Lectures on Dynamical Systems

Anatoly Neishtadt

Lectures for "Mathematics Access Grid Instruction and Collaboration" (MAGIC) consortium, Loughborough University, 2007

Part 3

LECTURE 14

NORMAL FORMS

Resonances near fixed points of maps

Consider a map

$$x \mapsto Ax + V(x),\ V = O(|x|^2),\ x \in \mathbb{R}^n$$

where A is a linear operator. Assume that V is analytic in some neighborhood of 0.
We use the notation: λ_j, $j = 1, 2, \ldots, n$ are eigenvalues of A (multipliers of fixed point 0), $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, $m = (m_1, m_2, \ldots, m_n)$,
$|m| = |m_1| + |m_2| + \ldots + |m_n|$, $\lambda^m = \lambda_1^{m_1} \lambda_2^{m_2} \cdots \lambda_n^{m_n}$.

Definition

The set of eigenvalues of the operator A is called *a resonant* one if a relation of the form

$$\lambda_s = \lambda^m,$$

is satisfied, where components of m are integer non-negative, $|m| \geq 2$. This relation is called *a resonance relation* or just *a resonance*. The value $|m|$ is called *an order of the resonance*.

Resonant monomials, resonant normal form near fixed point

Assume that eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the operator A are all different. So, the the eigenvectors $e_1, e_2, \ldots, e_n$ of the complexified operator A form a basis in $\mathbb{C}^n$.

Let some system $\mathcal{S}$ of resonance relations be given. We will assume that $\mathcal{S}$ contains all resonance relations which can be derived from any subsystem of $\mathcal{S}$.

Definition

A vector monomial $x^m e_s$ is called a *resonant* one for resonances in $\mathcal{S}$ if the resonance relation $\lambda_s = \lambda^m$ is presented in the system $\mathcal{S}$.

Definition

A map

$$x \mapsto Ax + \ldots$$

is said to be in the resonant normal form for resonances from $\mathcal{S}$ if the nonlinear part of its right hand side is a sum of resonant vector monomials.

Reduction to resonant normal form near fixed point

Theorem

If multipliers of a fixed point do not satisfy resonance relations up to an order N inclusively except, may be, resonances from $\mathcal{S}$, then by a polynomial real close to the identical transformation of variables

$$x = y + O(|y|^2)$$

the system is reducible to the form

$$y \mapsto Ay + w(y) + O(|y|^{N+1})$$

were w is a sum of resonant vector monomials of degrees not exceeding N.

Thus, the system without the term $O(|y|^{N+1})$ (also called *a truncated system*) is in a resonant normal form.

Corollary

If there are no resonances of any order, except, may be, resonances from $\mathcal{S}$, then a formal transformation of variables reduces the original system to a system in a formal resonant normal form.

Reduction to resonant normal form near fixed point

Procedure of reduction to resonant normal form near fixed point is analogous to that near an equilibrium
The map under consideration has the form

$$x \mapsto Ax + V(x),\ \ V(x) = v_2(x) + v_3(x) + \ldots + v_N(x) + O(|x|^{N+1})$$

where $v_r(x)$ is the homogeneous vector polynomial of x of degree r.
We are looking for a transformation of variables $x \mapsto y$ of the form

$$x = y + h(y),\ \ h(y) = h_2(y) + h_3(y) + \ldots + h_N(y)$$

which reduces the map to the form

$$y \mapsto Ay + w(y) + O(|y|^{N+1}),\ \ w(y) = w_2(y) + w_3(y) + \ldots + w_N(y)$$

where $h_r(y), w_r(y)$ are homogeneous vector polynomials of y of degree r, and $w_r(y)$ contains only resonant monomials.
Plugging the transformation of variables into original map, assuming that the transformed map has required form and equating terms of order r we get *a homological equation*

$$h_r(Ay) - Ah_r(y) = V_r(y) - w_r(y)$$

where V_r is the homogeneous vector polynomial of degree r whose coefficients are expressed through coefficients of $v_2, \ldots, v_r,\ h_2, \ldots, h_{r-1},\ w_2, \ldots, w_{r-1}$.
Take as $w_r(y, t)$ the sum of resonant monomials in $V_r(y, t)$.

Reduction to resonant normal form near near fixed point, continued

Lemma

For this choice of w_r the homological equation has a solution h_r in the form of a sum of non-resonant monomials. The solution in such form is a unique.

Proof.

Let $e_1, e_2, \ldots, e_n$ be eigenvectors of the complexified operator A, that correspond to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. The eigenvalues of A are all different, and so the eigenvectors form a basis in $\mathbb{C}^n$. Let $y_1, y_2, \ldots, y_n$ be coordinates of y in this basis. Denote $U_r(y) = V_r(y) - w_r(y)$. Then

$$U_r = \sum_{s=1,\ldots,n;\ |m|=r} U_{s,m} y^m e_s, \quad h = \sum_{s=1,\ldots,n;\ |m|=r} h_{s,m} y^m e_s$$

Equating in the homological equation the coefficients in front of $y^m e_s$, we get

$$(\lambda^m - \lambda_s) h_{s,m} = U_{s,m}$$

Thus, $h_{s,m} = U_{s,m}/(\lambda^m - \lambda_s)$. If y is real, then $h(y)$ is real. This completes the proof. $\square$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$.

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) =$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) = (-1+\delta)((-1+\delta)y+ay^3)+$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) = (-1+\delta)((-1+\delta)y+ay^3)+a((-1+\delta)y+ay^3)^3 =$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) = (-1+\delta)((-1+\delta)y+ay^3)+a((-1+\delta)y+ay^3)^3 =$
$(1 - 2\delta + \delta^2)y +$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) = (-1+\delta)((-1+\delta)y+ay^3)+a((-1+\delta)y+ay^3)^3 =$
$(1 - 2\delta + \delta^2)y + a(-2 + 2\delta)y^3 + O(y^5)$

Example: normal form for period-doubling bifurcation

The period-doubling bifurcation is a local bifurcation which takes place in generic ODEs and maps when a periodic trajectory (or a fixed point, for a map) loses stability as a real multiplier crosses the unit circle in the complex plane in the point -1 . For maps it is called also *the flip* bifurcation.

Consider a one-dimensional map ($n = 1$). General case reduces to this one by means of Shoshitaisvili reducton principle. Assume that $\lambda = -1 + \delta$, where δ is small.

The resonance relation $\lambda = \lambda^m$ for $\lambda = -1$ gives $m = 2k + 1,\ k = 1, 2, \ldots$.
Therefore, the map can be transformed to the form

$$y \mapsto (-1 + \delta)y + ay^3 + O(y^5)$$

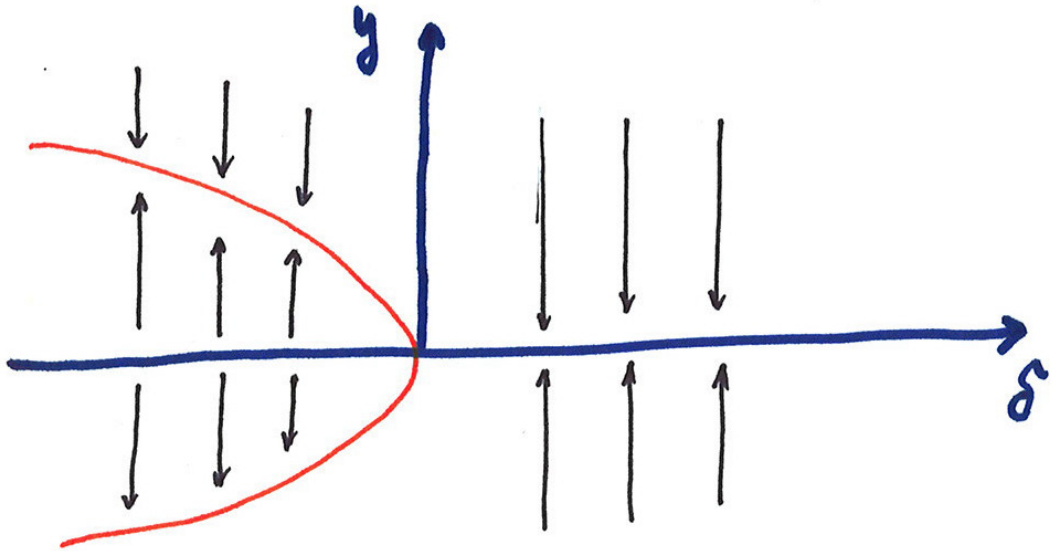
Denote M the truncated map, $M(y) = (-1 + \delta)y + ay^3$. Consider square of this map

$M^2(y) = M((-1+\delta)y+ay^3) = (-1+\delta)((-1+\delta)y+ay^3)+a((-1+\delta)y+ay^3)^3 = (1 - 2\delta + \delta^2)y + a(-2 + 2\delta)y^3 + O(y^5)$

We will consider generic case: $a \neq 0$.
The bifurcation diagram is the same as for the map $\tilde{M}^2 : y \to (1 - 2\delta)y - 2ay^3$.
Map $\tilde{M}^2$ has the fixed point at $y = 0$ and, when δ has sign opposite to that of a, two fixed points at $y = \pm\sqrt{-\delta/a}$. These two fixed points of the map $\tilde{M}^2$ correspond to the periodic trajectory of period 2 of the map M.

Example: normal form for period-doubling bifurcation, continued

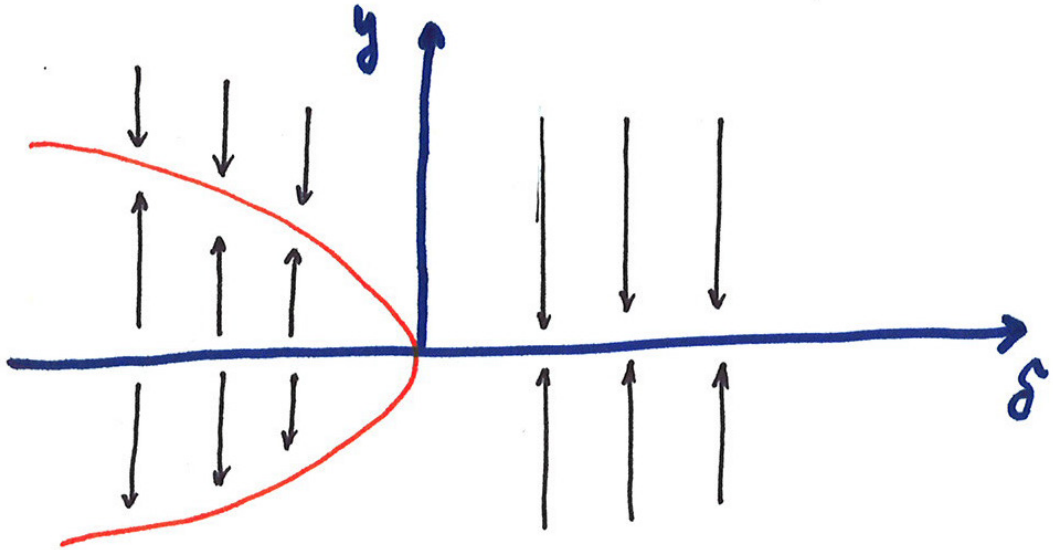
For $a > 0$ the phase portrait in the plane δ, y looks as follows:



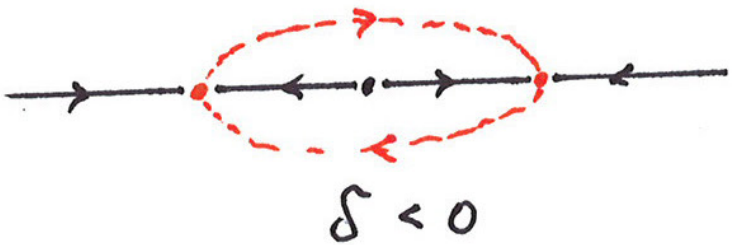
This is a supercritical case.

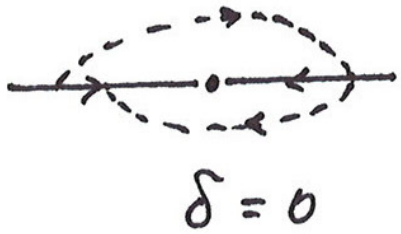
Example: normal form for period-doubling bifurcation, continued

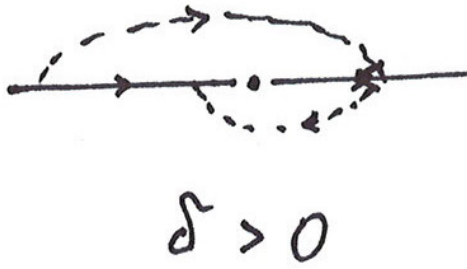
For $a > 0$ the phase portrait in the plane δ, y looks as follows:



This is a supercritical case. The bifurcation diagram looks as follows

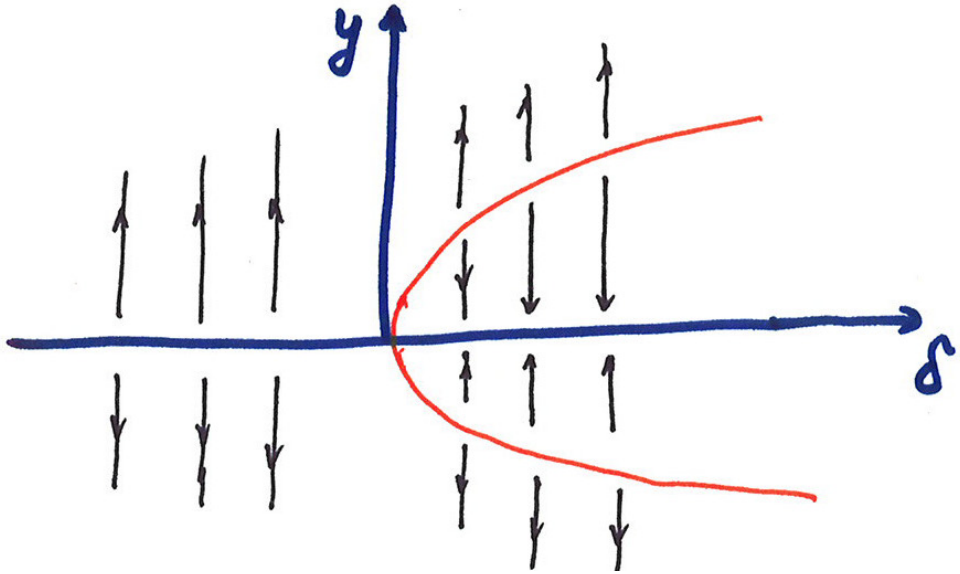






Example: normal form for period-doubling bifurcation, continued

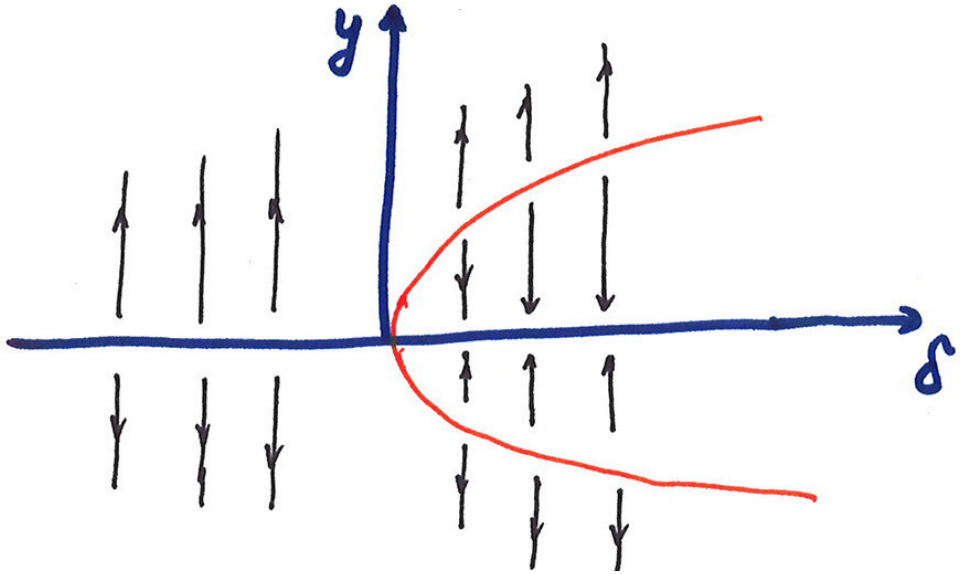
For $a < 0$ the phase portrait in the plane δ, y looks as follows:



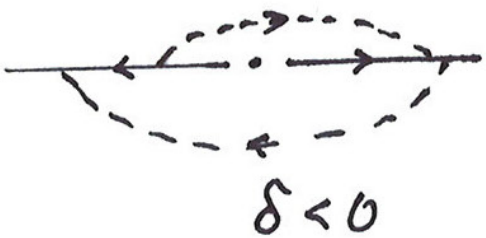
This is a subcritical case.

