# Applied Analytical Methods, Part 2: Evolution Equations with variational structure

150950

E. van Groesen
Applied Analysis & Mathematical Physics
Faculty of Mathematical Sciences,
University of Twente

Ourref : TW01/AAMP/EvG
Date : March 27, 2005

# Contents

# Chapter 1

# Introduction

In this course we consider problems with a variational structure and introduce specific methods to study these. The emphasis is on infinite-dimensional problems, since the origin of the problems lead to initial-boundary-value problems for partial or integral differential equations. Many results of finite dimensional case can and will be generalized.
Just as important as the introduction of the mathematical theory, is the second aim to illustrate where and how optimization/variational problems arise in nature and technical sciences, and how such a specific property can help to study and understand the problem[1].
Below we present a brief overview of the contents of the course.

## 1.1 General formulation of optimisation problems

Generally speaking, for an optimisation problem we have the following basic ingredients:

- a set of admissible elements $\mathcal{M}$, usually some subset of an (infinite dimensional) space $\mathcal{U}$;
- a functional $\mathcal{L}$, defined on $\mathcal{U}$ (or only on $\mathcal{M}$).

The optimisation problem of $\mathcal{L}$ on $\mathcal{M}$ concerns questions about an element $\hat{u}$ that minimizes the functional on the set of admissible elements, denoted by

$$\hat{u} \in \text{Min } \{\mathcal{L}(u) \mid u \in \mathcal{M}\},$$

which is by definition an element for which

$$\mathcal{L}(\hat{u}) \leq \mathcal{L}(u) \quad \text{for all} \quad u \in \mathcal{M}.$$

These questions may deal with the existence, the uniqueness, and the characterization and computation of the minimizer.

We will mainly deal with the *characterisation* of the minimizer (and more general critical points); instead of proving 'existence' of minimisers. We will concentrate on the equation(s) that have to be satisfied by such a critical point; a local investigation near the point will show that for density functionals on function spaces, the element $\hat{u}$ usually satifies some (ordinary or partial) differential equation, the *Euler-Lagrange equation*.
The actual calculation of minimising elements can be done by using numerical methods.

[1] In other disciplines, economy, life-sciences, etc. optimisation problems are also abundant; we will restrict ourselves and mainly deal with applications form the natural and technical sciences.

## 1.2 Mathematical method of local investigation

In real analysis courses at an introductory level, functions of one or more variables are considered. The definition of differentiation of functions is a vital part of such courses, and a standard result is the following

> **Algorithm of Fermat**, for 1-D optimisation problems.
> If the differentiable scalar function of one variable $f : \mathcal{R} \to \mathcal{R}$ attains a (local) extreme value at the point $\hat{x}$, then the derivative at that point vanishes:
>
> $$f'(\hat{x}) = 0.$$

Viewed as a condition for a point to be an extremal element, this condition is necessary but not sufficient; every point, including saddle points, that satisfy this property are called stationary, or critical, points.

Knowing the above result for functions of one variable, the generalisation to functions of more variables, $n$-dimensional problems, is remarkably simple using partial derivatives to reduce the problem to $n$ 1-D problems, as follows.
For $F : \mathcal{R}^n \to \mathcal{R}$, let $\nabla F$ be the gradient of the function, the column vector

$$\nabla F(x) = \begin{pmatrix} \partial_{x_1} F(x) \\ ... \\ \partial_{x_n} F(x) \end{pmatrix}.$$

Recall that the gradient is related to the (Frechet-) derivative using the standard innerproduct and is defined with the notion of directional derivative as follows:
At the point $x$ the directional derivative in the direction $\eta$ is found by differentiating the scalar function obtained by restricting $F$ to the line through $x$ in the direction $\eta$, i.e. the function

$$\varepsilon \to F(x + \varepsilon\eta),$$

and so

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} F(x + \varepsilon\eta) \equiv DF(x)\eta \equiv \nabla F(x) \cdot \eta.$$

If $x$ minimizes $F$ on $\mathcal{R}^n$, this point certainly minimizes the restriction (for $\varepsilon = 0$ ), and hence the directional derivative vanishes in every direction $\eta$:

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} F(x + \varepsilon\eta) = 0.$$

From $\nabla F(x) \cdot \eta = 0$ for all $\eta$, it then follows that

$$\nabla F(x) = 0.$$

This is the direct generalisation of Fermat's algorithm to $n$-dimensional optimization problems.

When the space of definition for a scalar function is infinite dimensional (the function is then usually called a function*al* ), the above can be generalised as follows:

- by restricting the functional to one dimensional lines the notion of directional derivative can be defined just as easily; it will be called the *first variation* in that case;

- when dealing with density-functionals, a generalisation of the gradient can be defined and will lead to the notion of *variational derivative*. The specific expression is related to the choice of the $L_2$ innerproduct for functions under consideration. *Lagrange's Lemma* is the result that enables the identification of the first variation with the variational derivative (modulo certain boundary conditions and some smoothness assumptions).

The typical notation to be used in the following for the variational derivative is $\delta\mathcal{L}(u)$, and Fermat's algorithm generalises to

$$\delta\mathcal{L}(u) = 0$$

as the condition for a minimizing element. This equation is most times a differential equation, replacing the algebraic equation $\nabla F(x) = 0$ that is obtained for a minimzer of a function of a finite number of variables.
Just as in finite dimensions, the second derivative may reflect minimisation properties, and in general provide insight into the character of a critical point.
In the Calculus of Variations these aspect are dealt with in the *theory of first and second variation*.

## 1.3 Optimality in the natural sciences

> *"..... je suis convaincu que par tout la nature agit selon quelque principe d'un maximum ou minimum."* (Euler, 1746)

This quotation of one of the greatest scientists that shaped the modern mathematical description and investigation of the natural sciences, expresses clearly the underlying expectation. The belief that optimization was important to describe natural phenomena was verified by Euler for various problems, and exploited to present a more thorough investigation of the problems. More far reaching conclusions were drawn by some other scientists:

> *"..... des loix du mouvement ou l'action est toujours employee avec la plus grande economie, demontreront l'existence de l'Etre supreme ... "*, (Maupertuis, 1757)

but this point of view belongs to metaphysics, and is as such not very fruitful for a deeper investigation[2].

Actually, optimization problems are known already from ancient times; well known is *Dido's problem*: the problem to find the plain domain of largest area given the circumference of the domain. Many other problems can also be formulated as *geodetic problems*, where one investigates those curves (or surfaces) with the property that a functional measuring the length (or the area) is as small as possible. A major example is the following

> **Fermat's principle**, 1662
> The actual trajectory of a light ray between two points in an inhomogeneous medium has the property that the time (or optical length) required to transverse the curve is as small as possible when compared to the time required for any other curve between the points.

[2]It should be noted, however, that modern theoretical physicists who look for "a theory of everything" (Grand Universal Theory) actually search for functionals (Lagrangians) that produce the desired unified field equations upon optimization, just as Einsteins general theory of relativity is based on a minimality principle.

In fact, the investigation of this principle led Fermat to the mathematical result stated above as Fermat's algorithm[3].
From Fermat's principle, *Snell's law* can be derived about the breaking of light between two media. A dual point of view (looking for the evolution of light fronts, the surfaces that can be reached by the light from a point source in a give time) was investigated by Huygens, 1695. *Huygen's principle*, of vital importance for the basic understanding of light propagation, can be considered as a major example of what later has become known as duality methods.
These historical remarks make it clear that, although the analytical methods of the classical Calculus of Variations were developed in the eighteenth century by scientists like Newton, Euler and Lagrange, some basic ideas can already be found in the seventeenth century. From a more closer historical investigation it becomes clear that practical problems from physics provided the initial motivation for the beautiful mathematical theory that has been developed since then[4].

## 1.4 Dynamical systems with a variational structure

Except from problems that have by their very nature an "obvious" formulation as a minimization problem (minimum length, minimum costs, etc), there are many problems for which such an extremizing property exists, but not so obvious. Important examples can be found in dynamical systems.
The *principle of minimum (potential) energy* leads to equilibrium states for which the total energy is minimal (the kinetic energy vanishes for equilibria). For nontrivial dynamic evolutions in certain systems, a less intuitive quantity, the "action" (see the quotation of Maupertuis), turns out to be an important functional; actual evolutions correspond to saddle points (not extremizers in general) of this functional. Formulations of such systems were studied by Lagrange, Hamilton etc., and the many results are collected in what is now called *Classical Mechanics*, a well structured set of methods and results to study dynamical systems of collection of mass points, mechanical (rigid) structures etc.
Nowadays, much effort is done to generalize these ideas to continuous systems, in particular to fluid dynamics, like water waves, and more general field theories. In this course we will deal with these modern aspects, particularly in Part II.
The systems referred to above, Lagrangian and Hamiltonian systems, are roughly speaking, conservative (the energy is conserved), and the dynamic motions have a variational nature. But even for systems that do not have such a structure, it is still possible that certain solutions have some extremizing property. This is true, for instance, for Poisson systems which are generalisations of Hamiltonian systems. When these systems are 'degenerate', no dynamic variational principle for the evolution can be found (in a simple way). However the equations do contain some variational struture (that may be be somewhat hidden), which makes it possible that special (but important)solutions can be characterised in a variational way. These solutions can be equilibrium (time independent) solutions, but can also be 'steady state solutions'. Often these are called *coherent structures* and are characteristic for such problems; examples are phenomena like 'travelling waves', 'solitons' and 'vortices'; owing to their variational nature, these can be found in a systematic way.
Even when the system is not conservative, but (mainly) dissipative, such as in gradient and thermodynamic systems, equilibrium solutions can be found by exploiting variational structures in the equations.
In all these cases that the special solutions can be characterised variationally, numerical algorithms can be used (or designed) to calculate the solutions approximately.

[3] Fermat did not write down the actual equation; he reasoned that small variations near a minimizer produces a higher order variation in the function, the fundamental idea that leads to the result and justifies to adhere his name to the mathematical algorithm. Fermat didn't know the concept of derivative of functions other than polynomials; it was Leibniz who introduced in 1684 the concept of derivative of arbitrary functions.

[4] The interested reader may consult such references like Goldstein, and Newman vol. 2.

## 1.5 Contents of the Lecture Notes

The basic methods and variational structures in the sciences are described in Part I.
Specifically, Chapter 2 deals with the standard local methods of first and second variation for problems without constraints, and applications are given in Chapter 3.
In Chapter 4 variational problems with constraints will be considered, with applications in Chapter 5. Lagrange's multiplier rule will be generalised for the case that the set of admissible functions is the (intersection of) level sets of given other functionals (equality constraints).
All these matters will be discussed; of course, the infinite dimensional setting will lead to various aspects that make the theory more difficult then the finite dimensional case, but the main ideas are comparable.
In Chapter 6 we will consider eigenvalue-problems for symmetric differential operators (the infinite dimensional analog of symmetric matrices) and characterize the eigenfunctions in a variational way; among others things, Fourier theory including the completeness of the set of eigenfunctons, will be treated.
The linear eigenvalue problems are related to minimisation of convex (quadratic) functionals. In Chapter 7 we deal with various other convex problems; when the functionals are no longer quadratic, but still convex, various advanced methods are available.
Part II mainly concerns more advanced variational theory for dynamical systems. Poisson structures are introduced, and generalisations to partial differential equations, describing various wave equations for instance, are treated.
In one Annex, MAPLE-commands for calculating directional derivatives, linearization and variational derivatives are described. In another Annex, a bird eye's view on functional analytic aspects is given.

# Part I

# Evolution equations with variational structure

# Chapter 2

# Evolution equations

In this chapter we recall some basic notions from the theory for finite dimensional dynamical systems, and then generalise these notions for socalled evolution equations, i.e. dynamical systems in infinite dimensional spaces (ususally function spaces).

## 2.1 Dynamical systems in finite dimensions, basic notions

Consider a system evolving in time, $t \in IR$, for which the *state* is determined by specifying at each $t$ a point $u$ from the *state space* $\mathcal{M}$. In general, $\mathcal{M}$ will be a linear space or some manifold embedded in a linear space. The dynamical system is said to be finite or infinite dimensional depending on whether $\mathcal{M}$ is finite of infinite dimensional.
Vaguely, the "state" is the collection of information that is necessary to find in a unique way the future states of the system. We start to consider finite dimensional systems for which the state space is $IR^n$, and we use notation that is standard in that case[1].

Dynamical systems are considered for which the evolution

$$t \mapsto x(t) \in IR^n$$

is described by a *differential equation* of the form

$$\dot{x} = f(x), \tag{2.1}$$

where $f$ is a given *vector field*, i.e. a mapping (in general nonlinear) from $IR^n$ into itself:

$$f : IR^n \to IR^n, IR^n \ni x \mapsto f(x) \in IR^n.$$

Observe specifically that we consider
• systems of equations that are of *first* order in time,
• autonomous systems: the vector field $f$ does not depend explicitly on time $t$.

[1] For efficiency reasons, later we will adjust notation to be able to deal with finite and infinite dimensional systems at the same time.

### 2.1.1 Flow as Lie-group of transformations

In the following it is assumed that the *initial value problem* for (2.1) is *well-posed*, meaning that for each $x_0 \in IR^n$ there exists a time interval $J$, with $0 \in J$, and a unique function $x : J \to IR^n$ such that

$$\begin{cases} \dot{x} &= f(x), \text{ for } t \in J \\ x(0) &= x_0. \end{cases} \tag{2.2}$$

(Note that the choice of the initial time as $t = 0$ is no restriction for the autonomous systems we consider.)
The *existence and uniqueness theorem* for ordinary differential equations guarantees the well-posedness if the vector field $f$ is *Lipschitz continuous* in a neighbourhood of $x_0$: for some $L$

$$|f(x) - f(y)| \leq L|x - y|,$$

for all pairs $x, y$ in a neighbourhood of $x_0$.
Differentiability of $f$ is sufficient for Lipschitz continuity; in that case the dependence on the initial value $x_0$ is in fact differentiable.

In general, the existence is only *local*: for sufficiently small values of $t$ around $t = 0$; the existence interval can be extended to a maximal existence interval $J$, a subset of $IR$. If the maximal existence interval $J$ is the whole real line $IR$ (or $IR_+$) the solution is said to be a *global solution*.
The existence theorem justifies to call $x$ the state of the system. This is the main reason to restrict the general considerations to systems that are of first order in time.

**Exercise 1** *1.* The simple non-linear scalar equation

$$\dot{x} = x^2$$

has no global solutions other than $x \equiv 0$.

*2.* The second order scalar equation (linear pendulum)

$$\ddot{x} + x = 0$$

can be written as a system of two first order scalar equations, e.g.

$$\dot{x} = y, \quad \dot{y} = -x;$$

the state is now the pair $(x, \dot{x})$ and not the scalar $x$ only.

The solution will typically be denoted by

$$x(t) = \Phi_t(x_0) \tag{2.3}$$

in the following; $\Phi_t$ is called the *flow-map* and can be interpreted in two different ways.

- For fixed $x_0$

  $$t \mapsto \Phi_t(x_0)$$

  describes the *trajectory* in the state space of the solution with initial data $x_0$ parameterised with the time $t$.

- For fixed $t \in IR$, $\Phi_t$ is a map from $IR^n$ into itself:

$$\Phi_t : IR^n \to IR^n.$$

**Exercise 2** *For given $n \times n$-matrix $A$, all solutions of the linear system $\dot{x} = Ax$ exist globally; the system has the exponential function as flow map:*

$$\dot{x} = Ax \text{ has flow } \Phi_t(x_0) = \exp[tA]x_0.$$

Assume from now on that $f$ is differentiable and that the solutions exists globally (for all $t$). Then the flow map has the special property that the map $\Phi_t$ is in fact one-to-one, with inverse given by $\Phi_t^{-1} = \Phi_{-t}$. Moreover, since the solution depends smoothly on the initial data, $\Phi_t$, and hence its inverse, is differentiable. From these observations it follows that $\Phi_t$ is a *one parameter family of transformations (diffeomorphisms) on $IR^n$*. From the unique solvability assumption it follows that $t \mapsto \Phi_t$ is in fact a group (see Fig. 2.1 and 2.2):

$$\Phi_{s+t} = \Phi_s \Phi_t \quad \text{for all } s. \text{ and } t$$

A group like this, depending differentiable on $t$, is an example of a *Lie group*. It is completely determined by the vector field $f$ since

$$f(x) = \frac{d}{dt}\Phi_t(x)\Big|_{t=0}. \tag{2.4}$$

$f$ is called the *infinitesimal generator* of $\Phi_t$.

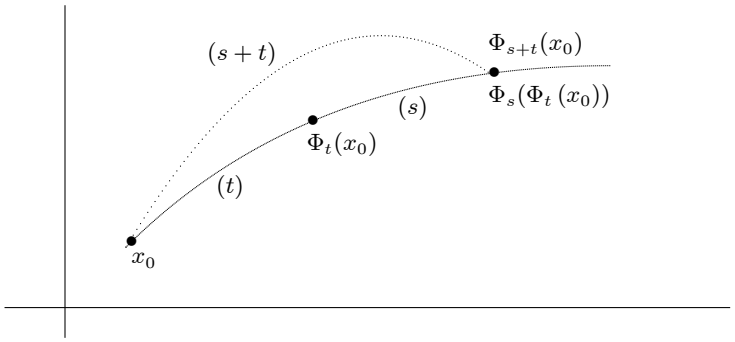


Figure 2.1: Illustration of the group property of the flow map $\Phi$ : $\Phi_{s+t}(x_0) = \Phi_s(\Phi_t(x_0))$ for each initial data $x_0$.

**Exercise 3**

1. *Verify the group property for the exponential map $\exp[tA]$ for an arbitrary $n \times n$-matrix $A$.*
2. *Calculate the flow map for the pendulum equation $\ddot{x} + x = 0$ and interpret its flow and the group property.*

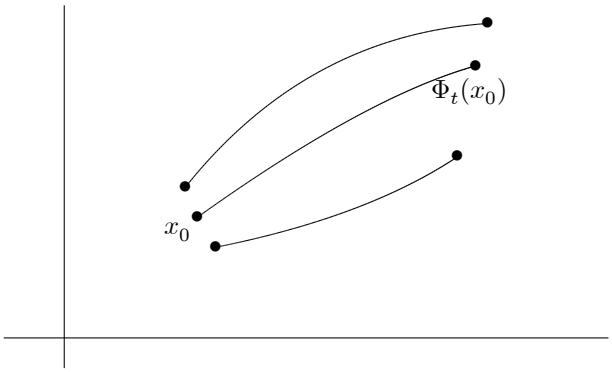


Figure 2.2: For fixed $x_0$, $t \mapsto \Phi_t(x_0)$ is the trajectory of the solution with initial data $x_0$; the vector field is tangent to this trajectory. For fixed $t$, $\Phi_t$ is a diffeomorphism on $IR^n$; a neighbourhood of $x_0$ is mapped under the flow onto a neighbourhood of $\Phi_t(x_0)$.

### 2.1.2 Equilibria and linearization

An *equilibrium solution*, shortly *equilibrium*, is a time-independent solution:

$\hat{x}$ such that $f(\hat{x}) = 0$.

It corresponds to a fixed point of the flow map:

$\Phi_t(\hat{x}) = \hat{x}$ for all time .

To find equilibria we have to find the zero's of the vector field. This is an algebraic problem: solving the $n$ components of $x$ from the $n$ equations (in general nonlinear) $f(x) = 0$. Although this may be a difficult problem in itself, it is the first attack to study a given diferential equation.

A next approach is to get some insight into the flow in a neighbourhood of an equilibrium is the procedure of linearization. Near an equilibrium $\hat{x}$, write a solution as

$$x(t) = \hat{x} + \varepsilon\xi(t) + \mathcal{O}(\varepsilon^2). \tag{2.5}$$

Expansion of the vectorfield (assuming differentiability) near $\hat{x}$ gives

$$f(x) = f(\hat{x}) + \varepsilon Df(\hat{x})\xi + \mathcal{O}(\varepsilon^2),$$

where $Df(\hat{x})$ is the Jacobi-matrix consisting of partial derivatives: for $f = (f_1, \ldots, f_n$

$$[Df(\hat{x})]_{j,k} = \frac{\partial f_j(\hat{x})}{\partial x_k}.$$

Since $f(\hat{x}) = 0$ for an equilibrium, the *linearized equation* is given by the linear system

$$\dot{\xi} = Df(\hat{x})\xi. \tag{2.6}$$

The solutions of this system define the *linearized flow*. As long as a solution $\xi$ of this system is of order 1, $\hat{x} + \varepsilon\xi(t)$ is a valid approximation of the solution $\Phi_t(\hat{x} + \xi(0))$ of the complete system.

**Exercise 4** Investigate the linearization at the equilibria of the pendulum equation

$$\ddot{x} + \sin x = 0.$$

Determine the linearized flow, and sketch the flow of the nonlinear equation.

### 2.1.3 Lyapunov stability

**Definition 5** Lyapunov stability of an equilibrium.
*With $\hat{x}$ be an equilibrium solution of $\dot{x} = f(x)$, i.e $f(\hat{x}) = 0$, $\hat{x}$ is said to be* stable in the sense of Lyapunov *if the following holds:*
*for each $\varepsilon > 0$ there exists $\delta > 0$ such that for each initial data $x_0$ the solution $\Phi_t(x_0)$ (is defined for all $t \geq 0$ and) remains within an $\varepsilon$-distance from $\hat{x}$ for all $t > 0$:*

$$|\Phi_t(x_0) - \hat{x}| < \varepsilon$$

*provided $x_0$ lies within a $\delta$-distance from $\hat{x}$: $|x_0 - \hat{x}| < \delta$. Shortly:*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x_0 \in \mathcal{M} \; \forall t \geq 0 : \; |x_0 - \hat{x}| < \delta \Rightarrow |\Phi_t(x_0) - \hat{x}| < \varepsilon.$$

In this definition, the value of $\varepsilon$ measures the a priori prescribed maximal allowed deviation of the solution from the equilibrium; the value of $\delta$ (depending on $\varepsilon$) measures the maximal allowable deviation of the initial disturbance.

**Exercise 6** *1.* Show directly (finding $\delta$ for given $\varepsilon$) the Lyapunov stability of the trivial equilibrium $\hat{x} = 0$ of the following systems

$$\dot{x} = -x; \quad \ddot{x} + x = 0; \quad \ddot{x} + sinx = 0.$$

*2.* Give (sufficient) conditions for the matrix $A$ in order that the origin $\hat{x} = 0$ is a Lyapunov stable equilibrium of the system $\dot{x} = Ax$.

### 2.1.4 Diagnostics of dynamics with functions, Lyapunov functions

In the following we will regularly be able to use the following simple, but essential, idea. The equations we are considering are too difficult to find the solutions explicitly. Then *any* information we can detect for the flow map is valuable. Sometimes it is possible to obtain useful global information from the way how a certain scalar function defined on the state space (for instance the energy) evolves when calculated at a solution. If we denote the function by $E : IR^n \to IR$, for given initial data $x_0$, this evolution is given by

$$t \to E(\Phi_t(x_0)).$$

Obviously this idea can give information provided (i) the evolution of $E$ can be found (or estimated), and (ii) the value of $E$ alone tells something about the possible states $x$ with that value.
We first addres the second issue in a general way; this is useful for the following to develop the correct intuitive ideas. Then we give as a specific example the idea of a Lyapunov function.

#### Squeezing property of level sets of scalar functions

Let $E$ be such that

$$E(0) = 0, \text{ and } E(x) > 0 \text{ for all } x \text{ in a ngbh of } 0.$$

Then intuitively, near 0 the value of $E$ is small positive (if $E$ is continuous), and, conversely, the smaller the value of $E(x)$, the closer the point $x$ should be to 0.
To make this more precise, denote the *level sets* of $E$ as follows:

$$E^{-1}(\sigma) := \{x \mid E(x) = \sigma\}.$$

**Definition 7** Geometric squeezing property
*The geometric squeezing property for the function $E$ means that there exist continuous functions $\rho_1, \rho_2 : IR_+ \to IR_+$, monotonically increasing with $\rho_1(0) = \rho_2(0) = 0$, such that*

$$\rho_2(|x|) \leq E(x) \leq \rho_1(|x|).$$

*Interpretation:* if $x$ has value $E(x) = \sigma$, i.e. $x$ lies on the $\sigma$ level set of $E$, then the distance to the origin is bounded above and below:

$$\rho_1^{-1}(\sigma) \leq |x| \leq \rho_2^{-1}(\sigma);$$

the $\sigma$-level set is "squeezed" between a ball with radius $\rho_1^{-1}(\sigma)$ and a ball with radius $\rho_2^{-1}(\sigma)$. Hence, the value of $E(x)$ gives some information where $x$ can be found.

**Lemma 8** *A continuous scalar function $E : IR^n \mapsto IR$ vanishing at 0 and sign definite in a neighbourhood, has the squeezing property.*

**Proof.** Define functions $\bar{\rho}_1$ as the largest value of $E$ on the ball, and $\bar{\rho}_2$ as the smallest value outside the ball:

$$\bar{\rho}_1(\delta) := \max\{E(x) \mid |x| \leq \delta\}, \quad \text{and} \quad \bar{\rho}_2(\varepsilon) := \min\{E(x) \mid |x| \geq \varepsilon\}.$$

Verify that the functions are well defined, non-decreasing, vanishing only at 0. Then construct strictly monotone functions $\rho_1, \rho_2$ such that $\bar{\rho}_1 \leq \rho_1, \bar{\rho}_2 \geq \rho_2$, and the proof is complete. □.

**Exercise 9** Let $S$ be a symmetric matrix, with all eigenvalues (necessarily real) positive. Consider the quadratic function $x \to x \cdot Sx$. Construct the functions $\rho_1, \rho_2$ explicitly in this case.

