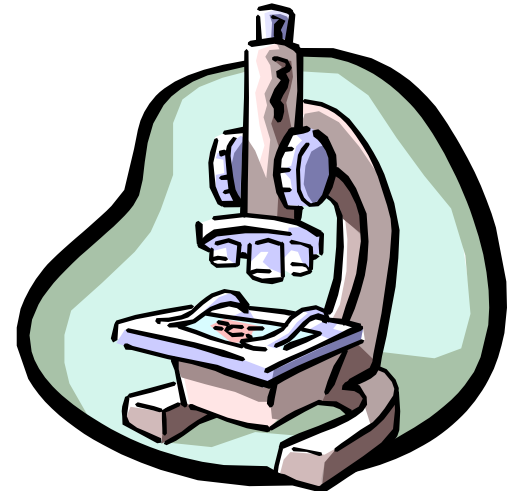


MATHEMATICAL AND COMPUTER MODELING IN SCIENCE AND ENGINEERING

Sergey Pankratov

**Lehrstuhl für Ingenieurwissenschaften in der
Informatik**



- **Introduction**
- **Part 1. Classical deterministic systems**
 - **Chapter 1. Methodological principles of modeling (5 -76)**
 - **Chapter 2. Mathematical methods for modeling (77 -132)**
 - **Chapter 3. Computational techniques (133 -172)**
 - **Chapter 4. Case studies (173 -334)**
 - **Literature (335 -338)**
- **Part 2. Quantum modeling**
- **Part 3. Stochastic modeling**
- **Appendices**

- This course contains examples of the application of selected mathematical methods, numerical techniques and computer algebra manipulations to mathematical models frequently encountered in various branches of science and engineering
- Most of the models have been imported from physics and applied for scientific and engineering problems.
- One may use the term “compumatical methods” stressing the merge of analytical and computer techniques.
- Some algorithms are presented in their general form, and the main programming languages to be used throughout the course are MATLAB and Maple.
- The dominant concept exploited throughout Part 1 of the course is the notion of local equilibrium. Mathematical modeling of complex systems far from equilibrium is mostly reduced to irreversible nonlinear equations. Being trained in such approach allows one to model a great lot of situations in science and engineering

Part 1

Classical deterministic systems

Chapter 1. Methodological Principles of Modeling

In this chapter basic modeling principles are outlined and types of models are briefly described. The relationship of mathematical modeling to other disciplines is stressed

Computer as a modeling tool

- The emphasis in this part is placed on dynamical systems. This is due to the fact that change is the most interesting aspect of models
- In modern mathematical modeling, various mathematical concepts and methods are used, e.g. differential equations, phase spaces and flows, manifolds, maps, tensor analysis and differential geometry, variational techniques, Lie groups, ergodic theory, stochastics, etc. In fact, all of them have grown from the practical needs and only recently acquired the high-brow axiomatic form that makes their direct application so difficult
- As far as computer modeling as a subset of mathematical modeling is concerned, it deals with another class of problems (algorithms, numerics, object-orientation, languages, etc.)
- The aim of this course is to bridge together computational methods and basic ideas proved to be fruitful in science and engineering

Training or education?

- One can probably design a good interactive program without any knowledge of analytical techniques or geometric transformations. However, for modeling problems, it would be difficult in such case to get outside the prescribed set of computer tricks. **This reflects a usual dichotomy between education and training**
- The present course has been produced with some educational purposes in mind, i.e. encompassing the material supplementary to computer-oriented training. For instance, discussion of such issues as symmetry or irreversibility is traditionally far from computer science, but needed for simulation of complex processes
- Possibly, not everything contained in the slides that follow will be immediately needed (e.g. for the exam). To facilitate sorting out the material serving the utilitarian purposes, the slides covering purely theoretical concepts are designated by the symbol §

Notes on exercises

- Exercises to the course are subdivided into two parts. The first part includes questions and problems needed to understand theoretical concepts. These questions/problems (**denoted c##**) are scattered throughout the text/slides. It is recommended not to skip them
- The second part consists of exercises largely independent from the course text/slides. They are usually more complicated than those encountered in the text. One can find both these problems and their solutions in the Web site corresponding to the course (**denoted ##**)
- Usually, exercises belonging to the second part represent some real situation to be modeled. Modeling can be performed in a variety of ways, with different aspects of the situation being emphasized and different levels of difficulty standing out. This is a typical case in mathematical modeling, implying that the situation to be modeled is not necessarily uniquely described in mathematical terms
- Use of computers is indispensable for the majority of exercises

What is mathematical modeling?

- Replacement of an object studied by its image – a mathematical model. A model is a simplification of reality, with irrelevant details being ignored
- In a mathematical model the explored system and its attributes are represented by mathematical variables, activities are represented by functions and relationships by equations
- Quasistatic models display the relationships between the system attributes close to equilibrium (e.g. the national economy models); dynamic models describe the variation of attributes as functions of time (e.g. spread of a disease)
- Modeling stages: 1) theoretical, 2) algorithmic, 3) software, 4) computer implementation, 5) interpretation of results
- Mathematical modeling is a synthetic discipline (superposition of mathematics, physics, computer science, engineering, biology, economics, sociology, ..). Enrichment due to interdisciplinarity

Basic types of mathematical models

- Qualitative vs. quantitative
- Discrete vs. continuous
- Analytical vs. numerical
- Deterministic vs. random
- Microscopic vs. macroscopic
- First principles vs. phenomenology
- In practice, these pure types are interpolated
- The ultimate difficulty: there cannot be a single recipe how to build a model
- For every phenomena – many possible levels of description
- One always has to make a choice: the art of modeling

See also: C. Zenger, Lectures on Scientific Computing,

http://www5.in.tum.de/lehre/vorlesungen/sci_comp/

Important principles of model building

- Models should not contradict fundamental laws of nature (e.g. the number of particles or mass should be conserved – the continuity equation)
- Testing (validation) of models against basic laws of nature
- Symmetry should be taken into account
- Scaling can be exploited to reduce the complexity
- To use analogies (e.g. chemical reactions and competition models)
- Universality: different objects are described by the same model (e.g. vibrations of the body of a car and the passage of signals through electrical filters)
- Hierarchical principle: each model may incorporate submodels – a step-like refinement
- Modularity and reusability
- From problem to method, not vice versa

Basic properties of mathematical models

- Causality: models based on dynamical systems are causal i.e. the effect cannot precede the cause and the response cannot appear before the input signal is applied. Causality is a result of human experience - non-causal systems would allow us to get the signals from the future or to influence the past.
- Causality is closely connected with time-reversal non-invariance (arrow of time). The time-invariance requires that direct and time-reversed processes should be identical and have equal probabilities. Most of mathematical models corresponding to real-life processes are time non-invariant (in distinction to mechanical models).
- A mathematical model is not uniquely determined by investigated object or situation. Selection of the model is dictated by accuracy requirements. Examples - a table: rectangular or not, ballistics: influence of the atmosphere, atom: finite dimensions of nucleus, military planning: regular armies (linear), partisans (nonlinear).

Basic models in physics

- Physical systems can be roughly separated into two classes: **particles and fields**, these are two basic models

Roughly – because there is an overlap, e.g. particles as field sources

- The main difference – in the number of degrees of freedom
- Any physical system consisting of a finite number N of particles has only finite number of degrees of freedom $n \leq 3N$

The number of degrees of freedom is defined as dimensionality of the configuration space of a physical system, e.g. a system with 2 degrees of freedom is $\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t)$ where \mathbf{F} is a plane vector field, $\mathbf{x} \in \mathbb{R}^2$

- The field is characterized by infinite number of degrees of freedom
- Classical dynamics is probably the most developed part of science, it studies the evolution of systems of material points – bodies that are so small that their inner structure is disregarded and the only characteristic is their position in space, $\mathbf{r} = \mathbf{r}_i(t)$

Theory, experiment and models

- What is the relationship between these three components of attaining physical knowledge?
- The interplay between theory and experiment induces the creation of models that are used to simulate the observations and predict new features that can be observed in new experiments
- Increasing complexity and expensive (sometimes prohibitively) experiments calls for simulation of both the theory and experiment
- Modeling gives the following results:
 - The theory is insufficient and must be modified, revised, improved
 - A new theory is needed
 - The accuracy of experiments is insufficient
 - New and better experiments are needed

One of the best examples of a model related to a theory is the Ernst Ising model of ferromagnetism

Physics: a collection of efficient mathematical models

- All physical laws are, in today's terminology, just mathematical models, although underlying ideas are not necessarily formulated in mathematical terms

- **10 worlds of physics:**

1. The classical world

2. Thermal world

3. Nonequilibrium world

4. Continuum world

5. Electromagnetic world

6. Plasma world

7. The quantum world

9. Relativistic world

8. High energy world

10. Cosmological world

Worlds of physics are just clusters of suitable models

The tree of mathematical modeling in physics, with branches, leaves and buds as individual models

There are links between “worlds” invoking substructures with repeatable, reusable patterns

- The Galilei's group (inertial systems)
- The Newton's law of motion (classical limit of special relativity and quantum mechanics)
- Newtonian gravity (classical limit of general relativity)
- The Kepler's problem (rotation of planets about the Sun)
- Potential fields, classical scattering
- Euler-Lagrange equations
- Variational schemes
- Noether's theorem and conservation laws , conservative systems
- Hamiltonian equations
- Hamilton-Jacobi equations
- Motion on manifolds, constraints
- The Liouville theorem

- Classical thermodynamics (equilibrium)
- The nature of heat, temperature, heat transfer
- Mechanical work and heat, interconversion, engines and cycles
- Heat capacity ($C=dQ/dT$)
- Laws of thermodynamics, thermodynamic potentials
- The concept of entropy, reversible and irreversible processes
- Entropy production
- Thermochemistry, chemical reactions
- Equation of state
- Phase transitions
- Heat as the particle motion, Maxwell distribution, statistical mechanics

Key figures: S.Carnot,
R.Clausius J.-B-J.Fourier,
J.W.Gibbs, J.P.Joule,
A.-L.Lavoisier,

- The Liouville equation, Gibbs distribution, open systems
- Kinetic equations, Boltzmann equation, Bogoliubov's hierarchy
- Diffusion, Langevin equation, Fokker-Planck equation
- Fluctuation-dissipation theorem (FDT)
- Linear response theory, Kubo formula, Onsager reciprocal relations
- Multiple scattering theories
- Classical stochastic models, nonlinear regime, branching and bifurcations, stability of nonequilibrium stationary states, attractors
- The Poincaré map, logistic model
- Dynamical chaos, indeterminism (impossibility of predictions)
- Dissipative structures, order through fluctuations, Turing structures
- Chiral symmetry breaking and life

- The Euler equation
- The Navier-Stokes equation
- Hyperbolic flow equations, shock and rarefaction waves
- Compressible gas dynamics and supersonic flows
- Self-similar models and explosions
- Turbulent flows and the models of turbulence
- Elastic solid models
- Viscoelasticity, plasticity, composites
- Seismic ray propagation and seismic ray theory
- Acoustics, sound wave/pulse excitation, propagation and scattering
- Detonation and flames, propagation of fires
- Superfluidity

- The Maxwell equations
- The Laplace and Poisson equations
- Interaction of electromagnetic (EM) fields with matter
- Electromagnetic response of material media
- Linear and nonlinear susceptibilities
- Linear and nonlinear optics
- Atoms and molecules in the EM field
- EM wave and pulse propagation
- Diffraction and scattering of EM waves
- Electromagnetic radiation
- Rays of light, asymptotic theories, coherence of light
- Photometry and colorimetry

- The plasma dielectric function, linear waves in plasma
- Screening in plasma, correlations of charged particles
- Hydrodynamic models of plasma
- Distribution functions
- Kinetic models of plasma: collision integrals of Boltzmann, Landau, Klimontovich, Lenard-Balescu, etc.
- Collisionless plasma, a self-consistent field model (the Vlasov equation)
- Plasma in external fields, the magnetized plasma
- Landau damping
- Theory of plasma instabilities
- Quasilinear and nonlinear models of plasma

Model hierarchies in physics

§

- The main ideas of model construction in physics are better grasped by discussing analytical models. Once we understand them, we can proceed to computational models and to specific cases, collectively named physical engineering
- In physics, there are always hierarchies of models with various degrees of generality. Most of the models created to describe physical phenomena are of the local character, produced within the framework of a more general model being applied to a restricted situation. **Example:** acoustics incorporate models that are specific cases of fluid dynamics and thermodynamics, under the assumption of small variations of pressure; shock pulses is the reverse case
- Models of heat (and mass) transfer were initially constructed as an independent area with its own *ad hoc* laws. Only after **J.W.Gibbs**, **J.C.Maxwell** and **L.Boltzmann** connected them to mechanics, the hierarchical structure of thermodynamic models became clear

Beyond physics

- The successful solution by Newton of the Kepler problem inspired thinkers and philosophers (e.g. Laplace) to develop the mechanistic model of the world.
- For centuries it was believed that all events are, in principle, predictable and can be described by differential equations similar to equations of motion.
- Seemingly unpredictable and irreversible phenomena, such as weather or human behavior, were believed unpredictable and irreversible only due to very large number of variables.
- It was hoped that with the advent of the powerful computers all long-range predictions would be possible (which is wrong).
- The biggest challenge of biology, medicine, society and economics is that randomness leads to fine-tuned processes (in time) and structures (in space). It means that the notion of the world as a machine is inadequate. In fact, fully mechanistic world would be incompatible with life (evolution – order through fluctuations)

New science: systems far from equilibrium

Classical science mostly studied systems and their states close to equilibrium. Such systems react on perturbations more or less predictably: they tend to return to equilibrium (evolve to a state that minimizes the free energy)

However, systems close to equilibrium can describe only a small fraction of phenomena in the surrounding world – in fact, it is a linear model. In contrast, non-equilibrium systems are ubiquitous in nature. Any system subject to a flow of energy and matter can be driven in the nonlinear mode, far from equilibrium (examples – open systems: the earth, living cell, public economy, social group)

Most of the processes are complex, interrelated, nonlinear, and irreversible. Often a tiny influence can produce a considerable effect. This is a typical feature of systems far from equilibrium. They can lose their stability and evolve to one of many states

To model the processes in the real world, one must learn how to describe systems far from equilibrium.

Nonlinear science

- Science in its contemporary form is predominantly nonlinear. This is clearly seen especially in modern interdisciplinary studies, e.g. in economics, ecology, financial mathematics, sociology, etc.
- Even modern physics started from nonlinear equations of motion – one can see it already from the Kepler problem, which contains typical features of nonlinear systems, e.g. dependence of the period on the amplitude and periodic orbits with many harmonics
- Three-body problem led to systematic study of nonlinear dynamics
- So, many-body problems must be, in principle, nonlinear. Indirect manifestation of this fact: existence of complex physical objects such as continuum media equations (Navier-Stokes, Burgers, nonlinear acoustics, gas flow, etc.) or Einstein gravitation equations
- The main feature: for typical nonlinear situations, it is impossible to make long-term predictions, even for slightly perturbed systems

Linear science

- Success of classical electrodynamics (pre-laser period), theory of oscillations, wave propagation, quantum mechanics and linear quantum field theory shifted the focus from nonlinear tasks, e.g. from the problem of dynamical system stability

- The main idea – linearization, the main principle – superposition:

$$L\Psi = L\sum_n a_n \Psi_n = \sum_n a_n L\Psi_n = \sum_n a_n \lambda_n \Psi_n$$

where L is the linear operator acting in some space containing Ψ_n

- A function $f(x)$ is linear if it satisfies the relationship

$$f\left(\sum_n \alpha_n x_n\right) = \sum_n \alpha_n f(x_n)$$

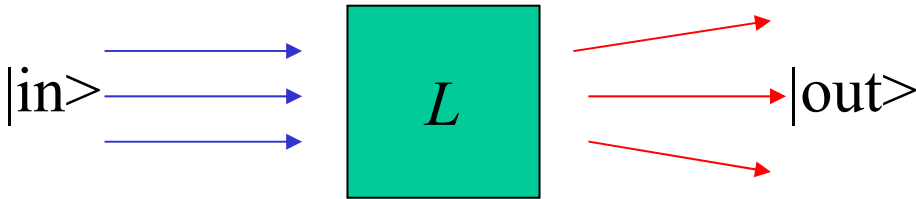
- **Exercise c1:** Is the function $f(z) = \sin z$ linear? The same for:
 $f(z) = \exp z$; $f(z) = \ln z$; $f(z) = \operatorname{arctg} z$; $f(z) = (1+az)^m$

How can one linearize these functions?

Are the following expressions linear? $\mathbf{ar}+c$; $b_i = m_{ij}x_j$; $\mathbf{F} = e[\mathbf{v},\mathbf{H}]$

Linear systems

- Linear transform: input \rightarrow output, linear map or linear response



Usually, time-invariant systems

- Integral transformations: Fourier, Laplace, Hankel, Weber, Hilbert (analytical signal), wavelets. Signal theory and image processing
- Classical linear system theory is mostly applied in electrical engineering e.g. for signal transmission, stationary noise reduction or removal, time-invariant filtering, predictive coding, etc.
- Classical optics is in fact a theory of linear systems and transforms
- Linear time-invariant operators: if the input $f(t)$ is delayed by τ , $f_\tau(t) = f(t - \tau)$, then the output is also delayed by τ :

$$g(t) = Lf(t) \Rightarrow g(t - \tau) = Lf_\tau(t)$$

This property is closely connected with energy conservation

Nonlinear trends in science and engineering

- All the fields started to develop their own nonlinear approaches
Examples: nonlinear optics, acoustics, radio-engineering, etc.
- The most versatile nonlinear medium is plasma. The study of plasma combines particle and fluid dynamics, electrodynamics, wave theory, stochastics, kinetics, stability and turbulence theories
- Simplification of nonlinear problems result in linearized versions

A majority of models should be unrestricted by linearity assumptions

- No simple unifying concepts analogous to vector spaces and linear operators exist in nonlinear theories
- However, new techniques emerged, specific for nonlinear science: approximate averaging (Krylov-Bogoliubov-Mitropolskii), KAM theory – conservation of invariants, dynamic entropy (Kolmogorov-Sinai). “Chaos” became the code word for nonlinear science
- **Use of computers is the key to modern nonlinear dynamics**

What equations are used for modeling

- As a rule, real systems and processes are described by nonlinear equations, e.g. differential equations with respect to time and spatial coordinates
- Such systems are distributed in space and correspond to an infinite number of degrees of freedom
- If the equations modeling the system do not contain spatial derivatives (ODE-based models), such a system is called point-like or having a null-dimension
- In the modeling with ODE, each degree of freedom is described by a second-order ODE. Differential equations of the first order correspond to $\frac{1}{2}$ degrees of freedom
- The equation $du/dt = f(u, t)$ gives an example of a dynamical system, i.e. the one, whose behavior is uniquely determined by its initial state (deterministic behavior)

Model simplification

- **How to simplify mathematical models?**
- A good way is to reduce the equations to the most primitive form, retaining the essence of the model
- **Usual tricks of the trade:**
 - disregarding small terms (in fact, power series)
 - using small or large parameters (as a rule - asymptotic expansions)
 - replacing geometrical forms by more symmetrical ones
 - substituting constants for functions (in fact, the average value theorem)
 - linearization
 - scaling
 - discretization and introduction of lattices (lattice models)
 - transition to dimensionless units (a specific case of scaling)

Ignoring spatial distribution of quantities,
e.g. transition to homogeneous (point) models,
 $\partial / \partial \mathbf{r} = 0$, leads to ODE instead of PDE

Application of dimensionless units

- Dimensionless form of mathematical models plays a special role: numerical values are independent of measurement units
- One can consider specific cases of the model not by comparing physical quantities, but by choosing numerical limits
- For instance, in numerical techniques being applied to dimensionless models, one can neglect terms that are small as compared to errors in other terms
- History: scaling and dimensionless combinations appeared first in the problems of heat and mass transfer – complex interactions of thermodynamical, fluid dynamical and electrodynamical processes
- Multiphysics processes, study of turbulence, multiphase system dynamics – scaling and dimension analysis is a powerful heuristical tool. Chernobyl heavy accident modeling – fluid with internal heat sources (computer code “Rasplav” and others)
- Modeling of the Earth and Venus climate (atmospheric circulation)

From Aristotle to Newton – evolution of models of motion

- Ancient thinkers: the motion of bodies is possible only in the presence of external forces produced by other bodies
- Aristotle (in the contemporary language): the state of a moving body (or particle) is described by three coordinates (x, y, z) changing under the influence of an external force:

$$\frac{d\mathbf{r}}{dt} = \mathbf{f}(\mathbf{r}), \quad \mathbf{r} = (x, y, z), \quad \mathbf{f} = (f_x, f_y, f_z)$$

- The difference of this first-order dynamical system with that described by the Newton equation is, primarily, in the irreversible character of the motion
- Aristotle's considerations were rooted in everyday experience: the trolley should be towed to be in motion; if the muscle force stop acting, the trolley comes to the rest
- **Even the most fundamental models are not unique!**

Was Aristotle always wrong?

- The trolley stops due to friction: $\mathbf{F}_R = -\alpha \mathbf{v}$
- this is the simplest model.

The Newton equations: $\frac{d\mathbf{r}}{dt} = \mathbf{v}, m \frac{d\mathbf{v}}{dt} = \mathbf{F} - \alpha \mathbf{v}$

In such model the state of a system would be given by positions alone – velocities could not be assigned freely

where \mathbf{F} is the towing force, m the mass of the trolley

In the case when \mathbf{F} is constant or slowly varies with time, the system is close to equilibrium, so the inertial term $m d\mathbf{v}/dt$ is small compared to other terms. The equilibrium solution:

$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{\mathbf{F}}{\alpha}$ has the Aristotle's form. Applicability area of such model: acceleration is negligible (almost uniform motion) and friction is substantially large,

$$|d\mathbf{v} / dt| \ll |\alpha \mathbf{v}|$$

For example, the Aristotle model would correspond to the Universe immersed in an infinite fluid with very low Reynolds number

Steady-state models

- The Aristotle model is, in fact, extensively used, an example is the Ohm's law: $\mathbf{j} = \sigma \mathbf{E}$ or $\mathbf{v} = \mathbf{E} / ne\rho$, where e is the electron charge, \mathbf{E} – electric field (acting force), n - the charge density, \mathbf{v} is the average velocity of the charge carriers.
- The Ohm's law is a typical macroscopic stationary model, when the driving force is compensated by resistance. Another example of such models is the steady-state traffic flow simulation (see Slide 000).
- Steady state models serve as a foundation of thermodynamics – the temperature may be defined only for equilibrium
- Steady-state models are typical of classical physics: in fact, it dealt only with slow and smooth motions, e.g. the planetary movement
- Models describing rapid and irreversible changes, resulting in multiple new states, appeared only in the XX century

Nonlinear events in the natural world

- It is widely said that the 20th century was the century of physics and the 21st one is the century of biology. The latter deals mostly with nonlinear phenomena, and the respective models should by necessity be nonlinear
- On a macroscopic scale, it has been illustrated by Lottka-Volterra model of the struggle for existence between two competing species. This dynamic situation (a predator-prey model) is described by nonlinear differential equations giving the time rate of evolution
- We are immersed in the natural world of nonlinear events. For instance, our emotional reactions – our likes and dislikes – are probably highly nonlinear. Our behavioral reactions to heat and cold, to colors and sounds, to local pressure and other stimuli are mostly subordinated to the Weber-Fechner law: the magnitude R of psychological response is proportional to logarithm of magnitude J of physical stimulus, $R = A \ln(J/J_0)$, $J \geq J_0$, $R = 0$ for $J < J_0$ (threshold)

Linearization

- The transition from $d\mathbf{x}/dt=\mathbf{F}(\mathbf{x})$ to $d\mathbf{y}/dt=A\mathbf{y}$ with $\mathbf{F}(\mathbf{x})=A\mathbf{x}+\mathbf{R}_2(x)$, $A=d\mathbf{F}(\mathbf{a})/d\mathbf{x}$
- Linearization is defined correctly: the operator A does not depend on the coordinate system
- The advantage of linearization: immediately obtained solution

$$y(t) = e^{At} y(0) = (I + At + A^2 t^2 / 2! + \dots) y(0)$$

- For any $T>0$ and any $\varepsilon>0$ there exists $\delta > 0$, so that if $|x(0)|<\delta$ then $|x(t)-y(t)|< \varepsilon\delta$ for all t , $0<t<T$. That is, for small deviations from the equilibrium $x=a$, the difference between the original and linearized systems $\mathbf{R}_2(x)$ is small compared to x : motions $x(t)$, $y(t)$ of the both systems with the initial condition $x(0)=y(0)=x_0$ are close to one another
- Linearization lies in the foundation of the oscillations theory: near the equilibrium point the behavior of the potential function coincides with that of its quadric part. Small variation of the potential only can shift the position of the equilibrium point

Dynamical systems

- The keyword – evolution. Let the evolution operator T^t transform some initial state of the system $P(t=0)=P_0$ into $P=P(t)$, $T^t P_0=P$
The dynamical system is defined as the one for which the evolution operator satisfies the relation: $T^t T^s = T^{t+s}$ (time is additive and the evolution operator is multiplicative)
- One more condition: $[T^t, T^s] = 0$ where $[.]$ denotes a commutator – evolution operators corresponding to different temporal intervals are commutative
- Definition of a dynamical system in terms of a differential operator allows one to generalize the representation of a dynamical system through specific equations (ODE, PDE, integro-differential, etc.)
- In fact, the dynamical system is equivalent to a Cauchy problem: $(x'(t) = f(t, x(t)), t \geq t_0 \geq 0; x(t_0) = x_0)$. A non-dynamical system is such, whose behavior is not uniquely determined by its initial conditions

The state of a dynamical system

- A simple example of a dynamical system is the mechanical system with a finite number n of freedoms (degrees of freedom).
- The state x_i , $i=1..n$ of a dynamical system is characterized by its configuration and by the latter's speed of change, which is called motion $\dot{x}_i = f_i(x_1..x_n)$, where $x=\{x_1..x_n\}$ can be interpreted as a point in \mathbb{R}^n
- Thus, the state of a dynamical system is represented geometrically as a moving point. The latter is called the phase point and a space of all such points is called the phase space. The change of states of a phase point may be represented as the motion of a phase point over a trajectory in the phase space. This motion is a one-parameter transformation $x(t) = g^t x$, $g^{t+s} = g^t \cdot g^s$ - a map of the phase space onto itself (a one-parameter group of diffeomorphisms). This group is also called a phase flux. The latter notion is of fundamental importance for many-body theories (statistical physics, kinetics)

The motivation – why bother?

- Dynamical systems contain geometric information in the form of a vector field. Interpretation of this geometrical information is the primary means of grasping evolutionary equations
Dynamics is the geometry of behavior
- Qualitative analysis of solutions of dynamical systems may provide better understanding than numerical and even analytical calculations
- Quantitative investigation of dynamical systems can be readily accompanied by visualization
- Asymptotic analysis and long-term numerical exploration: to find the maximal time step still ensuring the closure of the approximate solutions to the exact solution
- The exact solution exists (Cauchy-Peano theorem), is unique and continuous on rhs and initial conditions under general assumptions

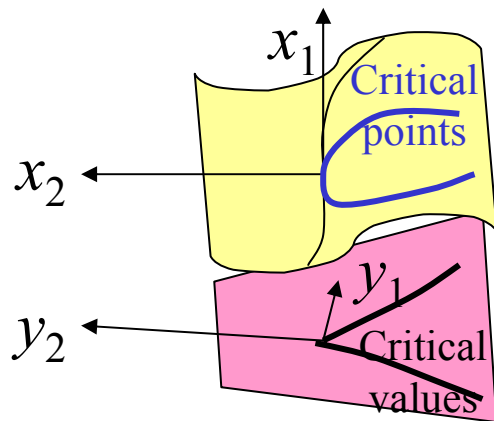
Changes and catastrophes

- Things change all the time, and mostly it is a normal behavior.
- In modern technology, engineers are trying to figure out typical patterns of evolutionary behavior both for individual system components and interaction (communication) links. For example: **if some system was operating and suddenly stopped working – it's because something has changed**
- To find out a problem, one must be capable to trace recent changes
- The user must be provided with data where to look for the problem
- Mathematical description of the world is based on interplay of continuous (smooth) transformations and discrete jumps.
- Theory of singularities (catastrophe theory) discusses critical point sets when several functions of several arguments are considered
- V.I.Arnold: **catastrophe - spasmodic change erupting in the form of a sudden response of a system to a smooth variation of external conditions**

Catastrophic behavior

- Catastrophe theory proved useful in applied science, e.g. biology, chemistry, macroscopic and laser physics, optics, even sociology – in those fields where bifurcations are observed. In such cases, catastrophe theory is an adequate language to describe nonlinearity
- Example (Whitney cusp): compare a smooth map of a plane (x_1, x_2)
 $y_1 = x_1^3 + x_1 x_2, y_2 = x_2$ with $y_1 = x_1^2 - x_2^2, y_2 = 2x_1 x_2$

The first one is structurally stable, the second structurally unstable (*structural stability* means that any perturbed map has the same singular points, at least locally – here near $(0,0)$):



$$y_1 = x_1^2 - x_2^2 + \alpha x_1, y_2 = 2x_1 x_2 - \alpha x_2$$

Equation for critical points:
$$\begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{vmatrix} = 0$$

In this case crit. points form a circle $x_1^2 + x_2^2 = \frac{\alpha^2}{4}$

Phase flux and phase fluid

- Evolution of a dynamical system is represented by the motion of a point in the phase space. Thus, a phase trajectory appears
- If the phase trajectory lies in the finite region, the motion is called finite, otherwise it is called infinite
- For physical systems, the phase flux (or phase flow) is usually given by the motion equations $\dot{q} = Q(q, p, t); \dot{p} = P(q, p, t)$, where q and p – vectors of coordinate and momentum. The solution $q = q(t, q_0, p_0), p = p(t, q_0, p_0)$ determines the phase trajectory.

In the operator form: $(q(t), p(t)) = T^t(q_0, p_0)$, the evolution operator T^t being a diffeomorphism. If we consider some finite region Γ_0 in the phase space and all the points inside this region as initial conditions, then we may think of these points as comprising some “phase fluid”

Phase flux makes the phase fluid flow, which means that $\Gamma_0 \Rightarrow \Gamma_t$

Invariants of motion

- The term “motion” for a dynamical system is understood in a generalized sense, e.g. evolution of a country’s economy is also motion of a bunch of phase points; competition of two interactive species is a motion, etc.
- If some symmetry can be observed during such motion, then the invariants of motion exist – quantities or combinations of variables that do not change in process of evolution. More accurate formulation of this fact is known as the Noether’s theorem:

If a dynamical (Lagrangian) system allows 1-parameter group of diffeomorphisms h^λ , then there exists the first integral of the system

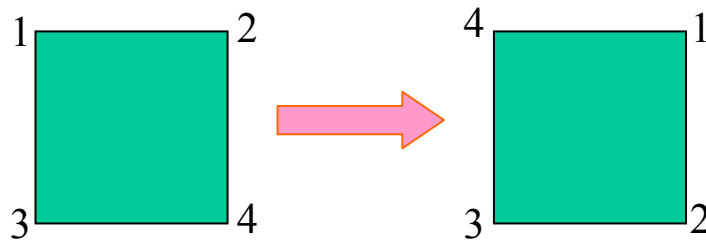
- To put it simply, one can construct the Lagrangian function and notice its symmetry properties, which immediately leads to the integrals of motion. All the conservation laws (energy, momentum, etc.) are just specific cases of the Noether’s theorem

Symmetry

- There are many ingenious techniques for obtaining efficient model representations, e.g. geometrical and graphical ones.
- The most efficient models have the common feature: they exploit symmetries of the objects. Roughly speaking, a symmetry of an object is a transform whose action leaves the object unchanged
- For instance, after a rotation by 120 degrees an equilateral triangle looks the same as before the rotation - this transformation is a symmetry. Rotations by 240 and 360 degrees are also symmetries
- Rotation by 360 degrees is equivalent to doing nothing: each point is mapped to itself - the trivial symmetry
- Symmetries are used to classify geometrical objects: e.g. the equilateral triangle has 6 symmetries (3 rotations and three flips - reflections with respect to 3 axes), an isosceles has 2 (trivial+flip), the general triangle has only the trivial symmetry

Symmetry and invariance

- Symmetry and invariance are closely related
- Symmetry is associated with operations on a system that transform it into itself – the transformed system should be indistinguishable from the one before transformation
- If the symmetry is perfect, then it should be experimentally impossible to distinguish initial and final states

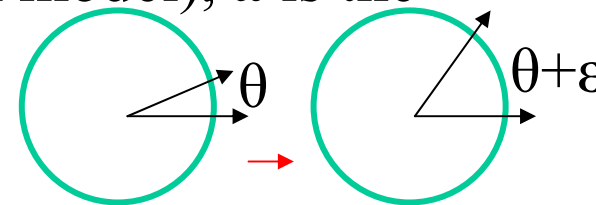


Rotation of a square through $\pi/4$

- The properties of the square are invariant under the rotation
- This transformation may be regarded as equivalent to a vertices permutation: $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 4, 4 \rightarrow 1$: there are 8 permutations that leave the square invariant – finite symmetry (8 group elements)

Symmetries as coordinate transformations

- Symmetry of a system is a set of transformations that leave the model of the system (e.g. expressed by differential equations) unchanged. In the simplest case, these transformations involve the system's coordinates – coordinate transformations
- For example, for an electron in crystal moving in the periodic field (this model is of fundamental importance for semiconductor theory and modern electronics), $U(\mathbf{r}) = \sum_a U(\mathbf{r} - \mathbf{r}_a)$, where \mathbf{r}_a denotes the position of atoms in crystal lattice. All the equations of any model describing the behavior of electron should be invariant under the transformation $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{n}\mathbf{a}$ (F. Bloch model), \mathbf{a} is the lattice period
- The unit circle $x^2 + y^2 = 1$ has a symmetry
 $\Gamma_\varepsilon: (x, y) \rightarrow (x', y') = (x \cos \varepsilon - y \sin \varepsilon, x \sin \varepsilon + y \cos \varepsilon), \varepsilon \in (-\pi, \pi]$



Or, in polar coordinates: $(\cos \theta, \sin \theta) \rightarrow (\cos(\theta + \varepsilon), \sin(\theta + \varepsilon))$

- Each symmetry has a unique inverse, which is itself a symmetry
- E.g. if G is a rotation of the equilateral triangle by 120 degrees, then G^{-1} is a rotation by 240 degrees
- The most interesting symmetries are smooth and invertible
- Smoothness: let x be the position of a general point of the object, if
$$G : x \rightarrow z(x)$$

is a symmetry, then z is assumed to be infinitely differentiable

- Since G^{-1} is also a symmetry, x is infinitely differentiable over z

This symmetry is called C^∞ -diffeomorphism, i.e. a smooth invertible mapping whose inverse is also smooth

- Symmetries in computer modeling, in distinction to computer graphics, are not required to preserve structure (e.g. for fluids)
- Symmetries of solids imply that the latter are of rigid material

Symmetry of equations

- Symmetry considerations is a crucial test demarcating physically-based modeling. What is the mathematical reason for it?
- All main equations of physics – basic mathematical models of Nature – (that of Newton, Euler-Lagrange, Hamilton-Jacobi, Laplace, D'Alembert, Helmholtz, Maxwell, Schrödinger, Klein-Gordon, Dirac, Einstein, etc.) can be classified according to their own symmetry
- Continuous symmetries vs. discrete symmetries: the former can imply arbitrarily small operations, e.g. we can rotate the sphere by an infinitesimal angle and nothing changes (see Slide 42)
- This fact leads to an entire branch of mathematics known as Lie groups (after a Norwegian mathematician Sophus Lie, 1842-1899)
- Lie theory - continuous group of transformations applied to DE
- The equations can be simplified by appropriate Lie transforms

Symmetry and conservation laws

- The meaning of Noether's theorem (Slide 43): all conservation laws are the consequence of some symmetry: a single-parameter symmetry group determines first integral of a dynamical system
- If the system can sustain even more symmetric transformations, then several integrals of motion arise
- Not all of these integrals are equally important: some are due to the fundamental properties of space-time, whereas others may be the consequence of the symmetry of specific model
- **Example:** energy conservation is due to time-invariance of closed physical systems – producing measurements over a system today, tomorrow or in T years should give the same result, differing only by measurement errors; laws of Nature are assumed to be constant
- If $\mathbf{r}(t)$ is the law of motion, then $\mathbf{r}(t+T)$ is the same law of motion: invariance under time shifts is a special case of Galilean invariance

How symmetry works

- A simple example: a system to be modeled is known to have a mirror symmetry, i.e. invariant under the transform $x \rightarrow -x$, x being some parameter, e.g. a coordinate. For instance, the butterfly is invariant under mirror reflection
- Then one can build a model disregarding the negative values of x
- In this case solutions (obtained by processing the model) must have a certain parity: they should be either even or odd with respect to $x \rightarrow -x$, i.e. expressed as even or odd functions of x
- If the solution does not satisfy this condition, then either the model is wrong or there is some computational error
- **Symmetry as a testing tool:** even this simple symmetry provides a powerful test for validation of the model
- The implication: symmetry of a modeled system should always be manifested in the solutions

Simplifications due to symmetry

- The example with mirror symmetry demonstrate how one can simplify modeling due to symmetry. In physical modeling, this can be seen on the formal level by constructing the Lagrangian and analyzing its invariance properties
- The resulting integrals of motion allow to easily obtain the solution
- **Typical examples:** what quantities should be conserved when a physical or engineering system has a symmetry of an infinite homogeneous cylinder (or can be roughly modeled by it)?

The answer: if the axis of the cylinder coincides with z , then z -components of momentum p_z and of angular momentum $m_z = x\partial_y - y\partial_x$ should be integrals of motion

What quantities are conserved in the field of an infinite plane (x,y) made of homogeneous material? **The answer:** from symmetry, we may conclude that p_x, p_y, m_z are integrals of motion

Simplifications due to symmetry-II

- What quantities should be conserved in the field of a cone, with its axis coinciding with z ? The field of two points located on z -axis?
The answer: m_z
- What quantities should be conserved if the system can be modeled by a homogeneous prism whose axis is parallel to z ? The answer: p_z as well as (possibly) a discrete analog of m_z
- Physics is always hunting for symmetry, because it is the powerful tool that allows physicists to reduce the seemingly kaleidoscopic world to a couple of dozens of fundamental models. Physicists are somewhat different from other people since they search and find new and more fundamental types of symmetry
- A curious fact: people start feeling and using symmetry long before they learn this word: night and day, a ball, left-right, flowers, a butterfly, a medieval cathedral, an Egyptian pyramid, etc.

Translational invariance

- A trivial example: a pencil (or a rod, a stick, a classroom pointer,...)
- These objects can be described by endpoint coordinates (x_1, y_1, z_1) , (x_2, y_2, z_2) in an arbitrary coordinate system (x, y, z) .
- Let us choose a Euclidean 3D orthogonal system. The pencil has a characteristic to be modeled – the length L :

$$L^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$$

- The spatial translations are given by $\mathbf{r}' = \mathbf{r} + \mathbf{a}$, where \mathbf{a} is a displacement vector. Substituting this expression into L gives $L' = L$, i.e. the length is invariant under spatial translations (**Exercise c2: prove it**)
- **Our model for the physical observable – the length of a pencil – should be translationally invariant**
- This example reveals a nontrivial fact about our world: in most cases, mathematical models should be expressed by equations invariant under arbitrary translations

Rotational invariance

- **Exercise c3:** prove that the rod length L is invariant under rotation $\mathbf{r}' = \mathbf{R}\mathbf{r}$, where \mathbf{R} is the rotation operator

This means that if the observer is using the rotated coordinates, the rod length for him will remain the same, and he will be working with the same mathematical model as an observer using initial coordinates

- This fact reflects a fundamental property of our space-time: isotropy, which means that laws of Nature are invariant under the group of rotations about all axes passing through a point (SO(3)group)
- Conservation of angular momentum is due to isotropy of our space, (likewise the conservation of momentum is the direct consequence of symmetry under 3-parameter group of spatial displacements)
- **The statement that physical laws should be invariant under spatial translations and rotations is rather strong: it is equivalent to the statement that all points in space are equivalent**

Symmetry of shapes and equations

§

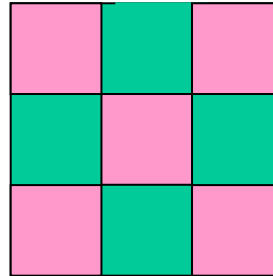
- **Example:** the perfect rotational invariance of a ball becomes a property of the gravitational law – there is no preferred direction in the equations of motion. Such model is the simplest one
- Instead of looking for the symmetry of shapes, one can analyze the laws governing the motions and configurations
- Symmetry of underlying equations does not necessarily result in the same symmetry of their solutions. **Example:** trajectories of the planets are ellipses, not circles – now we would call this effect a **spontaneously broken symmetry**
- Specific ways in which objects move and systems evolve are determined not only by differential equations, but also initial conditions – the former expressing the law and the latter the accidents in history. Therefore, symmetries may manifest themselves in the equations, but not necessarily in the solutions

Discrete and continuous symmetries

- Two distinct kinds of symmetries: **continuous** (e.g. with respect to rotations) and **discrete** (e.g. under parity transformation)
- Discrete symmetries have some **smallest steps** – operations that cannot be subdivided. For instance, a half of reflection $x \rightarrow -x$ is not a symmetry operation
- **In classical science**, mostly continuous symmetries were discussed – any system can be invariantly transformed “a little bit” (exception – point groups, crystallographic models)
- **In quantum science**, discrete symmetries are thoroughly studied – the quantum states are often classified with respect to discrete symmetries (e.g. lattice translations - Bloch functions, finite groups - molecular terms, time reversal – Kramers degeneracy, parity – space inversion, charge conjugation, CPT-invariance, particles - antiparticles, permutation symmetry - bosons and fermions, many-particle systems, quark models, supersymmetry, qubits, etc.)