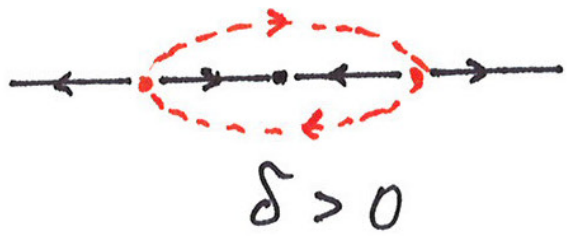
Example: normal form for period-doubling bifurcation, continued

For $a < 0$ the phase portrait in the plane δ, y looks as follows:



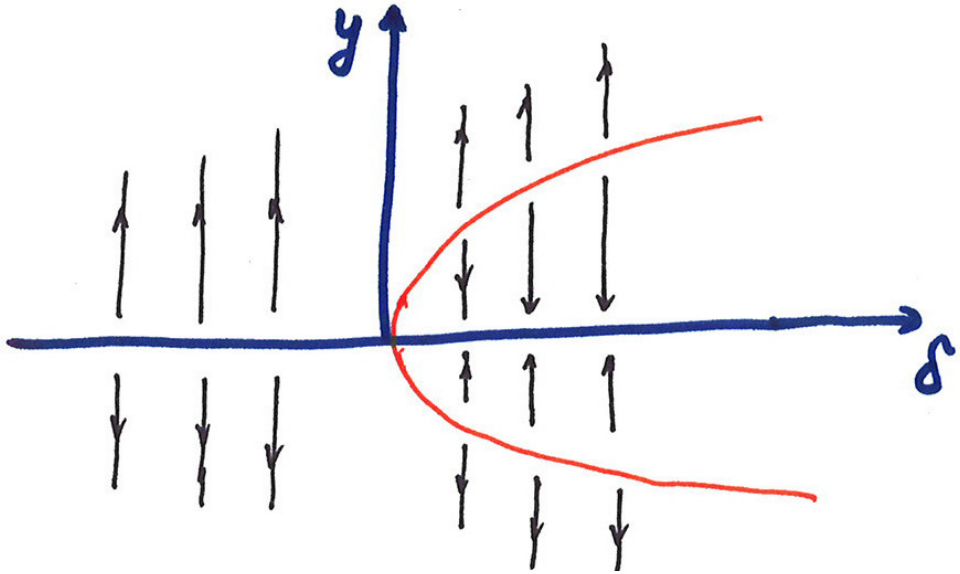
This is a subcritical case.
The bifurcation diagram looks as follows



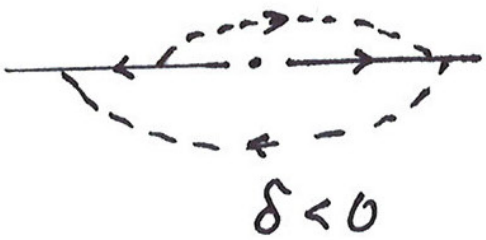


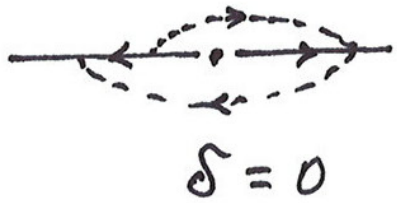
Example: normal form for period-doubling bifurcation, continued

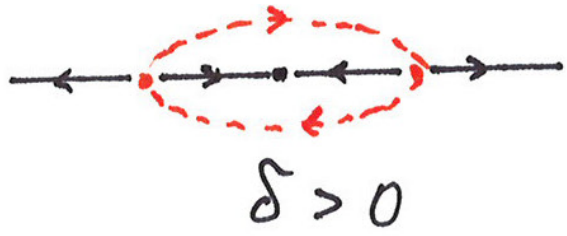
For $a < 0$ the phase portrait in the plane δ, y looks as follows:



This is a subcritical case.
The bifurcation diagram looks as follows



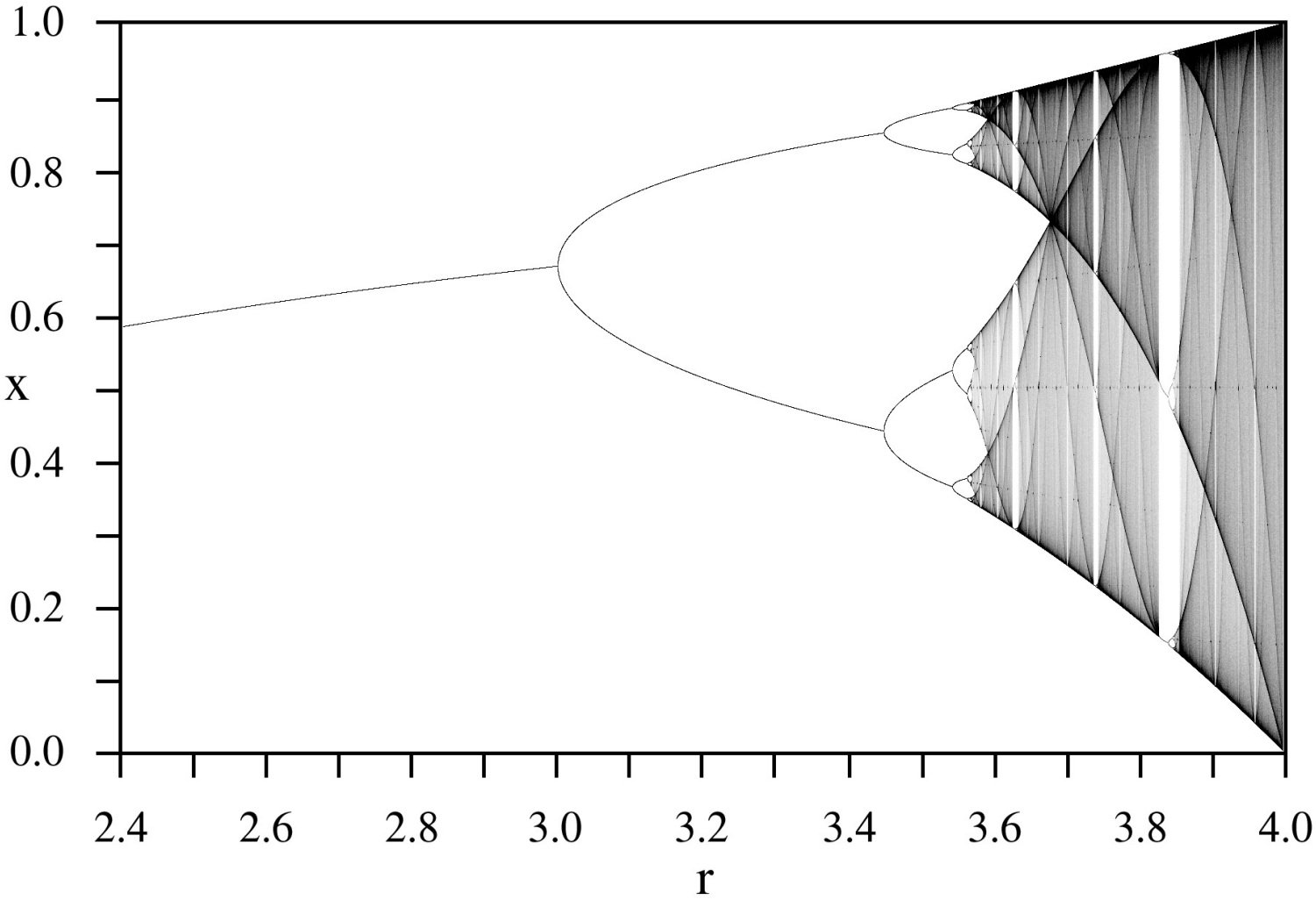




These diagrams describe also period-doubling bifurcation for periodic trajectories of ODEs.

Example: picture of period-doubling cascade

Here is picture of period doubling cascade for logistic map $x \mapsto rx(1-x)$ (source: Wikipedia). Point of accumulation of period doublings is $r \approx 3.57$.



About explanation of Feigenbaum's universality

The detailed explanation of existence of period-doubling cascade and the Feigenbaum universality was given first for a certain class $\mathcal{M}$ of non-invertible one-dimensional maps.

About explanation of Feigenbaum's universality

The detailed explanation of existence of period-doubling cascade and the Feigenbaum universality was given first for a certain class $\mathcal{M}$ of non-invertible one-dimensional maps.

M. Feigenbaum have introduced nonlinear non-invertible operator $J\colon \mathcal{M} \to \mathcal{M}$ with the following property: Jf is topologically equivalent to $f \circ f$ for any $f \in \mathcal{M}$. So, if Jf has a periodic orbit of a period N, then f has a periodic orbit of the period $2N$. This operator is called *the doubling operator*. This operator perform a *renormalisation* in $\mathcal{M}$.

About explanation of Feigenbaum's universality

The detailed explanation of existence of period-doubling cascade and the Feigenbaum universality was given first for a certain class $\mathcal{M}$ of non-invertible one-dimensional maps.

M. Feigenbaum have introduced nonlinear non-invertible operator $J\colon \mathcal{M} \to \mathcal{M}$ with the following property: Jf is topologically equivalent to $f \circ f$ for any $f \in \mathcal{M}$. So, if Jf has a periodic orbit of a period N, then f has a periodic orbit of the period $2N$. This operator is called *the doubling operator*. This operator perform a *renormalisation* in $\mathcal{M}$.

It turned out that this operator has a fixed point f_*. All multipliers of this fixed point are situated inside the unit circle but one which is real and bigger than 1; the value of this multiplier is the Feigenbaum constant μ_F.

About explanation of Feigenbaum's universality

The detailed explanation of existence of period-doubling cascade and the Feigenbaum universality was given first for a certain class $\mathcal{M}$ of non-invertible one-dimensional maps.

M. Feigenbaum have introduced nonlinear non-invertible operator $J\colon \mathcal{M} \to \mathcal{M}$ with the following property: Jf is topologically equivalent to $f \circ f$ for any $f \in \mathcal{M}$. So, if Jf has a periodic orbit of a period N, then f has a periodic orbit of the period $2N$. This operator is called *the doubling operator*. This operator perform a *renormalisation* in $\mathcal{M}$.

It turned out that this operator has a fixed point f_*. All multipliers of this fixed point are situated inside the unit circle but one which is real and bigger than 1; the value of this multiplier is the Feigenbaum constant μ_F.

Denote W^u and W^s the unstable and stable manifold of the fixed point f_*. Near f_* the action of J on W^u is approximately as follows:
$\mathcal{M}f = \mu_F f,\ f \in W^u$. Near W^u the operator $\mathcal{M}$ is invertible.

About explanation of Feigenbaum's universality

The detailed explanation of existence of period-doubling cascade and the Feigenbaum universality was given first for a certain class $\mathcal{M}$ of non-invertible one-dimensional maps.

M. Feigenbaum have introduced nonlinear non-invertible operator $J\colon \mathcal{M} \to \mathcal{M}$ with the following property: Jf is topologically equivalent to $f \circ f$ for any $f \in \mathcal{M}$. So, if Jf has a periodic orbit of a period N, then f has a periodic orbit of the period $2N$. This operator is called *the doubling operator*. This operator perform a *renormalisation* in $\mathcal{M}$.

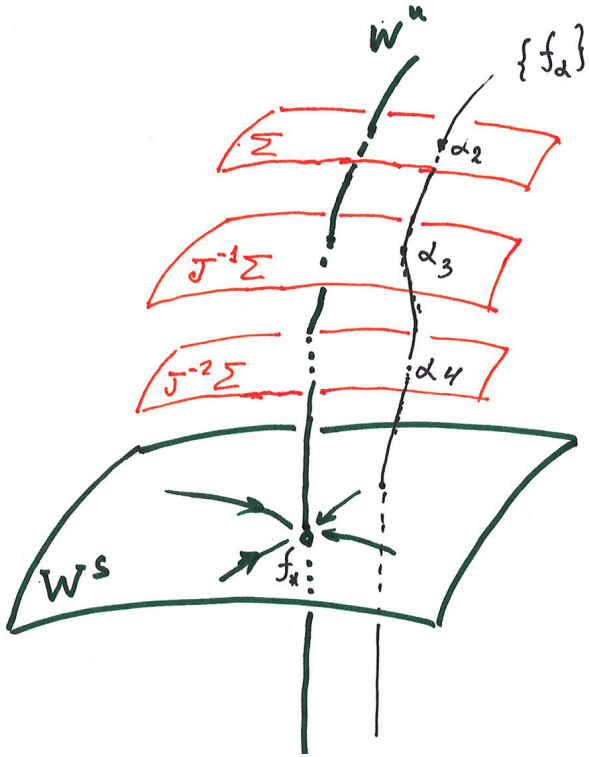
It turned out that this operator has a fixed point f_*. All multipliers of this fixed point are situated inside the unit circle but one which is real and bigger than 1; the value of this multiplier is the Feigenbaum constant μ_F.

Denote W^u and W^s the unstable and stable manifold of the fixed point f_*. Near f_* the action of J on W^u is approximately as follows: $\mathcal{M}f = \mu_F f,\ f \in W^u$. Near W^u the operator $\mathcal{M}$ is invertible.

On the other hand, in $\mathcal{M}$ there is a surface Σ of maps having a fixed point with multiplier (-1). This is a surface of codimension 1. It turns out that W^u transversally crosses Σ. The surfaces $J^{-k}\Sigma$ (preimages of Σ), accumulate on W^s as $k \to \infty$. The surface $J^{-k}\Sigma$ is composed of maps having a periodic orbit of period 2^k with multiplier (-1). Lengths of segments between successive intersections of W^u and $J^{-k}\Sigma$ decay in geometric progression with the common ratio μ_F as $k \to \infty$.

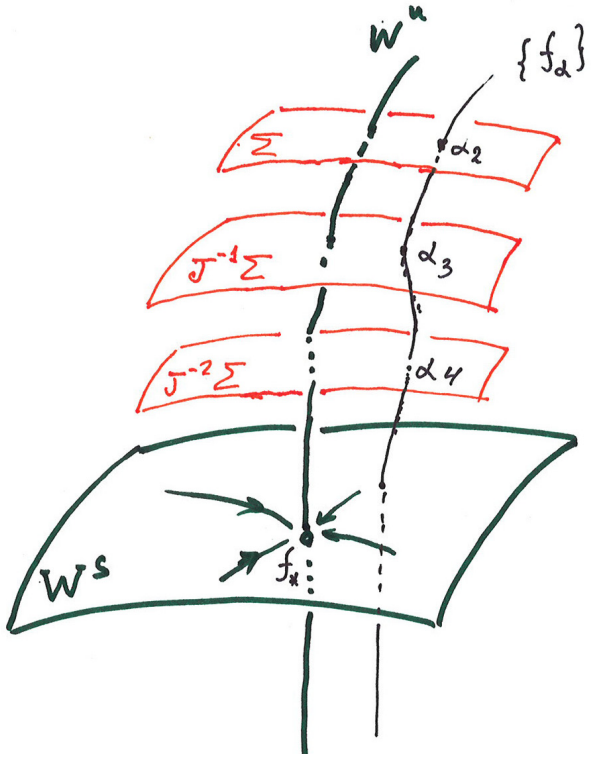
About explanation of Feigenbaum's universality, continued

The structure of $\mathcal{M}$ in a neighborhood of f_* looks as follows:



About explanation of Feigenbaum's universality, continued

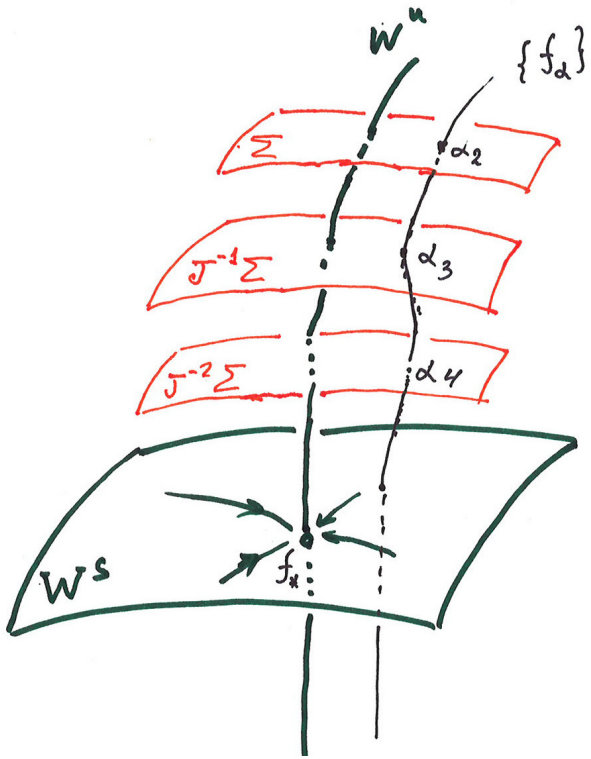
The structure of $\mathcal{M}$ in a neighborhood of f_* looks as follows:



Now consider a one parametric family of maps $f_\alpha \in \mathcal{M}$, $\alpha \in \mathbb{R}$. This family forms a curve in $\mathcal{M}$. Assume that this curve crosses W^s not far from f_*. Then it should cross all $J^{-k}\Sigma$ with big enough k. This implies existence of the infinite cascade of period doublings.

About explanation of Feigenbaum's universality, continued

The structure of $\mathcal{M}$ in a neighborhood of f_* looks as follows:



Now consider a one parametric family of maps $f_\alpha \in \mathcal{M}$, $\alpha \in \mathbb{R}$. This family forms a curve in $\mathcal{M}$. Assume that this curve crosses W^s not far from f_*. Then it should cross all $J^{-k}\Sigma$ with big enough k. This implies existence of the infinite cascade of period doublings.

Lengths of segments between successive intersections of this curve and $J^{-k}\Sigma$ decay in geometric progression with the common ratio $1/\mu_F$ as $k \to \infty$. This implies the universality.

LECTURE 15

NORMAL FORMS

C^∞ and finitely smooth normal forms

Consider an ODE

$$\dot{x} = Ax + O(|x|^2),\ x \in \mathbb{R}^n$$

where A is a linear operator. Assume that the right hand side of this ODE is of smoothness C^∞ in some neighborhood of 0. Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ be the vector of eigenvalues of A.
In this theory it is supposed that the equilibrium is *hyperbolic* (all eigenvalues have non-zero real parts).

Theorem (S.Sternberg)

If the equilibrium is hyperbolic and non-resonant, then by a real close to the identical C^∞ smooth transformation of variables

$$x = y + O(|y|^2)$$

the system is reducible to the form

$$\dot{y} = Ay$$

C^∞ and finitely smooth normal forms, continued

Let $\mathcal{S}$ be some system of resonance relations, i.e. relations of the form

$$\lambda_s = m_1\lambda_1 + m_2\lambda_2 + \ldots + m_n\lambda_n,$$

with integer non-negative $m_1, m_2, \ldots, m_n$ such that $\sum_{j=1}^{n} m_j \geq 2$.

Theorem (K.T.Chen)

If the equilibrium is hyperbolic and its eigenvalues do not satisfy resonance relations except, may be, resonances from $\mathcal{S}$, then by a real close to the identical C^∞ smooth transformation of variables

$$x = y + O(|y|^2)$$

the system is reducible to the form

$$\dot{y} = Ay + w(y)$$

were formal Tailor series for w is in the formal normal form for resonances from $\mathcal{S}$.

C^∞ and finitely smooth normal forms, continued

Consider an ODE

$$\dot{x} = Ax + O(|x|^2),\ x \in \mathbb{R}^n$$

where A is a linear operator. Assume that the right hand side of this ODE is of smoothness $\mathbb{C}^\infty$ in some neighborhood of 0.

Theorem (S.Sternberg)

If the equilibrium is hyperbolic, then for any natural k there exists natural $N = N(A, k)$ such that if the eigenvalues of the equilibrium do not satisfy any resonance relations up to order N inclusively, then by a real close to the identical C^k smooth transformation of variables

$$x = y + O(|y|^2)$$

the system is reducible to the form

$$\dot{y} = Ay$$

Remark

The same result is valid for systems of smoothness C^r, $r > N(A, k)$.

C^∞ and finitely smooth normal forms, continued

Let $\mathcal{S}$ be some system of resonance relations.

Theorem (G.R.Belitskii-V.S.Samovol)

If the equilibrium is hyperbolic, then for any natural k there exists natural $N = N(A, k)$ such that if the eigenvalues of the equilibrium do not satisfy any resonance relations up to order N inclusively, except, may be, resonances from $\mathcal{S}$, then by a real close to the identical C^k smooth transformation of variables

$$x = y + O(|y|^2)$$

the system is reducible to the form

$$\dot{y} = Ay + w(y)$$

were w is a vector polynomial of degree not exceeding N having normal form for resonances from $\mathcal{S}$.

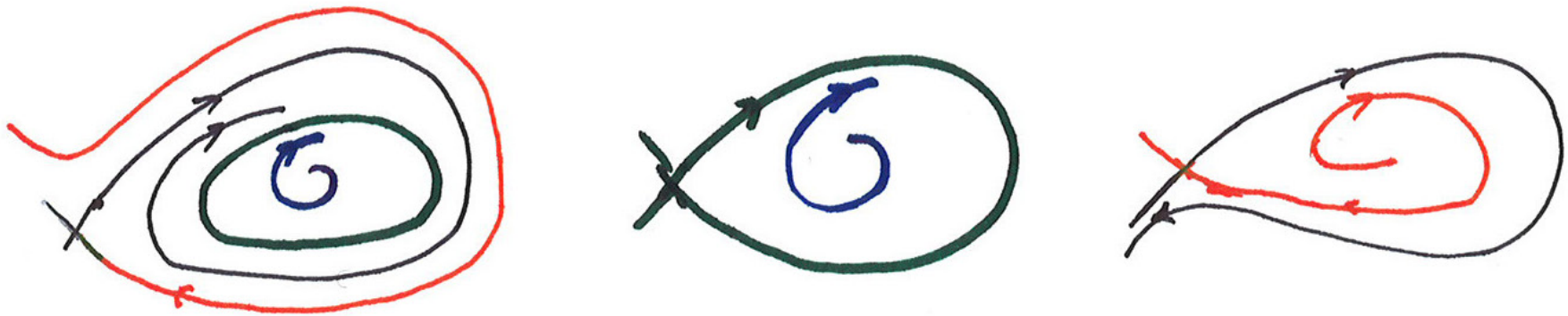
Remark

The same result is valid for systems of smoothness C^r, $r > N(A, k)$.