**Lyapunov functions**

**Definition 10** *For the d.e. $\dot{x} = f(x)$ with equilibrium $\hat{x}$, a function is called a Lyapunov function if the following two conditions are satisfied:*
*i. vanishes at $\hat{x}$ and is positive in a ngbh;*
*ii. evaluated at solutions starting near $\hat{x}$, the value of is nonincreasing.*

**Proposition 11** *If there exists a Lyapunov functional, the equilibrium is stable in the sense of Lyapunov.*

**Proof.** It is no restriction to assume that $\hat{x} = 0$. From the first condition it follows that the squeezing property holds for certain functions $\rho_1, \rho_2$. Given $\varepsilon > 0$, the monotonicity of $\rho_1$ and $\rho_2$ implies that there exists a $\delta > 0$ such that

$$\rho_2(\sigma) < \rho_1(\delta) \Rightarrow \sigma < \varepsilon \tag{2.7}$$

(see the figure).
For this $\delta = \delta(\varepsilon)$, the condition of stability can be proved as follows. For any $x_0$, with $|x_0| < \delta$ it holds that

$$\rho_2(|\Phi_t(u_0)|) \leq (\Phi_t(x_0)) \leq (x_0) \leq \rho_1(|x_0|)) < \rho_1(\delta)$$

and from $\rho_2(|\Phi_t(u_0)|) \leq \rho_1(\delta)$ and the definition of $\delta$: $|\Phi_t(u_0)| < \varepsilon$. □

**Exercise 12** *1.* Show that $(x) = \frac{1}{2}x \cdot Sx$ is a Lyapunov function for the linear system $\dot{x} = -Sx$ when $S$ is a symmetric, positive-definite matrix.

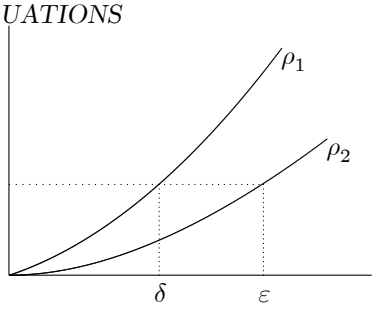


Figure 2.3:

*2.* Show that the same Lyapunov function can be used for the non-linear system

$$\dot{x} = Sx + r(x), \ \textit{with} \ \ r(x) = \mathcal{O}(x^2).$$

This is a special case of the Theorem of Poincaré-Lyapunov: the linearization around the equilibrium (here $\hat{x} = 0$) determines the stability of the equilibrium of the nonlinear system.

## 2.2 Evolution equations in infinite dimensional spaces

In contrast to finite dimensional systems for which the state is determined with a finite number of variables (the components of $x \in I\!R^n$ for some finite $n$), in an infinite dimensional system specification of the state requires an infinity of data: the state is an element from an infinite dimensional space. As in the finite dimensional case, we consider dynamical systems described by a differential equation of the form

$$\partial_t u = K(u). \tag{2.8}$$

Here $K$, again called a *vector field*, is a mapping from the state space in itself:

$$K : \mathcal{U} \to \mathcal{U}.$$

**Exercise 13** *1.* A simple example is $K(u)(x) = -c_0 \partial_x u$, ($c_0$ a given constant) in which case the dynamic eqn is the simplest partial differential equation

$$\partial_t u = -c_0 \partial_x u$$

of which the general solution can be written down explicitly:

$$u(x,t) = f(x - c_0 t), \ \textit{for any function} \ \ f.$$

With $u$ the elevation of the water surface, this represents a 'wave', with profile $f$ that translates with constant velocity $c_0$.

*2.* More complicated pde's are for instance the nonlinear diffusion equation

$$\partial_t u = \partial_x^2 u + u(1-u),$$

and the famous Korteweg - de Vries equation

$$\partial_t u + c_0 u_x + \alpha u_{xxx} + \beta u u_x = 0$$

that describes (approximately) uni-directional surface waves on a layer of fluid above a flat bottom.

Equations of the general form $\partial_t u = K(u)$ in infinite dimensions are often called *evolution equations*; infinite dimensional systems are also called *continous systems*, to distinguish from the *discrete systems* in finite dimensions.

### 2.2.1 Initial value problem

Local existence and uniqueness of the *initial value problem*

$$\begin{cases} \partial_t u &= K(u), u \in \mathcal{U}, \\ u(0) &= u_0 \end{cases} \tag{2.9}$$

can be proved in exactly the same way as in finite dimensions provided $\mathcal{U}$ is a Banach space and $K$ is Lipschitz continuous (of course with respect to the norm in the Banach space).

When all solution exist globally (for all $t$), the flow map can and will be defined as before, and defines a one parameter Lie-group. However, global existence is now more cumbersome than in finite dimensions. The next example shows that the solutions of a simple linear (diffusion) equation exist (in any reasonable space) only for $t \geq 0$; this is a motivation to study *semi-groups* (defined for $t \geq 0$ only), which we will not do.

**Exercise 14** *1.* Consider the initial boundary value problem for the linear diffusion equation:

$$\begin{cases} \partial_t u &= u_{xx} \\ u(0,t) &= u(\pi,t) = 0, \\ u(x,0) &= u^0(x) \equiv \sum u_k^0 \sin kx \end{cases}$$

The solution can be written down explicitly:

$$u(x,t) \equiv \Phi_t(u^0) = \sum u_k^0 e^{-k^2 t} \sin kx,$$

from which existence for all $t \geq 0$ follows, but not for $t \leq 0$ in general.

*2.* Note, however, that if $u^0 \in T_N$, we essentially have a finite dimensional problem and at the same time global existence: for all $t \in IR$.

### 2.2.2 Equilibria and linearization

An *equilibrium* (solution) of $\partial_t u = K(u)$ is any element $\hat{u}$ such that

$$\hat{u} \in \mathcal{U}, \quad K(\hat{u}) = 0.$$

Note that when dealing with partial differential equations, the equation for $\hat{u}$ will be a (partial) differential equation in the spatial variables.

**Exercise 15** Investigate (*phase plane analysis*) the equilibria of the nonlinear diffusion equation

$$\partial_t u = u_{xx} + u(1-u)$$

on the set of functions that approach constant values at infinity.

When thinking about linearization around an equilibrium, it appears that we need the notion of differentiability of the vector field $K$. This requires the specification of a norm, at least to specify the notion of a *Frechet differentiable* map. Instead of doing this, we use a more formal definition, and are satisfied when the usual limiting procedure leads to the definition of a linear map, the formal derivative.

**Definition 16** Formal Frechet derivative
*For a map $K : \mathcal{U} \to \mathcal{U}$, the formal Frechet derivative of $K$ at $\hat{u}$ is defined as the linear map $DK(\hat{u})$ such that*

$$\frac{d}{d\varepsilon} K(\hat{u} + \varepsilon v)\Big|_{\varepsilon=0} = DK(\hat{u})v \tag{2.10}$$

Note that for rigorous Frechet differentiability in a Banach space $U$ it is furthermore required that $DK(\hat{u})$ is a continuous (i.e. bounded) map. In that case, the linearization

$$K(\hat{u} + \varepsilon v) = K(\hat{u}) + \varepsilon DK(\hat{u})v + o(\varepsilon)$$

has a precise meaning for the remainder: $||\frac{o(\varepsilon)}{\varepsilon}|| \to 0$ for $\varepsilon \to 0$.

In any case, with a similar motivation as in the finite dimensional case, the linearization is defined.

**Definition 17** *For the evolution equation*

$$\partial_t u = K(u),$$

*the* linearized equation *at an equilibrium point $\hat{u}$ is defined as the linear evolution equation*

$$\partial_t \, \xi = \; DK(\hat{u}) \, \xi. \tag{2.11}$$

**Exercise 18** *1. Determine the linearization of the non-linear diffusion equation and of the KdV-equation around the trivial equilibrium solution.*

*2. Show that the linearization around the trivial state of a linear evolution eqn is the equation itself.*

### 2.2.3 Lyapunov stability

Formally, the definition of stability of an equilibrium can be formulated as in finite dimensions. However, since we may have various norms, we can generalize the notion of stability by measuring the initial disturbance with a different (stronger) norm than the measure for the evolution.

**Definition 19** Lyapunov stability of an equilibrium.
*Let $\hat{u}$ be an equilibrium solution: $K(\hat{u}) = 0$. Then $\hat{u}$ is said to be* stable in the sense of Lyapunov *with respect to the norms by $||\;\;||_1$ and $||\;\;||_2$ if:*

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall u_0 \in \mathcal{U} \;\; \forall_{t \geq 0} : \; ||u_0 - \hat{u}||_1 < \delta \Rightarrow ||\Phi_t(u_0) - \hat{u}||_2 < \varepsilon.$$

### 2.2.4 Diagnostics with functionals

Guided by the finite dimensional result, one can look at the level sets of functionals on the infinite dimensional space $\mathcal{U}$:

$$E : \mathcal{U} \to IR.$$

Specific functionals are the norms, $E(u) = ||u||^2$; the level sets then define the spheres in the space. Unlike in finite dimensions, it is not true that in inifinite dimensions a continuous function satisfies the squeezing property.

**Exercise 20** *1.* For $E(u) = ||u_x||^2$ on $U = \{u : [0,\pi] \to IR \mid u(0) = u(\pi) = 0\}$ we have $||u||^2 \leq E(u)$ (Poincare-Friedrichs), but not $E(u) \leq C||u||^2$ for some constant $C$.

*2.* On the other hand, on the same space, for $E(u) = ||u||^2$, we have $E(u) \leq ||u_x||^2$, but not $c||u_x||^2 \leq E(u)$ for some $c$.

Therefore we usually have to be satisified with a squeezing property of the level sets measured by two different norms.

**Definition 21** Geometric squeezing property
*The geometric squeezing property for the functional $E$ with respect to norms $||\ ||_1$ and $||\ ||_2$ means that there exist continuous functions $\rho_1, \rho_2 : IR_+ \to IR_+$, monotonically increasing with $\rho_1(0) = \rho_2(0) = 0$, such that*

$$\rho_2(||u||_2) \leq E(u) \leq \rho_1(||u||_1)$$

**Interpretation:** the squeezing property implies that the $\sigma$-level set of $E$ is "squeezed" between a $||\ ||_1$-ball with radius $\rho_1^{-1}(\sigma)$ and a $||\ ||_2$-ball with radius $\rho_2^{-1}(\sigma)$:

$$||u||_2 \leq \rho_2^{-1}(\sigma), \quad ||u||_1 \geq \rho_1^{-1}(\sigma). \tag{2.12}$$

Again, the value of $E(u)$ gives some information about the property of $u$.

### 2.2.5 Lyapunov functionals

In view of the foregoing, the following definition and result are easily motivated and proved.

**Definition 22** *A functional $: \mathcal{U} \to IR$ is called a* Lyapunov functional *for the evolution equation $\partial_t u = K(u)$ at the equilibrium $\hat{u}$ if satisfies:*
*(i)* non-negativity*: $(u) \geq 0$ for all $u \in \mathcal{U}$ and $(\hat{u}) = 0$,*
*(ii)* geometric squeezing property: *there exist continuous functions $\rho_1, \rho_2 : IR_+ \to IR_+$ monotonically increasing with $\rho_1(0) = \rho_2(0) = 0$, and norms $||\ ||_1$ and $||\ ||_2$ such that !!!!!!!! (iii)* the dynamic condition: $t \to (\Phi_t(u_0))$ *is non-increasing.*

**Proposition 23** *If there exists a Lyapunov functional with the properties listed in the definition above, the equilibrium solution $\hat{u}$ is Lyapunov stable with respect to the metrics $\rho_1$ and $\rho_2$.*

**Exercise 24** *1.* For the simple diffusion equation,

$$\begin{aligned} \partial_t u &= u_{xx}, \\ u(0,t) &= u(\pi,t) = 0, \end{aligned}$$

the trivial equilibrium is clearly stable in view of the explicit solution. We can prove Lyapunov stability with respect to the norm $||\ \ ||$ by using this norm itself as Lyapunov functional:

$$_0 = ||u||^2.$$

The squeezing property is immediate, and the dynamic condition follows from

$$\partial_{t0}(u(t)) = \partial_t \int u^2 = 2\int u\partial_t u = 2\int uu_{xx} = -2\int u_x^2 \leq 0$$

where we have performed a partial integration in the last step, using the homogeneous boundary condition.

2. Show that with this expression an estimate for the exponential decrease of $||u(t)||$ follows.

3. One can also try to use $_1 = ||u||_1^2$ as a Lyapunov functional to derive the stability in a stronger norm (for initial conditions $u_0$ for which $||u_0||_1$ is finite). Show that $_1$ is also non-increasing; find an estimate for the exponential decrease of this norm.

## 2.3 Evolution equations with variational structure

Two classes of equations (not disjoint) can be recognized; in each case a characteristic functional can be found that influences the dynamics, and that makes it possible to characterize one class as being 'dissipative equations', and the other class as being 'conservative'. One class are the Hamiltonian systems, considered already before, and the generalisation to *Poisson systems*; these systems are 'conservative' with the Hamiltonian ('energy') being conserved. In contrast, *gradient systems* are 'dissipative': during the evolution the value of a functional decreases in the direction of *steepest descent*; after discretising the time derivative, these dynamic equations are often used in a numerical code to determine the (local) minima of a given function.

## 2.4 Summary of conservative and dissipative structures

In the previous chapter we have considered various differential equations and investigated them in an ad hoc way. The main methods that have been summarized are the investigation of equilibria, and the investigation of the linearized equation. In fact, without further assumptions, not much more can be done.

In this chapter we consider evolution equations which have a special structure; these special structures are often found in equations from the physical and technical sciences. A much more rich theory can be developed by exploiting the specific structure in the equations. Essentially two large classes of equations will be considered; although not completely defined by the following criterion, it is usefull to characterize them a priori by the effect that the evolution in the state space has on the evolution of some suitably defined function or functional.
We will roughly distinguish between *dissipative equations* for which a (dissipation-)functional decreases monotonically, which implies that the evolution is driven to a much simpler set, for instance to an equilibrium. The most characteristic example is the (linear) diffusion equation.
For the other class of equations one specific functional remains constant during the evolution. In many cases this functional is an energy-like quantity $I$, just like for Hamiltonian systems. Then the dynamics is restricted to level sets of the functional, the value of which is determined by the initial value: $I(u(t)) = I(u_0)$ for all time $t$. From this "conservation" property, the name *conservative systems* is in

wide use. It should be noted that only the restriction to a level set does not give much information: the dynamics is merely restricted to a set of codimension one. However, the specific structure that is commonly encountered makes it possible to obtain more information.

To motivate the following, we start to summarize the general form of the equations we will encounter. As will be done from now on, we consider finite and infinite dimensional systems at the same time.

**Summary:** *Equations with a specific structure.*
The equations we will consider have, roughly, the following general form: for $u$ in a finite or infinite dimensional space $\mathcal{U}$

$$\partial_t u = -S\delta F(u) + \Gamma\delta H(u) \tag{2.13}$$

with:

- $F$ and $H : \mathcal{U} \to IR$ given functionals,
- $\delta F$ and $\delta H$ denoting the "variational derivatives" of $F$ and $H$ respectively, the direct generalization of the "gradient" of a function on a finite dimensional space, and considered as elements from the dual space $\mathcal{U}^*$,
- $S : \mathcal{U}^* \to \mathcal{U}$, a linear, self-adjoint (symmetric) operator

  $$\langle \zeta, S\eta \rangle = \langle S\zeta, \eta \rangle,$$

  and $\Gamma : \mathcal{U}^* \to \mathcal{U}$, a linear skew-symmetric operator

  $$\langle \zeta, \Gamma\eta \rangle = -\langle \Gamma\zeta, \eta \rangle$$

  (with certain additional properties, to be explained later on).

Then, roughly speaking, the difference between the dissipative and conservative structure is as follows:

$$\partial_t u = -S\delta F(u) \qquad \text{"dissipative"} \tag{2.14}$$
$$\partial_t u = \Gamma\delta H(u) \qquad \text{"conservative"} \tag{2.15}$$

Most, if not all, eqn's from Mathematical Physics can be formulated in such a way (possibly not in an unique fashion). We list some of the most important standard examples.

**Exercise 25** *1.* The pendulum equation $\ddot{x} + \sin x = 0$ can be written as as a conservative system

$$\partial_t u = J\nabla H(u), \quad \textit{with } u = \begin{pmatrix} q \\ p \end{pmatrix},$$

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad H(u) = \frac{1}{2}p^2 + (1 - \cos q), \quad \nabla H(u) = \begin{pmatrix} \partial_q H \\ \partial_p H \end{pmatrix}.$$

*2.* The *linear wave equation*: $\partial_t u = -c_0 \partial_x u$, for functions on $IR$ vanishing at infinity, can be written as

$$\partial_t u = \Gamma\delta H(u) \textit{ with } \Gamma = -c_0\partial_x, \quad H(u) = \int \frac{1}{2}u^2, \quad \delta H(u) = u.$$

3. The *KdV-equation*

$$\partial_t u = -c_0(u_x + \alpha u_{xxx} + \beta u u_x),$$

for functions on $IR$ vanishing at infinity, can be written as a conservative system

$$\partial_t u = \Gamma \delta H(u) \quad with \quad \Gamma = -c_0 \partial_x,$$

$$H(u) = \int (\frac{1}{2}u^2 - \frac{1}{2}\alpha u_x^2 + \frac{\beta}{6}u^3), \quad \delta H(u) = u + \alpha u_{xx} + \frac{1}{2}\beta u^2.$$

4. The *linear diffusion equation:* $\partial_t u = \partial_x^2 u$, for functions on $IR$ vanishing at infinity, or on a compact interval with Dirichlet or Neumann boundary conditions, can be written as a dissipative system $\partial_t u = -S\delta F(u)$ with

$$S = -\partial_x^2, \quad F = \int \frac{1}{2}u^2, \quad \delta F(u) = u,$$

or with

$$S = Id(entity), \quad F = \int \frac{1}{2}u_x^2, \quad \delta F(u) = -\partial_x^2 u.$$

5. The *Burgers equation* is the nonlinear diffusion equation

$$\partial_t u = u_{xx} + u u_x$$

and has a mixed structure of the form

$$\partial_t u = \partial_x^2 \delta F(u) + \partial_x \delta H(u), \quad for \quad F = H = \int \frac{1}{2}u^2$$

6. The *KdV-Burgers equation*

$$\partial_t u = u_{xx} - u_{xxx} + u u_x$$

has also a mixed structure:

$$\partial_t u = \partial_x^2 \delta F(u) + \partial_x \delta H(u), \quad for \quad F = \int \frac{1}{2}u^2, \quad H(u) = \int (\frac{1}{2}u_x^2 + \frac{1}{6}u^3).$$

From the general formulation it is clear that derivatives (gradient and varational derivative) of function(al)s play a fundamental role.

## 2.5 Gradient systems, Steepest descent

Gradient systems usually describe systems with some dissipative character. Since these systems can also be used in a constructive way to calculate minimizers of a given smooth functional, we will treat them briefly.

**Definition 26** *A* gradient system *is a dynamical system in the state space* $\mathcal{U}$ *of the form*

$$\partial_t u = -\delta H(u)$$

*where* $H : \mathcal{U} \to IR$ *a given functional.*

**Remark 27** It is possible to consider, somewhat more general, equations of the form

$$\partial_t u = -S\delta H(u)$$

with: $S : U^* \to U$, a linear and self-adjoint operator $\langle \zeta, S\eta \rangle = \langle S\zeta, \eta \rangle$, that is positive-definite: $\langle \zeta, S\zeta \rangle > 0$. Then much of the following is easily generalized to this case.

It is important to note that equilibrium solutions of a gradient system arise from a variational principle.

**Proposition 28** *For a gradient system* $\partial_t u = -\delta H(u)$*, the dynamic* equilibrium solutions $\hat{u}$ *are precisely the critical points of the functional* $H$*:*

$$\hat{u} \in Crit_u H(u) : \quad \delta H(\hat{u}) = 0.$$

In particular, when $H$ is an energy-like quantity that can be minimized, the principle of minimum energy applies, as we shall show now by the following investigation.

The role that the functional plays for the dynamics, can in fact be understood quite well by studying the evolution of $H$ on trajectories and explains the name *dissipative*: $H$ decreases monotonically outside equilibria:

$$\partial_t H(u) = \langle \delta H(u), -\delta H(u) \rangle = -|\delta H(u)|^2 \quad \begin{cases} \leq 0 \\ = 0 \quad \text{iff} \quad \delta H(u) = 0 \end{cases} .$$

This shows in particular that if $\bar{u}$ is an isolated local minimizer of $H$, trajectories starting close eventually apporach the point $\bar{u}$; it is said that the point $\bar{u}$ is an *asymptotically stable equilibrium solution*; the rate of convergence to the minimizer depends on the geometry of the level sets of $H$ near $\bar{u}$.

**Proposition 29** *Local minimizers of* $H$ *are asymptotically stable equilibrium solutions of the gradient system.*

The decrease of $H$ on solutions also clearly shows that the solutions define trajectories of *steepest descent*: at each point the trajectory is along the direction of steepest descent of the functional $H$. They define the most "efficient" way to reach lower values of $H$.
This observation can be exploited in a constructive way. Using a discretization for the time derivative, these equations are often used to find a (possibly local) minimizer of $H$ in a numerical way.
The dynamic behaviour near a local minimizer (where $H$ decreases to its lowest possible value) is completely determined by the topological properties of the level sets of $H$.

**Exercise 30** *1.* In finite dimensions with $x \in IR^n$ the equation is of the form

$$\partial_t x = -\nabla H(x).$$

Supposing that $x = 0$ is a (local) minimizer for $H$, determine the rate of convergence to 0 depending on the value of $\alpha$ ($> 1$) when $H(x) \approx |x|^\alpha$ near 0.

*2.* The simplest linear diffusion equation

$$\partial_t u = u_{xx}, \quad u(0) = u(\pi) = 0$$

is of this form with $H(u) = \int \frac{1}{2} u_x^2$; determine the rate of convergence to the zero state.

*3.* The non-linear diffusion equation ($\alpha \in IR$)

$$\partial_t u = u_{xx} + \alpha u(1-u), \quad u(0) = u(\pi) = 0.$$

is a gradient system too:

$$\partial u = -\delta H(u) \quad \textit{with} \quad H(u) = \int [\frac{1}{2} u_x^2 - \alpha(\frac{1}{2}u^2 - \frac{1}{3}u^3)].$$

Investigate for which values of $\alpha$ the trivial solution $u \equiv 0$ is the minimizer, and find the rate of convergence in that case.

## 2.6 Thermodynamic systems, Constrained steepest descent

We have seen that gradient systems can be used for numerical purposes to find the minimizer of a given functional; the dynamic trajectories are in the direction of steepest descent.

When looking for constrained minimizers, this method has to adapted to take the constraints into account. We briefly describe the modification; for simplicity we restrict ourselves to the case of one functional constraint.
These systems are called thermodynamic systems, since there is one integral that is conserved (the "energy") while another one decreases monotonically (the "entropy")[2].

**Definition 31** *A* thermodynamic system *is a dynamical system in the state space $\mathcal{U}$ that is of the form*

$$\partial_t u = -[\delta H(u) - \lambda(u)\delta I(u)], \quad \textit{with } \lambda(u) = \frac{\langle \delta I(u), \delta H(u) \rangle}{\langle \delta I(u), \delta I(u) \rangle}$$

*where $H : \mathcal{U} \to \mathcal{R}$ and $I : \mathcal{U} \to \mathcal{R}$ are functionals.*

To see the dynamic properties, observe that any functional $F$ evolves according to

$$\partial_t F(u) = \langle \delta F(u), [\delta H(u) - \lambda(u)\delta I(u)] \rangle.$$

Substituting the functional $I$ for $F$ and the expression for $\lambda(u)$, it follows that

$$\partial_t I(u) = \langle \delta I(u), [\delta H(u) - \lambda(u)\delta I(u)] \rangle = 0$$

Hence, for the dynamics the functional $I$ is conserved, a constant of the motion, a *first integral*:

$$I(u(t)) = I(u(0)) \text{ for all } t.$$

Geometrically this is seen since the vectorfield $\delta H - \lambda \delta I$ is perpendicular to $\delta I$, and so tangent to the level sets of $I$.

[2] From a mathematical point of view, the system can also be considered as a simple example of a dynamical system on a manifold: the level set of the conserved functional as the manifold on which a disspative system is defined.

On each level set of $I$, the system behaves like a dissipative system as treated in the foregoing chapter. In fact, for the evolution of $H$:

$$\partial_t H(u) = \langle \delta H(u), [\delta H(u) - \lambda(u)\delta I(u)] \rangle = \langle [\delta H(u) - \lambda(u)\delta I(u)], [\delta H(u) - \lambda(u)\delta I(u)] \rangle,$$

so

$$\partial_t H(u) \begin{cases} \leq 0 \\ = 0 \text{ iff } \delta H(u) = \lambda \delta I(u) \end{cases} .$$

From this it follows that $H$ decreases monotonically, except from the points that are the equilibria of the system. Indeed, an equilibrium solution satisfies the equation

$$\delta H(\hat{u}) = \lambda \delta I(\hat{u})$$

for some scalar $\lambda$. Recalling Lagrange's multiplier rule, this is the equation for constrained critical points of $H$ on a level set of $I$:

$$\hat{u} \in \text{Crit } \{H(u) \mid I(u) = \gamma\}.$$

From the above observations it is clear that the trajectories are in the direction of *constrained steepest descent*; hence the equation can be used to find constrained minimizers of $H$ on level sets of $I$ in a numerical way.

**Exercise 32** *1.* The dynamic system above provides a way to prove Lagrange's multiplier rule in an alternative way, different from the proof as given before. Give the detailed argumentation.

*2.* Write down the equation of constrained steepest descent to find the solution of

$$Min \{ \int u_x^2 \mid \int u^2 = 1, u(0) = u(\pi) = 0 \},$$

and investigate the convergence to the minimal element.

## 2.7 Hamiltonian and Poisson systems

In this section we introduce a generalization of Hamiltonian systems, socalled Poisson systems; these are systems that have a conservative structure and appear quite regularly is systems form classical mechanics and in in continuous systems from mathematical physics. The structure of the equation implies that there is an integral, the Hamiltonian of the system (which is mosttimes the total energy). Here we investigate only the simplest properties; in the next sections and chapters we give many examples and show that the coherent structures that appear in these systems are a direct, but deeper, consequence of the specific structure.

