Shao-Kai Jonathan Huang

January 15, 2026

Contents

Chapter 1

Introduction

Chapter 2

General Fenichel Theory

Chapter 3

Geometric Singular Perturbation Theory

Note. Some useful background information include: knowledge about the geometry of curves and surfaces and general smooth manifolds and an elementary understanding of classical bifurcation theory.

3.1 Fenichel's Theorem

Let's start with the general formulation of an (m, n)-fast-slow system:

Definition 3.1.1. For $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$, a slow-fast system is defined by the equations

$$\begin{aligned} \varepsilon \frac{\mathrm{d}x}{\mathrm{d}\tau} &= \varepsilon \dot{x} = f(x, y, \varepsilon), \\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= \dot{y} = g(x, y, \varepsilon), \end{aligned} \tag{3.1}$$

where $0 < \varepsilon \ll 1$ is a small parameter representing the ratio of time scales between the fast variable x and the slow variable y. Here, $f : \mathbb{R}^m \times \mathbb{R}^n \times [0, \varepsilon_0) \to \mathbb{R}^m$ and $g : \mathbb{R}^m \times \mathbb{R}^n \times [0, \varepsilon_0) \to \mathbb{R}^n$ are sufficiently smooth functions.

On the fast time scale $t = \tau/\varepsilon$, the system can be rewritten as

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= x' = f(x, y, \varepsilon), \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= y' = \varepsilon g(x, y, \varepsilon). \end{aligned} \tag{3.2}$$

1. Slow subsystem: Setting $\varepsilon = 0$ in the slow time scale equations (3.1) gives

$$\begin{aligned} 0 &= f(x, y, 0), \\ \dot{y} &= g(x, y, 0). \end{aligned} \tag{3.3}$$

   The slow flow is constrained to the *critical manifold* defined by

$$C_0 = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n : f(x, y, 0) = 0\}. \tag{3.4}$$

   A main goal of this chapter is to state Fenichel's theorem for perturbations of the critical manifold C_0.

2. Fast subsystem: Setting $\varepsilon = 0$ in the fast time scale equations (3.2) gives

$$\begin{aligned} x' &= f(x, y, 0), \\ y' &= 0. \end{aligned} \tag{3.5}$$

Definition 3.1.2 (normal hyperbolicity). A subset $S \subseteq C_0$ of the critical manifold is said to be *normally hyperbolic* if $(\mathrm{D}_x f)(p, 0) \in M_{m\times m}$ has no eigenvalues with zero real part for all $p \in S$.

Recall that an equilibrium point $p \in S$ is said to be *hyperbolic* if the linearization $(D_x f)(p, 0)$ has no eigenvalues with zero real part.

Proposition 3.1.1. A subset $S \subseteq C_0$ is normally hyperbolic if and only if for each equilibrium point $p = (x^*, y^*) \in S$, x^* is a hyperbolic equilibrium point of the fast subsystem $x' = f(x, y^*, 0)$.

Definition 3.1.3. A normally hyperbolic subset $S \subseteq C_0$ is called *attracting* if all eigenvalues of $(D_x f)(p, 0)$ have negative real parts for $p \in S$, *repelling* if all eigenvalues have positive real parts for all $p \in S$, and of *saddle type* otherwise.

We also need a notion of distance between sets.

Definition 3.1.4 (Hausdorff distance). Let $A, B \subseteq \mathbb{R}^n$ be two nonempty compact sets. The *Hausdorff distance* between A and B is defined as

$$\operatorname{dist}(A, B) = \max\left\{\sup_{a\in A}\inf_{b\in B}\|a-b\|,\ \sup_{b\in B}\inf_{a\in A}\|b-a\|\right\}. \tag{3.6}$$

Now we can state the *Fenichel-Tikhonov theorem*, in a form suitable for the study of fast-slow systems.

Theorem 3.1.1 (Fenichel-Tikhonov). Suppose $S = S_0$ is a compact normally hyperbolic submanifold of the critical manifold C_0 in equation (3.4), and $f, f \in C^r$ for $r < \infty$. Then, for sufficiently small $\varepsilon > 0$, we have

(i) There exists a locally invariant manifold S_ε diffeomorphic to S_0 (this implies trajectories can enter and leave S_0 only through its boundary).

(ii) S_ε and S_0 have Hausdorff distance $\operatorname{dist}(S_\varepsilon, S_0) = O(\varepsilon)$.

(iii) The flow on S_ε converges to the slow flow on S_0 as $\varepsilon \to 0$.

(iv) S_ε is C^r-smooth.

(v) S_ε is normally hyperbolic and has similar stability properties with respect to fast variables to S_0.

(vi) S_ε is in general not unique. At regions lying at a fixed distance from ∂S_ε, all manifolds satisfying (i)-(v) lie at a Hausdorff distance $O(e^{-K/\varepsilon})$ from each other for some $0 < K = O(1)$.

Definition 3.1.5 (slow manifold). The locally invariant manifold S_ε obtained in theorem 3.1.1 is called a *slow manifold*.

Remark. Even though theorem 3.1.1 makes clear that slow manifolds are in general not unique, any two slow manifolds are exponentially close to each other in ε away from the boundary ∂S_ε. Thus, for practical purposes, it is conventional to refer to *the* slow manifold S_ε when discussing dynamics away from ∂S_ε.

We can extend slow manifolds under the flow of the full system, where the extension may not have anything to do with a critical manifold.

Example (Finding the slow manifold). Consider a $(1, 1)$-fast-slow system given by

$$\begin{aligned} x' &= y^2 - x, \\ y' &= -\varepsilon y. \end{aligned} \tag{3.7}$$

The critical manifold is given by $C_0 = \{(x, y) \in \mathbb{R}^2 \mid x' = y^2 - x = 0\}$. Since $\partial/\partial x(y^2 - x) = -1 < 0$, C_0 is normally hyperbolic and attracting. By Fenichel's theorem, there exists a slow manifold S_ε for sufficiently small $\varepsilon > 0$. To find an expression for S_ε, notice that we can solve equation (3.7) explicitly as

$$(x(t), y(t)) = \left(\left[x(0) - \frac{y(0)^2}{1-2\varepsilon}\right] e^{-t} + \frac{y(0)^2}{1-2\varepsilon} e^{-2\varepsilon t},\ y(0)e^{-\varepsilon t}\right), \tag{3.8}$$

by solving for $y(t)$ first. In the case $x(0) = y(0)^2/(1-2\varepsilon)$, the system evolves only on the slow time scale $\tau = \varepsilon t$, and in particular

$$x(t) = \frac{1}{1-2\varepsilon} y(t)^2 \quad \text{for all } t \geq 0. \tag{3.9}$$

Hence,

$$S_\varepsilon = \left\{(x, y) \in \mathbb{R}^2 \,\middle|\, x = \frac{y^2}{1-2\varepsilon}\right\}. \tag{3.10}$$

Exercise. Show that for $0 < \varepsilon \ll 1$, the equilibrium point q is globally asymptotically stable.

Answer. To see this, first recall the definition of (Lyapunov) stability and asymptotic stability. An equilibrium point q is *(Lyapunov) stable* if for every neighborhood U of q, there exists a neighborhood $V \subseteq U$ such that every solution starting in V remains in U for all future time ($t \geq 0$). An equilibrium point q is *asymptotically stable* if it is stable and for any initial condition $(x(0), y(0))$, the solution trajectory $(x(t), y(t))$ approaches q as $t \to \infty$.

Since $\varepsilon > 0$, by the explicit solution, we have

$$\begin{aligned} x(t) &= \left[x(0) - \frac{y(0)^2}{1-2\varepsilon}\right] e^{-t} + \frac{y(0)^2}{1-2\varepsilon} e^{-2\varepsilon t} \to 0, \\ y(t) &= y(0)e^{-\varepsilon t} \to 0 \quad \text{as } t \to \infty, \end{aligned} \tag{3.11}$$

and hence $(x(t), y(t)) \to (0, 0) = q$ as $t \to \infty$, and q is globally attracting. Suppose $\|(x, y) - (0, 0)\|$ is small. By the triangle inequality, we have

$$\begin{aligned} x(t) &\leq |x(0)| + \left|\frac{y(0)^2}{1-2\varepsilon}\right| + \left|\frac{y(0)^2}{1-2\varepsilon}\right|, \\ y(t) &\leq |y(0)|. \end{aligned} \tag{3.12}$$

Hence, for small enough $x(0)$, $y(0)$, the solution trajectories $(x(t), y(t))$ are arbitrarily close to q. Together, q is globally asymptotically stable. ⊛

Remark. The general problem of computing slow manifolds analytically or numerically is highly nontrivial.

3.2 The Slow Flow

To obtain an analytical expression for the slow flow on the critical manifold, we define the weaker (with respect to normal hyperbolicity) concept of *regularity*.

Definition 3.2.1. Let C_0 be the critical manifold. We call $C_{0,s} = \{p \in C_0 \mid \det(\mathrm{D}_x f(p, 0)) \neq 0\}$ the *fast-slow singular points*, and $C_{0,r} = C_0 - C_{0,s}$ the *fast-slow regular points.*

Consider an equilibrium point $p^* \in C_0$ with fast subsystem linearization

$$\mathrm{D}_x f(p^*, 0) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad (\mathrm{D}_x f)(p^*, 0)^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.13}$$

Hence $p^* \in C_{0,r}$ is a (fast-slow) regular point in the sense of definition 3.2.1, but not normally hyperbolic since the eigenvalues of $\mathrm{D}_x f(p^*, 0)$ are purely imaginary.

