THE METHOD OF AVERAGING IS EXTENDED TO A GENERAL COORDINATE-FREE (POSSIBLY INFINITE-DIMENSIONAL) SETTING. WE CONSIDER PERTURBATIONS OF DYNAMICAL SYSTEMS WITH ONLY CLOSED ORBITS, AND STUDY THE PERSISTENCE OF SOME CIRCULAR SYMMETRY. SIMPLE TEST EXAMPLES SHOW THAT FOR FINITE PERTURBATION AN INFINITE NUMBER OF DISTINCT CIRCULAR SYMMETRIES MAY EXIST. THE THEORY IS THEN EXTENDED TO THE HAMILTONIAN SETTING, WHERE A CIRCULAR SYMMETRY LEADS TO A CONSERVED QUANTITY. WE EXPLORE THE CONDITIONS UNDER WHICH AN EXACT INVARIANT, GENERALIZING THE ADIABATIC INVARIANT, EXISTS FOR FINITE PERTURBATION. THE PROCESS OF REDUCTION MAY BE APPLIED TO OBTAIN A NEW HAMILTONIAN SYSTEM FOR THE DYNAMICS OF THE SLOW VARIABLE. THIS PROCEDURE IS APPLIED IN A UNIFORM MANNER TO OBTAIN THE GUIDING CENTER AND OSCILLATION CENTER DESCRIPTIONS OF PARTICLE MOTION IN MAGNETIC FIELDS AND OSCILLATING WAVES. THIS IS THEN RELATED IN THE INFINITE DIMENSIONAL CONTEXT TO THE EIKONAL DESCRIPTION OF WAVES.

*SUPPORTED BY U.S. DOE CONTRACTS # DE-AC03-76SF00098 AND W-7405-ENG-36.

Question: What do the guiding center description of a charged particle in a slowly varying magnetic field, the oscillation center description of a charged particle in a high-frequency wave with slowly varying amplitude, the eikonal description of high-frequency waves moving in a slowly varying medium, and various other physical examples that use the method of averaging have in common?

Answer: In each case a natural separation of space and time into scales of fast and slow variation occurs because of the presence of an approximate symmetry. This allows a simplified description of the dynamics to be given.

Aim: To develop a unifying geometric framework for these examples, making explicit connections with modern geometric dynamics.

Advantages of this Approach:

I. Coordinate-Free

- → easy to use in any geometry
- → results are explicitly independent of coordinates (other approaches are often not!)

II. Works in an entirely Hamiltonian Framework (for Hamiltonian systems)

- → automatic conservation laws
- → connections with Hamiltonian field theories

III. Works for infinite-dimensional systems

→ applicable to wave systems

IV. Provides connection between Noether's Theorem and averaging techniques

V. Gives insight into breakdown of approximation

→ geometric structure of resonances and bifurcations

→ topological obstructions to simplification

→ long-term behavior - eg. chaotic dynamics

VI. Gives insight into non-uniqueness of resulting system

→ explicit geometric examples for finite perturbation

→ connection with intrinsic non-uniqueness inherent at any level of perturbation theory that arises in Takens' theory of normal forms for vector fields

VII. Suggests general principles for finding a separation of scales in general systems with approximate symmetry

VIII. Unifies the framework of many physical processes

IX. Connects drifts and pseudo-forces that arise from averaging with "effective potentials" and "magnetic" terms in the Poisson bracket that result from reduction of systems with symmetry

X. Sheds light on the "geometric diffraction" limit

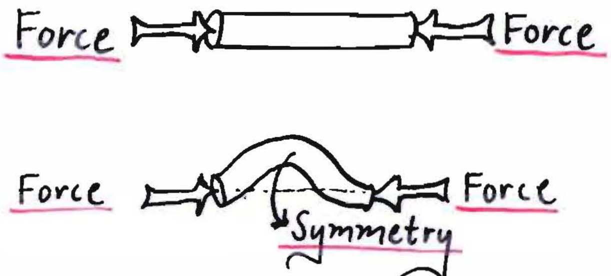
→ and its analogs in non-wave systems

→ connections with catastrophe theory + renormalization group

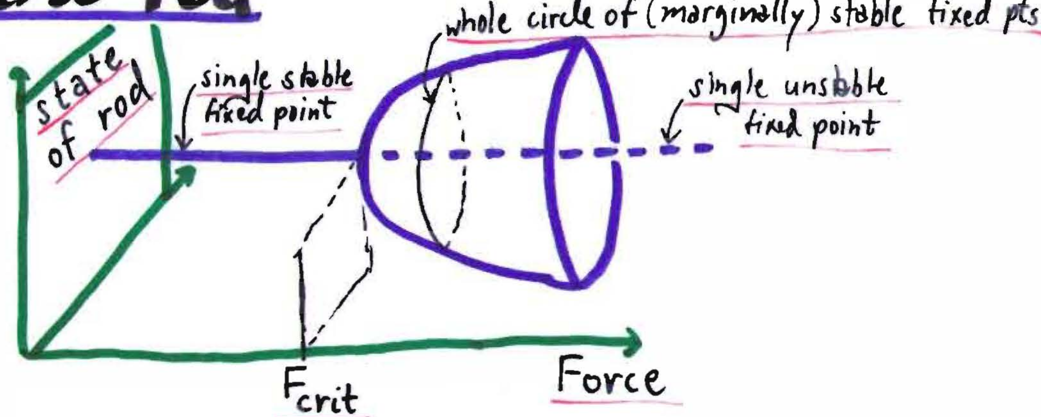
Heuristics of Motivating Physical Examples

I. Euler buckling and related plasma instabilities

a. Cylindrically symmetric rod



rod buckles when $F > F_{crit}$
but in any direction

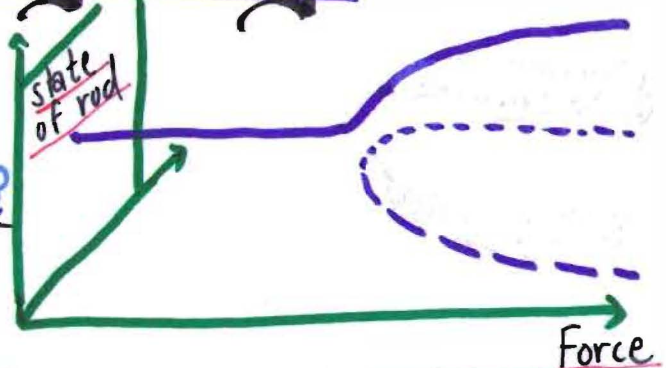


bifurcation diagram

b. Break the cylindrical symmetry

now ∃ preferred direction for buckling
→ diagram appears to change radically

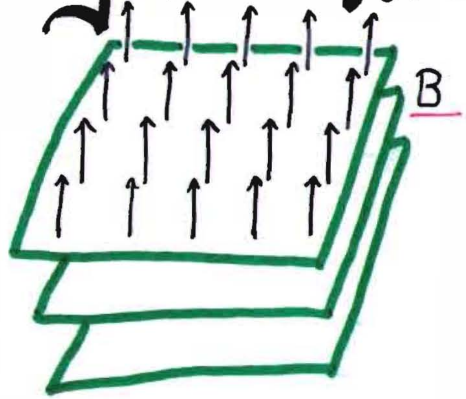
What really happens for small perturbation?
→ Still buckles in any direction, but
slowly drifts to preferred direction.



