Chapter 3

Bifurcation Theory (the minimum)

Continuing the analysis of the previous section we would like to put it on a more systematics ground: we worried only about hyperbolic fixed points; are more complex situations relevant? To answer to such questions it is first necessary to understand their meaning, that is: what does it mean for a scenario to be irrelevant.

3.1 Generic Vector fields

Let us consider a first order autonomous differential equation,

$$\dot{x} = V(x) \tag{3.1.1}$$

where $V \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ and $x \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}^d)$. We are interested in the *typical local behavior* of such systems. Unfortunately, before being able to even address such an issue, it is necessary to give a technical meaning to the three words *typical, local* and *behavior*.

3.1.1 Local behavior

We say that two vector fields have the same local *behavior* in $U \subset \mathbb{R}^d$ if they are equivalent (conjugated) in U. We thus say that we *undestand* the behaviour of a vector field if it is equivalent to a vector field that yields an ODE that can be explicitely solved.

By *local* understanding in a region K we mean that for each point $x \in K$ we are able to consider some neighborhood of x in which we understand the solutions of (3.1.1).[1]

If we could consider only neighborhoods U in which $V(x) \neq 0$ with, at most, the exception of one point where the linear part is hyperbolic, then we understand already the typical local behavior. In fact, either $V(\bar{x}) \neq 0$ and then the flow box Theorem tells us that the field has the same local behavior than a constant vector field; or, if $V(\bar{x}) = 0$, then Grobmann-Hartman Theorem tells us that the field has the same local behavior than its linear part.

Of course, this is not always possible (think of the case $V \equiv 0$), our claim is that *typically* it is.

3.1.2 Typical

To define *typical* let us consider the following.

Definition 3.1.1 *Given a topological space Ω, we say that a set $A \subset \Omega$ is* generic *if it is open and dense.*

Now $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ is a Banach space hence the topology is trivially determined by the norm. For now on *typical* will mean that it happens for a countable intersection of *generic* sets (this is also called a *residual* set).

Problem 3.1 *Prove that a residual set is dense.*

Problem 3.2 *Give an example of a typical set in $\mathbb{R}$ with zero Lebesgue measure.*

Next, for each $K \subset \mathbb{R}^d$, let us define

$$A_K := \{V \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n) \, : V(x) = 0 \text{ implies } \partial_x V \text{ hyperbolic } \forall\, x \in K\}$$

Remark 3.1.2 *In the following we will prove that, for K compact, A_K is generic, hence $A_{\mathbb{R}^d}$ is typical. Note that the same holds for*

$$\{V \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n) \, : V(x) = 0 \textit{ implies } \det(\partial_x V) \neq 0 \forall\, x \in K\}.$$

[1]Note that, if K is compact, then finitely many such neighborhoods will cover K. On the other hand if, for example, $K = \mathbb{R}^d$, then countably many neighborhoods will do the job.

Yet, the goal is to find the smallest possible generic set (see Problems 3.22, 3.23). This allows to obtain a generic understanding with the least effort.

Problem 3.3 *Prove that, for each compact set $K \subset \mathbb{R}^d$, if $V \in A_K$, then V has only finitely many zeroes in K.*

Problem 3.4 *Prove that, for each compact set $K \subset \mathbb{R}^d$, A_K is open.*

To prove that A_K is generic we need to establish the density, this is not entirely obvious and we need a result of independent interest.

Theorem 3.1.3 (Sard–baby version) *Let $F \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, and $A = \{x \in \mathbb{R}^d \ : \ \det(D_x F) = 0\}$, then $F(A)$ has zero Lebesgue measure.*

Proof. Let $Q_\delta(x) := \{z \in \mathbb{R}^d \ : \ |x_i - z_i| \leq \delta \quad \forall i \in \{1, \dots, d\}\}$, clearly it suffices to prove that for each $x \in \mathbb{R}^d$ the Lebesgue measure of $F(A \cap Q_1(x))$ is zero. Now, for each $n \in \mathbb{N}$ and $k \in \{-n, \dots, 0, \dots, n\}^d =: S_n$, let $x_k := \frac{1}{n} k$ and $\Delta_k := Q_{1/n}(x + x_k)$. Clearly $Q_1(x) = \cup_{k \in S_n} \Delta_n$. Of course we will consider only the Δ_k such that $\Delta_k \cap A \neq \emptyset$. For each such Δ_k let us chose $\xi_k \in \Delta_k \cap A$.

Next, for each $y \in \Delta_k$, $F(y) - F(\xi_k) = D_{\xi_k} F(y - \xi) + o(\|y - \xi\|)$. Since $\xi_k \in A$, there exists $v_k \in \mathbb{R}^d$, $\|v_k\| = 1$, such that $\langle v_k, D_{\xi_k} F w \rangle = 0$ for all $w \in \mathbb{R}^d$. Hence, setting $C = \|DF\|_\infty$ and for n large enough, $F(\Delta_k) \subset \{w + t v_k \in \mathbb{R}^d \ : \ \langle w, v_k \rangle = 0; \ \|w\| \leq 2Cn; \ |t| = o(\frac{1}{n})\}$.

Thus, calling λ the Lebesgue measure,

$$\lambda(F(\Delta_k)) \leq 4^{d-1} C^{d-1} n^{-d-1} o(n^{-1}) = \lambda(\Delta_k) \cdot o(1).$$

Thus

$$\lambda(F(A \cap Q_1(x))) \leq \sum_{k \in S_n} \lambda(\Delta_k) o(1) = 2^d o(1),$$

as announced. □

Problem 3.5 *Use Sard's Theorem to show that, for each compact set $K \subset \mathbb{R}^d$, A_K is dense in $\mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$. Prove that $A_{\mathbb{R}^d}$ is typical.*

Definition 3.1.4 *We say that two vector fields V, W are equivalent in the open set U, if, for each $t > 0$, there exists a homeomorphism $F : U \to U$ such that, calling ϕ_t^V, ϕ_t^W the flows generated by the vector fields, holds $\phi_t^v \circ F = F \circ \phi_t^W$.*

3.2 Generic families of vector fields

Our next aim is to consider a situation in which the system has a control parameter. That is, it is described by the equations

$$\dot{x} = V(x, \lambda) \tag{3.2.2}$$

where $x \in \mathbb{R}^d$ and $\lambda \in [-2, 2]$ is the parameter that, in principle, can be varied. Now by *local* understanding in a region K we mean that for each point $(\bar{x}, \bar{\lambda}) \in K \times [-1, 1] =: K^1$ we can find a neighborhood of the form $U \times (\lambda - \varepsilon, \lambda + \varepsilon)$ in which we are able to understand the behavior of the solutions of (3.2.2).