Example (parametrizing the fast system using slow variables)**.** Let C_0 be a manifold, and $p \in C_0$ a regular point. By the implicit function theorem, there exists a neighborhood $U \subseteq \mathbb{R}^n$ of y^* and a smooth function $h : U \to \mathbb{R}^m$ such that the critical manifold can be locally represented as the graph of h, i.e.,

$$C_0 \cap (\mathbb{R}^m \times U) = \{(h(y), y) \mid y \in U\}. \tag{3.14}$$

Hence, we can reduce the slow subsystem (3.3) to

$$\dot{y} = g(h(y), y, 0), \quad y \in U, \tag{3.15}$$

in some neighborhood U of p. Consider the unforced van der Pol oscillator given by

$$\begin{aligned} \varepsilon \dot{x} &= y - \frac{1}{3}x^3 + x, \\ \dot{y} &= -x. \end{aligned} \tag{3.16}$$

For this system, $C_0 = \{(x, y) \in \mathbb{R}^2 \mid y = x^3/3 - x\}$. Differentiating the slow-variable equation with respect to the slow time τ gives

$$\frac{\partial f}{\partial x}\dot{x} + \frac{\partial f}{\partial y}\dot{y} = 0 \Rightarrow (1 - x^2)\dot{x} + \dot{y} = 0. \tag{3.17}$$

Using $\dot{y} = -x$, we have an explicit expression for the *slow flow* on C_0:

$$\dot{x} = \frac{x}{1 - x^2}. \tag{3.18}$$

The slow flow has an equilibrium point at $x = 0$, which is unstable since $\dot{x} > 0$ for $x > 0$ and $\dot{x} < 0$ for $x < 0$. The points $x = \pm 1$ are singularities of the slow flow, and they separate C_0 into

$$\begin{aligned} C_0^{a-} &= C_0 \cap \{(x, y) \in \mathbb{R}^2 \mid x < -1\} \\ C_0^{r} &= C_0 \cap \{(x, y) \in \mathbb{R}^2 \mid |x| < 1\} \\ C_0^{a+} &= C_0 \cap \{(x, y) \in \mathbb{R}^2 \mid x > 1\}. \end{aligned} \tag{3.19}$$

We can compute $\mathrm{D}_x f|_{C_0} = -x^2 + 1$. Hence $\mathrm{D}_x f|_{C_0^{a-}}, \mathrm{D}_x f|_{C_0^{a+}} < 0$, $\mathrm{D}_x f|_{C_0^{r}} > 0$, and thus C_0^{a-} and C_0^{a+} are normally hyperbolic and attracting, while C_0^r is normally hyperbolic and repelling.

Remark. A problem with the Implicit Function Theorem formulation is that it does not describe a vector field in $\mathbb{R}^{n+m}$ **(why?)**. Therefore, we cannot compare the singular limit with the full system directly.

Fact. The slow subsystem must be tangent to the critical manifold, and hence one can always embed it into $\mathbb{R}^{m+n}$.

Proposition 3.2.1. Let the critical manifold $C_0 \subseteq \mathbb{R}^m \times \mathbb{R}^n$ be smooth, $p = (x_0, y_0) \in C_0$, and let $(\mathrm{D}_x f)(p, 0)$ have maximal rank (hence $p \in C_{0,r}$). Then there exists a neighborhood $V \subseteq C_0$ of p such that the slow flow on $C_0 \cap V$ is given by

$$\begin{aligned} \dot{x} &= -\left(\mathrm{D}_x f(q, 0)\right)^{-1} \left(\mathrm{D}_y f(q, 0)\right) g(q, 0), \\ \dot{y} &= g(q, 0), \quad q \in C_0 \cap V. \end{aligned} \tag{3.20}$$

Notation. We may write

$$\begin{aligned} \dot{x} &= -(\mathrm{D}_x f)^{-1}(\mathrm{D}_y f)\, g, \\ \dot{y} &= g, \end{aligned} \tag{3.21}$$

It is in general very difficult to obtain an explicit expression for the slow flow on C_0, since this involves

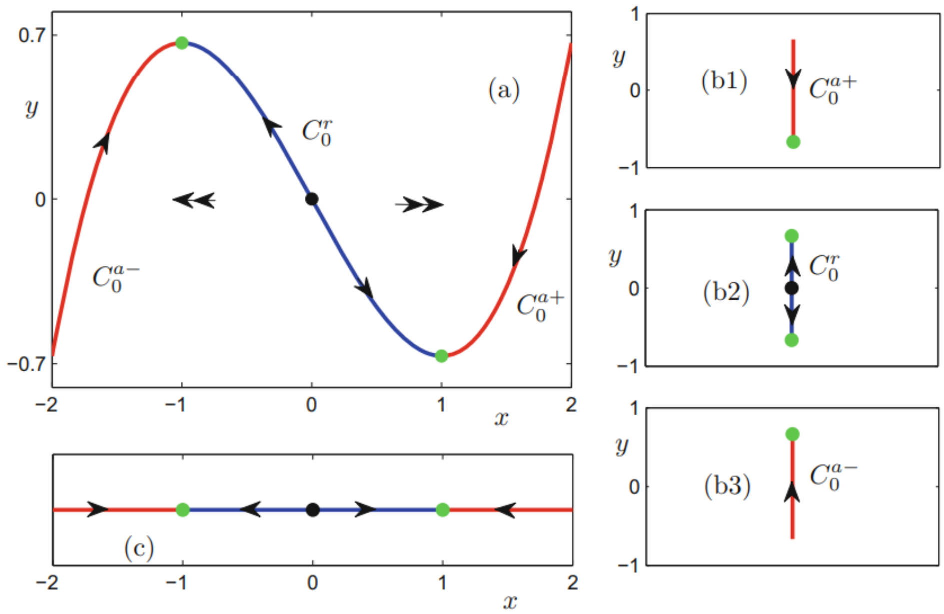

Figure 3.1: The critical manifold (blue curve) and the slow flow (black arrows) of the unforced van der Pol oscillator. The red dots indicate the singularities of the slow flow at $x = \pm 1$.

solving a nonlinear equation $f(x, y, 0) = 0$ in high dimensions.

3.3 Singularities

A large part of multiple time scale dynamics deals with loss of regularity and normal hyperbolicity. Therefore, let's consider a singular point $p \in C_{0,s}$. The simplest case is when $(\mathrm{D}_x f)(p, 0) : \mathbb{R}^m \to \mathbb{R}^m$ has rank $m - 1$ with zero eigenvalue of multiplicity one.

Example. Consider the $(1, 1)$-fast-slow system in fast time scale t:

$$\begin{aligned} x' &= f(x, y, \varepsilon) = -x^2 + y, \\ y' &= \varepsilon g(x, y, \varepsilon). \end{aligned} \tag{3.22}$$

We have $C_0 = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$, and $(\mathrm{D}_x f)(p, 0) = -2x = 0$ at the origin. Therefore, $p = (0, 0) \in C_{0,s}$ is not regular nor normally hyperbolic. Also, we can check that

$$\frac{\partial^2 f}{\partial x^2}(p, 0) = -2 \neq 0. \tag{3.23}$$

The fast subsystem given by

$$x' = -x^2 + y, \quad y' = 0 \tag{3.24}$$

gives rise to what is called a "fold bifurcation" at $y = 0$.

Definition 3.3.1 (fold bifurcation)**.** A point $p \in C_0$ such that $f(p, 0) = 0$ is called a *fold point* if $\operatorname{rank} \mathrm{D}_x f(p, 0) = m - 1$. It is said to be *nondegenerate* if for vectors w, v in the left and right nullspaces of $\mathrm{D}_x f(p, 0)$, respectively, we have

$$w^T \mathrm{D}_{xx} f(p, 0) v^2 \neq 0, \quad w^T \mathrm{D}_y f(p, 0) \neq 0. \tag{3.25}$$

Note. Fold points can be viewed as fold bifurcations of the fast subsystem. Fold bifurcations are also called *saddle-node bifurcation*, *turning point*, or *limit point*.

Although a fold bifurcation point can be made to disappear for $x = f(x)$ with $x \in \mathbb{R}$ by a perturbation, a nondegenerate fold bifurcation is stable under perturbations in 1-parameter families of vector fields given by the fast subsystem $x' = f(x, y)$ (it has "codimension one").

Definition 3.3.2 (codimension)**.** From linear algebra, the *codimension* of a subspace W of a finite-dimensional vector space V is defined as $\operatorname{codim}(W) = \dim(V) - \dim(W)$. More generally, the codimension of W in V is the (possibly infinite) dimension of the quotient space V/W.

Proposition 3.3.1. Fold bifurcations are *generic* in the topological sense in 1-parameter families of (sufficiently smooth) vector fields.

There are many other singularities and bifurcations that can occur in fast-slow systems, such as *cusp singularity*, *transcritical point*, and *Hopf bifurcation*.

Example (cusp singularity)**.** Consider the $(1, 2)$-fast-slow system

$$\begin{aligned} x' &= y_1 + y_2 x - x^3 = f(x, y, \varepsilon), \\ y_1' &= \varepsilon g_1(x, y, \varepsilon), \\ y_2' &= \varepsilon g_2(x, y, \varepsilon). \end{aligned}$$

The critical set

$$C_0 = \{(x, y) \in \mathbb{R}^2 : y_1 = -y_2 x + x^3\} \tag{3.26}$$

is a *manifold*, but it contains a curve of fold points given by

$$L = \left\{ (x, y) \in C_0 : \frac{\partial f}{\partial x} = y_2 - 3x^2 = 0 \right\}. \tag{3.27}$$

We can further compute the second partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = -6x. \tag{3.28}$$

Hence, by definition 3.3.1 the fold points are nondegenerate everywhere except at the origin, where we have a *cusp singularity*; see Figure 3.2(c).

Note. A cusp singularity is a codimension-2 bifurcation of the fast subsystem. From bifurcation theory, we know two parameters are needed to understand its dynamics.