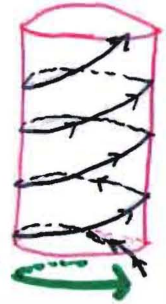
$\boxed{\text{~ Breaking symmetry has introduced a new slow time scale ~}}$

II. Gyromotion of a particle in a magnetic field

a. Homogeneous magnetic field

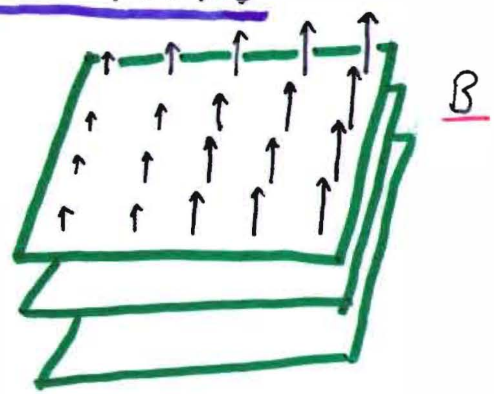


particle orbits are helices:



rotation of helix "cylinder" in phase space is circle symmetry

b. Slightly inhomogeneous magnetic field

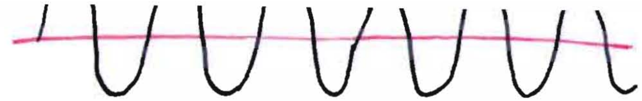


breaks the circle symmetry slightly

→ particles still approximately go in helices but slowly drift from helix to helix

III. Oscillation center motion of a charged particle in a high-frequency wave

a. Constant Amplitude wave



particle oscillates back and forth
→ circle symmetry in phase space

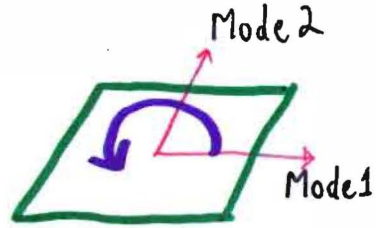
b. Modulated Amplitude

particle still oscillates, but center weakly accelerates as if in ponderomotive potential

IV. Linear Mode-Coupling

a. Degenerate energy state in wave system

relative mix of eigenstates is constant in time

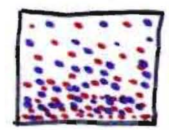


b. Slight splitting of eigenvalues

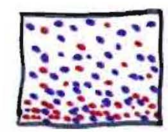
→ energy in modes slowly sloshes back & forth

V. Two-species gas (red and blue atoms) in gravitational field

a. Symmetry under red ↔ blue → homogeneous mixing



b. red atoms slightly heavier → atmosphere law leads to slow relative density gradient

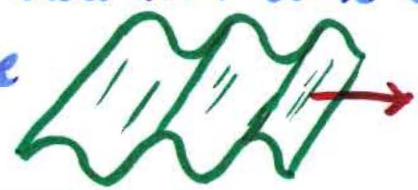


VI. Eikonal waves

a. Homogeneous medium

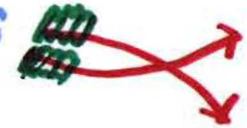
→ waves of frequency ω come back to initial state after time $\frac{2\pi}{\omega}$

→ circle symmetry in ∞-dim state space



b. Slowly varying medium

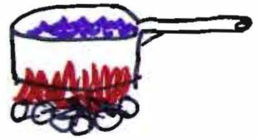
→ wave packets move according to ray equations



→ slow drift from plane-wave to plane-wave

VII. Dynamics of Dislocations in fluid instabilities

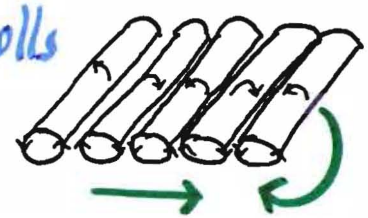
eg. Rayleigh-Benard



a. Exact homogeneity

→ superposition of perfect rolls

→ two-torus symmetry in state space

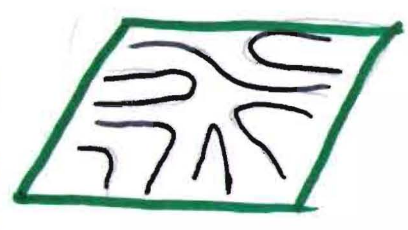


b. In reality

$\exists$ dislocations in roll structure classified by 1st homotopy group of two-torus

→ these drift slowly according to their own dynamics

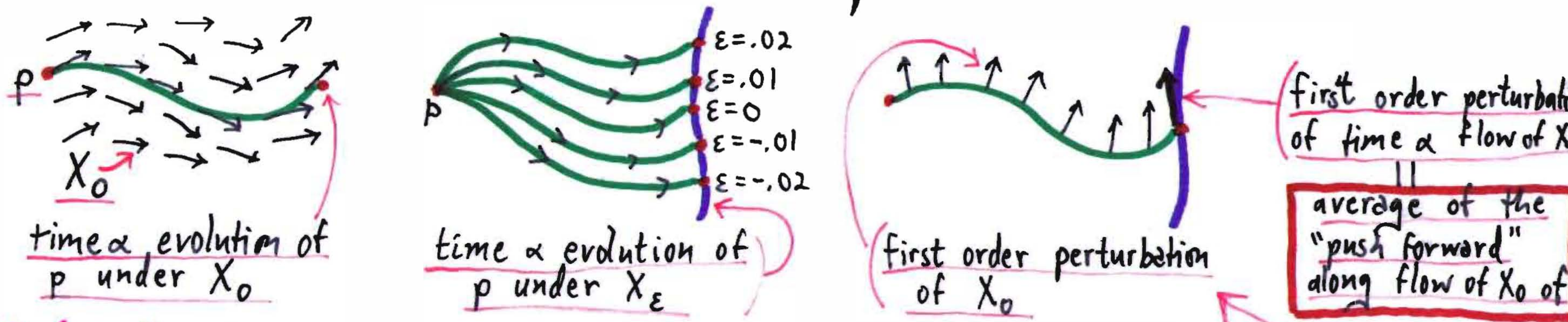
→ locally try to be plane rolls ⇒ "eikonal" description



Mathematical Setup (For exact statements and proofs see Mathematical Addenda)

I. The Perturbation Lemma

The first order perturbation of the time-α evolution of a dynamical system (described by a vector field X_ε which is a perturbation of X_0) is equal to the average along the unperturbed evolution of the first-order perturbation in X_ε.