SOME NONLOCAL BIFURCATIONS

Birth of cycle at homoclinic loop of a saddle in planar system

The bifurcation under consideration looks like this:

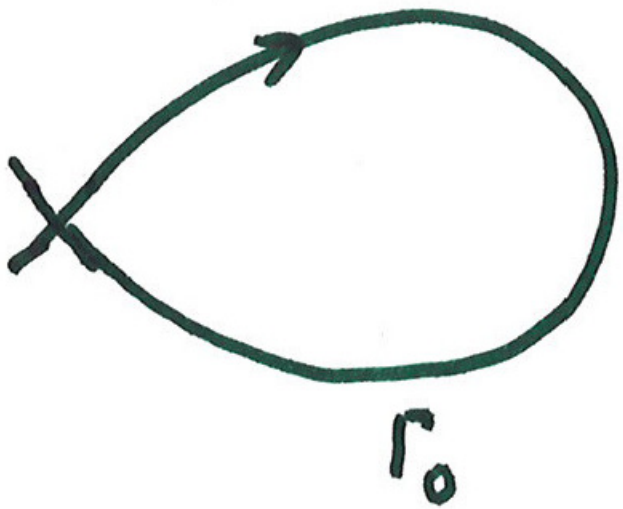


Birth of cycle at homoclinic loop of a saddle in planar system, continued

Consider a two-dimensional system

$$\dot{x} = v(x, \alpha), \ x \in \mathbb{R}^2, \ \alpha \in \mathbb{R}^1$$

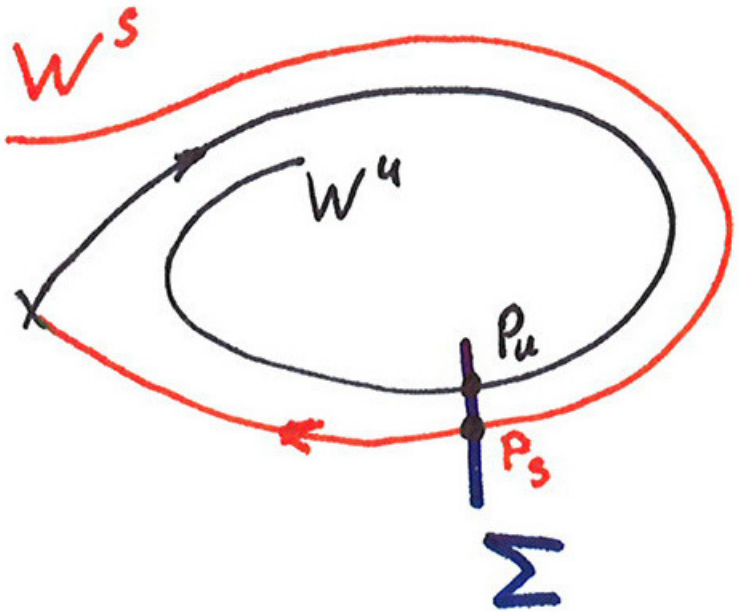
right hand side of the system is smooth enough.
Let for $\alpha = 0$ the the system has a saddle equilibrium at $x = 0$ and homoclinic loop Γ_0.



Then for all small α the system has the saddle equilibrium near $x = 0$. Without loss of generality we may assume that this equilibrium is at $x = 0$.
Let $\lambda_1(\alpha)$, $\lambda_2(\alpha)$ be eigenvalues of this equilibrium, $\lambda_1(\alpha) < 0 < \lambda_2(\alpha)$. Their sum $\sigma(\alpha) = \lambda_1(\alpha) + \lambda_2(\alpha)$ is called *the saddle quantity*.

Birth of cycle at homoclinic loop of a saddle in planar system, split function

Let Σ be an oriented segment transversal to Γ_0.
Denote $W^s = W^s(\alpha)$, $W^u = W^u(\alpha)$ the stable and the unstable manifolds of our saddle equilibrium (they exist according to Hadamard-Perron theorem).
Denote $p_s(\alpha)$, $p_u(\alpha)$ the points of the first intersection of Σ with W^s and W^u respectively.



Denote $\rho(\alpha)$ the oriented distance between $p_u(\alpha)$ and $p_s(\alpha)$. The picture corresponds to the case $\rho(\alpha) > 0$. The function $\rho(\cdot)$ is called *the split function*.

Birth of cycle at homoclinic loop of a saddle in planar system, A.A.Andronov and E.A.Leontovich theorem

Theorem (A.A.Andronov and E.A.Leontovich)

Consider a two-dimensional system

$$\dot{x} = v(x, \alpha),\ x \in \mathbb{R}^2,\ \alpha \in \mathbb{R}^1$$

Let for $\alpha = 0$ *the the system has a saddle equilibrium at* $x = 0$ *with eigenvalues* $\lambda_1(0) < 0 < \lambda_2(0)$ *and homoclinic loop* Γ_0. *Assume that the following genericity conditions hold:*
a) $\sigma(0) = \lambda_1(0) + \lambda_2(0) \neq 0$, *and*
b) $d\rho(0)/d\alpha \neq 0$, *where* $\rho(\cdot)$ *is the previously defined split function.*
Then, for all sufficiently small $|\alpha|$, *there exists a neighborhood of* Γ_0 *in which a unique limit cycle bifurcates from* Γ_0. *Moreover, the cycle is stable and exists for* $\rho(\alpha) > 0$ *if* $\sigma(0) < 0$, *and is unstable and exists for* $\rho(\alpha) < 0$ *if* $\sigma(0) > 0$.

Birth of cycle at homoclinic loop of a saddle in planar system, A.A.Andronov and E.A.Leontovich theorem

Theorem (A.A.Andronov and E.A.Leontovich)

Consider a two-dimensional system

$$\dot{x} = v(x, \alpha),\ x \in \mathbb{R}^2,\ \alpha \in \mathbb{R}^1$$

Let for $\alpha = 0$ *the the system has a saddle equilibrium at* $x = 0$ *with eigenvalues* $\lambda_1(0) < 0 < \lambda_2(0)$ *and homoclinic loop* Γ_0. *Assume that the following genericity conditions hold:*
a) $\sigma(0) = \lambda_1(0) + \lambda_2(0) \neq 0$, *and*
b) $d\rho(0)/d\alpha \neq 0$, *where* $\rho(\cdot)$ *is the previously defined split function.*
Then, for all sufficiently small $|\alpha|$, *there exists a neighborhood of* Γ_0 *in which a unique limit cycle bifurcates from* Γ_0. *Moreover, the cycle is stable and exists for* $\rho(\alpha) > 0$ *if* $\sigma(0) < 0$, *and is unstable and exists for* $\rho(\alpha) < 0$ *if* $\sigma(0) > 0$.

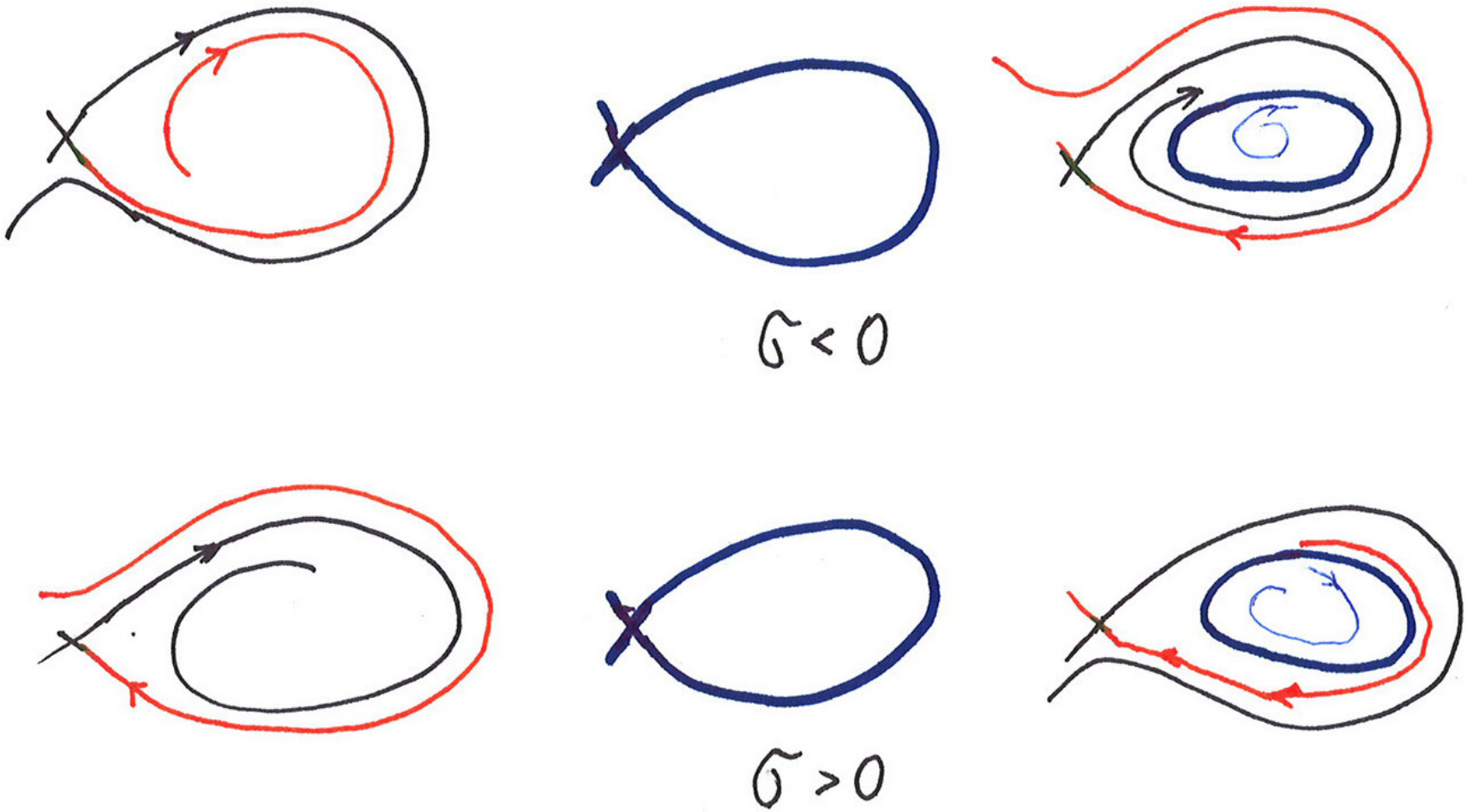
Remark

The condition $d\rho(0)/d\alpha \neq 0$ is equivalent to the condition

$$\int_{-\infty}^{+\infty} \exp\left[-\int_0^t \operatorname{div} v\, dt\right] \left(v_1 \frac{\partial v_2}{\partial \alpha} - v_2 \frac{\partial v_1}{\partial \alpha}\right) dt \neq 0$$

where all expressions involving $v = (v_1, v_2)$ are evaluated at $\alpha = 0$ along a solution corresponding to the homoclinic orbit Γ_0.

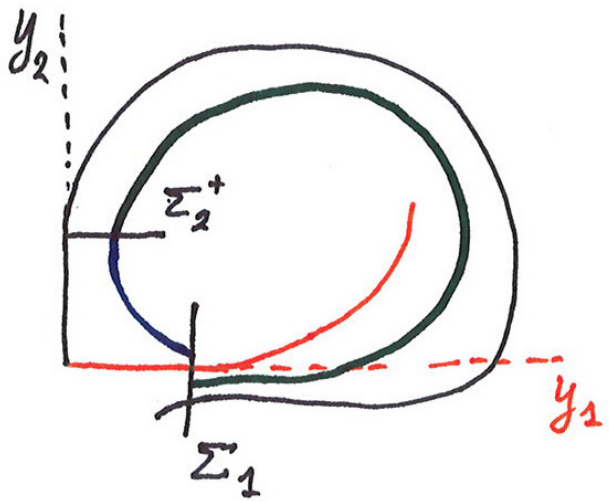
Birth of cycle at homoclinic loop of a saddle in planar system, A.A.Andronov and E.A.Leontovich theorem, continued



Scheme of proof of A.A.Andronov and E.A.Leontovich theorem

To find a limit cycle we will look for a fixed point of Poincaré return map. This map is the composition of two maps: *singular map*, which correspond to pass of trajectory near the saddle point, and *regular map* which correspond to remaining part of trajectory.

Near the saddle one can make smooth transformation of variables $x \mapsto y$ such that W^u and W^s will have equations $y_1 = 0$ and $y_2 = 0$ respectively.

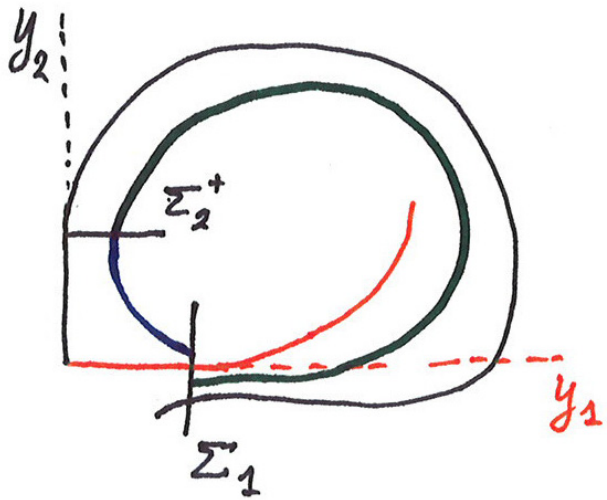


The equations in the new variables have the form

$$\dot{y}_1 = (\lambda_1 + O(|y_1| + |y_2|))y_1, \ \dot{y}_2 = (\lambda_2 + O(|y_1| + |y_2|))y_2$$

Scheme of proof of A.A.Andronov and E.A.Leontovich theorem, continued

Near the origin the behavior of trajectories of our system is close to behavior of a linear system The phase flow determines the map of a segment $\Sigma_1^+ = \{y_1 = \delta_1, 0 < y_2 < \kappa_1\}$ to a segment $\Sigma_2^+ = \{y_2 = \delta_2, 0 < y_1 < \kappa_2\}$. Here $\delta_{1,2}$, $\kappa_{1,2}$ are small enough positive constants. Denote corresponding map as Q_{sing}. Approximate formula for Q_{sing} is the formula for analogous map in the linearised system, $y_2 \mapsto y_1 = cy_2^{-\lambda_1/\lambda_2}$, $c = \mathrm{const} > 0$.



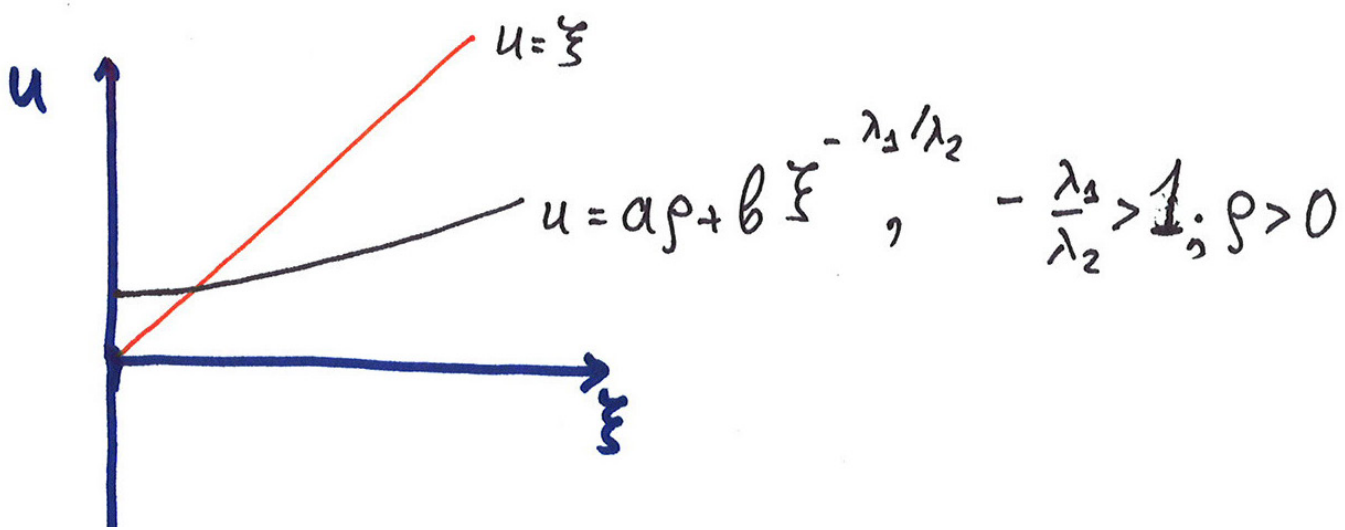
The phase flow determines also the map of Σ_2^+ to $\Sigma_1 = \{y_1 = \delta_1, |y_2| < \kappa_1\}$. Denote corresponding map as Q_{reg}. Approximate formula for Q_{reg} is $y_1 \mapsto y_2 = a\rho + dy_1$, where $a > 0$ and $d > 0$ are constants. So, the composition $Q = Q_{reg} \circ Q_{sing}$ maps Σ_1^+ to Σ_1, and the approximate formula for Q is $y_2 \mapsto a\rho + by_2^{-\lambda_1/\lambda_2}$, where a and b are positive constants.

Scheme of proof of A.A.Andronov and E.A.Leontovich theorem, continued

The approximate equation for fixed point of the map Q is

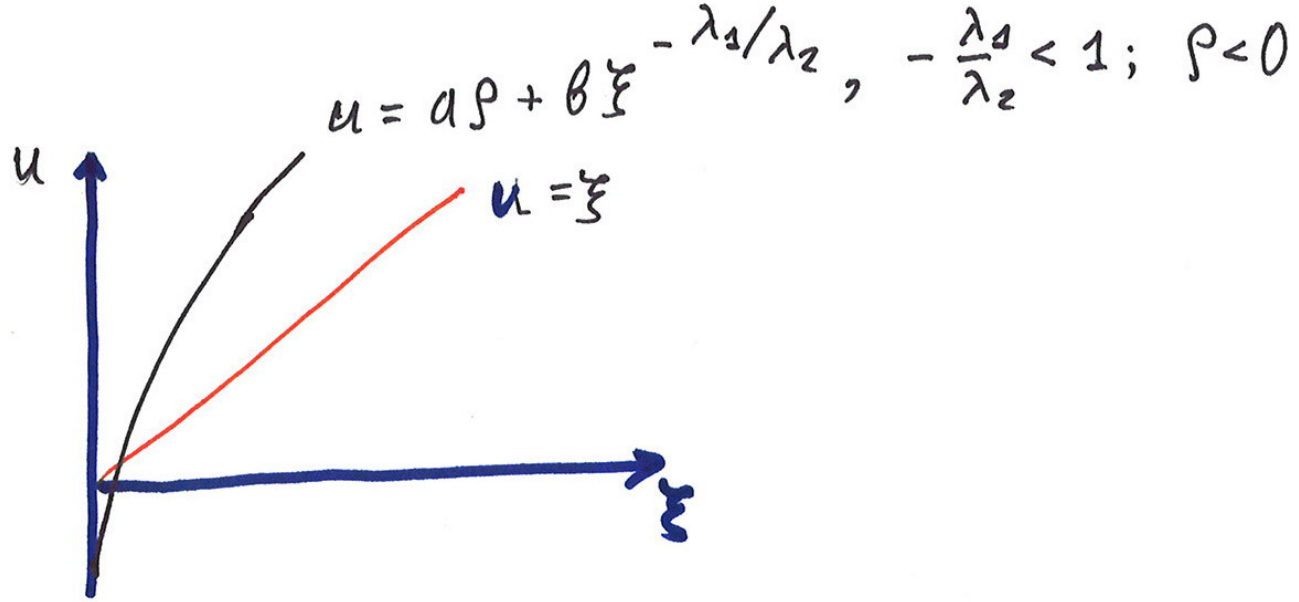
$$\xi = a\rho + b\xi^{-\lambda_1/\lambda_2},\ \xi > 0$$

If $\sigma(0) = \lambda_1(0) + \lambda_2(0) < 0$, then for small enough α we have $-\lambda_1/\lambda_2 > 1$. Then for small $\rho > 0$ the equation for fixed point has a unique solution in the domain of small ξ. In this case there are no such solutions for small $\rho < 0$. So, we have a unique limit cycle if α is small and $\rho(\alpha) > 0$. This cycle is stable.