Example (other singularities)**.** Consider the $(2, 1)$-fast–slow system

$$\begin{aligned} x_1' &= y x_1 - x_2 - x_1(x_1^2 + x_2^2) = f_1(x, y), \\ x_2' &= x_1 + y x_2 - x_2(x_1^2 + x_2^2) = f_2(x, y), \\ y' &= \varepsilon g(x, y, \varepsilon), \end{aligned} \tag{3.22}$$

where $f := (f_1, f_2)$. The critical manifold

$$C_0 = \{(x, y) \in \mathbb{R}^2 : x_1 = 0 = x_2\}$$

is simply the y-axis. The linearization with respect to the fast variables is

$$D_x f\big|_{C_0} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{pmatrix}_{C_0} = \begin{pmatrix} y - 1 & -1 \\ 1 & y \end{pmatrix}.$$

Therefore, C_0 consists only of regular points, but $D_x f|_{C_0}$ has a pair of complex eigenvalues $\pm i$ at

$y = 0$, so that C_0 is not normally hyperbolic at the origin, which is a Hopf bifurcation of the fast subsystem; see Figure 3.2(d).

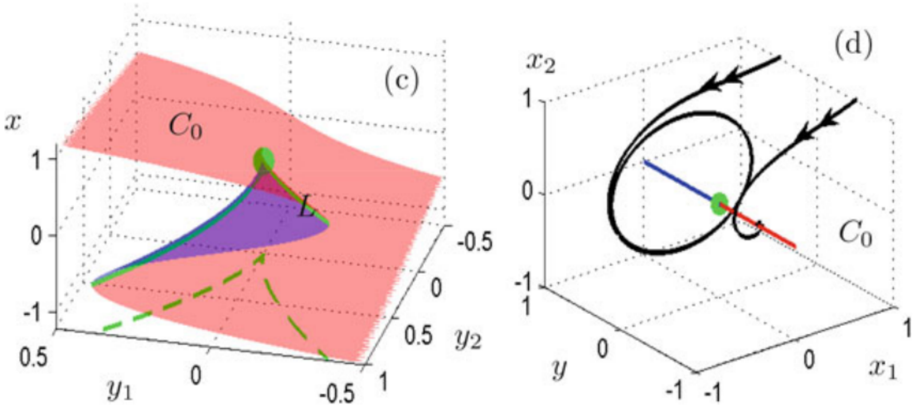

Figure 3.2: From Figure 3.3 (c) (d) of main text. (c) Cusp surface containing a curve of folds L (green, solid); a projection of the curve of folds onto the slow variables (green dashed curve). The cusp point itself (green dot) is at the origin. (d) Hopf bifurcation (green dot). Two fast subsystem trajectory segments starting at $(x_1, x_2, y) = (1, 1, \pm\frac{1}{2})$ are shown.

As most singularities are natural transition points from the slow to the fast subsystem or vice versa, we can define the idea of a *candidate orbit*.

Definition 3.3.3 (candidate orbit). A *candidate orbit* is a homeomorphic image $\gamma_0(t)$ of a real interval (a, b) with $a < b$ where

(i) The interval (a, b) is partitioned as $a = t_0 < t_1 < \cdots < t_m = b$.

(ii) The image of each subinterval (t_{i-1}, t_i) for $i = 1, \ldots, m$ is either a trajectory of the slow subsystem (3.3) or a trajectory of the fast subsystem (3.5).

(iii) The orbit $\gamma_0(a, b)$ has an orientation consistent with the that induced by the fast and slow flows on each subinterval $\gamma_0(t_{j-1}, t_j)$.

The trajectories are called *singular trajectories*.

Remark. Usually, An *orbit* is the set of points of the manifold, while a *trajectory* is a mapping whose image set is the orbit. However, in the text they seem to be used interchangeably.

what happens when the critical manifold is neutrally stable over a large subset, e.g. when it is elliptic?

3.4 Examples

Here we consider a few examples of fast-slow systems to illustrate the concepts introduced so far.

Example (slow manifold and Fenichel-Tikhonov). Consider the affine $(1, 1)$-fast-slow system

$$\begin{aligned} \varepsilon\dot{x} &= y - x, \\ \dot{y} &= 1, \end{aligned} \tag{3.29}$$

with critical manifold $C_0 = \{(x, y) \in \mathbb{R}^2 : y = x\}$. The solution of (3.29) can be calculated explicitly as

$$(x(\tau), y(\tau)) = \big(y(0) + \tau - \varepsilon + (x(0) - y(0) + \varepsilon)e^{-\tau/\varepsilon},\ y(0) + \tau\big). \tag{3.30}$$

Observe that since

$$x(\tau) - y(\tau) + \varepsilon = (x(0) - y(0) + \varepsilon)e^{-\tau/\varepsilon},$$

$x(0) - y(0) + \varepsilon = 0$ implies that $x(\tau) - y(\tau) + \varepsilon = 0$ for all times. Therefore, the curve

$$C_\varepsilon = \{(x, y) \in \mathbb{R}^2 : y = x + \varepsilon\}$$

is a slow manifold. Furthermore, $C_0 \neq C_\varepsilon$ for $\varepsilon > 0$ and $d_H(C_0, C_\varepsilon) = \mathcal{O}(\varepsilon)$ as $\varepsilon \to 0$ in the Hausdorff distance, since $x(\tau) - y(\tau) = \varepsilon$ for all τ.

Example. Consider the $(1, 1)$-fast–slow system

$$\begin{aligned} \varepsilon\dot{x} &= -(x + y^{1/\varepsilon}) = f(x, y, \varepsilon), \\ \dot{y} &= -y = g(x, y, \varepsilon), \end{aligned} \tag{3.24}$$

in a small fixed neighborhood

$$\mathcal{N} = \{\, |x| < \delta,\ |y| < \delta \,\}$$

of $(x, y) = (0, 0)$. For fixed $\delta \in (0, 1)$, we have that $y \in \mathcal{N}$ implies $|y|^{1/\varepsilon} < \delta^{1/\varepsilon} \to 0$ as $\varepsilon \to 0$. Therefore, the critical manifold of (3.24) is formally given by

$$C_0 = \{(x, y) \in \mathcal{N} : x = 0\}.$$

Certainly C_0 is a smooth curve, since it is just a line. For the full system (3.24), one has

$$y(\tau) = y(0)e^{-\tau},$$

so that

$$\dot{x} = -\frac{1}{\varepsilon}\big(x + y(0)^{1/\varepsilon}e^{-\tau/\varepsilon}\big) \Rightarrow x(\tau) = x(0)e^{-\tau/\varepsilon} - y(0)^{1/\varepsilon}\frac{\tau}{\varepsilon}e^{-\tau/\varepsilon}.$$

Hence, the x-axis is a smooth invariant manifold along which the fast dynamics

$$x(\tau) = x(0)e^{-\tau/\varepsilon} = x(0)e^{-t}$$

take place. Any slow manifold C_ε, which exists by Fenichel's theorem, near the origin is constructed from two curves. We substitute $\tau(y)$ from the above solution and eliminate τ to obtain

$$x = x(0)\left(\frac{y}{y(0)}\right)^r + ry^r \ln\left(\frac{y}{y(0)}\right), \quad r \equiv \frac{1}{\varepsilon}. \tag{3.31}$$

Example (FitzHugh-Nagumo)**.** Consider the three-dimensional *FitzHugh-Nagumo equation* (with parameter $I = 0$) given by

$$\begin{aligned} x_1' &= f_1(x, y) = x_2, \\ x_2' &= f_2(x, y) = sx_2 - x_1(x_1 - a)(1 - x_1) + y, \\ y' &= \varepsilon g(x, y) = \frac{\varepsilon}{s}(x_1 - \gamma y). \end{aligned} \tag{3.32}$$

The one-dimensional critical manifold is given by

$$C_0 = \{(x_1, x_2, y) \in \mathbb{R}^2 \times \mathbb{R} \mid x_2 = 0,\ y = x_1(x_1 - a)(1 - x_1)\}. \tag{3.33}$$

Solve

$$\frac{\partial}{\partial x_1}\big(x_1(x_1 - a)(1 - x_1)\big) = 3x_1^2 - 2(1 + a)x_1 + a = 0,$$

and we get the two nondegerate fold points of the critical manifold, with

$$x_{1,\pm} = \frac{1 + a \pm \sqrt{(1 + a)^2 - 3a}}{3}. \tag{3.34}$$

Although the 3D FitzHugh-Nagumo equation has no attractors, it has many interesting bounded invariant sets corresponding to traveling waves of the associated PDE.

Exercise (identifying candidate orbits). Identify different (classes of) candidate orbits for

(a) The unforced van der Pol equation and the van der Pol equation with constant forcing.

(b) The three-dimensional FitzHugh-Nagumo system.

Answer.

(a) The van der Pol equation with forcing term is given by

$$x'' + \mu(x^2 - 1)x' + x = a > 0. \tag{3.35}$$

Upon rescaling, we can put it into the fast-slow form

$$\begin{aligned} \varepsilon\dot{x} &= f(x, y) = y - \frac{x}{3} + x, \\ \dot{y} &= g(x, y) = a - x, \quad 0 < \varepsilon \ll 1. \end{aligned} \tag{3.36}$$

Setting $\varepsilon = 0$ in (3.36) gives the critical manifold

$$C_0 = \{(x, y) : f(x, y) = 0\} = \Big\{(x, y) : y = \frac{x^3}{3} - x\Big\}. \tag{3.37}$$

The fold points satisfy $f_x = 0$, i.e. $1 - x^2 = 0$, hence

$$x = \pm 1, \quad y = \frac{x^3}{3} - x = \begin{cases} -\frac{2}{3}, & x = 1, \\ \frac{2}{3}, & x = -1. \end{cases} \tag{3.38}$$

For the fast dynamics, stability of C_0 is determined by f_x:

$$f_x(x, y) = 1 - x^2 \begin{cases} < 0, & |x| > 1 \quad \text{(attracting)}, \\ > 0, & |x| < 1 \quad \text{(repelling)}. \end{cases} \tag{3.39}$$

Along C_0 we have $0 = f(x, y)$, so differentiating in slow time gives

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau} f(x(\tau), y(\tau)) = f_x\dot{x} + f_y\dot{y} = (1 - x^2)\dot{x} + (a - x),$$

hence, away from the folds $x = \pm 1$, the slow flow can be derived exactly as

$$\dot{x} = \frac{x - a}{1 - x^2}. \tag{3.40}$$

At the fold values $y = \pm 2/3$, the cubic $y = \frac{x^3}{3} - x$ has the additional roots

$$y = -\frac{2}{3}: \quad x^3 - 3x + 2 = 0 \;\Rightarrow\; x = 1 \text{ (double)}, \; x = -2, \tag{3.41}$$

$$y = \frac{2}{3}: \quad x^3 - 3x - 2 = 0 \;\Rightarrow\; x = -1 \text{ (double)}, \; x = 2. \tag{3.42}$$

Thus the singular jump targets are

$$(1, -2/3) \;\rightsquigarrow\; (-2, -2/3), \qquad (-1, 2/3) \;\rightsquigarrow\; (2, 2/3), \tag{3.43}$$

where $\rightsquigarrow$ denotes a fast jump at (approximately) constant y.