### 2.7.1 Definition of Poisson systems

**Definition 33** *A* Poisson system *is a dynamical system in state space $\mathcal{U}$ of the form*

$$\partial_t u = \Gamma(u)\delta H(u) \tag{2.16}$$

*with:*

- $H : \mathcal{U} \to I\!R$, *a functional called the* Hamiltonian,

- *for each* $u$, $\Gamma(u) : \mathcal{U}^* \to \mathcal{U}$ *is a linear and skew-symmetric operator* $\langle \zeta, \Gamma(u)\eta\rangle = -\langle \Gamma\zeta, \eta\rangle$,
- $\Gamma$ *is such that the following bracket* $\{\ ,\ \}$ *is a* Poisson-bracket*:*

$$\{F, G\}(u) := \langle \delta F(u), \Gamma(u)\delta G(u)\rangle. \tag{2.17}$$

*This requires in addition to being* bi-linear *and* skew-symmetric*, that* Jacobi's condition *is satisfied, meaning that for any three functionals* $F, G, H$*:*

$$\{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} = 0. \tag{2.18}$$

*Moreover, if the map* $\Gamma$ *is invertible, the Poisson system is called a* Hamiltonian system.

The meaning of the Jacobi condition may not become clear until the next chapter. From an algebraic point of view, it implies that we are dealing with a *Lie-algebra*.
Arbitrary skew-symmetric maps $\Gamma = \Gamma(u)$ that depend on the point under consideration, will not satisfy Jacobi's condition; a condition should be satisfied that is rather awkward to write down. Instead of doing that, we observe the following special property.

**Proposition 34** *If* $\Gamma$ *is a* constant *skew-symmetric map (not depending on* $u \in \mathcal{U}$*), Jacobi's condition is automatically satisfied, and the bracket is a Poisson bracket.*

**Proof.** For arbitrary functionals $F, G$ and $H$ consider $\{\{F, G\}, H\}$ and write $K = \{F, G\} = \langle \delta F, \Gamma\delta G\rangle$. The variational derivative of $K$ is needed and follows from

$$\langle \delta K(u), v\rangle = \langle D^2F(u)v, \Gamma\delta G(u)\rangle + \langle \delta F(u), \Gamma D^2G(u)v\rangle \tag{2.19}$$

Hence

$$\{K, H\}(u) = \langle D^2F(u)\Gamma\delta H, \Gamma\delta G\rangle + \langle \delta F(u), \Gamma D^2G(u)\Gamma\delta H(u)\rangle.$$

Now we use the skew-symmetry of $\Gamma$, and the symmetry of the second derivative of a functional (verify this!!) to rewrite this as

$$\{\{F, G\}, H\} = \langle D^2F(\Gamma\delta H), \Gamma\delta G\rangle - \langle D^2G(\Gamma\delta F), \Gamma\delta H\rangle. \tag{2.20}$$

Adding the expressions obtained by cyclic permutation, Jacobi's identity follows. □

### 2.7.2 Simplest dynamic properties

#### Equilibria

If $\Gamma$ is not degenerate, i.e. for Hamiltonian systems, the only equilibria are elements $\hat{u}$ with

$$\delta H(\hat{u}) = 0, \tag{2.21}$$

i.e. the critical points of $H$ are equilibrium solutions:

$$\hat{u} \in \text{Crit } \{H(u) \mid u \in \mathcal{U}\}. \tag{2.22}$$

Note that these are the same as for the gradient system $\partial_t u = -\delta H(u)$, which equation may be used to construct the special critical points that are the minimizers of $H$.

If $\Gamma$ is degenerate, elements $\bar{u}$ such that $\delta H(\bar{u}) \in \ker(\Gamma(\bar{u}))$ will be other equilibria. We will see examples in th enext section.

**Diagnostics**

Any functional $F$ evolves according to

$$\partial_t F(u) = \{F, H\}(u).$$

In particular, the next result holds.

**Proposition 35** *For the Poisson system $\partial_t u = \Gamma \delta H(u)$, a functional $I$ is a first integral iff $I$ Poisson commutes with $H$, meaning $\{I, H\} = 0$:*

$$\partial_t I(u) = 0 \quad \textit{iff} \quad \{I, H\} = 0.$$

*Since $\{H, H\} = 0$ from skew-symmetry, the Hamiltonian $H$ itself is a first integral:*

$$\partial_t H(u) = 0.$$

### 2.7.3 Canonical Hamiltonian systems

The standard example of a Poisson system is a canonical Hamiltonian system and will be resumed first. Then degenerate Hamiltonian systems are described, after which Darboux' theorem can be formulated and appreciated.
For $\mathcal{M} = I\!R^{2n}, n \in I\!N$, write $u = (q, p) \in I\!R^n \times I\!R^n$. For $F \in C^\infty(\mathcal{M})$ the derivative $dF$ is written like

$$dF(u) = (\partial_q F, \partial_p F)(q, p), \tag{2.23}$$

and can also be viewed as an element from $I\!R^{2n}$:

$$\nabla F = (\partial_q F, \partial_p F) = \begin{pmatrix} \partial_q F \\ \partial_p F \end{pmatrix}.$$

Let $J$ be the mapping such that

$$JdF = (\partial_p F, -\partial_q F). \tag{2.24}$$

In matrix form, $J$ is given by

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \tag{2.25}$$

where $I_n$ is the identity matrix in $I\!R^n$. $J$ is called the standard *symplectic matrix*. It is skew-symmetric and invertible:

$$J^* = -J, \quad J^{-1} = -J, \quad J^2 = -I_{2n}$$

Consider the bracket with $J$ as structure map:

$$\{F, G\}(u) := \langle dF(u), JdG(u) \rangle = \partial_q F \cdot \partial_p G - \partial_p F \cdot \partial_q G. \tag{2.26}$$

In matrix notation this can be written like

$$\{F, G\} = \nabla F \cdot J \nabla G,$$

and in components

$$\{F, G\} = \partial_q F \cdot \partial_p G - \partial_p F \cdot \partial_q G = \sum_{k=1}^{n} \left( \frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial F}{\partial p_k} \frac{\partial G}{\partial q_k} \right)$$

Since $J$ is a constant skew-symmetric matrix, (2.26) is a Poisson bracket. It is called the *canonical Poisson bracket* on $IR^{2n}$, and the variables $q$ and $p$ are a set of so-called *canonically conjugate variables*: $q$ the (generalized) *position variable* and $p$ the (generalized) *momentum variable*. According to the definition, the invertibility of $J$ implies that this Poisson structure is in fact a symplectic structure and that (2.26) defines a Hamiltonian system. The state equation for this bracket is

$$\partial_t u = JdH(u) \equiv X_H(u), \tag{2.27}$$

which, in the canonical variables, reads:

$$\partial_t \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \begin{pmatrix} \partial_q H \\ \partial_p H \end{pmatrix}, \text{ i.e. } \begin{cases} \partial_t q &= \partial_p H \\ \partial_t p &= -\partial_q H \end{cases}$$

These equations are known as the canonical *Hamilton's equations.* In Classical Mechanics the Hamiltonian of is often the total energy of the system.

**Second order Hamiltonian systems**
An important special class of equations is encountered when one considers for instance the motion of $N$ particles in $IR^d, d = 1, 2$ or 3, under the influence of a conservative force. Then the position of the particles is determined by a point $q$ in the so-called *configuration space* $Q = IR^{dN}$. Given the function $V$ on $Q$ as the potential energy which defines the conservative force $-\partial_q V(q)$, *Newton's equations* are given by

$$M\partial_t^2 q = -\partial_q V(q). \tag{2.28}$$

Here $M$ is a constant, positive matrix accounting for the mass of the particles. Since $M$ is invertible, the second order equation can be rewritten as a set of first order equations by introducing the momentum $p$ as a new independent variable:

$$\begin{aligned} \partial_t q &= M^{-1}p \\ \partial_t p &= -\partial_q V(q) \end{aligned} \tag{2.29}$$

With the Hamiltonian defined as

$$H(q,p) := \frac{1}{2}M^{-1}p \cdot p + V(q), \tag{2.30}$$

the equations (2.29) are recognized as Hamilton's equations. This shows that Newton's equations are equivalent to the canonical Hamilton equations on the *phase space* $IR^{dN} \times IR^{dN}$ with $q$ the position-variable, $p$ the momentum variable and $H$, being the sum of kinetic and potential energy, the total energy. A Hamiltonian of this form will be called a *natural Hamiltonian*, and equations of the form (2.28) are sometimes called *second order* Hamiltonian equations.

**Exercise 36** *1.* Consider the motion of one particle in the plane under the influence of some potential force $V$. Write down the equation in Cartesian coordinates.

*2.* Define the transformation $\psi$ from Cartesian to polar coordinates. Rewrite the dynamic equations in polar coordinates directly. Verify the general formula given in Section 3.5.

*3.* Observe that if $V$ is a "central" force, depending on the distance to the origin only, that there is a first integral different from the energy. What is the flow of this integral? Why will it be called angular momentum?

### 2.7.4 Complex canonical structure

Associated to the real canonical structure described by (2.26) there is a natural *complex structure*. For $z \in C^n$, let $\bar{z}$ denotes the complex conjugate:

$$\text{if } \; z = q + ip \in C, \; (q,p) \in IR^n \times IR^n, \quad \text{then } \; \bar{z} = q - ip.$$

The inner product of $z_1$, and $z_2$, with $z_k = q_k + ip_k$, reads

$$\langle z_1, z_2 \rangle_C = Real(z_1 \cdot \bar{z_2}) = q_1 \cdot q_2 + p_1 \cdot p_2$$

where *Real* denotes the real part.
Functions (real valued) $\hat{F}$ on $C^n$ are related to (real valued) functions $F$ on $IR^{2n}$ by

$$\hat{F}(z) = \hat{F}(q + ip) = F(q,p)$$

and for the derivative it holds

$$d\hat{F}(z) = \partial_q F + i\partial_p F.$$

The bracket

$$\{\hat{F}, \hat{G}\}(z) := \langle d\hat{F}(z), -id\hat{G}(z) \rangle_C \tag{2.31}$$

is a Poisson bracket and is naturally related to the real canonical bracket (2.26):

$$\{\hat{F}, \hat{G}\} = \partial_q F \cdot \partial_p G - \partial_p F \cdot \partial_q G = \{F, G\}.$$

The state equation for a Poisson system with Hamiltonian $\hat{H}$ is

$$\partial_t z = -id\hat{H}(z). \tag{2.32}$$

**Exercise 37** *A system of n uncoupled harmonic oscillators with (real) frequencies $\omega_1, \ldots, \omega_n$ is described in complex variables with a Hamiltonian*

$$H(z) = \sum_k \frac{1}{2} \omega_k |z_k|^2, \; z = (z_1, \ldots, z_n) \in C^n$$

*as*

$$\partial_t z_k = -i\omega_k z_k, \quad 1 \leq k \leq n.$$

### 2.7.5 Degenerate Hamiltonian systems

Now let $\bar{\mathcal{M}} = IR^n \times IR^n \times IR^s$ and write $u = (q,p,z) \in IR^n \times IR^n \times IR^s$, and for $F \in C^\infty(\bar{\mathcal{M}})$:

$$\nabla F(u) = (\partial_q F, \partial_p F, \partial_z F).$$

The explicit expression (2.26) defines also a Poisson structure on this space $\bar{\mathcal{M}}$, but now the bracket has a singular structure matrix:

$$\{F, G\}(u) = \nabla F(u) \cdot \bar{J} \nabla G(u) \;\; \text{with} \;\; \bar{J} = \begin{pmatrix} 0 & I_n & 0 \\ -I_n & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.33}$$

The dynamical equations for a system with Hamiltonian $H$ read

$$\begin{cases} \partial_t q &= \partial_p H(q,p,z) \\ \partial_t p &= -\partial_q H(q,p,z) \\ \partial_t z &= 0 \end{cases}. \tag{2.34}$$

The constancy of $z$ implies that the evolution from a given initial value $(q_0, p_0, z_0)$ takes place in the affine subspace $z = z_0$. In that $2n$-dimensional subspace the dynamics for $q$ and $p$ is described by the (non-degenerate) Hamiltonian-system with Hamiltonian $H(q, p, z_0)$ (in which $z_0$ plays the role of a parameter):

$$\begin{cases} z(t) &= z_0 \\ \partial_t q &= \partial_p H(q, p, z_0) \\ \partial_t p &= -\partial_q H(q, p, z_0) \end{cases} \tag{2.35}$$

This may serve as the simplest example of what is called a *Poisson submanifold*. In this simple case it is clear that the original dynamical equations in $IR^{2n+s}$ essentially reduce to a Hamiltonian system in $IR^{2n}$. This reduction from $2n + s$ variables in (2.34) to $2n$ essential variables in (2.35), is possible as a consequence of the fact that the $s$-component variable $z$ is a constant of the motion.
This motivates efforts to look for a similar reduction in the more difficult case when a given system has constants of the motion which are not simply certain components of the state variable $u$.

**Exercise 38** *Let $\Gamma$ be an arbitrary, skew-symmetric, matrix in $IR^m$, $m \geq 2$. Show that $m = 2n + s$ such that after a suitable transformation, $\Gamma$ is given by the (singular, if $s > 0$) structure matrix $\bar{J}$ given in (2.33).*

### 2.7.6 Darboux Theorem

The canonical bracket (2.26) and its degenerate form (2.33) are rather special. In both cases the manifolds are linear spaces, global coordinates are used and the structure map is constant.
In general, the situation may be much more complicated. However, one could hope for such a canonical description, maybe not on all of the manifold but at least locally in the neighbourhood of some point of $\mathcal{M}$. The difficulty may be that the rank of the Poisson structure can vary from point to point. If that is *not* the case the next result states that it is possible to define local coordinates such that the Poisson structure has locally a canonical form. This result makes clear why the word "canonical" has been used for the cases considered before.

**Theorem 39** (Darboux)
*Let $\mathcal{M}$ be a Poisson manifold of dimension $2n + s$ which is of constant rank $2n$ everywhere. Then at each point $u_0 \in \mathcal{M}$ there exists local coordinates $(q, p, z) \in IR^n \times IR^N \times^s$ in terms of which the Poisson bracket takes the canonical form (2.33) and the equations are given by (2.35).*

The proof of this theorem will not be given; see the references.

# Chapter 3

# Fluid dynamics

## 3.1 Finite dimensional systems from fluid dynamics

Many systems from Classical Mechanics can be described both as a Lagrangian system and as a Hamiltonian system: this is the case when a 'Legendre transformation' can be defined; see later. Now we give two examples from fluid dynamics for which the Hamiltonian formulation is the most natural one; a description as a Lagrangian system is not possible, or at least not easy. For instance, the variables used in the Hamiltonian description do not naturally split into pairs of canonical variables and the Hamiltonian is not a Legendre function.

### 3.1.1 Particle motion in plane fluid flows

Consider the two dimensional motion of an incompressible fluid. The velocity field $v : IR^2 \to IR^2, v(x,y) = (v_1(x,y), v_2(x,y))$ is assumed to be given, and time-independent for the moment. Incompressibility implies that $\text{div } v = 0$, and hence that Stokes stream function $\psi$ can be introduced, $\psi : IR^2 \to IR$, such that

$$v(x,y) = J\nabla\psi = \begin{pmatrix} \psi_y \\ -\psi_x \end{pmatrix}, \text{ with } J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{3.1}$$

Hence, *the Eulerian velocity is a Hamiltonian vector field with the stream function $\psi$ as Hamiltonian.* The corresponding set of Hamilton's equations:

$$\partial_t \begin{pmatrix} x \\ y \end{pmatrix} = J\nabla\psi(x,y) \tag{3.2}$$

describes the flow of the *particles* of the fluid: $(x(t), y(t))$ is the position at time $t$ of the particle that occupied the position $(x(0), y(0))$ at $t = 0$. Note that, by its meaning, the time derivative is the so-called *material time-derivative*, i.e. at fixed particle, also called the Lagrangian time-derivative.

For *steady flows* as those considered above, $\psi$ itself is a first integral in the description of the particle motion; this means that the particle trajectories coincide with the level curves of $\psi$.
In general, the flow will not be steady and $\psi$ will depend on time $t$ explicitly. Then, for instationary flows, $v$, and hence $\psi$, depend explicitly on time and provides an example of a non-autonomous Hamiltonian system:

$$\partial_t \begin{pmatrix} x \\ y \end{pmatrix} = J\nabla\psi(x,y,t); \tag{3.3}$$

the Hamiltonian is then no longer a conserved quantity.

**Remark 40** *Systems with both a Hamiltonian and a Gradient structure.*
First, consider the special case that the streamfunction is given by $\psi(x,y) = \frac{1}{2}(x^2 - y^2)$. Define a new function $\phi(x,y) := xy$, and observe that

$$J\nabla\psi(x,y) = -\nabla\phi(x,y).$$

This means that the Hamiltonian structure for the particle dynamics (for which $\psi$ is the Hamiltonian), is at the same time a gradient system:

$$\partial_t \begin{pmatrix} x \\ y \end{pmatrix} = -\nabla\phi(x,y).$$

This is a simple example of a system that has at the same time a Hamiltonian and a gradient structure: the notions of conservative and dissipative structure are clearly not excluding.
For plane flows this is more generally true if (and only if) the streamfunction is a harmonic function, i.e. if $\psi$ satisfies

$$\Delta\psi \equiv \psi_{xx} + \psi_{yy} = 0;$$

then $\psi$ can be interpreted as the real part of a complex analytic function $\psi(x,y)+i\phi(x,y)$ in $z = x+iy$, where $\phi$ is the harmonic conjugate of $\psi$. Then the *Cauchy-Riemann relations* are satisfied by $\psi$ and $\phi$, and these relations can precisely be expressed by $J\nabla\psi(x,y) = -\nabla\phi(x,y)$.
In the following subsection we will see that the necessary condition $\Delta\psi = 0$ precisely means that the flow is irrotational: the vorticity vanishes identically. Hence, *the particle dynamics of an irrotational plane flow has a Hamiltonian and a gradient structure at the same time.*

### 3.1.2 Kirchhoff's equations for vortex points

The *vorticity* $\omega$ of a three-dimensional velocity field $V$ is defined by

$$\omega = \text{curl } V.$$

For a flow in the $(x,y)$-plane (write $x = (x,y)$), the vorticity is is perpendicular to the plane: $\omega = (0,0,\omega)$, and the $z$-component is related to the stream function like

$$\omega(x) = -\Delta\psi(x) \equiv -[\psi_{xx} + \psi_{yy}],$$

where $\Delta$ is the Laplace operator.
A *vortex point* with strength $\gamma_k$ at position $r_k$ is a singular vorticity distribution

$$\omega(x) = \gamma_k \delta(x - r_k)$$

where $\delta$ is Dirac's delta function. The corresponding stream function is the fundamental solution of $-\Delta$:

$$\psi_k(x) = -\frac{1}{2\pi}\gamma_k \log|x - r_k|. \tag{3.4}$$

**Exercise 41** *1.* For a flow in the $(x,y)$-plane with velocity field $v = (y^2, 0)$, determine the stream function $\psi$ and the vorticity $\omega$; deterimine the particle motion.

*2.* A vortex point as above centered in the origin, so with $\omega(x) = -1/2\pi \log|x|$, is often called a *weak vortex*. Determine and sketch the corresponding velocity field.

*3.* A *rigid vortex* centered at the origin, is described by a streamfunction $\omega(x) = -|x|^2$; determine corresponding velocity field and the vorticity. Explain the name rigid vortex.

$N$ vortex points, with strength $\gamma_1, \ldots, \gamma_N$ at places ${}_1 \ldots r_N$, produce a stream function $\sum_k \psi_k(x)$. Neglecting self-interaction, the evolution of the position of the $j$-th vortex-point due to the presence of the other points is governed by (see (3.3))

$$\partial_t r_k = \sum\nolimits_j^* J\partial_{r_k}[-\frac{1}{2\pi}\gamma_j \log|r_j - r_k|], k = 1, \ldots, N. \tag{3.5}$$

(The asterisk in $\Sigma^*$ denotes that the self-interaction term $j = k$ is omitted). These are *Kirchhoff's equations* for the motion of $N$-vortex points.

We will now show that these equations are a Poisson system on

$$\mathcal{M} = \{(r_1, \ldots, r_N) \in IR^{2N} \mid r_j \neq r_k \text{ for } j \neq k\}.$$

Introduce the Hamiltonian

$$H(r_1, \ldots, r_N) = \sum\nolimits_{j,k}^* -\frac{1}{4\pi}\gamma_k\gamma_j \log|r_k - r_j|, \tag{3.6}$$

and for functions of the position ${}_k$ the bracket:

$$\{F, G\} := \sum_k \partial_{r_k} F \cdot \frac{1}{\gamma_k} J\partial_{r_k} G. \tag{3.7}$$

The structure map of this bracket is the $2N \times 2N$ matrix

$$\Gamma = \text{diag}\,(\frac{1}{\gamma_1}J, \ldots, \frac{1}{\gamma_N}J). \tag{3.8}$$

The state equation is the equation for the positions $r_1, \ldots, r_N$ and is given by

$$\partial_t r_k = \frac{1}{\gamma_k} J\partial_{r_k} H(r_1, \ldots, r_N), \quad 1 \leq k \leq N. \tag{3.9}$$

It is readily seen that, with $H$ given by (3.6), the equations (3.9) are precisely the equations (3.5). Hence Kirchhoff's equations are a Poisson system on $\mathcal{M}$ with bracket given by (3.7). In fact, since $\Gamma$ is invertible, it is a Hamiltonian system.

## 3.2 Wave equations

In this section we give several examples of infinite dimensional systems from fluid dynamics, in particular examples of wave equations (in Eulerian description).
Various wave equations in one space dimension are considered. Most of these equations model in some order of accuracy the evolution of surface waves on a layer of fluid, like water.

Two types of wave equations can be distinguished: equations that are of first order in time and equations of second order in time. The simplest characteristic examples are the linear (nondispersive) equations $\partial_t \eta = -c_0 \partial_x \eta$ describing waves running in one direction (of increasing $x$-values) with velocity $c_0$, and $\partial_t^2 \eta = c_0^2 \partial_x^2 \eta$ describing combinations of waves running with velocity $c_0$ in both directions. For that reason these equations are the simplest representations of what are called “uni-directional"

and “bi-directional" wave equations respectively. More general equations incorporate effects of dispersion and non-linearity; Korteweg-de Vries and Boussinesq equation are typical representatives.

In order to assure that the variational formulation of the equations can be nicely described, boundary conditions deserve some special attention. This becomes clear from the partial integrations that have to be performed at several places, most notably in the definition of variational derivative. Two physically important cases will be considered, each of which is such that all contributions from boundary effects vanish identically.
The two cases to be considered are waves on all of $IR$ and periodic waves (which can be interpreted as functions on the unit sphere $S^1$). In the first case the state space is

$$\mathcal{U} = C_0^\infty(IR),$$

where $\eta \in C_0^\infty$ are the test functions on $IR$, i.e. the infinitely differentiable functions $\eta$ that have compact support (vanish, together with all derivatives, sufficiently fast for $|x| \to \infty$). In the second case the state space is

$$\mathcal{U}_{per} = \left\{ \eta \in C^\infty(S^1 \to IR) \;\middle|\; \int \eta = 0 \right\}.$$

In both cases, partial integrations in density functionals will not produce boundary contributions. The normalisation $\int \eta = 0$ in the periodic case excludes (as in the other case) non-zero constant functions. We will employ the notation $\langle f, g \rangle$ to denote the $L_2$-innerproduct (over $\mathcal{R}$ and $S^1$ respectively) of functions $f$ and $g$. With respect to this innerproduct, the differential operator $\partial_x$ is skew-symmetric:

$$\langle f, \partial_x g \rangle = \langle -\partial_x f, g \rangle. \tag{3.10}$$

Note also that $\partial_x$ is invertible:

$$\partial_x^{-1} f = \int_{-\infty}^{x} f(\xi) d\xi, \text{ resp } .\partial_x^{-1} f = \int_0^x f(\xi) d\xi,$$

the last primitive being well defined because of the assumption $\int f = 0$.

### 3.2.1 Uni-directional wave equations

Consider for scalar functions $\eta \in \mathcal{U}$, or $\eta \in \mathcal{U}_{per}$, and a given functional $H(\eta)$ the evolution equation

$$\partial_t \eta = -c_0 \partial_x \delta H(\eta), \tag{3.11}$$

where $c_0$ is a constant (that can be taken equal to one by time-scaling). The functional $H(\eta)$ is called the Hamiltonian. The operator $-c_0\partial_x$ generalises the symplectic matrix, which motivates to call an equation of this kind a ‘generalised’ Hamiltonian system (or Poisson system). The skew -symmetry of the operator $-c_0\partial_x$ guarantees that the Hamiltonian is conserved, a constant of the motion.
Several specific examples will now be given.

1. For the *mass functional* $M(\eta) = \int \eta$ it holds that $\delta M(\eta) = 1$ (which does *not* belong to $U$). Consequently, with $M$ as the Hamiltonian, the state equation: $\partial_t \eta = 0$ leads to a trivial flow $\Phi^M$:

$$M(\eta) = \int \eta \quad \Rightarrow \quad \Phi_t^M(\eta) = \eta \;\text{ for all } t. \tag{3.12}$$

2. The momentum functional $I(\eta) = \frac{1}{2}\int \eta^2$ has $\delta I(\eta) = \eta$, and the state equation reads

$$\partial_t \eta = -c_0 \partial_x \delta I(\eta) = -c_0 \partial_x \eta. \tag{3.13}$$

This equation can be solved easily; the flow is nothing but a translation with fixed speed $c_0$:

$$I(\eta) = \frac{1}{2}\int \eta^2 \quad \Rightarrow \quad \Phi_t^I(\eta)(x) = \eta(x - c_0 t). \tag{3.14}$$

This equation can therefore be called the (uni-directional) *translation equation*. This simple example will play an important role later in analyzing the dynamics of more difficult equations.

3. *Linear dispersive equations*
Consider the Hamiltonian

$$H_0(\eta) = \int (\frac{1}{2}\eta^2 + \frac{1}{2}\alpha \eta_x^2)$$

where $\alpha$ is a given constant. Then $\delta H_0(\eta) = \eta - \alpha \eta_{xx}$ and equation (3.11) reads

$$\partial_t \eta = -c_0 \partial_x (\eta - \alpha \eta_{xx}).$$

This is an example of a linear dispersive wave equation: there are monochromatic waves of the form

$$\eta(x,t) = Ae^{i(kx - \omega t)}$$

for arbitrary amplitude $A$, provided the wave-number $k$ and the frequency $\omega$ satisfy the *dispersion relation*:

$$\omega = c_0(k + \alpha k^3).$$

*Dispersion* is the effect that the propagation speed of the monochromatic wave with wave number $k$, given by the so-called *phase velocity*

$$c_{phase} = \frac{\omega(k)}{k},$$

depends explicitly on the value of $k$. In the example, $c_{phase} = c_0(1 + \alpha k^2)$, which means that shorter waves ($|k|$ large) propagate faster than longer waves ($|k|$ small).
For waves on all of $IR$, $k$ can be any real number; for waves on a periodic interval of length $2\pi$, $k$ should be restricted to the integers.

