

Renormalization of planar analytic saddles

Peer-reviewed author version

BONCKAERT, Patrick & VERSTRINGE, Freek (2012) Renormalization of planar analytic saddles. In: Journal of Dynamics and Differential Equations, 24(3), p. 569-593.

Handle: <http://hdl.handle.net/1942/13110>

Renormalization of planar analytic saddles

P. Bonckaert
Universiteit Hasselt
Agoralaan Gebouw D
3590 Diepenbeek
Belgium
patrick.bonckaert@uhasselt.be

F. Verstringe
Universiteit Hasselt
Agoralaan Gebouw D
3590 Diepenbeek
Belgium
freek.verstringe@uhasselt.be

September 9, 2011

Abstract

Let $X = \lambda_1 x_1 \frac{\partial}{\partial x_1} + \lambda_2 x_2 \frac{\partial}{\partial x_2} + O(|x|^2)$ be an analytic vector field near $x = 0$. We suppose that the linear part of this vector field has real eigenvalues λ_1, λ_2 and that the ratio $\eta = -\frac{\lambda_1}{\lambda_2}$ is a positive irrational number. In a previous paper of the first author and P. De Maesschalck, it was shown that any analytic saddle can be conjugated analytically to a form ‘as close as desired’ to the formal normal form. In this paper we will iterate and renormalize these conjugacies. The iteration of this procedure will be strongly connected to the diophantine properties of η and we will establish the convergence of this process. A consequence of this convergence will be the by now classical linearization theorem of Bruno.

MSC classification: 34C20 , 39A99.

Keywords: Dynamical systems, normal forms, renormalization, majorant method.

1 Introduction and framework of the results

Let $X = \lambda_1 x_1 \frac{\partial}{\partial x_1} + \lambda_2 x_2 \frac{\partial}{\partial x_2} + O(|x|^2)$ be an analytic vector field near $x = 0$. Throughout this article we will suppose that the linear part of this vector field has real eigenvalues λ_1, λ_2 and that the ratio $\eta = -\frac{\lambda_1}{\lambda_2} > 0$. We recall from [1, 2] that any analytic saddle can be conjugated analytically to a form ‘as close as desired’ to the formal normal form (see below for precision). For example if η is irrational, then this formal normal form is linear. In this paper we will iterate and renormalize these conjugacies. According to the diophantine properties of

η , we will investigate the convergence of this process. A consequence of this will be the by now classical linearization theorem of Bruno, which we now state in any dimension:

Theorem 1. [3] *Let $X = A + f$ be an analytic vector field near the origin such that $X(0) = 0$, $A = DX(0)$ and $f = O(|x|^2)$. Suppose that the linear part A has real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ that satisfy the following condition*

$$-\sum_{k=1}^{\infty} 2^{-k} \log(\omega(2^k)) < \infty, \quad (1)$$

where $\omega(m) = \inf_{2 \leq |i| \leq m, 1 \leq j \leq n} |\langle \lambda, i \rangle - \lambda_j|$. Then there exists an analytic coordinate transformation $U = id + u$ that conjugates the given vector field X to its linear part A .

The holomorphic linearization results for saddles (i.e. at least one eigenvalue is positive and at least one is negative) were first proven by Siegel [13] and later improved by many authors, see e.g. [12, 3, 9]. The result stated in Theorem 1 is to our knowledge the best known so far and it is still unknown whether the Bruno condition is optimal in dimension higher than or equal to three. The result of Theorem 1 was shown to be optimal in dimension two by Pérez-Marco and Yoccoz in [10] by studying the holonomy map and using results of the classification of one-dimensional mappings by Yoccoz from [14]. Since this approach is rather indirect and depends explicitly on the continued fraction expansion of the number η , it is interesting to connect the analytic ‘near’ normal forms in [1, 2] to Theorem 1 where the continued fraction expansion of the number η plays a central role.

In the original article [3] by Bruno, it is explained that, concerning two dimensional systems, condition (1) can be translated as a condition on continued fractions of the ratio $\eta = -\lambda_1/\lambda_2$. If the continued fraction expansion of the number η is given as $\frac{q_n}{p_n}$, $n \in \mathbb{N}$, then the equivalent condition is expressed as:

$$\sum_{n \geq 2} \frac{\ln(p_{n+1})}{p_n} < \infty. \quad (2)$$

The proof we give is interesting from another point of view. It is essentially an iteration of the reduction procedure that was initially proven in [2]. Let us explain their result first, as we will need a more refined version of it. It should be stressed that this result is very stable: it holds under perturbations of a parameter, even if the parameter enters explicitly in the linear part. We will state this result not in its full generality, but restrict to what is relevant with respect to our problem. We will omit the parameter dependence and the corresponding results concerning formal Gevrey vector fields since we will not need it. The idea of the reduction is that it is possible to remove from a given vector field those monomials that have indices corresponding to a certain ‘good’ set

$$G_C = \{(k, j) \in \mathbb{N}^n \times \{1, 2, \dots, n\} \mid |\langle \lambda, k \rangle - \lambda_j| > C|k|, \text{ and } |k| \geq 2\}. \quad (3)$$

The idea is that monomials $x^k e_j$ corresponding to an index $(k, j) \in G_C$ are ‘sufficiently far away from resonance’. What remains after a suitable analytic

coordinate transformation is a reduced vector field with terms that have indices corresponding to a (cone-shaped) bad set

$$B_C = \{(k, j) \in \mathbb{N}^n \times \{1, 2, \dots, n\} \mid |\langle \lambda, k \rangle - \lambda_j| \leq C|k|, \text{ and } |k| \geq 2\}. \quad (4)$$

More precisely:

Theorem 2. [1] *Let and let $X = A + \sum_{(k,j)} f_{k,j} x^k \frac{\partial}{\partial x_j}$ be an analytic vector field. Here $A = DX(0)$, the linear part of X , is in diagonal form and has eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. There exists an analytic coordinate transform $U = x + \sum_{(k,j)} u_{k,j} x^k \frac{\partial}{\partial x_j}$ containing only terms in the good set, i.e. $u_{k,j} \neq 0 \implies (k, j) \in \mathbf{G}_C$, that transforms X into a analytic vector field $Y = A + \sum_{(k,j)} g_{k,j} x^k \frac{\partial}{\partial x_j}$ and Y contains only bad terms i.e. $g_{k,j} \neq 0 \implies (k, j) \in \mathbf{B}_C$.*

Both the good and the bad set depend on a parameter C . As C tends to zero, the opening angle of the ‘cone’ B_C tends to zero and the reduced systems has far fewer terms, although the radius of convergence of the normal form transformation may tend to zero. In the nonresonant case, the formal limit exists and the corresponding normal form is the linear system $Y = A$, it is however not clear that the corresponding normal form formal transformation actually converges. The intention of this text is to build a transformation $U = \text{id} + u$ as an infinite composition

$$U = (\text{id} + u_2) \circ (\text{id} + u_3) \circ \dots \quad (5)$$

The transformation $\text{id} + u_i$ transforms a vector field containing only bad terms corresponding to the parameter $C = C_i$ to a vector field containing only bad terms corresponding to the parameter $C = C_{i+1}$. The main idea is now to choose an appropriate rate at which C_i goes to zero. This rate depends explicitly on the continued fraction expansion of η . Moreover, at each step we will have to take the flatness of the given reduced form into account. To give the estimate of this flatness is one of the key points in the proof. This again depends explicitly on the continued fraction expansion of η . At each step we will renormalize the given equation. The renormalization operator consists essentially of a rescaling of the equation and the weight of the rescaling is directly related to the flatness of the reduced form. We give a more geometric approach of our ideas afterwards in Section 4. The approach is, as we will explain, completely equivalent.

The method of proof we provide can possibly be extended towards systems of larger dimension. In order to do so, one needs a higher dimensional alternative for the continued fractions. Such alternatives have been developed in relation with Hamiltonian dynamics, see e.g. [4]. Moreover, it is interesting to compare the proofs in [5, 6, 4] to our proofs: they use a similar idea of an iteration-renormalization technique in Fourier space. It is also worth to note that the authors of [4] recently succeeded in extending their scheme to a higher dimensional case by using a more dimensional equivalent of continued fractions.

The structure of this text is as follows: in Section 2 we provide a quantitative version of Theorem 2 in dimension 2. This is necessary because we need a good estimate on the radius of convergence of the reduced form as well as on the corresponding transformation to be able to control the infinite composition (5). In Section 3 we prove our main result. We proceed then by explaining its more

geometric interpretation in Section 4. We shall need some results on continued fractions, some of which are standard, and others are perhaps less known. In order to be self-contained, we add them in an Appendix.

2 One step in the conjugacy procedure

In this entire section we prove the subsequent theorem, that will serve as a tool to make a single step in the renormalization process that we build afterwards. It is essentially a quantified version of Theorem 2, that was originally proven by the use of the implicit function theorem: it will provide us estimates on how large the usually unspecified neighbourhoods are that appear in the implicit function theorem.

Let us first state the result that we prove in this section, the norms that are used in this statement are defined afterwards.

Theorem 3. *Let $C > 0$, $0 < \delta < 1$, $0 < \delta_1 < 2/5$, $\delta_2 = \delta_1/2$, $M = \frac{1}{C\delta}$. We put $f = (f_1, f_2)$, $u = (u_1, u_2)$ and $g = (g_1, g_2)$ and $A(x, y) = (\lambda_1 x, \lambda_2 y)$.*

Suppose furthermore that f is local analytic in the neighbourhood of the origin, $f = O(|x|^2)$ and $\|f\|_{\delta+\delta_1, \mathcal{A}} \leq \min\{\frac{1}{16M^2\delta_1}, \frac{\delta_1}{8M}\}$. Then the given vector field

$$X = (\lambda_1 x + f_1(x, y)) \frac{\partial}{\partial x} + (\lambda_2 y + f_2(x, y)) \frac{\partial}{\partial y} \quad (6)$$

can be conjugated to an analytic reduced form

$$Y = (\lambda_1 x_1 + g_1(x_1, y_1)) \frac{\partial}{\partial x_1} + (\lambda_2 y_1 + g_2(x_1, y_1)) \frac{\partial}{\partial y_1} \quad (7)$$

by means of an analytic near-identity transformation

$$(x, y) = U(x_1, y_1) = (x_1 + u_1(x_1, y_1), y_1 + u_2(x_1, y_1)). \quad (8)$$

Moreover the reduced form and the transformation satisfy the following properties, where B_C and G_C are defined in (4) and (3):

1. $g_i(x, y) = \sum_{(k,i) \in B_C} g_{k,i} x^k$,
2. $u_i(x, y) = \sum_{(k,i) \in G_C} u_{k,i} x^k$,
3. $\|g\|_{\delta, \mathcal{A}} \leq \delta_2/2$,
4. $\|u\|_{\delta, \mathcal{B}} \leq \delta_2/2$.

The proof will be a consequence of Proposition 11 that is proven at the end of this section. Throughout this section we will use the notations as in the statement of the above theorem. The definitions of the norms are as follows:

Definition 4. *Given $\delta > 0$, we define the following normed spaces of analytic functions:*

$$\begin{aligned} \mathcal{A}_\delta &:= \{f = \sum_{|k| \geq 2} a_k x^k \mid \|f\|_{\delta, \mathcal{A}} := \sum_{k \geq 2} |a_k| \delta^{|k|} < \infty\}, \\ \mathcal{B}_\delta &:= \{f = \sum_{|k| \geq 2} a_k x^k \mid \|f\|_{\delta, \mathcal{B}} := \max\{\|f\|_{\delta, \mathcal{A}}, \|Df\|_{\delta, \mathcal{A}}\} < \infty\}, \end{aligned}$$

we will drop the index \mathcal{A} or \mathcal{B} in the notation of the norms whenever no confusion is possible.

