1 Banach spaces and the contractive mapping principle

In rigorously proving local or asymptotic results about *solutions* of various problems, where a closed form solution does not exist or is awkward, the contractive mapping principle is a handy tool. Some general guidelines on how to construct this operator are discussed in §1.3. It is desirable to go through the rigorous proof, whenever possible — this should be straightforward when the asymptotic solution has been correctly found—, one reason being that this quickly signals errors.

In §1.0.1 we discuss, for completeness, a few basic facts about Banach spaces. There is of course a vast literature on the subject; see e.g. [80].

1.0.1 A brief review of Banach spaces

Familiar examples of Banach spaces are the n-dimensional Euclidian vector spaces $\mathbb{R}^n$. A norm exists in a Banach space, which has the essential properties of a length: scaling, positivity except for the zero vector which has length zero and the triangle inequality (the sum of the lengths of the sides of a triangle is no less than the length of the third one). Once we have a norm, we can define limits, by reducing the notion to that in $\mathbb{R}$: $x_n \to x$ iff $\|x - x_n\| \to 0$. A normed vector space $\mathcal{B}$ is a Banach space if it is complete, that is every sequence with the property $\|x_n - x_m\| \to 0$ uniformly in n, m (a Cauchy sequence) has a limit in $\mathcal{B}$. Note that $\mathbb{R}^n$ can be thought of as the space of functions defined on the set of integers $\{1, 2, ..., n\}$. If we take a space of functions on a domain containing infinitely many points, then the Banach space is usually infinite-dimensional. An example is $L^\infty[0,1]$, the space of bounded functions on $[0,1]$ with the norm $\|f\| = \sup_{[0,1]} |f|$. A function L between two Banach spaces which is linear, $L(x+y) = Lx + Ly$, is bounded (or continuous) if $\|L\| := \sup_{\|x\|=1} \|Lx\| < \infty$. Assume $\mathcal{B}$ is a Banach space and that S is a closed subset of $\mathcal{B}$. In the *induced topology* (i.e., in the same norm), S is a complete normed space.

1.0.2 Fixed point theorem

Assume $\mathcal{M} : S \mapsto \mathcal{B}$ is a (linear or nonlinear) operator with the property that for any $x, y \in S$ we have

$$\|\mathcal{M}(y) - \mathcal{M}(x)\| \leqslant \lambda \|y - x\| \tag{1}$$

with $\lambda < 1$. Such operators are called **contractive**. Note that if $\mathcal{M}$ is linear, this just means that the norm of $\mathcal{M}$ is less than one.

Theorem 1. *Assume $\mathcal{M} : S \mapsto S$, where S is a closed subset of $\mathcal{B}$ is a contractive mapping. Then the equation*

$$x = \mathcal{M}(x) \tag{2}$$

has a unique solution in S.

Proof. Consider the sequence $\{x_j\}_j \in \mathbb{N}$ defined recursively by

$$\begin{aligned} x_0 &= x_0 \in S \\ x_1 &= \mathcal{M}(x_0) \\ &\cdots \\ x_{j+1} &= \mathcal{M}(x_j) \\ &\cdots \end{aligned} \tag{3}$$

Exercise 1. *Show that if L is a linear operator from the Banach space $\mathcal{B}$ into itself and $\|L\| < 1$ then $I - L$ is invertible, that is $x - Lx = y$ has always a unique solution $x \in \mathcal{B}$. "Conversely," assuming that $I - L$ is not invertible, then in whatever norm $\|\cdot\|_*$ we choose to make the same $\mathcal{B}$ a Banach space, we must have $\|L\|_* \geqslant 1$ (why?).*

1.1 The abstract implicit function theorem

The *Fréchet derivative* of a possibly nonlinear operator A in a Banach space, if it exists, is defined to be the linear operator D_xA with the property

$$\|A(x+h) - A(x) - (D_xA)h\| = o(\|h\|) \text{ as } \|h\| \to 0 \tag{8}$$

Exercise 2 (Implicit function theorem in Banach spaces)**.** Prove the following:

"Let X, Y, Z be Banach spaces. Assume that the mapping $f : X \times Y \to Z$ be is continuously Fréchet differentiable. If $(a, b) \in X \times Y$ is s.t. $f(a, b) = 0$ and $Df(a, b)(0, y)$ is a Banach space isomorphism from Y onto Z, then there exist neighborhoods A of a, B of b and a Fréchet differentiable function $g : A \to B$ s.t. $f(x, g(x)) = 0$ and $f(x, y) = 0$ if and only if $y = g(x)$, for all $(x, y) \in A \times B$."

1.2 Fixed points and vector valued analytic functions

A theory of analytic functions with values in a Banach space can be constructed by almost exactly following the usual construction of analytic functions. For the construction to work, we need the usual vector space operations and a topology in which these operations are continuous. A typical setting is that of a Banach algebra, see §4.1. A detailed presentation is found in [62] and [72], but the basic facts are simple enough for the reader to redo the necessary proofs.

1.3 Choice of the contractive map

An equation can be rewritten in a number of equivalent ways. In solving an asymptotic problem, as a general guideline we mention:

- The operator $\mathcal{N}$ appearing in the final form of the equation, which we want to be contractive, should not contain derivatives of highest order, divided differences with small denominators, or other operations poorly behaved with respect to asymptotics, and it should only depend on the sought-for solution in a formally small way. The latter condition should be, in a first stage, checked for consistency: the discarded terms, calculated using the first order approximation, should indeed turn out to be small.

- To obtain an equation where the discarded part is manifestly small it often helps to write the sought-for solution as the sum of the first few terms of the approximation, plus an exact remainder, say δ. The equation for δ is usually more contractive. It also becomes, up to smaller corrections, linear.

- The norms should reflect as well as possible the expected growth/decay tendency of the solution itself and the spaces chosen should be spaces where this solution lives.

- All freedom in the solution has been accounted for, that is, we should make sure the final equation cannot have more than one solution.

Note 2. At the stage where the problem has been brought to a contractive mapping setting, it usually can be recast without conceptual problems, but perhaps complicating the algebra, to a form where the implicit function theorem applies (and vice versa). The contraction mapping principle is often more natural, especially when the topology, suggested by the problem itself, is not one of the common ones. But an implicit function reformulation might bring in more global information.

2 Examples

2.1 Linear differential equations in Banach spaces

Consider the equation

$$Y'(t) = L(t)Y(t); \quad Y(0) = Y_0 \tag{9}$$

in a Banach space X, where $L(t) : X \to X$ is linear, norm continuous in t and uniformly bounded,

$$\sup_{t\in[0,\infty)} \|L(t)\| < L \tag{10}$$

Then the problem (9) has a global solution on $[0,\infty)$, and $\|Y(t)\| \leqslant \|Y_0\|e^{(L+\varepsilon)t}$.

Proof. By comparison with the case when $X = \mathbb{R}$, the natural growth is indeed Ce^{Lt}, so we rewrite (9) as an integral equation, in a space where the norm reflects this possible growth. Consider the space of continuous functions $Y : [0,\infty) \mapsto X$ in the norm

$$\|Y\|_{\infty,L} = \sup_{t\in[0,\infty)} e^{-Lt/\lambda}\|Y(t)\| \tag{11}$$

with $\lambda < 1$ and the auxiliary equation

$$Y(t) = Y_0 + \int_0^t L(s)Y(s)ds =: \mathcal{A}[Y](t) \tag{12}$$

which is well defined on X and is contractive there since

$$e^{-Lt/\lambda}\left|\int_0^t L(s)Y(s)ds\right| \leqslant Le^{-Lt/\lambda}\int_0^t e^{Ls/\lambda}\|Y\|_{\infty,L}ds$$
$$= \lambda(1-e^{-Lt/\lambda})\|Y\|_{\infty,L} \leqslant \lambda\|Y\|_{\infty,L}, \tag{13}$$

and therefore in a ball of radius $(1+\gamma)\|Y_0\|$, for large enough γ (in fact, we need $\gamma(1-\lambda) > \lambda$),

$$\|\mathcal{A}[Y]\|_{\infty,L} \leqslant \|Y_0\| + \lambda(1+\gamma)\|Y_0\| < (1+\gamma)\|Y_0\|$$

while

$$\|\mathcal{A}[Y_1] - \mathcal{A}[Y_2]\|_{\infty,L} \leqslant \lambda\|Y_1 - Y_1\|_{\infty,L},$$

implying $\mathcal{A}$ to be contraction map; and so a unique solution exists for the initial value problem (9) with given exponential bounds for growth as given. We note that in linear problems, we do not need to restrict the analysis to a ball. $\square$

3 Local existence and uniqueness of solutions of nonlinear systems

Consider the system of equations (or one vector equation if you prefer)

$$y' = F(x,y); \quad y(x_0) = y_0 \tag{14}$$

where $y \in \mathbb{R}^n$, $x \in \mathbb{R}$. The second condition, the initial value, makes (18) an *initial value problem, IVP*. You see that by taking $y = \tilde{y} + y_0$, $x = \tilde{x} + x_0$ and $\tilde{F}(\tilde{x},\tilde{y}) = F(\tilde{x}+x_0, \tilde{y}+y_0)$, we can assume, without loss of generality that our IVP is

$$y' = F(x,y); \quad y(0) = 0 \tag{15}$$

For mere existence of not necessarily unique solutions, mere continuity of F suffices; see [31] pp. 2–7. Clearly *some* condition is needed, for the simple equation $y' = f(x)$ has no solution if f is, say, not Lebesgue measurable (why?)

Existence and uniqueness requires stronger properties. Indeed, the equation

$$y' = 2y^{1/2}, \quad y(0) = 0$$

has as solutions $y = 0, y = x^2$ (and many more; can you find them?)

Let $\mathcal{D} = \mathbb{D}_\varepsilon = \{z : |z| < \varepsilon\}$. We will assume that $F : \mathbb{D}_\delta \times \mathbb{D}_\varepsilon^n \mapsto \mathbb{R}^n$ is L^1 in x and Lipschitz continuous in y (for some C and all $(x,y) \in \mathcal{D}$ we have $|F(x,y_1) - F(x,y_2)| < C|y_1 - y_2|$).

By taking a slightly smaller ε if needed, we can assume that F is continuous up to the boundary, that is continuous in $\overline{\mathbb{D}_\varepsilon} \times \overline{\mathbb{D}_\varepsilon^n}$.

We consider the space of continuous functions y on $\mathbb{D}_\varepsilon$ with the sup norm, $\|y\|_\infty = \sup_{x\in\overline{\mathbb{D}_\delta}} |y|$ form a Banach space; call this Banach space $\mathcal{B}$.

We now consider a closed subspace of $\mathcal{B}$, the closed ball $B = \{y \in \mathcal{B} : \|y\| \leqslant \delta\}$.

Exercise. Check that the IVP (19) is equivalent to

$$y = \int_0^x F(s, y(s))ds \tag{16}$$

Let ε be small enough. How small that is, we'll calculate in a moment. We now consider the *nonlinear* operator $\mathcal{M}$ be defined on B with values in B, given by

$$\mathcal{M}(y) = \int_0^x F(s, y(s))ds \tag{17}$$

For $|\mathcal{M}(y)|$ to be bounded by δ, we need that $\varepsilon \max_{\overline{\mathbb{D}_\varepsilon} \times \overline{\mathbb{D}_\varepsilon^n}} |F(s, y(s)| < \delta$. Check that this can be arranged by taking ε small enough.

For $\mathcal{M}(y)$ to be contractive, check that it suffices to have

$$C\varepsilon < \alpha < 1$$

This ensures contractivity and therefore existence and uniqueness of solutions of the IVP.

4 Local existence and uniqueness of analytic solutions: contractive mapping approach

The study of analytic systems mirrors the analysis of §3.

Consider the system of equations (or one vector equation if you prefer)

$$y' = F(x, y); \quad y(x_0) = y_0 \tag{18}$$

where $y \in \mathbb{C}^n$, $x \in \mathbb{C}$. The second condition, the initial value, makes (18) an *initial value problem, IVP*. You see that by taking $y = \tilde{y} + y_0$, $x = \tilde{x} + x_0$ and $\tilde{F}(\tilde{x}, \tilde{y}) = F(\tilde{x} + x_0, \tilde{y} + y_0)$, we can assume, without loss of generality that our IVP is

$$y' = F(x, y); \quad y(0) = 0 \tag{19}$$

We must specify the properties of F. Let where $\mathbb{D}_\varepsilon = \{z : |z| < \varepsilon\}$. We will assume that $F : \mathbb{D}_\delta \times \mathbb{D}_\varepsilon^n \mapsto \mathbb{C}^n$ is analytic in $\mathbb{D}_\delta \times \mathbb{D}_\varepsilon^n$ for some $\delta > 0, \varepsilon > 0$. This means that F has a convergent Taylor series in $(x, y_1, ..., y_n)$ in $\mathbb{D}_\delta \times \mathbb{D}_\varepsilon^n$.

It is known (by Hartog's theorem: google it!) that if F is separately analytic in each variable (thinking therefore of the others as being "frozen"), then it is analytic in the stronger sense above.

By taking a slightly smaller ε if needed, we can assume that F is continuous up to the boundary, that is continuous in $\overline{\mathbb{D}_\delta} \times \overline{\mathbb{D}_\varepsilon^n}$.

Check that the functions y which are analytic in $\mathbb{D}_\delta$ and continuous in $\overline{\mathbb{D}_\delta}$ endowed with the sup norm, $\|y\|_\infty = \sup_{x \in \overline{\mathbb{D}_\delta}} |y|$ form a Banach space; call this Banach space $\mathcal{B}$.

We now consider a closed subspace of $\mathcal{B}$, the closed ball $B = \{y \in \mathcal{B} : \|y\| \leqslant \varepsilon\}$.

Exercise. Check that the IVP (19) is equivalent to

$$y = \int_0^x F(s, y(s))ds \tag{20}$$

Let ε be small enough. How small that is, we'll calculate in a moment. We now consider the *nonlinear* operator $\mathcal{M}$ be defined on B with values in B, given by

$$\mathcal{M}(y) = \int_0^x F(s, y(s))ds \tag{21}$$

For $|\mathcal{M}(y)|$ to be bounded by ε, we need that $\delta \max_{\overline{\mathbb{D}_\delta} \times \overline{\mathbb{D}_\varepsilon^n}} |F(s, y(s)| < \varepsilon$. Check that this can be arranged by taking δ small enough.

For $\mathcal{M}(y)$ to be contractive, check that it suffices to have

$$\sup_{|x|<\delta; |y|<\varepsilon} \left\| \frac{\partial F_i}{\partial y_j} \right\|_{\mathbb{R}^n \to \mathbb{R}^n} \delta < \alpha < 1$$

where the norm of the Jacobian is the usual matrix norm. This ensures contractivity and therefore existence and uniqueness of solutions of the IVP.

Exercise 1. Formulate and prove an analytic analog of the result in §2.1 in the case A is analytic in a connected, open and bounded domain in $\mathbb{C}$.

4.1 The exponential and the log of a matrix

A natural setting in which functions of a matrices and more generally of (say bounded) operators are analyzed is that of a Banach algebra.

This is a Banach space endowed with multiplication which is distributive, associative and continuous in the Banach norm.

Continuity of the addition and multiplication are spelled out as

$$\|x + y\| \leqslant \|x\| + \|y\|; \quad \|xy\| \leqslant \|x\|\|y\|, \quad \forall x, y \tag{22}$$

Note that $n-$dimensional matrices form a Banach space w.r.t. the usual norm, $\|A\| = \max_{\|x\|=1} \|Ax\|$.

We can consider the sum

$$e^M = \sum_{k=0}^{\infty} M^k/k! \tag{23}$$

Since $\|M^k\| \leqslant \|M\|^k$ and the series

$$\sum_{k=0}^{\infty} \|M\|^k/k! \tag{24}$$

converges, it follows that e^M is correctly defined, by a norm-convergent series. You can check the usual properties of the exponential. Careful though: $AB \neq BA$ in general, so we can't expect $e^{A+B} = e^A e^B$.

For the log, if M is diagonalizable and 0 is not in the spectrum of M, then define $\log M$ to be $A[\log \Lambda]A^{-1}$. Here, A is the diagonalization matrix, Λ is diagonal, and so is, by definition $\log \Lambda$, consisting of the logs of the diagonal elements of Λ. Of course, this log may be complex if the eigenvalues are not positive, and it is not uniquely defined.

Exercise: If M has a nontrivial Jordan normal form with no zero eigenvalue, it is enough to define the log block by block. Each block is of the form $\lambda I - N$, $\lambda \in \mathbb{C}$, I the identity matrix and N a nilpotent. Then, the sum

$$I \log \lambda - \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{N}{\lambda} \right)^k = I \log \lambda - \sum_{k=1}^{m} \frac{1}{k} \left(\frac{N}{\lambda} \right)^k \tag{25}$$

since $N^{m+j} = 0 \ \forall j \in \mathbb{N}$ for some $m \leqslant \dim(M)$ since N is nilpotent.

One can define, in a similar way, more general functions of matrices.

5 The fundamental solution of a linear system

Consider a linear system of differential equations of the form

$$w' = A(z)w \tag{26}$$

where $w \in \mathbb{C}^n$ and A is analytic near z_0; we first look at (26) near z_0; without loss of generality, we can take $z_0 = 0$.

In a neighborhood of a regular point (i.e., a point where A analytic) there exist n linearly independent vector solutions, $\{w_j : j = 1, ..., n\}$ of (202) by the existence and uniqueness theorems we have proven. We will also see shortly that $w_i(t)$ are linearly independent at *all* t.

Furthermore, you can choose initial conditions so that $w_j(z = 0) = e_j$, the unit vector in the direction j. If you construct a matrix M having as the $j - th$ column the vector w_j, you can check immediately that

$$M' = A(z)M; \quad M(0) = I \tag{27}$$

Inverse matrix. Let's now see what equation M^{-1} woud satisfy, assuming momentarily it has one (we will show this is the case). Since $MN = I$ we have $M'N + MN' = 0 \Rightarrow MN' = -AMN = -A$ and thus

$$\boxed{N' = -NA}; \ N(0) = I \tag{28}$$

thus N gives the backward evolution $z \to -z$.

Note 3. You can check that the matrix equation (149) is equivalent to the set of equations $w_i' = Aw_i, \ w_i(0) = e_i$, while (28) corresponds to the system $u_i' = -A^t u_i, \ u_i(0) = e_i$ where A^t is the transpose of A.

Proposition 2. *Consider the IVPs* (149) *and* (28)*, with A Lipschitz continuous in the interval J containing the origin. We have $MN = NM = I$ in J. If A is analytic in a connected open set $\mathcal{D}$ (a region) then M, N are analytic in $\mathcal{D}$ and $MN = -NM = I$ in $\mathcal{D}$.*

Proof. We treat the analytic case, the continuous case being very similar. We have

$$(MN)' = M'N + MN' = AMN - MNA; \quad (MN)(0) = I \tag{29}$$

Note that this is a linear equation for the matrix MN ($X \mapsto AX - XA$ is a linear, bounded operator on matrices). Thus the IVP has a unique solution. But I is a solution, thus $MN = I$ for all t. $\square$

The fact that the determinant of M is nonzero at all t also follows from the following proposition, important in its own right.

Proposition 3. *If M is a solution of* (149) *then*

$$\det M(t) = \det M(0) \exp\left(\int_0^t \mathrm{Tr}A(s)\right) ds$$

Proof. Let $d(t) = \det(M(t))$. As long as M is invertible, we have

$$M(t+\varepsilon) = M(t)+(AM)(t)\varepsilon+o(\varepsilon) = (I+A\varepsilon)M(t) \Rightarrow d(t+\varepsilon) = \det(I+\varepsilon A)d(t) \tag{30}$$

and now the result follows from the following lemma, which is easily proved by induction on the size of the matrix. $\square$

Lemma 4. *For any matrix A,* $\det(I + \varepsilon A) = 1 + \varepsilon \mathrm{Tr}A + o(\varepsilon)$.

Now either Proposition 2 or 3 shows that the vectors $w(t)$ are linearly independent at any t.

Take any z-independent vector w_0, and let $w = Mw_0$. We have

$$(Mw_0)' = M'w_0 = AMw_0 \Rightarrow w' = Aw; \quad w(0) = Iw_0 = w_0 \tag{31}$$

Thus, we see that the solution of the initial value problem $w' = Aw, w(0) = w_0$ is simply Mw_0. We will often work with the fundamental matrix solution M, as it often simplifies the calculations.

Note 4. There is a generalization in some sense of this result in the case of PDE evolution equations. The Schrödinger equation $i\phi_t = H\psi = -\Delta\psi + V(x)\psi$; $\psi(0) = \psi_0$ where Δ is the Laplacian has the solution

$$\psi(t) = U(x,t)\psi(0) \tag{32}$$

where $U(x,t)$ is a unitary family of operators, and in fact $U = e^{-iHt}$, and functional analysis arguments give a precise meaning to e^{-iHt} (as a unitary operator) whenever H is a self-adjoint operator. This form of U does *not* hold if V depends on t as well.

Lemma 5. *The matrix differential equation*

$$W' = AW \tag{33}$$

has the general solution is $W = MC$ where M is the fundamental solution and C is any matrix of constants.

Proof. Indeed, since M is invertible, we can define $Q = M^{-1}W$, which we write in the form $W = MQ$. We then have

$$M'Q + MQ' = AMQ \Leftrightarrow MQ' = 0 \Leftrightarrow Q' = 0 \tag{34}$$

(since $M' = AM$) which indeed means that Q is a constant matrix. $\square$

5.1 Domain of existence of regular solutions

As discussed, the differential system (26) is equivalent to the matrix differential equation (33) which, by Exercise1, has an analytic solution throughout the domain of analyticity of A. Solutions of linear systems can only be singular only at singularities of the coefficients, in this case the singularities of A.

6 Isolated singularities of linear systems

Consider the system

$$w' = A(z)w \tag{35}$$

where A is a matrix valued analytic function, but now with *an isolated singularity at* z_0. Clearly, by translating z we can take $z_0 = 0$, and by rescaling z, we can assume that A is analytic in $\mathcal{D} = \mathbb{D} \setminus \{0\}$ where $\mathbb{D}$ is the open unit disk. Though the equation is single-valued in $\mathcal{D}$, since $\mathcal{D}$ is not simply connected, the solutions may not be, as seen by solving the equation $y' = ay/z$ with $a \notin \mathbb{Z}$. We can take $z = e^{\zeta}$ and $\mathcal{D}$ becomes $\{\zeta : \mathrm{Re}\zeta \in \mathbb{R}^-\}$, a half plane. By the standard existence and uniqueness theorems, we find that there is a unique solution of the system, rewritten in ζ, and thus there is a fundamental solution of (35), in the form $M(\ln z)$, which shows once more that, in principle at least, the solution of (35) may not be single-valued.

7 Some general facts about solutions near isolated singularities

In the generality of the singular systems in §6 all we can say now, without a lot more theory, is the way the solution itself can be ramified. Once more, we consider that we rescaled everything so that $z = 0$ is the isolated singularity, and $\mathcal{D} = \mathbb{D} \setminus \{0\}$ is the domain of analyticity of A.

Theorem 2. *The general solution of* (35) *is of the form*

$$M(z) = S(z)z^P \quad (z^P := e^{\ln z\, P}) \tag{36}$$

where P is a constant matrix, and $S(z)$ is analytic in $\mathcal{D}$. With the price of changing the matrix M to MT, with T a constant matrix, we can write

$$MT = S_1 x^J \tag{37}$$

where J is the Jordan normal form of P.

Note 5. This implies in particular (and is implied by, as we will see) the relation

$$M(ze^{2\pi i}) = M(z)C \tag{38}$$

for some invertible constant matrix C. By (38), a rotation by 2π generates another solution (since MC is indeed a solution). The map $M \mapsto MC$ is a group, and it is the *the monodromy group* at the singularity (0). One also says that (38) is the monodromy at zero.

What this theorem says is that the solution itself is single-valued up to multiplication by z^P with P constant. Of course, there is no reason to expect that S is analytic at zero-just that 0 is an isolated singularity. For the proof we need the following result.

Lemma 6. *Assume M is any matrix analytic on the universal covering of $\mathcal{D}$ (that is, $M(z) = F(\ln(z))$ whereF is analytic in the left half plane) which satisfies*

$$M(ze^{2\pi i}) = MC \quad \text{where } C \text{ is a constant invertible matrix.} \tag{39}$$

Then

$$M(z) = S(z)z^P \tag{40}$$

where P is a constant matrix and $S(z)$ is analytic in $\mathcal{D}$. At the price of altering M by a constant matrix, P can be taken to be in Jordan normal form.

Proof of the lemma. Since C is invertible, we can define P (up to $2\mathbb{Z}\pi iI$) by $C = e^{2\pi iP}$. Let

$$S = Mz^{-P} \tag{41}$$

$$S(ze^{2\pi i}) = Me^{2\pi iP}e^{-P\ln z - 2\pi iP} = Me^{-P\ln z} = S(z) \tag{42}$$

since e^{aP} and e^{bP} commute, if a and b are scalars. Let now T be the change of basis that brings P to its Jordan normal form, that is $T^{-1}PT = J$. We then have

$$MT = STT^{-1}z^PT = STz^J \tag{43}$$

where ST is also ingle valued, as required. ☐

Proof of the theorem. We only need to show that the assumptions of the lemma above hold. Take $N(z) = M(ze^{2\pi i})$. That is, we use the fact that M exists on the universal covering of $\mathcal{D}$, and look at its value on the second Riemann sheet. We have

$$N(z)' = e^{2\pi i}M'(ze^{2\pi i}) = A(ze^{2\pi i})M(ze^{2\pi i}) = A(z)M(ze^{2\pi i}) = A(z)N \tag{44}$$

where we used the fact that M is already a solution, and A is single-valued. Thus, by Remark 5, we must have $N = MC$ where C is a constant matrix. ☐

Remark 6. *If S happens to be analytic, note also the emerging noninteger powers of z and $\ln z^j$ through the term z^J.*

Indeed, if J_1 is an elementary Jordan block in J, we have

$$z^J = z^{\lambda I+N} = z^\lambda e^{N\ln z} = z^\lambda(1 + N\ln z + \cdots \ln z^l N^l/l!) \tag{45}$$

where $N^{l+1} = 0$, and thus $l < n$, the degree of the system.

8 Regular singular points of differential equations, nondegenerate case

8.1 Example

Consider the hypergeometric equation

$$x(x-1)y'' + y = 0 \tag{46}$$

around $x = 0$. The indicial equation is $r(r-1) = 0$ (a *resonant case*: the roots differ by an integer). Substituting $y_0 = \sum_{k=0}^{\infty} c_k x^k$ in the equation and identifying the powers of x yields the recurrence

$$c_{k+1} = \frac{k^2 - k + 1}{k(k+1)} c_k \tag{47}$$

with $c_0 = 0$ and c_1 arbitrary. By linearity we may take $c_1 = 1$ and by induction we see that $0 < c_k < 1$. Thus the power series has radius of convergence at least 1. The radius of convergence is in fact exactly one as it can be seen applying the ratio test and using (47); the series converges exactly up to the nearest singularity of (46).

Exercise 1. *What is the asymptotic behavior of c_k as $k \to \infty$?*

We let $y_0 = y_0 \int g(s)ds$ and get, after some calculations, the equation

$$g' + 2\frac{y_0'}{y_0} g = 0 \tag{48}$$

and, by the previous discussion, $2y_0'/y_0 = 2/x + A(x)$ with $A(x)$ is analytic. The point $x = 0$ is a regular singular point of (48) and in fact we can check that $g(x) = C_1 x^{-2} B(x)$ with C_1 an arbitrary constant and $B(x)$ analytic at $x = 0$. Thus $\int g(s)ds = C_1(a/x + b\ln(x) + A_1(x)) + C_2$ where $A_1(x)$ is analytic at $x = 0$. Undoing the substitutions we see that we have a fundamental set of solutions in the form $\{y_0(x), B_1(x) + xB_2(x)\ln x\}$ where B_1 and B_2 are analytic.

8.2 Singularities of the first kind versus regular singularities

For the system of equations (35) a singularity is of the first kind if A is analytic in a punctured disk, say $\mathcal{D}$ and it has a first order pole at zero. Regular versus irregular singularities are classified according to the type of *solutions* the system admits. If in (40) $S(z)$ is meromorphic, then the singularity is *regular*, whereas if 0 is an essential singularity of S, the singularity is *irregular*. Sometimes even in the irregular singular case we may get formal power series for S, but they usually have zero radius of convergence.

We now show that at singularities of the first kind, S is meromorphic. The converse is not true, but something along these lines holds if the system is written as an nth order equation, as we shall see.

We then let

$$A(z) = z^{-1}A_0(z)$$

where A_0 is analytic in the full disk $\mathbb{D}$, and for the problem to be interesting, $A_0(0) \neq 0$. Since A is singular only at 0, we can find a fundamental matrix M (invertible, of course) s.t. any nontrivial solution of (35) can be written as Mc for some constant nonzero vector c. Now $Mc \neq 0$ for any $z \neq 0$ and is an analytic vector. We choose a direction $\gamma = e^{i\phi}$ and evolve towards the origin, taking $z = t\gamma$; let $f_i(z) = (Mc)_i(z)$. We have

$$|f_i(t\gamma + \varepsilon\gamma)|^2 = f_i(t\gamma + \varepsilon\gamma)\overline{f_i(t\gamma + \varepsilon\gamma)}$$
$$= f_i(t\gamma)\overline{f_i(t\gamma)} + 2\mathrm{Re}(f_i'\overline{f_i\gamma})\varepsilon + O(\varepsilon^2) \Rightarrow \frac{d}{dt}|f_i(t\gamma)|^2 = 2\mathrm{Re}(f_i'\overline{f_i\gamma}) \quad (49)$$

We have, by Cauchy-Schwarz

$$\left|\frac{d}{dt}\|Mc\|\right| = \left|\frac{d}{dt}\sqrt{\sum_{i=1}^{n} f_i\overline{f_i}}\right| = \frac{\sum_{i=1}^{n}\mathrm{Re}(f_i'\overline{f_i\gamma})}{\|Mc\|} \leqslant \frac{\sum_{i=1}^{n}|f_i||f_i'|}{\|Mc\|} \leqslant \frac{\|Mc\|\|M'c\|}{\|Mc\|}$$
$$= \|M'c\| = \|AMc\| \leqslant t^{-1}\max_{z\in\mathbb{D}}\|A_0(z)\|\|Mc\| =: \frac{\kappa}{t}\|Mc\| \quad (50)$$

In particular

$$-\frac{\kappa}{t} \leqslant \frac{\|Mc\|'}{\|Mc\|} \leqslant \frac{\kappa}{t} \Rightarrow \|Mc\| \leqslant Ct^{-\kappa} \quad (51)$$

Recalling that $M = S(z)z^B$ for some matrix B, we see that $\|S\| \leqslant Cz^{\|B\|+kappa}$. Thus, if $m = \lfloor|\kappa| + \|B\|\rfloor + 1$, then z^mS is analytic in $\mathbb{D}\setminus\{0\}$ and it is bounded at zero; as we know this means that z^mS has a removable singularity at zero thus S has (at most) a pole at 0.

This also means that we have

$$M = \tilde{S}z^{\tilde{B}} \quad (52)$$

where S is analytic at zero and $\tilde{B}$ is a constant matrix.

We also see that, if A has a higher order pole, say double pole, then the best estimate that we can get by the method above is $\|Mc\| \leqslant \exp(\kappa t^{-1})$. While this does not show that this *must* be the growth, it is not hard to construct simple examples in which the growth rate is indeed exponential (e.g., the scalar equation $f' = z^{-2}f$).

8.2.1 Singularities of first kind and power series solutions

Consider the following differential equations:

$$f' + (1+x)f = 1 \quad (53)$$

$$x(1-x)f'(x) + (1+x)^2f(x) = 1 \quad (54)$$

$$x^2f' - f = x \quad (55)$$

Note 7. (i) Of course, these first order equations that we use for illustration in this section can be solved in closed form, but this is not the point.

(ii) These equations can be made homogeneous (second order) by differentiating them.

The first equation has no singularities and the solution is thus entire. Let's see how this is reflected at the level of the recurrence relation for the coefficients. Plugging in $f = \sum_{k=0}^{\infty} c_k x^k$ we get

$$c_{k+1} = -\frac{1}{m+1}\left(c_k + c_{k-1}\right) \tag{56}$$

It is not hard to show inductively that $c_k \leqslant AB^k (k/2)!$ for suitable A and B (check). For (54) we get

$$c_k = -\frac{k-3}{k+1} c_{k-1} + \frac{1}{k+1} c_{k-2} \tag{57}$$

which, as in the hypergeometric example can be shown to lead to a series with radius of convergence 1.

In the third example however, we get

$$c_k = (k-1)c_{k-1} \Rightarrow c_k = A(k-1)! \tag{58}$$

At the level of the series, a negative power of x shifts the coefficient c_k to c_{k+1} while differentiation essentially results in multiplying c_k by k. Depending on the strength of the singularity relative to the order of the equation, this results in a balance of the type $c_{k+1} \sim c_k/k$, $c_{k+1} \sim c_k$ or in case the pole is of higher order than the order of the equation, $c_{k+1} \sim kc_k$. The behavior of solutions depends critically on this balance.

