0.1 Introduction

Birkhoff Normal form notes taken from an old set of Rajeev’s notes and these are based on Gustavson, F. G. 1966, A.J, 71, 680. A nice explanation and connection to KAM is given in an Introduction to dynamical systems by Arrowsmith and Place (in library, green book). In future expand using this book perhaps.

Use of Hamilton-Jacobi equation originally taken from the 3 body problem book by Inmanen? Transformation to Poincaré or Delauney coordinates is also illustrated in Morbidelli’s book. Rough outline of KAM theory is following Morbidelli’s book, but should be improved and expanded perhaps rolling Sylvio Ferraz-Mello’s book. Lie derivatives should be explained more carefully in future. Problems could be added on the Birkhoff normal form and with Lie derivatives. Lagrange’s equations and how they follow from the Hamiltonian in canonical coordinates could be illustrated.
1 The Birkhoff Normal form

Suppose we have a Hamiltonian

\[ H(q, p) = \frac{1}{2} \sum_i \omega_i (p_i^2 + q_i^2) + O(3) \]

that consists of \( N \) harmonic oscillators and some higher order terms that are cubic in \( p, q \).

We can regard the Hamiltonian as a power series

\[ H(q, p) = \sum_{r=2}^{\infty} H^r(q, p) \]

where

\[ H^r(q, p) = \sum_{|A|+|B|=r} H_{AB} q_1^{a_1} q_2^{a_2} q_3^{a_3} \ldots p_1^{b_1} p_2^{b_2} \ldots \]

and \( |A| = a_1 + a_2 + a_3 \ldots + a_n \), and \( |B| = \sum_i b_i \). In other words

\[ H^r(q, p) = \sum_{|A|+|B|=r} H_{AB} \prod_i q_i^{a_i} p_i^{b_i} \]

and \( H^r \) is a polynomial with degree \( r = |A| + |B| \).

We desire a way to perform a canonical transformation to remove orders of the Hamiltonian with \( r > 2 \).

It is convenient to introduce the operator

\[ D = \sum_i \omega_i \left( q_i \frac{\partial}{\partial p_i} - p_i \frac{\partial}{\partial q_i} \right) \]

This operator is reminiscent of the exterior derivative. Consider a one form

\[ f^r = H^r(p, q) \sum_i \omega_i (q_idq_i + p_idp_i) \]

and take its exterior derivative

\[ df^r = \sum_i \omega_i \left( q_i \frac{\partial H^r}{\partial p_i} - p_i \frac{\partial H^r}{\partial q_i} \right) dp_i \wedge dq_i \]

\[ = (DH^r) dp_i \wedge dq_i \]

If \( DH^r = 0 \) then the one form \( f^r \) is closed as \( df^r = 0 \).

The operator

\[ D = \sum_i \omega_i \frac{\partial}{\partial \theta_i} \]
where $\theta_i$ are normal to the coordinate $I_i = \frac{1}{2}(q_i^2 + p_i^2)$.

What types of polynomials of order $r$ are in the kernel of $D$? In other words what one forms in written like $f^r$ are closed? Let us operate on $I_i = \frac{1}{2}(p_i^2 + p_i^2)$

\[ DI_i = \frac{\omega_i}{2}(2q_ip_i - 2p_iq_i) = 0 \]

Likewise

\[ DI_i^n = n I_i^{n-1} DI_i = 0 \]

All powers of $I_i$ are in the kernel of $D$.

More generally we can transfer polynomials in $q, p$ to eigenvectors of $D$. The vector $z_j = q_j + i p_j$ is an eigenvector of $D$ with eigenvalue $i \omega_j$. The vector $\bar{z}_j = q_j - i p_j$ is an eigenvector of $D$ with eigenvalue $-i \omega_j$. So $D$ can be diagonalized in a basis where polynomials are written with respect to $z, \bar{z}$. The product $z_j \bar{z}_j = q_j^2 + p_j^2 = 2I_j$.

