Smale - Birkhoff Homoclinic Theorem

Newton's Equation

The starting point is the equation

$$\ddot{\phi} + g(\phi) = 0.$$ 

If $$G(\phi) = \int_0^\phi g(s) \, ds$$ is an antiderivative of $$g(s)$$, then we know that the energy

$$E(\phi, \dot{\phi}) = \frac{1}{2} \dot{\phi}^2 + G(\phi)$$

is conserved along solutions.

Our assumptions on the nonlinearity $$g$$ are:

1. $$g \in C^\infty$$
2. $$g(0) = 0$$, $$g'(0) < 0$$.
3. $$g(\phi) = -g(-\phi)$$
4. $$g''(\phi) > 0$$, $$\phi > 0$$.
5. $$g(\phi) \to +\infty$$ as $$\phi \to +\infty.$$
It follows that there exist unique numbers 

\[ 0 < \varphi_0 < \zeta_0 \quad \text{s.t.} \quad g(\varphi_0) = 0 \quad G(\zeta_0) = 0. \]

A typical example is \( g(\varphi) = -\varphi + \varphi^2 \).

Let's look at the phase portrait for the corresponding first order system. Put \( x_1 = \varphi \), \( x_2 = \dot{\varphi} \), then \( \dot{x} = f(x) \) with

\[
f(x) = \begin{pmatrix} +x_2 \\ -g(x_1) \end{pmatrix}.
\]

Since \( g'(0) < 0 \), the critical point at \((0,0)\) is a saddle. The other critical points at \((\pm \varphi_0, 0)\) are centers.
The key feature is the existence of a homoclinic orbit \( p(t) \) on \( E = 0 \). Take the positive one. It has the properties

\[
p(0) = 0, \quad p'(0) = 0, \quad |p(t)| \leq C e^{-\sqrt{g'}|t|/|\dot{g}|}.
\]
so it decays exponentially together with its derivatives as \( t \to \pm \infty \). The stable and unstable manifolds of \( (0,0) \) coincide. Solutions are defined for all \( t \in \mathbb{R} \), for all initial conditions.

Duffing's Equation

Now let's add a little damping

\[
\ddot{\phi} + \alpha \dot{\phi} + g(\phi) = 0, \quad \alpha > 0.
\]

The energy is no longer conserved, it is dissipated since
\[
\frac{d}{dt} E (\varphi(t), \dot{\varphi}(t)) = -\alpha \varphi(t)^2 = 0.
\]

This has the effect of breaking the homoclinic and periodic orbits. The centers at \((\pm \varphi_0, 0)\) become stable foci.

The orbits cross the level sets of \(E\) in the decreasing direction. All solutions are defined for all \(t \in \mathbb{R}\).

Next consider the effect of adding a small periodic forcing term

\[
\ddot{\varphi} + \alpha \dot{\varphi} + g(\varphi) = \varepsilon f(t)
\]

\(f \in C^\infty, \quad f(t+T) = f(t).\)
A simple Gronwall's inequality argument shows that $E(\phi, \phi')$ remains finite for all $t \in \mathbb{R}$, so solutions are still globally defined.

What happens to the hyperbolic fixed point at $(0, 0)$? The system has the form

$$\dot{x} = A x + h(x) + \varepsilon F(t)$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -g''(0) & 0 \end{bmatrix}, \quad h(x) = \begin{bmatrix} 0 \\ -g(x_1) + g'(0)x_1 \end{bmatrix}$$

$$F(t) = \begin{bmatrix} 0 \\ f(t) \end{bmatrix}, \quad h(0) = 0, \quad D h(0) = 0.$$ 

Computing the spectrum of $A$, we find

$$\sigma(A) = (\lambda_1, \lambda_2, \frac{g''(0)}{\lambda_1}) \quad \lambda_1 = \frac{1}{2} \left[ a + \sqrt{a^2 - 4g'(0)} \right] > 0.$$ 

Clearly $T \sigma(A) \not\in \mathbb{Z}$, so we know...
there exists a unique small amplitude $O(\varepsilon)$ periodic solution $x_\varepsilon(t)$.

To analyze the qualitative behavior near $x_\varepsilon(t)$ we need to calculate the Floquet multipliers. The linearized equation is

$$\dot{y} = Ay + Dh(x_\varepsilon(t))y = A_\varepsilon(t)y$$

$A_\varepsilon(t+T) = A_\varepsilon(t)$. Let $\overline{y}_\varepsilon(t)$ be the fundamental matrix of $A_\varepsilon(t)$ with $\overline{y}_\varepsilon(0) = I$.

For sufficiently small $\varepsilon$, we have $\overline{y}_\varepsilon(t) \approx e^{At}$ $0 \leq t \leq T$, since $Dh(x_\varepsilon(t)) = O(\varepsilon^2)$. Thus since $\sigma(\overline{y}_\varepsilon(t)) \approx \sigma(e^{AT})$ and $\sigma(A) \subset \mathbb{R}$ we have that $\sigma(\overline{y}_\varepsilon(t)) \subset (0, \infty)$. Thus Floquet theory tells us that

$$\overline{y}_\varepsilon(t) = P_\varepsilon(t)e^{B_\varepsilon t}$$
where \( P_\varepsilon(t) + B_\varepsilon \) are real and
\[
P_\varepsilon(t + T) = P_\varepsilon(t).
\]
Thus \( \sigma(B_\varepsilon) = \sigma(\Lambda T) \)
\[
= (T\lambda_1, T\lambda_2).
\]
Since \( \lambda_1 > 0, \lambda_2 < 0 \)
the Floquet multipliers \( e^{\lambda_i T} \) lie off the unit circle. It follows that \( x_\varepsilon(t) \)
has local stable and unstable manifolds \( W_{S,\varepsilon}^{loc}, W_{U,\varepsilon}^{loc} \)

These manifolds are invariant under the flow.