Scheme of proof of A.A.Andronov and E.A.Leontovich theorem, continued

If $\sigma(0) = \lambda_1(0) + \lambda_2(0) > 0$, then for small enough α we have $-\lambda_1/\lambda_2 < 1$. Then for small $\rho < 0$ the equation for fixed point has a unique solution in the domain of small ξ. In this case there are no such solutions for small $\rho > 0$. So, we have a unique limit cycle if α is small and $\rho(\alpha) < 0$. This cycle is unstable.

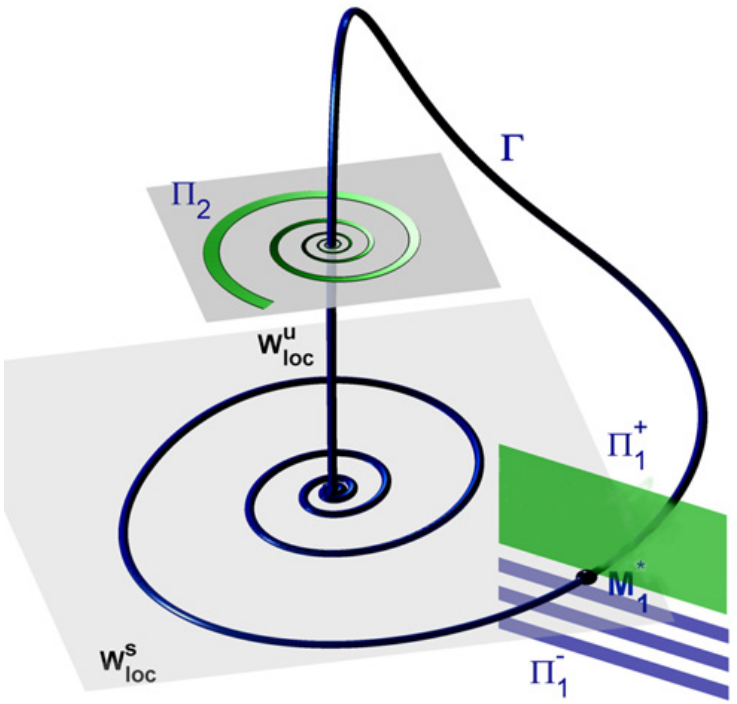


Shilnikov chaos

Consider a three-dimensional system

$$\dot{x} = v(x, \alpha), \ x \in \mathbb{R}^3, \ \alpha \in \mathbb{R}^1$$

right hand side of the system is smooth enough.
Let for $\alpha = 0$ the the system has a *saddle-focus equilibrium* at $x = 0$, i.e. for this equilibrium one eigenvalue $\lambda_1(0)$ is real and two eigenvalues $\lambda_{2,3}(0)$ are complex conjugate, $\lambda_{2,3}(0) = \gamma \pm i\omega$. Let for $\alpha = 0$ the the system has homoclinic trajectory Γ_0 (picture from the paper of L.P.Shilnikov and A.L.Shilnikov in Scholarpedia):



Shilnikov chaos, continued

Theorem (L.P.Shilnikov)

Suppose that a three-dimensional system

$$\dot{x} = v(x, \alpha),\ x \in \mathbb{R}^3,\ \alpha \in \mathbb{R}^1$$

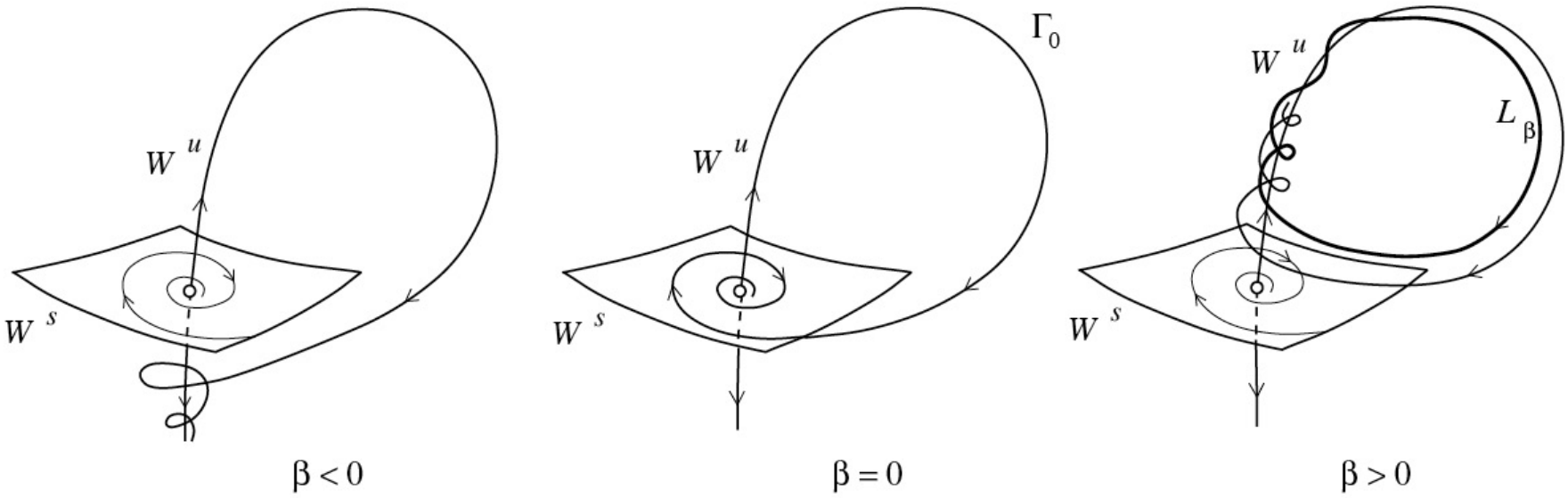
has at $\alpha = 0$ *a saddle-focus equilibrium point* $x = 0$ *with eigenvalues* $\lambda_1(0) > 0 > \mathrm{Re}\lambda_{2,3}(0)$ *and a homoclinic trajectory* Γ_0. *Assume that the following genericity conditions hold:*
a) $\sigma(0) = \lambda_1(0) + \mathrm{Re}\lambda_{2,3}(0) > 0$, *and*
b) $\mathrm{Im}\lambda_{2,3}(0) \neq 0$.
Then, the system has an infinite number of saddle periodic trajectories in a neighborhood of Γ_0 *for all sufficiently small* $|\alpha|$.

Remark

For saddle-focus equilibria in 3D the sum of the real eigenvalue and the real part of the complex eigenvalue is called *the saddle quantity*.

Shilnikov chaos, continued

Assume that instead of condition a) in the Shilnikov theorem we have $\sigma(0) = \lambda_1(0) + \mathrm{Re}\lambda_{2,3}(0) < 0$. If when α changes the unstable manifold of the equilibrium passes transversally through the stable manifold, then the bifurcation diagram looks like this (L.P.Shilnikov, picture from: Yu.A.Kuznetsov, "Elements of applied bifurcation theory", $\beta = \beta(\alpha)$ is *the split function*)



For $\beta < 0$ there are no periodic trajectories near Γ_0. For $\beta > 0$ there is a unique periodic trajectory near Γ_0.

Exercises

1. Prove that the condition $d\rho(0)/d\alpha \neq 0$ is equivalent to the condition

$$\int_{-\infty}^{+\infty} \exp\left[-\int_0^t \operatorname{div} v \, dt\right] \left(v_1 \frac{\partial v_2}{\partial \alpha} - v_2 \frac{\partial v_1}{\partial \alpha}\right) dt \neq 0$$

where all expressions involving $v = (v_1, v_2)$ are evaluated at $\alpha = 0$ along a solution corresponding to the homoclinic orbit Γ_0.

LECTURE 16

NORMAL FORMS OF HAMILTONIAN SYSTEMS

Comparison with normal forms theory for general systems

What is special in the normal forms theory for Hamiltonian systems with respect to the case of general systems?

- From the viewpoint of the "general" theory equilibria and periodic trajectories in Hamiltonian systems and fixed points of symplectic maps are all resonant ones. Indeed, e.g. for equilibria, if λ_1 is an eigenvalue of an equilibrium for a Hamiltonian system, then $\lambda_2 = -\lambda_1$ is also an eigenvalue. So, infinite number of resonance relations of the form $\lambda_s = \lambda_s + k(\lambda_1 + \lambda_2)$ are satisfied.
- For a Hamiltonian system we will perform reduction to normal forms by means of canonical symplectic transformations of variables and thus will work with Hamilton's function rather than with equations of motions
- If the Hamiltonian system is in a normal form, then the order reduction can be performed easily (order reduction for the case of general systems in normal form is also possible (A.D.Bryuno), but it is essentially more complicated).

Canonical transformations

The canonical Hamiltonian system with the Hamilton function H is the ODE system of the form

$$\dot{p} = -\left(\frac{\partial H}{\partial q}\right)^T, \quad \dot{q} = \left(\frac{\partial H}{\partial p}\right)^T$$

Here $p \in \mathbb{R}^n$, $q \in \mathbb{R}^n$, p and q are considered as vector-columns, H is a function of p, q, t, the superscript "T" denotes the matrix transposition. Variables p, q are called canonical variables. Components of p are called "impulses", components of q are called "coordinates".

Symplectic (or canonical symplectic) transformation of variables with the generating function $S(P, q, t)$, depending on the new impulses and the old coordinates, is the transformation of variables which is implicitly defined by the formulas

$$p = \left(\frac{\partial S(P, q, t)}{\partial q}\right)^T, \quad Q = \left(\frac{\partial S(P, q, t)}{\partial P}\right)^T$$

Behavior of new variables P, Q is described by the Hamiltonian system with the Hamilton function

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

Example

For $(p, q) \in \mathbb{R}^2$ the transformation $(p, q) \mapsto (\rho, \varphi)$ defined by formulas $p = \sqrt{2\rho}\cos\varphi$, $q = \sqrt{2\rho}\sin\varphi$ is a symplectic transformation.

Resonances near equilibria of autonomous system

Let the origin of coordinates be an equilibrium position of an analytic autonomous Hamiltonian system with n degrees of freedom. Suppose that the eigenvalues of the quadratic part of the Hamiltonian in a neighbourhood of the equilibrium position are all distinct and purely imaginary. Then the Hamilton function can be represented in the form

$$H = \tfrac{1}{2}\Omega_1(p_1^2 + q_1^2) + \tfrac{1}{2}\Omega_2(p_2^2 + q_2^2) + \ldots + \tfrac{1}{2}\Omega_n(p_n^2 + q_n^2) + H_3 + H_4 + \ldots$$

where H_m is a form of degree m in the phase variables p, q.

Definition

The eigenfrequencies $\Omega_1, \Omega_2, \ldots, \Omega_n$ satisfy a resonance relation of order $l > 0$ if there exist integers k_i such that $k_1\Omega_1 + \ldots + k_n\Omega_n = 0$ and $|k_1| + \ldots + |k_n| = l$.

For example, $\Omega_1 = \Omega_2$ is a resonance relation of order 2.

Normal form in absence of resonances

Definition

A Birkhoff normal form of degree L for the Hamiltonian is a polynomial of degree L in canonical phase variables P, Q that is actually a polynomial of degree $[L/2]$ in the variables $\rho_i = \frac{1}{2}(P_i^2 + Q_i^2)$.

Example

For a system with two degrees of freedom,

$$H = \Omega_1\rho_1 + \Omega_2\rho_2 + \tfrac{1}{2}(\Omega_{11}\rho_1^2 + 2\Omega_{12}\rho_1\rho_2 + \Omega_{12}\rho_2^2)$$

is a Birkhoff normal form of degree 4. The terms quadratic in ρ describe the dependence of the frequencies of the oscillations on the amplitudes.

Remark

Any system in the Birkhoff normal form is completely integrable.

Normal form in absence of resonances, the Birkhoff theorem

Theorem (G.D.Birkhoff)

Suppose that the eigenfrequencies Ω_i do not satisfy any resonance relation of order L or less. Then in a neighbourhood of the equilibrium position 0 there exists a symplectic close to the identical change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position 0 and such that in the new variables the Hamilton function is reduced to a Birkhoff normal form $\mathcal{H}_L(\rho)$ of degree L up to terms of degree higher than L:

$$H(p, q) = \mathcal{H}_L(\rho) + R, \;\; R = O(|P| + |Q|)^{L+1}$$

Normal form in absence of resonances, the Birkhoff theorem

Theorem (G.D.Birkhoff)

Suppose that the eigenfrequencies Ω_i do not satisfy any resonance relation of order L or less. Then in a neighbourhood of the equilibrium position 0 there exists a symplectic close to the identical change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position 0 and such that in the new variables the Hamilton function is reduced to a Birkhoff normal form $\mathcal{H}_L(\rho)$ of degree L up to terms of degree higher than L:

$$H(p, q) = \mathcal{H}_L(\rho) + R, \;\; R = O(|P| + |Q|)^{L+1}$$

The proof of this theorem will be obtained as a particular case of the proof of the theorem about *a resonant normal form*.

Normal form in absence of resonances, the Birkhoff theorem

Theorem (G.D.Birkhoff)

Suppose that the eigenfrequencies Ω_i do not satisfy any resonance relation of order L or less. Then in a neighbourhood of the equilibrium position 0 there exists a symplectic close to the identical change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position 0 and such that in the new variables the Hamilton function is reduced to a Birkhoff normal form $\mathcal{H}_L(\rho)$ of degree L up to terms of degree higher than L:

$$H(p, q) = \mathcal{H}_L(\rho) + R, \;\; R = O(|P| + |Q|)^{L+1}$$

The proof of this theorem will be obtained as a particular case of the proof of the theorem about *a resonant normal form*.

Discarding the non-normalized terms R we obtain an integrable system whose action-angle variables are the symplectic polar coordinates ρ_i, φ_i defined by

$$P_i = \sqrt{2\rho_i} \cos \varphi_i, \;\; Q_i = \sqrt{2\rho_i} \sin \varphi_i$$

and whose trajectories wind round the tori $\rho = \mathrm{const}$ with frequencies $\partial \mathcal{H}_l / \partial \rho$.

Normal form in absence of resonances, the Birkhoff theorem

Theorem (G.D.Birkhoff)

Suppose that the eigenfrequencies Ω_i do not satisfy any resonance relation of order L or less. Then in a neighbourhood of the equilibrium position 0 there exists a symplectic close to the identical change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position 0 and such that in the new variables the Hamilton function is reduced to a Birkhoff normal form $\mathcal{H}_L(\rho)$ of degree L up to terms of degree higher than L:

$$H(p, q) = \mathcal{H}_L(\rho) + R, \ R = O(|P| + |Q|)^{L+1}$$

The proof of this theorem will be obtained as a particular case of the proof of the theorem about *a resonant normal form*.

Discarding the non-normalized terms R we obtain an integrable system whose action-angle variables are the symplectic polar coordinates ρ_i, φ_i defined by

$$P_i = \sqrt{2\rho_i} \cos \varphi_i, \ Q_i = \sqrt{2\rho_i} \sin \varphi_i$$

and whose trajectories wind round the tori $\rho = \mathrm{const}$ with frequencies $\partial \mathcal{H}_l / \partial \rho$.

Remark

Most of similar tori, which are invariant under the phase flow, in the general case exist also in the original system; this follows from the results of Kolmogorov-Arnold-Moser (KAM) theory.

Resonant normal forms

The definition of a normal form must be modified for the case where the eigenfrequencies satisfy some resonance relations. The same modification is also appropriate for nearly resonant frequencies. Let $\mathcal{K}$ be a sublattice of the integer lattice $\mathbb{Z}^n$ defining the possible resonances.

Definition

A resonant normal form of degree L for the Hamiltonian for resonances in $\mathcal{K}$ is a polynomial of degree L in canonical symplectic variables P_i, Q_i which in the symplectic polar coordinates depends on the phases φ_i only via their combinations (k, φ) for $k \in \mathcal{K}$.

Resonant normal forms

The definition of a normal form must be modified for the case where the eigenfrequencies satisfy some resonance relations. The same modification is also appropriate for nearly resonant frequencies. Let $\mathcal{K}$ be a sublattice of the integer lattice $\mathbb{Z}^n$ defining the possible resonances.

Definition

A resonant normal form of degree L for the Hamiltonian for resonances in $\mathcal{K}$ is a polynomial of degree L in canonical symplectic variables P_i, Q_i which in the symplectic polar coordinates depends on the phases φ_i only via their combinations (k, φ) for $k \in \mathcal{K}$.

Theorem (T.M.Cherry)

Suppose that the eigenfrequencies do not satisfy any resonance relations of degree L or less, except, possibly, for relations $(k, \Omega) = 0$ for $k \in \mathcal{K}$. Then in a neighbourhood of the zero equilibrium position there exists a symplectic close to the identical change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position and such that in the new variables the Hamilton function reduces to a resonant normal form of degree L for resonances in $\mathcal{K}$ up to terms of degree higher than L.

Proof of the theorem about resonant normal form

In the system with Hamiltonian H we perform the change of variables with a generating function

$$P^T q + S(P, q), \ S = S_3 + \ldots + S_L$$

The new Hamiltonian has the form

$$\mathcal{H} = \frac{1}{2}\Omega_1(P_1^2 + Q_1^2) + \ldots + \frac{1}{2}\Omega_n(P_n^2 + Q_n^2) + \mathcal{H}_3 + \mathcal{H}_4 + \ldots$$

Here S_l and $\mathcal{H}_l$ are forms of degree l in P, q and in P, Q, respectively.

