Springer Series in
Computational
Mathematics

31

Ernst Hairer
Christian Lubich
Gerhard Wanner

# Geometric Numerical Integration

## Structure-Preserving Algorithms for Ordinary Differential Equations

Second Edition

With 146 Figures

Springer

Ernst Hairer
Gerhard Wanner
Section de Mathématiques
Université de Genève
2-4 rue du Lièvre, C.P. 64
CH-1211 Genève 4, Switzerland
email: Ernst.Hairer@math.unige.ch
Gerhard.Wanner@math.unige.ch

Christian Lubich
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen, Germany
email: Lubich@na.uni-tuebingen.de

Library of Congress Control Number: 2005938386

Mathematics Subject Classification (2000): 65Lxx, 65P10, 70Fxx, 34Cxx

ISSN 0179-3632
ISBN-10 3-540-30663-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-30663-4 Springer Berlin Heidelberg New York
ISBN-10 3-540-43003-2 1st Edition Springer Berlin Heidelberg New York



Springer is a part of Springer Science+Business Media
springer.com

Printed in The Netherlands



Typesetting: by the authors and TechBooks using a Springer LaTeX macro package

Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN: 11592242 46/TechBooks 5 4 3 2 1 0

# Preface to the First Edition

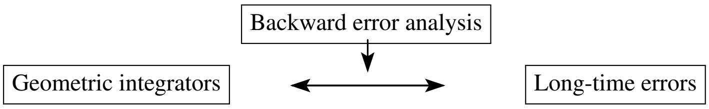
> They throw geometry out the door, and it comes back through the window.
> (H.G.Forder, Auckland 1973, reading new mathematics at the age of 84)

The subject of this book is numerical methods that preserve geometric properties of the flow of a differential equation: symplectic integrators for Hamiltonian systems, symmetric integrators for reversible systems, methods preserving first integrals and numerical methods on manifolds, including Lie group methods and integrators for constrained mechanical systems, and methods for problems with highly oscillatory solutions. Structure preservation – with its questions as to where, how, and what for – is the unifying theme.

In the last few decades, the theory of numerical methods for general (non-stiff and stiff) ordinary differential equations has reached a certain maturity, and excellent general-purpose codes, mainly based on Runge–Kutta methods or linear multistep methods, have become available. The motivation for developing structure-preserving algorithms for special classes of problems came independently from such different areas of research as astronomy, molecular dynamics, mechanics, theoretical physics, and numerical analysis as well as from other areas of both applied and pure mathematics. It turned out that the preservation of geometric properties of the flow not only produces an improved qualitative behaviour, but also allows for a more accurate long-time integration than with general-purpose methods.

An important shift of view-point came about by ceasing to concentrate on the numerical approximation of a single solution trajectory and instead to consider a numerical method as a *discrete dynamical system* which approximates the flow of the differential equation – and so the geometry of phase space comes back again through the window. This view allows a clear understanding of the preservation of invariants and of methods on manifolds, of symmetry and reversibility of methods, and of the symplecticity of methods and various generalizations. These subjects are presented in Chapters IV through VII of this book. Chapters I through III are of an introductory nature and present examples and numerical integrators together with important parts of the classical order theories and their recent extensions. Chapter VIII deals with questions of numerical implementations and numerical merits of the various methods.

It remains to explain the relationship between geometric properties of the numerical method and the favourable error propagation in long-time integrations. This




is done using the idea of *backward error analysis*, where the numerical one-step map is interpreted as (almost) the flow of a modified differential equation, which is constructed as an asymptotic series (Chapter IX). In this way, geometric properties of the numerical integrator translate into structure preservation on the level of the modified equations. Much insight and rigorous error estimates over long time intervals can then be obtained by combining this backward error analysis with KAM theory and related perturbation theories. This is explained in Chapters X through XII for Hamiltonian and reversible systems. The final Chapters XIII and XIV treat the numerical solution of differential equations with high-frequency oscillations and the long-time dynamics of multistep methods, respectively.

This book grew out of the lecture notes of a course given by Ernst Hairer at the University of Geneva during the academic year 1998/99. These lectures were directed at students in the third and fourth year. The reactions of students as well as of many colleagues, who obtained the notes from the Web, encouraged us to elaborate our ideas to produce the present monograph.

We want to thank all those who have helped and encouraged us to prepare this book. In particular, Martin Hairer for his valuable help in installing computers and his expertise in Latex and Postscript, Jeff Cash and Robert Chan for reading the whole text and correcting countless scientific obscurities and linguistic errors, Haruo Yoshida for making many valuable suggestions, Stéphane Cirilli for preparing the files for all the photographs, and Bernard Dudez, the irreplaceable director of the mathematics library in Geneva. We are also grateful to many friends and colleagues for reading parts of the manuscript and for valuable remarks and discussions, in particular to Assyr Abdulle, Melanie Beck, Sergio Blanes, John Butcher, Mari Paz Calvo, Begoña Cano, Philippe Chartier, David Cohen, Peter Deuflhard, Stig Faltinsen, Francesco Fassò, Martin Gander, Marlis Hochbruck, Bulent Karasözen, Wilhelm Kaup, Ben Leimkuhler, Pierre Leone, Frank Loose, Katina Lorenz, Robert McLachlan, Ander Murua, Alexander Ostermann, Truong Linh Pham, Sebastian Reich, Chus Sanz-Serna, Zaijiu Shang, Yifa Tang, Matt West, Will Wright.

We are especially grateful to Thanh-Ha Le Thi and Dr. Martin Peters from Springer-Verlag Heidelberg for assistance, in particular for their help in getting most of the original photographs from the Oberwolfach Archive and from Springer New York, and for clarifying doubts concerning the copyright.

Geneva and Tübingen, November 2001 The Authors

# Preface to the Second Edition

The fast development of the subject – and the fast development of the sales of the first edition of this book – has given the authors the opportunity to prepare this second edition. First of all we have corrected several misprints and minor errors which we have discovered or which have been kindly communicated to us by several readers and colleagues. We cordially thank all of them for their help and for their interest in our work. A major point of confusion has been revealed by Robert McLachlan in his book review in SIAM Reviews.

Besides many details, which have improved the presentation throughout the book, there are the following major additions and changes which make the book about 130 pages longer:

- a more prominent place of the Störmer–Verlet method in the exposition and the examples of the first chapter;
- a discussion of the Hénon–Heiles model as an example of a chaotic Hamiltonian system;
- a new Sect. IV.9 on geometric numerical linear algebra considering differential equations on Stiefel and Grassmann manifolds and dynamical low-rank approximations;
- a new improved composition method of order 10 in Sect. V.3;
- a characterization of B-series methods that conserve quadratic first integrals and a criterion for conjugate symplecticity in Sect. VI.8;
- the section on volume preservation taken from Chap. VII to Chap. VI;
- an extended and more coherent Chap. VII, renamed Non-Canonical Hamiltonian Systems, with more emphasis on the relationships between Hamiltonian systems on manifolds and Poisson systems;
- a completely reorganized and augmented Sect. VII.5 on the rigid body dynamics and Lie–Poisson systems;
- a new Sect. VII.6 on reduced Hamiltonian models of quantum dynamics and Poisson integrators for their numerical treatment;
- an improved step-size control for reversible methods in Sects. VIII.3.2 and IX.6;
- extension of Sect. IX.5 on modified equations of methods on manifolds to include constrained Hamiltonian systems and Lie–Poisson integrators;
- reorganization of Sects. IX.9 and IX.10; study of non-symplectic B-series methods that have a modified Hamiltonian, and counter-examples for symmetric methods showing linear growth in the energy error;

- a more precise discussion of integrable reversible systems with new examples in Chap. XI;
- extension of Chap. XIII on highly oscillatory problems to systems with several constant frequencies and to systems with non-constant mass matrix;
- a new Chap. XIV on oscillatory Hamiltonian systems with time- or solution-dependent high frequencies, emphasizing adiabatic transformations, adiabatic invariants, and adiabatic integrators;
- a completely rewritten Chap. XV with more emphasis on linear multistep methods for second order differential equations; a complete backward error analysis including parasitic modified differential equations; a study of the long-time stability and a rigorous explanation of the long-time near-conservation of energy and angular momentum.

Let us hope that this second revised edition will again meet good acceptance by our readers.

Geneva and Tübingen, October 2005 The Authors

# Table of Contents

# Chapter I.
# Examples and Numerical Experiments

This chapter introduces some interesting examples of differential equations and illustrates different types of qualitative behaviour of numerical methods. We deliberately consider only very simple numerical methods of orders 1 and 2 to emphasize the qualitative aspects of the experiments. The same effects (on a different scale) occur with more sophisticated higher-order integration schemes. The experiments presented here should serve as a motivation for the theoretical and practical investigations of later chapters. The reader is encouraged to repeat the experiments or to invent similar ones.

## I.1 First Problems and Methods

Numerical applications of the case of two dependent variables are not easily obtained. (A.J. Lotka 1925, p. 79)

Our first problems, the Lotka–Volterra model and the pendulum equation, are differential equations in two dimensions and show already many interesting geometric properties. Our first methods are various variants of the Euler method, the midpoint rule, and the Störmer–Verlet scheme.

### I.1.1 The Lotka–Volterra Model

We start with an equation from mathematical biology which models the growth of animal species. If a real variable $u(t)$ is to represent the number of individuals of a certain species at time $t$, the simplest assumption about its evolution is $du/dt = u\cdot\alpha$, where $\alpha$ is the reproduction rate. A constant $\alpha$ leads to exponential growth. In the case of more species living together, the reproduction rates will also depend on the population numbers of the *other* species. For example, for two species with $u(t)$ denoting the number of predators and $v(t)$ the number of prey, a plausible assumption is made by the *Lotka–Volterra model*

$$\begin{aligned} \dot u &= u(v-2) \\ \dot v &= v(1-u), \end{aligned} \tag{1.1}$$

where the dots on $u$ and $v$ stand for differentiation with respect to time. (We have chosen the constants 2 and 1 in (1.1) arbitrarily.) A.J. Lotka (1925, Chap. VIII) used

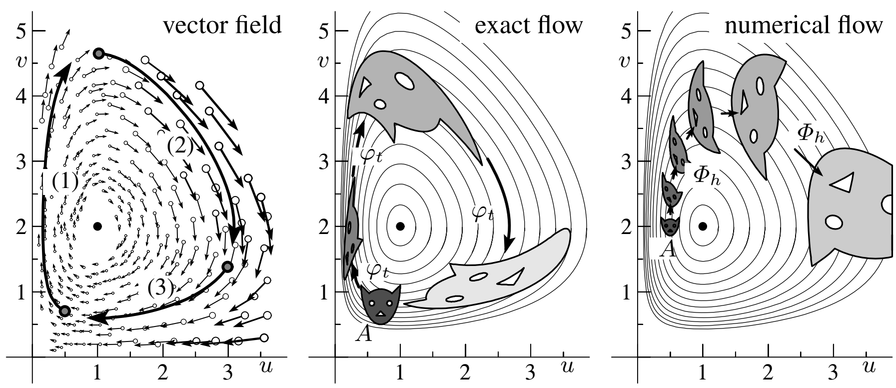



**Fig. 1.1.** Vector field, exact flow, and numerical flow for the Lotka–Volterra model (1.1)

this model to study parasitic invasion of insect species, and, with its help, V. Volterra (1927) explained curious fishing data from the upper Adriatic Sea following World War I.