8.3 Detailed analysis of singularities of the first kind

Consider the system

$$w' = \frac{1}{z}Bw + A_1(z)w; \quad \text{or, in matrix form,} \quad M' = \frac{1}{z}BM + A_1(z)M \tag{59}$$

where B is a constant matrix and A_1 is analytic at zero. Let J be the Jordan normal form of B and $T^{-1}BT = J$. Then, we see that

$$T^{-1}MT = \frac{1}{z}T^{-1}BTT^{-1}MT + T^{-1}A_1(z)TT^{-1}MT \tag{60}$$

with $\tilde{M} = T^{-1}MT$ we see that

$$\tilde{M}' = \frac{1}{z}J\tilde{M} + A_2(z)\tilde{M} \tag{61}$$

where clearly A_2 is also analytic. In other words,

Remark 8. *In* (130) *we can assume, without loss of generality that B is in its Jordan normal form, J. We will thus study equations of the form*

$$y' = \frac{1}{z}Jy + A(z)y \tag{62}$$

where $A(z)$ is analytic.

There are three cases leading to somewhat different analytic properties of M.

1. All eigenvalues are distinct and do not differ by integers. In this case, the elements of the fundamental matrix have the analytic structure $\sum z^{\lambda_k} A_k(z)$ where λ_k are the eigenvalues of J and A_k are analytic.

2. Some eigenvalues can be repeated, but no two eigenvalues differ by *positive* integers. Then, the elements of the fundamental matrix are of the form $\sum_{k,l\leqslant n} z^{\lambda_k} \ln z^l A_{kl}(z)$.

3. Some eigenvalues differ by *positive* integers. Then the powers of z may differ from the eigenvalues.

8.4 Nondegenerate case

Assumption. No two eigenvalues of B differ by a positive integer.

Theorem 3. *Under the assumption above, (62) has a fundamental matrix solution in the form $M(z) = Y(z)z^J$, where $Y(z)$ is a matrix analytic in $\mathcal{D}$.*

Exercise 2. *Check that, if we had not arranged for B to be in its Jordan normal form, the solution of (130) would be $M(z) = Z(z)z^B$, where $Z(z)$ is a matrix analytic at zero.*

Proof. Clearly, it is enough to prove the theorem for (62). We look for a solution of (62) in the form $M = Yz^J$, where

$$Y(z) = I + zY_1 + z^2Y_2 + \cdots \tag{63}$$

we get

$$Y'z^J + \frac{1}{z}YJz^J = \frac{1}{z}JYz^J + AYz^J \tag{64}$$

Multiplying by z^{-J} we obtain

$$Y' + \frac{1}{z}YJ = \frac{1}{z}JY + AY \tag{65}$$

or

$$Y' = \frac{1}{z}\Big(JY - YJ\Big) + AY \tag{66}$$

Using (331) we get

$$Y_1 + 2zY_2 + 3z^2Y_3 + \cdots = \Big[(JY_1 - Y_1J) + z(JY_2 - Y_2J) + \cdots\Big]$$
$$+ A_0J + zA_1J + \cdots + zA_0Y_1 + z^2(A_0Y_2 + A_1Y_1) + \cdots \tag{67}$$

The associated system of equations, after collecting the powers of z is

$$kY_k = (JY_k - Y_kJ) + A_{k-1}J + \sum_{j=1}^{k-1} Y_jA_{k-j-1}; \quad k \in \mathbb{N} \tag{68}$$

or

$$V_kY_k = A_{k-1}J + \sum_{j=1}^{k-1} Y_jA_{k-j-1}; \quad k \in \mathbb{N} \tag{69}$$

where

$$V_kM := kM - (JM - MJ) \tag{70}$$

is a linear operator on matrices $M \in \mathbb{R}^{n^2}$. As a linear operator on a finite dimensional space, $V_kX = Y$ has a unique solution for every Y iff $\det V_k \neq 0$ or, which is the same, $jX - JX + XJ = 0$ implies $X = 0$. We show that this is the case, by showing that $Xv = 0$ for all generalized eigenvectors of J.

Let v be one of the eigenvectors of J. If $V_kX = 0$ we obtain, since $Jv = \lambda v$,

$$k(Xv) - J(Xv) + X\lambda v = 0 \tag{71}$$

or

$$J(Xv) = (\lambda + k)(Xv) \tag{72}$$

Here we use our assumption: $\lambda + k$ is not an eigenvalue of J. This forces

$$Xv = 0 \tag{73}$$

We let $v_0 = v$ and take the next generalized eigenvector, v_1, in the same Jordan block as v, if any.

We remind that we have the following relations between these generalized eigenvectors:

$$Jv_i = \lambda v_i + v_{i-1} \tag{74}$$

where $v_0 = v$ is an eigenvector and $1 \leqslant i \leqslant m - 1$ where m is the dimension of the Jordan block. With $i = 1$ we get

$$k(Xv_1) - J(Xv_1) + X(\lambda v_1 + v_0) = 0 \tag{75}$$

and, using (73) (i.e., $Xv_0 = 0$), we get the same equation (76), now for Xv_1:

$$J(Xv_1) = (\lambda + k)(Xv_1) \tag{76}$$

and thus $Xv_1 = 0$. Inductively, we see that $Xv = 0$ for any generalized eigenvector of J, and thus $X = 0$.

Now, we claim that $V_k^{-1} \leqslant Ck^{-1}$ for some C. We let $\mathcal{C}$ be the commutator operator, $\mathcal{C}X = JX - XJ$ Now $\|JX - XJ\| \leqslant 2\|J\|\|X\|$ and thus

$$V_k^{-1} = k^{-1}\left(I - k^{-1}\mathcal{C}\right)^{-1} = k^{-1}(1 + o(1)); \quad (k \to \infty) \tag{77}$$

Therefore, the function kV_k is bounded for $k \in \mathbb{R}^+$.

We rewrite the system (68) in the form

$$Y_k = V_k^{-1}A_{k-1}J + V_k^{-1}\sum_{j=1}^{k-1} Y_j A_{k-j-1}; \quad k \in \mathbb{N} \tag{78}$$

or, in abstract form, with $\mathsf{Y} = \{Y_j\}_{j\in\mathbb{N}}$, $(\mathsf{LY})_k := V_k^{-1}\sum_{l=0}^{k-2} A_l Y_{k-1-l}$, where we regard Y as a function defined on $\mathbb{N}$ with matrix values, with the norm

$$\|\mathsf{Y}\| = \sup_{n\in\mathbb{N}} \|\mu^{-n}\mathsf{Y}(n)\|; \quad \mu > 1 \tag{79}$$

we have

$$\mathsf{Y} = \mathsf{Y}_0 + \mathsf{LY} \tag{80}$$

Exercise 3. *Show that (80) is contractive for μ sufficiently large, in an appropriate ball that you will find.*

The solution of this exercise is given in the appendix. $\square$

Note 9. *We have* **not** *seriously used the fact that J is a Jordan matrix in this proof. It follows that if J is replaced by any B with eigenvalues not differing by integers, we have $M = Y(z)z^B$.*

9 Scalar $n-$th order linear equations

These are equations of the form

$$y^{(n)} + a_1(z)y^{(n-1)} + \cdots + a_n(z)y = 0; \quad a_j \text{ analytic in } \mathbb{D}\setminus\{0\} \tag{81}$$

Definition 7. An equation of the form (81) has a singularity of the first kind at zero if

$$a_i(z) = b_i(z)/z^i \tag{82}$$

where b_i are analytic at zero. We will see shortly the reason for this terminology.

9.1 Connection between systems of equations and higher order scalar ones

There is an obvious way in which an n-th order equation can be transformed to an n-dimensional first order system. Take for simplicity a second order equation

$$y'' + a(z)y' + b(z)y = 0 \tag{83}$$

and we write the equivalent system

$$\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -b(z) & -a(z) \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} \Leftrightarrow \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -a(z) & -b(z) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{84}$$

with the notation $u_1 = y, u_2 = y'$. In general, of course, we take $v_0 = y, ..., v_k = y^{(k)}, ...$ and note that (81) is equivalent to

$$\begin{pmatrix} v_0 \\ v_1 \\ \cdots \\ v_{n-1} \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & \cdots & & \\ -a_n(z) & -a_{n-1}(z) & -a_{n-2}(z) & \cdots & -a_1(z) \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ \cdots \\ v_{n-1} \end{pmatrix} \tag{85}$$

A natural definition of a singularity of the first kind for n-th order equations is that the associated system has a first order singularity. But as mentioned, there is more than one way to arrive at a system.

For instance, the Euler equation

$$y'' = 6z^{-2}y \tag{86}$$

has as general solution

$$y = C_1 z^{-2} + C_2 z^3 \tag{87}$$

and we would expect it to correspond to a system with a singularity of the first kind, since the vector solution is bounded by a power of z near the origin. However, the system associated via (85) (or, here (202)) is

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = z^{-2} \begin{pmatrix} 0 & 1 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{88}$$

We can see the nature of the problem in the following way: in (86), if $y \sim z^m$ then $y' \sim mz^{m-1}$, $y/y' = O(z)$. This is quite general, certainly it is the case for equations admitting convergent series as solutions, as these can be differentiated term by term). In a system, no component should play a special role, but here we ended up with $u_1 = o(u_2)$. It is instead natural to take $u_1 = y/z$ and $u_2 = y'$, or, equivalently, $u_1 = y, u_2 = zy'$. We then get $u_2' = y' + zy'' = y' + 6z^{-1}y$ and thus the system

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \frac{1}{z} \begin{pmatrix} 0 & 1 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{89}$$

which is indeed a system with a singularity of first kind.

More generally, the natural substitution is

$$u_k = z^{k-1} y^{(k-1)}, \; k = 1, 2, ..., n \tag{90}$$

We then have

$$u_k' = (k-1)z^{k-2}y^{(k-1)} + z^{k-1}y^{(k)} \Rightarrow zu_k' = (k-1)u_k + u^{(k+1)} \tag{91}$$

while, as usual, u_n is special, since $y^{(n)}$ can be written in terms of lower order derivatives, using (81):

$$\begin{aligned} zu_n' = (n-1)u_n + z^n y^{(n)} = (n-1)u_n \\ - z^n(b_n z^{-n} y + b_{n-1} z^{-n+1} y' + ... + b_1 z^{-1} y^{(n-1)}) = (n-1)u_n \\ - b_n y - b_{n-1} z y' - ... - b_1 z^{n-1} y^{(n-1)} = (n-1)u_n - b_n u_1 - b_{n-1} u_2 - ... - b_1 u_n \end{aligned} \tag{92}$$

In matrix form, the end result is the system

$$\boldsymbol{u}' = z^{-1} B \boldsymbol{u} \tag{93}$$

where

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ & & \cdots & & & \\ -b_n(z) & -b_{n-1}(z) & -b_{n-2}(z) & -b_{n-3}(z) & \cdots & (n-1) - b_1(z) \end{pmatrix} \tag{94}$$

or, recalling (82),

$$\boldsymbol{u}' = z^{-1} B(0) \boldsymbol{u} + A(z) \boldsymbol{u} \tag{95}$$

where A is analytic at zero.

Corollary 8. *If an equation of the form (81) has a singularity of the first kind, then the singularity is regular (the general solution is a convergent combination of powers and logs).*

We will see that the converse is true also.

9.2 Possible solutions to the n-th order scalar equation

Assume that there is a set of n linearly independent of solutions of (82) in the form

$$y_j = \sum_{k,m=1}^{n'} H_{kmj}(z) z^{p_k} \ln^{l_m} z, \quad j = 1, ..., n \tag{96}$$

where the H_k are single-valued and l_m are integers. Since they are linearly independent, their Wronskian does not vanish at some point. From the general theory of systems we know then that it cannot vanish anywhere for $z \neq 0$. Thus the system of equations

$$\{y_j^{(n)} + a_1(z) y_j^{(n-1)} + \cdots + a_n(z) y_j = 0, j = 1, ..., n\} \tag{97}$$

allows for expressing the coefficients a_k as rational functions of $y_j^{(l)}$. Thus a_k grow at most algebraically at zero. Since they are single-valued in $\mathbb{D} \setminus \{0\}$, it follows that the coefficients are meromorphic.

Note also that if y is a solution of (81), then $y(ze^{i\beta})$ is a solution of the equation

$$y^{(n)}e^{-in\beta} + a_1(z)y^{(n-1)}e^{-i(n-1)\beta} + \cdots + a_n(z)y = 0 \tag{98}$$

As we proceeded in §7, we can define $\tilde{y}_j$ to be $y(ze^{2\pi i})$ (a notation for the analytic continuation of y on a circle around the origin). From (98) we get that $\tilde{y}_j$ are also solutions of (81) (which, of course, have to be linear combinations of y_j). In this way, we can eliminate the logs in the expressions (96), if there are any, see below.

Note 10. (i) For one term $f = H_{km}(z)z^{p_k} \ln^{l_m} z$ the difference between f and and its analytic continuation along a circle around the origin is

$$\hat{f} = H_{km}(z)z^{p_k}e^{2\pi i p_k}(\ln z + 2\pi i)^{l_m} - H_{km}(z)z^{p_k}(\ln z)^{l_m} \tag{99}$$

which is a sum of the form (96) with the powers of the logs reduced by one. So we can transform a solution with $\max l_m = M$ to one in which $\max l_m = M - 1$. Thus, applying M operations of the type (99) we get a solution of (81) with no logs. That is, there is at least one solution of the form

$$\hat{y}_0 = \sum_{k=1}^{n} H_k(z)z^{p_k} \tag{100}$$

Note also that if $\alpha = e^{2\pi i p_1}$ and $\beta = e^{2\pi i p_2}$ are different, then we can eliminate, say, the term $H_1(z)z^{p_1}$ by replacing $\hat{y}_0$ by $\hat{y}_0(z) - e^{-2\pi i p_1}\hat{y}_0(ze^{2\pi i})$. Proceeding in this manner, we see that there are solutions of the form

$$y_0 = H(z)z^a \tag{101}$$

where we can always assume $H(0) = 1$ since any other starting power than zero can be absorbed into a, and multiplicative constants don't matter since the equation is linear.

Exercise 1. (a) Substitute $y = y_1 g$ in (81) and show that g' satisfies an equation of type (81) of order lower by one.

(b)Show that the condition (82) is preserved by changes of dependent variable as in (a).

Note that if a first order scalar equation

$$y' + z^{-n}A(z)y = 0 \tag{102}$$

with A meromorphic admits a solution of the form

$$y(z) = z^a H(z)$$

with H analytic then

$$H' + az^{-1}H + AH = 0 \Rightarrow A = -az^{-1} - H'/H \tag{103}$$

and thus –since H'/H is has a pole of order at most one whenever H is meromorphic– A has a pole of order at most one.

(c) Use (a) and (b) to show that (81) has a complete set of solutions of the form (96), then (82) is satisfied.

9.3 Frobenius' theorem

Systematizing what we have obtained so far we have the following.

Definition 9. An equation of the form (81) has a regular singularity at zero if there exists a fundamental set of solutions in the form of finite combinations of functions of the form

$$y_i = z^{\lambda_i}(\ln z)^{m_i} f_i(z); \quad \text{(by convention, } f_i(0) \neq 0) \tag{104}$$

where f_i are analytic, $m_i \in \mathbb{N} \cup \{0\}$.

Theorem 4 (Frobenius)**.** *An equation of the form* (81) *has a regular singularity at zero iff the singular point is of the first kind, that is iff* (82) *holds.*

We see one advantage of the scalar formulation of a differential system: we have the Frobenius theorem as an "iff" statement (recall (88)).

Note 11. We could have allowed l_m to be noninteger, since, by converting the equation into a system and noting that the coefficients of the system are single-valued, Lemma 6 implies that l_m are necessarily $\in \mathbb{N} \cup \{0\}$.

9.4 The indicial equation

We saw in Note 10 that Frobenius type solutions can be found in the form $z^\lambda(1+o(1))$. We insert this into the differential equation and note that $y^{(j)} = \lambda(\lambda-1)\cdots(\lambda-j+1)z^{\lambda-j}(1+o(1))$ and also that $a_j z^{\lambda-n+j}(1+o(1)) = b_j(0)z^{\lambda-n})(1+o(1))$. Thus, the equation for the leading power of z is

$$\lambda(\lambda-1)\cdots(\lambda-n+1) + \lambda(\lambda-1)\cdots(\lambda-n+1)b_1(0) + \cdots + b_n(0) = 0 \tag{105}$$

This is the **indicial equation** and it determines all possible lowest powers("p_1") in a Frobenius-type solution–of the form (100).

9.4.1 Eigenvalues of $B(0)$ in (95)

The eigenvalue equation, $(B - \lambda I)x = 0$, is easy to solve. If we expand this out as a system, using the explicit form (149), we get

$$(0-\lambda)x_0 + x_1 = 0 \tag{106}$$
$$(1-\lambda)x_1 + x_2 = 0 \tag{107}$$
$$(2-\lambda)x_2 + x_3 = 0 \tag{108}$$
$$\ldots \tag{109}$$
$$-b_n(0)x_0 - b_{n-1}(0)x_1 - \cdots - (b_1 - [n-1-\lambda])x_{n-1} = 0 \tag{110}$$

Without loss of generality we can take $x_1 = 1$. Then

$$x_1 = \lambda, \; x_2 = \lambda(\lambda-1), \; \cdots, x_{n-1} = \lambda(\lambda-1)\cdots(\lambda-(n-2))$$

and thus (466) is equivalent to

$$-b_n(0)-\lambda b_{n-1}(0)-\cdots-(b_1-[(n-1)-\lambda])\lambda(\lambda-1)\cdots(\lambda-(n-2))=0 \quad (111)$$

which is precisely (105). We have shown

Proposition 10. *The eigenvalues of $B(0)$ are precisely the roots of the indicial equation.*

9.5 Equations of the form (81) with regular singularities; analysis using the system formulation

Note first that if we write (81) as a system, we get a fundamental matrix solution of the form (134). Singling out a Jordan block B , we see that we get

$$z^B = e^{\lambda z \ln B} = z^\lambda \left(1 + N\lambda \ln z + ... + \frac{\lambda^{m-1}(\ln z)^m}{(m-1)!}N^{m-1}\right)$$

which is a matrix of the form

$$z^\lambda \begin{pmatrix} 1 & \ln z & \frac{\ln^2 z}{2!} & \cdots & \frac{\ln^m z}{(m-1)!} \\ 0 & 1 & \ln z & \cdots & \frac{\ln^{m-1} z}{(m-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (112)$$

and the solution matrix is

$$z^\lambda \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1m} \\ S_{21} & S_{22} & \cdots & S_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ S_{m1} & S_{m2} & \cdots & S_{mm} \end{pmatrix} \begin{pmatrix} 1 & \ln z & \frac{\ln^2 z}{2!} & \cdots & \frac{\ln^m z}{(m-1)!} \\ 0 & 1 & \ln z & \cdots & \frac{\ln^{m-1} z}{(m-2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \quad (113)$$

By the substitution (90), the element 11 of the matrix product above is a solution y. Thus there is a solution of (81) of the form $z^\lambda S_{11}(z)$ where S_{11} is single-valued. On the other hand, this also needs to be a linear combination of the form (96). By the assumption on the nature of the solutions, S_{11} has a convergent series thus

$$y =: y_1 = z^{\lambda-(n-1)}F(z), \ n \in \mathbb{N} \quad (114)$$

where F is analytic. From this point on, we can proceed as in the previous section to conclude that the singular point is regular. Combining Proposition 10 with Theorem 3 and (331), we obtain the following result.

Proposition 11. *If the roots of the indicial equation do not differ by nonzero integers, then for a root λ of multiplicity m, there are m linearly independent solutions of* (81) *in the form*

$$y_{1,\lambda} = z^\lambda f_1(x), \ y_{2,\lambda} = z^\lambda f_1(x)\ln x + f_{11}(x), \ ...,$$
$$y_{m,\lambda} = z^\lambda f_1(x)\ln^m x + f_{m1}(x)\ln^{m-1} x + ... + f_{mm}(x) \quad (115)$$

9.5.1 Writing a system of equations as one higher order equation

Take first a simple case,

$$u' = au + bv, \quad v' = cu + dv; \quad (a, b, c, d \text{ analytic in some domain } \mathcal{D}) \tag{116}$$

and assume for simplicity that b does not vanish in $\mathcal{D}$. Then we write $v = b^{-1}(u' - au)$ and get

$$[b^{-1}(u' - au)]' = cu + db^{-1}(u' - au) \tag{117}$$

which expanded and normalized is an equation of the type (81), equivalent to (116). The same is true if $d \neq 0$. What if however this $b = d = 0$?

Let's look at the very simple system

$$u' = u; \quad v' = v; \quad \text{or, in matrix form,} \quad M' = M \tag{118}$$

Of course we can integrate it in closed form and the general solution is $u = C_1 e^z, v = C_2 e^z$. We would be tempted to say that there is no scalar equation equivalent to this. Indeed, a genuinely second order linear ODE should have two linear independent solution, since the existence of a solution to the IVP must exist and be unique. But we remember that in passing from an n-th order equation to a system we had a number of choice, far from equivalent. Here we can try something similar to (90). If we take $v = z^{-1}w$ in the second equation we get

$$u' = u; \quad w' = zv' + v = zv + v = w + z^{-1}w \tag{119}$$

and now we can find a second order system.

Exercise 2. Show that if $f_1, ..., f_n$ are analytic and the Wronskian does not vanish,

$$\begin{vmatrix} f_1^{(n-1)} & f_2^{(n-1)} & \cdots & f_n^{(n-1)} \\ f_1^{(n-2)} & f_2^{(n-2)} & \cdots & f_n^{(n-2)} \\ \cdots & \cdots & \cdots & \cdots \\ f_1 & f_2 & \cdots & f_n \end{vmatrix} \neq 0 \tag{120}$$

then there is an $n - th$ order equation of the form (81) having $f_1, f_2, ..., f_n$ as solutions. One way is to take a general equation of the form (81) and see if its coefficients can be determined from the fact that $f_1, f_2, ..., f_n$ are solutions.

Exercise 3. Show that by transformations of the type that led to (119) if necessary, we can find an n-th order scalar equation equivalent to a given system.

9.6 Examples of resonant and nonresonant second order equations

In view of what we know already about n-th order equations, it is clear that we can always arrange that one root of the indicial equation is zero, by a substitution of the form $y = z^\lambda y_1$. We look for simplicity at second order equations.

We assume that the second root, λ is nonzero real part, since a double root falls in the nondegenerate case that we looked at already.

We only need to look at the case $\mathrm{Im}\lambda \neq 0$ or if $\mathrm{Im}\lambda = 0$, then $\mathrm{Re}(n) > 0$.

Consider then the equation

$$z^2 y'' + (1-n)zy' + zy = 0 \tag{121}$$

The equation can be solved in terms of Bessel functions,

$$y = C_1 z^{n/2} J_n(2\sqrt{z}) + C_2 z^{n/2} Y_n(2\sqrt{z}) \tag{122}$$

The exact solution is not what we are looking for, and the equation is not very special otherwise. We use it as to illustrate the way the nature of the Frobenius solutions depend on the roots of the indicial equation,

$$\lambda(\lambda - n) = 0 \tag{123}$$

We look first for a power series solution starting with z^0. The recurrence relation for the coefficients is

$$c_1 = (n-1)c_0;\ m(m-n)c_m + c_{m-1} = 0, \quad m > 1 \tag{124}$$

The coefficient c_0 is arbitrary; we can take it for definiteness to be one. It is clear that if $n \notin \mathbb{N}$, the recurrence (124) has a solution and the solution is entire. There is a second solution, starting with z^n which, after division by z^n, is also entire.

If $n \in \mathbb{N}$ we see that the equation for c_n is $0 \cdot c_n = c_{n-1}$, that is, no power series solution with $c_0 = 1$ exists in this case (for exceptional equations, we may have $c_{n-1} = 0$, allowing for analytic solutions). The fact that there are solutions of the form $z^n Y(z)$ with $Y(z)$ analytic (in fact, entire) is at the origin of this phenomenon: for such a solution we have $c_{n-1} = 0$ and c_n is undetermined, as it should. This also suggests what we should try. The solution y_0 starting with $c_0 = 1$ is nonanalytic, and thus its monodromy is nontrivial. On the other hand by linearity and the usual arguments, $y_0(ze^{2\pi i}) - y_0(z)$ must be a solution of (121) which is $o(x^{n-1})$. Indeed, the solution must be a combination of powers and logs, the allowed powers are z^0 and z^n, and all coefficients up to c_{n-1} are determined uniquely. We expect then $y_0(ze^{2\pi i}) - y_0(z) = Cz^n Y(z)$ for some C which in turn suggests that $y_0(z) = c_0 + c_1 z + ... + c_{n-1}z^{n-1} + Cz^n \ln z + ...$

To keep things relatively simple, we take $n = 3$. If we substitute $y_0 = c_0 + c_1 z + c_2 z^2 + Cz^3 \ln z + c_4 z^4 + d_4 z^4 \ln z + ...$ in (121) we get

$$2c_1 = c_0,\ 2c_2 = c_1,\ 3C = -c_2,\ 4d_1 + 1 = 0, 4c_4 + 5Cd_1 + c_3 = 0$$

and so on, c_3 is *free* and the rest of the coefficients can now can be determined uniquely.

The general solution has thus the form

$$y_0 = C_1[A(z) + z^n \ln z Y(z)] + C_2 Y(z)$$

where A is analytic and Y is the second solution (entire).

Exercise 4. *Is $A(z)$ entire?*

10 Changing the eigenvalue structure of J by transformations

To solve the general case, in which eigenvalues *may* differ by positive integers, we find transformations which decrease one eigenvalue by one, leaving all others the same and without changing the structure of the ODE.

Write J in the form

$$J = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} \tag{125}$$

where J_1 is the Jordan block we care about, $\dim(J_1) = m \geqslant 1$, while J_2 is a Jordan matrix, consisting of the remaining blocks. The transformation we are looking for would change J into $J - I_1$ where

$$I_1 = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \tag{126}$$

where I is the identity matrix. That, in turn, would change the fundamental solution to

$$Y z^{J-I_1} \tag{127}$$

This suggests we try this change of variables in our equation. In matrix form,

$$M' = z^{-1} J M + A M \tag{128}$$

where we take $M = M_1 z^{I_1}$.

Exercise 1. *Show that, if $a \in \mathbb{C}$ and P is a projector, $P^2 = P$, then*

$$z^{aP} = P z^a + (I - P) \tag{129}$$

Is it true that $(z^B)' = z^{-1} B z^B$ for any matrix B?

We have

$$M_1' z^{I_1} + z^{-1} M_1 I_1 z^{I_1} = z^{-1} J M_1 z^{I_1} + A M_1 z^{I_1} \tag{130}$$

We can multiply to the right by z^{I_1} and get

$$M_1' = z^{-1} J M_1 - z^{-1} I_1 M_1 + A M_1 \tag{131}$$

which does not quite work, because of non-commutation. So it is natural to try

$$M = z^{I_1} M_1 = \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} M_1$$

since J and z^{I_1} commute. We then have

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} M_1 + \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} M_1' = z^{-1} J \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} M_1 + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} M_1 \tag{132}$$

We multiply to the left by

$$\begin{pmatrix} 1/z & 0 \\ 0 & 1 \end{pmatrix}$$

and get

$$M_1' = z^{-1} J M_1 + \begin{pmatrix} 1/z & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} z & A_{12} \\ A_{21} z & A_{22} \end{pmatrix} M_1$$
$$= z^{-1} J M_1 + \begin{pmatrix} A_{11} & A_{12}(0)/z + \tilde{A}_{12} \\ A_{21} z & A_{22} \end{pmatrix} M_1 = \frac{1}{z} \begin{pmatrix} J_1 & A_{12}(0) \\ 0 & J_2 \end{pmatrix} + \tilde{A} M \qquad (133)$$

where $\tilde{A}$ is analytic. We have thus obtained

Proposition 12. *By the change of variables $M = z^{I_1} M_1$, the equation for M_1 is of the form*

$$M_1' = z^{-1} R M_1 + \tilde{A} M_1 \qquad (134)$$

where R has eigenvalues $\lambda_1 - 1, ..., \lambda_m$.

Exercise 2. Use this procedure repeatedly to reduce any resonant system to a nonresonant one. That is done by arranging that the eigenvalues that differ by positive integers become equal.

Exercise 3. Use Exercise 2 to prove the following result.

Theorem 5. *Any system of the form*

$$y' = \frac{1}{z} B(z) y \qquad (135)$$

where B is an analytic matrix at zero, has a fundamental solution of the form

$$M(z) = Y(z) z^{B'} \qquad (136)$$

where B' is a constant matrix, and Y is analytic at zero. In the nonresonant case, $B' = B(0)$. In the resonant case, the eigenvalues of R do not differ by integers, and they are a subset of the eigenvalues of $B(0)$, precisely those that do not differ by integers of other eigenvalues, or, in the groups that do, the one which has the smallest real part.

Note that this applies even if $B(0) = 0$.

Exercise 4. *Find B' in the case where only two eigenvalues differ by a positive integer, where the integer is* 1.

10.1 Example

Let's consider again the equation

$$x(x-1) y'' + y = 0 \qquad (137)$$

We want to use the theory we have developed this far, to find the shape of the generic solution at $0, 1, \infty$ (the only singular points of the equation).

As usual, we write $u_1 = y, u_2 = xy'$.

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}' = \frac{1}{x}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ (x-1)^{-1} & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{138}$$

or

$$\mathbf{y} = x^{-1}B\mathbf{y} + A\mathbf{y} \tag{139}$$

where B has eigenvalues $0, 1$ (Tr$= 1$, det$= 0$), is resonant.

If instead we had tried the naive transformation $u_1 = y, u_2 = y'$ we get

$$\mathbf{y}' = x^{-1}B\mathbf{y} + A\mathbf{y}; \quad B := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad A := \begin{pmatrix} 0 & 1 \\ (1-x)^{-1} & 0 \end{pmatrix} \tag{140}$$

which is now nonresonant! This shows that resonance is not invariant under changes of variables, and that we may be able to reduce a resonant case to a nonresonant one by suitable transformations. The matrix B is brought to the Jordan normal form by the matrix

$$T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad T^{-1}BT = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{141}$$

$$B := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad A := \begin{pmatrix} 0 & 1 \\ 1/(1-x) & 0 \end{pmatrix} \tag{142}$$

It follows that the fundamental solution of this equation is

$$M = Y(x)x^B \tag{143}$$

where $Y(x)$ is analytic near zero (in this case, analytic in the unit disk, since $x = 1$ is the singular point closest to the origin (other than the origin itself).

Thus,

$$\begin{aligned} M = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}\left(I + \ln x \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) &= \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}\begin{pmatrix} 1 & \ln x \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} y_{11} & y_{11}\ln x + y_{12} \\ y_{21} & y_{21}\ln x + y_{22} \end{pmatrix} \end{aligned} \tag{144}$$

and thus, by applying M to some initial condition (a, b) we get that the general solution of (137) in a neighborhood of 0 is

$$y = Ay_{11} + B(y_{11}\ln x + y_{12}) \tag{145}$$

On the other hand, since the characteristic roots of (331) are $0, 1$, it must be that $y_{11}(0) = 0$ and the general solution is of the form

$$y = aA_1(x) + b(xA_2(x)\ln x + A_3(x)) = aA_3(x) + bx\ln xA_2(x)$$

where A_i are analytic.

10.2 Example:Bessel functions

The equation

$$f'' + \frac{3}{2x} f' + f = 0 \tag{146}$$

has the general solution

$$C_1 x^{-1/4} J_{1/4}(x) + C_2 x^{-1/4} Y_{1/4}(x) \tag{147}$$

where J and Y are Bessel functions.

The indicial equation is obtained by substituting $f(x) = x^\lambda$ in (146), and reads

$$a^2 + a/2 = 0 \Rightarrow a \in \{0, -\tfrac{1}{2}\} \tag{148}$$

The roots are nonresonant, and thus there

In matrix form, as in the example before, we have

$$\mathbf{f}' = M\mathbf{f} \tag{149}$$

where

$$M = \begin{pmatrix} 0 & 1 \\ -1 & -\dfrac{3}{2x} \end{pmatrix} = \frac{1}{x}\begin{pmatrix} 0 & x \\ -x & -\dfrac{3}{2} \end{pmatrix} = \frac{1}{x} B + A \tag{150}$$

where

$$B = \begin{pmatrix} 0 & 0 \\ 0 & -\dfrac{3}{2} \end{pmatrix}; \quad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{151}$$

and thus the fundamental matrix is

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} x^{\begin{pmatrix} 0 & 0 \\ 0 & -\frac{3}{2} \end{pmatrix}} = \begin{pmatrix} f_{11} & x^{-3/2} f_{12} \\ f_{21} & x^{-3/2} f_{22} \end{pmatrix} \tag{152}$$

This means that the general solution of the equation is of the form

$$f = C_1 f_1(x) + C_2 x^{-3/2} f_2(x) \tag{153}$$

with f_1 and f_2 analytic.

Exercise 5. *It also follows from the analysis above that $f_2(0) = 0$. Why?*

11 Some special functions and their regular singular points

Here is a good and up to date online source of information about special functions: http://dlmf.nist.gov/.