We want to transform equation (6) into (7) by means of the analytic coordinate transform (8). It is clear that at the formal level this transformation can be chosen such that $u_1(x_1, y_1) = \sum_{(k,1) \in G_C} u_{k,1} x_1^{k_1} x_2^{k_2}$ and $u_2(x_1, y_1) = \sum_{(k,2) \in G_C} u_{k,2} x_1^{k_1} x_2^{k_2}$. We explain that these series are actually convergent and we give a precise estimate on the radius of convergence of the transformation and their corresponding normal form (7). Let $\mathbf{x} = (x, y)$ and $\mathbf{x}_1 = (x_1, y_1)$. The initial equation (6) can be written in the form $\dot{\mathbf{x}} = A\mathbf{x} + f(\mathbf{x})$, the coordinate transform $\mathbf{x} = \mathbf{x}_1 + u(\mathbf{x}_1)$ turns this into equation $\dot{\mathbf{x}}_1 = (I + Du)^{-1} \cdot (A + f) \circ (I + u)(\mathbf{x}_1) = A\mathbf{x}_1 + g(\mathbf{x}_1)$. Hence

$$\begin{aligned} A + g &= (I + Du)^{-1} \cdot (A + f) \circ (I + u) \\ \Leftrightarrow (I + Du) \cdot (A + g) &= (A + f) \circ (I + u) \\ \Leftrightarrow \mathcal{F}(u, g, f) &= 0, \end{aligned}$$

where \mathcal{F} is the functional

$$\begin{aligned} \mathcal{F} : U \times V \times W &\longrightarrow \mathcal{A}_\delta : (u, g, f) \mapsto Du \cdot (A + g) - A \cdot u + g - f \circ (\text{id} + u), \\ X &:= \{u \in \mathcal{B}_\delta \mid u = \sum_{k \in G_C} u_{kj} x^k e_j\}; \\ U &:= \{u \in X \mid \|u\|_{\delta, \mathcal{B}} < \delta_1\}; \\ V &:= \{g \in \mathcal{A}_\delta \mid g = \sum_{k \in B_C} g_{kj} x^k e_j\}; \\ W &:= \mathcal{A}_{\delta + \delta_1}; \end{aligned} \tag{9}$$

The idea in [1] is to find for a fixed f a solution $u(f)$, $g(f)$ of this functional equation for some $\delta > 0$ by means of an implicit function theorem. In order to do so the authors of [1] proved a somewhat different version of the following lemma.

Lemma 5. *The functional \mathcal{F} is C^1 in a neighbourhood of the origin and the norm of $\mathcal{L} = D_{(u,g)}\mathcal{F}(0, 0, 0)^{-1}$ is bounded by $\frac{1}{C\delta}$, if $\delta < 1$.*

Proof. It follows from [1] that $(u, f, g) \mapsto f \circ (\text{id} + u)$ is C^1 . $Du \cdot A - A \cdot u$ is linear and continuous in u . $Du \cdot g$ is bilinear and continuous. The presence of the term $Du \cdot (A + g)$ is the main reason why a C^1 -norm is chosen for the transformation u . The boundedness of \mathcal{L} follows from:

$$\begin{aligned} \|Du \cdot g\| &= \left\| \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \dots & \frac{\partial u_1}{\partial x_n} \\ \vdots & \dots & \vdots \\ \frac{\partial u_n}{\partial x_1} & \dots & \frac{\partial u_n}{\partial x_n} \end{pmatrix} \cdot g \right\| \\ &= \sum_{\alpha=1}^n \sum_{j=1}^n \left\| \frac{\partial u_\alpha}{\partial x_j} g_j e_\alpha \right\|_\delta \\ &\leq n^2 \max_{\alpha,j} \left\| \frac{\partial u_\alpha}{\partial x_j} \right\|_\delta \cdot \|g_j\|_\delta \\ &\leq n^2 \|Du\|_\delta \cdot \|g\|_\delta. \end{aligned}$$

We calculate the Gateaux derivatives of \mathcal{F} at $(0, 0, 0)$ and find

$$\begin{aligned} D_u \mathcal{F}(0, 0, 0).u &= \lim_{t \rightarrow 0} \frac{tDu.A - tA.u}{t} = Du.A - A.u, \\ D_f \mathcal{F}(0, 0, 0).f &= \lim_{t \rightarrow 0} \frac{tf \circ \text{id}}{t} = f, \\ D_g \mathcal{F}(0, 0, 0).g &= \lim_{t \rightarrow 0} \frac{-tg}{t} = -g. \end{aligned} \tag{10}$$

Hence $D_{(u,g)} \mathcal{F}(0, 0, 0).(\hat{u}, \hat{g}) = D\hat{u}.A - A.\hat{u} - \hat{g}$. We invert this operator. Remark that $\mathcal{L} = D_{(u,g)} \mathcal{F}(0, 0, 0) : X \times V \rightarrow \mathcal{A}_\delta$. Take now $m = \sum_{j=1}^n \sum_{k \in \mathbb{N}^n} m_k x^k e_j \in \mathcal{A}_\delta$. Then $m = w + v$, where $w = \sum_{(k,j) \in \mathbf{G}_C} m_k x^k e_j$ and $v = \sum_{(k,j) \in \mathbf{B}_C} m_k x^k e_j$. Clearly $\mathcal{L}^{-1}.v = -v$, and $u = \mathcal{L}^{-1}.w = \sum_{(k,j) \in \mathbf{G}_C} \frac{m_{k,j}}{\langle \lambda, k \rangle - \lambda_j} x^k e_j$. We have:

$$\begin{aligned} \|u\|_\delta &= \left\| \sum_{(k,j) \in \mathbf{G}_C} \frac{k_i \cdot m_{k,j}}{\langle \lambda, k \rangle - \lambda_j} x^k e_j \right\|_\delta \\ &\leq \frac{1}{C} \left\| \sum_{(k,j) \in \mathbf{G}_C} m_{k,j} x^k e_j \right\|_\delta \\ &\leq \frac{1}{C} \|m\|_\delta. \end{aligned}$$

$$\begin{aligned} \left\| \frac{\partial u}{\partial x_i} \right\|_\delta &= \left\| \sum_{(k,j) \in \mathbf{G}_C} \frac{k_i \cdot m_{k,j}}{\langle \lambda, k \rangle - \lambda_j} x^{k-e_i} e_j \right\|_\delta \\ &\leq \frac{1}{C} \left\| \sum_{(k,j) \in \mathbf{G}_C} m_{k,j} x^{k-e_i} e_j \right\|_\delta \\ &\leq \frac{1}{C} \sum_{(k,j) \in \mathbf{G}_C} |m_{k,j}| \delta^{|k|-1} \\ &\leq \frac{1}{C\delta} \sum_{(k,j) \in \mathbf{G}_C} |m_{k,j}| \delta^{|k|} \\ &\leq \frac{1}{C\delta} \|m\|_\delta. \end{aligned}$$

□

Remark 6. *We have been a bit sloppy in the proof of the C^1 property; but it should be noted that we do not use it below.*

We will use the previous lemma to construct a solution of $\mathcal{F}(u, g, f) = 0$ for a fixed f by means of a fixed point construction.

Suppose therefore that we are given the functional

$$\mathcal{F}(u, g, f) = Du.(A + g) - A.u + g - f \circ (\text{id} + u),$$

defined as in (9) and we want to solve for $u(f), g(f)$ for $\|f\|$ small enough. We remark that the functional is well defined and C^1 in a neighbourhood of

the origin if $\max(\|Du\|_\delta, \|u\|_\delta) < \delta_1$, $\|g\|_\delta$ is well defined and $\|f\|_{\delta+\delta_1}$ is well-defined. Recall that $\mathcal{L} = d\mathcal{F}(0, 0, 0)$. We prove that the mapping

$$T_f(u, g) = -\mathcal{L}^{-1}(Du.g - f \circ (\text{id} + u))$$

has a fixed point $(u, g) = T_f(u, g)$ whenever f is small enough. This implies that $\mathcal{F}(u, g, f) = 0$. We shall use contractive properties of $T_f(u, g)$. Let now $0 < \delta_2 < \delta_1$, $\|u\|_\delta < \frac{\delta_2}{2}$, $\|u'\|_\delta < \frac{\delta_2}{2}$; we will fix the value of δ_2 later on to $\delta_1/2$. We use the mean value theorem and Cauchy estimate and find that

$$\begin{aligned} |f(x + u(x)) - f(x + u'(x))| &\leq \sup_{\|\eta\| \leq \delta + \delta_2} |Df(\eta)| \cdot |u(x) - u'(x)| \\ &\leq \frac{\|f\|_{\delta+\delta_1}}{\delta_1 - \delta_2} \|u - u'\|_\delta. \end{aligned} \quad (11)$$

Now let $\|(u, g)\|_\delta = \max\{\|g\|_\delta, \|u\|_\delta, \|Du\|_\delta\}$, then

$$\begin{aligned} \|T_f(u, g) - T_f(u', g')\|_\delta &\leq \|\mathcal{L}^{-1}\| \cdot (\|Du.g - Du'.g'\|_\delta + \|f \circ (\text{id} + u) - f \circ (\text{id} + u')\|_\delta) \\ &\leq \|\mathcal{L}^{-1}\| \cdot (\|Du\|_\delta \|g - g'\|_\delta + \|Du - Du'\|_\delta \|g'\|_\delta \\ &\quad + \|f \circ (\text{id} + u) - f \circ (\text{id} + u')\|_\delta) \\ &\leq \|\mathcal{L}^{-1}\| \cdot \left(\|Du\|_\delta \|g - g'\|_\delta + \|Du - Du'\|_\delta \|g'\|_\delta \right. \\ &\quad \left. + \frac{\|f\|_{\delta+\delta_1} \|u - u'\|_\delta}{\delta_1 - \delta_2} \right). \end{aligned}$$

We now introduce a majorating scheme. In order to do so, we define the following: $M = \|\mathcal{L}^{-1}\|$; $(u_0, g_0) = T_f(0, 0)$; $(u_{n+1}, g_{n+1}) := T_f(u_n, g_n)$; $\tilde{y}_n := \|T_f(u_n, g_n)\|_\delta$ and $\tilde{x}_n := \|(u_{n+1}, g_{n+1}) - (u_n, g_n)\|_\delta = \|T_f(u_{n+1}, g_{n+1}) - T_f(u_n, g_n)\|_\delta$. Then clearly we have that

$$\begin{aligned} \tilde{y}_{n+1} &= \|T_f(u_{n+1}, g_{n+1})\|_\delta = \|(u_{n+2}, g_{n+2})\|_\delta \\ &\leq \|(u_{n+2}, g_{n+2}) - (u_{n+1}, g_{n+1})\|_\delta + \|(u_{n+1}, g_{n+1})\|_\delta \leq \tilde{x}_{n+1} + \tilde{y}_n \\ \tilde{x}_{n+1} &= \|T_f(u_{n+1}, g_{n+1}) - T_f(u_n, g_n)\|_\delta \\ &\leq M \cdot \left((\|(u_{n+1}, g_{n+1})\|_\delta + \|(u_n, g_n)\|_\delta) \|(u_{n+1} - u_n, g_{n+1} - g_n)\|_\delta \right. \\ &\quad \left. + \frac{\|f\|_{\delta+\delta_1} \|u_{n+1} - u_n\|_\delta}{\delta_1 - \delta_2} \right) \\ &\leq M(\tilde{y}_n \tilde{x}_n + \tilde{y}_{n-1} \tilde{x}_n) + \frac{M\|f\|_{\delta+\delta_1}}{\delta_1 - \delta_2} \tilde{x}_n. \end{aligned} \quad (12)$$

Hence it is natural to consider the recursion

$$\begin{cases} y_0 = x_0 = \|T_f(0, 0)\|_\delta = \|f\|_\delta \\ x_{n+1} = \beta x_n + \alpha x_n y_n, \\ y_{n+1} = x_{n+1} + y_n \end{cases} \quad (13)$$

where $\alpha := 2M$ and $\beta := \frac{M\|f\|_{\delta+\delta_1}}{\delta_1 - \delta_2}$. Remark that whenever we find a convergent, positive solution (x_n, y_n) of the difference equation (13), it will be a majorant of $(\tilde{x}_n, \tilde{y}_n)$, due to the inequalities given by (12). Hence the corresponding

(u_n, g_n) is a Cauchy sequence in the Banach space $\mathcal{A}_\delta \times \mathcal{B}_\delta$ and converges to some (u, g) that is a solution of the fixed point problem $(u, g) = T_f(u, g)$. For sure we have to take into account that for all $n \in \mathbb{N}$ we maintain $y_n \leq \frac{\delta_2}{2}$ during the iteration process to make sure that the Cauchy estimate given by (11) and used in the estimate of equation (12) is valid. We first study the convergence of equation (13) by means of the following theorem. Afterwards we deal with the initial conditions and translate them in terms of the original problem.

