SMOOTH LINEARIZATION AND NORMAL FORMS

1. Formal normal form

Let

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} : \mathbb{R}^m \supset U \to \mathbb{R}^m$$

such that O is a hyperbolic fixed point, and let $g = df(O)$ be the linearization of f at O. We discuss the possibility of *smoothly* conjugate f to g. We start with a theory of formal power series.

We call $k = (k_1, \cdots, k_m) \in \mathbb{N}_0^m$ a multi-index, and given $x = (x_1, \cdots, x_m) \in \mathbb{R}^m$, we denote

$$x^k = x_1^{k_1} \cdots x_m^{k_m}.$$

We also denote $|k| = k_1 + \cdots k_m$. Multi-indices is convenient for writing multi-variable power series.

We say a vector $\lambda = (\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m$ (or $\mathbb{C}^m$) is non-resonant up to order N if for any $i = 1, \cdots, n$ and $k \in \mathbb{N}_0^m$ with $|k| \le N$, we have $\lambda_i \neq \lambda_1^{k_1} \cdots \lambda_m^{k_m}$ (except the trivial one $\lambda_i = \lambda_i$).

Proposition 1. *Assume that $\lambda \in \mathbb{R}^m$ is non-resonant up to all orders, and let $f = (f_1, \cdots, f_m)$ be a formal power series of the form*

$$f_i(x) = \sum_{k \in \mathbb{N}_0^m} f_{i,k} x^k,$$

such that $df(O) = \operatorname{diag}\{\lambda_1, \cdots, \lambda_m\}$, then there exists a formal power series $h = (h_1, \cdots, h_m)$

$$h_i(x) = \sum_{k \in \mathbb{N}_0^m} h_{i,k} x^k$$

such that

$$h \circ f = g \circ h$$

as formal power series, where $g = (\lambda_1 x_1, \cdots, \lambda_n x_n)$ is the linearization of f.

Proof. Assume that

$$h_i(x) = \sum_{j=1}^{m} a_{ij} x_j + \sum_{|k|>1} h_{i,k} x^k.$$

By definition

$$h_i(f(x)) = \sum_{i=1}^{m} a_{ij} f_j + \sum_{|k|>1} h_{i,k} x^k, \quad g_i(h(x)) = \lambda_i h_i(x).$$

Compare linear terms of the equation $h \circ f = g \circ h$, we get

$$\sum_{j=1}^{m} a_{ij} \lambda_j x_j = \sum_{j=1}^{m} \lambda_i a_{ij} x_j.$$

From the first order non-resonance relation $\lambda_i \neq \lambda_j$ for $i \neq j$, we obtain $a_{ij} = 0$ for $i \neq j$. Hence it suffices to assume

$$h_i(x) = a_i x_i + \sum_{|k|>1} h_{i,k} x^k.$$

We have

$$h_i \circ f = a_i(\lambda_i x_i + \sum_{|k|>1} f_{i,k} x^k) + \sum_{|k|>1} h_{i,k} (\lambda_1 x_1 + \sum_{|j_1|>1} f_{1,j_1} x^{j_1})^{k_1} \cdots (\lambda_m x_m + \sum_{|j_m|>1} f_{m,j_m} x^{j_m})^{k_m}$$

and

$$g_i \circ h = \lambda_i (a_i x_i + \sum_{|k|>1} h_{i,k} x^k).$$

Assume that $h_i \circ f = g_i \circ h$ already hold up to degree N, we show that one can choose $h_{i,k}$ for $|k| = N + 1$ such that $h_i \circ f = g_i \circ h$ up to order $N + 1$.

Comparing the terms of order $N + 1$ we have

$$\sum_{|k|=N+1} \left(a_i f_{i,k} x^k + h_{i,k} \lambda^k x^k + C_{i,k} x^k \right) = \sum_{|k|=N+1} \lambda_i h_{i,k} x^k,$$

where $C_{i,k}$ is a polynomial in terms of $h_{i,k}$ and $f_{i,k}$ with $|k| \leq N$. We then get

$$h_{i,k}(\lambda_i - \lambda^k) = f_{i,k} + C_{i,k},$$

which can be solved if $\lambda_i \neq \lambda^k$. The proposition follows by induction. □

The above proposition easily generalizes to the case when the non-resonance conditions are not satisfied.