Candidate invariant sets / orbits:

(1) Equilibrium: Solving $\dot{y} = 0$ gives $x = a$, and then $f(x, y) = 0$ gives

$$(x, y) = (a,\ a^3/3 - a). \tag{3.44}$$

Linearizing (3.36) at (3.44) gives

$$J = \begin{pmatrix} \frac{1-a^2}{\varepsilon} & \frac{1}{\varepsilon} \\ -1 & 0 \end{pmatrix}, \qquad \operatorname{tr} J = \frac{1-a^2}{\varepsilon}, \quad \det J = \frac{1}{\varepsilon},$$

so the equilibrium is unstable for $0 < a < 1$ and stable for $a > 1$.

(2) Relaxation periodic orbit ($a = 0$): For $a = 0$, the equilibrium is at $(0, 0)$ and lies on the repelling sheet. A singular closed orbit is obtained by concatenating slow drift on the attracting sheets of C_0 with the fast jumps (3.43):

$$(2, \tfrac{2}{3}) \xrightarrow{\text{slow on } C_0^a} (1, -\tfrac{2}{3}) \xrightarrow{\text{fast}} (-2, -\tfrac{2}{3}) \xrightarrow{\text{slow on } C_0^a} (-1, \tfrac{2}{3}) \xrightarrow{\text{fast}} (2, \tfrac{2}{3}). \tag{3.45}$$

This is the candidate singular relaxation oscillation that persists as a limit cycle for sufficiently small ε.

(3) Relaxation-type cycle ($0 < a < 1$): For $0 < a < 1$ the equilibrium (3.44) lies on the repelling sheet, so one again expects a relaxation-type singular cycle obtained from the same fold-to-fold jumps (3.43), but with slow drift determined by the slow flow.

(4) Canards near $a \approx 1$: Since the equilibrium crosses the fold at $a = 1$, for small ε there can be canard trajectories that track the repelling sheet $|x| < 1$ for $\mathcal{O}(1)$ time before jumping; these form additional periodic orbits in a narrow parameter window near $a = 1$. Refer to Figure 3.3 for a demonstration.

> **Note.** The term *canard*, french for "duck", was coined in the early 1980s by French mathematicians because these solutions are “ducks” in the sense of being strange, rare, and hard to catch.

(b) To be finished...

⊛

Example (turning point). To be finished...

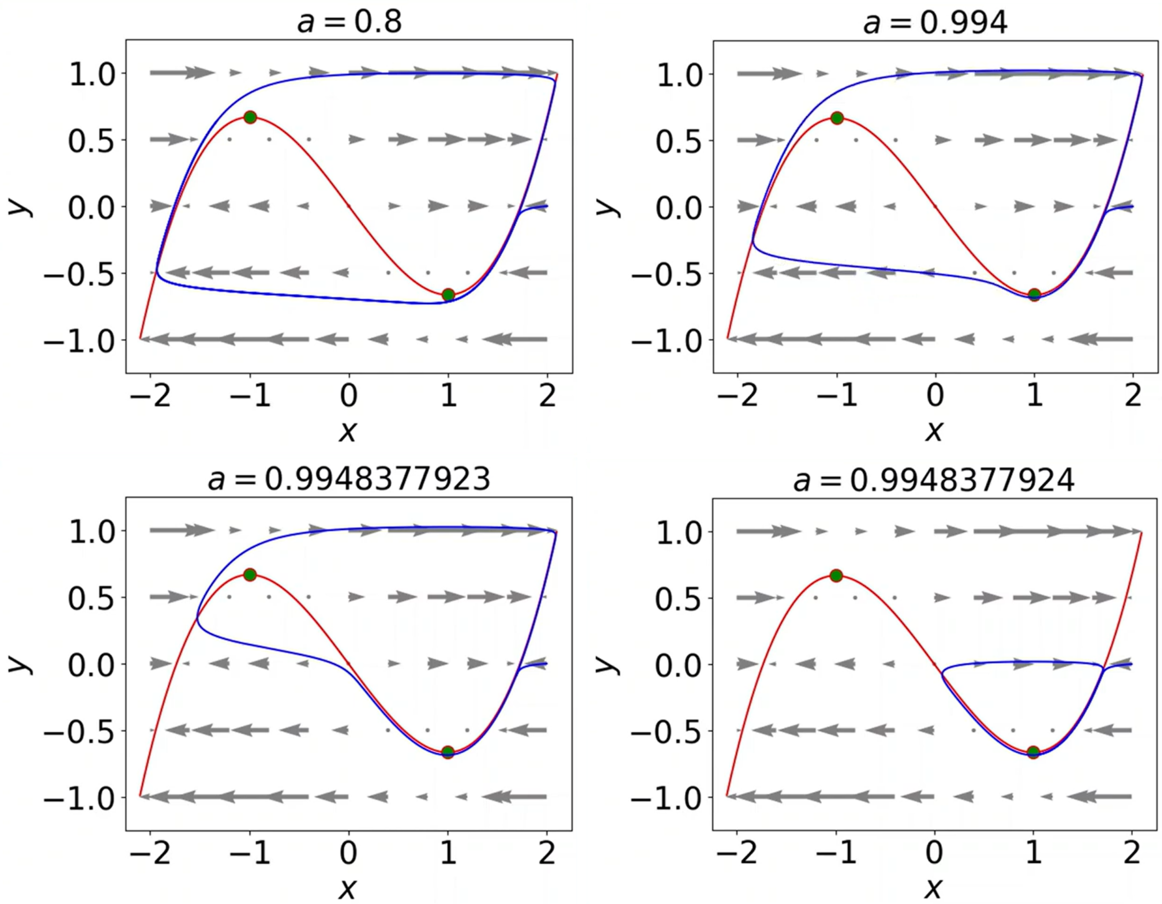

Figure 3.3: Canard solutions in the forced van der Pol equation for $\varepsilon = 0.04$. The blue line is the solution with $(x(0), y(0)) = (2, 0)$, and the red line is the critical manifold $y = x^3/3 - x$. The green dots mark the turning points from repeling to attracting branches. Various values of a near the fold value $a = 1$ are plotted. Grey arrows are the magnitude of $\dot{x}$. Even though the center branch is repelling, when a is sufficiently close to 1, solutions can track the repelling branch for a significant amount of time ($O(1)$) before jumping away. Source: https://www.youtube.com/watch?v=P1X2zkJDdUQ.

Chapter 4

Normal Forms

Introduction

Note.

(i) An important problem is now to know how to bring a fast-slow system into a **normal form**.

(ii) There is no complete general theory for what a "normal form" for a fast-slow system should be.

(iii) Some transformations can be used for specific classes of systems to bring them into fast-slow form.

4.1 The Normally Hyperbolic Case

4.2 Fold Points

Having demonstrated how to bring a normally hyperbolic system into fast-slow form, we now ask whether there are "normal forms" for singularities of the critical manifold. Here, we consider the case when the critical manifold contains a **fold point**, i.e., a point where normal hyperbolicity is lost.

4.3 Fold Curves

Next, we ask whether our previous normal form approach can be extended to folds in systems with more than one slow variable, that is, **fold curves**.

4.4 Systems of First Approximation

Remark. The next results truly separate the theory of *fast-slow systems* from *classical bifurcation theory*.

4.5 A Note On Linear Systems

We have applied various techniques to simplify fast-slow systems given in the **standard form**

$$\begin{aligned} \varepsilon\frac{\mathrm{d}x}{\mathrm{d}\tau} &= \dot{x} = f(x,y,\varepsilon), \\ \frac{\mathrm{d}y}{\mathrm{d}\tau} &= \dot{y} = g(x,y,\varepsilon), \end{aligned} \tag{4.1}$$

where $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, and $0 < \varepsilon \ll 1$. However, not all fast-slow systems can be transformed into this useful form, so we must consider techniques to analyze more general systems. To begin, consider a more general system of the form

$$\varepsilon \frac{\mathrm{d}z}{\mathrm{d}\tau} = \varepsilon \dot{z} = F(z,\varepsilon), \quad z \in \mathbb{R}^N, \quad F(\varepsilon) \in M_{N\times N}(\mathbb{R}). \tag{4.2}$$

We can make the decomposition $F(z) = F_0 + \varepsilon F_1(\varepsilon)$, where $F_0 = F(0)$. Our goal is to find a suitable transformation that brings this system into fast-slow form, conveniently written as

$$\begin{aligned} \varepsilon \dot{x} &= A_{11}x + \varepsilon A_{12}y, \\ \dot{y} &= A_{21}x + A_{22}y, \end{aligned} \tag{4.3}$$

where the matrix A given by

$$A = \begin{pmatrix} A_{11}(\varepsilon) & A_{12}(\varepsilon) \\ A_{21}(\varepsilon) & A_{22}(\varepsilon) \end{pmatrix} \in \mathrm{GL}(N;\mathbb{R}). \tag{4.4}$$

Example (Calculation of normal form). A concrete 2×2 example is given by

$$\frac{dz}{dt} = \underbrace{\begin{pmatrix} -1.2 & 1.6 \\ 0.6 & -0.8 \end{pmatrix}}_{=:F_0} z + \varepsilon \underbrace{\begin{pmatrix} 0.2 & -0.6 \\ 0.4 & -1.2 \end{pmatrix}}_{=:F_1} z, \tag{4.5}$$

(with $t = \tau/\varepsilon$). Although (**??**) is not in standard form, its phase portrait and time series show a clear fast transient followed by slow evolution.