4. *Linear waves above topography*
Introduce the physical variables $g$ for the accelaration of gravity, and $h(x)$ for the depth of the layer over a slowly varying bottom. Uni-directional waves with small amplitude and long wave length are described by an equation of the form

$$\partial_t \eta = -\Gamma(x)\eta$$

where

$$\Gamma(x) = \frac{1}{2}\left[c(x)\partial_x + \partial_x c(x)\right], \text{ with } c(x) = \sqrt{gh(x)}.$$

Verify that this is a Hamiltonian system.
Find the general solution of this equation; show that for solutions

$$\sqrt{c(x)}\eta(x,t) = \text{ constant on characteristics } \frac{dx}{dt} = c(x).$$

Show that for decreasing depth (i.e. a wave running up a shoal) the following characteristic properties are found:
the wave-height increases, the velocity decreases, the wavelength decreases.
The operator $\Gamma(x)$ reduces to the operator $c_0\partial_x$ above a flat bottom at height $h_0$. Show that if $\Gamma(x)$ is replaced by $c(x)\partial_x$, the equation is not Hamiltonian; investigate this equation, and observe that the three properties above do not hold any more.
The same is true if $\Gamma(x)$ is replaced by $\partial_x c(x)$; investigate.

5. *Nonlinear hyperbolic equations*
For a Hamiltonian of the form

$$H_1(\eta) = \int (\frac{1}{2}\eta^2 + N(\eta))$$

with $N : IR \to IR$ a given function, the variational derivative is

$$\delta H_1(\eta) = \eta + n(\eta), \text{ where } n = \frac{dN}{d\eta}$$

and equation eqG1.60) becomes a typical hyperbolic equation. For instance, taking $c_0 = 1$ and $N(\eta) = \frac{1}{6}\eta^3$, leads to the equation

$$\partial_t \eta = -(1+\eta)\partial_x \eta,$$

which is the simplest example of a nonlinear hyperbolic equation. For this equation, depending on the initial condition, the solution may become multivalued and develop into a *shock*.
Determine the solutions explicitly by verifying that

$$\eta(x,t) = \text{ constant on characteristics } \frac{dx}{dt} = 1 + \eta.$$

Sketch the characteristics in the $(x,t)$-plane; first for initial data of step-shape; for instance, $\eta_0 = 0$ for $x < 0$ and $\eta_0 = 1$ for $x > 0$, then for arbitrary initial profile.
Verify that there are no non-constant solutions that travel undisturbed in shape.

6. *KdV-type of equations*
For more general Hamiltonians, the flow of (3.11) cannot be written down explicitly in general. The Korteweg-de Vries equation is of the form (3.11). The equation incorporates both the

dispersive and the nonlinear effects considered above. The Hamiltonian consists of a linear combination of $H_0$and $H_1$. Introducing parameters $\alpha$and $\beta$, $H$is given by

$$H(\eta) = \int (\frac{1}{2}\eta^2 - \frac{1}{2}\alpha\eta_x^2 + \beta\eta^3). \tag{3.15}$$

Then $\delta H(\eta) = \eta + \alpha\eta_{xx} + 3\beta\eta^2$, and (3.11) reads

$$\partial_t\eta = -c_0\partial_x(\eta + \alpha\eta_{xx} + 3\beta\eta^2). \tag{3.16}$$

Introducing a coordinate frame moving with velocity $c_0$, and scaling the $x$and $\eta$variable, this equation can be brought to a standard form for the KdV-equation that reads

$$\partial_t\eta + \eta_{xxx} + 6\eta\eta_x = 0. \tag{3.17}$$

Each term in (3.15) has a specific physical meaning. However, in the following, we will often consider the KdV-equation in its normalised form (3.17), following common use in the literature (sometimes with a factor 2replacing the coefficient 6). It is then often written by introducing an additional minus-sign in the Hamiltonian:

$$\partial_t\eta = \partial_x\delta H(\eta), with H(\eta) = \int (\frac{1}{2}\eta_x^2 - \eta^3). \tag{3.18}$$

Special solutions, cnoidal waves and solitary waves, will be considered later.
The KdV-equation was formulated in 1895 by Korteweg and de Vries as an approximate equation describing the evolution of rather long, rather low waves. Waves of this form had been observed by Scott Russel more than fifty years before, 1845. The KdV-equation is an important example of a *completely integrable system*, one of the first partial differential equations recognized as such in the sixties.

**Remark 42** *Potential representation*
Sometimes the equations above are written with a different dependent variable. Here we will show how the equation transforms under such a transformation of variable. The general theory of transformations could be used; here we derive the result in an ad hoc way.
The new variable to be considered is a "potential-like" function $v$, related to the function $\eta$ by

$$\eta = v_x, \ i.e. \quad v = \partial_x^{-1}\eta \tag{3.19}$$

(an arbitrary constant can be normalised to zero without loss of generality). Then, let $\hat{H}$ be the Hamiltonian functional expressed in the variable $v$. That means that $\hat{H}$ is related to $H$ by

$$\hat{H}(v) = H(\eta). \tag{3.20}$$

For an arbitrary variation $\zeta$ in the potential $v$, it holds

$$\langle\delta_v\hat{H}(v), \zeta\rangle = \langle\delta_\eta H(\eta), \partial_x\zeta\rangle$$

from which it follows that

$$\delta_v\hat{H}(v) = -\partial_x\delta_\eta H(\eta), \ \text{or}\delta_\eta H(\eta) = -\partial_x^{-1}\delta_v\hat{H}(v). \tag{3.21}$$

Substituting this in the bracket (**??**) leads to a Poisson bracket for functionals of $v$:

$$\{\hat{F}, \hat{G}\}(v) \equiv \{F, G\}(v_x) \tag{3.22}$$

$$= \langle\delta_v\hat{F}(v), \partial_x^{-1}\delta_v\hat{G}(v)\rangle. \tag{3.23}$$

The state-equation for $v$ with this bracket reads

$$\partial_t v = \partial_x^{-1}\delta_v \hat{H}(v), \text{ i.e.} \partial_t v_x = \delta_v \hat{H}(v) \tag{3.24}$$

Of course, this result could also have been found by performing the transformation (3.19) directly in the state equation (3.11).

### 3.2.2 Bi-directional wave equations

Equations of second order in time for the wave elevation $\eta$ arise usually from two equations of first order in time for $\eta$ and a velocity-type variable $u$ by eliminating $u$. Therefore we will first consider more general systems of two scalar equations of first order in time.

From a dynamical point of view, first order in time equations provide the most convenient formulation to interpret the evolution from an initial data to the state at later times. The state space then consists of pairs of functions $(u, \eta) \in \mathcal{U} \times \mathcal{U}$. As an important example, consider the special Poisson bracket for functionals $F = F(u, \eta)$ given by

$$\{F, G\} = \langle \begin{pmatrix} \delta_u F \\ \delta_\eta F \end{pmatrix}, \Gamma \begin{pmatrix} \delta_u G \\ \delta_\eta G \end{pmatrix} \rangle \tag{3.25}$$

when for the structure map $\Gamma$ is taken:

$$\Gamma = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & 0 \end{pmatrix}. \tag{3.26}$$

This is a Poisson bracket since $\Gamma$ is skew-symmetric and constant. The bracket can be written like

$$\{F, G\} = \langle \delta_u F, -\partial_x \delta_\eta G \rangle + \langle \delta_\eta F, -\partial_x \delta_u G \rangle, \tag{3.27}$$

and the state equation for a system with Hamiltonian $H$ is given by

$$\partial_t \begin{pmatrix} u \\ \eta \end{pmatrix} = \begin{pmatrix} 0 & -\partial_x \\ -\partial_x & 0 \end{pmatrix} \begin{pmatrix} \delta_u H \\ \delta_\eta H \end{pmatrix}, \tag{3.28}$$

i.e.

$$\begin{cases} \partial_t u &= -\partial_x(\delta_\eta H), \\ \partial_t \eta &= -\partial_x(\delta_u H). \end{cases} \tag{3.29}$$

**Remark 43** *Canonical formulation*
There is a simple relation between the bracket (3.28) and the canonical bracket (**??**). This can be observed by introducing the potential $v$ for $u$ as in (3.19), and writing

$$\hat{F}(v, \eta) = F(u, \eta). \tag{3.30}$$

Then the bracket (3.27) becomes

$$\{\hat{F}, \hat{G}\} = \langle \delta_\eta \hat{F}, \delta_v \hat{G} \rangle - \langle \delta_v \hat{F}, \delta_\eta \hat{G} \rangle \tag{3.31}$$

which leads to the following state equations for $v$ and $\eta$:

$$\begin{cases} \partial_t v &= -\delta_\eta \hat{H} \\ \partial_t \eta &= \delta_v \hat{H}. \end{cases} \tag{3.32}$$

These are just the "canonical" bracket (**??**) and Hamilton's equations with $\eta$ and $v$ as the canonical position and momentum variable respectively. But note, different from the Lagrangian description in Section 5.1, now in the Eulerain description, the time derivative is the usual partial derivative (at fixed spatial position). □

As was done for uni-directional equations, we will consider several specific examples.

1. Both for $M_1(u;\eta) = \int u$ and $M_2(u,\eta) = \int \eta$ as Hamiltonian, the Hamiltonian vectorfield vanishes and the flow is trivial.

2. *Translation equation*
With $I(u,\eta) = \int u\eta$ as momentum functional, it holds that $\delta_u I(u,\eta) = \eta, \delta_\eta I(u,\eta) = u$ and the equations (3.28) become

$$\partial_t u = -\partial_x u,\ \partial_t \eta = -\partial_x \eta.$$

The flow is a translation, as in (3.14):

$$I(u,\eta) = \int u\eta \Rightarrow \Phi_t^I(u,\eta)(x) = (u(x-t), \eta(x-t)). \tag{3.33}$$

3. *Linear waves above topography*
For bi-directional waves above a varying bottom at $h(x)$, the simplest approximation for the energy is taken as the Hamiltonian. It is given by

$$H_0(u,\eta) = \int (\frac{1}{2}h(x)u^2 + \frac{1}{2}g\eta^2).$$

Here $\eta$ is the surface elevation, and $u$ is related to the velocity at the free surface. With this Hamiltonian, the equations (3.28) become

$$\left\{ \begin{array}{rcl} \partial_t u &=& -\partial_x g\eta \\ \partial_t \eta &=& -\partial_x [h(x)u]. \end{array} \right. \tag{3.34}$$

Eliminating one of the variables in favour of the other one, there results a second order wave equation for $u$ and $\eta$:

$$\partial_t^2 u = \partial_x^2[c^2(x)u] \quad \text{and} \quad \partial_t^2 \eta = \partial_x[c^2(x)\partial_x \eta]$$

where we have introduced $c^2(x) = gh(x)$. For waves above an even bottom at depth $h_0$, with $c_0^2 = gh_0$ the standard second order equations for $u$ and $\eta$ arise:

$$\partial_t^2 u = c_0^2 \partial_x^2 u \ and \partial_t^2 \eta = c_0^2 \partial_x^2 \eta.$$

4. *Boussinesq-type of equations*
We consider for simplicity a layer of constant depth, and rescale all physical variables to unity. To account for effects of dispersion and non-linearity, we add some terms to the quadratic Hamiltonian $H_0$ given above. In general the Hamiltonian can then be written like

$$H(u,\eta) = H_0(u,\eta) + h(u,\eta), \tag{3.35}$$

and the equations become

$$\left\{ \begin{array}{rcl} \partial_t u &=& -\partial_x[\eta + \delta_\eta h(u,\eta)] \\ \partial_t \eta &=& -\partial_x[u + \delta_u h(u,\eta)] \end{array} \right. . \tag{3.36}$$

Taking for the additional contribution to $H$ the expression

$$h(u,\eta) = \varepsilon \int (-\frac{1}{2}u_x^2 + \frac{1}{2}u^2\eta) \tag{3.37}$$

with $\varepsilon > 0$ small, leads to a model for surface waves for which dispersion and nonlinearity are present in the same way as in the KdV-equation (3.17). In fact, (3.37) provides equations of Boussinesq-type:

$$\left\{ \begin{array}{lcl} \partial_t u &=& -\partial_x[\eta + \frac{1}{2}\varepsilon u^2] \\ \partial_t \eta &=& -\partial_x[u + \varepsilon(u_{xx} + u\eta)]. \end{array} \right. \tag{3.38}$$

This particular set has the property that it is completely integrable, and as such has been considered in detail.

# Chapter 4

# Poisson dynamics

In this chapter we investigate specific properties of Poisson systems. These properties will be present if there is a first integral $I$ in addition to the Hamiltonian $H$. This additional integral foliates the state space by its level sets, and defines (in general) a symmetry group. As a consequence, equilibria that are critical points of the Hamiltonian can be generalized to the corresponding elements on the level sets of $I$, and define relative equilibrium solutions by translation with the symmetry flow.
We consider first a special simple example in detail, and then generalize some aspects. The applications in the next Chapter will show that the mathematical relative equilibria are often the "coherent structures" known in physics.

## 4.1 Hamiltonian systems with a cyclic variable

In order to illustrate the concepts and methods of the next sections, we consider the following simple system quite extensively.
For a classical Hamilton system with $(n+1)$ degrees of freedom, let

$$u = (q_0, q_1, \ldots, q_n,\ p_0, p_1, \ldots, p_n) \in I\!R^{2(n+1)}$$

be the canonical variables, and $H$ the Hamiltonian:

$$\partial_t u = J\ \partial_u H(u), \tag{4.1}$$

where $J$ is the standard symplectic matrix in $I\!R^{2(n+1)}$. For the following we will assume that the variable $q_0$ is a *cyclic variable*, meaning that $H$ does not depend on $q_0$:

$$\frac{\partial H}{\partial q_0} = 0. \tag{4.2}$$

The equation for the canonically conjugate variable $p_0$, i.e. $\partial_t p_0 = -(\partial H/\partial q_0)$, then shows that $p_0$ is constant during the evolution:

$$\partial_t p_0 = 0.$$

A more transparent formulation is obtained by adjusting notation. Write

$$u = (\gamma, \varphi; x) \ \text{ with } \ \varphi = q_0,\ \gamma = p_0 \ \text{ and } \ x = (q_1, \ldots, q_n;\ p_1, \ldots, p_n) \in I\!R^{2n}.$$

Using the notation

$$\hat{H}(\gamma; x) = H(u), \tag{4.3}$$

(4.1) can be written like

$$\begin{cases} \partial_t \gamma &= 0 \\ \partial_t \varphi &= \partial_\gamma \hat{H}(\gamma; x) \\ \partial_t x &= J \partial_x \hat{H}(\gamma; x) \end{cases} \tag{4.4}$$

where now $J$ is the symplectic matrix in $IR^{2n}$. The original integral $p_0$ can be written like

$$I(u) = p_0. \tag{4.5}$$

### 4.1.1 Equilibrium solutions

Time-independent solutions of (4.1) are precisely the critical points $\bar{u}$ of $H$:

$$\partial_u H(\bar{u}) = 0. \tag{4.6}$$

As a consequence of (4.2), each critical point (if any) is degenerate. This can be restated as follows: for any $\varphi \in IR$, $\bar{u} = (\bar{\gamma}, \varphi, \bar{x})$ is an equilibrium solution of (4.1) provided $(\bar{\gamma}, \bar{x})$ is a critical point of $\hat{H}(\gamma; x)$:

$$\partial_\gamma \hat{H}(\bar{\gamma}; \bar{x}) = 0 \ \text{ and } \ \partial_x \hat{H}(\bar{\gamma}, \bar{x}) = 0. \tag{4.7}$$

We will write

$$\bar{u}(\varphi) = (\bar{\gamma}, \varphi, \bar{x}) \in \text{Crit }_u H(u). \tag{4.8}$$

The fact that with each equilibrium there corresponds a continuous branch of equilibria $\varphi \mapsto \bar{u}(\varphi)$, a simple consequence of the property (4.2), is a first example of degeneracy that is characteristic for a system with a continuous symmetry.

### 4.1.2 Reduced dynamics

From a second glance at the equations (4.4) it may be noticed that the solution procedure for the initial value problem can proceed along the following lines. Given initial data $(\gamma_0, \varphi_0, x_0)$ at $t = 0$, it follows from $\partial_t \gamma = 0$ that $\gamma(t) = \gamma_0$ for all $t$. Then the equation for $x$ becomes

$$\partial_t x = J \partial_x \hat{H}(\gamma_0; x). \tag{4.9}$$

This is a Hamilton system in $x$, with a Hamiltonian $x \mapsto \hat{H}(\gamma_0; x)$ that depends on $\gamma_0$ as a fixed parameter. Since the solution of (4.9) can be sought without knowing the evolution of $\varphi$, the system (4.9) is called the *reduced dynamics*.
Assuming that the solution of (4.9) with $x(0) = x_0$ has been found, the evolution of $\varphi$ follows by a direct integration:

$$\varphi(t) = \varphi_0 + \int_0^t \partial_\gamma \hat{H}(\gamma; x(s)) ds. \tag{4.10}$$

In conclusion one can say that the original problem in $IR^{2(n+1)}$ reduces essentially to the problem (4.9), i.e. to a problem in $IR^{2n}$. This illustrates the more general fact to be proved lateron that for a Hamilton system the knowledge of one additional first integral (like $I$ here) reduces the system with *two* variables.

### 4.1.3 Equilibria of the reduced dynamics

For given $\gamma_0$, the Hamiltonian $x \mapsto \hat{H}(\gamma_0; x)$ may have none, one or more critical points. Assume that $x_0$ is a critical point, i.e. $x_0$ satisfies

$$\partial_x \hat{H}(\gamma_0; x_0) = 0. \tag{4.11}$$

Assume, moreover, that the following non-degeneracy condition is satisfied: the second derivative of $\hat{H}$ with respect to $x$, i.e. the $n \times n$ Hessian matrix

$$D_x^2 \hat{H}(\gamma_0; x_0), \tag{4.12}$$

is invertible. Then the Implicit Functions Theorem guarantees that in a full neighbourhood of $\gamma_0$ there exists a smooth branch of critical points $\hat{x}(\gamma)$:

$$\gamma \mapsto \hat{x}(\gamma) \quad \text{with} \quad \partial_x \hat{H}(\gamma; \hat{x}(\gamma)) = 0, \tag{4.13}$$

which branch is unique in a neighbourhood of $x_0$ and satisfies $\hat{x}(\gamma_0) = x_0$. Assuming the existence of such a smooth branch, consider the *value function*, defined as

$$h(\gamma) \ := \ \hat{H}(\gamma; \hat{x}(\gamma)). \tag{4.14}$$

From

$$\frac{dh}{d\gamma} = \partial_\gamma \hat{H} + \partial_x \hat{H} \circ \frac{d\hat{x}}{d\gamma}$$

and the critical point condition (4.11) it follows that

$$\lambda(\gamma) := \frac{dh}{d\gamma}(\gamma) = \partial_\gamma \hat{H}(\gamma; \hat{x}(\gamma)) \tag{4.15}$$

where $\lambda(\gamma)$ is the slope of the tangent to the value function. This value $\lambda$ will be equal to the Lagrange multiplier of the constrained critical point problem, as we shall see below.

### 4.1.4 Relative Equilibria

Dynamically spoken, the critical points of $\hat{H}(\gamma; \cdot)$ are precisely the equilibrium solutions of the reduced dynamics (4.9). Such an equilibrium $\hat{x}(\gamma)$ of the reduced dynamics leads one to consider the corresponding vectors in the original space:

$$\hat{u}(\gamma; \varphi) = (\gamma, \varphi; \hat{x}(\gamma)) \in I\!R^{2(n+1)}. \tag{4.16}$$

In view of generalizations lateron, it is illustrative to show the relevance of these elements $\hat{u}$ directly in terms of the functions $H$ and $I$. In fact, $\hat{u}$ is for each $\varphi$ a *constrained critical point* of $H$, with $H$ constrained to a specified levelset of $I$:

$$\hat{u}(\gamma, \varphi) \in \text{Crit}_{\ u}\{H(u) | I(u) = \gamma\}. \tag{4.17}$$

In particular, if $\hat{x}(\gamma)$ minimizes $x \mapsto \hat{H}(\gamma, x)$, then $\hat{u}$ minimizes $H$ on $I^{-1}(\gamma)$:

$$\hat{u}(\gamma, \varphi) \in \text{Min } \{H(u) \mid I(u) = \gamma\}. \tag{4.18}$$

Just as for the unconstrained critical points of $H$ - the equilibrium solutions of (4.1) - these constrained critical points are degenerate and appear as one parameter families, with $\varphi$ the parameter.
For the value at a critical point, it holds

$$h(\gamma) = H(\hat{u}(\gamma, \varphi)), \tag{4.19}$$

where $h$ is the value function defined in (4.14). The derivative of this function, $\lambda(\gamma) := (dh/d\gamma)(\gamma)$ can now be related to the Lagrange multiplier. To that end, note that the constrained critical point $\hat{u}$ satisfies according to *Lagrange's multiplier rule* the equation $\partial_u H(\hat{u}) = \sigma \partial_u I(\hat{u})$ for some Lagrange multiplier $\sigma \in IR$. In fact, the multiplier $\sigma$ is precisely the value $\lambda$, as is seen from

$$\partial_u H(\hat{u}) = \lambda \partial_u I(\hat{u}) \Longleftrightarrow \begin{cases} \partial_\gamma \hat{H}(\gamma, \hat{x}(\gamma)) = \lambda(\gamma) \\ \partial_\varphi \hat{H}(\gamma, \hat{x}(\gamma)) = 0 \\ \partial_x \hat{H}(\gamma, \hat{x}(\gamma)) = 0 \end{cases}$$

Indeed, these equations are recognized as the definition of $\lambda$, the cyclic-condition (4.2), and the equation that specifies $\hat{x}(\gamma)$ to be a critical point of $x \mapsto \hat{H}(\gamma; x)$, respectively.
Resuming these critical point considerations, we have:

$$\hat{u}(\gamma, \varphi) \in \text{Crit}_{\ u}\{H(u) \mid I(u) = \gamma\} \tag{4.20}$$

satisfies

$$\partial_u H(\hat{u}) = \lambda \partial_u I(\hat{u}) \tag{4.21}$$

where

$$\lambda(\gamma) = \frac{dh}{d\gamma}(\gamma), \quad \text{with} \quad h(\gamma) = H(\hat{u}(\gamma, \varphi)). \tag{4.22}$$

### 4.1.5 Relative equilibrium solutions

Concerning the dynamics, since $\lambda(\gamma) = \partial_\gamma \hat{H}(\gamma; \hat{x}(\gamma))$ is constant, it follows from (4.4) that $\varphi$ evolves linearly in time, and that for arbitrary initial value $\varphi_0$,

$$t \mapsto \hat{u}(\gamma; \varphi_0 + \lambda(\gamma)t) \tag{4.23}$$

is a special solution of (4.1). A solutions of this kind is called a *relative equilibrium solution.*
When $\lambda(\gamma) \neq 0$, it is a time-dependent solution: a translation with speed $\lambda(\gamma)$ along the one parameter family of constrained critical points $\{\hat{u}(\gamma; \varphi) \mid \varphi \in IR\}$. This can be denoted in a different way as follows. Let $\Phi_t^H(u_0)$ denote the solution of (4.1) at time $t$ which has initial value $u_0$. Besides $\Phi_t^H$, consider also the $I$-flow, i.e. $\Phi_\tau^I$ for the equation (4.1) with $H$ replaced by $I$:

$$\partial_\tau u = J \partial_u I(u). \tag{4.24}$$

Since, in the $(\gamma, \varphi; x)$-variables, $\partial_u I(u) = (1, 0; 0) \in IR \times IR \times IR^{2n}$, the equations (4.24) are written in full

$$\begin{cases} \partial_t \gamma = 0 \\ \partial_t \varphi = 1 \\ \partial_t x = 0 \end{cases},$$

and so the $I$-flow is just the translation along the $\varphi$-variable:

$$\Phi_\tau^I(\gamma_0, \varphi_0; x_0) = (\gamma_0, \varphi_0 + \tau; x_0). \tag{4.25}$$

Using this $I$-flow, the degeneracy of the critical points and the constrained critical points of $H$ can be written like (see (4.16) and (4.18))

$$\bar{u}(\varphi) = \Phi_\varphi^I(\bar{u}) \quad \text{and} \quad \hat{u}(\gamma, \varphi) = \Phi_\varphi^I(\hat{u}(\gamma)) \tag{4.26}$$

respectively, where $\bar{u} = (\bar{\gamma}, 0, \bar{x})$ and $\hat{u}(\gamma) = (\gamma, 0, \hat{x}(\gamma))$ are representatives of a critical point of $H$ and of a constrained critical point of $H$ on the $\gamma$-level of $I$ respectively. The relative equilibrium solution (4.23) can then be described like

$$\hat{u}(\gamma; \varphi_0 + \lambda(\gamma)t) = \Phi^I_{\varphi(t)}(\hat{u}(\gamma)), \quad \text{with} \quad \varphi(t) = \varphi_0 + \lambda(\gamma)t. \tag{4.27}$$

More generally one can observe that the flows $\Phi^I$ and $\Phi^H$ commute:

$$\Phi^H_t \Phi^I_s = \Phi^I_s \Phi^H_t \quad \text{for all} \quad s \quad \text{and} \quad t.$$

This can be used to write a solution $u(t) = (\gamma, \varphi; x)$ as

$$u(t) = \Phi^I_{\varphi(t)}(v(t)),$$

where $v = (\gamma, \psi; x)$.
Then the evolution for $\psi$ is determined by $\partial_t \psi = \partial_t \varphi - \partial_t \alpha$, and so theoretically if $\alpha$ is choosen such that $\partial_t \psi = 0$, $v$ represents the solution $u$ apart from a time-dependent translation along the $\varphi$ variable. Stated differently, $v$ remains at the same $\varphi$-value. This explains the usual saying that in the dynamics for $v$ the $I$-flow has been divided out. A same procedure will be used in Section **??** in a more general setting.