Proposition 2. *Assume that $\lambda \in \mathbb{R}^m$ satisfies $\lambda_i \neq \lambda_j$ for $i \neq j$. Then for the formal power series $f = (f_1, \cdots, f_m)$, $f(O) = O$, and $df(O) = \operatorname{diag}\{\lambda_1, \cdots, \lambda_m\}$, there exists formal power series h and g such that*

$$h \circ f = g \circ h$$

as formal power series, and $g_i = \sum g_{i,k} x^k$ satisfies $g_{i,k} = 0$ whenever $\lambda_i \neq \lambda^k$ (we say g contains only resonant terms). The power series g is a formal normal form of f near O.

Proof. We proceed with induction as in Proposition 1. Indeed, we get the following equation for $|k| = N + 1$:

$$h_{i,k}(\lambda_i - \lambda^k) = a_i f_{i,k} - g_{i,k} a^k + C'_{i,k},$$

where $C'_{i,k}$ depends on $h_{i,k}$, $f_{i,k}$ and $g_{i,k}$ with $|k| \leq N$. If $\lambda_i \neq \lambda^k$, we set $g_{i,k} = 0$ and solve for $h_{i,k}$ as before. If $\lambda_i = \lambda^k$, we set

$$g_{i,k} = \frac{a_i f_{i,k} - C'_{i,k}}{a^k}.$$

□

Remark. *We have the following characterization of power series consisting only of resonant terms: g consists of only resonant terms if it commutes with the linear part of f, i.e*

$$g \circ df(O) = df(O) \circ g.$$

We can go from formal power series to actual coordinate changes using the following lemma:

Lemma 3. *Given any formal power series h there exists a C^∞ function with the same derivatives at O.*

See K-H, Lemma 6.6.3 for a proof using bump functions.

Corollary 4. *Assume that g is the formal normal form of f. Then there exists a C^∞ map h such that*

$$h \circ f = \tilde{g} \circ f,$$

where $\tilde{g}$ agrees with g up to all orders at O.

2. Smooth normal forms for hyperbolic fixed points

In the case of a hyperbolic fixed point it is possible to conjugate a map to its normal form.

Theorem 5 (Sternberg's Theorem)**.** *Assume that f has a hyperbolic fixed point at O, and the spectrum of f is non-resonant up to all order. Then f is C^∞ conjugate to its linear map.*

Theorem 6 (Chen's Theorem)**.** *Assume that two C^∞ maps f and g with hyperbolic fixed points at O are formally conjugate, then they are C^∞ conjugate.*

Both theorem follow from Corollary 4 and the following:

Proposition 7. *Assume that two maps f and g with hyperbolic fixed points at O are C^∞ tangent, then they are C^∞ locally conjugate.*

Proof. First of all, we modify both f and g such that they are globally ε-close to their linear map. Then the stable/unstable manifold theorem applies, and we change coordinates such that $W^u_f(O) = \mathbb{R}^k \times \{0\}$ and $W^s_f(O) = \{0\} \times \mathbb{R}^{m-k}$.

There are two main ideas in this proof. The first is known as *Sternberg's wedge method*. Let $\alpha = f - g$, then α vanishes at O up to all orders. We then decompose

$$\alpha(x, y) = \alpha^+(x, y) + \alpha^-(x, y), \quad (x, y) \in \mathbb{R}^k \times \mathbb{R}^{m-k},$$

where α^+ vanishes up to all orders at $y = 0$ and α^- vanishes up to all orders at $x = 0$. This can be done using a "wedge mollifier", namely a function ρ satisfying $\rho = 1$ on $H_{1/2}$ (horizontal cone) and $\rho = 0$ on $V_{1/2}$ (vertical cone). Then

$$\alpha^+ = \alpha(1 - \rho), \quad \alpha^- = \alpha\rho.$$

We will show that f is conjugate to $f + \alpha^-$, and then $f + \alpha^-$ is conjugate to $f + \alpha$.