Proof of the theorem about resonant normal form

In the system with Hamiltonian H we perform the change of variables with a generating function

$$P^T q + S(P, q), \ S = S_3 + \ldots + S_L$$

The new Hamiltonian has the form

$$\mathcal{H} = \tfrac{1}{2}\Omega_1(P_1^2 + Q_1^2) + \ldots + \tfrac{1}{2}\Omega_n(P_n^2 + Q_n^2) + \mathcal{H}_3 + \mathcal{H}_4 + \ldots$$

Here S_l and $\mathcal{H}_l$ are forms of degree l in P, q and in P, Q, respectively. The old and new Hamiltonians are connected by the relation

$$H\left(P + \left(\frac{\partial S(P, q, t)}{\partial q}\right)^T, q\right) = \mathcal{H}\left(P, q + \left(\frac{\partial S(P, q, t)}{\partial P}\right)^T\right)$$

Proof of the theorem about resonant normal form

In the system with Hamiltonian H we perform the change of variables with a generating function

$$P^T q + S(P, q), \ S = S_3 + \ldots + S_L$$

The new Hamiltonian has the form

$$\mathcal{H} = \tfrac{1}{2}\Omega_1(P_1^2 + Q_1^2) + \ldots + \tfrac{1}{2}\Omega_n(P_n^2 + Q_n^2) + \mathcal{H}_3 + \mathcal{H}_4 + \ldots$$

Here S_l and $\mathcal{H}_l$ are forms of degree l in P, q and in P, Q, respectively. The old and new Hamiltonians are connected by the relation

$$H\left(P + \left(\frac{\partial S(P, q, t)}{\partial q}\right)^T, q\right) = \mathcal{H}\left(P, q + \left(\frac{\partial S(P, q, t)}{\partial P}\right)^T\right)$$

Equating here the forms of the same order in P, q we obtain

$$\sum_{j=1}^{n} \Omega_j \left(P_j \frac{\partial S_l}{\partial q_j} - Q_j \frac{\partial S_l}{\partial P_j}\right) = \mathcal{H}_l - F_l, \ l = 3, \ldots, L$$

The form F_l is uniquely determined if we know the S_ν, $\mathcal{H}_\nu$ for $\nu \leq l - 1$.

Proof of the theorem about resonant normal form, continued

In the symplectic polar coordinates ρ, φ the last equation takes the form

$$\Omega_1 \frac{\partial S_l}{\partial \varphi_1} + \ldots + \Omega_n \frac{\partial S_l}{\partial \varphi_n} = \mathcal{H}_l - F_l$$

We choose

$$S_l = \sum i \frac{f_k(\rho)}{(k, \Omega)} \exp(i(k, \varphi)), \quad k \in \mathbb{Z}^n \setminus \mathcal{K}$$

where the f_k are the coefficients of the Fourier series of F_l. Then $\mathcal{H}_l$ is in the required normal form. Thus we can successively determine all the S_l, $\mathcal{H}_l$. Returning to Cartesian coordinates we obtain the result.

Order reduction in the resonant normal form

Let $\mathcal{K}$ be a sublattice of $\mathbb{Z}^n$ defining the possible resonances. Suppose that the Hamiltonian is in the resonant normal form for resonances in $\mathcal{K}$. Let r be the the rank of $\mathcal{K}$.

Order reduction in the resonant normal form

Let $\mathcal{K}$ be a sublattice of $\mathbb{Z}^n$ defining the possible resonances. Suppose that the Hamiltonian is in the resonant normal form for resonances in $\mathcal{K}$. Let r be the the rank of $\mathcal{K}$.

Denote $\bar{\mathcal{K}}$ the minimal sublattice of $\mathbb{Z}^n$ that contains $\bar{\mathcal{K}}$ and is such that if some vector of the form $dk, d \in \mathbb{N}, k \in \mathbb{Z}^n$ belongs to $\bar{\mathcal{K}}$, then the vector k also belongs to $\bar{\mathcal{K}}$.

Lemma

$\mathbb{Z}^n$ has a basis such that first r vectors of this basis belong to $\bar{\mathcal{K}}$.

Let R be a matrix whose rows are vectors of this basis. Then R is an integer unimodular matrix, $\det R = \pm 1$.

Order reduction in the resonant normal form

Let $\mathcal{K}$ be a sublattice of $\mathbb{Z}^n$ defining the possible resonances. Suppose that the Hamiltonian is in the resonant normal form for resonances in $\mathcal{K}$. Let r be the the rank of $\mathcal{K}$.

Denote $\bar{\mathcal{K}}$ the minimal sublattice of $\mathbb{Z}^n$ that contains $\bar{\mathcal{K}}$ and is such that if some vector of the form $dk, d \in \mathbb{N}, k \in \mathbb{Z}^n$ belongs to $\bar{\mathcal{K}}$, then the vector k also belongs to $\bar{\mathcal{K}}$.

Lemma

$\mathbb{Z}^n$ has a basis such that first r vectors of this basis belong to $\bar{\mathcal{K}}$.

Let R be a matrix whose rows are vectors of this basis. Then R is an integer unimodular matrix, $\det R = \pm 1$.

Express original Hamiltonian via symplectic polar coordinates ρ, φ. Make canonical change of variables $\rho, \varphi \mapsto I, \psi$ by means of generating function $W = I^T R\varphi$. Then

$$\psi = R\varphi, \ \rho = R^T I$$

The Hamilton function in the new variables does not depend on the last $n - r$ components of ψ, and thus last $n - r$ components of I are the first integrals of the system. These first integrals are linear combinations with integer coefficients of the quantities $\rho_j = \frac{1}{2}(P_j^2 + Q_j^2)$. On a common level of these first integrals the system is reduced to a Hamiltonian system with r degrees of freedom for the first r components of the vectors I, ψ.

Order reduction in the resonant normal form, continued

Example

Let $n = 2$, and possible resonance relations are reducible to the form $2\Omega_1 - \Omega_2 = 0$. Then $\bar{\mathcal{K}} = \{d(2, -1),\ d \in \mathbb{Z}\}$. Any Hamilton function in the resonant normal form depend on phases via combination $2\varphi_1 - \varphi_2$: $\mathcal{H} = \mathcal{H}(\rho_1, \rho_2, 2\varphi_1 - \varphi_2)$. Take

$$R = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

Make canonical change of variables $\rho, \varphi \mapsto I, \psi$ by means of generating function $W = (2\varphi_1 - \varphi_2)I_1 + (-\varphi_1 + \varphi_2)I_2$:

$$\begin{aligned} \psi_1 &= 2\varphi_1 - \varphi_2,\ \psi_2 = -\varphi_1 + \varphi_2, \\ \rho_1 &= 2I_1 - I_2,\ \rho_2 = -I_1 + I_2 \end{aligned}$$

Therefore, $I_1 = \rho_1 + \rho_2$, $I_2 = \rho_1 + 2\rho_2$, and I_2 is the first integral of the system. Behavior of I_1, ψ_1 is described by the Hamiltonian system with the Hamilton function

$$\mathcal{H}(2I_1 - I_2, -I_1 + I_2, \psi_1)$$

Resonances near periodic trajectories

As we know, under rather general assumptions study of a dynamics in the neighborhood of a periodic trajectory on the energy level of an autonomous Hamiltonian system can be reduced to study of a dynamics in the neighborhood of an equilibrium of a time-periodic Hamiltonian system. Moreover, quadratic part of the Hamilton function is time-independent.

Suppose that the eigenvalues of the quadratic part of the Hamiltonian in a neighbourhood of the equilibrium position 0 are all distinct and purely imaginary. Then the Hamilton function can be represented in the form

$$H = \tfrac{1}{2}\Omega_1(p_1^2 + q_1^2) + \tfrac{1}{2}\Omega_2(p_2^2 + q_2^2) + \ldots + \tfrac{1}{2}\Omega_n(p_n^2 + q_n^2) + F(p, q, t).$$

where the expansion of F in the phase variables begins with terms of the third order of smallness, and F has period 2π in time t.

Definition

The eigenfrequencies $\Omega_1, \Omega_2, \ldots, \Omega_n$ satisfy a resonance relation of order $l > 0$ for 2π-periodic hamiltonian i if there exist integers k_i such that $k_1\Omega_1 + \ldots + k_n\Omega_n + k_0 = 0$ and $|k_1| + \ldots + |k_n| = l$.

For example, $\Omega_1 = \Omega_2 + 3$ is a resonance relation of order 2.

Normal form in absence of resonances for non-autonomous system

Theorem (G.D.Birkhoff)

Suppose that the characteristic frequencies Ω_i of the 2π-periodic system do not satisfy any resonance relation of order L or less. Then there is a symplectic change of variables that is 2π-periodic in time and reduces the Hamiltonian function to the same Birkhoff normal form of degree L as if the system were autonomous, with the only difference that the remainder terms of degree $L+1$ and higher depend 2π-periodically on time.

Non-autonomous resonant normal forms

The definition of a normal form must be modified for the case where the eigenfrequencies satisfy some resonance relations. The same modification is also appropriate for nearly resonant frequencies. Let $\mathcal{K}$ be a sublattice of the integer lattice $\mathbb{Z}^{n+1}$ defining the possible resonances.

Definition

A non-autonomous resonant normal form of degree L for the Hamiltonian for resonances in $\mathcal{K}$ is a polynomial of degree L in canonical symplectic variables P_i, Q_i which in the symplectic polar coordinates depends on the phases φ_i and time t only via their combinations $k_1\varphi_1 + k_2\varphi_2 + \ldots + k_n\varphi_n + k_0 t$ for $(k_1, k_2, \ldots, k_n, k_0) \in \mathcal{K}$.

Non-autonomous resonant normal forms

The definition of a normal form must be modified for the case where the eigenfrequencies satisfy some resonance relations. The same modification is also appropriate for nearly resonant frequencies. Let $\mathcal{K}$ be a sublattice of the integer lattice $\mathbb{Z}^{n+1}$ defining the possible resonances.

Definition

A non-autonomous resonant normal form of degree L for the Hamiltonian for resonances in $\mathcal{K}$ is a polynomial of degree L in canonical symplectic variables P_i, Q_i which in the symplectic polar coordinates depends on the phases φ_i and time t only via their combinations $k_1\varphi_1 + k_2\varphi_2 + \ldots + k_n\varphi_n + k_0 t$ for $(k_1, k_2, \ldots, k_n, k_0) \in \mathcal{K}$.

Theorem

Suppose that the eigenfrequencies do not satisfy any resonance relations of degree L or less, except, possibly, for relations $k_1\Omega_1 + k_2\Omega_2 + \ldots + k_n\Omega_n + k_0 = 0$ for $(k_1, k_2, \ldots, k_n, k_0) \in \mathcal{K}$. Then in a neighbourhood of the zero equilibrium position there exists a symplectic close to the identical 2π-periodic in time change of variables $(p, q) \mapsto (P, Q)$ fixing the equilibrium position and such that in the new variables the Hamilton function reduces to a resonant normal form of degree L for resonances in $\mathcal{K}$ up to terms of degree higher than L.

Order reduction in the non-autonomous resonant normal form

Let $\mathcal{K}$ be a sublattice of $\mathbb{Z}^{n+1}$ defining the possible resonances. Suppose that the Hamiltonian has the non-autonomous resonant normal form for resonances in $\mathcal{K}$. Let r be the the rank of $\mathcal{K}$. Then a system has $n - r$ independent integrals in involution which are linear combinations with integer coefficients of the quantities $\rho_j = \frac{1}{2}(P_j^2 + Q_j^2)$. On a common level of these first integrals the system is reduced to an autonomous Hamiltonian system with r degrees of freedom.

Order reduction in the non-autonomous resonant normal form

Let $\mathcal{K}$ be a sublattice of $\mathbb{Z}^{n+1}$ defining the possible resonances. Suppose that the Hamiltonian has the non-autonomous resonant normal form for resonances in $\mathcal{K}$. Let r be the the rank of $\mathcal{K}$. Then a system has $n-r$ independent integrals in involution which are linear combinations with integer coefficients of the quantities $\rho_j = \frac{1}{2}(P_j^2 + Q_j^2)$. On a common level of these first integrals the system is reduced to an autonomous Hamiltonian system with r degrees of freedom.

Example

Consider non-autonomous system with one degree of freedom near the equilibrium. Let all possible resonances are of the form $d(3\Omega + k_0) = 0$. Then truncated at the terms of the 3rd order Hamiltonian in the normal form is

$$\mathcal{H} = \omega\rho + B\rho^{3/2}\cos(3\varphi + k_0 t + \gamma),\ B = \mathrm{const},\ \gamma = \mathrm{const}$$

Introduce $\psi = \varphi + (k_0 t + \gamma)/3$ as a new variable. The Hamilton function for ρ, ψ is

$$\Phi = \delta\rho + B\rho^{3/2}\cos(3\psi)$$

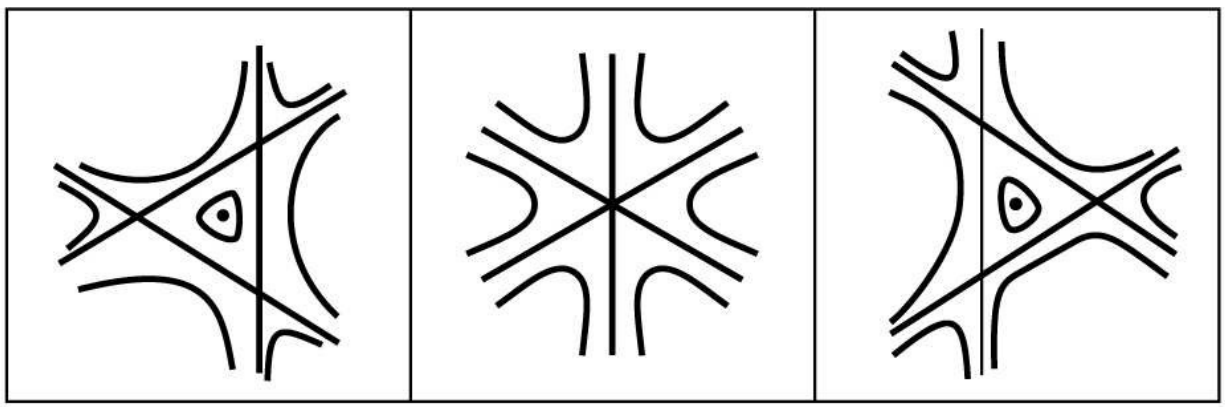
Here $\delta = \Omega + k_0/3$ is the resonance detuning.

Order reduction in the non-autonomous resonant normal form, continued

Consider the Hamilton function

$$\Phi = \delta\rho + B\rho^{3/2}\cos 3\psi,\ \ \delta = \Omega + k_0/3$$

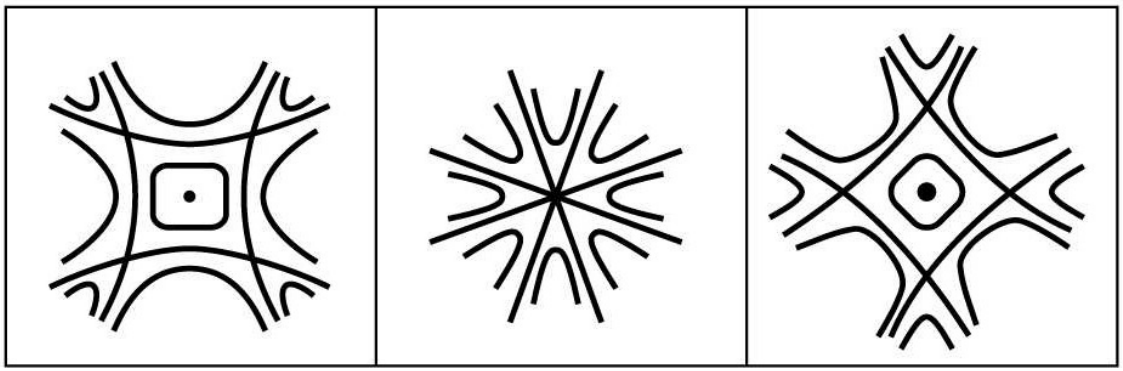
Assume that $B \neq 0$. Then in the plane with symplectic polar coordinates ρ, ψ the phase portrait for this Hamilton function for different δ looks as follows:



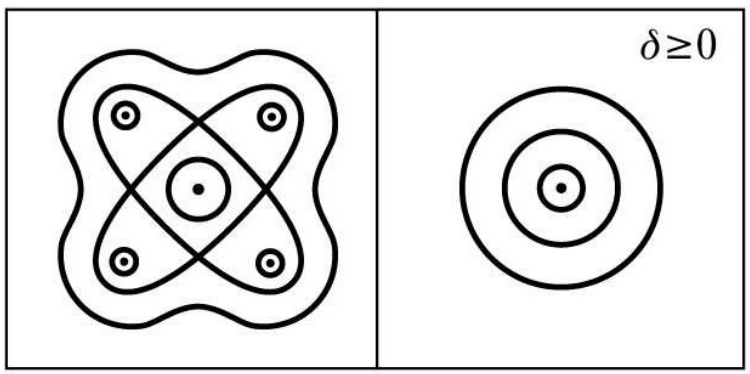
For this resonance the periodic trajectory of the original autonomous Hamiltonian system is unstable.

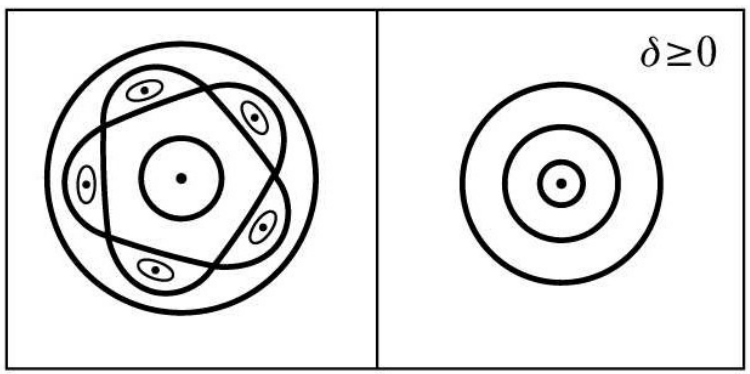
Order reduction in the non-autonomous resonant normal form, continued

In an analogous manner one can construct phase portraits of systems in normal form for resonances $4\Omega + k_0 = 0$ and $k\Omega + k_0 = 0,\ k \geq 5$. These phase portraits looks as follows:



b





For resonances of order great or equal 5 the periodic trajectory of the original autonomous Hamiltonian system is stable.

LECTURE 17

PERTURBATION THEORY FOR INTEGRABLE SYSTEMS

Perturbation theory

One of typical problems of perturbation theory can be formulated as follows. Consider system of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon), \ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

This system is called *the perturbed system*, term $\varepsilon v_1(x, \varepsilon)$ is called *the perturbation*.
Putting $\varepsilon = 0$ we get *the unperturbed system*, the dynamics in this system is assumed to be known.
The goal of the perturbation theory is studying of dynamics of perturbed system on **"long enough"** time intervals.

Perturbation theory

One of typical problems of perturbation theory can be formulated as follows. Consider system of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

This system is called *the perturbed system*, term $\varepsilon v_1(x, \varepsilon)$ is called *the perturbation*.
Putting $\varepsilon = 0$ we get *the unperturbed system*, the dynamics in this system is assumed to be known.
The goal of the perturbation theory is studying of dynamics of perturbed system on **"long enough"** time intervals.

The words **"long enough"** are very important here. One should consider typical time intervals on which the perturbation has essential effect on dynamics.

Unperturbed systems with rotating phases

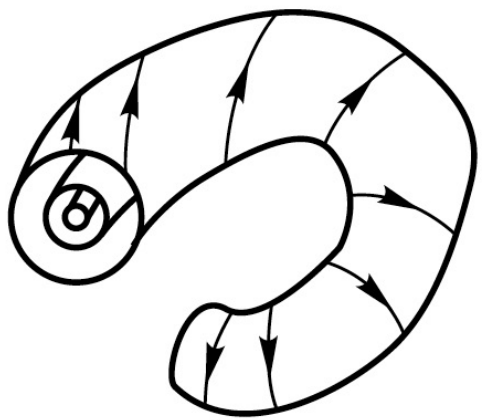
Assume that the unperturbed system is *integrable* .
There are different definitions of integrability. Here the system is called integrable in some domain of the phase space, if in this domain one can make transformation of variables

$$x \mapsto I, \varphi, \quad I \in D \subset \mathbb{R}^n, \ \varphi \in \mathbb{T}^m, \ m+n=N$$

such that in new variables the system has the form

$$\dot{I} = 0, \dot{\varphi} = \omega(I)$$

Thus the domain D is foliated by invariant tori. Variable I enumerates tori. For given torus (i. e. given I) variable φ defines coordinates (*phases*) on this torus. In these coordinates motion on torus is m-dimensional rotation with the frequency vector $\omega(I)$.