On \( W_{S,\varepsilon}^{loc} \) solutions tend exponentially fast to \( x_\varepsilon(t) \) as \( t \to +\infty \). Moreover,
there is a nghdr \( U \) of \( 0 \) (incl. of \( \varepsilon \))
s.t. any solution which remains in \( U \) for
all large times actually lies on $W_{s}^{10c}(\varepsilon)$. A similar statement holds for $W_{u}^{10c}(\varepsilon)$, $t \to -\infty$.

Let $x(t, t_0, x_0, \varepsilon)$ be the solution of

$$
\dot{x} = A x + h(x) + \varepsilon F(t), \quad x(t_0) = x_0.
$$

Define the Poincaré map

$$
\Pi_{\sigma_0,\varepsilon}(x_0) = x(T, \sigma, x_0, \varepsilon).
$$

$\Pi_{\sigma_0,\varepsilon}$ is a diffeomorphism on $\mathbb{R}^2$ since the flow is globally defined. Note that $x_\varepsilon(\sigma)$ is a fixed point of $\Pi_{\sigma_0,\varepsilon}$. The sets

$$
W_{s}^{10c}(\sigma, \varepsilon) = W_{s}^{10c}(\varepsilon) \cap \{ t = \sigma \}
$$

$$
W_{u}^{10c}(\sigma, \varepsilon) = W_{u}^{10c}(\varepsilon) \cap \{ t = \sigma \}
$$

are invariant under $\Pi_{\sigma_0,\varepsilon}$. If $x \in W_{s}^{10c}(\sigma, \varepsilon)$ then $\Pi_{\sigma_0,\varepsilon}^k(x) \to x_\varepsilon(\sigma)$ as $k \to \infty$, and

if $x \in W_{u}^{10c}(\sigma, \varepsilon)$ then $\Pi_{\sigma_0,\varepsilon}^{-k}(x) \to x_\varepsilon(\sigma)$ as $k \to \infty$. 

Now define the global invariant manifolds

\[ W^g_s(\sigma, \epsilon) = \bigcup_{k=0}^{\infty} \mathcal{T}^{-k}_{\sigma, \epsilon}(W^{loc}_s(\sigma, \epsilon)) \]

\[ W^g_u(\sigma, \epsilon) = \bigcup_{k=0}^{\infty} \mathcal{T}^k_{\sigma, \epsilon}(W^{loc}_u(\sigma, \epsilon)) \]

These are smooth curves in \( \mathbb{R}^2 \) since \( \mathcal{T}_{\sigma, \epsilon} \) is a diffeomorphism. Moreover,

\[ W^g_s(\sigma, \epsilon) = \{ x : \mathcal{T}^k_{\sigma, \epsilon}(x) \in U \text{ for large } k \} \]

\[ W^g_u(\sigma, \epsilon) = \{ x : \mathcal{T}^{-k}_{\sigma, \epsilon}(x) \in U \text{ for large } k \} \]

The Meaning of Life

We want to show that \( W^g_s(\sigma, \epsilon) \) and \( W^g_u(\sigma, \epsilon) \) intersect transversely at a point \( p \neq x_\epsilon(\sigma) \). Such a point is called a transverse homoclinic point.
Since $W_u^g$ and $W_s^g$ are invariant under $\Pi_{\sigma,\varepsilon}$ we have that

$$q^k = \Pi_{\sigma,\varepsilon}^k(q) \in W_u^g(\sigma,\varepsilon) \cap W_s^g(\sigma,\varepsilon)$$

for all $k \in \mathbb{Z}$. Thus we begin to see that the dynamics can be extremely complicated.

Study the dynamics near $q$. 
Existence of a Transverse Homoclinic Point

Thm: For small $\varepsilon, \delta$ there exists $0 \leq \sigma \leq T$

s.t. $W^{g, l}_{\delta}(\sigma, \varepsilon)$ and $W^{u, l}(\sigma, \varepsilon)$ intersect transversely at a point $\gamma = (2(\delta, \varepsilon), 0) \approx (\delta, 0)$.

proof: We will first show that for all $\delta, \varepsilon$ sufficiently small there is a $0 \leq \sigma \leq T$

and $2\delta, \varepsilon \approx \delta$ s.t. the IVP

$$\ddot{\varphi} + \omega^2 \dot{\varphi} + g(\varphi) = \varepsilon f(t)$$

$$\varphi(\sigma) = 2\delta, \omega(\sigma) = 0$$

a solution with $\varphi(t) \in U$ for $|t|$ large.

(The initial condition for $\varphi$ and the initial time have been exchanged for boundary conditions at $\pm \infty$.) Thus,
\( \varphi(t) \in W^s_{\text{loc}}(\varepsilon) \) for \( t \) large and
\( \varphi(t) \in W^u_{\text{loc}}(\varepsilon) \) for \( -t \) large.

We can therefore conclude that
\[ q = (\varphi(\varepsilon), \dot{\varphi}(\varepsilon)) = (z(0, \varepsilon), 0) \in W^s_{\text{loc}}(\varepsilon, \varepsilon) \cap W^u_{\text{loc}}(\varepsilon, \varepsilon). \]

We are going to look for \( q \) as a small \( O(\varepsilon) \) perturbation of the homoclinic orbit \( p(t) \) of the unperturbed equation:
\[ \varphi(t) = p(t - \varepsilon) + z(t - \varepsilon). \] So we are going to construct \( z \) which should be uniformly \( O(\varepsilon) \) with \( z(0) = 0 \) so that \( q \) solves the equation.