### 4.1.6 Stability

The simple example under consideration is also well suited to reflect upon the stability of relative equilibria.
First consider the reduced dynamics (4.9). As is well known, the stability of an equilibrium $\hat{x}(\gamma)$ for the reduced dynamics may in some cases be related to the character of $\hat{x}(\gamma)$ as a critical point of $\hat{H}(\gamma; \cdot)$. In fact, if $x_0 = \hat{x}(\gamma_0)$ is a strict (nondegenerate) local extremum of $\hat{H}(\gamma_0; x)$, maximal or minimal, then $x_0$ is a centre for (4.9) and Lyapunov stable. This stability of the relative equilibrium is called *reduced stability*.

In the case that the *Hessian matrix* $D^2_x \hat{H}(\gamma_0, x_0)$, which is non-negative definite for a minimizer $x_0$, is in fact positive definite, $D^2_x \hat{H}(\gamma; \hat{x}(\gamma))$ will, by continuity, also be positive definite for all $\gamma$ in a small neighbourhood of $\gamma_0$. Then for those values of $\gamma$, each point on the continuous branch $\gamma \mapsto \hat{x}(\gamma)$ will be a minimizer for $\hat{H}(\gamma; \cdot)$. Consequently, in a small neighbourhood of $\gamma_0$, all points $\hat{x}(\gamma)$ will be a centre and Lyapunov stable for the reduced dynamics (reduced stable).
It must be observed that for this analysis the dependence of $\hat{H}$ on $\gamma$ does not play any role: the second derivative is only with respect to the components of $x$, not with respect to $\gamma$ (nor with respect to $\varphi$, which is absent). To emphasize this more, consider the second derivative with respect to all variables. This is the $2(n+1)$ square matrix

$$D^2_{(\gamma,\varphi,x)} \hat{H} = \begin{pmatrix} \partial^2_\gamma \hat{H} & 0 & D^2_{(\gamma,x)} \hat{H} \\ 0 & 0 & 0 \\ D^2_{(x,\gamma)} \hat{H} & 0 & D^2_x \hat{H} \end{pmatrix}. \tag{4.28}$$

Clearly this matrix is singular as a consequence of the fact that the coordinate $\varphi$ is cyclic: the matrix is singular in the direction $col\ (0, 1, 0)$ of the $I$-flow.
Therefore consider the square $(2n+1)$ matrix

$$D^2_{(\gamma,x)} \hat{H} = \begin{pmatrix} \partial^2_\gamma \hat{H} & D^2_{(\gamma,x)} \hat{H} \\ D^2_{(x,\gamma)} \hat{H} & D^2_x \hat{H} \end{pmatrix}. \tag{4.29}$$

But also this matrix does not have to be positive definite, even if that is the case for $D_x^2\hat{H}$. In fact, if $D_x^2\hat{H}$ is positive definite, the sign of $\partial_\gamma^2\hat{H}$ is the determining quantity. If $\partial_\gamma^2\hat{H} < 0$, which means that (locally) the value function is *concave*, there is a negative eigenvalue. If $\partial_\gamma^2\hat{H} > 0$, in which case the value function is (locally) *convex*, all eigenvalues are positive, and matrix (4.29) is positive definite.
However, since the dynamics will remain on level sets of $I$, the stability of a relative equilibrium will not be affected by the dependence on $\gamma$. However, the flow along the cyclic coordinate will in general influence the stability properties. Indeed, in general $\lambda$ will change with $\gamma$ : $(d\lambda/d\gamma) \neq 0$ near $\gamma_0$. This implies that even if $\hat{x}(\gamma)$ is a stable solution of the reduced dynamics, the solution (4.23) will not be Lyapunov stable for (4.4), since an arbitrary small change in $\gamma$ will produce a deviation in the $\varphi-$variable that increases linearly in time. This effect is the same as in a one-degree of freedom system with a family of periodic solutions with continuously varying period. In that case the appropriate notion of stability is that of *orbital stability*. For the case considered here, *relative stability* is another name that is in use.

**Exercise 44** A simple example of the foregoing is the spherical pendulum, a mass point $m$ connected with a massless inextensible wire of length $\ell$ to a fixed point (that is taken as the origin of a coordinate system). Introducing spherical coordinates, with $\theta$ the angle from the $z$-axis pointing in the direction of the gravitational force, the Lagrangian is the difference of kinetic and potential energy. Up to a multiplicative factor $m\ell^2$, and with $\omega_0^2 = g/\ell$, it is given by

$$L(\varphi,\theta;\dot{\varphi},\dot{\theta}) = \frac{1}{2}(\dot{\varphi}^2\sin^2\theta + \dot{\theta}^2) - \omega_0^2(1-\cos:\theta).$$

With momentum variables canonically conjugate to $\varphi$ and $\theta$:

$$p_\varphi = \dot{\varphi}\sin^2\theta,$$

$$p_\theta = \dot{\theta}$$

, the Hamiltonian reads

$$H(u) = \frac{1}{2}(p_\theta^2 + \frac{p_\varphi^2}{\sin^2\theta}) + \omega_0^2(1-\cos:\theta)$$

where $u = (\varphi,\theta;p_\varphi,p_\theta)$. For this case, the $\varphi-$variable is cyclic and $p_\varphi$ is a first integral: the angular momentum. Writing $I(u) = p_\varphi \equiv \gamma$ and $x = (\theta, p_\theta)$,

$$\hat{H}(\gamma;\theta,p_\theta) = \frac{1}{2}(p_\theta^2 + \frac{\gamma^2}{\sin^2\theta}) + \omega_0^2(1-\cos:\theta).$$

An equilibrium solution corresponds to a critical point of $H$ and is the equilibrium state in the vertical position downwards or upwards: $\partial_u H(\bar{u}) = 0$ for $\gamma = 0,\ \theta = 0$ or $\theta = \pi$.
*Constrained critical points* of $H$ at given $\gamma$ follow from $\partial_u H = \lambda\partial_u I$:

$$\partial_\gamma\hat{H} = \lambda \quad \Rightarrow \quad \gamma = \lambda\sin^2\theta$$

$$\partial_x\hat{H} = 0 \quad \Rightarrow \quad p_\theta = 0 \ \ \textit{and}\ \gamma^2 = \frac{\sin^4\theta}{\cos:\theta},$$

and $\varphi$ arbitrary. For given $\gamma$ the last equation defines a unique angle $\bar{\theta}(\gamma) \in [0,\pi/2)$. The $I$-flow is a rotation about the vertical axis, and so the corresponding *relative equilibrium solution* is a rotation of the mass in the horizontal plane $z = -\cos:\bar{\theta} < 0$ with angular velocity $\lambda = \gamma/\sin^2\bar{\theta}$.

## 4.2 Relative equilibria

In this section we consider the general case of a Poisson system that has at least one additional integral. It is shown that the results of the foregoing section can be generalized to find special solutions that are the relative equilibria.

The standard notation will be employed: for the state equation

$$\partial_t u = \Gamma(u)\delta H(u) \equiv X_H(u),\ u \in \mathcal{M}, \tag{4.30}$$

and for the corresponding Hamiltonian $H$-flow: $\Phi_t^H$.
Besides this equation, we also have to consider for other functionals $I$ the state equation (replacing the time-variable $t$ by $\varphi$ for clarity):

$$\partial_\varphi u = \Gamma(u)\delta I(u) \equiv X_I(u) \tag{4.31}$$

and we will denote the $I$-flow consistently by $\Phi_\varphi^I$.
The *fixed point set* of this flow consists of the elements $\bar{u}$ for which

$$\Phi_\varphi^I(\bar{u}) = \bar{u} \ \text{ for all } \ \varphi.$$

In particular, critical points of $I$ are in the fixed point space of $\Phi^I$. If the structure map $\Gamma$ is invertible (i.e. when dealing with Hamilton systems), these are in fact the only elements in the fixed point set.

### 4.2.1 First integrals and Casimir functionals

**Definition 45** *A functional $I : \mathcal{M} \to IR$ is called a* first integral*, or a* constant of the motion*, for (4.30), if its value remains constant in time on solutions: for each initial data $u_0$:*

$$I(\Phi_t^H(u_0)) = I(u_0) \ \textit{ for all } \ t. \tag{4.32}$$

*A functional $C : \mathcal{M} \to IR$ is called a* Casimir functional *if $C$ is constant on any Hamiltonian flow: (4.32) holds for each Hamiltonian $H \in \mathbf{F}(\mathcal{M})$.*

As stated in the introduction to this chapter, the existence of a first integral $I$ implies that the state space is foliated by the level sets of $I$ which are invariant for the dynamics. (In the foregoing section these level sets are just the vectors with constant coordinate $\gamma$.)
From the fact that the Poisson dynamics is given by the evolution of functionals according to

$$\partial_t F(u) = \{F, H\}(u),$$

the next result follows easily.

**Proposition 46** *A functional $I : \mathcal{M} \to IR$ is an integral for (4.30) iff $I$ and $H$ satisfy*

$$\{I, H\} = 0. \tag{4.33}$$

Functionals that satisfy this relation are said to *Poisson commute*, or that they are *functionals in involution*.
Since a Casimir functional is a special integral, it certainly satisfies this property. But there is more.

**Proposition 47** *The functional $C$ is a Casimir functional iff $C$ Poisson commutes with each functional $F \in \mathbf{F}(\mathcal{M})$:*

$$\{F, C\} = 0 \quad \textit{for all} \quad F \in \mathbf{F}(\mathcal{M}). \tag{4.34}$$

*This is equivalent (if $\mathbf{F}(\mathcal{M})$ is complete) with the statement that the structure map annihilates the variational derivative of $C$:*

$$\Gamma(u)\delta C(u) = 0 \quad \textit{for all } u \in \mathcal{M}. \tag{4.35}$$

*The Hamilton flow of a Casimir functional is trivial:*

$$\Phi^C_\varphi = Identity \quad \textit{on } \mathcal{M}. \tag{4.36}$$

**Proof.** For a Casimir functional $C$, the necessary and sufficient condition (4.34) follows from the fact that $\partial_t C(u) = \{C, H\}(u) = 0$ for all Hamiltonians $H$. Then, from (4.34), $\langle \delta F(u), X_C(u)\rangle = 0$ for all $F$, and from completeness, $X_C(u) = 0$ for all $u$, i.e. (4.35). Then (4.36) follows directly. □

**Remark 48** Property (4.32) expresses the fact that the functional $I$ is invariant for the Hamilton $H$-flow. Conversely, it follows from the (skew-) symmetry in the condition (4.33) that if $I$ is an integral for (4.30), then $H$ is invariant for the $I$-flow:

$$H(\Phi^I_\varphi(u)) = H(u) \quad \textit{for all} \quad u \in \mathcal{M}. \tag{4.37}$$

Stated differently, $H$ is also a first integral for the dynamics (4.31). In the next section we will show that the equation (4.30) itself is "invariant" (more precisely: equivariant) under the $I$-flow, which is expressed by saying that $\Phi^I$ is a symmetry of (4.30). In the foregoing section the symmetry is just the flow along the cyclic coordinate. □

**Remark 49** The set of integrals for (4.30) forms an algebra: linear combinations, and products of integrals are also integrals for (4.30). Moreover, if $I_1$ and $I_2$ are any two integrals, then $\{I_1, I_2\}$ is also an integral:

$$\{I_1, H\} = \{I_2, H\} = 0 \Rightarrow \{\{I_1, I_2\}, H\} = 0. \tag{4.38}$$

This is a simple consequence of Jacobi's identity. If $\{I_1, I_2\}$ is functionally independent of $I_1$ and $I_2$ this provides a new integral and the process may be continued. In this way, the set of all integrals for (4.30) is closed for the Poisson-bracket operation. This shows that the set of all integrals of (4.30) is a *Lie-algebra*.

### 4.2.2 Constrained Hamiltonian extremizers

We first consider the case of a Poisson system with one additional integral. A direct proof of the main results will be given here in a simple way, without exploiting the commutativity property of the corresponding Hamiltonian flows (which will be treated in the next section).
Suppose that $H$ and $I$ are functionals that are functionally independent and in involution: $\{I, H\} = 0$. Then we consider the constrained Hamiltonian extremizers: the constrained critical points of $H$ on a $\gamma$-level set of $I$. Suppose that this variational problem has a solution, say $U(\gamma)$. We will denote this in the following by writing

$$U(\gamma) \in \text{Crit}_{\,u}\{H(u) \mid I(u) = \gamma\},$$

or shorter, with $I^{-1}(\gamma)$ the $\gamma$-level set of $I$:

$$U(\gamma) \in \text{Crit } \{H \mid I^{-1}(\gamma)\}. \tag{4.39}$$

Furthermore assume that $U(\gamma)$ is a *regular point* of $I^{-1}(\gamma)$ in the sense that $\delta I(U(\gamma)) \neq 0$. This assumption implies that the tangent space to the level set consists precisely of those elements $\eta$ for which the first variation vanishes:

$$T_U I^{-1}(\gamma) = \{\eta \mid \langle \delta I(U), \eta \rangle = 0\}.$$

For the critical point $U$ of $H$ on $I^{-1}(\gamma)$ the first variation of $H$ vanishes in all those tangent directions: $\langle \delta H(U), \eta \rangle = 0$. This leads to the next result which is Lagrange's multiplier rule.

**Proposition 50** Lagrange's multiplier rule.
*The constrained critical point $U(\gamma)$ of (4.39) satisfies for some* Lagrange multiplier $\lambda$ *the equation*

$$\delta H(U) = \lambda\, \delta I(U). \tag{4.40}$$

*This multiplier $\lambda$ depends on the critical point $U(\gamma)$.*

In the following we will often deal with branches of critical points $U(\gamma)$, parameterized by $\gamma$. On such a branch the multiplier is a function of $\gamma$, which is then denoted by $\lambda = \lambda(\gamma)$.
The invariance of both $H$ and $I$ for the $I-$flow may imply that a constrained critical point is not isolated, as expressed in the next result; see Fig. 4.1

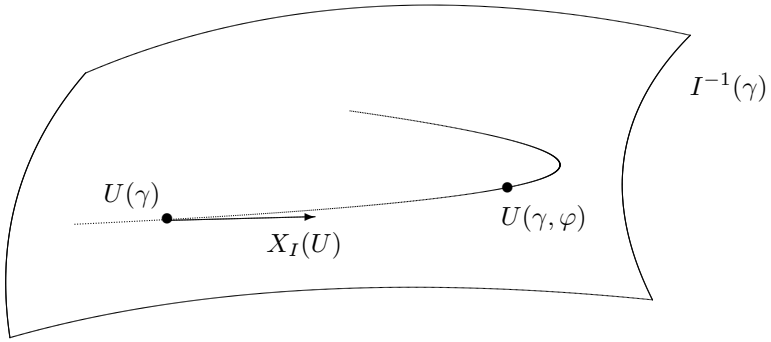


Figure 4.1: In a level set of $I$, the $I$-flow applied to a constrained critical point $U(\gamma)$ leads to a branch of constrained critical points $U(\gamma,\varphi) = \Phi^I_\varphi(U(\gamma))$; see Proposition 51. The tangent vector $X_I(U) = \Gamma(U)\delta I(U)$ to the $I$-flow is shown. This branch is also the trajectory of the relative equilibrium solution: the solution of the Poisson system with $U(\gamma)$ as initial value that transvers this branch with a speed determined by the multiplier $\lambda : \Phi^H_t(U(\gamma)) = \Phi^I_{\lambda t}(U(\gamma))$; see Proposition 54.

**Proposition 51** *Let $U(\gamma)$ be a constrained critical point of (4.39). Assume that $U(\gamma)$ is not in the fixed point set of the Hamiltonian $I$-flow $\Phi^I$. (In particular, $I$ is not a Casimir functional.) Then, for each $\varphi \in IR$, also the functions*

$$U(\gamma,\varphi) := \Phi^I_\varphi(U(\gamma)) \tag{4.41}$$

*are constrained critical points:*

$$U(\gamma,\varphi) \in Crit\ \{H \mid I^{-1}(\gamma)\}. \tag{4.42}$$

*Together they form a one-parameter branch of critical points, the $I$-flow of one (any) representative. Each $U(\gamma,\varphi)$ satisfies for the same multiplier $\lambda$ the same equation as $U(\gamma)$: $\delta H(U) = \lambda\, \delta I(U)$.*

**Proof.** Since the constrained critical point $U(\gamma)$ satisfies

$$\delta H(U) - \lambda \delta I(U) = 0,$$

for the multiplier $\lambda \in IR$, $U(\gamma)$ can also be viewed as an (unconstrained!) critical point of the functional $u \mapsto H(u) - \lambda I(u)$. Since both $H$ and $I$ are invariant for the $I$-flow (see Remark 48):

$$H(u) - \lambda I(u) = H(\Phi^I_\varphi u) - \lambda I(\Phi^I_\varphi u) \quad \text{for all} \quad \varphi,$$

it follows at once that also $U(\gamma, \varphi)$ defined by (4.41) is a critical point of $H - \lambda I$ for all $\varphi$, and therefore satisfies the same equation (see also the following Remark 52). Since $I(\Phi^I_\varphi U) = I(U) = \gamma$, each element $U(\gamma, \varphi)$ satisfies the constraint $I(u) = \gamma$. This shows that it is a constrained critical point of (4.39) for each $\varphi$. □

**Remark 52** A more algebraic proof that $U(\gamma, \varphi)$ satisfies the same equation as $U(\gamma)$ can be obtained as follows. Since $H$ is invariant for the $I$-flow it holds $H(u) = H(\Phi^I_\varphi u)$. Differentiation of this expression with respect to $u$ there results

$$\delta H(\Phi^I_\varphi u) = D\Phi^I_\varphi(u) \delta H(u).$$

Since the same holds for $I$, it follows that

$$\delta[H - \lambda I](\Phi^I_\varphi u) = D\Phi^I_\varphi \delta[H - \lambda I](u)$$

and hence the result.

**Remark 53** *Degeneracy of the dynamical critical point problem*
In the case considered in the Proposition above, the constrained critical point problem is clearly degenerate: a whole branch of critical points is found. This degeneracy reflects itself also in a property of the second variation. Indeed, differentiating the equation $\delta[H - \lambda I](\Phi^I_\varphi U) = 0$ with respect to $\varphi$, and then taking $\varphi = 0$ for simplicity, there results

$$[D^2 H(U) - \lambda D^2 I(U)] X_I(U) = 0. \tag{4.43}$$

This shows that the second variation operator is singular, with the Hamilton $I$-vector field $X_I(U)$ belonging to its kernel. This vector field is tangent to the $I$-level set. So, even when the critical points are constrained minimizers or maximizers, the (constrained) second variation is degenerate in this direction. This is a typical property of variational problems arising from such dynamical constructions[1].

### 4.2.3 Manifold of Relative Equilibria

A constrained Hamiltonian extremizer as considered above is called a *relative equilibrium*. Their relevance for the dynamics will be seen below, but first we consider the useful geometric picture that can be obtained for this set of constrained critical points if also $\gamma$ is interpreted as a parameter. This set is called the *Manifold of Relative Equilibria* determined by $H$ and $I$:

$$\text{MRE}\ (H, I) = \left\{ v \in \text{Crit}\ \{H \mid I^{-1}(\gamma)\} \,\middle|\, \gamma \in IR \right\}. \tag{4.44}$$

The word "manifold" is used here in a rather loose way. In fact the MRE may consist of several components, the different components corresponding to different critical points that are not connected by the $I-$flow (for instance, a component consisting of minimizers, one or more components consisting

[1] In that respect it may be noted that most results dealing with the second variation in standard variational calculus, assume the second variation to be *non*-degenerate, and therefore exclude these interesting dynamical problems.

of saddle points, etc).
In general it is very difficult to determine this set precisely. But the result of Proposition 51 shows that with each element $U$, a one-parameter branch belongs to MRE when the $I-$flow acts non-trivially on it:

$$\text{MRE } (H, I) \supset \{\Phi^I_\varphi U \mid \varphi \in IR\}. \tag{4.45}$$

Then, near a point $U(\gamma_0)$ for which there is a smooth branch $\gamma \mapsto U(\gamma)$ for $\gamma$ in a neighbourhood of $\gamma_0$, and for which the $I-$flow acts non-trivially, this manifold is at least two-dimensional and described locally with the parameters $\gamma$ and $\varphi$:

$$\text{MRE } (H, I) \supset \{U(\gamma, \varphi) \equiv \Phi^I_\varphi U(\gamma) \mid \gamma, \varphi \in IR\}. \tag{4.46}$$

Two tangent vectors to this manifold at a point $U$ are found to be

$$X_I(U) = \Gamma(U)\delta I(U) \quad \text{and} \quad n(U) = \frac{dU}{d\gamma}. \tag{4.47}$$

Here $n(U)$ is normal to the level set $I^{-1}(\gamma)$, while $X_I(U)$ is the tangent vector to the $I$-flow in the level set $I^{-1}(\gamma)$, see Fig. 4.2.

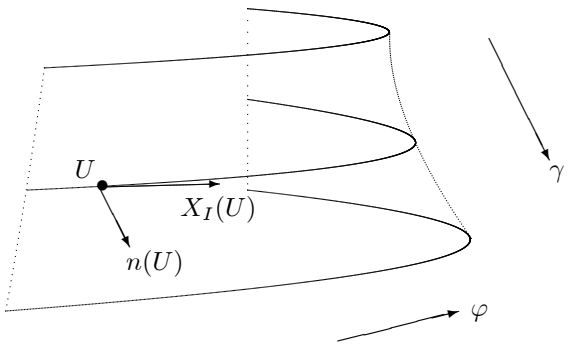


Figure 4.2: Schematic drawing of a two-dimensional (section of the) Manifold of Relative Equilibria at a point $U = U(\gamma, \varphi)$. The two tangent vectors $X_I(U) = \Gamma(U)\delta I(U)$ (the tangent along the $I$-flow in the levelset $I^{-1}(\gamma)$) and $n(U) = (dU(\gamma)/d\gamma)$ (normal to the levelset $I^{-1}(\gamma) : \langle \delta I(U), n(U)\rangle = 1$) are shown.

### 4.2.4 Relative equilibrium solutions

Now we come to the connection of constrained critical points and the dynamical equation under consideration. A relative equilibrium *solution* will be the solution that starts at a relative equilibrium. Its evolution will be shown to be a motion in the MRE, a motion at fixed $\gamma$ along the $I$-flow with speed determined by the multiplier $\lambda$.

**Proposition 54** Relative equilibrium solutions
*Let a constrained Hamiltonian extremizer $U(\gamma, \varphi_0)$ of (4.39) satisfy the equation (4.40) with multiplier $\lambda$. Then the application of the $I$-flow with speed $\lambda$:*

$$\Phi^I_{\lambda t}(U(\gamma, \varphi_0)) = U(\gamma, \varphi_0 + \lambda t) \tag{4.48}$$

*is a solution of the Poisson system (4.30):*

$$\Phi^H_t(U(\gamma,\varphi_0)) = \Phi^I_{\lambda t}(U(\gamma,\varphi_0)). \tag{4.49}$$

*A solution of this kind is called a* relative equilibrium solution*. If the action of the $I$-flow is trivial, in particular if $I$ is a Casimir functional, this solution is a time-independent solution: $\Phi^H_t U(\gamma,\varphi_0) = U(\gamma,\varphi_0)$.*

**Proof.** For ease of notation write $u(t) = U(\gamma, \varphi_0 + \lambda t)$. Then the following holds (we express in the notation that the structure map may depend explicitly on $u$):

$$\begin{aligned} \partial_t u &= \frac{d}{dt}\Phi^I_{\lambda t}(U(\gamma,\varphi_0)) = \lambda X_I(u(t)) \\ &= \lambda\Gamma(u(t))\delta I(u(t)) = \Gamma(u(t))\delta H(u(t)), \end{aligned}$$

the last equality from the fact that $u(t)$ satisfies $\delta H(u(t)) = \lambda\delta I(u(t))$ for each $t$ as expressed in Proposition 51. This is precisely the required result. □

**Remark 55** When $I$ is a Casimir functional, or when $U(\gamma)$ is in the fixed point set of the $I$-flow, the relative equilibrium solution is a time-independent solution. But also in another important case the solution may be a genuine equilibrium. This is the case if the multiplier $\lambda$ happens to be zero, say for $\gamma_0$. Then, although there exists a continuous branch $\{U(\gamma_0,\varphi) \mid \varphi\}$, and the critical points are degenerate, there is no motion along this branch. Since in general $\lambda(\gamma)$ changes with $\gamma$, i.e. $\partial_\gamma\lambda(\gamma_0) = \partial^2_\gamma h(\gamma_0) \neq 0$, the vanishing of $\lambda$ will generally occur only at isolated values of $\gamma$.

## 4.3 Symmetries and commuting flows

In this section we introduce several notions that are basic to understand the effect of the presence of integrals on the dynamics. The notions of symmetry and commuting flows are first formulated for arbitrary evolution equations. Then they are specialized to Poisson systems with additional integrals to produce the main result: the Lie-bracket of two Hamiltonian vector fields is the Hamiltonian vector field of the Poisson bracket of the two Hamiltonians. Consequently, functionals that Poisson commute, have Hamiltonian flows that commute.

### 4.3.1 Symmetries of evolution equations

Consider an evolution equation on a manifold $\mathcal{M}$

$$\partial_t u = K(u), \tag{4.50}$$

for which the flow map, to be denoted by $\Phi_t$, is well defined.

**Definition 56** *A map $\Psi : \mathcal{M} \to \mathcal{M}$ is called a* symmetry *for (4.50) if the equation is invariant under $\Psi$, meaning that $v(t) := \Psi(u(t))$ is a solution of (4.50) if $u$ is a solution. Equivalently, $\Psi$ is a symmetry for (4.50) if $\Psi$ commutes with the flow map $\Phi_t$:*

$$\Psi\Phi_t = \Phi_t\Psi \quad \textit{for all} \quad t. \tag{4.51}$$

Differentiating $v(t) = \Psi(u(t))$ with respect to time, there results

$$\partial_t v(t) = D\Psi(u(t))\partial_t u = D\Psi(u(t))K(u(t)).$$

Consequently, $\Psi$ is a symmetry iff

$$D\Psi(u)K(u) = K(\Psi(u)), \quad \text{for all} \quad u \in \mathcal{M}. \tag{4.52}$$

It is then said that the vectorfield $K$ is *equivariant* under the map $\Psi$.
A symmetry may be *discrete*, but in the following we will mainly deal with *continuous symmetries* which depend smoothly on one or more parameters. In particular, $\Psi_\varphi$ may represent the flow of another evolution equation, $\varphi$ denoting the evolution parameter (time).