Warning: One must push-forward the vectors along the unperturbed flow to the final point and then average. It is not correct to write the perturbation in some coordinate system and average component-wise.

II. Averaging Around Closed Orbits

Assume all orbits of X_0 are closed and of period 1 (this is no restriction in the Hamiltonian case: see V). Given any vector field Y, we may define its average around the loops: $\overline{Y}$ ($= \int_0^1 \mathcal{F}_\alpha X_{0*} \cdot Y \, d\alpha$ where $\mathcal{F}_\alpha X_0$ is the flow along X_0 for a time α).

The loop average has the following nice properties:

a) $\overline{X_0} = X_0$ the unperturbed vector field is its own average

b) $\mathcal{F}_\alpha X_{0*} \cdot \overline{Y} = \overline{Y}$ the loop average of a vector field is symmetric under the unperturbed flow

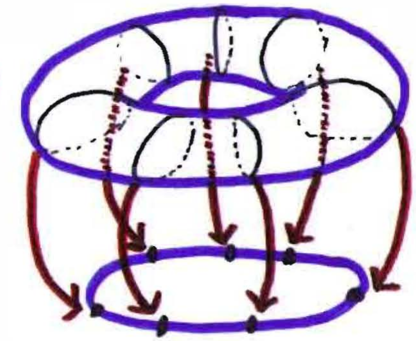
c) $[X_0, \overline{Y}] = 0$ the infinitesimal version of b) (connects w/ normal forms)

We will be interested in the case $Y = \left.\frac{\partial X_\varepsilon}{\partial \varepsilon}\right|_{\varepsilon=0}$ (i.e. Y is the first order perturbation to X_0).

III. Projecting to the Orbit Space

Because the orbits of X_0 are all topological circles with the same period, the space whose points are these circles is a smooth manifold. (Kruskal calls this the ring-space.) Dynamics on this manifold describes the drift from loop to loop.

eg.



Property II b says that $\overline{Y}$ gives rise to a well-defined vector field on the orbit space, which (from I) represents the first order drifts from orbit to orbit.

(Moser has shown that non-degenerate fixed points of the first-order dynamics on the orbit space represent loops from which periodic orbits of the exact non-linear X_ϵ branch off. Morse theory on the orbit space can guarantee a certain number of periodic orbits.)

IV. Hamiltonian Mechanics

In modern parlance, a Hamiltonian system is defined on a **Poisson Manifold**. This is a manifold with a Poisson bracket $\{\ ,\ \}$ which takes pairs of functions and gives back a function satisfying:

a) Bilinearity: $\{af_1+bf_2, g\} = a\{f_1,g\} + b\{f_2,g\}$

b) Anti-symmetry: $\{f,g\} = -\{g,f\}$

c) Jacobi identity: $\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$

d) Leibnitz rule: $\{fg,h\} = f\{g,h\} + g\{f,h\}$

Every function H then determines a dynamics via the vector field $X_H \equiv \{\cdot, H\}$ (in coordinates x^μ we have: $X_H^\mu = \{x^\mu, H\}$).

Thus the time derivative of an observable f is given by the usual: $\dot{f} = \{f, H\}$.

V. Hamiltonian Orbit Space

In the case where X_ε comes from a hamiltonian H_ε (so $X_\varepsilon = X_{H_\varepsilon} = \{\cdot, H_\varepsilon\}$), then in fact the first order perturbation Y and its average $\overline{Y}$ are hamiltonian.

The Hamiltonian for the averaged perturbation is the average around the loops of the Hamiltonian for the perturbation.

$$\boxed{\text{i.e. } \int_0^1 \frac{\partial H_\varepsilon}{\partial \varepsilon}\Big|_{\varepsilon=0} \circ \varphi_\alpha^{X_0} \, d\alpha}$$

Note that if all orbits are periodic in a Hamiltonian system, then the period is guaranteed to be constant on each energy surface.

VI. Noether's Theorem

Symmetries of a Hamiltonian system are generated by conserved quantities.

Proof: Let J generate a symmetry: i.e. $X_J \cdot H = 0$, but then
$\dot{J} = \{J, H\} = -\{H, J\} = -X_J \cdot H = 0$ so J is conserved. ■

VII. The Process of Reduction

Given a Hamiltonian system with a group of symmetries G acting on it, one may restrict attention to a lower dimensional Hamiltonian system which contains all the essential dynamics. In case G is abelian, one may reduce the dimension by 2 for every dimension of G.

eg. i) Translation symmetry: $G = \mathbb{R}^3$ The generator is the total momentum and is conserved. We choose a total momentum, removing 3 dimensions. We consider dynamics on the $\mathbb{R}^3$ action orbit space (i.e. move with the center of mass) to remove 3 more dimensions for a total of 6 ($\mathbb{R}^3$ is abelian).

eg. ii) Rotation symmetry: $G = SO(3)$ The generator is the total angular momentum. Fixing this removes 3 dimensions. We can go to the orbit space of the subgroup of rotations about the z-axis removing 1 more dimension (Jacobi's elimination of the node). But now x and y rotations are not symmetries (they change the z-axis) so we can only reduce by a total of 4 dimensions. This is because $SO(3)$ is non-abelian.

VIII. Effective Potentials and Magnetic Terms

in the Poisson bracket are often the result of reduction. In the ii^nd example, going to a rotating frame of reference introduces a centrifugal force (modification of the Hamiltonian) and a Coriolis force (modification of the Poisson bracket).