Equations (1.1) constitute an autonomous system of differential equations. In general, we write such a system in the form

$$\dot{y} = f(y) \,. \tag{1.2}$$

Every $y$ represents a point in the *phase space*, in equation (1.1) above $y = (u, v)$ is in the phase plane $\mathbb{R}^2$. The vector-valued function $f(y)$ represents a *vector field* which, at any point of the phase space, prescribes the velocity (direction and speed) of the solution $y(t)$ that passes through that point (see the first picture of Fig. 1.1).

For the Lotka–Volterra model, we observe that the system cycles through three stages: (1) the prey population increases; (2) the predator population increases by feeding on the prey; (3) the predator population diminishes due to lack of food.

**Flow of the System.** A fundamental concept is the *flow* over time $t$. This is the mapping which, to any point $y_0$ in the phase space, associates the value $y(t)$ of the solution with initial value $y(0) = y_0$. This map, denoted by $\varphi_t$, is thus defined by

$$\varphi_t(y_0) = y(t) \quad \text{if} \quad y(0) = y_0. \tag{1.3}$$

The second picture of Fig. 1.1 shows the results of three iterations of $\varphi_t$ (with $t = 1.3$) for the Lotka–Volterra problem, for a set of initial values $y_0 = (u_0, v_0)$ forming an animal-shaped set $A$.[1]

**Invariants.** If we divide the two equations of (1.1) by each other, we obtain a single equation between the variables $u$ and $v$. After separation of variables we get

$$0 = \frac{1-u}{u}\,\dot{u} - \frac{v-2}{v}\,\dot{v} = \frac{d}{dt}\,I(u,v)$$

[1] This cat came to fame through Arnold (1963).

where

$$I(u, v) = \ln u - u + 2\ln v - v, \tag{1.4}$$

so that $I(u(t), v(t)) = \mathit{Const}$ for all $t$. We call the function $I$ an *invariant* of the system (1.1). Every solution of (1.1) thus lies on a level curve of (1.4). Some of these curves are drawn in the pictures of Fig. 1.1. Since the level curves are closed, all solutions of (1.1) are periodic.

## I.1.2 First Numerical Methods

**Explicit Euler Method.** The simplest of all numerical methods for the system (1.2) is the method formulated by Euler (1768),

$$y_{n+1} = y_n + hf(y_n). \tag{1.5}$$

It uses a constant step size $h$ to compute, one after the other, approximations $y_1$, $y_2$, $y_3$, ... to the values $y(h)$, $y(2h)$, $y(3h)$, ... of the solution starting from a given initial value $y(0) = y_0$. The method is called the *explicit Euler method*, because the approximation $y_{n+1}$ is computed using an explicit evaluation of $f$ at the already known value $y_n$. Such a formula represents a mapping

$$\Phi_h : y_n \mapsto y_{n+1},$$

which we call the *discrete* or *numerical flow*. Some iterations of the discrete flow for the Lotka–Volterra problem (1.1) (with $h = 0.5$) are represented in the third picture of Fig. 1.1.

**Implicit Euler Method.** The *implicit Euler method*

$$y_{n+1} = y_n + hf(y_{n+1}), \tag{1.6}$$

is known for its all-damping stability properties. In contrast to (1.5), the approximation $y_{n+1}$ is defined implicitly by (1.6), and the implementation requires the numerical solution of a nonlinear system of equations.

**Implicit Midpoint Rule.** Taking the mean of $y_n$ and $y_{n+1}$ in the argument of $f$, we get the *implicit midpoint rule*

$$y_{n+1} = y_n + hf\Big(\frac{y_n + y_{n+1}}{2}\Big). \tag{1.7}$$

It is a *symmetric* method, which means that the formula is left unaltered after exchanging $y_n \leftrightarrow y_{n+1}$ and $h \leftrightarrow -h$ (more on symmetric methods in Chap. V).

**Symplectic Euler Methods.** For *partitioned* systems

$$\begin{aligned} \dot{u} &= a(u, v) \\ \dot{v} &= b(u, v), \end{aligned} \tag{1.8}$$

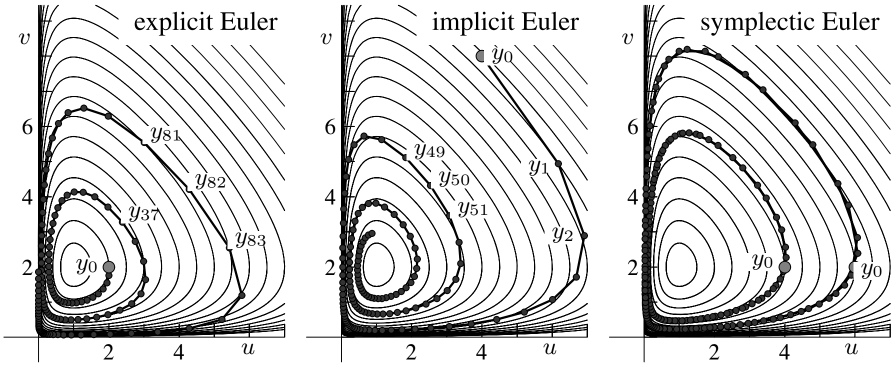



**Fig. 1.2.** Solutions of the Lotka–Volterra equations (1.1) (step sizes $h = 0.12$; initial values $(2, 2)$ for the explicit Euler method, $(4, 8)$ for the implicit Euler method, $(4, 2)$ and $(6, 2)$ for the symplectic Euler method)

such as the problem (1.1), we consider also *partitioned* Euler methods

$$\begin{aligned} u_{n+1} &= u_n + ha(u_n, v_{n+1}) \\ v_{n+1} &= v_n + hb(u_n, v_{n+1}), \end{aligned} \quad \text{or} \quad \begin{aligned} u_{n+1} &= u_n + ha(u_{n+1}, v_n) \\ v_{n+1} &= v_n + hb(u_{n+1}, v_n), \end{aligned} \tag{1.9}$$

which treat one variable by the implicit and the other variable by the explicit Euler method. In view of an important property of this method, discovered by de Vogelaere (1956) and to be discussed in Chap. VI, we call them *symplectic Euler methods*.

**Numerical Example for the Lotka–Volterra Problem.** Our first numerical experiment shows the behaviour of the various numerical methods applied to the Lotka–Volterra problem. In particular, we are interested in the preservation of the invariant $I$ over long times. Fig. 1.2 plots the numerical approximations of the first 125 steps with the above numerical methods applied to (1.1), all with constant step sizes. We observe that the explicit and implicit Euler methods show wrong qualitative behaviour. The numerical solution either spirals outwards or inwards. The symplectic Euler method (implicit in $u$ and explicit in $v$), however, gives a numerical solution that lies apparently on a closed curve as does the exact solution. Note that the curves of the numerical and exact solutions do not coincide.

## I.1.3 The Pendulum as a Hamiltonian System

A great deal of attention in this book will be addressed to Hamiltonian problems, and our next examples will be of this type. These problems are of the form

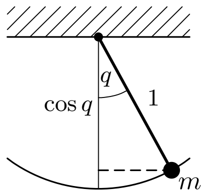
$$\dot{p} = -H_q(p, q), \qquad \dot{q} = H_p(p, q), \tag{1.10}$$

where the *Hamiltonian* $H(p_1, \ldots, p_d, q_1, \ldots q_d)$ represents the total energy; $q_i$ are the position coordinates and $p_i$ the momenta for $i = 1, \ldots, d$, with $d$ the number of degrees of freedom; $H_p$ and $H_q$ are the vectors of partial derivatives. One verifies easily by differentiation (see Sect. IV.1) that, along the solution curves of (1.10),

$$H\big(p(t), q(t)\big) = Const, \tag{1.11}$$

i.e., the Hamiltonian is an invariant or a *first integral*. More details about Hamiltonian systems and their derivation from Lagrangian mechanics will be given in Sect. VI.1.

**Pendulum.** The mathematical pendulum (mass $m = 1$, massless rod of length $\ell = 1$, gravitational acceleration $g = 1$) is a system with one degree of freedom having the Hamiltonian




$$H(p, q) = \frac{1}{2}\,p^2 - \cos q, \tag{1.12}$$

so that the equations of motion (1.10) become

$$\dot{p} = -\sin q, \qquad \dot{q} = p. \tag{1.13}$$

Since the vector field (1.13) is $2\pi$-periodic in $q$, it is natural to consider $q$ as a variable on the circle $S^1$. Hence, the phase space of points $(p, q)$ becomes the cylinder $\mathbb{R} \times S^1$. Fig. 1.3 shows some level curves of $H(p, q)$. By (1.11), the solution curves of the problem (1.13) lie on such level curves.

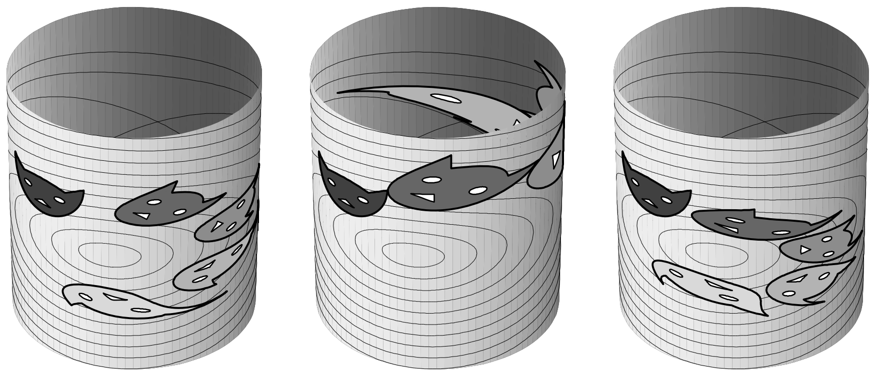



**Fig. 1.3.** Exact and numerical flow for the pendulum problem (1.13); step sizes $h = t = 1$

**Area Preservation.** Figure 1.3 (first picture) illustrates that the exact flow of a Hamiltonian system (1.10) is area preserving. This can be explained as follows: the derivative of the flow $\varphi_t$ with respect to initial values $(p, q)$,

$$\varphi_t'(p, q) = \frac{\partial\big(p(t), q(t)\big)}{\partial(p, q)},$$

satisfies the variational equation [2]

$$\dot{\varphi}_t'(p,q) = \begin{pmatrix} -H_{pq} & -H_{qq} \\ H_{pp} & H_{qp} \end{pmatrix} \varphi_t'(p,q) \ ,$$

where the second partial derivatives of $H$ are evaluated at $\varphi_t(p,q)$. In the case of one degree of freedom ($d=1$), a simple computation shows that

$$\frac{d}{dt} \det \varphi_t'(p,q) = \frac{d}{dt} \Big( \frac{\partial p(t)}{\partial p} \frac{\partial q(t)}{\partial q} - \frac{\partial p(t)}{\partial q} \frac{\partial q(t)}{\partial p} \Big) = \ldots = 0.$$

Since $\varphi_0$ is the identity, this implies $\det \varphi_t'(p,q) = 1$ for all $t$, which means that the flow $\varphi_t(p,q)$ is an *area-preserving* mapping.