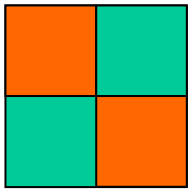
Finite symmetry transformations: decorations

- We can decorate a square:
- This figure still retains the full symmetry of the square



The square is symmetric with respect to rotation about the center, reflection in diagonals and bisectors. The square is also symmetric with respect to an inversion through its center

The figure below does not have the symmetry of a square:



- this figure does not transform into itself after a rotation by $\pi/4$ about its center

- We can extend the symmetry by introducing a more complicated transform: first rotation, then flip – turning orange into green and vice versa. Practical example: textile production (H.J.Woods, 1930 – “black and white groups” and “braids”); generalization – polychromatic groups; crystallography – Shubnikov groups; a physical example – magnetism (spin flip), L.D.Landau

The enforcement of symmetries

- If the coordinate frame, in which we try to describe the system to be modeled, is displaced, rotated or otherwise moved, the laws of the system evolution must remain the same
There is no preferred frame of reference
- **The consequence:** we must classify all the dynamical variables regarding their behavior under transformations (e.g. rotations) of the coordinate system in which these variables are expressed
 1. Those that remain unchanged are scalars
 2. Those that have components changing like Cartesian coordinates are vectors (there exist co- and contravariant vectors)
 3. Those that have components changing like ntuples – products – of Cartesian coordinates are tensors of rank n

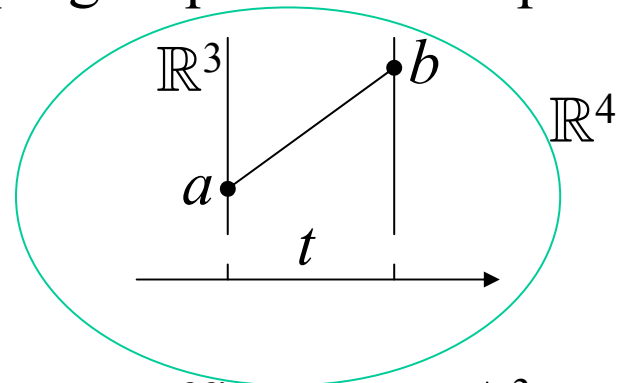
The rule: *never add quantities that do not transform in the same way*

Always ensure that both sides of equation transform equally

- The Galilei's space-time structure serves as fundamental symmetry of the classical world. It defines the class of so called inertial systems and includes the following components:
- 1) Our world – a 4D affine space \mathbb{A}^4 whose elements are events (called world points). Parallel transport of the world \mathbb{A}^4 gives 4D linear space \mathbb{R}^4

- 2) The classical time is a linear mapping of parallel transport of the world on real time axis: $t: \mathbb{R}^4 \rightarrow \mathbb{R}$

a, b are events, $t(a, b) = t(b - a)$ is a time Interval between events. If $t(a, b) = 0$, these events are simultaneous.



A set of simultaneous events constitutes 3D affine space \mathbb{A}^3

- 3) The distance between two simultaneous events: $\rho(a, b) = \|a - b\|$

The Galilean group §

- The distance defined through the scalar product converts \mathbb{A}^3 into 3D Euclidean space \mathbb{E}^3 , which we used to observe
- A group of all transformations preserving the Galilean structure is called the Galilean group whose elements are Galilean transforms – they are affine transformations preserving time intervals and distances between simultaneous events
- Three typical Galilean transforms are
 - a) $G_1(\mathbf{r}, t) = (\mathbf{r} + \mathbf{v}t, t)$ – motion with a constant velocity \mathbf{v}
 - b) $G_2(\mathbf{r}, t) = (\mathbf{r} + \mathbf{l}, t + s)$ – displacement of the coordinate origin
 - c) $G_3(\mathbf{r}, t) = (g\mathbf{r}, t)$ – rotation in the coordinate space \mathbb{R}^3 , $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is an orthogonal transformation, $g^T = g^{-1}$, $r'^i = G_{ij}r^j$

One can prove that any Galilean transformation can be uniquely presented as the product $G = G_1 G_2 G_3$, so the dimensionality of the Galilean group is $3 + 4 + 3 = 10$

Some consequences of the Galilean invariance

- If we differentiate the relation $a)$ (see the preceding slide) on time, we get the additive velocities: $\mathbf{v}=\mathbf{v}'+\mathbf{V}$, which means that the same point has different velocities in various coordinate systems – there exists neither absolute velocity nor absolute rest.

- However, acceleration does not depend on the coordinate system:

$$\mathbf{w}=\mathbf{w}'$$

- For a material point the Lagrangian – a function characterizing the state of a system – can depend only on \mathbf{r},\mathbf{v},t ; but due to Galilean invariance (uniformity of time, homogeneity of space, isotropy):

$$L = L(\mathbf{v}^2) = \alpha \mathbf{v}^2 := \frac{m}{2} \mathbf{v}^2$$

- The coefficient m is called the **mass** of a material point (particle).

Kinetic energy

- The concept of kinetic energy is associated with motion, it is the direct consequence of the Galilean invariance applied to a material point: all inertial systems are invariant under the Galilean group.
- If one can neglect interaction both between points and with the environment, the Lagrangian of a system of points becomes

$$L = \sum_i \frac{m_i \mathbf{v}_i^2}{2}$$

- This form is called the kinetic energy (usually denoted by T). In Cartesian coordinates and in inertial system, T is the quadric form over velocities in diagonal representation. In a general case:

$$q_i = q_i(x_k), \quad T = \frac{1}{2} a_{ik} q_i q_k$$

- Mass m is positive and the form T must be positively defined

Potential energy

- The concept of potential energy is associated with conservative forces (examples: gravity, electrostatics). For a system of N material points, the potential energy is the function of N variables

$$U(r_1, \dots, r_N)$$

- If we consider a closed system, which means that all the material points comprising the system are so far away from the rest of the world that they interact only with each other, then due to spatial homogeneity: $U(r_1, \dots, r_N) = U(\mathbf{r}_i - \mathbf{r}_j), i \neq j, i, j = 1, 2 \dots N$

This means that for a closed system the potential energy is a function of $(N-1)$ independent difference variables

- In most cases, interactions between elements of the system are independent (do not interfere), moreover, these interactions usually depend only on the distance between particles, which results in $U(r_1, \dots, r_N) = U(\mathbf{r}_i - \mathbf{r}_j) = \sum_{i \neq j}^N U(|\mathbf{r}_i - \mathbf{r}_j|)$

Potential fields

- A vector field is potential if and only if its work $A = \int_l \mathbf{f} d\mathbf{r}$

does not depend on the trajectory l (i.e. depends only on initial and final points). In such case, the function of the point \mathbf{x} is correctly defined, $U(\mathbf{x}) = -\int_{\mathbf{x}_0}^{\mathbf{x}} \mathbf{f} d\mathbf{r}$ and $\mathbf{f} = -\partial U / \partial \mathbf{x}$

The field is potential if and only if its work over the closed path vanishes. Another equivalent criterion of potential field is $\text{curl } \mathbf{f} = \mathbf{0}$. For example, all central fields, i.e. invariant under all motions leaving the center intact, are potential. This is true for Euclidean space of any dimensionality n . The Newtonian gravitational field and the Coulomb electrical field are potential ones.

Energy conservation

- A theorem: the total energy of a potential system is conserved in process of the system's evolution, i.e. $E(t)=E(t_0)$, $E=T+U$
- Proof: 1) $\frac{dT}{dt} = \sum_{i=1}^N (\dot{\mathbf{r}}_i, m_i \ddot{\mathbf{r}}_i) = \sum_{i=1}^N (\dot{\mathbf{r}}_i, \mathbf{f}_i)$
- 2) $T(t) - T(t_0) = \int_{\mathbf{r}(t_0)}^{\mathbf{r}(t)} \mathbf{f} d\mathbf{r} = U(\mathbf{r}(t_0)) - U(\mathbf{r}(t))$
- Here $\mathbf{r}=(\mathbf{r}_1..\mathbf{r}_N)$ may be interpreted as a radius in the configuration space of N points, $\mathbb{R}^{3N} = \mathbb{R}^3 \times \dots \times \mathbb{R}^3$; $\mathbf{f}=(\mathbf{f}_1...\mathbf{f}_N)$ is $3N$ -dimensional force. The quantity $L = T - U$ is the system's Lagrangian
- In Lagrangian terms, energy conservation law can be written as

$$E = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L = \text{const}$$

where q_i are coordinates and $p_i = \partial L / \partial \dot{q}_i$ are corresponding momenta.

The Hamiltonian function: $H(p, q, t) = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L$, $p = [p_i]$, $q = [q_i]$

Closed and open systems

For a closed system, it is assumed that all its parts do not interact with anything that is not included in the system (e.g. the system in question is far away from all other bodies)

For open systems, it is difficult to make any general statements. An important specific case, when the system in question interacts only with another one whose motion may be considered given (that is independent on the first system), is called **the motion in the external field**. The Earth is moving in the external field of the Sun, so far as the influence of the Earth on the Sun's motion is neglected. Thermodynamics, heat transfer, statistical and condensed matter theories are those that study many-body open systems.

The concept of **the temperature** is related to open systems – it describes the equilibrium of the given system with a thermostat.

Irreversibility is a characteristic feature of many-body open systems.

Systems in external field

- Example: if the system in external field consists of one material point, then its Lagrangian takes the form
$$L = \frac{1}{2} m \mathbf{v}^2 - U(\mathbf{r}, t),$$

The Lagrangian function is defined as $L=T-U$

where $U(\mathbf{r}, t)$ is the potential energy in an external field.

The most popular models are built according to the pattern of a system moving in an external field: Newton, Kepler – motion of celestial bodies; F. Bloch – electrons in crystals; plasma physics; chemical bonding; condensed matter response to electromagnetic (e.g. laser) radiation; interaction of radiation with biological objects, e.g. microwave radiation emitted by mobile devices, etc.

An important case: motion in a central field

Astrophysics and atomic physics are based on central fields

$$\ddot{\mathbf{r}} = -\frac{\partial U}{\partial \mathbf{r}}, U(\mathbf{r}) = U(r)$$

Applying the conservation laws, motion in a central field can be fully explored. The Coulomb field case $U(r)=const/r$ can be solved exactly

Usefulness of conservation laws

- The origin of conservation laws lies in the symmetry principles, but their applications are rather mundane and universal.
- For example, description of fluid flows, of heat and mass transfer, of shock, detonation, and rarefaction waves are based on phenomenological representation of conservation laws.
- Modern numerical models tend to exploit conservation laws to produce new algorithms and to test the obtained results.
- The fluid-dynamic models of heavy traffic on a road network are also grounded on conservation laws (see Case Studies).
- Thus, despite their fundamental microscopic origin, the conservation laws can be stated and utilized entirely in terms of macroscopic variables. This fact makes the conservation laws very instrumental in building mathematical models, since it is usually hardly possible to trace the evolution of all the system's elements

Hyperbolic conservation laws

- For numerical modeling, especially in fluid dynamics, hyperbolic systems of conservation laws are especially useful.
- In macroscopic modeling, e.g. in computational fluid dynamics, hyperbolic of conservation laws are represented as time-dependent systems of PDEs:

$$\frac{\partial u^i(\mathbf{r}, t)}{\partial t} + \frac{\partial f^{ij}(\mathbf{u}(\mathbf{r}, t))}{\partial x^j} = 0$$

where \mathbf{u} is a conserved quantity, which may be interpreted as a vector whose components are intensive state variables. The term *intensive* signifies that these variables do not depend on a system as a whole – they represent the corresponding densities. The tensor function f^{ij} is called a flux or a flow function

The above equation must be supplemented by initial conditions and – in a bounded region – boundary conditions, see e.g. *R. LeVeque. Numerical Methods for Conservation Laws, Birkhäuser, 1990*

Conservation laws as balance equations

- Hyperbolic conservation laws are usually produced by considering the temporal behavior of the integral
$$I(t) = \int_{(V)} d^3r u(\mathbf{r}, t)$$
 taken over the whole volume occupied by a system.

Here, for simplicity, a scalar case is presented.

The quantity $I(t)$ is called extensive, since its value depends on the system's volume. Thus, $u(\mathbf{r}, t)$ may be represented by a functional derivative:

$$u(\mathbf{r}, t) = \frac{\delta I}{\delta V}$$

The change of $I(t)$ is due to two factors: 1) production of I inside the volume V , 2) migration of the density u through the boundary (flux):

$$\frac{\partial I}{\partial t} = \int_{(V)} \rho_0(\mathbf{r}) d^3r + \int_{(\delta V)} f^i(\mathbf{r}) d\sigma^i$$
. If there are no sources within the system ($\rho_0 = 0$), then the balance leads to $\partial_t u + \partial_i f^i(u; \mathbf{r}, t) = 0$

– a scalar conservation law for u .

Conservation laws in CFD

- In the computational fluid dynamics the hyperbolic system of conservation laws may be written as

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{f}^i(\mathbf{U})}{\partial x^i} = \mathbf{Q}, \quad t \in (0, T), \mathbf{r} = \{x^i\} \in \mathbb{R}^3, i = 1, 2, 3,$$

$$\mathbf{U} = \begin{pmatrix} \rho \\ j^1 \\ j^2 \\ j^3 \\ \rho E \end{pmatrix}, \mathbf{f}^i = \begin{pmatrix} j^i \\ v^i j^1 + p \delta_1^i \\ v^i j^2 + p \delta_2^i \\ v^i j^3 + p \delta_3^i \\ j^i H \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} 0 \\ \rho f_M^1 \\ \rho f_M^2 \\ \rho f_M^3 \\ \rho q + \rho v^k f_M^k \end{pmatrix}$$

where $j^i = \rho v^i$ is the mass current, q is thermal source density, p – pressure, E – energy, $H = E + p/\rho$ – enthalpy and index M signifies mass forces. This is a general form of Euler equations for an ideal fluid.

They incorporate mass, momentum and energy conservation

Irreversibility

- **Why are the future and the past so different?**
- The problem of irreversibility consists in the fact that the Hamiltonian systems (underlying most mathematical models of the world) are time-reversible both at microscopic level (motion equations for individual degrees of freedom), and at macroscopic level (equations for phase flux density), whereas real processes are almost always irreversible
- From the mathematical viewpoint, time-invariance means that each solution remains a solution after we change t by $-t$. It is clear that the same function cannot satisfy both time-invariant and non-invariant equation.
- The first example describing an irreversible process – a privileged direction of time – was Fourier's law of heat propagation.
- If we wish to preserve the Hamiltonian approach, we have to pinpoint the place where the irreversibility has been introduced (explicitly or implicitly).

The arrow of time

§

- The unidirectional flow of time is one of the greatest puzzles, although it appears an obvious feature both of our consciousness and of physical/engineering/technological world
- There are major disagreements between scientists about how to treat this problem, which is usually denoted as “the arrow of time”
- In fact, there are at least **five arrows of time**:
 - 1) **The biological or cognitive arrow** – psychological awareness of passing time, pointing to the future
 - 2) **The causal arrow** – delay between cause and effect
 - 3) **The thermodynamic arrow** - the entropy growth
 - 4) **The cosmological arrow** – expansion of the Universe
 - 5) **The relativistic arrow** – time cannot be reversed by physically realizable Lorentz transformation, in distinction to space

The laws of unpredictability

- Classical science whose manifest was Newtonian Principia treated the world as a gigantic mechanism created by the God's design. The Universe was looking as a perfect automat where there was no place for randomness, and if the accident still trapped a person, it was owing to his/her mistakes, ignorance or negligence.
- *The Laplace Demon*: hypothetical creature that was able to instantly compute all the trajectories of all particles of the world; hence this creature would hold all the connections between past and future (including present and future). **Roles:** God is the main designer and lawmaker; the Laplace Demon is the God's secretary (knows all but can't change anything). A perfect computer.
- So the world in the classical rationalistic model is, in principle, fully predictable, and if someone could establish good relationships with the Demon secretary, all the past and future states of the world would be no secret any more.
- Gambling, football championships, lucky strikes would cease to exist; futurology and descriptive history would be unnecessary; astrologists, fortune tellers, etc. would lose their lucrative businesses.
- However experience shows (sometimes quite painfully) that Nature is more inclined to an unpredictable than to a regular behavior of an eternal automat.

The laws of unpredictability-II

- The weather cataclysms, unforeseen economic collapses, unpredictable social upheavals – all this was observed and experienced by huge masses of people
- Such observations did not testify to the Laplace determinism. One could say, nonetheless, that if one knows the distribution of mass in the dice, all forces acting on it from the outer world, its initial position, the velocity of the hand throwing the dice; if then one could integrate the motion equations on a super-computer, then *probably* the classical *probability theory* would be redundant
- Why should we talk about chances when everything can be computed?
- Even very simple systems (e.g. a couple of billiard balls) behave unpredictably. One cannot get rid of uncertainty, even if enormous data have been collected and processed. *Unpredictability is indispensable*
- Please note: only classical – non-quantum systems are considered. In quantum models uncertainty enters from the very beginning – it is the starting point of the theory and not the fact to be explained. Einstein could not accept it: “A real witchcraft calculus”. And: “God does not throw dice”
- Is classical mechanics, allowing unpredictability, also becoming the witchcraft calculus? What are the modern rules for determining future by the present?

The laws of unpredictability-III

- The unpredictability laws are the evolution equations of the dynamical systems theory (Newtonian equations of motion are the particular case of such theory)
- The behavior of the objects described by evolution equations may become totally unpredictable in some time. **Example:** the Earth atmosphere – a typical dynamical system described by deterministic evolution equations, however it is impossible to foresee its state, say, in a month's time, no matter how powerful a computer would be used. The long-term weather forecast is always frustrating
- **Two types of randomness:** 1) too many degrees of freedom: particles, events, objects – practically impossible to reckon (theoretically possible but useless); this is the statistical description. **Example:** gas in a can, 10^{23} molecules. Even if some fantastic supercomputer could integrate the corresponding system of equations, it would be practically impossible to insert all the initial conditions (think of the time and e.g. quantity of paper needed to record them)
- 2) deterministic chaos: transition to randomness due to sensitive dependence on initial conditions – experimentally or computationally indistinguishable initial states eventually evolve to states that are far apart (in the phase space). Chaotic dynamics bridges regular evolution of complex systems with the random one

The Bradbury butterfly

- In the story “A Sound of Thunder” by Ray Bradbury, a famous science fiction writer, an outstanding example of the dynamical chaos is given: the story tells of the 21st century people who learned how to travel back and forth in time. So a bunch of young men went to the Mesozoic era (as though to New York suburbs) to hunt dinosaurs
- There was, however, one rule that time-travelers should strictly observe: they were forbidden to leave a certain path laid in some other dimension and hence exerting no influence on the world evolution. But one of the hunters scared by a formidable dinosaur slipped off the trail and stomped on a butterfly
- When the travelers returned to their time, they discovered that a different political regime is established in their country, different orthography is accepted and in general something happened that they could not foresee. A negligible in the world scale event – a butterfly’s death – led to unpredictable consequences. **Close trajectories of the evolving world diverged, so the situation became unpredictable.** This unpredictable change occurs owing to instability
- Dynamical chaos and the related unpredictability arises out of the exponential sensitivity to initial conditions (or other parameters)

Scaling

- If the world population could be reduced to that of a village numbering 100 people, with all the current proportions among the Earth inhabitants remaining intact, then such a village would have the following approximate proportions:
 - 8 Africans
 - 14 Americans (including North, Central and South America)
 - 21 Europeans
 - 57 Asians
 - 52 women and 48 men
 - 30 whites and 70 non-whites
 - 30 Christians and 70 non-Christians
 - 89 heterosexuals and 11 homosexuals
 - 6 persons would possess 60% of the whole accumulated wealth

70 illiterates; 80 in shabby houses;
50 undernourished
1 is dying; 1 has a computer;
1 is the University graduate

Scaling in work

- This example demonstrates how scaling works – it allows you to immediately see the proportions and make use of them (e.g. here you may readily see that you are in much better position than many million other people)
- In financial politics: transformation of French Franc (1 to 100), Russian Ruble (1 to 10 in 1961 and 1 to 1000 in 1994) - change of scale due to inflation. The shape and size of the coin are preserved but its value is scaled – diminished. During the medieval gold coinage period, when the size of a coin manifested its value, such scaling would involve the change of coin dimensions
- The heat content of a sphere $\sim r^3$; heat radiated from a sphere $\sim r^2$ Thus (heat loss/heat content) $\sim r^{-1}$ – the larger the object the smaller the ratio. **The biological implication:** babies are expected to be more susceptible to temperature changes than adults; **socio-economic implication:** better insulating clothes must be provided

Scaling in science and engineering

- Well-known examples of scaling (power law) relationships:
 - Shock wave radius after a nuclear explosion (see below in detail).
 - Scaling law for the breathing rate of animals, $R=AM^\gamma$, R is the mass of oxygen absorbed per unit time, M is the body mass, A is proportional to absorptive capacity of respiratory organ, $\gamma=n/3$ is a scaling exponent, n is the dimensionality of the organ ($n=1,2,3$). Other examples of scaling in biology are given in Case Studies.
 - Ecology: spreading of contaminations, e.g. of liquid waste mound in a porous medium
 - Turbulence-related phenomena, shear flows, bubbles and bursts (a flow whose mean properties do not depend on the coordinate along the average velocity is called a *shear flow*)
 - Computational geophysics, atmospheric phenomena, plausible explanation of “flying saucers” – long-lived pancake-form patches

Chapter 2. Mathematical Methods for Modeling

Selected mathematical methods are described in this chapter, with the emphasis being placed on differential equations, both ODE and PDE

ODE-representation of dynamical systems

- Many problems (not only in mechanics and physics) lead to models described by ODE of the second order:

$$d^2x / dt^2 - f(x, dx / dt, t) = 0$$

- If one puts $dx/dt=v$, then one has a system of two equations of the first order
 $dx / dt = v, dv / dt = f(x, v, t)$
- In most cases, it is easier to analyze this system than the equation of motion. A more general case of the dynamical system is

$$dx / dt = F(x, v, t), dv / dt = G(x, v, t)$$

Here $F(x, v, t)$ does not necessarily coincide with v . In case variable t (interpreted as time) is absent, the dynamical system is called autonomous. The system may be autonomous but non-conservative (e.g. damped oscillator). An important class of dynamical systems is Hamiltonian systems (the first integral – the energy – does exist)

Autonomous systems

- The general form of continuous-time autonomous system is $d\mathbf{x}/dt = \mathbf{v}(\mathbf{x})$, where $\mathbf{x} = (x_1 \dots x_n)$; $\mathbf{v} = (v_1 \dots v_n)$

Any system can be reduced to autonomous if we increase the number of unknown functions (e.g. putting $t = x_{n+1}$) then instead of

$$d\mathbf{x}/dt = \mathbf{v}(t, \mathbf{x})$$

we get an autonomous system with $(n+1)$ variables:

$$d\mathbf{x}/dt = \mathbf{v}(\mathbf{x}, x_{n+1}), \quad dx_{n+1}/dt = 1$$

- If $\mathbf{x} = \mathbf{f}(t)$ is the solution of autonomous system, then it can be represented by a parametric curve in the space \mathbb{R}^n that is called the phase space of the system. Parametric curves representing solutions are called phase trajectories. Two phase trajectories either don't have joint points, or coincide. To characterize the system's behavior, it is enough to know only some special phase curves (e.g. equilibrium)

Properties of autonomous systems

- Phase space of an autonomous system is segmented by nonintersecting trajectories. This is not true for non-autonomous systems.
- The point \mathbf{a} is called the equilibrium point of an autonomous system if $\mathbf{v}(\mathbf{a})=0$. The equilibrium point is the solution, $\mathbf{x}(t)=\mathbf{a}$, of the autonomous system and, hence, the point $\mathbf{x}=\mathbf{a}$ is the phase trajectory. However, a phase trajectory not reduced to a point is a smooth curve, i.e. it has a non-zero tangent vector in each point.
- Phase trajectories determine a vector field $\mathbf{v}(\mathbf{x})$. Equilibrium points are critical (singular, fixed) points of $\mathbf{v}(\mathbf{x})$. The set of all possible trajectories is called the phase flow.
- The function $u(\mathbf{x})$ is called the integral of motion if $u = \text{const}$ along the phase trajectory, $u(\mathbf{x}(t)) = \text{const}$ for all t .
- For autonomous systems, solutions depend on the time and on the initial point t_0 only through differences $(t-t_0)$ – such systems are invariant under time translations

First-order autonomous systems

- The simplest dynamical systems are those whose state can be represented by a single variable x
- This variable may be interpreted as a coordinate in an abstract space also called **phase space**
- Examples of processes that can be represented by such models:
 - radioactive decay
 - a body falling in a viscous fluid
 - chemical reactions
 - reproductive behavior, models of population growth
 - metabolism and evolution
 - electrical networks (e.g. discharge of a capacitor C through a resistance R), etc.
 - econophysics, models of economic growth
 - transition to turbulent behavior

First-order autonomous systems-II

- The motion equation $dx/dt=v(x,t)$ has the solution:

$$t - t_0 = \int_{x_0}^x \frac{dx}{v(x)}$$

(if the integral exists). The inverse of $t(x)$ gives $x=x(t-t_0)$.

- Autonomous systems are invariant under time shifts – for conservative systems this leads to energy conservation.
- Phase space of first-order systems is not necessarily a real line.

E.g. for rotation about an axis the phase space is a circle, $\theta \in [-\pi, \pi]$

The velocity v must be 2π -periodic, $v(\theta + 2\pi n) = v(\theta)$

- If the motion is periodic, the period is given by:
$$T = \int_{-\pi}^{\pi} \frac{d\theta}{v(\theta)}$$

In fact, most of 1D problems can be solved by this method

First-order autonomous systems are extremely important because in many cases one can reduce a second-order system to a first-order one

How to evaluate the behavior of first-order dynamical system without solving ODE?

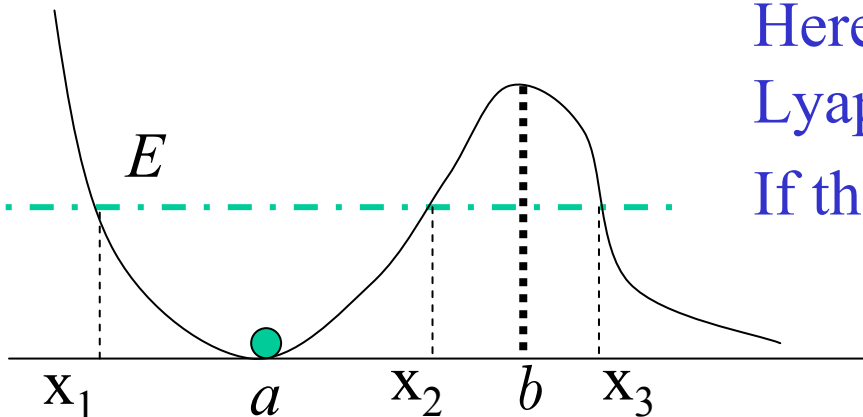
- The behavior of $x(t)$ (in autonomous case $x(t-t_0)$) is controlled by zeroes of the velocity function $v(x)=0$, $x=x_m$ (equilibrium or fixed points). If the system is initially at x_m , it remains there all the time. In all other initial points the state of the system must change.
- Open intervals between two equilibrium (fixed) points, together with intervals that go from fixed points to infinity, are invariant sets, since a system whose evolution started between two fixed points cannot pass either of them.
- There are two types of equilibrium points: stable and unstable. The equilibrium position a is stable when one can find the system near the equilibrium during the whole its evolution, provided it was sufficiently close to equilibrium at the initial moment (this is called Lyapunov stability)
- Asymptotic methods: expansion $x(t,\varepsilon)=x_0(t)+\varepsilon x_1(t)+\dots \varepsilon^n x_n(t) + R_{n+1}$

The Lyapunov stability

- A solution $x=\varphi_0(t)$, $y=\psi_0(t)$ is called to have Lyapunov stability if for any $\varepsilon>0$ one can find $\delta=\delta(\varepsilon)>0$, so that for all solutions $x=\varphi(t)$, $y=\psi(t)$ satisfying $|\varphi_0(t_0)-\varphi(t_0)|<\delta$, $|\psi_0(t_0)-\psi(t_0)|<\delta$, for all $t>t_0$ holds:
 $|\varphi_0(t)-\varphi(t)|<\varepsilon$, $|\psi_0(t)-\psi(t)|<\varepsilon$
- If the solution possess the Lyapunov stability and if for small $\delta>0$ the following conditions are satisfied

$$\lim_{t \rightarrow \infty} |\varphi_0(t) - \varphi(t)| = 0, \lim_{t \rightarrow \infty} |\psi_0(t) - \psi(t)| = 0,$$

then the solution $\varphi_0(t)$, $\psi_0(t)$ is said to be asymptotically stable



Here the equilibrium position a is Lyapunov-stable if there is no friction. If there is friction, then the ball motion will contract with time, i.e. the equilibrium a is asymptotically stable. Equilibrium b is unstable

Rough classification of equilibrium points

- There are only three basic kinds of equilibrium points of autonomous dynamical systems:
 - sinks – nearby solutions converge to the equilibrium point
 - sources – nearby solutions diverge
 - nodes – any other behavior

Let a be an equilibrium point for $dx/dt = f(x)$, then

- If $f'(a) < 0$, then a is a sink
- If $f'(a) > 0$, then a is a source
- If $f'(a) = 0$, then the equilibrium is indeterminate (node)

Example: the logistic population model has a source at $x=0$ and a sink at $x=1$, which means that any nonzero initial population eventually comes to the normalized value $x=1$

Exercise c4: find and classify equilibrium points of a pendulum

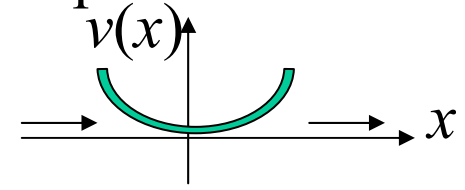
Stability of first-order dynamical systems

- A simple stability test is as follows: a fixed point is stable if the velocity $v(x)$ is decreasing when $x \rightarrow x_m$ i.e. $dv(x_m)/dx < 0$
- One can approximate velocity near a stable equilibrium point with linear function $v(x) = -b(x - x_m)$, so that the evolution here is described by $x(t) = x_m + (x_0 - x_m)e^{-bt}$ where $x_0 = x(0)$

This solution shows explicitly that the trajectory $x(t)$ does not reach the fixed point in finite time.

Structurally unstable points: $x = 0$, $v(x) = kx^2, k > 0$

If $x > 0$ then the system moves away from the equilibrium point and if $x < 0$ then $x(t)$ is moving towards it. The equilibrium (fixed) point is neither stable nor unstable: small perturbation may radically change evolution of the system.



Second-order dynamical systems

- Autonomous first-order systems have relatively simple solutions that are determined by fixed (equilibrium) points
- Most dynamical systems are more complicated than first-order systems. A time-dependent first-order dynamical system, with velocity function $v(x,t)$, can be treated as an autonomous second-order system by defining a new variable $y=t$ to produce a new equation of motion: $dx/dt=v(x,y)$; $dy/dt=1$
- The state of second-order dynamical system is represented by two real variables (x,y) that may be considered as coordinates of the system in its phase space, or by a single vector $\mathbf{r}=(x,y)$.

Exercise c5: Find the solution of the second-order system $dx/dt=-y$; $dy/dt=x$ with initial conditions $\mathbf{r}(0)=(a,b)$. Eliminate the time to find an equation for phase curves. Can the same expression be obtained from $dy/dx=-x/y$. What solution brings more information? What situations can be described by this system?

Nonlinear analysis – main ideas

- The study of stability of an autonomous system is based on the stability analysis of linear approximation. For a second-order:

$$\frac{dx}{dt} = P(x, y) \rightarrow ax + by + p(x, y), \frac{dy}{dt} = Q(x, y) \rightarrow cx + dy + q(x, y)$$

- Here p and q represent the nonlinear terms. By transforming the variables linearly, we can eliminate constant terms
- Looking for critical points: $P(x, y)=0$, $Q(x, y)=0$, transferring the origin of coordinates into a critical point (near the equilibrium)
- A linear system is obtained by suppressing p and q terms

$$\frac{dX}{dt} = AX, X = \begin{pmatrix} x \\ y \end{pmatrix}, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then the usual linear stability theory is applied: $\det|A - \lambda E| = 0$,
 λ is the characteristic exponent dependent on the problem parameters

Equilibrium states of second-order systems

- Conservation laws do not hold when one must regard friction (or amplification), so sometimes dynamical systems with dissipation are called dissipative, to distinguish them from Hamiltonian ones
- Analysis of dynamical systems with two variables

$$\frac{dx}{dt} = P(x, y), \quad \frac{dy}{dt} = Q(x, y) \quad (1)$$

starts from exploring equilibrium points (x_0, y_0) : $P=Q=0$

Near the equilibrium we can linearize the system $(x - x_0 \rightarrow x, y - y_0 \rightarrow y)$

$$\frac{dx}{dt} = ax + by, \quad \frac{dy}{dt} = cx + dy \quad (2)$$

The characteristic equation: $\lambda^2 - T\lambda + \Delta = 0$, $\lambda_{1,2} = (1/2)[T \pm (T^2 - 4\Delta)^{1/2}]$, $T = a + d$ is a trace and $\Delta = ad - bc$ is a determinant. The behavior of the system (phase trajectories) depends on signs of $T, \Delta, D := T^2 - 4\Delta$

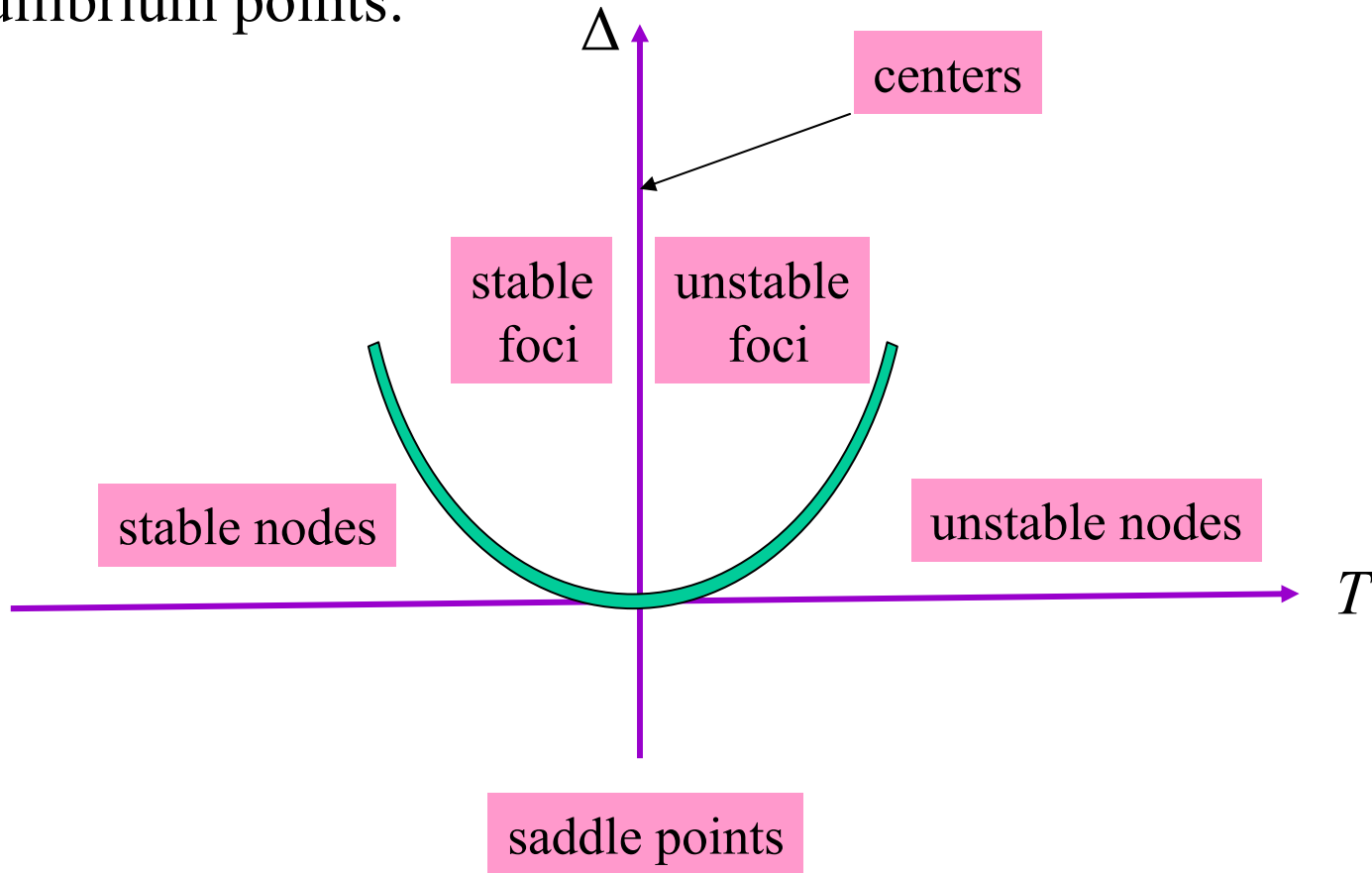
The meaning of this analysis is that one can determine the behavior of the given nonlinear dynamical system judging by its linearized version. The topological structure of phase curves near the equilibrium is the same for both systems

Classification of equilibrium points

1. Roots of the characteristic equation are different, real and have the same sign, $0 < \Delta < (T/2)^2$, $\lambda_1 \lambda_2 > 0$: the solution of (2) is parabolic.
If $\lambda_{1,2} > 0$, trajectories strive to infinity; if $\lambda_{1,2} < 0$, they tend to zero:
 - 1.1. $T < 0$: asymptotically stable node (attractor)
 - 1.2. $T > 0$: asymptotically unstable node (repeller)
2. Roots are real and confluent, $\lambda_1 = \lambda_2 \neq 0$: a degenerate node
3. Roots are real different, with different signs ($D > 0, \Delta < 0, \lambda_1 \lambda_2 < 0$): hyperbolic solutions, a saddle point
4. Complex conjugated roots, $\text{Re } \lambda_{1,2} \neq 0$ ($D < 0, T \neq 0$)
 - 4.1. $T < 0$: stable node (attractor)
 - 4.2. $T > 0$: unstable node (repeller)
5. Imaginary roots, $\text{Re } \lambda_{1,2} = 0$ ($D > 0, T = 0$): center (e.g. no dissipation)
6. One of the roots is zero ($\Delta = 0$): node (e.g. very large dissipation)

Equilibrium point diagram

- The graphical representation summarizes classification of equilibrium points:



Two special cases for equilibrium points §

1. *If the equilibrium point for the linearized system (2) is a center, then for the starting nonlinear system (1) it is either center or focus.*

The meaning of this fact is that geometric properties of the phase portrait are conserved in the process of linearization, except in the neighborhood of a center of the linearized system.

The reason for the center to present a special case:

For a center of a linearized system $T=0$, and any small perturbation, e.g. disregarded nonlinear terms or infinitely small dissipation, can make T nonzero and hence turn ellipses into spirals.

2. *If at least one of the roots of the linearized system (2) is zero, then a special investigation is needed to analyze the equilibrium point of the starting nonlinear system. Two different time scales may arise.*

Differential equations

- Ordinary (ODE) and partial (PDE) – the principal means to describe changing phenomena
- As a rule, real systems are described by nonlinear PDEs that are quite difficult to solve analytically. Hence, the following approaches are employed:
 - linearization
 - reduction of the distributed system to null-dimensional (and consequently PDE to ODE)
 - qualitative analysis (qualitative theory of ODE)
 - symmetry methods (Lie symmetries theory)
 - numerical methods

Literature: V.I. Arnold. Ordinary Differential Equations, MIT Press;
E.L. Ince. Ordinary Differential Equations, Dover Publ.