Dynamics of a charged particle in a magnetic field is most naturally represented (i.e. no unphysical vector potential) by a non-canonical Poisson bracket:

$$\boxed{\{f,g\} = \frac{\partial f}{\partial \underline{x}} \cdot \frac{\partial g}{\partial \underline{v}} - \frac{\partial f}{\partial \underline{v}} \cdot \frac{\partial g}{\partial \underline{x}} + \underline{B} \cdot \left(\frac{\partial f}{\partial \underline{v}} \times \frac{\partial g}{\partial \underline{v}}\right)}$$

with the Hamiltonian being just the kinetic energy: $H = \frac{1}{2}mv^2$.
Since the reduced Poisson brackets often look like this, the modifications are called magnetic terms.

A classical theorem of Larmor says that to first order in $\underline{B}$ the Coriollis force acts like a magnetic field.

A school of nineteenth century thought (eg. Felix Klein) felt that all potential energies arose in this fashion from the kinetic energy of hidden degrees of freedom. This idea gave great impetus to the kinetic theory description of thermodynamics.

IX. Averaging and Reduction

We may now look at the case discussed in V as a Hamiltonian system $X_0 + \overline{Y}$ with a circle symmetry generated by H_0. Going to the orbit space is then exactly the process of reduction. We also see that H_0 is a conserved quantity to this order in perturbation theory. For the example of gyromotion, H_0 may be taken to be a function of the magnetic moment which is then an adiabatic invariant.

In this setting, we see that the drifts and pseudoforces of averaging are exactly the effective potentials and magnetic Poisson brackets of reduction.

X. Approximate Noether's Theorem

Approximate symmetries lead to approximately conserved quantities.

Proof: Assume J generates a transformation under which the dynamics is approximately symmetric: i.e. $X_J \cdot H = \varepsilon$ is small. Then $\dot{J} = X_H \cdot J = \{J, H\} = -\{H, J\} = -X_J \cdot H = -\varepsilon$ is small ∎

This explains why a slight breaking of symmetries should lead to dynamics on a slow time scale.

XI. Lie Transforms from this Perspective

We are given a Poisson manifold with dynamics specified by a Hamiltonian H_ε. For H_0 we have a circle symmetry generated by J_0, so $\{H_0, J_0\} = 0$ and J_0 is conserved.

We would like to find a circle symmetry for H_ε with $\varepsilon \neq 0$ generated by J_ε. If we choose a circle action, it is hard to guarantee that it is Hamiltonian. If we choose a symmetry generator, it is hard to guarantee that it generates a circle action.

We may accomplish both goals by deforming the circles of J_0 by a canonical transformation.

As ε varies from 0 this transformation varies from the identity.

Thus, it is convenient to think of the transformation as the time-one flow of a Hamiltonian vector field X_{K_ε} generated by the Hamiltonian K_ε. We would like to find K_ε so that: $0 = \{H_\varepsilon, J_0 \circ \exp(-X_{K_\varepsilon})\}$. We may expand consistently in ε:

$$0 = \{H_0 + \varepsilon H^1 + \varepsilon^2 H^2 + \dots, J_0 - \varepsilon X_{K^1} \cdot J_0 + \varepsilon^2 (X_{K^1}^2 - X_{K^2}) \cdot J_0 + \dots\}$$

$$= \underbrace{\cancel{\{H_0, J_0\}}}_{=0} + \varepsilon(\{H^1, J_0\} - \{H_0, X_{K^1} \cdot J_0\}) + \varepsilon^2(\{H_0, (X_{K^1}^2 - X_{K^2}) \cdot J - \{H^1, X_{K^1} \cdot J\} + \{H^2, J_0\}) + \dots$$

To first order $0 = \{H^1, J_0\} - \{H_0, \{J, K^1\}\} = -\{J_0, \{K^1, H_0\} + H^1\}$

One choice is $\{K^1, H_0\} + H^1 = 0$ which leads to the method of averaging. We see, though, that we may set this to be anything that Poisson commutes with J_0, i.e. has the circle symmetry.

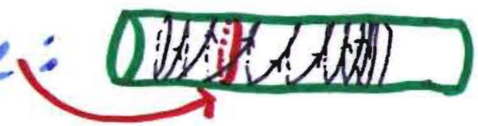
XII. Example of Non-Uniqueness of the circle symmetry in the general setting.

For the unperturbed flow

we just rotate a cylinder. Any perturbed flow typically has a discrete set of closed orbits separated by helices which limit on them.

Since Lie brackets are anti-commutative, if we want a symmetry generated by a vector field Z_ε, (i.e. $\mathcal{L}_{Z_\varepsilon}\cdot X_\varepsilon = 0 = [Z_\varepsilon, X_\varepsilon] = -\mathcal{L}_{X_\varepsilon}\cdot Z_\varepsilon$) then we need only define Z_ε at one point on each orbit of X_ε and then push it forward along the flow to the other points.

In the case at hand, we may pick **any** circle that crosses each orbit of the helical part exactly once:

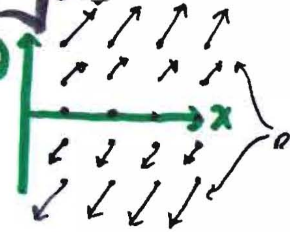


.

Take Z_ε to be tangent to this circle and carry it by the flow of X_ε to get an infinite family of symmetries.

XIII. Example of Rigorous Intrinsic Chaos in Gyromotion

Consider the 2-dimensional x-y motion of a charged particle in a linear magnetic field $B_z = \alpha y$.

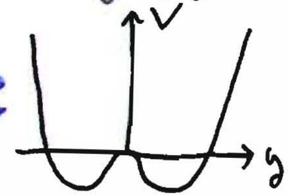


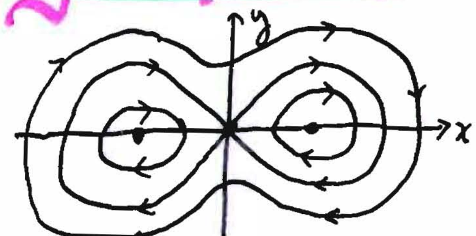
Define the vector potential by taking $A_y = 0$

so $B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} = -\frac{\partial A_x}{\partial y} = \alpha y$ so we take $A_x = -\frac{\alpha}{2} y^2$.

