Variational methods in mathematical physics

Variational Methods In Mathematical Physics

E. W. C. van Groesen

VARIATIONAL METHODS IN MATHEMATICAL PHYSICS

VARIATIONAL METHODS IN MATHEMATICAL PHYSICS

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE TECHNISCHE WETENSCHAPPEN AAN DE TECHNISCHE HOGESCHOOL EINDHOVEN, OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF.DR. P.VAN DER LEEDEN, VOOR EEN COMMISSIE AANGEWEZEN DOOR HET COLLEGE VAN DEKANEN IN HET OPENBAAR TE VERDEDIGEN OP VRIJDAG 8 DECEMBER 1978 TE 16.00 UUR

DOOR

EMBRECHT WILHELMUS CORNELIS VAN GROESEN

GEBOREN TE BREDA

DRUK: WIBRO HELMOND

Dit proefschrift is goedgekeurd
door de promotoren:
prof.dr. L.J.F. Broer
dr. F.W. Sluijter

aan mijn ouders

Contents

PART II: VARIATIONAL DYNAMICAL SYSTEMS

GENERAL INTRODUCTION.

In this thesis we deal with some topics from the theory which is classically called the calculus of variations. The motivation is the fact that a large class of problems from mathematical physics can be given a variational formulation.
To place some of the following in a more general context let us state some standard terminology first. Let $\mathcal{M}$ be a given set of functions ($\mathcal{M}$ is a subset of a metric linear space V) and let f be a functional defined on $\mathcal{M}$.
A *variational principle*, denoted by

$$\text{(1)} \qquad \operatorname*{stat}_{u \in \mathcal{M}} f(u)$$

is the problem of determining all those functions $\bar{u} \in \mathcal{M}$ for which the functional f has a stationary value on $\mathcal{M}$ (i.e. for which there exists some neighbourhood of $\bar{u}$ in $\mathcal{M}$ such that for every $u \in \mathcal{M}$ in this neighbourhood the difference $f(u) - f(\bar{u})$ is of smaller order that the distance from u to $\bar{u}$). Such points $\bar{u}$ are solutions of (1) and called stationary points of f on $\mathcal{M}$.
An *extremum principle*, say the minimum principle

$$\text{(2)} \qquad \inf_{u \in \mathcal{M}} f(u) \ ,$$

asks for the elements $\hat{u} \in \mathcal{M}$ (called minimal points of f on $\mathcal{M}$) for which f takes its smallest value on $\mathcal{M}$, i.e. $f(u) \geq f(\hat{u})$ for all $u \in \mathcal{M}$. To say that a specific problem is described by a variational (extremal) principle means that the solutions of the problem are in a one-to-one correspondence with the solutions of (1) ((2) respectively).
One of the basic problems in the classical theory of calculus of variations is to determine the equation satisfied by the solutions of (1). Assuming the existence of a sufficiently smooth solution, a local investigation (theory of first variation) leads to this

so-called Euler (Euler-Lagrange) equation (or set of equations). In general this is a (partial) differential equation, together with a set of boundary conditions.

Problems which are described by a variational principle are advantegeous above others for several reasons, of which we mention: (i) the notion of generalized solution of the Euler equation is defined in a natural way by bringing the solutions of (1) into a one-to-one correspondence with the generalized solution set of the Euler equation: (ii) a transformation of the Euler equation is usually easier performed via a transformation of the functional, and (iii) Noether's theorem provides us in a simple way, with every continuous group of transformations for which the functional f and the set $\mathcal{M}$ are invariant, with an identity between the Euler expression and a quantity which is a divergence (these identities reduce for stationary points to the "local conservation laws" of dynamical systems).

If it is known that a specific stationary point is a (local) extremum, some additional extremality properties can be derived (theory of second variation). A local analysis which gives the results stated above assumes the existence of a stationary point. As this is no minor point one looks for methods to prove the existence of stationary points for specific cases. There seems to be no unified way to get such results unless some additional information is known (or can be obtained) about the global character of the stationary points. In the simplest case when the problem is described by an extremum principle as (2), the proof of the existence of at least one stationary point may run along the following variational lines. Firstly one shows that the functional f is bounded from below on $\mathcal{M}$. Then one proves that the infimum of f on $\mathcal{M}$ is actually attained at some point $\hat{u} \in \mathcal{M}$. The existence of such a minimal point $\hat{u}$ being proved, a local analysis in the neighbourhood of $\hat{u}$ (if $\hat{u}$ is not isolated) shows that $\hat{u}$ is a stationary point of f on $\mathcal{M}$ (hence $\hat{u}$ is a solution of (1)), and, being a global minimal point of f on $\mathcal{M}$, $\hat{u}$ is also a local extremal point for which some extremality properties hold.

As is well known, apart from an existence statement, an extremum principle (2) may also allow the actual construction of a minimal

element as the limit of a minimizing sequence.
With this short general description we have indicated some important aspects of problems which can be given a variational formulation and emphasized the difference between a local variational principle as (1) and a global extremum principle as (2) with respect to the potential possibility to prove the existence of solutions.

So far we have not specified the set $\mathcal{M}$. The theory of first and second variation is completely standard if $\mathcal{M}$ is the whole linear space V or if $\mathcal{M}$ is an affine set of the form

$$\mathcal{M} = \{ u = u_o + v \mid v \in V_o \}, \tag{3}$$

wherein u_o is a fixed element from V (usually meant to satisfy specified boundary conditions) and V_o is a linear subspace of V (the "set of admissible variations"). In these cases the variational (extremum) principles are said to be unconstrained.
Matters are much more complicated if the set $\mathcal{M}$ is defined as the set of elements which satisfy a given operator equation, e.g.

$$\mathcal{M} = \{ u \in V \mid T(u) = y_o \},$$

wherein T is a (nonlinear) mapping defined on V and y_o is some element from the range of T.
In part I (chapters 1 and 2) we deal with these socalled constrained variational principles.
In chapter 1 we state conditions on f and T which assure that problem (2) has a solution and treat the local theory of first and second variation. The theory of first variation leads to the multiplier rule, a result which in its present generality is due to Lusternik. As a recipe to find this governing equation as the equation for the stationary points of a related unconstrained variational principle, this result is well known and often applied in mathematical physics. Nevertheless, it seems not to be possible to give a convenient reference to a thorough investigation of this local theory [See however the recent monograph of M.S. Berger, Nonlinearity and Functional Analysis, Academic Press 1977, where, in section 3.1 F, this local theory is dealt with in a Hilbert space context].

For the special case that the mapping T is a functional t on V, the multiplier rule states that the stationary points of the constrained extremum principle

$$\inf_{t(u)=p} f(u) \quad , \qquad p \in R\!\!\!Z \quad (u \in V) \tag{4}$$

are also stationary points of the unconstrained variational principle

$$\operatorname*{stat}_{u \in V} [f(u) - \mu t(u)] \tag{5}$$

for some multiplier $\mu \in R\!\!\!Z$. The actual equation for these stationary points can be envisaged as a (nonlinear) eigenvalue problem, with the multiplier μ playing the rôle of eigenvalue. For this reason these variational principles are important for bifurcation theory. In chapter 2 we show that in a number of interesting cases, solutions of (4) can be given several alternative formulations. Using some ideas and notions which stem from the theory of convex analysis, we shall show that with problem (4) there can be associated a dual variational principle which is closely related to unconstrained extremum principles

$$\inf_{u \in V} [f(u) - \mu t(u)] \quad , \quad \mu \in R\!\!\!Z \tag{6}$$

and with which a variational formulation for the multiplier μ of (4) can be given. Furthermore, we investigate when the solutions of (4) are in a one-to-one correspondence with solutions of one of the "inverse" extremum principles

$$\sup_{f(u)=r} t(u) \quad , \quad \inf_{f(u)=r} t(u) \; . \tag{7}$$

An important class of problems which can be formulated by (4) are problems for which a "principle of least energy" holds, with f denoting the energy and t being some constraint. For many specific systems the multiplier μ and the alternative formulations can be given a clear physical interpretation. Despite this fact, a precise investigation of these alternative global characterizations for

solutions of (4) as given here seems to be new.

In part II of this thesis (chapters 3-6) we deal with several classes of dynamical systems whose equations can be derived from a variational principle as (1) wherein $\mathcal{M}$ is essentially as in (3). From a physical point of view these problems are characterized by the fact that one special coordinate (viz. the time) plays a distinguished rôle. Mathematically speaking these problems have the property that no extremum principle of the form (2) is available as the functionals are usually unbounded from below and above on $\mathcal{M}$. Therefore it is not possible to prove the existence of solutions of (1) along the variational lines indicated above. [However, for a restricted class of solutions, such as stationary or steady-state solutions, it may be possible to transform the variational principle to an extremum principle of the form (2) and then prove the existence].
Two main types of variational dynamical systems are Lagrangian and Hamiltonian system, the equationsof which can be described as the stationary points of an action functional defined on configuration space and a canonical action functional on phase space respectively. These systems and some ideas from Classical Mechanics are described in chapter 3. Using the notion of polar functional we show that under some conditionsa Lagrangian system is also a Hamiltonian system and conversely. This result is usually obtained by applying a Legendre transformation to the respective Euler equations (equationsof motion), but using the variational formulation of a Legendre transformation (which is the idea of a polar functional) we derive this result from the variational principles. In this way one is immediately led to the notion of a modified action functional. The corresponding modified action principle is trivially equivalent to the action principle, but its specific form made it possible to recognize some well known variational principles from the theory of fluid dynamics to be of this form, and this led to a constructive way to derive from first principles all variational principles in this field which were previously found in an ad hoc way (see section 3.5 for a short description).
In chapter 4 we consider socalled first order Hamiltonian systems, and investigate the relation with the classical notion of Hamiltonian system. The canonical transformation theory for classical Hamiltonian

systems ceases to be valid for these first order Hamiltonian systems. In fact, we show that merely the requirement that a (non-linear) transformation maps one class of first order Hamiltonian systems into another class of first order Hamiltonian systems almost inevitable leads to the well known Miura transformation, a transformation mapping the (class of higher order) Korteweg-de Vries equation(s) into the (class of higher order)modified KdV equation(s).

In chapter 5 we deal with some problems of a more physical character. For one-dimensional dynamical systems (i.e. with one space variable), one often speaks about (unidirectional) wave propagation. For translational invariant classical Hamiltonian systems there is no preferred direction of propagation in the sense that if there is a solution which may be called unidirectionally propagative, then there exists also a corresponding solution running in the opposite direction. This symmetry is not present in translational invariant first order Hamiltonian systems, and these systems are often calledunidirectionally propagative (e.g. KdV- and BBM-equation). However, because this notion is not explicitly defined in literature, it is difficult to understand the meaning of such statements. Therefore we pose a definition of unidirectional propagativity. This definition has some physical evidence and leads to the acceptable result that for first order *linear* Hamiltonian systems the energy velocity (defined as the velocity of the centre of gravity of the energy density) is a weighted average of the group velocity. Surprisingly enough, for a restricted class of *nonlinear* first order Hamiltonian systems the group velocity of the linearized equations plays an equally important rôle in the exact expression for the energy velocity. With this result we are able to formulate in a precise way in which sense the BBM equation is unidirectionally propagative. Furthermore, in chapter 5 we describe how some classical Hamiltonian systems may approximately be separated into two (unidirectionally propagative) first order Hamiltonian systems, and investigate exact separation for linear systems.

In the final chapter 6 we consider the classical problem of surface waves on a two-dimensional inviscid layer of fluid over a horizontal bottom under influence of gravity. Leaning heavily on the Hamiltonian character of this system, we describe several approximations of the Boussinesq type, comment on their peculiarities and describe for some

of them the approximate separation into two first order Hamiltonian systems.

To conclude this general introduction we have to mention the introductionary chapter 0. This chapter is included to introduce the notation and to facilitate the reading for those who are not acquainted with those standard results from ("non-linear") functional analysis which will be used in the rest of this thesis.

CHAPTER 0: SOME TOPICS FROM FUNCTIONAL ANALYSIS.

0.1. BANACH SPACES AND DUALITY.

0.1.1. INTRODUCTION.

Here and in the rest of this chapter, V and W will stand for Banach spaces (B-spaces) over the scalar field of *real* numbers. The norm will be denoted by $||\ ||$ or, if there is a chance of misunderstanding by $||\ ||_V$ and $||\ ||_W$ respectively.

Convergence (in norm) of a sequence $\{u_n\} \subset V$ to some element $\hat{u} \in V$ will be denoted by $u_n \to \hat{u}$, thus $u_n \to \hat{u}$ in V means $||u_n - \hat{u}||_V \to 0$ for $n \to \infty$.

A mapping from V into W is said to be *bounded* if it maps bounded sets of V into bounded sets of W. The linear space (over the real numbers) consisting of all bounded, *linear* mappings from V into W will be denoted by B(V,W).

DUAL SPACE. Of particular importance is the space which consists of all bounded linear *functionals* defined on V, i.e. $B(V,R\ell)$, which will be denoted by V^*. Supplied with the norm

$$||\ell|| := \sup_{||u|| \leq 1} |\ell(u)| \quad \text{for } \ell \in V^*,\ u \in V,$$

it is a B-space (c.f. section 0.2.1.), and is called the *normed dual* of V. A typical element of V^* is often written as u^*, and its effect on some $u \in V$, $u^*(u)$, as $\langle u^*, u \rangle$. Thus we have for instance

$$(1.1) \qquad ||u^*|| = \sup_{||u|| \leq 1} |\langle u^*, u \rangle| \quad \text{for } u^* \in V^*,\ u \in V,$$

from which it follows that

$$|\langle u^*,u\rangle| \leq ||u^*||\cdot||u|| \quad \forall u \in V \ \forall u^* \in V^*. \tag{1.2}$$

The expression $\langle u^*,u\rangle$ is (by definition) linear in $u \in V$ for fixed $u^* \in V^*$, but it is also linear in $u^* \in V^*$ for fixed $u \in V$. This clarifies the notation $\langle u^*,u\rangle$ for $u^*(u)$ and the adverb "dual" in the term dual space.
The dual space V^* has the following fundamental

PROPERTIES 0.1.1. (i) *V^* separates points on V, i.e. if $u_1,u_2 \in V$ with $u_1 \neq u_2$, there exists $u^* \in V^*$ such that $\langle u^*,u_1\rangle \neq \langle u^*,u_2\rangle$.*

(ii) *For every $u \in V$, $u \neq 0$, there exists $u^* \in V^*$ such that $\langle u^*,u\rangle = 1$ and $||u^*||\cdot||u|| = 1$.*

PROOF: These properties are weak formulations of the Hahn-Banach theorem. See e.g. Rudin [1, theorems 3.3, 3.4, 3.5]. □

In many practical situations, e.g. when V is some function space, one looks for a representation of V^*.

DEFINITION 0.1.2. A *representation* of V^* is a space V_* with elements u_* say, together with a bilinear mapping $[,] : V_* \times V \to R\!\!\!l$ such that the elements of V^* are in a one-to-one correspondence with the functionals

$$[u_*, \ ] : V \to R\!\!\!l \quad , \quad u_* \in V_*. \tag{1.3}$$

In practice, this isomorphism between V^* and the functionals defined by (1.3) is used to identify V_* and V^*. However it must be emphasized that in this case the duality map $\langle,\rangle$ has got a definite meaning!

A very simple representation of V^* can be given if V is a Hilbert space.

THEOREM 0.1.3. *(Riesz representation theorem)*
Let H be a Hilbert space with innerproduct $(,)$. Then H^* can be identified with H if for the duality map the innerproduct is taken. Thus,

if ℓ is any bounded linear functional on H, there exists a unique element $u^* \in H$ such that $\ell(u) = (u^*,u)$ $\forall u \in H$.

PROOF. See e.g. Brown & Page [2, p.348] or Ljusternik & Sobolew [3, p.133]. □

0.1.2. WEAK CONVERGENCE.

The norm on a B-space V induces a topology on V, called the original or norm topology. With this topology, such notions as (norm-) closed and (norm-) compact sets can be defined. However, in many important situations, viz. when V is infinite dimensional, this original topology is too strong in many respects and one wants to deal with a coarser topology. The coarsest topology such that all the functionals

$$<u^*, > : V \to R\ell \quad , u^* \in V^*$$

are continuous (i.e. the topology on V induced by V^*) is called the *weak topology*. This weak topology is of extreme importance, and with it such notions as weak-closure and weak-compactness of a subset of V can be defined. However, because of the limited needs in the rest of this thesis (in fact, mainly dealing with convergence of sequences of elements from V) it is possible to describe the desired results in a somewhat simpler way.

DEFINITION 0.1.4. A sequence $\{u_n\} \subset V$ is said to converge *weakly* to some element $\hat{u} \in V$ if

$$<u^*,u_n> \to <u^*,\hat{u}> \text{ as } n \to \infty \quad , \quad \forall u^* \in V^*.$$

This weak convergence is written as $u_n \to \hat{u}$ in V.

The following results are easy consequences of the foregoing

PROPERTIES 0.1.5. (i) *If $u_n \to \hat{u}$ in V, then $u_n \to \hat{u}$ in V.*
(ii) *If $u_n \to \hat{u}$ in V, then $\{u_n\} \subset V$ is uniformly bounded in V, i.e. there exists a number $m > 0$ such that $||u_n|| \leq m$ $\forall n$.*

(iii) *Weak limits are unique, i.e. if* $u_n \rightharpoonup \hat{u}$ *and* $u_n \rightharpoonup \hat{v}$ *in* V, *then* $\hat{u} = \hat{v}$.

DEFINITION 0.1.6. Let M be a subset of V.

(i) M is *weakly sequentially closed* if for every weakly convergent sequence in M the weak limit belongs to M;

(ii) M is *weakly sequentially compact* if every sequence in M contains a subsequence which converges weakly to some element from M.

As will become clear in section 0.5., B-spaces for which the closed unit ball is weakly sequentially compact are of special importance. B-spaces with this propety are reflexive B-spaces, as shall be shown in the next subsection.

0.1.3. REFLEXIVE B-SPACES.

As we have seen in subsection 0.1.1., the expression

$$\langle u^*, u\rangle \quad , \quad u^* \in V^*, \; u \in V$$

is for fixed $u^* \in V^*$ (by definition) a bounded linear functional on V. With the estimate (1.2) it followsthat the mapping

$$V^* \ni u^* \mapsto \langle u^*, u\rangle \in \mathbb{R}l \quad , \quad u \in V$$

is for every $u \in V$ a bounded linear functional on V^*, i.e.

(1.4) $$\langle \cdot, u\rangle \in (V^*)^* \quad \text{for every } u \in V,$$

where $(V^*)^* = V^{**}$ is the dual space of V^* and is called the *second dual* of V. Functionals of the form (1.4) with u ranging over V define a subspace of V^{**}. If this subspace is the whole of V^{**}, V is called reflexive:

DEFINITION 0.1.7. The B-space V is called reflexive if the *canonical mapping* $\kappa : V \to V^{**}$ defined by

$$\langle \kappa(u), u^* \rangle = \langle u^*, u \rangle \qquad \forall u^* \in V^*$$

maps V *onto* all of V^{**}.

The following theorem can serve as an alternative definition and emphasizes the desired property.

THEOREM 0.1.8. *A B-space V is reflexive if and only if its closed unit ball is weakly sequentially compact.*

PROOF: From Rudin [1, theorem 3.1.2.] it follows that a convex and (norm-) closed set in an arbitrary B-space is closed in the weak topology. From Dunford & Schwartz [4, theorem 6.1] it follows that in an arbitrary B-space a set which is closed in the weak topology, is weakly sequentially compact if and only if it is compact in the weak topology. Hence, in an arbitrary B-space, the closed unit ball is weakly sequentially compact if and only if it is compact in the weak topology. The theorem then follows from Dunford & Schwartz [4, theorem 4.7]. □

As a useful consequence of this concept we state

COROLLARY 0.1.9 *In a reflexive B-space V every bounded sequence* $\{u_n\} \subset V$, *with* $||u_n|| \leq m \quad \forall n$, *has a weakly convergent subsequence, say* $u_{n'} \rightharpoonup \hat{u} \in V$ *and moreover* $||\hat{u}|| \leq m$.

EXAMPLE 0.1.10. It is an immediate consequence of Riesz representation theorem 0.1.3. that every Hilbert space is a reflexive B-space.

The following lemma plays a fundamental rôle in many applications.

DEFINITION 0.1.11. A subset Z^* of V^* is said to be a *complete* set of linear functionals if

$$[u \in V \,,\ \forall z^* \in Z^* \ \langle z^*, u \rangle = 0] \Rightarrow u = 0$$

LEMMA 0.1.12. *Let V be a reflexive B-space. Suppose* Z^* *is a complete set in* V^*. *Then* Z^* *is a dense subset of* V^*.

PROOF: Suppose Z^* is not dense in V^*. Then there exists some $v_o^* \in V^*$, $v_o^* \neq 0$, and a neighbourhood $\Omega(v_o^*) \subset V^*$ of v_o^* such that $\Omega(v_o^*) \cap Z^* = \emptyset$. According to the separation theorem of Hahn-Banach (cf. Rudin [1, theorem 3.5]) there exists $u^{**} \in V^{**}$ such that

$$u^{**}(v_o^*) = 1 \quad \text{and} \quad u^{**}(z^*) = 0 \quad \forall z^* \in Z^*.$$

As V is a reflexive B-space, with $u^{**} \in V^{**}$ there corresponds an element $u \in V$ such that

$$u^{**}(v^*) = \langle v^*, u \rangle \quad \forall v^* \in V^*$$

(cf. definition 0.1.7.). In particular,

$$u^{**}(z^*) = \langle z^*, u \rangle = 0 \quad \forall z^* \in Z^*.$$

As Z^* is a complete set, this implies that $u = 0$, which contradicts the result $u^{**}(v_o^*) = \langle v_o^*, u \rangle = 1$. Hence Z^* must be dense in V^*. □

The foregoing lemma makes it possible in many important situations to *construct* a representation of the dual space for a given reflexive B-space.

COROLLARY 0.1.13. *Let* V *be a reflexive* B*-space, and let* H *be a Hilbert space, with* $(,)_H$ *as innerproduct. Suppose* V *is continuously embedded in* H *(i.e.* $V \subset H$ *and there exists a constant* $c > 0$ *such that* $||v||_H \leq c \cdot ||v||_V$ *for all* $v \in V$; *c.f. subsection* 0.2.3). *Let* H_* *be the completion of* H *with respect to the norm* $||\ ||_{V^*}$:

$$||h||_{V^*} := \sup_{\substack{v \in V \\ v \neq 0}} \frac{|(h,v)_H|}{||v||_V}, \quad h \in H.$$

Then H_* *is a representation of* V^* *with* $(,)_H$ *as duality map.*

PROOF: Let $Z^* := \{(h,\cdot)_H : V \to \mathbb{R} \mid h \in H\}$. Then $Z^* \subset V^*$ as follows from

$$|(h,v)_H| \leq ||h||_{H^*} \cdot ||v||_H \leq c \cdot ||h||_{H^*} \cdot ||v||_V \quad \forall v \in V, \forall h \in H.$$

Moreover, Z^* is a complete set: if $(h,v)_H = 0 \quad \forall h \in H \Rightarrow v = 0$. From lemma 0.1.12 it follows that Z^* is dense in V^*, and the completing operation gives a representation for V^*. □

0.1.4. FUNCTIONSPACES.

We shall now briefly describe some function spaces which will be used in the sequel. Let Ω be an open domain of $R\!\!\!Z^n$. We consider real valued functions defined on Ω. For simplicity we shall restrict to the case $n = 1$ because that is all we shall need, but the following definition and results can be generalized to arbitrary $\Omega \subset R\!\!\!Z^n$ provided the boundary $\partial\Omega$ of Ω is sufficiently smooth.

$\underline{C^m\text{-spaces}}$, $0 \leq m \leq \infty$

$C^m(\Omega)$: the space of functions defined and m-times continuously differentiable in Ω;

$C^m(\bar{\Omega})$ (if Ω is bounded): subspace of $C^m(\Omega)$ consisting of functions all of whose derivatives of order $\leq m$ can be extended as continuous functions to $\bar{\Omega}$. Equiped with the norm

$$||u||_{C^m} := \sum_{k=1}^{m} \sup_{x \in \bar{\Omega}} |\partial_x^k u|$$

it is a B-space (if $m < \infty$);

$C_o^m(\bar{\Omega})$ (if Ω is bounded): subspace of $C^m(\bar{\Omega})$ consisting of functions with compact support in Ω.

With the $||\ ||_{C^m}$-norm this is also a B-space;

$C_o^m(R\!\!\!Z)$: Subspace of $C^m(R\!\!\!Z)$ of functions which have compact support.

$\underline{L_p\text{-spaces}}$, $1 < p < \infty$

$L_p(\Omega)$: space of measurable functions u for which the p-th power of $|u|$ is integrable over Ω. Equipped with the norm

$$||u||_{L_p} := \{\int_{\Omega} |u|^p dx\}^{1/p}.$$

it is a B-space. In particular:

$L_2(\Omega)$ is a Hilbert-space with innerproduct

$$(u,v) = \int_{\Omega} w(x) \cdot v(x)\, dx.$$

It is well known that with the L_2-innerproduct as duality map, the dual space of L_p is the space L_q for appropriate q:

$$(L_p)^* = L_q\ ,\ \frac{1}{p} + \frac{1}{q} = 1\ ,\quad 1 < p,q < \infty.$$

Consequently, L_p-spaces, $1 < p < \infty$, are reflexive B-spaces.

Sobolev-spaces H^m, $0 \leq m < \infty$

$H^m(\Omega)$: space of functions u in Ω such that $\partial_x^k u \in L_2(\Omega)$ for every k, $0 \leq k \leq m$, where ∂_x^k denotes the distributional derivative. Equipped with the innerproduct

$$(u,v)_{H^m} := \sum_{k=1}^{m} (\partial_x^k u, \partial_x^k v)$$

it is a Hilbert-space, and the corresponding norm will be denoted by $|| \; ||_{H^m}$;

$H_o^m(\Omega)$: closure is $H^m(\Omega)$ of $\{u \in C^m(\Omega) |$ u has compact support in $\Omega\}$. With $(,)_{H^m}$ this is also a Hilbertspace.

REMARKS 0.1.14 (i) Note that $H^o(\Omega) = H_o^o(\Omega) = L_2(\Omega)$.

(ii) If $\Omega = R\ell$, we have $H^m(R\ell) = H_o^m(R\ell)$ and $C_o^\infty(R\ell)$ is a dense subset for every $m \geq 0$ (see Treves [5, proposition 13.1]).

(iii) It is to be noted that the Sobolev spaces can also be obtained by a closure operation: $H_{(o)}^m(\Omega)$ is the closure of $C_{(o)}^\infty(\Omega)$ under the norm $|| \; ||_{H^m}$ (c.f. Treves [5, proposition 24.1]).

(iv) As $H_{(o)}^m(\Omega)$ is a Hilbert space, it is a reflexive B-space and the dual space $(H_{(o)}^m(\Omega))^*$ can be identified with $H_{(o)}^m$ itself if for the duality map the innerproduct $(,)_{H^m}$ is taken. However, in many applications it is necessary to consider Sobolev spaces of different order, which would cause to take different duality maps in each case. This inconvenience can be circumvented by taking a fixed bilinear form, usually the L_2-innerproduct $(,)$, as duality map. As $H_{(o)}^m(\Omega)$, $m \geq 0$, is clearly continuously embedded in $L_2(\Omega)$, a representation of $(H_{(o)}^m(\Omega))^*$ with $(,)$ as duality map may be constructed as described in corollary 0.1.13:

Writing $H^{-m} = (H_o^m(\Omega))^*$, H^{-m} is the completion of L_2 with respect to $|| \; ||_{H^{-m}}$:

$$||u^*||_{H^{-m}} = \sup_{\substack{u \in H^m \\ u \neq 0}} \frac{|(u^*,u)|}{||u||_{H^m}}$$

It can be proved that $H^{-m}(\Omega)$ is a space of distributions (Treves [5, proposition 24.2]):

$H^{-m}(\Omega)$, $m \geq 1$: space of distributions in Ω which can be written as finite sums of derivatives of order $\leq m$ of functions belonging to $L_2(\Omega)$.

0.2. OPERATORS ON BANACH SPACES

0.2.1. LINEAR OPERATORS.

The linear space consisting of all bounded, linear operators L from V into W (V and W are B-space) has already been denoted by B(V,W).

THEOREM 0.2.1. B(V,W) *is a* B*-space if equipped with the operator norm:*

$$(2.1)\qquad ||L|| = \sup_{\substack{u \neq 0 \\ u \in V}} \frac{||Lu||_W}{||u||_V} = \sup_{||u||_V = 1} ||Lu||_W, \quad L \in B(V,W).$$

PROOF: See Rudin [1; theorem 4.1] □

For given $L \in B(V,W)$ the expression

$$<w^*, Lv> \ , \ v \in V, \ w^* \in W^*$$

is defined for every $w^* \in W^*$, $v \in V$ and is, for fixed $w^* \in W^*$ a bounded, linear functional on V. This leads one to define the *adjoint* of L:

$$(2.2)\qquad L^* : W^* \to V^*, \ <L^* w^*, v> \ := <w^*, Lv>.$$

It is easily seen that $L^* \in B(W^*, V^*)$, that L^* is uniquely defined by (2.2) and

$$(2.3)\qquad ||L|| = ||L^*||.$$

If V is a reflexive B-space, and $L : V \to V^*$, then $L^* : V \to V^*$. In that case, L is said to be *selfadjoint* if $L = L^*$, i.e. if

$$(2.4)\qquad <v,Lu> = <Lv,u> \qquad \forall u \in V \quad \forall v \in V$$

For $L : V \to W$, the *null-space*

$$\mathcal{N}(L) = \{u \in V | Lu = 0\}$$

and the *range*

$$\mathcal{R}(L) = \{w \in W | \exists u \in V, \ Lu = w\}$$

are linear subspaces of V and W respectively.
Because we shall need these results in the next chapter, we shall des-

cribe here some relations between the null space and range of L and those of its adjoint L^*. Therefore we recall that if N is an arbitrary subspace of V, the *annihilator* $N^\perp$ of N is defined as

$$N^\perp := \{v^* \in V^* \mid \langle v^*, v\rangle = 0 \quad \forall v \in V\} \subset V^*. \tag{2.5}$$

[Note that if V is a Hilbert space H, and V^* is identified with H, $\langle,\rangle$ is the innerproduct of H and $N^\perp$ is the orthogonal complement of N. This specific situation may be a guide for the following manipulations]. If R is an arbitrary subspace of V^*, the annihilator ${}^\perp R$ of R is defined as

$${}^\perp R := \{v \in V \mid \langle v^*, v\rangle = 0 \quad \forall v^* \in R\} \subset V. \tag{2.6}$$

It is easily seen that in general

$$R \subset ({}^\perp R)^\perp,$$

and it can be proved (c.f. Rudin [1, theorem 4.7]) that

$$R = ({}^\perp R)^\perp \text{ if } R \text{ is a } \textit{closed} \text{ subspace of } V^*. \tag{2.7}$$

With these definitions, note that

$$\begin{aligned} {}^\perp\mathcal{R}(L^*) &= \{v \in V \mid \langle v^*, v\rangle = 0 \quad \forall v^* \in \mathcal{R}(L^*)\} \\ &= \{v \in V \mid \langle L^* w^*, v\rangle = 0 \ \forall w^* \in W^*\} \\ &= \{v \in V \mid Lv = 0\} \\ &= \mathcal{N}(L). \end{aligned}$$

Hence, in general

$${}^\perp\mathcal{R}(L^*) = \mathcal{N}(L), \tag{2.8}$$

and with (2.7) it follows that

$$\mathcal{R}(L^*) = \mathcal{N}(L)^\perp \textit{ if } \mathcal{R}(L^*) \textit{ is closed in } V^*. \tag{2.9}$$

Finally we shall need the following result:

$$\mathcal{R}(L) \textit{ is closed in } W \Leftrightarrow \mathcal{R}(L^*) \textit{ is closed in } V^*. \tag{2.10}$$

(c.f. Rudin [1; theorem 4.14]).

0.2.2. CONTINUITY OF OPERATORS.

Now let T be an arbitrary (not necessarily linear) operator from V into W. As we have introduced two concepts of convergence (viz. weak convergence and convergence in norm) there are several notions of continuity, of which we shall need the following ones:

DEFINITION 0.2.2.
(i) T is *continuous* at $\hat{u} \in V$ if for every sequence $\{u_n\} \subset V$ for which $u_n \to \hat{u}$ in V, it follows that $T(u_n) \to T(\hat{u})$ in W.
(ii) T is *strongly continuous* at $\hat{u} \in V$ if for every sequence $\{u_n\} \subset V$ for which $u_n \rightharpoonup \hat{u}$ in V it follows that $T(u_n) \to T(\hat{u})$ in W.
(iii) T is *weakly continuous* at $\hat{u} \in V$ if for every sequence $\{u_n\} \subset V$ for which $u_n \rightharpoonup \hat{u}$ in V it follows that $T(u_n) \rightharpoonup T(\hat{u})$ in W.

REMARK 0.2.3. As is well known, for linear operators the concepts of boundedness and continuity are equivalent. For non-linear operators this is no longer true. Furthermore for linear operators continuity implies weak continuity.

For functionals $f: V \to R\!\!\!l$, the definitions of strong continuity and weak continuity coincide as in $R\!\!\!l$ the concepts of convergence (in norm) and weak convergence coincide. According to custom we define

DEFINITION 0.2.4. The functional $f: V \to R$ is called *weakly continuous* (w.c) at $\hat{u} \in V$ if for every sequence $\{u_n\} \subset V$ with $u_n \rightharpoonup \hat{u}$ in V it follows that $f(u_n) \to f(\hat{u})$ (in $R\!\!\!l$).

In many applications functionals are met which are not w.c. but which have one of the following properties.

DEFINITION 0.2.5. $f: V \to R\!\!\!l$ is called *weakly lower semi-continuous* (w.l.s.c) at $\hat{u} \in V$ if for every sequence $\{u_n\} \subset V$ with $u_n \rightharpoonup \hat{u}$ in V the following inequality holds

$$f(\hat{u}) \leq \liminf f(u_n); \tag{2.11}$$

weakly upper semi-continuity at $\hat{u}$ is defined likewise with (2.11) replaced by

$$f(\hat{u}) \geq \limsup f(u_n).$$

REMARKS 0.2.6 (i) If f is w.c. at $\hat{u} \in V$, then f is both w.l.s.c. and w.u.s.c. at $\hat{u}$ and conversely.
(ii) It is well known that the norm in a Hilbert space H is w.l.s.c., but is not w.c. (if H is infinite dimensional). More generally, if $L : V \to V^*$ is a linear, selfadjoint operator on a reflexive B-space V which satisfies

$$\langle Lu,u\rangle \geq 0 \quad \forall u \in V,$$

the functional $f(u) := \langle Lu,u\rangle$ is w.l.s.c. at all of V.

As a last concept we state

DEFINITION 0.2.7. $F: V \to \mathbb{R}l$ is called *coercive* on V if

$$f(u) \to \infty \quad \text{if } ||u|| \to \infty \quad \text{(uniformly)}$$

$$\text{i.e.} \quad \forall M > 0 \ \exists R > 0 \quad \forall u \in V \quad [||u|| \geq R \Rightarrow f(u) > M].$$

The following peculiar properties show that a w.c. functional can not be coercive:

PROPERTY 0.2.8. *If* $t: V \to \mathbb{R}l$ *is w.c. then for arbitrary* $R > 0$:

$$\inf_{||u|| = R} t(u) = \inf_{||u|| \leq R} t(u)$$

$$\sup_{||u|| = R} t(u) = \sup_{||u|| \leq R} t(u)$$

PROOF: See Vainberg [9, theorem 8.3] □

0.2.3 EMBEDDINGTHEOREMS FOR FUNCTION SPACES.

In subsection 0.1.4 we have introduced some function spaces. At this place we shall describe how some of these spaces are related to each other. These properties can best be described with the aid of embed-

ding operators. If $V \subset W$, the *embedding operator* from V into W (the natural injection) is the identity operator

$$\mathrm{Id}: V \to W$$

which maps each element from V onto the same element considered as an element from W. If V and W are normed spaces, continuity properties of this embedding operator are of particular importance. E.g. if the embedding operator is continuous it is a bounded mapping, which means that there exists a constant $c > 0$ such that

$$||u||_W \leq c||u||_V \quad \forall u \in V.$$

EMBEDDING THEOREM 0.2.9 *Let Ω be a bounded or unbounded interval of $\mathbb{R}^1$.* (i) $H^m_{(o)}(\Omega)$ *is continuously embedded in* $H^k_{(o)}(\Omega)$ *if* $k \leq m$: *thus* $\mathrm{Id}: H^m_{(o)}(\Omega) \to H^k_{(o)}(\Omega)$ *for* $k \leq m$ *and* $||u||_{H^k} \leq ||u||_{H^m}$ $\forall u \in H^m_{(o)}(\Omega)$. *If Ω is bounded, the embedding operator is strongly continuous if* $k < m$: *if* $u_n \to u$ *in* $H^m_{(o)}(\Omega)$, *then* $u_n \to u$ in $H^k_{(o)}(\Omega)$. (ii) $H^m_{(o)}(\Omega)$, $m \geq 1$, *is continuously embedded in* $C^{m-1}_{(o)}(\Omega)$: *thus* $\mathrm{Id}: H^m_{(o)}(\Omega) \to C^{m-1}_{(o)}(\Omega)$ $m \geq 1$, *and* $||u||_{C^{m-1}} \leq c \cdot ||u||_{H^m}$ $\forall u \in H^m_{(o)}(\Omega)$ *for some constant* $c > 0$ *depending only on* m *and* Ω. *If Ω is bounded the embedding operator is strongly continuous: if* $u_n \to u$ *in* $H^m_{(o)}(\Omega)$, *then* $u_n \to u$ *in* $C^{m-1}_{(o)}(\Omega)$.