The second idea is the so-called *homotopy method*, a widely used technique due to Jurgen Moser. In order to find a conjugacy between f and $f + \alpha^-$, we instead try to find a family of conjugacies h_t between f and $f_t := f + t\alpha^-$ for all $t \in [0, 1]$.

We define a (time-dependent) vector field $v_t : \mathbb{R}^m \to \mathbb{R}^m$ by

$$v_t(h_t z) = \frac{d}{ds}|_{s=t} h_s(z) = \lim_{s\to t} \frac{h_s(z) - h_t(z)}{s - t},$$

then $v_t(z) = \frac{d}{ds}|_{s=t} h_s(h_{-t}z)$. Using the conjugate relation

$$f = h_t^{-1} \circ f_t \circ h_t,$$

for any s, t we have $h_s^{-1} \circ f_s \circ h_s = h_t^{-1} \circ f_t \circ h_t$, hence

$$f_s \circ h_s \circ h_{-t} = h_s \circ h_{-t} \circ f_t.$$

Differentiating at $s = t$, we get

$$\alpha^- + Df_t \circ v_t = v_t \circ f_t,$$

or

$$\alpha^- \circ f_{-t} + (f_t)_* v_t = v_t,$$

where $(f_t)_* v_t = (Df_t)v_t \circ f_{-t}$. The last equation can be written formally as $(Id - (f_t)_*)v_t = \alpha^- \circ f_t$, which has a formal solution

$$v_t = \sum_{n=0}^{\infty} (f_t)^n_* \alpha^{-1} \circ f_{-t} = \sum_{n=0}^{\infty} (Df_t^n)\alpha^- \circ f_t^{-(n+1)}. \tag{1}$$

We will show that this series in fact converges in C^∞ space, yielding a real solution.

Since $\alpha^-(x, y)$ vanishes at all order at $x = 0$, for any $K > 0$, there exists a constant $C_K > 0$ such that

$$\|\alpha^-(x, y)\| \le C_K \|x\|^K \|y\|,$$

and

$$\|D^\beta \alpha^-(x, y)\| \le C_K \|x\| \|y\|^{K-k}$$

for a multi-index $|\beta| = k$ and $k < K$.

Given any $(x, y) \in \mathbb{R}^m$, let $n(x, y) = \sup\{n : f_t^{-n} \notin V_{1/2}\}$. We also assume that (by choosing ε small), any (x, y) in $H_2 = V_{1/2}^c$ satisfies $\|\Pi_x f_t^{-1}(x, y)\| \leq \mu^{-1}\|x\|$. Since $\alpha^- \circ f_t^{-k}(x, y) = 0$ for all $k \geq n(x, y)$, we have

$$\sum_{n=0}^{\infty} \|(Df_t^n)\alpha^- \circ f_t^{-(n+1)}(x, y)\| = \sum_{n=0}^{n(x,y)} \|(Df_t^n)\alpha^- \circ f_t^{-(n+1)}(x, y)\|$$
$$\leq \sum_{n=0}^{n(x,y)} C_K \|Df_t\|^n \mu^{-(n+1)K} \|x\| \|\bar{y}\|,$$

where $\bar{y} = \Pi_y f_t^{-n(x,y)}(x, y)$. For (x, y) in any compact set $\|x\|, \|\bar{y}\|$ are uniformly bounded. Choose K such that $\|Df_t\|\mu^{-K} < 1$, then (1) converges uniformly over compact set.