11.1 Hypergeometric functions

The general solution of the equation

$$x(x-1)y'' + [(a+b+1)x - c]y' + aby = 0 \tag{154}$$

is

$$y = A \cdot {}_2F_1(a,b;c;x) + Bx^{1-c}\ {}_2F_1(a-c+1, b-c+1; 2-c; x), \tag{155}$$

With the substitution

$$y = u, y' = v/x$$

we get

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \frac{1}{x} B \begin{pmatrix} u \\ v \end{pmatrix} + A(x) \begin{pmatrix} u \\ v \end{pmatrix} \tag{156}$$

where

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 1-c \end{pmatrix} \tag{157}$$

and

$$A = \begin{pmatrix} 0 & 0 \\ -ab/(x-1) & -(a+1+b-c)/(x-1) \end{pmatrix} \tag{158}$$

The eigenvalues of B are clearly 0 and $1-c$. Note that they are resonant when $c \in \mathbb{Z} \setminus \{1\}$.

Exercise 1. *(a) Find the behavior near the origin of the general solution in the nonresonant case.*

(b) In the resonant case, show that there is always a solution of the form $x^{1-c}A(x)$ if $1-c > 0$ and $A(x)$ otherwise, where A is analytic. Use reduction of order (explained in general in the next section) to find the behavior of the second solution. Reduction of order in the first case would mean: look for $y(x)$ in the form $x^{1-c}A(x)g(x)$ where $x^{1-c}A(x)$ is already a solution. Solve the equation for g.

11.2 The exponential integral

This is defined by

$$\mathrm{Ei}_1(z) = \int_z^\infty \frac{e^{-t}}{t} dt \quad (z \neq 0) \tag{159}$$

where the path does not cross the negative real axis or pass through the origin. There is a cut along the interval $(-\infty, 0]$. The function is also defined on $\mathbb{R}^-$, in terms of the principal part of the integral, as a multivalued function, but we will not worry about this now.

We see that

$$\mathrm{Ei}_1(z)' = -\frac{e^{-z}}{z} \tag{160}$$

With the substitution $\mathrm{Ei}_1(z) = g(z)e^{-z}$ we get

$$zg' - zg + 1 = 0 \tag{161}$$

We transform this into a second order homogeneous equation by differentiating once more in z:

$$g'' + (1/z - 1)g' - g/z = 0 \tag{162}$$

Clearly, zero is the only singular point of this equation. We write as before $g = u, g' = v/z$ and we get

$$u' = g' = v/z; \tag{163}$$

We get

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \frac{1}{x} B \begin{pmatrix} u \\ v \end{pmatrix} + A(x) \begin{pmatrix} u \\ v \end{pmatrix} \tag{164}$$

where

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{165}$$

where clearly the eigenvalues of B are $0, 0$. Note that

$$z^B = \begin{pmatrix} 1 & \ln z \\ 0 & 1 \end{pmatrix} \tag{166}$$

Write the general solution of (162) in a neighborhood of zero. Here, it is easy enough to find the behavior of $\mathrm{Ei}_1(z)$ directly from the integral expression. How?

11.3 Bessel functions

The Bessel functions of the first kind satisfy the equation

$$x^2 y'' + xy' + (x^2 - \nu^2)y = 0 \tag{167}$$

or, in normal form,

$$y'' + \frac{1}{x} y' + \left(1 - \frac{\nu^2}{x^2}\right) y = 0 \tag{168}$$

The general solution of this equation is

$$y = C_1 J_\nu(x) + C_2 Y_\nu(x) \tag{169}$$

In this case, the system is

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \frac{1}{x} B \begin{pmatrix} u \\ v \end{pmatrix} + A(x) \begin{pmatrix} u \\ v \end{pmatrix} \tag{170}$$

where

$$B = \begin{pmatrix} 0 & 1 \\ \nu^2 & 0 \end{pmatrix} \tag{171}$$

with eigenvalues $\nu, -\nu$. In the nonresonant case there are two solutions which behave near zero like

$$x^{\pm\nu} A_{\pm}(x) \tag{172}$$

with $A_{\pm}$ analytic.

The algorithm is clear. I attach a Maple file with the general procedure, in which, instead of $e1$ you would insert any second order ODE. Finally, let's make a simple connection with equilibria. If we have a system of the form

$$x' = ax + by \tag{173}$$

$$y' = cx + dy \tag{174}$$

and the associated matrix is diagonalizable, then we can bring it to the form

$$u' = \lambda_1 u; \quad v' = \lambda_2 v \tag{175}$$

Of course, this can be easily solved in closed form. But we also note that we can write

$$\frac{dv}{du} = b\frac{v}{u}; \quad b = \frac{\lambda_2}{\lambda_1} \tag{176}$$

which perhaps the simplest case we can think of within Frobenius theory. Suppose first that $b \in \mathbb{R}$, then based on Frobenius theory, it is very easy to draw the phase portrait. Discuss also the case when b is complex, and the case when the Jordan form of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is nontrivial.

11.4 Reduction of order

Let λ_1 be a characteristic root such that $\lambda_1 + n$ is not a characteristic root. Then, there is a solution of (**??**) of the form $y_1 = z^{\lambda_1}\varphi(z)$, where $\varphi(z)$ is analytic and we can take $\varphi(0) = 1$.

We can assume without loss of generality that $\lambda_1 = 0$. Indeed, otherwise we first make the substitution $y = z^{\lambda_1} w$ and divide the equation by z^{λ_1}.

The general term of the new equation is of the form

$$z^{-\lambda_1} b_l z^{-l} (z^{\lambda_1} w)^{n-l} = z^{-\lambda_1} b_l z^{-l} \sum_{j=0}^{n-l} \binom{n-l}{j} w^{(j)} (z^{\lambda_1})^{(n-l-j)}$$

$$= z^{-\lambda_1} b_l z^{-l} \sum_{j=0}^{n-l} \binom{n-l}{j} w^{(j)} z^{\lambda_1 - n + l + j} = b_l \sum_{j=0}^{n-l} \binom{n-l}{j} w^{(j)} z^{-(n-j)} \tag{177}$$

which is of the same type as (**??**).

```
> with(linalg):
> e1:=z^2*diff(y(z),z,z)+z*diff(y(z),z)+(z^2-nu^2)*y(z) = 0;
```

$$e1 := z^2\left(\frac{d^2}{dz^2}y(z)\right) + z\left(\frac{d}{dz}y(z)\right) + \left(z^2 - \nu^2\right)y(z) = 0 \quad (1)$$

```
> d2:=diff(y(z), z, z)=solve(e1,diff(y(z), z, z));
```

$$d2 := \frac{d^2}{dz^2}y(z) = -\frac{z\left(\frac{d}{dz}y(z)\right) + y(z)\,z^2 - y(z)\,\nu^2}{z^2} \quad (2)$$

```
> s1:=y(z)=u(z);
```

$$s1 := y(z) = u(z) \quad (3)$$

```
> s2:=diff(y(z),z)=v(z)/z;
```

$$s2 := \frac{d}{dz}y(z) = \frac{v(z)}{z} \quad (4)$$

```
> diff(s2,z);
```

$$\frac{d^2}{dz^2}y(z) = \frac{\frac{d}{dz}v(z)}{z} - \frac{v(z)}{z^2} \quad (5)$$

```
> subs(d2,%);
```

$$-\frac{z\left(\frac{d}{dz}y(z)\right) + y(z)\,z^2 - y(z)\,\nu^2}{z^2} = \frac{\frac{d}{dz}v(z)}{z} - \frac{v(z)}{z^2} \quad (6)$$

```
> subs(s2,%);
```

$$-\frac{v(z) + y(z)\,z^2 - y(z)\,\nu^2}{z^2} = \frac{\frac{d}{dz}v(z)}{z} - \frac{v(z)}{z^2} \quad (7)$$

```
> subs(s1,%);
```

$$-\frac{v(z) + u(z)\,z^2 - u(z)\,\nu^2}{z^2} = \frac{\frac{d}{dz}v(z)}{z} - \frac{v(z)}{z^2} \quad (8)$$

```
> solve(%,diff(v(z),z));
```

$$\frac{u(z)\left(-z^2 + \nu^2\right)}{z} \quad (9)$$

```
> expand(%);
```

$$-u(z)\,z + \frac{u(z)\,\nu^2}{z} \quad (10)$$

```
> mm:=matrix([[0,1],[nu^2,0]]);
```

$$mm := \begin{bmatrix} 0 & 1 \\ \nu^2 & 0 \end{bmatrix} \quad (11)$$

```
> jordan(%);
```

Figure 1:

Thus we assume $\lambda_1 = 0$ and take $y = \varphi w$. As discussed, we can assume

$\varphi(0) = 1$. The equation for w is

$$\sum_{l=0}^{n} z^{-l} b_l \sum_{j=0}^{n-l} \binom{n-l}{j} w^{(j)} \varphi^{(n-l-j)} = 0 \tag{178}$$

or

$$\sum_{j=0}^{n} w^{(j)} \sum_{l=0}^{n-j} z^{-l} b_l \binom{n-l}{j} \varphi^{(n-l-j)} = 0 \tag{179}$$

or also

$$\sum_{j=0}^{n} w^{(n-j)} \sum_{l=0}^{j} z^{-l} b_l \binom{n-l}{j} \varphi^{(j-l)} = 0 \tag{180}$$

We note that this equation, after division by φ (recall that $1/\varphi$ is analytic) is of the same form as (**??**). However, now the coefficient of w is

$$\sum_{l=0}^{n} z^{-l} b_l \binom{n-l}{0} \varphi^{(n-l)} = \sum_{l=0}^{n} z^{-l} b_l \varphi^{(n-l)} = 0 \tag{181}$$

since this is indeed the equation φ is solving.

We divide the equation by φ (once more, remember $\varphi(0) = 1$), and we get

$$\sum_{j=0}^{n-1} w^{(1+(n-1-j))} \tilde{b}_j = 0 \tag{182}$$

where

$$\tilde{b}_j = \sum_{l=0}^{j} z^{-l} b_l \binom{n-l}{j} \frac{\varphi^{(j-l)}}{\varphi} \tag{183}$$

has a pole of order at most j, or

$$\sum_{j=0}^{n-1} g^{(n-1-j)} \tilde{b}_j = 0 \tag{184}$$

with $w' = g$. This is an $(n-1)$th order equation for g, and solving the equation for w reduced to solving a lower order equation, and one integration, $w = \int g$.

Thus, by knowing, or assuming to know, one solution of the nth order equation, we can reduce the order of the equation by one. Clearly, the characteristic roots for the g equation are $\lambda_i - \lambda_1 - 1$, $i \neq 1$. We can repeat this procedure until the equation for g becomes of first order, which can be explicitly solved. This shows what to do in the degenerate case, other than, working in a similar (in some sense) way with the equivalent nth order system.

12 Nonlinear systems

A point, say $z = 0$ is a singular point of the first kind of a nonlinear system if the system can be written in the form

$$y' = z^{-1}h(z, y) = z^{-1}(L(z)y + f(z, y)) \tag{185}$$

where h is analytic in z, y in a neighborhood of $(0, 0)$. We will not analyze these systems in detail, but much is known about them, [6] [2]. The problem, in general, is nontrivial and the most general analysis to date for one singular point is in [6], and utilizes techniques beyond the scope of our course now. We present, without proofs, some results in [2], which are more accessible. They apply to several singular points, but we will restrict our attention to just one, in the setting of (185). In the nonlinear case, a "nonlinear nonresonance" condition is needed, namely: if λ_i are the eigenvalues of $L(0)$, we need a *diophantine condition*: for some $\nu > 0$ we have

$$\inf \left\{ (|\mathbf{m}|+k)^{\nu} |k+\mathbf{m}\cdot\boldsymbol{\lambda}-\lambda_i| \;\middle|\; \mathbf{m} \in \mathbb{N}^n, |\mathbf{m}| > 1, k \in \mathbb{N}\cup\{0\}; i \leqslant n \right\} > 0 \tag{186}$$

Furthermore, $L(0)$ is assumed to be diagonalizable. (In [6] a weaker nonresonance condition is imposed, known as the Brjuno condition, which is known to be optimal.)

Proposition 13. *Under these assumptions, There is a change of coordinates $y = \Phi(z)u(z)$ where Φ is analytic with analytic inverse, so that the system becomes*

$$u' = z^{-1}h(z, u) = z^{-1}(Bu + f(z, u)) \tag{187}$$

where B is a constant matrix.

Proposition 14. *The system (187) is analytically equivalent* in a neighborhood of $(0, 0)$*, that is for small u as well as small z, to its linear part, namely to the system*

$$w' = z^{-1}Bw \tag{188}$$

In terms of solutions, it means that the general *small* solution of (185) can be written as

$$y = H(z, \Phi(z)z^B C) \tag{189}$$

where $H(u, v)$ is analytic as a function of two variables, C is an arbitrary constant vector. The diophantine, and more generally, Brjuno condition is generically satisfied. If the Brjuno condition fails, equivalence is still possible, but unlikely. The structure of y in an equation of the form (189) is

$$y_j(z) = \sum_{m,k} c_{k,\mathbf{m}} z^k z^{\mathbf{m}\cdot\boldsymbol{\lambda}} \tag{190}$$

13 Variation of parameters

As we discussed, a linear nonhomogeneous equation can be brought to a linear homogeneous one, of higher order. While this is useful in a theoretical quest, in practice, it is easier to solve the associated homogeneous system and obtain the solution to the nonhomogeneous one by integration. Indeed, if the matrix equation

$$Y' = B(z)Y \tag{191}$$

has the solution $Y = M(z)$, then in the equation

$$Y' = B(z)Y + C(z) \tag{192}$$

we seek solutions of the form $Y = M(z)W(z)$. We get

$$M'W + MW' = B(z)MW + C(z) \quad \text{or} \quad M(z)W' = C(z) \tag{193}$$

giving

$$Y = M(z)\int_a^z M^{-1}(s)C(s)ds \tag{194}$$

14 Equilibria

We start with the simple example of the physical pendulum. It is helpful in a number of ways, since we have a good intuitive understanding of the system. Yet, the ideal (frictionless) pendulum has nongeneric features.

We can use conservation of energy to write

$$\frac{1}{2}mv^2 + mgl(1-\cos x) = const \tag{195}$$

where x is the angle and $v = dx/dt$, so with $l = m = 1$ we get

$$x'' = -\sin x \tag{196}$$

14.1 Exact solutions

This equation can be solved exactly, in terms of Weierstrass elliptic functions. Integration could be based on (197), and also by multiplication by x' and integration, which leads to the same.

$$\frac{1}{2}x'^2 - \cos x = C \tag{197}$$

$$\int_0^x \frac{ds}{\sqrt{C + 2\cos s}} = t + t_0 \tag{198}$$

With the substitution $\tan(x/2) = u$ we get

$$\int_0^{\tan(x/2)} \frac{du}{\sqrt{1+u^2}\sqrt{C+1+(C-1)u^2}} = t + t_0 \tag{199}$$

Whenever a differential system can be reduced to mere integrations as above, we say that the system is integrable by quadratures. On the other hand, by definition the elliptic integral of the first kind, $F(z,k)$ is defined as

$$F(z,k) = \int_0^z \frac{ds}{\sqrt{1-s^2}\sqrt{1-k^2s^2}} \tag{200}$$

and we get, with $K = \sqrt{2}/\sqrt{1+C}$,

$$iKF\left(\cos(z/2), K\right)\Big|_0^x = t + t_0 \tag{201}$$

Inverting, we see that x is a Weierstrass elliptic function of t. At this point, we should study elliptic functions to proceed. They are in fact very interesting and worthwhile studying, but we'll leave that for later. For now, it is easier to gain insight on the system from the equation than from the properties of elliptic functions.

14.2 Discussion and qualitative analysis

Written as a system, we have

$$\begin{aligned} x' &= v \\ v' &= -\sin x \end{aligned} \tag{202}$$

The point $(0,0)$ is an equilibrium, and $x = 0, v = 0$ is a solution. So are the points $x = n\pi, v = 0$, $n \in \mathbb{N}$. In general, a point $\mathbf{x}_0$ is an equilibrium of the equation $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ if $\mathbf{F}(\mathbf{x}_0) = 0$, implying that $\mathbf{x}_0$ is a solution. The physical interpretation here is clear, if $x = 0, v = 0$ the system is in equilibrium.

Note that (202) is a *Hamiltonian system*, i.e., it is of the form

$$\begin{aligned} x' &= \frac{\partial H(x,v)}{\partial v} \\ v' &= -\frac{\partial H(x,v)}{\partial x} \end{aligned} \tag{203}$$

where $H(x,v) = \frac{1}{2}v^2 + 1 - \cos x$. For Hamiltonian systems, we see that H is a *conserved quantity*, that is $H(x(t), v(t)) = const$ along a given *trajectory* $\{(x(t), v(t)) : t \in \mathbb{R}\}$ (just calculate $\frac{d}{dt}H$). The trajectories are thus the level lines of H, that is

$$H(x,v) = \frac{1}{2}v^2 + 1 - \cos x = E \tag{204}$$

(we artificially added 1 to make $H \geqslant 0$, since H is defined by the differential system up to an additive constant). Of course, not every 2-dimensional system $y' = F(y)$ is Hamiltonian; one obvious condition is $\nabla F = 0$. We will return to this later.

Drawing the phase portrait of the system (say two-dimensional) means plotting the vector field F, its special points, trajectories of interest and so on. A trajectory is the set $\{(x(t), y(t)) : t \in A \subset \mathbb{R}\}$ where A is typically taken to be the domain of existence of the solution.

In our example each trajectory is associated with a given value of $E = H$ and the shape depends on E.

The trajectories are level sets of H; at all points where $\nabla H \neq 0$, that is where the right side of (203) is nonzero, by the implicit function theorem, the trajectories are analytic (more generally if H is smooth, so will be the trajectories); otherwise they are typically singular.

Let first $0 < E < 2$. We have $1 - \cos x = 2\sin^2(x/2) = E - v^2/2 < 2$ and thus $x \in (-\alpha, \alpha)$, $\alpha < \pi/2$). Also, $v^2/2 = E - 2\sin^2(x/2) < 2$ and thus $\{(x(t), v(t)) : t \in \mathbb{R}\}$ are compact sets; in fact the curves are closed (why?) and non-intersecting (since $\nabla H \neq 0$), smooth boundaries of the domains $H \leqslant E$. Since $\nabla H = 0$ only at $(0, 0)$ and $H > 0$ otherwise, the maximum of H occurs on the boundary of $\{(x, v) : H(x, v) \leqslant E\}$.

Physically, for initial conditions close to zero, the pendulum would periodically swing around the origin, with amplitude limited by the total energy.

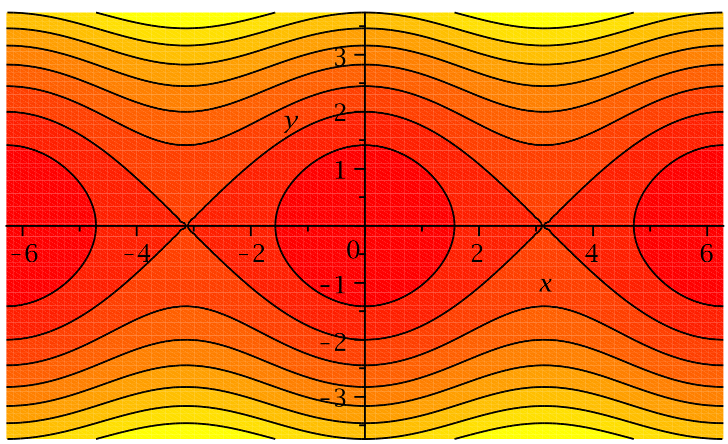


Figure 2: Contour plot of $v^2/2 - \cos x$

Fig. 12 represents a numerical contour plot of $v^2/2 - \cos x$. If we zoom in, we see that the program had difficulties at the critical points $\pm\pi$, showing once more that there is something singular there.

14.3 Linearization of the phase portrait

Let's first analyze the system approximately for small E. Along any such trajectory, x is small too and we have $E = H(x, v) \approx \frac{1}{2}v^2 + \frac{1}{2}x^2$. The trajectories are approximately circles (this needs a more serious discussion, coming shortly). The flow for this approximate Hamiltonian is $x' = v, v' = -x$ or $x'' = -x$, the harmonic oscillator.

Take

$$(1 - u^2/2) = \cos x; \quad u \in [-2, 2] \quad \text{or } u^2 = 4\sin(x/2)^2 \tag{205}$$

which defines two holomorphic changes of coordinates

$$u = \pm 2\sin(x/2) \tag{206}$$

These are indeed biholomorphic changes of variables until $\sin(x/2)' = 0$ that is, $x = \pm\pi$. With any of these changes of coordinates we get

$$\frac{u}{\sin x}u' = v \tag{207}$$

$$v' = -\sin x \tag{208}$$

or

$$uu' = v\sin x \tag{209}$$

$$v' = -\sin x \tag{210}$$

which would give the same trajectories family as

$$u' = v \tag{211}$$

$$v' = -u \tag{212}$$

for which the exact solution, $A\sin t, A\cos t$ gives rise to circles. The same could have been seen easily seen by making the same substitution, (215) in (204). We note again that in (215) we have $u^2 \in [0, 4]$, so the equivalence does not hold beyond $u = \pm 2$. The level sets $H \leqslant E < 2$ are analytically conjugated to circles.

What about the other equilibria, $x = (2k+1)\pi$? It is clear, by periodicity and symmetry that it suffices to look at $x = \pi$. If we make the change of variable $x = \pi - s$ we get

$$s' \quad = -v \tag{213}$$

$$v' \quad = -\sin s \tag{214}$$

In this case, the same change of variable, $u = 2\sin(s/2)$ gives a set of orbits equivalent to

$$u' = v \tag{215}$$

$$v' = u \tag{216}$$

implying $v^2 - u^2 = E$ as long as the change of variable is meaningful, that is, for $u < 2$, or $|s| < \pi$. So the curves associated to (213) are analytically conjugated to the hyperbolas $v^2 - u^2 = E$. The equilibrium is unstable, points starting nearby necessarily moving far away. The point $\pi, 0$ is a saddle point.

The trajectories starting at π are *heteroclinic*: they link different saddles of the system. Only "exceptional" systems have heteroclinic trajectories (or homoclinic ones, connecting a fixed point to itself.

In our case, heteroclinic trajectories correspond to $E = 2$ and this gives

$$v^2 = 2(1 + \cos(x)) \tag{217}$$

or

$$v^2 = 4\cos(x/2)^2 \tag{218}$$

that is, the trajectories are given explicitly by

$$v = \pm 2\cos(x/2) \tag{219}$$

This is a case where the elliptic function solution reduces to elementary functions: The equation

$$\frac{dx}{dt} = 2\cos(x/2) \tag{220}$$

has the solution

$$x = 2\arctan(\sinh(t + C)) \tag{221}$$

We see that the time needed to move from one saddle point to the next one is infinite.

Note that we can fully describe the trajectories –in terms of elementary functions. In the process however, the time dependence, which was the parametrization, is lost. It is a different matter to solve (202).

14.4 Connection to regularly perturbed equations

Note that at the equilibrium point $(\pi, 0)$ the system of equations is analytically equivalent, insofar as trajectories go, to the system

$$\begin{pmatrix} x \\ v \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ v \end{pmatrix} \tag{222}$$

The eigenvalues of the matrix are ± 1 with (unnormalized) eigenvectors $(1, 1)$ and $(-1, 1)$. Thus, the change of variables to bring the system to a diagonal form is $x = \xi + \eta, v = \xi - \eta$. We get

$$\xi' + \eta' = \xi - \eta \tag{223}$$
$$\xi' - \eta' = \xi + \eta \tag{224}$$

By adding and subtracting these equations we get the diagonal form

$$\xi' = \xi \tag{225}$$
$$\eta' = -\eta \tag{226}$$

or

$$\frac{d\xi}{d\eta} = -\frac{\xi}{\eta};\ or\ \xi_\eta + \frac{1}{\eta}\xi = 0 \tag{227}$$

a standard regularly perturbed equation. Clearly the solutions of (227) are $\xi = C/\eta$ with $C \in (-\infty, \infty)$, and insofar as the phase portrait goes, we could have written $\eta_\xi + \frac{1}{\xi}\eta = 0$, which means that the trajectories are the curves $\xi = C/\eta$ with $C \in [-\infty, \infty]$, hyperbolas and the coordinate axes. In the original variables, the whole picture is rotated by $45°$.

14.5 Completing the phase portrait

We see that, for $H > 2$ we have

$$v = \pm\sqrt{2h + 2\cos(x)} \tag{228}$$

where now $h > 2$. With one choice of branch of the square root (the solutions are analytic, after all), we see that $|v|$ is bounded, and it is an open curve, defined on the whole of $\mathbb{R}$. Note that the explicit form of the trajectories, given by (204) does not, in general, mean that we can solve the second order differential equation. The way the pendulum position depends on time, or the way the point moves along these trajectories, is still transcendental.

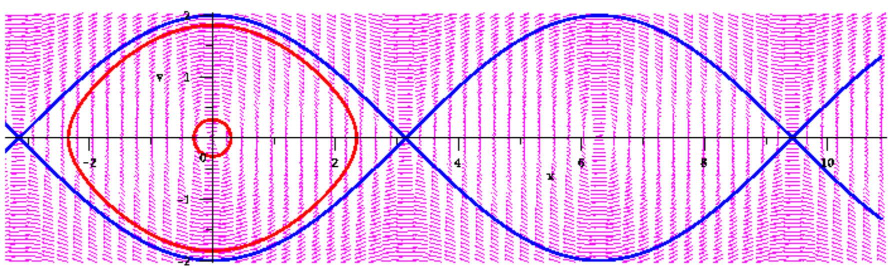


Figure 3: Contour plot of $y^2/2 - \cos x$

14.6 Local and asymptotic analysis

Near the origin, for $C = a^2$ small, we have

$$x' = v \tag{229}$$
$$v' = x - x^3/6 + ... \tag{230}$$

implying

$$x' = v \tag{231}$$
$$v' \approx x \tag{232}$$

which means

$$x \approx a \sin t \tag{233}$$

$$v \approx a \cos t \tag{234}$$

For C very large, we have

$$\frac{dx}{\sqrt{C + \cos x}} = dx(C + \cos x)^{-1/2} = dxC^{-1/2}(1 + \cos x/C)^{-1/2}$$

$$= dx(C^{-1/2} - \frac{1}{2} \cos x/C^{-3/2} + \cdots) \tag{235}$$

which means

$$C^{-1/2}x + \frac{1}{2} \sin x/C^{-3/2} + \cdots = t + t_0 \tag{236}$$

or

$$x = C^{1/2}(t + t_0) - \frac{1}{2} \sin(C^{1/2}t)/C^{-3/2} + \cdots = \tag{237}$$

The solutions near the critical point $(\pi, 0)$ can be analyzed similarly.

Local and asymptotic analysis often give sufficient qualitative, and sometimes quantitative information about all solutions of the equation.

15 Equilibria

In [7], Chapter 1.3 about nonlinear systems starts with the words:

"We must start by admitting that almost nothing beyond general statements can be made about most nonlinear systems. In the remainder of this book we will meet some of the delights and horrors about such systems, but the reader must bear in mind that the line of attack we develop in this text is only one and that any other tool in the workshop of applied mathematics, including numerical integration, perturbation methods, and asymptotic analysis, can and should be brought to bear on a specific problem". Since the book was written, there has been substantial progress, especially in using tools of asymptotic analysis, to find the behavior of nonlinear systems. We will see about these later but for now, we start with classical results and tools.

15.1 Flows

Consider the system

$$\frac{dx}{dt} = F(x) \tag{238}$$

where F is smooth enough. Such equations can be considered in $\mathbb{R}^n$ or, more generally, in Banach spaces.

As we know by now, if x_0 is a regular point for F, then there exists a unique local solution of (238) with $x(0) = x_0$.

Remark 12. *(a) Note that equilibria, defined as points where* $F(x_e) = 0$ *are singular points of the* field*.* Trajectories *can intersect there. But this does not mean that* flows *are singular there. Indeed, if we write*

$$x(t) = x_0 + \int_0^t F(x(s))ds$$

and the field is smooth at x_0*, the map above is contractive, and there is a unique solution. In particular, if* $x_0 = x_e$ *then* $x(t) \equiv x_e$*. See also Remark 14 below.*

The initial condition x_0 is mapped, by the solution of the differential equation (238) into $x(t)$ where $t \in (-a, b)$.

The map $x(0) \to x(t)$ written as $f^t(x_0)$ is the flow associated to F.

For $t \geqslant 0$ we note the (commutative) semigroup property $f^0 = I, f^{s+t} = f^s f^t$. This follows from uniqueness of solutions, giving $x(t+s; x_0) = x(s; x(t; x_0))$.

Fixed points, hyperbolic fixed points in $\mathbb{R}^n$. **Example**. If $F(x) = Bx$ where B **does not depend on** x, then the general solution is

$$x = e^{Bt} x_0 \tag{239}$$

where x_0 is the initial condition at $t = 0$. (Note again that a simple exponential formula does not exist, in general, if B depended on t.)

In this case, the flow f is given by the linear map

$$f^t(x_0) = e^{DF(0)t} x_0 \tag{240}$$

Note that $(D_x f)(0) = e^{Bt}$.

Note 13. Remember that the eigenvalues of $e^{B\alpha}$ are $e^{\lambda_i \alpha}$ where λ_i are the eigenvalues of B.

Definition 15. *The point* x_0 *is a fixed point of* f *if* $f^t(x_0) = x_0$ *for all* t.

Proposition 16. *If* f *is associated to* F*, then* x_0 *is a fixed point of* f iff $F(x_0) = 0$.

Proof. Indeed, we have $x(t + \Delta t) = x(t) + F(x_0)\Delta t + o(\Delta t)$ for small Δt. Then $x(t + \Delta t) = x(t)$ implies $F(x_0) + o(\Delta t) \Rightarrow F(x_0) = 0$. Conversely, it is obvious that $F(x_0) = 0$ implies that $x(t) = x_0$ is a solution of (238), and this solution is unique, by Remark 12. ☐

Remark 14. *This also shows that if closures of trajectories intersect (as sets) at an equilibrium, then along any nontrivial trajectory ending at* x_0 *(that is, a trajectory other that* $x(t) = x_0$*, we must have* $x(t) \neq x_0$ *for all* $t \in \mathbb{R}$ *and thus* $x_0 = \lim_{t\to\infty} x(t)$ *(possibly along a subsequence).*

Assume 0 is a fixed point of f, $F(0) = 0$. The flow f depends on two variables, x_0 and t. Since $x(t; x_0) = f^t(x_0)$, we clearly have

$$\frac{\partial f}{\partial t} = F(x(t; x_0)) = F(f_t(x_0)) \tag{241}$$

To see what $\dfrac{\partial f}{\partial x_0}$ is near 0, we write

$$x' = F(x) = F(0) + (DF)(0)x + o(x) = (DF)(0)x + o(x) \tag{242}$$

We thus expect to leading order

$$x' = (DF)(0)x \Rightarrow x = e^{t(DF)(0)}x_0 \Rightarrow \frac{\partial f}{\partial x_0} = e^{t(DF)(0)} \tag{243}$$

This is indeed the case, and it is shown below.

Proposition 17. *If f is associated to the C^1 field F with $F \in C^1$ differentiable, and x_0 is a fixed point of f, then $D_x f^t|_{x=x_0} = e^{DF(x_0)t}$.*

That is, the flow is tangent to the linear flow.

Proof. Without loss of generality we take $x_0 = 0$. Take the initial condition $x(t = 0) = x_0$ small enough. Let $DF(0) = B$. We have $F(x) = Bx + g(x)$ where $g(x) = o(x)$ for small x. Note that we also have for x_1, x_2 close to 0 and close to each other, because $F \in C^1$,

$$F(x_2) - F(x_1) = DF(x_2)(x_2 - x_1) + o(x_2 - x_1) \tag{244}$$

and thus

$$F(x_2) - F(x_1) - B(x_2 - x_1) = g(x_2) - g(x_1) = o(1)(x_2 - x_1) \tag{245}$$

We have $x' = Bx + g(x)$. Taking $x = e^{Bt}u$ we get

$$e^{Bt}u' + Be^{Bt}u = Be^{Bt}u + g(e^{Bt}u) \tag{246}$$

Thus

$$u = x_0 + \int_0^t e^{-Bs} g(e^{Bs}u(s))ds \tag{247}$$

or

$$x = e^{Bt}x_0 + \int_0^t e^{B(t-s)} g(x(s))ds \tag{248}$$

Take initial the initial condition in $\mathcal{N} = \{x_0 | |x_0| < \delta\}$. Let $\mathcal{B}$ be the Banach space of functions defined on $[0, T]$ with the sup norm, and the ball B_0 of radius $2e^{\|B\|T}|x_0|$.

We claim that (248) is contractive in B_0, if δ is small. Indeed, you can easily check that B_0 is preserved since $g(x) = o(x)$.