PROOF: See Sobolew [6, §8 - §11]; see also Treves [5; section 24] □

0.3. DIFFERENTIATION OF OPERATORS.

0.3.1 FRECHET-DERIVATIVE.

Let $T : V \to W$ be an arbitrary operator. The following notion of Frechet derivative is a direct generalization of the special case where $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$.

DEFINITION 0.3.1. The operator T is said to be *differentiable* at $\hat{u} \in V$ if there exist a bounded, linear operator (depending on $\hat{u}$ in general), denoted by $T'(\hat{u})$, from V into W such that

$$||T(\hat{u} + k) - T(\hat{u}) - T'(\hat{u}) \cdot h|| = o(||h||) \quad \text{for } ||h|| \to 0; \tag{3.1}$$

$T'(\hat{u})$ is called the (Frechet-) *derivative of* T *at* $\hat{u}$:

$$T'(\hat{u}) : V \to W \quad , \quad T'(\hat{u}) \in B(V,W). \tag{3.2}$$

If T is differentiable at every point of some set $A \subset V$, T is said to be differentiable on A, and the mapping

$$A \ni u \to T'(\hat{u}) \in B(V,W)$$

is called the *derivative of* T *on* A:

$$T' : A \to B(V,W). \tag{3.3}$$

If this mapping is continuous, T is said to be *continuously differentiable* on A, and we write $T \in C^1(A;W)$. T is said to be continuously differentiable at $\hat{u} \in V$ if there exists some neighbourhood $\Omega(\hat{u}) \subset V$ of $\hat{u}$ such that $T \in C^1(\Omega(\hat{u});W)$.

REMARKS 0.3.2. (i) The operator $T'(\hat{u})$, if it exists, is uniquely determined by (3.1) (c.f. Brown & Page [2, chapter 7]).
(ii) As B(V,W) itself is a B-space (c.f. Theorem 0.2.1) it makes sense to refer to continuity properties of the derivative T'.
(iii) It is easily seen that if T is differentiable at $\hat{u}$, then T is continuous at $\hat{u}$.
(iv) Example: if $T : R\ell^n \to R\ell^m$, let us write $T(x) = (t_1(x),..,t_m(x))$, where $x = (x_1,...,x_n) \in R\ell^n$ and $t_i : R\ell^n \to R\ell$, $i = 1,..,m$. Then T is (Frechet-) differentiable at $\hat{x}$ if t_i is differentiable at $\hat{x}$ for $i = 1,...,m$, and $T'(\hat{x})$ is the n x m Jacobian matrix with elements $[\frac{\partial t_i}{\partial x_k}(\hat{x})]$, $i = 1,..,m;k = 1,..,m$, which has to be envisaged as a bounded, linear mapping from $R\ell^n$ into $R\ell^m$. The derivative T' sends $x \in R\ell^n$ onto the Jacobian matrix evaluated at x.

For the explicit construction of the derivative of a given operator one may advantageously use the following lemma.

LEMMA 0.3.3 *Suppose there exists a bounded linear operator, which we shall again denote by* $T'(\hat{u})$, *such that*

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} [T(\hat{u} + \varepsilon h) - T(\hat{u})] = T'(\hat{u}) \cdot h, \quad \forall h \in V, \tag{3.4}$$

where the limit is taken for real ε *and convergence in the norm of* W *is meant. (This mapping* $T'(\hat{u})$ *is known as the Gateaux derivative of* T *at* $\hat{u}$.) *Furthermore, if* T' *exists in some neighbourhood of* $\hat{u}$ *and is continuous at* $\hat{u}$, *then* T *is (Frechet) differentiable at* $\hat{u}$ *and* $T'(\hat{u})$ *is infact the (Frechet-) derivative of* T *at* $\hat{u}$. *(In other words: a continuous Gateaux derivative is a Frechet derivative.)*

PROOF: See Vainberg [7, theorem 2.1]. □

Finally we note that the *chain-rule* for differentiable operators holds:

THEOREM 0.3.4. *Let* $T : V \to W$ *and* $S : W \to Z$, *where* Z *is another B-space. Suppose* T *is differentiable at* $\hat{u} \in V$ *and* S *is differentiable at* $\hat{w} = T(\hat{u}) \in W$. *Then the composite mapping* $SoT : V \to Z$ *is differentiable at* $\hat{u}$ *and we have*

$$(SoT)'(\hat{u}) = S'(T(\hat{u})) \cdot T'(\hat{u}). \tag{3.5}$$

PROOF: See Brown & Page [2, p.276]. □

0.3.2. HIGHER ORDER DERIVATIVES; TAYLOR EXPANSION.

As B(V,W) itself is a B-space (equipped with the operator norm) one may investigate the differentiability of the operator T' as given by (3.3). Let us suppose for simplicity that T' is defined on all of V:

$$T' : V \to B(V,W).$$

Then T' is differentiable at $\hat{u}$, with derivative which shall be denoted by $T''(\hat{u})$, if

$$T''(\hat{u}) : V \to B(V,W) \tag{3.6}$$

such that

$$||T'(\hat{u}+h) - T'(\hat{u}) - T''(\hat{u})\cdot k|| = o(||k||) \text{ for } ||k|| \to 0,\ k \in V.$$

By definition of operator norm this is equivalent to

$$\sup_{||h|| = 1} ||T'(\hat{u}+k)\cdot h - T'(\hat{u})\cdot h - T''(\hat{u})\cdot k\cdot h|| = o(||k||),$$

thus

$$||T'(\hat{u}+k)\cdot h - T'(\hat{u})\cdot k - T''(\hat{u})\cdot k\cdot h|| = ||h||\cdot o(||k||)$$

From these observations it follows that $T''(\hat{u})$ may also be considered as a *bilinear* mapping from V x V into W

$$(3.7) \qquad T''(\hat{u}) : V \times V \to W, \ V \times V \ni (h,k) \to T''(\hat{u})\cdot h\cdot k \in W$$

which is *symmetric*

$$(3.8) \qquad T''(\hat{u})\cdot h\cdot k = T''(\hat{u})\cdot k\cdot h \quad \forall h \in V, \ \forall k \in V.$$

Of course, $T''(\hat{u})$ is called the *second derivative* of T at $\hat{u}$, and one has the usual formula

$$(3.9) \qquad ||T(\hat{u}+h) - T(\hat{u}) - T'(\hat{u})\cdot h - \tfrac{1}{2}T''(\hat{u})\cdot h\cdot h|| = o(||h||^2)$$
$$\text{for } ||h|| \to 0.$$

Proceding along the same lines one may define higher order-derivatives: the *m-th order derivative* of T at $\hat{u}$, denoted by $T^{(m)}(\hat{u})$ is a *m-linear* operator:

$$T^{(m)}(\hat{u}) : \prod_{i=1}^{m} V \to W.$$

If $T^{(m)}$ exists and is continuous on some subset $A \subset V$, we write

$$T \in C^m(A;W)$$

For differentiable operators, *Taylor expansion* is possible:

THEOREM 0.3.5 *Let* $A \subset V$ *and* $T \in C^{n+1}(A;V)$. *Let* $\hat{u} \in A$ *and* $h \in V$ *such that* $\hat{u}+t\cdot h \in A$ *for every* $0 \le t \le 1$. *Then we have:*

$$(3.10) \qquad T(\hat{u}+h) - T(\hat{u}) = \sum_{m=1}^{n} \frac{1}{m!} T^{(m)}(\hat{u})\cdot \underbrace{h\cdot h\cdot \ldots \cdot h}_{m\text{-times}} + R,$$

where the remainder $R \in W$ *satisfies*

$$(3.11) \qquad ||R|| \le \frac{1}{(n+1)!} \cdot \sup_{0 \le t \le 1} ||T^{(n+1)}(\hat{u}+t\cdot h)||\cdot||h||^{n+1}.$$

PROOF: See Dieudonné [8, sec. 8.14]; see also Vainberg [7;sec.4.6]. □

0.3.3 DIFFERENTIATION OF FUNCTIONALS.

As a special case, the definitions and statements of the foregoing subsections hold equally well if $W = R\!\!\!1$, i.e. if we are dealing with functionals on V. If $f : V \to R\!\!\!1$ is differentiable at $\hat{u} \in V$, the derivative of f at $\hat{u}$ is written as $f'(\hat{u})$ and as $f'(\hat{u}) \in B(V,R\!\!\!1) = V^*$ we may write

$$(3.12) \qquad f(\hat{u}+h) - f(\hat{u}) = \langle f'(\hat{u}),h\rangle + o(||h||) \ , \ ||h|| \to 0.$$

If f is differentiable on a set $A \subset V$, the derivative of f on A

$$f' : A \to V^*$$

is often called the *gradient* of the functional and written as

$$f'(u) = \text{grad } f(u).$$

In a special context also the name functional derivative is used. In this respect we want to make the following remark about a point which may cause some confusion.

REMARK 0.3.6. As was noted before, for a given B-space V there may be several representations of V^*. Connected with this is the observation that for a given functional f the actual form of $f'(u)$ depends on the representation chosen. By way of example consider

$$f : H_o^1(\Omega) \to R\!\!\!1, \ f(u) = \int_\Omega (\tfrac{1}{2}u_x^2 + \tfrac{1}{2}u^2)\, dx.$$

Then f is differentiable at every $u \in H_o^1$ and we have

$$\langle f'(u),v\rangle = \int_\Omega (u_x \cdot v_x + u \cdot v)\, dx, \ u,v \in H_o^1$$

If we take $(H_o^1)^* = H_o^1$ with the innerproduct of H^1 as duality map we have

$$f' : H_o^1 \to H_o^1 \ , \ f'(u) = u.$$

But if we take the L_2-innerproduct as duality map, $(H_o^1)^* = H^{-1}$ (c.f. subsection 0.1.4) and then

$$f' : H_o^1 \to H^{-1} \ , \ f'(u) = -u_{xx} + u.$$

This very simple example expresses the necessity to specify the duality map in these situations.

In most applications from mathematical physics dealing with function spaces, it is custom to take the L_2-innerproduct as duality

map. Because in that case one often speaks about functional derivative, we shall restrict that name to this situation.

DEFINITION 0.3.7. Let V be a function space and V^* the dual space of V with respect to the L_2-innerproduct. If $f : V \to R\!l$ is differentiable at $\hat{u}$, the derivative of f at $\hat{u}$, $f'(\hat{u})$ considered as an element from V^* is called the *functional derivative*, so that we have

$$f(\hat{u}+h) - f(\hat{u}) = \int_\Omega f'(u)\cdot h \, dx + o(||h||) \text{ for } ||h|| \to 0. \tag{3.13}$$

[This functional derivative is often written as $\frac{\delta f}{\delta u}$ but, unfortunately, the same symbol is usually used to denote the functional derivative at the point $u \in V$. (This inadequate notation can be considered to be a straight forward generalization of the imperfect notation $\frac{df}{dx}$ for ordinary functions $f : R\!l \to R\!l$.)]

To complete the specialization to functionals, we note that if f is twice differentiable at $\hat{u}$ then

$$f(\hat{u}+h)-f(\hat{u}) = \langle f'(\hat{u}),h\rangle+\tfrac{1}{2}\langle f''(\hat{u})\cdot h,h\rangle+o(||h||^2) \quad \text{for } ||h|| \to 0 \tag{3.14}$$

and

$$f''(\hat{u}) : V \to V^*$$

may also be considered as a *bilinear* functional on $V \times V$

$$f''(\hat{u}) : V \times V \to R\!l$$

which is *symmetric*

$$\langle f''(\hat{u})\cdot h,k\rangle = \langle f''(\hat{u})\cdot k,h\rangle \quad \forall h \in V \quad \forall k \in V. \tag{3.15}$$

The following result shall be needed in the sequel

LEMMA 0.3.8. *Let* L *be a linear, bounded operator from* V *into* W, *with adjoint* $L^* : W^* \to V^*$ *and let* $f : W \to R\!l$ *be differentiable at* $\hat{w} = L\hat{u}$ *Then the mapping* $foL : V \to R\!l$ *is differentiable at* $\hat{u} \in V$ *and*

$$(foL)'(\hat{u}) = f'(L\hat{u})\cdot L = L^* f'(L\hat{u}) \in V^*. \tag{3.16}$$

PROOF: As L is a linear and bounded mapping, it is differentiable at every $u \in V$ and $L'(u)\cdot h = Lh$ for all $h \in V$. Then the result follows from the chain rule (theorem 0.3.4) and some manipulations with duality maps: if we use subscripts to distinguish between the duality maps of V and W we have:

$$\langle (foL)'(\hat{u}),\ h\rangle_V = \langle f'(L\hat{u})\cdot L'(\hat{u}),h\rangle_V = \langle f'(L\hat{u})\cdot L,h\rangle_V =$$

$$= \langle f'(L\hat{u}),\ L\ h\rangle_W = \langle L^* f'(L\hat{u}),h\rangle_V,$$

valid for arbitrary $h \in V$. Hence the result (3.16). □

0.4. POTENTIAL OPERATORS.

In classical mechanics when dealing with systems which have a finite number of degrees of freedom, one is sometimes interested in the question whether a given force-field $F : R\!\!\!\!\mathbb{l}^n \to R\!\!\!\!\mathbb{l}^n$ is a "conservative" field, i.e. whether there exist a function $f : R\!\!\!\!\mathbb{l}^n \to R\!\!\!\!\mathbb{l}$, usually called the potential, such that

$$f'(x) = F(x) \quad \forall x \in R\!\!\!\!\mathbb{l}^n. \tag{4.1}$$

(If F is represented as $(F_1,\ldots, F_n)$, $F_i : R\!\!\!\!\mathbb{l}^n \to R\!\!\!\!\mathbb{l}$, (4.1) is equivalent to

$$\frac{\partial f}{\partial x_i}(x) = F_i(x),\ i = 1,\ldots,n.).$$

In a more general setting this question is even more important and amounts to the question whether for a given operator $F : V \to W$ there exists a functional $f : V \to R$ such that

$$f'(u) = F(u) \quad \forall u \in V.$$

This question will be answered in the following, and it is shown that the necessary and sufficient condition for the finite dimensional case, viz.

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i} \quad i,j = 1,\ldots,n,$$

generalizes to the more general setting.

DEFINITION 0.4.1. An operator $F : V \to V^*$ is called a *potential operator* (or *gradient operator*) on (the set $A \subset$) V if there exists a differentiable functional $f : V \to R\!\!\!\!\mathbb{l}$ such that

(4.2) $\qquad F(\ddot{u}) = f'(u) \qquad \forall u \in (A \subset) V.$

This functional f is called the *potential* of F on $(A \subset) V$.

THEOREM 0.4.2. *Let* $F : V \to V^*$ *be continuously differentiable on all of* V, *with derivative* $F' : V \to B(V,V^*)$. *Then, in order that* F *be a potential operator it is necessary and sufficient that the bilinear functional*

$$\langle F'(u)\cdot , .\rangle : V \times V \to \mathbb{R}l : (h,k) \to \langle F'(u)\cdot h,k\rangle$$

is symmetric for every $u \in V$ *i.e. that*

(4.3) $\qquad \langle F'(u)\cdot h,k\rangle = \langle F'(u)\cdot k,h\rangle \quad \forall h \in V \quad \forall k \in V.$

Moreover, if (4.3) is satisfied, the potential f *of* F *on* V *is uniquely determined up to an arbitrary constant, and is given by*

(4.4) $$f(u) = f(u_o) + \int_0^1 ds\, \langle F(u_o + s(u-u_o)),\, u-u_o\rangle \quad \forall u \in V,$$

where $u_o \in V$ *is arbitrary. (If* u_o *is chosen to be the zero-element,* (4.4) *simplifies to*

(4.5) $$f(u) = f(0) + \int_0^1 ds\, \langle F(su),u\rangle \qquad \forall u \in V).$$

PROOF: See Vainberg [9, §5]. □

COROLLARY 0.4.3. *If* $T : V \to V^*$ *is a bounded and linear operator, it is a potential operator if and only if* T *is selfadjoint, i.e. if* $T = T^*$, *and in that case its potential up to an arbitrary constant is given by*

(4.6) $\qquad f(u) = \frac{1}{2}\langle Tu,u\rangle.$

PROOF: The requirement that T be selfadjoint, i.e. equation (2.4), is equivalent to the requirement (4.3). Then the potential can be found from (4.5) or verified by differentiation of (4.6). □

REMARK 0.4.4. When dealing with operator equations of the form

$$T(u) = 0, \qquad u \in V \tag{4.7}$$

one is often interested in the question whether this equation can be derived from a variational principle. Formulated in a fairly general way, this amounts to the question whether there exists a functional, say $f : V \to \mathbb{R}$, such that the stationary points of f, i.e. the solution of $f'(u) = 0$ (c.f. subsection 0.5.1), are in some sense related to the solutions of (4.7). The foregoing definition and theorem answer this question only in a very restricted sense. The limited applicability of these results is easily demonstrated: if $f : W \to \mathbb{R}$ is a functional and $L : V \to W$ a linear operator, the operator $L^* \circ f' \circ L : V \to V^*$ is a potential one (with potential $f \circ L$, c.f. lemma 0.3.8), whereas $f' \circ L : V \to W^*$ is not a potential operator (unless $L = c\,I$ for some constant c). Nevertheless, the solution sets of the equations

$$L^* f'(Lu) = 0 \quad \text{and} \quad f'(Lu) = 0$$

are the same if L is a one-to-one mapping.

We shall now describe a simple class of potential operators which will frequently be used in the following.

NEMYTSKY OPERATORS 0.4.5. Let $\gamma : \mathbb{R} \to \mathbb{R}$ be a continuous function and let V be a function space of functions u defined on $\Omega \subset \mathbb{R}$. Then the function $\gamma(u(x))$ is defined on Ω and the mapping

$$G : u \to G(u) \quad \text{where} \quad G(u)(x) := \gamma(u((x)) \qquad \forall x \in \Omega$$

is an operator on V into some function space W, consisting of functions defined on Ω. Operators of this kind are called *Nemytsky operators*, and it can be proved that if γ satisfies an estimate of the form

$$|\gamma(z)| \leq a + b\,|z|^r, \tag{4.8}$$

where a and b are positive constants and $r = p/q$ with $p,q \in [1,\infty)$, then G maps all of $L_p(\Omega)$ into $L_q(\Omega)$ and is continuous and bounded (and conversely, if G maps all of $L_p(\Omega)$ into $L_q(\Omega)$ for some $p,q \in [1,\infty)$, then G is necessarily continuous and bounded and γ satisfies an estimate of the form (4.8); see Vainberg [9; §19]).

For simplycity we shall consider the case where $V = H^1(\Omega)$ and where γ satisfies

(4.9) $\qquad \gamma \in C^1(R\!\!\!l)$ and $\gamma(0) = 0$ if Ω is unbounded.

Then, as the embedding operator $Id : H^1(\Omega) \to C^o(\Omega)$ is continuous (c.f. theorem 0.2.9), G is a mapping from $H^1(\Omega)$ into itself:

(4.10) $$G : H^1(\Omega) \to H^1(\Omega).$$

Moreover, it is easily seen that G is a potential operator on $H^1(\Omega)$ with potential g (chosen to satisfy $g(0) = 0$)

(4.11) $$g(u) = \int_\Omega dx \int_o^{u(x)} \gamma(z)\, dz, \quad u \in H^1(\Omega),$$

for which we have

(4.12) $$\langle g'(u), v\rangle = \int_\Omega dx\, G(u)(x) \cdot v(x) \quad \forall u \in H^1(\Omega)\ , \ \forall v \in H^1(\Omega).$$

Hence G is the functional derivative of g (the L_2-innerproduct as duality map) and the range of G is a subspace of $(H^1(\Omega))^*$, viz. $H^1(\Omega)$ itself.

0.5. FUNCTIONALS ON BANACH SPACES

0.5.1. EXTREME POINTS.

Let f be a real valued functional defined on all of a B-space V. We shall be interested in the range of the functional f, i.e. in the set

$$\{f(u) \mid u \in V\} \subset R\!\!\!l.$$

DEFINITION 0.5.1. A point $\hat{u} \in V$ is called a *local extremum* of f if there exists a neighbourhood $\Omega(\hat{u})$ of $\hat{u}$ in V such that

$f(u) \leq f(\hat{u})$ for all $u \in \Omega(\hat{u})$: f is *maximal* at $\hat{u}$

or $f(u) \geq f(\hat{u})$ for all $u \in \Omega(\hat{u})$: f is *minimal* at $\hat{u}$.

If for this neighbourhood $\Omega(\hat{u})$ the whole space V can be taken, $\hat{u}$ is called a global extremum.

If f is differentiable at $\hat{u}$, then $\hat{u}$ is called a *stationary point* (or a *critical point*) of f if $\quad f'(\hat{u}) = 0$.

The next theorem summarizes the results of the *theory of first and second variation* for functionals on B-spaces.

THEOREM 0.5.2. *Consider* f *on some subset* $\Omega \subset V$, *and let* $\hat{u}$ *be an interior point of* Ω. *Suppose* f *is minimal at* $\hat{u}$.
Then, if f *is differentiable at* $\hat{u}$, $\hat{u}$ *is a stationary point of* f:

$$f'(\hat{u}) = 0, \tag{5.1}$$

and if f *is twice differentiable at* $\hat{u}$, $f''(\hat{u}) : V \times V \to \mathbb{R}l$ *is a non-negative operator in the sense that*

$$\langle f''(\hat{u})\cdot h,h\rangle \geq 0 \qquad \forall h \in V. \tag{5.2}$$

PROOF: Let $h \in V$ be arbitrary and consider

$$\varphi(t) := f(\hat{u} + th).$$

As $\hat{u}$ is an interior point of Ω, φ is defined in some neighboorhood of $t = 0$. Moreover, φ is (twice) differentiable at $t = 0$ if f is (twice) differentiable at $\hat{u}$, and we have

$$\frac{d\varphi}{dt}(0) = \langle f'(\hat{u}),h\rangle, \quad \frac{d^2\varphi}{dt^2}(0) = \langle f''(\hat{u})\cdot h,h\rangle.$$

As f is minimal at $\hat{u}$, φ must be minimal at $t = 0$, and thus

$$\frac{d\varphi}{dt}(0) = 0, \quad \frac{d^2\varphi}{dt^2}(0) \geq 0.$$

Then (5.1) and (5.2) follow because $h \in V$ is arbitrary. □

The inequality (5.2) may be envisaged as a necessary condition for a stationary point $\hat{u}$ to be minimal. It is also possible to give a sufficient condition.

THEOREM 0.5.3. *Let* f *be twice continuously differentiable at the stationary point* $\hat{u}$, *and suppose that there exists some constant* $c > 0$ *such that*

$$\langle f''(\hat{u})\cdot h,h\rangle \geq c\|h\|^2 \qquad \forall h \in V.$$

Then $\hat{u}$ *is a minimal point of* f, *and moreover there exists a neighbourhood* $\Omega(\hat{u})$ *of* $\hat{u}$ *such that*

$$f(u) - f(\hat{u}) \geq \tfrac{1}{4}c||u-\hat{u}||^2 \qquad \forall u \in \Omega(\hat{u}).$$

PROOF: Writing $u = \hat{u}+h$ the statements immediately follow from

$$\begin{aligned} f(u)-f(\hat{u}) &= \langle f'(\hat{u}),h\rangle + \tfrac{1}{2}\langle f''(\hat{u})\cdot h,h\rangle + o(||h||^2) \text{ for } ||h|| \to 0 \\ &= \tfrac{1}{2}\langle f''(\hat{u})\cdot h,h\rangle + o(||h||^2) \\ &\geq \tfrac{1}{2}c||h||^2 + o(||h||^2) \\ &\geq \tfrac{1}{4}c||u-\hat{u}||^2 \text{ for } ||u-\hat{u}||=||h|| \text{ sufficiently small } \square. \end{aligned}$$

0.5.2. EXISTENCE OF EXTREME POINTS

If V is a finite dimensional space, Weierstrasz' theorem states that a continuous function on a bounded and closed subset, e.g. the closed unit ball, is bounded from above and from below and attains its maximum and minimum on that set. But if V is aninfinite dimensional space, a closed and bounded set needs not to be compact and Weierstrasz' theorem ceases to be valid. However, we know that in a reflexive B-space the closed unit ball is weakly (sequentially) compact (c.f. theorem 0.1.8). By requiring a functional to be continuous with respect to weak convergence, Weierstrasz' theorem may be generalized as shall be shown.

THEOREM 0.5.4. *Let* V *be a reflexive* B-*space and* $\Omega \subset V$ *a bounded and weakly sequentially closed subset. Let* $f : V \to \mathbb{R}l$ *be w.l.s.c. on* Ω. *Then* f *is bounded from below on* Ω *and attains its infinum at some point* $\hat{u} \in \Omega$.

PROOF: The proof is standard and will be given as an illustration of some concepts introduced earlier.
Suppose first that f is not bounded from below on Ω. Then there exists a sequence $\{u_n\} \subset \Omega$ such that $f(u_n) \to -\infty$ for $n \to \infty$. As Ω is bounded, this sequence is bounded and has a weakly convergent subsequence (corollary 0.1.9), say $u_{n'} \rightharpoonup \hat{u}$ in V. As Ω is weakly sequentially closed, $\hat{u} \in \Omega$. For this subsequence we also have $f(u_{n'}) \to -\infty$ for $n' \to \infty$. But as f is w.l.s.c., $f(\hat{u}) \leq \liminf f(u_{n'}) = -\infty$, which is impossible. Hence f is bounded from below on Ω.

Now, let $\alpha := \inf \{f(u) | u \in \Omega\}$. Then $\alpha > -\infty$, and there exists a minimizing sequence $\{u_n\} \subset \Omega$ for which $f(u_n) \to \alpha$ for $n \to \infty$. Again this sequence is bounded and has a weakly convergent subsequence, say $u_{n'} \rightharpoonup \hat{u}$ with $\hat{u} \in \Omega$. As $f(u_{n'}) \to \alpha$ for $n' \to \infty$, and as f is w.l.s.c., we have $f(\hat{u}) \leq \liminf f(u_{n'}) = \alpha$. By definition we also have $f(\hat{u}) \geq \alpha$. Consequently $f(\hat{u}) = \alpha$, which means that f attains its infimum at $\hat{u}$. This completes the proof. □

REMARK 0.5.5. From the proof it is easily seen that theorem 0.5.3. remains valid if the requirement Ω is bounded is replaced by the requirement that f is *coercive* on V, i.e.

$$f(u) \to \infty \quad \text{for } ||u|| \to \infty, \ u \in V.$$

THEOREM 0.5.6. *(Generalized Weierstrasz' theorem).*
Let V *be a reflexive* B*-space and* Ω *a bounded and weakly sequentially closed subset of* V. *Let* $t : V \to R\!\!l$ *be w.c. Then* t *is bounded from above and from below on* Ω *and attains its infimum and its supremum at points of* Ω.

PROOF: If g is a functional which is w.u.s.c., it follows from theorem 0.5.3., applied to $f = -g$, that g is bounded from above on Ω and attains its supremum at some point of Ω. With this observation the theorem easily follows: as t is w.c. it is both w.l.s.c. and w.u.s.c. (c.f. remark 0.2.6.). □

0.6. POLAR FUNCTIONS AND SUBDIFFERENTIABILITY.

In this section we shall briefly describe some notions from the theory of Convex Analysis which will be used in chapters 2 and 3. We consider the simplest case first (functions defined on $R\!\!l$); an extension to functionals on a reflexive B-space is then an easy generalization.

Let h be a function defined on $R\!\!l$

$$h : R\!\!l \to \overline{R\!\!l}.$$

Here $\overline{R\!\!l}$ is the extended real line, i.e. $\overline{R\!\!l} = R\!\!l \cup \{-\infty\} \cup \{\infty\}$.
(Allowing functions taking infinite values turns out to be useful in

what follows. For instance, if g is a function defined on some interval $I \subset R\ell$, we have

$$\inf_{x \in I} g(x) = \inf_{x \in R\ell} h(x)$$

if we agree to set

$$h(x) = \begin{cases} g(x) & \text{for } x \in I \\ \infty & \text{for } x \notin I \end{cases}.$$

DEFINITION 0.6.1. The *polar function* (or *conjugate* function, or *supporting* function) of h is a function $h^* : R\ell \to \overline{R\ell}$ defined by

$$h^*(\mu) = \sup_{p \in R\ell} [\mu p - h(p)] , \quad \mu \in R\ell. \tag{6.1}$$

If for some $\mu \in R\ell$, $h^*(\mu)$ is finite this number has a clear interpretation: it is the smallest value of α such that the linear function with slope μ, i.e. $\mu p - \alpha$, is dominated by h(p):

$$h(p) \geq \mu p - \alpha \qquad \forall p \in R\ell.$$

DEFINITION 0.6.2. If the linear function $\bar{\mu}p - h^*(\bar{\mu})$ is *exact* at some $\bar{p} \in R\ell$, i.e. if

$$\bar{\mu}\bar{p} - h^*(\bar{\mu}) = h(\bar{p}), \tag{6.2}$$

we shall say that h is *subdifferentiable* at $\bar{p}$ and $\bar{\mu}$ is called a *subgradient* of h at $\bar{p}$. The set of all subgradients at $\bar{p}$ is called the *subdifferential* of h at $\bar{p}$, and this set is written as $\partial h(\bar{p})$. We write $\bar{\mu} \in \partial h(\bar{p})$ if $\bar{\mu}$ is a subgradient of h at $\bar{p}$, and $\partial h(\bar{p}) = \emptyset$ if h is not subdifferentiable at $\bar{p}$ (i.e. if there exists no linear function which is exact at $\bar{p}$ and dominated by the function h).

REMARK 0.6.3. It must be noted that subdifferentiability of h at $\bar{p}$ has nothing to do with the smoothness of the function in a neighbourhood of $\bar{p}$ but depends crucially on the global behaviour of h. This in sharp contrast to the more familiar concepts of differentiation.

DEFINITION 0.6.4. The *bipolar* function h^{**} of h is defined as the po-

lar function of h^*:

$$h^{**}: R\ell \to \overline{R\ell},\ h^{**}(p) = \sup_{\mu \in R\ell} [\mu p - h^*(\mu)].$$

The polar function is defined (as a function into $\overline{R\ell}$) for arbitrary functions h, but this notion is especially useful when h is a convex function.

DEFINITION 0.6.5. The function $h : R\ell \to \overline{R\ell}$ is *convex* if h satisfies

$$h(\lambda p+(1-\lambda)q) \leq \lambda \cdot h(p)+(1-\lambda)\cdot h(q) \quad \text{for all } \lambda, 0<\lambda<1,$$

for all values of p and q for which both h(p) and h(q) are finite.

The following properties are direct consequences of the foregoing definition.

PROPERTIES 0.6.6. *Let* $h : R\ell \to \overline{R\ell}$ *be an arbitrary function.*

(a) *The polar function* $h^*: R\ell \to \overline{R\ell}$ *is a convex function.*

(b) $h^{**}(p) \leq h(p)$ *for all* $p \in R\ell$.

(c) *The polar function of* h^{**} *equals* h^*: $h^{***} = h^*$.

(d) $\bar{\mu} \in \partial h(\bar{p})$ *if and only if* (i) $h(\bar{p})$ *is finite*

(ii) $\bar{\mu}(q-\bar{p}) + h(\bar{p}) \leq h(q) \quad \forall q \in R\ell$.

(e) *if* $\partial h(\bar{p}) \neq \emptyset$, *then* $h(\bar{p}) = h^{**}(\bar{p})$.

(f) *if* $h(\bar{p}) = h^{**}(\bar{p})$ *then* $\partial h(\bar{p}) = \partial h^{**}(\bar{p})$.

(g) *if* $\partial h(\bar{p}) \neq \emptyset$, *and if* h *is differentiable at* $\bar{p}$ *then* $\partial h(\bar{p})=\{h'(\bar{p})\}$.

PROOF: See Rockafellar [10, part III, V]. □

REMARK 0.6.7. There is a close analogy between the *Legendre transform* of a smooth, convex function $h \in C^2(R\ell)$ and the polar function h^*: if $I := \{h'(p) \mid p \in R\ell\}$, then the restriction of h^* to I is the Legendre transform of h, and $h^{**} = h$ on all of $R\ell$ (See Rockafellar [10, section 26]; c.f. also section 2.2).

The foregoing definitions and properties are easily generalized for functionals defined on a reflexive B-space.
Let V be a reflexive B-space, V^* its dual and $\langle , \rangle$ the duality map.
Let $f : V \to \overline{R\ell}$ be a functional, possibly taking infinite values.

DEFINITION 0.6.8. The *polar-functional* of f is a functional $f^*: V^* \to \overline{R\ell}$ defined by

$$f^*(u^*) = \sup_{u \in V} [\langle u^*, u\rangle - f(u)] , \quad u^* \in V^*.$$

The *bipolar-functional* $f^{**} : V \to \overline{R}l$ is the polar functional of f^*:

$$f^{**}(u) = \sup_{u^* \in V^*} [\langle u^*, u\rangle - f^*(u^*)] , \quad u \in V.$$

The functional f is said to be *subdifferentiable* at $\bar{u} \in V$ with subgradient $\bar{u}^* \in V^*$ if

$$\langle \bar{u}^*, \bar{u}\rangle - f^*(\bar{u}^*) = f(\bar{u}).$$

The definition of *subdifferential* and the notation $\bar{u}^* \in \partial f(\bar{u})$ are defined as in definition 0.6.2.

With the obvious changes, the properties listed in 0.6.6. can be shown to hold in this case too. (c.f. Ekeland & Temam [11, Ch.1, section 4,5]; see also Vainberg [7; section 8.4]) As we shall not need these results in the rest of this thesis, we shall not pursue this subject any further here.

PART I: CONSTRAINED EXTREMUM PRINCIPLES.

CHAPTER 1: EXISTENCE AND LOCAL ANALYSIS.

1.1. INTRODUCTION.

In this chapter we shall deal with the following problem. Let V and Y be two reflexive B-spaces, and let $f : V \to R\!l$ be a real valued functional defined on all of V. Furthermore, let $T : V \to Y$ be an operator (generally non linear), and $y_o \in Y$. We shall be concerned with the range of the functional f on the set of points $u \in V$ for which $T(u) = y_o$. More particularly we shall study the existence and look for possible characterizations of solutions (c.f. definition 1.1.1. below) of what shall be called

Problem $\mathcal{P}$

$$(1.1)\ \mathcal{P}: \quad \inf \{f(u) \mid u \in V\ ,\ T(u)=y_o\}$$

Minimization problems of this kind are called *constrained extremum problems*, in contradistinction to unconstrained minimization problems for a functional f when f in considered on all of the space V. We shall use the name *manifold* for the set of points

$$(1.2) \qquad \mathcal{M} := \{u \in V \mid T(u) = y_o\}\ .$$

DEFINITION 1.1.1. An element $\hat{u} \in \mathcal{M}$ is said to be a *constrained global minimum point* of f with respect to $\mathcal{M}$ if

$$(1.3) \qquad f(\hat{u}) \leq f(u) \quad \text{for all } u \in \mathcal{M}.$$

An element $\hat{u} \in \mathcal{M}$ is said to be a constrained *local* minimum point of f with respect to $\mathcal{M}$ if there exists a neighbourhood $\Omega(\hat{u})$ of $\hat{u}$ in V such that

$$f(\hat{u}) \leq f(u) \quad \text{for all } u \in \mathcal{M} \cap \Omega(\hat{u}) \tag{1.4}$$

Furthermore, $\hat{u} \in \mathcal{M}$ is said to be a *solution of problem* $\mathcal{P}$ if and only if $\hat{u}$ is a constrained global minimum point of f on $\mathcal{M}$, and then the infimum of f on $\mathcal{M}$ is attained at $\hat{u}$ and we write

$$f(\hat{u}) = \inf \mathcal{P}. \tag{1.5}$$

In section 1.2 we shall state conditions for the functional f and the operator T which assure that problem $\mathcal{P}$ has at least one solution. Comparing the value of f at a solution $\hat{u}$ of problem $\mathcal{P}$ with the value of f at neighbouring points of $\hat{u}$ on the manifold lead in first order to an equation for $\hat{u}$ (the multiplier rule, section 1.4) and in second order to a statement about the non-negativity of a certain operator (section 1.5). A local investigation of this kind is an adapted version of the theory of first and second variation for unconstrained minimization problems (c.f. theorem 0.5.2). However, in order that such a local investigation is possible it is necessary that the manifold $\mathcal{M}$ is sufficiently "regular" in a neighbourhood of the point $\hat{u}$. This will be studied in more detail in section 1.3. Because these results are obtained by an investigation which is essentially local in character, they also hold for points which are local, but not global minimum points.
Because of its importance for the rest of this thesis, the foregoing results are specialized in section 1.6 to the case where $Y = R^l$, i.e. where the operator T is a functional $t : V \to R^l$. In section 1.7 some physical applications of the theory are presented.
The multiplier rule as stated in section 1.4. is originally due to Lusternik [12] (see also Maurin [13] and Curtain & Pritchard [14, section 12.4]). For functional constraints, Vainberg [7, sections 9.4 , 9.5] is a most familiar reference. For constraints described by certain differential equations, Klötzler [15] proves the multiplier rule and treats the theory of second variation.

1.2. AN EXISTENCE RESULT.

To assure that problem $\mathcal{P}$ makes sense, we demand that the manifold $\mathcal{M}$ is non-void, which means that y_o must be in the range of the operator T: $y_o \in \mathcal{R}(T)$.