Theorem 7. *Consider the recursively defined sequences*

$$\begin{cases} x_{n+1} = \beta x_n + \alpha x_n y_n, \\ y_{n+1} = y_n + \beta x_n + \alpha x_n y_n, \end{cases} \quad (14)$$

where $\alpha > 0$, $0 < \beta < 1$. Suppose that $x_0 = y_0 > 0$ and suppose that for certain $\beta < \beta' < 1$ we have that

$$x_0 \leq \frac{(\beta' - \beta)(1 - \beta')}{\alpha}. \quad (15)$$

Then the sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ are both convergent.

Proof. We show by induction that:

$$x_n \leq (\beta')^n x_0 \text{ and } y_n \leq \sum_{k=0}^n (\beta')^k x_0.$$

This is clearly true for $n = 0$. Suppose that this statement is true for all $0 \leq m \leq n$; then it is also true for $n + 1$. Indeed:

$$y_n \leq \sum_{k=0}^n (\beta')^k x_0 \leq \frac{x_0}{1 - \beta'}, \quad (16)$$

Hence

$$\begin{aligned} x_{n+1} &= x_n(\beta + \alpha y_n) \leq (\beta')^n x_0(\beta + \frac{\alpha x_0}{1 - \beta'}) \\ &\leq (\beta')^{n+1} x_0, \end{aligned}$$

and

$$\begin{aligned} y_{n+1} &= y_n + x_n \leq \sum_{k=0}^n (\beta')^k x_0 + (\beta')^{n+1} x_0 \\ &\leq \sum_{k=0}^{n+1} (\beta')^k x_0, \end{aligned}$$

This concludes the theorem. □

Corollary 8. *Consider the recursively defined sequences*

$$\begin{cases} x_{n+1} = \beta x_n + \alpha x_n y_n, \\ y_{n+1} = y_n + \beta x_n + \alpha x_n y_n, \end{cases}$$

where $\alpha > 0$, $0 < \beta < 1$. Suppose also that $x_0 = y_0 > 0$ and

$$x_0 \leq \frac{(1 - \beta)^2}{4\alpha};$$

Then the sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ are both convergent. Moreover

$$y_n \leq \frac{2x_0}{1 - \beta}. \quad (17)$$

Proof. The condition (15) is equivalent to

$$x_0 \leq \frac{(1 - \beta)^2}{4\alpha};$$

because

$$\frac{(\beta' - \beta)(1 - \beta')}{\alpha}$$

reaches its supremum at $\beta' = \frac{1 + \beta}{2}$. Inequality (17) clearly follows from substituting $\beta' = \frac{1 + \beta}{2}$ in inequality (16) in the proof of the previous lemma. \square

Proposition 9. $(T_f(u_n, g_n))_{n \in \mathbb{N}}$ converges if all of the following three properties hold:

- (i) $\beta < 1$
- (ii) $y_n \leq \frac{\delta_2}{2}$, for each $n \in \mathbb{N}$ (this is necessary to keep $\|u_n\|_\delta \leq \frac{\delta_2}{2}$ and $\|g_n\|_\delta \leq \frac{\delta_2}{2}$).
- (iii) $x_0 = y_0 = \|T_f(0, 0)\|_\delta = M\|f\|_\delta \leq \frac{(1 - \beta)^2}{4\alpha}$

Proof. The above three conditions ensure that the estimates that lead to the majorating equation (13) are valid and that corollary 8 can be applied. \square

Lemma 10. The three properties in Proposition 9 are valid when

- (i) $\|f\|_{\delta + \delta_1} < \frac{\delta_1 - \delta_2}{2M}$,
- (ii) $\|f\|_{\delta + \delta_1} \leq \frac{\delta_2(\delta_1 - \delta_2)}{4M\delta_1}$,
- (iii) $\|f\|_{\delta + \delta_1} \leq \frac{\frac{1}{2} + M(\delta_1 - \delta_2) - \sqrt{M^2(\delta_1 - \delta_2)^2 + M(\delta_1 - \delta_2)}}{M}$.

This simplifies if $\delta_1 = 2\delta_2$ to

- (i) $\|f\|_{\delta + \delta_1} \leq \frac{\delta_1}{8M}$,
- (ii) $\|f\|_{\delta + \delta_1} < \frac{\delta_1}{8M}$,
- (iii) $\|f\|_{\delta + \delta_1} \leq \frac{1 + M\delta_1 - \sqrt{M^2\delta_1^2 + 2M\delta_1}}{2M}$.

Proof. The first property in Proposition 9 is clearly equivalent to $\|f\|_{\delta+\delta_1} < \frac{\delta_1-\delta_2}{2M}$. For the second property we use (we choose $\beta' = \frac{1+\beta}{2}$, in the proof above)

$$y_n \leq x_0 \sum_{k=0}^{\infty} (\beta')^k \leq x_0 \frac{1}{1-\beta'} \leq x_0 \frac{2}{1-\beta}.$$

Hence the second property is valid when

$$\begin{aligned} M\|f\|_{\delta} &= x_0 \leq \frac{\delta_2(1-\beta)}{4} \\ \Leftrightarrow \|f\|_{\delta} &\leq \frac{\delta_2(1-\beta)}{4M} \\ \Leftrightarrow \|f\|_{\delta+\delta_1} &\leq \frac{\delta_2(1-\beta)}{4M} \\ \Leftrightarrow \|f\|_{\delta+\delta_1} &\leq \frac{\delta_2}{4M} \left(1 - \frac{2M\|f\|_{\delta+\delta_1}}{\delta_1-\delta_2}\right) \\ \Leftrightarrow \frac{4M}{\delta_2}\|f\|_{\delta+\delta_1} + \frac{2M\|f\|_{\delta+\delta_1}}{\delta_1-\delta_2} &\leq 1 \\ \Leftrightarrow \|f\|_{\delta+\delta_1} &\leq \frac{\delta_2(\delta_1-\delta_2)}{4M\delta_1}. \end{aligned}$$

Since $\|f\|_{\delta} \leq \|f\|_{\delta+\delta_1}$, the third property is clearly valid if

$$\begin{aligned} \|f\|_{\delta+\delta_1} &\leq \frac{(1-\beta)^2}{4M\alpha} \\ \Leftrightarrow 4M\alpha\|f\|_{\delta+\delta_1} &\leq \left(1 - \frac{2M\|f\|_{\delta+\delta_1}}{\delta_1-\delta_2}\right)^2 \\ \Leftrightarrow 8M^2\|f\|_{\delta+\delta_1} &\leq 1 - \frac{4M\|f\|_{\delta+\delta_1}}{\delta_1-\delta_2} + \frac{4M^2\|f\|_{\delta+\delta_1}^2}{(\delta_1-\delta_2)^2}. \end{aligned} \tag{18}$$

We compute the roots of the equation

$$\begin{aligned} 4M\alpha x &= 1 - \frac{4Mx}{\delta_1-\delta_2} + \frac{4M^2x^2}{(\delta_1-\delta_2)^2} \\ \Leftrightarrow 1 - 4M\alpha x - \frac{4Mx}{\delta_1-\delta_2} + \frac{4M^2x^2}{(\delta_1-\delta_2)^2} &= 0. \end{aligned}$$

We find two positive roots. The smallest is given by

$$\frac{\frac{1}{2} + M(\delta_1 - \delta_2) - \sqrt{M^2(\delta_1 - \delta_2)^2 + M(\delta_1 - \delta_2)}}{M}.$$

Hence condition (18) is valid if

$$\|f\|_{\delta+\delta_1} \leq \frac{\frac{1}{2} + M(\delta_1 - \delta_2) - \sqrt{M^2(\delta_1 - \delta_2)^2 + M(\delta_1 - \delta_2)}}{M}$$

The conditions that appear in case $2\delta_2 = \delta_1$ are obvious. \square

Summarizing the above, we have the following

Proposition 11. $(T_f(u_n, g_n))_{n \in \mathbb{N}}$ converges if $\delta_1 = 2\delta_2$, δ_1 is smaller than $\frac{2}{5}$ and the following properties hold:

$$\|f\|_{\delta+\delta_1} \leq \min\left\{\frac{1}{16M^2\delta_1}, \frac{\delta_1}{8M}\right\}. \quad (19)$$

Here, $M = \frac{1}{C\delta}$.

Proof. The Taylor series of $h(x) = 1 + x - \sqrt{x^2 + 2x}$ at infinity starts as $\frac{1}{8x} + \dots$. Hence it is natural to consider the function $g(x) = 1 + x - \sqrt{x^2 + 2x} - \frac{1}{8x}$. The derivative of this function is $g'(x) = \frac{8x^2\sqrt{x(x+2)} - 8x^3 - 8x^2 + \sqrt{x(x+2)}}{8\sqrt{x(x+2)}x^2}$.

Take now an arbitrary $x \geq 1$. Then we have the estimates

$$\begin{aligned} 8x^2\sqrt{x(x+2)} - 8x^3 - 8x^2 + \sqrt{x(x+2)} &\leq 8x^2(x+2) - 8x^3 - 8x^2 + x + 2 \\ &\leq -6x^3 - 6x^2 + x + 2 \\ &\leq -12x^2 + x + 2 \\ &\leq -11x + 2 \\ &\leq -9, \end{aligned}$$

from which it follows that $g'(x) < 0$; and hence g is decreasing on $[1, \infty[$. Because $\frac{1}{10} < g(1) = \frac{15}{8} - \sqrt{3} < \frac{1}{5}$ and since

$$\lim_{x \rightarrow \infty} g(x) = 1 - \lim_{x \rightarrow \infty} (x - \sqrt{x^2 + 2x}) = 1 - \lim_{x \rightarrow \infty} \frac{-2x}{x + \sqrt{x^2 + 2x}} = 0,$$

it follows that $g(x) \geq 0$ for all $x \geq 1$ i.e. $h(x) \geq \frac{1}{8x}$. Since $h'(x) = 1 - \frac{x+1}{\sqrt{x^2+2x}} = 1 - \frac{\sqrt{x^2+2x+1}}{\sqrt{x^2+2x}} < 0$, h is decreasing on the interval $[0, 1]$ and hence $\frac{1}{10} \leq g(1) \leq h(1) \leq h(y)$, for each $y \in [0, 1]$. We can summarize that for each $x \geq 0$ we have that $\min(\frac{1}{10}, \frac{1}{8x}) \leq h(x)$. Hence also $\min\{\frac{1}{16M^2\delta_1}, \frac{1}{20M}\} \leq \frac{1+M\delta_1-\sqrt{M^2\delta_1^2+2M\delta_1}}{2M}$. We conclude that all three inequalities in Proposition 10 are valid if $\|f\|_{\delta+\delta_1} \leq \min\{\frac{1}{16M^2\delta_1}, \frac{\delta_1}{8M}\}$ and δ_1 is smaller than $\frac{2}{5}$. \square

3 Narrowing the cone

We are now ready to prove the following theorem:

Theorem 12. Let $(\frac{q_n}{p_n})_{n \in \mathbb{N}}$ be the continued fraction approximants of $-\frac{\lambda_1}{\lambda_2}$. Suppose furthermore that

$$\sum_{n \geq 0} \frac{|\ln(C_{n+1})|}{p_{n+1} + q_{n+1}} < \infty, \text{ where } C_{n+1} = \frac{|\lambda_1 p_{n+1} + \lambda_2 q_{n+1}|}{p_{n+1} + q_{n+1} + 1}. \quad (20)$$

Then equation (6) can be linearized to $\lambda_1 x \frac{\partial}{\partial x} + \lambda_2 y \frac{\partial}{\partial y}$ by means of a convergent transformation.