To show contractivity, we note that

$$x_1(t) - x_2(t) = \int_0^t e^{B(t-s)}[g(x_1(s)) - g(x_2(s))]ds \tag{249}$$

where we know that

$$\|g(x_1(s)) - g(x_2(s))\| = o(1)\|x_1(s) - x_2(s)\| \tag{250}$$

The rest of the contractivity proof is straightforward. We have by the definition of B_0

$$\left|\int_0^t e^{B(t-s)}g(x)dx\right| \leqslant |x_0|o(1)e^{\|B\|T}\|B\|^{-1} = o(x_0) \tag{251}$$

and thus

$$x = e^{Bt}x_0 + o(x_0)$$

proving the statement $\square$

.

15.2 Linearizations. The Hartman-Grobman theorem

Definition 18. • *The fixed point $x = 0$ is hyperbolic if the matrix $D_x f|_{x=0}$ has no eigenvalue on the unit circle.*

• *Equivalently, if f is associated with F, the fixed point 0 is hyperbolic if the matrix $DF(0)$ has no purely imaginary eigenvalues.*

The following result generalizes to Banach space settings.

Let U and V be open subsets of $\mathbb{R}^n$. Let f be a diffeomorphism between U and V with a **hyperbolic** fixed point, that is there is $x_0 \in U \cap V$ so that $f(x_0) = x_0$ and $Df(x_0)$ has no spectrum on the unit circle. Without loss of generality, we may assume that $x = 0$.

The following result shows that there is a continuous change of coordinates which transforms f into its linear part.

Theorem 6 (Hartman-Grobman for maps)**.** *Under these assumptions, f and $Df(0)$ are topologically conjugate, that is, there are neighborhoods U_1, V_1 of zero, and a homeomorphism h from U_1 to V_1 so that $h^{-1} \circ f \circ h = Df(0)$.*

The proof is not very difficult, but it is preferable to leave it for later; we however illustrate it on some simple cases.

15.3 Conjugation of maps, a simple case: one-dimensional analytic maps

We note that the higher dimensional version of the analytic linearization result below is **not** a simple extension of the one-dimensional case.

A quadratic map, $f(x) = \alpha x + \beta x^2$ contains the core of the problem, while being algebraically easier to handle. Note that by the substitution $f(x) = a^{-1}\tilde{f}(ax)$ we get

$$\tilde{f}(t) = \alpha t + a^{-1}\beta t^2 = \alpha t + t^2 \quad \text{if } a = \beta \tag{252}$$

and we can assume without loss of generality that $\beta = 1$. We are looking for conjugation map, "tangent to the identity" ($h'(0) = 1$, to ensure that the existence of h means that the maps are close to each-other) with the property

$$h^{-1} \circ f \circ h = \alpha I, \quad \text{where } I \text{ is the identity} \tag{253}$$

In our example, it means

$$h(\alpha y) = \alpha h(y) + h(y)^2 \tag{254}$$

Let's first see why hyperbolicity is important.

First of all, obviously with $\alpha = 1$ we cannot have such an h. Neither can we find such a map if $\alpha = -1$. Indeed, in this case

$$h(-y) = -h(y) + h(y)^2 \Rightarrow h(-y) = h(-y) - 2h(-y)^3 + h(-y)^4 \tag{255}$$

and since $h = x + o(1)$ this is impossible as well. You can show that if ω is an n-th root of unity, we also get a contradiction. Now, for one-dimensional maps, it is not necessary that f be hyperbolic for h to exist. The Brjuno condition measuring how far an irrational angle is from the rationals is in fact necessary and sufficient for the existence of an analytic h.

Let's now take $|\alpha| \neq 1$. We are looking for $h(x) = x + O(x^2)$ and then it is natural to substitute $h(x) = x + x^2\delta(x)$ in the conjugation equation (260).

We get

$$\alpha x + \alpha^2 x^2 \delta(\alpha x) = \alpha x + \alpha x^2 \delta(x) + x^2 + 2x^3\delta(x) + x^4\delta(x)^2 \tag{256}$$

If $|\alpha| > 1$ we isolate $\delta(\alpha x)$,

$$\delta(\alpha x) = \alpha^{-2} + \alpha^{-1}\delta(x) + 2\alpha^{-2}x\delta(x) + x^2\alpha^{-2}\delta(x)^2 \tag{257}$$

or, with $\alpha = 1/t$, $\alpha x = z$,

$$\delta(z) = t\delta(zt) + t^2 + 2t^3 z\delta(zt) + t^4 z^2 \delta(zt)^2 \tag{258}$$

If $|\alpha| < 1$ we isolate $\delta(x)$ and write $t = \alpha$,

$$\delta(x) = -\frac{1}{t} + t\delta(tx) - 2(x/t)\delta(x) - (x^2/t)\delta(x)^2 \tag{259}$$

Both equations (258) and (259) are contractive for small x. Let's check for instance (259). To make the map "manifestly" contractive, we substitute $\delta(z) = -1/t + u(z/\varepsilon)$ in (259) and get

$$u(z) = [-\varepsilon + 2\varepsilon^2 z t^{-2} - \varepsilon^3 t^{-3}] + tu(tz) - 2\varepsilon z t^{-1} u(z) z^2 + 2t^{-2}\varepsilon^2 z^2 u(z) - t^{-1}\varepsilon z^2 u(z)^2 \tag{260}$$

You can now check that (260) is contractive in a disk of radius $2\varepsilon/(1 - |t|)$ in the space of functions analytic for $|z| < \varepsilon$ if ε is small enough.

The proof for (258) follows similar steps (the algebra is in fact slightly simpler).

In a general one-dimensional case, we write $f(y) = \alpha y + y^2 f_1(y)$. Let $|\alpha| > 1$. The equation for δ reads (we take again $\alpha = 1/t$)

$$\delta(y) = t\delta(ty) + t^2 y^2 (1 + ty\delta^2(ty))^2 f_1(ty + t^2 y^2 \delta(ty)) \tag{261}$$

This map is contractive given enough regularity of f, by similar arguments.

To apply this to flows, we would linearize the flow. A toy model is to take a discrete evolution

$$x(t+1) = \alpha x(t) + x(t)^2 \tag{262}$$

and substitute $x = h(v)$:

$$h(v(t+1)) = \alpha h(v(t)) + h(v(t))^2 = h(\alpha v(t)) \Rightarrow v(t+1) = \alpha v(t) \tag{263}$$

Consider

$$x' = F(x) \tag{264}$$

over a Banach space, where F is a C^1 vector field defined in a neighborhood of the origin 0 and $F(0) = 0$. As before,

$$DF(0) = B$$

Remember that, by Lemma 17, the flow f^t associated with F satisfies

$$D_x f^t|_0 = e^{Bt} \tag{265}$$

Note that the flow f^t is hyperbolic *iff* $\sigma(B) \cap i\mathbb{R} = \emptyset$. In this case we naturally say that F is hyperbolic.

Theorem 7 (Hartman-Grobman for flows, [10])**.** *Suppose that* 0 *is a hyperbolic fixed point of the flow described by* F *in* (264)*. Then there is a homeomorphism between the flows of* F *and* $DF(0)$*, that is a homeomorphism between a neighborhood of zero into itself so that*

$$f^t = h \circ e^{Bt} \circ h^{-1}; \quad B = DF(0) \tag{266}$$

In fact, Theorem 7 follows from Theorem 6. Indeed, then we take first $f(x) := f^1(x)$, the flow at time 1 which by the given assumptions has a hyperbolic fixed point at zero. Then, recalling (243), there is an h s.t.

$$f^1(h(x)) = h(e^B x) \tag{267}$$

We claim that this conjugation extends for all t, that is

$$\boxed{f^t(h(x)) = h(e^{tB} x)} \tag{268}$$

as long as $e^{Bt}x$ is in the domain of h.

For the proof note that, assuming (268) we would have

$$h(x) = f^t(h(e^{-tB} x)) \tag{269}$$

and this is what we will check. Fix $t = T$ and let

$$\hat{h}(x) = f^T[h(e^{-BT} x)]$$

and the goal is to show $\hat{h} = h$. Using (267) we get

$$\begin{aligned} f^1[\hat{h}(x)] &= f^1[f^T[h(e^{-BT}x)]] = f^T[f^1[h(e^{-BT}x)]] \\ &= f^T[h(e^B e^{-BT}x)] = f^T[h(e^{-BT}e^B x)] = \hat{h}(e^B x) \end{aligned} \tag{270}$$

That is (267) holds with $\hat{h}$ instead of h. By uniqueness $\hat{h}(x) = h(x)$ and thus (268) follows. See also [8].

Smoother linearizations The more regularity is needed, the more conditions are required. Let us now consider a two by two system,

$$\dot{u} = \lambda_1 u + A_0 v^2 + A_1 uv + A_2 u^2 + \cdots \tag{271}$$

$$\dot{v} = \lambda_2 u + B_0 u^2 + B_1 uv + B_2 v^2 + \cdots \tag{272}$$

We assumed without loss of generality that the linear part is diagonal (more generally, we should take a Jordan normal form), since this can be arranged by linear changes of variables.

We try, by a change of variables, to eliminate the quadratic correction (at the expense of course of introducing higher order terms).

$$u = U + a_0 V^2 + a_1 VU + a_2 U^2 \tag{273}$$

$$v = V + b_0 U^2 + b_1 VU + b_2 V^2 \tag{274}$$

Substituting (273) in (271) we get

$$\dot{U} = \lambda_1 U + (-\lambda_1 a_2 + A_2)U^2 + (-a_1\lambda_2 + A_1)UV + (\lambda_1 a_0 - 2\,a_0\lambda_2 + A_0)V^2 + \cdots \tag{275}$$

$$\dot{V} = \lambda_2 V + (\lambda_2 b_0 + B_0 - 2\,\lambda_1 b_0)U^2 + (-b_1\lambda_1 + B_1)UV + (-\lambda_2 b_2 + B_2)V^2 + \cdots \tag{276}$$

and require that the quadratic monomials in U, V vanish; we get a system of equations which we solve for a_i, b_i. The result is

$$a_0 = -\frac{A_0}{\lambda_1 - 2\lambda_2}, a_1 = \frac{A_1}{\lambda_2}, a_2 = \frac{A_2}{\lambda_1}, b_0 = \frac{B_0}{-\lambda_2 + 2\lambda_1}, b_1 = \frac{B_1}{\lambda_1}, b_2 = \frac{B_2}{\lambda_2} \tag{277}$$

we see that for more regularity, we need *nonresonance conditions*: so far, we need

$$\lambda_i \neq 0; \quad \lambda_1 \neq 2\lambda_2; \quad \lambda_2 \neq 2\lambda_1$$

Let $\mu_i = e^{\lambda_i}$ and $\boldsymbol{\mu}^{\mathbf{k}} = \mu_1^{k_1} \cdot \mu_2^{k_2} \cdots \mu_n^{k_n}$ Consider also the following example due to Euler [14]:

Theorem 8 (Sternberg-Siegel, see [10])**.** *Assume F is differentiable, with a hyperbolic fixed point at zero, and DF is Hölder continuous near zero. Assume further that $A = DF(0)$ satisfies*

$$\lambda_{i,j,k} \in \sigma(A) \Rightarrow \mathrm{Re}\lambda_i \neq \mathrm{Re}\lambda_j + \mathrm{Re}\lambda_k \tag{278}$$

when $\mathrm{Re}\lambda_j < 0 < \mathrm{Re}\lambda_k$ (for maps $\mu_i \neq \mu_j\mu_k$ if $|\mu_j| < 1 < |\mu_k|$). Then the functions h in Theorems 6 and 7 can be taken to be diffeomorphisms.

Note 15. *In one or two dimensions, there is obviously no further restriction besides* $|\lambda_i| \neq 1$.

Smooth linearizations

Theorem 9 (Sternberg-Siegel, see [10])**.** *Assume* $F \in C^\infty$ *and the eigenvalues of* $DF(0)$ *are nonresonant, that is*

$$\lambda_i - \mathbf{k}\boldsymbol{\lambda} \neq 0 \tag{279}$$

for any $\mathbf{k} \in (\mathbb{N} \cup \{0\})^n$ *with* $|\mathbf{k}| > 1$. *(For maps,* $\mu_i \neq \boldsymbol{\mu}^{\mathbf{k}}$.*) Then the functions* h *in Theorems 6 and 7 can be taken to be* C^∞ *diffeomorphisms.*

We will later prove the Hartman-Grobman theorem for maps.

15.3.1 Linearization proofs

Consider a nonlinear system in a neighborhood of 0 taken to be a fixed point:

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}(\mathbf{x}),\ \mathbf{F}(\mathbf{x}) = O(\mathbf{x}^2) \tag{280}$$

Assume first that A is diagonalizable, and let $S^{-1}AS = \Lambda$; with $\mathbf{x} = S\mathbf{y}$ we get

$$S\dot{\mathbf{y}} = AS\mathbf{y} + \mathbf{F}(S\mathbf{y}) \Rightarrow \dot{\mathbf{y}} = \Lambda\mathbf{y} + \mathbf{F}(S\mathbf{y});\ \mathbf{F}(S\mathbf{y}) = O(\mathbf{x}^2) \tag{281}$$

(similarly for a more general Jordan form). Thus, without loss of generality we can assume $A = \Lambda$. Then

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}(\mathbf{x});\quad \text{where } \mathbf{F} \text{ is at least quadratic} \tag{282}$$

where $\mathbf{x}$ is a vector in $\mathbb{R}^d$, and we assume that we have normalized the equation s.t. Λ is diagonal (or more generally a Jordan matrix; we will not yet deal with this extra layer of complication.) We want to find an h which conjugates (282) with the linear equation

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} \tag{283}$$

We also $h(0) = 0$ to preserve the fixed point, and that h be locally invertible. That means $Dh(0)$ is invertible. As we'll see first, without loss of generality we can assume that $B = Dh(0) = I$. We have

$$\dot{\mathbf{x}} = \frac{\partial\mathbf{h}}{\partial\mathbf{w}}\dot{\mathbf{w}} + \frac{\partial\mathbf{h}}{\partial\mathbf{w}}\Lambda\mathbf{w} \tag{284}$$

(because we want (283) to hold), and on the other hand,

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}(\mathbf{x}) = \Lambda\mathbf{h}(\mathbf{w}) + \mathbf{F}(\mathbf{w} + \mathbf{h}(\mathbf{w})) \tag{285}$$

and finally, we get the nonlinear PDE

$$\frac{\partial\mathbf{h}}{\partial\mathbf{w}}\Lambda\mathbf{w} = \Lambda\mathbf{h}(\mathbf{w}) + \mathbf{F}(\mathbf{w} + \mathbf{h}(\mathbf{w})) \tag{286}$$

Or, equivalently

$$L\mathbf{h} = \mathbf{F}(\mathbf{w} + \mathbf{h}) \tag{287}$$

where

$$L\mathbf{h} := \frac{\partial \mathbf{h}}{\partial \mathbf{w}}\Lambda - \Lambda\mathbf{h} \tag{288}$$

An equation of the form

$$\frac{\partial \mathbf{h}}{\partial \mathbf{w}}\Lambda - \Lambda\mathbf{h} = \mathbf{v} \tag{289}$$

is called a *homological equation*. On the other hand $\mathbf{h}\mathbf{w} = B\mathbf{w} + \mathbf{H}_2(\mathbf{w})$ where $\mathbf{H}_2$ is at least quadratic.

We then have

$$(B + \mathbf{H}_2(\mathbf{w}))\Lambda\mathbf{w} = \Lambda(B + \mathbf{H}_2) + \mathbf{F}(\mathbf{w} + \mathbf{h}(\mathbf{w}) \tag{290}$$

in the limit $\mathbf{w} \to 0$ we have

$$B\Lambda = \Lambda B + o(1) \tag{291}$$

which means B (which by choice is invertible) is in the commutative C^*-algebra generated by Λ. Note also that by taking $\mathbf{w} = B^{-1}\mathbf{w}_1$ we get as a linear part in the $\mathbf{w}$ variable, $\mathbf{h}\mathbf{w}_1 = BB^{-1}\mathbf{w} + \mathbf{H}_2(B^{-1}\mathbf{w}_1)$ and the new B is simply

$$\tilde{B} = I \tag{292}$$

Now we Taylor expand $\mathbf{h}$ for small $\mathbf{w}$ and organize as is often done the terms be homogeneous polynomials $\mathbf{h}_n$ of degree n (this means that any monomial in $\mathbf{h}_n$ is of the form $\mathbf{w}^{\mathbf{k}}$ where $|\mathbf{k}| = k_1 + ... + k_d = n$.):

$$\mathbf{h}(\mathbf{w}) = \mathbf{w} + \sum_{n=2}^{\infty} \sum_{|\mathbf{k}|=n} \mathbf{h}_{\mathbf{k}} \mathbf{w}^{\mathbf{k}} \tag{293}$$

Decomposing the PDE by homogeneous monomials, it splits into an infinite set of ODEs.

We see that if $M_n e_j$ where M_n is a monomial of total degree n, i.e. $\prod w_k^{n_k}$ with $\sum_{n_k} = n$ and e_j is the unit vector along the direction j of degree n then $\frac{\partial \mathbf{h}}{\partial w}\Lambda w = const M_n e_j$. It is simple to generalize this if we see it in a 2 by 2 case, with $e_j = e_1$, $\mathbf{h} = M_n = w_1^{k_1} w_2^{k_2} e_1$, $k_1 + k_2 = n$:

$$\frac{\partial M_n}{\partial w}\Lambda w = M \begin{pmatrix} n_1 \dfrac{1}{w_1} & n_2 \dfrac{1}{w_1} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 w_1 & 0 \\ 0 & \lambda_2 w_2 \end{pmatrix} = M \begin{pmatrix} \lambda_1 n_1 + \lambda_2 n_2 \\ 0 \end{pmatrix} \tag{294}$$

Likewise also,

$$\Lambda M e_1 = \lambda_1 M e_1 \tag{295}$$

15.3.2 Homogeneous polynomial decomposition

We write

$$\mathbf{h}(\mathbf{w}) = \mathbf{w} + \sum_{n\geqslant 2}\sum_{|\mathbf{k}|=n} \mathbf{h}_{\mathbf{k}}\mathbf{w}^{\mathbf{k}} = \sum_{n\geqslant 2} \mathbf{h}^{[n]}(\mathbf{w}) \tag{296}$$

where $\mathbf{h}^{[n]}$ are homogeneous polynomials of degree n. Similarly, in the equation for x we write

$$\mathbf{F}(\mathbf{w}) = \sum_{n\geqslant 2}\sum_{|\mathbf{k}|=n} \mathbf{F}_{\mathbf{k}}\mathbf{w}^{\mathbf{k}} = \sum_{n\geqslant 2} \mathbf{F}^{[n]}(\mathbf{w}) \tag{297}$$

with $\mathbf{F}^{[n]}(\mathbf{w})$ homogeneous polynomials of degree n.

15.3.3 Preservation of monomials

Let L be as defined in (288). Note that the monomials $M_{\mathbf{k}}e_j = \mathbf{w}^{\mathbf{k}}e_j$ where e_j is the j-th unit vector form a basis in the space of all vector polynomials of degree $\leqslant n$.

Lemma 19.

$$L\mathbf{w}^{\mathbf{k}}e_j = (\mathbf{k}\cdot\boldsymbol{\lambda} - \lambda_j)\,\mathbf{w}^{\mathbf{k}}e_j \tag{298}$$

Proof. We note that $\frac{\partial \mathbf{w}^{\mathbf{k}}e_j}{\partial w}$ is a matrix with only one nonzero row, the jth, and $\Lambda\mathbf{w}^{\mathbf{k}}e_j = \mathbf{w}^{\mathbf{k}}\lambda_j e_j$, and thus the jth component of $\frac{\partial \mathbf{w}^{\mathbf{k}}e_j}{\partial w}\Lambda\mathbf{w}^{\mathbf{k}}e_j$ is nonzero, and it equals $\mathbf{w}^{\mathbf{k}}\sum_{i=1}^{d}\lambda_i k_i$ and the result follows. □

Definition 20 (Nonresonance)**.** *the eigenvalues $\lambda_1, ..., \lambda_d$ are nonoresonant if $\lambda_j \neq (\mathbf{k}\cdot\boldsymbol{\lambda})$ for $j = 1, ..., d$ and $\mathbf{k}$ with $|\mathbf{k}| \geqslant 2$.*

We assume from now on that $\lambda_1, ..., \lambda_d$ are nonoresonant.

Lemma 21. *The equation*

$$L\mathbf{P}^{[n]} = \mathbf{F}^{[n]};\ \mathbf{F}^{[n]}(\mathbf{x}) = \sum_{j=1}^{d}\sum_{\mathbf{k},|\mathbf{k}|=n} f_{j,\mathbf{k}}\mathbf{x}^{\mathbf{k}}e_j$$

where $\mathbf{P}^{[n]}$ is a homogeneous polynomial, has the solution

$$\mathbf{P}^{[n]}(\mathbf{x}) = \sum_{j=1}^{d}\sum_{\mathbf{k},|\mathbf{k}|=n} \frac{f_{j,\mathbf{k}}}{\mathbf{k}\cdot\boldsymbol{\lambda} - \lambda_j}\mathbf{x}^{\mathbf{k}}e_j \tag{299}$$

Proof. This follows from (298). □

Now we start by eliminating the quadratic terms in $\mathbf{F}$.

We substitute $\mathbf{x} = \mathbf{w} + \mathbf{h}^{[2]}(\mathbf{w})$ in (284), where $\mathbf{h}^{[2]}(\mathbf{w})$, is a homogeneous polynomial of degree 2, $\mathbf{F}_2(\mathbf{x}) = \mathbf{F}^{[2]}(\mathbf{x}) + o(\mathbf{x}^2)$. We get

$$\begin{aligned}\dot{\mathbf{x}} &= \Lambda\mathbf{x} + \mathbf{F}_2(\mathbf{x}) = \Lambda\mathbf{x} + \mathbf{F}^{[2]}(\mathbf{x}) + o(\mathbf{x}^2)\\ &= \Lambda\mathbf{w}+\Lambda\mathbf{h}^{[2]}(\mathbf{w})+\mathbf{F}^{[2]}(\mathbf{w}+\Lambda\mathbf{h}^{[2]}(\mathbf{w}))+o(\mathbf{w}^2) = \Lambda\mathbf{w}+\Lambda\mathbf{h}^{[2]}(\mathbf{w})+\mathbf{F}^{[2]}(\mathbf{w})+o(\mathbf{w}^2)\end{aligned} \tag{300}$$

and on the other hand, combining the substitution with (300) we get

$$\dot{\mathbf{x}} = \dot{\mathbf{w}}+\frac{\partial\mathbf{h}^{[2]}}{\partial\mathbf{w}}\dot{\mathbf{w}} = \Lambda\mathbf{w}+\frac{\partial\mathbf{h}^{[2]}}{\partial\mathbf{w}}\Lambda\mathbf{w}+o(\mathbf{w}^2) = \Lambda\mathbf{w}+\Lambda\mathbf{h}^{[2]}(\mathbf{w}))+\mathbf{F}^{[2]}(\mathbf{w})+o(\mathbf{w}^2) \tag{301}$$

and thus, to $o(\mathbf{w}^2)$,

$$\frac{\partial\mathbf{h}^{[2]}}{\partial\mathbf{w}}\Lambda - \Lambda\mathbf{h}^{[2]} = L\mathbf{h}^{[2]} = \mathbf{F}^{[2]}(\mathbf{w}) \tag{302}$$

which, under the nonresonance condition has a solution, c.f. (299).

Theorem 10 (Poincaré)**.** *If the eigenvalues are nonresonant, then there is a formal series transformation linearizing the system.*

Proof. We let $\mathbf{x} = \mathbf{w}_1$, $\mathbf{w} = \mathbf{w}_2$. After eliminating the quadratic terms the equation becomes

$$\dot{\mathbf{w}}_2 = \Lambda\mathbf{w}_2 + \mathbf{F}^{[3]} + o(\mathbf{w}^3) \tag{303}$$

and the transformation $\mathbf{w}_2 = \mathbf{w} + \mathbf{h}^{[3]}(\mathbf{w})$ eliminates the cubic terms if

$$\frac{\partial\mathbf{h}^{[3]}}{\partial\mathbf{w}}\Lambda - \Lambda\mathbf{h}^{[3]} = L\mathbf{h}^{[3]} = \mathbf{F}^{[3]}(\mathbf{w}) \tag{304}$$

which has a unique solution by Lemma 21. Inductively we get at step n the equation

$$\dot{\mathbf{w}}_n = \Lambda\mathbf{w}_n + \mathbf{F}^{[n]}(\mathbf{w}_n) + o(\mathbf{w}_n^{n+1}) \tag{305}$$

in which the monomials of order $\leqslant n$ are eliminated by a transformation $\mathbf{w}_n = \mathbf{w} + \mathbf{h}^{[n]}(\mathbf{w})$.

The transformation leading from $\mathbf{w}_1$ to $\mathbf{w}_n$ is

$$\mathbf{w}_1 = \mathbf{h}(\mathbf{w}_2) = \cdots = \mathbf{h}_2 \circ \mathbf{h}_3 \cdots \mathbf{h}_n(\mathbf{w}_n) \tag{306}$$

Note that at a formal level the expansion converges in the sense that the coefficients of the monomial of degree $\leqslant n$ do not change by composition with $\mathbf{w}_{n+1}$ or higher, since

$$\mathbf{w}^{\mathbf{k}} \circ (\mathbf{w} + \mathbf{w}^{\mathbf{n}}) = (\mathbf{w} + \mathbf{w}^{\mathbf{n}})^{\mathbf{k}} = \mathbf{w}^{\mathbf{n}} + O(\mathbf{w}^{\mathbf{k}+\mathbf{n}})$$

□

Theorem 11 (Poincaré-Dulac). *An equation of the form*

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}_2(\mathbf{x})$$

is formally equivalent to

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} + \mathbf{R}_2(\mathbf{x})$$

where $\mathbf{R}_2(\mathbf{x})$ *is a series (possibly finite) containing the resonant monomials in* $\mathbf{F}_2$ *and only those.*

Proof. We simply follow the procedure in the proof of the Poincaré theorem, while keeping the monomials that cannot be eliminated because of resonannce. We leave the details as an exercise. □

15.3.4 Improved convergence rate

Also, in the proof of convergence, a much more rapid scheme can be used, namely, instead of writing $\mathbf{x} = \mathbf{w} + \mathbf{h}^{[n]}(\mathbf{w})$ (where as the notation indicates $\mathbf{h}^{[n]}$ is a homogeneous polynomial of degree n, one takes a general $\mathbf{h}_n$ of degree at least n and solve for all the monomials present in the nonlinearity up to a higher order nonlinearity, namely by solving

$$\frac{\partial \mathbf{h}^{[n]}}{\partial \mathbf{w}}\Lambda - \Lambda\mathbf{h}^{[n]} = L\mathbf{h}^{[n]} = \mathbf{F}_n(\mathbf{w}) \tag{307}$$

where $\mathbf{F}_n$ contains all monomials of degree $\in [n, 2n-1]$. Then, the error terms generated come from taking $\mathbf{F}_n(\mathbf{w})$ instead of $\mathbf{F}(\mathbf{w}+\mathbf{h}_n)$ which only introduces terms of order at least $2n-1$.

By following this procedure we have at step j a least power satisfying $p_j = 2p_{j-1} - 1$, that is $p_n \geqslant C2^n$

Note 16. Alternatively, if we do not aim for computational simplicity but rather for the simplicity of the result, we can solve (307) to all orders, to avoid introducing additional error terms at the next stage.

Note 17. In §15.3.4 at stage 1 we replace $\mathbf{F}(\mathbf{w}+\mathbf{h})$ by $\mathbf{F}(\mathbf{w})$, introducing an error term $\mathbf{F}(\mathbf{w}+\mathbf{h}) - \mathbf{F}(\mathbf{w}) = \mathbf{R}(\mathbf{w})$; at the next stage we should ideally solve $\mathbf{h}_2 = \mathbf{R}(\mathbf{w}+\mathbf{h_2})$ which is nonlinear, and we once more replace this with $\mathbf{h}_2 = \mathbf{R}(\mathbf{w})$, which we solve exactly, at the expense of introducing an error at the next order $\mathbf{R}(\mathbf{w}+\mathbf{h_2}) - \mathbf{R}(\mathbf{w}) = \mathbf{F}(\mathbf{w}+\mathbf{h}+\mathbf{h_2}) - \mathbf{F}(\mathbf{w}+\mathbf{h_2}) - \mathbf{F}(\mathbf{w}+\mathbf{h}) + \mathbf{F}$. If the coefficients of F satisfy $|f_{\mathbf{k},j}| < MR^{-|\mathbf{k}|}$, we choose $A > 1/R$ large enough to accommodate for the slight increase in the argument of $F : w + h + h_2$, etc. We can crudely estimate the coefficients of the error term after n successive substitutions by $2^n MA^{|\mathbf{k}|}$

*

15.4 Types of resonances

The resonances are classified according to their degree. This relates to the number of "partial linearizations" $\mathbf{h}_n$ can be performed. Clearly, if some λ is zero, this corresponds to $\lambda_j = 2\lambda_j$; it is also a resonance at higher orders, $\lambda_j = n\lambda_j$, for any $j \geqslant 2$. If λ_1, λ_2 are s.t.

$$n_1\lambda_1 = -n_2\lambda_2; \quad n_1, n_2 \in \mathbb{N}$$

then this constitutes a resonance since it implies

$$\lambda_2 = n_1\lambda_1 + (n_2 + 1)\lambda_2$$

Taking the opposite sign,

$$\lambda_1 = m\lambda_2$$

with $m \geqslant 2$ is a resonance since it is the same as

$$\lambda_1 = 0\lambda_1 + m\lambda_2$$

but it entails no higher order resonance since

$$\lambda_1 = n_1\lambda_1 + n_2\lambda_2 \Rightarrow m(1 - n_1)\lambda_1 = n_2\lambda_2 \Rightarrow (n_1 = 0 \text{ and } n_2 = m)$$

Furthermore, if p_1 and p_2 are relatively prime, then

$$p_1\lambda_1 = p_2\lambda_2$$

is not a resonance since, together with $\lambda_1 = n_1\lambda_1 + n_2\lambda_2$ it implies

$$p_1\lambda_1 = p_1n_1\lambda_1 + p_1n_2\lambda_2 \Rightarrow p_2\lambda_2 - p_2n_1\lambda_2 = p_1n_2\lambda_2 \tag{308}$$

and thus $n_1 = 0$ implying $p_2 = p_1n_2 \Rightarrow n_1 = 1 \Rightarrow p_1 = p_2$, contradiction.

15.4.1 Poincaré domains and Siegel domains

Definition 22. The eigenvalues $\lambda_1, ..., \lambda_n$ belong to the *Poincaré domain* (a subset of $\mathbb{C}^n$) if the convex hull of $\lambda_1, ..., \lambda_n$ in $\mathbb{C}$ does not contain zero inside. Otherwise it is said that they belong to the *Siegel domain*.

Theorem 12 (The convex separation theorem)**.** *A convex domain not containing the origin is contained in some half-plane.*

More generally, if in a vector space S is a convex set not containing a point y, then S and the point are separated by a hyperplane. A further generalization is the Hahn-Banach separation theorem.

Proof. Note that if three points in the set are not contained in a sector of opening at most π, then they form a triangle containing the origin. Take the line joining λ_1 and λ_2. If 0 and λ_3 are on opposite sides of the line, then the triangle is in the half plane containing that line and λ_3. If they are on the same side, then the origin is inside the triangle. Exercise: write a rigorous proof. $\square$

Theorem 13. *If the eigenvalues are in a Poincaré domain, then there are only finitely many (possibly zero) resonances.*

Proof. Indeed, without loss of generality, we can assume that $\lambda_1, ..., \lambda_n$ are all in the open right half plane. Then $\mathrm{Re}\lambda_i > 0$ for all i. This easily implies

$$\lim_{|\mathbf{k}|\to\infty} \mathrm{Re}\,(\mathbf{k}\cdot\boldsymbol{\lambda}) = \infty \tag{309}$$

$\square$

15.4.2 The case when the Jordan form of A is not diagonal

Note: the result is stated incorrectly in Arnold [14]. It is however stated accrurately and proved in [9].

Lemma 23. *For general $DF(0)$, there is a basis in which L is triangular, with the same eigenvalues as in the diagonal case.*

Theorem 14. *If $\boldsymbol{\lambda}$ is in the Siegel domain then either*
(i) There are infinitely many resonances
(ii) There exist sequences s.t.

$$\lambda_j - \mathbf{k}\cdot\boldsymbol{\lambda} \;\to 0 \textit{ with } k \in \mathbb{N} \textit{ as } |\mathbf{k}| \to \infty \tag{310}$$

Proof. As in the convex separation proof, there exists a triple $\lambda_{1,2,3}$ forming a triangle containing 0. Thus, for some $\alpha_{1,2,3} \in \mathbb{R}^+$ with $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and

$$\alpha_1\lambda_1 + \alpha_2\lambda_2 + \alpha_3\lambda_3 = 0 \tag{311}$$

Evidently, both conditions (i) and (ii) are invariant under linear changes of coordinates. We have in mind linear invertible transformations from $\mathbb{R}^2$ into $\mathbb{R}^2$, not only conformal ones.

There are two cases: (i) there are two linearly independent λs (over $\mathbb{R}$); (ii) the three λs are collinear.