### 4.3.2 Commuting flows

Consider a one parameter family of mappings $s \mapsto \Psi_s : \mathcal{M} \to \mathcal{M}$. More specifically we assume it to be a Lie group, and denote its infinitesimal generator by $N$. This means that $\Psi_s$ is the flow map of the equation

$$\partial_s v = N(v). \tag{4.53}$$

**Definition 57** *Two flows $\Phi_t$ and $\Psi_s$ are said to* commute *if*

$$\Phi_t \Psi_s = \Psi_s \Phi_t \quad \textit{for all } s,t \in IR. \tag{4.54}$$

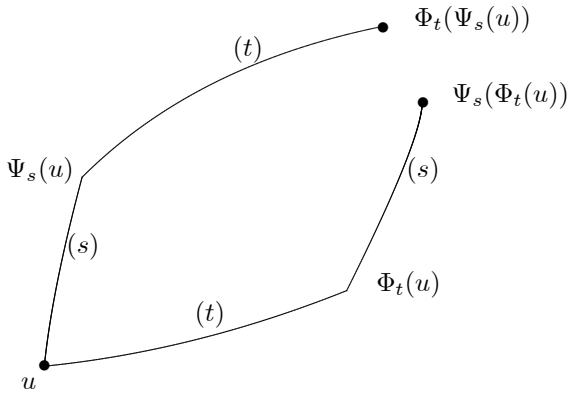


Figure 4.3: A successive application of two given flows $\Phi_t$ and $\Psi_s$ depends in general on the order in which the flows are applied; if that is not the case the flows are said to *commute*.

Differentiating for fixed $u \in \mathcal{M}$ each side of the identity $\Phi_t \Psi_s(u) = \Psi_s \Phi_t(u)$ once with respect to $t$ there results for the left hand side the expression

$$\partial_t \Phi_t(\Psi_s(u)) = K(\Phi_t(\Psi_s(u)))$$

and for the right hand side

$$\partial_t \Psi_s(\Phi_t(u)) = D\Psi_s(\Phi_t(u))\partial_t \Phi_t(u) = D\Psi_s(\Phi_t(u))K(\Phi_t(u)).$$

Taking $t = 0$, and equating both expressions, it follows that

$$K(\Psi_s(u)) = D\Psi_s(u)K(u) \qquad \text{for all } s \in IR, \tag{4.55}$$

which is precisely (4.52), i.e. $K$ is equivariant for $\Psi_s$.
In the same way it follows that the vectorfield $N$ is equivariant for the flow $\Phi_t$:

$$D\Phi_t(u)N(u) = N(\Phi_t(u)). \tag{4.56}$$

Now we continue the differentiation process: differentiating (4.55) with respect to $s$ at $s = 0$ (or (4.56) with respect to $t$ at $t = 0$) there results:

$$DK(u)N(u) - DN(u)K(u) = 0. \tag{4.57}$$

The expression on the left hand side is precisely the so-called Lie-bracket of the vector fields $K$ and $N$. This turns out to be an extremely important notion for the following.

**Definition 58** *On the set of (infinitely often) differentiable vector fields on $\mathcal{M}$, the* Lie-bracket *is a mapping from pairs of vector fields into a vector field defined by*

$$[K, N](u) := DK(u)N(u) - DN(u)K(u), \tag{4.58}$$

*for all $u \in \mathcal{M}$.*

From the definition (4.58) it readily follows that the Lie-bracket operation is *bilinear* and *skew-symmetric*. Moreover, it can be verified in a straightforward way that this bracket also satisfies *Jacobi's identity:*

$$[[K, N], M] + [[M, K], N] + [[N, M], K] = 0. \tag{4.59}$$

This is summarised by saying that the vectorfields on $\mathcal{M}$ form a *Lie-algebra* with the bracket (4.58). We have seen that if the flows $\Phi_t$ and $\Psi_s$ commute, that then (4.57) holds. Using the Lie-bracket notation, this can be written like

$$[N, K] = 0 \quad \text{on} \quad \mathcal{M}. \tag{4.60}$$

Hence (4.60) is a *necessary* condition for the flows to commute. In fact, it is true that (4.60) is not only necessary, but also a *sufficient* condition for the flows to commute. (The proof of the sufficiency essentially follows by integrating the local condition (4.60); it will not be reproduced here; see e.g. Arnold, 1989)

**Proposition 59** *The flows $\Phi_t$ and $\Psi_s$ for the vector fields $K$ and $N$ respectively, commute if and only if the Lie-bracket of the vector fields vanishes: $[K, N] = 0$.*

This result is of considerable practical importance. It allows one to decide whether two flows commute or not without having to calculate the flows explicitly (which would be impossible in general). We collect the results above as follows.

**Proposition 60** *The following statements are equivalent:*
*(i) The flows $\Phi_t$ and $\Psi_s$ of the vector fields $K$ respectively $N$ commute;*
*(ii) The vector field $K$ is equivariant under $\Psi_s$;*
*(iii) The vector field $N$ is equivariant under $\Phi_t$;*
*(iv) The Lie-bracket of the vector fields $K$ and $N$ vanishes identically on $\mathcal{M}$.*

**Remark 61** A somewhat different interpretation of the relevance of the commutativity property is obtained as follows. Consider the vector field that is the sum of two (known) vector fields:

$$\partial_t u = K(u) + N(u). \tag{4.61}$$

Then a natural question is to ask what can be said about the flow of (4.61) when the flows $\Phi_t$ and $\Psi_s$ for the separate vector fields $K$ and $N$ are known. In order to investigate that, start with a (time-dependent) change of variables form $u(t)$ to $v(t)$ according to:

$$u(t) = \Phi_t(v(t)). \tag{4.62}$$

Then

$$\partial_t u = K(u(t)) + D\Phi_t(v(t))\partial_t v.$$

This shows that $\Phi_t(v(t))$ is a solution of (4.61) provided that $v$ satisfies:

$$\partial_t v = D\Phi_t^{-1}(v)N(\Phi_t(v)). \tag{4.63}$$

Consequently, if $N$ is equivariant under $\Phi_t$, then $D\Phi_t^{-1}(v)N(\Phi_t(v)) = N(v)$ and (4.63) is precisely equation (4.53) for which the flow map $\Psi$ is assumed to be known. Hence, if $[K, N] = 0$, the solution of (4.61) with initial condition $u_0$ is given by $u(t) = \Phi_t\Psi_t(u_0)$ where the order of the (commuting!) flows $\Phi_t$ and $\Psi_t$ is arbitrary. This can be expressed as follows:*the flow of the sum of two vector fields $K$ and $N$ which have vanishing Lie-bracket, is the superposition of the two flows of $K$ and $N$.*
When the Lie-bracket does not vanish, the transformation (4.62) leads to the equation (4.63) in which the $\Phi-$flow is still present.

### 4.3.3 The Lie-bracket of Hamiltonian vector fields

From now on we restrict to a Poisson manifold and consider Hamiltonian vector fields $X_H$ and $X_I$ with corresponding flows denoted by $\Phi^H$ and $\Phi^I$ as usual:

$$\begin{aligned}\partial_t u &= \Gamma(u)\delta H(u) \equiv X_H(u)\\ \partial_t u &= \Gamma(u)\delta I(u) \equiv X_I(u).\end{aligned}$$

The following result provides a simple way to conclude in specific cases whether $\Phi^H$ and $\Phi^I$ commute. It relates in a fundamental way the Poisson bracket of functionals to the Lie-bracket of the corresponding Hamiltonian vector fields.

**Proposition 62** *On a Poisson manifold the Lie-bracket of Hamiltonian vector fields $X_H$ and $X_I$ is the Hamiltonian vector field of the Poisson bracket of $H$ and $I$:*

$$[X_H, X_I] = X_{\{H,I\}}. \tag{4.64}$$

**Proof.** For an arbitrary functional $F$ we have on the one hand

$$\langle \delta F, X_{\{H,I\}}\rangle = \{F, \{H, I\}\},$$

which, by Jacobi's identity, equals to

$$= \{\{F, H\}, I\} + \{\{I, F\}, H\}.$$

On the other hand, using the symmetry of $D^2F(u)$, we can derive the following chain of equalities:

$$\begin{aligned}\langle \delta F, [X_H, X_I]\rangle &= \langle \delta F, DX_H \circ X_I\rangle - \langle \delta F, DX_I \circ X_H\rangle\\ &= \langle \delta F, DX_H \circ X_I\rangle + \langle D^2F \circ X_I, X_H\rangle\\ &\quad -[\langle \delta F, DX_I \circ X_H\rangle + \langle D^2F \circ X_H, X_I\rangle]\\ &= \langle \delta\langle \delta F, X_H\rangle, X_I\rangle - \langle \delta\langle \delta F, X_I\rangle, X_H\rangle\\ &= \{\{F, H\}, I\} - \{\{F, I\}, H\}.\end{aligned}$$

Hence $\langle \delta F, X_{\{H,I\}}\rangle = \langle \delta F, [X_H, X_I]\rangle$ for all $F$, and the result follows. □

This result has some immediate consequences.

**Proposition 63** *Suppose that the Poisson bracket of two functionals $H$ and $I$ vanishes or is a Casimir functional: $\{H, I\} = C(asimir)$. Then the Hamiltonian flows $\Phi^H$ and $\Phi^I$ commute, and the other equivalences of Proposition 60 hold.*

**Proof.** With (4.64) it follows that the Lie-bracket of the two vector fields vanishes since the vector field of a Casimir functional vanishes: $[X_H, X_I] = X_{\{H,I\}} = X_C = 0$. Then Proposition 62 states that the flows commute. □

**Remark 64** Note that when $\{H, I\} = C(asimir)$ with $C$ non-trivial, despite the fact that the Hamiltonian flows commute, the $I-$flow will in general not lie in level sets of $H$ (and conversely). In fact, since

$$\partial_\varphi H(\Phi^I_\varphi(u)) = C(\Phi^I_\varphi(u)) = C(u),$$

the value of $H$ on $\Phi^I_\varphi(u)$ increases linearly in $\varphi$ with rate determined by $C(u)$.

# Chapter 5

# Coherent Structures as Relative Equilibria

In this chapter we consider some coherent structures that are observed and relevant in fluid dynamics. Although the notion of coherent structure is not defined precisely, we here refer to such phenomena as travelling waves and vortices. These phenomena are characterized by the fact that during their evolution they exhibit a well defined spatial structure, that they behave as a relatively simple, low-dimensional system and that, by the very nature that they are observable and computable numerically, these phenomena are stable with respect to (small) perturbations in initial data and in the equation that describes the phenomenon.

## 5.1 Vortex points

Kirchhoff's equations for vortex points were shown to be a Hamiltonian system and can be derived in a consistent way from 2D Euler equations. For convenience, we recall the governing equations and Poisson structure.
For $N$ vortex points, with strength $\gamma_1, \ldots, \gamma_N$ at places $r_1 \ldots r_N$, the Hamiltonian is given by

$$H(r_1, \ldots, r_N) = \sum\nolimits^{*} -\frac{1}{4\pi}\gamma_k\gamma_j \log|r_k - r_j|. \tag{5.1}$$

(The asterisk in $\sum^*$ denotes that the self-interaction term $j = kt$is omitted). With the bracket:

$$\{F, G\} := \sum_k \partial_{r_k} F \cdot \frac{1}{\gamma_k} J \partial_{r_k} G, \tag{5.2}$$

the governing equations, Kirchhoff's equations, are

$$\partial_t r_k = \frac{1}{\gamma_k} J \partial_{r_k} H(r_1, \ldots, r_N), \quad 1 \le k \le N, \tag{5.3}$$

explicitly,

$$\partial_t r_k = \sum\nolimits^{*} J\ [-\frac{1}{4\pi}\gamma_j \frac{r_k - r_j}{|r_j - r_k|^2}], \quad k = 1, \ldots, N. \tag{5.4}$$

### 5.1.1 Momenta integrals and their flows

We start to investigate the integrals and their flows. The system admits the following three integrals: *linear momenta:*

$$L = (L_1, L_2) = \sum_k \gamma_k r_k = (\sum_k \gamma_k x_k, \sum_k \gamma_k y_k). \tag{5.5}$$

and *angular momentum:*

$$P = \frac{1}{2}\sum_k \gamma_k |r_k|^2 = \frac{1}{2}\sum_k \gamma_k (x_k^2 + y_k^2). \tag{5.6}$$

Indeed, it is readily verified that

$$\begin{aligned} \{L, H\} &= J\sum_k \partial_{r_k} H = 0 \\ \{P, H\} &= \sum_k r_k \cdot J\partial_{r_k} H = 0. \end{aligned}$$

It must be noted, that the integrals $L_1, L_2$and $P$ do not Poisson commute in general. By introducing the total vorticity by $\gamma = \sum_k \gamma_k$, it holds

$$\{L_1, L_2\} = \gamma, \tag{5.7}$$
$$\{L_1, P\} = L_2, \tag{5.8}$$
$$\{L_2, P\} = -L_1. \tag{5.9}$$

In particular, this will imply that the Hamiltonian flows of $L$and $P$do not commute. The flows of $L_1$and $L_2$commute as a consequence of Proposition 63 since their bracket is a Casimir function, the constant total vorticity.
To describe these flows explicitly, let $e = (e_1, e_2)$ be a given unit vector. Then taking as Hamiltonian the integral $e \cdot L$, the equations become

$$\partial_t r_k = Je$$

From this it follows that *the Hamiltonian flow of $e \cdot L$ is a uniform translation in the direction $Je$* (which direction is perpendicular to $e$):

$$\Phi^{e\cdot L} r(t) = r(t) + tJe. \tag{5.10}$$

The Hamiltonian flow for $\alpha P, \alpha$ real, follows from the equations

$$\partial_t r_k = \alpha J r_k.$$

This is *a uniform rotation about the origin with angular velocity $\alpha$*. The rotation is clockwise if $\alpha > 0$. Explicitly, for each component:

$$r_k(t) = \mathcal{R}(\alpha t) r_k(0), \quad \text{with} \quad \mathcal{R}(\tau) = \begin{pmatrix} \cos\tau & \sin\tau \\ -\sin\tau & \cos\tau \end{pmatrix}. \tag{5.11}$$

Consistent with the non-commutativity of $L$and $P$, translations and rotations do not commute.

### 5.1.2 Relative Equilibria

Relative equilibria can be found by looking for critical points of $H$ on level sets of $L$ or $P$. Since $L$ and $P$ do not commute, it is not clear *a priori* that both $L$ and $P$ can be prescribed. However the following observations show that only one of these integrals provide essential information. Two cases have to be distinguished depending on whether the total vorticity $\gamma = \sum \gamma_k$is zero or not.
Under a translation $R$ of the origin of the coordinate system the integrals $L$ and $P$ change according to

$$L_R = L_0 - R\gamma, \quad \text{with} \quad L_R = \sum (r_k - R)\gamma_k,$$

and

$$P_R = P_0 - R \cdot L_0 + \frac{1}{2}\gamma |R|^2, \quad \text{with} \quad P_R = \frac{1}{2}\sum \gamma_k |r_k - R|^2.$$

From this it is seen that if $\gamma \neq 0, R$ can be chosen in such a way that $L_R = 0$:

$$\text{iff} \quad \gamma \neq 0 \quad \text{then} \quad L_R = 0 \quad \text{for} \quad R = L_0/\gamma.$$

This point $R$ is called the *center of the vorticity distribution.*
On the other hand, if $\gamma = 0$ then $R$ can be chosen such that $P_R = 0$. Now, the Hamiltonian $H$ under consideration, i.e. (1.37), is invariant for any translation of the origin and for a rotation about the origin:

$$\begin{aligned} H(r_1, \ldots, r_N) &= H(r_1 - R, \ldots, r_N - R) \\ &= H(\mathcal{R}(\tau) r_1, \ldots, \mathcal{R}(\tau) r_N). \end{aligned}$$

Hence, the only constrained critical point problems of interest are

$$\text{when} \quad \gamma = 0: \qquad \text{Crit} \, \{H \mid e \cdot L = \ell\}, \tag{5.12}$$

and

$$\text{when} \quad \gamma \neq 0: \qquad \text{Crit} \, \{H \mid P = p\}. \tag{5.13}$$

**Proposition 65** *A constrained critical point of (refeqG71.10) satisfies for some multiplier $V$ (which depends on $\ell$)*

$$\nabla H = V \, e \cdot \nabla L, \tag{5.14}$$

*and the corresponding relative equilibrium is a uniform translation along the line through the origin in the direction $Je$ with velocity $V$.*
*A critical point of (5.13) satisfies for some multiplier $\alpha$ (which depends on $p$)*

$$\nabla H = \alpha \nabla P, \tag{5.15}$$

*and the corresponding relative equilibrium is a rotation about the origin with angular velocity $\alpha$.*

Next we will look explicitly for relative equilibrium solutions for 2, and for special cases of 3 point vortices. The motion of two point vortices is simple as we shall see. In that case the system is completely integrable and all solutions correspond to relative equilibria. For more vortex points relative equilibria can also be found, although the dynamics may be very complicated, even in case of three vortex points.

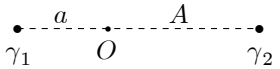


Figure 5.1: Relative equilibrium of two point vortices of equal sign: rotation around the center of vorticity $O$.

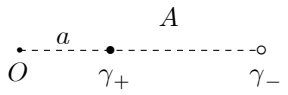


Figure 5.2: Relative equilibrium of two point vortices with different sign and total vorticity nonzero: rotation around the center of vorticity $O$.

### 5.1.3 Two point vortices

First consider two point vortices for which the total vorticity $\gamma = \gamma_1 + \gamma_2$is not zero. We place the center of vorticity in the origin. For given values of $p \in IR$, consider the variational problem (5.13).

**Proposition 66** *A relative equilibrium characterized by (5.13) exists if $p$ satisfies $p\gamma > 0$ for points of the same sign, and if $p\gamma < 0$ for points of different sign. The relative equilibrium is a steadily rotating configuration of the two point vortices positioned in alignment with the origin (see Fig. 5.1, 5.2). The points with strengths $\gamma_1, \gamma_2$ are at distance $a$ and $A$ from the origin, where:*

$$a^2 = \frac{2p\gamma_2}{\gamma\gamma_1}\ , \quad A^2 = \frac{2p\gamma_1}{\gamma\gamma_2}. \tag{5.16}$$

*The configuration rotates as a whole around the origin with angular velocity:*

$$\alpha = -\frac{\gamma_1\gamma_2}{4\pi p}. \tag{5.17}$$

*If the vortices have equal sign, they are positioned on opposite sides of the origin; if they have different sign, they are at the same side of the origin. In both cases, the rotation is counterclockwise if $\gamma > 0$ and clockwise if $\gamma < 0$.*

**Proof.** From the equation

$$\nabla H = \alpha \nabla P,$$

it readily follws that for all time

$$\gamma_1 r_1(t) + \gamma_2 r_2(t) = 0, \tag{5.18}$$

and that the angular velocity $\alpha$is given by

$$\alpha = -(\gamma/2\pi)|r_1 - r_2|^{-2}. \tag{5.19}$$

This last result shows that $\text{sign}(\alpha) = -\text{sign}(\gamma)$, giving the statements about the direction of rotation. For the value of $p$ it holds

$$p = \frac{1}{2}\frac{\gamma_1}{\gamma_2}\gamma|r_1|^2 = \frac{1}{2}\frac{\gamma_2}{\gamma_1}\gamma|r_2|^2.$$

From this follow the conditions on $p$, depending on the signs of the points. Expressing everything in terms of $p$, and using the definition of $a$ and $A$, the result follows easily.

This result can also be obtained by invoking *a priori* the configuration as depicted in Fig. 5.1 and 5.2, and expressing $H$ and $P$ directly in terms of $a$ and $A$:

$$H = -\frac{\gamma_1\gamma_2}{2\pi}\log(a+A), \tag{5.20}$$

$$P = \frac{1}{2}\gamma_1 a^2 + \frac{1}{2}\gamma_2 A^2. \tag{5.21}$$

Then variations with respect to the distances $a$ and $A$ yield:

$$\begin{pmatrix} \partial_a H \\ \partial_A H \end{pmatrix} = \alpha \begin{pmatrix} \partial_a P \\ \partial_A P \end{pmatrix} \tag{5.22}$$

from which the same results can be obtained.
The value function can be calculated explicitly:

$$h(p) = -\frac{\gamma_1\gamma_2}{4\pi}\log\frac{2p\gamma}{\gamma_1\gamma_2}; \tag{5.23}$$

its graph is shown in Fig. 5.3 and 5.4. Note that $\alpha$ as given above is indeed related to $h(p)$ according to $\alpha = dh/dp$.

figure=/users//mirande/brenny/diktaat-ads/figads1.id,height=3cm

Figure 5.3: The value function for two positive point vortices.

figure=/users/mirande/brenny/diktaat-ads/figads2.id,height=3cm

Figure 5.4: The value function for two point vortices of opposite sign. The picture shows the integral diagrams for $H$ and $P$ for different choices of the strengths $\gamma_1$ and $\gamma_2$, with different cases for negative or positive total vorticity $\gamma$. The case with $\gamma = 0$ does not give a steadily rotating solution (see also Proposition 67).

Next consider the equilibrium state of two point vortices of opposite sign, with total vorticity zero: $\gamma_1 = -\gamma_2 = \gamma$. We choose the origin such that $P = 0$, and use the characterization (5.12). It is possible to choose axes such that the $L_1$component of the linear momentum vanishes, leading to

$$\text{Crit } \{H \mid L_1 = 0;\ L_2 = \ell_2\ \} \tag{5.24}$$

for given value of $\ell_2 \in IR\backslash\{0\}$. Then the following result is obtained easily.

**Proposition 67** *For two equal point vortices of opposite sign, a relative equilibrium characterized by (5.24) consists of a steadily translating configuration of the two point vortices, positioned on opposite sides of the origin (see Fig. 5.5), both at distance $A$ from the origin, where:*

$$A = \frac{\ell_2}{2\gamma}. \tag{5.25}$$

*The translation is along the $x-$axis with speed given by:*

$$V_2 = \frac{\gamma^2}{2\pi\ell_2}. \tag{5.26}$$

*The value function in this case is given by*

$$h(\ell_2) = \frac{\gamma^2}{2\pi}\log\frac{\ell_2}{\gamma},$$

*and the value for $V_2$ is consistent with $V_2 = dh/d\ell_2$.*

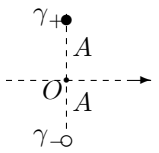


Figure 5.5: Relative equilibrium of two point vortices with opposite strengths: a uniform translation along the $x$-axis.

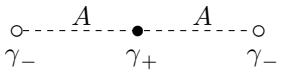


Figure 5.6: Relative equilibrium of one positive point vortex flanked by two equal negative vortices: rotation around the positive center vortex.

**Exercise 68** *1.* For the motion of two vortex points (a Hamiltonian system in $IR^4$), there are two constants of the motion (the energy and angular or linear momentum). The system is "completely integrable", which implies that in fact *every* solution is a relative equilibrium solution. Investigate this.

*2.* The motion of three vortex points is much more complicated: the dynamics can be chaotic. Consider the motion of three vortex points: two vortices, each with strength $\frac{1}{2}\gamma_-$, and one with strength $\gamma_+$. Investigate the relative equilibria that are such that the three points are aligned, as in Fig. 7.6. Distinguish between the cases $\gamma_+ + \gamma_- = 0$ and $\gamma_+ + \gamma_- \neq 0$.

## 5.2 Travelling waves

In this section we consider the wave equations encountered before. Using the general theory we investigate the relative equilibria that can be found using the additional integrals that are present.
For one-dimensional wave equations we investigate solutions decaying at infinity: $u \in C_0^\infty(IR)$, or solutions that are periodic with a period to be prescribed: $u \in C_{per}^\infty$.