THEOREM 1.2.1. *Suppose that the manifold $\mathcal{M}$ is weakly sequentially closed. Let* $f: V \to Rl$ *be weakly lower semi-continuous , and coercive on $\mathcal{M}$, i.e.*

$$f(u_n) \to \infty \quad \text{if } u_n \in \mathcal{M},\ \|u_n\| \to \infty. \tag{2.1}$$

Then f *is bounded from below on $\mathcal{M}$ and* f *attains its infimum, i.e. problem $\mathcal{P}$ has a solution.*

PROOF: This result is an easy consequence of theorem 0.5.4. and a slightly modified version of remark 0.5.5. □

REMARK 1.2.2. If $\mathcal{M}$ is known to be a bounded subset, condition (2.1) is void.

In the following lemma a condition for the operator T is given that assures that $\mathcal{M}$ is weakly sequentially closed. The formulation has been chosen to be directly applicable for a specific problem to be dealt with in section 1.7.

LEMMA 1.2.3. *Let* $T : V \to Y$ *with* Y *continuously embedded in a reflexive* B*-space* Z. *Suppose* T *satisfies*

$$\text{if } u_n \rightharpoonup \hat{u} \text{ in } V, \text{ then } T(u_n) \rightharpoonup T(\hat{u}) \text{ in } Z. \tag{2.2}$$

Then the manifold $\mathcal{M}$ is weakly sequentially closed.

PROOF. Let $\{u_n\} \subset \mathcal{M}$ be a sequence in $\mathcal{M}$, and $u_n \rightharpoonup \hat{u}$ in V. We have to show $\hat{u} \in \mathcal{M}$. As $u_n \in \mathcal{M}$, we have $T(u_n) = y_o \ \forall n$ and hence $T(u_n) \to y_o$ in y. By the continuous embedding of y into Z: $T(u_n) \to y_o$ in Z. Furthermore because of (2.2) we also have $T(u_n) \rightharpoonup T(\hat{u})$ in Z. Thus

$T(\hat{u}) = y_o$, which shows that $\hat{u} \in \mathcal{M}$. □

An immediate consequence of the foregoing lemma is:

COROLLARY 1.2.4. *If* $T: V \to Y$ *is weakly continuous,* $\mathcal{M}$ *is weakly sequentially closed.*

PROOF: Take Z = Y in lemma 1.2.3. □

1.3. REGULAR POINTS OF THE MANIFOLD

At this place we shall study the structure of the manifold

$$\mathcal{M} = \{u \in V \mid T(u) = y_o\} \tag{3.1}$$

in the neighbourhood of some point $\hat{u} \in \mathcal{M}$.
For the following we shall assume that $T: V \to Y$ is continuously differentiable at $\hat{u}$ and write

$$T(\hat{u}+h) - T(\hat{u}) = T'(\hat{u}).h + \omega(\hat{u};h), \tag{3.2}$$

where

$$||\omega(\hat{u};h)|| = o(||h||) \quad \text{for } ||h|| \to 0 \text{ in } V. \tag{3.3}$$

DEFINITION 1.3.1. A point $\hat{u} \in \mathcal{M}$ is said to be an *isolated point* of the manifold $\mathcal{M}$ if there exists a neighbourhood $\Omega(\hat{u})$ of $\hat{u}$ in V such that

$$\Omega(\hat{u}) \cap \mathcal{M} = \{\hat{u}\}. \tag{3.4}$$

LEMMA 1.3.2. *If* $T'(\hat{u}): V \to Y$ *is boundedly invertible, then* $\hat{u}$ *is an isolated point of manifold.*

PROOF: Suppose $\hat{u} + h \in \mathcal{M}$. Then by (3.2) we have

$$T'(\hat{u}).h + \omega(\hat{u};h) = 0.$$

Applying $T'(\hat{u})^{-1}$ to this equation, it is seen that h must satisfy the inequality

$$||h|| \leq ||T'(\hat{u})^{-1}||.||\omega(\hat{u};h)||.$$

With (3.3) it follows that there exists some positive number δ such that $h = 0$ is the only element which satisfies this inequality together with $||h|| < \delta$. Hence the set

$$\Omega(\hat{u}) := \{\hat{u} + h \mid h \in V, ||h|| < \delta\}$$

is a neighbourhood of $\hat{u}$ which satisfies (3.4). This proves the lemma. □

A local investigation of the manifold in a neighbourhood of a non-isolated point $\hat{u}$ means to characterize the set

$$\{h \in V \mid \hat{u} + h \in \mathcal{M}, \ ||h|| \text{ small}\}.$$

This amounts to a study of all small norm solutions of the non-linear operator equation

$$T(\hat{u} + h) - T(\hat{u}) = 0 . \tag{3.5}$$

To make any progress in this direction, and in view of the foregoing lemma, we suppose that the null space of the operator $T'(\hat{u})$ is non-trivial:

$$\hat{\mathcal{N}} := \mathcal{N}(T'(\hat{u})) \neq \{0\}.$$

Clearly, $\hat{\mathcal{N}}$ is a vector subspace of V. Hence there exists a *projection operator* P defined on V with $\mathcal{R}(P) = \hat{\mathcal{N}}$. (We recall that P is a projection operator if P is linear and $P^2 = P$). Moreover, denoting the null space of P by W, V is the *algebraic direct sum* of $\hat{\mathcal{N}}$ and W:

$$V = \hat{\mathcal{N}} + W,$$

which means that with every element $h \in V$ there correspond unique

elements $v(= Ph)$ and $w(= (I - P)h)$ such that

$$h = v + w, \quad v \in \hat{\mathcal{N}}, \quad w \in W. \tag{3.6}$$

Of course the mapping: $\hat{\mathcal{N}} \times W \ni (v,w) \to h \in V$ is continuous. If the inverse of this mapping is also continuous (which is equivalent to saying that the projection operator P is continuous), the decomposition of V is called a *topological* direct sum, and written as

$$V = \hat{\mathcal{N}} \oplus W. \tag{3.7}$$

For the analysis to follow it will be necessary to have this stronger concept of topological direct sum.
As $T'(\hat{u}) \in B(V,Y)$, the null space $\hat{\mathcal{N}}$ is a *closed* subspace of V. Closedness of $\hat{\mathcal{N}}$ being a necessary condition for the result (3.7) is in general (reflexive) B-spaces not sufficient (c.f. Dieudonné [8, section 5.4]). However, in many important situations there is such a topological direct sum.

HYPOTHESIS 1.3.3. At least one of the following statements is true:
(a) V is Hilbert space;
(b) dim. $\hat{\mathcal{N}} < \infty$;
(c) codim. $\hat{\mathcal{N}}$ (=dim. $V/\hat{\mathcal{N}}$) $< \infty$.

LEMMA 1.3.4. *If hypothesis* 1.3.3. *is satisfied, the closed subspace* $\hat{\mathcal{N}}$ *is a topological direct summand , i.e. there exists a (closed) subspace* W *such that* (3.7) *holds.*

PROOF: If statement (a) of hypothesis 1.3.3. holds, the contents of the lemma is the well-known *projection-theorem for Hilbertspaces* (see e.g. Rudin [1, theorem 12.4], Dieudonné [8, section 6.3]). In the other cases the lemma follows essentially from Rudin [1, lemma 4.21]. □

To proceed our local investigation of the manifold in a neighbourhood of $\hat{u}$ we substitute the decomposition (3.6) into (3.5). Then, using (3.2) and the fact that $T'(\hat{u}).v = 0$ if $v \in \hat{\mathcal{N}}$, there results the operator equation

(3.8) $$T'(\hat{u}).w + \omega(u;v + w) = 0 \quad , \quad v \in \hat{\mathcal{N}}, w \in W.$$

For given element $v \in \hat{\mathcal{N}}$ with $||v||$ small, this equation may or may not have small norm solution $w \in W$.

DEFINITION 1.3.5. *(Analytical).*
Suppose Hypothesis 1.3.3. is satisfied. A point $\hat{u} \in \mathcal{M}$ is said to be a *regular point of the manifold* $\mathcal{M}$ if there exists a neighbourhood $N = \{v \in \hat{\mathcal{N}} \mid ||v|| < \varepsilon_o\}$ such that for every $v \in N$ the equation

$$T(\hat{u} + v + w) - T(\hat{u}) = 0$$

has a unique solution $w \in W$, which depends on v and shall be denoted by $w = \phi(v)$, and such that the mapping $\phi: N \to W$ is continuous and satisfies

(3.9) $$||\phi(v)|| = o(||v||) \quad \text{if } ||v|| \to 0.$$

This analytical definition can be given a clear geometrical interpretation. Therefore the following notions turn out to be useful.

DEFINITION 1.3.6. The *tangent space* of $\mathcal{M}$ at $\hat{u}$ is defined as the null space of the operator $T'(\hat{u}) \in B(V,Y)$:

$$\hat{\mathcal{N}} = \{v \in V \mid T'(\hat{u}).v = 0\} .$$

The *tangent plane* of $\mathcal{M}$ at $\hat{u}$ is defined to be the set

$$\hat{u} + \hat{\mathcal{N}} = \{\hat{u} + v \in V \mid v \in \hat{\mathcal{N}}\}.$$

Intuitively speaking, the idea of a tangent plane at $\hat{u}$ will be that with every point $u \in \mathcal{M}$ in a sufficiently small neighbourhood of $\hat{u}$ there corresponds a point $\hat{u} + v$ in the tangent plane such that the distance from $\hat{u}$ to $\hat{u} + v$ differs only through higher order terms from the distance between $\hat{u}$ and u. In fact it is easily seen that definition 1.3.3. is such that the tangent plane as defined in 1.3.6. satisfies these intuitive requirements if $\hat{u}$ is a regular point of $\mathcal{M}$. More precisely speaking, defintition 1.3.5. is equivalent to

DEFINITION 1.3.7. *(Geometrical).*
Suppose Hypothesis 1.3.3. is satisfied. A point $\hat{u} \in \mathcal{M}$ is said to be a regular point of $\mathcal{M}$ if there exists a homeomorphism (i.e. a continuous one-to-one map which has a continuous inverse) from a neighbourhood $\hat{u} + N$ of $\hat{u}$ in the tangent plane onto a neighbourhood of $\hat{u}$ in the manifold $\mathcal{M}$, say

$$\hat{u} + N \ni \hat{u} + v \mapsto \hat{u} + v + \phi(v) \in \mathcal{M},$$

which satisfies

$$||\phi(v)|| = o(||v||) \qquad \text{if } ||v|| \to 0.$$

The following lemma gives a sufficient condition for a point $\hat{u}$ to be a regular point of $\mathcal{M}$.

LEMMA 1.3.8. *Let* T *be continuously differentiable at* $\hat{u} \in V$, *with* $\mathcal{N}(T'(\hat{u})) \neq \{0\}$. *Suppose* $T'(\hat{u})$ *maps* V *onto all of* Y, *i.e. suppose*

$$\mathcal{R}(T'(u)) = Y. \tag{3.10}$$

Then, if hypothesis 1.3.3. *is satisfied,* $\hat{u}$ *is a regular point of* $\mathcal{M}$.

PROOF. Define the operator $F : \hat{\mathcal{N}} \times W \to Y$ by

$$F(v,w) := T(\hat{u} + v + w) - T(\hat{u}) \ , \ v \in \hat{\mathcal{N}} \ , \ w \in W. \tag{3.11}$$

Then the proof consists of an application of the *implicit function theorem* (c.f. Dieudonné [8, section 10.2]). To verify the conditions of this theorem we note that: (i) $F(0,0) = 0$; (ii) F is continuously differentiable with respect to both variables in a neighbourhood of (0,0) because T is assumed to be continuously differentiable in a neighbourhood of $\hat{u}$; (iii) If $D_W F(v_o,w_o)$ denotes the derivative of F with respect to $w \in W$ at the point (v_o,w_o), then $D_W F(0,0) = T'(\hat{u}) \in B(W,Y)$. Now if condition (3.10) is satisfied, $T'(\hat{u})$ is a one-to-one mapping from W onto Y and hence has a continuous inverse. We shall denote this "pseudo-inverse" by

(3.12) $$[T'(\hat{u})]_{ps}^{-1} : Y \to W.$$

The conditions of the theorem being fulfilled, the implicit function theorem states that there exist neighbourhoods $N = \{v \in \mathcal{V} \mid ||v|| < \varepsilon_o\}$ and $M = \{w \in W \mid ||w|| < \delta_o\}$ such that

(i) for each fixed $v \in N$, the equation

(3.13) $$F(v,w) = 0$$

has a unique solution $w \in M$;
(ii) this solution can be given as $w = \phi(v)$, where ϕ is continuously differentiable on N and $\phi(0) = 0$. Finally, to obtain the required estimate for $\phi(v)$, we note that $\phi(v)$ satisfies $F(v,\phi(v)) = 0$, which is by (3.11) and (3.2) equivalent to

$$T'(\hat{u}).\phi(v) + \omega(\hat{u}; v+\phi(v)) = 0.$$

With the aid of the pseudo-inverse (3.12), which is a bounded operator, this leads to the following estimate

(3.14) $$||\phi(v)|| \leq ||[T'(\hat{u})]_{ps}^{-1}||.||\omega(\hat{u}; v + \phi(v))||.$$

From the continuity of ϕ, and from $\phi(0) = 0$ it then follows from (3.3) that

$$||\omega(\hat{u}; v + \phi(v))|| = o(||v||) \quad \text{for } ||v|| \to 0.$$

Hence ϕ satisfies the requirement (3,9), and the proof is complete. □

For the theory of second variation in section 1.5, we shall need the following lemma. Therefore we note that if T is twice continuously differentiable at $\hat{u}$, the remainder $\omega(\hat{u};h)$ defined in (3.2), can be specialized to

(3.15) $$\omega(\hat{u};h) = \frac{1}{2} T''(\hat{u}).h.h + \nu(\hat{u};h),$$

where

$$(3.16) \qquad ||\nu(\hat{u};h)|| = o(||h||^2) \quad \text{for } ||h|| \to 0.$$

LEMMA 1.3.9. *With the same conditions as in lemma* 1.3.8, *but with* T *twice continuously differentiable at* $\hat{u}$, *the operator* $\phi: N \to W$ *satisfies the stronger estimate*

$$(3.17) \qquad ||\phi(v)|| = 0(||v||^2) \quad \text{for } ||v|| \to 0.$$

PROOF: Using $||\phi(v)|| = o(||v||)$, it follows from (3.15) and (3.16) that

$$||\omega(\hat{u};v + \phi(v))|| \leq \frac{1}{2} ||T''(\hat{u}).v.v|| + o(||v||^2).$$

Together with (3.14) this leads to (3.17). □

1.4. MULTIPLIER RULE *(Theory of first variation).*

With the results of the foregoing section is possible to study the behaviour of a given functional f on the manifold $\mathcal{M}$ in a neighbourhood of a regular point $\hat{u}$. Therefore we assume that $f: V \to R\mathcal{l}$ is continuously differentiable at $\hat{u}$ and write

$$(4.1) \quad f(\hat{u}+h) - f(\hat{u}) = < f'(\hat{u}),h > + o(||h||) \quad \text{for } ||h|| \to 0,\ h \in V.$$

As $\hat{u}$ is assumed to be a regular point of $\mathcal{M}$, it follows from definition 1.3.5. that the mapping $u(\cdot,v) : (-\varepsilon_o,\varepsilon_o) \to \mathcal{M}$:

$$(4.2.) \quad u(\varepsilon,v) := \hat{u} + \varepsilon v + \phi(v) \quad \text{for } v \in \hat{\mathcal{N}}, ||v|| = 1,\ -\varepsilon_o < \varepsilon < \varepsilon_o$$

defines for every $v \in \hat{\mathcal{N}}$ a continuous curve on $\mathcal{M}$ through $\hat{u}$. (Moreover, this mapping is continuously differentiable at $\varepsilon = 0$).
Considering f on such a curve gives with (3.9) and (4.1):

$$(4.3) \quad f(u(\varepsilon;v)) - f(\hat{u}) = \varepsilon < f'(\hat{u}).v > + o(\varepsilon) \quad \text{for } \varepsilon \to 0.$$

From definition 1.1.1. it follows that if $\hat{u}$ is a constrained local minimum point of f with respect to $\mathcal{M}$, the sign of the expression

$$f(u(\varepsilon;v)) - f(\hat{u}) \tag{4.4}$$

must certainly be independent of the sign of ε. In that case it follows from (4.3) that we must have

$$\langle f'(u), v\rangle = 0 \quad \text{for every } v \in \mathcal{V}. \tag{4.5}$$

REMARK 1.4.1. Of course this same condition must hold if $\hat{u}$ is a constrained local *maximum* point of f with respect to $\mathcal{M}$ (defined in an obvious way). On the other hand, any regular point $\bar{u} \in \mathcal{M}$ which satisfies (4.5) is called a *constrained stationary point* of f with respect to $\mathcal{M}$. Thus, as usual, a necessary condition for a regular point $\hat{u}$ to be a constrained extremum (maximum or minimum) point is that $\hat{u}$ is a constrained stationary point.

It is possible to write condition (4.5) as an operator equation for $\hat{u}$ in which the operator $T'(\hat{u})$ appears explicitly. To that end we need the results as described in subsection 0.2.1.
With the definition of annihilator as in 0.(2.5), condition (4.5) may be written as

$$f'(\hat{u}) \in \mathcal{V}(T'(\hat{u}))^{\perp}. \tag{4.6}$$

If $\mathcal{R}(T'(\hat{u}))$ is closed in Y, it follows with 0.(2.9) and 0.(2.10) that (4.6) is equivalent to

$$f'(\hat{u}) \in \mathcal{R}(T'(\hat{u})^*),$$

which means that there exists some element $\lambda^* \in Y^*$ such that

$$f'(\hat{u}) = T'(\hat{u})^*.\lambda^*.$$

Note that this element λ^* is unique up to some element from $\mathcal{N}(T'(\hat{u})^*)$, i.e. up to an arbitrary element from $^{\perp}\mathcal{R}(T'(\hat{u}))$ (as follows from

0.(2.8)). In analogy with the simplest case, $V = \mathbb{R}^n, n > 1$ and $Y = \mathbb{R}$, λ^* will be called a *Lagrange multiplier*.

The results obtained so far can be formulated in the following

THEOREM 1.4.2. *(Multiplier rule).*
Let $f: V \to \mathbb{R}$ *and* $T: V \to Y$ *be continuously differentiable at* $\hat{u} \in \mathcal{M}$.
Suppose that (i) $\mathcal{N}(T'(\hat{u})) \neq \{0\}$
(ii) $\hat{u}$ *is a regular point of* $\mathcal{M}$, *and* $\mathcal{R}(T'(\hat{u}))$ *is closed in* Y. *Then, if* $\hat{u}$ *is a constrained local minimum point of* f *with respect to* $\mathcal{M}$, *there exists a Lagrange-multiplier* $\lambda^* \in Y^*$, *such that* $\hat{u}$ *satisfies the equation*

$$f'(\hat{u}) = T'(\hat{u})^* . \lambda^* .$$

Moreover, λ^* *is unique up to an arbitrary element from* $\mathcal{N}(T'(\hat{u})^*) = {}^{\perp}\mathcal{R}(T'(\hat{u}))$.

For practical applications, the following theorem may be more appropriate.

THEOREM 1.4.3. *Let* $f: V \to \mathbb{R}$ *and* $T: V \to Y$ *be continuously differentiable at* $\hat{u} \in \mathcal{M}$.
Suppose that (i) *Hypothesis* 1.3.3. *is satisfied.*
(ii) $\mathcal{N}(T'(\hat{u})) \neq \{0\}$;
(iii) $\mathcal{R}(T'(\hat{u})) = Y$.
Then, if $\hat{u}$ *is a constrained local minimum point of* f *with respect to* $\mathcal{M}$, *there exists a unique Lagrange multiplier* $\lambda^* \in Y^*$ *such that* $\hat{u}$ *satisfies*

$$f'(\hat{u}) = T'(\hat{u})^* . \lambda^* . \tag{4.7}$$

PROOF: Conditions (i), (ii) and (iii) imply that $\hat{u}$ is a regular point of $\mathcal{M}$ (lemma 1.3.8), and that $\mathcal{R}(T'(\hat{u}))$ is closed as Y itself is closed, Hence condition (ii) of theorem 1.4.2. is satisfied. The uniqueness of the Lagrange multiplier follows at once from the observation

$$\mathcal{N}(T'(\hat{u})^*) = {}^{\perp}\mathcal{R}(T'(\hat{u})) = {}^{\perp}Y = \{0\} . \qquad \square$$

1.5. EXTREMALITY PROPERTY. *(Theory of second variation).*

In this section we shall extend the theory of the foregoing section to include second order effects. Therefore we assume f and T to be twice continuously differentiable at $\hat{u}$ and write

$$(5.1)\quad f(\hat{u}+h) - f(\hat{u}) = \langle f'(\hat{u}),h\rangle + \frac{1}{2}\langle f''(\hat{u}).h,h\rangle + o(||h||^2) \quad \text{for } ||h|| \to 0.$$

and

$$(5.2)\quad T(\hat{u}+h) - T(\hat{u}) = T'(\hat{u}).h + \frac{1}{2}T''(\hat{u}).h.h + \nu(\hat{u};h)$$

where

$$(5.3)\quad ||\nu(u;h)|| = o(||h||^2) \qquad \text{for } ||h|| \to 0.$$

As in the foregoing section we consider f on curves on $\mathcal{M}$ through $\hat{u}$:

$$u(\varepsilon,v) = \hat{u} + \varepsilon v + \phi(\varepsilon v) \quad , v \in \mathcal{N}, \ ||v|| = 1, \ |\varepsilon| < \varepsilon_o,$$

and note that because of lemma 1.3.9., the mapping $\phi : \mathcal{N} \to W$ satisfies

$$(5.4)\qquad ||\phi(\varepsilon v)|| = 0(\varepsilon^2) \ \text{ for } \varepsilon \to 0, \quad ||v|| = 1, \ v \in \mathcal{N}.$$

More particularly, it follows from (3.8) and (3.15) that $\phi(\varepsilon v)$ satisfies

$$(5.5)\qquad T'(u).\phi + \frac{1}{2}\varepsilon^2 T''(u).v.v + \sigma(u;\varepsilon v,\phi) = 0,$$

where

$$\sigma(u;\varepsilon v,\phi) = \frac{1}{2}T''(u).\phi.\phi + \varepsilon T''(\hat{u}).v.\phi + \nu(\hat{u};\varepsilon v+\phi)$$

and thus

$$(5.6)\qquad ||\sigma(u;\varepsilon v,\phi)|| = o(\varepsilon^2) \quad \text{for } \varepsilon \to 0 \ .$$

With these results we find

(5.7) $f(u(\varepsilon;v)) - f(\hat{u}) = \varepsilon\langle f'(\hat{u}),v\rangle + \langle f'(\hat{u}),\phi\rangle + \frac{1}{2}\varepsilon^2\langle f''(\hat{u}).v,v\rangle + o(\varepsilon^2)$.

If $\hat{u}$ is a constrained stationary point of f with respect to $\mathcal{M}$ which satisfies equation (4.7), (5.7) may be written as

$$f(u(\varepsilon;v)) - f(\hat{u}) = \varepsilon\langle T'(u)^*.\lambda^*,v\rangle + \langle T'(\hat{u})^*.\lambda^*,\phi\rangle + \\ + \frac{1}{2}\varepsilon^2\langle f''(\hat{u}).v,v\rangle + o(\varepsilon^2) .$$

Note that

$$\langle T'(\hat{u})^*.\lambda^*,v\rangle = \langle \lambda^*, T'(\hat{u}).v\rangle = 0 \qquad \text{as } v \in \mathcal{N},$$

and from (5.5) together with (5.6) it follows that

$$\langle T'(\hat{u})^*.\lambda^*,\phi\rangle = \langle \lambda^*, T'(\hat{u}).\phi\rangle = -\frac{1}{2}\varepsilon^2\langle \lambda^*,T''(\hat{u}).v.v\rangle + o(\varepsilon^2).$$

Hence

(5.8) $f(u(\varepsilon;v))-f(\hat{u}) = \frac{1}{2}\varepsilon^2[\langle f''(\hat{u}).v,v\rangle - \langle \lambda^*,T''(\hat{u}).v.v\rangle] + o(\varepsilon^2)$.

The expression in square brackets will be called the *second variation* and denoted by

(5.9) $s(\hat{u},\lambda^*;v) := \langle f''(\hat{u}).v,v\rangle - \langle \lambda^*,T''(\hat{u}).v.v\rangle$.

If $\hat{u}$ is a constrained local minimum point of f with respect to $\mathcal{M}$, it follows from (5.8) that the second variation will be non-negative for every $v \in \hat{\mathcal{N}}$, so that we have obtained the following theorem.

THEOREM 1.5.1 (*Extremality property*)
Suppose $f: V \to Rl$ *and* $T: V \to Y$ *are twice continuously differentiable at* $\hat{u}$, *and suppose that conditions* (i), (ii) *and* (iii) *of theorem* 1.4.3 *are satisfied. Then, if* $\hat{u}$ *is a constrained local minimum point of* f *with respect to* $\mathcal{M}$, *with* λ^* *as the corresponding multiplier, the second variation* $s(\hat{u},\lambda^*;v) := \langle f''(\hat{u}).v,v\rangle - \langle \lambda^*,T''(\hat{u}).v.v\rangle$ *is non-negative on* $\hat{\mathcal{N}}$ *i.e.*

(5.10) $s(\hat{u},\lambda^*;v) \geq 0$ *for all* $v \in \hat{\mathcal{N}} = \mathcal{N}(T(\hat{u}))$.

The foregoing theorem may be looked at as a *necessary* condition for a constrained stationary point $\hat{u}$ to be a constrained local minimum point. It is possible to give also a *sufficient condition.*

THEOREM 1.5.2. *Let* $f: V \to Rl$ *and* $T: V \to Y$ *be twice continuously differentiable at* $\hat{u}$, *and let* $\hat{u}$ *be a constrained stationary point for which conditions* (i), (ii) *and* (iii) *of theorem* 1.4.3. *are satisfied. Suppose the second variation satisfies for some constant* $c > 0$:

$$s(\hat{u},\lambda^*;v) \geq c.||v||^2 \quad \textit{for all } v \in \hat{\mathcal{M}}. \tag{5.11}$$

Then $\hat{u}$ *is a constrained local minimum point. Moreover, there exists a neighbourhood* $M(\hat{u}) \subset \mathcal{M}$ *of* $\hat{u}$ *in* $\mathcal{M}$ *such that*

$$f(u) - f(\hat{u}) \geq \frac{1}{8}.c.||u-\hat{u}||^2 \quad \textit{for every } u \in M(\hat{u}) \subset \mathcal{M}. \tag{5.12}$$

PROOF: The set

$$M_{\varepsilon_o} := \{u(\varepsilon,v) \,|\, u(\varepsilon,v) = \hat{u}+\varepsilon v+\phi(\varepsilon v),\ v \in \hat{\mathcal{M}},\ ||v|| = 1,\ |\varepsilon| < \varepsilon_o\}$$

is a subset of $\mathcal{M}$ and a neighbourhood of $\hat{u}$ for every $\varepsilon_o > 0$ sufficiently small. We shall show that for ε_o properly chosen, every element $u \in M_{\varepsilon_o}$ satisfies the inequality (5.12). Therefore we note that from (5.4) it follows that

$$||u(\varepsilon,v) - \hat{u}||^2 = ||\varepsilon v+\phi(\varepsilon v)||^2 = \varepsilon^2 + 0(\varepsilon^3) \quad \text{for } \varepsilon \to 0.$$

Hence there exists some $\varepsilon_1 > 0$ such that

$$||u - \hat{u}||^2 < 2\varepsilon^2 \quad \text{for every } u \in M_{\varepsilon_1}.$$

From (5.8) together with (5.11) it follows that

$$f(u(\varepsilon,v)) - f(\hat{u}) \geq \frac{1}{2}\varepsilon^2.c + o(\varepsilon^2), \quad \text{for every } \varepsilon \to 0.$$

Hence there exists some $\varepsilon_2 > 0$ such that

$$f(u) - f(\hat{u}) \geq \frac{1}{4} \cdot \varepsilon^2 \cdot c \qquad \text{for every } u \in M_{\varepsilon_2} .$$

Taking $\varepsilon_o := \min(\varepsilon_1, \varepsilon_2)$ it follows that

$$f(u) - f(\hat{u}) \geq \frac{1}{4} \cdot \varepsilon^2 \cdot c \geq \frac{1}{4} \cdot \frac{1}{2} \cdot c \cdot ||u-\hat{u}||^2 \qquad \text{for every } u \in M_{\varepsilon_o} .$$

This proves the theorem. □

1.6. SPECIALIZATION TO FUNCTIONAL CONSTRAINTS.

In this section we specialize the results of the foregoing sections to the case that Y is a finite dimensional Euclidean space, say $Y = R\!\!\!\!\mathcal{l}^n$, and thus $T: V \to R\!\!\!\!\mathcal{l}^n$. If elements of $R\!\!\!\!\mathcal{l}^n$ are denoted by row vectors, the operator T can be described as

$$T(u) = (t_1(u), t_2(u), \ldots, t_n(u)),$$

where

$$t_i : V \to R\!\!\!\!\mathcal{l} \quad , \; i = 1, 2, \ldots, n$$

are functionals. The dual space $Y^* = (R\!\!\!\!\mathcal{l}^n)^*$ is the n-dimensional Euclidean space of column vectors, and the duality map between $R\!\!\!\!\mathcal{l}^n$ and $(R\!\!\!\!\mathcal{l}^n)^*$ is the usual matrix multiplication.

To investigate the conditions (i) and (ii) of theorem 1.4.3. we note that

$$T'(u).v = (\langle t_1'(\hat{u}), v \rangle, \langle t_2'(\hat{u}, v \rangle, \ldots, \langle t_n'(\hat{u}), v \rangle).$$

LEMMA 1.6.1. *If* V *is infinite dimensional, or if* dim $V \geq 2$, *then*

$$\hat{\mathcal{N}} = \mathcal{N}(T'(\hat{u})) \neq \{0\} \qquad \textit{for every } \hat{u} \in \mathcal{M},$$
and hypothesis 1.3.3. *is satisfied.*

PROOF: If dim V is infinite, $\hat{\mathcal{N}}$ is infinite dimensional, with codimension n. If dim V = m, then (if $T'(\hat{u}) \neq 0$) dim $\hat{\mathcal{N}} = m-1$, and thus dim $\hat{\mathcal{N}} \geq 1$ if $m \geq 2$. □

DEFINITION 1.6.2. A finite set of elements $\{\ell_i\}_{i=1,2,\ldots,n}$ from V^* is called *linearly independent*, if for $a_i \in R\!\!\!l$, i=1,..,n.

$$\sum_{i=1}^{n} a_i \ell_i = 0 \iff a_i = 0 \quad \text{for} \quad i = 1,..,n.$$

LEMMA 1.6.3. *For* $\hat{u} \in \mathcal{M}$ *we have*

$$\mathcal{R}(T'(\hat{u}) = R\!\!\!l^n \iff \{t_i'(\hat{u})\}_{i=1,\ldots,n} \text{ is a linearly independent set of } V^*.$$

PROOF: Suppose $\mathcal{R}(T'(\hat{u}) = R\!\!\!l^n$. Then there exist elements $v_i \in V$, i=1,..,n, such that

$$T'(\hat{u}).v_i = (<t_i'(\hat{u}),v_i>,\ldots,<t_n'(\hat{u}),v_i>) = \underline{e}_i ,$$

where $\underline{e}_i$ is the i-th unit vector of $R\!\!\!l^n$. Then

$$\sum_k <a_k t_k'(\hat{u}),v_i> = \sum_k a_k <t_k'(\hat{u}),v_i> = a_i \quad ,i = 1,..,n,$$

so that

$$\sum_k a_k\, t_k'(\hat{u}) = 0 \Rightarrow a_i = 0 \quad \text{for } i = 1,\ldots,n.$$

Now suppose $\mathcal{R}(T'(\hat{u})) \neq R\!\!\!l^n$. This means that there exists some column vector $\underline{\alpha}^* \in (R\!\!\!l^n)^*$, $\underline{\alpha}^* \neq 0$, such that

$$<T'(\hat{u}).v\ ,\ \underline{\alpha}^*> = 0 \quad \forall v \in V .$$

If α_i^* denote the components of $\underline{\alpha}^*$, this implies that $\Sigma \alpha_i^* t_i'(\hat{u}) = 0$. As $\underline{\alpha}^* \neq 0$, this means that the set $\{t_i'(\hat{u})\}_{i=1,\ldots,n}$ is not linearly independent, which completes the proof. □

With the foregoing results, theorems 1.4.3. and 1.5.1. are easily specialized to the case $Y = R\!\!\!l^n$. Because of its importance for the next chapter, we shall formulate the theorems for the special case n = 1, for which $T = t: V \to R\!\!\!l$. Therefore note that it follows from lemma 1.6.1. that $\hat{\mathcal{M}} \neq \{0\}$, provided dim $V \geq 2$, and from

lemma 1.6.3. that

$$\mathcal{R}(t'(\hat{u})) = Rl \iff t'(\hat{u}) \neq 0.$$

(If $V = Rl$, then either $t'(\hat{u}) = 0$ or $\hat{u}$ is an isolated point of the "manifold").

THEOREM 1.6.4. *Let* $f: V \to Rl$ *and* $t : V \to Rl$ *be two given functionals on a reflexive* B*-space* V. *Let for* $p \in Rl$

$$\mathcal{M} := \{u \in V \mid t(u) = p\} \neq \emptyset.$$

Then we have

(i) EXISTENCE: *If* f *is w.l.s.c. on* V, *and coercive on* V *and* t *is w.c. on* V,
then f *is bounded from below on* $\mathcal{M}$ *and attains its infimum on* $\mathcal{M}$
(Hence, problem $\mathcal{P}$: $\inf\{f(u) \mid u \in \mathcal{M}\}$ *has a solution).*
(ii) MULTIPLIER RULE: *Let* f *and* t *be continuously differentiable at* $\hat{u} \in \mathcal{M}$ *Suppose* $t'(\hat{u}) \neq 0$ *and* $\dim V \geq 2$. *Then, if* $\hat{u}$ *is a constrained local minimum point of* f *with respect to* $\mathcal{M}$ *there exists a unique Lagrange-multiplier* $\lambda \in Rl$ *such that* $\hat{u}$ *satisfies*

$$f'(\hat{u}) = \lambda . t'(\hat{u}) \quad . \tag{6.1}$$

(iii) EXTREMALITY PROPERTY: *If* f *and* t *are twice continuously differentiable at* $\hat{u}$, *then in the same situation as in* (ii) *above, the second variation*

$$s(\hat{u},\lambda;v) = \langle S(\hat{u},\lambda).v, v\rangle \text{ with } S(\hat{u},\lambda) = f''(\hat{u}) - \lambda t''(\hat{u}) \tag{6.2}$$

is non-negative on $\hat{\mathcal{V}}$:

$$s(\hat{u},\lambda;v) \geq 0 \quad \text{for all } v \in V \text{ which satisfy } \langle t'(\hat{u}), v\rangle = 0 \ . \tag{6.3}$$

(iv) SUFFICIENCY: *If* f *and* t *are twice continuously differentiable at* $\hat{u}$, *and if* $\hat{u}$ *is a constrained stationary point for which* $t'(\hat{u}) \neq 0$

and for which

(6.4) $s(\hat{u},\lambda;v) \geq c.||v||^2$ *for all* $v \in \hat{\mathcal{N}} = \{v \mid \langle t'(\hat{u}),v\rangle = 0\}$

for some constant $c > 0$, *then* $\hat{u}$ *is a constrained local minimum point and there exists a neighbourhood* $M(\hat{u}) \subset \mathcal{M}$ *of* $\hat{u}$ *in* $\mathcal{M}$ *such that*

(6.5) $f(u) - f(\hat{u}) \geq \frac{1}{8}.c.||u-\hat{u}||^2$ *for every* $u \in M(\hat{u}) \subset \mathcal{M}$.

1.7. THE ELASTIC LINE.

Consider an inextensible string of total length ℓ which has constant mass density. Suppose the string has finite bending stiffness, the energy density of which is proportional to the inverse square of the radius of curvature. The endpoints and the tangents to the string at these points are taken to be fixed. The effect of gravity forces is neglected.

The *principle of least energy* is assumed to hold, which means that the actual configuration which is preferred by the string is that configuration for which the total bending energy is as small as possible when compared with all other configurations that satisfy the boundary conditions and the total length and constant mass density conditions.