The last two pictures of Fig. 1.3 show numerical flows. The explicit Euler method is clearly seen not to preserve area but the symplectic Euler method is (this will be proved in Sect. VI.3). One of the aims of 'geometric integration' is the study of numerical integrators that preserve such types of qualitative behaviour of the exact flow.

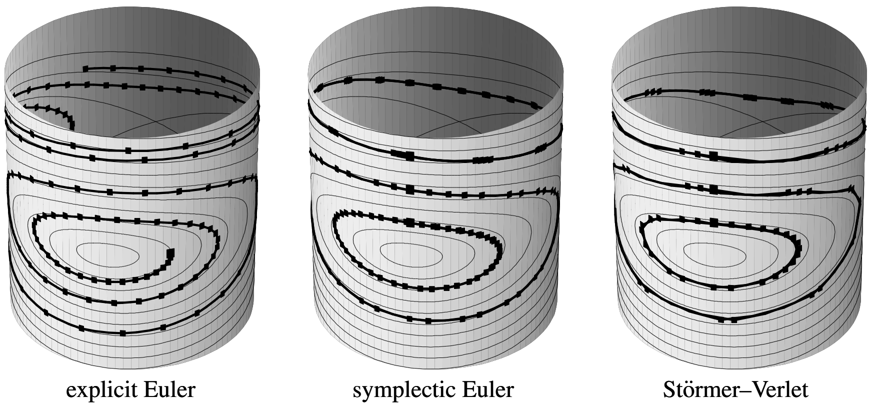



**Fig. 1.4.** Solutions of the pendulum problem (1.13); explicit Euler with step size $h=0.2$, initial value $(p_0,q_0)=(0,0.5)$; symplectic Euler with $h=0.3$ and initial values $q_0=0$, $p_0=0.7,1.4,2.1$; Störmer–Verlet with $h=0.6$

**Numerical Experiment.** We apply the above numerical methods to the pendulum equations (see Fig. 1.4). Similar to the computations for the Lotka–Volterra equations, we observe that the numerical solutions of the explicit Euler and of the implicit Euler method (not drawn in Fig. 1.4) spiral either outwards or inwards. The symplectic Euler method shows the correct qualitative behaviour, but destroys the left-right symmetry of the problem. The Störmer–Verlet scheme, which we discuss next, works perfectly even with doubled step size.

[2] As is common in the study of mechanical problems, we use *dots* for denoting time-derivatives, and we use *primes* for denoting derivatives with respect to other variables.

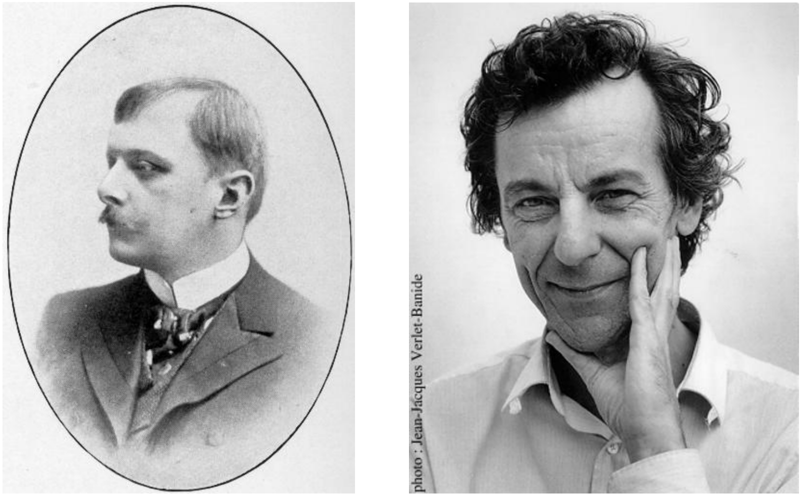



**Fig. 1.5.** Carl Störmer (left picture), born: 3 September 1874 in Skien (Norway), died: 13 August 1957.
Loup Verlet (right picture), born: 24 May 1931 in Paris

## I.1.4 The Störmer–Verlet Scheme

The above equations (1.13) for the pendulum are of the form

$$\begin{aligned} \dot p &= f(q) \\ \dot q &= p \end{aligned} \qquad \text{or} \qquad \ddot q = f(q) \tag{1.14}$$

which is the important special case of a second order differential equation. The most natural discretization of (1.14) is

$$q_{n+1} - 2q_n + q_{n-1} = h^2 f(q_n), \tag{1.15}$$

which is just obtained by replacing the second derivative in (1.14) by the central second-order difference quotient. This basic method, or its equivalent formulation given below, is called the *Störmer method* in astronomy, the *Verlet method*[3] in molecular dynamics, the *leap-frog method* in the context of partial differential equations, and it has further names in other areas (see Hairer, Lubich & Wanner (2003), p. 402). C. Störmer (1907) used higher-order variants for numerical computations concerning the aurora borealis. L. Verlet (1967) proposed this method for computations in molecular dynamics, where it has become by far the most widely used integration scheme.

Geometrically, the Störmer–Verlet method can be seen as produced by parabolas, which in the points $t_n$ possess the right second derivative $f(q_n)$ (see Fig. 1.6

[3] Irony of fate: Professor Loup Verlet, who later became interested in the history of science, discovered precisely "his" method in Newton's *Principia* (Book I, figure for Theorem I, see Sect. I.2.1 below).

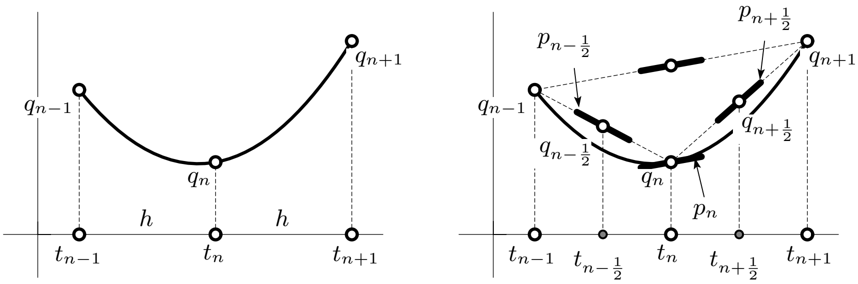



**Fig. 1.6.** Illustration for the Störmer–Verlet method

to the left). But we can also think of polygons, which possess the right slope in the midpoints (Fig. 1.6 to the right).

Approximations to the derivative $p = \dot{q}$ are simply obtained by

$$p_n = \frac{q_{n+1} - q_{n-1}}{2h} \qquad \text{and} \qquad p_{n+1/2} = \frac{q_{n+1} - q_n}{h}. \tag{1.16}$$

**One-Step Formulation.** The Störmer–Verlet method admits a one-step formulation which is useful for actual computations. The value $q_n$ together with the slope $p_n$ and the second derivative $f(q_n)$, all at $t_n$, uniquely determine the parabola and hence also the approximation $(p_{n+1}, q_{n+1})$ at $t_{n+1}$. Writing (1.15) as $p_{n+1/2} - p_{n-1/2} = hf(q_n)$ and using $p_{n+1/2} + p_{n-1/2} = 2p_n$, we get by elimination of either $p_{n+1/2}$ or $p_{n-1/2}$ the formulae

$$\begin{aligned} p_{n+1/2} &= p_n + \frac{h}{2}\, f(q_n) \\ q_{n+1} &= q_n + hp_{n+1/2} \\ p_{n+1} &= p_{n+1/2} + \frac{h}{2}\, f(q_{n+1}) \end{aligned} \tag{1.17}$$

which is an explicit one-step method $\Phi_h : (q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$ for the corresponding first order system of (1.14). If one is not interested in the values $p_n$ of the derivative, the first and third equations in (1.17) can be replaced by

$$p_{n+1/2} = p_{n-1/2} + h\, f(q_n).$$

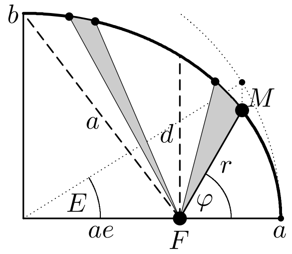
## I.2 The Kepler Problem and the Outer Solar System

I awoke as if from sleep, a new light broke on me. (J. Kepler; quoted from J.L.E. Dreyer, *A history of astronomy*, 1906, Dover 1953, p. 391)

One of the great achievements in the history of science was the discovery of the laws of J. Kepler (1609), based on many precise measurements of the positions of Mars by Tycho Brahe and himself. The planets move in *elliptic orbits* with the sun at one of the foci (Kepler's first law)

$$r = \frac{d}{1 + e\cos\varphi} = a - ae\cos E, \tag{2.1}$$

(where $a$ = great axis, $e$ = eccentricity, $b = a\sqrt{1-e^2}$, $d = b\sqrt{1-e^2} = a(1-e^2)$, $E$ = eccentric anomaly, $\varphi$ = true anomaly).




Newton (*Principia* 1687) then *explained* this motion by his general law of gravitational attraction (proportional to $1/r^2$) and the relation between forces and acceleration (the "Lex II" of the *Principia*). This then opened the way for treating arbitrary celestial motions by solving differential equations.

**Two-Body Problem.** For computing the motion of two bodies which attract each other, we choose one of the bodies as the centre of our coordinate system; the motion will then stay in a plane (Exercise 3) and we can use two-dimensional coordinates $q = (q_1, q_2)$ for the position of the second body. Newton's laws, with a suitable normalization, then yield the following differential equations

$$\ddot q_1 = -\frac{q_1}{(q_1^2+q_2^2)^{3/2}}, \qquad \ddot q_2 = -\frac{q_2}{(q_1^2+q_2^2)^{3/2}}. \tag{2.2}$$

This is equivalent to a Hamiltonian system with the Hamiltonian

$$H(p_1,p_2,q_1,q_2) = \frac{1}{2}\left(p_1^2+p_2^2\right) - \frac{1}{\sqrt{q_1^2+q_2^2}}, \qquad p_i = \dot q_i. \tag{2.3}$$

## I.2.1 Angular Momentum and Kepler's Second Law

The system has not only the total energy $H(p,q)$ as a first integral, but also the angular momentum

$$L(p_1,p_2,q_1,q_2) = q_1p_2 - q_2p_1. \tag{2.4}$$

This can be checked by differentiation and is nothing other than *Kepler's second law*, which says that the ray $FM$ sweeps equal areas in equal times (see the little picture at the beginning of Sect. I.2).