### 5.2.1 KdV-type of equations

For wave equations of the form

$$\partial_t u = -\partial_x \delta H(u) \tag{5.27}$$

there are two functionals that are obvious candidates to be integrals for (5.27). The first one is the *mass-functional*

$$M(u) = \int u. \tag{5.28}$$

Its variational derivative in $\delta M(u) = 1$, which belongs to $C_{per}^\infty$in the periodic case, but which is not well-defined on the set of decaying solutions: $\delta M(u) \notin C_0^\infty(IR)$.
Since $\{M, H\}(u) = \langle 1, -\partial_x\ \delta H(u)\rangle$, $M$commutes with $H$for periodic waves and formally also for decaying solutions on $IR$provided $\delta H(u)$vanishes if $u$and its derivatives vanish: $\delta H(u) \to 0$for $|x| \to \infty$.
With these restrictions, $M$is an integral, and in fact a Casimir functional. The $M$-flow is trivial.
The second obvious candidate is the *momentum functional*

$$I(u) = \int \frac{1}{2}u^2. \tag{5.29}$$

Since

$$\begin{aligned}\{I,H\}(u) &= \langle u, -\partial_x \delta H(u)\rangle = \langle \delta H(u), \partial_x u\rangle \\ &= \frac{d}{d\varepsilon} H(u(\cdot+\varepsilon))\Big|_{\varepsilon=0}, \end{aligned} \tag{5.30}$$

the vanishing of $\{I,H\}$ will be a consequence of translation invariance of $H$. Since the $I$-flow is nothing but translation, $\Phi^I_\varphi\ u(x) = u(x-\varphi)$, this is precisely the condition that $H$ is invariant for the $I$-flow. For a density functional, $H$ will be translation invariant if the density does not depend explicitly on the spatial variable $x$. Hence

$$\{I,H\} = 0 \ \text{ iff } \ H \ \text{ is translationally invariant}$$

For *decaying solutions* on $IR$, $u \in C^\infty_0(IR)$,the functional $I$can be used to find relative equilibria:

$$\mathrm{MRE}(H,I) = \{U(\gamma,\varphi) | \gamma > 0, \varphi \in IR\},$$

with

$$U(\gamma,\varphi) = \Phi^I_\varphi : U(\gamma) \in \text{Crit } \{H(u) | I(u) = \gamma\}.$$

The equation for a constrained critical point reads

$$\delta H(u) = \lambda \delta I(u), \ \text{ i.e. } \ \delta H(u) = \lambda u,$$

where the multiplier $\lambda$is related to the value function $\lambda = \partial_\gamma\ h(\gamma)$. The corresponding relative equilibrium solution is the solution given by

$$\Phi^I_{\lambda t}\ U(\gamma)(x) = U(\gamma)(x - \lambda t)$$

For *periodic solutions*, $u \in C^\infty_{per}$, both $I$and $M$can be used to find relative equilibria. Since the value of $M$determines the mean-value of $u$, there is no loss of understanding to restrict to the case $M = 0$, i.e. to functions with mean value zero:

$$\mathrm{MRE}_0(H,I) = \{U(\gamma,\varphi) | \gamma > 0, \varphi \in IR\}$$

with

$$U(\gamma,\varphi) = \Phi^I_\varphi(U(\gamma)) \in \text{Crit } \{H(u) \mid I(u) = \gamma, M(u) = 0\}. \tag{5.31}$$

The equation for a constrained critical point reads

$$\delta H(u) = \lambda \delta I(u) + \alpha \delta M(u), \ \text{ i.e. } \ \delta H(u) = \lambda u + \alpha. \tag{5.32}$$

The corresponding relative equilibrium is again given by (**??**) since the $M$-flow is trivial.
The relative equilibrium solution given by (**??**) is a *travelling wave* solution of (5.27): the function $x \mapsto U(\gamma)(x)$defines the wave profile, $\Phi^I_\varphi\ U(\gamma)$ is the translation of this profile over a distance $\varphi$, and (**??**) describes the translation with constant velocity $\lambda$. The Manifold of Relative Equilibria can thus be called the *Travelling Wave Manifold* in the state space.

**Remark 69** Travelling wave solutions of (5.27) can also be found directly by substituting the "Ansatz" $u(x,t) = v(x - \lambda t)$ in (5.27). Then there results the equation for the wave profile $v$:

$$\lambda \partial_x v = \partial_x \delta H(v).$$

Upon integrating this equation with respect to $x$, the equation (5.32) results with $\alpha$ a constant of integration. For solutions in $C^\infty_0(IR)$, and with $\delta H(u) \to 0$ as $|x| \to \infty$, it follows from (5.32) that $\alpha = 0$, and equation (**??**) is obtained.

### 5.2.2 KdV-solitons

The normalized KdV-equation (3.17):

$$\partial_t u + u_{xxx} + 6uu_x = 0$$

is written as a Poisson system with Hamiltonian $H$ like

$$\partial_t u = \partial_x \delta H(u), \quad H(u) = \int (\frac{1}{2}u_x^2 - u^3).$$

For solutions decaying at infinity, $u \in C_0^\infty(IR)$, equation (**??**) reads

$$-u_{xx} - 3u^2 = \lambda u.$$

For each $\lambda < 0$ this equation has a solution. The solution satisfying the constraint $I(u) = \gamma$ is given explicitly by

$$U(\gamma; x) = \frac{1}{2}\sigma \mathrm{sech}^2[\frac{1}{2}\sqrt{\sigma}x] \tag{5.33}$$

where

$$\sigma = -\lambda = (3\gamma)^{2/3} \tag{5.34}$$

and sech$= 1/\cosh$. These solutions are in fact minimizers of the constrained Hamiltonian:

$$U(\gamma) \in \text{Min } \{H(u) \mid I(u) = \gamma\}.$$

A direct calculation of the value function using the explicit expression (5.33) yields

$$h(\gamma) = -\frac{1}{5}(3\gamma)^{5/3} \tag{5.35}$$

and the expression for the multiplier $\lambda = \partial_\gamma h(\gamma)$ is in agreement with the result in (5.34).
The *relative equilibrium solution* is the translation of the waveprofile (5.33) with velocity $\lambda$:

$$\Phi^I_{\lambda t} U(\gamma)(x) = U(\gamma)(x + \lambda t) = \frac{1}{2}\sigma \mathrm{sech}^2[\frac{1}{2}\sqrt{\sigma}(x + \lambda t)]. \tag{5.36}$$

(Since the $I$-flow is translation to the left for the structure map $\partial_x$, the sign of $\lambda$ is negative, and the wave travels to the right.) This travelling wave is called a *soliton*, for reasons to be described below.

**Exercise 70** *1.* Regroup the terms in equation (**??**) such that a second order Hamiltonian system arises:

$$u_{xx} = -V'(\lambda, u)$$

for a suitable function V that depends on $\lambda$. Use *phase plane analysis* to show that this equation, for certain values of $\lambda$, has a *homoclinic orbit*.

*2.* Show that this homoclinic orbit corresponds to the solitary wave profile.

*3.* By eliminating $\lambda$ as parameter in favour of $\gamma$, verify the given results.

*4.* Perform the same programme for the BBM-equation:

$$(1 - \partial_x^2)\partial_t u = \partial_x[u + u^2];$$

find the solitary waves explicitly.

### 5.2.3 KdV-cnoidal waves

For periodic solutions, with normalized period $2\pi, x \in (-\pi, \pi)$, the constrained minimization problem (5.31):

$$\text{Min } \{H(u) \mid I(u) = \gamma,\ M(u) = 0\} \tag{5.37}$$

leads to the equation

$$-u_{xx} - 3u^2 = \lambda u + \alpha \tag{5.38}$$

with solutions that can be written down explicitly using a Jacobian elliptic function known as *cnoidal-function*. These constrained minimizers of $H$ have fundamental period $2\pi$. Other solutions of (5.38) have period $(2\pi/k)$, with $k \geq 2$ some integer, and can be found by rescaling. These solutions with non-fundamental period $2\pi$ are constrained *saddle points* for $H$.
The relative equilibrium solution is again a travelling wave: the translation of the wave profile $U$that is the solution of (5.37) with velocity $\lambda$:

$$\Phi^I_{\lambda t} U(\gamma)(x) = U(\gamma)(x + \lambda t).$$

These solutions are called *cnoidal waves*.
The value function $h$ is plotted in the Figure. It may be observed, noting that $\lambda = \partial_\gamma h(\gamma)$, that the propagation speed of the wave changes sign: the translation is to the left (where $\lambda > 0$) for small values of $\gamma$, and to the right (where $\lambda < 0$) for larger values of $\gamma$. At a critical value the wave is steady: a time-independent non-trivial equilibrium solution of the KdV equation.

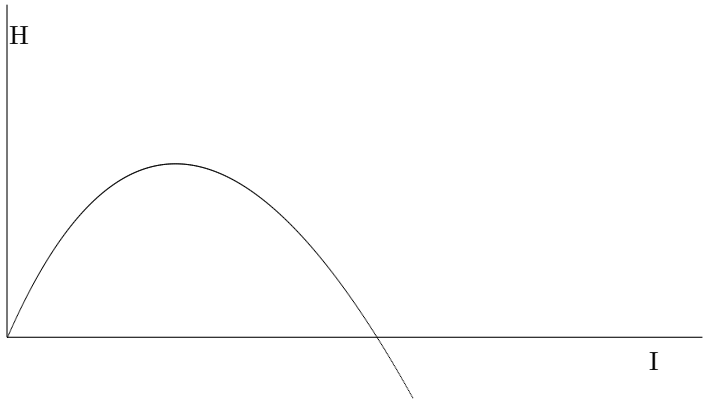


Figure 5.7: Diagram $H$ vs $I$ parametrized by $t$

**Exercise 71** *1.* Regroup the terms in equation (**??**) such that a second order Hamiltonian system arises:

$$u_{xx} = -V'(\lambda, \alpha, u)$$

for a suitable function V that depends on $\lambda$ and $\alpha$. Use *phase plane analysis* to show that this equation, for certain values of the parameters has solutions that corresponds to the cnoidal wave profiles.

*2.* Derive the extremum values for the parameter $\lambda$.

*3.* Perform the same programme for the BBM-equation:

$$(1 - \partial_x^2)\partial_t u = \partial_x[u + u^2].$$

# Part II

# Annex

# Chapter 6

# Modelling and discretizations

We have seen in the previous chapters that the specific structure of an equation leads to particular properties, like coherent structures/relative equilibria for Poisson systems. In all interesting cases the solutions of the dynamic equation cannot be found explicitly and one likes to construct and investigate simpler equations that have the same fundamental properties as the original problem. For Poisson systems one therefore looks for simpler models that retain the Poisson structure. This is an example of *consistent modelling*. A *discretization* of an equation that is constructed for numerical purposes can just as well be seen as a (discrete) model of the problem under consideration, and the same reasoning applies. In this chapter we investigate the characteristic properties and difficulties connected to discretizing partial differential equations. We restrict to *space discretizations*, keeping time continuous: the continuous evolution equation (pde) will be replaced by a discrete dynamical system (ode). We illustrate the general considerations to the specific case of Fourier truncation; the essential idea is that functions from the infinite dimensional state space $\mathcal{U}$ $\mathcal{U}$ are approximated by functions from a finite dimensional state space $\mathcal{V}$. This space $\mathcal{V}$ can often be identified with a space of parameters, for instance the Fourier coefficients. But we will explicitly identify the elements of $\mathcal{V}$ with functions: $\mathcal{V}$ will be embedded in $\mathcal{U}$, since this makes it possible to compare the original dynamic equation with its discretized form.
It should be noted that at this place we do not bother about the process of *approximation*: the precise properties of an approximation in $\mathcal{V}$ of an element from $\mathcal{U}$, is not our concern here and is studied in *approximation theory*.

In the first Section we deal with some general theory; in Section 8.2 we introduce a software programme WAVEPACK to illustrate the general ideas to the discretization of the (dissipative) KdV equation; in the last section several software experiments are described.

## 6.1 Restriction of the dynamics

Consider the continuous evolution equation in (infinite dimensional) state space $\mathcal{U}$ in the general form

$$\partial_t u = K(u).$$

Let $u \in \mathcal{U}$ be approximated by some (finite dimensional) $v \in \mathcal{V}$. The difference $u - v$ is then considered as an error (infinite dimensional); it will typically be denoted by $w$, and then $w$ belongs to some space, $\mathcal{W}$ say,

$$\mathcal{U} \ni u = v + w, \quad v \in \mathcal{V}, \quad w \in \mathcal{W}.$$

The equation for $u$ transforms to a set of coupled equations for $v, w$, say

$$\partial_t v = K_1(v,w), \tag{6.1}$$
$$\partial_t w = K_2(v,w). \tag{6.2}$$

To find a simplified equation to replace the full equation for $u$, the simplest idea would be to neglect the error $w$ by writing $w = 0$ and to consider

$$\partial_t v = K_1(v,0) \tag{6.3}$$

as the approximate equation for $v$. In general this will not be exact since $w \neq 0$:

$$w = 0 \text{ only if } K_2(v,0) \equiv 0 \text{ for all } v \in \mathcal{V}. \tag{6.4}$$

**Exercise 72** Show that for the following linear, homogeneous equations, Fourier truncation with $N$ modes is exact for initial data with only $N$ modes:

$$\partial_t u = c_0 u_x,$$
$$\partial_t u = u_{xx}.$$

Investigate the effect of the error for arbitrary initial data.

### 6.1.1 Mode coupling and generation

Let $\mathcal{V}$ be the space $\mathcal{F}_N$ of $N$-mode Fourier truncations (we use complex notation for simplicity):

$$\mathcal{F}_N = \{\sum_{-N}^{N} c_k e^{ikx} \mid (c_{-N}, \ldots, c_0, \ldots, c_N) \in C^N\}. \tag{6.5}$$

This is a linear space. However, the space is *not closed for multiplication*: the product of two elements from $\mathcal{F}_N$ contains higher order modes in general: for $f, g \in \mathcal{F}_N$

$$f(x) \cdot g(x) = [\sum_{-N}^{N} f_k e^{ikx}][\sum_{-N}^{N} g_\ell e^{i\ell x}] = \sum_{k=-N}^{N} \sum_{\ell=-N}^{N} f_k g_\ell e^{i(k+\ell)x} \in \mathcal{F}_{2N}.$$

This (unavoidable) property of the approximation is called *mode-generation*: higher order terms $e^{imx}$, with $N < |m| \leq 2N$ appear in the product that are not present in the original functions $f, g$.

**Observation.** *Nonlinearity and space-inhomogeneity are sources that give rise to mode generation*: for an initial condition $u_0 \in \mathcal{F}_N$ (for which $w_0 = 0$), the resulting solution will have $w(t) \neq 0$; *small scaled structures* (described by $w$) *are generated by the large scaled structures* that are described by $v \in \mathcal{F}_N$.

**Exercise 73** Verify the observation for the following equations

$$\partial_t u = c(x) u_x \quad \textit{inhomogeneous wave eqn}$$
$$\partial_t u = u_{xx} + u u_x \quad \textit{Burgers eqn, nonlinear}$$

### 6.1.2 Consistent discretization

There are various ways to approximate the full set of equations and to define a simplified model. To be specific, we consider the basic equation

$$\partial_t u = L\delta H(u); \tag{6.6}$$

combining the various cases that are obtained for symmetric and skew-symmetric operators $L$, we can find the procedure how to deal with more general equations that have gradients in the vector field. Below we will describe two approximations that may differ somewhat. Each is meant to give an approximation that describes the large-scale features, neglecting the contribution from the small scales in some way or another. It should be noted that there are almost no mathematical results about the precise relation between the exact and each of the approximated dynamics.

In the following we refer to the transformation described in Subsection 3.5.2.; there the exact equation for $u$ was transformed in the equivalent set of equations for $v, w$.

**Truncated dynamics**
One procedure to approximate the equation is to truncate $u$ to $N$-th order by taking $w = Qu = 0$. At the same time, the given functional $H$ is truncated accordingly:

$$\bar{H}(v) := \hat{H}(v, 0) = H(Pu). \tag{6.7}$$

Using this restricted functional, the equation for $v = Pu$ becomes the so-called truncated dynamics

$$\partial_t v = \bar{L}\delta_v \bar{H}, \quad \text{with} \quad \bar{L}(v) = PL(v)P^* \,. \tag{6.8}$$

In particular, observe that:
• the variational derivative is replaced by the gradient of the restricted functional;
• when $L$ is (skew-)symmetric, then so is $\bar{L}$; in particular: if the original equation is of steepest descent type or of conservative type, then so is the truncated equation.

In fact, it can be shown that if the full equation is a Poisson system, then so is the truncated system. (This is clear form skew-symmetry if $L$ is a constant operator; it remains true in the more general case.)

**Projected dynamics**
The truncated dynamics above differs in principle from the projected dynamics, that is obtained when the original equation is projected into the large scales, and then putting $w = 0$. This leads to

$$\partial_t v = PLP^*\delta_v \hat{H}(v, 0) + PLQ^*\delta_w \hat{H}(v, 0).$$

The main difference is the way the functional is dealt with: first restricting to large scales and then taking the derivative (for truncated dynamics), or taking the projection of the derivative (for the projected dynamics).

**Exercise 74** As a simple (almost trivial) example to see the difference, consider $U = IR^2, V = IR$, and for $u = (x, y), v = x$ the function

$$H(u) = H(x, y) := y \sin x.$$

Then

$$\bar{H}(x) = 0, \quad \textit{for which} \quad \nabla\bar{H}(x) = 0, \quad \textit{while} \quad \nabla H(x, 0) = (0, \sin x).$$

Interpret these formulae geometrically by drawing $\nabla H$ in the plane and considering $\bar{H}$ as defined on the plane.

**Equivalence of the various approximate dynamics**
When the off-diagonal elements vanish, $L_{1,2} = 0, L_{2,1} = 0$, the two equations formulated above coincide, and are given by the truncated dynamics: with $\bar{H}(v) := \hat{H}(v,0) = H(Pu)$, the equation for $v = Pu$ is

$$\partial_t v = \bar{L}\delta_v \bar{H}, \quad \text{with} \quad \bar{L}(v) = PL(v)P^* . \tag{6.9}$$

Of course the equations are also the same if

$$\delta_w \hat{H}(v,0) = 0,$$

but this will usually not be the case.

**Exercise 75** *1.* Show that if $L$ is a differential operator with constant coefficients, the off-diagonal terms, such as $PLQ$, vanish for Fourier decomposition. Show that this is related to the property that *the space of $N$-mode Fourier truncations $F_N$ is closed under differentiation.*

*2.* For $H_0(u) = \int u^2$ and $H_1(u) = \int u^3$ compare the expressions

$$P\delta_u H(u), \quad \textit{with} \quad \delta_v \bar{H}(v), \quad \textit{and with} \quad P\delta_v \bar{H}(v).$$

Is there a difference if one truncates $3u^2$ directly by projection, or if it is seen as the variational derivative of $H_1$? Is there any reason to prefer one of the choices?

**Restriction of first integrals**
Suppose $I$ is a first integral of the complete dynamics. Then it is natural to investigate what remains for the restricted dynamics. Unfortunately,

> *in general it is not true that the restriction of a first integral of the complete dynamics is a first integral of the restricted dynamics.*

However, we have the following exceptions of this general rule.

**Proposition 76** *When the original system is a Poisson system with Hamiltonian $H$, the truncated dynamics is again a Poisson system for which $\hat{H}$ is the Hamiltonian, and hence a first integral for the truncated dynamics.*

The following result gives in a special case the conservation of the Casimir property.

**Proposition 77** *Let $C$ be a Casimir functional for the complete dynamics. Then its restriction $\bar{C}(v) = C(v,0)$ is a Casimir function(al) when the off-diagonal operators vanish.*

**Proof.** Since $C$ is a Casimir for the complete dynamics it holds that

$$L_{11}\delta_v C(v,w) + L_{12}\delta_w C(v,w) = 0$$

for all $(v,w)$, and so certainly for $w = 0$. Assuming $L_{12} = 0$, the result follows.

## 6.2 WAVEPACK: description

WAVEPACK is developed by F. van Beckum & W. Djohan; the following is composed from the manual for version 2.1 (September 1994).
WAVEPACK is a software package under development that is operated on PCs and designed to illustrate and calculate wave phenomena governed by linear and nonlinear partial differential equations. The aims are both instructional and research oriented, with emphasis on *diagnostics* of the phenomena. WAVEPACK is operated through a userfriendly menu-driven interface. Initial data and coefficients in the equations can be varied, as well as the appearance and contents of the graphic display (wave forms at different times, mode dynamics, etc.).
The software is easy to operate and will not be explained further. The software contains several topics; the mathematics of each topic is explained in the "Mathematical Explanation"-command from the menu.

### 6.2.1 KdV Equation, dissipation optional

This section is about the Korteweg-de Vries (KdV) equation. The evolution of the wave profiles will be calculated numerically; each profile being represented by a finite Fourier series. A standard explicit fourth order Runge Kutta method will be used for the time integration.
The KdV equation will be taken in the following form:

$$u_t + u_{xxx} + 6uu_x = 0. \tag{6.10}$$

The term $u_{xxx}$ describes dispersion for infinitesimal waves (the dispersion relation is $\omega(k) = k^3$ ), and $uu_x$ is the nonlinear effect which causes steepening of the wave: wave tops go faster than lower parts.

To include dissipative terms and to enable investigations on the various terms separately, we extend the equation in this section to the following general form:

$$\partial_t u = -\partial_x \left(\alpha u + \beta\, \partial_x^2 u + 3\gamma u^2\right) + \theta\, \partial_x^2 u - \sigma u,$$

from which Korteweg-de Vries (in the above normalized form) is recovered by choosing the coefficients:

$$\alpha = 0, \quad \beta = 1, \quad \gamma = 1, \quad \theta = 0 \quad \text{and} \quad \sigma = 0.$$

The terms with $\theta$ and $\sigma$ represent the addition of dissipation: for $\theta > 0$ viscosity and for $\sigma > 0$ uniform damping is modelled. By changing signs, damping turns into excitation.

In the diagnostics of the dynamics we exploit the fact that KdV can be formulated as a Hamiltonian system

$$\partial_t u = \partial_x \delta_u H(u) \tag{6.11}$$

with $H$ the energy.
For KdV there are in fact infinitely many integrals that are constant during the evolution. Three of them (which will be used in the following) have a clear physical meaning:

$$\begin{aligned} M(u) &= \int u\, dx && \text{(mass)} \\ I(u) &= \int \tfrac{1}{2}u^2\, dx && \text{(horizontal momentum)} \\ H(u) &= \int \left(\tfrac{1}{2}u_x^2 - u^3\right) dx && \text{(energy)} \end{aligned}$$

These integrals are constant for waves decaying sufficiently fast at infinity, or for space periodic waves that we consider here.
The Hamiltonian formulation emphasizes the Hamiltonian structure of the equation. We use a similar description for the numerical discretization to ensure it has the same Hamiltonian structure.

### 6.2.2 Space discretization: Fourier truncation

For the space discretization we consider functions from $\mathcal{U}$:

$$\mathcal{U} = \left\{ u \,\middle|\, u \text{ is } 2\pi \text{ periodic with } \int u = 0 \right\}.$$

and apply Fourier truncation. We use the following notation:

$$\phi_k = \begin{cases} \sin(-kx) & k < 0 \\ 1 & k = 0 \\ \cos(kx) & k > 0 \end{cases}$$

but mean value zero implies that the component $\phi_0$ is never used.
We write the $n$-mode truncation of a function $u$ as

$$\text{Trun}_n\left(u(x,t)\right) = \sum_{k=-n}^{n} \hat{u}_k(t)\,\phi_k \equiv \hat{u}\phi \tag{6.12}$$

where $\hat{u} = (\hat{u}_{-n}, \hat{u}_{-n+1}, \cdots, \hat{u}_{n-1}, \hat{u}_n)$ , without $\hat{u}_0$ , can be seen as an element of $\Re^{2n}$, and $\phi = (\phi_{-n}, \phi_{-n+1}, \cdots, \phi_{n-1}, \phi_n)$ is the vector containing the base functions.
These functions define the truncated space $\mathcal{F}_n$.

### 6.2.3 Restriction of functionals, discretization of space derivative

Inserting (6.12) into the functionals $H$, $I$ and $M$ (i.e. restricting the functionals to $\mathcal{F}_n$), they become functions on $\Re^{2n}$ :

$$\begin{aligned} \hat{H}(\hat{u}) &:= H(\hat{u}\phi) \\ \hat{I}(\hat{u}) &:= I(\hat{u}\phi) = \pi \textstyle\sum_{k=-n}^{n} \frac{1}{2}\hat{u}^2 \\ \hat{M}(\hat{u}) &:= M(\hat{u}\phi) = 2\pi\hat{u}_0 \end{aligned}$$

The expansion for $H$ is rather complicated: it involves a triple summation and is not detailed here.
Note that $\hat{M}(\hat{u}) = 0$ since $u$ has mean value zero.

The variational derivative of $H$ is related to the gradient of $\hat{H}$ as follows.
The partial derivatives of $\hat{H}$ with respect to $\hat{u}_k$ form the vector

$$\partial\hat{H}(\hat{u}) \equiv \left( \frac{\partial\hat{H}}{\partial\hat{u}_{-n}}, \cdots, \frac{\partial\hat{H}}{\partial\hat{u}_n} \right).$$

This vector are the coefficients of the Fourier truncation of $\delta H(u)$:

$$\text{Trun}_n\left(\delta H(u)\right) = \partial\hat{H}(\hat{u})\phi \equiv \sum_{-n}^{n} \frac{\partial\hat{H}}{\partial\hat{u}_k}\phi_k$$

Let $D$ be the $2n$ x $2n$ skew-symmetic matrix that is the discretization of $\partial_x$:

$$\begin{aligned} D &= \text{skew diagonal } (-n, -n+1, \cdots, n-1, n) \\ &= \begin{bmatrix} & & & & & & -n \\ & & & & & -n+1 & \\ & & & & \cdot & & \\ & & & \cdot & & & \\ & & \cdot & & & & \\ & n-1 & & & & & \\ n & & & & & & \end{bmatrix} \end{aligned}$$

### 6.2.4 Spatial discretization

The space discretization of (6.11) becomes:

$$\partial_t \hat{u} = D\,\partial \hat{H}(\hat{u}) \tag{6.13}$$

Component-wise, it can be written as:

$$\partial_t \hat{u}_k = k \frac{\partial \hat{H}}{\partial \hat{u}_{-k}}, \quad -n \leq k \leq n.$$

The discretization above has again a Hamiltonian structure: $D$ is skew-symmetric, and the derivative of the (discretized) energy appears correctly. As a consequence, the discretized energy $\hat{H}(\hat{u})$ is conserved for (6.13).

**Exercise 78** Show that the discretized equation has been obtained by truncation as explained in Section 8.1; show that it is also obtained as the projected dynamics.

Now observe the following:

**Proposition 79** *Fourier truncation (6.12) respects the (continuous) translation symmetry: for any function $u_n \in \mathcal{F}_n$, its translate over an arbitrary distance $\varepsilon$: $u_n(x-\varepsilon)$ also belongs to $\mathcal{F}_n$.*

**Exercise 80** Show that the space spanned by the polynomials $\{1, x, x^2, \ldots, x^n\}$ has the same property. Are there more linear spaces with this property? Show that this property is equivalent to the invariance of $F_n$ for differentiation $\partial_x$.