It remains to show that for any k, the order-k derivatives of (1) converges. We prove this statement assuming Lemma 10 and 8. We have for $n < n(x, y)$,

$$\|(Df_t^n)\alpha^- \circ f_t^{-(n+1)}(x, y)\|_{C_k} \leq C_K C_{m,l} \|(Df_t^n)\|_{C_k} \|\alpha^-\|_{C_k} \|f^{-(n+1)}\|_{C^k}^k$$
$$\leq \sum_{n=0}^{n(x,y)} C_K \mu^{-(n+1)(K-k)} D^{(n+1)(k+1)}.$$

Choose $K \gg k$ such that $\mu^{-(K-k)} D^{k+1} < 1$, then the series (1) converges absolutely in C^k norm. Since k is arbitrary, the series converges in C^∞ topology. □

Lemma 8. *Let $f, g : \mathbb{R}^m \to \mathbb{R}$ be C^l functions, then there exists a constant $C'_{m,l}$ such that*

$$\|fg\|_{C^l} \leq C'_{m,l} \|f\|_{C^l} \|g\|_{C^l}.$$

Lemma 9. *Let $f : \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}^m$ be C^l functions, then there exists a constant $C_{m,l}$ such that for any multi-index β with $|\beta| = l$, we have*

$$\|f \circ g\|_{C^l} \leq C_{m,l} \|f\|_{C^l} (1 + \|g\|_{C_l}^l).$$

Lemma 10. *Let $f : \mathbb{R}^m \to \mathbb{R}^m$ be a C^l function. Then there exists a constant $D > 1$ depending on m, l, $\|f\|_{C^l}$ such that*

$$\|f^n\|_{C^l} \leq D^n.$$

3. Finitely smooth normal forms

All of the discussions has a counter part in finitely smoothness. Let $f : U \subset \mathbb{R}^m \to \mathbb{R}^m$ be a C^l function fixing O, and $df(O) = \text{diag}\{\lambda_1, \cdots, \lambda_m\}$. Assume that

$\lambda_i \neq \lambda_j$ for $i \neq j$. Then there exists a function $f^l = (f_1^l, \cdots, f_m^l)$, such that

$$f_i^l(x) = \sum_{k \in \{0, \cdots, l\}^m} f_{i,k}^l x^k$$

is a polynomial of degree l, whose derivatives up to order l coincide with those of f. (We call f a polynomial of degree l). Then exists polynomials $g^l = (g_1^l, \cdots, g_m^l)$, $h^l = (h_1^l, \cdots, h_m^l)$ such that

$$h^l \circ f^l = g^l \circ h^l \quad \text{up to order } l,$$

and $g_{i,k}^l$ contains only resonant terms. g^l is the normal form of f up to order l. We immediately obtain:

Proposition 11. *Assume that g^l is the normal form of f up to order l. Then there exists a C^∞ map h such that*

$$f \circ h = \tilde{g}^l \circ h,$$

where $\tilde{g}^l$ agrees with g^l up to order l at O.

Indeed, it suffices to choose $h = h^l$, a polynomial map. Suppose $(\lambda_1, \cdots, \lambda_m)$ has no resonances up to order l, then the normal form $g^l = (\lambda_1 x_1, \cdots, \lambda_m x_m)$ is linear.

There is also a finitely smoothness version of Proposition 7.

Theorem 12 (Belitskii-Samovol)**.** *Given $\lambda = (\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m$, satisfying $\lambda_i \neq \lambda_j$ for $i \neq j$, and $|\lambda_i| \neq 1$ for all i. Then for any $k \in \mathbb{N}$, there exists $l = l(\lambda, k)$ such that if f and g are C^l tangent functions with $df(O) = \mathrm{diag}\{\lambda_1, \cdots, \lambda_m\}$, then f and g are locally conjugate.*

Proof. We repeat the proof of Proposition 7, we note that in order to show the series (1) converge in C^k, we only need the functions f and g to be tangent up to order K, with K sufficiently large. The theorem follows. □

The following are the counter parts to Sternberg's and Chen's theorem in finite smoothness.

Proposition 13. *Given $\lambda = (\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m$, satisfying $\lambda_i \neq \lambda_j$ for $i \neq j$, and $|\lambda_i| \neq 1$ for all i. Then for any $k \in \mathbb{N}$, there exists $l = l(\lambda, k)$ such that if $f \in C^l$ with $df(O) = \mathrm{diag}\{\lambda_1, \cdots, \lambda_m\}$, then f is locally conjugate to its normal form g^l of order l.*

In particular, if λ is non-resonant up to order l, then f is conjugate to its linear part.