Linear differential equations

- The general linear differential equation is of the type:
 $Ly(x) = f(x)$ where L is the differential operator of n -th order

$$L = p_0(x) \frac{d^n}{dx^n} + p_1(x) \frac{d^{n-1}}{dx^{n-1}} + \dots + p_n(x)$$

It is usually assumed that the coefficients p_i are continuous and single-valued and p_0 does not vanish throughout the interval (a, b)

If y_i $i=1..m$ is a solution of the homogeneous equation ($f=0$), then
 $y = \sum_{i=1}^m C_i y_i(x)$ is also a solution, C_i being arbitrary constants

If n linearly independent solutions y_i , $i=1..n$ of the homogeneous equation are known, then $y = \sum_{i=1}^n C_i y_i(x)$ is its complete solution

Linear differential equations (continued)

- Let $y = u$ be any solution of the inhomogeneous equation and y_0 is the complete solution of the homogeneous equation, then the general solution of the inhomogeneous equation is $y = y_0 + u$

Example: $\frac{d^2 y}{dx^2} + y = kx$ The particular integral is $u = kx$ and the general solution is: $y = A \cos x + B \sin x + kx$

- The Wronskian.** Let y_i be n linearly independent solutions of the homogeneous equation of the n -th order. Then it is impossible to find non-zero constants C_i to satisfy the following condition identically in the interval (a,b) :
$$\sum_{i=1}^n C_i y_i(x) = 0$$

Differentiating $n-1$ times we get n equations to determine the constants C_i , and these equations are consistent if the determinant (Wronskian)

$$\Delta = \left\| y_i^{(j)} \right\| \neq 0, i = 1..n; j = 0..n-1$$

The fundamental system of solutions

- Any linearly independent set of n solutions of the equation $Lu = 0$ forms a fundamental system of solutions. The condition that any given set of solutions forms a fundamental system is that the Wronskian of these solutions is not zero
- There is an infinite number of possible fundamental sets of solutions
- If the Wronskian of n solutions vanishes at any point of the interval (a,b) these solutions are linearly dependent
- If the Wronskian of $k < n$ solutions vanishes identically in (a,b) these k solutions are linearly dependent

A pathological example: two functions $y_1 = x^3$ and $y_2 = |x|^3$ are linearly independent in $-1 < x < 1$ but their Wronskian $W(x) = y_1 dy_2/dx - y_2 dy_1/dx$ is identically zero

Solution of an inhomogeneous linear system

- Solution of an inhomogeneous linear system can be always obtained if the solution of the corresponding homogeneous system is known, e.g. with the help of variation of constants
- If $X(t)$ is the fundamental system of solutions of the homogeneous part, then the inhomogeneous solution is:

$$x(t) = X(t) \int_{t_0}^t X^{-1}(t') f(t') dt'$$

- The general solution for initial problem $x(t_0)=x_0$ is given by

$$x(t) = X(t) X^{-1}(t_0) x_0 + X(t) \int_{t_0}^t X^{-1}(t') f(t') dt'$$

- This solution can be generalized on space-dependent (e.g. periodic) input or control vector $f=f(x,t)$. In this case the solution may be obtained by iterative treatment of the resulting integral equation

Second-order linear differential equations

- If the equation $F(x, y, y', y'') = 0$ depends on the unknown function y and its derivatives linearly, then

$$a(x)y'' + b(x)y' + c(x)y = f(x)$$

The most important case is the autonomous system, when the factors a, b, c do not depend explicitly on x . Then we get the oscillator equation

$$(y := q, x := t, a := m, b := h, c := k) \quad m\ddot{q} + h\dot{q} + kq = f(t)$$

The general solution: $y = C_1 y_1(x) + C_2 y_2(x)$ or $q(t) = C_1 q_1(t) + C_2 q_2(t)$

How can one find two independent solutions? Ansatz $e^{\lambda t}$

The characteristic equation: $m\lambda^2 + h\lambda + k = 0$

Its solution
$$\lambda_{1,2} = \frac{-h \pm \sqrt{h^2 - 4mk}}{2m}$$

$$q(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

The general solution of the homogeneous equation

Nonlinear analysis

Principle of superposing solutions to obtain the general solution does not hold, so well developed linear response techniques fail

Usually it is very difficult not only to obtain a general solution, but also to write an exact particular solution

One has to resort to approximations using slow and rapid variables, small and large parameters. Perturbation methods are widely used

Nonlinear systems reveal properties unknown in linear analysis.

An example – nonlinear oscillator: restoring, damping and driving forces interact and cannot be studied separately. Typically, everything is mixed up, independent actions are an exception

Typical sources of nonlinearity: 1) the state of the object changes during the modeled process (e.g. saturation); 2) fundamental laws of nature are basically nonlinear; 3) complex geometry; 4) in biology, social and behavioral science response is usually not proportional to stimulus; 5) influence of the environment (e.g. viscosity, friction)

Adequate modeling techniques is the dynamical systems approach

Method of isoclines (a reminder)

- This geometric method applied to $dy/dx=f(x,y)$ consists in drawing curves in the plane where $f(x,y)=\text{const}$ – isoclines are curves along which the derivatives of the integral is constant. By varying this constant, we obtain the family of isoclines
- A short line segment can be drawn at each point along isocline, the segment slope is equal to the constant; thus a family of parallel short segments is produced and a gradual turn of slope is shown
- Such plot gives an idea how to draw integral curves with gradually turning tangents. This method indicates the large-scale behavior
- Method of isoclines can be improved if the second derivative is used to calculate the curvature radius at (x,y) :
$$R = \frac{(1 + y'^2)^{3/2}}{y''}$$

From here the center of curvature

is obtained and arcs are drawn
producing integral curves

$$\left(x - \frac{y'R}{(1 + y'^2)^{1/2}}, y + \frac{R}{(1 + y'^2)^{1/2}} \right)$$

PDE-based modeling

- Many important problems can be described by equations

$$a(x) \frac{\partial^2 u}{\partial x^2} + b(x) \frac{\partial u}{\partial x} + c(x)u = \alpha \frac{\partial u}{\partial t} + \beta \frac{\partial^2 u}{\partial t^2}$$

where α, β are constants

- This model includes Laplace, Schrödinger, Fokker-Planck, wave and diffusion equations. There are no general techniques available to construct **exact** solutions to such equations
- Approximate procedures , e.g. perturbation methods or short wavelength asymptotics (slowly varying coefficients; separation of evolution into slow and fast components) can be used.
- However, these techniques **always have limited validity regions**
- Symmetry and scaling are especially important for PDE-based modeling. Each type of equation should be invariant under certain transformations (has its own symmetry group)
- Applications of PDE-based modeling range from standard (e.g. heat and mass transfer) to rather exotic (land mine detection)**

PDE-based modeling – the description of media

- If a particle moves along the x -axis, then its motion is described by the equation $x=x(t)$. If the dynamical system consists of several particles, then $x_i=x_i(t)$, $i=1,2,..$
- But how to describe the motion if the number of particles N is very large, e.g. nearing the Avogadro number?
- It is useless to store particle numbers in such situation, much wiser is to characterize the whole array by continuous parameters, e.g. density, temperature, pressure, force and velocity fields
- Transition to continuous variables implies the replacement of a discrete system of a large number of constituents by a field (e.g. number of cars by their density or chain of masses by a string)
- In numerical methods, the reverse transition usually takes place
- Cooperative phenomena and collective motion can be correctly described only by PDE-based models

Partial differential equations

- The unknown function depends on several variables
- Distributed systems, infinite number of degrees of freedom
- Physical phenomena relate space and time derivatives
- Scalar (e.g. acoustics), vector (e.g. electrodynamics, hydrodynamics), and tensor (e.g. general relativity) PDEs
- PDEs describe not the evolution of a separate object, but the changing state of the medium
- Mathematical models of the medium are, in general, non-linear – it is necessary to know how to solve nonlinear PDEs
- Many important nonlinear PDEs admit linearizing transformations (the latter can usually be found by symmetry methods)
- Over the recent several years, symmetry-finding packages have been developed for computer algebra systems

How to solve PDEs?

0. PDE should be correctly formulated from the physical, economical, biological, engineering, etc. point of view (constructing the mathematical model)

Most of the natural laws are stated in terms of PDEs, e.g. Maxwell equations,, Schrödinger equation, Dirac equation, Einstein equations, Navier-Stokes equations, Newton's cooling law, etc.

1. Initial and boundary conditions are very important – in fact, they manifest external influence on the studied system
2. There is an arsenal of methods, the most important are those that convert PDEs into ODEs
3. Ten useful techniques:
 - 1) Separation of variables. This technique reduces a PDE with respect to n variables to n ODEs

Ten useful techniques to solve PDEs

- 2) Change of coordinate variables (rotation, scaling, affine transformation, etc.). This technique converts the original PDE into an easier PDE or into ODE
- 3) Transformation of the unknown function. Transition to a new function that is easier to find. Main fields of application: quantum mechanics, hydrodynamics, nonlinear wave theory
- 4) Integral transformations. This technique reduces the number of independent variable usually making the solution procedure much easier. The most important integral transformations are Fourier, Laplace, and Hankel transforms
- 5) Perturbation methods. This technique reduces a complex problem, e.g. a nonlinear one, to a series of much simpler or already solved problems. The main techniques in quantum mechanics and laser physics (nonlinear optics)

Ten useful techniques to solve PDEs-II

6) Eigenfunction expansion. This technique seeks the PDE solution as an infinite series of eigenfunctions with weighting coefficients. The eigenfunctions can be found by solving the eigenvalue problem corresponding to the original problem. This technique may be interpreted as bordering to integral transformations method (on the one hand) and to the Green's functions method (on the other). One of the main techniques in the classical heat transfer theory and in orthodox quantum mechanics, e.g. atom physics

7) The Green's functions method. The meaning of this technique is the decomposition of the input, e.g. source term or boundary conditions, into very short impulses (delta-functions) and finding the response to each impulse – the Green's function. The solution to the problem is represented as the total response – the sum or integral over elementary responses. One of the most powerful techniques in modern physics. Applied mainly to linear problems

Ten useful techniques to solve PDEs-III

8) Method of integral equations. This technique converts a PDE into an integral equation that can be solved by methods specific for integral equations (e.g. by Born iterations or numerically). The method of integral equations is especially important for electrical and electronic engineering

9) Calculus of variations. This technique seeks the solution of a PDE problem by reformulating it as a variational (extremal) problem. The fact is that the minimum of a certain functional (usually the latter represents the energy of the system) gives the solution to the PDE problem – usually it is called the weak form of the solution. The variational technique is close to the method of finite elements – one of the most important numerical methods of solving PDEs

Ten useful techniques to solve PDEs – numerical methods

10) Numerical methods. This technique represents approximately a PDE by a system of finite difference equations that can be solved on a computer, usually by iterations. Besides, there exist methods that tend to approximate solutions by polynomial expansions or representations (e.g. splines)

Some newly developed numerical methods based on hierarchical representations proved to be very efficient. These latter methods are using special classes of hierarchical basis functions. An example of the modern hierarchical techniques is the wavelet transformation

Hierarchy is one of the most important concepts in computer science (informatics). One of the pioneers of hierarchical techniques in numerical mathematics is Chr. Zenger

Classification of PDEs

- The general classification table

Linearity	Linear				Nonlinear		
Order	1	2	3	4	5	...	m
Coefficients	Constant				Variable		
Dimensionality ($n-1$, n = number of variables)	1	2	3	4	5	...	$n-1$
Homogeneity	Homogeneous				Inhomogeneous		
Basic type (for linear equations)	Elliptic			Parabolic		Hyperbolic	

Famous PDE-based models:

the Maxwell equations

- In vacuum:

$$\operatorname{rot} \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{j}; \operatorname{div} \mathbf{E} = 4\pi\rho$$

$$\operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0; \operatorname{div} \mathbf{H} = 0$$
- In material media:

$$\operatorname{rot} \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c} \mathbf{j}^e; \operatorname{div} \mathbf{E} = 4\pi\rho^e$$

$$\operatorname{rot} \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0; \operatorname{div} \mathbf{H} = 0$$

Here

$$\mathbf{P}(\mathbf{r}, t) = \int_{-\infty}^t \mathbf{j}(\mathbf{r}, t') dt', \mathbf{j} = \frac{\partial \mathbf{P}}{\partial t} \quad \mathbf{D}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) + 4\pi\mathbf{P}(\mathbf{r}, t)$$

The whole classical electrodynamics and optics can be derived from these equations. The Maxwell equations are closely connected with special relativity and can be written in covariant 4D form. One can easily obtain from them the wave equation describing EM waves

The Schrödinger equation

- In the nonrelativistic quantum theory the state of a system is fully described by a wave function Ψ . Similarly to dynamical systems theory, knowledge of this function at initial moment t_0 determines the system's behavior in all future moments. This evolution is governed by the Schrödinger equation:

$$i\hbar\partial_t\Psi = H\Psi$$

Here the operator H is the system's Hamiltonian,

$H=T+U$, where T and U are kinetic and potential energy respectively.

- The Schrödinger equation is a linear one, which reflects the fact that quantum mechanics is a linear theory, and each solution – state of a system – can be represented as a superposition of other states. This *superposition principle* allows one to invoke the mathematical techniques of linear algebra, matrix theory and functional analysis
- The Schrödinger equation, determining the behavior of a system in a stationary field, is

$$\frac{\hbar^2}{2m}\Delta\psi + [E - U(\mathbf{r})]\psi = 0, \Psi = \psi e^{-\frac{iEt}{\hbar}}$$

The Klein-Gordon (KG) equation

- To generalize quantum mechanics to relativistic problems (when particles can move with velocities close to the speed of light, c) is a nontrivial problem. For example, there is a critical dependence of the particle spin. The KG equation is a wave equation describing relativistic spin zero particles:
$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta + \left(\frac{mc}{\hbar} \right)^2 \right] \psi(\mathbf{r}, t) = 0$$
- In this form, it looks like a classical wave equation with an extra term $(mc/\hbar)^2 = \lambda_C^{-2}$, λ_C is called the Compton wavelength of a particle with mass m .
- Despite its apparent simplicity, the KG equation has some unusual features. Since it is second order in time, to predict the evolution of a system, one has to know not only the wave function ψ , but also its derivative $\partial\psi/\partial t$. This is connected with specifying the charge, which leads to the concepts of *antiparticle* and negative energy solutions and even negative density. All these concepts cannot be accurately treated using the KG equation (Klein's paradox).

The Dirac equation

- When we say “electronics”, “e-business”, “electronic media”, etc., we implicitly assume that the particle named “electron” plays an important role in many processes. But what are the properties of electron and how is it described?
- To answer these questions, one must find an equation describing free electron. Such an equation was obtained in 1928 by P.Dirac using general principles of mathematical modeling in physics – linearity (superposition), invariance (symmetry), first order in time (evolution is uniquely determined by the initial state).
- For a particle having spin $1/2$, only one system of linear equations can be shown to satisfy all these requirements:
$$\left(\gamma^\mu \partial_\mu + i \frac{mc}{\hbar} \right) \psi(x) = 0, \quad \partial_\mu := \partial / \partial x^\mu, \mu = 0, 1, 2, 3, \gamma^\mu \text{ are } 4 \times 4 \text{ matrices}$$

Four independent solutions of the Dirac equation describe both the state of a particle (electron) and antiparticle (positron)

The Hamilton-Jacobi (HJ) equation §

- A PDE for the function $S(\mathbf{r}, t) = S(q_i, t)$, called “action”, which is especially important both in classical and quantum mechanics:
$$\partial_t S + H(q_i, \partial S / \partial q_i, t) = 0, \quad i = 1..n$$

Here n is the number of degrees of freedom (independent coordinates needed to describe the physical or engineering system); the function S may be defined through the Lagrangian:
$$S = \int L dt + const$$

The HJ equation is very useful in finding analytically and numerically the motion of classical particles. Its value stems from the fact that it automatically gives all the constants of motion – conserved quantities. In fact, the solution itself is formulated in terms of these constants.

An intriguing feature of the HJ equation is that it can be derived from the Schrödinger equation using the form $\Psi(\mathbf{r}, t) = A(\mathbf{r}, t) \exp(iS(\mathbf{r}, t) / \hbar)$

This fact may indicate that classical and quantum models may be just different realizations of the same mathematical structure

The Boltzmann equation

- The equation describing the conservation of particles in the 6D phase space is called the Boltzmann equation:

$$\partial_t f + \operatorname{div}_{\mathbf{r}}(\mathbf{v}f) + \operatorname{div}_{\mathbf{v}}(\dot{\mathbf{v}}f) = \operatorname{St}(f)$$

This is the kinetic (balance) equation for the distribution function $f(\mathbf{r}, \mathbf{v}, t)$ giving by definition the average number of particles $f d^3r d^3v$ in the phase volume element. If there is no external influence, the distribution function does not depend on \mathbf{r} and t (Maxwell distribution):

$$f(r, v, t) = f_0(v) = n \left(\frac{m}{2\pi kT} \right)^{3/2} e^{\frac{-mv^2}{2kT}} \quad \text{where } n \text{ is the average density, } k \text{ is the Boltzmann constant}$$

The quantity $\operatorname{St}(f)$ is called the collision integral playing the role of a source term. Due to collisions, particles jump from one region of the phase space to another. Using the Boltzmann equation, one can obtain all the equations of hydrodynamics, heat and mass transfer, etc.

The Vlasov equation

- This is an important equation describing the behavior of a system of charged particles in the electromagnetic field created by them. For non-relativistic particles and neglecting the radiation, we have:

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \frac{\partial f}{\partial \mathbf{p}} = 0, f = f(\mathbf{r}, \mathbf{p}, t), \mathbf{p} = m\mathbf{v}, \mathbf{F} = \mathbf{F}(\mathbf{r}, t) = e\mathbf{E}(\mathbf{r}, t)$$

$$\operatorname{div} \mathbf{E} = 4\pi\rho(\mathbf{r}, t) = 4\pi \int f(\mathbf{r}, \mathbf{p}, t) d^3 p; \operatorname{rot} \mathbf{E} = 0$$

- This is the system of equations replacing the Boltzmann kinetic equation, when the collision integral is put to zero. It means that the pair correlations between particles are disregarded, and the whole system becomes non-dissipative and time-reversible
- The Vlasov equation is especially useful for describing waves and other collective motions in the system of charged particles (plasma)
- Analysis of collective motions (instabilities and damping) carried out using the Vlasov equation may play a decisive role for future power engineering**

The Euler equation

- Hydrodynamics of the *ideal* fluid is described by three equations: the continuity equation (conservation of mass) $\frac{\partial \rho}{\partial t} + \operatorname{div} \rho \mathbf{v} = 0$
- The Euler equation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \nabla) \mathbf{v} = -\frac{\nabla p}{\rho} + \mathbf{g}$$

Here $\mathbf{v}=\mathbf{v}(\mathbf{r},t)$ and $p=p(\mathbf{r},t)$ denote the pressure field in the fluid, $\rho(\mathbf{r},t)$ is the fluid density, and \mathbf{g} is the gravity acceleration.

- The liquid is called *ideal* when any possible dissipative processes in it, e.g. heat, mass and momentum transfer, can be neglected. This fact is manifested in one more conservation law – that of the entropy of each liquid fragment: $\frac{\partial s}{\partial t} + (\mathbf{v} \nabla) s = 0$

The equations of hydrodynamics are greatly simplified when the fluid density may be considered constant, $\operatorname{div} \mathbf{v} = 0, \frac{\nabla p}{\rho} = \nabla \left(\frac{p}{\rho} \right)$

The Navier-Stokes (NS) equation

- The NS equation is the main model of theoretical fluid dynamics. This equation takes into account irreversible momentum transfer between different fluid fragments, manifested as viscosity and described by phenomenological viscosity coefficients η and ζ .
- In the approximation when the fluid can be considered incompressible, the NS equation may be written as

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla p + \frac{\eta}{\rho} \Delta \mathbf{v} + \mathbf{g}$$

- The NS equation incorporates several physical processes, the main of them being convection, diffusion and viscosity (shear stress). This makes the NS equation quite universal. In different situations, particular terms of the equation may dominate resulting in different flow modes. The NS equation contains an empirical viscosity coefficient; its physical meaning can be understood only within the kinetic theory. One can get the NS from the Boltzmann equation.

The Fokker-Planck (FP) equation

- There are two types of models designed to describe classical non-equilibrium (irreversible) processes. One is based on Boltzmann equation and is valid for systems with dominant pair collisions (e.g. rarefied gases). Between the collisions the object is free.
- Another type of models is convenient for treatment of diffusive motion, when an object undergoes random influences all the time – it is never left alone. This is the situation of the Brownian motion
- The random motion of a system may be described by the transition probability $W(x, t | x_0, t_0)$ from one state $x := (\mathbf{r}_0, \mathbf{v}_0, t_0)$ into another $x := (\mathbf{r}, \mathbf{v}, t)$. When such a transition is determined only by the actual state (i.e. there is no memory of the preceding states – the Markov process), W satisfies the FP equation: (a_i and b_{ik} are so called

$$\frac{\partial W}{\partial t} = - \frac{\partial}{\partial x_i} [a_i(x, t)W] + \frac{1}{2} \frac{\partial^2 [b_{ik}(x, t)W]}{\partial x_i \partial x_k}$$

kinetic coefficients)

The Fokker-Planck (FP) equation-II

- The kinetic coefficient a_i and b_{ik} express the average velocity and displacement of a point representing the modeled system in its phase space. For a particle, this is a 6D space.
- However, the FP equation can be applied to various systems, not necessarily physical ones, whose phase space is not reduced to the direct product of coordinate and momentum spaces. This equation is a universal tool to describe the behavior of a system experiencing a stochastic force from the environment.
- The FP equation has become progressively important for modeling economical, financial, biological, sociological, etc. processes. For example, the distribution of the population with respect to income may be modeled by solving the respective FP equation.
- There are several forms, in which the FP equation can be written. One of the most popular is the form of a conservation law:

$$\frac{\partial f}{\partial t} + \frac{\partial j_\alpha}{\partial x_\alpha} = 0, \quad j_\alpha = a_\alpha f - \frac{1}{2} \frac{\partial}{\partial x_\alpha} (b_{\alpha\beta} f) \quad - \text{the “probability current”}$$

The telegraph equation (TE)

§

- This equation was traditionally very important in applied electrodynamics and electrical engineering. In fact, the telegraph equation is the linear wave equation with attenuation:

$$u_{tt} - a^2 \Delta u - pu_t - qu = f(\mathbf{r}, t), \quad x > 0, 0 < t < T$$

Usually, the 1D homogeneous case is considered:

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - a_1 \frac{\partial u}{\partial t} - a_2 u = 0$$

The particular solutions: $u(x, t) = e^{st \pm ikx}$ where $s = \sigma + i\omega$ is the root of:

$$\frac{1}{c^2} + a_1 s + (a_2 + k^2) = 0$$

Complex conjugated solutions of TE form a system of damped sinusoidal waves. These solutions embrace those for diffusion and wave equations as specific cases. TE may be conveniently treated using the Laplace transform, which gives the propagation of initial signal along the transmission line with dispersion and decay

The Green's functions (GF)

- The solution of the initial value problem

$$\frac{d^2 x}{dt^2} + \omega^2 x = f(t), x(0) = \dot{x}(0) = 0$$

can be written as an integral:

$$x(t) = \int_0^t G(t,s) f(s) ds, \quad G(t,s) = \frac{\sin \omega(t-s)}{\omega}$$

Here $G(t,s)$ does not depend on the source $f(t)$ and is determined by the properties of the differential operator and initial values.

- $G(t,s)$ is called the Green's function (GF). Green's functions, when they exist, may be regarded as the inverse of the differential (or integral, or integro-differential, etc.) operator.
- For linear differential equations (DE), the GF method is reduced to the transformation of DE into the integral equation.
- Green's functions exist both for ODE and PDE.

A new way to solve differential equations

Main ideas of the Green's functions theory may be illustrated on simple examples. The simplest of all differential equations:

$$\begin{aligned} -i \frac{dy}{dx} &= f(x) \\ y(a) &= y_0 \end{aligned} \quad \begin{aligned} &\text{(where } i, i^2=-1, \text{ is written for the convenience, though} \\ &\text{there is some sense in it because } id/dx \text{ is a} \\ &\text{translation operator)} \end{aligned}$$

may be solved as

$$y(x) = y_0 + i \int_0^x f(x') dx'$$

Suppose now that x changes within some closed interval $[a, b]$, then we can modify the solution:

$$y(x) = y_0 + i \int_a^b \theta(x - x') f(x') dx', \quad \text{where } \theta(t) = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases} \quad \begin{array}{l} \text{is a Heaviside} \\ \text{function.} \end{array}$$

Here $i\theta(x-x')$ is the kernel of the integral operator K in $y = L^{-1}f = Kf$

Solving the oscillator equation

Let us search a solution of the oscillator equation as a Fourier integral:

$$x(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} X(\omega) e^{-i\omega t}$$

Let us assume that the driving force can be represented in the same way: $f(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F(\omega) e^{-i\omega t}$ Then the motion equation becomes algebraic:

$$(\omega_0^2 - \omega^2) X(\omega) = F(\omega) + \underline{\text{const}(\omega_0^2 - \omega^2) \delta(\omega_0^2 - \omega^2)}$$

or

$$X(\omega) = \frac{F(\omega)}{(\omega_0^2 - \omega^2)} + \underline{\text{const} \delta(\omega_0^2 - \omega^2)}$$

This is the general solution of the inhomogeneous equation

where the underlined term is due to the division by the resonance factor $(\omega_0^2 - \omega^2)$ which can be zero. The product of the type $x\delta(x)$ can be interpreted as zero and one can add it.

However, the term with delta function in $X(\omega)$ does contribute – the general solution of the free equation

The particular solution of the driven equation

- One can write the solution as the response

$$x(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{-i\omega t}}{\omega_0^2 - \omega^2} \int_{-\infty}^{\infty} dt' e^{i\omega t'} f(t') = \int_{-\infty}^{\infty} dt' G(t, t') f(t'),$$

where the Green's function,
satisfies the equation

$$G(t, t') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{e^{-i\omega(t-t')}}{\omega_0^2 - \omega^2}$$

$$\left(\frac{d^2}{dt^2} + \omega_0^2 \right) G(t, t') = \delta(t - t')$$

Exercise c9:
Prove this relation!

Interpretation: the Green's function satisfies the oscillator equation with the pulsed driving force, therefore the solution of this equation with an arbitrary force is represented as the superposition of elementary responses to short pulses acting in various time points

Apparent difficulty: the integrand in $G(t, t')$ has poles in $\omega = \pm\omega_0$

So, to calculate the GF, one has to know how to handle these poles.

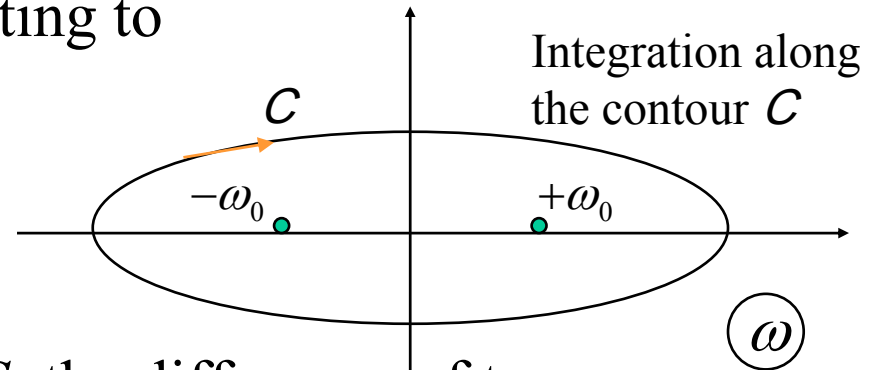
How to compute the Green's functions?

- The integral $G(t, t')$ does not exist in the usual sense.

The trick: first, to exclude poles from the integration region, then to exclude the exclusion (by transition to a limit)

- Deformation of the contour, exiting to complex plane, then returning (in the limit) to the real axis

- The question: how to deform the integration contour? Contour C : the difference of two integrals over infinite lines above and below the x axes.



The integration gives:

We obtained the solution of the homogeneous equation. The change of integration path results in adding this solution

$$G(t, t') = \frac{-1}{2\pi} \oint_C d\omega \frac{e^{-i\omega\tau}}{(\omega - \omega_0)(\omega + \omega_0)} =$$

$$= -\frac{-2\pi i}{2\pi} \left(\frac{e^{-i\omega_0\tau}}{2\omega_0} + \frac{e^{i\omega_0\tau}}{-2\omega_0} \right) = \frac{1}{\omega_0} \sin \omega_0 \tau, \tau := t - t'$$

A collection of Green's functions (GF) §

- The solution of a free equation obtained by the integration over the contour C is often called the Pauli D-function – very important in the quantum field theory

$$D(\tau) = D^{ret}(\tau) - D^{adv}(\tau)$$

- The retarded Green's function appears when the integration path lies above the real axes:

$$D^{ret}(\tau) = \frac{1}{2\pi} \int_{C^{ret}} d\omega \frac{e^{-i\omega\tau}}{\omega_0^2 - \omega^2} = \begin{cases} \frac{\sin \omega_0 \tau}{\omega_0}, & \tau > 0 \\ 0, & \tau < 0 \end{cases} = \theta(\tau) \frac{\sin \omega_0 \tau}{\omega_0}$$

- The Fourier-representation: $D^{ret}(\omega) = (\omega_0^2 - \omega^2 - i\varepsilon\omega)^{-1}$

Here an infinitely small displacement of both poles (attenuation)

is added: $\pm\omega_0 \rightarrow \pm\omega_0 - i\varepsilon$

- The advanced GF: $D^{adv}(\tau) = -\theta(-\tau) \frac{\sin \omega_0 \tau}{\omega_0}$

- Fourier image: $D^{adv}(\omega) = (\omega_0^2 - \omega^2 + i\varepsilon\omega)^{-1}$

- 4 variants of passing by the poles + V.P. = 5 GFs

Advanced: for the problems describing the influence from the future, e.g. final condition instead of initial one

The Green's functions (continued)

- GF are important since they allow one to express the unknown function in terms of known quantities – this may be sufficient for the investigation of solution properties
- Theory of the Green's functions has been developed in a variety of ways. Here we discuss it in an heuristic and non-rigorous manner
- Some analogies may be helpful, e.g. let us consider a matrix equation $\mathbf{A}\mathbf{y} = \mathbf{f}$, where \mathbf{A} is $n \times n$ non-singular matrix and \mathbf{f} is a n -dimensional vector. The solution: $\mathbf{y} = \mathbf{A}^{-1}\mathbf{f}$, where \mathbf{A}^{-1} does not depend on \mathbf{f} and can be used to find the solution for any vector \mathbf{f}
- Similarly, the differential equation $dy/dx = f(x)$ has an inverse $y(x) = \int dx f(x)$, so the integral is the inverse operator for d/dx and is made unique by specifying supplementary conditions
- The idea: to find the unknown function by inversion, then specify

Green's functions for initial values

We may note that the solution of the preceding equation has a form

$$y(x) = y_0 + Kf(x)$$

where K is an integral operator $Kf(x) := i \int_a^b \eta(x-x') f(x') dx'$

with the kernel $G(x, x') = i\eta(x-x')$

which is the Green's function of the problem defined by the simple ODE with the boundary condition $y(a)=y_0$

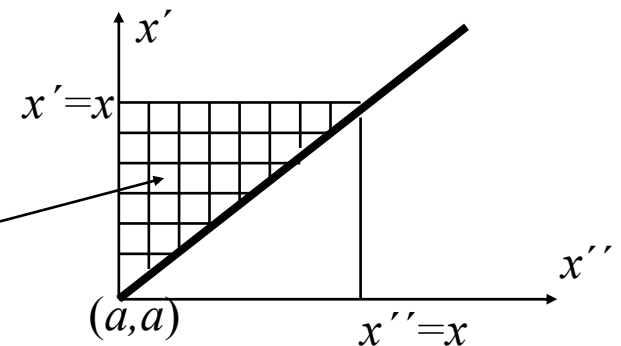
A less trivial example:

$$\frac{d^2 y}{dx^2} = f(x), y(a) = y_0, y'(a) = p$$

Integrating, we obtain

$$y(x) = y_0 + p(x-a) + \int_a^x dx' \int_a^{x'} dx'' f(x'')$$

Integration region



We integrate first over x'' then over x' but we can change the order

Green's functions for boundary values

We get: $y(x) = y_0 + p(x-a) + \int_a^x dx' f(x') \int_{x'}^x dx'' = y_0 + p(x-a) + \int_a^x (x-x') f(x') dx'$

Restricting x to the interval $[a, b]$, we may write

$$y(x) = y_0 + p(x-a) + \int_a^b (x-x') \eta(x-x') f(x') dx'$$

where

$$G_2(x, x') := (x-x') \eta(x-x')$$

is the Green's function for the operator d^2/dx^2 with the above supplementary (initial) conditions. The general solution:

$$y(x) = C_1 + C_2 x + \int_a^b G_2(x, x') f(x') dx'$$

where C_1, C_2 can be determined from boundary conditions

Exercise c10: Find a solution for boundary conditions $y(0)=y_0, y(1)=y_1$

Some useful mathematical tricks: elimination of the dissipative term

- Suppose we have a free dynamical system with dissipation:

$$\ddot{x} + a(t)\dot{x} + b(t)x = 0$$

- We can rewrite this equation in the form, which does not contain the dissipative term explicitly, making a substitution $x(t)=y(t)z(t)$
- Then $y\ddot{z} + (2\dot{y} + ay) + (\ddot{y} + a\dot{y} + by)z = 0$ and if $2\dot{y} + ay = 0$

we have

$$x(t) = \exp\left(-\frac{1}{2}\int_{t_0}^t a(t')dt'\right)z(t)$$

and

$$\ddot{z} + \left(b(t) - \frac{a^2(t)}{4} - \frac{\dot{a}(t)}{2}\right)z = 0$$

Functions $x(t)$ and $z(t)$ have the same zeros. So, for dynamical modeling we can always consider equations of the type:

$$y'' + q(x)y = 0 \quad \text{instead of three-term initial equation}$$

Useful mathematical tricks: estimating integrals

- Suppose we need to calculate $I = \int \exp(f(t)) dt$

Where the function $f(t)$ has a sharp maximum or oscillates rather slowly. Then we can simplify $f(t)$ near $t=a$ which is determined by $f'(a) = 0$. Thus $f(t) = f(a) + \frac{f''(a)}{2}(t-a)^2$ and using the Poisson integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

we get

$$I = \sqrt{\frac{2\pi}{-f''(a)}} e^{f(a)}$$

This trick can be used e.g. for the Bessel function:

$$J_\nu(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(x \sin t - \nu t)} dt \quad - \text{the extremum is determined by } \cos a = \nu/x$$

so a is real at $x > \nu$. Integrals of the above type are ubiquitous in optics, electrical and radio engineering, and communication theory.

Integration of oscillating functions

- Suppose we have to calculate $I(a, x) = \int_0^x f(t) \cos at dt$ e.g.

$$I(a, x) = \int_0^x e^{-t} \cos at dt$$

for large a , e.g. for $a=100$. The integral is difficult to compute numerically with high accuracy, since the integrand is rapidly oscillating and the result is the sum of a large number of nearly equal terms with opposite signs.

- In this case one can use the following trick: smooth function $f(t)$ is replaced by another smooth function $\tilde{f}(t)$ so that one could compute the integral $\tilde{I}(a, x) = \int_0^x \tilde{f}(t) \cos at dt$ analytically.

Then the computation of the initial integral is reduced to identity:

$$I(a, x) = \int_0^x f(t) \cos at dt = \int_0^x \tilde{f}(t) \cos at + \int_0^x [f(t) - \tilde{f}(t)] \cos at dt$$

Integration of oscillating functions-II

- Let us try to estimate $I(a, \pi) = \int_0^{\pi} f(t) \cos at dt$
- If we neglect the second term using the inequality $|f(t) - \tilde{f}(t)| \leq \varepsilon$, formally we get the disregarded term as $\delta I \leq \pi \varepsilon$
- However, the latter quantity may easily exceed the estimated integral.

- In fact, the integration error is much lower, since it is the integral of a smooth function multiplied by a rapidly oscillating function.
 - Therefore, one can expect the following estimate for the error:
- $$\delta I(a, \pi) / I(a, \pi) \approx |f - \tilde{f}| / f$$
- For $f(t) = e^{-t}$ the integral can be calculated analytically:

$$I(a, \pi) \approx \frac{1 - e^{-\pi} (\cos \pi a - a \sin \pi a)}{1 + a^2}$$

- The above examples show that in some cases analytical methods allow one to achieve the results easily and with fair accuracy

Equations with periodic coefficients

- There are many problems in science and engineering leading to equations of the type $y'' + q(x)y = 0$ where $q(x) = q(x+a)$
- For instance, the equation $y'' + (k^2 - q(x))y = 0$ models the motion of electron in the crystal lattice – the problem of fundamental importance for electronic technology (modern computers would not exist if this problem had not been solved)
- The equation $y'' + k^2 q(x)y = 0$ describes propagation of acoustical and electromagnetic waves in the media with periodic properties
- The physical requirement: solutions must be bounded everywhere
- Solutions of equations with periodic coefficients are of the type $y(x) = e^{\lambda x} u(x)$, where $u(x)$ is a periodic function with period a .
- In physics, such solutions are called Bloch functions, they describe the state of a particle in a translationally-invariant media. The dispersion law dictated by periodic structures leads to the fact that the particle in a crystal becomes *a quasiparticle*
- Equations with periodic coefficients give a good example of *universality* of mathematical models: the same equations describe totally different phenomena

Zones of stability

- The point k^2 belongs to stability zone, if all solutions of the equation with periodic coefficients are finite for $-\infty < x < \infty$, and to instability (forbidden) zone if solutions are infinite
- From the form of the solutions we see that k^2 lies in the forbidden zone if multipliers $|e^{\lambda a}| \neq 1$, i.e. $\text{Re } \lambda \neq 0$. The function $k^2(\lambda)$ represents the dispersion law of electronic, acoustic or EM waves
- For electrons in the crystal lattice k^2 has the meaning of energy and the dependence $k^2(\lambda)$ is called electronic spectrum (usually denoted in 3D case as $E(\mathbf{p})$ where \mathbf{p} is quasimomentum).
- **The type of electronic spectrum (band structure) determines the material properties (metal, dielectric, semiconductor).**
- For wave propagation problems k^2 has the meaning of frequency of waves that can propagate through a periodic structure, the respective dispersion law determines allowed frequencies.