A simple calculation shows that \( z(\varepsilon) \) must solve
\[ \ddot{z} + g'(p(t)) z = R(t, z, \lambda) - \alpha \dot{z} + \varepsilon f(t+\sigma) \]
\[ \dot{z}(0) = 0 \quad z = O(\varepsilon) \]

where

\[ R(t, z) = -g(p(t)+z) + g(p(t)) + g'(p(t))z \]
\[ = O(|z|^2) \]

uniformly in \( t \). \( z \) is going to be constructed via the Liapunov-Schmidt method. Consider the problem

\[ L z = N(z, \omega, \varepsilon, \sigma) \]
\[ L = \frac{d^2}{dt^2} + g'(p(t)) \alpha \dot{p}(t) \]
\[ N(z, \omega, \varepsilon, \sigma) = R(t, z, \lambda) - \alpha \dot{z} + \varepsilon f(t+\sigma) \]

on the spaces

\[ X = \{ z \in C^2(\mathbb{R}, \mathbb{R}) : \dot{z}(0) = 0 ; \sup |z| < \infty \} \]
\[ \sup |\dot{z}| < \infty, \sup |\ddot{z}| < \infty \} \]

\[ Y = \{ h \in C(\mathbb{R}, \mathbb{R}) : \sup |h| < \infty \} \]
Lemma:  
1. \( L : \mathbb{R} \to \mathbb{R} \)
2. \( \mathcal{K}(L) = \text{span} \{ \dot{\mathcal{P}}(t) \} \)
3. \( \mathcal{R}(L) = \{ h \in \mathcal{V} : \int_{-\infty}^{\infty} h(t) \dot{\mathcal{P}}(t) \, dt = 0 \} \)

proof:  1. is obvious.

Since \( \mathcal{P}(t) \) is smooth, we see from Newton's eqn. that \( L \dot{\mathcal{P}} = 0 \). Thus \( \text{span} \{ \dot{\mathcal{P}}(t) \} \subset \mathcal{K}(L) \). Let \( u_2 = \frac{1}{-g'(0)} \dot{\mathcal{P}}(t) \).

Then \( u_2(0) = 0, \dot{u}_2(0) = 1 \) and \( Lu_2 = 0 \).

Let \( u_1 \) solve \( Lu_2 = 0, u_2(0) = 1, \dot{u}_2(0) = 0 \).

Then \( \{ u_1, u_2 \} \) is a basis for \( \mathcal{K}(L) \) since \( u(t) = \begin{bmatrix} u_1(t) & u_2(t) \end{bmatrix} \) is a fundamental matrix for \( M(t) = \begin{bmatrix} 0 & 1 \\ -g'(\mathcal{P}(t)) & 0 \end{bmatrix} \).

All solutions of \( L \varphi = 0 \) are given by

\[ \varphi = c_1 u_1 + c_2 u_2 \]
We know that the Wronskian satisfies
\[ \det \mathbf{U}(t) = \mathbf{U}(0) \exp \int_0^t \mathbf{M}(s) \, ds = \mathbf{U}(0) = 1, \]
since \( \mathbf{M}(s) = 0 \). Recalling that
\[ |p(t)| \leq C e^{-at}, \quad a = \sqrt{-g''(0)} > 0 \]
it follows that \( u_1(t) \sim e^{at} \) for
\( |t| \) large. This means that \( u_2 \not\equiv 0 \)
and so \( K(L) \) in \( X \) is spanned by
\[ cp = u_2. \]

Given any continuous \( h \), it
follows by variation of parameters that the solution (in \( C^2 \)) of
\[ Lz = h, \quad z(0) = c_1, \quad z'(0) = c_2 \]
is given by
\[ z(t) = c_1 u_1(t) + c_2 u_2(t) + \int_0^t \{ u_1(s) u_2(t) - u_2(s) u_1(t) \} h(s) \, ds. \]
\[ = \left[ c_1 + \int_0^t u_2(s) h(s) \, ds \right] u_1(t) + \left[ c_2 + \int_0^t u_1(s) h(s) \, ds \right] u_2(t). \]

We ask: given \( \xi \), when does this formula determine \( z \in \mathbb{X} \)?

It is easy to show using the bounds on \( u_1 \) and \( u_2 \) that the second term and its derivatives are bounded uniformly. The first term remains bounded for large \( t \) iff

\[ c_1 + \int_0^\infty u_2(s) h(s) \, ds = 0 \]

and for large \( -t \) iff

\[ c_1 + \int_0^-\infty u_2(s) h(s) \, ds = 0 \]

which forces

\[ \int_{-\infty}^{\infty} p(s) h(s) \, ds = 0. \]
It follows from this lemma that \( \mathcal{K}(L) \)
has a closed complementary subspace \( M \) in \( \mathbb{X} \).

\[ \mathbb{X} = \mathcal{K} + M \quad \text{with} \]

\[ M = \{ z \in \mathbb{X} : \ddot{z}(0) = 0 \} \].

It also follows that \( \mathcal{R}(L) \) is closed in \( \mathbb{V} \) and

\[
\Phi \, h(t) = h(t) - \frac{\int h(s) \dot{p}(s) \, ds}{\int (\dot{p}(s))^2 \, ds} \dot{p}(t)
\]

is a projection onto \( \mathcal{R}(L) \).