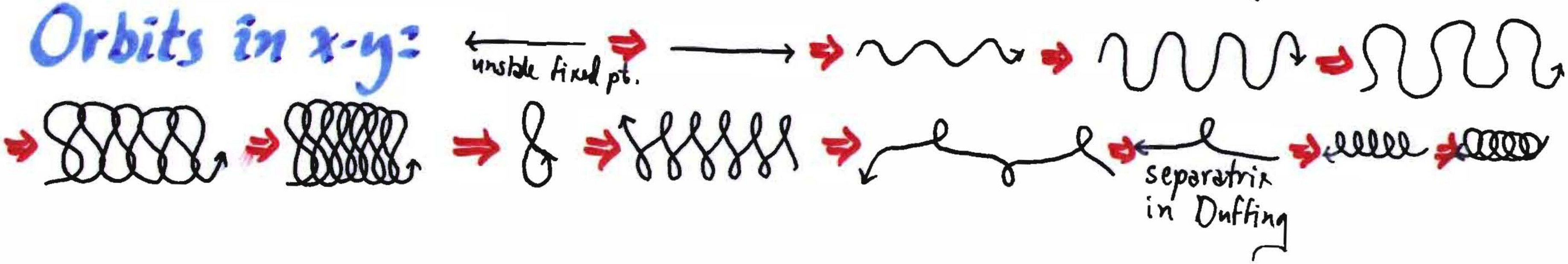
The Hamiltonian is $H = \frac{1}{2}\left(\left(p_x + \frac{\alpha}{2} y^2\right)^2 + p_y^2\right)$ and since it is x-translation symmetric, p_x is conserved. The equations of motion are: $\dot{x} = \frac{\partial H}{\partial p_x} = p_x + \frac{\alpha}{2} y^2$ $\quad \dot{y} = \frac{\partial H}{\partial p_y} = p_y$ $\quad \dot{p}_x = -\frac{\partial H}{\partial x} = 0$ $\quad \dot{p}_y = -\frac{\partial H}{\partial y} = \left(p_x + \frac{\alpha}{2} y^2\right)\alpha y$

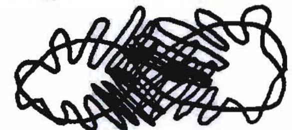
$\Rightarrow p_x = \text{constant} = p_x^0 \qquad \ddot{y} = \dot{p}_y = \alpha p_x^0 y + \frac{\alpha^2}{2} y^3$ i.e. **Duffing's Equation**

(Effective) Potential: 

Phase Portrait: 

Orbits in x-y:



Now perturb with low amplitude wave in the y direction to get periodically driven Duffing's eqn. This has been shown to **rigorously** have **Smale horseshoes** in the dynamics by the Meltnikov technique. This shows that the stable and unstable manifolds of the unstable fixed point cross transversally (and so must look like ). This forces a region of phase space () to be conjugate to a shift on sequences of 2 symbols.

> The consequence is that there are initial conditions for the particle with orbits like: 
> with **any** sequence of ... UUDDDU...

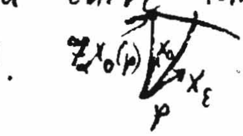
Mathematical Addendum

Exact Statements and Proofs of some of the results indicated previously.

Theorem: For ε in an interval I around 0 we are given a complete vector field X_ε on a manifold M.

We define $\Psi : \mathbb{R}\times I \times M \to M$ by $\Psi_{\alpha,\varepsilon}(p) = \mathcal{F}_\alpha X_\varepsilon(p)$, the flow for time α of X_ε.

Given α_0, $\Psi_{\alpha_0,\varepsilon}(p)$ defines a curve through $\Psi_{\alpha_0,\varepsilon_0}(p)$ as ε varies in I, (ε_0 in I).

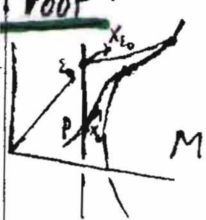


$\frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0}$ is a tangent vector in $T_{\Psi_{\alpha_0,\varepsilon_0}(p)}M$.

The claim is that
$$\boxed{\frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0} D\Psi_{\alpha_0-\alpha,\varepsilon_0}(\Psi_{\alpha,\varepsilon_0}(p)) \cdot \frac{\partial X_\varepsilon}{\partial\varepsilon}(\Psi_{\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=\varepsilon_0} d\alpha}$$

i.e. the infinitesimal perturbation of the time α_0 flow is the average of the infinitesimal perturbations along the integral curve of X_{ε_0} for time α_0.

Proof:



Define $\varphi_\alpha : I\times M \to I\times M$ by $(\varepsilon,p)\mapsto(\varepsilon, \mathcal{F}_\alpha X_\varepsilon(p)) = (\varepsilon, \Psi_{\alpha,\varepsilon}(p))$

i.e. it is the flow of $X = (0, X_\varepsilon)$

By definition of $\varphi_{\alpha *}$, it sends tangents to curves to tangents of the images of the curves.

So $\varphi_{\alpha_0 *}\left(\frac{\partial}{\partial\varepsilon}, 0\right)_{(\varepsilon_0,p)} = \frac{\partial}{\partial\varepsilon}\varphi_{\alpha_0}(\varepsilon,p)\Big|_{\varepsilon=\varepsilon_0}$ but by the definition of φ_α this is $= \left(\frac{\partial}{\partial\varepsilon}, \frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0}\right)$

By the fundamental formula relating flows and Lie derivatives:

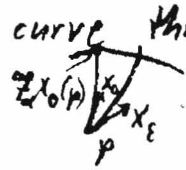
$$\frac{d}{d\alpha_0}\varphi_{\alpha_0 *}\left(\frac{\partial}{\partial\varepsilon},0\right) = -\varphi_{\alpha_0 *} L_X\left(\frac{\partial}{\partial\varepsilon},0\right) = \varphi_{\alpha_0 *} L_{\frac{\partial}{\partial\varepsilon}} X$$

$$= \frac{d}{d\alpha_0}\Big[\underbrace{\left(\frac{\partial}{\partial\varepsilon},0\right)}_{\text{constant}} + \int_0^{\alpha_0}\varphi_{\alpha *}\cdot L_{\frac{\partial}{\partial\varepsilon}} X \, d\alpha\Big]$$

But $\varphi_{0 *}\left(\frac{\partial}{\partial\varepsilon},0\right) = \left(\frac{\partial}{\partial\varepsilon},0\right) = \left(\frac{\partial}{\partial\varepsilon},0\right) + \int_0^0 \varphi_{\alpha *}\cdot L_{\frac{\partial}{\partial\varepsilon}} X\, d\alpha$

So, because they satisfy the same initial conditions and the same differential equation on $T(I\times M)$, uniqueness of solutions tells us

Given α_0, $\psi_{\alpha_0,\varepsilon}(p)$ defines a curve through $\psi_{\alpha_0,\varepsilon_0}(p)$ as ε varies in I, (ε_0 in I).