Proof. Since the proof is rather technical we explain the idea of proof first. The idea is to give an iterative approach of the transformation

$$x + u(x) = (x + u_2(x)) \circ (x + u_3(x)) \circ \dots$$

to the normal form by narrowing the cone a little further with each transformation. We will narrow the cones in close relation to the continued fraction expansion $\left(\frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ of the ratio $-\frac{\lambda_1}{\lambda_2}$. Indeed we will consider the subsequent bad sets (consisting of the indices (k, j) of corresponding terms $a_k x^k e_j$ that are not yet removed)

$$B_{C_n} = \{(k, j) \in \mathbb{N}^2 \times \{1, 2\} \mid |k_1 \lambda_1 + \lambda_2 k_2 - \lambda_j| \leq C_n |k_1 + k_2|, \text{ and } |k| \geq 2\}. \quad (21)$$

At each step of the procedure we have that $(x + u_2(x)) \circ (x + u_3(x)) \circ \dots \circ (x + u_{n-1}(x))$ transforms the original vector field determined by $\dot{x} = Ax + f(x)$ to $\dot{x} = Ax + g_{n-1}(x)$. We then put $f_n := g_{n-1}$ and consider at this iteration step n the functional equation:

$$\mathcal{F}_{C_{n+1}}(u_n, g_n, f_n) = 0.$$

Here we have a function f_n that contains only terms in B_{C_n} and we look for a solution (u_n, g_n) such that g_n contains only terms in $B_{C_{n+1}}$. This solution is obtained by applying Theorem 3. Roughly speaking, this proposition establishes two things.

- (i) It provides a bound to the solution

$$\|g_n\|_{\delta_n - \delta_1^{(n)}} \leq \frac{\delta_2^{(n)}}{2} = \frac{\delta_1^{(n)}}{4}, \quad \|u_n\|_{\delta_n - \delta_1^{(n)}} \leq \frac{\delta_2^{(n)}}{2} = \frac{\delta_1^{(n)}}{4}. \quad (22)$$

There is some freedom to choose the $\delta_1^{(n)}$, but one has to take into account a few constraints. First of all we will obtain that $\delta^{(n+1)} = \delta^{(n)} - \delta_1^{(n)}$, and since $\delta_{lim} := \lim_{n \rightarrow \infty} \delta^{(n)}$ will play the role of the radius of convergence of the transformation u_n , we want that $\delta_{lim} > 0$. Hence it makes sense to choose $\delta_1^{(n)} = \frac{\alpha}{2^n}$, where $\alpha < 2\delta$.

- (ii) It provides a solution of $\mathcal{F}_{C_{n+1}}(u_n, g_n, f_n) = 0$ provided that $\|f_n\|_{\delta^{(n)}}$ satisfies the corresponding bound (19).

It is however not always true that $\|f_n\|_{\delta^{(n)}}$ satisfies the bound (19). Hence, we can not apply Theorem 3 directly because $\|f_n\|_{\delta^{(n)}}$ is possibly too large. Instead we will need to rescale the equation as follows.

Define $R_\kappa : \mathbb{C}^n \rightarrow \mathbb{C}^n : x \mapsto \kappa x$. We will extensively use the rescaling operator $\mathcal{R}_\kappa(f) = R_\kappa^{-1} \circ f \circ R_\kappa$. We do this because if we want to find a solution (u, g) of the equation $\mathcal{F}(u, g, f) = 0$ where $\|f\|_{\delta + \delta_1}$ is too large to apply Theorem 3, we can solve $\mathcal{F}(\tilde{u}, \tilde{g}, \mathcal{R}_\kappa(f)) = 0$ instead. It is then readily verified that $u = \mathcal{R}_\kappa^{-1}(\tilde{u})$ and $g = \mathcal{R}_\kappa^{-1}(\tilde{g})$ is a solution of $\mathcal{F}(u, g, f) = 0$. Moreover $\|\mathcal{R}_\kappa(f)\|_{\delta + \delta_1} \leq \kappa^{w-1} \|f\|_{\delta + \delta_1}$, if f is an analytic function for which $Df(0) = 0, \dots, D^{w-1}f(0) = 0$; we call this number w the minimal order of f . This approach has the advantage that Theorem 3 can be applied if we choose κ small enough. Observe that if the minimal order w of f is big, the factor κ can be chosen rather large. To find a lower bound for w at each iteration step is actually the key to the solution. Such a bound is shown in a subsequent technical Lemma 13. We will also need the following technical remark concerning the numbers

C_n : using Lemma 22 of the appendix it follows that

$$\begin{aligned}
& \left| -\frac{\lambda_1}{\lambda_2} - \frac{q_{n+1}}{p_{n+1}} \right| < \left| -\frac{\lambda_1}{\lambda_2} - \frac{q_n}{p_n} \right| \\
\Rightarrow & \left| \frac{\lambda_1 p_{n+1} + \lambda_2 q_{n+1}}{\lambda_2 p_{n+1}} \right| < \left| \frac{\lambda_1 p_n + \lambda_2 q_n}{\lambda_2 p_n} \right| \\
\Rightarrow & \left| \frac{\lambda_1 p_{n+1} + \lambda_2 q_{n+1}}{p_{n+1}} \right| < \left| \frac{\lambda_1 q_n + \lambda_2 p_n}{q_n} \right| \\
\Rightarrow & C_{n+1} = \left| \frac{\lambda_1 p_{n+1} + \lambda_2 q_{n+1}}{q_{n+1} + p_{n+1} + 1} \right| < \left| \frac{\lambda_1 p_n + \lambda_2 q_n}{q_n + p_n + 1} \right| = C_n
\end{aligned}$$

Geometrically this means that $B_{C_{n+1}} \subset B_{C_n}$ for all $n \in \mathbb{N}$.

Let us start the proof of Theorem 12. We will start the procedure from a certain N_0 that will be specified in a subsequent technical Lemma 13. We define for $n \geq N_0$

$$\delta_1^{(n)} = \frac{\alpha}{2^{n-1}}, \quad \delta_2^{(n)} = \frac{\alpha}{2^n}, \quad \alpha < \frac{\delta}{2}.$$

We choose κ_{N_0} small enough in such a way that the norm of $\tilde{f}_{N_0} := \mathcal{R}_{\kappa_{N_0}}(f_{N_0})$ is small enough to apply Theorem 3 to the equation $\mathcal{F}_{C_{N_0+1}}(\tilde{u}_{N_0}, \tilde{f}_{N_0+1}, \tilde{f}_{N_0}) = 0$ with $\delta = \delta^{(N_0)}$ and $\delta_1 = \delta_1^{(N_0)}$. We define for later use

$$\delta^{(N_0)} = \delta - \delta_1^{(N_0)}, \quad \delta^{(n)} = \delta^{(n-1)} - \delta_1^{(n)},$$

for all $n > N_0$. It is readily verified that $\delta_{\lim} = \lim_{n \rightarrow \infty} \delta^{(n)} > 0$. Let $f_{N_0} := f \in \mathcal{A}_\delta$. We obtain a solution satisfying the bounds

$$\begin{aligned}
\|\tilde{u}_{N_0}\|_{\delta^{(N_0)}} &\leq \frac{\delta_2^{(N_0)}}{2} = \frac{\alpha}{2^{N_0+1}}, \\
\|\tilde{f}_{N_0+1}\|_{\delta^{(N_0)}} &\leq \frac{\delta_2^{(N_0)}}{2} = \frac{\alpha}{2^{N_0+1}}.
\end{aligned}$$

Define $u_{N_0} = \mathcal{R}_{\kappa_{N_0}}^{-1}(\tilde{u}_{N_0})$ and $f_{N_0+1} = \mathcal{R}_{\kappa_{N_0}}^{-1}(\tilde{f}_{N_0+1})$.

Suppose now that we have defined $(u_n, \kappa_n, \tilde{u}_n, \tilde{f}_{n+1}, \gamma_n)$ that satisfy

$$\begin{cases} \tilde{u}_n = \mathcal{R}_{\gamma_n}(u_n) \\ \tilde{f}_n = \mathcal{R}_{\gamma_n}(f_n) \\ \gamma_n = \prod_{i=N_0}^n \kappa_i = \gamma_{n-1} \kappa_n \\ \delta^{(n)} = \delta - \sum_{i=N_0}^n \delta_1^{(i)} \\ \|\tilde{u}_n\|_{\delta^{(n)}} \leq \delta_1^{(n)} = \frac{\alpha}{2^{n-1}} \\ \|\tilde{f}_{n+1}\|_{\delta^{(n)}} \leq \delta_1^{(n)} = \frac{\alpha}{2^{n-1}} \\ \mathcal{F}_{C_{n+1}}(\tilde{u}_n, \tilde{f}_{n+1}, \tilde{f}_n) = 0, \end{cases} \quad (23)$$

We define $(u_{n+1}, \kappa_{n+1}, \tilde{u}_{n+1}, \tilde{f}_{n+2}, \gamma_{n+1})$, and show that they satisfy the same equations where n is replaced by $n+1$. It is sufficient to determine κ_{n+1} , \tilde{u}_{n+1} , \tilde{f}_{n+2} , to check the two inequalities and the functional equation. One can define $\gamma_{n+1} = \gamma_n \kappa_{n+1}$, $\mathcal{R}_{\gamma_{n+1}}^{-1}(\tilde{u}_{n+1}) = u_{n+1}$ and $\mathcal{R}_{\gamma_{n+1}}^{-1}(\tilde{f}_{n+2}) = f_{n+2}$ afterwards.