For simplicity we consider only the case for which the boundary conditions are such that the string is known to lie in a plane. Then, taking a Cartesian coordinate system OXY with the origin at one endpoint and the other endpoint at the x-axis, a typical material point of the string can be described by a two component vector $\underline{r} = (x,y)$, and the complete configuration may be described as

(7.1) $\underline{r}(s) = (x(s),y(s))$.

where s is the *arc length*, running from 0 (the endpoint in the origin 0 say) to ℓ: $0 \leq s \leq \ell$. In using this particular parameter s to characterize the string, it must be remembered that the tangent vector is a unit vector:

(7.2) $$\underline{r}_s \cdot \underline{r}_s = 1$$

where $\underline{r}_s \cdot \underline{r}_s = x_s^2 + y_s^2$. The conditions at the endpoints may be described as

(7.3) $$\underline{r}(0) = 0, \quad \underline{r}(\ell) = (L,0) \quad \text{with } |L| < \ell$$
$$r_s(0) = (\cos\alpha, \sin\alpha), \quad \underline{r}_s(\ell) = (\cos\beta, \sin\beta) .$$

For a configuration described by (7.1) for which x(s) and y(s) are sufficiently smooth functions of s, the radius of curvature R satisfies

$$R^{-2} = \underline{r}_{ss} \cdot \underline{r}_{ss} ,$$

so that the total bending energy is up to some multiplicative constant

(7.4) $$\int_0^\ell \frac{1}{2}\, \underline{r}_{ss} \cdot \underline{r}_{ss}\, ds .$$

The principle of least energy then states that the actual configuration is described by that two component vector (7.1) which satisfies (7.2) and (7.3) and for which the functional (7.4) is minimal when compared to all other configurations described by (7.1) which satisfy (7.2) and (7.3). To obtain a sound mathematical description of the principle it is necessary to specify the function space in which the minimum is sought. Therefore we choose

$$\underline{H}^2 := \{\underline{r} = (x,y) \mid x \in H^2(0,\ell) \quad , y \in H^2(0,\ell)\} ,$$

where $H^2(0,\ell)$ is the second Sobolev space of functions defined on the interval $(0,\ell)$, and $\underline{H}^2$ is supplied with the norm

$$||\underline{r}||^2_{\underline{H}^2} := ||x||^2_{H^2} + ||y||^2_{H^2} = \int_0^\ell [\, \underline{r}_{ss} \cdot \underline{r}_{ss} + \underline{r}_s \cdot \underline{r}_s + \underline{r} \cdot \underline{r}\,]\, ds$$

(c.f. subsection 0.1.4.). Then the functional (7.4) is neatly defined on $\underline{H}^2$:

(7.5) $$f(\underline{r}) := \frac{1}{2}\int_0^{\ell} \underline{r}_{ss}\cdot\underline{r}_{ss}\,ds \quad , \; f : \underline{H}^2 \to R\ell \; ,$$

and the operator T defined by

(7.6) $$T(\underline{r}) := \underline{r}_s\cdot\underline{r}_s - 1$$

is easily seen (using the embedding result $H^2 \subset C^1$ from subsection 0.2.3.) to be a mapping from $\underline{H}^2$ into H^1:

(7.7) $$T : \underline{H}^2 \to H^1 .$$

Then the principle of least energy leads one to consider the following abstract constrained extremum problem

(7.8) $$\inf \{ f(\underline{r}) \mid \underline{T}(\underline{r}) = \underline{a} \; , \; \underline{r} \in \underline{H}^2 \} \; ,$$

where $\underline{T} : \underline{H}^2 \to H^1 \times (R\ell^2)^4$ is defined by

(7.9) $$\underline{T}(\underline{r}) := (T(\underline{r}) \; ; \; \underline{r}(0); \; \underline{r}(\ell); \; \underline{r}_s(0); \; \underline{r}_s(\ell))$$

and

(7.10) $$\underline{a} = (0 \; ; \; (0,0) \; ; \; (L,0) \; ; \; (\cos\alpha, \sin\alpha); \; (\cos\beta, \sin\beta)).$$

Formulated in this way, the problem has been brought into a form to which the abstract theory of sections 1.2 - 1.5 may be applied. Note that in this formulation the boundary conditions are considered as constraints. Although a fully satisfactory treatment is possible in this way, the boundary conditions can be dealt with in another way. This is done by choosing some function $\underline{r}_o \in \underline{H}^2$ which satisfies (7.2) and (7.3). Such a function is likely to exist if the boundary conditions are compatible with the constraints, i.e. if

(7.11) $|L| < \ell$, α and β arbitrary real numbers,

or

(7.12) $L = \ell, \; \alpha = \beta = 0$

or $L = -\ell, \; \alpha = \beta = \pi .$

Note that these conditions guarantee that the manifold

$$\mathcal{M} := \{\underline{r} \in \underline{H}^2 \mid \underline{T}(\underline{r}) = \underline{a}\} ,$$

where $\underline{a}$ is given by (7.10), is non-void. In the special case for which (7.12) holds, $\mathcal{M}$ consists of only one element, viz

$$\mathcal{M} = \{\underline{r}_o = (s,0)\} \quad \text{or} \quad \mathcal{M} = \{\underline{r}_o = (-s,0)\} .$$

Having chosen such a function $\underline{r}_o$, we note that every function $\underline{r} \in \underline{H}^2$ which satisfies the boundary conditions, can be written as

$$\underline{r} = \underline{r}_o + \underline{h} , \quad \text{with } \underline{h} \in \underline{H}_o^2 , \tag{7.13}$$

where

$$\underline{H}_o^2 = \{\underline{h} \in H^2 \mid \underline{h}(0) = \underline{h}(\ell) = \underline{h}_s(0) = \underline{h}_s(\ell) = \underline{0}\}$$

Defining a functional $f(\underline{r}_o;.)$ and an operator $T(\underline{r}_o;.)$ by

$$f(\underline{r}_o; \underline{h}) := f(\underline{r}_o + \underline{h}) , \quad f(\underline{r}_o;.) : \underline{H}_o^2 \to R\ell \tag{7.14}$$

$$T(\underline{r}_o;\underline{h}) := T(\underline{r}_o + \underline{h}) \qquad T(\underline{r}_o;.) : \underline{H}_o^2 \to H_o^1 , \tag{7.15}$$

the principle of least energy amounts to an investigation of the constrained extremum problem

$$\inf \{f(\underline{r}_o;\underline{h}) \mid T(\underline{r}_o;\underline{h}) = 0 , \ \underline{h} \in \underline{H}_o^2 \} . \tag{7.16}$$

By construction, the two constrained extremum problems (7.8) and (7.16) are equivalent.

EXISTENCE. To prove the existence of a solution of the constrained extremum problem, we apply theorem 1.2.1. to the problem (7.8). Therefore we have to verify the conditions of the theorem.
(i) The functional f given by (7.5) is w.l.s.c. on $\underline{H}^2$. This is an easy consequence of remark 0.2.6.(ii).
(ii) Although f is *not* coercive on all of $\underline{H}^2$, it is not too difficult to show that f is coercive on the manifold $\mathcal{M}$.

(iii) $\mathcal{M}$ is weakly closed. This is proved with the aid of lemma 1.2.3., and the proof uses extensively the embedding theorem 0.2.9. for Sobolev spaces. Suppose $\underline{r}_n \rightharpoonup \hat{\underline{r}}$ in $\underline{H}^2$. Then $\underline{r}_n \to \hat{\underline{r}}$ in $\underline{H}^1$, and hence $\underline{r}_n(0) \to \hat{\underline{r}}(0)$, $\underline{r}_n(\ell) \to \hat{\underline{r}}(\ell)$ and $\underline{r}_{n_s}(0) \to \hat{\underline{r}}_s(0)$, $\underline{r}_{n_s}(\ell) \to \hat{\underline{r}}_s(\ell)$ in R^2, which shows that the boundary operators are continuous with respect to weak convergence. Remains to investigate the operator T as given by (7.6). We shall show that $T(\underline{r}_n) \to T(\hat{\underline{r}})$ in L_2. To that end we note (using $||\ ||$ for the L_2-norm of both scalar and vector-functions)

$$||T(\underline{r}_n) - T(\hat{\underline{r}})|| = ||(\underline{r}_{n_s} + \hat{\underline{r}}_s).(\underline{r}_{n_s} - \hat{\underline{r}}_s)||$$

$$\leq ||\underline{r}_{n_s} + \hat{\underline{r}}_s|| \cdot ||\underline{r}_{n_s} - \hat{\underline{r}}_s|| \to 0 \text{ for } n \to \infty,$$

because $||\underline{r}_{n_s} + \hat{\underline{r}}_s||$ is uniformly bounded and $||\underline{r}_{n_s} - \hat{\underline{r}}_s|| \to 0$ as $\underline{r}_n \rightharpoonup \hat{\underline{r}}$ in $\underline{H}^1$. Hence $T(\underline{r}_n) \to T(\hat{\underline{r}})$ in L_2 and as H^1 is continuously embedded in L_2 we may apply lemma 1.2.3. to conclude that $\mathcal{M}$ is weakly closed. (In fact it can be shown that if $\underline{r}_n \rightharpoonup \hat{\underline{r}}$ in $\underline{H}^2$, then $T(\underline{r}_n) \to T(\hat{\underline{r}})$ in H^1 for some subsequence).
The conditions of theorem 1.2.1. being satisfied it follows that if $\mathcal{M} \neq \emptyset$, *problem (7.8) has at least one solution* $\hat{\underline{r}} \in \underline{H}^2$. Consequently, problem (7.16) has at least one solution, and if we take $\underline{r}_o = \hat{\underline{r}}$ (the existence of "some" element $\underline{r}_o$ satisfying (7.2) and (7.3) now being proved when $\mathcal{M} \neq \emptyset$), this solution is $\underline{h} = \underline{0}$.

REGULARITY of elements from $\mathcal{M}$. For the vectorfunctions from the manifold $\mathcal{M}$ we have $\underline{r} \in \underline{H}^2 \subset \underline{C}^1$ (each component of $\underline{r}$ is a C^1-function) and hence $\underline{r}_s \in \underline{C}^o$ and $\underline{r}_{ss} \in L_2$. Moreover, because of condition (7.2) we have $\underline{r}_s.\underline{r}_s \in C^\infty$ as $1 \in C^\infty$. By differentiating (7.2) with respect to s it is found that

$$\underline{r} \in \mathcal{M} \Rightarrow \underline{r}_s.\underline{r}_{ss} = 0 \qquad \forall s \in (0,\ell), \tag{7.17}$$

which result implies that discontinuities in the components of $\underline{r}_{ss}$ (if any) cancel in expressions like $\underline{r}_s.\underline{r}_{ss}$. We shall use this property in the following.

MULTIPLIER RULE. To obtain the equation satisfied by the minimum point $\hat{\underline{r}}$, it is somewhat simpler to use the formulation (7.16) with $\underline{r}_o = \hat{\underline{r}}$. Therefore we have first of all to study the null space and range of the operator

$$(7.18)\quad T'(\hat{\underline{r}};0):\underline{H}_o^2 \to H_o^1 \quad , \quad T'(\hat{\underline{r}};0).\underline{h} = T'(\hat{\underline{r}}).\underline{h} = \hat{\underline{r}}_s.\underline{h}_s \ .$$

We shall show that if $\hat{\underline{r}}(s) \neq (\pm s,0)$

$$(7.19)\qquad \hat{\mathcal{N}} := \mathcal{N}(T'(\hat{\underline{r}};0)) \neq \{0\} \quad , \quad \mathcal{R}(T'(\hat{\underline{r}};0)) = H_o^1 \ .$$

In other words, *if* $\hat{\underline{r}}$ *is not an isolated point of* $\mathcal{M}$ (and $\hat{\underline{r}}$ is isolated only if (7.12) holds) *then* $\hat{\underline{r}}$ *is a regular point of the manifold* $\mathcal{M}$.

To prove (7.19), note that $\underline{h} \in \hat{\mathcal{N}}$ if $\underline{h}$ satisfies

$$\underline{h}_s \cdot \underline{r}_s = 0 \quad , \quad h \in \underline{H}_o^2 \ .$$

Therefore we take

$$\underline{h}(s) = \int_o^s \underline{h}_s(\xi)d\xi \quad \text{with } \underline{h}_s = b\,\underline{r}_{ss}$$

wherein $b \in H^2 \cap H_o^1$ is chosen in such a way that $\underline{h}_s \in H_o^1$. Then $\underline{h} \in H_o^2$ if b satisfies

$$(7.20)\qquad \int_o^\ell b\,\underline{r}_{ss}\,ds = 0 \ .$$

Provided $\underline{r}_{ss} \neq \underline{0}$ there exists a function $b \neq 0$ which fulfills the requirements. This proves $\hat{\mathcal{N}} \neq \{0\}$. Now, if $\eta \in H_o^1$ is arbitrary, we look for a solution $\underline{h}$ of

$$\underline{h}_s \cdot \underline{r}_s = \eta \quad , \quad \underline{h} \in \underline{H}_o^2 \ .$$

Such a solution exists, provided $\hat{\underline{r}}_{ss} \neq 0$, which can be seen by taking

$$\underline{h}(s) = \int_o^s h_s(\xi)d\xi \ , \quad \text{with} \quad \underline{h}_s = \eta\underline{r}_s + b\,\underline{r}_{ss} \ ,$$

and where $b \in H^2 \cap H_o^1$ is chosen such that $\underline{h}_s \in H_o^1$ and

$$\int_0^{\ell} b\, \underline{r}_{ss}\, ds = -\int_0^{\ell} \eta\, \underline{r}_s\, ds\ .$$

This shows $\mathcal{R}(T'(\hat{\underline{r}};\underline{0})) = H_o^1$ if $\hat{r} \neq (\pm s,0)$.

The conditions of theorem 1.4.3. being verified (hypothesis 1.3.3. is clearly satisfied), we get the governing equation for $\hat{\underline{r}}$ as follows. Take the L_2-inner product (,) as duality map. Then $(H_o^1)^* = H^{-1}$, $(\underline{H}_o^2)^* = \underline{H}^{-2}$ (c.f. subsection 0.1.4.) and we have

$$\langle f'(\hat{\underline{r}};\underline{0}),\underline{h}\rangle = (\underline{r}_{ss},\underline{h}_{ss}) = (\underline{r}_{ssss},\underline{h}) \qquad \forall \underline{h}\in \underline{H}_o^2$$

$$\langle T'(\hat{\underline{r}};\underline{0}).\underline{h},\sigma\rangle = (\underline{r}_s.\underline{h}_s,\sigma) = (-(\sigma \underline{r}_s)_s,\underline{h}) \quad \forall \underline{h}\in \underline{H}_o^2\ ,\ \forall \sigma\in H^{-1}\ .$$

(Hence the adjoint operator $T'(\hat{\underline{r}};\underline{0})^* : H^{-1} \to \underline{H}^{-2}$ is defined by

$$T'(\hat{\underline{r}};\underline{0})^*.\sigma = -(\sigma \underline{r}_s)_s \quad \forall \sigma\in H^{-1}\ ).$$

Then the multiplier rule states that there exists a unique element $\sigma\in H^{-1}$ such that $\hat{\underline{r}}$ satisfies

$$\hat{\underline{r}}_{ssss} = (\sigma\, \hat{\underline{r}}_s)_s\ . \tag{7.21}$$

Although this is essentially an equation for elements from $\underline{H}^{-2}$, it can be shown that $\hat{\underline{r}} \in C^\infty$ and $\sigma \in C^\infty$.
[Note that for the particular situation that (7.12) holds, the only vectorfunction which satisfies the constraint (7.2) and the boundary conditions, viz $\hat{\underline{r}} = (\pm s,0)$, *does* satisfy an equation of the form (7.21). However, then σ is not unique: every arbitrary constant will do].

EXTREMALITY PROPERTY. From section 1.5 the following result immediately follows:

$$s(\hat{r}\ ,\sigma;\underline{h}) = (\underline{h}_{ss},\underline{h}_{ss}) - (\sigma,\underline{h}_s.\underline{h}_s)$$
$$= \int_0^{\ell} (\underline{h}_{ss}.\underline{h}_{ss} - \sigma\, \underline{h}_s.\underline{h}_s)\, ds \geq 0 \quad \text{for all } \underline{h}\in \underline{H}_o^2 \text{ with } \underline{h}_s.\underline{r}_s = 0.$$

The foregoing results can be summarized to give the following

THEOREM 1.7.1. *If the boundary conditions* (7.3) *satisfy* (7.11) *there exists an infinitely smooth solution* $\hat{\underline{r}}$ *of the principle of least energy. This solution satisfies the equations*

$$\underline{r}_{ssss} = (\sigma\, \underline{r}_s)_s$$

$$\underline{r}_s \cdot \underline{r}_s = 1$$

for some unique Lagrangemultiplier $\sigma \in C^\infty$. *Moreover, the non-negativity result* (7.22) *holds.*

REMARKS 1.7.2.

(i) Concerning the *uniqueness* of the solution of problem (7.8) (or (7.16)), we note that at least for some choises of the boundary conditions there will be at least two solutions : e.g. if $|L| < \ell$, and $\underline{r}_s(0) = \underline{r}_s(\ell) = (1,0)$, $(\hat{x},-\hat{y})$ is a solution of (7.8) if $(\hat{x},\hat{y})$ is a solution.

(ii) An alternative description of the foregoing system is possible by writing $\underline{r}_s = (\cos\theta, \sin\theta)$ where now $\theta \in H^1$ is a scalarfunction of $s \in [0,\ell]$. Then the constraint (7.2) is satisfied for arbitrary $\theta \in H^1$, and the principle of least energy leads to the following constrained extremum problem

$$(7.23)\quad \inf\{\frac{1}{2}\int_0^\ell \theta_s^{\,2}\, ds \,|\, \theta \in H^1; \theta(0) = \alpha; \theta(\ell) = \beta; \int_0^\ell \cos\theta\, ds = L; \int_0^\ell \sin\theta\, ds = 0\}.$$

This formulation is equivalent to (7.8) in the sense that if $\hat{\theta}$ is a solution of (7.23), the solution of (7.8) is given by

$$\hat{\underline{r}}(s) = \int_0^s (\cos\theta(\xi), \sin\theta(\xi))\, d\xi.$$

Note that instead of the constraint (7.2), problem (7.23) has two functional constraints (apart from the boundary conditions), and the theory as described in section 1.6., may be applied. A somewhat modified version of (7.23) will be more extensively studied in the next chapter.

(iii) The system considered above may be envisaged as a description of the time-independent states of an elastic line which is able to move in a plane. With the foregoing results it is an easy matter to find the governing *dynamical equations* from Hamiltons principle. For shortness we shall only outline the method. (See also chapter 3.) Consider vectorfunctions $\underline{r} = \underline{r}(s,t)$ of $s \in [0,\ell]$ and the time t, $t \in [t_1,t_2]$ say, from the B-space V:

$$V = \{\underline{r} \mid \underline{r}(s,.) \in \underline{H}^1(t_1,t_2)\ \forall s\in(0,\ell),\ \underline{r}(.,t)\in\underline{H}^2(0,\ell)\ \forall t\in(t_1,t_2)\}.$$

The kinetic energy density is given by $\frac{1}{2}\rho\, \underline{r}_t \cdot \underline{r}_t$, and the *action functional* is defined as the difference of the total kinetic energy and potential energy

$$A(\underline{r}) := \int_{t_1}^{t_2} dt \int_0^{\ell} ds\left[\frac{1}{2}\rho\, \underline{r}_t \cdot \underline{r}_t - \frac{1}{2} B\, \underline{r}_{ss} \cdot \underline{r}_{ss}\right]. \tag{7.24}$$

(Here B is some material constant and ρ is the constant mass density). Then *Hamiltons principle* states that the actual description of the dynamical system is such that it is a stationary point of the actionfunctional A with respect to the set of elements from V which satisfy the constraint (7.2) together with specified boundary conditions at $s = 0$, $s = \ell$ for all $t \in [t_1,t_2]$ and at $t = t_1$, $t = t_2$ for all $s \in [0,\ell]$. *Assuming* the existence of a solution $\hat{\underline{r}}$ of this constrained variational problem, the theory of sections 1.3., 1.4. may be applied to result into:

There exists a unique function $\sigma(s,t)$ *such that the governing equations for* $\hat{\underline{r}}(s,t)$ *are given by*

$$\begin{aligned} \rho\, \underline{r}_{tt} &= (\sigma\, \underline{r}_s)_s - B\, \underline{r}_{ssss}. \\ \underline{r}_s \cdot \underline{r}_s &= 1\ . \end{aligned} \tag{7.25}$$

CHAPTER 2: DUAL AND INVERSE VARIATIONAL PRINCIPLES.

2.1. INTRODUCTION.

As an important result of the foregoing chapter, we have obtained in section 1.4 the multiplier rule for constrained extremum problems of the form

$$(1.1) \qquad \mathcal{P} : \inf_{T(u)=y_o} f(u) ,$$

where $f: V \to R\!l$ and $T: V \to Y$.

Actually, the equation that must be satisfied by every solution of problem $\mathcal{P}$, i.e. equation 1.(4.7), is well known and is used extensively in many applications from mathematical physics. In most cases this equation will probably have been found with the aid of the following

RECIPE 2.1.1. *Let $\hat{u}$ be a solution of problem $\mathcal{P}$. Then there exists some element $\lambda^* \in Y^*$ such that the functional ℓ, defined by*

$$(1.2) \qquad \ell : V \times Y^* \to R\!l \quad , \quad \ell(u,y^*) = f(u) - \langle y^*, T(u) - y_o \rangle$$

has $(\hat{u},\lambda^*) \in V \times Y^*$ as stationary point.

Indeed, the stationary point $(\hat{u},\lambda^*)$ of the unconstrained functional ℓ satisfies (c.f. section 0.5.1.)

$$(1.3) \qquad \ell'_u(\hat{u},\lambda^*) = f'(\hat{u}) - T'(\hat{u})^*.\lambda^* = 0$$

$$(1.4) \qquad \ell'_{y^*}(\hat{u},\lambda^*) = T(\hat{u}) - y_o = 0,$$

where ℓ'_u and ℓ'_{y^*} denote the partial derivatives of ℓ with respect to u and y^* respectively, and eq. (1.3) is precisely eq. 1.(4.7), whereas (1.4) expresses the fact that $\hat{u}$ satisfies the constraint. Hence, provided the conditions of theorem 1.4.3. are satisfied, there is complete agreement between the multiplier rule and the recipe 2.1.1.

In this chapter we shall investigate the rôle played by the multiplier λ^* and its dependence on the actual minimal point $\hat{u}$ somewhat more extensively. In particular we shall be interested in the relation between problems of the kind (1.1) and the unconstrained minimization problems

$$\inf_{u \in V} \quad \ell(u, y^*)$$

for $y^* \in V^*$ fixed.

Although the results to be obtained in the following sections 2.2. and 2.3. can be generalized to answer several of these questions, we shall from now on restrict ourselves to extremum problems with functional constraints (i.e. $Y = Y^* = R\!\!I$, and $T = t : V \to R\!\!I$ a functional). This restriction will not only simplify the presentation and the interpretation of the results, but seems to be also a most interesting case for many present-day problems from non-linear analysis.

Let f and t be two functionals on the reflexive Banach space V, and define the functional ℓ (somewhat different from (1.2)) by

$$\ell : V \times R\!\!I \to R\!\!I \ , \ \ell(u,\mu) := f(u) - \mu t(u). \tag{1.5}$$

Then, according to the multiplier rule (section 1.6.), if $\bar{u}$ is a solution of

$$\mathcal{P}_{\bar{p}} : \inf_{t(u)=\bar{p}} f(u) \ , \tag{1.6}$$

and $t'(\bar{u}) \neq 0$, $\bar{u}$ is a stationary point of the functional $\ell(.,\bar{\mu})$, i.e. $\bar{u}$ satisfies

$$f'(\bar{u}) = \bar{\mu}\, t'(\bar{u}), \tag{1.7}$$

where $\bar{\mu}$ is the unique multiplier corresponding to the minimal point $\bar{u}$.

To investigate when $\bar{u}$ is an actual minimal point of $\mathcal{L}(\cdot,\bar{\mu})$ and to study the relation between $\bar{p}$ and the "corresponding" multiplier $\bar{\mu}$, we shall consider $\bar{p}$ in (1.6) as a parameter. Hence, for given functionals f and t, we shall study the *family* of constrained extremum problems

$$(1.8) \qquad \mathcal{P}_p : \inf_{t(u)=p} f(u) \quad , \quad p \in R\mathcal{l}.$$

In the same way, regarding μ as a parameter, we shall investigate the family of unconstrained extremum problems

$$(1.9) \qquad \mathcal{K}_\mu : \inf_{u\in V} \mathcal{L}(u,\mu) = \inf_{u\in V} [f(u) - \mu t(u)] \quad , \quad \mu \in R\mathcal{l}.$$

In section 2.3. it shall be shown that (1.9) is closely related to a dual formulation of problems $\mathcal{P}_p$, which implies (among other things) that for specific values of p the solution of $\mathcal{P}_p$ is also a solution of $\mathcal{K}_\mu$ for some $\mu \in R$.

If $\hat{u}$ is a solution of $\mathcal{P}_p$, f is minimal at $\hat{u}$ with respect to the manifold $\{u \mid t(u) = p\}$. It is interesting to investigate when $\hat{u}$ is also an extremal point of the functional t on the manifold $\{u \mid f(u) = f(\hat{u})\}$. This will be the subject of section 2.4., where we compare the solution sets of $\mathcal{P}_p$ with the solution sets of problems $\mathcal{S}_r$ and $\mathcal{Q}_r$:

$$\mathcal{S}_r : \sup_{f(u)=r} t(u) \quad , \quad \mathcal{Q}_r : \inf_{f(u)=r} t(u) \quad , r \in \mathcal{R}(f).$$

As we have seen in section 1.5 (and in subsection 0.5.1.), from the fact that an element $\hat{u}$ is a solution of an extremum principle, it follows that the second variation has some positivity properties. The (sign of the) second variation plays an important rôle in several applications. We shall try to describe this briefly for a specific situation.

One is sometimes interested in the complete set of stationary points of the functional $\ell(.,\mu)$, i.e. in the solution set of the operator equation (1.7). The study of this solution set and its dependence on μ (e.g. the number of solutions and their properties) is the subject of what is commonly called *bifurcation theory*. The reason for this interest is often that an equation of the form (1.7) is the equation for the stationary states (i.e. time-independent solutions) of an *evolution equation*. For instance, for evolution equations of the form

$$w_t = f'(w) - \bar{\mu}t'(w)$$

(which are parabolic equations in general) or

$$w_{tt} = f'(w) - \bar{\mu}t'(w)$$

(conservative wave equations), time independent solutions $\bar{u}$ satisfy (1.7). The *stability* (in the sense of Lyapunov) of a particular stationary state $\bar{u}$ for such an evolution equation is often directly related to extremality properties of $\bar{u}$ for the functional $\ell(.,\bar{\mu})$. As follows already from theorem 0.5.3. the second variation of the functional $\ell(.,\bar{\mu})$ at $\bar{u}$:

$$s(\bar{u},\bar{\mu};v) := \langle\{f''(\bar{u}) - \bar{\mu}\, t''(\bar{u})\}.v,v\rangle \qquad \text{for } v \in V$$

can play a fundamental rôle in such a stability analysis. This is even more true if a "principle of linearized stability" is known to hold, in which case merely from the positivity of the second variation (i.e. $s(\bar{u},\bar{\mu};v) > 0$ for all $v \in V$, $v \neq 0$) stability of $\bar{u}$ for the evolution equation can be deduced. As the only reason was to make plausible the importance of the second variation, we shall not pursue this subject any further here.[For an introduction to general evolution equation and the theory of Lyapunov stability see Zubov [16]. For applications of this theory to parabolic equations, and to see the rôle played by the second variation of (Lyapunov-) functionals see the contributions of Diekmann (Chapter 1.) and Koornwinder (Chapter 5.) in Diekmann & Temme [17]. See also Gelfand & Fomin [18, chapter 5] for an elementary introduction into the

relation between the theory of second variation and the conjugate point theory for the linearized operator equation].

In the next section we shall deal with some local investigations of a more or less heuristic character, which are mainly meant to illuminate the lines of the following sections. In section 2.5. we shall apply and demonstrate the derived abstract results to two specific problems from mathematical physics.

2.2. HEURISTIC CONSIDERATIONS.

Let us start to consider problem $\mathcal{P}_{\bar{p}}$ as the *primal problem* to be studied, where $\bar{p} \in \mathcal{R}(t)$, such that the manifold $\mathcal{M}_{\bar{p}} := \{u \in V \mid t(u) = \bar{p}\}$ is non-void. As announced, we consider a family of *perturbed problems* $\mathcal{P}_p$, defined in (1.8), where for the first instance, the parameter p may be thought to lie in a small neighbourhood $\mathcal{J}_{\bar{p}}$ of $\bar{p}$. For the first part of this section we assume the following hypothesis to hold.

HYPOTHESIS 2.1.1. There exists a neighbourhood $\mathcal{J}_{\bar{p}} \subset R\mathcal{l}$ of $\bar{p}$ such that:
(i) for every $p \in \mathcal{J}_{\bar{p}}$ there exists at least one solution of problem $\mathcal{P}_{\bar{p}}$;
(ii) there exists a solution branch $\{U(p) \mid U(p)$ is solution of $\mathcal{P}_p$, $p \in \mathcal{J}_{\bar{p}}\}$ such that the mapping $p \to U(p)$ is continuously differentiable. The derivative of this mapping at p will be denoted by $U'(p)$ and will be identified with its effect at 1 (thus the mapping $U'(p): R\mathcal{l} \to V$, $\alpha \to U'(p).\alpha = \alpha.U'(p)$ and the element $U'(p).1 \in V$ are identified as usual).
(iii) $t'(U(p)) \neq 0$ for $p \in \mathcal{J}_{\bar{p}}$.
(iv) $f \in C^2(V, R\mathcal{l})$ and $t \in C^2(V, R\mathcal{l})$.

With respect to this hypothesis we remark that condition (i) is satisfied if $\mathcal{J}_{\bar{p}} \subset \mathcal{R}(t)$ and if f and t satisfy conditions of theorem 1.6.4. (i). Condition (iii) will be satisfied if $t'(U(\bar{p}) \neq 0$ and if t' is continuous in a neighbourhood of $U(\bar{p})$, provided $\mathcal{J}_{\bar{p}}$ is taken sufficiently small. Condition (ii) is satisfied in many applications for almost every $\bar{p} \in \mathcal{R}(t)$.

Let us define a function h on $\mathcal{Y}_{\bar{p}}$ by

$$h : \mathcal{Y}_{\bar{p}} \to R\!\!\!l \ , \quad h(p) = \inf_{t(u)=p} f(u) \ . \tag{2.1}$$

As hypothesis 2.2.1. is assumed to hold, h is continuously differentiable on $\mathcal{Y}_{\bar{p}}$. Differentiating the expressions

$$h(p) = f(U(p)) \quad \text{and} \quad t(U(p)) = p, \quad p \in \mathcal{Y}_{\bar{p}}$$

with respect to p gives

$$h'(p) = \langle f'(U(p)), U'(p) \rangle \tag{2.2}$$

and

$$1 = \langle t'(U(p)), U'(p) \rangle \ . \tag{2.3}$$

By the multiplier rule there exists a number $\mu \in R\!\!\!l$ such that

$$f'(U(p)) = \mu \, t'(U(p)) \ , \tag{2.4}$$

where μ depends on U(p). Substituting (2.4) into (2.2) and using (2.3) we find

$$h'(p) = \mu \ (U(p)) \ .$$

From this result we obtain the following

PROPOSITION 2.2.2. *If hypothesis* 2.2.1. *is satisfied, every solution of* $\mathcal{P}_p$ *has the same unique multiplier* μ, *which may therefore be considered as a function of* p, *and which is related to the function* h(p) *by*

$$h'(p) = \mu(p) \ . \tag{2.5}$$

As f and t are assumed to be twice continuously differentiable, it follows from (2.4) that μ is a C^1-function, and we get by differentiating (2.4) with respect to p:

$$\{f''(U(p)) - \mu\, t''(U(p))\}.U'(p) = \mu'(p).t'(U(p))$$

and with (2.3) we find for the second variation

(2.6) $\quad s(U(p),\mu(p);U'(p)) := \langle\{f''(U(p)) - \mu t''(U(p))\}.U'(p),U'(p)\rangle = \mu'(p).$

As $U(p)$ is a solution of $\mathcal{P}_p$ we have as extremality property (c.f. section 1.6.)

(2.7) $\quad s(U(p),\mu(p);v) \geq 0$ for every $v \in V$ with $\langle t'(U(p)),v\rangle = 0$.

From these results we immediately obtain the following

PROPOSITION 2.2.3. *Assume that hypothesis* 2.2.1. *is satisfied and write* $\bar{\mu} = \mu(\bar{p})$ *and* $\bar{u} = U(\bar{p})$. *Then we have:*

(i) *if* $\mu'(\bar{p}) \geq 0$ *then*

$$s(\bar{u},\bar{\mu};v) \geq 0 \text{ for all } v \in V,$$

which means that $\bar{u}$ *satisfies the necessary conditions to be a minimal point of the functional* $\ell(\cdot,\bar{\mu}) = f - \bar{\mu}t$;
(ii) *if* $\mu'(\bar{p}) < 0$ *then* $\bar{u}$ *is not a minimal point of the functional* $\ell(\cdot,\bar{\mu})$. *In fact,* $\bar{u}$ *is a saddle point of* $\ell(\cdot,\bar{\mu})$ *in the following sense:* $\ell(\cdot,\bar{\mu})$ *is minimal at* $\bar{u}$ *with respect to all curves through* $\bar{u}$ *on the manifold* $\mathcal{M}_{\bar{p}}$, $\ell(\cdot,\bar{\mu})$ *is maximal at* $\bar{u}$ *with respect to the curve* $u(\varepsilon) = \bar{u} + \varepsilon\, U'(\bar{p})$, $\varepsilon \in R\ell$.

In the rest of this section we shall formulate in a canonical way an alternative variational formulation for the minimum value $h(\bar{p})$ of problem $\mathcal{P}_{\bar{p}}$. This alternative principle will be shown to hold if some local requirements (i.e. for p in some neighbourhood of $\bar{p}$) for $h(p)$ are fulfilled. However, this alternative turns out to be a useful device only if also some global requirements (i.e. for all $p \in R\ell$) are satisfied by the function $h(p)$.

The local requirement (which is only necessary to facilitate the heuristic analysis) is that h is a smooth curve in a neighbourhood

of $\bar{p}$ with

(2.9) $$h''(\bar{p}) \neq 0.$$

Then h is either convex or concave in a neighbourhood of $\bar{p}$.

HYPOTHESIS 2.2.4. Suppose h is defined for all p in some neighbourhood $\mathcal{V}_{\bar{p}}$ of $\bar{p}$, $h \in C^2(\mathcal{V}_{\bar{p}})$ and *strictly convex* on $\mathcal{V}_{\bar{p}}$, i.e. $h''(p) > 0$ for all $p \in \mathcal{V}_{\bar{p}}$.

With this hypothesis it is possible to define the *Legendre transform* of the function h. If we denote this Legendre transform by $^*h$, we have

(2.10) $$^*h(\mu) := \mu.p - h(p) \quad \text{for} \quad \mu \in \mathcal{V}_{\bar{\mu}} ,$$

where in the right hand side p has to expressed as a function of μ according to

(2.11) $$\mu = h'(p) ,$$

and where $\bar{\mu} = h'(\bar{p})$ and

(2.12) $$\mathcal{V}_{\bar{\mu}} = \{ h'(p) \mid p \in \mathcal{V}_{\bar{p}} \} .$$

The Legendre transform of $^*h$, to be denoted by $^{**}h$, is analogously given by

(2.13) $$^{**}h(q) = \mu.q - {}^*h(\mu) \quad \text{for } q \in \mathcal{V}_{\bar{p}}$$

where

(2.14) $$q = {}^*h'(\mu) .$$

Moreover, $^{**}h$ and h coincide on $\mathcal{V}_{\bar{p}}$:

(2.15) $$^{**}h(p) = h(p) \quad \text{for } p \in \mathcal{V}_{\bar{p}} .$$

Now an important observation to reach our goal is to note that $^*h$, as given by (2.10), (2.11) and (2.12) can also be described as

$$(2.16) \qquad {}^*h(\mu) = \sup_{q \in \mathcal{Y}_{\bar{p}}} [\mu \cdot q - h(q)] \quad \text{for} \quad \mu \in \mathcal{Y}_{\bar{\mu}}$$

and $^{**}h$ as

$$(2.17) \qquad {}^{**}h(p) = \sup_{\mu \in \mathcal{Y}_{\bar{\mu}}} [\mu \cdot p - {}^*h(\mu)] \quad \text{for } p \in \mathcal{Y}_{\bar{p}}$$

Substituting (2.16) into (2.17) and using (2.15) we find

$$(2.18) \qquad h(p) = {}^{**}h(p) = \sup_{\mu \in \mathcal{Y}_{\bar{\mu}}} \inf_{q \in \mathcal{Y}_{\bar{p}}} [h(q) + \mu(p-q)] \text{ for } p \in \mathcal{Y}_{\bar{p}} ,$$

valid for arbitrary function h satisfying hypothesis 2.2.4. Specializing to $p = \bar{p}$ and inserting the variational formulation for $h(\bar{p})$, i.e.

$$(2.19) \qquad h(\bar{p}) = \inf_{t(u)=p} f(u) ,$$

there results

$$(2.20) \qquad h(\bar{p}) = \sup_{\mu \in \mathcal{Y}_{\bar{\mu}}} \inf_{q \in \mathcal{Y}_{\bar{p}}} \inf_{t(u)=q} [f(u)-\mu q + \mu \bar{p}]$$

or equivalently

$$(2.21) \qquad h(\bar{p}) = \sup_{\mu \in \mathcal{Y}_{\bar{\mu}}} \inf_{u \in \Omega} [f(u)-\mu t(u)+\mu \bar{p}] ,$$

where

$$(2.22) \qquad \Omega = \{u \in V \mid t(u) = q, \quad q \in \mathcal{Y}_{\bar{p}}\} .$$

It is clear that we have obtained in this way an alternative variational characterization for the value $h(\bar{p})$ differing from (2.19), which we shall call the *formal dual problem*

$$(2.23) \qquad {}^*\mathcal{P}_{\bar{p}} : \sup_{\mu \in \mathcal{Y}_{\bar{\mu}}} \inf_{u \in \Omega} [f(u)-\mu(t(u)-\bar{p})] .$$

By construction, the supremum is attained for

$$\mu = \bar{\mu} = h'(\bar{p}), \tag{2.24}$$

and if problem $\mathcal{P}_{\bar{p}}$ has $\bar{u}$ as a solution, it follows from (2.23) that $\bar{u}$ satisfies

$$t(\bar{u}) = \bar{p} \tag{2.25}$$

and that $\bar{u}$ is a solution of

$$\inf_{u\in\Omega} [f(u) - \bar{\mu}t(u)] \tag{2.26}$$

(c.f. also theorem 2.3.5.).