Let us now try to understand the local picture for typical families of vector fields. In analogy with the previous section, for any $K \subset \mathbb{R}^d$, let us consider

$$\begin{aligned} \bar{A}_K := \{V \in \mathcal{C}^1 \; : & \; V(x, \lambda) = 0 \text{ implies } \partial_x V(x, \lambda) \text{ hyperbolic} \\ & \forall\, (x, \lambda) \in K^1\} \end{aligned}$$

Problem 3.6 *Prove that if $V \in \bar{A}_K$, then for each $(\bar{x}, \bar{\lambda}) \in K \times [-1, 1]$ there exists an open set of the form $U \times (-\varepsilon + \bar{\lambda}, \varepsilon + \bar{\lambda}) =: U \times I$ such that either $V(x, \lambda) \neq 0$ or there exists $X \in \mathcal{C}^1(I, K)$ such that $V(X(\lambda), \lambda) = 0$ for each $\lambda \in K$ and there are no other zeroes in $U \times I$. Then, prove that $\bar{A}_K$ is open.*

Clearly the above situations can be treated exactly as we did in the previous section and are therefore locally understandable. Unfortunately, the above does not exhaust all the possibilities.

Lemma 3.2.1 *For each K with non empty interior $\bar{A}_K$ is not generic.*

PROOF. Since $\bar{A}_K$ is open, the problem must be the density. To see this let us consider, for example, the case $d = 1$, a compact set K with interior containing zero and the family

$$V(x, \lambda) = \lambda a + \lambda x + bx^2. \tag{3.2.3}$$

Now let us consider any $W \in \mathcal{C}^1(\mathbb{R} \times [-1, 1], \mathbb{R})$ and look at $\tilde{V}(x, \lambda, \mu) := V(x, \lambda) + \mu W(x, \lambda)$. The claim is that for each μ sufficiently small, then $\tilde{V}(x, \lambda, \mu) \notin \bar{A}_K$. In fact, there exists $(x(\mu), \lambda(\mu)) \in K$ such that both

$\tilde{V}(x(\mu), \lambda(\mu), \mu) = 0$ and $\partial_x \tilde{V}(x(\mu), \lambda(\mu), \mu) = 0$. To see this we define the function $F : \mathbb{R}^3 \to \mathbb{R}^2$

$$F(x, \lambda, \mu) := \begin{pmatrix} \lambda a + \lambda x + bx^2 + \mu W(x, \lambda) \\ \lambda + 2bx + \mu \partial_x W(x, \lambda) \end{pmatrix} = \begin{pmatrix} \tilde{V} \\ \partial_x \tilde{V} \end{pmatrix},$$

clearly we are looking for $(x(\mu), \lambda(\mu))$ such that $F(x(\mu), \lambda(\mu)) = 0$. Since $F(0, 0, 0) = 0$ we can apply the implicit function theorem provided

$$\begin{pmatrix} 0 & a \\ 2b & 1 \end{pmatrix}$$

is invertible, that is if $ab \neq 0$. We have thus seen that the family has an open neighborhood disjoint from $\bar{A}_K$, hence the latter set cannot be dense. □

Thus, to have a generic situation we need to consider a larger set.

Looking at the above example it is natural to ask that $\partial_\lambda V \neq 0$ if $\det(\partial_x V) = 0$. This is a good idea but it does not suffice to have a nice theory. As we will see shortly it is useful to add some condition on the second derivative. It is then natural to consider at least $\mathcal{C}^2$ vector fields, since we will see in the following that higher derivatives can play a role, we will consider $\mathcal{C}^r$ vector fields, $r \geq 2$. Accordingly from now on the genericity will be according to the $\mathcal{C}^r$ topology. This would not have changed the previous discussion, see 3.29.

The above can be made precise in many ways, before stating a simple possibility we need a bit of notation.

Given a $d \times d$ matrix an a vector $w \in \mathbb{R}^d$ we consider the $d \times (d+1)$ matrix $(A\ w)$, by rank $(A\ w)$ we mean the number of linearly independent column. Also let us call $Z(A)$ the number of eigenvalues, counted with multiplicity, with zero real part and Π_A^0 the eigenprojector on the eigenspace associate to the eigenvalues with zero real part. Finally by $\mathrm{Tr}(A)$ we denote the trace of A. For $K \subset \mathbb{R}^d$ let $K^1 := K \times [-1, 1]$.

$$\begin{aligned} B_K = \Big\{ V \in \mathcal{C}^r\ &: \forall\, (x, \lambda) \in K^1\ V(x, \lambda) = 0 \Longrightarrow \big[\mathrm{rank}\,(\partial_x V\ \partial_\lambda V) = d\ ; \\ &\det(\partial_x V) \neq 0 \Longrightarrow (Z(\partial_x V) \leq 2, \text{ and} \\ &Z(\partial_x V) = 2 \Longrightarrow \mathrm{Tr}(\Pi^0_{\partial_x V}(\partial_{xx} V (\partial_x V)^{-1} \partial_\lambda V - \partial_{x\lambda} V)) \neq 0)\ ; \\ &\partial_x V v = 0 \Longrightarrow \mathrm{rank}\, \big(D_x V\ \partial_x^2 V(v, v)\big) = d\big] \Big\} \end{aligned}$$

Let us understand how the vector fields in B_K look like.

Lemma 3.2.2 *If $V \in B_K$ and $V(\bar{x}, \bar{\lambda}) = 0$, then there exists $\varepsilon > 0$ and a neighborhood $U \ni \bar{x}$ such that the set of zeroes of the vector field $V(x, \lambda)$ in $U \times (\bar{\lambda} - \varepsilon, \bar{\lambda} + \varepsilon)$ consists of a smooth curve. Moreover, the vector fields in $(\bar{\lambda} - \varepsilon, \bar{\lambda} + \varepsilon)$ belong to $A_{\bar{U}}$ with, at most, one exception.*

PROOF. First suppose, without loss of generality, that $(\bar{x}, \bar{\lambda}) = (0, 0)$.

If $\det(\partial_x V(0,0)) \neq 0$, then we can argue as in Problem 3.4. The implicit function theorem yields $\varepsilon > 0$ a neighborhood U of zero and $x \in \mathcal{C}^r([-\varepsilon, \varepsilon], \mathbb{R}^d)$ such that $V(x(\lambda), \lambda) = 0$ are the only zeroes of the vector fields $V(\cdot, \lambda)$, $\lambda \in [-\varepsilon, \varepsilon]$, in U. If $V(\cdot, 0) \notin A_{\overline{U}}$, hence $Z(A(0)) = 2$, then we have to show that is the only one not in $A_{\overline{U}}$. Let $A(\lambda) := \partial_x V(x(\lambda), \lambda)$, then by perturbation theory if ε is small enough all the eigenvalues of $A(0)$ with real part different from zero will stay so. The problem reduces then to show that the two eigenvalues with zero real part must acquire a non zero real part for $\lambda \neq 0$. Problem 3.26 shows that $Z(A(\lambda)) = 2$ implies $\mathrm{Tr}(\Pi^0_{A(\lambda)} A(\lambda)) = 0$. Hence if the λ for which $Z(A(\lambda)) = 2$ accumulate to zero, then $\mathrm{Tr}(\Pi^0_{A(0)} \dot{A}(0)) = 0$ (see Problem 3.27) which contradicts the hypothesis $V \in B_K$.