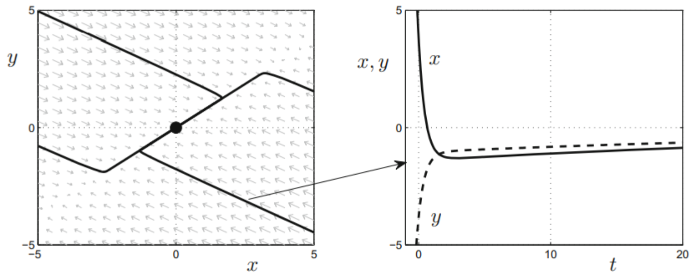

Figure 4.1: Phase portrait and time series for

Note (Splitting assumption). Let $\dim\big(\ker(F_0)\big) = n$ and set $N = m + n$. Assume that there exists the direct-sum decomposition

$$\mathrm{im}(F_0) \oplus \ker(F_0) = \mathbb{R}^{m+n}. \tag{4.6}$$

This is the linear analogue of the Fenichel-type splitting for normally hyperbolic fast-slow systems: every vector $z \in \mathbb{R}^{m+n}$ can be uniquely decomposed as

$$z = z_{\text{fast}} + z_{\text{slow}}, \quad z_{\text{fast}} \in \ker(F_0), \quad z_{\text{slow}} \in \mathrm{im}(F_0).$$

it separates the *fast directions* coming from $\mathrm{im}(F_0)$ from the *slow directions* coming from $\ker(F_0)$.

How can we construct fast and slow coordinates?

- Choose m linearly independent row vectors orthogonal to $\ker(F_0)$ and use them as rows of a matrix $Q \in \mathbb{R}^{m\times(m+n)}$. Then $\ker(F_0) = \ker(Q)$, and thus

$$\nu \in \ker(F_0) \Leftrightarrow F_0\nu = 0 \Leftrightarrow Q\nu = 0,$$

  so $x := Qz$ is a natural candidate for the fast variable(s). If z lies purely in the slow directions, then $x = Qz = 0$. For this, see the following example.

- Choose $P \in \mathbb{R}^{n\times(m+n)}$ from the *left nullspace* of F_0 so that

$$PF_0 = 0.$$

Then $y := Pz$ is the natural candidate for the slow variable(s). This is because if the linear fast dynamics is $\dot{x} = F_0 z$, then

$$\dot{y} = \frac{\mathrm{d}}{\mathrm{d}t}(Pz) = P\dot{z} = PF_0 z = 0.$$

Therefore, y is constant under the fast flow, and thus it is a slow variable.

- Stack the two maps into an invertible matrix (since the rows of Q and P span $\mathbb{R}^{m+n}$):

$$T := \begin{pmatrix} Q \\ P \end{pmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = Tz.$$

T is the linear analogue of *Fenichel coordinates.*

Example. For the same F_0, one computes

$$\ker(F_0) = \operatorname{span}\left(\begin{pmatrix} -0.8 \\ -0.6 \end{pmatrix}\right).$$

A vector orthogonal to this nullspace is $(0.6, -0.8)^\top$, hence one can take

$$Q = (0.6,\ -0.8) : \mathbb{R}^2 \to \mathbb{R}, \qquad x := Qz,$$

so that $x = 0$ on the line corresponding to the critical manifold candidate.

Theorem 4.5.1 (Kuehn 4.5.3 normal form for linear systems)**.** Suppose $F(0) = F_0$ satisfies the decomposition

$$\operatorname{im}(F_0) \oplus \ker(F_0) = \mathbb{R}^{m+n} = \mathbb{R}^N.$$

as mentioned above. Then the coordinate change

$$x = Qv, \quad y = Pv, \quad T := \begin{pmatrix} Q \\ P \end{pmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}$$

transforms the system (**??**) into the normal form

$$\begin{aligned} \varepsilon\dot{x} &= A_{11}(\varepsilon)x + \varepsilon A_{12}(\varepsilon)y, \\ \dot{y} &= A_{21}(\varepsilon)x + A_{22}(\varepsilon)y, \end{aligned} \tag{4.7}$$

where $A_{11}(0) \in \mathbb{R}^{m\times m}$ is nonsingular, $A_{12}(\varepsilon) \in \mathbb{R}^{m\times n}$, $A_{21}(\varepsilon) \in \mathbb{R}^{n\times m}$, and $A_{22}(\varepsilon) \in \mathbb{R}^{n\times n}$.

Proof. One convenient way to express the blocks is as follows: let the columns of V and W be bases of $\operatorname{im}(F_0)$ and $\ker(F_0)$, respectively, so that $T^{-1} = (V\ W)$. Then the blocks can be chosen as

$$\begin{aligned} A_{11}(\varepsilon) &:= QF_0V + \varepsilon\, QF_1(\varepsilon)V, \\ A_{12}(\varepsilon) &:= QF_1(\varepsilon)W, \\ A_{21}(\varepsilon) &:= PF_1(\varepsilon)V, \\ A_{22}(\varepsilon) &:= PF_1(\varepsilon)W. \end{aligned}$$

By construction $A_{11}(0) = QF_0V$ is invertible, and it is the only nonzero block of TF_0T^{-1}.

Explanation:

(1) Assume the splitting $\operatorname{Im}(F_0) \oplus \ker(F_0) = \mathbb{R}^{m+n}$, or equivalently, $\operatorname{Im}(F_0) \cap \ker(F_0) = \{0\}$. Notice that the restriction $F_0|_{\operatorname{Im}(F_0)} : \operatorname{Im}(F_0) \to \operatorname{Im}(F_0)$ is injective, for if $w \in \operatorname{Im}(F_0)$ and $F_0 w = 0$, then $w \in \operatorname{Im}(F_0) \cap \ker(F_0) = \{0\}$, hence $w = 0$. Since $\dim(\operatorname{Im}(F_0)) = m$ is finite, $F_0|_{\operatorname{Im}(F_0)}$ is a bijection, and hence an isomorphism.

Next, choose $Q \in \mathbb{R}^{m\times(m+n)}$ such that $\ker(Q) = \ker(F_0)$. Then the restriction $Q|_{\operatorname{Im}(F_0)}$:

$\mathrm{Im}(F_0) \to \mathbb{R}^m$ is also injective. Hence, $Q|_{\mathrm{Im}(F_0)}$ is an isomorphism onto $\mathbb{R}^m$. Consequently,

$$\mathrm{Im}(F_0) \xrightarrow{F_0} \mathrm{Im}(F_0) \xrightarrow{Q} \mathbb{R}^m$$

is an isomorphism. Writing this map in the basis V of $\mathrm{Im}(F_0)$ yields the matrix

$$A_{11}(0) = QF_0V \in \mathbb{R}^{m\times m},$$

which must therefore be invertible.

(2) Let

$$T := \begin{pmatrix} Q \\ P \end{pmatrix} \in \mathbb{R}^{(m+n)\times(m+n)}, \qquad T^{-1} = \begin{pmatrix} V & W \end{pmatrix},$$

where the columns of V form a basis of $\mathrm{Im}(F_0)$ and the columns of W form a basis of $\ker(F_0)$. Then

$$TF_0T^{-1} = \begin{pmatrix} Q \\ P \end{pmatrix} F_0 \begin{pmatrix} V & W \end{pmatrix} = \begin{pmatrix} QF_0V & QF_0W \\ PF_0V & PF_0W \end{pmatrix}. \tag{4.8}$$

Since $W \subset \ker(F_0)$, we have $F_0W = 0$, hence $QF_0W = 0$. Moreover, by construction P lies in the left nullspace of F_0, so $PF_0 = 0$, which implies $PF_0V = PF_0W = 0$. Therefore,

$$TF_0T^{-1} = \begin{pmatrix} QF_0V & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.9}$$

■

Note (summary). How do we know whether a (possibly nonlinear) system has a multiple time scale structure? Theorem 4.5.3 provides a partial answer for linear systems of the form (**??**). We should:

(i) Consider a linearized system.

(ii) Identify a small parameter.

(iii) Consider the eigenvalue structure of the singular limit system.

It often helps to identify the fast and slow variables and rewrite the system in fast-slow normal form.