Perturbed systems with rotating phases. Averaging principle

In the variables I, φ the perturbed system has the form

$$\dot{I} = \varepsilon f(I, \varphi, \varepsilon), \ \dot{\varphi} = \omega(I) + \varepsilon g(I, \varphi, \varepsilon)$$

The functions f and g have period 2π in components of φ. The variables I here are called *the slow variables*, and the phases φ are called *the fast variables*. The system in this form is called a *perturbed system with rotating phases*.

Perturbed systems with rotating phases. Averaging principle

In the variables I, φ the perturbed system has the form

$$\dot{I} = \varepsilon f(I, \varphi, \varepsilon), \ \dot{\varphi} = \omega(I) + \varepsilon g(I, \varphi, \varepsilon)$$

The functions f and g have period 2π in components of φ. The variables I here are called *the slow variables*, and the phases φ are called *the fast variables*. The system in this form is called a *perturbed system with rotating phases*.

In applications one is usually interested mainly in the behaviour of the slow variables on the time intervals of the length at least of order $1/\varepsilon$. To describe this behavior *the averaging principle* prescribes to replace the first equation in the perturbed system with rotating phases by *the averaged equation*

$$\dot{J} = \varepsilon F(J), \ F(J) = \frac{1}{(2\pi)^m} \int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} f(J, \varphi, 0)\, d\varphi_1 d\varphi_2 \dots d\varphi_m.$$

(Here φ_j are components of φ.)

Perturbed systems with rotating phases. Averaging principle

In the variables I, φ the perturbed system has the form

$$\dot{I} = \varepsilon f(I, \varphi, \varepsilon), \ \dot{\varphi} = \omega(I) + \varepsilon g(I, \varphi, \varepsilon)$$

The functions f and g have period 2π in components of φ. The variables I here are called *the slow variables*, and the phases φ are called *the fast variables*. The system in this form is called a *perturbed system with rotating phases*.

In applications one is usually interested mainly in the behaviour of the slow variables on the time intervals of the length at least of order $1/\varepsilon$. To describe this behavior *the averaging principle* prescribes to replace the first equation in the perturbed system with rotating phases by *the averaged equation*

$$\dot{J} = \varepsilon F(J), \ F(J) = \frac{1}{(2\pi)^m} \int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} f(J, \varphi, 0)\, d\varphi_1 d\varphi_2 \dots d\varphi_m.$$

(Here φ_j are components of φ.)
Let $I(t)$ be a slow motion in the original system, and $J(t)$ in the averaged one, $J(0) = I(0)$. According to the averaging principle, $I(t)$ is replaced by $J(t)$.
To justify this recipe (which does not always yield a correct answer) we must find conditions which ensure that $|I(t) - J(t)| \to 0$ for $0 \le t \le 1/\varepsilon$ as $\varepsilon \to 0$.
If the last relation holds, then it is desirable to have an upper estimate of $|I(t) - J(t)|$ for $0 \le t \le 1/\varepsilon$

Averaging principle, example

Example

Consider the system of equations

$$\dot{I} = \varepsilon(a + b\cos\varphi),\ \dot{\varphi} = \omega$$

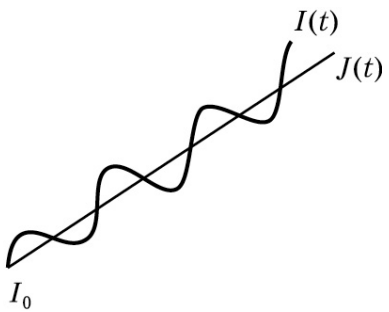
and the corresponding averaged system

$$\dot{J} = \varepsilon a$$

Here

$$I(t) = I(0) + \varepsilon a t + \varepsilon b[\sin(\omega t + \varphi_0) - \sin\varphi_0]/\omega,\ J(t) = I_0 + \varepsilon a t$$

The solutions of the exact system oscillate about the solutions of the averaged system with an amplitude of order ε and with frequency ω.



Averaging principle, example

Example

Consider the system of equations

$$\dot{I} = \varepsilon(a + b\cos\varphi),\ \dot{\varphi} = \omega$$

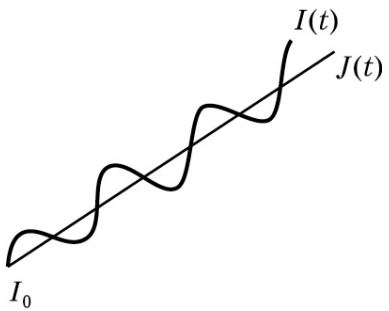
and the corresponding averaged system

$$\dot{J} = \varepsilon a$$

Here

$$I(t) = I(0) + \varepsilon a t + \varepsilon b[\sin(\omega t + \varphi_0) - \sin\varphi_0]/\omega,\ J(t) = I_0 + \varepsilon a t$$

The solutions of the exact system oscillate about the solutions of the averaged system with an amplitude of order ε and with frequency ω.



Averaging amounts to dropping the purely periodic term on the right-hand side of the equation. This term has the same order as the remaining term. But it oscillates and causes only small oscillations of the solution. The remaining term causes a drift which over time $1/(\varepsilon a)$ changes I by 1.

Averaging principle, example

Example

Consider the system of equations

$$\dot{I} = \varepsilon(a + b\cos\varphi),\ \dot{\varphi} = \omega$$

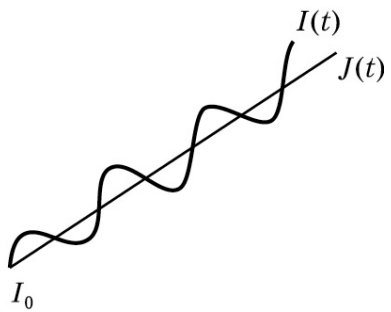
and the corresponding averaged system

$$\dot{J} = \varepsilon a$$

Here

$$I(t) = I(0) + \varepsilon a t + \varepsilon b[\sin(\omega t + \varphi_0) - \sin\varphi_0]/\omega,\ J(t) = I_0 + \varepsilon a t$$

The solutions of the exact system oscillate about the solutions of the averaged system with an amplitude of order ε and with frequency ω.



Averaging amounts to dropping the purely periodic term on the right-hand side of the equation. This term has the same order as the remaining term. But it oscillates and causes only small oscillations of the solution. The remaining term causes a drift which over time $1/(\varepsilon a)$ changes I by 1.

The averaging principle is based on the idea that in the general case, too, the oscillating terms discarded in averaging cause only small oscillations, which are superimposed on the drift described by the averaged system.

Averaging principle, example, continued

Example (The most standard example for the averaging method)

The van der Pol equation
$\ddot{x} = -x + \varepsilon(1 - x^2)\dot{x}$
describes oscillations with a small nonlinear "friction" which is negative for large amplitudes, and positive for small ones.

Averaging principle, example, continued

Example (The most standard example for the averaging method)

The van der Pol equation
$\ddot{x} = -x + \varepsilon(1 - x^2)\dot{x}$
describes oscillations with a small nonlinear "friction" which is negative for large amplitudes, and positive for small ones. Make transformation to symplectic polar coordinates: $x = \sqrt{2I}\sin\varphi,\ \dot{x} = \sqrt{2I}\cos\varphi$. The unperturbed equation take the form $\dot{I} = 0, \dot{\varphi} = 1$.

Averaging principle, example, continued

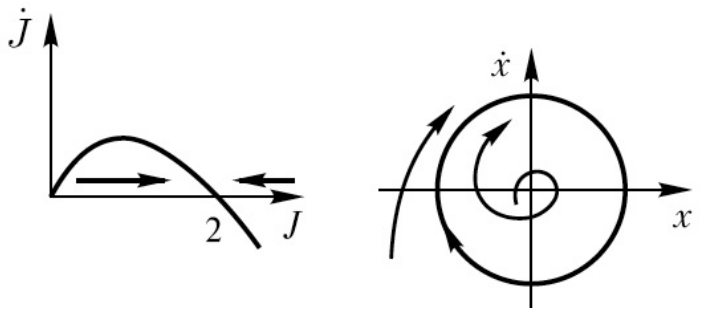
Example (The most standard example for the averaging method)

The van der Pol equation
$\ddot{x} = -x + \varepsilon(1 - x^2)\dot{x}$
describes oscillations with a small nonlinear "friction" which is negative for large amplitudes, and positive for small ones. Make transformation to symplectic polar coordinates: $x = \sqrt{2I}\sin\varphi,\ \dot{x} = \sqrt{2I}\cos\varphi$. The unperturbed equation take the form $\dot{I} = 0, \dot{\varphi} = 1$. The equation for $I = (x^2 + \dot{x}^2)/2$ in the perturbed motion has the form
$\dot{I} = \varepsilon(1 - x^2)\dot{x}^2 = 2\varepsilon I(1 - 2I\sin^2\varphi)\cos^2\varphi$

Averaging principle, example, continued

Example (The most standard example for the averaging method)

The van der Pol equation
$\ddot{x} = -x + \varepsilon(1 - x^2)\dot{x}$
describes oscillations with a small nonlinear "friction" which is negative for large amplitudes, and positive for small ones. Make transformation to symplectic polar coordinates: $x = \sqrt{2I}\sin\varphi,\ \dot{x} = \sqrt{2I}\cos\varphi$. The unperturbed equation take the form $\dot{I} = 0, \dot{\varphi} = 1$. The equation for $I = (x^2 + \dot{x}^2)/2$ in the perturbed motion has the form
$\dot{I} = \varepsilon(1 - x^2)\dot{x}^2 = 2\varepsilon I(1 - 2I\sin^2\varphi)\cos^2\varphi$
The averaged equation is
$\dot{J} = \varepsilon J(1 - \frac{1}{2}J)$
It has a repelling equilibrium $J = 0$, and an attracting one $J = 2$. To the equilibrium $J = 0$ there corresponds the equilibrium $x = 0$, $\dot{x} = 0$ of the original equation. To the equilibrium $J = 2$ there corresponds a stable limit cycle of the original equation, which is close to the circle $x^2 + \dot{x}^2 = 4$.



Procedure for eliminating fast variables and decoupling of slow and fast motions. Non-resonant case

The important role in all problems connected with the averaging principle is played by changes of variables that enable one to eliminate, with the required accuracy, the fast phases from the equations of perturbed motion and thus separate the slow motion from the fast one. These changes of variables reduce the original system of equations to the averaged one in the first approximation.

Procedure for eliminating fast variables and decoupling of slow and fast motions. Non-resonant case

The important role in all problems connected with the averaging principle is played by changes of variables that enable one to eliminate, with the required accuracy, the fast phases from the equations of perturbed motion and thus separate the slow motion from the fast one. These changes of variables reduce the original system of equations to the averaged one in the first approximation.

The elimination of fast variables is performed as follows. The desirable transformation of variables $(I, \varphi) \mapsto (J, \psi)$ is sought as a formal series

$$I = J + \varepsilon u_1(J, \psi) + \varepsilon^2 u_2(J, \psi) + \ldots, \quad \varphi = \psi + \varepsilon v_1(J, \psi) + \varepsilon^2 v_2(J, \psi) + \ldots,$$

where functions u_j, v_j are 2π-periodic in ψ. This transformation should be chosen in such a way that in the new variables the right hand sides of equations of motion do not contain fast variables, i.e. equations of motion should have the form

$$\dot{J} = \varepsilon F_0(J) + \varepsilon^2 F_1(J) + \ldots, \quad \dot{\psi} = \omega(J) + \varepsilon G_0(J) + \varepsilon^2 G_1(J) + \ldots$$

Procedure for eliminating fast variables, continued

Substituting formulas for transformation into equations of motion, and equating the terms of the same order in ε, we obtain the following chain of relations:

$$\begin{aligned} F_0(J) &= f(J,\psi,0) - \frac{\partial u_1}{\partial \psi}\omega,\ G_0(J) = g(J,\psi,0) + \frac{\partial \omega}{\partial J}u_1 - \frac{\partial v_1}{\partial \psi}\omega, \\ F_i(J) &= X_i(J,\psi) - \frac{\partial u_{i+1}}{\partial \psi}\omega,\ G_i(J) = Y_i(J,\psi) + \frac{\partial \omega}{\partial J}u_{i+1} - \frac{\partial v_{i+1}}{\partial \psi}\omega,\ i \geq 1. \end{aligned}$$

Functions X_i, Y_i are uniquely determined by the terms u_1, $v_1, \ldots, u_i$, v_i in expansion of variables transformation.

Procedure for eliminating fast variables, continued

Substituting formulas for transformation into equations of motion, and equating the terms of the same order in ε, we obtain the following chain of relations:

$$\begin{aligned} F_0(J) &= f(J,\psi,0) - \frac{\partial u_1}{\partial \psi}\omega,\ G_0(J) = g(J,\psi,0) + \frac{\partial \omega}{\partial J}u_1 - \frac{\partial v_1}{\partial \psi}\omega, \\ F_i(J) &= X_i(J,\psi) - \frac{\partial u_{i+1}}{\partial \psi}\omega,\ G_i(J) = Y_i(J,\psi) + \frac{\partial \omega}{\partial J}u_{i+1} - \frac{\partial v_{i+1}}{\partial \psi}\omega,\ i \geq 1. \end{aligned}$$

Functions X_i, Y_i are uniquely determined by the terms u_1, $v_1, \ldots, u_i$, v_i in expansion of variables transformation. The first equation in this chain implies that

$$F_0(J) = f_0(J),\ u_1(J,\psi) = \sum_{k\neq 0} \frac{f_k}{i(k,\omega)} \exp(i(k,\psi)) + u_1^0(J),$$

where f_k, $k \in \mathbb{Z}^m$ are Fourier coefficients of function f at $\varepsilon = 0$, and u_1^0 is an arbitrary function of J. It is assumed that the denominators in this formula do not vanish, and that the series converges and determines a smooth function. In the same way from the other equations of the chain one can sequentially determine $F_0, v_1, \ldots, G_i, u_{i+1}, F_i, v_{i+1}, i \geq 1$.

Procedure for eliminating fast variables and decoupling of slow and fast motions. Non-resonant case, continued

If the series for the change of variables are truncated at the terms of order $r \geq 1$, then we obtain a change of variables which reduces the equations of the perturbed motion to the form

$$\dot{J} = \varepsilon F_0(J) + \varepsilon^2 F_1(J) + \ldots + \varepsilon^r F_{r-1}(J) + \varepsilon^{r+1}\alpha(J, \psi, \varepsilon)$$
$$\dot{\psi} = \omega(J) + \varepsilon G_0(J) + \ldots + \varepsilon^r G_{r-1}(J) + \varepsilon^{r+1}\beta(J, \psi, \varepsilon)$$

Discarding here the terms of order ε^{r+1} we obtain a truncated system of the r-th approximation. The equation for J is decoupled from the other equations and can be solved separately. Then the behavior of ψ is determined by the means of the quadrature. The behavior of original variable I in this approximation is a slow drift (described by the equation for J), on which small oscillations (described by transformation of variables) are superimposed. The behavior of φ can be represented as a rotation with slowly varying frequency, on which oscillations are also superimposed.

Procedure for eliminating fast variables and decoupling of slow and fast motions. Non-resonant case, continued

If the series for the change of variables are truncated at the terms of order $r \geq 1$, then we obtain a change of variables which reduces the equations of the perturbed motion to the form

$$\dot{J} = \varepsilon F_0(J) + \varepsilon^2 F_1(J) + \ldots + \varepsilon^r F_{r-1}(J) + \varepsilon^{r+1}\alpha(J, \psi, \varepsilon)$$
$$\dot{\psi} = \omega(J) + \varepsilon G_0(J) + \ldots + \varepsilon^r G_{r-1}(J) + \varepsilon^{r+1}\beta(J, \psi, \varepsilon)$$

Discarding here the terms of order ε^{r+1} we obtain a truncated system of the r-th approximation. The equation for J is decoupled from the other equations and can be solved separately. Then the behavior of ψ is determined by the means of the quadrature. The behavior of original variable I in this approximation is a slow drift (described by the equation for J), on which small oscillations (described by transformation of variables) are superimposed. The behavior of φ can be represented as a rotation with slowly varying frequency, on which oscillations are also superimposed.

We have been assuming that in the formulae for the change of variables the denominators $(k, \omega(J)) = k_1\omega_1(J) + \ldots + k_m\omega_m(J)$ do not vanish in the domain considered. This assumption holds for single-frequency systems with non-vanishing frequency, for systems with constant incommensurable frequencies, for systems with finitely many harmonics in the perturbation. But for general multi-frequency systems this condition fails. This is the famous *small denominators* problem.

Averaging in one-frequency systems

Consider the system of equations of the perturbed motion which has only one phase. Assume that the frequency $\omega(I)$ does not vanish: $\omega(I) > c^{-1} > 0$, $c = \mathrm{const}$. Assume that the solution $J(t)$ of the averaged system for $0 \leq t \leq 1/\varepsilon$ does not approach too closely the boundary of the domain where the system is defined.

Theorem (P.Fatou)

The difference between the slow motion $I(t)$ in the exact system and $J(t)$ in the averaged system remains small over time $1/\varepsilon$:

$$|I(t) - J(t)| < c_1\varepsilon \text{ if } I(0) = J(0),\ 0 \leq t \leq 1/\varepsilon$$

where $c_1 = \mathrm{const} > 0$.

Averaging in one-frequency systems

Consider the system of equations of the perturbed motion which has only one phase. Assume that the frequency $\omega(I)$ does not vanish: $\omega(I) > c^{-1} > 0$, $c = \mathrm{const}$. Assume that the solution $J(t)$ of the averaged system for $0 \leq t \leq 1/\varepsilon$ does not approach too closely the boundary of the domain where the system is defined.

Theorem (P.Fatou)

The difference between the slow motion $I(t)$ in the exact system and $J(t)$ in the averaged system remains small over time $1/\varepsilon$:

$$|I(t) - J(t)| < c_1\varepsilon \text{ if } I(0) = J(0),\ 0 \leq t \leq 1/\varepsilon$$

where $c_1 = \mathrm{const} > 0$.

Proof.

The change of variables $I = J + \varepsilon u_1(J, \varphi)$ differs from the identity by a quantity of order ε. It reduces the exact system to the averaged one with an addition of a small (of order ε^2) perturbation. Over time $1/\varepsilon$ this perturbation can change the value of the slow variable, compared to its value in the averaged system, only by a quantity of order ε. Returning to the original variables we obtain the result of the theorem.

This proof, which is based on the elimination of the fast phase by a change of variables, is due to N.N.Bogolyubov.

Averaging in one-frequency systems, continued

Theorem

Suppose that the averaged system has a non-degenerate equilibrium position. Then the exact system has a periodic trajectory along which the slow variables vary within a neighbourhood of this equilibrium position of size of order ε. If all the eigenvalues of the averaged system linearized about this equilibrium position have negative real parts, then the periodic trajectory is orbitally asymptotically stable. If the real part of one of the eigenvalues is positive, then the trajectory is unstable.

Averaging in one-frequency systems, continued

Theorem

Suppose that the averaged system has a non-degenerate equilibrium position. Then the exact system has a periodic trajectory along which the slow variables vary within a neighbourhood of this equilibrium position of size of order ε. If all the eigenvalues of the averaged system linearized about this equilibrium position have negative real parts, then the periodic trajectory is orbitally asymptotically stable. If the real part of one of the eigenvalues is positive, then the trajectory is unstable.