If they are collinear, since they do not lie in a common half plane, then 0 is between, say λ_1 and λ_2. By a rotation we make both λ's real; one is positive and the other one is negative. If λ_1 and λ_2 are linearly dependent over $\mathbb{Q}$ then clearly there are infinitely many resonances. If not, it is known that the set of $k_1|\lambda_1| + k_2|\lambda_2|$ with $(k_1, k_2) \in \mathbb{Z}$ is dense in $\mathbb{R}$ from which the property follows trivially.

Case (ii). Through a linear invertible matrix we can make $\lambda_1 = 1$ and $\lambda_2 = i$; we let $\lambda_3 = \lambda$. By assumption, 0 is inside the triangle $\Delta(1, i, \lambda)$. There is a convex combination satisfying (311), or

$$-\,\alpha_3\lambda_3 = \alpha_1 + \alpha_2 i \tag{312}$$

or

$$-\,\lambda_3 = \beta_1 + \beta_2 i \tag{313}$$

where β_1, β_2 are positive, or, $\lambda = -\lambda_3$ is in the first quadrant. By another rescaling we make $\beta_1 = 1, \beta_2 = 1$. Then, considering all linear combinations of $-n_3\lambda_3$, n_1 and $n_2 i$ is the same as looking at $-n_3\lambda_3$ for all $n_3 \in \mathbb{N}$ on a square lattice in the first quadrant, and determining whether $-n_3\lambda_3$ goes through, or passes near, a node. This in turn is equivalent to the question of the evolution $X \mapsto X - \lambda_3$ mod the unit square, a discrete rotation on the torus $\mathbb{T}$. Thus, either the rotation is rational in which case there are infinitely many resonances, or else the trajectory is dense. This follows from the Poincaré recurrence theorem. □

15.5 Analytic equivalence

15.5.1 The Poincaré-Dulac theorem

Theorem 15 (Poincaré)**.** *Assume the eigenvalues of Λ are in the Poincaré domain and are nonresonant and in the system*

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}(\mathbf{x}); \quad \mathbf{F} = O(\mathbf{x}^2) \tag{314}$$

the function $\mathbf{F}$ is analytic in a polydisk of radius R containing the origin. Then the system (314) *is analytically equivalent to*

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} \tag{315}$$

for small enough $\mathbf{w}$.

Theorem 16 (Poincaré-Dulac)**.** *Under the same assumptions as in Theorem 15 except nonresonance, the system* (314) *is analytically equivalent to*

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} + \mathbf{R}(\mathbf{w}) \tag{316}$$

where $\mathbf{R}$ contains only resonant monomials.

Note 18. *The proof of Theorem 16 is very similar to that of Theorem 15 so we focus on Theorem 15. The C^∞ and C^r cases are similar and in fact simpler.*

15.6 Proof of Theorem 15

Proof. We look for an equivalence map

$$\mathbf{x} = \mathbf{w} + \mathbf{h}(\mathbf{w}); \quad \mathbf{h} = O(\mathbf{w}^2) \tag{317}$$

where $\mathbf{h}$ is analytic, reducing (314) to (315). The equation that $\mathbf{h}$ satisfies is then, see (286),

$$\frac{\partial \mathbf{h}}{\partial \mathbf{w}}\Lambda\mathbf{w} - \Lambda\mathbf{h} = \mathbf{F}(\mathbf{w} + \mathbf{h}(\mathbf{w})) \tag{318}$$

15.6.1 Spaces of vector valued analytic functions

For the proof of the theorem, we consider the space of vector valued functions $\mathbf{f}$,

$$\mathbf{f}(\mathbf{w}) = \sum_{j=0}^{d} \sum_{\mathbf{k}\geqslant 0} f_{\mathbf{k},j} \mathbf{w}^j e_j \tag{319}$$

analytic in a *polydisk* $\mathbb{D}_R = \{\mathbf{w} : |w_i| < R, i = 1, ..., d\}$ and continuous on $\overline{\mathbb{D}}_R$ with the norm

$$|||\mathbf{f}|||_R = |||\mathbf{f}||| = \sum_{j,\mathbf{k}} |f_{\mathbf{k},j}| \; R^{|\mathbf{k}|} \tag{320}$$

We denote by $C_a^\omega(\mathbb{D}_R)$ the space of functions with finite norm (320).

Note 19. Associate to a function $\mathbf{f}$ analytic in the polydisk $\mathbb{D}_R$ the function $\mathbf{f}_{\text{abs}}$ by

$$\mathbf{f} = \sum_{\mathbf{k}} f_{\mathbf{k}} \mathbf{w}^{\mathbf{k}} \Rightarrow \mathbf{f}_{\text{abs}} = f_{\text{abs}} = \sum_{\mathbf{k},j} |f_{\mathbf{k},j}| \mathbf{w}^{\mathbf{k}} \tag{321}$$

($\mathbf{f}_{\text{abs}}$ is in fact a scalar). The operator $\mathbf{f} \to \mathbf{f}_{\text{abs}}$ is unbounded in the usual sup norm.

Proposition 24. *(i) If $\mathbf{f}_{abs}$ is analytic in $\mathbb{D}_R$ and $\sup_{\mathbb{D}_R} |\mathbf{f}_{abs}| < \infty$, then $\mathbf{f}$ is analytic in $\mathbb{D}_R$ and continuous in $\overline{\mathbb{D}}_R$.*

(ii) If $\mathbf{f}$ is analytic $\mathbb{D}_R$ and $R' < R$, then the sup norms of $|\mathbf{f}_{abs}^{(n)}|$ over $\mathbb{D}_{R'}$ are finite for all n. (Note that $\mathbf{f}_{abs}^{(n)} = (\mathbf{f}^{(n)})_{abs}$.)

Proof. (i) is obvious.

(ii) Inductively on d we can see that $\sum_{|\mathbf{k}|=n} 1 < dn^d$. If $\mathbf{f}$ is analytic in a polydisk of radius R and continuous up to the boundary with sup norm M, then by Cauchy's formula we have

$$|f_{\mathbf{k},j}| \leqslant \frac{M}{(2\pi)^d R^{|\mathbf{k}|}} \tag{322}$$

then for any $R' < R$ we have

$$\begin{aligned} \sum_{\mathbf{k},j} |f_{\mathbf{k},j}|(R')^k &\leqslant \frac{Md}{(2\pi)^d} \sum_{n=0}^{\infty} \left(\frac{R'}{R}\right)^n \sum_{|\mathbf{k}|=n} 1 \\ &\leqslant \frac{d^2 M}{(2\pi)^d} \sum_{n=0}^{\infty} \left(\frac{R'}{R}\right)^n n^d < MC\left(d, \frac{R'}{R}\right) < \infty \end{aligned} \tag{323}$$

The proof is the same for any derivative. □

Proposition 25. *$C_a^\omega(\mathbb{D}_R)$ is a Banach space of analytic functions in $\mathbb{D}_R$ continuous in $\overline{\mathbb{D}}_R$. If $\mathbf{f} \in C_a^\omega(\mathbb{D}_R)$, then $|||\mathbf{f}||| = \sup_{\mathbf{w}\in\mathbb{D}_R} |\mathbf{f}_{abs}|$.*

Proof. Analyticity is clear. Continuity follows from the uniform convergence of the power series in $\mathbb{D}_R$. The fact that this is a Banach space can be seen either directly or as follows: the space $C_a^\omega(\mathbb{D}_R)$ is isomorphic to a weighted ℓ^1 space (with weight $R^{|\mathbf{k}|}$). The last part is clear since $f_{\rm abs}$ is an analytic function with positive coefficients whose sup, if finite, is reached at $w_i = R, i = 1, ..., d$. $\square$

Proposition 26. *In the scalar case, $d = 1$, the space $C_a^\omega(\mathbb{D}_R)$ is a Banach algebra.*

Note 20 (Reminder)**.** *A Banach algebra is a normed algebra in which the product "$*$" is continuous: $\|f * g\| \leqslant \|f\|\|g\|$.*

Proof. It is straightforward to check that

$$|||fg||| = \sup|(fg)_{\rm abs}| \leqslant \sup(|f_{\rm abs}|\ |g_{\rm abs}|) \leqslant \sup|f_{\rm abs}|\ \sup|g_{\rm abs}| = |||f|||\ |||g|||$$

$\square$

Definition 27. *If $\mathbf{H} = (H_1, H_2, ..., H_d)$, we naturally write*

$$\mathbf{H}^{\mathbf{k}}(\mathbf{w}) = H_1^{k_1}(\mathbf{w})H_2^{k_2}(\mathbf{w})\cdots H_d^{k_d}(\mathbf{w}) \tag{324}$$

Corollary 28.

$$|||\mathbf{H}^{\mathbf{k}}||| \leqslant |||H_1|||^{k_1}\cdots|||H_d|||^{k_d}$$

Proposition 29. *If $\mathbf{f}, \mathbf{g}$ are in $C_a^\omega(\mathbb{D}_R)$ and if $\mathbf{H}(0) = 0$ then for R' small enough $\mathbf{f}(\mathbf{H}) \in C_a^\omega(\mathbb{D}_{R'})$ and we have*

$$\mathbf{f}(\mathbf{H}(\mathbf{w})) = \sum_{\mathbf{k}} f_{\mathbf{k},j}\mathbf{H}^{\mathbf{k}}e_j \tag{325}$$

and

$$|||\mathbf{f}(\mathbf{H})||| \leqslant \sup_{|\mathbf{w}|<R'} |\mathbf{f}_{abs}(\mathbf{H}_{abs})| \tag{326}$$

Proof. This follows immediately from Corollary 28. $\square$

15.7 The analytic equation

Recall that the eigenvalues of Λ are assumed nonresonant and that this implies $|\boldsymbol{\lambda}\mathbf{k} - \lambda_j| \to \infty$ as $|\mathbf{k}| \to \infty$. Then there is a lower bound

$$|\boldsymbol{\lambda}\cdot\mathbf{k} - \lambda_j| > (1/a) > 0 \Rightarrow \frac{1}{|\boldsymbol{\lambda}\mathbf{k} - \lambda_j|} < a;\ \forall\mathbf{k}, j \tag{327}$$

Proposition 30. *If $\boldsymbol{\lambda}$ is nonresonant and a is as in* (327)*, then the operator*

$$L := \mathbf{H} \mapsto \frac{\partial\mathbf{H}}{\partial\mathbf{w}}\Lambda\mathbf{w} - \Lambda\mathbf{H} \tag{328}$$

is invertible from C_a^ω into and C_a^ω and the norm of the inverse is $|||L^{-1}||| \leqslant a$.

Proof. The inverse of L (say densely defined on a dense set, on polynomials) is, see Lemma 19,

$$L^{-1}\mathbf{f} = \sum_{j,|\mathbf{k}|\geqslant 2} \frac{f_{\mathbf{k},j}}{\boldsymbol{\lambda}\cdot\mathbf{k}-\lambda_j}\mathbf{w}^{\mathbf{k}}e_j \tag{329}$$

and by (327) the norm is less than a. L^{-1} extends thus by continuity to the whole of C_a^ω with the same norm (check the details: polynomials are dense in C_a^ω, the equation LL^{-1} holds on polynomials, etc.) □

We write (318) in the equivalent form

$$\mathbf{h}(\mathbf{w}) = L^{-1}\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w})) \tag{330}$$

Proposition 31. *For small enough ε there exists a unique solution* $\mathbf{h}$ *of* (330) *in $C_a^\omega(\mathbb{D}_\varepsilon)$.*

Proof 1. It is easily checked that $\mathbf{h} \mapsto= L^{-1}\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w}))$ is well defined in a small ball $|||\mathbf{h}||| < \varepsilon$. The Fréchet derivative

$$D_0\Big[\mathbf{h}(\mathbf{w})-L^{-1}\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w}))\Big] = I - D_0[L^{-1}\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w}))] = I - L^{-1}\mathbf{F}'(0)I = I$$

and the implicit function theorem applies. □

Proof 2. We take ε sufficiently small, in particular smaller than R'. We work in $B_{2\varepsilon}$, a ball of size 2ε in the space of functions $C_a^\omega(\mathbb{D}_\varepsilon)$. Since $|||\mathbf{F}||| = O(\mathbf{w}^2)$ and $|||L^{-1}||| \leqslant a$ we can easily check that $\mathbf{h} \in B_{2\varepsilon}$ implies that, for small ε, $L^{-1}\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w})) \in B_{2\varepsilon}$. (Here we used Proposition 29). By the assumptions on $\mathbf{F}$ and Proposition 24, we also have $\mathcal{L}^{-1}\mathbf{F}' \in C_a^\omega(\mathbb{D}_\varepsilon)$ and $|||\mathcal{L}^{-1}\mathbf{F}'||| = O(\mathbf{w})$. Now it is easy to check that (330) is contractive in $B_{2\varepsilon}$, with contractivity factor $O(\varepsilon)$. □

We now note that $\mathbf{F} = O(\mathbf{w}^2)$ and thus $\mathbf{h}(\mathbf{w}) = O(\mathbf{w}^2)$, and this means that $\mathbf{w}+\mathbf{h}(\mathbf{w})$ is a local analytic diffeomorphism near $\mathbf{w} = 0$. This finishes the proof of the theorem. □

15.8 The Poincaré domain resonant case and the extended system

In this case the Poincaré Dulac theorem shows that we are generically left with resonant monomials, cf. (316). Ideas going back to Dulac [3] and developed by Kazhdan, Kostant and Sternberg [5], Walcher [11] and Gaeta [4], show that the system can be extended so that it becomes linear. The idea is essentially to take each resonant monomial as a new dependent variable. We illustrate this on a number of examples, following relatively closely [4] (modulo notation and small typos in [4]).

Take $d = 2$, $\mathbf{w} = (x, y)$ and

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$$

with k a positive integer; here $\boldsymbol{\lambda} = (1, k)$ is in the Poincaré domain. There is only one resonance $\mathbf{k} = (k, 0)$ (with $j = 2, n = k$ (that is, $\lambda_2 = k\lambda_1 + 0\lambda_2$, $k_1 + k_2 = n = 2$ and the only resonant monomial is $\mathbf{w}^{\mathbf{k}} = x^k$. The Poincaré-Dulac normal form is

$$\begin{aligned} \dot{x} = & \quad x \\ \dot{y} = & \quad ky + \beta x^k \end{aligned}$$

with β an arbitrary coefficient. Let $q = x^k$, $\dot{q} = k(x^{k-1})\dot{x} = kq$. The extended system is

$$\begin{aligned} \dot{x} = & \quad x \\ \dot{y} = & \quad ky + \beta q \\ \dot{q} = & \quad kq \end{aligned}$$

The matrix of the system has a nontrivial Jordan block corresponding to the resonant k:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & \beta \\ 0 & 0 & k \end{pmatrix} \tag{331}$$

The solution is

$$x(t) = x_0 e^t \ , \ y(t) = (y_0 + y_1 t)e^{kt} \ , \ q(t) = q_0 e^{kt}; \ y_1 = \beta k q_0$$

and the submanifold $\mathcal{M}$ given by the constraint $q = x^k$ is invariant under this flow. Note also that replacing ky by $k'y$, for any $k' \neq k$ the matrix in (331) is diagonalizable, since in that case the only nondiagonal block has distinct eigenvalues. For (331) however, the normal Jordan form is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 1 \\ 0 & 0 & k \end{pmatrix} \tag{332}$$

Example 2.

Take $d = 3$, $\mathbf{w} = (x, y, z)$,

$$\Lambda \ = \ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

and $\boldsymbol{\lambda} = (1, 2, 5)$ in the Poincaré domain. There are four resonances: $[\mathbf{k}_1 = (2, 0, 0),\ j = 2, n = 2]$, $[\mathbf{k}_2 = (1, 2, 0), j = 3, n = 3]$, $[\mathbf{k}_3 = (3, 1, 0), j = 3, n = 4]$ and $[\mathbf{k}_4 = (5, 0, 0), j = 3, n = 5]$. and correspondingly we have, with $q_i = \mathbf{w}^{\mathbf{k}_i}$,

$$q_1 = x^2 \ , \ q_2 = xy^2 \ , \ q_3 = x^3 y \ , \ q_4 = x^5$$

The normal form is

$$\begin{aligned} \dot{x} = & \quad x \\ \dot{y} = & \quad 2y + c_1 x^2 \\ \dot{z} = & \quad 5z + c_2 xy^2 + c_3 x^3 y + c_4 x^5 \end{aligned}$$

with c_i arbitrary real coefficients.

$$\dot{q_1} = 2x\dot{x} = 2x^2 = 2q_1, \tag{333}$$

$$\dot{q_2} = \dot{x}y^2 + 2xy\dot{y} = xy^2 + 2xy(2y + c_1x^2) = 5q_2 + 2c_1q_3; \tag{334}$$

$$\dot{q_3} = 3x^2y\dot{x} + x^3\dot{y} = 3x^3y + x^3(2y + c_1x^2) = 5q_3 + c_1q_4 \tag{335}$$

$$\dot{q_4} = 5x^5 = 5q_4 \tag{336}$$

We take $q_1 = x^2, q_2 = xy^2, q_3 = x^3y, q_4 = x^5$ and the system extends to a linear system in $\mathbb{R}^7$,

$$\dot{\begin{pmatrix} x \\ y \\ z \\ q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 2c_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & c_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} \tag{337}$$

with Jordan normal form

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} = D + N \tag{338}$$

where N is a nilpotent, $N^4 = 0$. The general solution to the system (retaining only x, y, z) is

$$x(t) = x_0e^t, y(t) = (y_0 + ty_1)e^{2t}, z(t) = (z_0 + z_1t + z_2t^2 + z_3t^3)e^{5t} \tag{339}$$

(**Note.** There seem to be typos in the solution in [4].)

15.9 Connection with regular singular points and Frobenius theory

A system of the form

$$\dot{\boldsymbol{u}} = B(0)\boldsymbol{u} + zA(z)\boldsymbol{u}; \ \dot{z} = z; \ (\dot{f} := \frac{df}{dt}) \tag{340}$$

implying

$$\boldsymbol{u}' = \frac{d\boldsymbol{u}}{dz} = \frac{d\boldsymbol{u}}{dt}\frac{dt}{dz} = \frac{1}{z}\left(B(0)\boldsymbol{u} + zA(z)\boldsymbol{u}\right) \tag{341}$$

which is the same as (95). In the form (340) the system is nonlinear (in z only) but free of "explicit" singularities while (95) is linear but singular.

What resonances can we have? Note that the only nonzero monomials in (340) are of the form

$$\mathbf{w}^{\mathbf{k}} = u_1^{k_1} \cdots u_n^{k_n} z^p = u_i z^p e_j \Rightarrow \lambda_j = \lambda_i + p \tag{342}$$

which is exactly the Frobenius resonance condition, of which we can only have finitely many (regardless of whether we are in a Poincaré or Siegel domain)!

The space of functions whose all Taylor monomials are of the type $u_i z^p e_j$ forms a Banach space preserved by the right side of (340) (which is linear in u_i). Thus the Poincaré-Dulac proof applies with minor changes, and Frobenius theory is subsumed by Poincaré-Dulac linearization theory. Note also that after linearization and possibly extension of the system, Poincaré-Dulac reduces the study of a regular singularity, analytically, to a system with **constant coefficients**, analyzed at $t = \infty$, or to an Euler system by taking $t = \ln s$ to analyze it at $s = 0$.

16 Newton's method

Consider a simple contractive mapping setting, a linear one,

$$X = X_0 + LX \tag{343}$$

where $\|L\| = \lambda < 1$. If we iterate $X_{n+1} = X_0 + LX_n$ we have

$$\|X_{n+1} - X_n\| \leqslant \lambda \|X_n - X_{n-1}\| \Rightarrow \|X_{n+1} - X_n\| \leqslant C\lambda^n \tag{344}$$

and this is all that is guaranteed in full generality; indeed we can look at the one-dimensional case $x_{n+1} = a + \lambda x_n$, $|\lambda| < 1$ to see that this is optimal.

The actual convergence rate, depending on L may however may be faster. An example is the equation for the exponential $f' = f$; $f(0) = 1$ written in integral form, and iterated,

$$f(x) = 1 + \int_0^x f(s)ds; \quad f_{n+1} = 1 + \int_0^x f_n(s)ds; \quad f_0 = 0 \tag{345}$$

A simple induction argument shows that

$$f_n = \sum_{j=0}^{n-1} \frac{x^n}{n!} \tag{346}$$

and since $f_{n+1} - f_n = O(x^n/n!)$ and f_n approaches f factorially rather than geometrically.

Note 21. *If we are in a Hilbert space setting and the contractive operator is self-adjoint (or normal), then the convergence of the iterates is necessarily geometric since, then,* $\|A^n\| = \|A\|^n$.

In some cases, the convergence rate can be improved in the following way. To solve

$$X = F(X), \ \|F(X) - F(Y)\| \leqslant \lambda \|X - Y\| \tag{347}$$

where F is defined on a Banach space and has a continuous Fréchet derivative, we can write

$$\begin{aligned} X_{n+1} = F(X_{n+1}) = F(X_n) + DF(X_n)(X_{n+1} - X_n) + O((X_{n+1} - X_n)^2 \\ \Rightarrow X_{n+1} = [1 - DF(X_n)]^{-1}[F(X_n) - DF(X_n)] + O((X_{n+1} - X_n)^2 \end{aligned} \tag{348}$$

as the operator $(1 - DF(X_n))$ is invertible because of the contractivity of F. The error goes down like a^{2^n}. Equivalently, we can write this for the zero of an operator, $G(Z) = 0$, giving the iteration

$$Z_{n+1} = Z_n - [DG(Z_n)]^{-1}G(X_n) \tag{349}$$

which is known as the Newton-Kantorovich iteration. This is useful in at number of cases. For "practical purposes" in "actual calculations": at times we can explicitly write the inverse operator, and then the convergence is improved. An example is calculating the square root (the method was known to the Babylonians): $G = x^2 - a$, $G' = 2x$,

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \tag{350}$$

It is easy to check that the error satisfies $\varepsilon_{n+1} \leqslant \varepsilon_n^2/2$, $\varepsilon_n = 2(\varepsilon_0/2)^{2^n}$.

In a theoretical setting, this is useful when in a fixed point problem when the underlying operator in *not contractive* but has bounded norm when acting from a ball into a smaller ball, and the radius of the ball risks to shrink to zero if we straightforwardly iterate the operator. This is the case in the Siegel domain, due to small denominators.

16.1 The Siegel and Brjuno conditions

Definition 32. *(a) A point $\boldsymbol{\lambda} \in \mathbb{C}^n$ is of* **Siegel** *type (C, ν) where C and ν are positive constants, if for all $j = 1, 2, ..., d$ and $\mathbf{k}$ with $k_i + 1 \in \mathbb{N}, |\mathbf{k}| \geqslant 2$ we have*

$$|\lambda_j - \mathbf{k}\boldsymbol{\lambda}| \geqslant C|\mathbf{k}|^{-\nu} \tag{351}$$

(b) The optimal condition, beyond which convergence is not expected in general, is the **Brjuno condition**. *Let*

$$\omega_k = \inf\{|\lambda_j - \mathbf{k}\boldsymbol{\lambda}| : k_i + 1 \in \mathbb{N}, i, j = 1, ..., d, |\mathbf{k}| \in [2, 2^k]\} \tag{352}$$

The condition is

$$-\sum_{k=0}^{\infty} \frac{\ln \omega_k}{2^k} < \infty \tag{353}$$

Note 22. It can be shown quite straightforwardly that the Brjuno condition implies the Siegel condition. It can also be shown that the Siegel condition holds on a set of full measure if $\nu > (d-2)/2$.

In our problem, for a Newton iteration, the straightforward approach is to write

$$L\mathbf{h}_{n+1} = \mathbf{F}(\mathbf{w}+\mathbf{h_{n+1}}) = \mathbf{F}(\mathbf{w}+\mathbf{h_n}) + D\mathbf{F}(\mathbf{w}+\mathbf{h_n})(\mathbf{h}_{n+1}-\mathbf{h}_n) + O((\mathbf{h}_{n+1}-\mathbf{h}_n)^2) \tag{354}$$

and discard at each stage $O((\mathbf{h}_{n+1}-\mathbf{h}_n)^2)$. The precision of the iteration becomes quadratic, as it should in a Newton method, provided of course we invert the linear operator and write

$$\mathbf{h}_{n+1} = (L - D\mathbf{F}(\mathbf{w}+\mathbf{h_n}))^{-1}\Big[\mathbf{F}(\mathbf{w}+\mathbf{h_n}) - D\mathbf{F}(\mathbf{w}+\mathbf{h_n})\mathbf{h}_n\Big] \tag{355}$$

This is certainly a possible approach, but a quite awkward one, because $L - D\mathbf{F}(\mathbf{w}+\mathbf{h_n})$ is not diagonal anymore, and the coefficient of the monomial $\mathbf{w}^{\mathbf{k}}$ depends now on the set of coefficients of lower order monomials in a rather messy way. A better way is to use the procedure described in §15.3.4, which produces a convergence with rate α^{2^n} only using L and already calculated $\mathbf{h}_n$s, seen below.

Proposition 33. *If* $\mathbf{F}$ *has a zero of order* n*, then* $\mathbf{h}$ *has a zero of order* n *and* $\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w})) - \mathbf{F}(\mathbf{w})$ *has a zero of order* $2n-1$.

Proof. The order of the zero of $\mathbf{h}$ is obvious from (358). We have

$$\mathbf{F}(\mathbf{w}+\mathbf{h}(\mathbf{w})) - \mathbf{F}(\mathbf{w}) = D\mathbf{F}(\mathbf{w})\mathbf{h}(\mathbf{w})(1+o(1)) \tag{356}$$

where $D\mathbf{F}$ has a zero of order $n-1$ and $\mathbf{h}(\mathbf{w})$ has a zero of order n, and the statement follows. ☐

This suggests that the conjugation map, iterated in this way, has a convergence comparable to Newton's method.

16.1.1 The iteration under the Siegel condition

We could work with a condition closer to Brjuno's, essentially in the same way, but at the price of complicating the algebra quite a bit. We assume instead that $\boldsymbol{\lambda}$ is of Siegel (C,ν) type.

Note 23. *In the following, as usual, the symbol* $\lesssim$ *means less equal, up to a multiplicative constant the value of which is irrelevant.*

Proposition 34. *If*

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}_{\mathrm{n}}(\mathbf{x}); \;\; \textit{and } \mathbf{x} = \mathbf{w} + \mathbf{h}_{\mathrm{n}}(\mathbf{w}) \tag{357}$$

where

$$L\mathbf{h}_{\mathrm{n}} = \mathbf{F}_{\mathrm{n}}(\mathbf{w}) \tag{358}$$

then, the equation for $\mathbf{w}$ *is*

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} + \mathbf{F}_{n+1} := \Lambda\mathbf{w} + \big[I + (D\mathbf{h}_{\rm n})(\mathbf{w})\big]^{-1}\big[\mathbf{F}_{\rm n}(\mathbf{w} + \mathbf{h}_{\rm n}(\mathbf{w})) - \mathbf{F}_{\rm n}(\mathbf{w})\big] \quad (359)$$

Proof. Straightforward verification. □

Lemma 35. *Assume* $\mathbf{f} \in C_a^\omega(\mathbb{D}_R)$ *and* $|\hat{K}\mathbf{f}_{\mathbf{k}}\mathbf{w}^{\mathbf{k}}e_j| \lesssim |\mathbf{k}^\mu||\mathbf{f}_{\mathbf{k}}\mathbf{w}^{\mathbf{k}}|(\mu > 0)$*. Then, if* δ_n *is small enough (smaller than a constant depending on* μ, R, C*) and* $R'/R = e^{-\delta_n}$ *we have*

$$|||\hat{K}\mathbf{f}|||_{R'} \lesssim \delta_n^{-b}|||\mathbf{f}|||_R;\ b = \mu + d + 1 \quad (360)$$

Proof. We can check by induction that $\sum_{|\mathbf{k}|=n} 1 \leqslant dn^d$. If $|||\mathbf{f}||| = M$, then, in particular, $|f_{\mathbf{k}}| \leqslant M/R^{|\mathbf{k}|}$. Then,

$$\sum_{\mathbf{k}>0} |f_{\mathbf{k}}||\mathbf{k}|^\mu R'^{|\mathbf{k}|} \leqslant M\sum_{\mathbf{k}>0}|\mathbf{k}|^\mu\left(\frac{R'}{R}\right)^{|\mathbf{k}|} \leqslant dM\sum_{k=0}^{\infty} k^{\mu+d}e^{-k\delta_n}$$
$$\leqslant dM\int_0^\infty k^{\mu+d}e^{-k\delta_n}dk = d\,M\,\Gamma(\mu+d+1)\,\delta_n^{-(\mu+d+1)} \lesssim \delta_n^{-b} \quad (361)$$

□

Corollary 36. *We have*

$$|||\mathbf{h}_n|||_{R'} \lesssim \delta_n^{-b}|||\mathbf{F}_{\rm n}(\mathbf{w})|||_R \quad (362)$$

Proposition 37.

$$|||\mathbf{h}_{n+1}|||_{R'} \lesssim \delta_n^{-\gamma}|||\mathbf{h}_n|||_R^2 \quad \gamma = \nu + d + 3 \quad (363)$$

Proof. Recall the definition (359). On a disk of radius $< R$ we write

$$|\mathbf{F}_{\rm n}(\mathbf{w} + \mathbf{h}_{\rm n}(\mathbf{w})) - \mathbf{F}_{\rm n}(\mathbf{w})| \leqslant \sup|D\mathbf{F}_n|\sup|\mathbf{h}_n| \leqslant \sup|DL\mathbf{h}_n|\sup|\mathbf{h}_n| \quad (364)$$

here we used $L\mathbf{h}_n = \mathbf{F}_n$ and the sup is taken on any disk of radius $< R$. On the other hand $\mathbf{F}_n = L\mathbf{h}_n$ and we use $\mathbf{h}_{n+1} = L^{-1}\mathbf{F}_{n+1}$. On the disk of radius R' we have, using Lemma 35,

$$|||\mathbf{h}_{n+1}|||_{R'} \lesssim |||L^{-1}DL\mathbf{h}_n|||_{R'}\,|||\mathbf{h}_n|||_R \lesssim \delta_n^{-\gamma}|||\mathbf{h}_n|||_R^2; \quad (365)$$

if, say, $\sup\left|\big[I + (D\mathbf{h}_{\rm n})(\mathbf{w})\big]^{-1}\right| < 3/2$. □

Let $R_n = R_0\prod_{i=1}^n(1 - \exp(-\delta_n))$.

Corollary 38. *For some* C_1 *we have*

$$|||\mathbf{h}_{n+1}|||_{R_0\prod e^{-\delta_n}} \leqslant \left(C_1^n\prod\delta_n^{-\gamma}\right)|||\mathbf{h}_0|||^{2^n} \quad (366)$$

We choose a decreasing sequence of δs s.t. $\prod_{n=0}^{\infty}(1-\delta_n) > 0$, for instance $\delta_n = n^{-2}$. Then, in view of the telescopic nature of the product,

$$\prod_{n=2}^{\infty}(1-n^{-2}) = 1/2 \tag{367}$$

implying

$$|||\mathbf{h}_n|||_{R_n} \leqslant C^n (n!)^{2\gamma} |||\mathbf{h}_0|||_{R_0}^{2^n} \tag{368}$$

Corollary 39. *If $|||\mathbf{h}_0|||$ is small then $R_n \geqslant R_0/const.$ (see Note 24) and for large n,*

$$|||\mathbf{h}_n|||_{R_0/const} \lesssim 2^{-2^n} \tag{369}$$

The composition

$$(I+\mathbf{h}_1)\circ(I+\mathbf{h}_2)\circ\cdots\circ(I+\mathbf{h}_2)\cdots \tag{370}$$

is convergent and maps the equation

$$\dot{\mathbf{x}} = \Lambda\mathbf{x} + \mathbf{F}(\mathbf{x}) \tag{371}$$

to

$$\dot{\mathbf{w}} = \Lambda\mathbf{w} \tag{372}$$

Note 24. It is now a matter of algebra to show that one can consistently choose $|||\mathbf{h}_0|||$ small enough s.t. all the inequalities above hold *inductively*, for all n ;"const" above is not necessarily 2 since in the first few iterations, we may have to shrink the ball by more than what is suggested by (367).

Note 25. We could cast this in a contractive mapping setting by rescaling $\mathbf{w}$ at every iteration.