Mathieu and Hill equations

- The Mathieu equation $y'' + (a - 2q \cos 2x)y = 0$ first appeared as a result of solution of the 2D Helmholtz equation, describing a membrane oscillations in elliptic coordinates.
- Now it has become very important for the models describing parametric resonance phenomena in various branches of science and engineering – from usual child's swing to molecular and free electron lasers (FEL). The motion of electrons in a periodic array of atoms can also be modeled by the Mathieu equation
- The latter is the simplest case of a more general Hill equation $y'' + (a + p(x))y = 0$ where p is a periodic function of x . This equation was named after an American astronomer who derived it while investigating the stability of lunar motion
- Study of stability of periodic orbits in general produce Mathieu and Hill equations. These systems are non-autonomous.
- **Theory of Mathieu and Hill equations is a specific case of a more general Floquet theory (the Floquet-Lyapunov theorem).**

Chapter 3. Computational Techniques

The main computer modeling methods are overviewed, including numerical approaches and computer algebra techniques. The accent is put on Maple and MATLAB

The paradigm change

- During the last two decades, computers have radically changed the image of science and engineering.
- In 70ies, numerical techniques gradually became an accepted tool even in physics, traditionally dominated by analytical concepts
- Today, numerical modeling is applied to nearly all scientific areas, including those that are very difficult to formalize (e.g. biology)
- There are even cases when the problem can be modeled computationally without any empirical input (e.g. nuclear explosion)
- As the computers have become more powerful, a new research area – *scientific computing* – emerged and with the further increase of computer power, hybrid techniques combining symbolic solutions with previously developed numerical algorithms can be used
- To test such new environment, computer algebra systems (CAS) seem to provide the best functionality

Available mathematical software

- The volume and diversity of mathematical software is constantly rising, so the selection of appropriate tools has become a problem
- **Rough classification:**
 1. General purpose tools – Maple, Mathematica, MATLAB, MATHCAD, REDUCE, OCTAVE, etc. To create a general purpose tool (e.g. CAS) for all mathematical problems would be a very hard task. The basic idea is to design a kernel for solving comparatively simple problems and to call special libraries and procedures thereof
 2. Dedicated software – Differential Systems, Ode Workbench, Chaotic Mapper, Chaotic Dynamics Workbench, etc. Usually built with special-purpose databases valid for specific problems
 3. Problem solving environment (PSE) – contains the user interface allowing to specify the problem, then the software calls the method

Computer algebra systems (CAS)

- More intelligent mathematical assistants than computer languages
 - Integrated mathematical modules, can be used by non-professionals
 - Allow non-numerical – symbolic – computations. Representatives: AXIOM, DERIVE, MACSYMA, POWERMATH, REDUCE, muMATH, SYMBAL – previous generation of CAS
 - The most popular nowadays are **Maple** and **Mathematica**
 - Current improvements – a well-developed GUI, maximization of the use of computing resources available, more features
 - Feature groups: analytical (symbolic), numerical, graphical, GUI, communicational, conceptual, mathematical expertise (packages)
 - Three most important tasks to be jointly solved by CAS:
1) symbolic manipulations; 2) numerical computations; 3) visualization
- Mathematical models can be readily formulated in terms of CAS

- In fact, CAS were developed by physicists who usually dislike the conventional programming. Even today, though CAS are used in many fields, large application packages are mostly restricted to areas close to physics, e.g. high energy physics, cosmology, relativity, many-body theory, circuit/device design, etc.
- **M. Veltman** (1999 Nobel Prize in Physics for classical works on gauge theories): first general purpose CAS named SCHOONSHIP, 1967 (CERN): could perform complex algebraic calculations
- 1968 – REDUCE (**Anthony Hearn**, also a theoretical physicist): the motivation was to solve problems typical for particle physics – Feynman diagrams, cross-sections, transition amplitudes, etc. The first commercial distribution.
- Now: over 20 CAS are known with various functionalities
- Now: the main paradigm is to provide the user with a single problem solving environment (PSE). Advanced applications:
 1. Linear algebra, coordinate transforms; 2. Behavior of dynamical systems;
 3. Automatic search for symmetries and appropriate coordinate changes;
 4. Simplifications of quantum mechanical expressions; 5. Computational biology;
 6. Systems of differential equations; 7. Signal processing; 8. Mathematical physics

- Based on matrix calculus, advanced engineering software
- Main categories of functions (e.g. plotxy, plotxyz, elfun, elmat, specfun, specmat). Line editing and recall keys (different for UNIX/Linux, Windows, VMS, Mac)
- General purpose commands (e.g. managing commands and functions, controlling the command window, working with files and operating environment)
- **Simulink** – a MATLAB extension, especially designed to simulate dynamical systems. A blockset structure, ready-made graphical and functional units. Especially useful for block diagram design, e.g. in signal and image processing. DSP and communication design
- **Modeling with MATLAB:** PDE, nonlinear systems of equations, optimization, financial mathematics, statistics, curve fitting

Maple

- A huge number of commands and functions; there is a lot to remember. Maple Release 8.0 is current default for all platforms
- <http://www.maplesoft.com/>
- Maple allows to:
 - perform symbolic integration
 - find closed-form solutions for ODE and PDE
 - develop hybrid symbolic-numeric algorithms for scientific computing
 - carry out algebraic manipulation over differential and partial differential operators
 - design and implement rational approximation algorithms
- Extensive facilities for drawing curves and surfaces, plotting procedures may be conveniently used to graph the solutions of differential equations (e.g. use of **with plots** to access routines)

Document organization in Maple

- One should familiarize oneself with Maple GUI
- The best way is to learn use and syntax of Maple simultaneously and by examples
- A worksheet (*.mws) consists of logical groups called „execution groups“ Each group consists of several parts :
 - Input Region
 - Output Region
 - Text Region (e.g. comments)
- The principal objects in Maple: numbers, constants, names
 - numbers: integer, real, rational, radicals
 - constants: Pi, I, E, Infinity, gamma, true, false
 - names: sequences of symbols starting with literal (e.g. new, New, new_old, etc.)

Syntax and expressions

- After typing a command, you should inform the computer that the command has ended and must be accepted. There are two types of terminators: ; and : The colon is needed to suppress Return – the output is not printed
- The symbol % is called “ditto”, it is used to refer to previous return. Twofold ditto (%%) refers to the second last return, etc. In releases before Maple V.5, the double quote “ was used for ditto. Remark: this notation may be convenient for a succession of operations, but can be confusing when the commands are executed in arbitrary order
- Expressions are the main object for the most of commands. Expressions are composed of constants, variables, signs of arithmetical and logical operations
- Getting help: a “?” in front of a command, e.g. ?plot on graphics

Typical syntax errors and remedies

- Forgetting a bracket or an operator – the error message comes after Return (sometimes it appears meaningless)
- It is easier to type the brackets and some operators first, and then insert functions, variables and parameters, e.g.

```
> ( )*(sqrt( ) + ( )^( ) )* exp( );
```
- It is better to start a new worksheet with the command **restart** – this resets the system and variables used in a previous worksheet do not cast their value into the new one
- Not all combinations of symbols are allowed in Maple (e.g. **Pi** and **I** are already used by the system). In older versions of Maple **E** was reserved for the constant **e=2.71828..**; in later versions the value of the natural logarithm base is given by **e=exp(1)**; but **exp (1.0)** gives directly the number **2.718281828**

Assignment of variables in Maple

- In contrast to many programming languages, variables in Maple can be free (unbounded). For instance in Fortran or Basic (and to some extent in C), variables can be used only after being assigned a specific numerical value
- A variable can be made bound by using the assignment symbol :=
- Example: to find the solution of a quadratic equation

> solve(a*x^2 + b*x + c=0,x);

$$\frac{1}{2} \frac{-b + \sqrt{b^2 - 4ac}}{a}, \frac{1}{2} \frac{-b - \sqrt{b^2 - 4ac}}{a}$$

Here x is the unknown variable; a,b,c are parameters – all these letters are free. If we assign the variable y the value $ax^2 + bx + c$:

> y:=(a*x^2 + b*x + c=0,x); # then until y is cleared or assigned
another value, y is $ax^2 + bx + c$

Different manners of assignment

- An example: to solve the system of equations

$$ax + by = A; cx + dy = B \quad \text{with respect to } x \text{ and } y$$

First – the ‘restart’ command, since x and y could be already assigned

- Then we may assign e.g. z to the whole set of equations and u to the set of unknown variables

> **restart;**

> **z:={a*x+b*y=A,c*x+d*y=B}:**

> **u:={x,y}:**

> **sol:=solve(z,u);**

Pressing ‘return’ we obtain the Maple solutions, now we can assign these solutions to be x and y

> **assign(sol);**

> **x,y;**

Different manners of assignment-II

- After the 'assign' command variables are not free
- The used manner of assignment is convenient when there is only one solution
- Curly brackets signify declaration of a set
- An example of another manner of assignment. A problem:

To find the distance of the intersection point of two straight lines defined by the above equations from the origin of coordinates

```
> dist:=subs(sol,sqrt(x^2+y^2));
```

Here the substitute command 'subs' is conveniently used, and in this manner the variables x and y remain free

If we attempted to call the distance 'd' or 'D', Maple would produce an error (**Exercise c5: why?**)

Some sources of confusion

- The difference between “assign” $:=$ and “is equal to” $=$
- By assigning $x := a$ we are instructing the computer to replace x by a wherever the former is met
- Let us take a linear function, $y = kx + b$. This equation can be rearranged e.g. in the manner $y - b = kx$ or $x = (y - b)/a$
- Now suppose we assign a variable, $y := k * x + b$ then we are already unable to rearrange this assignment
- Suppose that one assigns a variable to an equation:
> $y := a * x^2 + b * x + c$

so that one can solve the quadric equation with the command

> solve (y,x). If one wishes now to solve a system of equation in variables (x,y), one should be careful because y has already been assigned (e.g. one would need to clear variables with **restart**)

Substitutions

- The substitute command is very important, since it is used to change variables. Its formal structure may be designated as:
`subs (old=new, expression),`
where the variable 'old' is replaced by the expression 'new' in the object 'expression' (actually, this object may be not only an expression but also a set, a list, etc.)
- It is allowed to use a number of substitutions 'old=new', but in this case multiple substitutions should be formulated as a set (see two preceding slides)
- Substitution does not change the initial object, but results in a new expression. Thus, some simplifications may be needed. Example:

> **restart;**

> **s:=subs(x=0,(sin(x)-1)*cos(x-Pi/3));**

> **factor(s),expand(s),simplify(s),eval(s);**

Exercise c6: Try it!

Usual difficulties with assignments

- The following example: we have assigned $y := -3 * x^2 + 7x + 9$

Now we want to know $y(2)$, we can get this value by assigning 2 to x , so y automatically becomes:

```
> y:=-3*x^2+7*x+9;
```

$$y := -3x^2 + 7x + 9$$

Exercise c7: Try it!

```
> x:=2;y;
```

$$x := 2$$

And now we need to integrate y , i.e. to compute $\int y dx = \int dx(-3x^2 + 7x + 9)$

Using the Maple notation, we might write

```
> z :=int(y,x) # but this is wrong and we will receive an error message
```

The matter is that both x and y are already numbers, whereas ‘int’ expects at least x to be a free variable. How can we set $x=2$ locally?

Reserved names in Maple

all the standard mathematical functions, e.g. sin, cos, exp, sqrt, sinh, cosh have predefined names, and so do Maple commands and procedures: int, diff, plot, sum, coeff, rhs, lhs,.. The same applies to data types: array, list, set, matrix,..

and	by	do	done	elif	RETURN
else	end	fi	for	from	break
if	in	intersect	local	minus	next
mod	not	od	option	options	
or	proc	quit	read	save	
stop	then	to	union	while	

Repetitions and loops

- If we want to obtain a better approximation, we need to repeat the same statements many times. Moreover, we need to determine when we have reached a sufficient accuracy
- This is provided by a while statement, for example:

```
> f:=(theta)->sin(theta)-theta*cos(theta)-evalf(Pi/2);
> pos:=3.0;
> neg:=1.0;
> while (abs(pos-neg))>=1e-5 do
>   avg:=(pos+neg)/2.0;
>   print(avg);
>   if (f(avg)>=0 then
>     pos:=avg;
>   else
>     neg:=avg;
>   end if;
> end do;
```

for, while and do loops are programming devices to perform repetitions.

The general structure: for *counter* from *first counter value* by *increment* to *last counter value* do;

while *condition* do; *commands* od;

Exercise: calculate the product of all odd numbers between -5 and 5

Solving equations with Maple

- Two different methods of finding roots:
1) analytical (symbolic); 2) numerical
- Maple provides two respective built-in functions:
1) `solve`; 2) `fsolve`
- Analytical approach: the equation is manipulated up to the form when the unknown is expressed in terms of the knowns – `solve` operates in this way. The variety of equations to be analytically solved is determined by the set of techniques available in Maple
- Each new Maple version – new commands and packages.
Multipurpose commands of the `xsolve` type: `dsolve` for ODE, `pdsolve` for PDE, `rsolve` for difference equations
- New libraries: `DEtools`, `PDEtools`, `LREtools`. Special packages for implicit solutions: `DESol`, `RESol` (difference), `RootOf` (algebra)

Analytical solutions

- The solve command – multipurpose, can be applied to systems of algebraic equations, inequalities, functional equations, identities. The form: `solve (eqn, var)`, where `eqn` is an equation or a system of equations, `var` is a variable or a group of variables. If the latter parameter is absent, `solve` will look for solutions with respect to all symbolic variables
- The system of equations and the group of variables are given as sets (in curly brackets and separated by commas)
- Example: `> s:=solve({6*x+y=sqrt(a), 3*x-2*y=4},{x,y});`
- If the equation is written without the equality sign it is viewed as one part of the equality, the other being considered zero:

`> sl:=solve(x^7+8*x,x);`

Numerical treatment of equations in Maple

Many equations encountered in science and engineering are hard to treat analytically – they have to be solved numerically

To solve differential equations numerically with the help of `dsolve` command one must declare the parameter `numeric`. This can be done by two manners:

- a) `dsolve(ODE, var, type=numeric, method)`
- b) `dsolve(ODE, var, numeric, method)`

Equations to be solved and initial (or boundary) conditions are defined as a set; `var` are lists of unknown functions; `method` parameters define the numerical techniques invoked to solve ODE. By default, the method is understood as Runge-Kutta-Fehlberg 4/5 (`rkf45`)

Also by default, as a result of applying `dsolve` with the parameter `numeric` a procedure is created to compute specific values

An example of numerical solution of ODE

- The dynamical system – the Cauchy problem for a nonlinear oscillator with linear friction:
> `s:=D(x)(t)=y(t),D(y)(t)=-a*y(t)-.8*sin(x(t));`
> `ic:=x(0)=0,y(0)=2;`
> `F:=dsolve({s,ic},{x(t),y(t)},numeric);`
> `a:=.2;evalf(F(1.5),5);`
> `plots[odeplot](F,[[t,x(t)],[t,y(t)]],0..30,labels=[t,"x,y"],axes=boxed);`
- To visualize numerical solutions of this Cauchy problem, one may use the graphics commands from the DEtools package
- The obtained solution may serve as a source of all the necessary information about the system

Exercise c8: Solve numerically the Cauchy problem for the following dynamical system:
Draw its phase portrait $\rightarrow \frac{dx}{dt} = y(t), \frac{dy}{dt} = x(t) - x^3(t), x(0) = 0, y(0) = 0.1$

Numerical methods for modeling

Some generalities:

- There are tons of books on numerical methods

Roughly, two groups of information sources and their recipients

- 1) numerical mathematics as a scientific subject
- 2) numerical mathematics as a working tool

This chapter is oriented towards 2): using general numerical methods to solve applied problems

- Therefore – ultimate generality and rigour are unnecessary

Functions are assumed smooth, meaning “finite together with all the derivatives of the order to be used in the specific case”, where needed

From a physicist's point of view, numerical techniques seem to be closer to experimental science than to idealized mathematical constructions: 1) it's an art where the result justifies the means; 2) it's a team game – interdisciplinary sport

Numerical integration of the Cauchy problem

- Let us consider the Cauchy problem:

$$\frac{dx}{dt} = f(x, t), x(0) = x_0, 0 \leq t \leq T, \quad \text{where } x \text{ is a } p\text{-dimensional vector,}$$

$f(x, t)$ is a given vector-function of
the same dimensionality, x_0 is the initial point

This problem corresponds to a non-autonomous dynamical system describing e.g. the missile or aircraft motion, evolution of biological and economical systems, etc.

The grid: $0 = t_0 < t_1 < \dots < t_N = T$

(see the lectures by M. Mehl-
Chr. Zenger, lecture 5)

For simplicity, the uniform grid:

$$t_n = n\tau, \quad \tau = T / N$$

Approximate solution as a grid function – a function of the discrete argument, x_n , $n=0..N$, each x_n being a p -dimensional vector. Mapping:
 $x_n \approx x(t_n)$ (x_n approximately represent the values of $x(t)$ on a grid)

Difference equations

- The discrete (grid) function x_n cannot satisfy any differential equation, so one has to construct some other equations for it
- Afterwards, one has to check that x_n , provided it could be found, is an approximate solution to the Cauchy problem and really describes the dynamical system in question
- A variety of ways to construct difference equations:

$$\frac{x_{n+1} - x_n}{\tau} = f(x_n, t_n) - \text{the explicit Euler scheme,}$$

In all these schemes the values x_n are determined sequentially: from right to left

$$\frac{x_{n+1} - x_n}{\tau} = f(x_{n+1}, t_{n+1}) - \text{the implicit Euler scheme,}$$

$$\frac{x_{n+1} - x_{n-1}}{2\tau} = f(x_n, t_n) - \text{central difference scheme}$$

- There are also more complicated difference schemes

The Euler's method

- Explicit scheme: $x_{n+1} = x_n + \tau f(x_n, t_n), n = 0, 1 \dots (N-1)$

Here x_0 is known, and new values are used in rhs to produce updates

- Implicit (backward) scheme – a little bit more complicated:

$$x_{n+1} = x_n + \tau f(x_{n+1}, t_{n+1})$$

Here, if the value x_n has already been known, x_{n+1} can be found from this nonlinear equation (with respect to x_{n+1}). However, the non-linearity is weak due to the small factor τ

The presence of a small parameter greatly facilitates the solution process. The meaning: x_{n+1} is only slightly different from x_n , so x_n can be taken as a good initial approximation

The general idea: it is enough to define x_0 to determine x_1, x_2, \dots

Central difference scheme

- In the Euler's method the sequence $\{x_n\}$, $n=1,2,\dots$ is uniquely determined by the initial value x_0
- In the central difference scheme this is insufficient: only declaring x_0 and x_1 we can determine $x_2 = x_0 + \tau f(x_1, t_1)$, etc.
- While the initial value is defined by the Cauchy problem, x_1 can – formally – possess any value
- In fact, however, x_1 should be carefully selected, for instance calculated by the explicit Euler's scheme: $x_1 = x_0 + \tau f(x_0)$
- The solution of the difference equation $\frac{x_{n+1} - x_{n-1}}{2\tau} = f(x_n, t_n)$ can be looked for in the standard form
(for homogeneous equations): $x_n = A_1 q_1^n + A_2 q_2^n$ where q_1, q_2 are the roots of characteristic equation; A_1, A_2 are constants determined from initial values (in this case x_0 and x_1)

Central difference scheme (an example)

- Let us consider the equation $\dot{x} = ax$. The characteristic equation for $\frac{x_{n+1} - x_{n-1}}{2\tau} = ax_n$ is $q^2 - 2a\tau q - 1 = 0$

(the characteristic equation is obtained by inserting q^n into the difference equation). The solution: $q_{1,2} = a\tau \pm \sqrt{1 + (a\tau)^2}$

The first root (we may assume $|a\tau| \ll 1$)

$$q_1 = 1 + a\tau + \frac{1}{2}(a\tau)^2 + O(\tau^3) \approx e^{a\tau} + O(\tau^3)$$

The solution of the difference equation: $x_n = q_1^n = e^{an\tau} (1 + O(\tau^3))$ strives in the limit to the solution of the ODE ($n \approx O(1/\tau)$)

Exercise c11: What kind of solution corresponds to the second root of the characteristic equation?

Centered derivatives

- One may ask, what might be the motivation to use central difference schemes to solve ODE, while the Euler methods are already here?
- Let us look at the derivatives: the Euler's method is based on the usual forward derivative. An equivalent definition:

$$f'(t) = \lim_{\tau \rightarrow 0} \frac{f(t + \tau) - f(t - \tau)}{2\tau} \quad \begin{array}{l} \text{- this formula is centered} \\ \text{with respect to } t \end{array}$$

- There is an important difference when τ is finite:

$$f(t + \tau) = f(t) + \tau f'(t) + \frac{1}{2} \tau^2 f''(t) + \frac{1}{6} \tau^3 f'''(\zeta_+);$$

$$f(t - \tau) = f(t) - \tau f'(t) + \frac{1}{2} \tau^2 f''(t) - \frac{1}{6} \tau^3 f'''(\zeta_-)$$

Which gives that the centered discretization error is quadratic in τ -improvement as compared to forward derivative methods

Numerical algorithms for Newton's equations

- To choose an appropriate method of integration of a system of differential equations, the following criteria are relevant:

- Stability
- Accuracy
- Reliability
- Efficiency
- Simplicity
- Possibility to use supercomputers (vectorization, parallelization)

The time-step τ should be properly chosen too. Usually, its choice is the result of a compromise between accuracy and computational cost. However, in modeling additional limitations may arise: e.g. τ should be small compared with the time of the processes in the system
Example: collision – the target object must pass negligible distance during τ

The most popular algorithms for numerical treatment of Newton eqns:

- a) Central differences; b) Average force; c) Euler-Cauchy method;
d) Predictor-corrector; e) Verlet method; f) Adams-Nordsieck method

One can find the rigorous treatment of these numerical algorithms in the courses of numerical analysis and scientific computing

Central differences for motion equations

- The velocity is defined in the middle of the time-step $\tau/2$:

$$\mathbf{x}_i(t + \tau) = \mathbf{x}_i(t) + \tau \dot{\mathbf{x}}_i\left(t + \frac{\tau}{2}\right)$$

\mathbf{F}_i is the force acting on the i -th object (particle) having the mass m_i

$$\dot{\mathbf{x}}_i\left(t + \frac{\tau}{2}\right) = \dot{\mathbf{x}}_i\left(t - \frac{\tau}{2}\right) + \tau \frac{\mathbf{F}_i(t)}{m_i}$$

See the discussion of central difference schemes in the book A.L. Garcia “Numerical Methods for Physics”, Prentice-Hall, 2000

- The average force method is a modification of central differences:

$$\mathbf{x}_i(t + \tau) \simeq \mathbf{x}_i(t) + \dot{\mathbf{x}}_i(t)\tau + \frac{1}{2}\ddot{\mathbf{x}}_i(t)\tau^2 \approx \mathbf{x}_i(t) + \tau\left(\dot{\mathbf{x}}_i(t) + \frac{\langle \mathbf{F}_i \rangle}{2m_i}\tau\right)$$

$$\dot{\mathbf{x}}_i(t + \tau) \simeq \dot{\mathbf{x}}_i(t) + \frac{\langle \mathbf{F}_i \rangle}{m_i}\tau, \text{ where } \langle \mathbf{F}_i \rangle := \frac{\mathbf{F}_i(\tilde{\mathbf{x}}_i(t + \tau)) + \mathbf{F}_i(\mathbf{x}_i(t))}{2},$$

$$\tilde{\mathbf{x}}_i(t + \tau) = \mathbf{x}_i(t) + \dot{\mathbf{x}}_i(t)\tau + \frac{\mathbf{F}_i(\mathbf{x}_i(t))}{2m_i}\tau^2$$

The average force method allows one to compute \mathbf{x}_i and $\dot{\mathbf{x}}$ simultaneously

Euler-Cauchy algorithm

- An iteration procedure:
- Iterations stop when the variations of phase point

$$\begin{aligned}\mathbf{x}^{(1)}(t + \tau) &= \mathbf{x}(t) + \dot{\mathbf{x}}(t)\tau + \frac{1}{2}\ddot{\mathbf{x}}(t)\tau^2 \\ \dot{\mathbf{x}}^{(1)}(t + \tau) &= \dot{\mathbf{x}}(t) + \ddot{\mathbf{x}}(t)\tau\end{aligned}$$

$\mathbf{x}, \dot{\mathbf{x}}$ become sufficiently small:

$$\mathbf{x}^{(n)}(t + \tau) = \mathbf{x}(t) + (\dot{\mathbf{x}}(t) + \dot{\mathbf{x}}^{(n)}(t + \tau))\frac{\tau}{2}$$

The predictor-corrector method

is, in fact, the same algorithm

$$\dot{\mathbf{x}}^{(n)}(t + \tau) = \dot{\mathbf{x}}(t) + (\ddot{\mathbf{x}}(t) + \ddot{\mathbf{x}}^{(n)}(t + \tau))\frac{\tau}{2}$$

with an element of prognosis

$$\ddot{\mathbf{x}}^{(n)}(t + \tau) = \mathbf{F}(\mathbf{x}^{(n-1)}(t + \tau)) / m$$

that is corrected at the following stage, then the iteration procedure is

$$\mathbf{x}_p(t + \tau) = \mathbf{x}(t - \tau) + 2\dot{\mathbf{x}}(t)\tau,$$

repeated:

$$\ddot{\mathbf{x}}_p^{(1)}(t + \tau) = \mathbf{F}(\mathbf{x}_p(t + \tau)) / m;$$

$$\mathbf{x}_c^{(n)}(t + \tau) = \mathbf{x}(t) + (\dot{\mathbf{x}}_c^{(n)}(t + \tau) + \dot{\mathbf{x}}(t))\frac{\tau}{2}$$

$$\mathbf{x}_c^{(1)}(t + \tau) = \mathbf{x}(t) + (\dot{\mathbf{x}}_c^{(1)}(t + \tau) + \dot{\mathbf{x}}(t))\frac{\tau}{2}$$

$$\dot{\mathbf{x}}_c^{(n)}(t + \tau) = \dot{\mathbf{x}}(t) + (\ddot{\mathbf{x}}_p^{(n)}(t + \tau) + \ddot{\mathbf{x}}(t))\frac{\tau}{2}$$

$$\dot{\mathbf{x}}_c^{(1)}(t + \tau) = \dot{\mathbf{x}}(t) + (\ddot{\mathbf{x}}_p^{(1)}(t + \tau) + \ddot{\mathbf{x}}(t))\frac{\tau}{2}$$

$$\ddot{\mathbf{x}}_p^{(n)}(t + \tau) = \mathbf{F}(\mathbf{x}_c^{(n-1)}(t + \tau)) / m$$

The Verlet scheme

- L.Verlet. “*Computer Experiments on Classical Fluids*”. Phys.Rev. (The Physical Review) v.159, p.98-103, 1967
- Utilizing intuitive mechanical notations, we have, using central differences:

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}(t), \quad \frac{d^2\mathbf{r}}{dt^2} = \frac{\mathbf{F}(\mathbf{r}, t)}{m} := \mathbf{w}(\mathbf{r}, t(\mathbf{r})) = \mathbf{w}(\mathbf{r})$$

$$\mathbf{v}_n = \frac{\mathbf{r}_{n+1} - \mathbf{r}_{n-1}}{2\tau} + O(\tau^2), \quad \mathbf{w}_n = \frac{\mathbf{r}_{n+1} + \mathbf{r}_{n-1} - 2\mathbf{r}_n}{\tau^2} + O(\tau^2), \quad \mathbf{w}_n := \mathbf{w}(\mathbf{r}_n)$$

- Rearranging terms, we arrive at the standard Verlet method:

$$\mathbf{v}_n = \frac{\mathbf{r}_{n+1} - \mathbf{r}_{n-1}}{2\tau} + O(\tau^2), \quad \mathbf{r}_{n+1} = 2\mathbf{r}_n - \mathbf{r}_{n-1} + \mathbf{w}_n \tau^2 + O(\tau^4)$$

- The “velocity” Verlet scheme: See the discussion of Verlet methods in the book by A.L.Garcia

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n \tau + \frac{1}{2} \mathbf{w}_n \tau^2, \quad \mathbf{v}_{n+1} = \mathbf{v}_n + \frac{1}{2} (\mathbf{w}_n + \mathbf{w}_{n+1}) \tau$$

Advection equation: simple numerics

- Transport of some material along x -axis, e.g. some pollutant is spilled into a fast moving stream

$$u_t + au_x = 0, 0 < t < \infty, -\infty < x < \infty, a = \text{const}$$

(u is the concentration of pollution). Before the polluted fluid arrives, $u(x,t)=0$ at x , then u increases and decreases back to 0.

The initial distribution (Cauchy problem):

$$u(x, 0) = u_0(x), -\infty < x < \infty$$

Finite difference method: the finite intervals of $X [a,b]$ and $T [0,t]$:

$$a = x_0 < x_1 < \dots < x_M = b; 0 = t_0 < t_1 < \dots < t_N = T$$

Equidistant grids: $x_{j=a+jh}, j = 0, 1..M; t_n = n\tau, n = 0, 1..N$

$$h = \frac{b-a}{M} \quad - \text{spatial step}; \quad \tau = \frac{T}{N} \quad - \text{time step};$$

$u_j^n = u(x_j, t_n)$ - grid function defining the table (matrix): $j=0, 1..M;$
 $n=0, 1..N$

Advection equation numerics-II

• Taylor expansions:
(and analogous for $x - h, t - \tau$)

$$u(x + h, t) = u(x, t) + h \frac{\partial u}{\partial x} + \frac{h^2}{2!} \frac{\partial^2 u}{\partial x^2} + O(h^3);$$

$$u(x, t + \tau) = u(x, t) + \tau \frac{\partial u}{\partial t} + \frac{\tau^2}{2!} \frac{\partial^2 u}{\partial t^2} + O(\tau^3)$$

Thus:

and $\frac{\partial u}{\partial x} = \frac{u(x, t) - u(x - h, t)}{h} + O(h), u(x, t) - u(x - h, t) = h \frac{\partial u}{\partial x} + O(h^2)$

So: $\frac{\partial u}{\partial t} = \frac{u(x, t + \tau) - u(x, t)}{\tau} + O(\tau)$

$$\frac{u(x, t + \tau) - u(x, t)}{\tau} + a \frac{u(x, t) - u(x - h, t)}{h} = 0$$

or

$$\frac{u_j^{n+1} - u_j^n}{\tau} + a \frac{u_j^n - u_{j-1}^n}{h} = 0, x = x_j, t = t_n$$

- a discretized form of the advection equation

Various difference schemes for advection

- Explicit forward solution (for u_j^{n+1}):

$$u_j^{n+1} = u_j^n - \left(\frac{a\tau}{h} \right) (u_j^n - u_{j-1}^n), j = 1..M, \frac{a\tau}{h} \equiv k$$

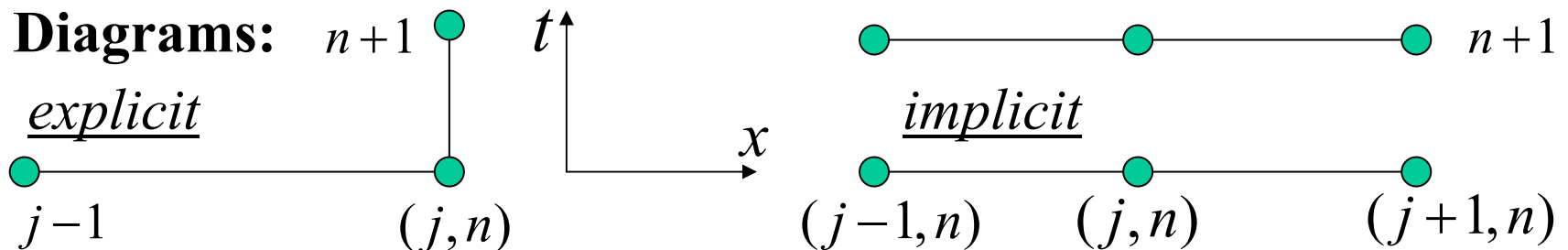
k (the Courant number) should be within the interval $[0,1]$

Outside of the stability region $0 \leq k \leq 1$ the error grows rapidly.

- Implicit difference scheme:

$$\frac{u_j^{n+1} - u_j^n}{\tau} + a \frac{(u_{j+1}^{n+1} - u_{j-1}^{n+1}) + (u_{j+1}^n - u_{j-1}^n)}{4h} = 0, j = 0, 1..M; n = 0, 1..N$$

This difference equation cannot be solved explicitly to determine u_j^{n+1} (there is more than one value for $t = t_{n+1}$)



The MacCormack scheme

The difference advection equation can be written with the help of the **predictor-corrector scheme**.

Predictor: $\bar{u}_j = u_j^n - k(u_{j+1}^n - u_j^n)$

Corrector: $u_j^{n+1} = \frac{1}{2}[u_j^n + \bar{u}_j - k(\bar{u}_j - \bar{u}_{j-1})]$

2D advection equation: $u_t + Au_x + Bu_y = 0$, $A=A(x,y)$, $B=B(x,y)$ – transport (advection) velocity components. The degenerate case:

$$u_t + Au_x = 0, u_t + Bu_y = 0$$

MacCormack 2D scheme:

$$\begin{aligned}\bar{u}_{ij} &= u_{ij}^n - \tau A \left(\frac{u_{i+1,j}^n - u_{i,j}^n}{h_1} \right); \bar{\bar{u}}_{ij}^{n+1} = \frac{1}{2} \left[u_{ij}^n + \bar{u}_{ij} - \tau A \left(\frac{\bar{u}_{ij} - \bar{u}_{i-1,j}}{h_1} \right) \right] \\ \bar{\bar{u}}_{ij} &= \bar{u}_{ij}^n - \tau B \left(\frac{u_{i,j+1}^n - u_{ij}^n}{h_2} \right); u_{ij}^{n+1} = \frac{1}{2} \left[u_{ij}^n + \bar{\bar{u}}_{ij} - \tau B \left(\frac{\bar{\bar{u}}_{ij} - \bar{\bar{u}}_{i,j-1}}{h_2} \right) \right]\end{aligned}$$

PDE modeling with MATLAB

- **MATLAB PDE Solver:** `pdepe` – solves initial/boundary value problems for parabolic and elliptic PDEs containing one space variable and time
- **PDE Helper:** `pdeval` – evaluates numerical solutions using the output of `pdepe`
- MATLAB expects PDE to be written in the form:

$$c\left(x, t, u, \frac{\partial u}{\partial x}\right) \frac{\partial u}{\partial t} = x^{-m} \frac{\partial}{\partial x} \left(x^m f\left(x, t, u, \frac{\partial u}{\partial x}\right) \right) + s\left(x, t, u, \frac{\partial u}{\partial x}\right)$$

where $m=0,1,2$ for slab, cylindrical or spherical symmetry respectively
 $f(x, t, u, u_x)$ is a flux term, $s(x, t, u, u_x)$ is a source term. Starting from this form, PDE can be coded to be used by `pdepe`:

`[c,f,s]=pdefun(x,t,u,dudx)`

The package PDE Toolbox comprises the region definition commands, the region discretization (grid generation), solving the respective SLE, and service functions

Example of PDE representation in MATLAB

- The 1D diffusion: $\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, 0 \leq x \leq 1, t \geq 0, a = 1 / \pi$
- Initial condition: $t = 0 : u(x, 0) = \sin(x / a)$
- Boundary conditions: $u(0, t) = 0; \frac{\partial u}{\partial x}(1, t) = -\frac{e^{-t}}{a}$
- The MATLAB-style equation: $\frac{1}{a^2} \frac{\partial u}{\partial t} = x^0 \frac{\partial}{\partial x} \left(x^0 \frac{\partial u}{\partial x} \right) + 0$

Here $m=0$ and

$$c = 1 / a^2, f = \partial u / \partial x, s = 0$$

Code:

```
function [c,f,s] = pdex1pde(x,t,u,DuDx)
```

```
c = pi^2;
```

```
f = DuDx;
```

```
s = 0
```

For the Poisson equation,
there is a special function
`poisolv`

The Fourier transform is
a very important tool for
solving PDEs. MATLAB
contains a set of functions
both for direct and inverse
Fourier transform

FEMLAB

- A product designed to assist engineers to solve complex PDE-based problems, helps to model heat transfer, fluid dynamics, chemical and electromagnetic phenomena, processes in structural mechanics
- Multiphysics: FEMLAB helps to model phenomena that involve simultaneously all the above disciplines, using the same GUI
- An example: analysis of a fuel cell might involve chemical reactions, electrical current distribution, heat transfer and fluid flow – all the facets of the core problem can be simulated together
- A synthetic modeling approach to engineering physics with a complex geometry is a strong point of FEMLAB
- Built on top of MATLAB, can be incorporated in MATLAB family
- An attempt to convert physics to code: one can start a model in FEMLAB, convert it to code, connect to other applications or tools
- FEMLAB is a constantly evolving system, trainings are offered
- www.femlab.com; www.comsol.com; www.comsol.se

FEMLAB: extended multiphysics modeling

- Application areas and respective modules:
 - Chemical Engineering: 3D and solid oxide fuel cells, catalytic converters, industrial reactors (porous reactor, turbulent reactor)
 - Fluid Dynamics: automotive, aeronautics, oil/gas, heat transfer, Navier-Stokes equation, two-phase flows, particle tracing
 - Electromagnetism: electric field simulations, waveguides, antenna design, electronics, optoelectronics, photonics, THz waves
 - Structural Mechanics: stress/strain analysis, engineering design
- New physical and mathematical modeling:
 - Nonlinear response, Landau-Ginzburg equation
 - Semiconductor device modeling
 - DNA chip modeling
 - Tokamak, particle tracing

Totally new areas such as nanotechnology are under consideration

Chapter 4. Case studies

The focus of this chapter is on application of mathematical methods and computer techniques studied in the preceding chapters to specific modeling situations. The main mathematical method used is solving differential equations

The logistic model

- The logistic model, despite its simplicity, was able to adequately represent the population dynamics in various countries, e.g. in England, Scotland and some parts of USA.
- In biology and ecology this model is used to describe various evolution scenarios, when the future population of species depends linearly on on the present population:

$$x_i(t+1) = \sum_j M_{ij} x_j(t), x_j(t) \geq 0$$

The diagonal terms of the matrix M represent individual growth rates, while the off-diagonal terms represent the interaction between species. Fixed amount of available resources corresponds to conservation laws. Evolution gets really competitive when the total population reaches its maximum value, $\sum_i x_i(t) = X$

The bugs are coming: population models

- Imagine a bunch of insects reproducing generation after generation. Initially there were x_0 bugs, after i generations there will be x_i of them (to bug us: the total mass of insects on the Earth grows faster than that of humans)
- When a population in a given region is sufficiently large, it may be represented by a real continuous variable, $x(t) > 0$.
- If we assume that there is a maximum sustainable population, $x = X$, (an equilibrium state), and that the population dynamics can be described by a single equation $dx/dt = v(x)$, then we can have a simple model $v(x) = ax(X-x)$, $a > 0$ and $dx/dt = ax(X-x)$ – this is called the logistic model. The discrete alternative is the logistic map:
 $x_{i+1} = ax_i(1-x_i)$, $0 < x_i < 1$, $a > 0$, $i = 0, 1, 2, \dots$

The logistic map is very generous, since any function having a nondegenerate extremum behaves in its neighborhood like this map near $x = 1/2$

- The relationship $x_{i+1}=f(x_i, a)$ is usually called the Poincaré map; the discrete variable i plays the role of time; generations may be interpreted as periods of motion. Fixed point of the map does not change with generations, $x_{i+1}=x_i$, so $\mu=dx_{i+1}/dx_i$ is a multiplier
 - In one-dimensional Poincaré recurrences f and a can be selected and normalized in such a way that if the initial data x_0 are taken from a finite interval $P=(\alpha, \beta)$, then the iterates $(x_1..x_n)$ belong to P
 - Thus the function f maps the interval P into itself – *endomorphism*
- The term *into* means that the iterates may not fill the entire interval.
- Endomorphism with a smooth inverse is a *diffeomorphism* (Slide 38)
- Usually, 1D endomorphisms are not invertible - in contrast to diffeomorphisms, so the past cannot be uniquely reconstructed. One might say that in such cases we deal with systems with unpredictable past (isn't it typical for human history?)

The logistic model: example of nonlinear growth

- The logistic equation can be solved by separating variables:

$$\int_{x_0}^x dx \frac{1}{x(X-x)} = a(t-t_0)$$

The quantity a is called control or growth parameter

Integration gives: $\frac{x}{X-x} = \frac{x_0}{X-x_0} \exp(aX(t-t_0))$
where $x_0 = x(t_0)$

- Usually, the solution is written in the form:

$$x(t) = \frac{Xx_0}{x_0 + (X-x_0)\exp(-aX(t-t_0))}$$

- The graph of this solution produces an S-shaped curve, which is called the logistic curve. In dimensionless units, it is represented by the function $x(t) = 1/(1+\exp(-t))$

Exercise c12. Using Maple and/or MATLAB, visualize the S-curve

Logistic models result in chaos

- A discrete model for bugs (or other insects) population expressed as a difference equation leads to somewhat unusual behavior.
- The initial population (known as the seed) produces only little effect on the population dynamics, which is mostly controlled by the growth parameter a . For some values of this parameter the population of insects changes chaotically, and for the chaotic regime there is an extreme sensitivity to the initial population

$$x_0 = x(t_0)$$

- For large enough values of the growth parameter, in our simple model $a > 3$ the population bifurcates into two (period doubling). Because the colony of insects then acts as if it were “attracted” by two populations, these solutions are called attractors.
- From a more general viewpoint, the problem consists in the evolution of an autonomous system described by vector equations $d\mathbf{x}/dt = \mathbf{F}(a, \mathbf{x})$, where the vector field \mathbf{F} depends on a parameter a , whose variation may radically change the system’s motion

How the disaster strikes: population extinction

Suppose that K individuals are removed each year, then the logistic equation becomes:

$$dx / dt = x(1 - x) - K = f(x, K)$$

where for simplicity the growth parameter $a = 1$. For $K < 1/4$ this equation has two equilibrium points; when $K > 1/4$ both equilibria disappear, but what is more important, $f(x, K) < 0$ for all values of the population x – *the population necessarily becomes extinct*.

On the contrary, for $K < 1/4$ the population never dies out. For the bifurcation point $K = 1/4$ the function $f(x, K) = 1/2$, which ensures the steady rise of the population.

This simple example demonstrates the importance of bifurcation phenomena. For $a \neq 1$, $K \rightarrow K / a$ and bifurcation occurs at $K = a/4$. Even in this 1D system, one can observe very complicated behavior: for large control parameter, both fixed points become unstable (chaos)

Computer implementation of logistic maps

- One can predict the population for two-cycle attractor by requiring that generation $(i+2)$ has the same number of species as generation i :
$$x = x_i = x_{i+2} = ax_{i+1}(1 - x_{i+1})$$

- Combining it with the main logistic map, we get

$$x = \frac{1 + a \pm \sqrt{a^2 - 2a - 3}}{2a}$$

- Exercise c13:** Produce a sequence of population values x_i

Plot the bifurcation diagram, i.e. generations x_i versus growth rate a . Observe the transition to chaos. Notice, whether parts of the first bifurcation diagram contain similar plots (self-similarity).