**Lemme:**

1. \( N : \overline{\mathbb{X}} \times \mathbb{R}^3 \to \overline{\mathbb{V}} \) is C1
2. \( N (0, 0, 0, \sigma) = 0 \)
3. \( D_2 \, N (0, 0, 0, \sigma) = 0 \).
Following the Lyapunov-Schmidt technique, we consider the problem

\[ L \tilde{z} = PN (\lambda \dot{\rho} + \tilde{z}, \lambda, \varepsilon, \sigma) \]

on \( M \). Since \( N(0, 0, 0, \sigma) = 0 \) and \( D_{2} N(0, 0, 0, \sigma) = 0 \), it follows from the implicit function theorem that there exists a \( C^1 \) solution \( \tilde{z}(\cdot, \lambda, \alpha, \varepsilon, \sigma) \in M \) for \((\lambda, \alpha, \varepsilon)\) small and \( \sigma \in \mathbb{R} \). By uniqueness,

\[ \tilde{z}(0, 0, 0, 0, \sigma) = 0 \]

and hence

\[ \| \tilde{z}(\lambda, \alpha, \varepsilon, \sigma) \|_{\mathcal{X}} = 0 (|\lambda| + |\alpha| + |\varepsilon|) \].

We also have that

\[ \tilde{z}(\cdot, \lambda, \alpha, \varepsilon, \sigma + \tau) = \tilde{z}(\cdot, \lambda, \alpha, \varepsilon, \sigma) \].
Since we are looking for solutions of the original equation in $M$, i.e. $\dot{z}(t) = 0$, it suffices to take $\lambda = 0$.

We will now show that for $\sigma, \varepsilon$ small and suitably related, there exists a $\tau \in [0, T]$ s.t.

$$b(\sigma, \varepsilon, \sigma) = (I - P) N(\hat{z}(\cdot, 0, \sigma, \varepsilon, \sigma), 0, \varepsilon, \sigma) = 0.$$ 

Hence, $\hat{z}(\cdot, 0, \sigma, \varepsilon, \sigma)$ represents a solution of the original equation and

$$(\hat{z} + \hat{z}(0), \dot{\hat{z}}(0)) = (\hat{z} + O(\varepsilon), 0)$$

a transverse homoclinic point for $T\sigma$.

The bifurcation equation is

$$c_0 b(\sigma, \varepsilon, \sigma) = \int_{-\infty}^{\infty} N(\hat{z}(t, 0, \sigma, \varepsilon, \sigma), 0, \varepsilon, \sigma) \dot{p}(t) \, dt = 0,$$

with $c_0 = \int_{-\infty}^{\infty} |\dot{p}(t)|^2 \, dt$. 
Thus,
\[ b_0 (a, \varepsilon, \sigma) = \int_{-\infty}^{\infty} \left[ R(t, \varepsilon) - a \right] - a \dot{\varphi}(t) + \varepsilon f(t+\sigma) \right] \dot{\varphi}(t) \, dt \]

Note that \( b_0 (0, 0, \sigma) = 0 \). The last two terms are dominant. Since \( R(t, \varepsilon) = O(\varepsilon^2) \) we get
\[ b \varepsilon (0, 0, \sigma) = \frac{1}{c_0} \int_{-\infty}^{\infty} f(t+\sigma) \dot{\varphi}(t) \, dt \equiv h(\sigma) \]

(Melnikov function)

and \( h(\sigma + \tau) = h(\sigma) \). Also
\[ b (a, \varepsilon, \sigma) = -a - \varepsilon h(\sigma) + O(\sigma^2 + \varepsilon^2) \]

Fix \( \sigma_0 \in [0, T] \) and set \( k = h(\sigma_0) \neq 0 \).

Then
\[ F(\varepsilon, \sigma) = \frac{1}{\varepsilon} b(k \varepsilon, \varepsilon, \sigma) = -k + h(\sigma) + \rho(a, \sigma) \]

where \( \rho(0, \sigma) = 0 \), \( \frac{\partial}{\partial \sigma} \rho(0, \sigma) = 0 \).
Now
\[ F(0, \sigma_0) = 0 \quad F\sigma(0, \sigma_0) = h'(\sigma_0) \]
so if \( h'(\sigma_0) \neq 0 \) we get a curve \( \epsilon(\sigma) \) defined near \( \sigma_0 \) s.t.
\[ b(h(\sigma_0) \epsilon(\sigma), \epsilon(\sigma), \sigma) = 0. \]

Of course, we consider \( \sigma \) so that
\[ \alpha = h(\sigma_0) \epsilon > 0. \]

\[ h(\sigma) \]

\[ k \quad \sigma_0 \quad \sigma \]

\{ Actually, \( \alpha \) is irrelevant — reversing \( t \) changes the sign of \( \alpha \) ! \}
We must show that our homoclinic point is transverse. Fix \((q, \xi, \sigma_0)\) s.t. \(b(q, \xi, \sigma_0) = 0\).

For \(\sigma\) near \(\sigma_0\), solve

\[
L_{\pm} = N(\varepsilon, a_0, \xi_0, \sigma)
\]

\[
\varepsilon_{\pm}(0, \sigma) = O(\varepsilon) \quad \varepsilon_{\pm}(0, \sigma) = 0
\]

and \(|\varepsilon_{\pm}|, |\dot{\varepsilon}_{\pm}|, |\ddot{\varepsilon}_{\pm}| = O(\varepsilon)\)

uniformly for \(\chi \in [-T, \infty)\) \((+)\)

\[
\chi \in (-\infty, T]\)

This can be done using the implicit function theorem, since \(L: X \to Y_\pm\) is an isomorphism on one-sided bounded functions (see the lemma). In fact by variation of parameters we have
\[ z_\pm(t) = c_\pm u_1(t) + \int_0^t [u_1(s)u_2(t) - u_2(s)u_1(t)] N(z_\pm(s), \sigma) \, ds \]

with \( \dot{z}_\pm(0) = 0 \)

\[ c_\pm + \int_0^{\pm \infty} u_2(s) N(z_\pm(s), \sigma) \, ds = 0. \]

It follows that for each \( \sigma \),

\[ q_\pm(\sigma) = (z_\pm(\sigma), 0) \in W_{\sigma}^{\text{glob}}(\sigma) \]

\[ r_\pm(\sigma) = (z_\mp(\sigma), 0) \in W_{\sigma}^{\text{glob}}(\sigma) \]

When \( \sigma = \sigma_0 \) we have \( z_- = z_+ \).