As a consequence, the discretized momentum $\hat{I}(\hat{u})$ is conserved for the discretized equation too:

$$\frac{d}{dt}\hat{H}(\hat{u}) = 0 \quad \text{and} \quad \frac{d}{dt}\hat{I}(\hat{u}) = 0.$$

**Exercise 81** *Investigate the momentum flow both as an action on the coefficients and on the corresponding elements of $\mathcal{F}_n$.*

### 6.2.5 KdV Equation with Dissipation

The KdV equation with dissipation can be written as follows

$$u_t = -u_{xxx} - 6uu_x + \theta u_{xx} - \sigma u. \tag{6.14}$$

The last two terms in the righthand side describe the dissipation. They consist of a viscous term $\theta u_{xx}$ and a damping term $-\sigma u$. For $\sigma = 0$ and $\theta > 0$, this is the KdV-Burgers equation.
The equation (6.14) can be rewritten in the form

$$\partial_t u = \partial_x \delta H(u) - \delta S(u) \tag{6.15}$$

where $S$ is the functional

$$S(u) = \frac{1}{2}\int \left(\theta u_x^2 + \sigma u^2\right) dx.$$

The Fourier truncation for this equation reads

$$\partial_t \hat{u} = D\partial \hat{H}(\hat{u}) - \partial \hat{S}(\hat{u}),$$

where $\hat{S}$ is the restriction of $S$ to $\mathcal{F}_n$:

$$\hat{S}(\hat{u}) = \frac{1}{2} \sum_{k=-n}^{n} \left(\theta k^2 + \sigma\right) \hat{u}_k^2.$$

The effect of the additional term on the dynamics can be seen from the evolution of integrals like $H$ and $I$ that are no longer constant; for instance for $I$ we have

$$\partial_t I(u) = -2S(u)$$

and similar for $\hat{I}$. This shows the *dissipative* effect when, for instance, both $\theta$ and $\sigma$ are positive. Other choices for the parameters $\theta$ and $\sigma$ may lead to interesting phenomena: *blow up* if both $\theta < 0$ and $\sigma < 0$, and *self-excitation* if lower order modes are unstable ($\theta k^2 + \sigma < 0$ for $k \leq K$, for some $K$) and higher order modes are stable ($\theta k^2 + \sigma > 0$ for $k > K$).

The evolution of both $H$ and $I$ can be seen on the $+$ and the $-$ screen. On the $+$ screen, in an integral diagram $H$ vs $I$ (parametrized by $t$) the relevance of the KdV cnoidal waves can be envisaged. Also depicted is the evolution of the tangent to this curve: $-\frac{dH}{dI}$ is a measure for the propagation speed.

### 6.2.6 KdV Solitons, Optimization Procedure

This topic calculates for the discretized KdV-equation the shape of a travelling wave and its speed. The data generated here, can be written to file in order to be used as input for software experiments as initial condition for the dynamic KdV-topic.

Starting point is that for the continuous system the wave profiles follow as relative equilibria from the constrained critical point problem for $H$ on level sets of $I$ and $M$:

$$\text{Min}_{\ u} \ \{H(u) \,|\, I(u) = I_0, M(u) = 0\}\,. \tag{6.16}$$

Lagranges multiplier rule learns that the critical points satisfy for certain multipliers $c$ and $\mu$ the equation:

$$\delta H(U) = c\delta I(U) + \mu \delta M, \tag{6.17}$$

explicitly:

$$-\partial_x^2 U - 3U^2 = cU + \mu. \tag{6.18}$$

With the wave shape $U$ found in this way, the relative equilibrium solution is given by $U(x - ct)$: the wave travelling at speed $c$.

The solution is depending on only one parameter $I_0$ , which determines the amplitude of the cnoidal wave; by varying $I_0$ a one-parameter family of cnoidal waves is obtained. In general, the multipliers $c$ and $\mu$ depend on $I_0$. In an integral diagram, $H$ vs $I$, the solution of (6.16) as function of $I_0$ determines a curve $h = h(I_0)$ (see figure (6.1)). It may be remarked that the multiplier $c$ (which is the propagation speed of the cnoidal wave) is the tangent to this function $h$:

$$-c = \frac{dh}{dI_0}$$

Hence, (in this case with $M_0 = 0$) the velocity for small amplitude cnoidal waves ($I_0$ small) is negative i.e. to the left, and for large $I_0$ positive, to the right.

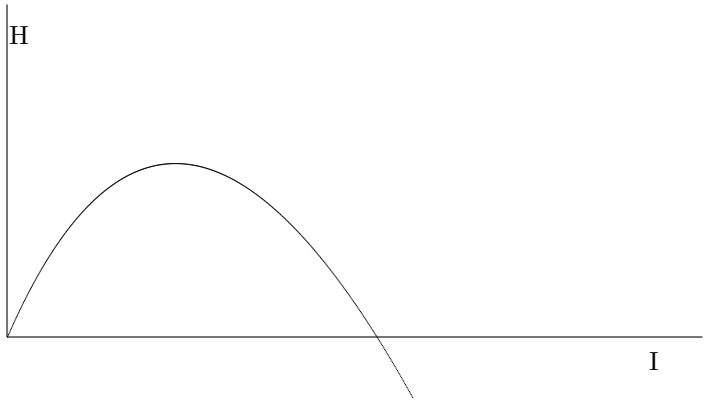


Figure 6.1: Diagram $H$ vs $I$ parametrized by $t$

### 6.2.7 Cnoidal waves for the discretized equation

As a consequence of the fact that for the discretized equation the discretized momentum function $\hat{I}$ is an exact integral, we can investigate relative equilibria for equation (6.13) by extremizing $\hat{H}$ on level sets of $\hat{I}$. The relative equilibrium solution (represented in $\mathcal{U}_n$) will be a solution travelling undisturbed in shape at constant velocity.
In this way, the existence of exact travelling waves for the discretized equation is inherited from the exact KdV equation that admit the *cnoidal waves*.

The "discrete" cnoidals (more precisely: the exact cnoidals of the discrete model) can be calculated with the next WAVEPACK item and then, through file-transport, be used as input for the dynamics.

The discretized version of the constrained optimization problem reads:

$$\text{Min}_{\ \hat{u}} \ \left\{ \hat{H}(\hat{u}) \,|\, \hat{I}(\hat{u}) = \sigma \right\}. \tag{6.19}$$

($\hat{M}(\hat{u}) = 0$ is automatically satisfied by setting $\hat{u}_0$ identically equal to 0). The optimization problem will be solved numerically (see next subsection).

**Remark 82** Once again, a solution of this Fourier truncated optimization problem (6.19),

$$u^{(n)} = \sum_{k=-n}^{n} \hat{u}_k^{(n)} \, \phi_k \tag{6.20}$$

will be an approximation of the corresponding cnoidal wave solution of the full KdV equation (6.10); this approximation will be better for increasing $n$. But at the same time, for any fixed finite value of $n$, the solution (6.20) is a travelling wave for the *discretized equation* with the same value of $n$, of course. It is important to realize, that the wave (6.20) has this property not because it approximates the cnoidal of the exact KdV equation, but because it is itself an exact relative equilibrium of the truncated dynamic equation. □

### 6.2.8 Constrained steepest descent procedure

To find a solution of the (finite dimensional) constrained minimization problem, we use the the constrained steepest descent equation (Section 3.3)

$$d_t\hat{u} = -\nabla\hat{H}(\hat{u}) + \lambda\nabla\hat{I}(\hat{u}) \tag{6.21}$$

with $\lambda$ such that the solution $\hat{u}(t)$ remains on the level set $\hat{I}(\hat{u}) = \sigma$, i.e.

$$\lambda = \frac{\nabla\hat{I}\cdot\nabla\hat{H}}{|\nabla\hat{I}|^2}.$$

When solving this equation numerically, the integration error will make $\hat{u}$ to deviate from the level set $\hat{I}(\hat{u}) = \sigma$. Therefore, after each time step we will make a correction in this sense, and the numerical algorithm for the minimization problem is as follows:

**1.** Set $\sigma$ to the desired value.

**2.** Find a vector $\hat{u}$ such that $\hat{I}(\hat{u}) = \sigma$.

**3.** Repeat until a certain accuracy is obtained:

- Compute $\lambda$.
- Solve $d_t\hat{u} = -\nabla\hat{H}(\hat{u}) + \lambda\cdot\nabla\hat{I}(\hat{u})$ over one time step to find a new $\hat{u}$.
- Scale $\hat{u}$ such that $\hat{I}(\hat{u}) = \sigma$.
- Plot $\hat{H}(\hat{u})$ and $\hat{I}(\hat{u})$ in the $\hat{I},\hat{H}$-plane.

**4.** Display the final $\hat{u}$ and plot the graph of $u(x) = \sum_k \hat{u}_k\phi_k(x)$.

## 6.3 Software Experiments

For the KdV equation the display shows much information for a good diagnostics. An introduction to all this is included in the default example.

### 6.3.1 The default example

Enter the topic KdV equation. The screen you see is called the KdV-screen. In frame 2 you can see: `Fourier modes:4`; this means that the calculation will be done in four mode Fourier truncation. The initial profile is specified in frame 3. It is specially prepared: it represents a cnoidal wave, and you can convince yourself by making a run (through command 1). You see the profile propagating undisturbed in shape and at constant speed. This experiment shows the existence of a travelling wave which is an *exact solution* of the discretized equations.
As the dialogue line says, by pressing "1", you can activate frame 3 to display the current values of Fourier coefficients, Hamiltonian and Momentum, and the time. We see a decay of $H$ and $I$ in the order of $10^{-8}$ per plot step. This is solely due to the numerical time integration method (fourth order explicit Runge Kutta): with smaller time step the error will drastically diminish. *Try time step = 0.002. You may try more values but afterwards please reset to 0.01.*

**The − screen**

The data in frame 3 can be seen in graphic display on an extra screen, the − screen. Stop the run by

pressing <Esc> and press "−" <Enter>. The − screen is displayed and you see three square frames and a long rectangular one. The square frames are for displaying the mode dynamics. Since four Fourier modes are active and only three frames are available you have to select the modes you want to see displayed, as the dialogue line is asking you. (Default is: frame 1, 2, 3 for mode 1, 2, 3.)
As soon as you have given three answers, the calculation starts and you see three perfect circles being drawn (indicating that there is no mode interaction, as can be expected for a cnoidal wave). In the meanwhile the values of $H$ and $I$ are plotted against time in frame 4 (clearly $H$ and $I$ are conserved). You can suppress the display (e.g. to save time) with key "1" for the mode dynamics and key "2" for the $H$, $I$ vs $t$ graph. With key "3" you can toggle the display of $H$ and $I$ values at the bottom of frame 4. (Default they are not refreshed and key "3" only works if key "2" (`plot H&I`) is on. So if you stop the graph you also stop the display of the values.)
You can change the scale of the frames (for zooming in or for making a long time $H, I$ graph) by pressing "6". This will pause the calculation. You select the frame you want to change and you type in the new coordinate bounds in the same way as in earlier topics. Impossible bounds like `X_max < X_min` are not accepted. After having set new coordinate bounds (or not) the calculation can be continued by pressing "6". Or you can start anew by pressing "5".
Return to the KdV-screen (*press <Esc>* <Enter>) and *set* $\theta = 0.05$ (in command 3). Type "−" <Enter> to get the − screen. After pressing <Enter> three times (to get the default allocation of the mode dynamics frames) you see that the circles now change into shrinking spirals; $I$ and $H$ are decreasing ($I$ monotonously, while $H$ first has an overshoot and then approaches zero from above). To repeat the calculation for closer look: *interrupt by pressing "6" and then restart by pressing "5" <Enter>*. You see the rotation speed of the spirals slowing down and even reversing. This indicates a reversal of the wave speed. Check this on the original KdV-screen: *stop the run with <Esc> then <Enter> to get the KdV-screen and start the run by command 1.* (You might like to set `notplot = 25` by command 5.)

**The + screen**

Stop the run by pressing <Esc> and the press "+" <Enter>. The so-called + screen appears and the calculation is started immediately.
Your attention is drawn to the wave in frame 1, which is here at a different scale than on the KdV screen. Next to it (frame 2) you recognize the $H, I$ vs $t$ graph. In frame 3 (bottom left) a curve is created by plotting $H(t)$ vs $I(t)$ at each moment they are available.
The derivative $\frac{dH}{dI}$ of this graph is depicted in frame 4. You can see the relation of it with the graph in frame 2 and 3; e.g. they have an extremum the moment that $\frac{dH}{dI}$ passes zero. There is also relation with frame 1; as mentioned in the theory, the derivative $\frac{dH}{dI}$ equals the wave speed (with minus sign). So when it passes a zero, the wave in frame 1 changes its direction of propagation.

**Remark 83** Frame 3 is in a certain sense a very interesting graph in the diagnostics. At first sight it is very similar to figure 6.1. Note however, that the yellow curve is a trajectory, an evolution, while figure 6.1 is not: it is just a collection of points, each point $Q$ representing one cnoidal wave; and if this cnoidal wave is running according to the KdV equation, the $I$ and $H$ values are constant and $Q$ will not move. The yellow curve on the screen however is a trajectory: $H$ and $I$ change because of dissipation. So if you start in a point $Q$, it will move and an interesting question is: what can we say about the motion ?
Now about the trajectory on the screen we only know for sure that it starts from a cnoidal wave (the default initial condition is chosen to be a cnoidal wave) and that it ends in the origin. So it has two points in common with the cnoidal curve in figure 6.1. In between it is hard to predict the way it goes. Will it keep close to the cnoidal curve ? For reasons of comparison we have in frame 3 prepared the location of the cnoidal curve by a few dots under toggle "C". press "C" and convince yourself

that the dissipative trajectory is indeed very very close to the cnoidal curve. This motivates to study dissipative solutions as approximated by a sequence of cnoidal wave profiles.
Warning. The location of the eight points of the cnoidal wave curve has been prepared for $\beta = 1$ and $\gamma = -1$ only. It would be to much of a job to do so for general $\beta$ and $\gamma$. Because of the restricted applicability we did not mention the toggle "C" in the dialogue line.

### 6.3.2 Viscous decay and uniform damping

This experiment shows the different damping effects of the dissipation parameters $\theta$ and $\sigma$ on the wave profile and on the evolution of $H$ and $I$.

**Exercise 84** Reset all data to the default by quitting from the KdV topic, then enter again.

1. Using command 2, change the initial condition to $u(x, 0) = \cos x + 0.3 \cos 3x$.
2. Take $\theta = 0.2$ , $\sigma = 0$, so KdV-Burgers equation with viscous effect. Note the asymptotic decay to the *first* mode: profiles on screen 1, screen + in $H$ vs $I$ diagram, $H$ and $I$ becoming equal (enlarge the scale). Each initial condition has this property: the cnoidal wave set is asymptotically attractive: the asymptotic states approach the zero state tangent to the cnoidal wave curve. The equation shows the phenomenon of *self-organization* in this sense.
3. Take $\theta = 0$ , $\sigma = 0.3$, so the KdV equation with uniform damping. Note the asymptotic decay: all modes contribute in the final states: the profile shows dispersive interaction for all times; the quotient of $H$ and $I$ is approximately 2. This asymptotic behaviour depends very strongly on the initial data. See, however, the following experiment.
4. Using command R, take the file KDV.DAT as an initial condition. This initial condition is an exact travelling wave solution of the KdV equation in discrete form (you can check it by executing command 1 or +). Set $\theta = 0$ , $\sigma = 0.3$. Using command +, observe the evolution. Observe the change in direction of propagation when $H$ reaches a maximum value (and $dH/dI$ becomes zero). At each time the solution looks roughly like a cnoidal wave, with varying value of $I$, tending to zero. In fact it can be proved that the difference between the dissipative solution and the cnoidal wave with the current value of $I$, divided by the norm of the solution (which tends to zero), is of the order $\sigma$. (See for instance the quotient of $H$ and $I$ in the asymptotics). The same can be proved (with a stronger result for the asymptotics) for KdV-Burgers.

### 6.3.3 Self-excitation.

In the previous experiment you have already seen that positive values of $\theta$ or $\sigma$ will give damping of the wave profile. Contrarily, for negative values these parameters will give an excitation (consider the linearized equation and observe that the first mode is unstable). An interesting case happens if one of these parameters is positive and the other is chosen to be negative. This will prevent the wave profile from decaying to zero. It turns out that the solution is attracted to a state which is a travelling wave. For a travelling wave the values of $I$ and $H$ are constant, and there are no mode interactions. To investigate this phenomenon, do the following experiment.

**Exercise 85** Reset all data to the default by quitting the KdV topic and enter again.

1. Take the cnoidal wave as an initial profile. You can prepare the data by yourself or just take it from file KDV.DAT via command R.

2. Using command 3, set $\theta = 0.2$ and $\sigma = -0.45$.

3. Execute command + (press "+" followed by pressing <Enter>). Using command 6 within the + screen, set the coordinate bounds as follows;
frame 2 : Xmin= -1, Xmax = 30, Ymin = -5.5, Ymax = 5.5,
frame 3 : Xmin= -1, Xmax = 5.5, Ymin = -5.5, Ymax = 1,
frame 4 : Xmin= -1, Xmax = 30, Ymin = -5.5, Ymax = 5.5,
then restart the calculation by command 5.

As you may see on the screen, in the beginning the value of the momentum $I$ is increasing and the Hamiltonian $H$ is decreasing. In frame 1 you can see the wave profile is not a travelling wave anymore since it has deformed in shape and becomes steeper. But after quite a time ($t$ larger than 25), you see in frame 2 the values of $H$ and $I$ tend to become constant. In frame 1 the wave profile is not much changed anymore and becomes a travelling wave, but be aware that the attracting travelling wave is **not** a cnoidal wave of the pure KdV (as you can see from the equation).
You can also investigate the above phenomenon in the − screen as follows.

**Exercise 86** 1. Exit from the + screen by pressing <Esc> followed by <Enter>.

2. Take the − screen (press "−" <Enter>). First you have to allocate mode numbers to the frames 1, 2 and 3; you just type <Enter> three times to accept the default, then the calculation will start immediately.

3. Interrupt the calculation by pressing "6" (without pressing <Enter>), and change the coordinate bounds as follows:
frame 1 : Xmin= -2, Xmax = 2, Ymin = -2, Ymax = 2,
frame 4 : Xmin= -1, Xmax = 50, Ymin = -5.5, Ymax = 5.5,
then restart the calculation by pressing "5" followed by <Enter>.

You see on the screen the mode dynamics are not circles anymore. This indicates that there are interactions between wave modes. But after quite a long time, $t > 20$, (you can view the time value by pressing "3"), the mode dynamics shows perfect circles. This phenomenon can be seen more clearly when you refresh the screen as follows: interrupt the calculation by pressing "6" and then continue by pressing "6" again followed by <Enter>. Soon after the last command, the previous picture will be erased and you get a clean screen and WAVEPACK will continue the calculation.

### 6.3.4 Modulated travelling waves

In the last experiment, when for fixed $\theta = 0.2$ the negative value of $\sigma$ is decresed further, crossing a critical value $\sigma_0$, a Hopf bifurcation takes place. Slightly above this value, there are stable *modulated travelling waves*. This phenomenon can be viewed as follows.

**Exercise 87** Reset all data to the default: exit from the KdV topic and enter again.

1. Using command R, take the initial data from file KDV.DAT.

2. Using command 3, set $\theta = 0.2$ and $\sigma = -0.6845$.

3. Using command 4, set the time step to 0.01.Execute command − (press "−" followed by pressing <Enter>) and adjust the coordinate axes as follows:
frame 1 : Xmin= -4, Xmax = 4, Ymin = -4, Ymax = 4,
frame 2 : Xmin= -3, Xmax = 3, Ymin = -3, Ymax = 3,

frame 3 : Xmin= -2, Xmax = 2, Ymin = -2, Ymax = 2,
frame 4 : Xmin= -1, Xmax = 300, Ymin = -150, Ymax = 30,
then restart the calculation by commmand 5.

In the meanwhile, the mode dynamics frame shows that all modes move as a spiral with radius getting larger and larger. After an initial transition period ($t \geq 35$) the solution seems to be settled down to a travelling wave. (If you need information about the current time of calculation, you can press "3". The use of this key is a toggle.) The first mode dynamics tends to a circle. You can see this more clearly by refreshing the screen via command 6; *pause the calculation by pressing "6", then continue by pressing "6" again followed by <Enter>.*
However, continuing the calculation, the values of both $H$ and $I$ start to oscillate and the mode dynamics (refresh via command 6) shows the appearance of a closed curve consisting of three slightly disturbed circles, rotated somewhat to close precisely. This shows that the solution is not an exact travelling wave, but a travelling wave that is modulated. This modulation is difficult to observe in the profile, but is clear in the diagnostics as described here.

**Remark 88** The phenomenon described here is qualitatively independent of the number of Fourier modes, provided it is greater than 3, and is a property of the excited continuous equation itself.

### 6.3.5 KdV cnoidals

Consider the KdV solitons topic of WAVEPACK.
Frame 1 is used to draw the graph of $(\hat{I}(\hat{u}), \hat{H}(\hat{u}))$, parameterized by $t$, during the calculation of the preceding algorithm. The horizontal axis represents $\hat{I}$ and the vertical axis represents $\hat{H}$.
Frame 3 is used to display numerical information during the minimization process, i.e. the values of the momentum $\hat{I}(\hat{u})$, the Hamiltonian $\hat{H}(\hat{u})$, the Fourier coefficients of the wave profile and the error, which is the Euclidean norm of the difference between two subsequent iterations.

In this experiment, we will explain how to produce a cnoidal wave for the n-mode Fourier truncated the KdV equation. Since the purpose of several commands in this topic is rather different from other topics, we will give the explanation in detail.

**Exercise 89** From the root menu of WAVEPACK, choose the the KdV solitons topic and press <Enter>.

1. Type <Enter> to display the main menu of this topic. Look to command 7 in this menu. This command is used to specify the desired value of the momentum (the constraint $\sigma$ in equation 6.19). Several cnoidal waves, with different value of momentum, can be produced at the same time by giving a list of values for $\sigma$. For example, in the sequel we will calculate two cnoidal waves which have momentum value 1.5 and 4.

2. Give command 7 to the computer. As soon as you give that command, the dialogue line will show a message: number of $\sigma$ :. It means you have to enter the number of cnoidal waves to be calculated. Type 2 followed by pressing <Enter>. Now you have to enter the values of the momentum (the constraint $\sigma$) one by one; being asked the value for $\sigma 1$ you type 1.5 <Enter> and for $\sigma 2$ you type 4 <Enter>. Having finished the data input, WAVEPACK will display the result in frame 2.

3. Now the minimization process can start. Give command 1 to the computer; type "1" <Enter>. As soon as you type the command, some dots will be drawn in frame 1 and the iteration result

is printed in frame 3. You can suppress the iterative display in frame 3 by pressing "1" (see dialogue line). The advantage is that you will get faster result since displaying takes time. Data in frame 3 can be refreshed by pressing "1" once again: it is a toggle.

4. If the calculation has been finished the dialogue line will show its standard invitation for giving a command, i.e. Type your choice or press <Enter> to see menu.

5. Having finished the minimization process, you can display the shape of the cnoidal waves via command 2. Type "2" followed by pressing <Enter>. The display in frame 1 will be changed and now it is prepared for displaying the cnoidal waves shape. The dialogue line offers the following options:
[1] select cnoidal [2] clear screen [3] change coo. bounds [Q] quit

6. Since you want to see a cnoidal wave profile you press "1" <Enter>. As in this experiment we have two cnoidal waves (with momentum $\sigma 1 = 1.5$ and $\sigma 2 = 4$), the dialogue line asks you which one should be displayed:
enter index for $\sigma$ (1..2) :
This means you have the choice between the first and the second cnoidal. Type "1" <Enter>. You see the cnoidal wave profile is shown in frame 1 and the dialogue line returns to the list of options. Now you want to see the second cnoidal. Type "1" <Enter> and then type 2 <Enter>. Now the second cnoidal wave is displayed in a different color. You see it is a more peaked shape: in general, higher momentum values give more sharply peaked cnoidals.

7. If you want to see only one cnoidal on the screen you can clear frame 1 using option 2 and then select the only one you want.

If you have seen enough, leave this session by typing q <Enter>. You will return to the main menu of KdV solitons topic. Data of the cnoidal waves can also be saved as an ASCII (text) file via command S. Later, this data can be used as an initial condition in other topics. Give command S to the computer: type "S" <Enter>. The dialogue line asks you a name for the data file:
Enter file name (<space-bar>=cancel)
Here WAVEPACK is waiting for you to give a filename or to cancel the command. For example, just type cnoidal.dat <Enter>. Now in your WAVEPACK directory there is an additional file called CNOIDAL.DAT which contains the data from the minimization process.

# Bibliography


[1] V.I. Arnol'd, *Mathematical methods of classical mechanics*, Springer 1989 (revised edition).

[2] F. van Beckum & W. Djohan, *WAVEPACK, software package and manual*; version 2.1. September 1994.

[3] C. Carathéodory, *Calculus of Variations and partial differential equations*, Chelsea, New York, 1982.

[4] F.H. Clarke, *Optimization and nonsmooth analysis*, Wiley, New York, 1983.

[5] R. Courant & D. Hilbert, *Methods of Mathematical Physics*, vol 1, 2; Interscience Publishers, New York, 11953, 1962.

[6] I. Ekeland, *Convexity methods in Hamiltonian Mechanics*, Springer, Berlin, 1990.

[7] I. Ekeland and R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam, 1976.

[8] B.A. Finlayson, *The method of weigthed residuals and variational principles*, Academic Press, New York, 1972.

[9] A. Friedman, *Variational principles and free boundary problems*, Wiley, New York, 1982.

[10] I.M. Gelfand & S.V. Fomin, *Calculus of Variations*, Prentice Hall, 1963.

[11] H. Goldstein, *Classical Mechanics*, Addison Wesley, 1980.

[12] Goldstein, *History of the Calculus of Variations*, Springer Verlag, 1980

[13] E. van Groesen & E.M. de Jager, *Mathematical structures in continuous dynamical systems*, North-Holland, Amsterdam, 1994.

[14] M.R. Hestenes, *Optimization theory, The finite dimensional case*, Wiley-Interscience, 1975.

[15] Ioffe & V.M. Tikhomirov, *Theory of Extremal problems*, North-Holland, 1979

[16] C. Lanczos, *The variational principles of mechanics*, Univ. Toronto Press, 1970.

[17] P. Lax, Integrals of nonlinear equations of evolution and solitary waves, *Comm. Pure Appl. Math.* **21**(1968)467-490

[18] V.M. Tikhomirov, *Stories about Maxima and Minima*, American Mathematical Society, 1990.

[19] J.L. Troutman, *Variational calculus with elementary convexity*, Springer, 1983.

[20] *Variatie-rekening*, Vakantiecursus 1985, CWI-syllabus, no. 7, Amsterdam, 1985

[21] Whitham, *Linear and nonlinear waves*, Wiley, 1976.

[22] E. Zeidler, *Nonlinear functional analysis and its applications*, part III, Variational methods and optimization, Springer, 1985.