- It is remarkable that the logistic map can serve as a rough model of transition to turbulence. Like the eventual population in logistic models, turbulence practically does not depend on the initial state. The Reynolds number R plays the role of the growth parameter a

Computer implementation of logistic maps-II

- An example of a MATLAB code for a bifurcation diagram corresponding to discrete time dynamical system:

$$x_{i+1} = f(x_i), f(x) = ax(1-x), i = 0, 1, 2, \dots, 0 \leq x \leq 1$$

- The algorithm: for each value of a calculate the steady-state mode. Produce a sequence of iterations x_i as a function of a

```
>> % growth rate parameter range
```

```
>> ab=2.9; ae=4.0
```

```
>>% starting point
```

```
>>y=1-1/ab; if y<0 y=0.5; end; y
```

```
>>% preliminary steps to reach the steady state
```

```
>>M=300; N=500; da=(ae-ab)/(N-1); Y=zeros(M,N);  
B=[ab:da:ae]
```

Computer implementation of logistic maps-III

```
>> % Preparing starting points for growth parameter a
>> for n=1:N
>>     for: j=1:256, y=B.*y.*(1-y); end
>>     Y(1,n)=y;
>> end
>> % Calculation of dynamics for each value of parameter a
>> for m=2:M
>>     Y(m,:)=B.*Y(m-1,:).*(1-Y(m-1,:));
>> end
>> plot (B,Y, 'k', axis[ab ae 0 1])
>> xlabel('a'),ylabel('y')
```

You can find another variant in:

<http://www.math.usu.edu/~powell/biomath/mlab3-02/node4.html>

Applying Euler's method to logistic equation

- Recall that the explicit Euler's method is a recursive application of:
 $y_{n+1} = y_n + hf(t_n, y_n)$, where $t_{n+1} = t_n + h$ (see Slide 141)

How well does the Euler's method predict the solution of the logistic equation $dx/dt = ax(X-x)$? Without loss of generality, we may consider $dy/dt = ry(1-y)$, where $r = aX^2$, $y = x/X$, so that the max. population = 1

Applying the Euler's method, we get successive approximations:

$$y_{n+1} = y_n + hry_n(1-y_n), \quad n=0, 1, 2, \dots$$

Exercise c14: Using Maple, show that for $hr \ll 1$ the Euler solution smoothly approaches the analytic solution (logistic growth). Compare the cases $hr=0.2$ and $hr=0.8$

Exercise c15: Using Maple, show that when $hr > 1$, the population can exceed the maximum value (damped oscillations about $x=X$)

Applications of logistic models

- What will be the total profit of a company over several years, if the annual profit is a ?

The profit for the year i will be $x_{i+1}=ax_i$. For rather high profits, the latter are restrained and in a simplest case we get the logistic model

- What will be the output of some goods, e.g. cars by a certain firm?

The simplest model describing the yearly variation of the output, with the imposed constraint of limited resource would be a logistic map

- A generalization: any human activity subordinated to a constraint of limited resources and/or external factors is described by the logistic model. The limiting term may be interpreted as some influence of the future
- Study of turbulence: the regular, almost periodic character of motion is destroyed for some critical value of parameter – the Reynolds number; the multiplier μ passes the value $+1$

Commercial and engineering use of logistic maps

- If we make the transform $x_i = \lambda y_i + \mu, \lambda = -1/r, \mu = 1/2$ then we get from the logistic equation: $y_{i+1} = y_i^2 + c, c = a/2 - a^2/4$
This equation is used in graphical industry to obtain rather beautiful ornaments, which are in fact *fractals*. This is an example of direct market value of mathematics.
- The logistic model reflects such features of discrete algorithms as stability and chaotic behavior. From this viewpoint, it is interesting for numerical techniques and computational science
- The equation describing the energy evolution of a nonlinear oscillator with a self-excitation has the form (in dimensionless units):
$$\frac{dE}{dt} = aE(1-E)$$
 which is a continuous logistic map. Here the growth (multiplication) parameter a is close to unity

Engineering applications of dynamical systems

- Control theory: purposefully change some parameters to reach the required regime. Automatic pilot systems may be described by equation $d\mathbf{x}/dt=\mathbf{f}(\mathbf{x},t,\mathbf{c},\mathbf{p})$, \mathbf{x} is a vector depicting the phase state of the control system, \mathbf{c} is a vector defining construction parameters of the pilot, \mathbf{p} is a random vector characterizing external influence (e.g. wind, adversary actions, explosions, etc.)
- For a given trajectory of the aircraft, one can put $\mathbf{f}(0,t,\mathbf{c},0)=0$
- A perturbation at $t=t_0$ (e.g. wind) leads to deflection, $x(t_0)=x_0$
- Stability means that for $t_0 < t < T$ $x(t) \rightarrow 0$, T – the allowed reaction time. For $T \rightarrow \infty$ - asymptotic stability
- In the linear approximation, $d\mathbf{x}/dt=A\mathbf{x}$ and asymptotic stability condition is $(\mathbf{x},A\mathbf{x}) < 0$
- The dynamical system for an aircraft moving in the vertical plane can be written as: $d\varphi/dt=\rho-\cos\varphi$; $d\rho/dt=2\rho(\lambda-\mu\rho-\sin\varphi)$, $\lambda>0$, $\mu>0$

Contemporary applications of low-order dynamical systems

- Economics: demand-supply equilibrium - market clearance states with prices and quantities to be calculated. Stability of market equilibrium. Permanent revision and adjustment – tâtonnement processes. Physical economics (econophysics), financial and marketing models. Corporate planning.
- Investigation of chaos and dissipative structures
- System analysis, models of weakly formalized processes
- Competition models, military planning and battle analysis
- Study of morphogenesis and spatial order
- Exploration of stability and bifurcations; phase transitions
- Cooperative effects, collective excitation, Kuramoto model
- Meteorology and climate modeling (slow-fast decomposition)
- Laser research, optimization of laser parameters

Contemporary applications of low-order dynamical systems - II

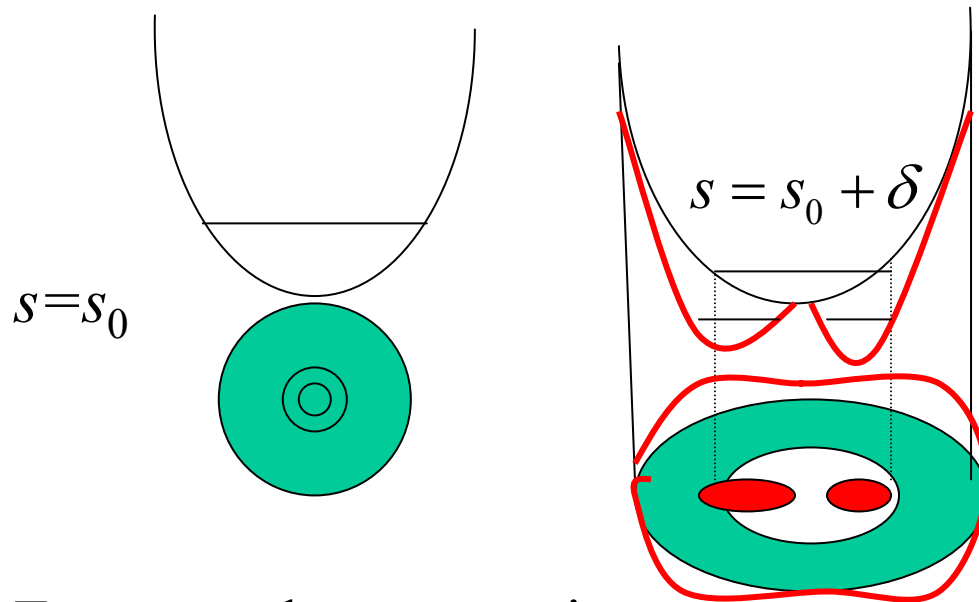
- Brain activity models, mathematical biology
- Demographic prognoses, “limits of growth” models, the world dynamics, modeling of ecological processes
- Epidemiological models, spread of infectious diseases
- Modern political analysis, dynamics of power distribution in hierarchical systems, mathematical sociology
- Information spread in the society, models of higher education
- Inertial confinement, laser initiation of thermonuclear reactions, thermal structures in various media
- Missile technology: non-stationary control theory, optimal control
- Aerodynamics, study of aircraft behavior
- Analysis of chemical processes and reaction rates, interaction of medicines and drugs with human body
- Extended (irreversible) thermodynamics, currents and flows

Classical mechanics

- The classical mechanics is almost exclusively based on the study of second-order dynamical systems
- A simple example: the motion of a particle in a potential well: $U(x,s)$, s – parameter (see next slide). This model is important in semiconductor technology and for molecular design
- The force acting on the particle is $F = -\frac{\partial U(x,s)}{\partial x}$, which means that the force profile changes with s .
- Equations of motion: $dx/dt=v$, $mdv/dt=F(x,s)$ have the first integral $E = mv^2 / 2 + U(x,s)$, representing the total energy, and a particular solution $x=0$, $v=0$.
- Thus, the model is reduced to the investigation of a family of curves depending on two parameters: s and E .
- One can judge about the stability of the equilibrium solutions by the behavior of the trajectories nearby. All closed curves surrounding singular points may be characterized by a so called Poincaré index

Phase portraits of a classical system

- The equilibrium solution is represented by a point $x=0, v=0$



Equilibrium positions of the system are points (a, s_0) with $\partial U(x, s) / \partial x|_{x=a} = 0$. The gradient of the first integral $E(x, s)$ is $(\partial U / \partial x, 0)$; it is zero only at equilibrium points.

For $s=s_0$ phase portraits are represented by ellipses, their size is growing with increased energy.

With the change of the parameter s , the form of the potential well is modified and new equilibrium positions appear. The equilibrium state $(x=0, v=0, s=s_0)$ becomes unstable.

Loss of stability

- With the change of the parameter s (see the preceding slide), the equilibrium state $x=v=0$ becomes unstable. Any small fluctuation results in transition to some other trajectory, which escapes from the equilibrium very rapidly (exponentially)
- Three types of trajectories near the unstable point: left minimum, right minimum, both minima – divided by a separatrix
- The idea of stability loss as a universal property of dynamical systems is generalized over the systems with many degrees of freedom: phase transitions (order parameter), turbulence (Re), lasers (feedback minus damping), social and economical models
- How many straws (N) can a camel carry? N is finite, than for $N+1=N_{crit}$ the camel's spine will probably break – loss of stability, a new state. This ancient example illustrates the transition from a gradual quantitative to abrupt qualitative change as a result of parameter variation. **The catastrophe theory** is a discipline engaged in studying such transitions (see R.Thom; V.I.Arnold)

Mechanics as modeling tool: a submarine trajectory

- Classical mechanics is not a boring scholarly discipline – it is primarily a powerful modeling tool
- A military action: let a submarine drifting at the depth h with a constant horizontal velocity v at the moment $t=0$ was given an order to surface urgently. The question is: what will be the trajectory of the submarine?
- Physics of the process: cisterns of the submarine must be rapidly freed from water and filled up with air, in order the average submarine density ρ should become less than the water density ρ_0
- If the time of the water purge phase is small compared with the characteristic period of motion, we can assume that at $t=0$ the lifting force F exceeding the gravity force P is instantly applied. The horizontal coordinate x changes with constant speed

$$dx / dt = v$$

Trajectory of a submarine-II

- Vertical motion of the submarine is governed by

$$\rho V \frac{d^2 y}{dt^2} = F - P = gV(\rho_0 - \rho)$$

- Solution:

$$x(t) = vt, y(t) = \frac{\rho_0 - \rho}{\rho} gt^2$$

- Surface time $t=s$:
$$y(s) = \frac{\rho_0 - \rho}{\rho} gs^2 = h, s = \left(\frac{\rho h}{g(\rho_0 - \rho)} \right)^{1/2}$$

- During this time the submarine would cover the distance:

$$L = vs = v \left(\frac{\rho h}{g(\rho_0 - \rho)} \right)^{1/2} \quad y = \frac{\rho_0 - \rho}{\rho v^2} gx^2$$

- The trajectory $y(x)$ is a parabola
- The main assumption here is that there are no more forces acting on the submarine, except F and P . This is true only for small speed, when we can neglect drag (see the book of A.P. Mikhailov and A.A.Samarski for details)

Trajectories of bodies

Trajectories are obtained from the law of motion $\mathbf{r}(t)=\{x_i(t)\}$ – a parametric curve, by excluding t .

Parabolic trajectory is always the result of a plain motion of a body, when it has the constant speed along one direction and is under the action of a constant force along other coordinate. **Examples:**

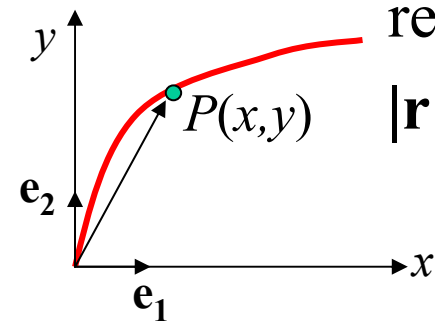
- Ballistics – motion of a projectile (with small drag);
- Flight of a body dropped from the height H with horizontal speed v
- Deflection of electrons passing between condenser plates, e.g. in image tubes, image converters, etc.

One can use geometric parameters to characterize the trajectory of a moving body: arc length, tangent vector, normal, curvature, binormal.

It is convenient to parameterize trajectory by the arc length s instead of time t , since $\left|\frac{d\mathbf{r}}{ds}\right|=1$, $\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{ds} \frac{ds}{dt} = v\boldsymbol{\tau}$, $\boldsymbol{\tau}$ - tangent vector.

Mathematical description of trajectories

- Let us consider a plane trajectory. Any point P on it may be represented by a vector $\mathbf{r} = x\mathbf{e}_1 + y\mathbf{e}_2$. The length of \mathbf{r} is



$|\mathbf{r}| = (\mathbf{r} \cdot \mathbf{r})^{1/2} = (x^2 + y^2)^{1/2}$, so the length of the curve is

$$s = \int_a^b ds = \int_a^b (\dot{x}^2 + \dot{y}^2)^{1/2} dt, \quad ds = |\mathbf{v}| dt, \quad |\mathbf{v}| = (\mathbf{v} \cdot \mathbf{v})^{1/2},$$

$\mathbf{v} = \dot{x}\mathbf{e}_1 + \dot{y}\mathbf{e}_2$ Here t is an arbitrary parameter,

e.g. time. In the natural parameter $t=s$, we have $|\mathbf{v}|=1$ ($d\mathbf{v}/dt := \mathbf{w}$)

and $\frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) = \frac{d}{dt}(|\mathbf{v}|^2) = 2\mathbf{v}\dot{\mathbf{v}} = 0$ i.e. velocity \mathbf{v} and acceleration \mathbf{w} are mutually perpendicular

- The curvature of a plane curve is defined as the absolute value of acceleration $k=|\mathbf{w}(s)|$, if the parameter is natural, $t=s$. Then we have

$$\frac{d\mathbf{v}}{ds} = \frac{d^2\mathbf{r}}{ds^2} = k\mathbf{n}, \quad \mathbf{n} = \frac{\mathbf{w}}{|\mathbf{w}|} = \left(\frac{d^2x}{ds^2}\mathbf{e}_1 + \frac{d^2y}{ds^2}\mathbf{e}_2 \right) \left[\left(\frac{d^2x}{ds^2} \right)^2 + \left(\frac{d^2y}{ds^2} \right)^2 \right]^{-1/2}$$

Stability of the Bohr-Rutherford atom

- One of the most popular models of the atom considers the heavy positive nucleus, with electrons moving in circular orbits around it.
- According to classical electrodynamics, accelerated charge must radiate electromagnetic waves, which means that rotating electrons would loose energy and eventually “fall down” to nucleus. The question is: when will it happen?
- The motion of the electron is described by the system of equations:

$$m\ddot{r} - m\omega^2 r = -e^2 / r^2; \frac{1}{r} \frac{d}{dt}(m\omega r^2) = -f$$

where the “radiative friction” is

$$\mathbf{f} = \frac{2e^2}{3c^3} \ddot{\mathbf{v}} = -\mathbf{e}_\varphi \frac{2e^2 \omega^3 r}{3c^3}$$

since for circular motion $\mathbf{v} = \mathbf{e}_\varphi \omega r$, $\dot{\mathbf{v}} = -\mathbf{e}_r \omega^2 r$, $\ddot{\mathbf{v}} = -\mathbf{e}_\varphi \omega^3 r$

So we have $\omega^2 \approx e^2 / mr^3$, $\dot{r} \approx -4cr_0^2 / 3r^2$, where $r_0 = e^2 / mc^2$ is the classical electron radius. This relationship reflects the Kepler’s law for planets: the rotation period $T = 2\pi/\omega = \text{const} \cdot r^{3/2}$

Stability of the Bohr-Rutherford atom-II

- The main assumption in this model is: the motion of electron is only weakly perturbed by the radiative force (it is slow compared to rotation, $|\ddot{r}| \ll \omega^2 r$). This is true down to very small wavelengths
- Integrating the approximate equations, we get $r^3(t) = a^3 - 4cr_0^2 t$ and the lifetime of the atom is ($a = 0.5 \cdot 10^{-8} \text{ cm}$, $r_0 = 3 \cdot 10^{-13} \text{ cm}$)

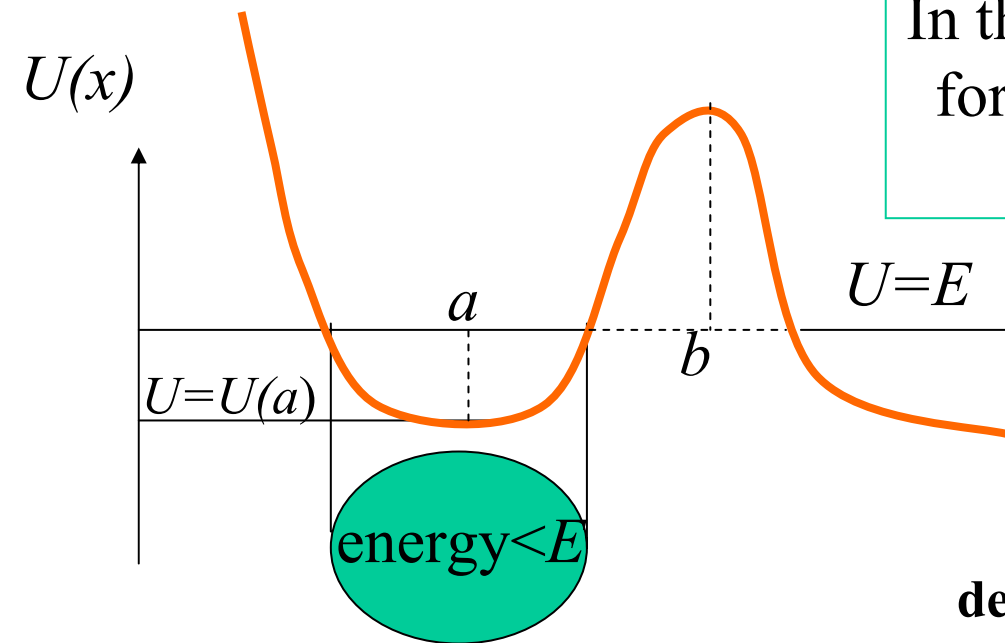
$$\tau = a^3 / 4cr_0^2 \approx 10^{-11} \text{ s}$$

- **This result demonstrates that atoms built according to the Bohr-Rutherford model could not exist for an observable time, so the model is unrealistic. Discrepancy of this mathematical model stimulated the appearance of quantum theory of atom - calculations of atomic structures must be based on Schrödinger equation (mathematical representation: the Hilbert space rather than phase space)**

Static models of atom made of point charges are equally impossible - electrostatic forces make such a configuration collapse

Oscillations

- Oscillations are finite motions that occur in the vicinity of equilibrium points
- If a is a local minimum of the potential $U(x,s)$, then $x=a$ brings the Lyapunov stability, i.e. for initial conditions $\{p(0),x(0)\}$ close to $\{0,a\}$ the whole phase trajectory $\{p(t),x(t)\}$ is close to $\{0,a\}$



In the vicinity of a, b : $y=(1/2)kx^2$;
for $k>0$ stable, for $k<0$ unstable
 $k=0$ – indifferent

Mathematical model of almost all systems is a multidimensional generalization of this situation:
 $x_{n+1}=Q(x_1..x_n)$, if Q is positive definite, then any small motion can be represented as a superposition of oscillations along main axes

Small oscillations, one degree of freedom

- Let $q=q_0$ be stable equilibrium, i.e. $\partial U / \partial q|_{q_0}=0, \partial^2 U / \partial q^2|_{q_0} > 0$

Then the solution of the system:

$$T = (1/2)a(q)\dot{q}^2, U = U(q)$$

is periodic for (q,p) near $(q_0, p=0)$. What is the period?

The period T_0 near equilibrium, with the decrease of the amplitude tends to the limit: $T_0 = 2\pi / \omega_0$, where

$$\omega_0^2 = \frac{b}{a}, b = \frac{1}{2} \frac{\partial^2 U}{\partial q^2} \Big|_{q=q_0}, a = a(q_0)$$

since for a linearized system $U = \frac{1}{2}b(q - q_0)^2$ and the solution is $q = A \cos \omega_0 t + B \sin \omega_0 t$

The qualitative picture of 1D small oscillations in the phase space (q, \dot{q}) can be plotted without solving equations: the relationship

$$\dot{q}^2 / 2 + bq^2 / 2 = \text{const} \quad \text{represents the ellipse}$$

Phase trajectories of the damped oscillator

- Free oscillations: $\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = 0$
- This system is autonomous (coefficients do not depend on time), but not conservative (damped oscillator)
- To obtain the phase portrait, let us introduce a new variable y :
 $\dot{x} = y, \dot{y} = -2\gamma y - \omega_0^2 x$
- The equation for integral curves: $dy / dx = -(2\gamma y + \omega_0^2 x) / y$

There is only one equilibrium point, $x=y=0$. To integrate, we can change variables, $y=xz$:

For small damping, $\gamma^2 < \omega_0^2$

the relation between y and x

$$\frac{zdz}{z^2 + 2\gamma z + \omega_0^2} = -\frac{dx}{x}$$

$$\sqrt{y^2 + 2\gamma xy + \omega_0^2 x^2} = C \exp\left(\frac{\gamma}{\sqrt{\omega_0^2 - \gamma^2}} \arctg \frac{y + \gamma x}{\sqrt{\omega_0^2 - \gamma^2} x}\right)$$

Phase trajectories of the damped oscillator-II

The speed of the phase point $\sqrt{\dot{x}^2 + \dot{y}^2} = \sqrt{y^2 + (2\gamma y + \omega_0^2 x)}$ is nowhere zero, except the origin of coordinates

To better understand the behavior of the curves, let us make a change of variables, $u = \sqrt{\omega_0^2 - \gamma^2} x, v = y + \gamma x$

Then the equation of phase trajectories can be written as

$$y^2 + 2\gamma xy + \omega_0^2 x^2 = u^2 + v^2 = C^2 \exp\left(\frac{2\gamma}{\sqrt{\omega_0^2 - \gamma^2}} \operatorname{arctg} \frac{u}{v}\right)$$

and, introducing polar coordinates

$$u = \rho \cos \varphi, v = \rho \sin \varphi,$$

we obtain $\rho = C \exp(\gamma \varphi / \sqrt{\omega_0^2 - \gamma^2})$

So, the integral curves of the free damped linear oscillator are logarithmic spirals about the equilibrium point, $u=v=0$ ($x=y=0$). Such equilibrium point is called a focus (which may be stable or unstable). For strong damping, $\omega_0^2 < \gamma^2$, the equilibrium point becomes a node

Example of oscillator-based model: a car wheel motion

- It is customary to model the behavior of a car wheel by the oscillator equation in the form:

$$\ddot{x} + 2\lambda\omega\dot{x} + \omega^2 = \omega^2 F(t), \text{ where } 2\lambda\omega = D/M, \omega^2 = K/M$$

M is the mass, K is the stiffness of the spring, D is the damping factor of the shock absorber, x is the vertical displacement of the suspended wheel under the assumption that the automobile body is immobile in vertical direction

- The solution would produce areas of satisfactory frequencies and damping, then the above relations show how to select the spring and shock absorber to get the required type of motion (e.g. the condition of oscillation suppression requires $D^2 \geq 4MK$ or $\lambda \geq 1$)

Exercise c16: Using Maple, find the solution to this problem, in particular the dependence of the amplitude of wheel vibrations on their frequency and parameters M , K , D

Normal oscillations

- A dynamical system supporting small oscillations can be decomposed into a direct product of oscillating systems with one degree of freedom
- From the algebraic point of view, the problem of small oscillations is that of a pair of quadratic forms, one of which - the kinetic energy - is always non-negative. These two forms can be simultaneously diagonalized (reduced to canonical form) by a linear transformation:

$$Q=Cq, \quad Q=(Q_1..Q_n)$$

- The geometrical meaning of such change of variables is: at first the Euclidean structure is introduced, where the kinetic energy form is a scalar product

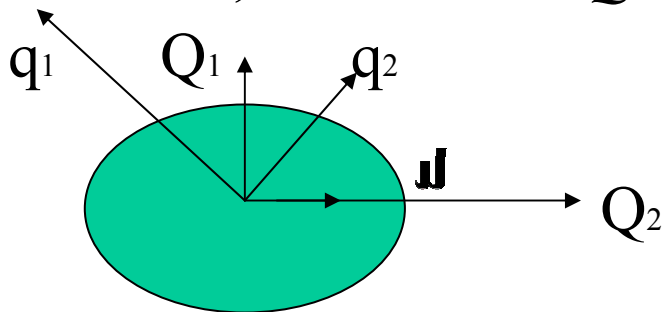
then the potential energy form is diagonalized by an orthogonal transformation:

$$\frac{1}{2}(A\dot{\mathbf{q}}, \dot{\mathbf{q}}) = \frac{1}{2}(\dot{\mathbf{Q}}, \dot{\mathbf{Q}}) \qquad U = \frac{1}{2} \sum_{i=1}^n \lambda_i Q_i^2$$

Natural modes

- The weights λ_i in the potential energy form are Eigenvalues. They satisfy the equation: $\det |B - \lambda A| = 0$. All the solutions of this equation are real, since both kinetic and potential energy matrices are symmetric.

- There may be three cases for each one-dimensional system:
 - $\lambda = \omega^2 > 0$, the solution $Q = C_1 \cos \omega t + C_2 \sin \omega t$ - oscillations
 - $\lambda = 0$, the solution $Q = C_1 + C_2 t$ - “null space”, $v = \text{const}$
 - $\lambda = -\beta^2 < 0$, the solution $Q = C_1 \cosh \beta t + C_2 \sinh \beta t$ - instability



$$\mathbf{q}(t) = \text{Re} \sum_{k=1}^n C_k e^{i\omega_k t} \boldsymbol{\eta}_k$$

The system has n natural modes alongside orthogonal direction Q_i . Any small oscillation may be represented as a sum of natural (normal) modes:

Example: two linked pendulums

Let two equal mathematical pendulums (length $l_1=l_2=l$, mass $m_1=m_2=m$) be connected by a massless spring whose length is equal to the distance between pivots. Let q_1, q_2 be the generalized coordinates denoting the angle to the vertical direction (that of gravity \mathbf{g}). Then $T = (\dot{q}_1^2 + \dot{q}_2^2) / 2$, $U = (1/2)[(q_1^2 + q_2^2) + \alpha(q_1 - q_2)^2]$ (we can choose the unit system where $g=1$) Let us find normal modes. If we put $Q_1 = \frac{q_1 + q_2}{\sqrt{2}}$, $Q_2 = \frac{q_1 - q_2}{\sqrt{2}}$, which gives the inverse transform

$$q_1 = \frac{Q_1 + Q_2}{\sqrt{2}}, q_2 = \frac{Q_1 - Q_2}{\sqrt{2}}$$

This linear transform is very helpful in many branches of physics and applied mathematics

then both forms would be diagonalized:

$$T = \frac{1}{2}(\dot{Q}_1^2 + \dot{Q}_2^2), U = \frac{1}{2}(\omega_1^2 Q_1^2 + \omega_2^2 Q_2^2), \omega_1 = 1, \omega_2 = \sqrt{1 + 2\alpha}$$

Two linked pendulums-II

- Analysis of normal modes:
 - 1) $Q_1 = 0$ i.e. $q_1 = q_2$ - synchronous motion of pendulums with the initial frequency $\omega = 1$, the coupling spring does not work
 - 2) $Q_2 = 0$ i.e. $q_1 = -q_2$ - pendulums are moving in opposite phase with the frequency $\omega_2 = \sqrt{1 + 2\alpha} > 1$, which is increased due to the spring action
- If the spring is very weak, i.e. $\alpha \ll 1$, then the energy is migrating from one pendulum to the other - “beats”.
Concurrent motion of pendulums: after time $T \approx \pi / \alpha$ only the pendulum that was initially at rest is moving, after $2T$ – again the first, etc. This is an example of competition of modes
- The problem of competing modes in two coupled oscillators is very important, it appears in many areas of engineering and science (laser resonators, circuitry, convective flows, biology)

Many linked pendula

- Consider N identical mathematical pendula whose pivots are placed equidistant along x -axis; all the pendula being coupled to each other by elastic springs. The system of motion equations:

$$ml^2 \frac{d^2 \varphi_i}{dt^2} = -mgl \sin \varphi_i + k(\varphi_{i+1} - \varphi_i) + k(\varphi_{i-1} - \varphi_i)$$

where φ_i is the angle with the vertical axis (y), m is the pendulum mass, l is its length, k is the elastic coupling constant. For simplicity, all these parameters are taken equal for all pendulums.

If $h=L/N$ (L is the system's length) tends to zero, we have

$$(\varphi_{i+1} - \varphi_i) - (\varphi_{i-1} - \varphi_i) = h^2 \frac{\partial^2 \varphi}{\partial x^2} \quad (\text{the continual limit})$$

and we get the nonlinear PDE for the angle φ :

$$\frac{\partial^2 \varphi}{\partial t^2} = -\omega^2 \sin \varphi + c^2 \frac{\partial^2 \varphi}{\partial x^2}, \quad \omega := \sqrt{\frac{g}{l}}, \quad c := \frac{h}{l} \sqrt{\frac{k}{m}}$$

This is the sine-Gordon equation
(c is the speed of sound wave)

The soliton solutions

- The sine-Gordon equation is one of the most important nonlinear models resulting in solitary wave solutions – solitons, “kinks”
- A number of physical effects, especially in condensed matter physics, lead to the sine-Gordon based models
- The sine-Gordon equation has solutions describing usual sound waves, but the most interesting is the soliton solution. The latter may be constructed using the map $\varphi = 4 \arctan u$. Differentiating, we get:

$$\frac{\partial \varphi}{\partial x} = \frac{4}{1+u^2} \frac{\partial u}{\partial x}, \quad \frac{\partial^2 \varphi}{\partial x^2} = \frac{4}{1+u^2} \frac{\partial^2 u}{\partial x^2} - \frac{8u}{(1+u^2)^2} \left(\frac{\partial u}{\partial x} \right)^2, \quad \frac{\partial^2 \varphi}{\partial t^2} = \frac{4}{1+u^2} \frac{\partial^2 u}{\partial t^2} - \frac{8u}{(1+u^2)^2} \left(\frac{\partial u}{\partial t} \right)^2$$

separating variables $u(x,t)=f(x)/g(t)$ we can seek solutions in the form $f(x)=\exp(x/a)$, $g(t)=\exp(Vt/a)$, where a and V are some constants

Finally, we obtain the soliton solution:

$$\varphi = 4 \arctan \left[\exp \frac{x - Vt}{a} \right]$$

The physical meaning of natural modes

- Introduction of normal coordinates and transition to normal (natural) modes has become a very important method in the XX century physics. The physical meaning of normal modes is collectively excited motions in the medium.
- Quasi-particles, elementary excitations, physics of condensed media, phonons, photons, etc. are the main notions spawning modern physical models. Without these notions, the description of macroscopic properties of materials would be impossible
- Quantization of an ensemble of independent oscillators led to the development of modern quantum theory, both in its mechanical and statistical parts.
- Study of motion instabilities has led to advanced theories of cooperative behavior, phase transitions, bifurcations, evolution, and self-organization. Applications of these theories are much wider than physics (biology, econophysics, financial modeling, social dynamics, structure formation, etc.)

- Transition to normal modes allowed P.A.M.Dirac, W.Pauli, P.Jordan, H.A.Kramers and other masters of modern physics to construct quantum mechanics and especially quantum electrodynamics – *a consistent theory of interaction of light with matter*.
What was the idea?
- In the Hamiltonian form, the motion equations of a harmonic oscillator are: $dq / dt = \partial \mathcal{H} / \partial p = p, dp / dt = -\partial \mathcal{H} / \partial q = -\omega^2 q$, where $\mathcal{H} = (p^2 + \omega^2 q^2) / 2$ is the Hamiltonian. If we multiply the motion equations by $\pm i\omega$ and combine them, we get $da / dt = -i\omega a, da^* / dt = i\omega a^*$, where complex amplitudes $a = (\omega q + ip) / \sqrt{2\omega}, a^* = (\omega q - ip) / \sqrt{2\omega}$ may be interpreted as vectors having the same length but rotating in opposite directions, $a(t) = a(0)e^{-i\omega t}, a^*(t) = a^*(0)e^{i\omega t}$. This representation was extensively studied in modern mathematics (Lee-Heisenberg-Weyl algebra). Now it is one of the main techniques.

- The Hamiltonian function may be expressed via normal modes a and a^* : $\mathcal{H} = (p^2 + \omega^2 q^2) / 2 = \omega a^* a$,

since $p = i\sqrt{\omega/2}(a^* - a), q = 1/\sqrt{2\omega}(a^* + a)$

In quantum mechanics, the normal modes are replaced by Hermite-conjugated operators a and a^+ , which are called annihilation and creation operators, since the Hamiltonian

$$\mathcal{H} = \sum_k (p_k^2 + \omega_k^2 q_k^2) / 2 = \sum_k \hbar \omega (a^+ a + 1/2)$$

consists of $3n$ identical independent quanta

$$\mathcal{H} = \sum_{k=1}^n \sum_{\lambda} (p_{k,\lambda}^2 + \omega_{k,\lambda}^2 q_{k,\lambda}^2) / 2 = \sum_{k=1}^n \sum_{\lambda} \hbar \omega_{k,\lambda} (a_{k,\lambda}^+ a_{k,\lambda} + 1/2)$$

Here parameter λ enumerates spatial modes, e.g. polarization

If a^+ is applied to the oscillator, then the number of quanta increases by 1 (creation), whereas a reduces their number (annihilation)

Systems with two degrees of freedom

- Models with already two degrees of freedom are much more complicated than those with one degree of freedom
- An autonomous system with two degrees of freedom can be described by a differential equation $\ddot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in E^2$, where \mathbf{f} is a 2D vector field
- The system is called **potential** if there exists a function $U : E^2 \rightarrow \mathbb{R}$ so that $\mathbf{f} = -\partial U / \partial \mathbf{x}$. Thus, equations of motion are $\ddot{\mathbf{x}} = -\partial U / \partial \mathbf{x}$

It is important that in the one-degree system one can always find the potential energy:

$$U(x) = -\int_{x_0}^x f(\xi) d\xi$$

For two-degree systems, one cannot, in general, find $U(\mathbf{x}) = U(x_1, x_2)$. For instance, the vector field $f_1 = x_2$, $f_2 = -x_1$ is not a potential one.

[A remark on terminology:](#) in classical mechanics the number of degrees of freedom is understood as a number of pairs (q, p) , whereas in the dynamical systems theory each first-order equation corresponds, by definition, to a degree of freedom

An example of a plane field

- Let \mathbf{f} lie in the plane (x,y) and be perpendicular to radius-vector \mathbf{r} in each point (being directed counterclockwise)
- Let the above force field be proportional to $r=|\mathbf{r}|$, i.e. $|\mathbf{f}|=ar$ and let a particle be in circular motion due to this force
- The elementary work, $dA = |\mathbf{f}| r d\varphi = ar^2 d\varphi$ gives the total work for the whole closed circle: $A = 2a\pi r^2$ proportional to the encircled area. Thus, A is non-zero for the closed path, and the tested field is not potential, $\mathbf{f} = -ay\mathbf{i} + ax\mathbf{j}$. This plane field has a rotational (rotor) character: $\text{curl}\mathbf{f} = 2a\mathbf{k}$ (sometimes *curl* is denoted as *rot*)
- If the medium is rotating as a whole with the angular velocity ω , then the velocity vector field is $\mathbf{v} = -\omega y\mathbf{i} + \omega x\mathbf{j}$ and $\text{curl}\mathbf{v} = 2\omega\mathbf{k} = 2\boldsymbol{\omega}$
- For plane fields, the complex analysis is the most convenient tool, with any vector being represented as $\boldsymbol{\xi} = \xi^\alpha \mathbf{e}_\alpha$; here $\xi^\alpha = x^\alpha + iy^\alpha$ are complex coordinates, \mathbf{e}_α - basis. The scalar product: $\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle = \sum_{\alpha=1}^n \xi^\alpha \bar{\eta}^\alpha$

Two-degree systems: phase space

- The motion equations in the plane can be written as

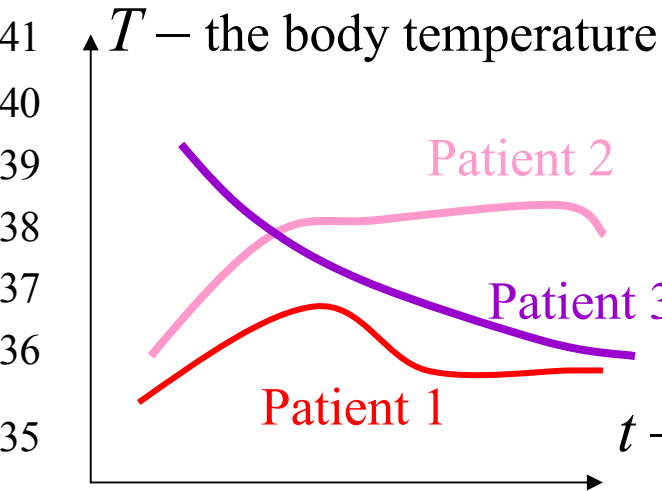
$$\dot{x}_1 = y_1, \dot{x}_2 = y_2 \quad \dot{y}_1 = -\frac{\partial U}{\partial x_1}, \dot{y}_2 = -\frac{\partial U}{\partial x_2}$$

This system determines the phase flux – phase space velocity in 4D space with coordinates (x_1, x_2, y_1, y_2) , which is called the phase space of the system with two degrees of freedom

- Phase curves of this system are subsets of the phase space, with the latter being subdivided by such curves whose projections onto the plane (x_1, x_2) produce trajectories (orbits) lying in the plane (x_1, x_2) . Phase curves cannot be intersected, while orbits, in general, can. Quantities (y_1, y_2) form the velocity space - a subspace of the phase space. From the mathematical point of view, configuration space (formed by coordinates) and velocity (momentum) space are equivalent, their equivalence is especially important in statistical physics

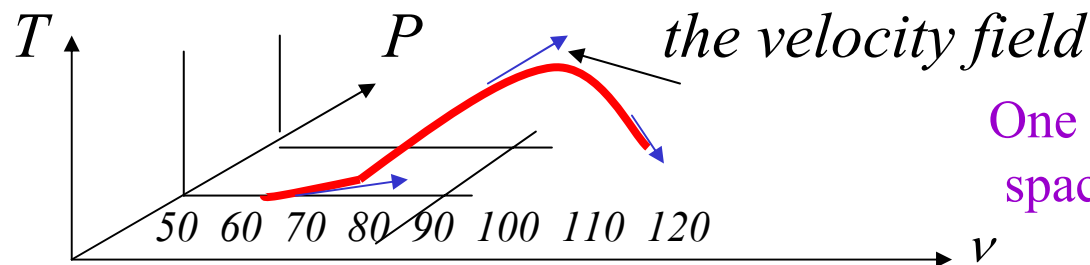
Example of a phase space: medical description of a patient

- 1D space describing the state of observed system – a patient



n patients can be described simultaneously by a point in $n \times 1D = nD$ phase space. This is a general idea: multiple objects may be represented by a single point

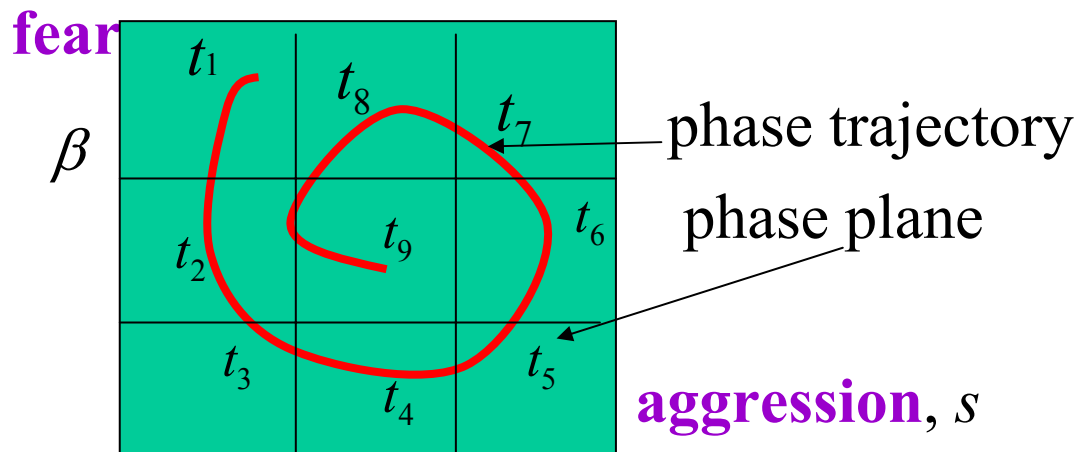
- 3D space of the patient states: the body temperature T , arterial pressure P , and heartbeat (pulse) frequency ν



One can always enhance the phase space by including new variables

Phase space of a dog

- The well-known ethologist Konrad Lorenz suggested to describe the emotional state of a dog by two parameters: mouth opening, s (measured in cm), proportional to aggression level and ear inclination angle, β , correlated with fear. To predict the animal behavior – from passive defense to vicious attack – 2D phase space proved to be sufficient.