To first order \( \lambda \) we see that

\[ z_\pm(0, \sigma) = -\int_0^\pm u_2(s) [-\alpha \dot{p}(s) - \varepsilon f(s+\sigma)] \, ds \]

\[ = I \cdot C \begin{bmatrix} \alpha & \varepsilon \phi(\sigma) \end{bmatrix}. \]
The separation of the stable and unstable manifolds is measured to first order by the Melnikov function

\[ \mu(\sigma) = \mathcal{E}_+ (0, \sigma) - \mathcal{E}_- (0, \sigma) \in C [ \sigma - \epsilon, h(\sigma)] . \]

Let \( \gamma_\pm (\xi) \) be a local parametrization of \( W^s_\pm (\sigma) = W^s_\pm (0, \sigma) \), \( W^u_\pm (\sigma_0) = W^u_\pm (0, \sigma_0) \),

with \( \gamma_\pm (0) = q_\pm (\sigma_0) \). The \( C^4 \) closeness of \( W^u_\pm (\sigma) \) to the homoclinic orbit \((p(\xi), \dot{p}(\xi))\) implies that for small \( \sigma, \epsilon \), \( \gamma_\pm (0) \) is not parallel to \((1, 0)\).

Denote by \( \Phi_0 (\cdot) \) the global diffeomorphism induced by the flow \( x(\sigma, \sigma_0, \cdot) \). Thus,

\[ \Phi_0 (W^u_\pm (\sigma)) = W^u_\pm (\sigma) \quad \text{and} \quad \Phi_0 \circ \gamma_\pm (\xi) \]
is a parametrization of \( W^g \ell (0) \). Since these manifolds cross the axis transversely, there is an \( C^1 \) invertible curve \( \Phi^\pm (0) \)

\[
\Phi^0 \circ \gamma^\pm \circ \Phi^\pm (0) = \Phi^\pm (0)
\]

\[
\Phi^\pm (0) = 0.
\]

Compute

\[
\frac{d}{d\sigma} \left[ \Phi^+ (\sigma) - \Phi^- (\sigma) \right] \bigg|_{\sigma = \sigma_0}
\]

\[
= \frac{d}{d\sigma} \left[ \gamma (\sigma, \sigma_0, \gamma^+ \circ \Phi^+ (\sigma_0) - \gamma (\sigma, \sigma_0, \gamma^- \circ \Phi^- (\sigma_0)) \right] \bigg|_{\sigma = \sigma_0}
\]

\[
= \gamma (\sigma_0, \sigma_0, \gamma^+ \circ \Phi^+ (\sigma_0)) - \gamma (\sigma_0, \sigma_0, \gamma^- \circ \Phi^- (\sigma_0))
\]

\[
+ \gamma^+ \circ \Phi^+ (\sigma_0) \Phi^+ (\sigma_0) - \gamma^- \circ \Phi^- (\sigma_0) \Phi^- (\sigma_0)
\]

\[
= \gamma^+ (0) \Phi^+ (0) - \gamma^- (0) \Phi^- (0)
\]

since \( D_{\Phi} x (\sigma_0, \sigma_0, \rho) = I \), and \( \Phi^\pm (0) = 0 \), \( \gamma^+ (0) = \gamma^- (0) \).
Let \( \mathbf{\gamma}^+(0) = (v_1, v_2) \) \( \mathbf{\gamma}^-(0) = (w_1, w_2) \)

\[ A = \mathbf{\gamma}^+(0) \quad B = \mathbf{\gamma}^-(0). \]

Note that \( v_2, w_2 \neq 0 \). Our equation above shows that

\[ \mu'(\phi_0) = A v_1 - B w_1 \]

\[ 0 = A v_2 - B w_2 \]

from which we get

\[ \mu'(\phi_0) = \frac{A}{w_2} \left[ v_1 w_2 - v_2 w_1 \right]. \]

To first order \( \mu'(\phi_0) = c h'(\phi_0) \).

So if \( h'(\phi_0) \neq 0 \) (as was already assumed) and \( a, c \) are small then \( \mathbf{\gamma}^+(0) \)

and \( \mathbf{\gamma}^-(0) \) are transverse.
The Smale Horseshoe

Let \( f: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \) be a diffeomorphism.

Let \( S = [0,1] \times [0,1] \subset \mathbb{R}^2 \) be the unit square.

Suppose that \( f | S \) is the composition of a contraction in \( x \) by an amount \( \lambda < \frac{1}{2} \), followed by expansion in \( y \) by an amount \( \mu > 2 \), followed by a folding and a translation.

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
1 \quad 2 \\
3 \quad 4
\end{array}
\end{array}
\end{array}
\xrightarrow{f}
\begin{array}{c}
\begin{array}{c}
\begin{array}{c}
1 \quad 2 \\
3 \quad 4 \\
\mu
\end{array}
\end{array}
\end{array}
\]
Note that $f(s)$ intersects $S$ in two disjoint vertical strips of width $\lambda$: $V_0$, $V_1$.