On the contrary, if $\det(\partial_x V(0,0)) = 0$ then the approach based on a direct application of the implicit function theorem fails. The problem is that the curve of the fixed points it is not a graph over λ so one need to change variables before applying the implicit functions theorem, let us see how.

The null space of $\partial_x V(0,0)$ must have dimension one, otherwise $\mathrm{rank}\,(\partial_x V(0,0)\ \ \partial_\lambda V(0,0)) < d$, let $v \in \mathbb{R}^d$, $\|v\| = 1$, be the unique vector such that $\partial_x V(0,0)v = 0$. Consider a vector $v \in \mathbb{R}^d$, $\|v\| = 1$, and the change of variables $(\lambda, x) = F_v(\xi, \tau)$ defined by

$$\begin{aligned} x &= \xi - \tau v \\ \lambda &= \langle \xi, v \rangle. \end{aligned} \tag{3.2.4}$$

It is easy to check that F^{-1} is defined by

$$\begin{aligned} \tau &= \lambda - \langle x, v \rangle \\ \xi &= \lambda v + x - \langle x, v \rangle v. \end{aligned}$$

Then define the field $\tilde{V} := V \circ F$. Since $F(0,0) = 0$, $\tilde{V}(0,0) = 0$. To apply the implicit function theorem in the new variable we need $\partial_\xi \tilde{V}$ to be invertible, but $\partial_\xi \tilde{V}(x,\lambda) = \partial_x V(x,\lambda) + \partial_\lambda V(x,\lambda) \otimes v$.[2] It follows that $\partial_\xi \tilde{V}(0,0)$ must be invertible, otherwise there would exists $w \in \mathbb{R}^d$ such that, for all $\eta \in \mathbb{R}^d$,

$$0 = \langle w, \partial_\xi \tilde{V}(0,0)\eta\rangle = \langle w, \partial_x V(0,0)\eta\rangle + \langle w, \partial_\lambda V(0,0)\rangle\langle v, \eta\rangle.$$

Choosing $\eta = v$ follows $\langle w, \partial_\lambda V(0,0)\rangle = 0$ and hence $\partial_x V(0,0)^T w = 0$. But this would mean that all the column of the rectangular matrix $(\partial_x V(0,0)\ \ \partial_\lambda V(0,0))$ are orthogonal to w contradicting the definition of B_K.

So we can apply the implicit function theorem and obtain (for ξ, τ in a neighborhood of zero) a $\mathcal{C}^1$ function $\xi(\tau)$ such that $\tilde{V}(\xi(\tau), \tau) = 0$, with $\xi'(\tau) = (\partial_x V + \partial_\lambda V \otimes v)^{-1}\partial_x V v$. Then, setting $(x(\tau), \lambda(\tau)) := F(\xi(\tau), \tau)$ we have a $\mathcal{C}^1$ curve and a neighborhood of zero in $\mathbb{R}^{d+1}$ such that $V(x(\tau), \lambda(\tau)) = 0$ and no other zero is present in the neighborhood. Note that

$$x'(\tau) = \frac{dx(\tau)}{d\tau} = -(\partial_x V + \partial_\lambda V \otimes v)^{-1}\partial_\lambda V. \tag{3.2.5}$$

To conclude, we note that if $\lambda(\tau)$ were invertible, then we could have parametrized the curve as $(x(\lambda), \lambda)$ without the above change of coordinates. It is then natural to investigate the points for which $\frac{d\lambda}{d\tau} = 0$.

$$\frac{d\lambda}{d\tau} = \langle \xi'(\tau), v\rangle = 1 - \langle (\partial_x V + \partial_\lambda V \otimes v)^{-1}\partial_\lambda V, v\rangle = 1 + \langle x'(\tau), v\rangle. \tag{3.2.6}$$

Note that $\partial_x V x' + \partial_\lambda V \langle v, x'\rangle = -\partial_\lambda V$. Thus, if $\frac{d\lambda}{d\tau} = 0$, then $\langle x', v\rangle = -1$, hence $\partial_x V x' = 0$, i.e. $x'(0) = -v$ and $\det(\partial_x V) = 0$. On the other hand, if $\det(\partial_x V) = 0$, then there exists $w \in \mathbb{R}^d$ such that $\partial_x V^T w = 0$. Multiplying the above relation by w yields $\langle w, \partial_\lambda V\rangle\langle v, x'\rangle = \langle w, \partial_\lambda V\rangle \neq 0$, since $V \in B_K$. Hence $\frac{d\lambda}{d\tau} = 0$.

To conclude the proof of the Lemma we must show that the points for which $\det(\partial_x V(x(\tau), \lambda(\tau)) = 0$ do not accumulate at zero. Indeed,

[2] Given two vectors $v, w \in \mathbb{R}^d$, by $v \otimes w$ we mean the matrix with elements $(v \otimes w)_{ij} = v_i w_j$.

if they accumulate at zero, then $\frac{d^2\lambda}{d\tau^2}(0) = 0$. But then (3.2.6) implies $\langle v, x''(0)\rangle = 0$ and, differentiating $\partial_x V x' + \partial_\lambda V \langle v, x'\rangle = -\partial_\lambda V$ yields $\partial_x^2 V(v, v) + \partial_x V x''(0) = 0$. This means $\langle w, \partial_x^2 V(v, v)\rangle = 0$ but his contradicts the requirement on the rank of $(\partial_x V \partial_x^2 V(v, v))$ contained in the definition of B_K. □

Problem 3.7 *Show that $B_{\mathbb{R}^d}$ is typical.*

Definition 3.2.3 *Given $V \in \mathcal{C}^r$ we call a point $(\bar{x}, \bar{\lambda}) \in \mathbb{R}^{d+1}$ such that $V(\bar{x}, \bar{\lambda}) = 0$ and $V(\cdot, \bar{\lambda}) \notin A_{\overline{U}}$, for all neighborhood U of $\bar{x}$, a* bifurcation point*. If $\partial_{xx} V(\bar{x}, \bar{\lambda}) \neq 0$, then we call the bifurcation point* regular*.*

Thus to achieve a typical local understanding of the behavior of one parameter families of vector fields we have to worry only about families with, at most, one regular bifurcation point. Let us suppose, without loss of generality, that the regular bifurcation point is at $(0, 0)$, then by Taylor expansion

$$V(x, \lambda) = Ax + b\lambda + \frac{1}{2}\langle x, B, x\rangle + \lambda C x + \mathcal{O}(\lambda^2) + o(\|x\|^2), \quad (3.2.7)$$

where B is a vector of $d \times d$ symmetric matrices.