A beautiful *geometric* justification of this law is due to I. Newton[4] (*Principia* (1687), Book I, figure for Theorem I). The idea is to apply the Störmer–Verlet scheme (1.15) to the equations (2.2) (see Fig. 2.1). By hypothesis, the diagonal of the parallelogram $q_{n-1}q_nq_{n+1}$, which is $(q_{n+1}-q_n)-(q_n-q_{n-1}) = q_{n+1}-2q_n+q_{n-1} = Const\cdot f(q_n)$, points towards the sun $S$. Therefore, the altitudes of the triangles $q_{n-1}q_nS$ and $q_{n+1}q_nS$ are equal. Since they have the common base $q_nS$, they also have equal areas. Hence

$$\det(q_{n-1}, q_n - q_{n-1}) = \det(q_n, q_{n+1}-q_n)$$

and by passing to the limit $h \to 0$ we see that $\det(q,p) = Const$. This is (2.4).

[4] We are grateful to a private communication of L. Verlet for this reference

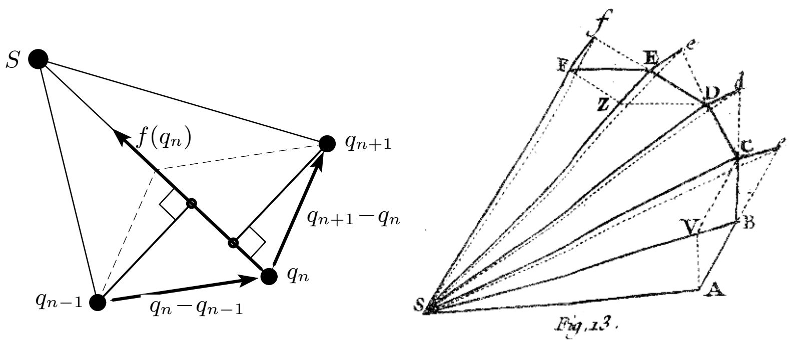



**Fig. 2.1.** Proof of Kepler's Second Law (left); facsimile from Newton's *Principia* (right)

We have not only an elegant proof for this invariant, but we also see that *the Störmer–Verlet scheme preserves this invariant for every* $h > 0$.

## I.2.2 Exact Integration of the Kepler Problem

> Pour voir présentement que cette courbe $ABC$ ... est toûjours une Section Conique, ainsi que Mr. Newton l'a supposé, *pag. 55. Coroll.I.* sans le démontrer; il y faut bien plus d'adresse: (Joh. Bernoulli 1710, p. 475)

It is now interesting, inversely to the procedure of Newton, to prove that *any* solution of (2.2) follows either an elliptic, parabolic or hyperbolic arc and to describe the solutions analytically. This was first done by Joh. Bernoulli (1710, full of sarcasm against Newton), and by Newton (1713, second edition of the *Principia*, without mentioning a word about Bernoulli).

By (2.3) and (2.4), every solution of (2.2) satisfies the two relations

$$\frac{1}{2}\left(\dot q_1^2+\dot q_2^2\right)-\frac{1}{\sqrt{q_1^2+q_2^2}}=H_0, \qquad q_1\dot q_2-q_2\dot q_1=L_0, \tag{2.5}$$

where the constants $H_0$ and $L_0$ are determined by the initial values. Using polar coordinates $q_1 = r\cos\varphi$, $q_2 = r\sin\varphi$, this system becomes

$$\frac{1}{2}\left(\dot r^2+r^2\dot\varphi^2\right)-\frac{1}{r}=H_0, \qquad r^2\dot\varphi=L_0. \tag{2.6}$$

For its solution we consider $r$ as a function of $\varphi$ and write $\dot r = \frac{dr}{d\varphi}\cdot\dot\varphi$. The elimination of $\dot\varphi$ in (2.6) then yields

$$\frac{1}{2}\left(\Big(\frac{dr}{d\varphi}\Big)^2+r^2\right)\frac{L_0^2}{r^4}-\frac{1}{r}=H_0.$$

In this equation we use the substitution $r = 1/u$, $dr = -du/u^2$, which gives (with $' = d/d\varphi$)

$$\frac{1}{2}(u'^2+u^2)-\frac{u}{L_0^2}-\frac{H_0}{L_0^2}=0. \tag{2.7}$$

This is a "Hamiltonian" for the system

$$u'' + u = \frac{1}{d} \quad \text{i.e.,} \quad u = \frac{1}{d} + c_1 \cos\varphi + c_2 \sin\varphi = \frac{1 + e\cos(\varphi - \varphi^*)}{d} \tag{2.8}$$

where $d = L_0^2$ and the constant $e$ becomes, from (2.7),

$$e^2 = 1 + 2H_0 L_0^2 \tag{2.9}$$

(by Exercise 7, the expression $1+2H_0L_0^2$ is non-negative). This is precisely formula (2.1). The angle $\varphi^*$ is determined by the initial values $r_0$ and $\varphi_0$. Equation (2.1) represents an elliptic orbit with eccentricity $e$ for $H_0 < 0$ (see Fig. 2.2, dotted line), a parabola for $H_0 = 0$, and a hyperbola for $H_0 > 0$.

Finally, we must determine the variables $r$ and $\varphi$ as functions of $t$. With the relation (2.8) and $r = 1/u$, the second equation of (2.6) gives

$$\frac{d^2}{\big(1 + e\cos(\varphi - \varphi^*)\big)^2}\, d\varphi = L_0\, dt \tag{2.10}$$

which, after an elementary, but not easy, integration, represents an implicit equation for $\varphi(t)$.

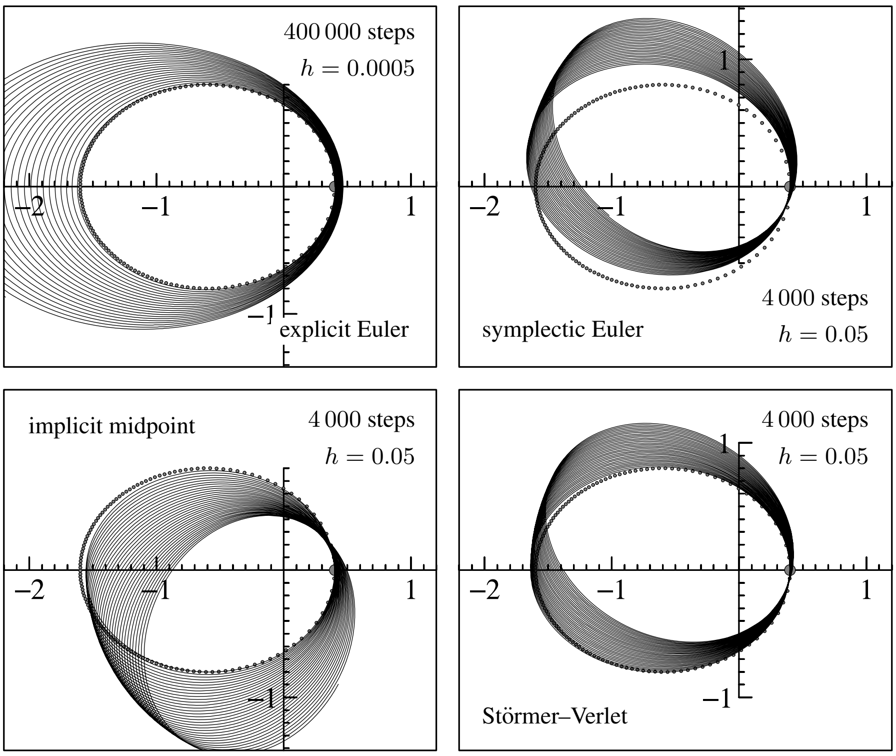



**Fig. 2.2.** Numerical solutions of the Kepler problem (eccentricity $e = 0.6$; in dots: exact solution)

### I.2.3 Numerical Integration of the Kepler Problem

For the problem (2.2) we choose, with $0 \le e < 1$, the initial values

$$q_1(0) = 1 - e, \quad q_2(0) = 0, \quad \dot{q}_1(0) = 0, \quad \dot{q}_2(0) = \sqrt{\frac{1+e}{1-e}}. \tag{2.11}$$

This implies that $H_0 = -1/2$, $L_0 = \sqrt{1-e^2}$, $d = 1 - e^2$ and $\varphi^* = 0$. The period of the solution is $2\pi$ (Exercise 5). Fig. 2.2 shows some numerical solutions for the eccentricity $e = 0.6$ compared to the exact solution. After our previous experience, it is no longer a surprise that the explicit Euler method spirals outwards and gives a completely wrong answer. For the other methods we take a step size 100 times larger in order to "see something". We see that the nonsymmetric symplectic Euler method distorts the ellipse, and that all methods exhibit a *precession* effect, clockwise for Störmer–Verlet and symplectic Euler, anti-clockwise for the implicit midpoint rule. The same behaviour occurs for the exact solution of *perturbed* Kepler problems (Exercise 12) and has occupied astronomers for centuries.

Our next experiment (Fig. 2.3) studies the conservation of invariants and the global error. The main observation is that the error in the energy grows linearly for the explicit Euler method, and it remains bounded and small (no secular terms) for the symplectic Euler method. The global error, measured in the Euclidean norm, shows a quadratic growth for the explicit Euler compared to a linear growth for the symplectic Euler. As indicated in Table 2.1 the implicit midpoint rule and the Störmer–Verlet scheme behave similar to the symplectic Euler, but have a smaller

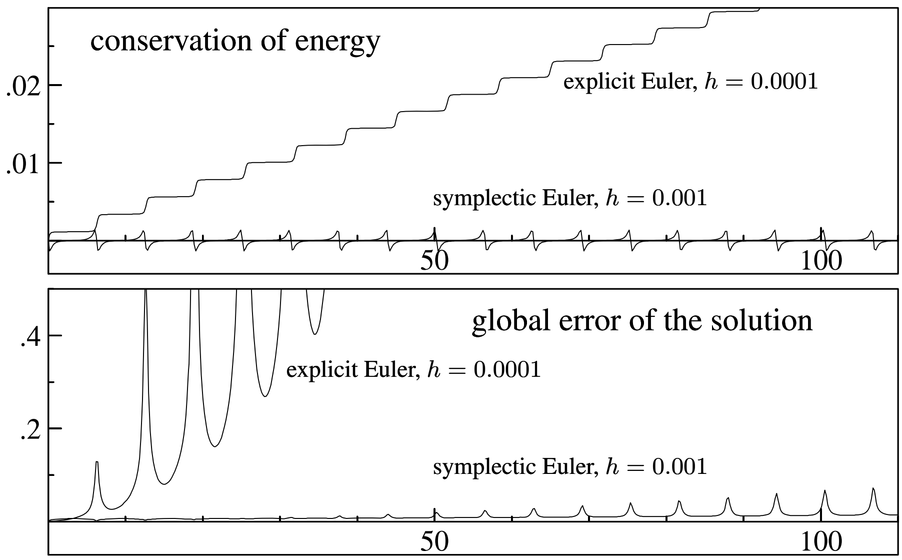



**Fig. 2.3.** Energy conservation and global error for the Kepler problem

**Table 2.1.** Qualitative long-time behaviour for the Kepler problem; $t$ is time, $h$ the step size

| method | error in $H$ | error in $L$ | global error |
|---|---|---|---|
| explicit Euler | $\mathcal{O}(th)$ | $\mathcal{O}(th)$ | $\mathcal{O}(t^2h)$ |
| symplectic Euler | $\mathcal{O}(h)$ | 0 | $\mathcal{O}(th)$ |
| implicit midpoint | $\mathcal{O}(h^2)$ | 0 | $\mathcal{O}(th^2)$ |
| Störmer–Verlet | $\mathcal{O}(h^2)$ | 0 | $\mathcal{O}(th^2)$ |

error due to their higher order. We remark that the angular momentum $L(p,q)$ is exactly conserved by the symplectic Euler, the Störmer–Verlet, and the implicit midpoint rule.