We rescale \tilde{f}_{n+1} to $\mathcal{R}_{\kappa_{n+1}}(\tilde{f}_{n+1})$ in order to be able to apply Theorem 3, with $\delta = \delta^{(n)} = \delta^{(n+1)} + \delta_1^{(n+1)}$, $\delta_1 = \delta_1^{(n+1)} = \frac{\alpha}{2^n}$, $C = C_{n+2}$, $f = \mathcal{R}_{\kappa_{n+1}}(\tilde{f}_{n+1})$. Let w_{n+1} be the minimal order of \tilde{f}_{n+1} . We want κ_{n+1} to be chosen such that

$$\begin{aligned} \|\mathcal{R}_{\kappa_{n+1}}(\tilde{f}_{n+1})\|_{\delta^{(n+1)}} &\leq \|\mathcal{R}_{\kappa_{n+1}}(\tilde{f}_{n+1})\|_{\delta^{(n)}} \leq \kappa_{n+1}^{w_{n+1}-1} \frac{\alpha}{2^{n+2}} \\ &\leq \min \left\{ \frac{C_{n+2}^2 (\delta^{(n)})^2}{16\delta_1^{(n+1)}}, \frac{C_{n+2}\delta^{(n)}\delta_1^{(n)}}{8} \right\}. \end{aligned}$$

A sufficient choice is

$$\kappa_{n+1} := \left(\min \left\{ \frac{2^n C_{n+2}^2 \delta_{lim}^2}{4\alpha}, \frac{2^n C_{n+2} \delta_{lim}}{2\alpha} \right\} \right)^{\frac{1}{w_{n+1}-1}}.$$

We apply Theorem 3 with this choice and find solutions $u = \tilde{u}_{n+1}$, $g = \tilde{f}_{n+2}$ of $\mathcal{F}_{C_{n+2}}(\tilde{u}_{n+1}, \tilde{f}_{n+2}, \mathcal{R}_{\kappa_{n+1}}(f_{n+1})) = 0$, that are bounded by

$$\begin{aligned} \|\tilde{u}_{n+1}\|_{\delta^{(n)}} &\leq \frac{\delta_2^{(n+1)}}{2} = \frac{\delta_1^{(n+1)}}{4} = \frac{\alpha}{2^{n+1}} \leq \frac{\alpha}{2^n} \\ \|\tilde{f}_{n+2}\|_{\delta^{(n)}} &\leq \frac{\delta_2^{(n+1)}}{2} = \frac{\delta_1^{(n+1)}}{4} = \frac{\alpha}{2^{n+1}} \leq \frac{\alpha}{2^n}. \end{aligned}$$

It will be sufficient to have that $\lim_{n \rightarrow \infty} \gamma_n = \gamma$ exists and $\gamma > 0$ in order to finish the proof. We will check afterwards that this condition is equivalent to condition (20). We explain why the formal power series

$$x + u(x) := \lim_{n \rightarrow \infty} v_n(x) := \lim_{n \rightarrow \infty} (x + u_{N_0}(x)) \circ (x + u_{N_0+1}(x)) \circ \dots \circ (x + u_{N_0+n}(x))$$

converges in a neighbourhood of the origin. Remark that $x + u(x)$ is well-defined as a formal power series because the minimal order of u_n tends to infinity whenever n tends to infinity. We show that $\|\mathcal{R}_\gamma(v_n)\|_{\delta_{lim}} \leq D$ for each n and a constant D independent of n . Remark first that

$$\mathcal{R}_\gamma(v_n)(x) = (x + \mathcal{R}_\gamma(u_{N_0})(x)) \circ (x + \mathcal{R}_\gamma(u_{N_0+1})(x)) \circ \dots \circ (x + \mathcal{R}_\gamma(u_{N_0+n})(x)).$$

We use the estimate

$$\begin{aligned} \|(x + \mathcal{R}_\gamma(u_n)(x))\|_{\delta_{lim}} &\leq \|(x + \mathcal{R}_{\gamma_n}(u_{N_0+n})(x))\|_{\delta^{(N_0+n)}} \\ &\leq \delta^{(N_0+n)} + \|\mathcal{R}_{\gamma_n}(u_{N_0+n})(x)\|_{\delta^{(N_0+n)}} \\ &\leq \delta^{(N_0+n)} + \|\tilde{u}_{N_0+n}\|_{\delta^{(N_0+n)}} \\ &= \delta^{(N_0+n)} + \frac{\alpha}{2^{N_0+n-1}} \leq \delta^{(N_0+n-1)}, \end{aligned}$$

and observe that

$$\begin{aligned} \|(x + \mathcal{R}_\gamma(u_{N_0})(x)) \circ (x + \mathcal{R}_\gamma(u_{N_0+1})(x)) \circ \dots \circ (x + \mathcal{R}_\gamma(u_{N_0+n})(x))\|_{\delta^{(N_0+n)}} \\ \leq \|(x + \mathcal{R}_\gamma(u_{N_0})(x))\|_{\delta^{(N_0)}} \leq \delta_{N_0} + \frac{\alpha}{2^{N_0-1}} =: D. \end{aligned}$$

It follows that $\|\mathcal{R}_\gamma(v_n)\|_{\delta_{lim}} \leq \|\mathcal{R}_\gamma(v_n)\|_{\delta^{(N_0+n)}} \leq D$ for each $n \geq 0$. Hence also $\|\mathcal{R}_\gamma(v)\|_{\delta_{lim}} \leq D$ showing that v is analytic on $\bar{B}(0; \delta_{lim})$.

We still need to show that $\lim_{n \rightarrow \infty} \gamma_n = \gamma$ exists and $\gamma > 0$ if condition (20) holds. We will need a sharp bound on w_n first, proven in the subsequent lemma. The existence of the limit is then proven afterwards.

Lemma 13. *Let $l \geq 1$ and consider the formal power series $f_l = \sum_{(k,j) \in \mathbf{B}_{C_l}} a_k x^k e_j$. Define w_l to be the smallest natural number for which $D^{w_l} f_l \neq 0$; then there exists an $N_0 \in \mathbb{N}$ such that $w_l \geq \frac{p_{l+1} + q_{l+1} + 1}{5}$ for each $l \geq N_0$.*

Proof. Throughout the proof we will suppose that $\lambda_1 > 0$ and $\lambda_2 < 0$; the case where $\lambda_1 < 0$ and $\lambda_2 > 0$ is analogous. We give a proof for l odd, i.e. $l = 2n + 1$ for some $n \in \mathbb{N}$. The even case is similar. We define the following lines:

$$\begin{cases} L_1 : & \lambda_1(k_1 - 1) + \lambda_2 k_2 = C_{2n+1}(k_1 + k_2) \\ L_2 : & \lambda_1(k_1 - 1) + \lambda_2 k_2 = -C_{2n+1}(k_1 + k_2), \end{cases}$$

or, equivalently,

$$\begin{cases} L_1 : & k_2 = \frac{C_{2n+1} - \lambda_1}{\lambda_2 - C_{2n+1}} k_1 + \frac{\lambda_1}{\lambda_2 - C_{2n+1}} \\ L_2 : & k_2 = \frac{C_{2n+1} + \lambda_1}{-\lambda_2 - C_{2n+1}} k_1 - \frac{\lambda_1}{\lambda_2 - C_{2n+1}}. \end{cases}$$

L_1 has slope $\frac{C_{2n+1} - \lambda_1}{\lambda_2 - C_{2n+1}} < -\frac{\lambda_1}{\lambda_2}$ and intersection with the line $k_2 = 0$ at the point $P_1 = \left(\frac{-\lambda_1}{C_{2n+1} - \lambda_1}, 0 \right)$. Remark that $\frac{-\lambda_1}{C_{2n+1} - \lambda_1} = \frac{\lambda_1}{-C_{2n+1} + \lambda_1} > 1$. L_2 has slope $\frac{C_{2n+1} + \lambda_1}{-\lambda_2 - C_{2n+1}} > -\frac{\lambda_1}{\lambda_2}$ and intersection with the line $k_2 = 0$ at the point $P_2 = \left(\frac{\lambda_1}{C_{2n+1} - \lambda_1}, 0 \right)$. Remark that $\frac{\lambda_1}{C_{2n+1} - \lambda_1} < 1$. The lines L_1 and L_2 determine the bounds of the set

$$\begin{aligned} \mathbf{B}_{C_{2n+1},1} &= \{k \in \mathbb{N}^2 \mid |(k_1 - 1)\lambda_1 + \lambda_2 k_2| < C_{2n+1}|k_1 + k_2|, \text{ and } |k| \geq 2\}, \\ C_{2n+1} &= \frac{|\lambda_1 p_{2n+1} + \lambda_2 q_{2n+1}|}{p_{2n+1} + q_{2n+1} + 1}. \end{aligned}$$

We define the lines

$$\begin{cases} M_2 : & k_2 = \frac{q_{2n+1}}{p_{2n+1}}(k_1 - 1) \\ M_1 : & k_2 = \frac{q_{2n}}{p_{2n}}(k_1 - 1). \end{cases}$$

M_1 is the line passing through $(1, 0)$ and $P_3 = (p_{2n} + 1, q_{2n})$ and M_2 is the line passing through $(1, 0)$ and $P_4 = (p_{2n+1} + 1, q_{2n+1})$, see also Figure 1. If $n \geq N_0$, then the interior of the triangle with corner points $(1, 0)$, P_2 , P_4 does not contain any point of \mathbb{Z}^2 : indeed, following remark 19 there is no positive integral point (this is a point $P = (x, y)$, where both x, y are positive integers) in the interior of this triangle if the determinant

$$\det \begin{vmatrix} p_{2n+1} & \frac{\lambda_1}{C_{2n+1} + \lambda_1} - 1 \\ q_{2n+1} & 0 \end{vmatrix} = \frac{q_{2n+1} C_{2n+1}}{C_{2n+1} + \lambda_1} < 1.$$

Because $q_{2n+1} C_{2n+1} = q_{2n+1} \frac{|\lambda_1 p_{2n+1} + \lambda_2 q_{2n+1}|}{p_{2n+1} + q_{2n+1} + 1} = |\lambda_1 p_{2n+1} + \lambda_2 q_{2n+1}| \frac{1}{1 - \frac{\lambda_1}{\lambda_2} + \frac{1}{q_{2n+1}}} \rightarrow 0$, if n tends to ∞ , this is true if $n \geq N_0$ for a certain N_0 large enough.

Using the same argumentation one shows that there exists no positive integral point in the interior of the triangle determined by $(1, 0)$, P_1 , P_3 , and hence a fortiori there exists no integral point in the interior of the triangle determined by $(1, 0)$, P_1 and P_6 . Consequently, any positive integral point in $B_{C_{2n+1},1}$ is can be written as

$$(1, 0) + a(p_{2n}, q_{2n}) + b(p_{2n+1}, q_{2n+1}), \text{ where } a, b \in \mathbb{N} \setminus \{0\}.$$

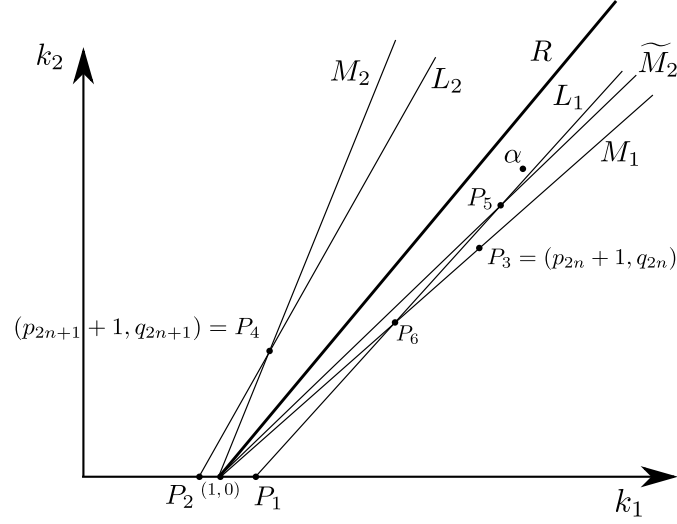


Figure 1: The cones

We define \widetilde{M}_2 to be the reflection of M_2 with respect to the line R ; and $P_5 = \widetilde{M}_2 \cap R$. Hence each positive integral point that lies in the set $B_{C_{2n+1},1}$ that is determined by the lines L_1 and L_2 lies either in the triangle Δ with corner points $(1,0), P_5, P_1$ or in the cone K_2 determined by the lines M_2 and \widetilde{M}_2 . Let now $P_5 = (a, b)$. Consider now any point $P = (a_1, b_1)$ in Δ ; then, because the slopes of \widetilde{M}_2 and L_1 are positive, we clearly have that $a_1 + b_1 \leq a + b$. We compute \widetilde{M}_2 and P_5 . Since \widetilde{M}_2 is the reflection of M_2 with respect to R , it is sufficient to compute the reflexion of $(p_{2n+1} + 1, q_{2n+1}) \in M_2$. Define $\eta = -\frac{\lambda_1}{\lambda_2}$, then $R : k_2 = \eta(k_1 - 1)$ and $n = (-\eta, 1)/\sqrt{1 + \eta^2}$ is a unit vector that is perpendicular to R . Hence the reflexion of the point $(p_{2n+1} + 1, q_{2n+1})$ is given by

$$\left(p_{2n+1} + 1 + \frac{2\eta}{1 + \eta^2} (-p_{2n+1}\eta + q_{2n+1}), q_{2n+1} - \frac{2}{1 + \eta^2} (-p_{2n+1}\eta + q_{2n+1}) \right).$$