Operating on a product (and only taking into account one degree of freedom)

\[ Dz^a \bar{z}^b = \omega(q \partial_p - p \partial_q)(q + ip)^a(q - ip)^b \]
\[ = \omega\left[ qai(q + ip)^{a-1}(q - ip)^b - qbi(q + ip)^a(q - ip)^{b-1} - pa(q + ip)^{a-1}(q - ip)^b - bp(q + ip)^a(q - ip)^{b-1} \right] \]
\[ = \omega(q + ip)^a(q - ip)^b - bi(q + ip) - pa(q - ip) - pb(q + ip) \]
\[ = \omega_i(a - b)z^{a-1}z^{b-1}z\bar{z} \]
\[ = \omega_i(a - b)z^a \bar{z}^b \]

Taking into account the other degrees of freedom

\[ Dz^A \bar{z}^B = i \sum I_i \omega_i(a_i - b_i)z^A \bar{z}^B \]

where $z^A = z_1^a z_2^a \ldots$ and likewise for $\bar{z}^B$. If $\sum_i \omega_i(a_i - b_i) = 0$ then $Dz^A \bar{z}^B = 0$. Polynomials in the kernel of $D$ can be expressed as polynomials in $z, \bar{z}$ such that $\sum_i \omega_i(a_i - b_i) = 0$.

A pseudo inverse of a polynomial for the part that is not in the kernel can be constructed with

\[ D^{-1}z^A \bar{z}^B = \frac{z^A \bar{z}^B}{i \sum \omega_i(a_i - b_i)} \]

if the denominator is not zero and $D^{-1}z^A \bar{z}^B = 0$ otherwise. If $H^s$ is transformed to depend on $z, \bar{z}$ we then can set

\[ w^s = D^{-1}H^s \]

Operating on $w^s$ with $D$ we find $Dw^s = H^s$. 

3
How do we make a canonical transformation to remove order by order in the Hamiltonian? Let us use a generating function of old coordinates and new momenta

\[ S(q, P) = \sum_i q_i P_i + w^s(q, P) \]

where \( w^s(q, P) \) is a polynomial in \( q, P \) that is of order \( s \).

New coordinates are given by

\[ Q_i = q_i + \frac{\partial w^s}{\partial P_i} \]

And momenta are related by

\[ p_i = P_i + \frac{\partial w^s}{\partial q_i} \]

The new Hamiltonian is

\[ H(q, p) = K(Q, P) \]

Recall our original Hamiltonian has

\[ H = H_0 + H_1 \]

with

\[ H_0 = \sum_i \frac{\omega_i}{2} (p_i^2 + q_i^2) \]

and \( H_1 \) has polynomials that are higher order (order 3). Our canonical transformation, to first order in the changes can be written

\[ q_i = Q_i - \frac{\partial w^s}{\partial P_i}(P, Q) \]

\[ p_i = P_i + \frac{\partial w^s}{\partial q_i}(P, Q) \]

Inserting these \( H_0 \) gives to first order in the changes

\[ H_0(q(Q, P), p(Q, P)) = \sum_i \left[ \frac{\omega_i}{2} (P_i^2 + Q_i^2) + \omega_i (P_i \frac{\partial w^s}{\partial Q} - Q_i \frac{\partial w^s}{\partial P}) \right] + .. \]

\[ = \sum_i \left[ \frac{\omega_i}{2} (P_i^2 + Q_i^2) - Dw^s \right] + .. \]

Meanwhile the higher order \( H_1 \) term can be written solely in terms of \( P, Q \). But because it is higher order, its term should match the first order term above. Taking only the lowest order terms

\[ H(q(Q, P), p(Q, P)) = \sum_i \left[ \frac{\omega_i}{2} (P_i^2 + Q_i^2) - Dw^s \right] + H_1(P, Q) \]
If we have chosen \( w^s \) so that \( Dw^s = H_1 \) then the two terms cancel and we are left with a Hamiltonian that only depends on \( P_i^2 + Q_i^2 \).

As long as there are no resonances out to order \( s \) this procedure can be done consecutively.

By iterating this procedure a polynomial Hamiltonian can be reduced to normal form to any order (as long as there is no resonance condition). If the frequencies are linearly independent to all orders then we can construct a complete set of conserved quantities \( I_i = p_i^2 + q_i^2 \). However, the series often does not converge. If the series always converged then all systems would be integrable. Even with a non-integrable system it may be useful to compute the first few orders achieving a good approximate solution.