Hence, if hypothesis 2.2.4. is satisfied, (2.23) gives a variational characterization for $\bar{\mu}$, which is the multiplier corresponding to $\bar{u}$ because of (2.24), and an alternative variational principle for $\bar{u}$, viz (2.26). Moreover, as $\bar{u}$ is an interior point of Ω, $\bar{u}$ satisfies (c.f. theorem 0.5.2.):

$$f'(\bar{u}) = \bar{\mu}t'(\bar{u}) \tag{2.27}$$

and

$$s(\bar{u},\bar{\mu};v) \geq 0 \quad \text{for all } v \in V. \tag{2.28}$$

Note that (2.27) agrees with the multiplier rule for $\bar{u}$ as a solution of $\mathcal{P}_{\bar{p}}$, and (2.28) agrees with property 2.2.3. (i)(as $\mu'(\bar{p}) = h''(\bar{p}) > 0$).

Let us now consider a concave function h.

HYPOTHESIS 2.2.5. Suppose h is defined for all p in some neighbourhood $\mathcal{V}_{\bar{p}}$ of $\bar{p}$, $h \in C^2(\mathcal{V}_{\bar{p}})$ and h is strictly concave on $\mathcal{V}_{\bar{p}}$, i.e. $h''(p)<0$ for all $p \in \mathcal{V}_{\bar{p}}$.

Then the Legendre transform $^*h$ is again given by (2.10), (2.11) on $\mathcal{V}_{\bar{\mu}}$, but instead of (2.16) one has the characterization

(2.29) $$ {}^{*}h(\mu) = \inf_{q \in \mathcal{Y}_{\bar{p}}} [\mu q - h(q)]. $$

This results in an expression for ${}^{**}h(\bar{p})$:

$$ {}^{**}h(\bar{p}) = \inf_{\mu \in \mathcal{Y}_{\bar{\mu}}} \sup_{q \in \mathcal{Y}_{\bar{p}}} \inf_{t(u)=q} [f(u) - \mu t(u) + \mu \bar{p}], $$

which gives rise to the following formal dual problem

(2.30) $$ {}^{*}\mathcal{P}_{\bar{p}} : \inf_{\mu \in \mathcal{Y}_{\bar{\mu}}} \sup_{q \in \mathcal{Y}_{\bar{p}}} \inf_{t(u)=q} [f(u) - \mu(t(u) - \bar{p})] . $$

The infimum being attained at $\mu = \bar{\mu}$, the remaining maxi-mini principle

$$ \sup_{q \in \mathcal{Y}_{\bar{p}}} \inf_{t(u)=q} [f(u) - \bar{\mu} t(u)] $$

expresses the saddle point property of $\bar{u}$ as a stationary point of the functional $\ell(\cdot, \bar{\mu}) = f - \bar{\mu} t$, in agreement with property 2.2.3.,(ii).

Concerning the relevance of the foregoing results, one may say that if it were known *a priori* that hypothesis 2.2.4 (or 2.2.5) is satisfied, the dual formulation (2.23) (2.30 respectively) gives another characterization for the value $h(\bar{p})$, a variational characterization and interpretation for $\bar{\mu}$ and, in case h is convex, a different variational principle for the constrained minimal solution of $\mathcal{P}_{\bar{p}}$. However, if the starting point is an investigation of problem $\mathcal{P}_{\bar{p}}$, it will be unknown whether hypothesis 2.2.4. or 2.2.5 is satisfied. Moreover, even if it were known that one of these hypothesis is fulfilled, and if $\mathcal{Y}_{\bar{p}}$ and $\mathcal{Y}_{\bar{\mu}}$ are known (such that Ω as defined by (2.22) is known in principle), problems ${}^{*}\mathcal{P}_{\bar{p}}$ and ${}^{*}\mathcal{P}_{\bar{p}}$ are still constrained extremum problems which, in general, will be as difficult as the original problem $\mathcal{P}_{\bar{p}}$.

Nevertheless, the foregoing treatment may illuminate somewhat the manipulations in the next section: naively speaking, one gets rid of the constraints in the formulation ${}^{*}\mathcal{P}_{\bar{p}}$ by defining as dual formulation

$$ \mathcal{P}_{\bar{p}}^{*} : \sup_{\mu \in R^{Z}} \inf_{u \in V} [f(u) - \mu(t(u) - \bar{p})] . $$

By doing so a much simpler dual problem results, but the relation with the original problem $\mathcal{P}_{\bar{p}}$ is no longer warranted in advance and has to be studied in detail. This will be done in the next section, but using the notion of *polar function* as defined in section 0.6 it will be clear (by comparing $h^*(\mu)$ with $^*h(u)$ and $h^{**}(p)$ with $^{**}h(p)$) that if hypothesis 2.2.4. is satisfied, the dual formulation $\mathcal{P}^*_{\bar{p}}$ and $^*\mathcal{P}_{\bar{p}}$ are the same provided $h^{**}(\bar{p}) = {}^{**}h(\bar{p})$, which in this case requires h to be subdifferentiable at $\bar{p}$. Hereafter it shall be shown that without any requirement as in hypothesis 2.2.4., subdifferentiability of h at $\bar{p}$, together with the existence of a solution of $\mathcal{P}_{\bar{p}}$, suffices to guarantee that $\mathcal{P}^*_{\bar{p}}$ is a meaningful dual formulation for $\mathcal{P}_{\bar{p}}$.

2.3. DUALITY PRINCIPLE.

In this section we shall consider the family of problems

$$(3.1) \qquad \mathcal{P}_p : \inf_{t(u)=p} f(u)$$

for all $p \in R\ell$ by setting

$$(3.2) \qquad \inf \mathcal{P}_p = \infty \quad \text{if } p \notin \mathcal{R}(t),$$

and we define

$$(3.3) \quad h : R\ell \to \bar{R}\ell,\ h(p) = \begin{cases} \inf \mathcal{P}_p & \text{if } p \in \mathcal{R}(t) \\ \infty & \text{else} \end{cases}$$

LEMMA 2.3.1. *An equivalent formulation of problem* $\mathcal{P}_p$ *is*

$$(3.4) \qquad \mathcal{P}_p : \inf_u \sup_\mu [f(u) - \mu(t(u) - p)]$$

where here and in the following $\inf_u$ *denotes* $\inf_{u \in V}$ *and* $\sup_\mu$ *denotes* $\sup_{\mu \in R\ell}$.

PROOF: Immediate from

$$\sup_\mu [f(u) - \mu(t(u)-p)] = \begin{cases} f(u) & \text{if } t(u) = p \\ \infty & \text{if } t(u) \neq p \end{cases} \qquad \square$$

Now we define - guided by the observations from the foregoing section - a variational problem $\mathcal{P}_p^*$.

DEFINITION 2.3.2. The *dual problem* $\mathcal{P}_p^*$ of the primal problem $\mathcal{P}_p$ is defined as

$$(3.5) \qquad \mathcal{P}_p^* : \sup_{\mu} \inf_{u} [f(u) - \mu(t(u) - p)].$$

The supremum of problem $\mathcal{P}_p^*$ will be denoted by $\sup \mathcal{P}_p^*$, and $\bar{\mu}$ is said to be a solution of $\mathcal{P}_{\bar{p}}^*$ if $\sup \mathcal{P}_{\bar{p}}^*$ is finite and

$$(3.6) \qquad \sup \mathcal{P}_{\bar{p}}^* = \inf_{u} [f(u) - \bar{\mu}(t(u) - \bar{p})].$$

The following lemma relates the polar and bipolar function of h to the functional $f - \mu t$ and to the dual problem respectively.

LEMMA 2.3.3. $\quad h^*(\mu) = - \inf_u [f(u) - \mu t(u)]$
$\qquad\qquad\quad h^{**}(p) = \sup \mathcal{P}_p^*$

PROOF: By defintion

$$h^*(\mu) = \sup_p [\mu p - h(p)] = \sup_p \sup_{t(u)=p} [-f(u) + \mu p]$$

$$= \sup_p \sup_{t(u)=p} [-f(u) + \mu t(u)] = \sup_u [- f(u) + \mu t(u)] ,$$

the last equality because of the fact that $\{u \in V \mid t(u)=p, \ p \in R\} = V$. With this result we immediately obtain

$$h^{**}(p) = \sup_{\mu} [p - h^*(\mu)] = \sup_{\mu} \inf_{u} [f(u) - \mu(t(u) - p)] = \sup \mathcal{P}_p^*. \qquad \square$$

LEMMA 2.3.4. $\quad \sup \mathcal{P}_p^* \leq \inf \mathcal{P}_p$

PROOF: As $h(p) = \inf \mathcal{P}_p$ and $h^{**}(p) = \sup \mathcal{P}_p^*$, the lemma states that $h^{**}(p) \leq h(p)$, which is property 0.6.6., (b). Without reference to this result, it is clear from the following steps:

$$\inf_u [f(u)-\mu(t(u)-p)] \leq f(u)-\mu(t(u)-p) \qquad \forall u \in V \quad \forall \mu \in R\ell$$

$$\sup_\mu \inf_u [f(u) - \mu(t(u)-p)] \leq \sup_\mu [f(u)-\mu(t(u)-p)] \qquad \forall u \in V$$

$$\sup_p \mathcal{P}_p^* \leq \inf_u \sup_\mu [f(u)-\mu(t(u)-p)] = \inf_p \mathcal{P}_p. \qquad \square$$

The following theorem shows that, under the stated conditions, problem $\mathcal{P}_{\bar{p}}^*$ is a variational characterization for the multiplier $\bar{\mu}$ corresponding to the solution $\bar{u}$ of $\mathcal{P}_{\bar{p}}$.

THEOREM 2.3.5. *Suppose* $\mathcal{P}_{\bar{p}}$ *has a solution, say* $\bar{u}$, *suppose* $\mathcal{P}_{\bar{p}}^*$ *has a solution* $\bar{\mu}$ *and suppose*

$$\inf \mathcal{P}_{\bar{p}} = \sup \mathcal{P}_{\bar{p}}^* . \tag{3.7}$$

Then $\bar{u}$ *is a solution of the unconstrained extremum problem*

$$\inf_u [f(u) - \bar{\mu}t(u)]; \tag{3.8}$$

consequently, if $f,t \in C^2(V,R\ell)$, *then* $\bar{u}$ *satisfies*

$$f'(\bar{u}) = \bar{\mu}\, t'(\bar{u}) \tag{3.9}$$

$$s(\bar{u},\bar{\mu};v) = \langle (f''(\bar{u})-\bar{\mu}\, t''(u)).v,v\rangle \geq 0 \quad \textit{for all } v \in V. \tag{3.10}$$

PROOF: By definition of $\bar{\mu}$:

$$\sup \mathcal{P}_{\bar{p}}^* = \inf_u [f(u) - \bar{\mu}(t(u) - \bar{p})],$$

and by definition of $\bar{u}$

$$\inf \mathcal{P}_{\bar{p}} = f(\bar{u}) = f(\bar{u}) - \bar{\mu}(t(\bar{u}) - \bar{p}) \quad \text{(because } t(\bar{u}) = \bar{p}).$$

From this it follows with (3.7) that $\bar{u}$ is a solution of (3.8), and the remaining part of the theorem follows from theorem 0.5.2.. $\square$

REMARK 2.3.6. A consequence of this theorem is that if problem $\mathcal{P}_{\bar{p}}^*$ admitsmore than one solution, then $t'(u) = 0$ for every solution of $\mathcal{P}_{\bar{p}}$ and hence $f'(\bar{u}) = 0$: every solution of $\mathcal{P}_{\bar{p}}$ is then a stationary point of the unconstrained functional f. Furthermore, if $\bar{\mu}$ is the unique solution of $\mathcal{P}_{\bar{p}}^*$, every solution of $\mathcal{P}_{\bar{p}}$ has this same value $\bar{\mu}$ as multiplier.

Because of the nice correspondence between solutions of $\mathcal{P}_{\bar{p}}$ and those of $\mathcal{P}_{\bar{p}}^*$ as expressed by theorem 2.3.5., we define

DEFINITION 2.3.7. Problem $\mathcal{P}_p$ is said to be *dual stable* if

(i) $\mathcal{P}_p$ has a solution

(ii) $\mathcal{P}_p^*$ has a solution

iii) $\sup \mathcal{P}_p^* = \inf \mathcal{P}_p$.

With the next lemma it will be possible to give an equivalent definition of dual-stability.

LEMMA 2.3.8. *The solution set of problem* $\mathcal{P}_p^*$ *coincides with the subdifferential of* h^{**} *at* p, *i.e.*

$$\mu \text{ is a solution of } \mathcal{P}_p^* \iff \mu \in \partial h^{**}(p).$$

PROOF: By definition

$$\mu \in \partial h^{**}(p) \iff h^{**}(p) + h^{***}(\mu) = \mu p$$

Using property 0.6.6. (c), i.e. $h^{***} = h^*$, and lemma 2.3.3. we find

$$\mu \in \partial h^{**}(p) \iff \sup \mathcal{P}_p^* = \inf_u [f(u) - \mu(t(u) - p)] \iff$$

$$\iff \mu \text{ is a solution of } \mathcal{P}_p^* \qquad \square$$

PROPOSITION 2.3.9. *Problem* $\mathcal{P}_p$ *is dual-stable if and only if*
(i) $\mathcal{P}_p$ *has a solution,*
(ii) h *is subdifferentiable at* p $\quad (\partial h(p) \neq \emptyset)$.

PROOF: Suppose $\mathcal{P}_p$ is dual stable. Then sup $\mathcal{P}_p^*$ = inf $\mathcal{P}_p$, i.e. $h^{**}(p) = h(p)$, and from property 0.6.6. (g) it follows that $\partial h(p) = \partial h^{**}(p)$. As $\mathcal{P}_p^*$ has a solution, according to lemma 2.3.8. $\partial h^{**}(p) \neq \emptyset$, and thus $\partial h(p) \neq \emptyset$ which means that h is subdifferentiable at p. On the other hand, suppose $\partial h(p) \neq \emptyset$. From property 0.6.6. (e) it then follows that $h(p) = h^{**}(p)$, and consequently $\partial h(p) = \partial h^{**}(p)$. Hence sup $\mathcal{P}_p^*$ = inf $\mathcal{P}_p$ and $\partial h^{**}(p) \neq \emptyset$ which implies with lemma 2.3.8. that $\mathcal{P}_p^*$ has a solution. This completes the proof. □

The aim is now to find a *stability-criterion,* i.e. conditions for the functionals f and t such that problem $\mathcal{P}_p$ is dual-stable. In the following we shall seek for a stability criterion which gives, for fixed functionals f and t, the set of values p for which $\mathcal{P}_p$ is stable, i.e. we shall characterize the dual-stable p-interval.

DEFINITION 2.3.10 The *dual-stable p-interval* is defined as

$$\mathcal{Y}_p := \{p \in \mathbb{R} \mid \mathcal{P}_p \text{ is dual-stable}\}.$$

It turns out to be possible to characterize this interval $\mathcal{Y}_p$ completely from some knowledge of the family of extremum problems $\mathcal{K}_\mu$.

DEFINITION 2.3.11. For the class of unconstrained extremum problems

$$\mathcal{K}_\mu : \inf [f(u) - \mu t(u)], \quad \mu \in \mathbb{R}$$

we define the intervals

$$\mathcal{Y}_\mu := \{\mu \in \mathbb{R} \mid \inf \mathcal{K}_\mu > -\infty\}$$

and

$$\hat{\mathcal{Y}}_\mu := \{\mu \in \mathcal{Y}_\mu \mid K_\mu \neq \emptyset\},$$

where

$$K_\mu := \{u \in V \mid u \text{ is a solution of } \mathcal{K}_\mu\}.$$

LEMMA 2.3.12. *$\mathcal{J}_\mu$ is a simply connected interval of $R\ell$, and $-\inf \mathcal{K}_\mu = h^*(\mu)$ is a finite, convex function on this interval.*

PROOF: In fact this lemma is nothing else than property 0.6.6. (a). To give a direct proof, we shall show that the function $k(\mu) := \inf \mathcal{K}_\mu$ satisfies for arbitrary $\lambda, \nu \in R\ell$

$$k(\lambda\mu + (1-\lambda)\nu) \geq \lambda.k(\mu) + (1-\lambda)k(\nu) \qquad \text{for all } \lambda,\ 0 < \lambda < 1,$$

from which the two statements follow. Therefore:

$$k(\lambda\mu + (1-\lambda)\nu) = \inf_u [f(u) - (\lambda\mu + (1-\lambda)\nu).\ t(u)]$$

$$= \inf_u \{\lambda[f(u) - \mu t(u)] + (1-\lambda)\ [f(u) - \nu t(u)]\}$$

$$\geq \inf_u \lambda.\ [f(u) - \mu t(u)] \ + \inf_u (1-\lambda)\ [f(u) - \nu t(u)]$$

$$= \lambda.k(\mu) + (1-\lambda).k(\nu)\ , \quad \text{for every } 0 < \lambda < 1. \qquad \square$$

The following lemma characterizes the interval $\hat{\mathcal{J}}_\mu$ for an important class of functionals.

LEMMA 2.3.13. *Let* $f: V \to R\ell$ *be* w.l.s.c. *and coercive on* V, *and let* $t: V \to R\ell$ *be* w.c. *Then we have* $\hat{\mathcal{J}}_\mu \equiv \mathcal{J}_\mu$. *Moreover,*
(i) *if* t *is bounded from above and from below on* V, *then* $\hat{\mathcal{J}}_\mu = R\ell$
(ii) *if* t *is bounded from above, then* $\hat{\mathcal{J}}_\mu \supset R\ell^+ = \{\mu \in R\ell \mid \mu \geq 0\}$
(iii) *if* t *is bounded from below, then* $\hat{\mathcal{J}}_\mu \supset R\ell^- = \{\mu \in R\ell \mid \mu \leq 0\}$

PROOF: Note that $f-\mu t$ is a w.l.s.c. functional on V for every $\mu \in R\ell$ (remark 0.2.6.(i)). Hence if $f-\mu t$ is bounded from below on V then $f-\mu t$ attains its infimum: see (proof of) theorem 0.5.3. and remark 0.5.4. Moreover, it follows that

$$\mathcal{J}_\mu \equiv \hat{\mathcal{J}}_\mu \supset \{\mu \in R\ell \mid f-\mu t \text{ is coercive on } V\}.$$

With this result, together with the coercivity of f on V, the three

statements follow. □

THEOREM 2.3.14. *Let* $\bar{\mu} \in \hat{\mathcal{J}}_\mu$. *Then* $\mathcal{P}_{\bar{p}}$ *is dual stable for* $\bar{p} \in \{t(u) \mid u \in K_{\bar{\mu}}\}$.

PROOF: For $\bar{p} \in \{t(u) \mid u \in K_{\bar{\mu}}\}$ we shall show that $\bar{\mu} \in \partial h(\bar{p})$ and that $\mathcal{P}_{\bar{p}}$ has a solution. The result then follows from proposition 2.3.9. Let $\bar{p} = t(\bar{u})$, with $\bar{u} \in K_{\bar{\mu}}$. Then (using lemma 2.3.3. in the first equality)

$$-h^*(\bar{\mu}) = \inf_u [f(u) - \bar{\mu}t(u)] = f(\bar{u}) - \bar{\mu}t(\bar{u}) = f(\bar{u}) - \bar{\mu}\bar{p} .$$

As $\bar{p} = t(\bar{u})$, we have by definition of $h(\bar{p})$: $h(\bar{p}) \leq f(\bar{u})$.

These results together imply that

$$h^*(\bar{\mu}) = \bar{\mu}\bar{p} - f(\bar{u}) \leq \bar{\mu}\bar{p} - h(\bar{p}).$$

On the other hand, by definition of $h^*(\bar{\mu})$:

$$h^*(\bar{\mu}) = \sup_p [\bar{\mu}p - h(p)] \geq \bar{\mu}\bar{p} - h(\bar{p}).$$

From these results it follows that $h(\bar{p}) = f(\bar{u})$, which means that $\bar{u}$ is a solution of $\mathcal{P}_{\bar{p}}$, and that $h^*(\bar{\mu}) + h(\bar{p}) = \bar{\mu}\bar{p}$, which means that $\bar{\mu} \in \partial h(\bar{p})$ (c.f. definition 0.6.2). This completes the proof. □

THEOREM 2.3.15. *(dual-stability criterion)*
The dual-stability interval is completely characterized by

$$\mathcal{J}_p = \{t(u) \mid u \in K_\mu, \mu \in \hat{\mathcal{J}}_\mu\}.$$

PROOF. The inclusion $\mathcal{J}_p \subset \{t(u) \mid u \in K_\mu, \mu \in \hat{\mathcal{J}}_\mu\}$ is an immediate consequence of definition 2.3.7. and theorem 2.3.5. The reversed inclusion is the contents of theorem 2.3.14. This completes the proof. □

The foregoing results show that there is a one-to-one correspondence between the solutionsof the constrained extremum problems $\{\mathcal{P}_p\}_{p \in \mathcal{J}_p}$ and the solutions of the unconstrained extremum problems $\{\mathcal{U}_\mu\}_{\mu \in \hat{\mathcal{J}}_\mu}$. This means that: (i) if $\bar{u} \in K_{\bar{\mu}}$, then $\bar{u}$ is a

solution of $\mathcal{P}_{\bar{p}}$ and $\bar{\mu}$ is a solution of $\mathcal{P}_{\bar{p}}^*$ where $\bar{p} = t(\bar{u})$ and
(ii) if $\bar{u}$ is a solution of $\mathcal{P}_p$, where $p \in \mathcal{Y}_p$, and if $\bar{\mu}$ is a solution of $\mathcal{P}_p^*$, then $\bar{\mu} \in \hat{\mathcal{Y}}_\mu$ and $\bar{u} \in K_{\bar{\mu}}$.

2.4. INVERSE EXTREMUM PRINCIPLES.

We start this section with some local investigations. Suppose $\hat{u} \in V$ satisfies for some $\lambda \in R$

$$f'(\hat{u}) = \lambda t'(\hat{u}) \tag{4.1}$$

Then, if $t'(\hat{u}) \neq 0$, $\hat{u}$ is a regular point of the manifold

$$\mathcal{T} := \{u \in V \mid t(u) = t(\hat{u})\} ,$$

and $\hat{u}$ is a constrained stationary point of f with respect to $\mathcal{T}$. Moreover, if $\lambda \neq 0$ then $f'(\hat{u}) \neq 0$, and $\hat{u}$ is a regular point of the manifold

$$\mathcal{F} := \{u \in V \mid f(u) = f(\hat{u})\}$$

and is a constrained stationary point of t with respect to $\mathcal{F}$. From these observations the following result is easily obtained.

PROPOSITION 2.4.1. *If* $f'(\hat{u}) \neq 0$ *and* $t'(\hat{u}) \neq 0$ *then*

$\hat{u}$ *is a constrained stationary point of* f *with respect to* $\mathcal{T}$ *if and only if*

$\hat{u}$ *is a constrained stationary point of* t *with respect to* $\mathcal{F}$

Note that if $\hat{u}$ satisfies (4.1) and $f'(\hat{u}) \neq 0$, $t'(\hat{u}) \neq 0$, then the tangent spaces at $\hat{u}$ to $\mathcal{T}$ and to $\mathcal{F}$ coincide:

$$\{v \in V \mid \langle t'(\hat{u}), v\rangle = 0\} = \mathcal{N}(t'(\hat{u})) = \mathcal{N}(f'(\hat{u})) = \{v \in V \mid \langle f'(\hat{u}), v\rangle = 0\}. \tag{4.2}$$

From this key observation it follows that some extremality properties of f on the manifold $\mathcal{T}$ at $\hat{u}$ may be transferred to extremality proper-

ties of t on the manifold $\mathcal{F}$ at $\hat{u}$:

PROPOSITION 2.4.2. *Let* $f \in C^2(V,Rl)$, $t \in C^2(V,Rl)$. *Suppose* $\hat{u} \in \mathcal{T}$ *satisfies* (4.1) *with* $f'(\hat{u}) \neq 0$ *and* $t'(\hat{u}) \neq 0$ *Suppose there exists a neighbourhood* $M_t(\hat{u}) \subset \mathcal{T}$ *of* $\hat{u}$ *in* $\mathcal{T}$ *and a constant* $c > 0$ *such that*

$$f(u) - f(\hat{u}) \geq c.||u-\hat{u}||^2 \quad \text{for every } u \in M_t(\hat{u}) \subset \mathcal{T} \tag{4.3}$$

(hence $\hat{u}$ *is a constrained minimum of* f *with respect to* $\mathcal{T}$ *).*
Then there exists a neighbourhood $M_f(\hat{u}) \subset \mathcal{F}$ *of* $\hat{u}$ *in* $\mathcal{F}$ *and a constant* $d > 0$ *such that*
(i) *if* $\lambda > 0$:

$$t(\hat{u}) - t(u) \geq d.||u-\hat{u}||^2 \quad \text{for every } u \in M_f(\hat{u}) \subset \mathcal{F} \tag{4.4}$$

(hence $\hat{u}$ *is a constrained maximum of* t *with respect to* $\mathcal{F}$ *)*
(ii) *if* $\lambda < 0$:

$$t(u) - t(\hat{u}) \geq d.\ ||u-\hat{u}||^2 \quad \text{for every } u \in M_f(\hat{u}) \tag{4.5}$$

(hence $\hat{u}$ *is a constrained minimum of* t *with respect to* $\mathcal{F}$ *).*

PROOF: Consider points in a neighbourhood of $\hat{u}$ at :

$$u(\varepsilon,v) = \hat{u} + \varepsilon v + \phi(\varepsilon v) \quad , \ ||v|| = 1, \quad v \in \mathcal{N}(t'(\hat{u})).$$

From lemma 1.3.7. we have

$$||\phi(\varepsilon v)|| = O(\varepsilon^2) \qquad \text{for } \varepsilon \to 0$$

and hence

$$||u-\hat{u}||^2 = ||\varepsilon v + \phi(\varepsilon v)||^2 = \varepsilon^2 + O(\varepsilon^3) \text{ for } \varepsilon \to 0 .$$

Equation 1.(5.8), specialized to $T = t: V \to Rl$ gives

$$f(u(\varepsilon,v)) - f(\hat{u}) = \frac{1}{2}\varepsilon^2 < \{f''(\hat{u}) - \lambda t''(\hat{u})\}.v,v> + o(\varepsilon^2) \quad \text{for } \varepsilon \to 0.$$

From these result, together with (4.3) it follows that for $|\varepsilon|$ sufficiently small:

$$<\{f''(\hat{u}) - \lambda t''(\hat{u})\}.v,v> \geq 2c\ \varepsilon^2 \quad \forall\, v \in \mathcal{N}(t'(\hat{u})),\ ||v|| = 1.$$

From this we deduce with (4.2):

$$-\lambda\langle\{t''(\hat{u}) - \lambda^{-1}.f''(\hat{u})\}.v,v\rangle \geq 2c||v||^2 \qquad \forall v \in \mathcal{M}(f'(\hat{u})).$$

The results then follow from theorem 1.6.4.(iv). □

With local investigations as described above it is not possible to relate global extrema of f with respect to $\mathcal{T}$ to global extrema of t with respect to $\mathcal{F}$. To study this relationship we shall consider three classes of constrained extremum problems.

DEFINITION 2.4.3.

(4.6) $\mathcal{P}_p$: $\inf_{t(u)=p} f(u)$ $\quad h(p) := \inf \mathcal{P}_p$, $p \in \mathcal{R}(t)$.

(4.7) $\mathcal{S}_r$: $\sup_{f(u)=r} t(u)$ $\quad s(r) := \sup \mathcal{S}_r$, $r \in \mathcal{R}(f)$.

(4.8) $\mathcal{Q}_r$: $\inf_{f(u)=r} t(u)$ $\quad q(r) := \inf \mathcal{Q}_r$, $r \in \mathcal{R}(f)$.

Specifically we shall investigate for which values of p the solutions of $\mathcal{P}_p$ can also be obtained from $\mathcal{S}_r$ or $\mathcal{Q}_r$ for some $r \in \mathcal{R}(f)$.

PROPOSITION 2.4.4. (i) *For every* $p \in \mathcal{R}(t)$ *for which* $\mathcal{P}_p$ *has a solution we have* $s(h(p)) \geq p$.
(ii) *If* $\mathcal{P}_p$ *has a solution, and if* $s(h(p)) = p$, *then the solution sets of* $\mathcal{P}_p$ *and* $\mathcal{S}_{h(p)}$ *coincide.*

PROOF: Let $\bar{u}$ be a solution of $\mathcal{P}_p$, then $f(\bar{u}) = h(p)$ and $t(\bar{u}) = p$.

Then

$$s(h(p)) = \sup_{f(u)=h(p)} t(u) \geq t(\bar{u}) = p,$$

which proves (i). Moreover, if $s(h(p)) = p$, then $\bar{u}$ is clearly a solution of $\mathcal{S}_{h(p)}$. On the other hand, if $\hat{u}$ is a solution of $\mathcal{S}_{h(p)}$, then

$$s(h(p)) = t(\hat{u}) \quad \text{and} \quad h(p) = f(\hat{u}).$$

Hence if $s(h(p)) = p$, then $\hat{u}$ is a solution of $\mathcal{P}_p$. This proves (ii).□

THEOREM 2.4.5. *Suppose* $p \in \mathcal{R}(t)$ *is such that* $\mathcal{P}_p$ *has a solution and that*

$$h(\zeta) > h(p) \quad \textit{for every} \quad \zeta > p,\ \zeta \in \mathcal{R}(t). \tag{4.9}$$

Then

$$s(h(p)) = p.$$

PROOF: The proof goes by contradiction. Suppose $s(h(p)) = p + \alpha$ for some $\alpha > 0$. (because of lemma 2.4.4.(i) we need not to investigate the possibility $\alpha < 0$). Then

$$\sup_{f(u)=h(p)} t(u) = p + \alpha,$$

which means that there exists an element $\hat{u} \in V$ and $\zeta \in Rl$ with $p < \zeta < p + \alpha$ sucht that $t(\hat{u}) = \zeta$ and $f(\hat{u}) = h(p)$. From this it follows that

$$h(\zeta) := \inf_{t(u)=\zeta} f(u) \leq f(\hat{u}) = h(p).$$

As $\zeta > p$ this contradicts the assumption (4.9). Hence $s(h(p)) = p$. □

The following corollaries follow immediately from the foregoing theorem.

COROLLARY 2.4.6. *Suppose* $p_o \in Rl$ *is such that* $\mathcal{P}_p$ *has a solution for every* $p > p_o$, $p \in \mathcal{R}(t)$, *and such that* h *is monotonically increasing for* $p > p_o$, $p \in \mathcal{R}(t)$. *Then* $s(h(p)) = p$ *for every* $p \in \mathcal{R}(t)$, $p > p_o$. *In other words: in that case the function* $s(r)$ *on* $\{h(p) \mid p > p_o = p \in \mathcal{R}(t)\}$ *is the inverse of the function* $h(p)$ *on* $\{p \mid p > p_o,\ p \in \mathcal{R}(t)\}$.

COROLLARY 2.4.7. *Suppose* $\mathcal{P}_p$ *has a solution and suppose* $\mu \in \partial h(p) \neq \emptyset$ *with* $\mu > 0$. *Then* $s(h(p)) = p$.

In the same way as the foregoing results relate the problems $\mathcal{J}_r$ to $\mathcal{P}_p$, it is possible to relate the problems $\mathcal{Q}_r$ to $\mathcal{P}_p$. We

merely state the results.

THEOREM 2.4.8. (i) *For every* $p \in \mathcal{R}(t)$ *for which* $\mathcal{P}_p$ *has a solution we have* $q(h(p)) \leq p$.
(ii) *If* $\mathcal{P}_p$ *has a solution and if* $q(h(p)) = p$ *then the solution sets of* $\mathcal{P}_p$ *and* $\mathcal{Q}_{h(p)}$ *coincide.*
(iii) *Suppose* $p \in \mathcal{R}(t)$ *is such that* $\mathcal{P}_p$ *has a solution and that*

$$h(\zeta) > h(p) \quad \textit{for every } \zeta < p, \ \zeta \in \mathcal{R}(t).$$

Then $q(h(p)) = p$.
(iv) *If* $p_o \in \mathbb{R}$ *is such that* $\mathcal{P}_p$ *has a solution for every* $p < p_o$, $p \in \mathcal{R}(t)$ *and such that* h *is monotonically decreasing for* $p < p_o$, *then* $q(h(p)) = p$ *for every* $p < p_o$, $p \in \mathcal{R}(t)$: $q(r)$ *on* $\{h(p) \mid p < p_o, p \in \mathcal{R}(t)\}$ *is the inverse of the function* $h(p)$ *on* $\{p \mid p < p_o, p \in \mathcal{R}(t)\}$.
(v) *If* $\mathcal{P}_p$ *has a solution and if* $\mu \in \partial h(p) \neq \emptyset$ *with* $\mu < 0$ *then* $q(h(p)) = p$.

REMARK 2.4.9. Because of the results described in corollary 2.4.6. and theorem 2.4.8. (iv), we want to call the problems $\mathcal{J}_r$ and $\mathcal{Q}_r$ *inverse* extremal problems corresponding to $\mathcal{P}_p$.

REMARK 2.4.10. In the applications of the next section f and t satisfy the conditions of theorem 1.6.4.(i). Then the existence of a solution of $\mathcal{P}_p$ for every $p \in \mathcal{R}(t)$ is warranted in advance, but $\mathcal{J}_r$ and $\mathcal{Q}_r$ need not to have solutions as we shall see below. Nevertheless, if it *is* known that e.g. $\mathcal{Q}_r$ has solutions for every $r > r_o$, properties of the function $q(r)$ can be used to obtain information about the function $h(p)$ on $\{q(r) \mid r > r_o\}$. Therefore it is only necessary to replace the rôle of f and t in theorem 2.4.8.

It is illustrative to consider the variational principles $\mathcal{J}_r$ and $\mathcal{Q}_r$ also from another point of view. To that end we consider the sets

$$\bar{B}_r := \{u \in V \mid f(u) \leq r\},$$

and we compare problems $\mathcal{P}_r$ and $\mathcal{Q}_r$ with the extremum problems

$$\bar{\mathcal{P}}_r \quad : \quad \sup_{f(u)\leq r} \; t(u) \; ,$$

$$\bar{\mathcal{Q}}_r \quad : \quad \inf_{f(u)\leq r} \; t(u) \; .$$

The idea is that the manifold $\{u \mid f(u) = r\} \subset V$ is the boundary of the set $\bar{B}_r$, which implies that if t is known to have a maximum or minimum at $\bar{B}_r$ which does not lie in the interior then this extremal element must lie on the boundary and hence be a solution of $\mathcal{P}_r$ or $\mathcal{Q}_r$. To make any progress in this direction we assume that

f is w.l.s.c. and coercive on V,

t is w.c. on V.

Then f is bounded from below on V and attains its infimum (c.f. theorem 0.5.4. and remark 0.5.5.). Therefore it is no restriction to assume that f satisfies

$$f(0) = 0, \quad f(u) \geq 0 \quad \forall u \in V \qquad (f'(0) = 0, \; \mathcal{R}(f) = R\mathcal{l}^+).$$

Moreover, we have the following

LEMMA 2.4.11. *The set* $\bar{B}_r$, *for* $r > 0$, *is bounded and weakly sequentially closed. Consequently,* t *is bounded from below and above and attains its maximum and minimum at every* $\bar{B}_r$, $r > 0$, *say at the points* M_r *and* m_r *respectively.*

PROOF: The boundedness of $\bar{B}_r$ follows from the coercivity of f. Moreover, if $u_n \rightharpoonup \hat{u}$ in V with $\{u_n\} \subset \bar{B}_r$, then $f(u_n) \leq r \;\; \forall n$. As f is w.l.s.c. we have $f(\hat{u}) \leq \liminf f(u_n) \leq r$. Hence $\hat{u} \in \bar{B}_r$ which shows that $\bar{B}_r$ is weakly sequentially closed. The rest of the lemma follows from theorem 0.5.6. □

Of course it is possible that both the maximal and the minimal

value of t are attained at interior points of $\bar{B}_r$. Then we have $t'(M_r) = 0 = t'(m_r)$, and problems $\mathcal{Q}_r$ and $\mathcal{J}_r$ need not to have a solution. Hence if it is known that t has only one stationary point on V, then at least one of these two extremal points must lie on the boundary of $\bar{B}_r$:

COROLLARY 2.4.12 *Suppose* $t(0) = 0$ *and* $t'(u) = 0$ *if and only if* $u = 0$. *Then, if* t *takes positive values at every neighbourhood of* $u = 0$ *we have*

$$\sup_{f\leq r} t(u) = \sup_{f=r} t(u) = t(M_r) \quad \textit{and } f(M_r) = r,$$

i.e. M_r *lies at the boundary of* $\bar{B}_r$ *and* $\bar{M}_r$ *is a solution of* $\mathcal{J}_r$ *for every* $r > 0$. *Moreover, if for some* $r_o > 0$, t *takes also negative values at the boundary of* $\bar{B}_{r_o}$, *then*

$$\inf_{f\leq r} t(u) = \inf_{f=r} t(u) = t(m_r) \quad \textit{and } f(m_r) = r,$$

i.e. m_r *lies at the boundary of* $\bar{B}_r$ *and* m_r *is a solution of* $\mathcal{Q}_r$ *for every* $r > r_o$.

2.5. APPLICATIONS.

In this section we shall show how the abstract results of the foregoing sections can be applied to two specific problems. For each of these problems, the unconstrained extremum problems $\mathcal{K}_\mu$ have been extensively studied in literature and we advantageously use the obtained results for the investigation of the problems $\mathcal{P}_p$.