Exercise (4.5.4). We start from Example 4.5.1:

$$\frac{dz}{dt} = F_0z + \varepsilon F_1z, \qquad F_0 = \begin{pmatrix} -1.2 & 1.6 \\ 0.6 & -0.8 \end{pmatrix}, \quad F_1 = \begin{pmatrix} 0.2 & -0.6 \\ 0.4 & -1.2 \end{pmatrix}, \tag{4.10}$$

with $t = \tau/\varepsilon$. Hence

$$\dot{z} = \frac{dz}{d\tau} = \frac{1}{\varepsilon}F_0z + F_1z, = \left(\frac{1}{\varepsilon}F_0 + F_1\right)z. \tag{4.11}$$

Then, we bring this system into fast-slow form by following the steps in the proof of Theorem 4.5.3:

(1) Compute $\ker(F_0)$ and pick Q: A direct computation (Example 4.5.2) gives

$$\ker(F_0) = \mathrm{span}\left(\begin{pmatrix} -0.8 \\ -0.6 \end{pmatrix}\right). \tag{4.12}$$

A vector orthogonal to $(-0.8, -0.6)^\top$ is $(0.6, -0.8)^\top$, so we take

$$Q = \begin{pmatrix} 0.6 & -0.8 \end{pmatrix}, \qquad x := Qz. \tag{E4}$$

(2) Pick P from the left nullspace and define y: Choose P so that $PF_0 = 0$ (a basis for the left nullspace of F_0). Let $P = (a \;\; b)$. Then

$$(a \;\; b)\begin{pmatrix} -1.2 & 1.6 \\ 0.6 & -0.8 \end{pmatrix} = (0 \;\; 0) \quad \Longrightarrow \quad -1.2a + 0.6b = 0 \Longrightarrow b = 2a.$$

We may take $a = 1$, hence

$$P = \begin{pmatrix} 1 & 2 \end{pmatrix}, \qquad y := Pz. \tag{4.13}$$

(3) Form T and its inverse: Define

$$T := \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} 0.6 & -0.8 \\ 1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = Tz. \tag{4.14}$$

One checks $\det(T) = 2 \neq 0$, so T is invertible and

$$T^{-1} = \begin{pmatrix} 1 & 0.4 \\ -0.5 & 0.3 \end{pmatrix} = \begin{pmatrix} V & W \end{pmatrix}, \qquad V = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix}, \quad W = \begin{pmatrix} 0.4 \\ 0.3 \end{pmatrix}. \tag{4.15}$$

(Indeed, V spans $\operatorname{im}(F_0)$ and W spans $\ker(F_0)$.)

(4) Compute the normal-form blocks: From the proof of Theorem 4.5.3, in (x, y)-coordinates

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = T\left(\frac{1}{\varepsilon}F_0 + F_1\right)T^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\varepsilon}QF_0V + QF_1V & QF_1W \\ PF_1V & PF_1W \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{4.16}$$

Therefore the standard fast-slow form

$$\varepsilon\dot{x} = A_{11}(\varepsilon)x + \varepsilon A_{12}(\varepsilon)y, \qquad \dot{y} = A_{21}(\varepsilon)x + A_{22}(\varepsilon)y \tag{4.17}$$

is obtained with

$$A_{11}(\varepsilon) = QF_0V + \varepsilon QF_1V, \quad A_{12}(\varepsilon) = QF_1W, \quad A_{21}(\varepsilon) = PF_1V, \quad A_{22}(\varepsilon) = PF_1W. \tag{4.18}$$

Now compute each term (here $F_1(\varepsilon) = F_1$ is constant):

$$QF_0V = -2, \qquad QF_1V = -\frac{1}{2}, \qquad QF_1W = \frac{1}{10}, \tag{4.19}$$

$$PF_1V = \frac{5}{2}, \qquad PF_1W = -\frac{1}{2}. \tag{4.20}$$

Hence

$$A_{11}(\varepsilon) = -2 - \frac{\varepsilon}{2}, \qquad A_{12}(\varepsilon) = \frac{1}{10}, \qquad A_{21}(\varepsilon) = \frac{5}{2}, \qquad A_{22}(\varepsilon) = -\frac{1}{2}. \tag{4.21}$$

The final answer is given as:

$$\boxed{\varepsilon\dot{x} = \left(-2 - \frac{\varepsilon}{2}\right)x + \varepsilon\left(\frac{1}{10}\right)y, \qquad \dot{y} = \left(\frac{5}{2}\right)x - \left(\frac{1}{2}\right)y.} \tag{4.22}$$

Chapter 5

Direct Asymptotic Methods

Chapter 6

Tracking Invariant Manifolds

Chapter 7

The Blowup Method

Main idea: Introduce geometric desingularization of nonhyperbolic equilibrium points using the so-called blowup method.

7.1 Basics

Note. There are two unrelated notions of "blow up" in mathematics.

- Finite-time existence of solution in the context of ODEs/PDEs.
- Geometric desingularization of nonhyperbolic equilibrium points using the so-called blowup method.

Main idea: Blowup is a geometric *desingularization* tool for nonhyperbolic equilibria, where we "replace" a singular point by a manifold (e.g., a circle/sphere) so that the induced dynamics becomes partially or fully hyperbolic.

Example (nonhyperbolic equilibrium points)**.** Consider a planar system

$$\begin{aligned} \dot{z}_1 &= z_1^2 - 2z_1 z_2 := F_1(z_1, z_2), \\ \dot{z}_2 &= z_2^2 - 2z_1 z_2 := F_2(z_1, z_2). \end{aligned} \tag{7.1}$$

Let $F(z_1, z_2) = (F_1(z_1, z_2), F_2(z_1, z_2))$. The origin $(0, 0)$ is a nonhyperbolic equilibrium point since the Jacobian matrix

$$\mathrm{D}F(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.2}$$

The simplest examples come from planar systems with nice analytical properties. Consider $\dot{z} = F(z)$, with $z \in \mathbb{R}^2$, $F \in C^\infty$, and $F(0) = 0$.

Definition 7.1.1 (polar blowup)**.** Consider the polar coordinate transformation $\phi : S^1 \times I \to \mathbb{R}^2$ given by

$$\phi(\theta, r) = (r\cos\theta, r\sin\theta), \tag{7.3}$$

where I is a possibly infinite interval in $\mathbb{R}$ containing 0. The *polar blowup* of the vector field F is the map $\hat{F}$ given by

$$\hat{F}(\theta, r) = \left(\mathrm{D}\phi^{-1}_{(\theta,r)} \circ F \circ \phi\right)(\theta, r) \tag{7.4}$$

for $r \neq 0$ and extended continuously to $r = 0$. $\hat{F}$ is a vector field on the manifold $S^1 \times I$.

Remark. The polar coordinate change is not one-to-one at the origin. Therefore, we have to show that this is a good definition for $r = 0$.

Remark. Since ϕ is a diffeomorphism for $r \neq 0$, the vector fields F and $\hat{F}$ are equivalent away from the origin, i.e. on $S^1 \times (0, \infty)$. The idea is to study the dynamics of $\hat{F}$ on the "blown-up" space $S^1 \times I$, especially on the set $r = 0$, called the *blowup sphere* or *blowup circle* in this case.

Example (Finding $\hat{F}$)**.** We have

$$F \circ \phi(\theta, r) = \begin{pmatrix} r^2 \cos^2\theta - 2r^2 \cos\theta \sin\theta \\ r^2 \sin^2\theta - 2r^2 \cos\theta \sin\theta \end{pmatrix}, \tag{7.5}$$

and

$$\begin{aligned} \hat{F}(\theta, r) &= \begin{pmatrix} -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \\ \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} r^2 \cos^2\theta - 2r^2 \cos\theta \sin\theta \\ r^2 \sin^2\theta - 2r^2 \cos\theta \sin\theta \end{pmatrix} \\ &= \begin{pmatrix} 3r \sin\theta \cos\theta \left(\sin\theta - \cos\theta\right) \\ \frac{1}{4} r^2 \left(\cos\theta + 3\cos(3\theta) + \sin\theta - 3\sin(3\theta)\right) \end{pmatrix}. \end{aligned} \tag{7.6}$$

Define

$$\overline{F} = \frac{1}{r} \hat{F} = \begin{pmatrix} 3 \sin\theta \cos\theta \left(\sin\theta - \cos\theta\right) \\ \frac{1}{4} r \left(\cos\theta + 3\cos(3\theta) + \sin\theta - 3\sin(3\theta)\right) \end{pmatrix}. \tag{7.7}$$

The phase portrait now has six equilibrium points on the blowup circle $r = 0$, located at $\theta = 0, \frac{\pi}{3}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{5\pi}{3}$. All of them are hyperbolic saddles as we can verify:

$$\mathrm{D}\overline{F}(\theta, 0) = \begin{pmatrix} ? & 0 \\ 0 & \frac{1}{4} \left(\cos\theta + 3\cos(3\theta) + \sin\theta - 3\sin(3\theta)\right) \end{pmatrix}. \tag{7.8}$$

Then

$$\begin{aligned} \mathrm{D}\overline{F}(0, 0) &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ \mathrm{D}\overline{F}_{\frac{\pi}{2}, 0} &= \begin{pmatrix} -3 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned} \tag{7.9}$$

Note. Dividing by r will not change the qualitative structure of the phase portrait of $\hat{F}$ away from $r = 0$. This is further pursued in Section 7.7.

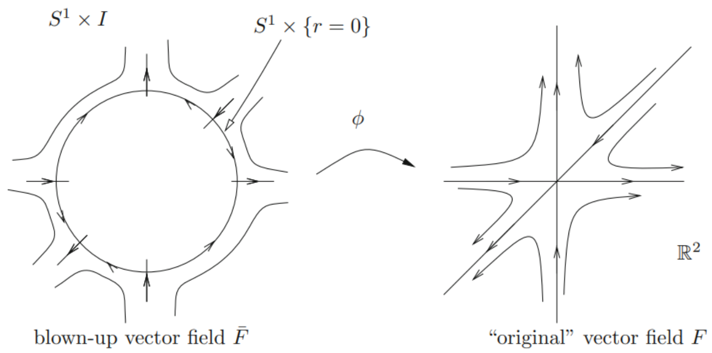

Figure 7.1: Polar blownup at $(0, 0)$.

Exercise. Complete the phase portrait of $\overline{F}$ on $S^1 \times I$.

Answer. ⊛

Result: All equilibrium points of new vector field are hyperbolic saddles. However, a single equilibrium is blown up to a full circle, which has six equilibrium points on it.