## I.2.4 The Outer Solar System

> The evolution of the entire planetary system has been numerically integrated for a time span of nearly 100 million years[5]. This calculation confirms that the evolution of the solar system as a whole is chaotic, . . .
>
> (G.J. Sussman & J. Wisdom 1992)

We next apply our methods to the system which describes the motion of the five outer planets relative to the sun. This system has been studied extensively by astronomers. The problem is a Hamiltonian system (1.10) ($N$-body problem) with

$$H(p,q) = \frac{1}{2}\sum_{i=0}^{5}\frac{1}{m_i}\,p_i^T p_i - G\sum_{i=1}^{5}\sum_{j=0}^{i-1}\frac{m_i\,m_j}{\|q_i - q_j\|}. \tag{2.12}$$

Here $p$ and $q$ are the supervectors composed by the vectors $p_i, q_i \in \mathbb{R}^3$ (momenta and positions), respectively. The chosen units are: masses relative to the sun, so that the sun has mass 1. We have taken

$$m_0 = 1.00000597682$$

to take account of the inner planets. Distances are in astronomical units (1 [A.U.] = 149 597 870 [km]), times in earth days, and the gravitational constant is

$$G = 2.95912208286 \cdot 10^{-4}.$$

The initial values for the sun are taken as $q_0(0) = (0,0,0)^T$ and $\dot q_0(0) = (0,0,0)^T$. All other data (masses of the planets and the initial positions and initial velocities) are given in Table 2.2. The initial data is taken from “Ahnerts Kalender für Sternfreunde 1994”, Johann Ambrosius Barth Verlag 1993, and they correspond to September 5, 1994 at 0h00.[6]

[5] 100 million years is not much in astronomical time scales; it just goes back to “Jurassic Park”.

[6] We thank Alexander Ostermann, who provided us with this data.

**Table 2.2.** Data for the outer solar system

| planet | mass | initial position | initial velocity |
|---|---|---|---|
| Jupiter | $m_1 = 0.000954786104043$ | $-3.5023653$<br>$-3.8169847$<br>$-1.5507963$ | $0.00565429$<br>$-0.00412490$<br>$-0.00190589$ |
| Saturn | $m_2 = 0.000285583733151$ | $9.0755314$<br>$-3.0458353$<br>$-1.6483708$ | $0.00168318$<br>$0.00483525$<br>$0.00192462$ |
| Uranus | $m_3 = 0.0000437273164546$ | $8.3101420$<br>$-16.2901086$<br>$-7.2521278$ | $0.00354178$<br>$0.00137102$<br>$0.00055029$ |
| Neptune | $m_4 = 0.0000517759138449$ | $11.4707666$<br>$-25.7294829$<br>$-10.8169456$ | $0.00288930$<br>$0.00114527$<br>$0.00039677$ |
| Pluto | $m_5 = 1/(1.3 \cdot 10^8)$ | $-15.5387357$<br>$-25.2225594$<br>$-3.1902382$ | $0.00276725$<br>$-0.00170702$<br>$-0.00136504$ |

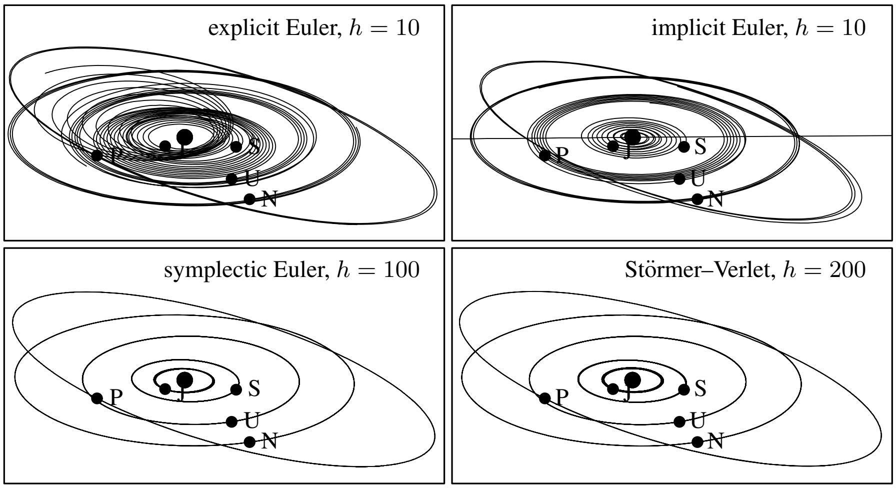



**Fig. 2.4.** Solutions of the outer solar system

To this system we apply the explicit and implicit Euler methods with step size $h = 10$, the symplectic Euler and the Störmer–Verlet method with much larger step sizes $h = 100$ and $h = 200$, repectively, all over a time period of 200 000 days. The numerical solution (see Fig. 2.4) behaves similarly to that for the Kepler problem. With the explicit Euler method the planets have increasing energy, they spiral outwards, Jupiter approaches Saturn which leaves the plane of the two-body motion. With the implicit Euler method the planets (first Jupiter and then Saturn) fall into the sun and are thrown far away. Both the symplectic Euler method and the Störmer–Verlet scheme show the correct behaviour. An integration over a much longer time of say several million years does not deteriorate this behaviour. Let us remark that Sussman & Wisdom (1992) have integrated the outer solar system with special geometric integrators.

## I.3 The Hénon–Heiles Model

> . . . because: (1) it is analytically simple; this makes the computation of the trajectories easy; (2) at the same time, it is sufficiently complicated to give trajectories which are far from trivial . (Hénon & Heiles 1964)

The Hénon–Heiles model was created for describing stellar motion, followed for a very long time, inside the gravitational potential $U_0(r,z)$ of a galaxy with cylindrical symmetry (Hénon & Heiles 1964). Extensive numerical experimentations should help to answer the question, if there exists, besides the known invariants $H$ and $L$, a *third* invariant. Despite endless tentatives of analytical calculations during many decades, such a formula had not been found.

After a reduction of the dimension, a Hamiltonian in two degrees of freedom of the form

$$H(p,q) = \frac{1}{2}(p_1^2 + p_2^2) + U(q) \tag{3.1}$$

is obtained and the question is, if such an equation has a *second* invariant. Here, Hénon and Heiles put aside the astronomical origin of the problem and choose

$$U(q) = \frac{1}{2}(q_1^2 + q_2^2) + q_1^2 q_2 - \frac{1}{3}q_2^3 \tag{3.2}$$

(see citation). The potential $U$ is represented in Fig. 3.1. When $U$ approaches $\frac{1}{6}$, the level curves of $U$ tend to an equilateral triangle, whose vertices are saddle points of $U$. The corresponding system

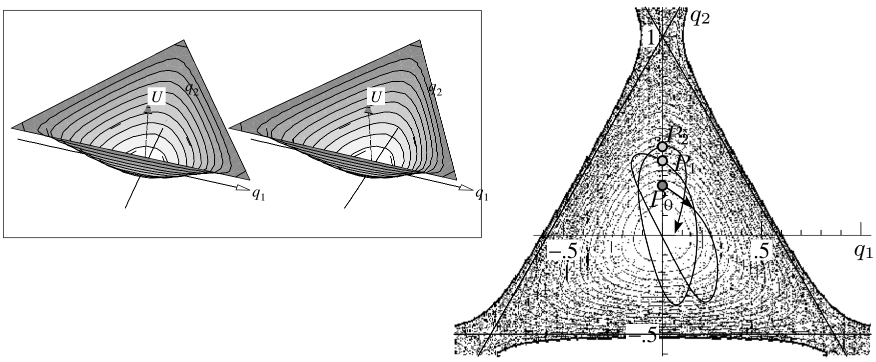



**Fig. 3.1.** Potential of the Hénon–Heiles Model and a solution

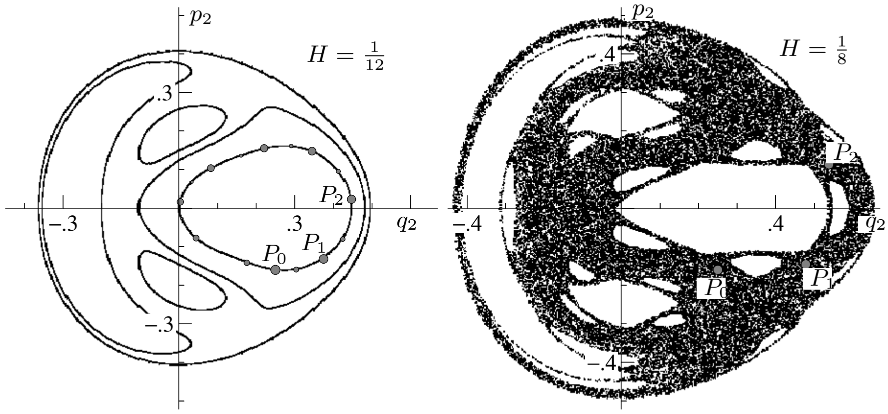



**Fig. 3.2.** Poincaré cuts for $q_1 = 0$, $p_1 > 0$ of the Hénon–Heiles Model for $H = \frac{1}{12}$ (6 orbits, left) and $H = \frac{1}{8}$ (1 orbit, right)

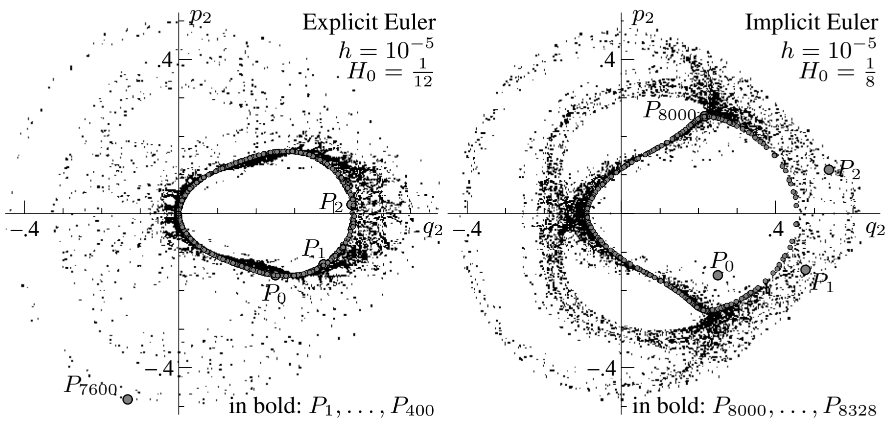



**Fig. 3.3.** Poincaré cuts for numerical methods, one orbit each; explicit Euler (left), implicit Euler (right). Same initial data as in Fig. 3.2

$$\ddot{q}_1 = -q_1 - 2q_1q_2\,, \qquad \ddot{q}_2 = -q_2 - q_1^2 + q_2^2 \tag{3.3}$$

has solutions with nontrivial properties. For given initial values with $H(p_0, q_0) < \frac{1}{6}$ and $q_0$ inside the triangle $U \leq \frac{1}{6}$, the solution stays there and moves somehow like a mass point gliding on this surface (see Fig. 3.1, right).

