using no bounds but exact values. This will produce the best results for this problem.

For that purpose, we have implemented the algorithm used in the proof of the theorem as a C program for a (given) fixed value of $\varepsilon$. The program computes and performs a finite number of the changes of variables used to prove the theorem. As a result, the reduced system (including the remainder) as well as the final change of variables are written.

To simplify and make the program more efficient, all of the coefficients have been stored as double-precision variables. During all of the operations, all of the coefficients less than $10^{-20}$ have been dropped in order to control the size of the Fourier series that appears during the process. Of course, this introduces some (small) numerical error in the results.[1]

After four changes of variables, (15) is transformed into

$$(16) \qquad \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \left[ \begin{pmatrix} 0.0 & b_{12} \\ b_{21} & 0.0 \end{pmatrix} + R(t) \right] \begin{pmatrix} x \\ y \end{pmatrix},$$

where $b_{12} = 1.000000366251255$ and $b_{21} = -0.992421151834871$. The remainder $R$ is very small: the largest coefficient it contains is less than $10^{-16}$. Note that the accuracy (relative error) of this remainder is very poor due to the use of double-precision arithmetic (15–16 digits) for the coefficients. During the computations, $M$ has not been given a value. Instead, we have tried to cancel all the frequencies with amplitude larger than $10^{-16}$. (It turns out from the computations that all of these frequencies satisfy $|k| \leq 20$.) It is also possible to obtain a better accuracy in the result, using a multiple-precision arithmetic.

Finally, to check the software, we have tabulated a solution of (16) for a timespan of 10 time units. We have transformed this table by means of the (quasi-periodic) change of variables given by the program. Then we have taken the first point of the transformed table as initial condition of (15) to produce (by means of numerical integration) a new table. The differences between these two tables are less than $10^{-13}$, as expected.

Therefore, for practical purposes, this is an "effective" Floquet theorem in the sense that it allows to compute the reduced matrix as well as the change of variables with the usual accuracy used in numerical computations.

REFERENCES

[1] V. I. ARNOL'D, *Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian*, Russian Math. Surveys, 18 (1963), pp. 9–36.

[2] N. N. BOGOLJUBOV, JU. A. MITROPOLISKI, AND A. M. SAMOILENKO, *Methods of Accelerated Convergence in Nonlinear Mechanics*, Springer-Verlag, New York, 1976.

[3] L. CHIERCHIA, *Absolutely continuous spectra of quasiperiodic Schrödinger operators*, J. Math. Phys., 28 (1987), pp. 2891–2898.

---

[1] If one wants to control that error, it is possible to use intervalar arithmetic for the coefficients and carry a bound of the remainder for each Fourier series.

[4]  E. I. DINABURG AND J. G. SINAI, *The one-dimensional Schrödinger equation with quasiperiodic potential*, Funct. Anal. Appl., 9 (1975), pp. 8–21.

[5]  L. H. ELIASSON, *Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation*, Comm. Math. Phys., 146 (1992), pp. 447–482.

[6]  A. M. FINK, *Almost Periodic Differential Equations*, Lecture Notes in Math. 377, Springer-Verlag, Berlin, 1974.

[7]  R. A. JOHNSON AND G. R. SELL, *Smoothness of spectral subbundles and reducibility of quasi-periodic linear differential systems*, J. Differential Equations, 41 (1981), pp. 262–288.

[8]  A. JORBA AND C. SIMÓ, *On the reducibility of linear differential equations with quasiperiodic coefficients*, J. Differential Equations, 98 (1992), pp. 111–124.

[9]  A. JORBA AND C. SIMÓ, *On quasi-periodic perturbations of elliptic equilibrium points*, SIAM J. Math. Anal., 27 (1996), pp. 1704–1737.

[10] J. MOSER AND J. PÖSCHEL, *On the stationary Schrödinger equation with a quasiperiodic potential*, Phys. A, 124 (1984), pp. 535–542.

[11] J. MOSER AND J. PÖSCHEL, *An extension of a result by Dinaburg and Sinai on quasi-periodic potentials*, Comment. Math. Helv., 59 (1984), pp. 39–85.

[12] A. I. NEISHTADT, *The separation of motions in systems with rapidly rotating phase*, J. Appl. Math. Mech., 48 (1984), pp. 133–139.

[13] H. RÜSSMANN, *On the one-dimensional Schrödinger equation with a quasi-periodic potential*, Ann. New York Acad. Sci., 357 (1980), pp. 90–107.

[14] C. SIMÓ, *Averaging under fast quasiperiodic forcing*, in Hamiltonian Mechanics: Integrability and Chaotic Behaviour, NATO ASI Series B: Physics, Plenum Press, New York, 1994, pp. 13–34.

[15] D. TRESHCHEV, *An estimate of irremovable nonconstant terms in the reducibility problem*, Amer. Math. Soc. Transl. Ser. 2, 168 (1995), pp. 91–128.