Euler-buckling.

The first system to be considered deals with the stationary states of an elastic line and serves as a model for the buckling of a thin, inextensible rod. In fact this problem was studied in section 1.7., but here we describe the configuration with $\theta(s)$, which is the angle between the tangent to the elastic line and the positive x-axis

(c.f. remark 1.7.2. (ii). For simplicity we take as boundary conditions

$$\theta(0) = \theta(\ell) = 0,$$

which means that, taking $\underline{r}(0) = \underline{0}$, the other endpoint of the line is no longer restricted to lie on the x-axis. Looking for configurations which have least bending energy for given distance in x-direction of the endpoints, may be formulated as follows

$$(5.1)\quad \mathcal{P}_p: \quad \inf_{t(\theta)=p} f(\theta) \quad \text{with} \quad \left| \begin{array}{l} f(\theta) = \frac{1}{2}\int_0^{\ell} \theta_s^{\,2}\, ds \\ \\ t(\theta) = \int_0^{\ell} (1-\cos\theta)\, ds \end{array} \right. \quad \text{for } \theta \in H_o^1(0,\ell).$$

It may be noted in advance that f and t satisfy hypothesis 2.4.4. Moreover

$$(5.2)\qquad \mathcal{R}(t) = [o,\ p_o) \qquad \text{with } p_o = 2\ell,$$

and

$$(5.3)\qquad t'(\theta) = \sin\theta = 0 \iff \theta = 0 \iff p = 0$$

From this it follows that $\mathcal{P}_p$ has a solution for every $p \in \mathcal{R}(t)$, and for $p \in (0,p_o)$ this solution satisfies for some unique $\mu \in R\ell$:

$$(5.4)\qquad f'(\theta) = \mu\, t'(\theta) \iff \theta_{ss} + \mu \sin\theta = 0.$$

[Note that although (5.4) is formally an equation in $(H_o^1)^* = H^{-1}$, every solution of (5.4) is actually a C^∞-function]. The multiplier μ has a physical interpretation in this case: it is proportional to the horizontal component of the compressive load necessary to maintain the rod in the required position. The unconstrained extremum problems $\mathcal{K}_\mu$ are :

(5.5) $$\mathcal{K}_\mu : \inf \int_0^\ell ds \left[\frac{1}{2}\theta_s^2 - \mu(1-\cos\theta)\right],$$

and from lemma 2.3.13 it follows with (5.2) that $\hat{\mathcal{J}}_\mu = R\ell$.
It is easily seen that

(5.6) for $\mu < 0$: $\inf \mathcal{K}_\mu = 0$, $K_\mu = \{0\}$.

For $\mu \geq 0$ the extremum problems $\mathcal{K}_\mu$ are well known in literature. Considered as a non-linear eigenvalue problem, $\mu > 0$ and fixed, the solutions of (5.4) were investigated: this problem serves as an example in almost every introduction to bifurcation theory. Moreover, the solutionscan be explicitely expressed in terms of Jacobi elliptic functions(see e.g. Stakgold [19], Reiss [20] and van der Varst [21]). From the available information we emphasize the following results.

PROPOSITION 2.5.1.

(5.7) (i) *for* $0 \leq \mu \leq \mu_1$, $\inf \mathcal{K}_\mu = 0$ *and* $K_\mu = \{0\}$, *where*

$\mu_1 = \pi^2/\ell^2$ *is the first eigenvalue of the linearized (around* $\theta = 0$*) eigenvalue problem corresponding to* (5.4), *i.e.* $\theta_{ss} + \mu\theta = 0$, $\theta(0) = \theta(\ell) = 0$;

(5.8) (ii) *for* $\mu > \mu_1$, $\inf \mathcal{K}_\mu < 0$ *and* $K_\mu = \{ \pm \theta_\mu \}$,

where θ_μ *is a solution of* (5.4) *uniquely determined (apart from sign) by the fact that it has no zero's in the interval* $(0,\ell)$. *The solutions* $\{\theta_\mu\}_{\mu>\mu_1}$ *form the socalled*

first bifurcation branch (first buckling modes) and on this branch t *and* f *are monotone increasing functions of* μ, t *running from* 0 *to* p_o *and* f *from* 0 *to* ∞.

With these results we may apply the theory as developed in section 2.3. and 2.4.:

COROLLARY 2.5.2. *The dual stable interval of* (5.1) *is given by*

(5.9) $\mathcal{I}_p = [o, p_o)$;

the first bifurcation branch may also be parameterized with $p \in (0, p_o)$, *and the solutions of* $\mathcal{P}_p$, $p \geq 0$ *are in a one-to-one correspondence with the solutions of* $\mathcal{K}_\mu$, $\mu \geq 0$. *Moreover, these solutions can also be characterized by*

(5.10) $$\mathcal{J}_r : \sup_{f(\theta)=r} t(\theta) \quad , \quad r \geq 0$$

and $s(r) := \sup \mathcal{J}_r$ *on* $[0,\infty)$ *is the inverse of the function* $h(p)$ *on* $[0, p_o)$:

(5.11) $$s(h(p)) = p \quad \textit{for} \quad p \in [0, p_o)$$

REMARKS 2.5.4. (i) See figures 2.5.1., 2.5.2. and 2.5.3.
(ii) In agreement with (5.6) and (5.7) we have

(5.12) $$\partial h(0) = (-\infty, \mu_1),$$

whereas $\partial h(p) = \{ h'(p)\} = \{\mu(p)\}$ for $p \in (0, p_o)$.
(iii) From the corollary, especially from (5.10), it follows that the solutions on the first bifurcation branch have also the property that for given bending energy the distance in x-direction of the endpoints is as large as possible.
(iv) The problems $\mathcal{Q}_r$ have no solution for $r > 0$: the infimum of the functional t on the set $\bar{B}_r = \{\theta \in H_o^1 \mid f(\theta) \leq r\}$ is attained for every $r > 0$ at the interior point $\theta \equiv 0$:

$$q(r) = \inf_{f(\theta)=r} t(\theta) = \inf_{f(\theta)\leq r} t(\theta) = t(0) = 0$$

(the second equality follows from property 0.2.8. as $f(\theta)$ is a norm equivalent to the H^1-norm on H_o^1).

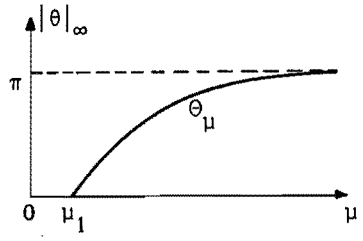

fig. 2.5.1.

The first branch in a bifurcation diagram $|\theta|_\infty$ v.s. μ.

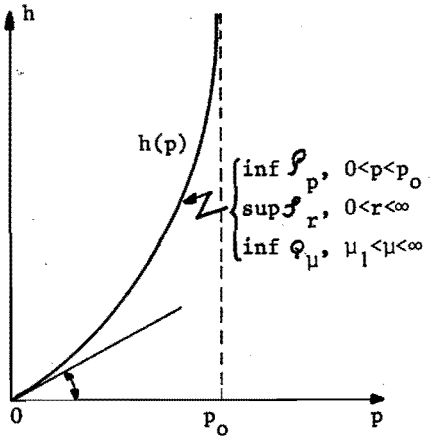

fig. 2.5.2.

Graph of the function h(p); indicated are several variational characterizations for the extremal elements of $\mathcal{P}_p$.

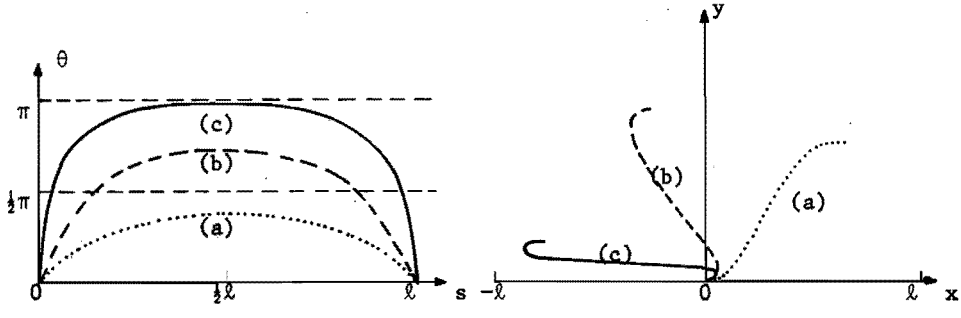

fig. 2.5.3.

Qualitative pictures of the function Θ_μ for three different values of μ and corresponding graphs of the buckled rod (with $\underline{r}(\underline{0}) = \underline{0}$). The values of μ correspond to values of p with (a) $0<p<\ell$, (b) $\ell<p<2\ell = p_o$, (c) p close to p_o ($p<p_o$).

As a last example we consider the constrained extremum problems

$$(5.13)\qquad \mathcal{P}_p : \inf_{t(u)=p} f(u) \quad \text{with} \quad \left| \begin{array}{l} f(u) = \frac{1}{2}\int_0^{\ell} u_x^2\, dx \\ t(u) = \int_0^{\ell} \Gamma(u)dx \end{array} \right. \quad \text{on } H_0^1(0,\ell)$$

Here Γ is a C^3-function on $R\ell$ with $\Gamma(0) = 0$ for which $\gamma(u) := \frac{d\Gamma(u)}{du}$ satisfies

(i) $\gamma(0) = 0, \quad \gamma'(0) = 1$

(ii) $\lim_{|u|\to\infty} \frac{\gamma(u)}{u} \leq 0$

(iii) $\gamma''(u).u < 0$ for every $u \neq 0$.

To simplify matters we shall consider as a specific example

$$(5.14)\qquad \gamma(u) = u - u^3, \quad t(u) = \int_0^{\ell} dx\, [\tfrac{1}{2}u^2 - \tfrac{1}{4}u^4].$$

For these functionals f and t hypothesis 2.4.4. is satisfied. The functional t is bounded from above but not from below

$$(5.15)\qquad \mathcal{R}(t) = (-\infty, p_o) \quad \text{where } p_o = \tfrac{1}{4}\ell.$$

For every $p \in \mathcal{R}(t)$, problem $\mathcal{P}_p$ has at least one solution, and as

$$(5.16)\qquad t'(u) = 0 \text{ on } H_0^1 \iff u = 0 \implies p = 0$$

(note that $u = 0$ is an isolated, but not the only, point of the manifold $\{u \mid t(u) = 0\}$), it follows that for $p \neq 0$ every solution satisfies for some unique multiplier $\mu \in R\ell$

$$(5.17)\qquad f'(u) = \mu\, t'(u) \quad , \quad -u_{xx} = \mu(u - u^3).$$

The class of unconstrained extremum problems $\mathcal{U}_\mu$ is now given by

(5.18) $$\mathcal{K}_\mu : \inf_u \int_0^\ell dx\, [\tfrac{1}{2} u_x^2 - \mu(\tfrac{1}{2} u^2 - \tfrac{1}{4} u^4)].$$

It is not difficult to show that

(5.19) $$\inf \mathcal{K}_\mu = -\infty \quad \text{if } \mu < 0.$$

From this result, together with lemma 2.3.13 it follows that

(5.20) $$\hat{\eta}_\mu = [0, p_o).$$

The non-linear eigenvalue problem (5.17) was studied for $\mu > 0$ by Chafee & Infante [22] and Henry [23], who were interested in the stability of these solutions for the evolution equation

(5.21) $$w_t = w_{xx} + \mu(w - w^3).$$

(See also Ambrosetti & Rabinowitz [24] who give variational characterizations for every solution of (5.17).
From these references it follows that problem (5.18) has for $\mu \geq 0$ the same qualitative properties as described in proposition 2.5.1. The extremal solutions, to be denoted by U_μ instead of Θ_μ, now satisfy $|U_\mu|_\infty = \max_{o \leq x \leq \ell} |U_\mu(x)| \to 1$ for $\mu \to \infty$. Apart from this aspect, the qualitative behaviour of the first bifurcation branch is as in figure 2.5.1.[For the specific example under consideration, i.e. for t given by (5.14), the solutions of (5.17) may again be expressed in terms of elliptic functions. In fact, problems (5.4) and (5.17) are known to be related by a non-linear transformation]. Consequently, the same results as desribed in corollary 2.5.3. are valid for problems $\mathcal{P}_p$ given in (5.13) with $p \geq 0$. Moreover it has been proved in the given references that the solutions U_μ on the first bifurcation branch are stable stationary solutions of (5.21).

For $p < 0$ matters are somewhat more complicated. We know in advance that $\mathcal{P}_p$ has at least two solutions $\pm U_p$ for every $p < 0$, and

that such a solution U_p satisfies (5.17) for a unique multiplier $\mu(p)$. Let us start with some technical results which are not difficult to prove.

LEMMA 2.5.4. (i) *If* $u \in H_o^1$ *satisfies* $t(u) = p < 0$, *then* $|u|_\infty^2 > 2$.
(ii) *If* u *is a solution of* (5.17) *for which* $t(u) = p < 0$, *then necessarily* $\mu < 0$. *In particular*

$$p < 0 \Rightarrow \mu(p) < 0. \tag{5.22}$$

(iii) *Every solution* u *of* (5.17) *satisfies*

$$< \{f''(u) - \mu t''(u)\}.u,u > = 2\mu \int_0^\ell u^4 dx \tag{5.23}$$

Consequently, every solution of (5.17) *with* $\mu < 0$ *is not a minimal point of the functional* $f-\mu t$ *on* H_o^1.
In particular, the solutions $\pm U_p$ *of* $\mathcal{P}_p$ *with* $p < 0$ *are saddle points of the functional* $f-\mu t$ *on* H_o^1.
(iv) *For the function* $h(p)$ *we have*

$$h(p) \to \infty \text{ for } p \to -\infty \tag{5.24}$$

and

$$\mu(p) \uparrow 0 \text{ for } p \to -\infty . \tag{5.25}$$

REMARKS 2.5.6. (i) From property (i) above it follows that the solutions of (5.17) with $\mu < 0$ does not bifurcate from the zero solution $u \equiv 0$.
(ii) From (iii) above it can be shown that every solution of (5.17) with $\mu < 0$ is an unstable stationary solution of the evolution equation (5.21).

Up to now nothing has been said about the continuity of the function $h(p)$ for $p < 0$. This is most easily investigated via the problems $\mathcal{Q}_r$. By considering the functional t on the closed balls $\bar{B}_r = \{u \mid f(u) \leq r\} \subset H_o^1$ it can be shown that for the problems

$$\mathcal{Q}_r \quad : \quad \inf_{f=r} \; t(u)$$

there exists some $r_o > 0$ such that

(i) $\mathcal{Q}_r$ has no solution if $0 < r < r_o$: the infimum of t on $\bar{B}_r$ is attained at the interior point $u \equiv 0$.

(ii) $\mathcal{Q}_r$ has a solution if $r > r_o$, and the function $q(r) := \inf \mathcal{Q}_r$ for $r > r_o$ is monotonically decreasing from 0 to $-\infty$ and is continuous.

From this last result it follows with theorem 2.4.8 (iii) and remark 2.4.10 that

$$h(q(r)) = r \qquad \text{for} \quad r > r_o,$$

from which we deduce that h(p) is a continuous function of $p \in (-\infty, 0)$ and is monotonically decreasing (in agreement with (5.25) ; see figure 2.5.4.). Moreover, the solutions of $\mathcal{P}_p$ for $p < 0$ are in a one-to-one correspondence with the solutions of $\mathcal{Q}_r$, $r > r_o$.

REMARK 2.5.7. It has not been found possible to specify the precise values of r_o and of $\mu(0^-) = \lim_{p \uparrow o} \mu(p)$. However it can be shown that these numbers satisfy the following estimates

$$2/\ell < r_o < \infty$$

$$-\infty < \mu(0^-) < -\mu_1 \quad , \quad \text{where } \mu_1 = \frac{\tilde{\pi}^2}{\ell^2} \, .$$

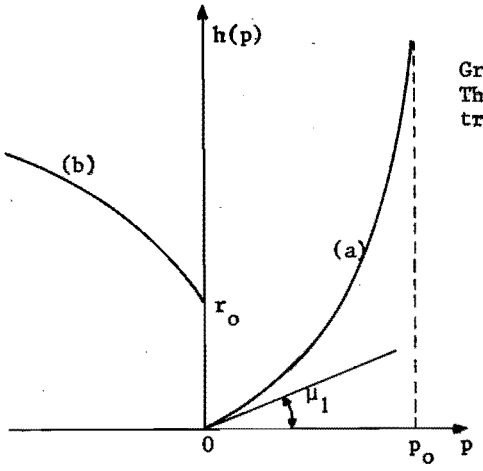

fig. 2.5.4.

Graph of the function h(p), $-\infty < p < p_o$. The two branches correspond to extremal elements of

(a) $\begin{cases} \inf \mathcal{P}_p, & 0 < p < p_o \\ \sup \mathcal{S}_r, & 0 < r < \infty \\ \inf \mathcal{K}_\mu, & \mu_1 < \mu < \infty \end{cases}$

(b) $\begin{cases} \inf \mathcal{P}_p, & -\infty < p < 0 \\ \inf \mathcal{Q}_r & r_o < r < \infty \end{cases}$

PART II: VARIATIONAL DYNAMICAL SYSTEMS

CHAPTER 3: CLASSICAL MECHANICS OF CONTINUOUS SYSTEMS

3.1. INTRODUCTION.

This chapter starts with a summary of some notions which will be used frequently in the rest of this thesis. In section 3.2. we define Lagrangian and Hamiltonian systems. The variational character of these definitions allows one to perform the "Legendre-transformation" in a more systematic way than is usually done. Moreover, this transformation then leads in a natural way to a so called modified action principle. This variational principle will greatly simplify the presentations of the results of section 3.5. In sections 3.3 and 3.4 we recall some terminology and results connected with canonical transformations and invariant integrals.

In section 3.5 we deal shortly with the problem how one can find variational principles describing the evolution of a continuum in the Eulerian setting. This has been a long standing difficulty until Lin [25] proposed a correct variational principle. However, this variational principle (as most others) was derived in an *ad hoc* way and, despite a lot of literature on this subject, it remained somewhat mysterious. We shall derive a variational principle for general evolutionary continua from first principles. This principle is then clearly understood and all known variational principles for fluid mechanics can be derived from it. We shall specialize this result to describe the motion of irrotational flow of a layer of incompressible fluid under the influence of gravity. The final result is a Hamiltonian system describing the surface waves. This description shall be the basis of the presentation in chapter 6.

3.2. LAGRANGIAN AND HAMILTONIAN SYSTEMS.

Let us start to recall some notions from Classical Mechanics for a system with a finite number (say n) of degrees of freedom. If Q denotes the configuration space of the dynamical system, then $q(t) \in Q$ represents one and only one state of the system and q is called (a set of) generalized coordinates. In the simplest case $Q = R\mathcal{l}^n$. The evolution of the system can be described as a trajectory in Q, i.e. a mapping from an interval of $R\mathcal{l}$ into Q : $t \mapsto q(t)$. If the system under consideration is a Lagrangian system, the Lagrangian $\mathcal{l}$ is a function on the tangent bundle of Q, i.e. $\mathcal{l}$: $(q,v) \mapsto \mathcal{l}(q,v) \in R\mathcal{l}$, where $q \in Q$ and v is an element from the tangent space to Q at q. If $Q = R\mathcal{l}^n$, the tangent space to Q is independent of the point $q \in Q$ and may be identified with $R\mathcal{l}^n$. If we denote this tangent space by V (the elements of which can be called generalized velocities), $\mathcal{l}$ is a function defined on Q x V. In many classical texts a typical element of V is denoted by $\dot{q}$, and one writes indifferently $\mathcal{l}(q,\dot{q})$ for the value of $\mathcal{l}$ at the arbitrary point $q \in Q$, $\dot{q} \in V$ and also for the value at $t \in R\mathcal{l}$ of the function $\mathcal{l}(q(t), \partial_t q(t))$ associated to a smooth trajectory $t \to q(t)$. To prevent this ambiguity we have introduced the velocity space V.

Many Lagrangian systems from Classical Mechanics can also be described with a Hamiltonian h which is then a function on the cotangent bundel of Q, i.e. h: Q x V^* → $R\mathcal{l}$. A typical element from V^* is usually denoted by p (the momentum variable): $h(q,p) \in R\mathcal{l}$. If $Q = R\mathcal{l}^n$, then V^* can again be identified with $R\mathcal{l}^n$. When both the Lagrangian and the Hamiltonian formulation are valid for a specific system, these formulations are related by a Legendre transformation. But the possibility to apply this transformation depends on convexity properties of $\mathcal{l}$ (or h), which properties have to be investigated in each case. Therefore, in the following we shall independently define the notion of a Lagrangian system and that of a Hamiltonian system and investigate thereafter the possible relationship.

For continuous systems the configuration manifold Q is infinitely dimensional and is some function space consisting of functions q defined on some region $\Omega \subset R\mathcal{l}^m$ in m-dimensional Euclidian space (m = 1 in the following except in section 3.5 where m = 3). The state of the system at time t will be denoted by $q(t) \in Q$ and its value at

$x \in \Omega$ by $q(x,t)$. We have used the term configuration *manifold* because even in the simplest examples Q is often not a linear space: it may be a set of functions satisfying the same prescribed value at the boundary $\partial\Omega$ of Ω. Then we suppose that

$$Q = \tilde{q} + Q_o, \tag{2.1}$$

where $\tilde{q}$ is some function defined on Ω satisfying the boundary conditions and Q_o is a linear space of functions with compact support in Ω. The velocity space V is also some function space of functions v defined on Ω. Concerning this velocity space we make the following remark. Let $q : t \to q(t)$ define a path on the configuration manifold. This mapping is differentiable (in the sense of Frechet, c.f. section 0.3) at $\bar{t}$ if there exists an element $\partial_t q(\bar{t}) \in Q$ such that

$$||q(\bar{t}+\varepsilon) - q(\bar{t}) - \varepsilon\cdot\partial_t q(\bar{t})||_Q = o(\varepsilon) \qquad \text{for } \varepsilon \to o$$

(the derivative of q at $\bar{t}$, i.e. the continuous linear mapping $q'(\bar{t}) : R\mathcal{l} \to Q$ has been identified with the element $\partial_t q(\bar{t}) \in Q : q'(\bar{t})\cdot\varepsilon = \varepsilon\cdot\partial_t q(\bar{t})$). When considering only such paths, the velocity space might be identified with Q_o. However, for many continuous systems the trajectories are not differentiable in this sense ($q'(\bar{t})$ is not a bounded mapping into Q, i.e. $\partial_t q(\bar{t}) \notin Q$). Nevertheless, in general Q is continuously embedded in a space V such that an element $\partial_t q(\bar{t}) \in V$ can be defined for which

$$||q(\bar{t}+\varepsilon) - q(\bar{t}) - \varepsilon\cdot\partial_t q(\bar{t})||_V = o(\varepsilon) \qquad \text{for } \varepsilon \to o.$$

Then, if $\partial_t q(\bar{t}) \in V\backslash Q$ this is a generalization of differentiability which must be allowed to make any progress.

LAGRANGIAN SYSTEMS.

DEFINITION 3.2.1. A *Lagrangian system* $(\mathcal{l},Q,V)$ is a dynamical system with *configuration manifold* Q, *velocity space* V, with Q continuously embedded in V, and *Lagrangian* $\mathcal{l} \in C^1(Q \times V, R\mathcal{l})$ such that the evolution of the system can be described with the following *action principle:* if $\hat{q} \in C^o(I,Q)$ with $\partial_t \hat{q} \in C^o(I,V)$ represents an actual evolution of the system over the time-interval $I = [t_o,t_1] \subset R\mathcal{l}$, then $\hat{q}$ is a stationary point of the *action functional*

$$A(q) = \int_I dt\ \mathcal{l}(q(t), \partial_t q(t)) \tag{2.2}$$

on the set

$$\{q \in C^0(I,Q) \mid \partial_t q \in C^0(I,V); q(t_0) = \hat{q}(t_0),\ q(t_1) = \hat{q}(t_1)\}. \tag{2.3}$$

From the action principle we can find the equations of motion for the dynamical system along familiar lines. Therefore we suppose that the configuration manifold Q can be written as in (2.1). The functional derivatives of $\mathcal{l}$ with respect to q and v at the point $(\hat{q}, \hat{v}) \in Q \times V$ are denoted by $\frac{\delta \mathcal{l}}{\delta q}(\hat{q},\hat{v})$ and $\frac{\delta \mathcal{l}}{\delta v}(\hat{q},\hat{v})$ respectively, and we have

$$\mathcal{l}(\hat{q}+\xi,\hat{v}+w)-\mathcal{l}(\hat{q},\hat{v}) = \int_\Omega dx[\tfrac{\delta \mathcal{l}}{\delta q}(\hat{q},\hat{v})\cdot\xi+\tfrac{\delta \mathcal{l}}{\delta v}(\hat{q},\hat{v})\cdot w]+o(||\xi||+||w||),$$

for $\xi \in Q_o$, $w \in V$, $||\xi||_\Omega \to 0$, $||w|| \to 0$.

Here $\frac{\delta \mathcal{l}}{\delta q}(\hat{q},\hat{v})$ must be interpreted as an element from Q_o^* (the dual space of Q_o with the L_2-innerproduct as duality map), and $\frac{\delta \mathcal{l}}{\delta v}(\hat{q},\hat{v}) \in V^*$.

The derivative of the action functional A at $\hat{q}$ with respect to an element ξ from the *space of admissible variations*

$$\{\xi \in C^0(I,Q_o) \mid \partial_t \xi \in C^0(I,V)\ ;\ \xi(t_0) = \xi(t_1) = 0\} \tag{2.4}$$

is easily seen to be

$$A'(\hat{q})\cdot\xi = \int_I dt \int_\Omega dx[\tfrac{\delta \mathcal{l}}{\delta q}(\hat{q}(t),\partial_t\hat{q}(t))\xi(t) + \tfrac{\delta \mathcal{l}}{\delta v}(\hat{q}(t),\partial_t\hat{q}(t))\partial_t\xi(t)],$$

which may be written after partial integration with respect to t of the last term (using $\xi(t_0) = \xi(t_1) = 0$ which causes the integrated term to vanish)

$$A'(\hat{q})\cdot\xi = \int_I dt \int_\Omega dx\{[\tfrac{\delta \mathcal{l}}{\delta q}(\hat{q}(t),\partial_t\hat{q}(t))-\partial_t\tfrac{\delta \mathcal{l}}{\delta v}(\hat{q}(t),\partial_t\hat{q}(t))]\cdot\xi(t)\}.$$

From the requirement that $A'(\hat{q})\cdot\xi = 0$ for every ξ from the set of admissible variations (2.4) we deduce from the action principle:

PROPOSITION 3.2.2. *If $\hat{q}(t)$ represents an actual evolution of a Lagrangian system $(\mathcal{l},Q,V)$, then $\hat{q}$ satisfies the following Euler-Lagrange equation:*

$$\partial_t\tfrac{\delta \mathcal{l}}{\delta v}(q(t),\partial_t q(t)) - \tfrac{\delta \mathcal{l}}{\delta q}(q(t),\partial_t q(t)) = 0. \tag{2.5}$$

HAMILTONIAN SYSTEMS.

DEFINITION 3.2.3. A *Hamiltonian system* (h,Q,P) is a dynamical system with configuration manifold Q, *momentum space* P, with Q continuously embedded in P^*, and *Hamiltonian* $h \in C^1(Q \times P, R\!l)$ such that the evolution of the system can be described with the following *canonical action principle*: if $\hat{q} \in C^0(I,Q)$ with $\partial_t \hat{q} \in C^0(I,P^*)$ represents an actual evolution of the system over the time-interval $I = [t_0,t_1] \subset R\!l$, then to $\hat{q}$ there corresponds an element $\hat{p} \in C^0(I,P)$ such that $(\hat{q},\hat{p})$ is a stationary point of the *canonical action functional*

$$(2.6) \qquad CA(q,p) = \int_I dt\ [\langle p, \partial_t q\rangle - h(q,p)]$$

on the set

$$\{(q,p) \in C^0(I,Q)\times C^0(I,P) \mid \partial_t q \in C^0(I,P^*); q(t_0)=\hat{q}(t_0), q(t_1)=\hat{q}(t_1)\}.$$

Denoting the functional derivatives of h with respect to q and p at the point $(\hat{q},\hat{p})$ by $\frac{\delta h}{\delta q}(\hat{q},\hat{p})$ $(\in Q_o^*)$ and $\frac{\delta h}{\delta p}(\hat{q},\hat{p})$ $(\in P^*)$ respectively, we obtain along familiar lines:

PROPOSITION 3.2.4 *If* $(\hat{q}(t),\hat{p}(t))$ *represents an actual evolution of a Hamiltonian system* (h,Q,P), *then* $(\hat{q},\hat{p})$ *satisfy the set of Hamilton equations*

$$(2.7) \qquad \partial_t q(t) = \frac{\delta h}{\delta p}(q(t),p(t)), \quad \partial_t p(t) = -\frac{\delta h}{\delta q}(q(t),p(t)).$$

Just as the stationary points of (2.2), i.e. the solutions q(t) of (2.5), define trajectories $t \mapsto q(t)$ in the configuration manifold Q, the stationary points of (2.6), i.e. the solutions q(t), p(t) of (2.7), define a mapping in the *phase space* $Q \times P : t \mapsto (q(t), p(t))$. Such a mapping completely characterizes the evolution of the system and is called a Hamilton flow for the Hamiltonian h or shortly *Hamilton flow* h. The variables $(q,p) \in Q \times P$ are said to be a pair of canonically conjugate variables and p is the momentum variable.

RELATION BETWEEN LAGRANGIAN AND HAMILTONIAN SYSTEMS.

We shall now show that a large class of Lagrangian systems are Hamiltonian systems (over the same configuration manifold Q) and conversely. This is done by constructing for given l the Hamiltonian h (and for given h the Lagrangian l) via a generalized Legendre-transformation and showing that the action principle goes over in the canonical action principle (and conversely). The generalized Legendre transform shall be described with the aid of polar functionals as introduced in section 0.6.

Starting with a Lagrangian system (l,Q,V) define the functional

$$l_q : V \to R l, \quad l_q(v) := l(q,v) \qquad \forall q \in Q,\ v \in V.$$

HYPOTHESIS 3.2.5. For each $q \in Q$ the functional l_q defined on the reflexive B-space V is *strictly convex*, i.e.

$$l_q(\lambda v+(1-\lambda)w) < \lambda \cdot l_q(v)+(1-\lambda)\cdot l_q(w) \qquad \forall \lambda \in (0,1),\ \forall v,w \in V,\ v \neq w,$$

and the functional derivative maps V *onto* V^*:

$$\mathcal{R}\left(\frac{\delta l_q}{\delta v}\right) = V^*.$$

Assuming this hypothesis to hold, the polar- and bipolar functional of l_q:

$$l_q^*(p) := \sup_{v \in V} [\langle p,v\rangle - l_q(v)] \qquad \text{for } p \in V^*,$$

$$l_q^{**}(v) := \sup_{p \in V^*} [\langle p,v\rangle - l_q^*(p)] \qquad \text{for } v \in V,$$

can be shown to have the following properties (c.f. section 0.6)

(i) for arbitrary $p \in V^*$ the functional $v \mapsto \langle p,v\rangle - l_q(v)$ has exactly one stationary point which is the solution of

$$p = \frac{\delta l_q}{\delta v}(v).$$

Therefore, l_q^* is defined on all of V^* and we may write

$$l_q^*(p) = \operatorname{stat}_{v \in V} [\langle p,v\rangle - l_q(v)]. \tag{2.8}$$

Moreover, $l_q^* : V^* \to R l$ is again a strictly convex functional and $l_q^* \in C^1(V^*, R l)$.

(ii) for arbitrary $v \in V$ the functional $p \mapsto \langle p,v\rangle - l_q^*(p)$ has exactly one stationary point which is the solution of

$$v = \frac{\delta l_q^*}{\delta p}(p).$$

Therefore we have

$$(2.9) \qquad l_q^{**}(v) = \operatorname{stat}_{p \in V^*} [\langle p,v\rangle - l_q^*(p)].$$

Moreover, l_q^{**} agrees with l_q on V:

$$(2.10) \qquad l_q^{**} = l_q \text{ on } V$$

From these observations it follows that if we define

$$(2.11) \qquad h : Q \times V^* \to R\mathcal{l},\ h(q,p) := l_q^*(p) \quad q \in Q,\ p \in V^*,$$

we have according to (2.8), (2.9) and (2.10) the relations

$$(2.12) \qquad h(q,p) = \operatorname{stat}_{v \in V} [\langle p,v\rangle - l(q,v)],$$

$$(2.13) \qquad l(q,v) = \operatorname{stat}_{p \in V^*} [\langle p,v\rangle - h(q,p)].$$

Substituting the characterization (2.13) into the action functional (2.2) we obtain a variational principle for the functional

$$CA(q,p) = \int_I dt\ [\langle p, \partial_t q\rangle - h(q,p)],$$

which leads to the canonical action principle for the Hamiltonian system (h,Q,V^*). Hence

THEOREM 3.2.6. *If the Lagrangian system* (l,Q,V) *satisfies hypothesis* 3.2.5 , *it is a Hamiltonian system* (h,Q,V^*) *where* $h : Q \times V^* \to R\mathcal{l}$ *is defined by* (2.11).
Consequently, the Euler-Lagrange equation (2.5) *and Hamilton equations* (2.7) *are equivalent.*

Starting with a Hamiltonian system (h,Q,P) define the functional

$$h_q : P \to R\mathcal{l} \qquad h_q(p) := h(q,p) \quad q \in Q,\ p \in P.$$

HYPOTHESIS 3.2.7. For each $q \in Q$ the functional h_q defined on the reflexive B-space P is strictly convex and has

$$\mathcal{R}\left(\frac{\delta h_q}{\delta p}\right) = P^*.$$

In the same way as above we have that $h_q^* \in C^1(Q \times P^*, R\mathcal{l})$ is strictly convex and that $h_q^{**} = h_q$ on P. Hence, defining

$$\mathcal{l} : Q \times P^* \to R\mathcal{l} \quad , \quad \mathcal{l}(q,v) := h_q^*(v) \quad q \in Q,\ v \in P^*, \tag{2.14}$$

we have once again the relations (2.12) and (2.13). Substituting the characterization (2.12) into the canonical action functional (2.6) we obtain a variational principle for the functional

$$MA(q,v,p) = \int_I dt\ [\langle p, \partial_t q - v\rangle + \mathcal{l}(q,v)]. \tag{2.15}$$

This functional will be called the *modified action functional*; taking stationary points of this functional on the set $\{(q,v,p) \in C^0(I,Q) \times C^0(I,P^*) \times C^0(I,P) \mid \partial_t q \in C^0(I,P^*); q(t_0) = \hat{q}(t_0), q(t_1) = \hat{q}(t_1)\}$ will be called the *modified action principle*, and the stationary points are easily seen to satisfy the equations

$$\begin{aligned} &\partial_t p - \frac{\delta \mathcal{l}}{\delta q}(q,v) = 0, \\ &\partial_t q - v = 0 \quad , \\ &p = \frac{\delta \mathcal{l}}{\delta v}(q,v) \quad . \end{aligned} \tag{2.16}$$

The modified action principle is clearly equivalent to the constrained variational principle

$$\operatorname*{stat}_{q,v} \int_I dt\ \mathcal{l}(q,v) \quad \text{subject to the constraint } \partial_t q = v, \tag{2.17}$$

(c.f. chapter 1 and recipe 2.1.1) and the variable p in (2.15), which equals the momentum canonically conjugate to q for stationary points, plays the rôle of a Lagrange multiplier. On the other hand, eliminating v explicitely in the variational principle (2.17) gives precisely the action principle for the action functional (2.2). Hence the modified action principle is equivalent to the action principle and we have obtained

THEOREM 3.2.8. *Suppose the Hamiltonian system* (h,Q,P) *satisfies hypothesis* 3.2.7. *Then it is a Lagrangian system* $(\mathcal{l},Q,P^*)$ *where* $\mathcal{l}: Q \times P^* \to R\mathcal{l}$

is defined by (2.13). *Consequently, the Euler–Lagrange equation* (2.5) *and Hamilton equations* (2.7) *are equivalent.*

3.3. CANONICAL TRANSFORMATIONS.

In this section we recall some results from the theory of canonical transformations. In chapter 5 we shall consider an important class of transformations on phase space which are not canonical.

Let Q x P be the phase space of some Hamiltonian system. We shall investigate differentiable transformations

$$(3.1) \qquad Q \times P \ni (q,p) \to (\bar{q},\bar{p}) \in \bar{Q} \times \bar{P}.$$

A transformation of this kind will be called *regular* if locally it is a one-to-one mapping (with the implicit function theorem this means that the first derivative of the mapping (3.1) must be boundedly invertible at every point).

DEFINITION 3.3.1. A regular transformation (3.1) is said to be a (time-independent) *canonical transformation* if there exists a functional $f \in C^1(Q \times \bar{Q}, R\!\!\!l)$ such that

$$(3.2) \qquad \langle p, \partial_t q\rangle = \langle \bar{p}, \partial_t \bar{q}\rangle + \partial_t f(q,\bar{q})$$

for arbitrary flow (q(t),p(t)) in phase space Q x P. The functional f is called the *generating functional* of the canonical transformation.

REMARK 3.3.2. It is easily seen that if for $f \in C^1(Q \times \bar{Q}, R\!\!\!l)$ the transformation given by

$$(3.3) \qquad p = \frac{\delta f}{\delta q}(q,\bar{q}), \quad \bar{p} = -\frac{\delta f}{\delta \bar{q}}(q,\bar{q})$$

defines a regular transformation, it is a canonical transformation which has f as generating functional.
This explains the name generating functional.