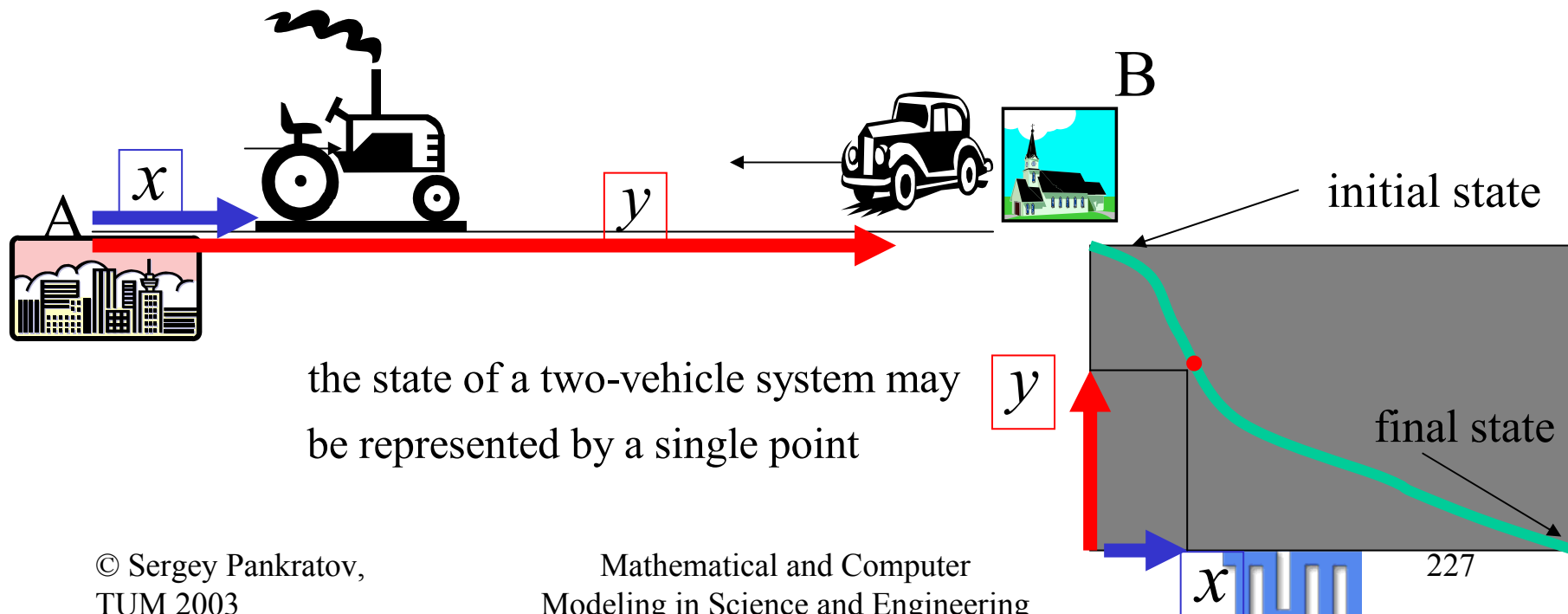


The next step would be to use the phase space methods for automatic human affective state recognition, See Proceedings IEEE, Sept.2003

The above picture may be said to represent the phase portrait of a dog – it describes the sequence of a dog's states.

Phase space leads to great simplification

- In a number of cases already the introduction of a phase space allows one to greatly simplify the problem. An example: the phase space of a system of two cars moving towards each other between points A and B. When the vehicles are speeding towards their destinations, the phase point is moving along the phase trajectory



The medium represented on the phase plane

- One can determine the state of N -particle system by indicating coordinates and momenta of each particle whose evolution is described by equation
$$\mathbf{r}_i = \frac{\mathbf{p}_i}{m_i}, \mathbf{p}_i = \mathbf{F}\left(\frac{\mathbf{p}_i}{m_i}, \mathbf{r}_i, t\right), i = 1, 2..N$$
- So the state of a system may be represented by a point in the $6N$ phase space. For 1D motion this space is reduced to $2N$ phase plane
- For non-interacting particles – though they may be submerged in an external field - this representation can be greatly simplified. The state of N -particle system is the direct product of N one-particle systems and consequently may be represented as “the cloud” of N points on the same phase plane (phase space)
- The mathematical reason for such simplification is that the laws of motion of different particles are particular solutions of the same equation
$$m\ddot{\mathbf{r}}_i = F(\dot{\mathbf{r}}_i, \mathbf{r}_i, t)$$

The medium on the phase plane-II

- The phase space representation is the most convenient when the external field is conservative and the system is autonomous. In this case the phase space (plane) is layered into non-intersecting phase trajectories along which N points representing the system are moving
- The Hamiltonian of such system is decomposed into a sum of separate Hamiltonians for each individual particle:

$$H(\mathbf{p}_1, \mathbf{r}_1; \mathbf{p}_2, \mathbf{r}_2 \dots \mathbf{p}_N, \mathbf{r}_N) = \sum_{i=1}^N H(\mathbf{p}_i, \mathbf{r}_i)$$

with the total energy being conserved

If the number of particles becomes very large, the mechanical system may be described as continuous medium. The transition to this description is effected by the introduction of phase density $f(\mathbf{r}, \mathbf{p}, t)$

The number of particles at the moment t between the phase points (\mathbf{r}, \mathbf{p}) and $(\mathbf{r}+d\mathbf{r}, \mathbf{p}+d\mathbf{p})$ is $dN=f(\mathbf{r}, \mathbf{p}, t)d^3r d^3p$

- The phase density or the distribution function in the phase space is the most important characteristic of the medium.

Distribution function

- Knowledge of the distribution function allows one to find all other characteristics of the medium. Evolution of the distribution function (for each sort of particles α) is governed by the kinetic equation:

$$\frac{\partial f_\alpha}{\partial t} + \frac{\mathbf{p}}{m_\alpha} \frac{\partial f_\alpha}{\partial \mathbf{r}} + \mathbf{F}(\mathbf{r}, \mathbf{p}, t) \frac{\partial f_\alpha}{\partial \mathbf{p}} = I_\alpha$$

where \mathbf{F} is the force acting on particles of given sort, I is the source of such particles in the phase space region $d^3r d^3p$. I is usually treated as “collision integral” in Boltzmann, Fokker-Planck or other form.

The distribution function $f(\mathbf{r}, \mathbf{p}, t)$ is, in fact, the average of the exact microscopic phase density

The equations for the exact

$$N_\alpha(\mathbf{r}, \mathbf{p}, t) = \sum_{i=1}^{N_\alpha} \delta(\mathbf{r} - \mathbf{r}_{i,\alpha}(t)) \delta(\mathbf{p} - \mathbf{p}_{i,\alpha}(t))$$

phase density $N_\alpha(\mathbf{r}, \mathbf{p}, t)$ are important for microscopic models of electromagnetic processes and are called the Klimontovich equations

Idealistic example of a distribution function §

Let us consider simple (1D) examples of phase space distributions.

1. All particles have the same energy E_0 (microcanonical ensemble)

There is only one phase space trajectory (shell) occupied

$$f(\mathbf{r}, \mathbf{p}, t) = f(E) = A\delta(E - E_0)$$

Let us find the mass density of the medium: $\rho(\mathbf{r}, t) = m \int f(\mathbf{r}, \mathbf{p}, t) d^3 p$

$$\rho(\mathbf{r}, t) = \rho(\mathbf{r}) = 2m^2 \int f(E) \frac{dE}{|p|} = \sqrt{2}m^{3/2} \int_{U(\mathbf{r})}^{E_{\max}} \frac{f(E)dE}{\sqrt{E - U(\mathbf{r})}} = \frac{\sqrt{2}Am^{3/2}}{\sqrt{E_0 - U(\mathbf{r})}}$$

the medium is spread over such spatial regions where the interaction potential is lower than E_0 . In the points (on the surface) $U(\mathbf{r}) = E_0$

the medium density tends to infinity. This is an indication to the fact that microcanonical distribution is a physical idealization

A realistic equilibrium distribution is given by the Gibbs ensemble, a more detailed discussion of these important subjects can be found in statistical theory

The Boltzmann distribution

- The canonical distribution (Maxwell-Boltzmann-Gibbs) is defined as $f(E) = Ce^{-\beta E}$. For 1D case we have

$$\rho(x) = mCe^{-\beta U(x)} \int_{-\infty}^{\infty} e^{-\beta \frac{p^2}{2m}} dp = Cm(2\pi m / \beta)^{1/2} e^{-\beta U(x)}$$

- The Maxwell-Boltzmann distribution is singled out in collision processes, when the energy of particles is redistributed. Any phase space distribution in the collisional media becomes non-stationary, except this one – it remains stationary (the collision functional $I=0$)
 - Consider the weak gravitation field
- We can find the mass density distribution
- $$U(x) = \begin{cases} gx, & 0 \leq x < \infty \\ \infty, & -\infty < x < 0 \end{cases}$$
- in the atmosphere submerged in such field:

$$\rho(x) = m \int_{-\infty}^{\infty} Ce^{-\beta mgx - \beta \frac{p^2}{2m}} dp = 2\sqrt{\frac{\pi}{\beta}} Cm^{3/2} e^{-\beta mgx}$$

Barometric relations

- The last expression is called **the barometric formula**. Let us find the pressure $P(h)$ at the height h . It is equal to the force acting on a column $h < x < \infty$:

$$P(h) = g \int_h^{\infty} \rho(x) dx = 2C \sqrt{\pi m} \beta^{-3/2} e^{-\beta m g h} = \frac{\rho(h)}{\beta m}$$

So, as usual, pressure is proportional to the local density. We can also find the average kinetic energy of atmospheric particles at the height h

$$\bar{T} = \frac{C e^{-\beta m g x}}{\rho(x)} \int_{-\infty}^{\infty} \frac{p^2}{2m} e^{-\beta \frac{p^2}{2m}} dp = -\frac{C e^{-\beta m g x}}{\rho(x)} \frac{\partial}{\partial \beta} G(\beta), G(\beta) = \int_{-\infty}^{\infty} e^{-\beta \frac{p^2}{2m}} dp = (2m)^{3/2} \frac{\sqrt{\pi}}{\sqrt{\beta}}$$

or $\bar{T} = \frac{1}{2\beta}$.

The average kinetic energy is independent of the height h and is determined solely by the parameter $\beta = 1/\theta$ of the distribution.

Very important parameter θ is called the temperature, β is called the inverse temperature. θ has the dimensionality of energy and can be measured in any energy units. Pressure is proportional both to ρ and θ

Working example of a distribution function:

Fermi distribution

§

- Consider $f(E) = B\eta(E_0 - E)$, $\eta(x) = \begin{cases} 1, x > 0 \\ 0, x < 0 \end{cases}$ - the Heaviside function
Calculation of the density gives

$$\rho(\mathbf{r}) = \int_{U(\mathbf{r})}^{E_0} \frac{dE}{\sqrt{E - U(\mathbf{r})}} = 2\sqrt{2}Bm^{3/2}(E_0 - U(\mathbf{r}))^{1/2}$$

for 1D case and

$$\rho(\mathbf{r}) = \frac{4\pi B}{3} m (2m(E_0 - U(\mathbf{r})))^{3/2} \quad \text{in 3D case}$$

In this example the medium is spread over the same region as for microcanonical ensemble but the density tends to zero on the energy surface. The Fermi distribution plays a crucial role in the condensed matter theory.

The constant B is inversely proportional to the Planck constant h

Example of two-degree systems: spherical pendulum

- The potential energy $U = \frac{1}{2}(x_1^2 + x_2^2)$ gives a plane motion
- Isolines on (x_1, x_2) -plane – concentric circles
- Equations of motion: $\ddot{x}_1 = -x_1, \ddot{x}_2 = -x_2$ are equivalent to a system $\dot{x}_1 = y_1, \dot{y}_1 = -x_1$ and $\dot{x}_2 = y_2, \dot{y}_2 = -x_2$
- This system breaks out into two independent subsystems, which means that the motion along each coordinate x_1, x_2 is equivalent to that of one-degree system
- Solutions:
$$x_1 = a_1 \cos t + b_1 \sin t, y_1 = -a_1 \sin t + b_1 \cos t$$
$$x_2 = a_2 \cos t + b_2 \sin t, y_2 = -a_2 \sin t + b_2 \cos t$$
- The energy conservation:

$$E = \frac{1}{2}(y_1^2 + y_2^2) + \frac{1}{2}(x_1^2 + x_2^2) = \text{const}$$

One more example of the plane motion: the Lissajou figures

- Small oscillations with two degrees of freedom:

$$\ddot{x}_1 = -x_1, \ddot{x}_2 = -\omega^2 x_2$$

- The potential energy: $U = \frac{1}{2}(x_1^2 + \omega^2 x_2^2)$

- The energy conservation: $\frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2) + U(x_1, x_2) = E$

It means that the motion is contained within the ellipse $U(x_1, x_2) \leq E$

The system is also broken into two independent 1D systems.

The conservation law should be fulfilled separately for each of them:

$$E_1 = \frac{1}{2}(\dot{x}_1^2 + x_1^2), E_2 = \frac{1}{2}(\dot{x}_2^2 + \omega^2 x_2^2), E_1 + E_2 = E$$

which means that the motion along each coordinate is confined by

$|x_1| \leq \sqrt{2E_1}, |x_2| \leq \sqrt{2E_2} / \omega$ - a rectangular inside the ellipse:

$$U(x_1, x_2) \leq E$$

Lissajou figures – a MATLAB code

- The solution: $x = a_1 \cos(\omega_1 t + \varphi_1), y = a_2 \cos(\omega_2 t + \varphi_2)$
- Free oscillations in mutually perpendicular directions – solutions of a free undamped oscillation equation:
- $X = a_1 \cos(\omega_1 t)$ Here, for simplicity, the phases are put to zero
- $Y = a_2 \cos(\omega_2 t)$
- From the mathematical point of view – one has to obtain the graph of a function written in parametric form
- If ω_1/ω_2 is a rational number then the produced graph is called a Lissajou figure
- The form of the Lissajou figures crucially depends on frequency and phase difference. If the frequencies of the two oscillations are close, the system strives to the spherical pendulum
- The circle is the simplest Lissajou figure ($\omega_1 = \omega_2, a_1 = a_2, \varphi_1 - \varphi_2 = \pi/2$)

2D Lissajou patterns

- **A MATLAB code:**
- % amplitudes and frequencies
- >>a1=1.2;
- >>a2=1.0
- >>w1=1.5
- >>w2=1.0

For simple operations, the interactive MATLAB mode is sufficient, but for repeated (looped) computations or for “what if” analysis, one should use m-files

Here m-files allow one to create a series of circles, ellipses and Lissajou figures by varying the parameters a1, a2, w1, w2

- % Operator **whos** allows one to obtain the full information on all % active variables. All variables remain active if they are not removed by the operator **clear**
- >>whos % Try it. Pay attention to the fact that constants are % interpreted as 1x1 matrices
- % Using the two 1st harmonics would produce ellipses, circles, lines (degenerate ellipses), 2nd harmonics produce “eights”, etc.

2D Lissajou patterns (continued)

- % Operator of t-values vector
- >>t=0:0.1:3.2
- % computation of vectors x,y:
- >>x=a1*cos(w1*t);
- >>y=a2*cos(w2*t);
- % Making a plot y(x):
- >>plot(x,y)
- % Change the time range:
- >>t=0:0.1:10;

A more general form of parametric equations producing the Lissajou figures would be:

$$x=a1*\cos(f(t)); y=a2*\sin(g(t))$$

One can also construct the n-dimensional generalization of Lissajou figures, which would produce spirals and cyclic attractors. Lissajou patterns are important for studies of oscillations and waves in vector field models

- % You can use the up arrow to spare yourself extra typing
- **EXERCISE Matlab1:** Compute and plot Lissajou patterns for three different ratios $w1/w2$ and for two sets of amplitudes $a1$ and $a2$
- **EXERCISE Matlab2:** Include phases into cosine arguments
- **EXERCISE Matlab3:** Generalize Lissajou patterns on 3D case
- **Exercise c17:** What pattern would be displayed on a screen for 180 degrees difference between x and y channels?

Beats – an example

- A periodically driven oscillator: $m\ddot{x} + kx = f(t) = A \sin \omega t$
- If the driving force frequency is near resonance, we have:

$$\begin{aligned} x(t) &= 1/m\omega_0 \int_0^t \sin \omega_0(t-\tau) f(\tau) d\tau = \frac{2A \sin \frac{\omega_0 - \omega}{2} t}{m(\omega_0 - \omega)(\omega_0 + \omega)} \cos \frac{\omega_0 + \omega}{2} t + \frac{A \sin \omega_0 t}{2m\omega_0^2} \\ &\quad - \frac{A \sin \frac{\omega_0 - \omega}{2} t}{m\omega_0(\omega_0 - \omega)} \cos \omega_0 t + \frac{A \sin \omega_0 t}{2m\omega_0^2} \\ &\approx \frac{2A \sin \frac{\omega_0 - \omega}{2} t}{m\omega_0(\omega_0 - \omega)} \cos \omega_0 t + \frac{A \sin \omega_0 t}{2m\omega_0^2} \end{aligned}$$

The first term may be interpreted as a harmonic oscillation with a slowly varying amplitude from 0 to $\frac{A}{m\omega_0(\omega_0 - \omega)}$ with a period $T = 2\pi / |\omega_0 - \omega|$

Such oscillations are called beats, they are always the result of interference of two or more oscillations. Beats lead to periodic energy migration from one oscillation to another. Beats are important for models based on coupled oscillators (competing modes). The concept of beats leads to quasiperiodic motion, with small amplitude change

Beats (a MATLAB code)

$$y(t) = a_1 \cos w_1 t + a_2 \cos w_2 t$$

```
>>a1=1.0; % amplitudes of harmonic
```

```
>>a2=1.0 % oscillations
```

```
>>w1=1.0; % frequencies of harmonic
```

```
>>w2=1.2; % oscillations
```

```
>>T0=0; % the initial moment
```

```
>>tm=20; % the final moment
```

```
>>N=600; % the number of computed points
```

```
>>T=tm-t0; % observation time
```

```
>>dt=T/N; % discretization step
```

```
>>t=t0:dt:tm; % time vector
```

```
>>y=a1*cos(w1*t)+a2*cos(w2*t); % the beat function
```

```
>>plot(t,y); % graphical output
```

% Despite apparent simplicity, this problem is very interesting.

% One can study the influence of higher harmonics, $2w_1$, $3w_1$,...

Beats (continued)

- How to change the axis scale?
- For instance, we would like to see the details in the other time range, $1 < t < 2$
- The command **axis([xmin xmax ymin ymax])**
- **Exercise Matlab4:** Write the equation of the modulating curve passing through the maximums when
a) $a_1 = a_2$; b) $a_1 = c \cdot a_2$, $c > 1$. Analyze the dependence on c .
- **Exercise Matlab5:** Plot the beats for $w_1 \ll w_2$. How does the plot depend on $u = w_1/w_2$?
- **Exercise Matlab6:** Use MATLAB to study the influence of second and third harmonics on the shape of oscillations?

One can hear beats using a musician's tuning fork. Now the so called quantum beats are extensively studied in relation to nanotechnology. Beats are ubiquitous in laser physics and nonlinear optics. Beats are also an important factor in EEG studies of the brain. Beating is closely connected with a fruitful idea of separation of any motion into slow and fast components; such separation is the main tool of nonlinear dynamics

Pendulum motion (numerical approach)

- The equations of motion of a simple pendulum:

$$d\omega/dt = \alpha(\theta) \quad d\theta/dt = \omega$$

where $\alpha(\theta) = -(g/l)\sin\theta$

is the angular acceleration. For numerical treatment, one may use the Euler method (iterations):

$$\omega_{n+1} = \omega_n + \tau\alpha_n \quad \theta_{n+1} = \theta_n + \tau\omega_n$$

If one is interested only in the evolution of the coordinates (angle), then the Verlet method can be applied:

$$\theta_{n+1} = 2\theta_n - \theta_{n-1} + \tau^2\alpha_n$$

It is convenient to use dimensionless units measuring the period (time scale) in $\sqrt{l/g}$. The two parameters l and g appear in the oscillator problem only as their ratio, which can be set to unity, so the harmonic amplitude period is 2π .

Pendulum as a dynamical system

- The equations of pendulum motion can be written as

$$dp / dt = -F \sin \theta, \quad d\theta / dt = Gp$$

where $F = mgl$, $G = 1 / ml^2$ F is the maximum value of potential energy. The deflection angle θ and the angular momentum p are canonically conjugated quantities that must satisfy Hamiltonian equations. The Hamiltonian is: $H(p, q) = \frac{1}{2} Gp^2 - F \cos \theta$

Since the system is autonomous and non-dissipative, $H(p, q) = E$ is the integral of motion. This fact means that the momentum depends on coordinate but not on time, $p = p(q, E)$, where E is the constant whose value determines the total energy of the system. Here $q = \theta$ – one can see that this system may be strongly nonlinear, and for different E one would observe totally different phase trajectories. Regions with different types of trajectories are divided by *separatrix*

Pendulum motion as a dynamical problem

- The parametric dependence of p and q on time can be found from the Hamiltonian equation:

$$dt = \frac{\partial q}{\partial H / \partial p}$$
or

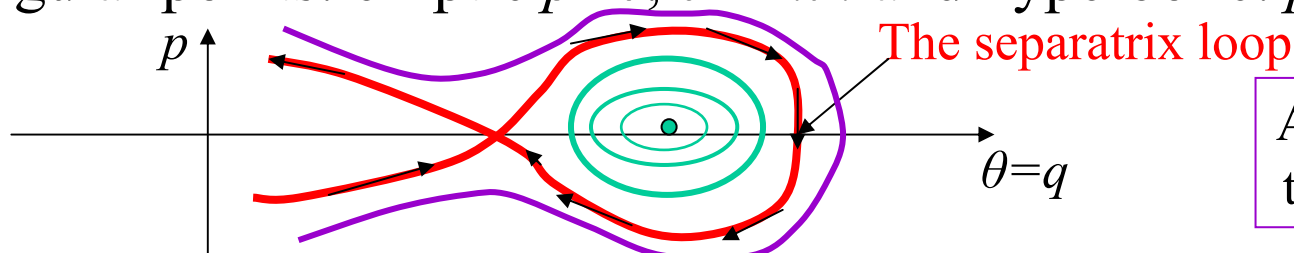
$$t = \int_{q_0}^q \frac{dq}{\partial H / \partial p}$$
- integral of this type always gives the period of a finite motion
- Since $H(p,q)$ depends only on p and q , which are interrelated by $p=p(q,E)$, finding the solution of a dynamical problem is reduced to numerical computation of the above integral.
- However, one can analyze the behavior of autonomous system even without integration. In the case of pendulum, if $E > F$, the momentum p never equals to zero, which results in always rising θ i.e. rotation. For $p > 0$ the pendulum rotates clockwise. If $E < F$, the motion is finite within the potential well – oscillations. Thus, the parameter E (energy) dictates the character of behavior.

Pendulum motion as a dynamical problem-II

What happens when the pendulum total energy equals mgl ?

- In this case $E=F$. Does the pendulum oscillate or rotate?
- The period of oscillations becomes infinite – motion over a separatrix (a curve that separates finite and infinite motion)
- The phase portrait contains two equilibrium points ($p=0$): one at the coordinate origin ($\theta = 0$), which is stable (elliptical) another at $\theta=\pm\pi$, which is unstable (hyperbolic)
- Phase trajectories look like ellipses in the neighborhood of an elliptic point and like hyperbolas near a hyperbolic one.

Pendulum as a dynamical system contains an infinite number of singular points: elliptic $p=0$, $\theta=2\pi n$ and hyperbolic: $p=0$, $\theta=(2n+1)\pi$



Arrows determine the path direction

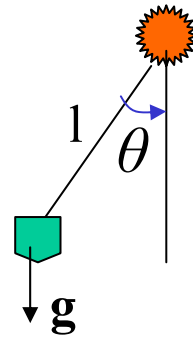
Pendulum (a MATLAB code to compute the time evolution of an oscillator)

- Select the numerical method to use: Euler or Verlet
- Set the initial position θ_0 and the initial velocity ω_0
- Set the constants and variables
- Take one backward step (Verlet)
- Make a loop with given time step:
 - record angle and time
 - compute new position and velocity
 - test the passage of equilibrium position $\theta = 0$
 - if yes, estimate the period
- Plot the oscillations as $\theta = \theta(t)$. In many cases it is convenient to replace the procedure **plot** by that of **comet** – the latter allows one to draw the “running” point on the trajectory. The ultimate result may be represented as a set of curves in a phase plane (θ, ω)

It is convenient to use for oscillation analysis dimensionless variables, e.g. time $t = (l/g)^{1/2} \tau$, see below

MATLAB code for pendulum oscillations

```
% Select the numerical method to use: Euler or Verlet
>> numericalMethod = menu ('Select a numericalMethod:','Euler','Verlet',...)
% Set initial position and velocity
>> theta0 = input ('Enter initial angle (in degrees):');
>> theta = theta0*pi/180; %Convert angle to radians
>> omega = 0; % Set the initial velocity
% Set the constants and variables
>> g_over_l = 1;
>> time = 0; % initial time
>> irev = 0; % count number of reversals
>> tau = input ('Enter time step:');
```



Similar codes may be written, of course, for any kind of oscillating system, e.g. for a spring-block with or without damping (friction). The same applies to electrical and electronic circuitry

% The angle theta is sufficient to be defined in certain finite limits, for instance varying from $-\pi$ to $+\pi$, these points being identified

MATLAB code for pendulum-II

```
% Take one step backward to start Verlet
>> accel = -g_over_1*sin(theta); %gravitational acceleration
>> theta_old = theta - omega*tau + 0.5*tau^2*accel;
% loop over steps with given time step and numerical method;
>> nstep = input ('Enter number of time steps: ');
>> for istep=1:nstep
% Angle and time for plotting
>> t_plot (istep) = theta*180/pi; %convert to degrees
>> time = time + tau
% Compute new position and velocity. The main idea of numerical
% calculations is simple: knowing the actual values of position and
% velocity, one can find their values in small interval tau
>> accel = -g_over_1*sin(theta); % gravitational acceleration
```

**For this case an analytic solution exists, see e.g.
L.D.Landau,E.M.Lifshits. Classical Mechanics**

MATLAB code for pendulum-III

```
>> if(NumericalMethod == 1)
>>     theta_old = theta; % save previous angle
>>     theta = theta + omega*tau; % Euler method
>>     omega = omega + accel*tau
>> else
>>     theta_new = 2*theta - theta_old + accel*tau^2;
>>     theta_old = theta; % Verlet method
>>     theta = theta_new;
>> end
```

For a specific case $d^2\theta/dt^2 + \sin\theta = 0$ (in dimensionless variables), the main algorithm is: $\theta_2 = \theta_1 + \omega_1 * dt$, $\omega_2 = \omega_1 - \sin(\theta_2) * dt$, then the initial conditions should be updated: $\theta_1 = \theta_2$, $\omega_1 = \omega_2$. It would be interesting to explore, from what time step dt the phase curve becomes distorted

Oscillations: test of passing $\theta=0$ and estimating the period

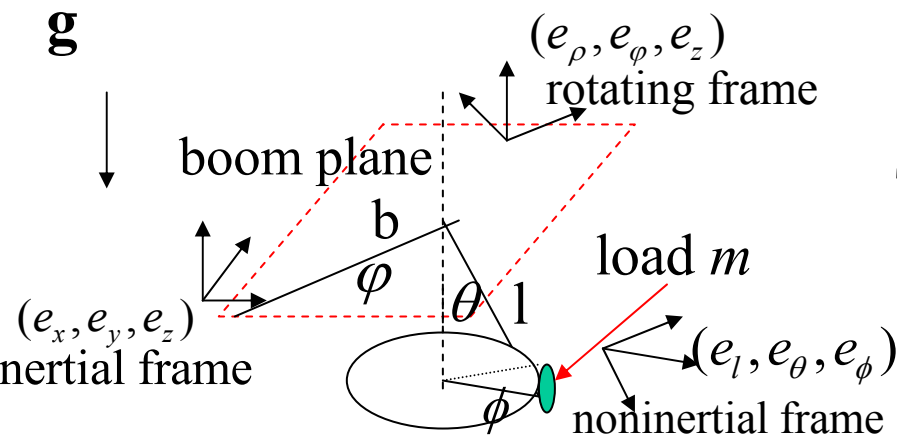
```
if(theta*theta_old<0) % test position
    fprintf('Turning point at time t= %f\n', time);
    if( irev == 0 ) % this is the first reversal
        time_old = time;
    else
        period(irev) = 2*(time - time_old);
        time_old = time;
    end
    irev = irev + 1
end
end
AvePeriod = mean (period);
Error = std(period)/sqrt(irev);
fprintf('Average period = %g+/- %g\n', AvePeriod, Error);
plot(t_plot,th_plot,'+');
```

See the book by A.L. Garcia for details

It would be instructive to test the accuracy of energy conservation for any grid point. It is convenient to plot the relationship $E=E(x)$ in a subplot in a dedicated window containing map (E,x) Such techniques demonstrate the value of conservation laws for numerical modeling

An example of pendulum modeling: tower crane

- Rotary (tower) cranes are ubiquitous in construction, their stability being the crucial element of safety.
- The dynamics of a crane may be modeled by a point mass (load) suspended by a massless cable from a horizontally moving support. Such system is a spherical pendulum, but in general in non-inertial reference frame



The potential energy of the load is

$$U = -mgl \cos \theta$$

The solution is obtained by writing the Lagrange equations for $L=T-U$ (a regular method), which gives differential oscillation equations,

and analyzing limiting cases. The stability of the crane (linearization!) critically depends on the parameter b/l

The driven pendulum

Damped oscillations under periodic external force (in dimensionless form): $d^2x/dt^2 + 2\lambda dx/dt + \sin x = f \cos x \cos \Omega t$

Small (harmonic) oscillations: $\sin x \approx x$

$$d^2x/dt^2 + 2\lambda dx/dt + x = f \cos \omega t$$

Consider a stationary mode (no transients)

Solution (in complex form):

The amplitude as a function of frequency: $x = Ae^{i\omega t} + A^*e^{-i\omega t}$

The real part of the solution:

$$A = \frac{f}{2(1 - \omega^2 + 2i\lambda\omega)}$$

$$x = a \cos(\omega t + \varphi), a = 2|A| = f / \sqrt{(1 - \omega^2)^2 + 4\lambda^2\omega^2}$$

$$\varphi = \arctg \frac{\omega^2 - 1}{2\lambda\omega}$$

Here the solution is written in a scaled form, $\omega \rightarrow \omega/\omega_0$.
To restore the explicit dependence on natural frequency,
one has to write $(\omega_0)^2 - \omega^2$ in place of $(1 - \omega^2)$

When damping tends to zero we get the quasi-periodic motion (beats)

$$x = f(\omega_0^2 - \omega^2)^{-1} (\cos \omega t - \cos \omega_0 t), x = \dot{x} = 0, t = 0$$

Driven anharmonic oscillations

The next term in $\sin x$: $d^2x/dt^2 + 2\lambda dx/dt + x - x^3/6 = f \cos \omega t$
(the Duffing equation) leads to additional functions

$$x^3/6 = (A^3 e^{3i\omega t} + 3A^2 A^* e^{i\omega t} + 3AA^{*2} e^{-i\omega t} + A^{*3} e^{-3i\omega t})/6$$

The terms containing $\exp(\pm 3i\omega t)$ result in small corrections of higher order. This can be interpreted as the appearance of a new force proportional to $\cos(3\omega t)$. This is an effect of harmonic generation.

The frequency of third harmonics is far from the resonance, therefore their contribution is small as compared with the principal frequency.

Now, instead of linear case for the amplitude $A(\omega)$, one has nonlinear relationship $(1 - \omega^2 + 2i\lambda\omega + |A|^2/2)A = f/2$ and

$$[(1 - \omega^2 + |A|^2/2)^2 + 4\lambda^2\omega^2]|A|^2 = f^2/4$$

The latter equation may be solved with respect to ω^2 , and afterwards one can explore the relationship $a(\omega)$, $a=2|A|$, e.g. using MAPLE

The nonlinear resonance

- The resonance equation on the preceding slide is cubic with regard to $|A|$, and its real roots determine the amplitude of driven oscillations as a function of external force f .
- For small f the amplitude A is also small and we have a symmetric resonance curve with a maximum at $\omega=\omega_0$ (or $\omega=1$ in dimensionless units). This corresponds to only one real root of the resonance equation. In general, the function $a(\omega)$ critically depends on λ, f
- Starting from some critical $f=f_c$ there exists a frequency region, when the resonance equation has three real roots. The boundaries of this region are determined by the condition $dA/d\eta=\infty$, where $\eta=\omega-\omega_0$ is a detuning
- A frequency region exists with two (in fact, up to six) different amplitudes for one driving frequency - jumps become possible. This phenomenon is called hysteresis, some amplitudes can be reached only by increasing or decreasing the frequency

The resonance curve

- The usual question of the vibration theory is: “How does the amplitude depend on the frequency?”
- If one can solve the equation $F(a,\omega)=0$, where $a=2|A|$ is the real amplitude, then the explicit function $a=a(\omega)$ is called a *resonance curve*. Strictly speaking, the notion of resonance is valid only in the linear case, for nonlinear oscillations one can talk of nonlinear amplitude response. One must remember that in the nonlinear case the solution is represented by a wide spectrum of frequencies
- The nonlinear amplitude $a=a(\omega)$ is a parametric function of damping coefficient λ and driving force amplitude f
- **Exercise c18:** plot the resonance curve.
- **Exercise c19:** plot the amplitude response for the Duffing equation (nonlinearity coefficient $\varepsilon = -1/6$). One can use MAPLE or MATLAB graphics for these two problems.

Perturbation expansions

- Often one can encounter problems that do not have a direct solution or the solution may exist in principle, but is very difficult to obtain.
- Then one can try to find an approximate solution of the problem by an expansion near the solution of a simplified (auxiliary) problem.
- The key word here is **a small parameter**, with the problem being reduced to auxiliary one when this parameter is striving to zero.
- A large parameter can be also used, since its inverse is a small one
- A typical example is the motion of the Earth, when the influence of other planets is considered as a perturbation ($m_{\text{planet}}/M_{\text{Sun}} \sim 10^{-3}$)
- To illustrate the perturbation method, let us consider a harmonic oscillator with a weak nonlinear dissipation: $\ddot{x} + \omega^2 x + \varepsilon \dot{x}^2 = 0$
- When ε is small, the solution of the dissipative equation should not be too far from the harmonic (frictionless) motion. So, we might try to expand the solution in powers of small parameter

Perturbation expansions-II

- We can insert the expansion $x(t) = x_0(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + O(\varepsilon^3)$ into the initial problem and collect the terms of the same order of ε
 Let us find the first-order solution $\ddot{x}_1 + \omega^2 x_1 = \omega^2 x_0^2 = \omega^2 e^{2i\omega t}$
 We may write the general solution as a sum of $A \exp(i\omega t) + B \exp(-i\omega t)$ and a particular solution $x_1 = C e^{2i\omega t}$ which gives $C = -1/3$
 So the first-order correction to the unperturbed solution is

$$x_1(t) = A e^{i\omega t} + B e^{-i\omega t} - \frac{1}{3} e^{2i\omega t} \quad \text{and the first-order solution:}$$

$$x(t) = (1 + \varepsilon A) e^{i\omega t} + \varepsilon B e^{-i\omega t} - \frac{\varepsilon}{3} e^{2i\omega t}$$
- The meaning of this solution is that nonlinearity has introduced a new harmonic. This is a typical and important result; for instance the major part of nonlinear optics – a discipline that studies interaction of powerful laser radiation with matter – is based on this effect. Nonlinear optics is a foundation for modern IT devices
- Perturbation method is also the main tool of quantum theory

Scaling an oscillator

- In order to understand better how scaling works, let us consider the simple example of a damped linear oscillator driven by an external harmonic force $f(t)$:

$$\ddot{q} + 2\lambda + \omega_0^2 q = f \exp(i\omega t), \omega_0^2 = k / m, \lambda = \gamma / m$$

- As usual, we can seek the solution in the form $q = B \exp(i\omega t)$
- In principle, the complex amplitude B is related to all the parameters of the problem $\{m, \gamma, k, f, \omega\}$, which makes the representation of B rather difficult
- To improve the situation, let us introduce some characteristic quantities x_0, t_0 (by change of variables $q = x_0 \xi, t = t_0 \tau$), that can be specified later. In some particular cases they may coincide with the measurement units (e.g. cm, s)
- This means that we have a certain freedom in choosing the units, and sometimes the measurement units are chosen from purely arbitrary or historical considerations (e.g. the SI system)

Dimensionless variables

- Usually, in engineering applications of the vibration theory, the most interesting effect is the dependence of the amplitude on the frequency. To analyze it, we may e.g. choose units in such a way as to achieve $k\tau^2/m=1$ (k – elasticity)
- Likewise, we can measure the driving force in “inertial quanta” $m\xi/\tau^2$. In dimensionless variables (ξ, τ) we have

$$\frac{d^2\xi}{d\tau^2} + \chi \frac{d\xi}{d\tau} + \xi = \exp(i\varpi\tau)$$

This equation contains two dimensionless (numerical) parameters:
 $\chi = \gamma / \sqrt{km}$, $\varpi = \omega \sqrt{m/k}$ - instead of five on the preceding slide.

- The dependence of the dimensionless complex amplitude on the dimensionless frequency (in general also complex) can be graphically represented for various values of dimensionless friction (see previous slides on forced oscillations, resonance and amplitude response)

Advantages of dimensionless variables (DV)

- Knowing the relationship between numerical combinations, e.g. $\bar{B} = F(\bar{\omega}, \bar{\chi})$, one can get the relationship between dimensional quantities:
$$B = \frac{f}{k} F(\omega \sqrt{m/k}, \gamma / \sqrt{km})$$
- The example of oscillator shows how advantageous it is to work in dimensionless units – the number of variables is diminished
- DV are especially useful in numerical treatment of models - determined not by physical quantities but by numbers
- When models are formulated in analytical form, they are assumed to have the uniform dimensionality throughout equation, $s = gt^2 / 2$
- The same formula in numerical treatment $s=4.90t^2$, requiring s in meters and t in seconds, otherwise – error. Here, one implicitly uses the SI system - a drastic narrowing of the scaling map
- It is important to remember that only dimensionless combinations can stand under the function symbol (logarithm is an exception)

Mathematical background of DV and scaling

- Π -theorem – allows one to reduce the number of quantities needed to describe an object
 - Provides a general prescription for the representation of the quantity $A=F(A_1,A_2,..)$ in terms of dimensionless combinations A_i
 - It is important that such prescription be insensitive to specific form of the function F ; it should only be sufficiently smooth
 - Lagrangian formulation (mechanics, electrodynamics, relativity, etc.): equations of motion do not change if we multiply the Lagrangian by any constant.
 - Scaling properties of equations of motion: let $\mathbf{r}(t)$ satisfy the Newton's equation $m d^2 \mathbf{r} / dt^2 = -\partial U / \partial \mathbf{r}$. If we put $t_1 = \alpha t$, $m_1 = \alpha^2 m$, we get for the scaled motion the same equation: $m_1 \frac{d^2 \mathbf{r}(t_1)}{dt_1^2} = -\frac{\partial U}{\partial \mathbf{r}}$
- This trick illustrates the transition to scaled units

See: [G.I.Barenblatt. *Scaling, self-similarity and intermediate asymptotics*, Cambridge, 2002](#)

Effects of scaling

- Scaling properties of the Newton's law (see preceding slide) show that if we increase the mass of the body (e.g. of a car) four times, it will pass the same trajectory twice as slow
- Here we implicitly assumed that the potential U does not depend on the inertial mass m . This fact contradicts our intuition, e.g. if we recall falling bodies. But in the gravitational field U is proportional to gravity mass (in Newtonian approximation), so the period does not depend on m
- For a pendulum $\{m, l, g, \theta\}$, the period is uniquely determined by scaling: the full system of dimensionless parameters consists of $S = lT^{-2}g^{-1}$ and θ , so there must be a relation of the type $S^{-1/2} = f(\theta)$ or $T = \sqrt{\frac{l}{g}} f(\theta)$ - which is really the case. So we managed to establish the relation for the period without solving equations

Effects of scaling-II

- The scaling function $f(\theta)$ can be determined either numerically, or by ODE integration, or from experiment (measuring the period or frequency, all other parameters being fixed)
- Modeling is largely based on scaling: e.g. aircraft and car design
- What is the pendulum frequency on the Moon?