Due to the previous discussion we need to consider only the following cases

a) A has a zero eigenvalue;

b) A has two purely imaginary conjugated eigenvalues.

Note that the condition $V \in B_{\mathbb{R}^d}$ imposes some conditions on b, B, C, we will detail them later as needed.

3.3 One dimension

In the one dimensional case (b) cannot take place. Then in (3.2.7) we have $A = 0$, $B \neq 0$.

Then $V(x, \lambda)$ has no solutions if $ac > 0$, while for $ac < 0$ there are the two solutions $x = \pm\sqrt{-\frac{\lambda b}{B}} + \mathcal{O}(\lambda)$. We have therefore the generic picture: either two points collide and kill each other or there is a creation of two zeroes of the vector field.

Problem 3.8 *Study the solutions of*

$$\dot{x} = \frac{B}{2}x^2 + g(x)$$

near zero when $g(0) = g'(0) = g''(0) = 0$.

Problem 3.9 *Prove that the two equilibrium points of the vector field are one attractive and the other repulsive.*

The above scenario is called a *saddle-node* bifurcation.

A natural question is if there exists a simpler standard form of the above bifurcation. Indeed we can try to kill some of the terms in 3.2.7 by a change of variable.

Problem 3.10 *Show that with a change of variables of the type* $x = \alpha\lambda + z$, $\mu = \rho\lambda$ *one can change the vector field* (3.2.7) *to the from* $\tilde{V}(z, \mu) = \mu + bz^2 + \mathcal{O}(\mu^2) + o(z^2)$.

The above is the *normal form* of the saddle node bifurcation. This type of reduction can be made for each bifurcation and give rise to the large field of normal form theory which, unfortunately, goes beyond the scopes of the present notes.

3.4 Two dimensions

3.4.1 A zero eigenvalue

In this case the vector field must have the form (possibly after a linear change of variable to put $\partial V_x(0,0)$ in normal form)

$$V(x, \lambda) = \begin{pmatrix} 0 & 0 \\ 0 & \nu \end{pmatrix} x + b\lambda + \frac{1}{2}\begin{pmatrix} \langle x, B_1 x\rangle \\ \langle x, B_2 x\rangle \end{pmatrix} + \lambda C x + \mathcal{O}(\lambda^2) + o(\|x\|^2), \tag{3.4.8}$$

with $b_1, B_1 \neq 0$. It easy to show that the scenario is exactly the same than in the one dimensional case. We leave the details to the reader.

3.4.2 Two purely imaginary conjugated eigenvalues: Hopf bifurcation

In this case the vector field must have the form (possibly after a linear change of variable to put $\partial V_x(0,0)$ in chosen form, see Problem 3.30)

$$V(x,\lambda) = Ax + R(x,\lambda), \tag{3.4.9}$$

with $A = \begin{pmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{pmatrix}$ for some $\omega_0 > 0$, $R(0,0) = \partial_x R(0,0) = 0$ and $\mathrm{Tr}(\partial_{xx} R(0,0) A^{-1} \partial_\lambda R(0,0) - \partial_{x\lambda} R(0,0)) \neq 0$.

In the above situation no new fixed point can appear, yet one expects something to happen. We will see that, depending on λ, a *periodic orbit* circling the fixed point is created. This is called an *Hopf bifurcation*.

To see how such an orbit is created some work is needed. To minimize it, we start by performing some changes of variables that reduces the ODE to a simpler one.

Problem 3.11 *Show that, with a change of coordinates of the type $x = \xi + \alpha(\lambda)$, the remainder R in (3.4.9) can be made to satisfy $R(0,\lambda) = 0$, for each λ small enough, $\partial_\xi R(0,0) = 0$ and $\mathrm{Tr}(\partial_{\xi\lambda} R(0,0)) \neq 0$.*

Problem 3.12 *Show that with a further change of variables $x = D(\mu)z$, $\lambda = \mu\rho(\mu)$ one can put (3.4.9) in the form*

$$\dot z = [\omega(\mu) J + \mu \mathbb{1}]\, z + R(z,\mu), \ \textit{where } J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{3.4.10}$$

with $\omega(0) = \omega_0$ and $R(0,\mu) = \partial_z R(0,\mu) = 0$.

Problem 3.13 *Find the solution of (3.4.10) in the case $R \equiv 0$.*

Given that the solutions of the linear part of (3.4.10) rotate around zero almost in circles, it may occur the idea to treat the problem in polar coordinates. In fact this point of view is quite advantageous and we will carry it out in order to show how a problem may simplify if viewed in different coordinates.

The polar coordinates can be written as $x = \rho v(\theta)$, where $\rho \in \mathbb{R}_+$, $\theta \in \mathbb{R}$ and $v(\theta) := (\cos\theta, \sin\theta)$.

Remark 3.4.1 *Note that such a change of coordinates is singular for* $\rho = 0$. *In addition, it is not globally one-one. Yet, to consider* θ *in the universal cover of* S^1 *rather than in* S^1 *will be very useful in the following.*

If we substitute such coordinates in (3.4.10), we obtain

$$\dot{\rho} v(\theta) + \rho n(\theta)\dot{\theta} = \mu\rho v(\theta) + \omega(\mu)\rho n(\theta) + R(\rho v(\theta), \mu),$$

where $n(\theta) := (-\sin\theta, \cos\theta)$. That is

$$\begin{aligned} \dot{\rho} &= \mu\rho + \langle v(\theta), R(\rho v(\theta), \mu)\rangle =: \mu\rho + a(\theta, \rho, \mu) \\ \dot{\theta} &= \omega_\mu + \rho^{-1}\langle n(\theta), R(\rho v(\theta), \mu)\rangle =: \omega(\mu) + b(\theta, \rho, \mu), \end{aligned} \tag{3.4.11}$$

where $a(\theta, 0, \mu) = \partial_\rho a(\theta, 0, \mu) = b(0, \mu) = 0$. In addition, note for later use that, $\partial_\rho^2 a(\theta, 0, 0)$ and $\partial_\rho b(\theta, 0, 0)$ are homogeneous trigonometric polynomials of degree three, of degree two, while $\partial_\rho^3 a(\theta, 0, 0)$ and $\partial_\rho^2 b(\theta, 0, 0)$ are of degree four. By 3.32 it follows that we can write $a(\theta, \rho, \mu) = a_0(\theta, \mu)\rho^2 + a_1(\theta, \rho, \mu)\rho^3$ and $b(\theta, \rho, \mu) = b_0(\theta, \mu)\rho + b_1(\theta, \rho, \mu)\rho^2$. Finally, the reader can easily verify that $a \in \mathcal{C}^r$, while $b \in \mathcal{C}^{r-1}$.

Note that the equation (3.4.11) is well defined also for $\rho = 0$ but in such a case, instead of a fixed point, it has the periodic orbit $(\rho(t), \theta(t)) = (0, \omega_0 t)$, this shows that the dynamics is indeed well adapted to these coordinates.