**Poincaré Cuts.** We fix first the energy $H_0$ and put $q_{10} = 0$. Then for any point $P_0 = (q_{20}, p_{20})$, we obtain $p_{10}$ from (3.1) as $p_{10} = \sqrt{2H_0 - 2U_0 - p_{20}^2}$, where we choose the positive root. We then follow the solution until it hits again the surface $q_1 = 0$ in the positive direction $p_1 > 0$ and obtain a point $P_1 = (q_{21}, p_{21})$; in the same way we compute $P_2 = (q_{22}, p_{22})$, etc. For the same initial values as in Fig. 3.1 and with $H_0 = \frac{1}{12}$, the solution for $0 \leq t \leq 300\,000$ gives $46\,865$ Poincaré cuts which are all displayed in Fig. 3.2 (left). They seem to lie exactly on a curve, as do the orbits for 5 other choices of initial values. This picture thus shows “convincing

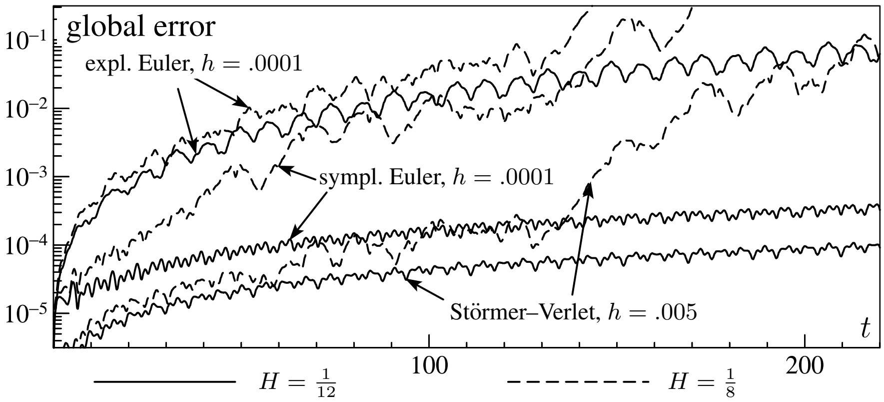



**Fig. 3.4.** Global error of numerical methods for nearly quasiperiodic and for chaotic solutions; same initial data as in Fig. 3.2

evidence" for the existence of a second invariant, for which Gustavson (1966) has derived a formal expansion, whose first terms represent perfectly these curves.

"But here comes the surprise" (Hénon–Heiles, p. 76): Fig. 3.2 shows to the right the same picture in the $(q_2, p_2)$ plane for a somewhat higher Energy $H = \frac{1}{8}$. The motion turns completely to chaos and all hope for a second invariant disappears. Actually, Gustavson's series does not converge.

**Numerical Experiments.** We now apply numerical methods, the *explicit* Euler method to the low energy initial values $H = \frac{1}{12}$ (Fig. 3.3, left), and the *implicit* Euler method to the high energy initial values (Fig. 3.3, right), both methods with a very small step size $h = 10^{-5}$. As we already expect from our previous experiences, the explicit Euler method tends to *increase* the energy and turns order into chaos, while the implicit Euler method tends to *decrease* it and turns chaos into order. The Störmer–Verlet method (not shown) behaves as the exact solution even for step sizes as large as $h = 10^{-1}$.

In our next experiment we study the *global error* (see Fig. 3.4), once for the case of the nearly quasiperiodic orbit ($H = \frac{1}{12}$) and once for the chaotic one ($H = \frac{1}{8}$), both for the explicit Euler, the symplectic Euler, and the Störmer–Verlet scheme. It may come as a surprise, that only in the first case we have the same behaviour (linear or quadratic growth) as in Fig. 2.3 for the Kepler problem. In the second case ($H = \frac{1}{8}$) the global error grows exponentially for all methods, and the explicit Euler method is worst.

**Study of a Mapping.** The passage from a point $P_i$ to the next one $P_{i+1}$ (as explained for the left picture of Fig. 3.2) can be considered as a *mapping* $\Phi : P_i \mapsto P_{i+1}$ and the sequence of points $P_0, P_1, P_2, \ldots$ are just the iterates of this mapping. This mapping is represented for the two energy levels $H = \frac{1}{12}$ and $H = \frac{1}{8}$ in Fig. 3.5 and its study allows to better understand the behaviour of the orbits. We see no significant difference between the two cases, simply for larger $H$ the deformations are more violent and correspond to larger eigenvalues of the Jacobian of $\Phi$. In

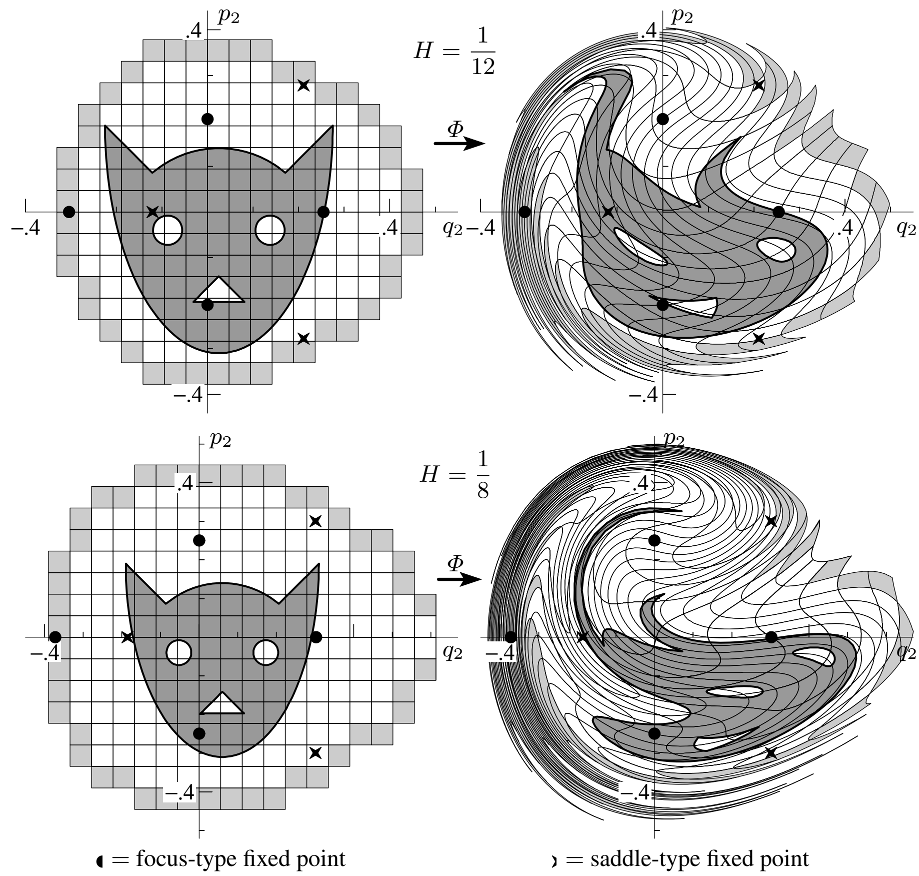



**Fig. 3.5.** The Poincaré map $\Phi : P_0 \to P_1$ for the Hénon–Heiles Model

both cases we have seven fixed points, which correspond to periodic solutions of the system (3.3). Four of them are stable and lie inside the white islands of Fig. 3.2.

# I.4 Molecular Dynamics

> We do not need exact classical trajectories to do this, but must lay great emphasis on energy conservation as being of primary importance for this reason. (M.P. Allen & D.J. Tildesley 1987)

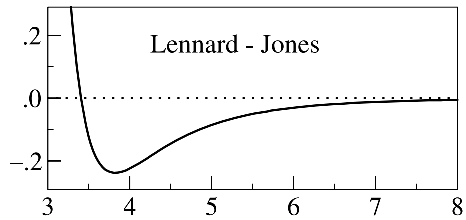
Molecular dynamics requires the solution of Hamiltonian systems (1.10), where the total energy is given by

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{N} \frac{1}{m_i}\, p_i^T p_i + \sum_{i=2}^{N}\sum_{j=1}^{i-1} V_{ij}\Big(\|q_i - q_j\|\Big), \tag{4.1}$$

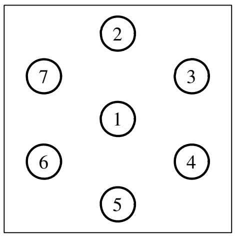
and $V_{ij}(r)$ are given potential functions. Here, $q_i$ and $p_i$ denote the positions and momenta of atoms and $m_i$ is the atomic mass of the $i$th atom. We remark that the outer solar system (2.12) is such an $N$-body system with $V_{ij}(r) = -Gm_i m_j / r$. In molecular dynamics the Lennard–Jones potential

$$V_{ij}(r) = 4\varepsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right) \tag{4.2}$$

is very popular ($\varepsilon_{ij}$ and $\sigma_{ij}$ are suitable constants depending on the atoms). This potential has an absolute minimum at distance $r = \sigma_{ij}\sqrt[6]{2}$. The force due to this potential strongly repels the atoms when they are closer than this value, and they attract each other when they are farther away.




**Numerical Experiments with a Frozen Argon Crystal.** As in Biesiadecki & Skeel (1993) we consider the interaction of seven argon atoms in a plane, where six of them are arranged symmetrically around a centre atom. As a mathematical model we take the Hamiltonian (4.1) with $N = 7$, $m_i = m = 66.34 \cdot 10^{-27}$ [kg],




$$\varepsilon_{ij} = \varepsilon = 119.8\, k_B \text{ [J]}, \qquad \sigma_{ij} = \sigma = 0.341 \text{ [nm]},$$

where $k_B = 1.380658 \cdot 10^{-23}$ [J/K] is Boltzmann's constant (see Allen & Tildesley (1987), page 21). As units for our calculations we take masses in [kg], distances in nanometers (1 [nm] $= 10^{-9}$ [m]), and times in nanoseconds (1 [nsec] $= 10^{-9}$ [sec]). Initial positions (in [nm]) and initial velocities (in [nm/nsec]) are given in Table 4.1. They are chosen such that neighbouring atoms have a distance that is close to the one with lowest potential energy, and such that the total momentum is zero and therefore the centre of gravity does not move. The energy at the initial position is $H(p_0, q_0) \approx -1260.2\, k_B$ [J].