If (h,Q,P) is a Hamiltonian system, define the functional $\bar{h} : \bar{Q} \times \bar{P} \to R\!\!\!l$ under the transformation (3.1) by

(3.4) $$\bar{h}(\bar{q},\bar{p}) := h(q(\bar{q},\bar{p}),p(\bar{q},\bar{p}))$$

Then, if (3.2) is satisfied, we have

$$CA(q,p) = \overline{CA}(\bar{q},\bar{p}) + f(q(t_1),\bar{q}(t_1)) - f(q(t_o),\bar{q}(t_o)),$$

where

$$CA(q,p) = \int_{t_o}^{t_1} dt\ [<p,\partial_t q> - h(q,p)]$$

$$\overline{CA}(\bar{q},\bar{p}) = \int_{t_o}^{t_1} dt\ [<\bar{p},\partial_t \bar{q}> - \bar{h}(\bar{q},\bar{p})].$$

As in the canonical action principle variations of the canonical action functional have to be considered on the set of coordinate functions satisfying the same values at the end points of the considered time interval, we immediately obtain

PROPOSITION 3.3.3. *Under a canonical transformation* (3.1) *any Hamiltonian system* (h,Q,P) *transforms into a Hamiltonian system* $(\bar{h},\bar{Q},\bar{P})$ *where* $\bar{h}$ *is given by* (3.4). *Consequently, the class of Hamilton's equations is invariant under a canonical transformation.*

REMARK 3.3.4. As was shown in the foregoing section, if h satisfies some convexity conditions then (h,Q,P) is a Lagrangian system $(\mathcal{L},Q,P^*)$. Under a canonical transformation (h,Q,P) transforms into a Hamiltonian system $(\bar{h},\bar{Q},\bar{P})$. However, in general there is no evidence at all that $\bar{h}$ satisfies the convexity condition. Therefore, $(\bar{h},\bar{Q},\bar{P})$ needs not to be a Lagrangian system over $\bar{Q} \times \bar{P}^*$. As an example, let the generating functional be given by

$$f(q,\bar{q}) = <q,\bar{q}>.$$

This functional defines the simple canonical transformation

$$\bar{q} := p\ ,\ \bar{p} = -q \quad (\bar{Q} = P\ ,\ \bar{P} = Q).$$

Then the convexity of $\bar{h}(\bar{q},\cdot) : \bar{P} \to R\mathcal{L}$ for fixed $\bar{q} \in \bar{Q}$ depends on the convexity of $h(\cdot,p) : Q \to R\mathcal{L}$ for fixed $p \in P$.

3.4 CONSERVED DENSITIES AND INVARIANT INTEGRALS.

In this section we consider Hamiltonian systems (h,Q,P) where Q and P consist of functions defined over some region $\Omega \subset R\mathcal{L}^n$.

Before stating the actual definition and results of this section, we have to make some precautions because several expressions that follow will not be defined on (arbitrary flows in) QxP. These difficulties stem from the fact that for an arbitrary trajectory in QxP, say $I \ni t \mapsto (q(t),p(t)) \in QxP$, the expression $(\partial_t q(t),\partial_t p(t))$ needs not to be an element of QxP (compare with the remarks made in section 3.2). Therefore we define:

DEFINITON 3.4.1. A trajectory $I \ni t \mapsto (q(t),p(t)) \in QxP$ is called a *smooth trajectory* in QxP if $(\partial_t q(t),\partial_t p(t)) \in QxP$ for every $t \in I$.

As a consequence of this definition, for arbitrary functional $g \in C^1(QxP,R\!l)$, the expression

$$(4.1)\quad \partial_t g(q(t),p(t)) = \langle \frac{\delta g}{\delta q}(q(t),p(t)),\partial_t q(t)\rangle + \langle \frac{\delta g}{\delta p}(q(t),p(t)),\partial_t p(t)\rangle$$

is well defined only for smooth trajectories.
In general, solutions of Hamiltons equations

$$\partial_t q = \frac{\delta h}{\delta p}(q,p) \in P^*$$

$$\partial_t p = -\frac{\delta h}{\delta q}(q,p) \in Q^*$$

does not define a smooth trajectory in QxP (as $Q \subset P^*$ and thus $P \subset Q^*$, but generally $Q \neq P^*$). Therefore we say that the evolution of the system is smooth if it is described by a *smooth solution* of Hamiltons equations, i.e. by a solution which defines a smooth trajectory in QxP. In many practical situations for which there exist smooth evolutions, this may be explained as follows. For given hamiltonian h define the subset $\tilde{Q} \times \tilde{P} \subset Q \times P$ by

$$(4.2)\quad \tilde{Q} \times \tilde{P} := \{(q,p) \in Q \times P \mid \frac{\delta h}{\delta q}(q,p) \in P,\ \frac{\delta h}{\delta p}(q,p) \in Q\}.$$

Then it is easily seen that if there exists a solution of Hamiltons equations which is a trajectory in $\tilde{Q}x\tilde{P}$, this is a smooth trajectory in QxP. This result may also be stated in the following way.

PROPOSITION 3.4.2. *Any smooth evolution* $(\hat{q},\hat{p}) \in C^0(I,\tilde{Q}x\tilde{P})$ *of a Hamiltonian system* (h,Q,P) *is a stationary point of the canonical action functional* CA(q,p) *on the set*

$$(4.3)\quad \{(q,p) \in C^0(I,\tilde{Q} \times \tilde{P}) \mid \partial_t q \in C^0(I,P^*); q(t_o)=\hat{q}(t_o), q(t_1)=\hat{q}(t_1)\}.$$

A similar difficulty is encountered if one wants to define the *Poisson bracket* as

(4.4) $\{f,g\}(q,p) := \langle \frac{\delta f}{\delta q}(q,p), \frac{\delta g}{\delta p}(q,p) \rangle - \langle \frac{\delta f}{\delta p}(q,p), \frac{\delta g}{\delta q}(q,p) \rangle$.

For arbitrary functionals $f,g \in C^1(Q \times P, R\!\!\!l)$, the right hand side will not be defined on all of $Q \times P$. To circumvent this difficulty one may consider this expression only on a subset of $Q \times P$. Therefore, for given subset $\tilde{Q} \times \tilde{P} \subset Q \times P$ (for instance implicitely defined by (4.2)), put

(4.5) $\mathcal{F}^n(\tilde{Q} \times \tilde{P}) := \{f \in C^n(Q \times P, R\!\!\!l) \mid (\frac{\delta f}{\delta q}, \frac{\delta f}{\delta p})\Big|_{\tilde{Q} \times \tilde{P}} \subset P \times Q\}$.

Then, $\{,\}$ is neatly defined on elements of this class of functionals: the Poisson bracket

(4.6) $\{,\} : \mathcal{F}^1(\tilde{Q} \times \tilde{P}) \times \mathcal{F}^1(\bar{Q} \times \hat{P}) \to \mathcal{F}^0(\tilde{Q} \times \tilde{P})$

is an antisymmetric, bilinear mapping.
Moreover, it is well known that the Poissonbracket satisfies *Jacobi's relation:*

(4.7) $\{\{f,g\},k\} + \{\{k,f\},g\} + \{\{g,k\},f\} = 0 \quad \forall f,g,k \in \mathcal{F}^2(\tilde{Q} \times \tilde{P})$.

After these precautions we come to the main ideas of this section.

DEFINITION 3.4.3. An operator E defined on phase space $Q \times P$ is said to be a *conserved* density for the Hamilton flow h if a corresponding n-component *flux vector* $\underline{F}$ can be found such that

(4.8) $\partial_t E(q,p) + \operatorname{div} \underline{F}(q,p) = 0$

for every smooth solution of Hamilton's equations. An expression like (4.8) is called a *local conservation law*. Upon integrating (4.8) over the domain Ω (assuming the integrals to exist) there results a *global* conservation law:

(4.9) $\partial_t \int_\Omega E(q,p)\, dx + \int_\Omega \operatorname{div} \underline{F}\, dx = 0$

or

(4.10) $\partial_t \int_\Omega E(q,p)\, dx + \int_{\partial\Omega} \underline{F} \cdot \underline{n}\, dx = 0$

where $\partial\Omega$ denotes the boundary of the domain Ω and $\underline{n}$ the outward pointing unit normal at $\partial\Omega$. In particular, if $\Omega = R\!\!\!l^n$ and $\underline{F} \to 0$ for

$|x| \to \infty$ for every $(q,p) \in Q \times P$, we arrive at an invariant integral: a functional $e \in C^1(Q \times P, R\ell)$ is called an *invariant functional* for the Hamilton flow h if

$$\partial_t e(q,p) = 0 \tag{4.11}$$

for every smooth solution of Hamilton's equations.

For arbitrary functional $g \in C^1(Q \times P, R\ell)$ we have (4.1), which for smooth solutions of Hamiltons equations results into

$$\begin{aligned} \partial_t g(q,p) &= \langle \frac{\delta g}{\delta q}(q,p), \frac{\delta h}{\delta p}(q,p) \rangle - \langle \frac{\delta g}{\delta p}(q,p), \frac{\delta h}{\delta q}(q,p) \rangle \\ &= \{g,h\}(q,p), \end{aligned} \tag{4.12}$$

where we have used the definition of Poisson bracket (4.4). Consequently

PROPOSITION 3.4.4. *A functional $g \in C^1(Q \times P, R\ell)$ is an invariant fuctional for the Hamilton flow* h *if and only if*

$$\{g,h\} = 0 \quad \text{on } \tilde{Q} \times \tilde{P} \tag{4.13}$$

Moreover, g *is an invariant functional for the Hamilton flow* h *if and only if* h *is an invariant functional for the Hamilton flow* g.

With these results, it follows with Jacobi's relation (4.7) that if f and g are invariant integrals for a Hamilton flow h, then the same is true for their Poisson bracket $\{f,g\}$.

EXAMPLES 3.4.5.

(i) For a Hamiltonian system (h,Q,P), the functional h itself is an invariant integral. It is usually called the *energy* as in many systems from classical mechanics it can be interpreted as such.

(ii) Suppose $\Omega = R\ell$ and suppose that h does not depend explicitely on $x \in R\ell$. Then h is *translational invariant*, which means that for all $\varepsilon \in R\ell$

$$h(T_\varepsilon q, T_\varepsilon p) = h(q,p) \quad \forall q \in Q \quad \forall p \in P, \tag{4.14}$$

where the shift operator T_ε is defined by $T_\varepsilon u(x) := u(x + \varepsilon)$. Taking the limit for $\varepsilon \to 0$ in (4.14) formally gives

$$\langle \frac{\delta h}{\delta q}(q,p), \partial_x q \rangle + \langle \frac{\delta h}{\delta p}(q,p), \partial_x p \rangle = 0 \tag{4.15}$$

(Note that this expression is again only defined on a subset of $Q \times P$)

With this result it is not difficult to verify that the *momentum* functional

$$m(q,p) := \langle p, \partial_x q \rangle \tag{4.16}$$

is an invariant functional for any translationally invariant Hamilton flow h.

3.5. VARIATIONAL PRINCIPLES FOR FLUID DYNAMICS.

To demonstrate the foregoing theory we shall summarize in this section some results from a forthcoming paper [26]. The aim of that paper is to clarify and unify known variational principles (var.pr.'s) for continuous systems (in particular from fluid dynamics) described in the Eulerian setting. The main tools to that end are (i) the fact that in the Lagrangian setting a Lagrangian for such systems is usually relatively easy to find by a direct generalization of the theory for systems consisting of a finite number of degrees of freedom and (ii) the fact that the transition from the Lagrangian setting to an Eulerian description can be performed in the variational formulation. This last observation was already present in the work of Broer & Kobussen [27] (c.f. also Kobussen [28]) who described the transition as a canonical transformation. However, starting with a Lagrangian, it is somewhat easier to describe this transition as a transformation of the coordinate functions. Especially when the modified action functional is used the necessary work to be done is minimal and the resulting expression are very transparent. Starting with a Lagrangian for arbitrary conservative continua in the Lagrangian setting, the modified action principle in the Eulerian description then turns out to be the basic variational principle. This var.pr. is in fact closely related to a var.pr. proposed by Lin [25] (See also Serrin [29] who announced this result). Restricting to specific systems or inserting certain *a-priori* potential representations for the Eulerian velocity (such as Clebsch representation) leads to the var.pr.'s known in literature. Moreover, by investigating the canonical transformation more closely it becomes possible to describe free surface problems too. Because of its relevance for chapter 6 we shall show here how this procedure may lead to a var.pr. for surface waves over a layer

of incompressible fluid.

In the *Lagrangian setting* the evolution of a continuous system is described with the time t and the initial position of the continuum as independent variables. To describe the position of the system in space, we use a fixed Cartesian coordinate system with base vectors $\underline{c}_i$, i= 1,2,3, which coincide with the reciprocal (dual) base vectors $\underline{c}^i$, and consider vectors with respect to these bases

$$\underline{x} = x^i \underline{c}_i = x_i \underline{c}^i \quad , \quad x^i = x_i \qquad i = 1,2,3$$

(Here and in the following the summation convention is used.)
If the continuum occupies a region $\sum$ at t = 0, the position of the continuum at subsequent times can be described with a time-dependent operator (the evolution operator)

$$X(t) : \sum \to \Omega(t), \tag{5.1}$$

where $\Omega(t)$ is the region of space occupied by the continuum at time t. If $\underline{\xi}$ is a typical element from $\sum$, the effect $X(t)\cdot\underline{\xi}$ denotes the position at time t of the element which was originally at position $\underline{\xi}$ (the "particle" characterized by $\underline{\xi}$) and is usually written as

$$\underline{x}(\underline{\xi},t) = x^i(\underline{\xi},t)\underline{c}_i.$$

By definition X(0) = Id(identity map), thus $\underline{x}(\underline{\xi},0) = \underline{\xi}$ and $\Omega(0) = \sum$.

If the initial density is $\rho_o(\underline{\xi})$, assumed to be positive, a direct generalization of a system consisting of a finite number of particles leads one to consider the expression

$$\mathcal{L}(x^i,v^i) = \int_{\sum} d^3\xi \; \tfrac{1}{2} \rho_o v^i v_i - V(x^i) \tag{5.2}$$

as the Lagrangian of a wide class of continuous systems. Here V denotes the potential energy functional which needs not to be specified at this moment. Taking the action principle as described in section 3.2. to be valid for such continuous systems described in the Lagrangian setting, the equations of motion can be found from the action functional

$$A(x^i) = \int_I dt \; \mathcal{L}(x^i,\bar{\partial}_t x^i), \tag{5.3}$$

where $\bar{\partial}_t = \left.\frac{d}{dt}\right|_{\underline{\xi} \text{ fixed}}$, and $\bar{\partial}_t x^i$ are the components of the "particle"-

velocity. Instead of working with this formulation, we prefer to use the modified action functional, which shall be written as (for convenience we write the multiplier as $-\rho_o\lambda_i$):

$$(5.4) \qquad MA(x_i,v^i,\lambda_i) = \int_I dt \left[\int_\Sigma d^3\xi\rho_o(\underline{\xi})\lambda_i(v^i - \bar{\partial}_t x^i) + \mathcal{L}(x^i,v^i)\right].$$

With this formulation we have completely described the system in the Lagrangian setting.

In the *Eulerian setting* the evolution of the systems is described with the time t and the position of the continuum at time t as independent variables.
Assuming the existence of the inverse of the mapping X(t)

$$(5.5) \qquad X^{-1}(t) : \Omega(t) \to \Sigma$$

we write $\underline{\xi} = \underline{\xi}(\underline{x},t)$ for the element with position $\underline{x}$ at time t which had initial position $\underline{\xi}$. With this mapping it is posible to define with every field variable in the Lagrangian setting a field variable in the Eulerian setting and conversely. Denoting such corresponding variables with the same symbol gives no difficulties provided we distinguish clearly between differentiation with respect to components of $\underline{x}$ and with respect to components of $\underline{\xi}$. Therefore we denote the components of $\underline{\xi}$ by ξ^a, a = 1,2,3 and write

$$f_{,i} = \frac{\partial f}{\partial x^i} \text{ and } f_{,a} = \frac{\partial f}{\partial \xi^a} .$$

In particular, the Jacobian matrices of the transformation X(t) and its inverse $X^{-1}(t)$ have elements which are denoted by $x^i{}_{,a}$ and $\xi^a{}_{,i}$ respectively and as these matrices are each others inverse we have

$$(5.6) \qquad x^i{}_{,a}\cdot\xi^b{}_{,i} = \delta^b_a \quad , \quad x^i{}_{,a}\cdot\xi^a{}_{,j} = \delta^i_j$$

where δ^b_a and δ^i_j are Kronecker symbols. The *mass-density* $\rho(\underline{x},t)$, implicitely defined by $\rho(\underline{x},t)d^3x = \rho_o(\underline{\xi})d^3\xi$, can be written with the determinant of a Jacobian matrix as

$$(5.7) \qquad \rho(\underline{x},t) = \rho_o(\underline{\xi}) \cdot \det \ [\xi^a_{,i}].$$

In the Eulerian setting the time derivative is denoted by $\bar{\partial}_t : \bar{\partial}_t = \frac{d}{dt}\Big|_{\underline{x} \text{ fixed}}$. Differentiating the identity $x^i(\underline{\xi}(\underline{x},t),t) = x^i$ with respect to t gives

$$\bar{\partial}_t x^i + x^i_{,a}\, \partial_t \xi^a = 0 \tag{5.8}$$

[Defining basevectors $\underline{e}_a$ as the tangent vectors to the parameter-curves: $\underline{e}_a = x^i_{,a}\, \underline{c}_i$, (5.8) may be written as

$$\partial_t \underline{x} = -\partial_t \xi^a \underline{e}_a. \tag{5.9}$$

This expression shows that $-\partial_t \xi^a$ are the components with respect to $\underline{e}_a$ of the *Eulerian velocity* $\underline{v}(\underline{x},t)$:

$$\underline{v}(\underline{x},t) = \bar{\partial}_t\, \underline{x}(\underline{\xi},t)\big|_{\underline{\xi} = \underline{\xi}(\underline{x},t)}]. \tag{5.10}$$

To describe the evolution of the continuum in the Eulerian setting, we transform the modified action principle (5.4) into an equivalent var.pr. in the Eulerian description. Therefore we note that the Lagrangian as given by (5.2) can be considered as a functional in the Eulerian setting:

$$\mathcal{L}(x^i,v^i) = \int_{\mathcal{L}} d^3\xi \tfrac{1}{2}\rho_o v_i v^i - V(x^i) = \int_{\Omega(t)} d^3x \tfrac{1}{2}\rho v_i v^i - \bar{V}(\xi^a) =: \bar{\mathcal{L}}(\xi^a, v^i), \tag{5.11}$$

where $\bar{V}(\xi^a)$ denotes the transformed potential energy functional. Moreover, using (5.6) and (5.8) we also have

$$\begin{aligned} \int_{\mathcal{L}} d^3\xi \rho_o \lambda_i (v_i - \bar{\partial}_t x^i) &= \int_{\Omega(t)} d^3x\, \rho\lambda_i (v^i + x^i_{,a}\, \partial_t \xi^a) \\ &= \int_{\Omega(t)} d^3x\, \rho\lambda_j (\partial_t \xi^a + v^i \xi^a_{,i}) x^j_{,a}. \end{aligned} \tag{5.12}$$

Writing $\lambda_a = \lambda_j\, x^j_{,a}$ (λ_a are the components of the vector $\underline{\lambda} = \lambda_i \underline{c}^i$ with respect to the basevectors $\underline{e}^a$, where $\underline{e}^a$ is the base reciprocal (dual) to the system $\underline{e}_a$, thus $\underline{e}_a \cdot \underline{e}^b = \delta^b_a$), the modified action functional is seen to be equivalent to

$$(5.13)\quad \overline{MA}(\xi^a,v^i,\lambda_a) = \int_I dt\ [\int_{\Omega(t)} d^3x\ \rho\lambda_a(\partial_t\xi^a + v^i\xi^a_{,i}) + \bar{\mathcal{L}}(\xi^a,v^i)]$$

where

$$(5.14)\quad \bar{\mathcal{L}}(\xi^a,v^i) = \int_{\Omega(t)} dx\ \tfrac{1}{2}\rho v_i v^i - \bar{V}(\xi^a).$$

In this var.pr. the variables ξ^a, v^i and λ_a are considered as independent variables, and it is of the form of a modified action principle in the Eulerian setting. With respect to this fundamental var.pr. some remarks can be made.

REMARKS 3.5.1.

(i) The action functional corresponding to (5.13) reads

$$(5.15)\quad \bar{A}(\xi^a) = \int_I dt\ \bar{\mathcal{L}}(\xi^a, -x^i_{,a}\,\partial_t\xi^a)$$

This result is of course also obtained if the original action functional (5.3) is expressed in the Eulerian setting with the aid of (5.8).

(ii) By its nature, the modified action principle (5.13) expresses the fact that when looking for stationary points of

$$(5.16)\quad \int_I dt\ \bar{\mathcal{L}}(\xi^a,v^i),$$

the variables v^i have to satisfy the constraints

$$(5.17)\quad \partial_t\xi^a + v^i\xi^a_{,i} = 0,$$

or equivalently

$$v^i = -x^i_{,a}\partial_t\xi^a.$$

Hence these constraints explicitely express the fact that v^i has to be considered as the components of the Eulerian velocity in (5.16). This interpretation completely clarifies the constraints as proposed by Lin.

(iii) If we write

$$(5.18)\quad p_i := -\rho_o\lambda_i,\ \bar{p}_a := +\rho\lambda_a,$$

the above derived results show that

$$\int_{\Sigma} d^3\xi \; p_i \bar{\partial}_t x^i = \int_{\Omega(t)} d^3x \; \bar{p}_a \partial_t \xi^a. \tag{5.19}$$

From this it follows with the theory of section 3.3. that the transformation $(x^i, p_i) \to (\xi^a, \bar{p}_a)$ is a *canonical transformation* (with vanishing generating functional). This transformation has been described in a somewhat different way by Broer & Kobussen [27].

Resuming the foregoing results we can say that (5.13), (5.14) describe in the Eulerian setting the evolution of a continuous system for which the action principle is assumed to hold in the Lagrangian setting. By specifying the potential energy functional $\bar{V}(\xi)$ the system is completely defined. In fluid dynamics, for ideal fluids, $\bar{V}$ depends only through the mass density ρ on the variables ξ^a : $\bar{V} = \bar{V}(\rho)$ In the original paper [26] it is shown how several known var.pr.'s for fluid dynamics can be derived from (5.13), (5.14). At this place we shall restrict to one specific model which will be studied in chapter 6. Therefore, let us first write down the equations obtained from (5.13), (5.14) by varying the variables λ_a and v^i:

$$\delta\lambda_a : \partial_t \xi^a + v^i \xi^a_{,i} = 0 \tag{5.20}$$

$$\delta v^i : v_i + \lambda_a \xi^a_{,i} = 0 \tag{5.21}$$

As has already been remarked, equations (5.20) define v^i to be the Cartesian components of the Eulerian velocity, and (5.21) expresses the multipliers λ_a in the variables v^i.

In the following we shall only be interested in a subclass of all possible flows, viz. the class of *irrotational flows*. For irrotational flow rot $\underline{v} = 0$ and thus for some potential φ

$$\underline{v} = \nabla\varphi \qquad (\text{i.e. } v_i = \varphi_{,i}) \tag{5.22}$$

To obtain a var.pr. for such flows we specialize (5.13), (5.14) by inserting the representation (5.22) together with (anticipating on (5.21)):

(5.23) $$\lambda_a = -\varphi_{,i}\, x^i_{,a} = -\varphi_{,a}$$

and consider the variable φ instead of the variables v^i and λ_a as fieldvariables. Then we find

(5.24) $$\mathrm{IrF}\mathcal{l}(\xi^a,\varphi)=\int_I dt\Big[\int_{\Omega(t)} d^3x\{-\rho\varphi_{,a}\partial_t\xi^a-\rho\varphi_{,j}x^j_{,a}\varphi_{,k}\xi^a_{,i}\delta^{ik}+\tfrac{1}{2}\rho(\nabla\varphi)^2\}-\bar{V}(\rho)\Big]$$

To rewrite this result into a more manageable form note that

$$-\rho\varphi_{,j}x^j_{,a}\varphi_{,k}\xi^a_{,i}\delta^{ik} = -\rho\varphi_{,i}\varphi_{,j}\delta^{ij} = -\rho(\nabla\varphi)^2$$
$$-\rho\varphi_{,a}\partial_t\xi^a = \rho\bar{\partial}_t\varphi - \rho\partial_t\varphi$$

Moreover, as

$$\int_{\Omega(t)} d^3x\rho\bar{\partial}_t\varphi=\int_{\Sigma} d^3\xi\rho_o(\xi)\cdot\bar{\partial}_t\varphi=\frac{d}{dt}\int_{\Sigma} d^3\xi\rho_o\cdot\varphi \quad (=\frac{d}{dt}\int_{\Omega(t)} d^3x\ \rho\cdot\varphi)$$

(where we have used the fact that the domain Σ is independent of t), this term gives an uninteresting contribution at the endpoints of the considered t-interval and may be omitted. With these observations it is found that the resulting expressions for (5.24) does not depend explicitely on the variables ξ^a: they merely appear via the variable ρ which may therefore be considered as the field variable instead of the variables ξ^a. This leads us to the var.pr.

(5.25) $$\mathrm{IrF}\mathcal{l}(\rho,\varphi) = \int dt\ \Big[\int_{\Omega(t)} d^3x\ \{-\rho\partial_t\varphi - \tfrac{1}{2}\rho(\nabla\varphi)^2\} - \bar{V}(\rho)\Big],$$

which var.pr. (for fixed domain $\Omega(t)$) was already studied by Bateman [30] (c.f. also Bateman [31]). Of course, the var.pr (5.25) can also be found directly from the canonical action functional corresponding to (5.13) by substituting the representation (5.23)

The var.pr. (5.25) gives the correct set of equations describing the flow in the interior of the domain $\Omega(t)$. In many important situations, $\Omega(t)$ is not (completely) prescribed *a-priori*. If only part of the boundary $\partial\Omega(t)$ of $\Omega(t)$ is fixed, say

$$\partial\Omega_o(t) = \partial\Sigma_o \quad \text{for all } t,$$

a generalized form of the action principle in the Lagrangian setting may be formulated which does not restrict the variations $\delta\underline{x}(\underline{\xi},t)$ at the "free boundary" $\partial\Omega(t) - \partial\Sigma_o$. As a consequence of this generalization a complete description of the system, including equations which describe the evolution of the free boundary, is obtained if also variations of the free boundary are considered.

To demonstrate this idea let us describe irrotational flow of a layer of fluid over a horizontal bottom (y = 0) under the influence of gravity (pointing in the negative y-direction) and ignoring surface tension. Considering two-dimensional flow for simplicity, the free surface may be described by $y = 1 + \eta(x,t)$, where $\eta(x,t)$ denotes the height of the fluid at place x, $-\infty < x < \infty$, and time t, measured from the equilibrium configuration y = 1. In the equilibrium configuration $\rho = \rho_o$ say, and this situation is taken as the zero level of the potential energy. Then the potential energy functional is

$$\bar{V}(\rho) = \int_{-\infty}^{\infty} dx \int_{0}^{1+\eta(x,t)} dy \, \{\rho\, e(\rho) - \rho_o\, e(\rho_o) + \rho g(y-1)\},$$

wherein $e(\rho)$ is the specific potential energy density of the fluid. The var.pr. (5.25) becomes

$$\text{(5.26)}\quad \text{IrF}l(\rho,\varphi,\eta) = \int dt \int_{-\infty}^{\infty} dx \int_{0}^{1+\eta(x,t)} dy \, \{-\rho\partial_t\varphi - \rho e(\rho) + \rho_o e(\rho_o) - \rho g(y-1) - \tfrac{1}{2}\rho(\nabla\varphi)^2\},$$

wherein variations of η account for variations of the free boundary. The var.pr. (5.26) was found by Luke [32] and it gives the correct set of equations:

$$\begin{aligned}
&\delta\rho: && \partial_t\varphi + g(y-1) + \tfrac{1}{2}(\nabla\varphi)^2 + \frac{d}{d\rho}(\rho e(\rho)) = 0 && \text{for } 0 < y < 1+\eta(x,t)\\
&\delta\varphi: && \partial_t\rho + \operatorname{div}(\rho\nabla\varphi) = 0 && \text{for } 0 < y < 1+\eta(x,t)\\
& && \partial_t y + \eta_x\cdot\varphi_x - \varphi_y = 0 && \text{at } y = 1+\eta(x,t)\\
& && \varphi_y = 0 && \text{at } y = 0\\
&\delta y: && \rho\partial_t\varphi + \rho e(\rho) - \rho_o e(\rho_o) + \rho q(y-1) + \\
& && \quad + \tfrac{1}{2}\rho(\nabla\varphi)^2 = 0 && \text{at } y = 1+\eta(x,t).
\end{aligned}$$

These equations are immediately recognized as Bernoulli's equation in the interior, the continuity equation, the kinematical relation for the free surface, the condition that no fluid flows through the bottom and Bernoulli's equation at the free surface respectively.

To obtain the system which shall be considered in chapter 6 we further specialize to *incompressible, irrotational flow* with constant density $\rho = \rho_o$. Then the corresponding variational principle is found from (5.26):

$$\text{(5.27)}\quad \text{IncIrF}l(\eta,\varphi) = \int dt \int_{-\infty}^{\infty} dx(-\rho_o) \int_{0}^{1+\eta(x,t)} dy[\partial_t\varphi + g(y-1) + \tfrac{1}{2}(\nabla\varphi)^2].$$

By partial integration $\int_0^{1+\eta(x,t)} dy\ \partial_t\varphi = \partial_t \int_0^{1+\eta(x,t)} dy\ \varphi - \psi\cdot\partial_t\eta$,

where

$$\psi(x,t) := \varphi(x,y = 1+\eta(x,t),t) \tag{5.28}$$

denotes the value of the velocity potential φ at the free surface. Then, omitting an uninteresting term, (5.27) becomes

$$\text{IncIrFl}(\eta,\varphi)=\int dt \int_{-\infty}^{\infty} dx[\rho_o\psi\cdot\partial_t\eta - \rho_o\{\tfrac{1}{2}g\eta^2+ \int_0^{1+\eta(x,t)} dy\tfrac{1}{2}(\nabla\varphi)^2\}]. \tag{5.29}$$

It may now be observed that, for stationary points of (5.29), φ satisfies $\Delta\varphi = 0$ in the interior, such that φ as a solution of

$$\begin{aligned} \Delta\varphi &= 0 \quad \text{for } 0 < y < 1+\eta(x,t) \\ \varphi &= \psi \quad \text{at } y = 1+\eta(x,t) \\ \varphi_y &= 0 \quad \text{at } y = 0, \end{aligned} \tag{5.30}$$

is uniquely determined by $\psi(x,t)$ and $\eta(x,t)$. Then the kinetic energy term in (5.29) may be envisaged as an implicitely defined function of ψ and y:

$$\underline{k}(\psi,\eta)\ (x,t) := \int_0^{1+\eta(x,t)} dy\ \tfrac{1}{2}(\nabla\varphi)^2, \ \varphi \text{ is the solution of (5.30)} \tag{5.31}$$

Then (5.29) may be written as

$$\text{IncIrFl}(\eta,\psi) = \int dt \int_{-\infty}^{\infty} dx\ [\rho_o\psi\cdot\partial_t\eta - \rho_o\{\tfrac{1}{2}g\eta^2+\underline{k}(\psi,\eta)\}]. \tag{5.32}$$

It may be observed that (5.32) is in fact a canonical action functional for the system, with η and $\rho_o\psi$ as a pair of canonically conjugate variables and Hamiltonian

$$h(\eta,\psi) = \int_{-\infty}^{\infty} dx\ \rho_o\{\tfrac{1}{2}g\eta^2+\underline{k}(\psi,\eta)\}. \tag{5.33}$$

In fact this result has been found by Broer [33] (c.f. also Broer et al [34]). At the same time Benjamin [35] used the expression (5.33) and the variables ψ and η to derive the basic equations, without ex-

plicitely referring to the Hamiltonian character of the system. Both Broer and Benjamin found these results without using Luke's var.pr. (5.26) as a starting point. More recently, Miles [36] derived (5.32) from Luke's var.pr. in the same way as described above (c.f. also Milder [37]). As was shown by some of these authors, this var.pr. can be used advantageously to find satisfactory approximations for the awkward exact equations corresponding to (5.32), (5.31) by looking for suitable approximate expressions for the Hamiltonian h as given by (5.33). This will be more fully investigated in chapter 6.

CHAPTER 4: FIRST ORDER HAMILTONIAN SYSTEMS.

4.1. INTRODUCTION.

A Hamiltonian system as described in the foregoing chapter (which we shall sometimes call a *classical* Hamiltonian system to distinguish from first order Hamiltonian systems to be introduced below) is defined by two first order (in time) equations for the two canonically conjugate variables (Hamilton's equations). If desired it is sometimes possible to extract from these equations one equation of second order in time for one of the variables. However, in many applications equations are met which are of first order in time for one scalar variable, which equation describes a conservative system. By way of example we mention two equations which will play an important rôle in the rest of this thesis: the Korteweg-de Vries equation (c.f. [38], abbriviated KdV equation)

$$\partial_t u = -\partial_x (u+\tfrac{1}{2}u^2+u_{xx}) \tag{1.1}$$

and an equation proposed by Benjamin, Bona & Mahony (BBM equation, c.f. [39])

$$(1-\partial_x^2)\partial_t u = -\partial_x (u+\tfrac{1}{2}u^2). \tag{1.2}$$

Both equations were derived as approximate descriptions for the evolution of unidirectionally propagating surface waves on a layer of fluid under the influence of gravity, where u denotes the height of the waves measured from an equilibrium (c.f. chapter 6). These equations describe a conservative system in the sense that there exist functionals of u which are independent of time as u evolves according

to such an equation. More generally, every equation of the form

$$\partial_t u = A \frac{\delta h}{\delta u}(u), \tag{1.3}$$

in which A is some antisymmetric operator and h is some functional, has this property. Moreover, it turns out that such equations are closely related to a set of Hamilton equations for classical Hamiltonian systems: the flow in u-space described by (1.3) has the same structure as the flow in phase space of a classical Hamiltonian system. We shall not completely investigate this relationship but restrict to some formal observations in this direction.

Therefore let us start with the canonical action principle for a classical Hamiltonian system (h,Q,P)

$$CA(q,p) = \int_I dt\, [\langle p, \partial_t q\rangle - h(q,p)]. \tag{1.4}$$

Observing that

$$\langle p,\partial_t q\rangle = \tfrac{1}{2}\langle p,\partial_t q\rangle - \tfrac{1}{2}\langle q,\partial_t p\rangle + \tfrac{1}{2}\partial_t \langle p,q\rangle,$$

and writing the two variables q and p as one two component vector $\binom{q}{p}$ with the usual $R\ell^2$-structure, (1.4) may be written as

$$CA(q,p) = \int_I dt\, [\tfrac{1}{2}\langle \binom{q}{p}, \partial_t S\binom{q}{p}\rangle - h(q,p)] \tag{1.5}$$

where $S : R\ell^2 \to R\ell^2$ is given by $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and where we have omitted a term which reduces to a contribution at the endpoints of the time interval. The structure of the variational principle (1.5) in (q,p) space is the same as that of a variational principle in u space of the form

$$A(u) = \int_I dt\, [\tfrac{1}{2}\langle u, B\partial_t u\rangle - h(u)] \tag{1.6}$$

wherein B is some antisymmetric operator and h a functional. Stationary points of (1.6) satisfy

$$\partial_t Bu = \frac{\delta h}{\delta u}(u), \tag{1.7}$$

which is precisely (1.3) with $B = A^{-1}$.

[The related structure of (1.5) and (1.6) is of course also reflected in the equations for the stationary points: Hamilton's equations corresponding to (1.4) may be written as

$$S\,\partial_t \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \frac{\delta h}{\delta q}(q,p) \\ \frac{\delta h}{\delta p}(q,p) \end{pmatrix} \tag{1.8}$$

which has to be compared with (1.7)].

Another way to arrive at an equation of the form (1.3) is to consider the canonical action principle (1.4) not on all of the phase space, but to restrict the canonical action functional to manifolds at which q and p are linearly related. For instance, if we write

$$p = u \quad \& \quad q = \tfrac{1}{2}Bu$$

where B is some antisymmetric operator, (1.4) transformsinto a functional of the form (1.6). This method and the underlying idea will be investigated in detail in the next chapter.