16.1.2 Simple model planar system

Consider the system

$$\dot{x} = ax;\ \dot{y} = by + xy^2 \tag{373}$$

The substitution $y = 1/z$ linearizes the system and it can be solved in closed form,

$$x(t) = C_2 \mathrm{e}^{at};\ y(t) = \mathrm{e}^{bt}\left(-\frac{\mathrm{e}^{bt+at}C_2}{b+a} + C_1\right)^{-1} \tag{374}$$

To find the linearization transformation explicitly, we solve for e^{at}, e^{bt} in terms of $x(t)$ and $y(t)$. Indeed, in this case the equations in the new variables should be $x_1' = ax_1, y_1' = by_1$. We take $C_1 = C_2 = 1$ and get

$$e^{at} = x(t);\ e^{bt} = \frac{y(t)(a+b)}{x(t)y(t)+a+b} \tag{375}$$

if $a+b \neq 0$; the linearizing transformation is then

$$x_1 = x;\ y_1 = \frac{y}{1+(a+b)^{-1}xy} \tag{376}$$

which is analytic at zero if $a+b \neq 0$. If $b=-a$ the system is resonant ($\lambda_2 = n\lambda_1 + (n+1)\lambda_2$) and xy^2 is a resonant monomial. The general solution is

$$x(t) = C_2 \mathrm{e}^{at};\ y(t) = \frac{\mathrm{e}^{-at}}{C_1 - C_2 t} \tag{377}$$

The system is hyperbolic, unless $a \in i\mathbb{R}$. The linearizing transformation is

$$x_1 = x;\ y_1 = \exp(-a^{-1}W(-ay^{-1}e^{-a})) \tag{378}$$

where W is the Lambert function, with the convergent expansion at the origin

$$y_1 = y\Big(1 - \frac{\ln(\ln z)}{\ln(z)} + \frac{\ln(\ln z)}{(\ln z)^2} + \frac{(\ln(\ln z))^2}{2(\ln(z))^2} - \frac{(\ln(\ln z))^2}{2(\ln(z))^3} - \frac{\ln(\ln z)}{(\ln z)^3} - \frac{(\ln(\ln z))^3}{6(\ln z)^3} + \cdots\Big) \tag{379}$$

where $z = -ay^{-1}e^{-a}$. Note that the transformation is not defined in a neighborhood of the origin, but on a Riemann surface. A real valued system with eigenvalues $\pm i$ and with the resonant monomial xy^2 after diagonalization is

$$x' = y;\ y' = -x - x^3 + x^2 y + xy^2 - y^3 \tag{380}$$

Note that in the real valued case, if $\lambda_1 = -ai, a \in \mathbb{R}$, then $\lambda_2 = ai, a \in \mathbb{R}$, and the system is always resonant, in the Siegel domain.

16.1.3 A 0 eigenvalue resonant example

The Euler system

$$\dot{x} = -x^2 \tag{381}$$

$$\dot{y} = y - x \tag{382}$$

with linearized part

$$\begin{pmatrix} 0 & 0 \\ -1 & 1 \end{pmatrix} \tag{383}$$

and $\lambda_1 = 1, \lambda_2 = 0$. The equations for the trajectories near $x=0, y=0$ is

$$y' + x^{-2}y - x^{-1} = 0 \tag{384}$$

which we have seen before: the singularity at 0 is irregular. The general solution of (381) is

$$x = \frac{1}{t+C_1};\ y(t) = e^t\left[e^{C_1}\mathrm{Ei}(1, C_1+t) + C_2\right] \tag{385}$$

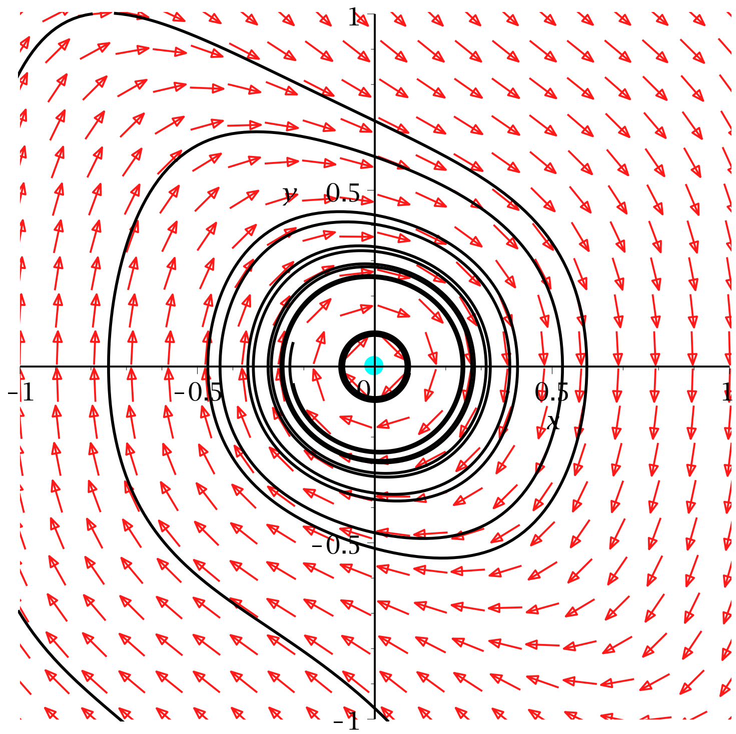


Figure 4: Contour plot of $x' = y;\ y' = -x - x^3 + x^2y + xy^2 - y^3$

while the solution of (384) is

$$y(x) = e^{1/x}\mathrm{Ei}(-1/x) + Ce^{1/x} \tag{386}$$

The Hamiltonian

$$H(x, y) = ye^{-1/x} - \mathrm{Ei}(-1/x) \tag{387}$$

has the same trajectories, as can be checked by writing the associated system (203). The trajectories near zero have exponential behavior and divergent series. We can linearize the system, but not in a very useful way: we can write $\mathrm{Ei}(-1/x) = \xi^2; y = \eta^2 e^{1/x(\xi)}$, but this is not close to the identity, neither does it help very much (except that here we have an explicit solution).

17 Planar systems

Assuming now we are studying a hyperbolic system in a neighborhood of an equilibrium. In a small neighborhood of the equilibrium, the system is equivalent by changes of coordinates to a linear system. So the local behavior is dictated by the types of flows associated to linear systems.

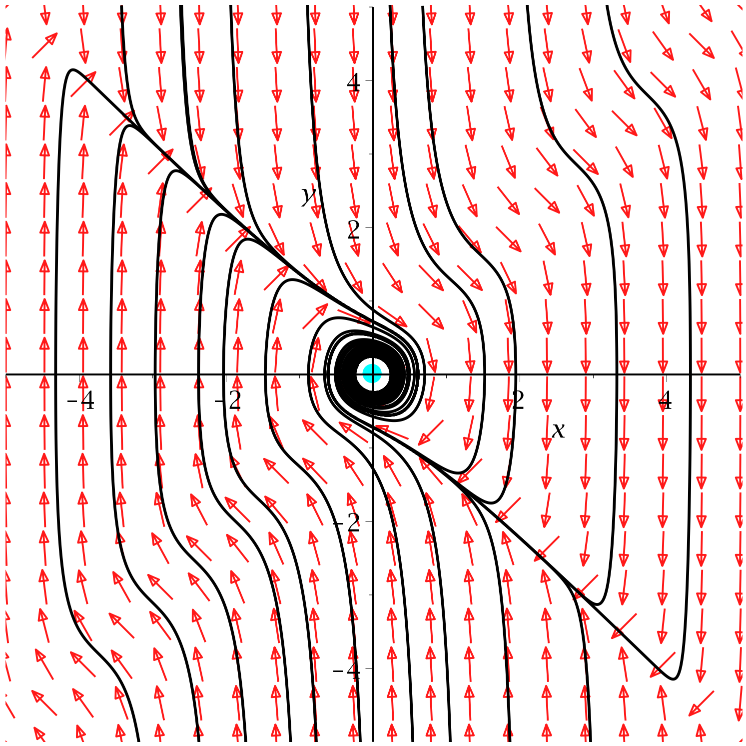


Figure 5: Contour plot of $x' = y;\ y' = -x - x^3 + x^2y + xy^2 - y^3$

Let

$$x' = Bx \tag{388}$$

where B is a 2×2 matrix with constant coefficients.

17.1 Distinct eigenvalues

In this case, the system can be diagonalized, and it is equivalent to a pair of trivial first order ODEs

$$x' = \lambda_1 x \tag{389}$$

$$y' = \lambda_2 y \tag{390}$$

17.1.1 Real eigenvalues

The change of variables that diagonalizes the system has the effect of rotating and rescaling the phase portrait of (389). The phase portrait of (389) can be fully described, since we can solve the system in closed form, in terms of simple

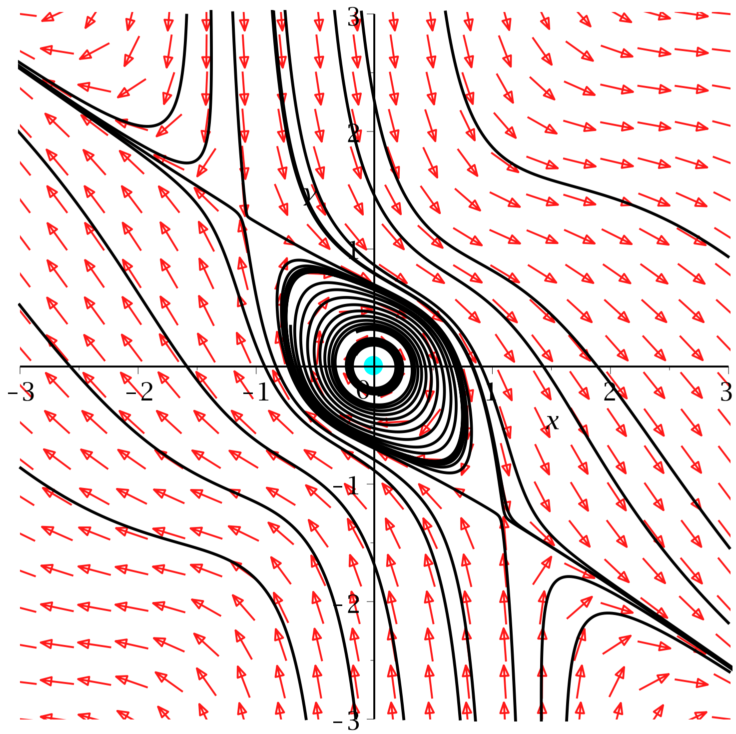


Figure 6: Contour plot of $x' = y + x^3$; $y' = -x - x^3 + x^2y + xy^2 - y^3$

functions:

$$x = x_0 e^{\lambda_1 t} \tag{391}$$

$$y = y_0 e^{\lambda_2 t} \tag{392}$$

On the other hand, we have

$$\frac{dy}{dx} = \frac{\lambda_2}{\lambda_1}\frac{y}{x} = a\frac{y}{x} \Rightarrow y = C|x|^a \tag{393}$$

where we also have as trajectories the coordinate axes: $y = 0$ ($C = 0$) and $x = 0$ ("$C = \infty$"). These trajectories are generalized parabolas. If $a > 0$ then the system is either (i) **a sink**, when both λ's are negative, in which case, clearly, the solutions converge to zero. See Fig. 8, or (ii) **a source**, when both λ's are positive, in which case, the solutions go to infinity.

The other case is that when $a < 0$; then the eigenvalues have opposite sign. Then, we are dealing with a **saddle**. The trajectories are generalized hyperbolas,

$$y = C|x|^{-|a|} \tag{394}$$

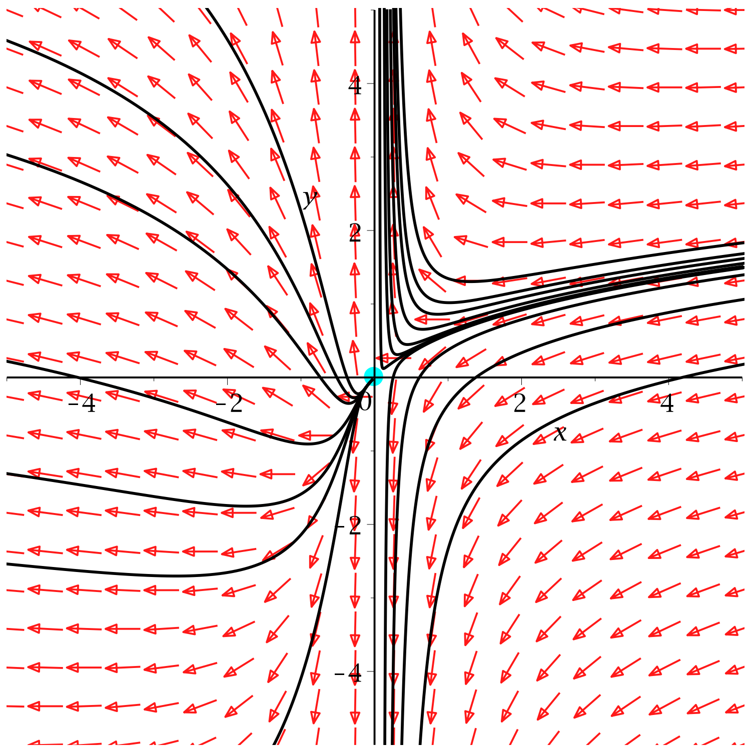


Figure 7: Contour plot of $x' = -x^2;\ y' = y - x$

Say $\lambda_1 > 0$. In this case there is a **stable manifold** the y axis, along which solutions converge to zero, and an **unstable manifold** in which trajectories go to zero as $t \to -\infty$. Other trajectories go to infinity both forward and backward in time. In the other case, $\lambda_1 < 0$, the figure is essentially rotated by $\pi/2$.

17.1.2 Complex eigenvalues

In this case we just keep the system as is,

$$x' = ax + by \tag{395}$$
$$y' = cx + dy \tag{396}$$

We solve for y, assuming $b \neq 0$ (check the case $b = 0$!), introduce in the second equation and we obtain a second order, constant coefficient, differential equation for x:

$$x'' - (a + d)x' + (ad - bc)x = 0 \quad or \tag{397}$$
$$x'' - \mathrm{tr}(B)x' + \det(B)x = 0 \tag{398}$$

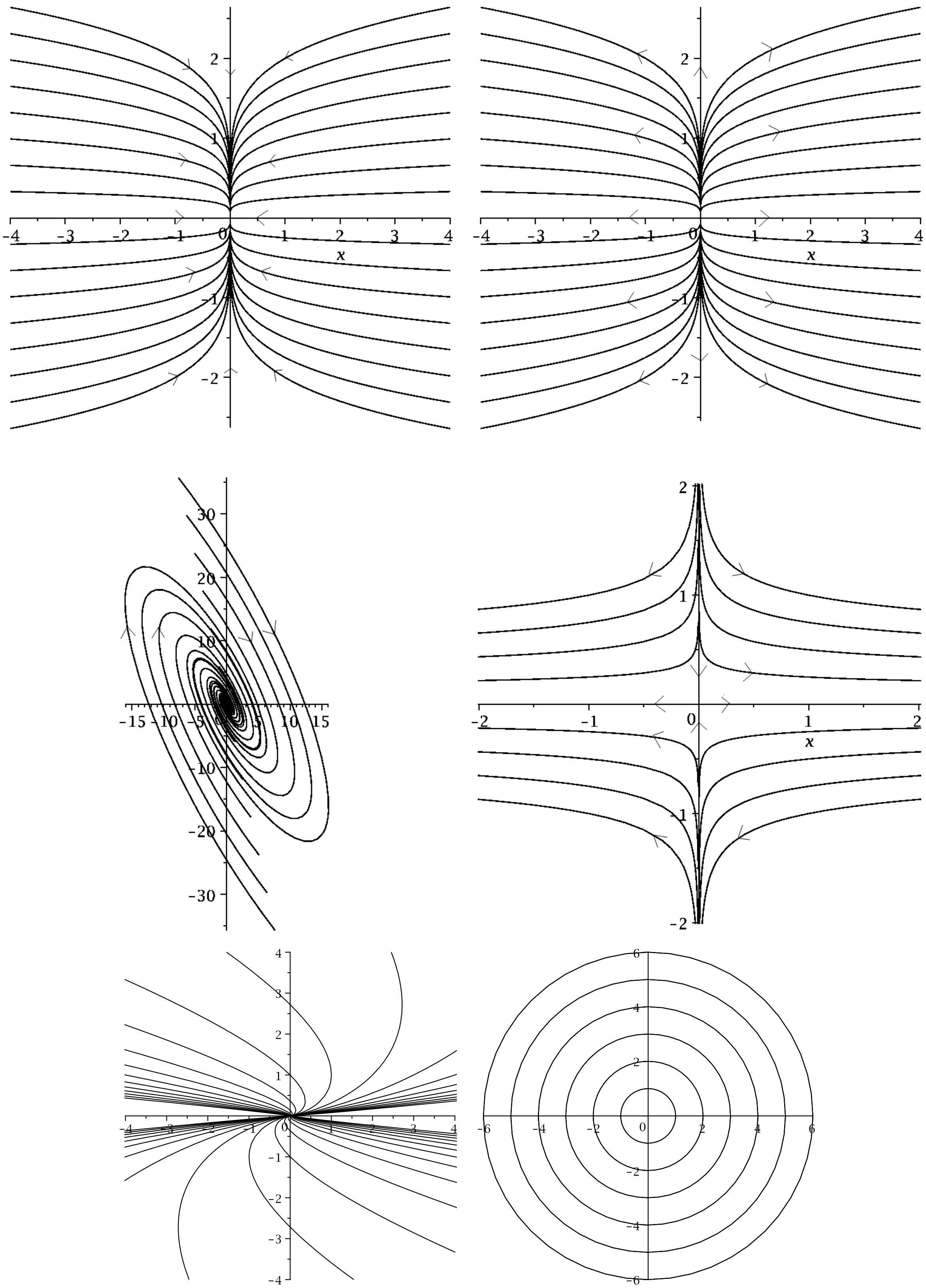


Figure 8: All types of linear equilibria in 2d, modulo euclidian transformations and rescalings: sink, source, spiral sink, saddle, nontrivial Jordan form, center resp. In the last two cases, the arrows point according to the sign of λ or ω, resp.

If we substitute $x = e^{\lambda t}$ in (397) we obtain

$$\lambda^2 - \mathrm{tr}(B)\lambda + \det(B) = 0 \tag{399}$$

and, evidently, since $\lambda_1 + \lambda_2 = \mathrm{tr}(B)$ and $\lambda_1\lambda_2 = \det(B)$, this is the same equation as the one for the eigenvalues of B. The eigenvalues of B have been assumed complex, and since the coefficients we are working with are real, the roots are complex conjugate:

$$\lambda_i = \alpha \pm i\omega \tag{400}$$

The real valued solutions are

$$x = Ae^{\alpha t}\sin(\omega t + \varphi) \tag{401}$$

where A and φ are free constants. Substituting in

$$y = b^{-1}x' - ab^{-1}x \tag{402}$$

we get

$$y(t) = Ae^{\alpha t}b^{-1}[(\alpha - 1)\cos(\omega t + \varphi) - \omega\sin(\omega t + \varphi) \tag{403}$$

which can be written, as usual,

$$y(t) = A_1e^{\alpha t}\sin(\omega t + \varphi_1) \tag{404}$$

If $\alpha < 0$, then we get a **spiral sink**. If $\alpha > 0$ then we get a spiral source, where the arrows are reverted.

A special case is that when $\alpha = 0$. This is the only non-hyperbolic fixed point with distinct eigenvalues. In this case, show that for some c we have $x^2 + cy^2 = A^2$, and thus the trajectories are ellipses. In this case, we are dealing with a **center**. We need more information about a nonlinear system to determine the nonlinear behavior.

17.2 Repeated eigenvalues

In 2d this case there is exactly one eigenvalue, and it must be real, since it coincides with its complex conjugate. Then the system can be brought to a Jordan normal form; this is either a diagonal matrix, in which case it is easy to see that we are dealing with a sink or a source, or else we have

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{405}$$

In this case, we obtain

$$\frac{dx}{dy} = \frac{x}{y} + \frac{1}{\lambda} \tag{406}$$

with solution

$$x = ay + \lambda^{-1}y\ln|y| \tag{407}$$

As a function of time, we can write

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t} = e^{\lambda t} \left[I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} t \right] \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \tag{408}$$

$$x(t) = (At + B)e^{\lambda t} \tag{409}$$

$$y(t) = Ae^{\lambda t} \tag{410}$$

We see that, in this case, only the x axis is a special solution (the y axis is not), and thus, all solutions approach (as $t \to \infty$ or $t \to -\infty$ for $\lambda < 0$ or $\lambda > 0$ respectively) the x axis.

Note 26. The eigenvalues of a matrix depend continuously on the coefficients of the matrix. In two dimensions you can see this by directly solving $\lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = 0$. Thus, if a linear or nonlinear system depends on a parameter α (scalar or not) and the equilibrium is hyperbolic when $\alpha = \alpha_0$, then the real parts of the eigenvalues will preserve their sign in a neighborhood of $\alpha = \alpha_0$. The type of equilibrium is the same and the local phase portrait changes smoothly unless the real part of an eigenvalue goes through zero.

Note 27. When conditions are met for a diffeomorphic local linearization at an equilibrium, then we have

$$\begin{pmatrix} x \\ y \end{pmatrix} = \varphi \begin{pmatrix} u \\ v \end{pmatrix} \tag{411}$$

where the equation in (u, v) is linear and the matrix φ is a diffeomorphism. We then have

$$\begin{pmatrix} x \\ y \end{pmatrix} = (D\varphi) \begin{pmatrix} u \\ v \end{pmatrix} + o(u, v) \tag{412}$$

which implies, in particular that the phase portrait very near the equilibrium is changed through a linear transformation.

17.3 Stable and unstable manifolds in 2d

Assume that g is differentiable, and that the system

$$\begin{pmatrix} x \\ y \end{pmatrix}' = g \begin{pmatrix} x \\ y \end{pmatrix} \tag{413}$$

has an equilibrium at zero, which is a saddle, that is, the eigenvalues of $(Dg)(0)$ are $-\mu$ and λ, where λ and μ are positive. We can make a linear change of variables so that $(Dg)(0) = \mathrm{diag}(-\mu, \lambda)$. Consider a linearization tangent to the identity, that is, with $Dg(0) = I$. We call the linearized variables (u, v).

Theorem 17. *Under these assumptions, in a disk of radius $\varepsilon > 0$ near the origin there exist two functions $y = f_+(x)$ and $x = f_-(y)$ passing through the origin, tangent to the axes at the origin and so that all solutions with initial conditions $(x_0, f_+(x_0))$ converge to zero as $t \to \infty$, while the initial conditions $(f_-(y_0), y_0)$ converge to zero as $t \to -\infty$. The graphs of these functions are called the* **stable and unstable manifolds, resp**. *All other initial conditions necessarily leave this disk as time increases, and also if time decreases.*

Proof. We show the existence of the curve f_+, the proof for f_- being the same, by reverting the signs. We have

$$\begin{aligned} x(t) &= \varphi_1(u(t), v(t)) \\ y(t) &= \varphi_2(u(t), v(t)) \end{aligned} \tag{414}$$

where (u, v) satisfy $u' = -\mu u$ and $v' = \lambda v$.

Consider a point $(\varphi_1(u_0, 0), \varphi_2(u_0, 0))$. There is a unique solution passing through this point, namely $(\varphi_1(u_+(t), 0), \varphi_2(u_+(t), 0))$ where $u_+(0) = u_0, v_+(0) = 0$. Since $u_+(t) \to 0$ as $t \to \infty$ and φ is continuous, we have

$$(\varphi_1(u_+(t), 0), \varphi_2(u_+(t), 0)) \to 0$$

as $t \to \infty$. We now write $(u, v) = \Phi(x, y)$. Along the decaying solution, we have $v = 0$. Since $\Phi = I + o(1)$, we have $\partial\Phi_2/\partial y = 1$ at $(0, 0)$, and the implicit function theorem shows that $\Phi_2(x, y) = 0$ defines a differentiable function $y = f(x)$ near zero, and $y'(0) = 0$ (by implicit differentiation, check). Note that $y = f_+(x)$ is equivalent to $v = 0$ and initial conditions with $v_0 = 0$ evolve to the origin, implying the conclusion for $(x_0, f_+(x_0))$. The proof for $(f_-(y_0), y_0)$ is similar. For other solutions we have, from (414), that x, y exits any small enough disk (check). □

17.4 Further examples, [8],[14]

17.4.1 Stable and unstable manifolds in an exactly solvable model

Consider the system

$$x' = x + y^2 \tag{415}$$
$$y' = -y \tag{416}$$

The linear part of this system at $(0, 0)$ is

$$x' = x \tag{417}$$
$$y' = -y \tag{418}$$

The associated matrix is simply

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{419}$$

with eigenvalues 1 and -1. They are resonant with the lowest degree of resonance 3. Then, the conditions of a differentiable homeomorphism, Theorem 8 are satisfied (but not, of course, those of analytic equivalence. Nonetheless, it will turn out that the system can be analytically linearized.)

Locally, near zero, the phase portrait of the system (419) is thus the prototypical saddle.

Insofar as the field lines go, we have

$$\frac{dx}{dy} = -\frac{x}{y} - y \tag{420}$$

a linear inhomogeneous equation that can be solved by variation of parameters, or more easily noting that, by homogeneity, $x = ay^2$ must be a particular solution for some a, and we check that $a = -1/3$. The general solution of the homogeneous equation is clearly $xy = C$. It is interesting to make it into a homogeneous second order equation by the usual method. We write

$$\frac{1}{y}\frac{dx}{dy} = -\frac{x}{y^2} - 1 \tag{421}$$

and differentiate once more to get

$$\frac{d^2x}{dy^2} = -2\frac{x}{y^2} \tag{422}$$

which is an Euler equation, with indicial equation $(\lambda - 2)(\lambda + 1) = 0$, and thus the general solution is

$$x(y) = ay^2 + \frac{b}{y} \tag{423}$$

where the constants are not arbitrary yet, since we have to solve the more stringent equation (420). Inserting (423) into (420) we get $a = -1/3$. Thus, the general solution of (421) is

$$3xy + y^3 = C \tag{424}$$

which can be, of course, solved for x.

To linearize the system we note that

$$y(t) = C_1 e^{-t},\ x(t) = -\frac{1}{3}e^{-2t} + C_2 e^t = -\frac{1}{3C_1}y(t)^2 + C_2 e^t \tag{425}$$

where we solve, as in the previous sections, for e^t and e^{-t} in terms of x and y:

$$e^{-t} = y/C_1;\ e^t = \frac{1}{C_1}x(t) + \frac{1}{3}y(t) \tag{426}$$

and thus we expect that the transformation

$$x_1 = x + \frac{1}{3}y^2,\ y_1 = y \tag{427}$$

linearizes the system. Indeed, we have

$$x_1' = x_1,\ y_1' = -y_1 \tag{428}$$

Note 28 (Global linearization)**.** The linearizing change of coordinates is thus

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = I \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} y^2/3 \\ 0 \end{pmatrix} \tag{429}$$

and in particular we see that the transformation is, to leading order, the identity. The unstable manifold is $y_1 = 0 = y$ and the stable one is $x_1 = 0$, the parabola $x = -y^2/3$.

Without using this explicit solution, the phase portrait can be obtained in

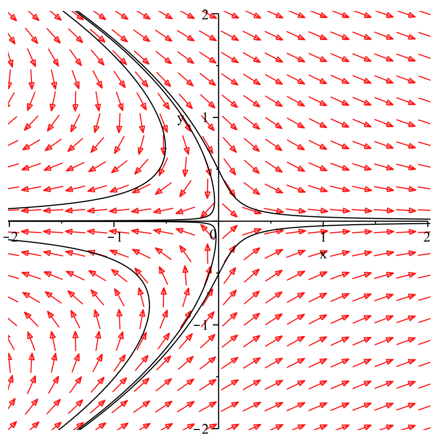


Figure 9: Phase portrait of (415)

the following way: we note that near the origin, the system is diffeomorphic to the linear part, thus we have a saddle there. There is a particular solution with $x = -1/3y^2$ and the field can be completed by analyzing the field for large x and y. This separates the initial conditions for which the solution ends up in the right half plane from those confined to the left half plane.

17.5 A limit cycle

Up to now we looked at equilibria, fixed points of the flow, which, along some direction(s), attract solutions as $t \to \infty$ or $t \to -\infty$. Fixed points are of course special, degenerate, trajectories. In nonlinear systems, solutions may be attracted by more structured trajectories: limit cycles.

We follow again [8], but with a different starting point. Let's look at the simple system

$$r' = r(1 - r^2)/2 \tag{430}$$
$$\theta' = 1 \tag{431}$$

where, later, we will think of (r, θ) as polar coordinates.

Obviously, we can solve (430) in closed form. The flow clearly has no fixed point, since $\theta' = 1 \neq 0$.

To solve the first equation, note that if we multiply by $2r$ we get

$$2rr' = r^2(1 - r^2) \tag{432}$$

or, with $u = r^2$,

$$u' = u(1 - u) \tag{433}$$

The exact solution is

$$r = \pm(1 + Ce^{-t})^{-1/2};\ \text{and also } r = 0;\ \pm 1, \text{ as special constant solutions} \tag{434}$$

$$\theta = t + t_0 \tag{435}$$

We see that all solutions that start away from zero converge to one as $t \to \infty$. We now interpret r and θ as polar coordinates and write the equations for x

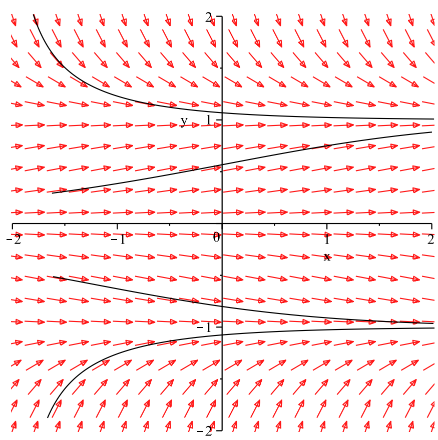


Figure 10: Phase portrait of (430)

and y. We get

$$x' = r'\cos\theta - r\sin\theta\theta' = \frac{1}{2}r(1 - r^2)\cos\theta - r\sin\theta$$
$$= \frac{1}{2}x - y - \frac{1}{2}(x^3 + xy^2) \tag{436}$$

$$y' = r'\sin\theta + r\cos\theta\theta' = \frac{1}{2}r(1 - r^2)\sin\theta + r\cos\theta$$
$$= x + \frac{1}{2}y - \frac{1}{2}(x^2y + y^3) \tag{437}$$

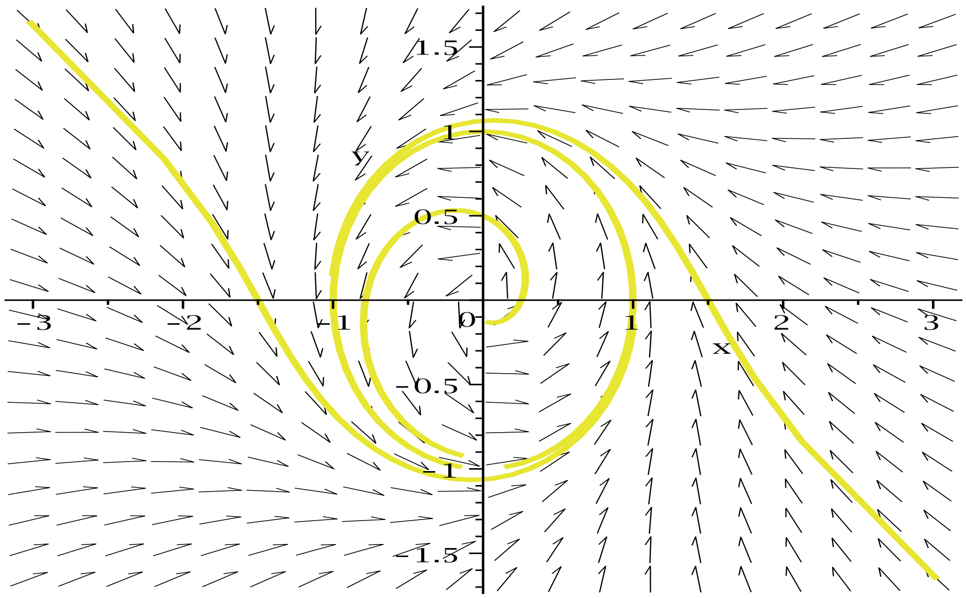


Figure 11: Phase portrait of (438), (439).

thus the system

$$x' = \frac{1}{2}x - y - \frac{1}{2}(x^3 + xy^2) \tag{438}$$

$$y' = x + \frac{1}{2}y - \frac{1}{2}(x^2y + y^3) \tag{439}$$

which looks rather hopeless, but we know that it can be solved in closed form.

To analyze this system, we see first that at the origin the matrix is

$$\begin{pmatrix} \frac{1}{2} & -1 \\ 1 & \frac{1}{2} \end{pmatrix} \tag{440}$$

with eigenvalues $1/2 \pm i$. Thus the origin is a spiral source.

Exercise 1. (a) Show that the only equilibrium of (438), (439) is $(0, 0)$.

(b) Although solvable in polar coordinates, show that there is no regular enough expression for the solutions of (438), (439), say as an implicit representation $F(x, y) = C$ with $F \in C(\mathbb{R}^2)$. (Note however that the system $\dot{y} = 2y, \dot{x} = x$ does not have a continuous implicit representation, but it has a meromorphic one, $y/x^2 = C$.)

Now we know the solution globally, by looking at the solution of (430) and/or its phase portrait.

We note that $r = 1$ is a solution of (430), thus the unit circle is a trajectory of the system (438). It is a closed curve, all trajectories tend to it asymptotically. This is a **limit cycle**.

17.6 Application: constant real part, imaginary part of analytic functions

Assume for simplicity that f is entire. The transformation $z \to f(z)$ is associated with the planar transformation $(x, y) \to (u(x, y), v(x, y))$ where $f = u + iv$. The grid $x = const, y = const$ is transformed into the grid $u = const, v = const$. We can first look at what this latter grid is transformed back into, by the transformation.