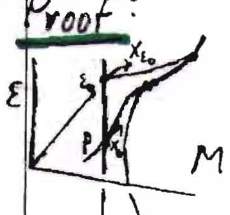
$\frac{\partial}{\partial\varepsilon}\psi_{\alpha_0,\varepsilon}(p)\big|_{\varepsilon=\varepsilon_0}$ is a tangent vector in $T_{\psi_{\alpha_0,\varepsilon_0}(p)}M$.

The claim is that

$$\boxed{\frac{\partial}{\partial\varepsilon}\psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0} D\psi_{\alpha_0-\alpha,\varepsilon_0}(\psi_{\alpha,\varepsilon_0}(p))\cdot\frac{\partial X_\varepsilon}{\partial\varepsilon}(\psi_{\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=\varepsilon_0} d\alpha}$$

i.e. the infinitesimal perturbation of the time α_0 flow is the average of the infinitesimal perturbations along the integral curve of X_{ε_0} for time α_0.

Proof:



Define $\varphi_\alpha : I\times M \to I\times M$ by $(\varepsilon,p)\mapsto(\varepsilon,\psi^{X_\varepsilon}_\alpha(p)) = (\varepsilon,\psi_{\alpha,\varepsilon}(p))$

i.e. it is the flow of $X=(0,X_\varepsilon)$

By definition of $\varphi_{\alpha*}$, it sends tangents to curves to tangents of the images of the curves.

So $\varphi_{\alpha_0*}\left(\frac{\partial}{\partial\varepsilon},0\right)_{(\varepsilon_0,p)} = \frac{\partial}{\partial\varepsilon}\varphi_{\alpha_0}(\varepsilon,p)\big|_{\varepsilon=\varepsilon_0}$ but by the definition of φ_α this is $=\left(\frac{\partial}{\partial\varepsilon}, \frac{\partial}{\partial\varepsilon}\psi_{\alpha_0,\varepsilon}(p)\big|_{\varepsilon=\varepsilon_0}\right)$

By the fundamental formula relating flows and Lie derivatives:

$$\frac{d}{d\alpha_0}\varphi_{\alpha_0*}\left(\frac{\partial}{\partial\varepsilon},0\right) = -\varphi_{\alpha_0*}L_X\left(\frac{\partial}{\partial\varepsilon},0\right) = \varphi_{\alpha_0*}L_{\frac{\partial}{\partial\varepsilon}}X$$

$$= \frac{d}{d\alpha_0}\Big[\underset{\uparrow\,\text{constant}}{\left(\frac{\partial}{\partial\varepsilon},0\right)} + \int_0^{\alpha_0}\varphi_{\alpha*}\cdot L_{\frac{\partial}{\partial\varepsilon}}X\,d\alpha\Big]$$

But $\varphi_{0*}\left(\frac{\partial}{\partial\varepsilon},0\right) = \left(\frac{\partial}{\partial\varepsilon},0\right) = \left(\frac{\partial}{\partial\varepsilon},0\right) + \int_0^0 \varphi_{\alpha*}\cdot L_{\frac{\partial}{\partial\varepsilon}}X\,d\alpha$

So, because they satisfy the same initial conditions and the same differential equation on $T(I\times M)$, uniqueness of solutions tells us that: $\varphi_{\alpha_0*}\left(\frac{\partial}{\partial\varepsilon},0\right) = \left(\frac{\partial}{\partial\varepsilon},0\right) + \int_0^{\alpha_0}\varphi_{\alpha*}\cdot L_{\frac{\partial}{\partial\varepsilon}}X\,d\alpha$

So $\left(\frac{\partial}{\partial\varepsilon}, \frac{\partial}{\partial\varepsilon}\psi_{\alpha_0,\varepsilon}(p)\big|_{\varepsilon=\varepsilon_0}\right)_{(\varepsilon_0,\psi_{\alpha_0,\varepsilon_0}p)} = \left(\frac{\partial}{\partial\varepsilon}, \int_0^{\alpha_0} D\psi_{\alpha_0-\alpha,\varepsilon_0}(\psi_{\alpha,\varepsilon_0}(p))\cdot\frac{\partial X_\varepsilon}{\partial\varepsilon}(\psi_{\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=0} d\alpha\right)$

which implies the desired result. ■

Corollary 1 For ε in an interval I around 0 we are given a complete vector field X_ε on a manifold M such that the orbits of $X_{\varepsilon=0}$ are all periodic of period 1.
We may define a one-parameter family of diffeomorphisms $\Psi: I\times M \to M$ by $\Psi_\varepsilon(p) = \varphi_{t=1}X_\varepsilon(p)$, the time one flow of X_ε. Ψ_0 is the identity by hypothesis.
$\frac{d}{d\varepsilon}\Psi_\varepsilon(p)\big|_{\varepsilon=0}$ is a tangent vector in T_pM.

The claim is ① $\boxed{\frac{d}{d\varepsilon}\Psi_\varepsilon(p)\big|_{\varepsilon=0} = \int_0^1 D(\varphi_\alpha X_0)_{(\varphi_{-\alpha}X_0)(p)}\cdot \tilde{X}_{(\varphi_{-\alpha}X_0)(p)}\, d\alpha}$

where $\tilde{X}(p) = \frac{\partial}{\partial\varepsilon}X_\varepsilon(p)\big|_{\varepsilon=0}$ is the infinitesimal perturbation of X_0, identifying TTM with TM.
i.e. to first order in ε the drift from loop to loop is the average of the perturbation.

Furthermore ② $\boxed{\varphi_\alpha X_{0*}\left(\frac{d}{d\varepsilon}\Psi_\varepsilon\big|_{\varepsilon=0}\right) = \frac{d}{d\varepsilon}\Psi_\varepsilon\big|_{\varepsilon=0}}$

i.e. the average drift is the same everywhere on a loop.