For computations in molecular dynamics one is usually not interested in the trajectories of the atoms, but one aims at macroscopic quantities such as temperature, pressure, internal energy, etc. Here we consider the total energy, given by the Hamiltonian, and the temperature which can be calculated from the formula (see Allen &

**Table 4.1.** Initial values for the simulation of a frozen argon crystal

| atom | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| position | 0.00 | 0.02 | 0.34 | 0.36 | −0.02 | −0.35 | −0.31 |
| | 0.00 | 0.39 | 0.17 | −0.21 | −0.40 | −0.16 | 0.21 |
| velocity | −30 | 50 | −70 | 90 | 80 | −40 | −80 |
| | −20 | −90 | −60 | 40 | 90 | 100 | −60 |

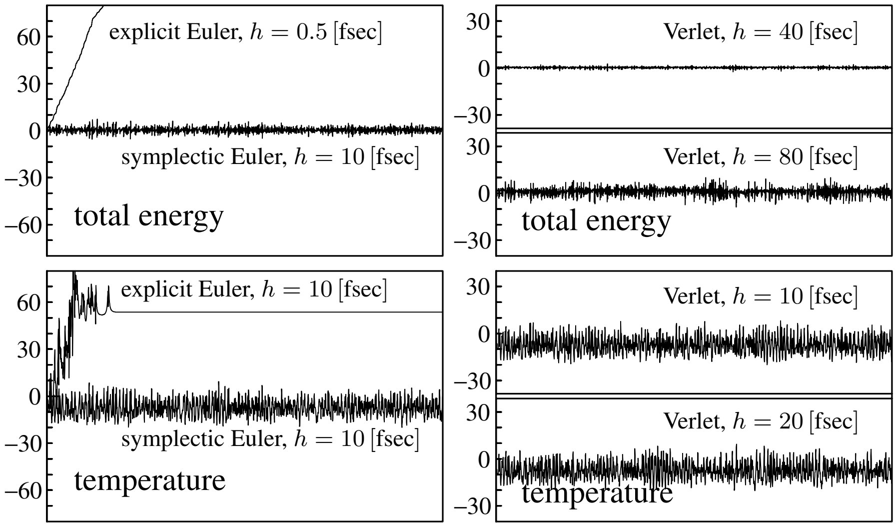



**Fig. 4.1.** Computed total energy and temperature of the argon crystal

Tildesley (1987), page 46)

$$T = \frac{1}{2Nk_B}\sum_{i=1}^{N} m_i\|\dot q_i\|^2. \tag{4.3}$$

We apply the explicit and symplectic Euler methods and also the Verlet method to this problem. Observe that for a Hamiltonian such as (4.1) all three methods are explicit, and all of them need only one force evaluation per integration step. In Fig. 4.1 we present the numerical results of our experiments. The integrations are done over an interval of length $0.2$ [nsec]. The step sizes are indicated in femtoseconds (1 [fsec] $= 10^{-6}$ [nsec]).

The two upper pictures show the values $\big(H(p_n,q_n)-H(p_0,q_0)\big)/k_B$ as a function of time $t_n = nh$. For the exact solution, this value is precisely zero for all times. Similar to earlier experiments we see that the symplectic Euler method is qualitatively correct, whereas the numerical solution of the explicit Euler method, although computed with a much smaller step size, is completely useless (see the citation at the beginning of this section). The Verlet method is qualitatively correct and gives much more accurate results than the symplectic Euler method (we shall see later that the Verlet method is of order 2). The two computations with the Verlet method show that the energy error decreases by a factor of $4$ if the step size is reduced by a factor of 2 (second order convergence).

The two lower pictures of Fig. 4.1 show the numerical values of the temperature difference $T - T_0$ with $T$ given by (4.3) and $T_0 \approx 22.72$ [K] (initial temperature). In contrast to the total energy, this is not an exact invariant, but for our problem it fluctuates around a constant value. The explicit Euler method gives wrong results, but the symplectic Euler and the Verlet methods show the desired behaviour. This time a reduction of the step size does not reduce the amplitude of the oscillations, which indicates that the fluctuation of the exact temperature is of the same size.

## I.5 Highly Oscillatory Problems

In this section we discuss a system with almost-harmonic high-frequency oscillations. We show numerical phenomena of methods applied with step sizes that are not small compared to the period of the fastest oscillations.

### I.5.1 A Fermi–Pasta–Ulam Problem

> . . . dealing with the behavior of certain nonlinear physical systems where the non-linearity is introduced as a perturbation to a primarily linear problem. The behavior of the systems is to be studied for times which are long compared to the characteristic periods of the corresponding linear problems. (E. Fermi, J. Pasta, S. Ulam 1955)

> In the early 1950s MANIAC-I had just been completed and sat poised for an attack on significant problems. ... Fermi suggested that it would be highly instructive to integrate the equations of motion numerically for a judiciously chosen, one-dimensional, harmonic chain of mass points weakly perturbed by nonlinear forces. (J. Ford 1992)

The problem of Fermi, Pasta & Ulam (1955) is a simple model for simulations in statistical mechanics which revealed highly unexpected dynamical behaviour. We consider a modification consisting of a chain of $2m$ mass points, connected with alternating soft nonlinear and stiff linear springs, and fixed at the end points (see Galgani, Giorgilli, Martinoli & Vanzini (1992) and Fig. 5.1). The variables $q_1, \ldots, q_{2m}$

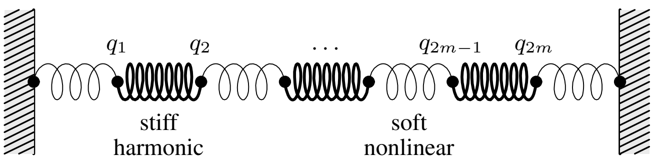



**Fig. 5.1.** Chain with alternating soft nonlinear and stiff linear springs

($q_0 = q_{2m+1} = 0$) stand for the displacements of the mass points, and $p_i = \dot{q}_i$ for their velocities. The motion is described by a Hamiltonian system with total energy

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{m}\big(p_{2i-1}^2 + p_{2i}^2\big) + \frac{\omega^2}{4}\sum_{i=1}^{m}(q_{2i} - q_{2i-1})^2 + \sum_{i=0}^{m}(q_{2i+1} - q_{2i})^4,$$

where $\omega$ is assumed to be large. It is quite natural to introduce the new variables

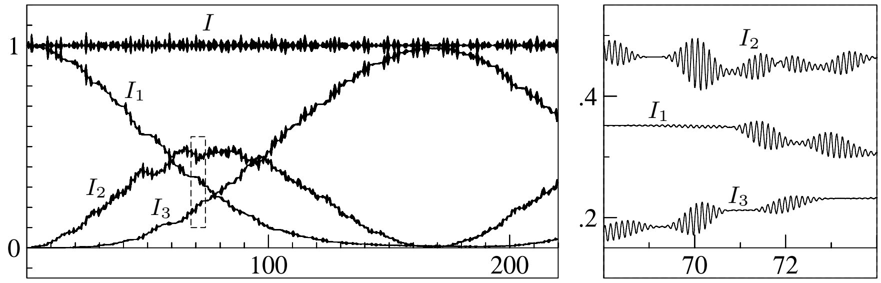



**Fig. 5.2.** Exchange of energy in the exact solution of the Fermi–Pasta–Ulam model. The picture to the right is an enlargement of the narrow rectangle in the left-hand picture

$$\begin{aligned} x_{0,i} &= \big(q_{2i} + q_{2i-1}\big)/\sqrt{2}, \qquad & x_{1,i} &= \big(q_{2i} - q_{2i-1}\big)/\sqrt{2}, \\ y_{0,i} &= \big(p_{2i} + p_{2i-1}\big)/\sqrt{2}, \qquad & y_{1,i} &= \big(p_{2i} - p_{2i-1}\big)/\sqrt{2}, \end{aligned} \tag{5.1}$$

where $x_{0,i}$ $(i = 1, \ldots, m)$ represents a scaled displacement of the $i$th stiff spring, $x_{1,i}$ a scaled expansion (or compression) of the $i$th stiff spring, and $y_{0,i}, y_{1,i}$ their velocities (or momenta). With this change of coordinates, the motion in the new variables is again described by a Hamiltonian system, with

$$\begin{aligned} H(y,x) \;=\; & \frac{1}{2}\sum_{i=1}^{m}\Big(y_{0,i}^2 + y_{1,i}^2\Big) + \frac{\omega^2}{2}\sum_{i=1}^{m} x_{1,i}^2 + \frac{1}{4}\Big((x_{0,1} - x_{1,1})^4 + \\ & + \sum_{i=1}^{m-1}\big(x_{0,i+1} - x_{1,i+1} - x_{0,i} - x_{1,i}\big)^4 + (x_{0,m} + x_{1,m})^4\Big). \end{aligned} \tag{5.2}$$

Besides the fact that the equations of motion are Hamiltonian, so that the total energy is exactly conserved, they have a further interesting feature. Let

$$I_j(x_{1,j}, y_{1,j}) = \frac{1}{2}\big(y_{1,j}^2 + \omega^2 x_{1,j}^2\big) \tag{5.3}$$

denote the energy of the $j$th stiff spring. It turns out that there is an exchange of energy between the stiff springs, but the total oscillatory energy $I = I_1 + \ldots + I_m$ remains close to a constant value, in fact, $I\big((x(t), y(t)\big) = I\big((x(0), y(0)\big) + \mathcal{O}(\omega^{-1})$. For an illustration of this property, we choose $m = 3$ (as in Fig. 5.1), $\omega = 50$,

$$x_{0,1}(0) = 1, \quad y_{0,1}(0) = 1, \quad x_{1,1}(0) = \omega^{-1}, \quad y_{1,1}(0) = 1,$$

and zero for the remaining initial values. Fig. 5.2 displays the energies $I_1, I_2, I_3$ of the stiff springs together with the total oscillatory energy $I = I_1 + I_2 + I_3$ as a function of time. The solution has been computed very carefully with high accuracy, so that the displayed oscillations can be considered as exact.

## I.5.2 Application of Classical Integrators

Which of the methods of the foregoing sections produce qualitatively correct approximations when the product of the step size $h$ with the high frequency $\omega$ is relatively large?

**Linear Stability Analysis.** To get an idea of the maximum admissible step size, we neglect the quartic term in the Hamiltonian (5.2), so that the differential equation splits into the two-dimensional problems $\dot{y}_{0,i} = 0$, $\dot{x}_{0,i} = y_{0,i}$ and

$$\dot{y}_{1,i} = -\omega^2 x_{1,i}, \quad \dot{x}_{1,i} = y_{1,i}. \tag{5.4}$$

Omitting the subscripts, the solution of (5.4) is

$$\begin{pmatrix} y(t) \\ \omega x(t) \end{pmatrix} = \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix} \begin{pmatrix} y(0) \\ \omega x(0) \end{pmatrix}.$$

The numerical solution of a one-step method applied to (5.4) yields

$$\begin{pmatrix} y_{n+1} \\ \omega x_{n+1} \end{pmatrix} = M(h\omega) \begin{pmatrix} y_n \\ \omega x_n \end{pmatrix}, \tag{5.5}$$

and the eigenvalues $\lambda_i$ of $M(h\omega)$ determine the long-time behaviour of the numerical solution. Stability (i.e., boundedness of the solution of (5.5)) requires the eigenvalues to be less than or equal to one in modulus. For the explicit Euler method we have $\lambda_{1,2} = 1 \pm ih\omega$, so that the energy $I_n = (y_n^2 + \omega^2 x_n^2)/2$ increases as $(1 + h^2\omega^2)^{n/2}$. For the implicit Euler method we have $\lambda_{1,2} = (1 \pm ih\omega)^{-1}$, and the energy decreases as $(1 + h^2\omega^2)^{-n/2}$. For the implicit midpoint rule, the matrix $M(h\omega)$ is orthogonal and therefore $I_n$ is exactly preserved for all $h$ and for all times. Finally, for the symplectic Euler method and for the Störmer–Verlet scheme we have

$$M(h\omega) = \begin{pmatrix} 1 & -h\omega \\ h\omega & 1 - h^2\omega^2 \end{pmatrix}, \quad M(h\omega) = \begin{pmatrix} 1 - \frac{h^2\omega^2}{2} & -\frac{h\omega}{2}\Big(1 - \frac{h^2\omega^2}{4}\Big) \\ \frac{h\omega}{2} & 1 - \frac{h^2\omega^2}{2} \end{pmatrix},$$

respectively. For both matrices, the characteristic polynomial is $\lambda^2 - (2 - h^2\omega^2)\lambda + 1$, so that the eigenvalues are of modulus one if and only if $|h\omega| \le 2$.