Since for small ρ we have $\dot{\theta} > 0$, it is convenient to use θ rather than t to parameterize the motion (here is now evident the advantage of using the universal cover of S^1). Calling again ρ the distance from the origin as a function of θ we have

$$\frac{d\rho}{d\theta} = \frac{\mu\rho + a(\theta, \rho)}{\omega + b(\theta, \rho)} =: \frac{\mu}{\omega}\rho + \beta(\theta, \mu)\rho^2 + \gamma(\theta, \rho, \mu)\rho^3, \tag{3.4.12}$$

where

$$\begin{aligned} \beta(\theta, \mu) &= \omega^{-1} a_0(\theta, \mu) - \mu\omega^{-2} b_0(\theta, \mu) \\ \gamma(\theta, 0, \mu) &= \mu\omega^{-3} b_0^2 + a_0 b_0 \omega^{-2} - \mu b_1 \omega^{-2} + a_1 \omega^{-1}. \end{aligned}$$

Note, that $\beta(\theta, 0)$ is a trigonometric homogeneous polynomial of third degree while $\gamma(\theta, 0, 0)$ is the sum of two monomial, one of degree four and one of degree six.

It is now convenient to perform a last change of variables: $\rho = \nu r$, $\mu = \pm\nu^2$, $\nu \geq 0$.[3] Under such changes of variables (3.4.12) becomes

$$\frac{dr}{d\theta} = \pm\frac{\nu^2}{\omega(\pm\nu^2)}r + \beta(\theta, \pm\nu^2)\nu r^2 + \nu^2\gamma(\theta, \nu r, \pm\nu^2)r^3, \tag{3.4.13}$$

Remark 3.4.2 *The reader my wonder what is going on: if the coefficients would not depend on θ, then the periodic orbit would be circular and would correspond to a zero in the above vector field. Such a zero would occur for $r = \mathcal{O}(\nu^{-1}\beta\gamma^{-1})$, thus it seems that I have just done the wrong scaling. The point is that the above naive analysis is correct only if we consider the average (with respect to θ) of the coefficients, but the average of β is zero! This is a very simple instance of a general theory called* averaging.

Remark 3.4.3 *In the following we will choose the case in which $\lambda > 0$, hence the change of variable with the plus is selected. The computations for $\lambda < 0$ are completely analogous and are left to the reader.*

Let us call $r(\theta, \xi, \nu)$ the solution of (3.4.13) with initial condition ξ and parameter ν.

Problem 3.14 *Prove that, for each $\theta \in [0, 2\pi]$ the function $r(\theta, \cdot, \cdot)$ are $\mathcal{C}^{r-1}$.*

We are finally ready to prove the existence of a periodic orbit. Clearly, an orbit is periodic if and only if $r(0, \xi, \nu) = r(2\pi, \xi, \nu)$. In other words, if we look at the motion only when it crosses the $\{\theta = 0 \mod 2\pi\}$ line, then we see the orbit always at the same point. To look at the motion only at some specified moment is a general strategy often used in Dynamical Systems of which we see here the first, but not the last, occurrence. It is called a *Poincaré section*.

In concrete, if we consider the map $S : \mathbb{R}^2_+ \to \mathbb{R}_+$ defined by $S(\xi, \nu) := r(2\pi, \xi, \nu)$, then the periodic orbits of the flow correspond to the fixed points of the maps $S(\cdot, \nu)$.[4]

[3] In fact, we have two different changes of variable according to the sign of μ.

[4] I mean the non trivial ones, since zero is always a trivial fixed point by construction.

Our last task is thus to study such a maps. The right idea is to develop them in power series of ν. For $\nu = 0$, $r(\theta, \xi, \nu)$ satisfies the Cauchy problem

$$\begin{aligned}&\frac{dr}{d\theta} = 0\\ &r(0, \xi, 0) = \xi.\end{aligned}$$

Thus $S(\xi, 0) = \xi$. To compute the derivative we must compute $\eta := \partial_\nu r(\theta, \xi, \nu)$. Such a derivative satisfies the equation obtained by differentiating (3.4.13) (see Theorem 1.1.14)

$$\begin{aligned}\frac{d\eta}{d\theta} &= \frac{2\nu}{\omega}r - \frac{2\nu^3\omega'}{\omega^2}r + \frac{\nu^2}{\omega}\eta + \beta r^2 + 2\nu r\eta\beta + 2\nu^2 r^2 \partial_{\nu^2}\beta\\ &\quad + 2\nu\gamma r^3 + 3\nu^2 r^2 \eta\gamma + 2\nu^3 r^3 \partial_{\nu^2}\gamma + \nu^3 r^3 (r + \nu\eta)\partial_{\nu r}\gamma\\ \eta(0, \xi, \nu) &= 0.\end{aligned} \tag{3.4.14}$$

Setting $\nu = 0$ in the above equation yields $\eta(\theta, \xi, 0) = \xi^2 \int_0^\theta \beta(\varphi, 0)d\varphi$. Accordingly, $\partial_\nu S(\xi, 0) = 0$ (see Problem 3.33).

To conclude we need to compute the second derivative at $\nu = 0$. Setting $\zeta(\theta, \xi) = \partial_\nu \eta(\theta, \xi, 0)$ and differentiating (3.4.14), yields

$$\begin{aligned}&\frac{d\zeta}{d\theta} = \frac{2}{\omega_0}\xi + 4\beta\xi\eta(\theta, \xi, 0) + 2\gamma(\theta, 0, 0)\xi^3\\ &\zeta(0, \xi, 0) = 0.\end{aligned}$$

which yields

$$\zeta(\theta, \xi) = \frac{2\theta}{\omega_0}\xi + 4\xi \int_0^\theta \beta(\varphi, 0)\eta(\varphi, \xi, 0)d\varphi + 2\xi^3 \int_0^\theta \gamma(\varphi, 0, 0)d\varphi.$$

Next, note that $\frac{d\eta(\varphi,\xi,0)}{d\varphi} = \xi^2\beta(\varphi, 0)$, thus

$$\int_0^\theta \beta(\varphi, 0)\eta(\varphi, \xi, 0)d\varphi = \frac{\eta(\theta, \xi, 0)^2}{2\xi^2} = \frac{\xi^2}{2}\left(\int_0^\theta \beta(\varphi, 0)d\varphi\right)^2.$$

Thus, setting $\bar\gamma = \int_0^{2\pi} \gamma(\varphi, 0, 0)d\varphi$, we have[5]

$$S(\xi, \nu) = (1 + \frac{2\pi}{\omega_0}\nu^2)\xi + \xi^3\bar\gamma\nu^2 + \nu^3\xi\Gamma(\xi, \nu) \tag{3.4.15}$$

[5] Since $S(0, \nu) = 0$, the coefficient of ν^3 must have the form $\xi\Gamma$.