Proof: ① Follows from the theorem taking $\alpha_0 = 1$, $\varepsilon_0 = 0$ and changing variables $\alpha \to -\alpha$ and using $\varphi_{1+\alpha}X_0 = \varphi_1 X_0 \circ \varphi_\alpha X_0 = \mathbb{I}\circ\varphi_\alpha X_0 = \varphi_\alpha X_0$

② $\varphi_\alpha X_{0*}\left(\frac{d}{d\varepsilon}\Psi_\varepsilon\big|_{\varepsilon=0}\right)(p) = D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\cdot\left(\frac{d}{d\varepsilon}\Psi_\varepsilon(\varphi_{-\alpha}X_0(p))\big|_{\varepsilon=0}\right)$

$= D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\cdot\left(\int_0^1 D(\varphi_\beta X_0)_{(\varphi_{-\beta}X_0)(\varphi_{-\alpha}X_0(p))}\cdot\tilde{X}_{(\varphi_{-\beta}X_0)(\varphi_{-\alpha}X_0(p))}\, d\beta\right)$

$= \int_0^1 D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\circ D\varphi_\beta X_{0\,(\varphi_{-\beta-\alpha}X_0(p))}\cdot\tilde{X}_{(\varphi_{-\beta-\alpha}X_0(p))}\, d\beta$ use chain rule

$= \int_0^1 D\varphi_{\beta+\alpha}X_{0\,(\varphi_{-(\beta+\alpha)}X_0\, p)}\cdot\tilde{X}_{(\varphi_{-(\beta+\alpha)}X_0\, p)}\, d\beta$ change variables to $\beta+\alpha$

$= \int_0^1 D\varphi_\beta X_{0\,(\varphi_{-\beta}X_0\, p)}\cdot\tilde{X}_{(\varphi_{-\beta}X_0\, p)}\, d\beta = \frac{d}{d\varepsilon}\Psi_\varepsilon\big|_{\varepsilon=0}$ ∎

We may rewrite theorem 1 by changing $\alpha \leftrightarrow \alpha_0 - \alpha$

: $\boxed{\frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0} D\Psi_{\alpha,\varepsilon_0}(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p))\cdot\frac{\partial X_\varepsilon}{\partial\varepsilon}(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=\varepsilon_0} d\alpha}$

talking about vector fields instead of vectors:

the orbits of $X_{\varepsilon=0}$ are all periodic of period 1.
We may define a one-parameter family of diffeomorphisms $\Psi : I\times M \to M$ by $\Psi_\varepsilon(p) = \varphi_{t=1}X_\varepsilon(p)$, the time one flow of X_ε. Ψ_0 is the identity by hypothesis.

$\frac{d}{d\varepsilon}\Psi_\varepsilon(p)\big|_{\varepsilon=0}$ is a tangent vector in T_pM.

The claim is ① $$\boxed{\frac{d}{d\varepsilon}\Psi_\varepsilon(p)\Big|_{\varepsilon=0} = \int_0^1 D(\varphi_\alpha X_0)_{(\varphi_{-\alpha}X_0)(p)}\cdot \tilde{X}_{(\varphi_{-\alpha}X_0)(p)}\, d\alpha}$$

where $\tilde{X}(p) = \frac{\partial}{\partial\varepsilon}X_\varepsilon(p)\big|_{\varepsilon=0}$ is the infinitesimal perturbation of X_0, identifying TTM with TM.
i.e. to first order in ε the drift from loop to loop is the average of the perturbation.

Furthermore ② $$\boxed{\varphi_\alpha X_{0*}\left(\frac{d}{d\varepsilon}\Psi_\varepsilon\Big|_{\varepsilon=0}\right) = \frac{d}{d\varepsilon}\Psi_\varepsilon\Big|_{\varepsilon=0}}$$

i.e. the average drift is the same everywhere on a loop.

Proof: ① Follows from the theorem taking $\alpha_0=1$, $\varepsilon_0=0$ and changing variables $\alpha\to-\alpha$ and using $\varphi_{1+\alpha}X_0 = \varphi_1X_0\circ\varphi_\alpha X_0 = \mathrm{I}\circ\varphi_\alpha X_0 = \varphi_\alpha X_0$

② $$\varphi_\alpha X_{0*}\left(\frac{d}{d\varepsilon}\Psi_\varepsilon\Big|_{\varepsilon=0}\right)(p) = D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\cdot\left(\frac{d}{d\varepsilon}\Psi_\varepsilon(\varphi_{-\alpha}X_0(p))\Big|_{\varepsilon=0}\right)$$

$$= D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\cdot\left(\int_0^1 D(\varphi_\beta X_0)_{(\varphi_{-\beta}X_0)(\varphi_{-\alpha}X_0(p))}\cdot\tilde{X}_{(\varphi_{-\beta}X_0)(\varphi_{-\alpha}X_0(p))}\,d\beta\right)$$

$$= \int_0^1 D\varphi_\alpha X_{0\,(\varphi_{-\alpha}X_0(p))}\circ D\varphi_\beta X_{0\,(\varphi_{-\beta-\alpha}X_0(p))}\cdot\tilde{X}_{\varphi_{-\beta-\alpha}X_0(p))}\,d\beta \quad \text{use chain rule}$$

$$= \int_0^1 D\varphi_{\beta+\alpha}X_{0\,(\varphi_{-(\beta+\alpha)}X_0\,p)}\cdot\tilde{X}_{(\varphi_{-(\beta+\alpha)}X_0\,p)}\,d\beta \quad \text{change variables to } \beta+\alpha$$

$$= \int_0^1 D\varphi_\beta X_{0\,(\varphi_{-\beta}X_0\,p)}\cdot\tilde{X}_{(\varphi_{-\beta}X_0\,p)}\,d\beta = \frac{d}{d\varepsilon}\Psi_\varepsilon\Big|_{\varepsilon=0}$$ $\blacksquare$

We may rewrite theorem 1 by changing $\alpha\leftrightarrow\alpha_0-\alpha$:

$$\boxed{\frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0} D\Psi_{\alpha,\varepsilon_0\,(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p))}\cdot\frac{\partial X_\varepsilon}{\partial\varepsilon}(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=\varepsilon_0}d\alpha}$$

or talking ~~about~~ vector fields instead of vectors:

$$\boxed{\frac{\partial}{\partial\varepsilon}\Psi_{\alpha_0,\varepsilon}\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0}\Psi_{\alpha,\varepsilon_0 *}\frac{\partial X_\varepsilon}{\partial\varepsilon}\Big|_{\varepsilon=\varepsilon_0}d\alpha}$$