Proof.

Make change of variables $I = J + \varepsilon u_1(J, \varphi)$. The Poincaré return map P for the plane $\{\varphi = 0 \mod 2\pi\}$ has the form

$$P\colon J \to J + \varepsilon 2\pi F(J)/\omega(J) + O(\varepsilon^2)$$

The truncated map (i.e. map without last term $O(\varepsilon^2)$) has a non-degenerate fixed point J_*: $F(J_*) = 0$, $\det(\partial F(J_*)/\partial J) \neq 0$. By the implicit function theorem, for sufficiently small ε the map P has a fixed point $J = J_* + O(\varepsilon)$, which serves as the initial condition for the periodic trajectory. Approximate calculation of multipliers of this fixed points provides the required stability properties. $\square$

Example: Stability of the upper position of a pendulum with vibrating suspension point

The equation of motion of a pendulum whose point of suspension performs vertical sinusoidal oscillations has the form

$$\ddot{\theta} + (g - a\Omega^2 \sin \Omega t)\ell^{-1} \sin \theta = 0$$

Here θ is the angle of deviation of the pendulum from the vertical, a and Ω are the amplitude and the frequency of the oscillations of the suspension point, ℓ is the length of the pendulum, g the acceleration of gravity. Assume that $a = a_0\varepsilon$ and $\Omega = \Omega_0/\varepsilon$. We write down the equations of motion in the canonical Hamiltonian form:

$$\theta' = \varepsilon \partial H/\partial p,\ p' = -\varepsilon \partial H/\partial \theta,\ \ H = \tfrac{1}{2}(p/\ell - a_0\Omega_0 \sin\tau \sin\theta)^2 - g\ell\cos\theta$$

Here $\tau = t/\varepsilon$ is the new time, prime denotes differentiation with respect to τ, and p is the momentum canonically conjugate to the angle θ. The system averaged over τ has the form

$$\theta' = \varepsilon \partial \mathcal{H}/\partial p,\ p' = -\varepsilon \partial \mathcal{H}/\partial \theta,\ \ \mathcal{H} = \tfrac{1}{2}p^2/\ell^2 + V,\ \ V = \frac{1}{4}a_0^2\Omega_0^2 \sin^2\theta - g\ell\cos\theta$$

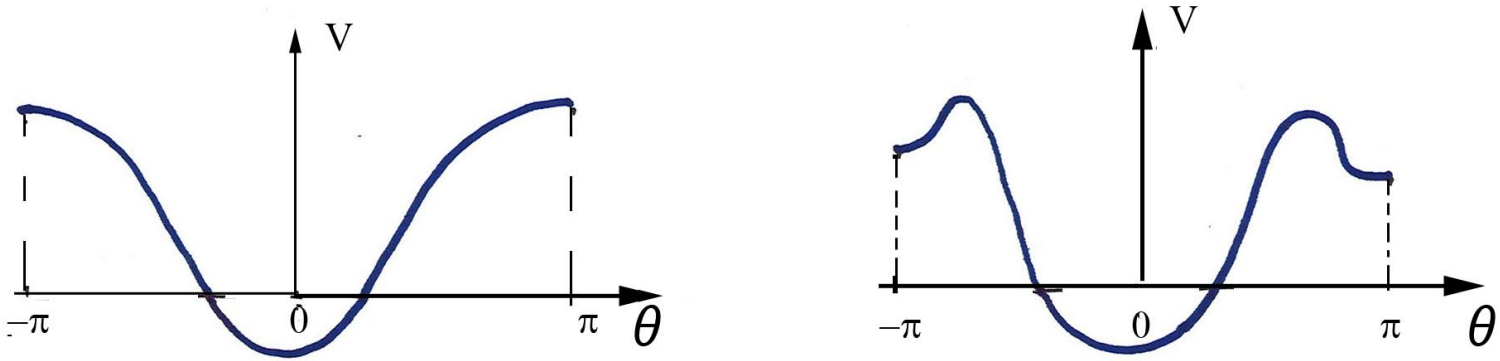
The function V is called the effective potential energy.

Example: Stability of the upper position of a pendulum with vibrating suspension point, continued

The effective potential energy

$$V = \frac{1}{4} a_0^2 \Omega_0^2 \sin^2 \theta - g\ell \cos \theta$$

at the point $\theta = \pi$ corresponding to the upper equilibrium position of the pendulum has a non-degenerate maximum for $\sqrt{2g\ell}/(\Omega_0 a) > 1$ and a non-degenerate minimum for $\sqrt{2g\ell}/(\Omega_0 a) < 1$.



For sufficiently small ε, the upper position of the original pendulum for these values of $\sqrt{2g\ell}/(\Omega_0 a)$ is unstable or stable, respectively. The instability follows from a previous Theorem. To prove stability one should apply Kolmogorov-Arnold-Moser (KAM) theory.

Exercises

1. Pass to the symplectic polar coordinates in the van der Pol equation. By means of a transformation of variables eliminate dependence of the fast phase up to terms $O(\varepsilon^2)$ inclusively from the equation for slow variable.

2. Write down Hamiltonian equations of motion of a pendulum with the vibrating suspension point.

AVERAGING AND ELIMINATION OF FAST PHASES IN HAMILTONIAN SYSTEMS

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian.

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

The Hamiltonian H_0 of the unperturbed system depends only on the action variables: $H_0 = H_0(I)$. The equations of the unperturbed motion have the usual form:

$$\dot{I} = 0,\ \dot{\varphi} = \partial H_0/\partial I$$

Suppose that the system is subjected to a small Hamiltonian perturbation. The perturbed motion is described by the system with Hamiltonian $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$:

$$\dot{I} = -\varepsilon \partial H_1/\partial \varphi,\ \ \dot{\varphi} = \partial H_0/\partial I + \varepsilon \partial H_0/\partial I$$

The perturbing Hamiltonian $H_1(I, \varphi, \varepsilon)$ has period 2π in components of φ.

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

The Hamiltonian H_0 of the unperturbed system depends only on the action variables: $H_0 = H_0(I)$. The equations of the unperturbed motion have the usual form:

$$\dot{I} = 0,\ \dot{\varphi} = \partial H_0/\partial I$$

Suppose that the system is subjected to a small Hamiltonian perturbation. The perturbed motion is described by the system with Hamiltonian $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$:

$$\dot{I} = -\varepsilon \partial H_1/\partial \varphi,\ \ \dot{\varphi} = \partial H_0/\partial I + \varepsilon \partial H_0/\partial I$$

The perturbing Hamiltonian $H_1(I, \varphi, \varepsilon)$ has period 2π in components of φ. This form of equations is standard for applying the averaging principle.

Application of the averaging principle

Suppose that the unperturbed frequencies $\partial H_0/\partial I$ do not satisfy identical linear relations with integer coefficients. In accordance with the averaging principle for approximate description of the evolution of the variables I we average equations of motion over the phases φ.

Theorem

In a Hamiltonian system with n degrees of freedom and n frequencies there is no evolution of the slow variables in the sense that the averaged system has the form $\dot{J} = 0$.

Proof.

When calculating the integral of $\partial H_0/\partial \varphi_j$ over the n-dimensional torus we can first integrate with respect to the variable φ_j. This single integral is equal to the increment of the periodic function H_1 over the period, that is, to zero. $\square$

Application of the averaging principle

Suppose that the unperturbed frequencies $\partial H_0/\partial I$ do not satisfy identical linear relations with integer coefficients. In accordance with the averaging principle for approximate description of the evolution of the variables I we average equations of motion over the phases φ.

Theorem

In a Hamiltonian system with n degrees of freedom and n frequencies there is no evolution of the slow variables in the sense that the averaged system has the form $\dot{J} = 0$.

Proof.

When calculating the integral of $\partial H_0/\partial \varphi_j$ over the n-dimensional torus we can first integrate with respect to the variable φ_j. This single integral is equal to the increment of the periodic function H_1 over the period, that is, to zero. □

Remark

To preserve the Hamiltonian form of the equations we slightly generalize the averaging principle: we average also the equation describing the variation of the angles (phases) φ. The resulting averaged system has the Hamiltonian $\mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J)$, where $\mathcal{H}_1(J)$ is the average of $H_1(J, \varphi, 0)$ over φ. Hence the phases undergo the uniform rotation with frequencies $\partial H_0/\partial J + \partial \mathcal{H}_1/\partial J$.

Averaging in systems with proper degeneration

One often encounters problems with *proper degeneracy*, in which the unperturbed Hamiltonian depends not on all the action variables and, correspondingly, some of the unperturbed frequencies are identically equal to zero: $H = H_0(I_1, \ldots, I_\ell) + \varepsilon H_1(I, \varphi, \varepsilon)$, $\ell < n$. The phases φ_j, $j > \ell$, are slow variables. According to the averaging principle, for approximate description of the evolution we must average the equations of the perturbed motion over the fast phases φ_i, $i \leq \ell$.

Theorem

In a Hamiltonian system with n degrees of freedom and ℓ frequencies, $\ell < n$, the variables conjugate to the fast phases are integrals of the averaged system.

Averaging in systems with proper degeneration

One often encounters problems with *proper degeneracy*, in which the unperturbed Hamiltonian depends not on all the action variables and, correspondingly, some of the unperturbed frequencies are identically equal to zero: $H = H_0(I_1, \ldots, I_\ell) + \varepsilon H_1(I, \varphi, \varepsilon)$, $\ell < n$. The phases φ_j, $j > \ell$, are slow variables. According to the averaging principle, for approximate description of the evolution we must average the equations of the perturbed motion over the fast phases φ_i, $i \leq \ell$.

Theorem

In a Hamiltonian system with n degrees of freedom and ℓ frequencies, $\ell < n$, the variables conjugate to the fast phases are integrals of the averaged system.

Example (Laplace-Lagrange theorem on the stability of the Solar System)

Consider the n-body problem under the assumption that the mass of one body (the Sun) is much larger than the masses of the other bodies (the planets). The unperturbed system is the one in which the planets do not interact with each other, and the Sun is at rest. The unperturbed system decomposes into $n-1$ Keplers problems. We suppose that the unperturbed orbits of the planets are Keplerian ellipses. In the problem under consideration there are $n-1$ fast phases – the mean longitudes of the planets. Their conjugate variables $\Lambda_j = \sqrt{\mu_j a_j}$, $j = 1, \ldots, n-1$, are integrals of the system averaged over the fast phases. Here the a_j are the major semiaxes of the Keplerian elliptic orbits of the planets, and the μ_j are factors depending on the masses.

Symplectic elimination of fast variables: Lindstedt's method

Consider a perturbed Hamiltonian $H = H_0(I) + H_1(I, \varphi, \varepsilon)$ with n degrees of freedom. Suppose that the frequencies do not satisfy any identical resonance relations. We try to find a symplectic near-identity change of variables $I, \varphi \mapsto J, \psi$ so that the new Hamiltonian H depends only on the slow variables: $\mathcal{H} = \mathcal{H}(J, \varepsilon)$. We seek a generating function of the change of variables and the new Hamiltonian in the form of formal series in ε:

$$I = J + \varepsilon \partial S/\partial \varphi, \ \psi = \varphi + \varepsilon \partial S/\partial J,$$

$$S(J, \varphi, \varepsilon) = S_1(J, \varphi) + S_2(J, \varphi) + \dots, \ \mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J) + \dots.$$

The functions S_i must have period 2π in φ. The old and new Hamiltonians satisfy the relation

$$\mathcal{H}(J, \varepsilon) = H_0(J + \varepsilon \partial S/\partial \varphi) + \varepsilon H_1(J + \varepsilon \partial S/\partial \varphi, \varphi, \varepsilon)$$

Equating here the terms of the same order in ε we obtain the system of equations

$$\mathcal{H}_1(J) = \frac{\partial H_0}{\partial J}\frac{\partial S_1}{\partial \varphi} + H_1(J, \varphi, 0), \ \mathcal{H}_j(J) = \frac{\partial H_0}{\partial J}\frac{\partial S_j}{\partial \varphi} + F_J(J, \varphi)$$

The function F_j is a polynomial in $\partial S_1/\partial \varphi, \dots, \partial S_{j-1}/\partial \varphi$.

Symplectic elimination of fast variables: Lindstedt's method, continued

Take as $\mathcal{H}_1(J)$ the average of $H_1(J, \varphi, 0)$ over φ (i.e. average over all components of φ). Then

$$S_1 = -\sum_{k \neq 0} \frac{h_k}{i(k, \omega)} \exp(i(k, \psi)) + S_1^0(J), \quad \omega(J) = \frac{\partial H_0(J)}{\partial J}$$

where h_k, $k \in \mathbb{Z}^m$ are Fourier coefficients of function $H_1(J, \varphi, 0)$ and S_1^0 is an arbitrary function of J. It is assumed that the denominators in this formula do not vanish, and that the series converges and determines a smooth function. In a similar way we can determine all $\mathcal{H}_j$, S_j.

LECTURE 18

AVERAGING AND ELIMINATION OF FAST PHASES IN HAMILTONIAN SYSTEMS

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian.

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

The Hamiltonian H_0 of the unperturbed system depends only on the action variables: $H_0 = H_0(I)$. The equations of the unperturbed motion have the usual form:

$$\dot{I} = 0,\ \dot{\varphi} = \partial H_0/\partial I$$

Suppose that the system is subjected to a small Hamiltonian perturbation. The perturbed motion is described by the system with Hamiltonian $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$:

$$\dot{I} = -\varepsilon \partial H_1/\partial \varphi,\ \ \dot{\varphi} = \partial H_0/\partial I + \varepsilon \partial H_0/\partial I$$

The perturbing Hamiltonian $H_1(I, \varphi, \varepsilon)$ has period 2π in components of φ.

Hamiltonian perturbation theory

Hamiltonian perturbation theory is dealing with the systems of the form

$$\dot{x} = v_0(x) + \varepsilon v_1(x, \varepsilon),\ x \in \mathbb{R}^N, 0 < \varepsilon \ll 1$$

in the case when both the unperturbed system and the perturbation are Hamiltonian. Suppose that the unperturbed Hamiltonian system is completely integrable, some domain of its phase space is foliated into invariant tori, and the action angle variables I, φ are introduced in this domain:

$$I = (I_1, \ldots, I_n) \in D \subset \mathbb{R}^n,\ \varphi = (\varphi_1, \ldots, \varphi_n) \bmod 2\pi \in \mathbb{T}^n$$

The Hamiltonian H_0 of the unperturbed system depends only on the action variables: $H_0 = H_0(I)$. The equations of the unperturbed motion have the usual form:

$$\dot{I} = 0,\ \dot{\varphi} = \partial H_0/\partial I$$

Suppose that the system is subjected to a small Hamiltonian perturbation. The perturbed motion is described by the system with Hamiltonian $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$:

$$\dot{I} = -\varepsilon \partial H_1/\partial \varphi,\ \ \dot{\varphi} = \partial H_0/\partial I + \varepsilon \partial H_0/\partial I$$

The perturbing Hamiltonian $H_1(I, \varphi, \varepsilon)$ has period 2π in components of φ. This form of equations is standard for applying the averaging principle.

Application of the averaging principle

Suppose that the unperturbed frequencies $\partial H_0/\partial I$ do not satisfy identical linear relations with integer coefficients. In accordance with the averaging principle for approximate description of the evolution of the variables I we average equations of motion over the phases φ.

Theorem

In a Hamiltonian system with n degrees of freedom and n frequencies there is no evolution of the slow variables in the sense that the averaged system has the form $\dot{J} = 0$.

Proof.

When calculating the integral of $\partial H_0/\partial \varphi_j$ over the n-dimensional torus we can first integrate with respect to the variable φ_j. This single integral is equal to the increment of the periodic function H_1 over the period, that is, to zero. $\square$

Application of the averaging principle

Suppose that the unperturbed frequencies $\partial H_0/\partial I$ do not satisfy identical linear relations with integer coefficients. In accordance with the averaging principle for approximate description of the evolution of the variables I we average equations of motion over the phases φ.

Theorem

In a Hamiltonian system with n degrees of freedom and n frequencies there is no evolution of the slow variables in the sense that the averaged system has the form $\dot{J} = 0$.

Proof.

When calculating the integral of $\partial H_0/\partial \varphi_j$ over the n-dimensional torus we can first integrate with respect to the variable φ_j. This single integral is equal to the increment of the periodic function H_1 over the period, that is, to zero. □

Remark

To preserve the Hamiltonian form of the equations we slightly generalize the averaging principle: we average also the equation describing the variation of the angles (phases) φ. The resulting averaged system has the Hamiltonian $\mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J)$, where $\mathcal{H}_1(J)$ is the average of $H_1(J, \varphi, 0)$ over φ. Hence the phases undergo the uniform rotation with frequencies $\partial H_0/\partial J + \varepsilon \partial \mathcal{H}_1/\partial J$.

Averaging in systems with proper degeneration

One often encounters problems with *proper degeneracy*, in which the unperturbed Hamiltonian depends not on all the action variables and, correspondingly, some of the unperturbed frequencies are identically equal to zero: $H = H_0(I_1, \ldots, I_\ell) + \varepsilon H_1(I, \varphi, \varepsilon)$, $\ell < n$. The phases φ_j, $j > \ell$, are slow variables. According to the averaging principle, for approximate description of the evolution we must average the equations of the perturbed motion over the fast phases φ_i, $i \leq \ell$.

Theorem

In a Hamiltonian system with n degrees of freedom and ℓ frequencies, $\ell < n$, the variables conjugate to the fast phases are integrals of the averaged system.

Averaging in systems with proper degeneration

One often encounters problems with *proper degeneracy*, in which the unperturbed Hamiltonian depends not on all the action variables and, correspondingly, some of the unperturbed frequencies are identically equal to zero: $H = H_0(I_1, \ldots, I_\ell) + \varepsilon H_1(I, \varphi, \varepsilon)$, $\ell < n$. The phases φ_j, $j > \ell$, are slow variables. According to the averaging principle, for approximate description of the evolution we must average the equations of the perturbed motion over the fast phases φ_i, $i \leq \ell$.

Theorem

In a Hamiltonian system with n degrees of freedom and ℓ frequencies, $\ell < n$, the variables conjugate to the fast phases are integrals of the averaged system.

Example (Laplace-Lagrange theorem on the stability of the Solar System)

Consider the n-body problem under the assumption that the mass of one body (the Sun) is much larger than the masses of the other bodies (the planets). The unperturbed system is the one in which the planets do not interact with each other, and the Sun is at rest. The unperturbed system decomposes into $n-1$ Keplers problems. We suppose that the unperturbed orbits of the planets are Keplerian ellipses. In the problem under consideration there are $n-1$ fast phases – the mean longitudes of the planets. Their conjugate variables $\Lambda_j = \sqrt{\mu_j a_j}$, $j = 1, \ldots, n-1$, are integrals of the system averaged over the fast phases. Here the a_j are the major semiaxes of the Keplerian elliptic orbits of the planets, and the μ_j are factors depending on the masses.