To find the preimage of these vertical strips we must contract in $y$ by $\mu^{-1}$ and expand in $x$ by $\lambda^{-1}$.
Thus $f^{-1}(s)$ consists of two disjoint horizontal strips $H_0, H_1$ of thickness $\mu^{-1}$.

If $f: H_i \xrightarrow{i-1} V_i$ and on $H_i$

$$Df\big|_{H_i} = (-1)^i \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix}.$$ 

In the same way, $f^{-1}: V_i \xrightarrow{i-1} H_i$ with

$$Df^{-1}\big|_{V_i} = (-1)^i \begin{bmatrix} 1^{-1} & 0 \\ 0 & \mu^{-1} \end{bmatrix}.$$ 

Now we want to describe the set

$$\Lambda = \{ x \in S : f^k(x) \in S , \forall k \in \mathbb{Z} \}$$

$$= \bigcap_{k=-\infty}^{\infty} f^k(S)$$

If nonempty, $\Lambda$ is invariant under $f$.

$$S \cap f(s) \cap f^2(s) = f(f^{-1}(s) \cap S \cap f(s))$$

\[\begin{array}{c}
\begin{array}{c}
H_0 \\
H_1
\end{array}
\end{array} \xrightarrow{f} \begin{array}{c}
\begin{array}{c}
V_0^2 \\
V_1^2
\end{array}
\end{array} \xrightarrow{f} \begin{array}{c}
\begin{array}{c}
V_0^2 V_1^2 \\
V_2^2 V_3^2
\end{array}
\end{array} \xrightarrow{f} \begin{array}{c}
\begin{array}{c}
S \cap f(s) \cap f^2(s)
\end{array}
\end{array} \xrightarrow{f} \begin{array}{c}
\begin{array}{c}
S \cap f(s) \cap f^2(s)
\end{array}
\end{array} \]
For \( k > 0 \), \( \sigma \cap f_0(s) \cap \ldots \cap f_k(s) \) consists of \( 2^k \) disjoint vertical strips \( V_{j}^{k} \), \( j = 0, 1, \ldots, 2^{k-1} \) of thickness \( \lambda^k \), labelled in the indicated manner. These strips are nested in the sense that

\[
V_{j}^{k-1} \subseteq V_{2j}^{k}, \quad V_{2j+1}^{k}, \quad j = 0, 1, \ldots, 2^{k-1} - 1
\]

or

\[
V_{j}^{k} \subseteq V_{\left\lfloor \frac{j}{2} \right\rfloor}^{k-1}, \quad j = 0, 1, \ldots, 2^{k-1} - 1
\]

Moreover, we have

\[
f : H_0^1 \cap V_{j}^{k-1} \xrightarrow{1-1} V_{j}^{k}, \quad j = 0, 1, \ldots, 2^{k-1} - 1
\]

\[
f : H_1^1 \cap V_{j}^{k-1} \xrightarrow{1-1} V_{2^{k-1} + j}^{k}, \quad j = 0, 1, \ldots, 2^{k-1} - 1
\]
Similarly, for $k > 0$, $f^{-k}(s) \cap \ldots \cap f(s) \cap s$ consists of $2^k$ horizontal strips $H_j^k$ for $j = 0, 1, \ldots, 2^k - 1$ of thickness $\lambda^{-k}$. We have the nesting

$$H_j^{k-1} \supset H_j^k, \quad H_j^{k-1} \supset H_{2j+1}^k, \quad j = 0, 1, \ldots, 2^{k-1} - 1$$

$$H_j^k \subseteq H_{\lceil j/2 \rceil}^{k-1}, \quad j = 0, 1, \ldots, 2^k - 1.$$

We also have the correspondences

$$f^!: V_0 \cap H_j^{k-1} \xrightarrow{1-1} H_j^k, \quad j = 0, 1, \ldots, 2^{k-1} - 1$$

$$f^{-1}: V_1 \cap H_j^{k-1} \xrightarrow{1-1} H_{2^{k-1}+j}^k, \quad j = 0, 1, \ldots, 2^{k-1} - 1.$$

Or

$$f: H_j^k \rightarrow \begin{cases} H_j^{k-1}, & j = 0, 1, \ldots, 2^{k-1} - 1 \\ H_j^{k-1} \cup H_{j+2^k-1}^{k-1}, & j = 2^{k-1}, \ldots, 2^k - 1 \end{cases}$$
Thus

\[ \Lambda = \bigcap_{k=-\infty}^{\infty} f^k(s) \]

\[ = \bigcap_{k=1}^{\infty} \left[ s \cap \ldots \cap f^{-k}(s) \right] \cap \left[ s \cap \ldots \cap f^k(s) \right] \]

\[ = \bigcap_{k=1}^{\infty} \left[ \bigcup_{i=1}^{2^{k-1}} H_i^k \right] \cap \left[ \bigcup_{j=1}^{2^{k-1}} V_j^k \right] \]

We see that \( \Lambda \) is a nonempty Cantor set invariant under \( f \).

\[
\left( \bigcup_{i=0}^{3} H_i^2 \right) \cap \left( \bigcup_{j=0}^{3} V_j^2 \right)
\]
Given \( x \in A \), for each \( k = 1, 2, \ldots \), there is a unique \( i(k) \) s.t. \( x \in H^k_{i(k)} \) and a unique \( j(k) \) s.t. \( x \in V^k_{j(k)} \).