How can we write an arbitrary polynomial in terms of \( z, \bar{z} \)?

\[
q = \frac{1}{2}(z + \bar{z}) \quad p = \frac{1}{2i}(z - \bar{z})
\]

Suppose \( H^3 = q^3 \)

\[
H^3 = q^3 = \frac{1}{8}(z + \bar{z})^3 = z^3 + \bar{z}^3 + 3z^2\bar{z} + 3z\bar{z}^2
\]

Operating with \( D \) on this

\[
DH^3 = \frac{i\omega}{8}(3z^3 - 3\bar{z}^3 + 3(2 - 1)z^2\bar{z} + 3(1 - 2)z\bar{z}^2)
\]

\[
= \frac{i\omega}{8}(3z^3 - 3\bar{z}^3 + 3z^2\bar{z} - 3z\bar{z}^2)
\]

\[
D^{-1}H^3 = \frac{1}{8\omega i} \left( \frac{z^3}{3} + \frac{\bar{z}^3}{3} + 3 \frac{z^2\bar{z}}{1} + 3 \frac{z\bar{z}^2}{1} \right)
\]

\[
= \frac{1}{8\omega i} \left( \frac{z^3}{3} - \frac{\bar{z}^3}{3} + 3z^2\bar{z} - 3z\bar{z}^2 \right)
\]

We now set \( w^3 = D^{-1}H^3 \) and use this to make the canonical transformation.

Not well explained here is how the resonance condition makes problems and why this works for a Hamiltonian that depends on powers of the action variable. See Arnold appendix 7 for a terse explanation.

2 The Hamilton-Jacobi equation

The Hamilton-Jacobi equation is a first order differential equation that when solved gives you conserved canonical momenta. If a complete set of momenta can be found, then the Hamiltonian system is integrable.

5
We consider $H(q, p)$

We perform a canonical transformation using a generating function that is a function of old coordinates and new momenta $S_2(q, P, t)$

The transformation gives a relation between new and old coordinates

$$\frac{\partial S(q, P, t)}{\partial P_i} = p_i$$

$$\frac{\partial S(q, P, t)}{\partial q_i} = Q_i$$

and new Hamiltonian

$$K(Q, P, t) = H(q(P, Q), p(P, Q), t) + \frac{\partial S}{\partial t}$$

The Hamilton-Jacobi equation is this equation set to $K = 0$. In other words

$$K(Q, P, t) = H(q, \frac{\partial S}{\partial P_i}, t) + \frac{\partial S}{\partial t} = 0$$

If $K = 0$ then Hamilton’s equations imply that all momenta are conserved. Hence the new momenta are conserved quantities. The idea is to solve this equation finding new momenta that are conserved. All momenta and coordinates found this way will be canonical and obey Poisson brackets.

2.1 Finding conserved quantities in the Kepler problem using the Hamilton-Jacobi equation

Recall the Kepler Hamiltonian in cylindrical coordinates

$$H(p_r, L; r, \theta) = \frac{p_r^2}{2} + \frac{L^2}{2r^2} - \frac{k}{r}$$

The Hamilton Jacobi equation is

$$H\left(r, \frac{\partial S}{\partial r}, \frac{\partial S}{\partial \theta}\right) + \frac{\partial S}{\partial t} = 0$$

(here $H$ does not depend on $\theta$ so this is omitted. The new momenta are

$$P_1 = \frac{\partial S}{\partial r} \quad P_2 = \frac{\partial S}{\partial \theta}$$
Let us assume that $S$ is separable.

$$S(r, \theta, P_1, P_2) = S_r(r) + S_\theta(\theta) + S_t(t)$$

where each of these functions also will depend on the new momenta. The old momenta

$$p_r = \frac{\partial S_r}{\partial r} \quad L = \frac{\partial S_\theta}{\partial \theta}$$

Using the Hamiltonian and substituting in for $p_r, L$, the Hamiltonian Jacobi equation

$$K = \frac{1}{2} \left( \frac{\partial S_r}{\partial r} \right)^2 + \left( \frac{\partial S_\theta}{\partial \theta} \right) \frac{1}{2r^2} - \frac{k}{r} + \frac{\partial S_t}{\partial t} = 0$$

(1)

Since the system is separable we can define constants

$$P_1 = -\frac{\partial S_t}{\partial t} \quad P_2 = \frac{\partial S_\theta}{\partial \theta}$$

Integrating these constants

$$S_\theta(\theta) = P_2 \theta + \text{constant} \quad S_t(t) = -P_1 t + \text{constant}$$

We will identify the constants with the new momenta. Inserting these new momenta into the Hamilton-Jacobi equation (equation 1)

$$S_r^2 = 2 \left( P_1 + \frac{k}{r} \right) - \frac{P_2^2}{r^2}$$

and back into our function for $S$

$$S = -P_1 t + P_2 \theta + \int^r dr' \sqrt{2P_1 + \frac{2k}{r'} - \frac{P_2^2}{r'^2}}$$