Symplectic elimination of fast variables: Lindstedt's method

Consider a perturbed Hamiltonian $H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$ with n degrees of freedom. Suppose that the frequencies do not satisfy any identical resonance relations. We try to find a symplectic near-identity change of variables $I, \varphi \mapsto J, \psi$ so that the new Hamiltonian H depends only on the slow variables: $\mathcal{H} = \mathcal{H}(J, \varepsilon)$. We seek a generating function of the change of variables and the new Hamiltonian in the form of formal series in ε:

$$I = J + \varepsilon \partial S/\partial \varphi, \ \psi = \varphi + \varepsilon \partial S/\partial J,$$

$$S(J, \varphi, \varepsilon) = S_1(J, \varphi) + \varepsilon S_2(J, \varphi) + \dots, \ \mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J) + \dots.$$

The functions S_i must have period 2π in φ. The old and new Hamiltonians satisfy the relation

$$\mathcal{H}(J, \varepsilon) = H_0(J + \varepsilon \partial S/\partial \varphi) + \varepsilon H_1(J + \varepsilon \partial S/\partial \varphi, \varphi, \varepsilon)$$

Equating here the terms of the same order in ε we obtain the system of equations

$$\mathcal{H}_1(J) = \frac{\partial H_0}{\partial J}\frac{\partial S_1}{\partial \varphi} + H_1(J, \varphi, 0), \ \mathcal{H}_j(J) = \frac{\partial H_0}{\partial J}\frac{\partial S_j}{\partial \varphi} + F_J(J, \varphi)$$

The function F_j is a polynomial in $\partial S_1/\partial \varphi, \dots, \partial S_{j-1}/\partial \varphi$.

Symplectic elimination of fast variables: Lindstedt's method, continued

Take as $\mathcal{H}_1(J)$ the average of $H_1(J, \varphi, 0)$ over φ (i.e. average over all components of φ). Then

$$S_1 = -\sum_{k \neq 0} \frac{h_k}{i(k, \omega)} \exp(i(k, \psi)) + S_1^0(J), \quad \omega(J) = \frac{\partial H_0(J)}{\partial J}$$

where h_k, $k \in \mathbb{Z}^m$ are Fourier coefficients of function $H_1(J, \varphi, 0)$ and S_1^0 is an arbitrary function of J. It is assumed that the denominators in this formula do not vanish, and that the series converges and determines a smooth function. In a similar way we can determine all $\mathcal{H}_j$, S_j.

Symplectic elimination of fast variables: Lindstedt's method, continued

Take as $\mathcal{H}_1(J)$ the average of $H_1(J, \varphi, 0)$ over φ (i.e. average over all components of φ). Then

$$S_1 = -\sum_{k \neq 0} \frac{h_k}{i(k, \omega)} \exp(i(k, \psi)) + S_1^0(J), \quad \omega(J) = \frac{\partial H_0(J)}{\partial J}$$

where h_k, $k \in \mathbb{Z}^m$ are Fourier coefficients of function $H_1(J, \varphi, 0)$ and S_1^0 is an arbitrary function of J. It is assumed that the denominators in this formula do not vanish, and that the series converges and determines a smooth function. In a similar way we can determine all $\mathcal{H}_j$, S_j.

If the series for function S is truncated at the terms of order $r \geq 1$, then we obtain a change of variables which reduces the perturbed Hamiltonian to the form

$$\mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J) + \ldots + \varepsilon^r \mathcal{H}_r + \varepsilon^{r+1} \alpha(J, \psi, \varepsilon)$$

Discarding here the term of order ε^{r+1} we obtain a truncated system of the r-th approximation: $\dot{J} = 0$, $\dot{\psi} = \partial \mathcal{H} / \partial J$. So, in this approximation $J = \mathrm{const}$, which means that phase space is foliated by invariant tori. Motion on each torus is a rotation with constant frequency vector. These tori are are obtained by small deformation from the unperturbed tori $I = \mathrm{const}$. This deformation is described by formula $I = J + \varepsilon \partial S / \partial \varphi$.

Symplectic elimination of fast variables: Lindstedt's method, continued

Take as $\mathcal{H}_1(J)$ the average of $H_1(J, \varphi, 0)$ over φ (i.e. average over all components of φ). Then

$$S_1 = -\sum_{k \neq 0} \frac{h_k}{i(k, \omega)} \exp(i(k, \psi)) + S_1^0(J), \quad \omega(J) = \frac{\partial H_0(J)}{\partial J}$$

where h_k, $k \in \mathbb{Z}^m$ are Fourier coefficients of function $H_1(J, \varphi, 0)$ and S_1^0 is an arbitrary function of J. It is assumed that the denominators in this formula do not vanish, and that the series converges and determines a smooth function. In a similar way we can determine all $\mathcal{H}_j$, S_j.

If the series for function S is truncated at the terms of order $r \geq 1$, then we obtain a change of variables which reduces the perturbed Hamiltonian to the form

$$\mathcal{H}(J, \varepsilon) = H_0(J) + \varepsilon \mathcal{H}_1(J) + \ldots + \varepsilon^r \mathcal{H}_r + \varepsilon^{r+1} \alpha(J, \psi, \varepsilon)$$

Discarding here the term of order ε^{r+1} we obtain a truncated system of the r-th approximation: $\dot{J} = 0$, $\dot{\psi} = \partial \mathcal{H} / \partial J$. So, in this approximation $J = \mathrm{const}$, which means that phase space is foliated by invariant tori. Motion on each torus is a rotation with constant frequency vector. These tori are are obtained by small deformation from the unperturbed tori $I = \mathrm{const}$. This deformation is described by formula $I = J + \varepsilon \partial S / \partial \varphi$.

Kolmogorov-Arnold-Moser theory explains to which extent the above picture is valid.

Some remarks

1. For perturbed Hamiltonian systems with *proper degeneration* there exists procedure of symplectic elimination of fast phases. The "actions" canonically conjugate to these phases are first integrals of corresponding truncated system.

Some remarks

1. For perturbed Hamiltonian systems with *proper degeneration* there exists procedure of symplectic elimination of fast phases. The "actions" canonically conjugate to these phases are first integrals of corresponding truncated system.

2. If the unperturbed frequencies are close to resonance, then instead of Lindstedt's method von Zeipel's method is used. This is a procedure of symplectic elimination of fast phases from a Hamiltonian, which allows to keep in a Hamiltonian the dependence on the resonant phases. The resulting Hamiltonian system is not, in general, completely integrable, but has some additional integrals which are close to linear combinations (with integer coefficients) of original "actions".

Some remarks

1. For perturbed Hamiltonian systems with *proper degeneration* there exists procedure of symplectic elimination of fast phases. The "actions" canonically conjugate to these phases are first integrals of corresponding truncated system.

2. If the unperturbed frequencies are close to resonance, then instead of Lindstedt's method von Zeipel's method is used. This is a procedure of symplectic elimination of fast phases from a Hamiltonian, which allows to keep in a Hamiltonian the dependence on the resonant phases. The resulting Hamiltonian system is not, in general, completely integrable, but has some additional integrals which are close to linear combinations (with integer coefficients) of original "actions".

3. One often encounters problems in which the perturbation depends periodically also on time t. This case reduces to the one considered above by introducing the new phase φ_{n+1} and its conjugate variable I_{n+1}. The variation of the extended set of phase variables is described by a system of equations with the Hamiltonian

$$H = I_{n+1} + H_0(I_1, \ldots, I_n) + \varepsilon H_1(I_1, \ldots, I_n, \varphi_1, \ldots, \varphi_n, \varphi_{+1}, \varepsilon).$$

KOLMOGOROV-ARNOLD-MOSER (KAM) THEORY

Non-degeneracy conditions for the unperturbed motion

Consider an unperturbed integrable Hamiltonian system with Hamiltonian $H_0(I)$. Its phase space is foliated into the invariant tori $I = \mathrm{const}$. The motion on a torus is conditionally periodic with frequency vector $\omega(I) = \partial H_0/\partial I$. A torus on which the frequencies are rationally independent is said to be *non-resonant*. A trajectory fills such a torus everywhere densely (as one says, it is a winding of the torus). The other tori $I = \mathrm{const}$ are said to be *resonant*. They are foliated into invariant tori of lower dimension.

Non-degeneracy conditions for the unperturbed motion

Consider an unperturbed integrable Hamiltonian system with Hamiltonian $H_0(I)$. Its phase space is foliated into the invariant tori $I = \mathrm{const}$. The motion on a torus is conditionally periodic with frequency vector $\omega(I) = \partial H_0/\partial I$. A torus on which the frequencies are rationally independent is said to be *non-resonant*. A trajectory fills such a torus everywhere densely (as one says, it is a winding of the torus). The other tori $I = \mathrm{const}$ are said to be *resonant*. They are foliated into invariant tori of lower dimension.

The unperturbed system is said to be non-degenerate if its frequencies are functionally independent:

$$\det(\partial\omega/\partial I) = \det(\partial^2 H_0/\partial I^2) \neq 0$$

In a non-degenerate system the non-resonant tori form an everywhere dense set of full measure. The resonant tori form a set of measure zero, which, however, is also everywhere dense.

Non-degeneracy conditions for the unperturbed motion, continued

The unperturbed system is said to be *isoenergetically non-degenerate* if one of the frequencies does not vanish and the ratios of the other $n-1$ frequencies to it are functionally independent on the energy level $H_0 = \mathrm{const}$. The condition of isoenergetic non-degeneracy can be written in the form

$$\det \begin{pmatrix} \partial\omega/\partial I & \omega \\ \omega & 0 \end{pmatrix} = \det \begin{pmatrix} \partial^2 H_0/\partial I^2 & \partial H_0/\partial I \\ \partial H_0/\partial I & 0 \end{pmatrix} \neq 0$$

In an isoenergetically non-degenerate system both the set of non-resonant tori and the set of resonant tori are dense on each energy level; but, as above, the first set has full measure, whereas the second has measure zero.

Non-degeneracy conditions for the unperturbed motion, continued

The unperturbed system is said to be *isoenergetically non-degenerate* if one of the frequencies does not vanish and the ratios of the other $n-1$ frequencies to it are functionally independent on the energy level $H_0 = \mathrm{const}$. The condition of isoenergetic non-degeneracy can be written in the form

$$\det \begin{pmatrix} \partial\omega/\partial I & \omega \\ \omega & 0 \end{pmatrix} = \det \begin{pmatrix} \partial^2 H_0/\partial I^2 & \partial H_0/\partial I \\ \partial H_0/\partial I & 0 \end{pmatrix} \neq 0$$

In an isoenergetically non-degenerate system both the set of non-resonant tori and the set of resonant tori are dense on each energy level; but, as above, the first set has full measure, whereas the second has measure zero.

Remark

The system can be non-degenerate but not isoenergetically non-degenerate, and vice versa.

Invariant tori of the perturbed system

Now consider a perturbed system with Hamiltonian

$$H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$$

Invariant tori of the perturbed system

Now consider a perturbed system with Hamiltonian

$$H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$$

Theorem (Kolmogorovs theorem)

If the unperturbed Hamiltonian system is non-degenerate or isoenergetically non-degenerate, then under a sufficiently small Hamiltonian perturbation most of the non-resonant invariant tori do not disappear but are only slightly deformed, so that in the phase space of the perturbed system there also exist invariant tori filled everywhere densely with phase curves winding around them conditionally periodically with the number of frequencies equal to the number of degrees of freedom. These invariant tori form a majority in the sense that the measure of the complement of their union is small together with the perturbation. In the case of isoenergetic non-degeneracy the invariant tori form a majority on each energy-level manifold.

Invariant tori of the perturbed system

Now consider a perturbed system with Hamiltonian

$$H = H_0(I) + \varepsilon H_1(I, \varphi, \varepsilon)$$

Theorem (Kolmogorovs theorem)

If the unperturbed Hamiltonian system is non-degenerate or isoenergetically non-degenerate, then under a sufficiently small Hamiltonian perturbation most of the non-resonant invariant tori do not disappear but are only slightly deformed, so that in the phase space of the perturbed system there also exist invariant tori filled everywhere densely with phase curves winding around them conditionally periodically with the number of frequencies equal to the number of degrees of freedom. These invariant tori form a majority in the sense that the measure of the complement of their union is small together with the perturbation. In the case of isoenergetic non-degeneracy the invariant tori form a majority on each energy-level manifold.

The proof of the theorem is based on the converging procedure for eliminating the fast phases. The invariant tori constructed in this theorem are called Kolmogorov tori. The frequency vectors of the motions on the Kolmogorov tori belong to the Cantor set

$$\Xi_\varkappa = \{\xi\colon\ \xi \in \mathbb{R}^n,\ |(k,\xi)| > \varkappa |k|^{-\nu},\ k \in \mathbb{Z}^n,\ k \neq 0\}$$

where $\varkappa \sim \sqrt{\varepsilon},\ \nu > n-1$.

Absence of evolution in systems with two degrees of freedom

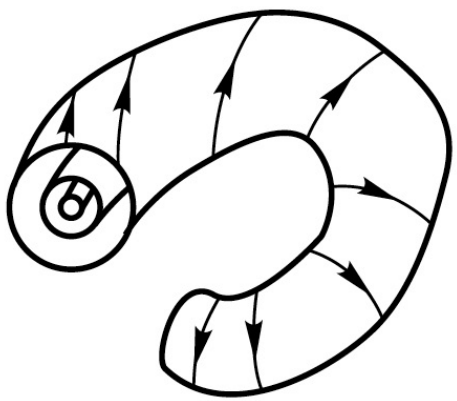
In systems with two degrees of freedom the existence of a large number of invariant tori implies the absence of evolution for all (and not just for most) initial conditions.

Theorem (V.I.Arnold)

In an isoenergetically non-degenerate system with two degrees of freedom, for all initial conditions the action variables remain forever near their initial values.

Proof.

In the system under consideration the phase space is four-dimensional, the energy levels are three-dimensional, and the Kolmogorov tori are two-dimensional and fill a larger part of each energy level. A two-dimensional torus divides a three-dimensional energy level. A phase curve starting in a gap between two invariant tori of the perturbed system remains forever trapped between these tori. The corresponding action variables remain forever near their initial values. □



Diffusion of slow variables in multidimensional systems and its exponential estimate

If the number of degrees of freedom n is greater than two, then the n-dimensional invariant tori do not divide a $(2n-1)$-dimensional energy level manifold, but are situated in it similar to points on a plane or lines in a space. In this case the gaps corresponding to different resonances are connected with each other. Therefore the invariant tori do not prevent a phase curve originated near a resonance going far away.

Diffusion of slow variables in multidimensional systems and its exponential estimate

If the number of degrees of freedom n is greater than two, then the n-dimensional invariant tori do not divide a $(2n-1)$-dimensional energy level manifold, but are situated in it similar to points on a plane or lines in a space. In this case the gaps corresponding to different resonances are connected with each other. Therefore the invariant tori do not prevent a phase curve originated near a resonance going far away.

Numerical experiments show that the evolution of the action variables apparently is not of directional nature, but is a more or less random walk along resonances around the invariant tori. This process is called *Arnold diffusion*.

Diffusion of slow variables in multidimensional systems and its exponential estimate

If the number of degrees of freedom n is greater than two, then the n-dimensional invariant tori do not divide a $(2n-1)$-dimensional energy level manifold, but are situated in it similar to points on a plane or lines in a space. In this case the gaps corresponding to different resonances are connected with each other. Therefore the invariant tori do not prevent a phase curve originated near a resonance going far away.

Numerical experiments show that the evolution of the action variables apparently is not of directional nature, but is a more or less random walk along resonances around the invariant tori. This process is called *Arnold diffusion*.

For generic systems diffusion happens exponentially slowly. The corresponding genericity condition is called *the steepness condition*. An analytic function is said to be steep if it has no stationary points and its restriction to any plane of any dimension has only isolated stationary points.

Theorem (Nekhoroshev.)

If the unperturbed Hamiltonian $H_0(I)$ is a steep function, then there exist a, b, c such that in the perturbed Hamiltonian system for a sufficiently small perturbation we have

$$|I(t) - I(0)| < \varepsilon^b \text{ for } 0 \leq t \leq exp(c/\varepsilon^a)$$

Here a, b, c are positive constants depending on the characteristics of the unperturbed Hamiltonian.

Invariant tori in the theory of small oscillations

KAM theory allows to construct invariant tori near equilibria and periodic trajectories in Hamiltonian systems, near fixed points of symplectic maps.

Invariant tori in the theory of small oscillations

KAM theory allows to construct invariant tori near equilibria and periodic trajectories in Hamiltonian systems, near fixed points of symplectic maps.

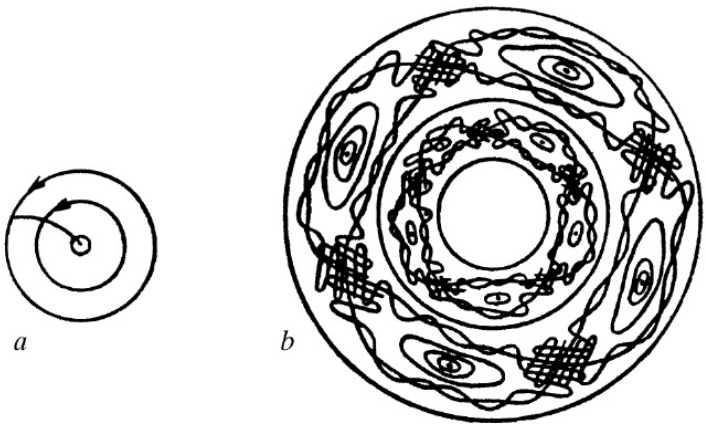
If there are no resonances up to 4-th order inclusively, then the normal form of the 4-th order plays a role of an unperturbed system, and the high order terms can be considered as a perturbation.

Invariant tori in the theory of small oscillations

KAM theory allows to construct invariant tori near equilibria and periodic trajectories in Hamiltonian systems, near fixed points of symplectic maps.

If there are no resonances up to 4-th order inclusively, then the normal form of the 4-th order plays a role of an unperturbed system, and the high order terms can be considered as a perturbation.

Here a Poincaré section near a periodic trajectory on an energy level for a Hamiltonian system with two degrees of freedom (or, a phase plane near a fixed point of a two-dimensional symplectic map) is shown ((a) - for normal form, (b) - for original system).

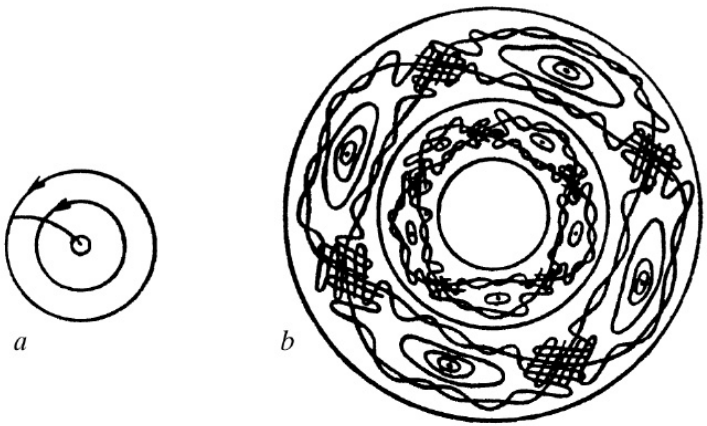


Invariant tori in the theory of small oscillations

KAM theory allows to construct invariant tori near equilibria and periodic trajectories in Hamiltonian systems, near fixed points of symplectic maps.

If there are no resonances up to 4-th order inclusively, then the normal form of the 4-th order plays a role of an unperturbed system, and the high order terms can be considered as a perturbation.

Here a Poincaré section near a periodic trajectory on an energy level for a Hamiltonian system with two degrees of freedom (or, a phase plane near a fixed point of a two-dimensional symplectic map) is shown ((a) - for normal form, (b) - for original system).



For Hamiltonian systems with two degrees of freedom existence of of big number of invariant tori implies stability of equilibria and periodic trajectories (correspondingly, for two-dimensional symplectic maps existence of big number of invariant curves implies stability of fixed points).