In order that the intersection of these strips be non-empty, they must be nested. Therefore

\[
H^{k-1}_{i(k-1)} \subsetneq H^k_{i(k)} \quad \text{and} \quad V^{k-1}_{j(k-1)} \subsetneq V^k_{j(k)}.
\]

This means that

\[
2i(k-1) = i(k) \quad \text{or} \quad 2i(k-1) + 1 = i(k)
\]

\[
2j(k-1) = j(k) \quad \text{or} \quad 2j(k-1) + 1 = j(k).
\]

Since \( \text{diam } H^k_{i(k)} = \lambda^{-k} \) and

\( \text{diam } V^k_{j(k)} = \mu^k \)

we see that

\( \bigcap_k H^k_{i(k)} \) is a horizontal line and

\( \bigcap_k V^k_{j(k)} \) is a vertical line. Hence,
there is a 1-1 correspondence between points in $\Lambda$ and nested sequences of horizontal and vertical strips.

Given $x \in \Lambda$ and corresponding strips $H_{i(k)}^k$, $V_{j(k)}^k$ we can associate a sequence $a = (\ldots, a_{-2}, a_{-1}, a_1, a_2, \ldots)$ by the formula $a = h(x)$

$$a_k = \begin{cases} 0 & j(k) \text{ is even} \\ 1 & j(k) \text{ is odd} \end{cases}$$

$$a_{-k} = \begin{cases} 0 & i(k) \text{ is even} \\ 1 & i(k) \text{ is odd} \end{cases}$$

Let $\Sigma$ be the set of all such sequences

$$\Sigma = \{ a : \{ \pm 1, \pm 2, \ldots \} \rightarrow \{0,1\} \}.$$
$\Sigma$ is a metric space with

$$d(a, b) = \sum_{k=1}^{\infty} 2^{-k} \left[ |a_k - b_k| + |a_{-k} - b_{-k}| \right].$$

**Thm**: $h: \Lambda \to \Sigma$ is a homeomorphism and $h \circ f|\Lambda = \tau \circ h$, where $\tau$ is the right shift, i.e., if $b = \tau(a)$ then

$$b_i = \begin{cases} a_{i+1}, & i \neq 1, \\ a_1, & i = 1. \end{cases}$$

This is usually summarized by saying "$f|\Lambda$ is topologically conjugate to the shift on two symbols."

**proof**: There is a 1-1 correspondance between nested families of strips and elements of $\Sigma$. For example
given \( a \in \Sigma \), put

\[
\begin{align*}
  j(1) &= a_1, \\
  j(2) &= 2j(1) + a_2 = 2a_1 + a_2 \\
  &\vdots \\
  j(k) &= \sum_{l=1}^{2^{k-l}} \alpha_l = 2j(k-1) + \alpha_k
\end{align*}
\]

end

\[
i(k) = \sum_{l=1}^{k} 2^{k-l} \alpha - l.
\]

This gives nested families of horizontal and vertical strips, which determine a unique \( x \in \Lambda \). Thus, \( h \) is onto.

If \( h(x) = h(y) \) then \( x \) and \( y \) determine the same family of nested strips i.e. \( x = y \). So \( h \) is 1-1.

To see that \( h \) is continuous, let \( \varepsilon > 0 \) be given. Choose \( k_0 \) s.t. \( 2^{-k_0+1} < \varepsilon \).
Choose $\delta > 0$ so that

$$B_\delta(x_0) = \{ x : ||x - x_0|| < \delta \}$$

is contained in

$$\bigcap_{k=0}^{K_0} \bigcup_{j(k)} V_{j(k)}^k \cap H_{i(k)}^k$$

where $V_{j(k)}^k$, $H_{i(k)}^k$ are the strips. This is possible since $x = \lim_{j(k)} V_{j(k)}^k \cap H_{i(k)}^k$ containing $x_0$. Note that for all $x \in B_\delta(x_0)$ we have

$$h(x)_k = h(x_0)_k \quad k = \pm 1, \ldots, \pm K_0$$

Thus,

$$d(h(x), h(x_0)) \leq \sum_{k=K_0+1}^{\infty} 2^{-k} [2^2] = 2^{-K_0+1} < \varepsilon$$

This proves continuity.

Finally, we must show the shift property. Let $x \in \Lambda$ with
$x \in V^k_{j(k)}, H^k_{i(k)}$. By the mapping properties of $f$, we have

$$f(x) \in V^{k+1}_{j(k)} \quad \text{or} \quad V^{k+1}_{j(k)+2k} \quad k = 1, 2, \ldots$$

and

$$f(x) \in H^{k-1}_{i(k)} \quad \text{or} \quad H^{k-1}_{i(k)-2^{k-1}} \quad k = 2, 3, \ldots$$

and

$$f(x) \in V^1_{i(1)}$$

So if $a = h(x)$ and $b = h \circ f(x)$, we have $b = \tau(a)$, i.e., $h \circ f = \tau \circ h$. 

Theorem: $f$ leaves a invariant and

(a) $A$ contains a countable set of periodic orbits of arbitrarily long periods.

(b) $A$ contains an uncountable set of (bounded) non-periodic orbits.

(c) $A$ contains a dense orbit.