Since $K = 0$ and $K = H + \frac{\partial S}{\partial t} = H - P_1 = 0$ we know that $P_1$ is the energy of the orbit;

$$P_1 = -\frac{k}{2a}$$

Let us check our old momenta

$$L = \frac{\partial S}{\partial \theta} = L = P_2$$

One of our new momenta is just the angular momentum. We already knew that the angular momentum was conserved so it should have come out of the Hamilton Jacobi equation. We can find our new coordinates

$$Q_1 = \frac{\partial S}{\partial P_1} = -t + \int^r dr' \left[ 2P_1 + \frac{2k}{r'} - \frac{P_2^2}{r'^2} \right]^{-1/2}$$
Using
\[ r = a(1 - e \cos E) \quad dr = ae \sin E dE \]
and \[ L = \sqrt{ka(1 - e^2)} \]

\[
Q_1 = -t + \frac{a^{3/2}}{k} \int (1 - e \cos E) dE
\]
\[
= -t + \frac{1}{n} (E - e \sin E)
\]
\[
= -t + \frac{M}{n}
\]

where we have used mean motion \( n = a^{3/2}/\sqrt{k} \). The angle \(-t + M/n\) can be associated with the time of perihelion \( \tau \). Performing a similar integral it is possible to show that

\[ Q_2 = \theta - f \]

with \( f \) the true anomaly. We associate the angle, \( \theta - f = \omega \), the angle of perihelion. To summarize

\[
Q_1 = -t + \frac{M}{n} = \tau
\]
\[
Q_2 = \theta - f = \omega
\]
\[
P_1 = -\frac{k}{2a} = \text{energy}
\]
\[
P_2 = \sqrt{ka(1 - e^2)} = L
\]

These coordinates and momenta are not only conserved but canonical. By redoing the problem in 3D and adding and subtracting these, it is possible to find more popular sets of coordinates, the Delauney, modified Delauney and Poincaré coordinates.

3 Some Perturbation Theory

KAM theory begins with a near integrable Hamiltonian that depends on a small parameter \( \epsilon \).

\[ H(p, q) = H_0(p) + \epsilon H_1(p, q) \]

Using a perturbative expansion, for each order of \( \epsilon \) an attempt is made to construct a canonical transformation that transforms the Hamiltonian to integrable form. Near identity canonical transformations are used to do this. Rather than describe canonical transformations in terms of generating functions we describe here a more direct Lie-derivative approach.
3.1 Lie derivatives and generating near-identity canonical transformations

Recall that time derivatives of a function can be written in terms of a Poisson bracket with a Hamiltonian

\[ \frac{df}{dt} = \{f, H\} \]

We can define a derivative function, known as the Lie derivative

\[ \mathcal{L}_H f = \{f, H\} \]

Here the Lie derivative evaluates the change of a scalar function along the flow of another vector field. In other words \( H \) generates a flow and \( \mathcal{L}_H \) gives the change of \( f \) along this flow.

If we apply the Lie derivative twice

\[ \mathcal{L}_H^2 f = \{\{f, H\}, H\} \]

and likewise \( i \) times \( \mathcal{L}_H^i f \). The function as a function of time to first order

\[ f(t) = f(0) + t \frac{df}{dt} + ... = f(0) + t \mathcal{L}_H f(0) + ... \]

and to higher order

\[ f(t) = f(0) + \sum_{i=0}^{\infty} \frac{t^i}{i!} \mathcal{L}_H^i f(0) \]

The Taylor series generates \( f \) as a function of time from a sum of Lie derivatives.

Previously we showed that a Hamiltonian flow (the flow made by a Hamiltonian using the equations of motion) generated canonical transformations. In other words, the transformation \( q(0), p(0) \rightarrow q(t), p(t) \) is a canonical transformation. In the same way other functions (other than the Hamiltonian) can be used to make a smooth series of canonical transformations. An advantage of this approach is that near identity transformations are constructed.