To study the solution of $S(\xi,\nu) = \xi$ for $\nu \neq 0$ it is convenient to introduce the function $F(\xi,\nu) = \nu^{-2}(S(\xi,\nu)-\xi) = \frac{2\pi}{\omega_0}\xi+\xi^3\bar{\gamma}+\nu\Gamma(\xi,\nu)$.

If $\bar{\gamma} > 0$, then $F(\xi, 0)$ has no solutions different from zero and the same must hold for small ν.

If $\bar{\gamma} < 0$, then $\xi_0 = \sqrt{-\frac{2\pi}{\omega_0\bar{\gamma}}}$ is the only positive solution of $F(\xi, 0) = 0$. We can then apply the implicit function theorem since $F(\xi_0, 0) = 0$ and

$$\partial_\xi F(\xi_0, 0) = \frac{2\pi}{\omega_0} + 3\xi^2\bar{\gamma} = -\frac{4\pi}{\omega_0} \neq 0.$$

Has a conclusion we have a unique $\xi(\nu) = \xi_0+\mathcal{O}(\nu)$ such that $S(\xi(\nu),\nu) = \xi(\nu)$ for $\nu \neq 0$.

Problem 3.15 *Compute, in terms of the Tailor coefficients of V, what it means $\bar{\gamma} = 0$ and shows that it is not generic.*

3.5 The Hamiltonian case

It is important to note that non generic situations may appear due to symmetries or other type of constraints. To give an example of such a situation let us consider an Hamiltonian vector field, that is a vector field of the type $V(x, p) = (\partial_p H, -\partial_x H)$ for some function $H(x, p)$. In this case

$$DV = \begin{pmatrix} \partial_{xp}H & \partial_{pp}H \\ -\partial_{xx}H & -\partial_{xp}H \end{pmatrix}.$$

Note that the trace of DV is always zero, hence if one eigenvalue is null also the other must be. It is thus clear that in this case the situation with two complex conjugate eigenvalues is generic for a vector field while two, not one, zero eigenvalues is generic for a vector field family. In fact, for meachnical systems, the Hamiltonial has ofter the form $H(x, p) = \frac{1}{2}p^2 + U(x)$, for some function U. Hence, $V(x, p) = (p, -\partial_x U)$, which means that the zeroes of the vector field are the critical points of U. Let us discuss Hamiltonian systems in which the Hamiltonian is of the above type.

We start with the so called *one degree of freedom*, i.e. $x, p \in \mathbb{R}$.

Problem 3.16 *Show that if U has a minimum, then the fixed point is a* center*, while if U has a maximum, then the corresponding fixed point is hyperbolic.*

We have thus a new phenomena: a center that is stable under small perturbations!

Let us consider the case in which a one parameter family of potentials $U(x, \lambda)$ has a degenerate minimum at zero, i.e. $U(0, \lambda) = 0, \partial_x U(0, \lambda) = 0, \partial_x^2 U(0, 0) = 0$. This means that $U(x) = \lambda x^2 + a(\lambda, x)x^3$ and

$$V(x, \lambda) = (p, 2\lambda x + a_1(x, \lambda)x^2)$$

Problem 3.17 *Show that in the above family we have the collision of two fixed point (a center and a saddle) that collide and exchange type.*

This meas that the zeroes of the vector fields are $p = 0$, $x(\lambda) = 0$ and $x(\lambda) \sim -\frac{2\lambda}{a_1(0,0)}$. We then have a new phenomena: two fixed point that cross and exchange type.[6]

Even more syngular situations may happen if more constraints are present. Cosider, for example the above situation when, for some reason, the Hamiltonian is constraiend to being symmetrical: $H(x, p) = H(-x, p)$. Then it would have the form $U(x) = \lambda x^2 + a(\lambda, x)x^4$.

Problem 3.18 *Show that in the above case one has one fixed point that evolves into three fixed points. Moreover show that if when only one fixed point is present, the fixed point is unstable, then of the three fixed point two are unstable and one stable. This is called a* peach fork *bifurcation.*

Next let us consider the case of two degree of freedom, i.e. $x, p \in \mathbb{R}^2$. Limited to the case of a minimum. In such a situation, at the point of minimum, we have

$$\partial_x V(x, p) = \begin{pmatrix} 0 & \mathbb{1} \\ -\partial_x^2 U & 0 \end{pmatrix}. \tag{3.5.16}$$

where $\partial_x^2 U$ is a positive symmetric matrix, let ω_1^2, ω_2^2 be its eigenvlues.

Problem 3.19 *Show that the eigenvalues of $\partial_x V$, at the fixed point, are $\pm\omega_i$.*

[6] Hence the set of fixed points no longher forms a smooth curve in the x, λ space.

Another surprise: a stable situation with four immaginary eigenvalue (an higher dimensional center).

Problem 3.20 *Consider the linear equation (obtained by the matrix* (3.5.16) *after a change of variables)*

$$\dot{x} = p$$
$$\dot{p} = \begin{pmatrix} -\omega_1^2 & 0 \\ 0 & -\omega_2^2 \end{pmatrix} x$$

Show that $p_i^2 + x_i^2$ *are invariant of the motion, i.e. the motion takes place on* two-dimensional tori.

Remark 3.5.1 *Contrary to the case of one degree of freedom, in which the conservation of the Hamiltonian iplies that the center is stable for the full motion, in higher dimension it is not clear if the center is stable or not for the full dynamics. Indeed this is a rather complex matter at presetn not completely clarified. Part of the answer is the subject of the so called KAM theory. We will discuss some aspects of it in the following.*

Problems

3.21. Compute $\tilde{V} = V \circ F$ where V is given by (3.2.3) and F by (3.2.4), i.e. $F(\xi, \tau) = (\xi - \tau, \xi)$. Show, by direct computation, that $\tilde{V}(\xi, \tau) = 0$ has solution $\xi(\tau) = -\frac{b}{a}\tau^2 + \mathcal{O}(\tau^3)$.

3.22. Prove that the set $\{A \in GL(n, \mathbb{R}) \ : \ \det(A) \neq 0\}$ is generic with respect to the topology induced by the norm.

3.23. Prove that the set $\{A \in GL(n, \mathbb{R}) \ : \ A \text{ is hyperbolic}\}$ is generic.

3.24. Prove that $\{A \in \mathcal{C}^0([-1, 1], GL(n, \mathbb{R})) \ : \ \operatorname{rank}(A(\lambda)) \geq n - 1 \ \forall \lambda \in [-1, 1]\}$ is generic.

3.25. Prove that the set $\{A \in GL(n, \mathbb{R}) \ : \ A \text{ is hyperbolic and has only simple eigenvalues}\}$ is generic (i.e. Jordan blocks are atypical).