**Numerical Experiments.** We apply several methods to the Fermi–Pasta–Ulam (FPU) problem, with $\omega = 50$ and initial data as given in Sect. I.5.1. The explicit and implicit Euler methods give completely wrong solutions even for very small step sizes. Fig. 5.3 presents the numerical results for $H$, $I$, $I_1$, $I_2$, $I_3$ obtained with the implicit midpoint rule, the symplectic Euler, and the Störmer–Verlet scheme. For the small step size $h = 0.001$ all methods give satisfactory results, although the energy exchange is not reproduced accurately over long times. The Hamiltonian $H$ and the total oscillatory energy $I$ are well conserved over much longer time intervals. The larger step size $h = 0.03$ has been chosen such that $h\omega = 1.5$ is close

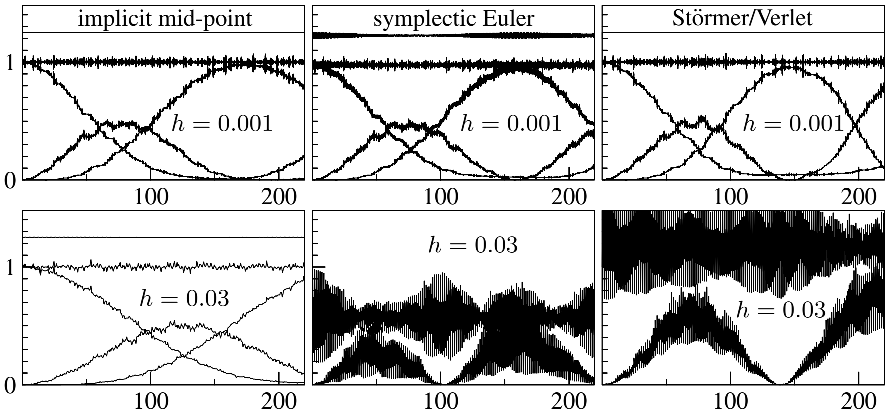



**Fig. 5.3.** Numerical solution for the FPU problem (5.2) with data as in Sect. I.5.1, obtained with the implicit midpoint rule (left), symplectic Euler (middle), and Störmer–Verlet scheme (right); the upper pictures use $h = 0.001$, the lower pictures $h = 0.03$; the first four pictures show the Hamiltonian $H - 0.8$ and the oscillatory energies $I_1, I_2, I_3, I$; the last two pictures only show $I_2$ and $I$

to the stability limit of the symplectic Euler and the Störmer–Verlet methods. The values of $H$ and $I$ are still bounded over very long time intervals, but the oscillations do not represent the true behaviour. Moreover, the average value of $I$ is no longer close to 1, as it is for the exact solution. These phenomena call for an explanation, and for numerical methods with an improved behaviour (see Chap. XIII).

# I.6 Exercises

1. Show that the Lotka–Volterra problem (1.1) in logarithmic scale, i.e., by putting $p = \log u$ and $q = \log v$, becomes a Hamiltonian system with the function (1.4) as Hamiltonian (see Fig. 6.1).

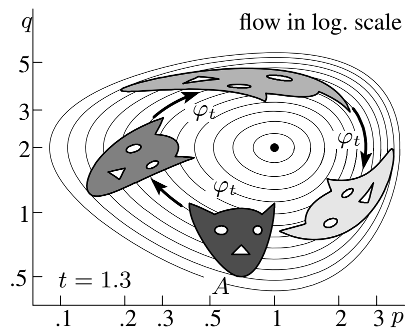



**Fig. 6.1.** Area preservation in logarithmic scale of the Lotka–Volterra flow

2. Apply the symplectic Euler method (or the implicit midpoint rule) to problems such as
$$\begin{pmatrix} \dot u \\ \dot v \end{pmatrix} = \begin{pmatrix} (v-2)/v \\ (1-u)/u \end{pmatrix}, \qquad \begin{pmatrix} \dot u \\ \dot v \end{pmatrix} = \begin{pmatrix} u^2 v(v-2) \\ v^2 u(1-u) \end{pmatrix}$$
with various initial conditions. Both problems have the same first integral (1.4) as the Lotka–Volterra problem and therefore their solutions are also periodic. Do the numerical solutions also show this behaviour?
3. A general two-body problem (sun and planet) is given by the Hamiltonian
$$H(p, p_S, q, q_S) = \frac{1}{2M}\, p_S^T p_S + \frac{1}{2m}\, p^T p - \frac{GmM}{\|q - q_S\|},$$
where $q_S, q \in \mathbb{R}^3$ are the positions of the sun (mass $M$) and the planet (mass $m$), $p_S, p \in \mathbb{R}^3$ are their momenta, and $G$ is the gravitational constant.
a) Prove: in heliocentric coordinates $Q := q - q_S$, the equations of motion are
$$\ddot Q = -G(M+m)\,\frac{Q}{\|Q\|^3}.$$
b) Prove that $\frac{d}{dt}\big(Q(t) \times \dot Q(t)\big) = 0$, so that $Q(t)$ stays for all times $t$ in the plane $E = \{q\,;\, d^T q = 0\}$, where $d = Q(0) \times \dot Q(0)$.
*Conclusion.* The coordinates corresponding to a basis in $E$ satisfy the two-dimensional equations (2.2).
4. In polar coordinates, the two-body problem (2.2) becomes
$$\ddot r = -V'(r) \qquad \text{with} \qquad V(r) = \frac{L_0^2}{2r^2} - \frac{1}{r}$$
which is independent of $\varphi$. The angle $\varphi(t)$ can be obtained by simple integration from $\dot\varphi(t) = L_0/r^2(t)$.
5. Compute the period of the solution of the Kepler problem (2.2) and deduce from the result Kepler's "third law".
*Hint.* Comparing Kepler's second law (2.6) with the area of the ellipse gives $\frac{1}{2}L_0 T = ab\pi$. Then apply (2.7). The result is $T = 2\pi(2|H_0|)^{-3/2} = 2\pi a^{3/2}$.
6. Deduce Kepler's first law from (2.2) by the elegant method of Laplace (1799).
*Hint.* Multiplying (2.2) with (2.5) gives
$$L_0 \ddot q_1 = \frac{d}{dt}\Big(\frac{q_2}{r}\Big), \qquad L_0 \ddot q_2 = \frac{d}{dt}\Big(-\frac{q_1}{r}\Big),$$
and after integration $L_0 \dot q_1 = \frac{q_2}{r} + B$, $L_0 \dot q_2 = -\frac{q_1}{r} + A$, where $A$ and $B$ are integration constants. Then eliminate $\dot q_1$ and $\dot q_2$ by multiplying these equations by $q_2$ and $-q_1$ respectively and by subtracting them. The result is a quadratic equation in $q_1$ and $q_2$.
7. Whatever the initial values for the Kepler problem are, $1 + 2H_0 L_0^2 \geq 0$ holds. Hence, the value $e$ is well defined by (2.9).
*Hint.* $L_0$ is the area of the parallelogram spanned by the vectors $q(0)$ and $\dot q(0)$.

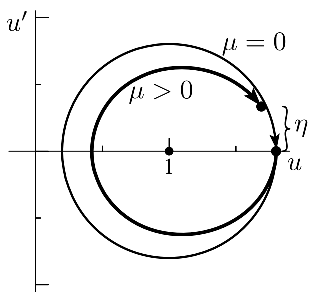
8. *Implementation of the Störmer–Verlet scheme.* Explain why the use of the one-step formulation (1.17) is numerically more stable than that of the two-term recursion (1.15).
9. *Runge–Lenz–Pauli vector.* Prove that the function

$$A(p,q) = \begin{pmatrix} p_1 \\ p_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ q_1p_2 - q_2p_1 \end{pmatrix} - \frac{1}{\sqrt{q_1^2+q_2^2}} \begin{pmatrix} q_1 \\ q_2 \\ 0 \end{pmatrix}$$

is a first integral of the Kepler problem, i.e., $A\big(p(t),q(t)\big) = Const$ along solutions of the problem. However, it is not a first integral of the perturbed Kepler problem of Exercise 12.
10. Add a column to Table 2.1 which shows the long-time behaviour of the error in the Runge–Lenz–Pauli vector (see Exercise 9) for the various numerical integrators.
11. For the Kepler problem, eliminate $(p_1, p_2)$ from the relations $H(p,q) = Const$, $L(p,q) = Const$ and $A(p,q) = Const$. This gives a quadratic relation for $(q_1, q_2)$ and proves that the solution lies on an ellipse, a parabola, or on a hyperbola.
12. Study numerically the solution of the perturbed Kepler problem with Hamiltonian

$$H(p_1,p_2,q_1,q_2) = \frac{1}{2}\left(p_1^2+p_2^2\right) - \frac{1}{\sqrt{q_1^2+q_2^2}} - \frac{\mu}{3\sqrt{(q_1^2+q_2^2)^3}},$$

where $\mu$ is a positive or negative small number. Among others, this problem describes the motion of a planet in the Schwarzschild potential for Einstein's general relativity theory[7]. You will observe a precession of the perihelion, which, applied to the orbit of Mercury, represented the historically first verification of Einstein's theory (see e.g., Birkhoff 1923, p. 261-264).




The precession can also be expressed analytically: the equation for $u = 1/r$ as a function of $\varphi$, corresponding to (2.8), here becomes

$$u'' + u = \frac{1}{d} + \mu u^2, \tag{6.1}$$

where $d = L_0^2$. Now compute the derivative of this solution with respect to $\mu$, at $\mu = 0$ and $u = (1 + e\cos(\varphi - \varphi^*))/d$ after one period $t = 2\pi$. This leads to $\eta = \mu(e/d^2)\cdot 2\pi \sin\varphi$ (see the small picture). Then, for small $\mu$, the precession after one period is

$$\Delta\varphi = \frac{2\pi\mu}{d}. \tag{6.2}$$

[7] We are grateful to Prof. Ruth Durrer for helpful hints about this subject.