Algebraically, $\hat{F}$ is the map that makes the diagram Fig 7.2 induced by ϕ commute.

$$\begin{tikzcd}
T(S^1 \times I) \arrow[r, "D\phi"] & T\mathbb{R}^2 \\
S^1 \times I \arrow[u, dashed, "\hat{F}"] \arrow[r, "\phi"'] & \mathbb{R}^2 \arrow[u, "F"']
\end{tikzcd}$$

Figure 7.2: Commutative diagram relating ϕ, $D\phi$, and their lifts

Definition 7.1.2 (generalized polar blowup). Let F be a smooth vector field on $\mathbb{R}^n$ with $F(0) = 0$. Consider the generalized polar coordinate transformation $\phi : S^{n-1} \times I \to \mathbb{R}^n$ given by

$$\phi(\overline{z}_1, \overline{z}_2, \dots \overline{z}_n, r) = (r\overline{z}_1, r\overline{z}_2, \dots r\overline{z}_n), \tag{7.10}$$

where $\sum_{i=1}^n \overline{z}_i^2 = 1$, and I is a possibly infinite interval in $\mathbb{R}$ containing 0. The (generalized) *polar blowup* of the vector field F is the map $\hat{F}$ given by

$$\hat{F}(\overline{z}_1, \overline{z}_2, \dots \overline{z}_n, r) = \left(\mathrm{D}\phi^{-1}_{(\overline{z}_1, \overline{z}_2, \dots \overline{z}_n, r)} \circ F \circ \phi\right)(\overline{z}_1, \overline{z}_2, \dots \overline{z}_n, r). \tag{7.11}$$

For example, the ambient coordinates could be explicitly parametrized as

$$\begin{aligned}
\overline{z}_1 &= \cos\theta_1, \\
\overline{z}_2 &= \sin\theta_1 \cos\theta_2, \\
\overline{z}_3 &= \sin\theta_1 \sin\theta_2 \cos\theta_3, \\
&\vdots \\
\overline{z}_{n-1} &= \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{n-2} \cos\theta_{n-1}, \\
\overline{z}_n &= \sin\theta_1 \sin\theta_2 \cdots \sin\theta_{n-2} \sin\theta_{n-1},
\end{aligned} \tag{7.12}$$

where $\theta_i \in [0, \pi]$ for $1 \le i \le n-2$ and $\theta_{n-1} \in [0, 2\pi)$. In physics, this is known as the *hyperspherical coordinates*.

Theorem 7.1.1 (well-definedness). Let F be a smooth vector field on $\mathbb{R}^n$ with $F(0) = 0$. Then the map $\hat{F}$ in defnition 7.1.2 is a well-defined smooth vector field on $S^{n-1} \times I$.

Proof. To show this, we first define two vector fields:

$$R = \sum_{i=1}^n \overline{z}_i \frac{\partial}{\partial \overline{z}_i}, \quad \text{and} \quad TS^{n-1} = \{v \in T\mathbb{R}^n : v \perp R\}. \tag{7.13}$$

Here R represents the *radial* component and V_{ij} the *rotational* components of the vector field F. ■

Note.

- Equation () gives an explicit way to calculate polar blowup.
- Dividing it by r^k with k as large as possible appears to be useful to desingularize a vector field.

For the following discussion, it is useful to understand the definition of *jets*.

Note.

Example (When blowing up fails to yield hyperbolic points). Consider the following system of ODEs:

$$\dot{z}_1 = z_2 \qquad\qquad = F_z(z_1, z_2), \tag{7.14}$$
$$\dot{z}_2 = z_1^2 + z_1 z + 2 \qquad\qquad = F_2(z_1, z_2), \tag{7.15}$$

with $F(z_1, z_2) = (F_1(z_1, z_2), F_2(z_1, z_2))$. We have $F(0) = 0$, and the Jacobian matrix at the origin is

$$\mathrm{D}F(0,0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{7.16}$$

which has a double zero eigenvalue. Thus, the origin is a nonhyperbolic equilibrium. We attempt a blow up at $(0,0)$:

$$\begin{aligned}
\hat{F}(\theta, r) &= \left(\mathrm{D}\phi^{-1}_{(\theta,r)} \circ F \circ \phi\right)(\theta, r) \\
&= \begin{pmatrix} -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \\ \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} r\sin\theta \\ r^2\cos^2\theta + r^2\cos\theta\sin\theta \end{pmatrix} \\
&= \begin{pmatrix} r\cos^3\theta + r\cos^2\theta\sin\theta - \sin^2\theta \\ r\cos\theta\sin\theta\,(1 + r\cos\theta + r\sin\theta) \end{pmatrix}.
\end{aligned} \tag{7.17}$$

Remark. This is expected since $\mathrm{D}F_{(0,0)} \neq 0$ the zero matrix. Hence, the lowest order terms in the Taylor expansion of F at $(0,0)$ are linear ($j_1(F)(0) \neq 0$), and so $\hat{F} = \overline{F}$.

The two singular points of $\overline{F}$ on $S^1 \times \{r = 0\}$ are $(\pi, 0)$ and $(-\pi, 0)$. We calculate that

$$\mathrm{D}\overline{F}(\theta, 0) = \left(hi\right) \tag{7.18}$$

Therefore,

$$\mathrm{D}\overline{F}(\pi, 0) = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

The two new equilibrium points are still not hyperbolic.

The Figure 7.2 in the textbook shows the same dynamics rewritten on a space where $(0,0)$ has been replaced by a circle of directions. It would be great if we could blowup again at the points $(\pi, 0)$ and $(-\pi, 0)$ and see what we get.

Proposition 7.1.1. Define

$$T_1 \colon S^1 \times \left(-\tfrac{1}{2}, \infty\right) \longrightarrow S^1 \times \left(\tfrac{1}{2}, \infty\right), \qquad T_1(\alpha, r) = (\alpha, r + 1),$$

and

$$\psi \colon \mathbb{R}^2 \to \mathbb{R}^2, \qquad \psi(z) = z - \frac{z}{\|z\|}.$$

Then the diagram

$$\begin{tikzcd}
 & \mathbb{R}^2 \\
\{z : \|z\| > \tfrac{1}{2}\} \arrow[ur, "\psi"] & S^1 \times \left(-\tfrac{1}{2}, \infty\right) \arrow[u, "\phi"'] \arrow[l, "\phi^{-1}\circ T_1"]
\end{tikzcd}$$

is commutative, i.e.

$$\psi \circ \phi^{-1} \circ T_1 = \phi.$$

Furthermore, blowing up by ϕ and dividing by r^k is equivalent to blowing up by ψ and dividing by $(\|z\| - 1)^k$.

The next exercise asks us to prove Proposition 7.1.1.

Proof. Write $z = \|z\|\alpha$, where $\alpha = \exp(i \arg z) \in S^1$ is the phase. Then

$$\psi(z) = z - \frac{z}{\|z\|} = (\|z\| - 1)\alpha$$

sends the radius $\|z\| \mapsto \|z\| - 1$. The blowup map $\phi(\alpha, r) = r\alpha$ has an inverse given by

$$\phi^{-1}(z) = \left(\frac{z}{\|z\|}, \|z\|\right).$$

Then, we can easily check commutativity of the diagram:

$$\psi \circ \phi^{-1} \circ T_1(\alpha, r) = \psi \circ \phi^{-1}(\alpha, r+1) = \psi((r+1)\alpha) = r\alpha = \phi(\alpha, r),$$

for all $(\alpha, r) \in S^1 \times \left(-\frac{1}{2}, \infty\right)$. Finally, the radius r in the ψ-picture is $\|z\| - 1$, so dividing by r^k in the ϕ-picture is equivalent to dividing by $(\|z\| - 1)^k$ in the ψ-picture. ■

After we pull back the vector field by ψ and desingularize, the new equilibrium points lie on the unit circle, since

$$\psi(z) = 0 \Leftrightarrow z = \frac{z}{\|z\|} \Leftrightarrow \|z\| = 1.$$

Note (repeated blowups). Now we can do repeated blowups in $\mathbb{R}^2$ at nonhyperbolic equilibrium points using ψ: Let z_1 be an equilibrium point on the unit circle $\{z \in \mathbb{R}^2 \mid \|z\| = 1\}$ after the first blowup. Then blowup using $T_{z_0} \circ \psi$, where $T_{z_0}(z) = z + z_0$.

At the n-th step we choose a point z_n, which is typically an equilibrium point, and define the next blowup map $\psi_n = T_{z_n} \circ \phi$ as a translate of ψ.

Example. Refer to Figure 7.3. There is a degenerate equilibrium at the origin. First blow up with ψ to get equilibrium points on the exceptional circle: $\psi^{-1}(0) = \{z \mid \|z\| = 1\}$. Then we shift the origin to $(1, 0)$ and blow up again with $\psi_2 = T_{(1,0)} \circ \psi$. Now the preimage becomes $(\phi_1 \circ \phi_2)^{-1}(0)$. The dashed lines represent the inner annulus of radius $\frac{1}{2}$, and the solid lines represent the unit circle.

Notation. Denote the sequences of blown-up and rescaled blown-up vector fields by $\hat{F}_{[n]}$ and $\overline{F}_{[n]}$, respectively, i.e.

$$\hat{F}_{[n]} = \left(\mathrm{D}(\psi_1 \circ \cdots \circ \psi_n)^{-1} \circ F \circ (\psi_1 \circ \cdots \circ \psi_n)\right), \quad \overline{F}_{[n]} = \frac{1}{r^k} \hat{F}_{[n]}, \quad k \in \mathbb{N}.$$

Proposition 7.1.2. Let $\Gamma_n = (\psi_1 \circ \cdots \circ \psi_n)^{-1}(0)$. Then the following hold:

(R1) There exists only one connected component of $\mathbb{R}^2 \setminus \Gamma_n$ that is unbounded. Call it A_n.

(R2) $\partial A_n \subset \Gamma_n$ and ∂A_n is homeomorphic to S^1.

(R3) A_n consists of finitely many smooth arcs meeting transversally.