Corollary 2 Same setting as in theorem 1 if M is in addition poisson manifold with bracket $\{,\}$ and X_ε are globally hamiltonian with hamiltonian H_ε (so $X_\varepsilon = X_{H_\varepsilon} \equiv \{\cdot, H_\varepsilon\}$)

then in fact $\frac{\partial}{\partial \varepsilon} \Psi_{\alpha_0,\varepsilon}\big|_{\varepsilon=\varepsilon_0}$ is hamiltonian with hamiltonian

$$\boxed{= \int_0^{\alpha_0} \frac{\partial H_\varepsilon}{\partial \varepsilon}\Big|_{\varepsilon=0} \circ \Psi_{-\alpha,\varepsilon_0} \, d\alpha}$$

i.e. the hamiltonian generating the perturbation of the α_0 flow to first order in ε is the average along the flow of the hamiltonian generating the infinitesimal perturbation at a point

Proof: First note that $\frac{\partial X_{H_\varepsilon}}{\partial \varepsilon}(p)\Big|_{\varepsilon=0} \cdot f = \frac{\partial}{\partial \varepsilon}\left(X_{H_\varepsilon(p)} \cdot f\right)\Big|_{\varepsilon=0}$

$= \frac{\partial}{\partial \varepsilon}\left(\{f, H_\varepsilon\}(p)\right)\Big|_{\varepsilon=0} = \{f, \frac{\partial H_\varepsilon}{\partial \varepsilon}\Big|_{\varepsilon=0}\}(p) = X_{\frac{\partial H_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0}} \cdot f\ (p)$

So $\frac{\partial X_{H_\varepsilon}}{\partial \varepsilon}\Big|_{\varepsilon=0} = X_{\frac{\partial H_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0}}$ Call $\frac{\partial H_\varepsilon}{\partial \varepsilon}\Big|_{\varepsilon=0} \equiv \widetilde{H}$

Next note that if Ψ is a poisson map (i.e. $\{f\circ\Psi, g\circ\Psi\} = \{f,g\}\circ\Psi$)

then $\left(D\Psi_{(p)} \cdot X_{\widetilde{H}(p)}\right)(f) = X_{\widetilde{H}(p)}(f\circ\Psi) = \{f\circ\Psi, \widetilde{H}\}(p)$

$= \{f\circ\Psi, \widetilde{H}\circ\Psi^{-1}\circ\Psi\}(p) = \{f, \widetilde{H}\circ\Psi^{-1}\}(\Psi(p)) = X_{\widetilde{H}\circ\Psi^{-1}}(\Psi(p))\, f$

So $D\Psi_{(p)} \cdot X_{\widetilde{H}(p)} = X_{\widetilde{H}\circ\Psi^{-1}}(\Psi(p))$

Thus $\frac{\partial}{\partial \varepsilon} \Psi_{\alpha_0,\varepsilon}(p)\Big|_{\varepsilon=\varepsilon_0} = \int_0^{\alpha_0} D\Psi_{\alpha,\varepsilon_0}(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p)) \cdot X_{\widetilde{H}}(\Psi_{\alpha_0-\alpha,\varepsilon_0}(p))\Big|_{\varepsilon=\varepsilon_0} d\alpha$

$= \int_0^{\alpha_0} X_{\widetilde{H}\circ\Psi_{-\alpha,\varepsilon_0}}(\Psi_{\alpha_0,\varepsilon_0}(p))\, d\alpha$ but X_H is linear in H so

$= X_{\int_0^{\alpha_0} \widetilde{H}\circ\Psi_{-\alpha,\varepsilon_0} d\alpha}(\Psi_{\alpha_0,\varepsilon_0}(p))$ as desired ∎

Space $= x, y, p_x, p_y$ Bracket: $\{f,g\} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial p_x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial p_y} - \frac{\partial f}{\partial p_x}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial p_y}\frac{\partial g}{\partial y}$

ven $B_z(x,y) = B_z^0 + \varepsilon B_z^1(x,y)$ $\quad B_z^1(0,0)=0$; $A_x = 0$ $\quad B_z = \frac{\partial A_y}{\partial x}$

ake $A_y^0(x,y) = B_z^0 x$ $\quad A_y^1(x,y) = \int_0^x B_z^1(x',y)\,dx'$

amiltonian $H = \frac{1}{2}(\underline{p} - \underline{A}_0 - \varepsilon \underline{A}_1)^2$

ns. of Motion: $\dot{x} = p_x$ $\quad \dot{y} = p_y - A_y^0 - \varepsilon A_y^1$

$$\dot{p}_x = (p_y - A_y^0)B_z^0 + \varepsilon\left[(p_y - A_y^0)B_z^1 - A_y^1 B_z^0\right] - \varepsilon^2 A_y^1 B_z^1$$

$$\dot{p}_y = \varepsilon(p_y - A_y^0)\frac{\partial A_y^1}{\partial y} - \varepsilon^2 A_y^1 \frac{\partial A_y^1}{\partial y}$$

nperturbed eqns: $\dot{x} = p_x$ $\quad \dot{y} = p_y^0 - B_z^0 x$ $\quad \dot{p}_x = p_y^0 B_z^0 - B_z^{0\,2} x$ $\quad \dot{p}_y = 0$

n to Unperturbed eqn. w/ initial conditions: (x^0, y^0, p_x^0, p_y^0)

$$x(t) = \frac{p_y^0}{B_z^0} + \left(x^0 - \frac{p_y^0}{B_z^0}\right)\cos B_z^0 t + p_x^0 \sin B_z^0 t$$

$$y(t) = (p_x^0 + y^0) - \left(x^0 - \frac{p_y^0}{B_z^0}\right)\sin B_z^0 t - p_x^0 \cos B_z^0 t$$

$$p_x(t) = -B_z^0\left(x^0 - \frac{p_y^0}{B_z^0}\right)\sin B_z^0 t + B_z^0 p_x^0 \cos B_z^0 t$$

$$p_y(t) = p_y^0$$

nerator of circle action for $\varepsilon = 0$:

$$J_{\varepsilon=0} = \frac{1}{B_z^0} H_{\varepsilon=0} = \frac{1}{2B_z^0}(p_x^2 + p_y^2) + \frac{1}{2}B_z^0 x^2 - p_y x$$

$$X_J = \frac{p_x}{B_z^0}\frac{\partial}{\partial x} + \frac{p_y - A_y^0}{B_z^0}\frac{\partial}{\partial y} + (p_y - A_y^0)\frac{\partial}{\partial p_x}$$

miltonian in terms of $J_{\varepsilon=0}$: $H_\varepsilon = B_z^0 J - \varepsilon(p_y - A_y^0)A_y^1 + \frac{1}{2}\varepsilon^2 A_y^{1\,2} = H_0 + \varepsilon H_1 + \varepsilon^2 H_2$