It follows that $\widetilde{M}_2 : k_2 = \beta(k_1 - 1)$, where

$$\beta = \frac{q_{2n+1}(1 + \eta^2) - 2(-p_{2n+1}\eta + q_{2n+1})}{p_{2n+1}(1 + \eta^2) + 2\eta(-p_{2n+1}\eta + q_{2n+1})}.$$

We find now

$$P_5 = (a, b) = \left(\frac{-\lambda_2\beta - \lambda_1 + C_{2n+1}\beta}{-\lambda_2\beta + C_{2n+1} - \lambda_1 + C_{2n+1}\beta}, \frac{-C_{2n+1}\beta}{-\lambda_2\beta + C_{2n+1} - \lambda_1 + C_{2n+1}\beta} \right).$$

A straightforward computation results in

$$a + b = \frac{(p_{2n+1} + q_{2n+1} + 1)(\lambda_1^2 + \lambda_2^2)}{\lambda_1^2 + \lambda_2^2 + 2\lambda_1(p_{2n+1}\lambda_1 + q_{2n+1}\lambda_2) + 2\lambda_2(p_{2n+1}\lambda_1 + q_{2n+1}\lambda_2)}.$$

It follows that

$$\lim_{n \rightarrow \infty} \frac{a+b}{(p_{2n+1} + q_{2n+1} + 1)} = 1.$$

As a consequence of this it follows that there exists an $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ and for all $P = (a_1, a_2) \in \Delta$ we have that

$$a_1 + a_2 \leq a + b < 2p_{2n+1} + 2q_{2n+1} + 1. \quad (24)$$

We consider two cases:

Case 1: There exist positive integral points in Δ .

Consider such an arbitrary point. We already observed that such a point equals

$$(1 + ap_{2n+1} + bp_{2n}, aq_{2n+1} + bq_{2n}),$$

for some $a, b \in \mathbb{N} \setminus \{0\}$. The weight (=the sum of the coordinates) of this point is given by $1 + ap_{2n+1} + bp_{2n} + aq_{2n+1} + bq_{2n}$. Since this point lies in Δ , (24) is valid, and hence $a = 1$. Hence the point with lowest weight in Δ is given by $(1 + p_{2n+1} + p_{2n}, q_{2n+1} + q_{2n})$. Clearly $Q = (1 + 2p_{2n+1} + p_{2n}, 2q_{2n+1} + q_{2n}) \notin \Delta$. Because the slope of the line determined by $(p_{2n} + 1, q_{2n})$ and \mathbb{Q} is strictly larger than the slope of \widetilde{M}_2 , every point $(\widetilde{c}p_{2n+1} + p_{2n}, \widetilde{c}q_{2n+1} + q_{2n})$ with $\widetilde{c} \geq 1$ is either in Δ or in the cone K_1 determined by the lines M_2 and \widetilde{M}_2 . As a consequence Q must lie in the cone K_1 (since it is not in Δ). Using Lemma 26 from the appendix, we know that if $c \in \mathbb{N}$ and $2c > a_{2n+2} + 1$, then $\left| \frac{cq_{2n+1} + q_{2n}}{cp_{2n+1} + p_{2n}} - \eta \right| \leq \left| \frac{q_{2n+1}}{p_{2n+1}} - \eta \right|$ and hence $(cp_{2n+1} + p_{2n} + 1, cq_{2n+1} + q_{2n}) \in K_1$. Again from Lemma 26 it follows that if $c \in \mathbb{N}$ and $2c \leq a_{2n+2} - 1$, then $(cp_{2n+1} + p_{2n} + 1, cq_{2n+1} + q_{2n}) \notin K_1$. Because $Q = (1 + 2p_{2n+1} + p_{2n}, 2q_{2n+1} + q_{2n}) \in K_1$, we must have that $2.2 > a_{2n+1} - 1$, equivalent with $a_{2n+1} < 5$. The point with lowest weight in $B_{C_{2n+1},1}$ is in this case $\alpha = (p_{2n+1} + p_{2n} + 1, q_{2n+1} + q_{2n})$ and has weight $p_{2n+1} + p_{2n} + q_{2n+1} + q_{2n} + 1$. This weight satisfies

$$\begin{aligned} p_{2n+1} + p_{2n} + q_{2n+1} + q_{2n} + 1 &\geq \frac{1}{5} (5p_{2n+1} + p_{2n} + 5q_{2n+1} + q_{2n}) + 1 \\ &\geq \frac{1}{5} (p_{2n+2} + q_{2n+2}) + 1. \end{aligned}$$

Case 2: There are no positive integral points in Δ .

All positive integral points in $\mathbf{B}_{C_{2n+1},1}$ lie in the open cone K_1 determined by M_2 and \widetilde{M}_2 . Since clearly $K_1 = \left\{ (a+1, b) \in \mathbb{N}^2 \mid \left| \eta - \frac{b}{a} \right| < \left| \eta - \frac{q_{2n+1}}{p_{2n+1}} \right| \right\}$. The cone K_1 is contained in the cone $K_2 = \{(1, 0) + a(p_{2n}, q_{2n}) + b(p_{2n+1}, q_{2n+1}) \mid a, b \geq 1\}$ determined by the lines M_2 and M_1 . Using Lemma 26 we see that each point $(1, 0) + (cp_{2n+1} + p_{2n}, cq_{2n+1} + q_{2n})$ is in K_1 if $2c > a_{2n+2} + 1$ and is not in K_1 if $2c \leq a_{2n+2} - 1$. It follows that the point with lowest weight in K_1 is given by $(1, 0) + (cp_{2n+1} + p_{2n}, cq_{2n+1} + q_{2n})$, for a certain $c \in \mathbb{N}$ that is larger than $\max\{1, \frac{a_{2n+2}-1}{2}\} \geq \frac{a_{2n+2}}{3}$. Hence

$$\begin{aligned} w_{2n+1} &\geq 1 + \frac{a_{2n+2}}{3} p_{2n+1} + p_{2n} + \frac{a_{2n+2}}{3} q_{2n+1} + q_{2n} \\ &\geq \frac{p_{2n+2} + q_{2n+2}}{3} + 1 \geq \frac{p_{2n+2} + q_{2n+2}}{5} + 1. \end{aligned}$$

Hence the set $\mathbf{B}_{C_{2n+1},1}$ has minimal weight larger than $\frac{1+p_{2n+2}+q_{2n+2}}{5}$. Completely analogous one shows that the same is true for the set $\mathbf{B}_{C_{2n+1},2} = \{(k, 2) \in \mathbf{B}_{C_{2n+1}}\}$, finishing the proof of Lemma 13. \square

We finish the proof of Theorem 12 by showing that $\lim_{n \rightarrow \infty} \gamma_n = \gamma > 0$. Note that $\gamma_n = \prod_{k=N_0}^n \kappa_k$, and $\ln(\gamma_n) = \sum_{k=N_0}^n \ln(\kappa_k)$. κ_n is defined by

$$\kappa_n = \left(\min \left\{ \frac{2^{n-1} C_{n+1}^2 (\delta_{lim})^2}{4\alpha^2}, \frac{2^{n-1} C_{n+1} \delta_{lim}}{2\alpha} \right\} \right)^{\frac{1}{w_n-1}}.$$

We have that for n large enough

$$\kappa_n \leq \left(\frac{2^{n-1} C_{n+1}^2 (\delta_{lim})^2}{4\alpha^2} \cdot \frac{2^{n-1} C_{n+1} \delta_{lim}}{2\alpha} \right)^{\frac{1}{w_n-1}}.$$

It is hence sufficient to show that the two sums in the right hand side of

$$\sum_{k=N_0}^{\infty} \ln(\kappa_k) = \sum_{n \geq N_0} \frac{1}{w_n-1} \ln \left(\frac{4^n \delta_{lim}^3}{2\alpha^3} \right) + \sum_{n \geq N_0} \frac{1}{w_n-1} \ln(C_{n+1}^3) \quad (25)$$

converge if condition (20) holds. We use the previous lemma to see that $w_n - 1 \geq p_n + q_n$ in both sums. We consider the first sum. Since $w_n - 1 \geq \frac{p_{n+1} + q_{n+1}}{5}$ and p_{n+1}, q_{n+1} go to infinity at least as fast as the Fibonacci sequence, the first sum is obviously convergent. For the second sum we have

$$\left| \sum_{n \geq N_0} \frac{1}{w_n-1} \ln(C_{n+1}^3) \right| \leq 5 \sum_{n \geq N_0} \frac{|\ln(C_{n+1})|}{p_{n+1} + q_{n+1}},$$

which converges by supposition. \square

We finish this section by showing that the condition on $-\frac{\lambda_1}{\lambda_2}$ appearing in the heading of Theorem 12 holds if the Bruno condition holds.

Proposition 14. *The Bruno condition (2) implies*

$$\sum_{n \geq 0} \frac{|\ln(C_{n+1})|}{p_{n+1} + q_{n+1}} < \infty, \text{ where } C_{n+1} = \frac{|\lambda_1 p_{n+1} + \lambda_2 q_{n+1}|}{p_{n+1} + q_{n+1} + 1}.$$

Proof. We have

$$\begin{aligned} - \sum_{n \geq 0} \frac{\ln(C_{n+1})}{p_{n+1} + q_{n+1}} &= - \sum_{n \geq 0} \frac{\ln \left(\frac{|\lambda_1 p_{n+1} + \lambda_2 q_{n+1}|}{p_{n+1} + q_{n+1} + 1} \right)}{p_{n+1} + q_{n+1}} \\ &= - \sum_{n \geq 0} \frac{\ln(|\lambda_1 p_{n+1} + \lambda_2 q_{n+1}|)}{p_{n+1} + q_{n+1}} + \sum_{n \geq 0} \frac{\ln(p_{n+1} + q_{n+1} + 1)}{p_{n+1} + q_{n+1}}. \end{aligned}$$

We investigate the two sums on the right hand side of this equation separately. The last sum converges because $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ increase at least as fast as the Fibonacci series. For the first sum we use Lemma 21 which implies that

$$|\lambda_1 p_{n+1} + \lambda_2 q_{n+1}| \leq \frac{|\lambda_2|}{p_{n+2}}.$$

We also use Lemma 20 and obtain

$$\frac{1}{p_{n+1} + q_{n+1}} = \frac{1}{p_{n+1} \left(1 + \frac{q_{n+1}}{p_{n+1}}\right)} \leq \frac{1}{p_{n+1} \left(1 + \frac{q_0}{p_0}\right)}.$$