3.26. Show that if $A \in GL(2, \mathbb{R})$ and its eigenvalues have zero real part, then $\operatorname{Tr}(A) = 0$.

3.27. If $A \in \mathcal{C}^1([-1,1], GL(n,\mathbb{R}))$ and $\Pi \in \mathcal{C}^1([-1,1], GL(n,\mathbb{R}))$ is an eigenprojector, show that $\frac{d}{d\lambda}\operatorname{Tr}(\Pi A) = 2\operatorname{Tr}(\Pi \frac{d}{d\lambda} A)$.

3.28. Show that the set $\{A \in \mathcal{C}^1([-1,1], GL(n,\mathbb{R}))$: at most two eigenvalues have zero real part$\}$ is generic.

3.29. Prove that the set

$$A_K := \{V \in \mathcal{C}^r(\mathbb{R}^n, \mathbb{R}^n) \;:\; V(x) = 0 \text{ implies } \partial_x V \text{ hyperbolic } \forall\, x \in K\}$$

is generic in the $\mathcal{C}^r$ topology.

3.30. Show that any matrix $A \in GL(2,\mathbb{R})$ with two eigenvalue with zero real part is conjugate to a matrix of the form

$$\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$$

for some $\omega > 0$.

3.31. Let $f \in \mathcal{C}^r(\mathbb{R}^{d+1})$ and write the elements of $\mathbb{R}^{d+1}$ as $(\xi_1, \dots, \xi_d, t)$. If $f(\xi, 0) = \partial_t^k f(\xi, 0) = 0$ for all $k \le s < r$, then there exists $g \in \mathcal{C}^{r-s}$ such that $f(\xi, t) = t^s g(\xi, t)$.

3.32. Let $f \in \mathcal{C}^r(\mathbb{R}^{d+1})$ and write the elements of $\mathbb{R}^{d+1}$ as $(\xi_1, \dots, \xi_d, t)$. Then, for all $s < r$, there exists $g \in \mathcal{C}^{r-s}$ such that $f(\xi, t) = \sum_{k=0}^{s-1} f^k(\xi, 0) t^k + t^s g(\xi, t)$.[7]

3.33. Show that if $p(\theta)$ is a product of an odd number of functions equal either to $\sin\theta$ or $\cos\theta$, then $\int_0^{2\pi} p(\theta) = 0$.

Hints to solving the Problems

3.3 Let $\bar{x} \in K$ such that $V(\bar{x}) = 0$. Then, by assumption $D_{\bar{x}} V$ is invertible, so $V(\bar{x} + \xi) = 0$ can be written as

$$D_{\bar{x}} V^{-1}(D_{\bar{x}} V \xi - V(\bar{x} + \xi)) = \xi.$$

[7] Essentially this is Taylor formula where one controls the smoothness of the remainder. This issue is relevant in the applications, but often not investigated in standard textbooks.

Since $D_{\bar{x}}V\xi - V(\bar{x}+\xi) = o(\|\xi\|)$, it follows that the above equation has the unique solution $\xi = 0$ in a sufficiently small neighborhood of zero. Hence there exists a neighborhood of $\bar{x}$ in which there are no other zeroes. Next, for each point in K consider a neighborhood as follows: if the V is different from zero at such a point, then consider a neighborhood for which the vector field is different from zero. If the vector field is zero at the point then consider the above neighborhood in which the point is the only zero. In such a way we have a covering of K, we can then extract a finite subcover hence proving the statement.

3.4 Let $V \in A_K$ and $\{x_i\}_{i=1}^M$ be the zeroes of V. Then for each vector field $W \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, $\|W\| \leq 1$, consider the family $V(x,\mu) := V(x) + \mu W(x)$. For each $i \in \{1, \dots, M\}$, use the implicit function theorem to show that there exists $\varepsilon_i, \delta_i > 0$ and $X_i \in \mathcal{C}^1([-\varepsilon_i, \varepsilon_i], \mathbb{R}^n) \to \mathbb{R}^d$, $X_i(x_i) = 0$, such that $V(X_i(\mu), \mu) = 0$ and $V(x,\mu) = 0$, $\|x - x_i\| \leq \delta_i$, $|\mu| \leq \varepsilon_i$ implies that $x = X_i(\mu)$. Verify (using perturbation theory) that, for μ small enough $\partial_x V(X(\mu), \mu)$ is hyperbolic. Next, set $\delta = \min \delta_i$ and $\rho := \inf_{|x-x_i| \geq \delta} \|V(x)\|$. Clearly $V(x,\mu) \neq 0$ if $|x - x_i| \geq \delta$ and $\mu < \rho$. Hence a neighborhood of V of size $\min\{\varepsilon_i, \rho\}$ belongs to A_K, hence A_K is open.

3.5 If $Z_K = \{z \in K \ : \ \det(D_x V) = 0\}$, then $V(Z_K)$ is a zero measure set by Sard's Theorem. Let $Z \subset \mathbb{R}^d$ be a zero measure set and, for each $v \in \mathbb{R}^d$, define $Z(v) = \{z \in \mathbb{R}^d \ : \ z - v \in Z\}$. Show that for each $\varepsilon > 0$ there exists $v \in \mathbb{R}^d$, $\|v\| \leq \varepsilon$, such that $0 \notin Z(v)$. Given $V \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$, use this to show that for each $\varepsilon > 0$ there exists $v \in \mathbb{R}^d$, $\|v\| = 1$ such that $V_\varepsilon(x) := V(x) + \varepsilon v$ has the property that $\det(D_x V_\varepsilon) = \det(D_x V) = 0$ implies $V_\varepsilon(x) \neq 0$. An application of the implicit function theorem then shows that the zeroes of V_ε are isolated. Finally, construct $\tilde{V}_\varepsilon$, $\|V_\varepsilon - \tilde{V}_\varepsilon\|_{\mathcal{C}^1} \leq \varepsilon$, such that the zeroes are unchanged but the derivative is hyperbolic, hence $\tilde{V}_\varepsilon \in A_K$. This last step can be performed locally so it suffices to show how to perform it around one single point. First of all note that, by continuity, there exists $\alpha > 0$ such that $V_\varepsilon(x) = 0$ implies $\|(D_x V_\varepsilon)^{-1}\| \leq \alpha^{-1}$. Next, let $x_0 \in K$ such that $V_\varepsilon(x_0) = 0$. Then $V_\varepsilon(x) = D_{x_0} V_\varepsilon (x - x_0) + o(x - x_0)$.

Thus there exists $\delta > 0$ such that, for all $\|x - x_0\| \leq \delta$,[8]

$$\|V_\varepsilon(x)\| \geq \frac{\alpha}{2}\|x - x_0\|.$$

Finally, consider the vector field $\tilde{V}_t(x) = V_\varepsilon(x)+t(x-x_0)\varphi(x-x_0)$. Where $\varphi \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$ is some fixed function such that the support of φ is contained in the ball of radius δ, $\varphi(0) = 1$, $\nabla\varphi(0) = 0$ and $\|\varphi\|_\infty \leq 1$. Then

$$\|\tilde{V}_t(x)\| \geq (\frac{\alpha}{2} - t)\|x - x_0\|$$

so if $t < \frac{\alpha}{2}$, the filed $\tilde{V}_t(x)$ has the same zeroes than V_ε. Moreover, $D_{x_0}\tilde{V}_t = D_{x_0}V_\varepsilon + t\mathbb{1}$ which is hyperbolic and

$$\|V_\varepsilon - \tilde{V}_t\|_{\mathcal{C}^1} \leq 2t\delta + t\|\varphi\|_{\mathcal{C}^1}$$

which can be made smaller than ε by choosing t sufficiently small.