Thus we can easily compute it, provided we know $T_E / T_M = \sqrt{\frac{l_E g_M}{g_E l_M}}$ the value on the Earth. The same applies to other quantities.

- Scaling of all the coordinates by the same factor means the transition to new (homothetic) trajectories, different from the initial ones only in their linear dimensions
- Often the potential energy is homogeneous function of the coordinates, i.e. $U(\alpha \mathbf{r}_i) = \alpha^k U(\mathbf{r}_i)$, α - any constant, k is the degree of homogeneity. In this case, motion times between respective trajectory points are scaled as

$$\frac{t'}{t} = \left(\frac{l'}{l} \right)^{1-k/2}$$

Scaling in biological modeling

*Un fourmi de dix-huit mètres
Avec un chapeau sur la tête,
Ça n'existe pas;
Pourquoi pas?* – a French song

An observation: small animals maintain much higher activity and metabolize food faster than large animals. Why?

- “The Jurassic Park” movie – the efficiency manifestation of mathematical zoology. Why do dinosaurs have small heads? A model: small sphere (head) connected by a cylinder to a big sphere (body)
- What happens if a dino grows and all dimensions are scaled? The head mass $\sim L^3$, the neck strength $\sim L^2$ (cross-section)
- Development of living species is subordinated to physical laws
- Can the mammals be as small as insects? No, since there would be not enough mass to support the constant body temperature $M \sim L^3$, $q \sim L^2$, q is the thermal flux into environment, L is characteristic dimensions of the body. Small creatures like gnomes could hardly exist on a human time-scale. Babies should be coated with blanket

Scaling in biological modeling-II

- Why do mammals have to eat periodically and not all food at once and in big quantities – like almost all reptiles? Because the consumption of the amount of food comparable with the body mass would disturb the thermal balance, which lies in a very narrow temperature range. Excessive food can freeze us to death
- How big can terrestrial creatures be? Two limitations: 1) the mass contained within the body should be efficiently cooled; 2) gravity force should not hinder the motion of this mass
- How does the animal speed v depend on its size L ? Animal speed is practically independent of L for a flat terrain: mechanical power is proportional to heat production $q \sim L^2$, air resistance $F \sim v^2 S \sim v^2 L^2$ Power needed to overwhelm this force is $P \sim Fv \sim v^3 L^2$ Thus, speed practically does not depend on L : speed of the horse, rhinoceros, cat, hare, human, even elephant is of the same order of magnitude.
- Computer biomechanics: better to model dinosaurs than nuclear warheads; modern spin-off tends to entertainment industry

Scaling in biological modeling-III

- The speed of a living creature while crossing hills is inversely proportional to its size, since in such conditions it is not the air resistance that the animal should overcome but gravity force, so $P \sim Mv \sim L^3v \sim L^2$, which means that $v \sim L^{-1}$

A cat can easily run uphill, while humans or horses have to slow down
In order not to be compacted by a promenading dinosaur, it is better to escape uphill: big size makes it difficult for the creature to catch you.

- Here is a situation when dimensions help to survive: when the animal must move from one water source to another, the maximum run time is $T \sim L$.
- Proof:** T is proportional to water stored in the body and inversely proportional evaporation, so $T \sim L^3L^{-2} \sim L$. Recalling that $v \sim L^0$, we conclude that the safe distance in the desert is $\sim L$.

Mathematical models taking into account heat exchange between mammals and environment exist for various body architectures and hair (body surface features)

Jumping dinosaurs

- In Hollywood movies, tyrannosaurs jump and fall going after the pray. This is physically unrealistic, even if dinosaurs would live somewhere among us.
- 1) It is highly improbable that a dinosaur's speed $v > 25$ km/h, and such speed can be easily achieved by a human at a short distance;
2) Had a dino fallen down, it would hardly survive after the very first fall – the energy output would result in serious damage for the skull and inner organs
- How high can a dinosaur jump? This height h practically does not depend on dimensions L . Energy E needed to jump to h is $E = mgh$ and, hence, $E \sim L^3$. Muscles produce work $A \sim FL$, where F is the muscle force proportional to limb cross-section, i.e. $F \sim L^2$. Thus $h \sim L^0$, i.e. the jump height does not depend on species size: cats, dogs, kangaroos and humans jump to about the same height.
- **One may ask: what does it have to do with computer-based mathematical modeling? The answer: scaling helps to understand solution in advance**

Nuclear explosion

- This is an example of *self-similarity*: a solution $f(x, t_1)$ is similar to $f(x, t_0)$ at an earlier moment $t_0 < t_1$.
- A very strong spherical shock wave due to instant release of energy by a point source (L.Sedov, 1946; J.von Neumann, 1947)
- The whole picture is determined by just two parameters: released energy E and initial gas (air) density ρ . From these variables plus r, t one can construct only one dimensionless combination $r(\rho / Et^2)^{1/5}$
- The whole motion is self-similar: the shock wave position R is $R = \alpha(Et^2 / \rho)^{1/5}$ where α is a constant. The speed of the shock wave $u = dR / dt = 2R / 5t = \frac{2\alpha E^{1/5}}{5\rho^{1/5} t^{3/5}}$ - so the law of shock wave motion is determined by scaling.

Pressure, density and velocity of the gas can be expressed via u using known fluid dynamics formulae. All these quantities are determined by the only dimensionless variable $\xi = r / R(t) = \frac{r}{\alpha} \left(\frac{\rho}{Et^2} \right)^{1/5}$

Models in chemical engineering

- A chemical reaction in the homogeneous media: in a certain volume there is a substance A , which can be converted into substance X . This reaction has a constant rate k_0 and the decrease of the initial quantity of A is small compared to it (excessive A)
 - Then X is transformed into Y , with the rate rising with the concentration of Y
 - Finally, molecules Y are irreversibly dissociated forming the substance B
- $$A \xrightarrow{k_0} X \xrightleftharpoons[k_2]{k_1} Y \xrightarrow{k_2} B$$
- One can state that for a given temperature the rate of a chemical reaction is proportional to the product of reagent concentrations. This fact (which can be proved in statistical physics) serves as a starting point for building a chemical engineering model
 - This is a competition model, similar to that of Lottka-Volterra (Volterra used the Lottka equations for an ecological model)

Models in chemical engineering-II

- The mathematical model of the process:

$$\dot{X} = k_0 - k_1XY, \dot{Y} = k_1XY - k_2Y, \dot{B} = k_2Y$$

- One is usually interested in such a regime, when the production rate (output) of the substance B is constant in time. The concentrations of X and Y are also kept constant: $\dot{X} = \dot{Y} = 0$
- Such regime corresponds to chemical equilibrium:

$$k_0 - k_1X_0Y_0 = 0, k_1X_0Y_0 - k_2Y_0 = 0$$

where $X_0 = k_2/k_1$ and $Y_0 = k_0/k_2$ are equilibrium concentrations.

Small deviations from equilibrium concentrations: $x(t), y(t)$

From the main equation, we have

$$\dot{x} = -k_2y - (k_0k_1 / k_2)x, \dot{y} = (k_0k_1 / k_2)x$$

This is the damped oscillator equation with $k_0k_1 / k_2 = 2\gamma, k_0k_1 = \omega_0^2$

One may see once more that the oscillator discussed previously is not just a toy model, but has numerous practical applications

Thermal explosion and chemical physics

- Suppose an exothermal chemical reaction occurs in some volume, with the heat escaping to the environment. The reaction speed depends on the local temperature $T(\mathbf{r}, t)$; for simplicity we may consider the average temperature at a given moment t :

$$T(t) = \frac{1}{V} \int_V d^3r T(\mathbf{r}, t)$$

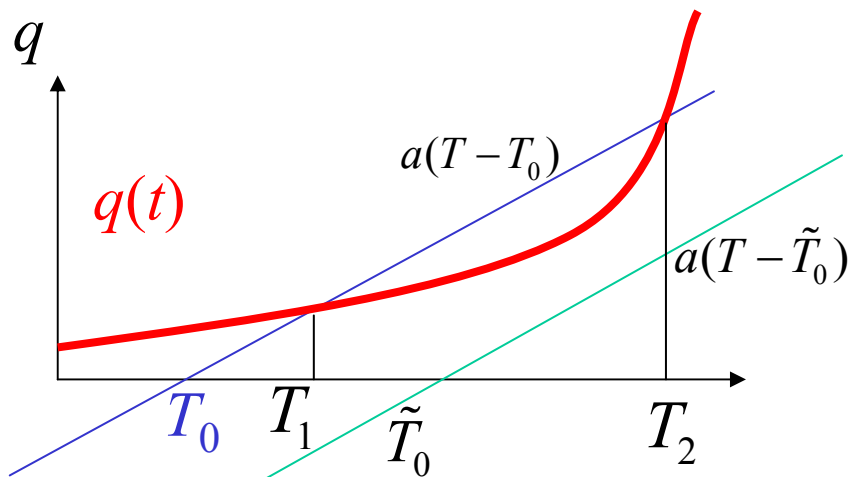
- So, the rate of heat production q depends on $T(t)$: $q=q(T)$.
- The process is described by the model

$$\dot{Q} = \frac{d}{dt}(cT) = q(T) - a(T - T_0)$$

where $Q(t)$ is the heat quantity in the reaction volume, c is the specific heat (we may consider $c=const$), a is the heat loss coefficient, T_0 is the ambient temperature. The stationary state, when $T=const$ in the process of reaction, corresponds to equilibrium points T_i

Thermal explosion and chemical physics-II

- The usual thermal effect of a chemical reaction:



Stationary state is achieved when the graph of reaction yield $q(t)$ crosses the line $a(T - T_0)$

Instability due to positive feedback

If the ambient temperature T_0 is sufficiently high ($T = \tilde{T}_0$), the stationary state is not achieved: heat production always surpasses heat dissipation and the volume is heated up. For $T = T_0$ two equilibrium states, T_1 and T_2 , are possible. Near T_1 the rhs of the reaction equation falls (transits from positive to negative values

- Thus, the state characterized by T_1 is stable: if T becomes lower than T_1 the thermal balance is positive and the volume heats up. Unrestrained heating due to instability leads to explosion; therefore chemical physics is usually an euphemism for study of explosives

Thermal explosion and chemical physics-III

- If the temperature T grows higher than T_1 , then the heat is lost faster than produced, and the volume is cooled down to T_1
- Analogously, one can show that the equilibrium point T_2 is unstable
- Thus, the whole process of thermal explosion essentially depends on the ambient temperature T_0 : if it is lower than T_2 then the volume temperature T strives with time to the equilibrium value T_1 ; if the initial temperature was greater than T_2 then the temperature would rise catastrophically – the thermal explosion
- Such considerations led to the creation in 1927 of the explosion theory by the Nobel Prize winner N.N. Semenov. Later this theory was generalized to incorporate chain processes. The discipline that explores explosion and combustion kinetics, developing respective modeling techniques is called chemical physics.
- Results obtained by chemical physics are of crucial importance for engineering and military applications.

Nuclear reactor modeling

- The working reactor is sustained in the critical state, when the number of fission neutrons corresponds to stationary power. In the subcritical mode, the number of neutrons is lower than equilibrium and the reaction is quickly damped. In the supercritical mode, the neutron flux is too high and the fission can lead to intense heating and explosion, especially in failures accompanied by coolant loss
- In former-generation reactors, critical state is physically unstable and is sustained with the aid of a complicated control system that is based on movable control rods regulating the neutron flux
- The complete mathematical model of a reactor is based on time-dependent 3D neutron kinetics in inhomogeneous media. Such model should be supplemented by heat transfer and fuel burn-out processes. For simplification, the latter are often treated separately
- Nuclear reactor modeling has developed into a vast discipline

Nuclear reactor modeling-II

- Elementary mathematical model of a nuclear reactor is based on two-group neutron diffusion. Two-group approximation means that from the whole spectrum of neutrons only fast and slow neutron peaks (e.g. 1MeV - fission and 0.025eV – thermal) are selected. The reactor is supposed to have a simple geometry, e.g. 3D cube Ω
- The distribution of neutrons in the reactor (neutron field) is described by a system of equations

$$\frac{1}{v_1} \frac{\partial \Phi_1}{\partial t} = \operatorname{div}(D_1 \nabla \Phi_1) - A_{11} \Phi_1 + A_{12} \Phi_2; \quad \frac{1}{v_2} \frac{\partial \Phi_2}{\partial t} = \operatorname{div}(D_2 \nabla \Phi_2) - A_{22} \Phi_2 + A_{21} \Phi_1$$

with boundary conditions $\Phi_1 = \Phi_2 = 0$ at $\partial\Omega$ Here the functions $\Phi_{1,2}(\mathbf{r}, t)$ (flux) correspond to fast and slow neutrons, respectively.

- More sophisticated mathematical models of a nuclear reactor are based on the Boltzmann kinetic equation. For nuclear systems, such kinetic (transport) models may include up to 50 species (fuel, intermediate and final fission products,)

Nuclear reactor modeling-III

The system of equations describing evolution of the neutron field in the reactor takes account of the following physical processes:

1. Diffusion (terms $\operatorname{div}(D_i \nabla \Phi_i)$)
2. Creation (birth) of fast neutrons due to absorption of slow ones ($A_{12}\Phi_2$) and vice versa ($A_{21}\Phi_1$)
3. Neutron absorption (death), diagonal terms (proportional to A_{ii})

Parameters A_{ij} and D_i are some given functions of coordinate \mathbf{r} determined by material properties. In this approximation the problem is linear, so it can be written as $\partial \varphi / \partial t = A \varphi$ where $\varphi = \{\Phi_i\}$, A is an elliptical differential operator (in a more complex case integro-differential), whose properties are close to those of Laplacian

4. In some models, advection terms $(\mathbf{v} \nabla) \mathbf{v}$ leading to quasilinear problems are also important. Solutions crucially depend on geometry

Elementary theory of nuclear reactors

- The linear theory of nuclear reactors is based on the treatment of the diffusion equation $\varphi_t = A\varphi$ (see the preceding slide)
- In the normal working mode of the reactor, operator A does not vary with time, so the particular solutions have the time-invariant form: $\varphi(\mathbf{r}, t) = e^{\lambda t} u(\mathbf{r})$. Thus $Au = \lambda u$ i.e. u should be the Eigenfunction of A . The spectrum of A is discrete, $\lambda = \lambda_k$, and the respective functions u_k make up the full system forming the basis for the solution: $\varphi(\mathbf{r}, t) = \sum_k a_k e^{\lambda_k t} u_k(\mathbf{r})$, where the coefficients a_k may be found from the expansion of the initial function:

$$\varphi(t=0, \mathbf{r}) = \varphi_0(\mathbf{r}) = \sum_k a_k u_k(\mathbf{r})$$

If we enumerate the point-like spectrum in such a way as

$\operatorname{Re} \lambda_k \rightarrow -\infty, k \rightarrow \infty$ than for large t : $\varphi(t) \approx a_1 e^{\lambda_1 t} u_1$

which gives the asymptotic solution to the problem

Neutron multiplication and criticality models

- The reactor behavior depends on the term with $\lambda_1 = \max_k \operatorname{Re} \lambda_k$
If $\lambda_1 < 0$, the reactor is *subcritical*, there is a decrease in neutron populations and generated power, reactor eventually shuts down
If $\lambda_1 > 0$, the reactor is *supercritical*, the power grows exponentially and reactor eventually explodes

The working regime of the reactor is $\lambda_1 \approx 0$, (critical mode). For a reactor to be critical, neutron multiplicative factor (determined by off-diagonal terms in the operator A) should be unity: a constant number of neutrons is produced in each generation and a constant fission energy is released.

- The principal Eigenvalue λ_1 depends on reactor geometry and physical composition. This value is usually regulated with the help of a complex system of rods containing neutron-absorbing material. The purpose of such regulation is to ensure $\lambda_1 \approx 0$.
- This problem becomes even more mathematically complex, when delay neutrons should be considered (the delay time \sim one minute)

Numerical methods in reactor modeling

(how to compute the point Au by the point u)

- One of the main tasks is to compute the boundary point of the spectrum of A , $\lambda = \lambda_1$. In numerical practice, A is not a differential operator but a finite-difference one. Usually the dimensionality of the finite-dimension space is very high ($10^3 \sim 10^6$), so it is hard to represent the matrix A in explicit form.
- Let us consider, for simplicity, 2D case ($\mathbf{r} = x, y$). If, in order to discretize the problem, one builds up a mesh with nodes

$$x_m = mh, y_n = nh, (m, n = 0, 1..N)$$

then one has for functions $\Phi_{1,2}^{m,n}$ defined on the lattice

$$\begin{aligned} \frac{\lambda}{v_i} \Phi_i^{m,n} = & \frac{1}{h} \left(D_i^{m+\frac{1}{2},n} \frac{\Phi_i^{m+1,n} - \Phi_i^{m,n}}{h} - D_i^{m-\frac{1}{2},n} \frac{\Phi_i^{m,n} - \Phi_i^{m-1,n}}{h} \right) + \\ & \frac{1}{h} \left(D_i^{m,n+\frac{1}{2}} \frac{\Phi_i^{m,n+1} - \Phi_i^{m,n}}{h} - D_i^{m,n-\frac{1}{2}} \frac{\Phi_i^{m,n} - \Phi_i^{m,n-1}}{h} \right) + A_{i,i}^{m,n} \Phi_i^{m,n} + A_{i,j}^{m,n} \Phi_j^{m,n} \end{aligned}$$

Heavy nuclear accidents

- After the Chernobyl accident, much more attention has been paid to nuclear safety issues. Problems of nuclear safety are very diverse:
 - prevention of heavy reactor accidents (e.g. with core meltdown)
 - description and analysis of various accident scenarios
 - nuclear waste management
 - armament plutonium efficient use

One of the main today's safety tools is modeling of complex physical phenomena that occur as a nuclear (NPP) accident progresses with time. Modern software packages model the whole power plant starting from physical interactions between fuel rods and containment vessel to cooling system plumbing.

The principal physical process in any reactor is neutron fission leading to energy production in the core, using heat to generate electricity

Heavy nuclear accident scenarios

How a catastrophe might occur?

First:

- A pipe might break
- The safety system might fail to respond

LOCA=Loss Of Coolant Accident
is the primary reason for hazard

Second:

- Core-cooling system might fail (e.g. loss of coolant)

**Core-Melt Accident
is dangerous, since
it may result in the**

Third:

- The fuel might melt
- The fuel might react with water
- The fuel mixture might melt through the containment

containment melt-down

Fourth:

- The containment might fail (from the melt or overpressurization)
- Activity might spread to the public

Earthquakes, acts of violence, terrorists, missile attack, ...

Heavy nuclear accident modeling

- Various scenarios of core melt accidents at nuclear power plants: to analyze radiation hazard and predict the consequences
- A crucial problem – core melt retention in the reactor vessel
- Mathematical model of corium – a fluid with internal heat sources interacting with the reactor vessel: **a new class of hydrodynamic problems**
- The study of convective heat transfer in a heat-generating fluid within a closed volume
- Inverse temperature distribution –condition for Rayleigh-Bénard convection. Thermal flux in corium depends on modified Rayleigh number

$$Ra\tilde{\alpha} = \frac{\alpha g Q H^5}{\lambda \nu \chi}$$
 where $\alpha, \nu, \chi, \lambda$ are thermal expansion, kinematic viscosity, thermal diffusivity and conductivity, g – gravity, H – vertical dimension (height), Q – power release density (the theory: P.Kondratenko, L.Bolshov)

- How to retain the melted core in the reactor vessel?
- Theory+experiment+numerical simulation
- 3-d heat & mass transfer models and respective computer codes
- Heat flux distribution, residual heat removal deficiency, stratification of a heat-generating fluid, temperature field picture, boundary layer properties, laminar and turbulent regimes
- What is an effective trap for the melted core? Corium-concrete interaction should be described as a physical problem
- Code validation using extensive experimental data
- Project «RASPLAV» (Nuclear Safety Institute, Russia)
- The popular mathematical approach consists of a the heat transport analysis, taking into account two-phase flow phenomena (e.g. convection). Relevant engineering modules such as channels, pipes, pumps,, vessels may be considered as “lump” parameters

Main reactor safety issues

- One cannot **prove** that a failure is impossible
 - Ideally: N entirely independent safety systems working on totally different principles
 - Complex sequences: loss of coolant → injections of emergency coolant → will the fuel rods balloon and block coolant channels?
Will chemical reactions take place? Will fuel cladding crumble allowing radioactive fuel pellets to fall out?
 - Probabilities are small ($\sim 10^{-8}$ per reactor per year), comparable with the probability of perpetuum mobile to be built
 - Very hard to compute, a lot of uncertainties. Strong trans-scientific element - based on judgments
 - The response of nuclear reactors and power plants to earthquakes
 - Waste disposal and management
 - Transport of radioactive materials
 - Social factors
 - Nuclear people vs. society?
- Modeling socio-psychological response appears to be not less important for nuclear safety issues than purely engineering problems**

Computer code development in nuclear safety analysis

- the major part of codes are designed to model core destruction during the initial stage of a heavy accident
- the predominant processes: physical & chemical interaction of construction materials accompanying core heating due to excessive heat release
- also include: coolant fluid loss
- melted components drain-down
- fission products & hydrogen escape
- mechanical deformation of reactor elements
- incorporated in standard codes: ICARE - France, CDAP/RELAP - US, ATHLET - Germany, SVECHA (CANDLE) - Russia
- 2D-3D mathematical description underlying codes is necessary

Some codes have been experimentally validated: Sandia Labs(US), Tver NPP (Russia), etc.

An example: «RASPLAV» software package §

- From mathematical modeling to a computer code: core melt accident – a new class of complex phenomena
- Heat & mass transfer for a fluid with internal energy sources; inhomogeneous heat release by a fluid
- Thermo-hydraulic behavior of the melted core
- Interaction of melt with the reactor vessel at the final stage of heavy accident (melt-down scenario)
- Residual heat removal
- Mechanical behavior of the construction material under temperature load (over 2600 °C)
- 3D modeling of large-scale experimental setup
- 3D CFD modeling purposed to quantitatively predict flow and heat transfer on the physical basis of mass, momentum and energy transport, leading to exchange between the phases

3-d codes (The «RASPLAV» project) §

- Examples: CONV2D and CONV3D codes: analyzing convection, diffusion, binary mixtures, taking melting processes into account (Nuclear Safety Institute, Moscow)
- Convection in a cavity having walls with different temperature ($Ra \sim 10^4 - 10^{12}$)
- Rayleigh-Bénard convection in a cavity with hot lower and cold higher wall at $Ra \sim 10^4 - 10^{10}$
- Convection of a heat-generating fluid at $Ra \sim 10^6 - 10^{14}$
- Convection/diffusion accompanied by nonstationary melting
- Convection at maximum possible Rayleigh numbers
- Validation of codes both for laminar and turbulent modes
- The open question: are 3D CFD models scalable? Can computer codes be validated with high-resolution measurements in small scale test facilities? Scaling is expected but not proved

Reactor vessel destruction

- sustainability of a reactor vessel under high pressure and temperature (e.g. for PWR = pressurized water reactor)
- heavy accident scenario with core melting and melt (corium) flow into the vessel: analysis of vessel carrying capacity during a heavy melt-down accident accompanied by the vapor explosion
- destruction criteria from dynamic stress/strain analysis
- thermal stress due to nonstationary temperature field
- innovative reactor safety concepts based on passive devices for heat removal, no need for pumps. Modeling is necessary
- material properties (elasticity, plasticity, melting point) as a function of temperature and pressure
- construction mechanics and equipment sustainability; predominantly numerical analysis (e.g. of safety devices, protective barriers and NPP constructions under the influence of explosions, shock, and thermal load)

Reactor vessel destruction: modeling results

- **Methods, mathematical models and software modules to analyze the mechanical state of vital reactor systems, in the normal and accident mode, up to the destruction stage**
- Temperature stress distribution is very inhomogeneous, with concentration zones possibly located far from melted metal – very dangerous!
- Maximal shear takes place in the intermediate area over the melt - it is in this area cracks and other mechanical defects leading to the vessel total destruction might arise
- As the vessel walls are partly melted, one may expect their destruction near the vessel bottom
- Modeling allows one to forecast the moment of destruction (taking into account uncertainties in mechanical properties)
- Three Mile Island accident was modeled using this approach

Linear and nonlinear models in economics and military planning

- The starting nonlinear model of the chemical engineering process is richer and more versatile than the linear oscillations model.
- The latter was obtained due to an assumption of small deviations from equilibrium mode (small perturbations of the concentration)
- Similar models appear in economics and in military planning, specifically in the case when non-regular armies (e.g. partisans) are participating in military activities. The solutions in such models show the temporal evolution of the military conflict and can predict which party can be defeated
- For economic planning, analogous models allow one to divide the produced output into consumed and accumulated parts – one of the crucial problems of economic growth
- These models are unified by a common feature: essentially non-negative character of variables – a supplementary constraint

The arms race

- Let us imagine two hypothetical adversaries accumulating e.g. weapons of mass destruction, W_1 and W_2 being the total quantities of such weapons owned by each country. By W_1 and W_2 one can also imply the quantities of fission material (Plutonium), number of warheads, or deadly chemical/biological agents, $W_{1,2} \geq 0$
- The simplest model is based on assumption that the temporal variation of W_1 and W_2 is proportional to the following factors:
 - intelligence data on the accumulated weapons by the adversary
 - ageing of already produced and deployed weapons with speed b_i
 - the degree of mutual distrust and political tension, characterized by a two-component function $c_i(t)$, can be obtained from another model (describing political interaction and competition, not necessarily a dynamical one). Here in the notations $i=1,2$.
- **This model is rather limited, e.g. it does not take into account the time delay and random factors, nevertheless it can lead to realistic inferences.**

Models of this type are used in modern political analysis

The arms race model

$$dW_1/dt = a_1(t)W_2 - b_1(t)W_1 + c_1(t)$$

$$a_{1,2}(t) > 0, b_{1,2}(t) > 0$$

$$dW_2/dt = a_2(t)W_1 - b_2(t)W_2 + c_2(t)$$

$$c_{1,2}(t) \text{ may be negative}$$

The simplest variant of the model: $a_{1,2}, b_{1,2}, c_{1,2}$ do not vary with time. We can start with exploring the equilibrium (fixed) points: it allows one to find the qualitative behavior of $W_{1,2}$

$$W_1^0 = \frac{a_1 c_2 + b_2 c_1}{b_1 b_2 - a_1 a_2} \quad W_2^0 = \frac{a_2 c_1 + b_1 c_2}{b_1 b_2 - a_1 a_2}$$

a_i are sometimes called fear parameters

The first important implication: for non-negative $W_{1,2}$, the equilibrium is possible only under the condition $b_1 b_2 > a_1 a_2$. The meaning of this condition is that equilibrium is violated when any of the two countries continues stockpiling armaments, irrespective of the counterpart's behavior and depreciation (amortization) rate. One may observe that the non-negativity condition is important here

The arms race model-II

- One may notice that if grievance parameters $c_{1,2}$ are equal to zero, then the equilibrium point corresponds to the absence of armament in both parties. The reciprocity rule, usual in inter-national relations, tends to equalize the political tension, $c_1=c_2=c$
- With growing t functions $W_{1,2}(t)$ are striving to equilibrium values $W_{1,2}^0$. The equilibrium is stable (asymptotically): any deviation from it becomes negligible after some time. The equilibrium point is a stable node (see below).
- From this model one can analyze possible strategies and behavior of both parties. For instance, let the rate of armament production change by a small quantity da (for simplicity, $da_{1,2}=da$). It is clear that the amount of armaments is also changed, the respective increase (or decrease) being $dW_{1,2}^0$. The both parties desire that the increases be equal, so that the balance should be preserved.
- Political factors may result in nonlinear response to a production change; such nonlinear competition models admit many solutions

The arms race model-III

- Increased stockpiled armaments:

$$dW_1^0 = \frac{a_1 a_2 c_2 + a_2 b_2 c_1 + a_1^2 c_2 + a_1 b_2 c_1}{(b_1 b_2 - a_1 a_2)^2} da$$

$$dW_2^0 = \frac{a_1 a_2 c_1 + a_1 b_1 c_2 + a_2^2 c_1 + a_2 b_1 c_2}{(b_1 b_2 - a_1 a_2)^2} da$$

- If, for simplicity, we assume that both parties have equal mistrust to each other, $c_1=c_2=c$, then from the equality $dW_1=dW_2$ we get the stable parity condition:

$$a_1(a_1 + b_2 - b_1) = a_2(a_2 + b_1 - b_2)$$

- The latter condition may be considered as a foundation for a treaty between countries, if the respective parameters $a_{1,2}$ (production rate) and $b_{1,2}$ (depreciation) are controllable. For a nonlinear model these parameters, especially $a_{1,2}$, may depend on $W_{1,2}$

The arms race negotiation strategy

- The negotiation strategy may be based on quantitatively assessing the balance between deployment and depreciation rates
- Let, for instance, $a_2 = \mu a_1, \mu > 0$. Then we have

$$a_1(1 - \mu) = \Delta b \equiv b_1 - b_2$$

This simple formula dictates the strategy: if $\mu < 1$, that is the production and deployment rate of the Country 2 is lower than for Country 1, then the depreciation and put-out rate by the Country 2 should be slower than in Country 1

- The arms race model can be made more sophisticated by subsequent inclusion of other negotiating parties and other factors (e.g. a variety of armaments), but the idea of the model remains the same as described.
- This model was developed by English physicist L. Richardson and was initially based on the assumption that large arsenals make the conflict more probable. Can a small incident start a large conflict?

Phase plane analysis – stability behavior of two nuclear countries

Here we treat only the linear case; the nonlinear case, involving bifurcations, requires more sophisticated analytical methods

- The autonomous system of two equations

$$dW_1/dt = a_1W_2 - b_1W_1 + c_1 \quad a_{1,2} > 0, b_{1,2} > 0$$

$$dW_2/dt = a_2W_1 - b_2W_2 + c_2$$

in the linear approximation represents an affine map of a plane onto a plane, with a nonsingular transform matrix. An affine transformation can always be factored into a product of translation, rotation, scaling (stretching and contracting), shear (elongation and compression), and reflection. This model may be represented graphically (see next slide)

- A more general linear case is represented by a system of equations $dx/dt = A(t)x + f(t)$, where $A(t)$ is the matrix function and $f(t)$ is a vector

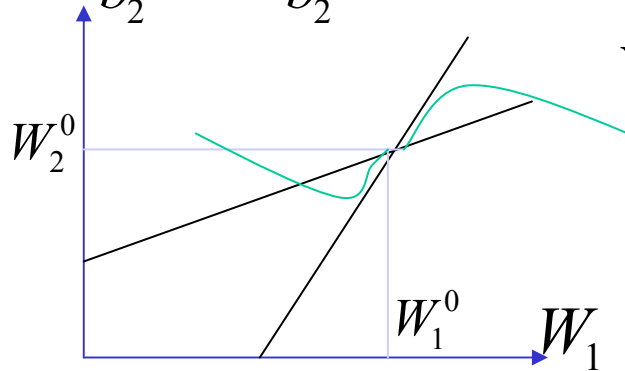
Stability analysis of military equilibrium

- Is the equilibrium of two rival countries stable or unstable (under the condition $b_1 b_2 > a_1 a_2$, see preceding slides)?
- This condition ensures that the tangent of zero-isocline:

is smaller than that of the ∞ -isocline

$$W_2 = \frac{a_2}{b_2} W_1 + \frac{c_2}{b_2}$$

Functions $W_{1,2}$ tend to equilibrium values with $t \rightarrow \infty$



The equilibrium is stable (node):

any disturbance becomes very small after some time. This may not be true

when the condition $b_1 b_2 > a_1 a_2$ is reversed

$$W_2 = \frac{b_1}{a_1} W_1 + \frac{c_1}{a_1}$$

There may be four cases: 1. Trajectories approach a stable point; 2. Trajectories go to infinity (a runaway arms race); 3. Trajectories go to disarmament ($W_{1,2}=0$); 4. Evolution of the system depends on the initial state

A model of labor market

- How does the number of employees change? When does the unemployment strike? These questions are of rising importance
- A two-component labor market where employers and employees interact. Variables characterizing the labor market: salary $p(t)$ and the number of employees $N(t)$
- Equilibrium: $N_0 > 0$ people are ready to work for payment $p_0 > 0$. If this equilibrium is perturbed, e.g. due to financial problems, intervention of the state, or demographic factors than functions $p(t)$ and $N(t)$ are deflected from equilibrium values
- *Assumption 1* of the model: employers change salary of employees proportionally to such deflection:
$$\frac{dp}{dt} = -a(N - N_0), a > 0$$
- *Assumption 2*: the number of employees grows or falls in proportion to pay increase or decrease about the equilibrium value p_0

A model of labor market - II

- $\frac{dN}{dt} = b(p - p_0), b > 0$ For large deviations from equilibrium pay, the model becomes nonlinear
- The linear model produces salary and employment oscillations:
$$\frac{d^2(p - p_0)}{dt^2} + ab(p - p_0) = 0, \frac{d^2(N - N_0)}{dt^2} + ab(N - N_0) = 0$$
- This system of equations (note: oscillations again!) may be written in the form which is convenient to readily obtain the first integral:
$$\frac{(N - N_0)}{b} \frac{d(N - N_0)}{dt} + \frac{(p - p_0)}{a} \frac{d(p - p_0)}{dt} = 0$$
- The first integral of the model: $a(N - N_0)^2 + b(p - p_0)^2 = \text{const}$

This expression takes the form of a **conservation law**.

It is seen that in the moments when $p=p_0$ (equilibrium payment), there are more employees than needed for equilibrium, $N>N_0$

Unemployment cycles

Contrariwise, when $N=N_0$ (equilibrium employment), the payment is over the equilibrium, $p>p_0$. The payment budget $B=pN$ at such period exceeds its equilibrium value $B_0=p_0N_0$.

- However, during the other phase of employment oscillations, this budget is lower than B_0 , so in the linear model the average salary budget (over the cycle) is:

$$\frac{1}{T} \int_{-T/2}^{T/2} B(t) dt = B_0 = p_0 N_0$$

- This simple model

nevertheless allows one to trace

the labor market development for perturbed equilibrium, e.g. when a part of employees are retired (there may be a fluctuation of the number of pensioners). The labor market response is a sensitive issue.

- A great number of models, e.g. arms race, interactive populations, employment, etc. are mathematically similar: *universality*

Synergetics

- So far, our attention was focused on the evolution of spatially uniform systems, so that we could model the behavior by ODEs.
- More complicated models correspond to the emergence of structures from a disordered initial state. Such “order out of chaos” is called self-organization.
- The discipline that studies complex structures (mostly arising in nonlinear media) is called **synergetics**. This terminology was introduced by H.Haken, a professor of physics in Stuttgart
- The problem of a set of forms sustained by a given medium is very ancient and complicated (old philosophy, elementary particle physics, thermal structures, diffusion, flocks of birds, etc.)
- This problem is impossible to treat without computer modeling (PDE-based modeling). Mathematical models in social and life sciences are closely connected with synergetics (for instance, evolution of population density in spatially non-uniform models)

Modeling of traffic flows

- The flow of traffic is described by dynamical models evolving over the spatial substrate having a complicated topology.
- In other words, traffic systems are complex dynamical systems, whose evolution unfolds on networks.
- The rough macroscopic picture of the traffic flow is similar to quasi-stationary electric current flow in electrical networks.
- A bird's view: if one takes a look at the traffic system (e.g. within a certain area) from the airplane, one can similarize the flow of traffic to quasi-stationary particle flux in a network of 1D wires, channels or vessels. A particular topology = network realization
- A physicist's view: traffic flow is a many-body system of strongly interacting particles – vehicles. A typical phenomenon is the phase transition from the free traffic state at low densities to congested traffic at high densities (the jamming transition).

Basic types of traffic models

1. **Microscopic**: based on dynamical equations for each vehicle (e.g. car following – modeling the response to preceding car actions)
2. **Macroscopic**: based on fluid dynamics equations, the coarsest level of description, phenomenological steady-state approach
3. **Kinetic**: intermediate between 1 and 2, based on Boltzmann equation. Kinetic models justify macroscopic those and produce more accurate results; on the other hand, the former are less computationally intensive than microscopic models
4. **Numerical**: cellular automata (e.g. with $n=3$ states on a square lattice); psycho-physiological (discrete regions of behavior)

For the realistic vehicular traffic modeling, three different regimes should be considered: city, highway and intercity road.

The main relationship to be modeled is **current vs. density** – the fundamental diagram. The problem is essentially nonlinear.

The science of transportation

- Transportation science follows the steps of physics: from phenomenological to progressively more microscopic modeling
- Microscopic modeling resembles molecular dynamics (see below), which is a quantitative, mostly computer-based method of describing macroscopic phenomena starting from individual particle motion (an example: fluid dynamics can be derived from particle collisions)
- Traffic flow microscopic modeling (microsimulation): molecular dynamics of intelligent particles flowing through a graph.
- A large part played by social, economic, cultural and other “soft” factors influencing the decisions of transportation agents (intelligent particles). Route planning – creation of underlying network – is also affected by “soft” factors
- New methods of steering traffic, e.g. telematics, play the role of external field in physics-based traffic flow modeling

Primitive dynamical models of traffic

- Before attempting to discuss traffic flow models beyond steady-state approach, one needs first to have regular methods for 1D smooth collective car motion (along the direction x)
- Such methods are intuitive and may seem naïve, but their purpose is to provide an insight for more sophisticated models
- Let us consider a flow of traffic on a highway (die Autobahn). We may characterize the traffic by the average density of vehicles ρ (e.g. cars per km) and by the average velocity u . The density is restrained by some maximum ρ_m when cars are moving bumper-to-bumper $0 \leq \rho \leq \rho_m$
- Since the number of cars is conserved, we have the continuity equation (conservation law):
$$\partial \rho / \partial t + \partial(\rho u) / \partial x = 0$$
- To obtain the equation for ρ , we must make an assumption about the relationship $u(\rho)$.

One-dimensional models of traffic

- On a highway, when the density is low, people tend to drive as close to the speed limit $u=u_m$ as possible, or if this limit does not exist, as in Germany, the maximum speed u_m is determined by the car and driver capacity.
- In heavy traffic one must slow down, with velocity decreasing as density rises. The simplest model is the linear one:

$$u(\rho) = u_m(1 - \rho / \rho_m)$$

- While on real highways the traffic flow may not exactly satisfy this equation, the latter was shown to be a good approximation (see e.g. R. Haberman, Mathematical Models, Prentice Hall, N.J. 1977). For $\rho \rightarrow 0$, $u \rightarrow u_m$, at $\rho \rightarrow \rho_m$, $u \rightarrow 0$
- The equation for the density evolution may be written as the Burgers equation

$$\partial \rho / \partial t + \partial / \partial x \left\{ \left(a + \frac{1}{2} b \rho \right) \rho \right\} = 0$$

One-dimensional models of traffic-II

- Another model for 1D traffic flux behavior: $j = \rho u \ln(\rho_m / \rho)$
- We can rewrite the traffic model as

$$\frac{\partial \rho}{\partial t} = - \left(\frac{d}{d\rho} \rho v(\rho) \right) \frac{\partial \rho}{\partial x} = -c(\rho) \frac{\partial \rho}{\partial x},$$

Here $c(\rho)=dj/d\rho$ is the effective velocity: the speed with which disturbances in the traffic flow would travel. It is analogous to the group velocity in the wave theory and to velocity of quasiparticles in many-body theories. The collective speed $c(\rho)$ may be both positive and negative, which means that disturbances can propagate opposite to the traffic flow direction. $c(\rho)$ in the linear model is also linear in the car density and takes the values $c(0) = v_m$ and $c(\rho_m) = -v_m$

- Since $c(\rho) \leq v(\rho)$, cars move faster than collective excitations
- The concept of collective excitation is ubiquitous, when one tries to describe the processes occurring in material media

Method of characteristics

- For $c(\rho) = \text{const}$ the initial density distribution $\rho(x, t = t_0) = \rho_0(x)$ propagates unchanged with the speed c . In other words, the line $dx/dt=c$ is the line of $\rho = \text{const}$. However, even in the nonlinear case the same holds true, since along the characteristic

$$\frac{d\rho(x(t), t)}{dt} = \frac{\partial \rho}{\partial t} + \frac{dx}{dt} \frac{\partial \rho}{\partial x} = \frac{\partial \rho}{\partial t} + c(\rho_0) \frac{\partial \rho}{\partial x}$$

and $\frac{d\rho(x(t), t)}{dt} = 0$ due to conservation. Each characteristic is a line of constant density (or of other local parameter, e.g. temperature)

- For small nonlinearity $\rho(x, t) = \rho_0 + \varepsilon \rho_1(x, t)$ we have

$$\frac{\partial \rho_1}{\partial t} + c(\rho_0) \frac{\partial \rho_1}{\partial x} = 0$$

- the disturbance propagates with the speed $c(\rho_0)$. This reflects the situation when cars

are nearly evenly spaced. The above equation describes a density wave propagating with a constant speed – typical for linear models

What is the optimal velocity?