(R4) $(\psi_1 \circ \cdots \circ \psi_n)|_{A_n}$ is an analytic diffeomorphism sending A_n onto $\mathbb{R}^2 \setminus \{0\}$.

(R5) $\widehat{F}_{[n]}$ restricted to A_n is analytically equivalent to F restricted to $\mathbb{R}^2 \setminus \{0\}$.

(R6) Up to rescaling by a positive function, the analytic equivalence on A_n holds also for $\overline{F}_{[n]}$.

Definition 7.1.3 (Lojasiewicz inequality). A vector field F on $\mathbb{R}^2$ is said to satisfy a *Lojasiewicz*

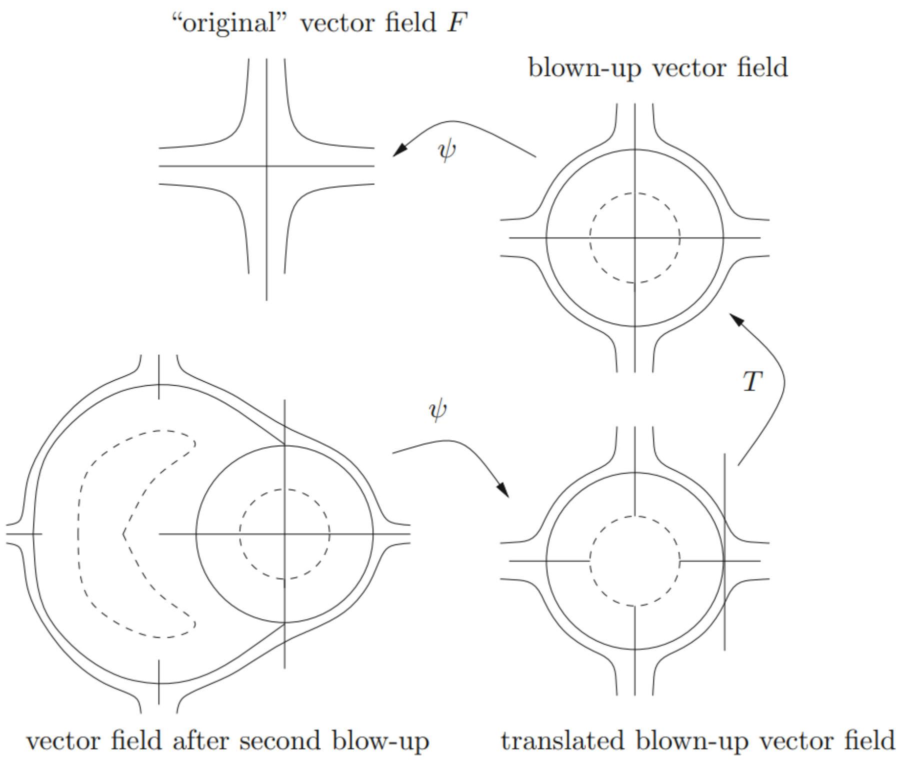

Figure 7.3: Repeated blow-ups.

inequality if there exist constants $c > 0$, $k \in \mathbb{N}$, and a neighborhood U of the origin such that

$$\|F\| \geq c\|z\|^k, \quad \text{for all } z \in U. \tag{7.19}$$

Note (Machine learning). A special case of the Lojasiewicz inequality, called the *Polyak inequality*, which is commonly used to prove linear convergence of gradient descent algorithms.

Definition 7.1.4 (Polyak inequality). f is a function of type $\mathbb{R}^d \to \mathbb{R}$, and has a continuous derivative ∇f.

$X^\star$ is the subset of $\mathbb{R}^d$ on which f achieves its global minimum (if one exists). Throughout this section we assume such a global minimum value $f^\star$ exists, unless otherwise stated. The optimization objective is to find some point $x \in X^\star$.

$\mu, L > 0$ are constants.

∇f is L-Lipschitz continuous iff

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \qquad \forall x, y.$$

f is μ-strongly convex iff

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{\mu}{2}\|y - x\|^2, \qquad \forall x, y.$$

f is μ-PL (where "PL" means "Polyak–Lojasiewicz") iff

$$\frac{1}{2}\|\nabla f(x)\|^2 \geq \mu\big(f(x) - f(x^\star)\big), \qquad \forall x.$$

Gradient descent is an important unconstrained optimization algorithm for minimizing the loss function in machine learning. The following theorem shows its convergence.

Theorem 7.1.2 (Linear convergence of gradient descent)**.** If f is μ-PL and ∇f is L-Lipschitz, then gradient descent with constant step size η

$$x_{k+1} = x_k - \eta \nabla f(x_k)$$

converges linearly as

$$f(x_k) - f(x^\star) \leq \big(1 - 2\mu\eta(1 - L\eta/2)\big)^k \big(f(x_0) - f(x^\star)\big), \qquad \eta \in (0, 2/L).$$

The convergence is the fastest when $\eta = 1/L$, at which point

$$f(x_k) - f(x^\star) \leq \big(1 - \mu/L\big)^k \big(f(x_0) - f(x^\star)\big).$$

Proof. Since ∇f is L-Lipschitz, we have the parabolic upper bound

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{L}{2}\|x_{k+1} - x_k\|^2.$$

Plugging in the gradient descent step,

$$\begin{aligned} f(x_{k+1}) - f(x_k) &\leq \langle \nabla f(x_k), -\eta \nabla f(x_k) \rangle + \frac{L}{2}\| - \eta \nabla f(x_k)\|^2 \\ &= \left(\frac{L\eta^2}{2} - \eta\right) \|\nabla f(x_k)\|^2 \\ &\leq 2\mu \left(\frac{L\eta^2}{2} - \eta\right) \big(f(x_k) - f(x^\star)\big), \end{aligned}$$

where the last inequality uses the μ-PL condition. Thus,

$$f(x_k) - f(x^\star) \leq \big(1 - 2\mu\eta(1 - L\eta/2)\big)^k \big(f(x_0) - f(x^\star)\big).$$

■

Analytic vector fields satisfy a Lojasiewicz inequality at an isolated equilibrium point.

Theorem 7.1.3 (Dum78, Dum93)**.** If F is a vector field on $\mathbb{R}^2$ that satisfies a Lojasiewicz inequality, then there exists a finite sequence of blowups desingularizing F. Mainore precisely, there exists a sequence

$$\psi_1 \circ \cdots \circ \psi_n$$

defining a rescaled blown-up vector field $\widehat{F}_{[n]}$ such that the equilibrium points of $\widehat{F}_{[n]}$ on ∂A_n are either

- hyperbolic or partially hyperbolic isolated equilibrium points $\bar{z}$, such that

  $$j_\infty\big(\widehat{F}_{[n]}|_{W^c}\big)(\bar{z}) \neq 0,$$

  where W^c is a center manifold for $\widehat{F}_{[n]}$ at $\bar{z}$, or
- regular smooth closed curves with boundary along which $\widehat{F}_{[n]}$ is normally hyperbolic.

Definition 7.1.5. The rescaled blown-up vector field $\overline{F}_{[n]}$ in the above theorem is called the *desingularization* of F.

7.2

7.3

7.4

7.5

7.6

7.7 Remarks On Rescaling

Note. After rescaling, trajectories follow the same curves but at different speeds; anything depending on time-of-flight is not preserved.

Multiplication transformation for desingularization:

$$\dot{r} = r^k(\cdots), \qquad \dot{\theta} = r^k(\cdots) \;\longrightarrow\; \dot{r} = (\cdots), \qquad \dot{\theta} = (\cdots),$$

or

$$\dot{x} = x^k(\cdots), \qquad \dot{y} = x^k(\cdots) \;\longrightarrow\; \dot{x} = (\cdots), \qquad \dot{y} = (\cdots).$$

Example (van der Pol equation)**.** Consider the van der Pol equation (again)

$$\varepsilon \frac{dx}{d\tau} = \varepsilon \dot{x} = y - \frac{x^3}{3} + x, \qquad \frac{dy}{d\tau} = \dot{y} = -x. \tag{7.20}$$

Setting $\varepsilon = 0$ yields $C_0 = \{y = \frac{x^3}{3} - x\}$. Differentiating the algebraic constraint gives $\dot{y} = (x^2 - 1)\dot{x}$. Therefore, the slow subsystem can be written as

$$\dot{x} = \frac{x}{1 - x^2}. \tag{7.21}$$

Away from the two fold points at $x = \pm 1$, one expects that the ODE is related to the ODEs $\dot{x} = x$ or $\dot{x} = -x$ obtained from multiplication by $\pm(1 - x^2)$.

Theorem 7.7.1 (Chi10)**.** Let $J \subset \mathbb{R}$ be an interval with $0 \in J$ and suppose $\gamma = \gamma(t)$ solves (**??**). Then the function $B\colon J \to \mathbb{R}$ defined by

$$B(t) := \int_0^t \frac{1}{G(\gamma(s))}\, ds \tag{7.52}$$

is invertible on its range $K \subset \mathbb{R}$. Let $\beta\colon K \to J$ denote the inverse of B. Then

$$\beta'(t) = G(\gamma(\beta(t))),$$

holds for all $t \in K$. Furthermore, $\tilde{\gamma}(t) := \gamma(\beta(t))$ solves the above equation.

Proof. Note that $B'(t) = 1/G(\gamma(t))$ is a continuous positive function. Hence, $B(t)$ is invertible. For the inverse $\beta(t)$, one obtains

$$\beta'(t) = \frac{1}{B'(\beta(t))} = G(\gamma(\beta(t))).$$

To check the last statement involves another direct calculation:

$$\tilde{\gamma}'(t) = \beta'(t)\gamma'(\beta(t)) = G(\gamma(\beta(t)))F(\gamma(\beta(t))) = G(\tilde{\gamma}(t))F(\tilde{\gamma}(t)).$$

■

Remark. It is usually not possible to find an explicit formula for the time rescaling we need.