3.7 If suffices to show that B_K is generic for each compact $K \subset \mathbb{R}^d$. The openness comes from the fact that a small perturbations cannot change the condition on the rank nor the transversality conditions, nor increase the number of eigenvalues with zero real part (due to perturbation theory). For the density, consider the set $W := \{(x, \lambda) \in K^1 : \text{rank}(\partial_x V\ \partial_\lambda V) < d\}$. Using the same strategy as in Theorem 3.1.3 show that $V(W)$ has zero Lebesgue measure.[9] This means that, for each $\varepsilon > 0$ there exists $v \in \mathbb{R}^d$, $\|v\| \leq \varepsilon$, such that for each $(x, \lambda) \in K$ such that $V(x, \lambda) = -v$ holds $\text{rank}(\partial_x V\ \partial_\lambda V) = d$. We can then consider the vector field $V_\varepsilon = V + v$. The proof of Lemma 3.2.2 shows that the zeroes of V_ε are either isolated or isolated curves. It is then sufficient to modify the vector field in a neighborhood of such curves in order to obtain an element of B_K. Since the argument will be local, one can limit the discussion to one such curve.

Suppose that $(\bar{x}, \bar{\lambda})$ is such that $V_\varepsilon(\bar{x}, \bar{\lambda}) = 0$ and $Z(\partial_x V(\bar{x}, \bar{\lambda})) > 2$. Then we know that the zeroes of V near such a point consist

[8] Note that, by the uniform continuity of the derivative on K, δ can be chosen independent on the point.

[9] In fact, this is nothing else than another special case of the general Sard Theorem.

in a curve $x(\lambda)$. Then consider $V_{\varepsilon,\mu}(x,\lambda) = V_\varepsilon(x,\lambda) + \mu\varphi(x - x(\lambda))A(\lambda)(x - x(\lambda))$, where φ is a function supported in a small neighborhood of zero. Verify that, for μ small enough $V_{\varepsilon,\mu}$ and V_ε have the same zeroes. By Problem 3.28 one can choose the matrices $A(\lambda)$ so that $Z(\partial_x V_{\varepsilon,\mu}(x(\lambda),\lambda)) \leq 2$ for λ close enough to $\bar\lambda$. Next, by adding a second order term to the vector field, one can satisfy also the second order condition.

Suppose instead that $\det(\partial_x V_\varepsilon(\bar x,\bar\lambda)) = 0$. In this case the zeroes belong to a curve $(x(\tau),\lambda(\tau))$ with $(x(0),\lambda(0)) = (\bar x,\bar\lambda)$ and $\lambda'(0) = 0$. Let us consider then the vector fields $V_{\varepsilon,\mu}(x,\lambda) = V_\varepsilon(x,\lambda) + \mu\langle x-\bar x, A(x-\bar x)\rangle\varphi(x-\bar x)$. For μ small enough the set of zeroes will change little and it will remain tangent to the original one in $(\bar x,\bar\lambda)$ since $\partial_x V_\varepsilon(\bar x,\bar\lambda) = \partial_x V_{\varepsilon,\mu}(\bar x,\bar\lambda)$. The condition on the second derivative can then be satisfied by choosing the vector of matrices A appropriately.

3.11 We know from the discussion Lemma 3.2.2 that there exists $x(\lambda)$ such that $V(x(\lambda),\lambda) = 0$, we can then set $\alpha(\lambda) = x(\lambda)$. We get then the wanted equation with the new remainder given by $R(\xi+x(\lambda),\lambda)-R(x(\lambda),\lambda)$. The other properties of R are obtained by direct computation.

3.12 Remember that the change of variable must be performed on the equation $\dot x = V(x,\lambda)$, so the vector field changes as $D^{-1}V(Dz)$. In addition, since $\partial_\xi R(0,0) = 0$, Problem 3.32 implies that we can write $\partial_\xi R(0,\lambda) = C(\lambda)\lambda$ for some $\mathcal{C}^{r-1}$ matrix C. Choose $D(\lambda) = D_0(\lambda)D_1(\lambda))$. Since we do not want to change the form of $\partial_x V$ at first order in λ we impose $[D_0, A] = 0$. Show that this implies $D_0(\lambda) = \begin{pmatrix} 1 & -a(\lambda) \\ a(\lambda) & 1\end{pmatrix}$. Show that one can choose a such that

$$D_0^{-1}\partial_x V(0,\lambda)D_0 = A + \lambda H(\lambda)$$

with $H_{11} = H_{22} \geq 0$. Note then that $H_{ii}(0) \neq 0$ since $\operatorname{Tr} H(0) \neq 0$ by hypothesis. Next, choose $D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1+\lambda b(\lambda)\end{pmatrix}$. Show that b can be chosen so that

$$D_1^{-1}D_0^{-1}\partial_x V(0,\lambda)D_0 D_1 = A + \lambda\tilde H(\lambda)$$

with $\tilde{H}_{ii} = H_{ii}$ and $\tilde{H}_{12} = -\tilde{H}_{21}$. The problem is then solved by choosing ρ.

3.27 Using a "dot" to mean differentiation holds $\frac{d}{d\lambda}\operatorname{Tr}(\Pi A) = \operatorname{Tr}(\dot{\Pi} A + \Pi \dot{A})$. If B is the portion of spectrum associated to $\Pi(0)$ and γ a curve surrounding it and no other part of the spectrum, then

$$\dot{\Pi}(0) = \frac{1}{2\pi i}\int_\gamma (z - A(0))^{-1}\dot{A}(0)(z - A(0))^{-1} dz$$

Thus

$$\begin{aligned}\operatorname{Tr}(\dot{\Pi} A) &= \operatorname{Tr}(\Pi\dot{A}) + \frac{1}{2\pi i}\int_\gamma z \operatorname{Tr}\left((z - A(0))^{-1}\dot{A}(0)(z - A(0))^{-1}\right) dz \\ &= \operatorname{Tr}(\Pi\dot{A}) + \frac{1}{2\pi i}\int_\gamma z \operatorname{Tr}\left((z - A(0))^{-1}\dot{A}(0)\right) dz = \operatorname{Tr}(\Pi\dot{A}).\end{aligned}$$

Notes

The present discussion is intended only to give a flavor of the subject and of how it can be systematically developed. For a more complete (and advanced) treatment of bifurcation theory see [Arn83, CH82].