Hence

$$-\sum_{n \geq 0} \frac{\ln(|C_{n+1}|)}{p_{n+1} + q_{n+1}} \leq \left(\frac{1}{1 + \frac{q_0}{p_0}}\right) \sum_{n \geq 1} \frac{\ln(p_{n+2})}{p_{n+1}} + \ln(|\lambda_2|) \sum_{n \geq 1} \frac{1}{p_{n+1}}.$$

It follows that if the Bruno condition $\sum_{n \geq 1} \frac{\ln(p_{n+2})}{p_{n+1}} < \infty$ holds, then

$$\sum_{n \geq 0} \frac{|\ln(C_{n+1})|}{p_{n+1} + q_{n+1}} < \infty.$$

□

4 A more geometric interpretation of the proof

Let $X = (\lambda_1 x + f_1(x, y)) \frac{\partial}{\partial x} + (\lambda_2 y + f_2(x, y)) \frac{\partial}{\partial y}$. We already showed that it is possible to find a coordinate transformation that transforms X to

$$X_{C_n} = (\lambda_1 x + f_1^{C_n}(x, y)) \frac{\partial}{\partial x} + (\lambda_2 y + f_2^{C_n}(x, y)) \frac{\partial}{\partial y},$$

where

$$f_i^{C_n}(x, y) = \sum_{(k, i) \in B_{C_n}} f_{k, i}^{C_n} x^k, B_{C_n} = \{(k, i) \in \mathbb{N}^2 \times \{1, 2\} \mid |\langle \lambda, k \rangle - \lambda_j| < C_n(k_1 + k_2)\},$$

and C_n is defined as in (20). Moreover we already constructed the coordinate transformations $\text{id} + u_n$ that transform X_{C_n} into $X_{C_{n+1}}$ within the proof of Theorem 12. Let us first rewrite the vector fields X_{C_n} . Let $B_{C_n, i} = \{k \mid (k, i) \in B_{C_n}\}$, for $i \in \{1, 2\}$ and consider $B_{C_n, 1}$. This set has a conical structure that is determined by two lines. Using the same argumentation as in Lemma 13, one can show that $B_{C_n, 1} \subset (1, 0) + \{a(p_n, q_n) + b(p_{n+1}, q_{n+1}) \mid a, b \in \mathbb{N}\}$, for $n \geq N_0$. Analogous, it follows that $B_{C_n, 2} \subset (0, 1) + \{a(p_n, q_n) + b(p_{n+1}, q_{n+1}) \mid a, b \in \mathbb{N}\}$. As a consequence we can write

$$X_{C_n} = \lambda_1 x (g_1^{C_n}(u, v)) \frac{\partial}{\partial x} + \lambda_2 y (g_2^{C_n}(u, v)) \frac{\partial}{\partial y},$$

where $u = x^{p_n} y^{q_n}$ and $v = x^{p_{n+1}} y^{q_{n+1}}$. Let us look at the vector field as a set of differential equations for a moment.

$$X_{C_n} : \begin{cases} \dot{x} &= \lambda_1 x (g_1^{C_n}(u, v)) \\ \dot{y} &= \lambda_2 y (g_2^{C_n}(u, v)). \end{cases}$$

It is interesting to look at $u = x^{q_n} y^{p_n}$, $v = x^{q_{n+1}} y^{p_{n+1}}$ as a ‘coordinate transformation’. Note that (at points where it is defined) we have that

$$\frac{u^{q_{n+1}}}{v^{q_n}} = x^{p_n q_{n+1} - p_{n+1} q_n} = x^{\pm 1}, \frac{u^{p_{n+1}}}{v^{p_n}} = y^{p_{n+1} q_n - p_n q_{n+1}} = y^{\pm 1}.$$

The sign depends on the parity of n . Moreover we can compute \dot{u} , \dot{v} and find

$$Z_{C_n} : \begin{cases} \dot{u} &= (\lambda_1 p_n + \lambda_2 q_n) u \left(\frac{\lambda_1 p_n f^{C_n}(u, v) + \lambda_2 q_n g^{C_n}(u, v)}{\lambda_1 p_n + \lambda_2 q_n} \right) \\ \dot{v} &= (\lambda_1 p_{n+1} + \lambda_2 q_{n+1}) v \left(\frac{\lambda_1 p_{n+1} f^{C_n}(u, v) + \lambda_2 q_{n+1} g^{C_n}(u, v)}{\lambda_1 p_{n+1} + \lambda_2 q_{n+1}} \right) \end{cases} \quad (26)$$

Using the definition of continued fractions, one verifies that $\eta_n = (\lambda_1 p_n + \lambda_2 q_n)/(\lambda_1 p_{n+1} + \lambda_2 q_{n+1})$ is a positive real but non-rational number. Moreover, one checks that if the continued fraction of $-\lambda_1/\lambda_2$ is given as $[a_0, a_1, a_2, \dots]$, then $\eta_n = [a_n, a_{n+1}, a_{n+2}, \dots]$. In short we have the following diagram:

$$\begin{array}{ccccccc} X & \xrightarrow{Tr} & X_{C_1} & \xrightarrow{Tr} & X_{C_2} & \xrightarrow{Tr} & X_{C_2} \dots \\ & & \downarrow R & & \downarrow R & & \downarrow R \\ & & Z_{C_1} & \xrightarrow{Tr} & \bullet & \xrightarrow{Tr} & \dots \\ & & & & \downarrow R & & \downarrow R \\ & & & & Z_{C_2} & \xrightarrow{Tr} & \dots \\ & & & & & & \downarrow R \\ & & & & & & Z_{C_3} \dots \end{array} \quad (27)$$

In this scheme Tr stands for a reduction by applying Theorem 2 and R stands for a ‘coordinate transformation’ of the type $u = x^{p_n} y^{q_n}$, $v = x^{p_{n+1}} y^{q_{n+1}}$. One then verifies afterwards that every square in this diagram commutes. Our original idea was to try to lift the transformations from the lower lines of these squares to the top lines and keeping track of radius of convergence at each step and trying to show convergence of the linearization procedure (provided some diophantine condition is valid). It is possible to provide a direct proof of the convergence of the linearization procedure with the use of this diagram, but the computations are a bit more lengthy. We provided this interpretation because it enlightens the real renormalization nature: the used transformations deliver an equations that has exactly the same nature as the starting equation!

5 Appendix

In order to be self-contained, we collect some more or less standard results about continued fractions that we have used above, see also [11, 7, 14, 8]. Let $\eta \in \mathbb{R} \setminus \mathbb{Q}$. The continued fraction expansion of $\eta = -\frac{\lambda_1}{\lambda_2} = [a_0; a_1, a_2, \dots]$ is defined as follows:

$$-\frac{\lambda_1}{\lambda_2} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots}}$$

More formally we have $a_0 := \lfloor \eta \rfloor$, $b_0 := \eta - \lfloor \eta \rfloor$, $a_1 := \lfloor \frac{1}{b_0} \rfloor$, $b_1 := \frac{1}{b_0} - \lfloor \frac{1}{b_0} \rfloor$, ..., $a_n := \lfloor \frac{1}{b_{n-1}} \rfloor$, $b_n := \frac{1}{b_{n-1}} - \lfloor \frac{1}{b_{n-1}} \rfloor$. We define the ratio's $\frac{q_n}{p_n} =$

$[a_0; a_1, \dots, a_n, 0, 0, \dots] \in \mathbb{Q}$. For example we have that $\frac{q_0}{p_0} = \frac{a_0}{1}$, $\frac{q_1}{p_1} = \frac{a_0 a_1 + 1}{a_1}$, $\frac{q_2}{p_2} = \frac{a_0 a_1 a_2 + a_0 + a_2}{a_1 a_2 + 1}$ and so on. We have recursively

$$(q_{n+1}, p_{n+1}) = (a_{n+1}q_n + q_{n-1}, a_{n+1}p_n + p_{n-1}). \quad (28)$$

For any $n \in \mathbb{N}$ and any $x \in \mathbb{R}$ we define

$$[a_0, a_1, \dots, a_{n-1}, x] = -\frac{\lambda_1}{\lambda_2} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \dots \frac{1}{a_{n-1} + \frac{1}{x}}}}$$

One says that the eigenvalues (λ_1, λ_2) satisfy a τ -diophantine condition if the inequality

$$|\lambda_1 k_1 + \lambda_2 k_2| \geq \frac{C_0}{(k_1 + k_2)^\tau}$$

holds for all non-zero $(k_1, k_2) \in \mathbb{N}^2$.

We state some useful lemmas concerning continued fractions. These can for instance be found in [11, 7, 14, 8] with proofs.

Lemma 15. *Let $x > 0$, $n \in \mathbb{N} \setminus \{0\}$, then*

$$[a_0, a_1, \dots, a_{n-1}, x] = \frac{xq_{n-1} + q_{n-2}}{xp_{n-1} + p_{n-2}}. \quad (29)$$

Lemma 16. *The set of ratio's $-\frac{\lambda_1}{\lambda_2}$ that do not satisfy the Bruno condition has zero Lebesgue measure in $\{(\lambda_1, \lambda_2) \in \mathbb{R}^2 \mid \lambda_1 > 0, \lambda_2 < 0\}$.*

Lemma 17. *Let $\left(\frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ be the continued fraction expansion of $-\frac{\lambda_1}{\lambda_2}$ then the sequences $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ increase at least as fast as the Fibonacci sequence.*

Lemma 18. *The continued fraction expansion $\left(\frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ of $-\frac{\lambda_1}{\lambda_2}$ satisfies the recurrence relations*

$$q_n p_{n-1} - p_n q_{n-1} = (-1)^n. \quad (30)$$

Remark 19. *Geometrically this means that the parallelogram spanned by the vectors (q_n, p_n) and (q_{n-1}, p_{n-1}) has surface area 1. Hence there are, apart from the corner points, no other integer couples in the closure of the parallelogram spanned by these two vectors.*

Lemma 20. *The following sequence of inequalities is true:*

$$\frac{q_0}{p_0} \leq \dots \leq \frac{q_{2n-2}}{p_{2n-2}} \leq \frac{q_{2n}}{p_{2n}} \leq \dots \leq -\frac{\lambda_1}{\lambda_2} \leq \dots \leq \frac{q_{2n+1}}{p_{2n+1}} \leq \frac{q_{2n-1}}{p_{2n-1}} \leq \dots \leq \frac{q_1}{p_1}.$$

Geometrically, this means that the slope of the resonant line is approached alternately from above by the continued fractions with an odd index, and from below by the continued fractions with an even index.

Lemma 21. Let $\left(\frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ be the continued fraction expansion of $\eta \in \mathbb{R}^+ \setminus \mathbb{Q}$. Then

$$\frac{1}{p_n + p_{n+1}} < |p_n \eta - q_n| < \frac{1}{p_{n+1}}. \quad (31)$$

Lemma 22. $\left(-\frac{\lambda_1}{\lambda_2} - \frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ is a strictly decreasing sequence.

Let $\eta \in \mathbb{R} \setminus \mathbb{Q}$, $\eta \geq 0$. We say that $(a, b) \in \mathbb{N}^2$, $b \neq 0$ is a best rational approximation to η if $|b\eta - a| < |p\eta - q|$ for all $q, p \in \mathbb{N}$ for which $0 < p \leq b$.

Lemma 23. Let $\eta \in \mathbb{R} \setminus \mathbb{Q}$, $\eta \geq 0$. Then every best rational approximation (a, b) has a ratio $\frac{a}{b}$ that is a continued fraction approximant of η (i.e. there exists a $k \in \mathbb{N}$ such that $\frac{q_k}{p_k} = \frac{a}{b}$).

We continue with some results on best rational approximation of η of the second kind. Because we only found a proof written in german by Perron [11], we thought it could be useful to add a proof.