4. The volume preserving and symplectic cases

If the map f preserves some structure, such as the volume form and the symplectic form, there are necessary resonances associated with those structures.

If f is volume preserving, then $\det df(O) = 0$, which means $\lambda_1 \cdots \lambda_m = 1$. Then the resonance $\lambda_i = \lambda_i(\lambda_1 \cdots \lambda_m)^l$ is always present. Assuming no other resonances are present, then the normal form g of f takes the form

$$g_i(x_1, \cdots, x_m) = \lambda_i x_i + \sum_{l=1}^{\infty} g_{i,l} x_i (x_1 \cdots x_m)^l.$$

A map $f : \mathbb{R}^{2m} \to \mathbb{R}^{2m}$ is called symplectic (under the standard symplectic form) if

$$(df(q,p))_* \omega = \omega,$$

where $\omega = dq_1 \wedge dp_1 + \cdots + dq_m \wedge dp_m$. A matrix M satisfying $M^T J M = J$ is called a symplectic matrix. f is a symplectic map if and only if $df(q,p)$ is a symplectic matrix for all (q,p).

Lemma 14. *χ is an eigenvalue of a symplectic matrix M if and only if χ^{-1} is also an eigenvalue.*

Proof. $\det(M - \lambda I) = 0$ is equivalent to

$$\begin{aligned}
&\det((M - \lambda I)^T J (M - \lambda I)) = 0 \Leftrightarrow \\
&\det(J - \lambda J M - \lambda M J + \lambda^2 J) = 0 \Leftrightarrow \\
&\det(I - \lambda M + \lambda J M J + \lambda^2 I) = 0 \\
&\det(\lambda^{-2} I - \lambda^{-1} M + \lambda^{-1} J M J + I) = 0 \Leftrightarrow \\
&\det(M - \lambda^{-1} I) = 0.
\end{aligned}$$

□

Therefore, the eigenvalues $(\lambda_1, \cdots, \lambda_{2m})$ of a symplectic matrix satisfies the relation $(\lambda_i = \lambda_{i+m}^{-1})$, after reordering. This implies the resonances $\lambda_i = \lambda_i \prod_{j=1}^{m} (\lambda_j \lambda_{j+m})^{k_j}$ are always present. Assuming no other resonances are present, then the normal form of a symplectic f takes the form

$$g_i(q_1, \cdots, q_m, p_1, \cdots, p_m) = \lambda_i x_i + \sum_{k \in \mathbb{N}_0^m} g_{i,k} x_i \prod_{j=1}^{m} (p_j q_j)^{k_j},$$

where $x_i = q_i$ if $i \leq m$, and $x_i = p_{i-m}$ if $i > m$.

5. Normal forms for equilibrium points of a vector field

In this section, we describe (without proof) the counter part of our theory for flows near a equilibrium point. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field and let φ_t be the flow defined by $\frac{d}{dt}\varphi_t(x) = F(\varphi_t(x))$. Assume that O is an equilibrium point of $F(x)$ and assume that $\mathrm{sp}(DF(x)) = \chi = \{\chi_1, \cdots, \chi_m\}$. We say that χ is resonant if for some $i = 1, \cdots, m$ and $k \in \mathbb{N}_0^m \neq e_i$ the following hold

$$\chi_i = \chi_1 k_1 + \cdots + \chi_m k_m.$$

Proposition 15. *Let $F(x)$ be a formal power series such that $DF(0) = \mathrm{diag}\{\chi_1, \cdots, \chi_m\}$. Then there exists a formal power series h such that*

$$(h)_* F(x) = G(x)$$

where $G(x)$ is a formal power series containing only resonant terms.

In particular, if χ is non-resonant up to all orders, then $G(x)$ coincide with the linear part of F.

Theorem 16 (Sternberg's theorem for flows)**.** *Let $F(x)$ be a C^∞ vector field such that $DF(0) = \mathrm{diag}\{\chi_1, \cdots, \chi_m\}$. Assume that χ is non-resonant up to all orders. Then F is locally C^∞ conjugate to its linear part.*