We take first $v(x(t), y(t)) = const$. We have

$$\frac{\partial v}{\partial x} x'(t) + \frac{\partial v}{\partial y} y'(t) = 0 \tag{441}$$

which we can write, for instance, as the system

$$x' = \frac{\partial v}{\partial y} \tag{442}$$

$$y' = -\frac{\partial v}{\partial x} \tag{443}$$

which, in particular, is a Hamiltonian system. We have a similar system for u. We can draw the curves $u = const, v = const$ either by solving this implicit equation, or by analyzing (442), or even better, by combining the information from both. Let's take, for example $f(z) = z^3 - 3z^2$. Then, $v = 3x^2y - y^3 - 6xy$. It would be rather awkward to solve $v = c$ for either x or y. The system of equations reads

$$x' = -6x + 3x^2 - 3y^2 \tag{444}$$

$$y' = 6y - 6xy \tag{445}$$

Note that $\nabla u = 0$ is equivalent to $z' = 0$ and so is $\nabla v = 0$. For equilibria, we thus solve $3z^2 - 6z = 0$ which gives $z = 0; z = 2$. Near $z = 0$ we have

$$x' = -6x + o(x, y) \tag{446}$$

$$y' = 6y + o(x, y) \tag{447}$$

which is clearly a saddle point, with x the stable direction and y the unstable one. At $x = 2, y = 0$ we have, denoting $x = 2 + s$,

$$s' = 6s + o(s, y) \tag{448}$$

$$y' = -6y + o(s, y) \tag{449}$$

another saddle, where now $y = 0$ is the stable direction. We note that $y = 0$ is, in fact, a special trajectory, and it belongs to the nonlinear unstable/stable manifold at the equilibrium points. Note also that the nonlinear stable manifold at zero is the same as th unstable one at 2: this is a heteroclinic orbit, or a heteroclinic connection.

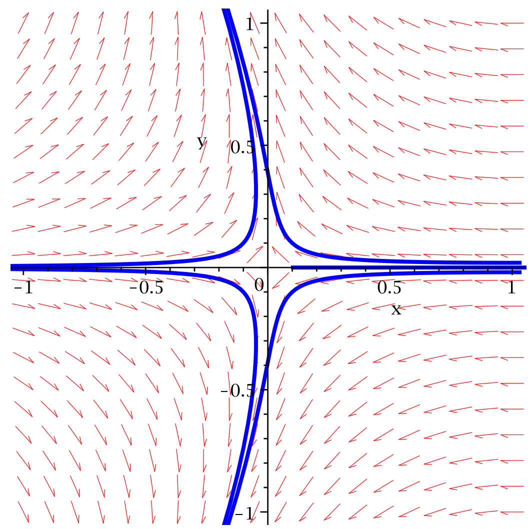


Figure 12: Phase portrait of (444) near $(0, 0)$.

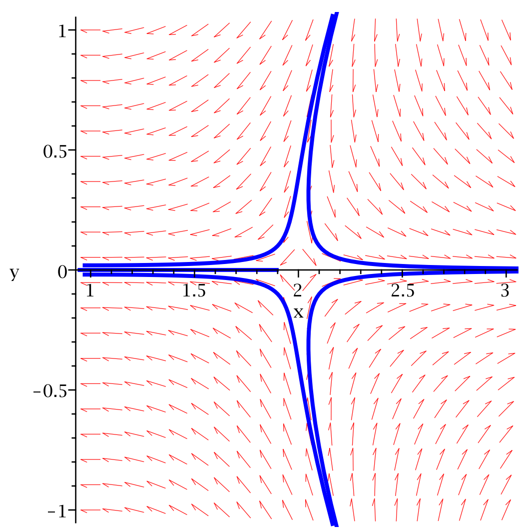


Figure 13: Phase portrait of (444) near $(0, 2)$.

We draw the phase portraits near $x = 0$ or near $x = 2$, we mark the special trajectory, and look at the behavior of the phase portrait at infinity. Then we "link" smoothly the phase portraits at the special points, and this should suffice for having the phase portrait of the whole system.

For the behavior at infinity, we note that if we write

$$\frac{dy}{dx} = \frac{y(1-6x)}{-6x+3x^2-3y^2} \tag{450}$$

we have the special solution $y = 0$, and if $|x| \gg 1, |y| \gg 1$, then the nonlinear terms dominate and we have

$$\frac{dy}{dx} \approx \frac{-6yx}{3x^2-3y^2} \tag{451}$$

By homogeneity, we look for special solutions of the form $y = ax$ (which would be asymptotes for the various branches of $y(x)$. We get, to leading order,

$$a = \frac{-6a}{3-3a^2} \tag{452}$$

We obtain

$$a = 0, a = \pm\sqrt{3} \tag{453}$$

We also see that, if $x = o(y)$, then $y' = o(1)$ as well. This would give us information about the whole phase portrait, at least qualitatively.

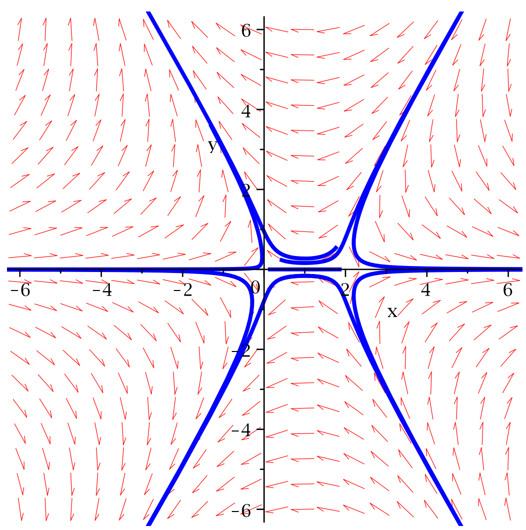


Figure 14: Phase portrait of (444), $v = const$.

Exercise 2. Analyze the phase portrait of $u(x, y) = const$.

The two phase portraits, plotted together give Note how the fields intersect at right angles, except at the saddle points. The reason, of course, is that $f(z)$ is a conformal mapping wherever $f' \neq 0$.

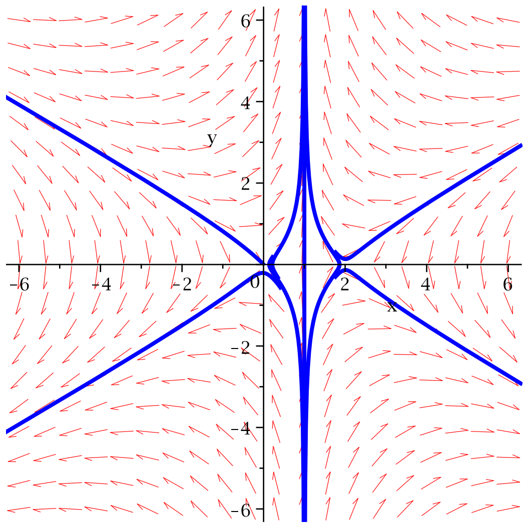


Figure 15: Phase portrait of $u = const$

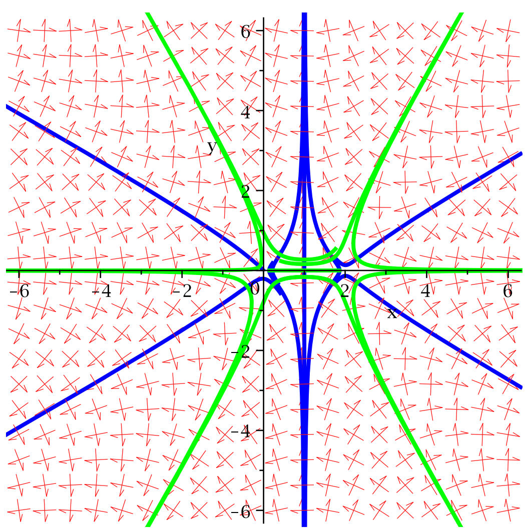


Figure 16: Phase portrait of $u = const$, and $v = const$.

Exercise 3. Draw the global phase portrait of the approximations of the pendulum, $x'' + x - x^3/6 = 0$, $x'' + x - x^3/3 + x^5/5 = 0$. Find the equilibria, local and global behavior. Find out if there are limit cycles. Discuss the conection sith the physical pendulum, $x'' + \sin(x) = 0$.

Exercise 4. Draw the global phase portrait of the damped pendulum, $x'' +$

$ax' + \sin x = 0$, where $a > 0$ is the air friction coefficient. Discuss what happens as $a \to 0$ and how this relates to the undamped pendulum, $a = 0$. Discuss also the bifurcation that occurs at $a = 0$ and $x = 0$ (for $a < 0$, the physical interpretation could be that we are looking backwards in time. Also, note any global bifurcations, that is changes in the global topology.

18 Bifurcations

Bifurcations occur when a change in a parameter induces topological changes in the phase portrait. These can be local, global (or both). There are many physical systems that are modeled by bifurcating dynamical systems, including reaction-diffusion equations, pattern formation, laser dynamics and so on.

Local bifurcations refer to the situations when a parameter crosses a value where the stability of in a neighborhood of a local equilibrium (or coalescing ones) changes. Global bifurcations affect higher (than zero...) dimensional attractors, such as limit cycles.

As we know, the phase portrait of a system depending on a parameter changes its **topological** structure near an equilibrium only if at least one eigenvalue becomes purely imaginary. This may happen if one eigenvalue (or both, but generically one) becomes zero, or else they pass through a point where they are nonzero, imaginary and complex conjugate to each other (since we are dealing with real-valued equations).

The classification is made by looking at the *normal form.* We keep all terms that cannot be eliminated topologically (using Hartman-Grobman) when the parameter changes in a neighborhood of the bifurcation point.

For instance, near $a = 0$,

$$x' = x^2 + r; \quad y' = -y$$

will represent a general system of the form

$$x' = a + bx^2 + O(x^3); \quad y' = -y + O(y^2)$$

18.1 Some types of bifurcations

We study the normal forms that are quadratic, or when the quadratic term i missing due, say, to a symmetry, cubic. They are classified according to this degree, and also according to the position of the "external" parameter.

$$x' = x^2 + r; \quad y' = -y; \quad \textbf{saddle-node bifurcation}$$

Here, for $r < 0$ there are two equilibria (a saddle and a node) that coalesce when $r = 0$ (when we have a "saddle-node"), while there is no equilibrium when $r > 0$.

$$x' = rx - x^2; \quad y' = -y; \quad \textbf{transcritical bifurcation}$$

When r goes through zero, two equilibria coalesce, and after the coalescence we still have two equilibria.

$$x' = rx - x^3; \quad y' = -y; \quad \textbf{supercritical pitchfork bifurcation}$$

One unstable equilibrium ($r < 0$) bifurcates into an unstable one and two stable ones.

$$x' = rx + x^3; \quad y' = -y; \quad \textbf{subcritical pitchfork bifurcation}$$

We see that in the pitchfork bifurcation, the quadratic term is absent (this is not so rare in applications due to symmetries of the system).

Of course, in the systems above, much of the information is contained in the x part, and we may in some sense ignore the y one, since the equations are decoupled.

$$x' = \beta x - y + \lambda x(x^2 + y^2); \quad y' = x + \beta y + \lambda y(x^2 + y^2) \ \textbf{Hopf bifurcation}$$

where $\beta = 0$ is the bifurcation point. Here, eigenvalues become imaginary, but nonzero.

18.2 Normal form of the saddle-node bifurcation

Consider a simple system which illustrates the first case, an eigenvalue going through zero, prototypical for *saddle-node bifurcations*,

$$x' = x^2 + r \tag{454}$$

$$y' = -y \tag{455}$$

Of course, we can solve this explicitly, but we choose not to, because solvable equations are infrequent. We first note that the only possible equilibria are $(\pm\sqrt{-r}, 0)$. Clearly, there are two of them if $r < 0$, one if $r = 0$ and none if $r > 0$. For $r = 0$, the equilibrium is non-hyperbolic and needs to be studied separately. For $r < 0$, at $x = \pm r$, we see that the linearized system is

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} \pm 2\sqrt{-r} & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{456}$$

Thus the point $(-\sqrt{-r}, 0)$ is a node, while $(\sqrt{-r}, 0)$ is a saddle.

Let's draw the complete phase portrait in the three regimes, $r < 0$, $r = 0$ and $r > 0$. Again, the portrait is determined by the set of equilibria, limit cycles, and by the behavior at infinity. The three lines $x = \pm\sqrt{-r}$ and $y = 0$ are special solutions of the system. We see that there are no limit cycles, since trajectories do not cross except at the equilibria, and the lines $x = \pm\sqrt{-r}$, never crossed, delimit regions where the sign of x' is constant.

Behavior at infinity: for x very large, we have $x' \approx x^2$ and $y' = -y$, and thus

$$\frac{dy}{dx} \approx \frac{-y}{x^2} \tag{457}$$

with the solution $y = Ce^{1/x}$. Thus in the far x field, the trajectories are expected to approach horizontal lines. How do we prove this rigorously? One way is to note that for any $\alpha > 1$ $(x^2 - r)^{-1} \leqslant \alpha x^{-2}$ if x is large enough.

Thus, we can write

$$\frac{y'(x)}{y(x)} \geqslant -\frac{\alpha}{x^2} \tag{458}$$

where we can integrate both sides and get

$$y(x) \geqslant C_{x_0,y_0} e^{\alpha/x} \tag{459}$$

where C_{x_0,y_0} is a constant depending on the initial condition (x_0, y_0). Similarly,

$$y(x) \leqslant C_{x_0,y_0} e^{\alpha'/x} \tag{460}$$

If instead x is bounded and $y \to \infty$, the direction field points straight to the origin, so there the trajectories essentially vertical lines. Piecing all this together, we get the phase portrait depicted below.

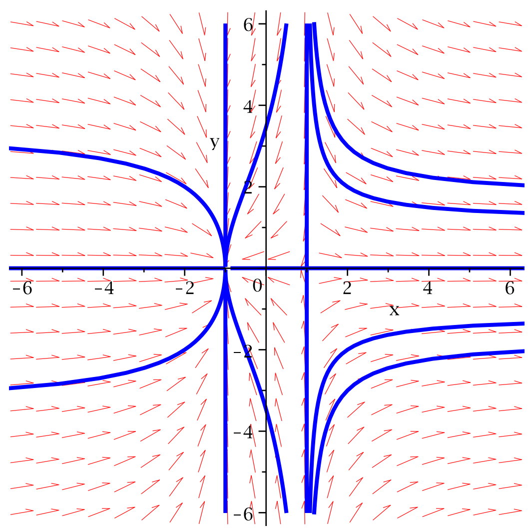


Figure 17: Phase portrait of (461) for $r = -1$

For $r = 0$ the system simply becomes

$$x' = x^2 \tag{461}$$
$$y' = -y \tag{462}$$

Clearly, the line $x = 0$, a special solution, is attracting, while the line $y = 0$ is repelling for $x > 0$ and attracting (since the field points towards the origin) for $x < 0$. So we see that, in some sense, the origin is now half-node, half saddle.

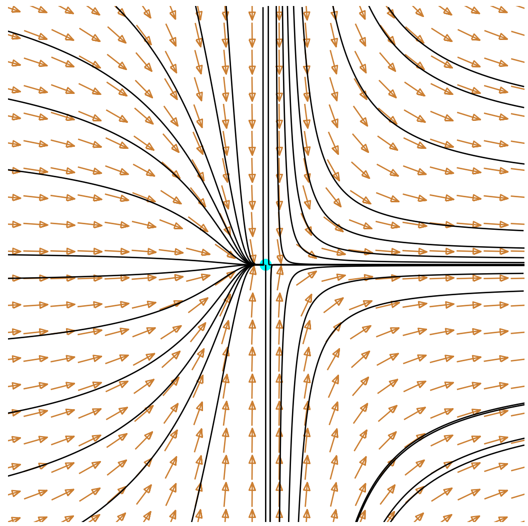
[h!] 

Figure 18: Phase portrait of (461) for $r = 0$

All nearby trajectories are attracted to zero if they start in the closed left half plane, and repelled otherwise.

The far-field picture is clearly the same as in the case $r < 0$, so we can piece together these informations to draw the phase portrait. We note that in this case, of course, the explicit solutions $y = Ce^{1/x}$, $C \in \mathbb{R}$ and $x = 0$, can be easily used to draw the phase portrait. This type of behavior as a function of r, at least in this example, explains the choice of name, saddle-node bifurcation.

Stable and center manifold

For hyperbolic systems, say near a saddle point, we have defined stable and unstable manifolds; these are manifolds invariant under the flow and tangent to the positive/negative eigenvectors.

Here one eigenvalue is zero. What manifolds are invariant under the flow and tangent to the eigenvectors? We have the y axis as a stable manifold, tangent to the direction $(0, 1)$, the eigenvector corresponding to the eigenvalue 1. The eigenvalue zero has $(1, 0)$ as eigenvector. $Ce^{1/x}$ are exact trajectories. *All of them* are tangent to Ox if $x < 0$, and one, with $C = 0$, that is Ox itself is tangent for $x > 0$. There is a continuum of invariant manifolds for $x < 0$, and thus in general since they can all be continued by $x = 0$ in the right half plane. These are called center manifolds, and in general none is privileged, except in an analytic setting like this one we could pick the analytic manifold $y = 0$. However, from the point of view of solution behavior, it is not distinguished.

“Linearization”

Note that the exact solution is $x = 1/(A - t), y = Be^{-t}$. Solving $y = e^{-t}$ for t, inserting in the first equation and solving for $1/A$ we get

$$x_1 = \frac{x}{1 - x \ln y} \tag{463}$$

with the property $x_1' = 0$. In the coordinates x_1, y the system is linear. The transformation from (x, y) to (x_1, y) has to be done separately in each quadrant,

$$x_1 = \frac{|x|}{1 - |x| \ln |y|} \tag{464}$$

and is quite singular, however.

*

Finally, for $r > 0$ there are no equilibria. We see that $x' > 0$ for all x. Trajectories extend from $-\infty$ to $+\infty$ in x. The behavior in the far field is the same as in the previous examples. The trajectories have horizontal lines as asymptotes for $x \to \pm\infty$ and, in the upper half plane, the asymptote for $x < 0$ lies above the one for $x > 0$, since $y' < 0$ there. We can now draw the phase portrait.

As we see, the node in the left half plane approaches the saddle, touches it at which time we have a half-node half-saddle picture, and then the equilibrium vanishes and the curves in the left half plane “spill over” in the right half plane.

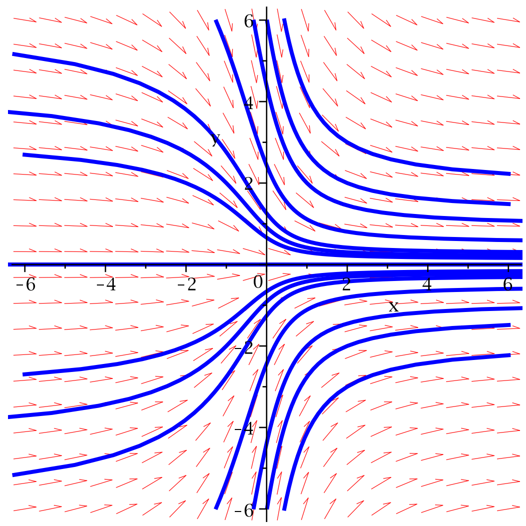


Figure 19: Phase portrait of (461) for $r > 0$

18.3 Transcritical bifurcation

In this type of bifurcation, there are two equilibria for all values of $r \neq 0$, a saddle and a node and one for $r = 0$. The node and saddle are interchanged when r changes sign.

Typically, for this and some other bifurcations, the y part is ignored, and for a good reason, as we mentioned there is effectively no y participation. The reason for which this type of bifurcation is called transcritical is from the way things look as a function of the parameter for the x-only system. To have however a unified picture in mind, and to recall that we are after all dealing with two dimensional systems for which it does *happen* that the normal form makes y "idle" we will look at the two dimensional system,

$$\begin{aligned} x' &= rx - x^2 \\ y' &= -y \end{aligned} \tag{465}$$

For $r \neq 0$ there are two equilibria, and for $r = 0$ only one; the two equilibria collide as before, but the outcome is different.

Take $r < 0$. Clearly, the origin, marked in blue, is a node. The other equilibrium, $x = r, y = 0$ is a saddle ($r - 2r = -r > 0$).

The global picture is obtained as before: the rays: $\{(-t, 0) : t < r\}$, $\{(t, 0) : t \in (r, 0)\}$,$\{(t, 0) : t > 0\}$, $\{(0, \pm t^2) : t > 0\}$ are special trajectories; in the far field, the trajectories are almost horizontal.

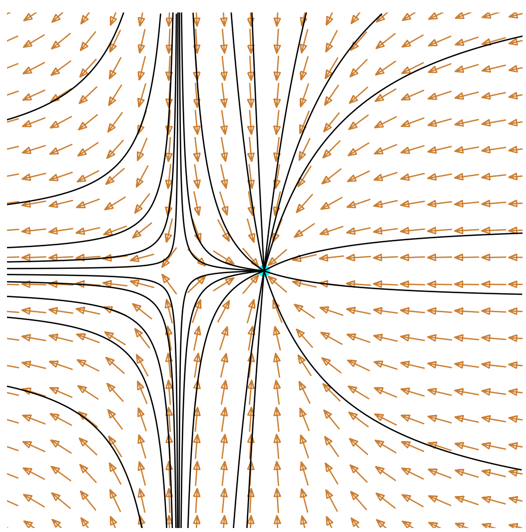


Figure 20: Transcritical phase portrait, $r < 0$

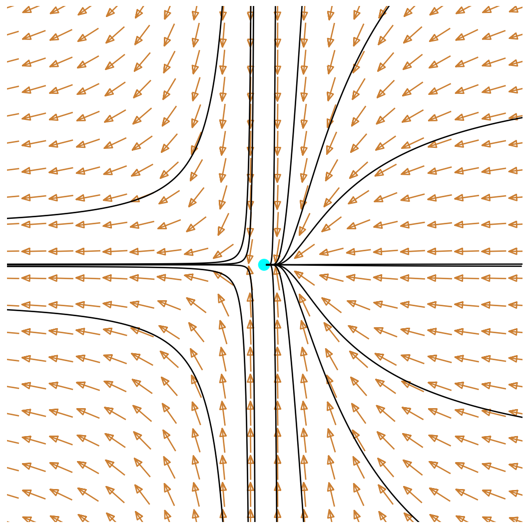


Figure 21: Transcritical phase portrait, $r = 0$

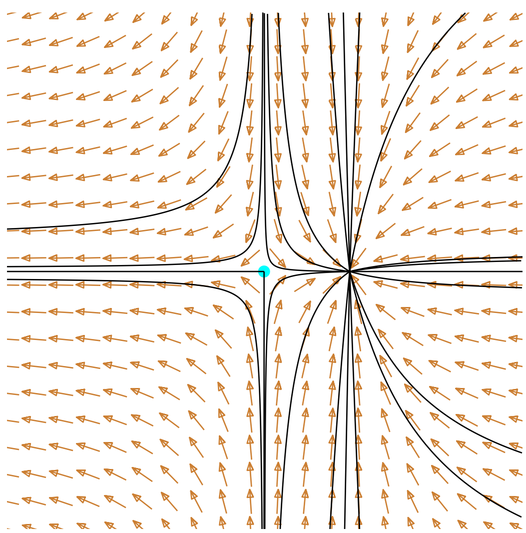


Figure 22: Transcritical phase portrait, $r > 0$

Here, we see that a saddle-node becomes a "half saddle-half node" and then it becomes a node-saddle and the types of equilibria are interchanged.

18.4 Normal form of the pitchfork bifurcation

Here we are dealing with a system with symmetries, the quadratic term is missing. The normal form of the supercritical pitchfork bifurcation is

$$\begin{aligned} x' &= rx - x^3 \\ y' &= -y \end{aligned} \tag{466}$$

whereas the subcritical one has the normal form

$$\begin{aligned} x' &= rx + x^3 \\ y' &= -y \end{aligned} \tag{467}$$

We look only at the supercritical case, the subcritical one being analyzed similarly.

18.4.1 Supercritical case

The field is an odd function of x and y, and stays odd for all (or only small, maybe) values of r. The name "pitchfork" will become clear in a moment.

In case 1) $r > 0$, we have three equilibria, $x = 0$ and $x = \pm\sqrt{r}$.

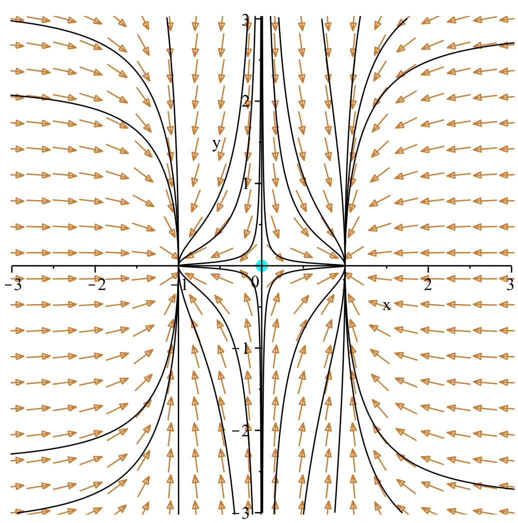


Figure 23: Pitchfork phase portrait, $r > 0$

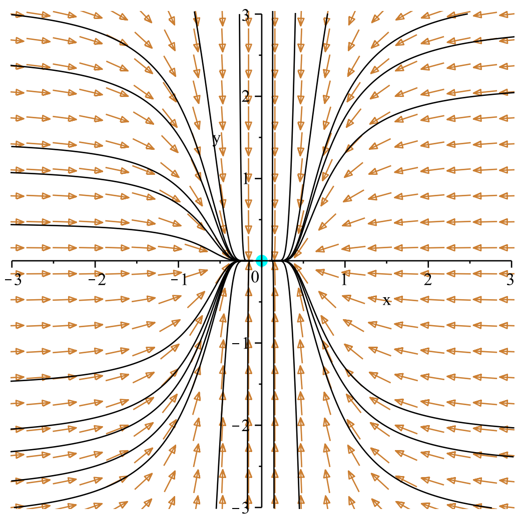


Figure 24: Pitchfork phase portrait, $r = 0$

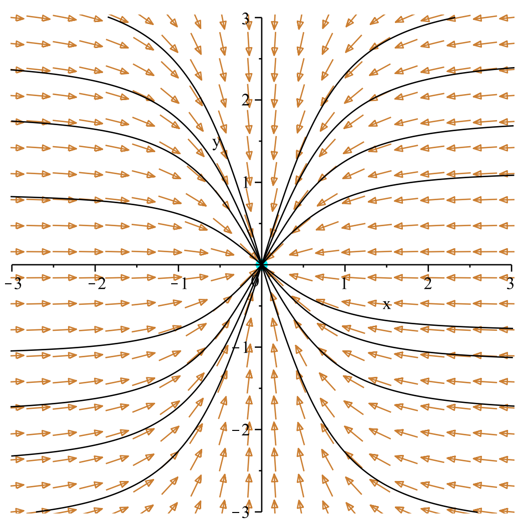


Figure 25: Pitchfork phase portrait, $r < 0$

It is clear that the origin is a saddle whereas the other two equilibria, symmetric, are nodes (sinks).

If, 2), $r = 0$, clearly we only have one equilibrium, and it is a node because $-x^3$ always points towards the origin.

By explicit solution, we see that the trajectories are given by $y = Ce^{-1/(2x^2)}$, which explains the fact that the phase portrait almost seems to have a continuum of nodes near zero.

Finally, in case 1) $r < 0$, we have only one equilibrium and it is a node. the number of equilibria changes.

Note that we can extend artificially the number of variables, to transform the two dimensional parameter-dependent problem into a three-dimensional parameter-free one,

$$x' = rx - x^3 \tag{468}$$
$$y' = -y \tag{469}$$
$$r' = 0$$

Clearly now the change in behavior is seen as a chenge in the 3d phase portrait, as a function of the initial condition in r.

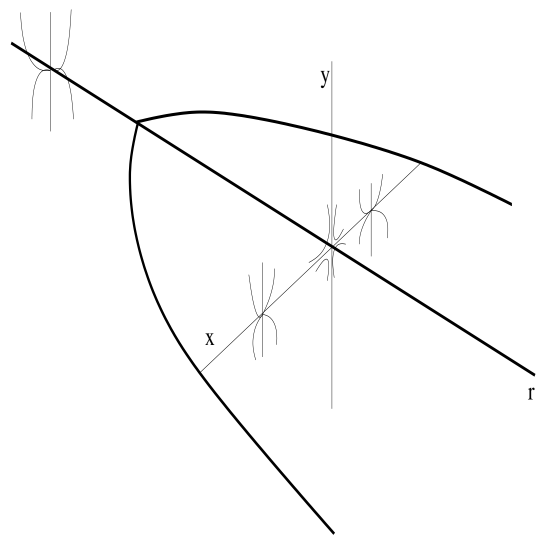


Figure 26: Pitchfork 3d phase portrait.

18.5 Normal form of the Hopf bifurcation

In this case, we are looking at a system for which passage through the critical value of the parameter (β) implies nonzero, purely imaginary eigenvalues. Take first $\sigma = -1$:

$$x' = \beta x - y - x(x^2 + y^2) \tag{470}$$
$$y' = x + \beta y - y(x^2 + y^2) \tag{471}$$

The origin is an equilibrium for all β, and it is the only one (this is best seen in polar coordinates, (473) below) where $\theta' > 0$. At the origin, the linearized system is

$$\begin{pmatrix} x \\ y \end{pmatrix}' = B \begin{pmatrix} x \\ y \end{pmatrix}; \quad B = \begin{pmatrix} \beta & -1 \\ 1 & \beta \end{pmatrix} \tag{472}$$

where the equation for the eigenvalues of the matrix B is $(\beta - \lambda)^2 + 1 = 0$, and thus $\lambda_\pm = \beta \pm i$. At $\beta \neq 0$ the equilibrium is hyperbolic, a spiral sink if $\beta < 0$ and a spiral source for $\beta > 0$. A change of phase portrait, a bifurcation, should occur at $\beta = 0$.

For a simple analysis of the phase portrait, we rewrite the system in polar coordinates.

$$r' = \beta r - r^3 \tag{473}$$

$$\theta' = 1 \tag{474}$$

For $\beta < 0$, $\beta r - r^3 = 0$ has only one solution, $r = 0$. In (x, y), all solutions converge to $(0, 0)$ (since $r' < 0$) while spiraling.

In the far field, we have

$$r' \approx -r^3 \tag{475}$$

$$\theta' = 1 \tag{476}$$

with solution

$$r = (2\theta + 2C)^{-1/2} \tag{477}$$

For r to be very large, we must have θ very close to $-C$. That is, asymptotically the curves in the far field (x, y) plane have radial lines as asymptotes. The spiraling ceases there.

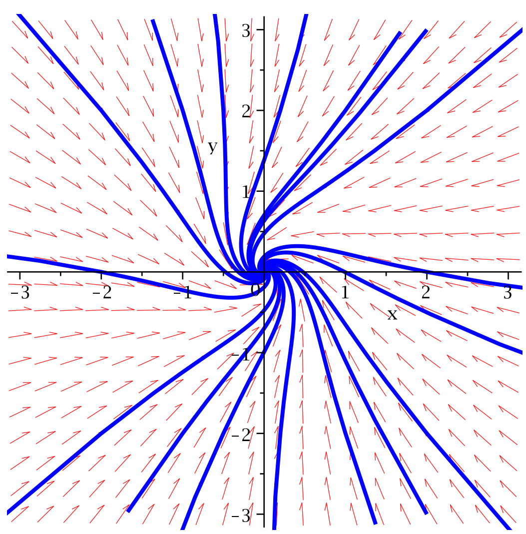


Figure 27: Phase portrait of (472) for $\beta = -1$

When $\beta = 0$, $r = 0$ is still the only solution of $\beta r - r^3 = 0$. Since again $r' < 0$, all trajectories go to the origin. The origin is approached at a very slow rate, $O(r^3)$, there is a lot of spiraling going on in that region. We see a tendency of a limit cycle being born.