If $\frac{A}{B}$ has the property that every rational number that lies between η and $\frac{A}{B}$ has a larger denominator, then we say that $\frac{A}{B}$ is a best rational approximation of η of the second kind.

Suppose now that $\eta = [a_0, a_1, \dots]$. We define $\xi_\nu = [a_\nu, a_{\nu+1}, \dots]$. One can verify that

$$\eta = \frac{q_{\nu-1}\xi_\nu + q_{\nu-2}}{p_{\nu-1}\xi_\nu + p_{\nu-2}},$$

$$\xi_\nu = \frac{p_{\nu-1}\eta - q_{\nu-2}}{p_{\nu-1}\eta - q_{\nu-2}}.$$

Let $p_{-1} = 0$, $q_{-1} = 1$. We define the following two approximation series of η , here is $\left(\frac{q_n}{p_n}\right)_{n \in \mathbb{N}}$ the continued fraction expansion of η :

$$\frac{q_0}{p_0}, \frac{q_1 + q_0}{p_1 + p_0}, \frac{2q_1 + q_0}{2p_1 + p_0}, \dots, \frac{a_2 q_1 + q_0}{a_2 p_1 + p_0} = \frac{q_2}{p_2}, \frac{q_3 + q_2}{p_3 + p_2}, \dots, \frac{a_4 q_3 + q_2}{a_4 p_3 + p_2}, \dots \quad (32)$$

$$\frac{q_0 + q_{-1}}{p_0 + p_{-1}}, \frac{2q_0 + q_{-1}}{2p_0 + p_{-1}}, \dots, \frac{a_1 q_0 + q_{-1}}{a_1 p_0 + p_{-1}} = \frac{q_1}{p_1}, \frac{q_2 + q_1}{p_2 + p_1}, \dots, \frac{a_3 q_2 + q_1}{a_3 p_2 + p_1}, \dots \quad (33)$$

Each member of (32) is smaller than η and each member of (33) is larger than η . Both sequences (32) and (33) are sorted by growing denominator. We show a few approximation lemmas.

Lemma 24. Let $\nu \in \mathbb{N}$, $\nu \geq 2$. Every positive rational number $\frac{A}{B}$ that is different from each of the members of (32) and (33) that is as close or closer to η as

$$\frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}},$$

for a certain $0 \leq c \leq a_\nu$, has a larger denominator (i.e. $B > cp_{\nu-1} + p_{\nu-2}$).

Proof. Let $\frac{A}{B}$ be such that

$$\left| \eta - \frac{A}{B} \right| \leq \left| \eta - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right|,$$

and suppose that $\frac{A}{B}$ is not one of the members of (32) or (33). Then

$$\left| \frac{A}{B} - \frac{q_{\nu-1}}{p_{\nu-1}} \right| = \left| \left(\eta - \frac{q_{\nu-1}}{p_{\nu-1}} \right) - \left(\eta - \frac{A}{B} \right) \right| \quad (34)$$

$$\stackrel{(a)}{\leq} \left| \left(\eta - \frac{q_{\nu-1}}{p_{\nu-1}} \right) \right| + \left| \left(\eta - \frac{A}{B} \right) \right| \quad (35)$$

$$\stackrel{(b)}{\leq} \left| \left(\eta - \frac{q_{\nu-1}}{p_{\nu-1}} \right) \right| + \left| \left(\eta - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right) \right| \quad (36)$$

$$\stackrel{(c)}{\leq} \left| \left(\eta - \frac{q_{\nu-1}}{p_{\nu-1}} \right) - \left(\eta - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right) \right|. \quad (37)$$

Remark that both quantities between the absolute values in (36) have an opposite sign. Furthermore, it is impossible that we have an equality in both (a) and (b), because this would mean that $\frac{A}{B}$ lies between η and $\frac{q_{\nu-1}}{p_{\nu-1}}$ and between η and $\frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}}$, which is contradictory since $\eta - \frac{q_{\nu-1}}{p_{\nu-1}}$ and $\eta - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}}$ have a different sign. Consequently

$$\left| \frac{A}{B} - \frac{q_{\nu-1}}{p_{\nu-1}} \right| < \left| \frac{q_{\nu-1}}{p_{\nu-1}} - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right| = \frac{1}{p_{\nu-1}(cp_{\nu-1} + p_{\nu-2})},$$

and

$$|Ap_{\nu-1} - q_{\nu-1}B| < \frac{B}{cp_{\nu-1} + p_{\nu-2}}.$$

Since $\frac{A}{B} \neq \frac{q_{\nu-1}}{p_{\nu-1}}$ it follows that $|Ap_{\nu-1} - q_{\nu-1}B| \geq 1$, since it is a natural number. As a consequence $B > cp_{\nu-1} + p_{\nu-2}$, which finishes the proof. \square

Lemma 25. *If $\frac{A}{B}$ is a best rational approximation of η of the second kind, then $\frac{A}{B}$ is one of the members of (32) or (33).*

Proof. Suppose that $\frac{A}{B}$ is a rational that is different from all members of (32) and (33). We consider two cases.

Case 1: $\frac{A}{B} < \eta$.

In this case $\frac{A}{B}$ lies either between two subsequent members of 32 or $\frac{A}{B} < \frac{q_0}{p_0}$. If $\frac{A}{B} < \frac{q_0}{p_0} = \frac{a_0}{1} < \eta$, we have a contradiction, since the ratio $\frac{a_0}{1}$ has a denominator that is not strictly smaller than the denominator of $\frac{A}{B}$. If $\frac{(c-1)q_{\nu-1} + q_{\nu-2}}{(c-1)p_{\nu-1} + p_{\nu-2}} < \frac{A}{B} < \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} < \eta$, then also

$$\begin{aligned} 0 &< \left| \frac{A}{B} - \frac{(c-1)q_{\nu-1} + q_{\nu-2}}{(c-1)p_{\nu-1} + p_{\nu-2}} \right| < \left| \frac{(c-1)q_{\nu-1} + q_{\nu-2}}{(c-1)p_{\nu-1} + p_{\nu-2}} - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right| \\ &= \frac{1}{((c-1)p_{\nu-1} + p_{\nu-2})(cp_{\nu-1} + p_{\nu-2})}. \end{aligned}$$

This implies

$$0 < |A((c-1)p_{\nu-1} + p_{\nu-2}) - B(cp_{\nu-1} + p_{\nu-2})| < \frac{B}{cp_{\nu-1} + p_{\nu-2}}.$$

It follows that $B > cp_{\nu-1} + p_{\nu-2}$, because $|A((c-1)p_{\nu-1} + p_{\nu-2}) - B(cp_{\nu-1} + p_{\nu-2})|$ is a nonzero natural number.

Case 2: $\frac{A}{B} > \eta$.

The proof is analogous and is left to the reader. \square

In fact we can be more precise on which members of (32) and (33) are best rational approximation of η of the second kind. Remark that the following lemma is one of the key ingredients to obtain the equivalence with the Bruno condition.

Lemma 26. *Consider the ratio*

$$\frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}}, \quad 0 < c \leq a_\nu.$$

We have that if $2c > a_\nu + 1$ then this ratio is a best approximation of the second kind; and if $2c \leq a_\nu - 1$ then this ratio is not a best approximation of the second kind.

It is clear that $\frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}}$ is a best rational approximation of the second kind if

$$\left| \eta - \frac{cq_{\nu-1} + q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right| < \left| \eta - \frac{q_{\nu-1}}{p_{\nu-1}} \right|,$$

which is equivalent with

$$\left| \frac{c(p_{\nu-1}\eta - q_{\nu-1}) + p_{\nu-2}\eta - q_{\nu-2}}{cp_{\nu-1} + p_{\nu-2}} \right| < \left| \frac{\eta p_{\nu-1} - q_{\nu-1}}{p_{\nu-1}} \right|.$$

We now use $\xi_\nu = \frac{p_{\nu-1}\eta - q_{\nu-1}}{p_{\nu-1}\eta - q_{\nu-2}}$, to obtain the equivalence with

$$\left| \frac{(c - \xi_\nu)(p_{\nu-1}\eta - q_{\nu-1})}{cp_{\nu-1} + p_{\nu-2}} \right| < \frac{\eta p_{\nu-1} - q_{\nu-1}}{p_{\nu-1}},$$

and simplifies to

$$p_{\nu-1} |c - \xi_\nu| < cp_{\nu-1} + p_{\nu-2}.$$

We have that $c < a_\nu \leq \xi_\nu$ hence this is equivalent to

$$p_{\nu-1}(\xi_\nu - c) < cp_{\nu-1} + p_{\nu-2}. \quad (38)$$

Suppose that $2c > a_\nu + 1$, then

$$2cp_{\nu-1} + p_{\nu-2} \geq (a_\nu + 1)p_{\nu-1} + p_{\nu-1} = a_\nu p_{\nu-1} \leq \xi_\nu B_{\nu-1},$$

and (38) is fulfilled. Suppose now that $2c \leq a_\nu - 1$, then

$$2cp_{\nu-1} + p_{\nu-2} \leq (a_\nu - 1)p_{\nu-1} + p_{\nu-1} \leq b_\nu p_{\nu-1} \leq \xi_\nu p_{\nu-1},$$

and (38) is false.

References

- [1] P. Bonckaert and P. De Maesschalck. Gevrey and analytic local models for families of vector fields. *Discrete Contin. Dyn. Syst. Ser. B*, 10(2-3):377–400, 2008.
- [2] P. Bonckaert and F. Verstringe. On the flat remainder in normal forms of families of analytic planar saddles. *C. R. Math. Acad. Sci. Paris*, 346(9-10):553–558, 2008.
- [3] A. D. Bruno. Analytic form of differential equations. I, II. *Trudy Moskov. Mat. Obšč.*, 25:119–262; *ibid.* 26 (1972), 199–239, 1971.
- [4] K. Khanin, J. Lopes Dias, and J. Marklof. Multidimensional continued fractions, dynamical renormalization and KAM theory. *Comm. Math. Phys.*, 270(1):197–231, 2007.
- [5] H. Koch and J. Lopes Dias. Renormalization of Diophantine skew flows, with applications to the reducibility problem. *Discrete Contin. Dyn. Syst.*, 21(2):477–500, 2008.
- [6] J. Lopes Dias. Renormalization scheme for vector fields on \mathbb{T}^2 with a Diophantine frequency. *Nonlinearity*, 15(3):665–679, 2002.
- [7] S. Marmi, P. Moussa, and J.-C. Yoccoz. Complex Brjuno functions. *J. Amer. Math. Soc.*, 14(4):783–841, 2001.
- [8] I. Niven. *Irrational numbers*. The Carus Mathematical Monographs, No. 11. The Mathematical Association of America. Distributed by John Wiley and Sons, Inc., New York, N.Y., 1956.
- [9] R. Pérez-Marco. Linearization of holomorphic germs with resonant linear part. 2000.
- [10] R. Pérez Marco and J.-C. Yoccoz. Germes de feuilletages holomorphes à holonomie prescrite. *Astérisque*, (222):7, 345–371, 1994. Complex analytic methods in dynamical systems (Rio de Janeiro, 1992).
- [11] O. Perron. *Die Lehre von den Kettenbrüchen*. Chelsea Publishing Co., New York, N. Y., 1950.
- [12] H. Rüssmann. Stability of elliptic fixed points of analytic area-preserving mappings under the Bruno condition. *Ergodic Theory Dynam. Systems*, 22(5):1551–1573, 2002.
- [13] L. Stolovitch. Sur un théorème de Dulac. *Ann. Inst. Fourier (Grenoble)*, 44(5):1397–1433, 1994.
- [14] J.-C. Yoccoz. Théorème de Siegel, nombres de Bruno et polynômes quadratiques. *Astérisque*, (231):3–88, 1995.