How to obtain the maximum flux of traffic, $j = \rho v + j_D = \max$?

Here j is the total flux, j_D is the random (diffusive) flux. In the simple 1D steady-state model we can neglect j_D

Since $\frac{\partial \rho}{\partial v} = -\frac{\rho_m}{v_m} \neq 0$ the extremality criterion $\frac{\partial j}{\partial v} = \frac{\partial j}{\partial \rho} \frac{\partial \rho}{\partial v} = 0$

leads to $\frac{\partial j}{\partial \rho} = 0$ For our model, this means:

$$\frac{\partial [\rho v_m (1 - \rho / \rho_m)]}{\partial \rho} = -2v_m \rho / \rho_m + v_m = 0$$

or

$$\rho_{opt} = \rho_m / 2$$

and the optimal velocity is half of the maximum, $v_{opt} = v_m / 2$

The speed $\partial j / \partial \rho$ is approximately that with which disturbance propagates. Thus, if density is higher than ρ_{opt} , the small rise in density propagates backwards: subsequent drivers are slowing down

Traffic jams, rarefaction and shock waves

- Stability of a uniform traffic flow (UTF). UTF is defined as the state with constant density ρ_0 and constant velocity v_0 . Small deviations may grow uncontrollably resulting in congestion.
- In actual traffic there may be discontinuities. For instance, a traffic light may be described by a step (Heaviside) function:

$$\rho(x, t = t_0) = -\rho_m \theta(x), \quad \theta(x) = 0, x < 0, \theta(x) = 1, x \geq 0$$

Behind the light, cars are bumper-to-bumper, when the light turns green not all cars start moving at once: the density decrease propagates backwards with a finite speed – rarefaction wave

- The square pulse initial condition $\rho(x, t = t_0) = \rho_m, -a < x < 0$ gives rise to a shock front at $x = -a$ (ρ is multivalued, low and high density characteristics intersect). Shock waves – drastic changes in traffic density - are dangerous, since drivers have finite reaction time. Jams are, in fact, phase transitions due to instability

More complicated model: bi-directional traffic

- Let ρ_+, ρ_- denote the vehicle densities in the right and left lanes, average velocities $u \geq 0$ and $v \geq 0$, then we have

$$\frac{\partial \rho_+}{\partial t} + \frac{\partial(\rho_+ u)}{\partial x} = 0, \frac{\partial \rho_-}{\partial t} + \frac{\partial(\rho_- v)}{\partial x} = 0$$

- the system of 2 equations with 4 unknowns; it may be closed by analogy with unidirectional model, $u = u_m - \alpha \rho_+, v = v_m - \beta \rho_-$
- A generalization of this linear model ($\alpha > 0, \beta > 0$):

$$u = Au_m + Bv_m - \alpha \rho_+ - \beta \rho_-, v = Cu_m + Dv_m + \alpha \rho_- + \beta \rho_+$$

- We may require anti-symmetry between the two lanes, so they are physically identical: $u(\rho_+, \rho_-) = -v(\rho_-, \rho_+)$ This condition may be satisfied by putting $A=D=0, B=-C=1$.

Here we have a two-component macroscopic compressible fluid.

Bi-directional traffic-II

The system of model relations for two lanes can be written as

$$u = v_m - \alpha\rho_+ - \beta\rho_-, v = -u_m + \alpha\rho_- + \beta\rho_+$$

and inserting it into continuity equations, one obtains the condition

$$(\rho_+^2 + \rho_-^2)(2\alpha + \beta)^2 + 2\rho_+\rho_-(4\alpha^2 - \beta^2 + 4\alpha\beta) + (u_m + v_m)^2 - 2(\rho_+ + \rho_-)(\beta + 2\alpha)(u_m + v_m) \geq 0$$

under which the roots (eigenvalues) of the characteristic equation

$$\det(B - \lambda A) = 0$$

are real. Here A and B are matrices in the first order conservation law:

$$A\rho_t + B\rho_x = 0, \rho = \begin{pmatrix} \rho_+ \\ \rho_- \end{pmatrix} \quad \text{This condition geometrically represents the region outside an ellipse in the plane}$$

$$(\rho_+, \rho_-): fP^2 + gQ^2 \geq E$$

The quantities E, f, g can be obtained from the parameters α, β, u_m, v_m

$$P = \frac{1}{\sqrt{2}} \left(\rho_+ + \rho_- - \frac{(2\alpha + \beta)(u_m + v_m)}{4\alpha(\alpha + \beta)} \right), Q = \frac{\rho_- - \rho_+}{\sqrt{2}}$$

Molecular dynamics

- An obvious way to simulate condensed media on a computer is to model the collective motion of individual molecules that make up the media. However, this is an enormous task for a macroscopic ($N \sim N_A$) system of strongly interacting molecules
- If the molecular system is rarefied, the collective motion of molecules produces the fluid-like behavior
- If it is possible to “sum up” all the molecules and, provided their interactions are modeled correctly, properties of fluids can be analyzed on a microscopic level
- Various situations and properties of the fluid can be modeled by changing the parameters of inter-molecular interactions, average energy E and separation of molecules $n^{1/3}a$
- In the case of classical particles, modeling corresponds to variation of input parameters U/E and $n^{1/3}a$, where U is the interaction energy, a is the effective radius of a molecule

Example of PDE modeling: spread of a pollutant

- Suppose that a mass M of some hazardous chemical or radioactive substance is suddenly released into the environment. The model is aimed to determine the subsequent concentration of the pollutant at various distances from its source.
- The model is greatly simplified if we consider the spherically symmetric medium; influence of boundaries can be discussed at the next stage of modeling
- The PDE governing the pollution spread:
$$\frac{\partial C}{\partial t} = D \left(r^{-2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial C}{\partial r} \right) \right)$$

where $C(r, t)$ is the pollutant concentration and D is its diffusivity in the environment. At long distances $C(r, t) = 0$ which serves as boundary condition. The initial condition:

where M is the total mass of the pollutant

$$M = 4\pi \int_0^{\infty} C(r, 0) r^2 dr$$

Spread of a pollutant: solution

- We can use scaling, normalizing $M=1$ and $D=1$ and putting $r = e^\alpha \tilde{r}, t = e^\beta \tilde{t}, C = e^\gamma \tilde{C}$ where α, β, γ are scaling constants to be determined. These constants define 3-parameter stretching or contracting transformation under which the PDE-model is invariant
- Substituting the scaling-Ansatz, we have
$$e^{2\alpha-\beta} \frac{\partial \tilde{C}}{\partial \tilde{t}} = \frac{1}{\tilde{r}^2} \frac{\partial}{\partial \tilde{r}} \left(\tilde{r}^2 \frac{\partial \tilde{C}}{\partial \tilde{r}} \right)$$

Thus the PDE is invariant if and only if $2\alpha - \beta = 0$

Substituting the scaling transform into initial condition, we get for $\tilde{t} = 0$

$$4\pi e^{3\alpha+\gamma} \int_0^\infty \tilde{C} \tilde{r}^2 d\tilde{r} = 1 \quad \text{which is scaling-invariant when } 3\alpha + \gamma = 0$$

So, due to scaling, the 3-parameter transform is reduced to

a 1-parameter transform

$$r = e^\alpha \tilde{r}, t = e^{2\alpha} \tilde{t}, C = e^{-3\alpha} \tilde{C}$$

Spread of a pollutant: solution-II

- Now we can find dimensionless (invariant) combinations of variables: $\mu \equiv e^\alpha = r / \tilde{r} = (t / \tilde{t})^{1/2} = (C / \tilde{C})^{-1/3}$
 - Introducing the variables $\xi = \tilde{C}\tilde{t}^{3/2} = Ct^{3/2}$, $\eta = \tilde{r}\tilde{t}^{-1/2} = rt^{-1/2}$, we notice that in order the solution to be scaling-invariant, these two variables should be related as $\xi = f(\eta)$ or $C(r, t) = t^{-3/2} f(rt^{-1/2})$
- The latter expression means that we can reduce PDE to an equivalent ODE depending on the variable η :
- $$\frac{d}{d\eta} \left(\eta^2 \frac{df}{d\eta} \right) + \frac{1}{2} \left(\eta^3 \frac{df}{d\eta} + 3\eta^2 f \right) = 0$$
- which leads to a family of solutions for the concentration:
- $C(r, t) = At^{-3/2} e^{-r^2/4t}$, $A = \text{const}$ The constant A can be found from the normalization to total mass M :
- $$4\pi \int_0^\infty dr r^2 C(r) = 4\pi \int_0^\infty d\eta \eta^2 f(\eta) = 1$$
- So the concentration:

$$C(r, t) = (1 / 4\pi t)^{3/2} e^{-r^2 / 4t}$$

Spread of a pollutant: interpretation

- To obtain the solution of the pollutant spread problem, we used the mass conservation $\int C(\mathbf{r}, t) d^3 r = M = M_0 = \text{const}$
Formally, the integral should depend on the parameter t , but in fact it does not depend on t . One can also explore the temporal behavior of moments $M_1^i = \int x^i C(\mathbf{r}, t) d^3 r, M_2^{ik} = \int x^i x^k C(\mathbf{r}, t), \dots$ which give information of the asymmetric concentration distribution
- At $r=0$ the concentration has a singularity, $C \sim t^{-3/2}$, which shows that the whole mass of pollutant is initially highly concentrated. For small t the concentration is close to zero everywhere, except near the origin, but for large times ($t \gg r^2 / 4$) the concentration is practically independent on distance (almost uniform). This means that diffusion tends to rapidly smooth out sharp peaks and flatten all distributions (entropy rise).
- **Spread of a pollutant in this model is a self-similar phenomenon, i.e. spatial distributions at various moments can be produced by a similarity transform**

Spread of infectious disease

- The world is becoming increasingly vulnerable to the outbreaks of terrorist activities and sudden epidemics. Infectious diseases are caused by a variety of pathogens (viruses, bacteria, fungi, protozoa) that can pass from host to host. Mathematical and computer modeling of this process is usually based on subdivision of the total population into 3 classes (stocks):
 1. Individuals susceptible to the disease but not yet infected (S)
 2. Persons who are infected and capable of transmitting the disease (I)
 3. Removed persons: those who are no longer capable of transmitting, due to recovery, quarantine or death (R)

The disease is transmitted horizontally by interaction between S and I classes. The transfer of I to R class is assumed to occur at constant rate. The model is close to that of chemical kinetics (see Slides 270 - 274). This 1-2-3 model is often called the SIR model. Every person within this model should be located inside one of these stocks

Spread of infectious disease-II

- The spatially homogeneous (ODE) model:
- The spatial spread of infectious disease can be modeled by reaction-diffusion system. More sophisticated models include population demographics, e.g. by introduction of logistic growth terms (see Slides 184+).

$$\begin{aligned}\frac{dS}{dt} &= -\alpha I(t)S(t) \\ \frac{dI}{dt} &= \alpha I(t)S(t) - \beta I(t) \\ \frac{dR}{dt} &= \beta I(t)\end{aligned}$$

Demographics can be also modeled by introducing a special variable q representing age (age structured models), so that $\rho(x, q, t)$ gives the population density at the location x , having age q at the moment $t > 0$. Spatial population density between $[a, b]$ years of age is given by

$$P(x, t; a, b) = \int_a^b \rho(x, q, t) dq$$

Here $x \in \overline{\Omega}, \Omega \in \mathbb{R}^2$, with smooth boundary $\partial\Omega$

Spread of infectious disease-III

- Disease-free population model:

where $\beta > 0$ is the birth rate and $\mu > 0$ is the mortality rate.

At the boundary of the region,

we have the Neumann conditions: $\partial \rho(x, q, t) / \partial n = 0, x \in \partial \Omega, t \geq 0, q \geq 0$

Then the total population is subdivided into 4 classes: susceptibles (S)

$u(x, q, t)$, exposed (E) $v(x, q, t)$, infected/infectious (I) $z(x, q, t)$, removed (R) $w(x, q, t)$

- This model is called in mathematical epidemiology SEIR model

The period of exposure is assumed to be fixed, $\tau = \text{const} > 0$

The transfer of the disease is the result of interaction between infected and susceptibles, i.e. can be modeled by “collision frequency” $v(u, z)$

$$v[u, z] = \gamma_1(q, t)z(x, q, t) + \int_{\tau}^{\infty} \gamma_2(q', t)z(x, q', t)dq' \quad \text{so that the incidence term is } -vu$$

$$\partial \rho / \partial t + \partial \rho / \partial q = D \Delta \rho - \mu \rho$$

$$\rho(x, 0, t) = \int_0^{\infty} \beta(q, t) \rho(q, x, t) dq$$

$$\rho(x, q, 0) = \rho_0(x, q)$$

Spread of infectious disease-IV

- Thus we arrive at the following PDE model for people susceptible to infectious disease:

$$\partial u / \partial t + \partial u / \partial q = D\Delta u - \mu u - \nu u$$

$$u(x, q, 0) = u_0(x, q); u(0, t) = \int_0^\infty \beta(q, t)u(q, t)dq + \int_\tau^\infty \beta(q, t)w(q, t)dq$$

$$\partial u / \partial n = 0, x \in \partial\Omega, q \geq 0, t \geq 0$$
- Analogous PDE models can be written for classes v (exposed), w (removed/recovered), and z (infected). The solution of these PDE problems may be obtained using the Green's function techniques
- Models of spatio-temporal evolution based on PDEs for epidemic dynamics still require detailed field studies. Robust prediction of the geographic spread of infectious disease can be validated by comparison with actual data.
- The basic use of epidemiological modeling is to determine under which conditions a disease will become an epidemic (i.e. the infection rate surpasses the recovery rate)**

Spread of infectious disease – ODE modeling

- A simplified model of the infection spread can be produced by using logistic equation. Let p be the infected fraction of the population, then $(1 - p)$ is the susceptible fraction, and the time variation of p can be modeled as
$$\frac{dp}{dt} = \alpha p(1 - p) - \gamma p$$
where α is the transmission rate of the disease and γ reflects the recovery (removal) of the infected population. These parameters are specific for the disease and, in general, for different populations.
- Investigation of this logistic equation may be performed in the same fashion as on Slide 187 (separation of variables, integration and inversion). Rates α and γ are control parameters.
- Discrete analogs and critical parameters can be obtained, describing the transition to catastrophic amplification of a disease, in the same way as for the general logistic model. The transmission rate α can be reduced by quarantining; γ can be raised by immunization

Forest and factory fires

- When studying thermal explosions (Slides 272-274), we have found “hot” and “cold” equilibrium temperatures, T_1 and T_2
- If the ambient temperature exceeds the “hot” point, spontaneous ignition may occur, with the temperature quickly rising to very high values. The physical reason for such instability is that the heat rate produced by the chemical reaction begins to surpass the heat loss from the surface, and the whole volume is heated still more (positive feedback)
- The detailed theory should include heat conductivity, which makes the model rather complicated (nonlinear PDE modeling)
- A number of factories producing or processing combustible stuff caught unexplained fires. A typical accident is spontaneous ignition of grain in elevators and fires in wood processing factories
- Spontaneous ignition of peat beds, forests and even villages around Moscow (1972, 2002); Spain. Portugal and South of France (2003)

Spontaneous ignition: a mathematical model

- The chemical reaction of mildly combustible material with oxygen (sometimes inflammable lubricants are also present) produces internal heat. The higher the temperature, the more heat is produced
- If the heat cannot diffuse and escape fast enough risky temperatures may result. The thicker the layer, the higher the hazard of ignition.
- The crucial task is to determine the critical volume of material
- A 1D model may be stated as
$$\rho c \frac{\partial u}{\partial x} = k \frac{\partial^2 u}{\partial x^2} + q$$

where q is the heat (per unit volume) produced by chemical reaction, c is the specific heat, k is heat conductivity and ρ density. One may take the boundary conditions as, e.g., $u(0,t)=T_2$ (contact with hot surface) and $-k\partial u(l,t)/\partial x = \alpha(u(l,t) - T_1)$ - convection cooling, T_1 is the air temperature, α is the respective coefficient

Since q is a nonlinear function of u , the whole problem is nonlinear

Spontaneous ignition: mathematical model-II

- The simplest heat generation model is reduced to: $q=BKq_0$, where B is the concentration of the reacting substance, K - the reaction rate and q_0 is the elementary heat of reaction. Using the Arrhenius law (in fact, following from the Gibbs distribution), we have:

$$q(u) = q_0 B A e^{-E/k_B u}$$

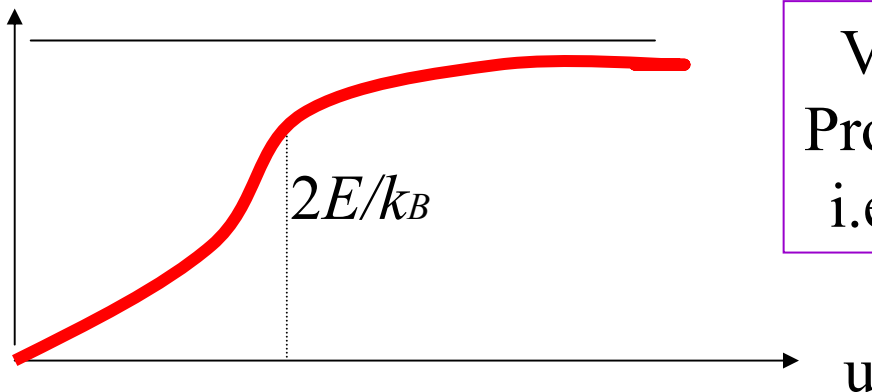
k_B is the Boltzmann constant. We can approximate near $u=T_1$:

$$u^{-1} = T_1^{-1} - (u - T_1) / T_1^2 \text{ and } E / k_B u \simeq 2E / k_B T_1 - Eu / k_B T_1^2$$

so that

$$q(u) = b e^{\beta u}, b = B q_0 e^{-2\beta T_1}, \beta = E / k_B T_1^2$$

q



Various possible relationships $q(u)$
Produce different nonlinear solutions,
i.e. different regimes of combustion

Simplified regime: no heat diffusivity

- It is instructive to find correspondence between a PDE and ODE modeling concepts. The local model is described by ODE:

$$\dot{u} = -u + \gamma e^u$$

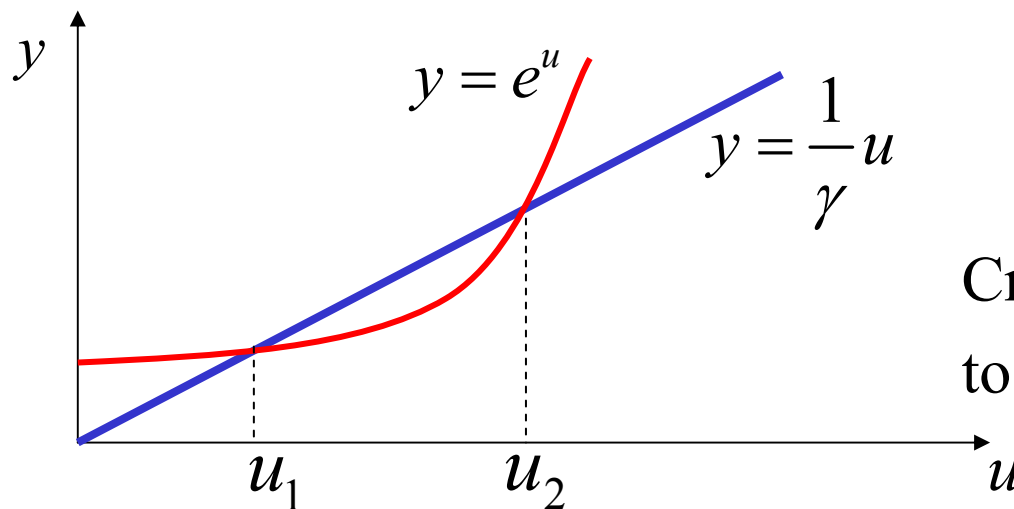
where γ is a positive constant and u is dimensionless temperature.

- This equation is similar to the PDE on the two preceding slides, but the heat conduction term is missing. Usually, it is necessary to know how the solution changes when the parameter γ is varied.
- Physically, this parameter is proportional to the quantity of the reactant and to the heat of reaction. One can use Maple command **dsolve** to analyze numerical solutions of this model for various values of γ .
- Between $\gamma = 0.3$ and 0.4 the critical change of behavior is observed – **a bifurcation**. Critical values u_* , γ_* can be found analytically, numerically or graphically from the equilibrium condition

$$du / dt = 0$$

Steady-state solutions and bifurcations

- Since bifurcations between different steady-state (equilibrium) solutions are very important, e.g. for phase transition models, we will show how to determine critical parameters u_*, γ_* (see also Slide 273). We have already met bifurcations in dynamical systems
- One can find the number of equilibrium solutions by setting $\dot{u} = 0$ which gives $u / \gamma = e^u$. Graphically, equilibrium solutions correspond to intersection points of two curves

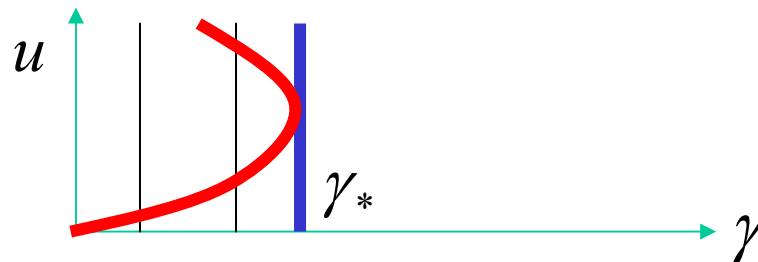


u_1 is “cold” and u_2 is “hot” equilibrium temperature.

Critical point corresponds to confluence of u_1 and u_2

Critical values and critical behavior

- With the increase of parameter γ , a state is reached when there are no solutions. Critical values γ_* and u_* can be obtained by solving simultaneously equations $\frac{1}{\gamma_*} u_* = e^{u_*}$ and $\frac{1}{\gamma_*} = e^{u_*}$ (tangents must coincide)
- From these equations we get $u_* = 1$ and $\gamma_* = 1/e \approx 0.368$
- The critical value γ_* corresponds to critical behavior, in the case considered here it means that spontaneous ignition occurs
- Critical behavior may be represented by a plot - a bifurcation diagram, $u = f(\gamma)$. This graph shows how two solutions (“cold” and “hot”) confluence, so the number of solutions progresses from 2 to 1 and then to 0.



Forest fire propagation

- The real PDE-based model of forest fire should include its moving with the wind (as inhabitants of Moscow region were reminded in 2002, those of France, Spain, Portugal in 2003). So it is necessary to model spontaneous ignition together with atmospheric variables (velocity, pressure, temperature, entropy, humidity, etc.)
- When modeling climate as a physical system and weather as the actual state of the atmosphere, a coarse grid is used, with cell side 30 to 50 km. This is sufficient for predicting rain within a limited time interval (~ 3 days), but if the fire behavior is to be predicted, higher resolution modeling (~ 1 km) should be employed.
- The *Farsite* model (Finney, 1993) – behavior of fires as a function of weather variables. Coupling of this model with high-resolution atmospheric model is needed (weather as input) in order to obtain fire spread forecast. However, physics of the fire spread is hard to simulate, because of numerous variables involved (e.g. vegetation)

Pulse propagation models

- Any communication system is using a transmission medium (e.g. copper wires, air, fiber optics). Information is transmitted through such medium with a certain distortion or spread, which limits the communication capabilities (e.g. in networking and telecom).
- This problem is of utmost importance and has been tackled by a number of outstanding scientists (Raleigh, Sommerfeld, Brillouin, Mandelstam, Ginzburg). A lot of books have been devoted to this subject where complicated models of pulse evolution are given.
- With the advent of femtosecond technology, fiber optics, and pulse compression techniques, exact description of ultrashort pulse spread is increasingly valuable, e.g. for computer optics. The problem of pulse deformation is also important for quantum engineering limits
- For simplicity, let us consider 1D scalar pulse propagation in homogeneous media. This model describes all the main propagation features, including dispersion and attenuation. Negative attenuation corresponds to amplification in active media (lasers pulses)

Pulse propagation-II

- Mathematically, the task is close to the advection problem:

$$\frac{\partial w}{\partial t} + a \frac{\partial u}{\partial x} = 0 \quad (*)$$

where w and u are functionally dependent $w = w(u), u = u(w)$ and can be obtained by the method of characteristics:

$u = F(x - at), w = G(t - x/a)$. Here F corresponds to the evolution of initial field distribution, whereas G corresponds to its propagation

- Equation (*) corresponds to ray kinematics (geometrical optics), which ignores the wave nature of the pulse. This equation always describes the transport of some material along x-axis and serves as a model to sharpen numerical methods.
- The same equation was used to analyze traffic flows (see Slides 303-313), specifically kinematic shock waves. Numerical schemes for the advection equation are reviewed on Slides 175-178

Electromagnetic (EM) pulse propagation

- More detailed propagation problem can be formulated as follows: one must find the field $E(x,t)$, $x > 0$, $-\infty < t < \infty$ if the field $E(0,t)$ is known for all t . The problem can be solved by a Fourier transform.

$$\frac{\partial^2 E(x,t)}{\partial x^2} - \frac{\hat{\varepsilon}}{c^2} \frac{\partial^2 E(x,t)}{\partial t^2} = 0$$
 - a 1D wave equation where $\hat{\varepsilon}$ is the dielectric operator accounting for dispersive properties of the medium. In Fourier space $\frac{d^2 E(x,\omega)}{dx^2} + k^2(\omega)E(x,\omega) = 0$ and

$$E(x,t) = \int_{-\infty}^{\infty} E(0,\omega) e^{i(k(\omega)x - \omega t)} \frac{d\omega}{2\pi} = \int_{-\infty}^{\infty} dt' E(0,t') \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t') + ik(\omega)x} =$$

$\int_{-\infty}^{\infty} dt' g(x,t-t') E(0,t')$ Here $g(x,t)$ is space-time Green's function, $g(0,t) = \delta(t)$ where $\delta(t)$ is the Dirac δ -function.

$g(x,t)$ defines the pulsed media response at a point x . From the formal point of view, the problem is solved (reduced to an integral). However technical difficulties of computing the resulting integral remain serious

Pulse propagation - interpretation

- At the preceding slide propagation of electromagnetic pulse was discussed. However, the same method can be applied to acoustic, hydrodynamic, even gravitational pulses in linear approximation.
- Pulses may have both deterministic and random component. In modern radio engineering (RE) both are important for information transmission. In RE pulses are usually represented by a complex amplitude: $E(x, t) = A(x, t)e^{i(k(\omega_0)x - \omega_0 t)}$ ω_0 - carrier frequency.
- One the crucial problems, e.g. of fiber optics communication, is the pulse spread, which may be calculated from the convolution integral at the preceding slide: $\tau^2(x) = \sigma^2 + \tau^2(0)$, where τ , σ are spread variances of the pulse and medium response function. From the practical viewpoint, only such distances x are important where the incoming pulse is similar to initial one and can be registered in the frequency range close to the sending device.
- In modern technology, such an uncommon object as the pulse with duration $\tau \approx T$ (wave period) appeared (Fourier is inadequate)

Nonlinear pulse and wavetrain models – basic equations

- The radical difference between linear and nonlinear pulse propagation is that in the nonlinear mode the properties of the medium $k(\omega)$ are determined not only by the medium, but also by the sheer presence of the pulse. Physically, it means that the pulse field is strong, as e.g. the EM field in the laser pulse $u_t + Lu = \varepsilon Mu$
- From the mathematical point of view, it means that e.g. the advection equation acquires a nonlinear term:
where M can be nonlinear. **Examples:**

- a) no dispersion, no dissipation (traffic flow) $u_t + v(u)u_x = 0;$
- b) Burgers equation (dissipative media) $u_t + v(u)u_x - \alpha u_{xx} = 0;$
- c) Korteweg-de Vries (dispersive media) $u_t + v(u)u_x + \beta u_{xxx} = 0;$
- d) Kadomtsev-Petviashvili (2D solitons) $\partial_x (u_t + uu_x - \gamma u_{xxx}) = \varepsilon u_{yy}$
- e) nonlinear Schrödinger equation $(A_t + vA_x) - \frac{i}{2} \omega_{kk} A_{xx} + i\alpha |A|^2 A = 0$
- f) sine-Gordon equation (see Slides 217,218) $u_{tt} - u_{xx} + \sin u = 0$

Nonlinear pulse/wave propagation – discussion

- *Latest motivation:* nonlinear optics (new devices); geophysics (oil and gas, fluid flows, natural catastrophes); telecommunications
- *Original motivation:* plasma, fluid dynamics (generalized solutions)
- *Transition to nonlinear mode:* what happens when gradients at the wavefront tend to infinity? The answer depends on the physical situation: in fluids the wavefront is destroyed to splashes; for a flux of noninteracting particles discontinuity leads to non-uniqueness (multiflow mode – many fluxes moving with different velocities); for acoustic and EM fields non-uniqueness is unacceptable, so the nonlinear evolution depends on the competition between dispersion and dissipation.
- *For lumped systems* the main model is nonlinear oscillator:
 $\ddot{x} + f(x) = 0$ *For distributed systems*, there are many nonlinear wave models. Large resource on *numerical methods* for nonlinear waves:
<http://acs366.phys.msu.su/isna16/Abstracts/section1.htm>

Computer modeling of waves: visualization of functions of two variables

- How to locate the disturbance and determine its speed?
- Four popular visualization methods: 1) xt -diagrams; 2) slice plots; 3) surface plots; 4) animation.
- 1) xt -diagram (density plot) - $u(x,t)$ is described by shades of gray or color to indicate the magnitude of u . MATLAB: **pcolor (x,t,u)**;
In Maple: first define $u(x,t)$ then load the plots package:
- ```
> u :=(x,t) -> u(x,t);
> with(plots)
> densityplot(u(x,t), x=a..b, t=0..T);
```
- MATLAB contains a package Wavelet Toolbox, which is useful for short pulse analysis
- 2) Slice plot – MATLAB: **waterfall (x,t,u)**;  
Maple: **> slices := {seq([x,t,u(x,t)], t={0,2,4..T})};**  
**> spacecurve(slices, x=a..b)**

# Surface plots and wave animation

3) Surface plot – MATLAB: `surf(x,t,u);`

Maple: `>plot3d(u(x,t),x=a..b, t=0..T);`

4) Animation - In MATLAB: function **movie**, in Maple: command **animate**

- MATLAB M-file:  
**M = moviein(length(t));**  
**for j=1:length(t),**  
**plot(x,u(j,:), M(:,j)=getframe;**  
**end;**  
**movie(M)**

- Maple:  
`> animate(u(x,t), x=a..b,t=0..T);`

For visualisation of 3D vector field the MATLAB commands:  
coneplot – vector arrowfield;  
streamline gives the flux;  
streamribbon adds the rotation;  
streamtubes – the tube width depends on the vector field divergence

# Virtual prototypes

- Complex products like aircraft, cars, trucks, locomotives, ships, power generators, nuclear plants, etc. include interacting electrical, mechanical, hydraulic, electronic, and software components
- Virtual prototyping: computer model replaces a physical prototype. Virtual prototyping software allows to build a computer model with moving parts, then simulate its behavior and optimize the design
- Benefits of this approach: reduced time-to-market, lower costs, rapid evaluation of large number of design alternatives, survival of optimal design long before building the physical prototype
- The actual state: many commercial simulation tools in a number of application areas. Most of the tools are strong in particular fields and do not cover the whole range of applications.
- Engineers use a number of different tools to model various aspects of the same product. This is far from ideal: too much time is spent on training, installation, support, negotiations, etc. Manufacturers often require an integrated tool, e.g. supplementary to CAD/CAM

# Co-simulation

- The approach: software tools call COM objects and vice versa
- E.g. **Simulink** can model control aspects, whereas a **co-simulator** models hydraulic and thermal parts of a complex system
- **Example:** in today's aircraft the fuel system not only supplies fuel to engines, but is also used as movable ballast to control the center of gravity, with avionics controlling valves and pumps
- Exact amount of fuel is transferred between tanks, with flow-meters providing necessary feedback
- Control and hydraulic systems must be modeled together to ensure optimal system design – **co-simulation** may be useful
- **Co-simulation** like multiphysics can provide universal modeling environment, with geometric and topological flexibility
- When **co-simulation software** models the behavior of a system, it automatically computes reactions (mechanical, thermodynamical, etc.) at each sensitive unit or point. Co-simulation is successfully used for modeling complex nonlinear dynamical systems

# Model implementation in the software:

## 1 - abstraction

- World  $\rightarrow$  Model  $\rightarrow$  Program
- Constructing an abstract concept of a system focusing on relevant aspects (ignoring details). **Examples:**
  - **Mendeleev's periodic system** – table of elements, details like availability or combinations frequency are omitted
  - **In astrophysics**, objects are classified and identified by their spectral properties, mostly in the infrared (IR) region. Chemical composition, volume, even location are largely ignored
  - **In biology**, classification of species ignores their behavior
  - **In behavioral sciences**, classification ignores habitus
  - **Gravitation** is irrelevant for **nuclear collision** processes
  - **Quantum theory** is irrelevant for calculation of **asteroid motion**
  - **Photon** properties are irrelevant (so far) for **car headlight** design

# Model implementation in the software:

## 2 – Object Oriented Approach

- Object is defined through its state and behavior
- Object state is specified by **fields** (data): they say what an object is
- Object behavior (functionality) is specified by **methods**: the latter say what an object does
- Object is distinguishable from other objects (has an identity)
- Object is “encapsulated”: its state can be changed only due to interaction with other objects
- Object is an instance of a class, different objects of the same class have similar fields and methods, but values of the fields (data) in general differ. E.g. all humans have hair color but the color of an individual’s hair can be different from others

**Questions:** Can the world be fully described through objects? Is everything an object? E.g. is a number an object? Are you an object?

# Model implementation in the software:

## 3 - UML

### The Unified Modeling Language

- Class and object modeling
- Component modeling
- Business process modeling with use cases
- Distribution and deployment modeling
- Engineering, reengineering, interface engineering, greenfield (from scratch) engineering

UML helps to specify, visualize and document models of software systems

**Diagram techniques:** use case diagrams; class diagrams; package diagrams; interaction diagrams; activity diagrams; state charts; sequence diagrams; deployment diagrams;

**UML modeling is the designing of software before coding**

UML Web sites: [www.uml.org](http://www.uml.org); [www.rational.com/uml](http://www.rational.com/uml); [www.omg.org](http://www.omg.org)

# Contemporary trends

- One can see that the modern science drifts more and more towards mathematical modeling. Even physics, which has always been based on experiment, is now developing similarly to mathematics, increasingly employing assumptions, relying on “inner beauty” instead of experimental validation and using computer simulations.
- One can also see that the methods applied in totally different areas are essentially quite similar. This fact reflects the progressive trend towards interdisciplinarity.
- A number of subjects in science (e.g. string theory, modern gauge theories, cosmological models, etc.) cannot be experimentally verified, at least in the foreseeable future. It is hard to imagine any experimental device - to test any statement - operating at energy over 100 GeV and respectively inaccessible distances. So one has to rely on modeling.
- In some field of engineering modeling (e.g. nuclear engineering) modeling tends to replace actual experiments



# Future trends

- A large class of models, namely quantum and stochastic modeling, requires a separate treatment (Parts 2 and 3).
- Stochastic modeling can combine simulation and data analysis.
- Stochastic modeling is becoming increasingly popular for the description of biochemical processes. In fact, if life can be treated by mathematical modeling, it may be based on stochastic models.
- A new field of application of stochastic modeling is “physical economy” or “econophysics” – a discipline applying methods from physics to economic systems. Most of the conventional economic stereotypes can acquire precise meaning in physical economy. The same applies to modeling of social and behavioral processes.
- Complex systems can be modeled not only by stochastic equations. For instance, new computing schemes, such as quantum computer, are based on mathematical models that are inherited from quantum mechanics. The latter has evolved into an engineering discipline providing a hardware basis for computation.

# Concluding remarks

- This course is not a threaded textbook, rather a collection of tasks from various fields of science and engineering.
- The course was intended to demonstrate how mathematical models can be processed analytically or numerically and implemented on computers. **Computers are indispensable for modeling.**
- Some models could not be treated profoundly enough, simply due to the sheer size of calculations. So the accent by necessity was put on problem formulation and main ideas vs. computational details. The slides may seem to contain too much material – the idea was to make them self-sufficient for the initial study of subject models.
- Some instructive models had to be omitted, in most cases due to lengthy computation procedures, inadmissible in a short course. **Examples:** military actions accompanied by terrorist activities; flight of an aircraft; missile dynamics; passive tracer propagation in ecology; black hole models; climate models; nuclear winter, etc.
- Model building becomes an essential step in constructing a theory

# Literature to Part 1 of the course

- **V.I. Arnol'd.** *Ordinary Differential Equations*. Springer 1992. – Probably the best known book on ODE.
- **P. Blanchard, R.L. Devaney, G.R. Hall.** *Differential Equations* (with CD-ROM). Brooks/Cole 2002 – especially Ch.5 on dynamical systems.
- **H.D. Abarbanel, M.I. Rabinovich, M.M. Sushchik.** *Introduction to Nonlinear Dynamics*. World Scientific 1993. - Useful examples and mathematical models. Clear, interesting and simply written.
- **N. Gershenfeld.** *The Nature of Mathematical Modeling*. Cambridge University, 2001. – A good interdisciplinary source.
- **H.Haken.** *Synergetics*. Springer, 1996. - A classical work.
- **R.Z. Sagdeev, G.M. Zaslavsky.** *Nonlinear Physics*. Harwood, 1988. – A book by top-class Russian-US scientists, very instructive.
- **J.M. Smith.** *Mathematical Ideas in Biology*. Cambridge, 1977. – Enjoyable reading, many nontrivial examples.

# Literature-II

- **L.D. Landau, E.M. Lifshits.** *Classical Mechanics*. – The classical work, the first part of the best known course of theoretical physics.
- **H. Goldstein.** *Classical Mechanics*. Addison-Wesley, 1982. – A well known textbook, a classical work.
- **A.J. Lichtenberg, M.A. Lieberman.** *Regular and Stochastic Motion*. Springer 1992 – Non-trivial and comprehensive, can be a little difficult for the first reading.
- **J.C. Polking.** *ODE using MATLAB*. Prentice Hall, NJ 2002. – Many useful tricks.
- *Using MATLAB, Version 6*. MathWorks 2001, The chapter on PDE.
- **D. Richards.** *Advanced Mathematical Methods with Maple*. Cambridge, 2001. - A very comprehensive volume, a successful merge of computer algebra and analytical methods.
- **G.R. Fulford, P. Broadbridge.** *Industrial Mathematics*. Case studies in the diffusion of heat and matter. Cambridge 2002

# Literature-III

- **A.A. Samarski, A.P. Mikhailov.** *Mathematical Modeling.* Moscow, Nauka, 1997 (In Russian). A comprehensive book containing many original results obtained by well-known Russian mathematicians. An interesting collection of models
- **G.R.Fulford, P. Broadbridge.** *Industrial Mathematics.* Case studies in the diffusion of heat and matter. Cambridge, 2002.
- **E.Bohl.** *Mathematik in der Biologie.* Springer, 2001. Many subjects from the present course are treated in this book.
- **L. Edelstein-Keshet.** *Mathematical Models in Biology,* McGraw-Hill, 1988.
- <http://www.bio.brandeis.edu/biomath>
- **M.T.Heath.** *Scientific Computing.* An introductory survey. McGraw-Hill, 1996.
- **R.N. Mantegna, H.E. Stanley.** *An Introduction to Econophysics.* Cambridge University Press, 1999.

# Literature-IV

- **B. Bruegge, A.H. Dutoit.** *Object-Oriented Software Engineering*. Prentice Hall, NJ 2000. - A book on UML-based modeling containing many real-world examples. May be useful also for computer science and engineering (CSE) students.
- **H.-J. Bungartz, S. Zimmer (Eds.).** *Numerische Simulation als interdisziplinäre Herausforderung* (in German). Springer 2002. - A collection of concise overviews on various subjects of computer and mathematical modeling. One can find non-trivial observations and witty remarks in this book.
- **J.M.T.Thompson.** *Instabilities and Catastrophes in Science and Engineering*. Wiley-Interscience, 1982. – A number of non-trivial examples, mostly from structural mechanics and civil engineering.
- **A.L. Garcia.** *Numerical Methods for Physics*. Prentice Hall, NJ 2000 (2nd Edition). – A useful book for practical applications
- **T.L. Saaty, J.Bram.** *Nonlinear Mathematics*. Dover, N.Y. 1981. – One of the first books on nonlinear models