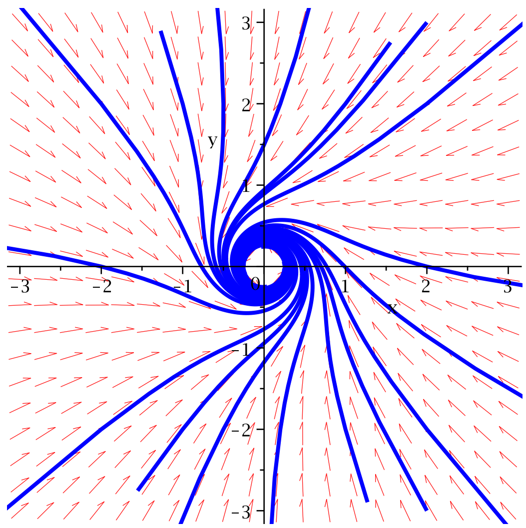


Figure 28: Phase portrait of (472) for $\beta = 0$

For $\beta > 0$, we have three solutions of $\beta r - r^3 = 0$, 0 and $\pm\sqrt{\beta}$ (the minus solution is "unphysical" for us). $r = 0$ is repelling and $r = \sqrt{\beta}$ is attracting. This means, in (x, y) that $x^2 + y^2 = \beta$ is *limit cycle*. We note that it approaches the origin as $\beta \to 0$. The spiral sink changes into a spiral source plus a limit cycle. It is probably worth looking at the exact solution, which can be obtained

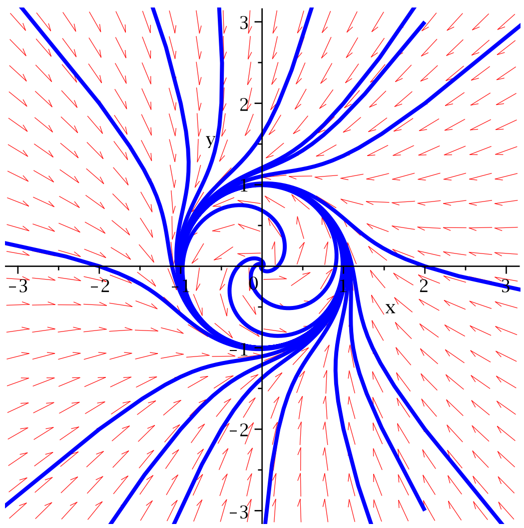


Figure 29: Phase portrait of (472) for $\beta = 1$

from the polar representation. For $\beta > 0$ we have

$$x(t) = \frac{\sqrt{\beta}\cos(t)}{\sqrt{1 + Ce^{-2\beta t}}};\ y(t) = \frac{\sqrt{\beta}\sin(t)}{\sqrt{1 + Ce^{-2\beta t}}} \tag{478}$$

For $\beta < 0$ the solution is

$$x(t) = \frac{\sqrt{|\beta|}\cos(t)}{\sqrt{e^{2|\beta|(t)} + C}};\ y(t) = \frac{\sqrt{|\beta|}\sin(t)}{\sqrt{e^{2|\beta|(t)} + C}} \tag{479}$$

while for $\beta = 0$ we get

$$x(t) = \frac{\cos t}{\sqrt{2t + C}};\ y(t) = \frac{\sin t}{\sqrt{2t + C}} \tag{480}$$

19 Bifurcations in more general systems. The central manifold theorem

Here we follow [7]. The setting is that of differential systems depending on a parameter,

$$\mathbf{x}' = \mathbf{f}_\mu(x), \quad x \in \mathbb{R}^n,\ \mu \in \mathbb{R}^m \tag{481}$$

where we assume sufficient smoothness of $\mathbf{f}_\mu$. To simplify the notation we will drop the boldface fonts wherever this is not confusing. Equilibrium solutions are given by the (constant) solutions of the equation

$$\mathbf{f}_\mu(x) = 0 \tag{482}$$

The equilibrium points depend smoothly on μ, by the implicit function theorem, as long as $D_x f_\mu$ is invertible, that is, as long as it has no zero eigenvalue. If $(\det D_x f_\mu)(x_0, \mu_0) = 0$, several branches of equilibria may form/disappear. These points (x_0, μ_0) are bifurcation points. For example, in the pitchfork bifurcation example, $\mu = r$ and $(0,0)$ is the only bifurcation point. In that case, the equilibria coalesce.

A crucial notion here is that of *transversality*. In one dimension, $y = f(x)$ crosses the x axis transversally at x_0 if $f(x_0) = 0$ and $f'(x_0) \neq 0$. In d dimensions, two manifolds intersect transversally if the tangent spaces at the intersection point span $\mathbb{R}^d$ (there is no loss in dimension). It is clear that transversal intersections are generic. In particular, two manifolds Σ_1 and Σ_2 of dimensions d_1 and d_2 intersect transversally along a manifold of dimension $d_1 + d_2 - d$. Equivalently, the codimension of $\Sigma_1 \cap \Sigma_2$ is $(d - d_1) + (d - d_2)$. Two surfaces in 3d intersect generically along a line, two generic curves do not intersect, and a curve and a manifold generically intersect at a point, etc.

For the vector field $\mu + x^2$ thought of as a family of curves in $\mathbb{R}$, the curve for $\mu = 0$ intersects the x axis non-transversally at $x = 0$.

However, if we lift the number of dimensions to include μ in the picture, we have a *transversal* intersection of the surface $F(x, \mu) = x^2 + \mu$ with the (x, μ) coordinate plane.

Also, it is clear that transversal intersections are stable in the following sense. If two manifolds intersect transversally, then any small perturbation of the manifolds will also have a transversal intersection. On the contrary, if two manifolds intersect non-transversally, then their generic perturbations will intersect transversally. Let f be a C^r vector field on $\mathbb{R}^n$ vanishing at the origin ($f(0) = 0$) and let $A = (DF)(0)$. We denote as usual by $\sigma_{u,c,s}$ the parts of the spectrum (eigenvalues) for which $\mathrm{Re}\lambda > 0$, $= 0$, < 0 respectively.

Denote the generalized eigenspaces of $\sigma_{u,c,s}$ by $E^{u,c,s}$ respectively. By definition, the stable manifold is a set invariant under the flow which is tangent to E^s, the unstable one is tangent E^u whereas the center manifold *also invariant under the flow* is tangent to E^c.

We remember that, in hyperbolic systems (for which therefore the center manifold is absent) the stable/unstable manifolds are unique.

We see that center manifolds need not be unique (typically they are not) on the simple example (461),

$$x' = x^2 \tag{483}$$

$$y' = -y \tag{484}$$

Clearly $(0,0)$ is a non-hyperbolic fixed point, with 0 eigenvalue in the x direction. We have a unique stable manifold at $(0,0)$: here we look for an invariant set tangent to the vertical axis, and in this case it is the vertical axis itself. How about sets tangent to the center direction, $x = 0$? See Figure 18.2. We can solve for the trajectories

$$\frac{dy}{dx} = -\frac{y}{x^2} \tag{485}$$

i.e, $y = Ce^{1/x}$. We see that there is no such trajectory for $x > 0$, but *for all* C, the trajectories in the left half plane are tangent to the real line. Any of these would be a center manifold.

Theorem 18 (Center manifold theorem for flows)**.** *There exist C^r stable and unstable manifolds (invariant under the flow and tangent to E^s, E^u) W^s and W^u respectively, and these are unique. There is a (generally nonunique) center manifold W^c, and it is C^{r-1}.*

Corollary 40. *We can take a set of local coordinates, $\tilde{x}, \tilde{y}, \tilde{z}$, corresponding to the local splitting $\mathbb{R}^d = W^c \times W^s \times W^u$, so that, topologically, the general system is equivalent to*

$$\tilde{x}' = \tilde{f}(\tilde{x}) \tag{486}$$

$$\tilde{y}' = -\tilde{y} \tag{487}$$

$$\tilde{z}' = \tilde{z} \tag{488}$$

Let us take the special case when W^s is empty. We bring the linear part at the equilibrium of our general system to the block diagonal form

$$x' = Cx + f(x, y) \tag{489}$$

$$y' = Hy + g(x, y) \tag{490}$$

where C is the part of the matrix whose eigenvalues have zero real part while H is the rest of the matrix, the "hyperbolic" part. The center manifold is tangent to E^c, and we can thus write it in locally in the form of the graph of a function, $y = h(x)$. Indeed, at the equilibrium $Hy + g(x, y) = 0$ and H has no zero eigenvalue, thus the implicit function theorem applies. Substituting into (489) we get

$$x' = Cx + f(x, h(x)) \tag{491}$$

Q: Does this give us the center manifold?

On the other hand, $h(x) = o(x)$ for small x, since it $Dh = 0$ there. Thus, we expect, and shall prove later, that the flow provided by (491)) is a good approximation of $\tilde{x}' = \tilde{f}(\tilde{x})$, which would evolve inside the center manifold. The following holds.

Theorem 19 (Henry, Carr)**.** *If the origin $x = 0$ of* (491) *is locally asymptotically stable/unstable, then the origin of* (489) *is also locally asymptotically stable/unstable.*

19.1 The saddle-node bifurcation: general case

We follow [7]. We remember that the normal form we were aiming at was $a + x^2$, or, of course, more generally $\mu - \mu_0 \pm (x - x_0)^2$. Consider now the system (481), and assume that at $\mu = \mu_0, x = x_0$ there is an equilibrium in which one eigenvalue is zero and nondegenerate. The center manifold theorem would then allow us to reduce the study to the case where the system is one-dimensional. More precisely, there is a 2d center manifold Σ in $\mathbb{R}^n \times \mathbb{R}$ through (x_0, y_0) so that (1) Σ is tangent to the plane spanned by the 0 eigenvector and the direction of μ,

(2) For any r, Σ is C^r in a neighborhood of (x_0, y_0),

(3) The vector field of (481) is tangent to Σ,

and

(4) There is a neighborhood U of (x_0, y_0) in Σ which is invariant under the flow.

If we restrict (481) to Σ, we get a one-parameter family of equations on the one dimensional curves $\Sigma_\mu := \{z \in \Sigma : \mu = const =: \mu\}$. This is the reduction of the bifurcation problem. We now need to impose conditions that imply that the bifurcation type of this one-dimensional system is the same as that for the normal form $\mu - \mu_0 \pm (x - x_0)^2$. These are: $\frac{\partial f}{\partial \mu}(x_0, \mu_0) \neq 0$ (transversality in the μ direction), and $\frac{\partial^2 f}{\partial x^2}(x_0, \mu_0) \neq 0$, that is the equilibrium is quadratic.

More precisely, the following theorem holds.

Theorem 20. *Consider the setting above, under the following assumptions:*

(SN1) $M = D_x f(x_0, \mu_0)$ has a simple eigenvalue 0 with right eigenvector v and left eigenvector w ($wM = 0 \leftrightarrow M^T w = 0$). M has k eigenvalues with negative real parts and $(n - k - 1)$ with positive real parts.

(SN2) $w \cdot D_\mu f(x_0, \mu_0) \neq 0$.

(SN3) $w \cdot (v \cdot D_x^2 f(x_0, \mu_0) v) \neq 0$. (Note that $v \cdot D_x^2 f(x_0, \mu_0) v$ is a vector since f is a vector.)

Then there is a smooth curve of equilibria in $\mathbb{R}^n \times \mathbb{R}$ passing through (x_0, μ_0) and tangent to the hyperplane $\mathbb{R}^n \times \{\mu_0\}$. Depending on the signs in (SN1), (SN2) there are no equilibria near (x_0, μ_0) for $\mu < \mu_0$ ($\mu > \mu_0$ resp.). The two equilibria near (x_0, μ_0) are hyperbolic, and have stable manifolds of dimension k and $k + 1$, resp. The conditions (SN1) and (SN2) are generic, in the sense of forming an open dense set in the family of vector fields with an equilibrium with zero eigenvalue at (x_0, μ_0).

19.2 Transcritical and pitchfork bifurcations

We need appropriate changes in the assumptions. They are natural, if you think of the shape of the normal form:

(A) Transcritical bifurcation. Here we must have $f_\mu(0) = 0$ for all μ, and thus $D_\mu f$ cannot be nonzero anymore. This condition is replaced by (SN2') $w \cdot (\partial^2 f / \partial \mu \partial x) v \neq 0$ at $\mu = \mu_0$.

(B) Pitchfork bifurcation (one dimension). Here we are dealing with systems with symmetry in which f is odd. Thus, now we cannot have $D_x^2 f \neq 0$. Then, (SN3) is replaced by (SN3'), $D_x^3 f \neq 0$

Under these assumptions a theorem similar to the one in the previous section holds.

19.3 Hopf bifurcations

Consider now a system of the form (481) for which, at some (x_0, y_0) $D_x f$ has exactly one pair of nonzero imaginary eigenvalues, and the systems is hyperbolic otherwise, near (x_0, y_0). Then, by the implicit function theorem, the equilibrium position varies smoothly with μ, unlike in most other bifurcations. We expect however, by looking at what we called the normal form, a qualitative change in the structure of the equilibrium to occur at μ_0: a spiral sink is transformed into a spiral source plus a limit cycle.

By changes of variables (straightforward but rather lengthy [7]), the block affected by the bifurcation can be brought to the form

$$x' = (d\mu + a(x^2 + y^2)x - (\omega + c\mu + b(x^2 + y^2))y + \text{higher order terms} \quad (492)$$

$$y' = (\omega + c\mu + b(x^2 + y^2)x) + (d\mu + a(x^2 + y^2))y + \text{higher order terms} \quad (493)$$

(essentially, the quadratic terms can be eliminated). If we momentarily discard the higher order terms, this takes the following form in polar coordinates

$$r' = (d\mu + ar^2)r \quad (494)$$

$$\theta' = (\omega + c\mu + br^2) \quad (495)$$

The phase portrait of (494) does not differ substantially from the one we used before, where br^2 was missing. If a, d are nonzero, then there are periodic orbits of the (x, y) system lying along the parabola $\mu = -ar^2/d$; the surface of periodic orbits has quadratic tangency with the plane $\mu = 0$ in $\mathbb{R}^2 \times \mathbb{R}$.

The Hopf bifurcation theorem essentially says that the higher order terms do not change this picture locally.

Theorem 21 (Hopf, 1942). *Suppose that the system $x' = f_\mu(x)$, $(x, \mu) \in \mathbb{R}^n \times \mathbb{R}$, has an equilibrium at (x_0, μ_0) so that the following properties are satisfied.*

(H1) $D_x f(\mu_0, x_0)$ has a unique pair of purely imaginary nonzero eigenvalues.

Then, there exists a smooth curve of equilibria $(x(\mu), \mu)$ with $x(\mu_0) = x_0$. The two eigenvalues which are imaginary at (x_0, μ_0), $\lambda(\mu)$ and $\overline{\lambda}(\mu)$ vary smoothly with μ.

Assume furthermore that

$$\frac{d}{d\mu}(\text{Re}(\lambda(\mu))\Big|_{\mu=\mu_0} = d \neq 0 \quad (496)$$

Then, there exists a unique three dimensional center manifold passing through (x_0, μ_0) in $\mathbb{R}^n \times \mathbb{R}$, and a smooth change of coordinates preserving the planes

$\mu = const.$ for which the Taylor expansion on the center manifold is given by (492). If $a \neq 0$, then there is a surface of periodic solutions in the center manifold which has quadratic tangency with the eigenspace of $\lambda(\mu_0)$ and $\overline{\lambda}(\mu)$ agreeing to second order with the paraboloid $\mu = -(a/d)(x^2+y^2)$. If $a < 0$, the periodic solutions are repelling.

20 Appendix

20.1 Solution to Exercise 3

The equation for Y_k is

$$kY_k + (Y_k J - JY_k) = R_k + \sum_{j=1}^{k-1} Y_j A_{k-j-1}; \quad R_k = A_{k-1}J \tag{497}$$

We consider the family of Banach spaces indexed by $\mu > 0$,

$$\mathcal{B}_\mu = \{\mathsf{Y} = (Y_l)_{l\in\mathbb{N}} : |||\mathsf{Y}|||_\mu := \sup_{j\in\mathbb{N}} \mu^{-j}\|Y_j\| < \infty\}$$

Note that, since $A(z)$ is analytic, the series $\sum_{l\in\mathbb{N}} A_l z^l$ converges, implying that, for some $C > 0$,

$$\sup_{j\in\mathbb{N}} \|A_j C^{-j}\| < \infty \tag{498}$$

Thus the vector $\mathsf{R} := (R_l)_{l\in\mathbb{N}}$ is in $\mathcal{B}_\mu$ for all $\mu > C$.

The function $\mathcal{C}$ given by $\mathcal{C}X = XJ - JX$ is evidently a linear function on $\mathbb{C}^{n^2}$, thus given by a matrix; since $\|\mathcal{C}X\| \leqslant 2\|J\| \|X\|$ by the triangle inequality, its norm is bounded by

$$\|\mathcal{C}\| \leqslant 2\|J\| \tag{499}$$

The function M_k given by

$$M_k X =: kX + \mathcal{C}X$$

is a linear function on $\mathbb{C}^{n^2}$, and thus it is also given by a matrix. We have shown that M_k is invertible, since $M_k X = 0 \Leftrightarrow X = 0$. Thus, for every k, M_k^{-1} exists (and evidently has finite norm).

We now also note that, if $k > 2\|J\|$ we have

$$\|M_k\|^{-1} \leqslant \frac{1}{k - 2\|J\|} \tag{500}$$

Indeed,

$$M_k^{-1} = k^{-1}(1 - k^{-1}\mathcal{C})^{-1} \tag{501}$$

Thus the series

$$\sum_{l=0}^{\infty} \mathcal{C}^l / k^l \tag{502}$$

converges for all $k > 2\|J\|$. This is called a Neumann series, and you can check that it converges to $(1-k^{-1}\mathcal{C})^{-1}$.

Thus,

$$\|(1-k^{-1}\mathcal{C})^{-1}\| \leqslant \sum_{l=0}^{\infty} k^{-l}(2\|J\|)^l = \frac{1}{1-2k^{-1}\|J\|} \tag{503}$$

and (500) follows.

Therefore,

$$\begin{aligned}\sup_{k\in\mathbb{N}} \|M_k^{-1}\| &= \max\{\max_{k\leqslant 2\|J\|+1} \|M_k^{-1}\|, \sup_{k\geqslant 2\|J\|+2} (1-2k^{-1}\|J\|)^{-1}\} \\ &= \max\{\max_{k\leqslant 2\|J\|+1} \|M_k^{-1}\|, 1/2\} = a_1 < \infty \end{aligned}\tag{504}$$

Then the operator $\hat{\mathsf{T}}$ defined by

$$(\hat{\mathsf{T}}\mathsf{Y})_j = M_j^{-1} Y_j \tag{505}$$

is bounded in $\mathcal{B}_\mu$, and

$$\|\hat{\mathsf{T}}\| = a_1 \tag{506}$$

We define the (linear) operator $\hat{\mathsf{L}}$ on $\mathcal{B}_\mu$, $\mu > C$, by

$$(\hat{\mathsf{L}}\mathsf{Y})_j = \sum_{j=1}^{k-1} Y_j A_{k-j-1}; \quad j \geqslant 1 \tag{507}$$

This is well defined on $\mathcal{B}_\mu$ and

$$\|\hat{\mathsf{L}}\| \leqslant \frac{1}{\mu - C} \tag{508}$$

Indeed, since $\|Y\|_j \leqslant \mu^j |||Y||| =: N\mu^j$, we have

$$\|\sum_{j=1}^{k-1} Y_j A_{k-j-1}\| \leqslant N \sum_{j=1}^{k-1} \mu^j C^{k-j-1} \leqslant NC^{k-1} \frac{\mu^k}{C^k(\mu/C-1)} = \frac{N\mu^k}{\mu - C} \tag{509}$$

Now, the system (497) can be written compactly as

$$\mathsf{Y} = \hat{\mathsf{T}}\mathsf{A} + \hat{\mathsf{T}}\hat{\mathsf{L}}\mathsf{Y} \tag{510}$$

This is a linear nonhomogeneous equation for Y. For it to be contractive, we need $\|\hat{\mathsf{T}}\hat{\mathsf{L}}\| \leqslant \|\hat{\mathsf{T}}\|\ \|\hat{\mathsf{L}}\| < 1$.

This is the case if

$$\frac{a_1}{\mu - C} < 1 \tag{511}$$

i.e., if $\mu > \mu_1 = C + a_1$. Thus $\mathsf{Y} \in \mathcal{B}_{\mu_1}$, implying that $\|Y_j\| \leqslant N\mu_1^j$ for some N and all j, and therefore the series

$$\sum_{j=1}^{\infty} Y_j z^j \tag{512}$$

converges (obviously to an analytic function) for $|z| < 1/\mu_1$, and therefore $Y(z)$ is analytic at zero as required.

20.2 Solution to Exercise 1

The definition of z^{aP} is $\exp(aP\ln z)$ Now, since $P^2 = P$ we have

$$\exp(P\ln z) = I + \sum_{k=1}^{\infty}(a\ln z)^k P^k = I + P\sum_{k=1}^{\infty}(a\ln z)^k/k!$$
$$= I + P(z^a - 1) = Pz^a + (I - P) \quad (513)$$

References

[1] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, (1955).

[2] R.D. Costin, Nonlinear perturbations of Fuchsian systems, *Nonlinearity* 9, pp. 2073–2082 (2008).

[3] H. Dulac, "Solution d'un systéme d'équations differéntiélles dans le voisinage des valeurs singuliérs", *Bull. Soc. Math. France* **40** (1912), 324-383

[4] , G. Gaeta, Giuseppe "Resonant normal forms as constrained linear systems", Modern Phys. Lett. A 17 (2002), no. 10, 583597.

[5] D. Kazhdan, B. Kostant and S. Sternberg, "Hamiltonian group actions and dynamical systems of Calogero type" *Comm. Pure Appl. Math.* **31** (1978), 481-508

[6] J. Ecalle and B. Valet Correction and linearization of resonant vector fields and diffeomorphisms, *Mathematische Zeitschrift*, 2, pp. 249–318 (1998)

[7] , J. Guckenheimer and P. Holmes, Nonlinear oscillations, dynamical systems and bifurcations of vector fields, Springer Verlag (1986).

[8] M.W.Hirsch, S. Smale and R.L. Devaney, *Differential Equations, Dynamical Systems & An Introduction to Chaos*, Academic Press, New York (2004).

[9] Y. Ilyashenko and S. Yakovenko, Lectures on Analytic Differential Equations, Graduate Studies in Mathematics, vol. 86, ISBN: 0-8218-3667-6 (2007).

[10] D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation theory*, Academic Press, Ney York , (1989)

[11] S. Walcher, "On differential equations in normal form", *Math. Ann.* **291** (1991), 293-314

21 Extended bibliography

[12] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing. New York: Dover, pp. 804–806, (1972).

[13] L. Ahlfors, *Complex Analysis*, McGraw-Hill, New York, (1979).

[14] V.I. Arnold *Geometrical Method in the Theory of Ordinary Differential Equations*, Springer-Verlag (1988).

[15] V.I. Arnold *Mathematical Methods of Classical Mechanics*, Springer, (1989).

[16] M. Aschenbrenner and L. van den Dries, Asymptotic differential algebra, in *Analyzable Functions and Applications, pp. 49–85*, Contemp. Math., 373, Amer. Math. Soc., Providence, RI, (2005).

[17] W. Balser, B.L.J. Braaksma, J-P. Ramis and Y. Sibuya, *Asymptotic Anal. 5, no. 1, pp. 27–45*, (1991).

[18] W. Balser, *From Divergent Power Series to Analytic Functions*, Springer-Verlag, (1994).

[19] C. Bender and S. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, (1978), Springer-Verlag, (1999).

[20] M.V. Berry and C.J. Howls, *Proceedings: Mathematical and Physical Sciences*, vol. 430, No. 1880, pp. 653–668, (1990).

[21] M.V. Berry and C.J. Howls, *Proc. Roy. Soc. London* Ser. A 443, no. 1917, pp. 107–126, (1993).

[22] M.V. Berry, *Proc. R. Soc. Lond.* A 422, pp. 7–21, (1989).

[23] M.V. Berry, *Proc. R. Soc. Lond.* A 430, pp. 653–668, (1990).

[24] M V. Berry, *Proc. Roy. Soc. London* Ser. A 434 no. 1891, pp. 465–472, (1991).

[25] E. Borel, *Leçons sur les Series Divérgentes*, Gauthier-Villars, Paris, (1901).

[26] B.L.J. Braaksma, *Ann. Inst. Fourier*, Grenoble, 42, 3, pp. 517–540, (1992).

[27] A. Cauchy, *Oeuvres Completes d'Augustin Cauchy*, publiées sous la direction scientifique de l'Academie de sciences et sous les auspices du m. le ministre de l'instruction publique. Gauthier-Villars, Paris, (1882–90).

[28] J. D. Cole, *On a quasilinear parabolic equation occurring in aerodynamics*, Quart. Appl. Math., Vol. 9, No. 3, pp. 225236, (1951).

[29] M.D. Kruskal and P.A. Clarkson, *Studies in Applied Mathematics*, 86, pp 87–165, (1992).

[30] E. Hille *Ordinary differential equations in the complex domain*, John Wiley & sons, 1976

[31] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, (1955).

[32] Costin O., Asymptotics and Borel summability, Chapmann & Hall, New York (2009). ++

[33] O. Costin, R.D. Costin and M. Huang *Asymptotically conserved quantities for nonlinear ODEs and connection formulae* preprint, arXiv:1205.0775.

[34] O. Costin, M. Huang and F. Fauvet, *Global Behavior of Solutions of Nonlinear ODEs: First Order Equations* Int Math Res Notices first published online October 31, 2011 doi:10.1093/imrn/rnr203 ++

[35] O. Costin and M.D. Kruskal, *Proc. R. Soc. Lond. A*, 452, pp. 1057–1085, (1996).

[36] O. Costin and M.D. Kruskal, *Comm. Pure Appl. Math.*, 58, no. 6, pp. 723–749, (2005).

[37] O. Costin and M. D. Kruskal, *Proc. R. Soc. Lond. A*, 455, pp. 1931–1956, (1999).

[38] NIST Digital Library of Mathematical Functions http://dlmf.nist.gov/

[39] P A Deift and X Zhou *A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation* Ann. of Math. (2) 137 no 2 pp. 295–368 (1993).

[40] H Flashka and A C Newell *Monodromy and spectrum preserving transformations* Commun. Math. Phys. **76** pp. 65–116 (1980).

[41] A.S. Fokas *Invariants, Lie-Bäcklund Operators, and Bäcklund transformations* Ph.D. Thesis (1979).

[42] A. P. Fordy and A. Pickering, Analysing negative resonances in the Painlev test,. Phys. Lett. A 160 347-354 (1991).

[43] L Fuchs *Sur quelques équations différentielles linéaires du second ordre* C. R. Acad. Sci., Paris **141** pp. 555-558 (1905).

[44] E. Hopf, The partial differential equation $u_t + uu_x = \mu u_{xx}$, Comm. Pure and Appl. Math., Vol. 3, pp. 201230 (1950).

[45] A R Its, A S Fokas and A A Kapaev *On the asymptotic analysis of the Painlevé equations via the isomonodromy method* Nonlinearity 7 no. 5, pp. 1921–1325 (1994).

[46] A. V. Kitaev, *Elliptic asymptotics of the first and the second Painlev transcendents* Russ. Math. Surv. 49 81, (1994).

[47] S.V. Kovalevskaya *Sur le problème de la rotation d'un corps solide autour d'un point fixe*, Acta Math., **12** H.2, pp. 177–232 (1889). CRM Proc. Lecture Notes, 32, *The Kowalevski property* (Leeds, 2000), 315372, Amer. Math. Soc., Providence, RI, 2002

[48] S.V. Kovalevskaya, *Sur une propriété du système déquations differentielles qui definit la otation d'un corps solide autour d'un point fixe*, Acta Mathematica, vol. 14, pp. 81–93 (1890).

[49] S.V. Kovalevskaya *Mémoire sur un cas particulier du problème d'un corps pesant autour d'un point fixe, où l'integration s'effectue à l'aide de fonctions ultraelliptiques du temps* Mémoires présentés par divers savants à l'Academie des Sciences de l'Institut National de France, Paris 31 pp. 1–62 (1890).

[50] P Painlevé *Mémoire sur les équations différentielles dont l'integrale générale est uniforme* Bull. Soc. Math. France **28** pp 201–261 (1900).

[51] P Painlevé *Sur les équations différentielles du second ordre et d'ordre supérieur dont l'integrale générale est uniforme* Acta Math. **25** pp. 1–85 (1902).

[52] O. Costin and R.D. Costin, SIAM J. Math. Anal. 27, no. 1, pp. 110–134, (1996).

[53] O. Costin and S. Tanveer, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 24, no. 5, pp. 795–823, (2007).

[54] O. Costin, *Analyzable functions and applications*, pp. 137–175, Contemp. Math., 373, Amer. Math. Soc., Providence, RI, (2005).

[55] O. Costin, *Duke Math. J.*, 93, No. 2, (1998).

[56] O. Costin and R.D. Costin, *Inventiones Mathematicae*, 145, 3, pp. 425–485, (2001).

[57] O. Costin, *IMRN* 8, pp. 377–417, (1995).

[58] O. Costin, G. Luo and S. Tanveer, *Integral Equation representation and longer time existence solutions of 3-D Navier-Stokes, submitted.* http://www.math.ohio-state.edu/∼tanveer/ictproc.ns.5.pdf (last accessed: August 2008).

[59] O. Costin and S. Garoufalidis *Annales de L'institut Fourier*, vol. 58 no. 3, pp. 893–914, (2008).

[60] E. Delabaere and C.J. Howls, *Duke Math. J.* 112, no. 2, 199–264, (2002).

[61] R.B. Dingle, *Asymptotic Expansions: Their Derivation and Interpretation*, Academic Press, London-New York, (1973).

[62] N. Dunford and J.T. Schwartz, *Linear Operators, Part I: General Theory*, Interscience, New York, (1960).

[63] J. Écalle and F. Menous, *Publicacions Matemàtiques*, vol. 41, pp. 209–222, (1997).

[64] J. Écalle, Preprint 90-36 of Université de Paris-Sud, (1990).

[65] J. Écalle, *Fonctions Resurgentes*, Publications Mathématiques D'Orsay, (1981).

[66] J. Écalle, in *Bifurcations and periodic orbits of vector fields*, NATO ASI Series, vol. 408, (1993).

[67] G.A. Edgar, *Transseries for Beginners*, preprint, (2008). http://www.math.ohio-state.edu/~ edgar/WG_W08/edgar/transseries.pdf (last accessed: August 2008).

[68] L. Euler, De seriebus divergentibus, *Novi Commentarii academiae scientiarum Petropolitanae* (1754/55) 1760, pp. 205–237, reprinted in *Opera Omnia Series*, I vol. 14, pp. 585–617. Available through The Euler Archive at www.EulerArchive.org (last accessed, August 2008).

[69] E. Goursat, *A Course in Mathematical Analysis, vol. 2: Functions of a Complex Variable & Differential Equations*, Dover, New York, (1959).

[70] C.G. Hardy, *Divergent Series*, Oxford, (1949).

[71] E. Hille *Ordinary Differential Equations in the Complex Domain*, Dover Publications (1997).

[72] E. Hille and R.S. Phillips, *Functional Analysis and Semigroups*, AMS Providence, R.I. (1957).

[73] J. van der Hoeven, *Transseries and Real Differential Algebra*, Lecture Notes in Mathematics 1888, Springer, New York, (2006).

[74] A.S.B. Holland, *Introduction to the Theory of Entire Functions*, Academic Press, US, (1973).

[75] I. Kaplansky, *Differential Algebra*, Hermann, Paris, (1957).

[76] T. Kawai and Y. Takei, *Advances in Mathematics*, 203, 2,pp. 636–672 (2006).

[77] D.A. Lutz, M. Miyake and R. Schäfke, *Nagoya Math. J.* 154, 1, (1999).

[78] J. Martinet and J-P. Ramis, *Annales de l'Institut Henri Poincaré (A) Physique théorique*, 54 no. 4, pp. 331–401, (1991).

[79] A.B. Olde Daalhuis, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* 454, no. 1968, pp. 1–29, (1998).

[80] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Academic Press, New York, (1972).

[81] J-P. Ramis, *Asterisque*, V. 59-60, pp. 173–204, (1978).

[82] J-P. Ramis, *C.R. Acad. Sci. Paris*, tome 301, pp. 99–102, (1985).

[83] J.F. Ritt *Differential Algebra*, American Mathematical Society, New York, (1950).

[84] W. Rudin, *Principles of Mathematical Analysis* p. 110, McGraw-Hill, New York, (1974).

[85] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, (1987).

[86] D. Sauzin, *RIMS Kokyuroku* 1493, pp. 48–117, (31/05/2006). http://arxiv.org/abs/0706.0137v1. (last accessed: August 2008).

[87] Y. Sibuya, *Global Theory of a Second Order Linear Ordinary Differential Equation with a Polynomial Coefficient*, North-Holland, Amsterdam-New York, (1975).

[88] A. Sokal, *J. Math. Phys.*, 21 pp. 261–263, (1980).

[89] G.G. Stokes *Trans. Camb. Phil. Soc.*, 10, pp. 106–128. Reprinted in *Mathematical and Physical papers by late sir George Gabriel Stokes,* Cambridge University Press, vol. IV, pp. 77–109, (1904).

[90] P. Suppes, *Axiomatic Set Theory*, Dover, (1972).

[91] E.C. Titchmarsh, *The Theory of Functions*, Oxford University Press, USA; 2nd edition, (1976).

[92] J. Weiss, M. Tabor and G. Carnevale, *The Painlevé property for partial differential equations* J. Math. Phys. 24, 522 (1983).

[93] W. Wasow *Asymptotic Expansions for Ordinary Differential Equations*, Interscience Publishers, New York, (1968).

[94] J. Weiss, M. Tabor, and G. Carnevale *The Painlev property for partial differential equations*, J. Math. Phys. 24, 522 (1983).

[95] G.B. Whitham, *Linear and Non-Linear waves*, Wiley, N.Y. (1974).

[96] E. Zermelo, *Mathematische Annalen*, 65, pp. 261–281, (1908).