Given a function \( \chi \), called a Hamiltonian generating function, we can generate an infinitesimal transformation for small parameter \( s \) of the scalar function \( f \) with

\[ f(s) = f(0) + s\{f, \chi\} = f(0) + \mathcal{L}_\chi f(0) \]

Here \( s \) serves like time along the flow generated by \( \xi \). For larger \( s \)

\[ f(s) = f(0) + \sum_{i} \frac{s^i}{i!} \mathcal{L}_\chi^i f(0) \]

The function can be a smooth function of \( q, p \) and \( s \).
It is possible to take a Lie derivative of a vector. A canonical transformation transforms $q,p$ to $Q,P$ with generating Hamiltonian function $\xi$ via

$$
Q(s) = q + \sum_i \frac{s^i}{i!} L^i_\chi q \\
P(s) = p + \sum_i \frac{s^i}{i!} L^i_\chi p
$$

The result is a one parameter family of canonical transformations, depending on $s$, where for small $s$ the transformation is near the identity.

We can use for short hand $S^\epsilon_\chi$ for the transformation and let it act on coordinates or on a Hamiltonian.

### 3.2 First order canonical transformation of the near integrable Hamiltonian

Starting with a Hamiltonian

$$H(p,q) = H_0(p) + \epsilon H_1(p,q)$$

we can try to find a series of near identity canonical transformations that allow the Hamiltonian to become $\tilde{H}(\tilde{p})$, only dependent on momenta.

Suppose we transform a Hamiltonian $H$ with a canonical transformation generated with the Hamiltonian generating function $\chi$ and using small $\epsilon$ to describe how distant the transformation is from the identity. We get a new Hamiltonian

$$K = S^\epsilon_\chi H$$

Writing this out to second order in $\epsilon$

$$K = H + \epsilon \{H, \chi\} + \frac{\epsilon^2}{2} \{\{H, \chi\}, \chi\}$$

Now we sub in for $H = H_0 + \epsilon H_1$

$$K = H_0 + \epsilon H_1 + \epsilon \{H_0, \chi\} + \frac{\epsilon^2}{2} \{\{H_0, \chi\}, \chi\} + \frac{\epsilon^3}{2} \{\{H_1, \chi\}, \chi\}$$

Now let us expand $K$ to first order in $\epsilon$,

$$K = K_0 + \epsilon K_1$$

Taking only the first order terms we find that

$$K_1 = H_1 + \{H_0, \chi\}$$
This equation is known as the homologic equation. We want to design \( \chi \) so that \( K_1 \) only depends on momenta. By solving for \( \chi \) and \( K_1 \) simultaneously we can strive to remove the first order terms that depend on \( q \). We can then repeat the procedure at higher orders. This is reminiscent of using Newton’s method to converge to a root, but here we are repeatedly removing the lowest order term in \( \epsilon \).

Using a Fourier series for \( H_1 \)

\[
H_1 = \sum_k c_k(p)e^{ik\cdot q}
\]

and a Fourier series for \( \chi \)

\[
\chi = \sum_k d_k(p)e^{ik\cdot q}
\]

\[
\{H_0, \chi\} = -i \sum_k d_k(p)k \cdot \omega_0(p)e^{ik\cdot q}
\]

where the frequencies \( \omega_0 = \frac{\partial H_0}{\partial p} \). We can solve the homologic equation with \( d_0 = 0 \) and

\[
d_k(p) = \frac{-ic_k(p)}{k \cdot \omega_0(p)}
\]

and with

\[
K_1 = c_0(p)
\]

We see the small divisor problem again in the values of \( d_k \).

**Remark** xxx Ferraz-Mello in his book on Canonical perturbation theory, has a nice comparison of perturbation theory using Lie-derivatives (also known as the Hori method?) with perturbation theory using standard generating functions. Could be illustrated here in the future.

### 3.3 Ways out of the small divisor problem and convergence

**Using the analyticity of the function:**

When taking a Fourier transform of \( H_1 \), the analyticity of the function ensures that the coefficients decay exponentially above some \(|k|\). In other words as long as the function is smooth past some small scale size, the coefficients will decay exponentially above a \( k \) related to that scale size.

**Restricting by a Diophantine condition**

By choosing an initial condition, \( p_* \) that satisfies a Diophantine condition

\[
|k \cdot \omega(p_*)| > \frac{\gamma}{|k|^\alpha} \quad \text{for every} \quad k \in \mathbb{Z}^n, k \neq 0
\]
we can ensure that the divisors are always smaller than something. A trick is to restrict
the setting to an extremely small neighborhood of $p_*$ and with size dependent on $\epsilon$.

By combining the exponential drop of coefficients $c_k$ with the Diophantine condition,
we can ensure that the procedure, when repeated, converges.

**Remark** The illustration that the perturbation expansion converges in this setting is part
of what is known as KAM theory.