Example: $a = (-a_2, a_1, a_0, a_2, \ldots)$

$= (\ldots, 1, 0, 1, 1, \ldots)$

Periodic of period 4

$a = h(x)$

$x \in H_0^1, H_1^2, V_1^1, V_2^2$

$\tau(a) = h(f(x)) = (\ldots, 1, 1, 0, 1, \ldots)$

$f(x) \in H_1^1, H_3^2, V_0^1, V_1^2$
\[ \tau_2(\alpha) = h(f^2(\alpha)) = (\ldots, i, i, i, 0, \ldots) \]

\[ f^2(x) \in H_1, H_3, V_1, V_2 \]

\[ \tau_3(\alpha) = h(f^3(\alpha)) = (\ldots, 0, i, i, i, 0, \ldots) \]

\[ f^3(x) \in H_1, H_2, V_1, V_3 \]

\[ H_2^2 \quad \Box \quad \Box \quad \Box \quad \Box \]

\[ H_3^2 \quad \Box \quad \Box \quad f^2(x) \rightarrow \Box \quad \Box \]

\[ H_1^2 \quad \Box \quad \Box \quad x \rightarrow \Box \quad \Box \]

\[ H_0^2 \quad \Box \quad \Box \quad \Box \quad \Box \]

\[ V_2^2 \quad V_2^2 \quad V_3^2 \quad V_4^2 \]

Q: Where exactly is \( x \) in \( H_1^2 \cap V_3^2 \)?

A: Imbed the picture self-similarly in \( H_1^2 \cap V_3^2 \), etc.
The Smale horseshoe is really just a model. In practice, a result of the following sort is necessary. First, a definition.

**Defn:** Let \( S = [0, 1] \times [0, 1] \) be the unit square. Let \( \varphi_1(x), \varphi_2(x) \) be continuous functions with 

\[
| \varphi_i(x_j) - \varphi_i(x_k) | \leq \mu |x_j - x_k| \quad \alpha \mu < 1,
\]

s.t. \( 0 < \varphi_1(x) < \varphi_2(x) < 1 \). A horizontal strip is a set of the form

\[
H = \{ (x, y) \in S : 0 \leq x \leq 1, \varphi_1(x) \leq y \leq \varphi_2(x) \}
\]

Vertical strips are defined in the obvious manner.
Let \( f: \mathbb{R}^2 \to \mathbb{R}^2 \) be a diffeomorphism.

Suppose that

\[ H_1: \text{S contains a pair of disjoint horizontal strips } H_0, H_1 \text{ and a pair of disjoint vertical strips } V_0, V_1. \text{ s.t.} \]

\[ f: H_i \xrightarrow{1:1} V_i \quad i = 0, 1. \]

Moreover, it is assumed that the vertical boundaries of \( H_i \) correspond to the vertical boundaries of \( V_i \), and likewise for the horizontal boundaries.

\[ H_2: \text{The set } B_+ = \{(f, \eta) \in \mathbb{R}^2: |f| < \sqrt{1 + |\eta|^2}\} \]

\( 0 < \eta < 1 \) is mapped into itself by \( Df(x) \), for all \( x \in H_i \). Moreover, if \((f_0, \eta_0) \in B_+\) and \((f_1, \eta_1)\) is its image point under \( Df(x) \),
then \(|\eta, 1| \leq v^{-1} |\eta_0, 1|\). Similarly, the set \(B_- = \{(\xi, \eta) \in \mathbb{R} : |\eta| < v \xi\}\) is mapped into itself by \(D_f^{-1}(x)\) for all \(x \in V_i\).

If \((\xi_0, \eta_0) \in S_-\) and \((\xi_1, \eta_1)\) is its image under \([D_f(x)]^{-1}\) then \(|\xi_0| > v^{-1} |\xi_1|\).

**Thm:** If \(H_1, H_2\) hold for \(f\) with \(0 < v < \frac{1}{2}\), then \(f\) has an invariant Cantor set \(\Lambda\) and \(f|\Lambda\) is topologically equivalent to a shift on two symbols.

**Proof:** \(H_2\) is necessary to get nested strips to shrink to a line. See Moser, Jürgen. "Stable and random motions in dynamical systems."

Princeton. \(\text{(Thm 3.2.)}\).

**Remark:** The Smale horseshoe fulfills \(H_1, H_2\) with \(v > 2, \frac{1}{2}\).
The Birkhoff-Smale Homoclinic Thm.

Thm: Let $f: \mathbb{R}^2 \to \mathbb{R}^2$ be a diffeomorphism with a hyperbolic fixed point. Suppose that $W^s(p)$ and $W^u(p)$ intersect transversely at a point $q \neq p$. Then $f$ has an invariant set $A$ on which $f$ is topologically conjugate to a shift on two symbols.

Sketch of proof: Show that for $k$ sufficiently large $f$ looks like a horseshoe map near $p$. Starting with a small neighborhood $U$ of $p$, iterate $f$ and $f^{-1}$. 
Let $V = f^k(U) \cap f^{-l}(U)$, let $\tilde{U} \circ f^{-k}(V)$

Then $f^{k+l}(U) \cap U$ looks like:
Moreover, \( f^{-(k+\ell)}(U) \cap U \) looks like

\[
\begin{array}{c|c|c|c|c}
1 & 1 & 1 & 1 & \\
\hline
1 & 1 & 1 & 1 & \\
\hline
\end{array}
\]

Thus \( f^{k+\ell} \) maps a pair of disjoint horizontal strips homeomorphically onto a pair of disjoint vertical strips with the proper boundary correspondence.

\( Df \) satisfies the estimates of \( H_2 \) near \( p \), since w.l.o.g. we may assume by hyperbolicity that \( \sigma(Df(p)) = \{ \lambda, \mu \} \)

with \( 0 < \lambda < 1 \), \( \mu > 1 \), and \( e \)-vectors \( e_1, e_2 \).

So \( Df^{k+\ell}(p) = \begin{bmatrix} \lambda^{k+\ell} & 0 \\ 0 & \mu^{k+\ell} \end{bmatrix} \), and in a small neighborhood of \( p \), \( H_2 \) will hold.