It is also possible to start with an equation of the form (1.3) and then construct a set of canonically conjugate variables such that (1.3) is equivalent to a classical Hamiltonian system. For instance, if $A = -\partial_x$, equation (1.3) reads

$$u_t = -\partial_x \frac{\delta h}{\delta u}(u) \tag{1.9}$$

If we restrict to 2π-periodic solutions of (1.9) with zero mean value, u may be written as a Fourier series

$$u(x,t) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi}} \left(a_n(t)\cdot\sqrt{n}\cdot\sin nx + b_n(t)\cdot\sqrt{n}\cdot\cos nx\right), \tag{1.10}$$

and it is not difficult to verify that (1.9) transforms into

$$\begin{aligned} \partial_t a_n &= \frac{\partial \bar{h}}{\partial b_n}(a_i,b_i) \\ \partial_t b_n &= -\frac{\partial \bar{h}}{\partial a_n}(a_i,b_i), \end{aligned} \qquad n \geq 1 \tag{1.11}$$

where the function $\bar{h}(a_i,b_i)$ is defined under the transformation (1.10) by

$$\bar{h}(a_i,b_i) = h(u).$$

Hence, for the considered class of solutions, (1.9) is equivalent to

a discrete classical Hamiltonian system with an infinite number of degrees of freedom: a_n and b_n are canonically conjugate position and momentum variables and $\bar{h}$ is the Hamiltonian. This idea is due to Gardner [40]. Broer [41] considered the same problem, without restricting to periodic solutions: an arbitrary L_2-function $u(x,t)$ can be written as

$$u(x,t)= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{u}(k,t)e^{ikx}dk= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} [\hat{u}_+(k,t)\cos kx+\hat{u}_-(k,t)\sin kx]dk,$$

where $\hat{u}_+$ and $\hat{u}_-$ are the even and odd parts of the Fouriertransform $\hat{u}$ of u:

$$\hat{u}_+(k,t)=\tfrac{1}{2}[\hat{u}(k,t)+\hat{u}(-k,t)], \quad \hat{u}_-(k,t)=\tfrac{1}{2}[\hat{u}(k,t) - \hat{u}(-k,t)].$$

If we restrict ourselves to solutions of (1.9) for which

$$\int_{-\infty}^{\infty} u(x,t)dx = \hat{u}(0,t) = 0 \text{ for all } t \geq 0$$

(c.f. Remark 4.2.2) we may define functions $q(x,t)$ and $p(x,t)$ by

$$q(x,t) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \frac{1}{\sqrt{k}} \cdot \hat{u}_-(k,t) \sin kx \, dk;$$

(1.12)

$$p(x,t) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \frac{1}{\sqrt{k}} \cdot \hat{u}_+(k,t) \cos kx \, dk.$$

Defining $\bar{h}(q,p) = h(u)$ under the transformation (1.12), equation (1.9) can be shown to be equivalent to

$$\partial_t q = \frac{\delta \bar{h}}{\delta p} (q,p)$$

$$\partial_t p = -\frac{\delta \bar{h}}{\delta q} (q,p),$$

i.e. a classical Hamiltonian system with q and p as canonically conjugate variables and $\bar{h}$ as Hamiltonian.

So far about the formal relations between classical Hamiltonian systems and systems of the form (1.3). In the next section we shall define these first order evolution equations as first order Hamiltonian systems and briefly introduce the standard terminology and results for these equations. In section 4.3. we shall deal with a well known non-linear transformation, viz. Miura's transformation. It is shown how this transformation turns up as the simplest non-linear transformation which transforms a class of first order Hamiltonian systems into another class of such systems. Moreover, the construction of these classes will generate the infinite sequence of invariant functionals (of polynomial type) for the KdV equation (1.1).

4.2. DEFINITIONS AND GENERAL RESULTS.

We consider scalar functions defined on the whole real axis which are sufficiently smooth to allow the necessary differentiations and which tend to zero, together with the derivatives, sufficiently rapidly as $|x| \to \infty$. This space will be denoted by U.

A first order evolution equation on U is an equiation of the form

$$\partial_t u = K(u), \tag{2.1}$$

where K is some (non-linear) operator on U. With the usual nomenclature a particular solution of (2,1) is represented by a trajectory $t \to u(t) \in U$, and (2.1) defines a flow in U. We shall be interested in a restricted class of evolution equations, viz. those for which there exists a functional h such that

$$K(u) = -\partial_x \frac{\delta h}{\delta u}(u).$$

DEFINITION 4.2.1. If h is a functional defined on U which satisfies

$$h(0) = \frac{\delta h}{\delta u}(0) = 0 \tag{2.2}$$

the evolution equation

$$\partial_t u = -\partial_x \frac{\delta h}{\delta u}(u) \tag{2.3}$$

is called a *first order Hamiltonian system*, and the functional h is called the *Hamiltonian*.

REMARKS 4.2.2.

(i) If the functional h satisfies (2.2) then the functional $\bar{h}$ defined on U by

$$\bar{h}(u) := h(u) + \alpha \int_{R\ell} u\,dx + \beta. \tag{2.4}$$

where α and β are arbitrary constants, gives rise to the same evolution equation:

$$(2.5) \qquad \partial_t u = -\partial_x \frac{\delta h}{\delta u}(u) \iff \partial_t u = -\partial_x \frac{\delta \bar{h}}{\delta u}(u)$$

Therefore, requirement (2.2) may be envisaged as a normalization condition to assure that with every evolution equation of the form

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

wherein H is a potential operator (c.f. section 0.4), there corresponds a *unique* Hamiltonian h.

(ii) If equation (2.3) is a local equation, c.f. remark 4.2.5. and definition 5.6.1., we have for functions $u \in U$:

$$\frac{\delta h}{\delta u}(u) \to 0 \quad \text{for } |x| \to \infty.$$

In that case, integrating (2.3) over the whole real axis there results

$$(2.6) \qquad \partial_t \int_{R\ell} u(x,t)\, dx = 0.$$

In order to describe the variational principle for (2.3), we define the inverse of the operator ∂_x on U by

$$(2.7) \qquad \partial_x^{-1} u(x) := \int_{-\infty}^{x} u(\xi) d\xi.$$

The adjoint of this operator will be denoted by ∂_x^{-*}:

$$(2.8) \qquad \langle u, \partial_x^{-1} v\rangle = \langle \partial_x^{-*} u, v\rangle \qquad \forall u,v \in U.$$

Note that

$$(2.9) \qquad \partial_x^{-*} u(x) = \int_{x}^{\infty} u(\xi) d\xi = \int_{R\ell} u(\xi) d\xi - \partial_x^{-1} u(x).$$

Consequently

$$\partial_x \partial_x^{-*} = \partial_x^{-*} \partial_x = -I \quad \text{and} \quad \partial_x^{-*} u = -\partial_x^{-1} u \iff \int_{R\ell} u dx = 0.$$

PROPOSITION 4.2.3. *Let* $\hat{u} \in C^1(I,U)$ *be an actual evolution of the first order Hamiltonian system described by* (2.3) *over the time interval* $I = [t_o, t_1]$. *Then* $\hat{u}$ *is a stationary point of the functional*

$$(2.10) \qquad A(u) := \int_I dt\, [\tfrac{1}{2}\langle u, \partial_x^{-1} \partial_t u\rangle + h(u)]$$

over the set

(2.11) $$\{u \in C^1(I,U) \,|\, u(t_o) = \hat{u}(t_o);\ u(t_1) = \hat{u}(t_1)\}.$$

PROOF. Consider an arbitrary variation $v \in C^1(I,U)$, satisfying $v(t_o) = v(t_1) = 0$. Then

$$A'(u)\cdot v = \int_I dt\ [\tfrac{1}{2}\langle v, \partial_x^{-1}\partial_t u\rangle + \tfrac{1}{2}\langle u, \partial_x^{-1}\partial_t v\rangle + \langle \frac{\delta h}{\delta u}(u), v\rangle]$$

$$= \int_I dt\ [\tfrac{1}{2}\langle v, \partial_t(\partial_x^{-1} - \partial_x^{-*})u\rangle + \langle \frac{\delta h}{\delta u}(u), v\rangle + \tfrac{1}{2}\partial_t \langle u, \partial_x^{-1} v\rangle].$$

With $v(t_o) = v(t_1) = 0$ this results in the following equation for the stationary points of A:

(2.12) $$\tfrac{1}{2}(\partial_x^{-1} - \partial_x^{-*})\partial_t u = -\frac{\delta h}{\delta u}(u).$$

Applying the operator ∂_x to this expression, equation (2.3) is obtained. □

The following definition is not completely standard.

DEFINITON 4.2.4. A functional $e \in C^1(U,R\ell)$ is said to be an *invariant integral* for (2.3) if

(2.13) $$\partial_t\, e(u) = 0$$

for every solution of (2.3). Any operator E on U for which

$$e(u) = \int_{R\ell} E(u)\, dx$$

is an invariant integral, will be called a *conserved density* for (2.3).

REMARK 4.2.5. From this definition it follows that if E is a conserved density then there exists a *flux density* T such that

(2.14) $$\partial_t\, E(u) + \partial_x\, T(u) = 0$$
$$\& \quad T(u) \to 0 \text{ for } |x| \to \infty$$

for every solution of (2.3). The expression (2.14) is of the form of a *local conservation law*. However, only if T is a local operator

(c.f. definition 5.6.1.) we obtain the usual result that for arbitrary interval $(a,b) \subset R\!\!\!/$, $\partial_t \int_a^b E(u)\,dx$ depends only on the value of u and its derivatives with respect to x, at the points x = a and x = b. In general, we have for solutions of (2.3)

$$\partial_t E(u) = E'(u)\cdot\partial_t u = -E'(u)\cdot\partial_x \frac{\delta h}{\delta u}(u) = -\partial_x[\partial_x^{-1} E'(u)\cdot\partial_x \frac{\delta h}{\delta u}(u)],$$

from which it follows that E is a conserved density for (2.3), with flux density $T(u) = \partial_x^{-1} E'(u)\cdot\partial_x \frac{\delta h}{\delta u}(u)$, if and only if

$$\int_{R\!\!\!/} dx\, E'(u)\cdot\partial_x \frac{\delta h}{\delta u}(u) = 0$$

for arbitrary $u \in U$. (Compare this with (2.15) and proposition 4.2.6. below.)

For a functional $g \in C^1(U,R\!\!\!/)$ we have for solutions of (2.3)

$$\partial_t g(u) = \langle \frac{\delta g}{\delta u}(u),\ \partial_t u\rangle = \langle \frac{\delta g}{\delta u}(u),\ -\partial_x \frac{\delta h}{\delta u}\rangle. \tag{2.15}$$

Therefore we define the *Poissonbracket* $\{,\} : C^1(U,R\!\!\!/) \times C^1(U,R\!\!\!/) \to C^0(U,R\!\!\!/)$ by

$$\{f,g\}(u) = \langle \frac{\delta f}{\delta u}(u),\ -\partial_x \frac{\delta g}{\delta u}(u)\rangle. \tag{2.16}$$

(c.f. Broer [41]). It is easily seen that $\{,\}$ is a bilinear and antisymmetric mapping, and it can be shown that it satisfies the *Jacobi relation* :

$$\{\{f,g\},k\} + \{\{k,f\},g\} + \{\{g,k\},f\} = 0. \tag{2.17}$$

From (2.15) and (2.16) we obtain

PROPOSITION 4.2.6. *A functional* $g \in C^1(U,R\!\!\!/)$ *is an invariant integral for* (2.3) *if and only if*

$$\{g,h\} = 0 \tag{2.18}$$

EXAMPLES 4.2.7.

(i) As $\{h,h\} = 0$, the Hamiltonian h itself is an invariant integral

for (2.3). In many cases this functional can be interpreted as the total *energy* of the system under consideration.
(ii) If h is *translation invariant*, then $\frac{1}{2}\langle u,u\rangle$ is an invariant integral for (2.3), the *momentum functional*. This follows simply from

$$\partial_t \tfrac{1}{2}\langle u,u\rangle = \langle u,\partial_t u\rangle = \langle u,-\partial_x \frac{\delta h}{\delta u}(u)\rangle = \langle u_x, \frac{\delta h}{\delta u}(u)\rangle = 0.$$

(iii) As has already been observed, remark 4.2.2. (ii), if (2.3) is a local equation then $\int_{R\ell} u\,dx$ is an invariant integral. We shall sometimes call this linear functional the *mass-functional*.

The KdV equation (1.1) is an important example of a first order Hamiltonian system:

$$(2.19)\qquad \partial_t u = -\partial_x \frac{\delta h}{\delta u}(u) \text{ with } h(u) = \int_{R\ell} dx\,[\tfrac{1}{2}u^2 - \tfrac{1}{2}u_x^2 + \tfrac{1}{6}u^3].$$

The BBM equation (1.2) is not of the form (2.3), but it can be brought to this form by a simple linear transformation. More generally, we shall consider equations of the form

$$(2.20)\qquad \partial_t Dv = -\partial_x \frac{\delta k}{\delta v}(v),$$

wherein D is some selfadjoint operator on U commuting with ∂_t and ∂_x. An equation of this type may be obtained from a variational principle: stationary points of

$$(2.21)\qquad \bar{A}(v) = \int_I dt\,[\tfrac{1}{2}\langle v,\partial_t D\partial_x^{-1} v\rangle + k(v)]$$

on the set (2.11) satisfy equation (2.20).

PROPOSITION 4.2.8.
(i) *Let* L *be a regular operator on* U, *commuting with* ∂_x *(and* ∂_t*). Under the transformation* $v := L^{-1}u$, *equation* (2.3) *transforms into* (2.20) *with*

$$(2.22)\qquad D = L^* L$$

and

$$(2.23)\qquad k(v) := h(Lv).$$

(ii) *If* D *admit the representation* (2.22) *for some regular operator* L, *the transformation* u : = Lv *transforms equation* (2.20) *into* (2.3) *with* h(u) : = k(L^{-1}u).

PROOF. This proposition is an easy consequence of the fact that if k and h satisfy (2.23) then (c.f. Lemma 0.3.8.)

$$L^* \cdot \frac{\delta h}{\delta u}(L\,v) = \frac{\delta k}{\delta v}(v).$$

With this result the statements can be obtained either by manipulating with the equations, such as

$$L^* \cdot [\partial_t u + \partial_x \frac{\delta h}{\delta u}(u)] = L^* L \partial_t v + \partial_x \frac{\delta k}{\delta v}(v),$$

or by using the variational characterizations (2.10) and (2.21):

$$A(L\,v) = \bar{A}(v). \qquad \square$$

The BBM equation (1.2) is of the form (2.20)

$$(2.24) \qquad \partial_t Dv = -\partial_x \frac{\delta k}{\delta v} \text{ with } D = 1-\partial_x^2 \text{ and } k(v) = \int dx\, [\tfrac{1}{2}v^2 + \tfrac{1}{6}v^3].$$

As follows from proposition 4.2.8. it can be brought to the form (2.3) by the linear transformation u : = $D^{\frac{1}{2}}$v, where $D^{\frac{1}{2}}$ is the selfadjoint, positive definite square root of the operator D. Another possibility is to take D as in (2.22) with L = I - ∂_x (c.f. Broer [50]). However, in many situations it is simpler to deal directly with the form (2.20) than with the transformed equation (c.f. section 5.6).

4.3 NON LINEAR TRANSFORMATIONS BETWEEN FIRST ORDER HAMILTONIAN SYSTEMS.

In this section we shall study transformations which map first order Hamiltonian systems into other first order Hamiltonian systems. Reasoning along classical lines, it will follow that we are almost inevitable be led to the wellknown Miura-transformation and to the two classes of equations which are of "KdV-type" and of "modified KdV-type". Although the results of our investigation are known in literature, the way in which they are derived seems to be not completely standard.

The starting point is a first order Hamiltonian system on a function space V

$$(3.1) \qquad \partial_t v = -\partial_x \frac{\delta g}{\delta v}(v)$$

where g is some translation invariant functional. Now consider a (non-linear) transformation $\mathcal{F}$ which maps V into some other function space, U say:

$$\mathcal{F} : V \to U , \quad u = \mathcal{F}(v). \tag{3.2}$$

A natural question to consider is to find the evolution equation satisfied by u if v evolves according to (3.1). We shall be interested in a somewhat differentquestion: for which transformation $\mathcal{F}$ is the evolution of u described by a first order Hamiltonian system? [In section 3.3 we saw that for classical Hamiltonian systems a specific class of transformations on phase space, viz. canonical transformations, map *every* classical Hamiltonian system into another classical Hamiltonian system]. From proposition 4.2.8. it follows that every linear transformation

$$u = L\,v \quad \text{with } LL^* = L^*L = I$$

maps every first order Hamiltonian system into another one. However, for non linear transformations it may not be expected that every equation of the form (3.1) is mapped into another first order Hamiltonian system. Therefore, a more precise formulation of the question under consideration is:

Find a non-linear transformation $\mathcal{F}$ and the class of functionals g for which there exists a functional h such that if $\partial_t v = -\partial_x \frac{\delta g}{\delta v}(v)$ then $\partial_t u = -\partial_x \frac{\delta h}{\delta u}(u)$ for $u = \mathcal{F}(v)$.

[Note that we do not require the transformation $\mathcal{F}$ to be invertible].

If the evolution of u is described by

$$\partial_t u = -\partial_x \frac{\delta h}{\delta u}(u) \tag{3.3}$$

then, inserting $u = \mathcal{F}(v)$, we find

$$R(v) \cdot \partial_t v + \partial_x \frac{\delta h}{\delta u}(\mathcal{F}(v)) = 0, \tag{3.4}$$

where we have introduced the linear operator R(v) which is defined as the derivative of the operator $\mathcal{F}$:

$$R(v) := \mathcal{F}'(v) : V \to U.$$

Comparing (3.1) with (3.4) we see that in order that

(3.5) $$R(v)\cdot\partial_t v + \partial_x \frac{\delta h}{\delta u}(\Phi(v)) = R(v)\cdot\left[\partial_t v + \partial_x \frac{\delta g}{\delta v}(v)\right]$$

it is necessary that

(3.6) $$\partial_x \frac{\delta h}{\delta u}(\Phi(v)) = R(v)\cdot\partial_x \frac{\delta g}{\delta v}(v).$$

Now note that with a given functional $f : U \to R\!\!\!l$ and given transformation (3.2) there is associated in a natural way a functional $\bar{f} : V \to R\!\!\!l$ defined by

$$\bar{f}(v) := f(\Phi(v)) \quad \text{for } v \in V.$$

(The converse is in general not true if Φ is not invertible). Then we have

$$\frac{\delta \bar{f}}{\delta v}(v) = R^*(v)\cdot\frac{\delta f}{\delta u}(\Phi(v)).$$

in which $R^*(v)$ is the adjoint of the operator $R(v)$. Therefore it is natural to restrict the study of (3.6) to functionals $g : V \to R\!\!\!l$ which can be obtained from functionals defined on U. Hence, if it is assumed that

$$g(v) = \bar{\zeta}(v) = \zeta(\Phi(v)),$$

then (3.6) may be written as

(3.7) $$\partial_x \frac{\delta h}{\delta u}(u) = R(v)\cdot\partial_x\cdot R^*(v)\cdot\frac{\delta \zeta}{\delta u}(u) \quad \text{for } u = \Phi(v).$$

Resuming these considerations we may say: if a transformation Φ and functionals h and ζ on U can be found such that (3.7) holds, then we have

(3.8) $$\partial_t u + \partial_x \frac{\delta h}{\delta u}(u) = R(v)\cdot\left[\partial_t v + \partial_x \frac{\delta \bar{\zeta}}{\delta v}(v)\right] \quad \text{for } u = \Phi(v).$$

In other words: if (3.7) is satisfied, the transformation Φ maps every solution v of

(3.9) $$\partial_t v + \partial_x \frac{\delta \bar{\zeta}}{\delta v}(v) = 0$$

onto a solution u of (3.3)

Concerning the transformation Φ to be found, we can formulate

two *a-priori* requirements: (i) if $\int_{RZ} u\,dx$ is an invariant integral for (3.3), then $\int_{RZ} \mathfrak{F}(v)\,dx$ must be an invariant integral for (3.9); (ii) the operator

$$R(v)\cdot\partial_x\cdot R^*(v) \tag{3.10}$$

as it enters in (3.7) must be expressible in terms of $u = \mathfrak{F}(v)$. [As $\int_{RZ} u^2\,dx$ is an invariant integral for every equation of the form (3.3) (when h is translation invariant), $\int_{RZ} \mathfrak{F}^2(v)\,dx$ must also be an invariant integral for (3.9). However, this fact is not used in the construction of appropriate operators $\mathfrak{F}$ but will be a consequence of the construction below].

The only *a-priori* known common invariant integrals for a class of equations (3.9) are $\int_{RZ} v\,dx$ and $\int_{RZ} v^2\,dx$. Therefore, in order to satisfy requirement (i), it is tempting (although other choices cannot be excluded by this reasoning) to take for $\mathfrak{F}$:

$$\mathfrak{F}(v) = av^2 + bv + \partial_x N(v) \tag{3.11}$$

where a and b are constants and N a (possibly non-linear) operator. For simplicity let us first consider the more simple transformation

$$\mathfrak{F}(v) = v^2 + \partial_x N(v), \tag{3.12}$$

where N has to be determined to satisfy requirement (ii) above. For $\mathfrak{F}$ given by (3.12) we have

$$R(v) = 2v + \partial_x\cdot N'(v) \quad \text{and} \quad R^*(v) = 2v - N'(v)^*\cdot\partial_x,$$

such that the operator (3.10) becomes

$$\begin{aligned} R(v)\cdot\partial_x\cdot R^*(v)\cdot\varphi &= 2v\cdot\partial_x[R^*(v)\cdot\varphi] + \partial_x[N'(v)\cdot\partial_x R^*(v)\cdot\varphi] \\ &= \partial_x[2v\cdot R^*(v)\cdot\varphi + N'(v)\cdot\partial_x R^*\cdot\varphi] - (R^*(v)\cdot\varphi)\cdot 2v_x. \end{aligned}$$

Observing that the last term may be written as

$$\begin{aligned} -2v_x\cdot(2v\cdot\varphi - N'(v)^*\cdot\varphi_x) &= -2(v^2)_x\cdot\varphi + 2v_x\cdot N'(v)^*\cdot\varphi_x \\ &= \partial_x[2v_x\cdot N'(v)^*\cdot\varphi] - 2[v^2 + v_x\cdot N'(v)^*\cdot]_x\cdot\varphi \end{aligned}$$

we obtain

(3.13) $R(v)\cdot\partial_x\cdot R^*(v)\cdot\varphi = \partial_x[2v\cdot R^*(v)\cdot\varphi + N'(v)\cdot\partial_x R^*(v)\cdot\varphi + 2v_x\cdot N'(v)^*\cdot\varphi] -$
$- 2[v^2 + v_x\cdot N'(v)^*]_x\cdot\varphi,$

valid for arbitrary function φ. From (3.7) it follows that for $\varphi = \frac{\delta\xi}{\delta u}(u)$, the right hand side of (3.13) has to be expressible as a derivative with respect to x. Remembering that for every translation in variant function ζ: $\frac{\delta\zeta}{\delta u}(u)\cdot u_x = \partial_x\Phi(u)$ for some Φ, we simply require that $u = v^2 + v_x\cdot N'(v)^*$. Comparing this with (3.12) it follows that $N(v) = c.v$ for arbitrary constant c. Taking $c = 1$ for simplicity gives as transformation

(3.14) $$u = \$(v) = v^2 + v_x,$$

and with this transformation (3.13) can be expressed in terms of u:

(3.15) $$R(v)\cdot\partial_x\cdot R^*(v)\cdot\varphi = \partial_x[4u\cdot\varphi - \partial_x^2\varphi] - 2u_x\cdot\varphi =: T(u)\cdot\varphi$$

In this way we have found that the transformation (3.14) satisfies the requirements (i) and (ii) above.

REMARK 4.3.1. It may be thought that the requirements for the transformation admit a much more general transformation then (3.14). In fact, it is easily shown that

(3.16) $$\$(v) := av^2 + bv + cv_x$$

is a transformation which satisfies requirements (i) and (ii) above for arbitrary constants a, b and c (and for which the construction to be outlined below can be adapted). However, if we restrict ourselves to polynomial expressions for $\$(v)$, i.e. S is a polynomial in v and derivatives of v with respect to x, then (3.16) is the most general one. This can be seen from the following simple reasoning: if n is the degree of $\$$, then $R(v)\cdot\partial_x\cdot R^*(v)$ is of degree $2\cdot(n-1)$. Assuming T(u) to be a polynomial of degree m in u, $T(\$(v))$ is of degree $n\cdot m$ in v. Hence we must have $2\cdot(n-1) = n\cdot m$, which has as nontrivial solution only $n = 2$ and $m = 1$. This corresponds precisely to (3.16) (and (3.15)).

Remains to investigate for which functionals $\zeta : U \to R\ell$ there exist functionals $h : U \to R\ell$ such that

$$(3.17) \qquad \partial_x^{-1} \cdot T(u) \cdot \frac{\delta \zeta}{\delta u}(u) = \frac{\delta h}{\delta u}(u).$$

With some trial and error it was found that the following pairs of functionals satisfy (3.17):

$$(3.18) \qquad \zeta_0(u) = \tfrac{1}{2} \int u dx \qquad h_0(u) = \tfrac{1}{2} \int u^2 dx$$

$$(3.19) \qquad \begin{aligned} \zeta_1(u) &= \tfrac{1}{2} \int u^2 dx & h_1(u) &= \tfrac{1}{2} \int dx\, [2u^3 + u_x^2] \\ \zeta_2(u) &= h_1(u) & h_2(u) &= \tfrac{1}{2} \int dx\, [5u^4 + 10uu_x^2 + u_{xx}^2] \\ \zeta_3(u) &= h_2(u) & h_3(u) &= \tfrac{1}{2} \int dx [14u^5 + 70u^2u_x^2 + 14uu_{xx}^2 + u_{xxx}^2]. \end{aligned}$$

Hence we see that for $1 \leq n \leq 3$

$$(3.20) \qquad \partial_x^{-1}\, T(u) \cdot \frac{\delta h_{n-1}}{\delta u}(u) = \frac{\delta h_n}{\delta u}(u),$$

or, for $0 \leq n \leq 3$

$$(3.21) \qquad \tfrac{1}{2} \cdot (\partial_x^{-1}\, T(u))^{n+1} \cdot 1 = (\partial_x^{-1} T(u))^n \cdot u = \frac{\delta h_n}{\delta u}(u).$$

It turns out that this process can be continued, as follows from

LEMMA 4.3.2. *For arbitrary* $n \geq 1$ *the operator*

$$(3.22) \qquad (\partial_x^{-1}\, T(u))^n \cdot 1$$

is a potential operator.

PROOF. In principle, the lemma can be proved with the theory of section 0.4 by showing that the derivative of the operator (3.22) is a symmetric operator. However, the proof is rather difficult and lengthy and gives no deeper insight in the underlying problem. Therefore we prefer to refer to Flaschka & Newell [42, p. 411], whose results imply the foregoing lemma: the operator L_F^+ defined by their formula

(3.92) equals $\frac{1}{4}\partial_x^{-1}\, T(-q)$ and then the lemma follows from their formula (3.96). □.

Resuming the foregoing results we obtain

THEOREM 4.3.3. *Define a sequence of functionals* $h_n : U \to R\mathcal{I}$ *by*

$$h_{-1}(u) := \tfrac{1}{2}\int_{R\mathcal{I}} u\, dx$$

(3.23)

$$h_n(u) : h_n(0) = 0,\ \frac{\delta h_n}{\delta u}(u) = \tfrac{1}{2}(\partial_x^{-1}\, T(u))^{n+1}\cdot 1 \quad \textit{for } n \geq 0$$

Then we have for

$$g_n(v) := \bar{h}_{n-1}(v) = h_{n-1}(\mathcal{G}(v)) \qquad n \geq 0 \tag{3.24}$$

that

$$R(v)\cdot\partial_x \frac{\delta g_n}{\delta v}(v) = \partial_x \frac{\delta h_n}{\delta u}(u) \quad \textit{for } u=\mathcal{G}(v) = v^2+v_x. \tag{3.25}$$

Consequently

$$R(v)\cdot [\partial_t v+\partial_x \frac{\delta g_n}{\delta v}(v)] = \partial_t u+\partial_x \frac{\delta h_n}{\delta u}(u) \quad \textit{for } u=\mathcal{G}(v),\ n \geq 0. \tag{3.26}$$

In this way we have obtained two infinite sequences of first order Hamiltonian systems over the spaces U and V which are related in the sense of (3.26). Because of (3.25) it is also possible to take linear combinations of such Hamiltonians:

if $g(v) = \sum_n \alpha_n g_n(v)$ and $h(u) = \sum_n \alpha_n h_n(u)$

then

$$R(v)\cdot [\partial_t v+\partial_x \frac{\delta g}{\delta v}(v)] = \partial_t u + \partial_x \frac{\delta h}{\delta u}(u).$$

For n = 1, (3.26) amounts to

$$R(v)\cdot [\partial_t v + \partial_x (2v^3 - v_{xx})] = \partial_t u + \partial_x (3u^2 - u_{xx}). \tag{3.27}$$

As the right hand side equated to zero is related via a simple transformation to the standard form of the KdV equation (1.1) and the expression in square brackets corresponds to the so-called "modified KdV equation", (3.27) is exactly the result as was discovered by

Miura [43].

The next lemma shows that each functional h_m (g_m) is an invariant integral for every Hamilton flow h_n in U space (for every Hamilton flow g_n in V space respectively.)

LEMMA 4.3.4.

$$\{h_\ell, h_j\} = 0 \text{ for all } \ell, j \geq -1 \tag{3.28}$$

$$\{g_\ell, g_\ell\} = 0 \text{ for all } \ell, j \geq 0 \tag{3.29}$$

PROOF. The result (3.29) follows from (3.28): with (3.24) and (3.25) it follows that

$$\{g_j, g_\ell\}(v) = \left\langle \frac{\delta g_j}{\delta v}(v), \partial_x \frac{\delta g_\ell}{\delta v}(v) \right\rangle = \left\langle R^*(v) \cdot \frac{\delta h_{j-1}}{\delta u}(\sharp(v)), \partial_x \frac{\delta g_\ell}{\delta v}(v) \right\rangle$$

$$= \left\langle \frac{\delta h_{j-1}}{\delta u}(\sharp(v)), R(v) \cdot \partial_x \frac{\delta g_\ell}{\delta v}(v) \right\rangle = \left\langle \frac{\delta h_{j-1}}{\delta u}(u), \partial_x \frac{\delta h_\ell}{\delta u}(u) \right\rangle$$

$$= \{h_{j-1}, h_\ell\}(u).$$

To prove (3.28) we use the definition (3.23) of $h_\ell(u)$:

$$\{h_\ell, h_j\}(u) = \left\langle \frac{\delta h_\ell}{\delta u}(u), \partial_x \frac{\delta h_j}{\delta u}(u) \right\rangle = \tfrac{1}{4} \langle (\partial_x^{-1} T(u))^{\ell+1} \cdot 1, \partial_x \cdot (\partial_x^{-1} T(u))^{j+1} \cdot 1 \rangle$$

Observing that $T^*(u) = -T(u)$ and $\langle \partial_x^{-1}\psi, \varphi\rangle = -\langle \psi, \partial_x^{-1}\varphi\rangle$, provided $\int \psi dx = 0$ or $\int \varphi dx = 0$, we obtain

$$\{h_\ell, h_j\}(u) = \begin{cases} \tfrac{1}{4}\langle (\partial_x^{-1} T(u))^{j+m+1} \cdot 1, \partial_x \cdot (\partial_x^{-1} T(u))^{j+m+1} \cdot 1 \rangle = 0 \text{ if } \ell = j+2m \\ \tfrac{1}{4}\langle (\partial_x^{-1} T(u))^{j+m+1} \cdot 1, T(u) \cdot (\partial_x^{-1} T(u))^{j+m+1} \cdot 1 \rangle = 0 \text{ if } \ell = j+2m+1. \end{cases}$$

This proves the lemma. □

The next lemma relates subsequent functionals h_n:

LEMMA 4.3.5. *For* $n \geq 0$ *we have*

$$\int \frac{\delta h_{n+1}}{\delta u}(u) dx = 2 \cdot (2n+3) h_n(u). \tag{3.30}$$

PROOF: Let $m(u) := \int \frac{\delta h_{n+1}}{\delta u}(u)dx = \int (\partial_x^{-1}T(u))^{n+1}\cdot u\,dx$.

With lemma 4.3.2. it follows that

$$<\frac{\delta m}{\delta u}(u),v> = \int [(\partial_x^{-1}T(u))^{n+1}\cdot u]'\cdot v\,dx = <v,[(\partial_x^{-1}T(u))^{n+1}\cdot u]'\cdot 1>.$$

With $[\partial_x^{-1}T'(u)\cdot 1]\cdot\varphi = 4\varphi$ for arbitrary $\varphi \in U$, we obtain

$$\begin{aligned}<\frac{\delta m}{\delta u}(u),v> &= <v,\sum_{j=0}^{n}(\partial_x^{-1}T(u)^j\cdot(\partial_x^{-1}T'(u)\cdot 1)\cdot(\partial_x^{-1}T(u)^{n-j}\cdot u> \\ &\quad + <v,(\partial_x^{-1}T(u))^{n+1}\cdot 1> \\ &= (4\cdot(n+1)+2)\cdot <v,(\partial_x^{-1}T(u))^n\cdot u> \\ &= 2\cdot(2n+3)\,<v,\frac{\delta h_n}{\delta u}(u)>.\end{aligned}$$

Hence $\frac{\delta m}{\delta u} = 2\cdot(2n+3)\frac{\delta h_n}{\delta u}$, and then the result follows from $m(0) = h_n(0) = 0$ □

We shall end this section with some remarks.

REMARK 4.3.6. *Conserved flux property.*

For the equation

$$\partial_t u = -\partial_x \frac{\delta h_n}{\delta u}(u) \tag{3.31}$$

the density u is a conserved density (it is easily seen that $\frac{\delta h_n}{\delta u}$ is a local operator). From lemma 4.3.4. and 4.3.5. it follows that its flux $\frac{\delta h_n}{\delta u}$ is also a conserved density for (3.31). Then it is easily verified that the functional

$$C_n(u) := \int dx[x\cdot u - t\frac{\delta h_n}{\delta u}(u)] = \int dx(x\cdot u) - 2(2n+1)\cdot t\cdot h_{n-1}(u) \tag{3.32}$$

is an invariant integral for (3.31). (c.f. remark 5.4.2. (iv) for an interpretation). Taking the Poissonbracket with a functional 1 there results

$$\{C_n,\ell\}(u) = \int \frac{\delta \ell}{\delta u}(u)dx - 2(2n+1)\cdot t\cdot\{h_{n-1},\ell\}(u). \tag{3.33}$$

In particular, with (3.28) and (3.30):

$$\{C_n,h_m\}(u) = \int \frac{\delta h_m}{\delta u}(u)dx = 2\cdot(2m+1)h_{m-1}(u). \tag{3.34}$$

As

(3.35) $\{h_o,\ell\}(u) = -\langle u, \partial_x \frac{\delta \ell}{\delta u}(u)\rangle = \langle \frac{\delta \ell}{\delta u}(u), u_x\rangle$ (=0 if ℓ is translation invariant)

it follows from (3.33) for n=1:

(3.36) $\{C_1,\ell\}(u) = \int \frac{\delta \ell}{\delta u}(u)\,dx$ for every translation invariant functional ℓ.

Property (3.34) means that the operator $\{C_n,\cdot\}$, $n \geq 1$, maps each functional h_m from the infinite sequence onto its preceding one h_{m-1}. Using Jacobi's relation

$$\{h_m,\{C_n,\ell\}\} + \{\ell,\{h_m,C_n\}\} + \{C_n,\{\ell,h_m\}\} = 0$$

it follows from (3.34) that

(3.37) $\{h_m,\{C_n,\ell\}\} = 0$ if $\{\ell,h_{m-1}\} = \{\ell,h_m\} = 0$

This result seems to be interesting only for $m = n = 1$: then it follows with (3.35) that for every translation invariant functional ℓ:

(3.38) $\{h_1,\{C_1,\ell\}\} = 0$ if $\{\ell,h_1\} = 0$

This means that if ℓ is translation invariant and an invariant integral for the KdV equation (equation (3.31) with n = 1), then the same is true for $\int \frac{\delta \ell}{\delta u}(u)\,dx$.

REMARK 4.3.7. *Relation with inverse scattering theory.*
Although the transformation (3.14) and the result (3.27) discovered by Miura is interesting in its own right, it would not have had so much attention if it didn't form the basis of a remarkable theory developed by Gardner & Green & Kruskal & Miura [44]. They showed how it is possible to solve the initial value problem for the KdV equation by *linear* operations only. Although this is not the place to give a detailed account of this "inverse scattering theory", we shall indicate the most important ideas to show why the result (3.26) is essential for this method to be applicable. We start with the more general transformation

(3.39) $$u = \phi_\lambda(v) = v^2 + v_x + \lambda$$

where $\lambda \in R\mathcal{Z}$. Then $R(v) = \phi'_\lambda(v)$ is independent of λ and the same sequence of functionals $h_m : U \to R\mathcal{Z}$ is obtained as before and

(3.40) $$\partial_t u + \partial_x \frac{\delta h_n}{\delta u}(u) = R(v)\cdot[\partial_t v + \partial_x \frac{\delta}{\delta v}[h_{n-1}(\phi_\lambda(v))]].$$

It is well known that when we define a new variable ψ by

$$v := \partial_x \log \psi,$$

the transformation (3.39) becomes a linear mapping between ψ and u:

(3.41) $$u = \frac{\psi_{xx}}{\psi} + \lambda$$

or

(3.42) $$\psi_{xx} + (\lambda - u)\psi = 0.$$

Then the right hand side of (3.40) may be expressed in ψ:

$$\partial_t u + \partial_x \frac{\delta h_m}{\delta u}(u) = R(v)\cdot[\partial_t v + \partial_x R^*(v)\cdot\frac{\delta h_{n-1}}{\delta u}(\phi_\lambda(v))]$$
$$= R(v)\cdot\partial_x \frac{1}{\psi}[\partial_t \psi + (2\psi_x - \psi\partial_x)\frac{\delta h_{n-1}}{\delta u}(u = \frac{\psi_{xx}}{\psi} + \lambda)]$$

Hence: if ψ evolves according to an equation of the form

(3.43) $$\partial_t \psi + (2\psi_x - \psi\partial_x)\cdot\frac{\delta h_{n-1}}{\delta u}(u = \frac{\psi_{xx}}{\psi} + \lambda) = C(\lambda)\cdot\psi$$

for arbitrary constant $C(\lambda)$, then under the transformation (3.41), u evolves according to

(3.44) $$\partial_t u + \partial_x \frac{\delta h_n}{\delta u}(u) = 0.$$

If we tried to use the transformation (3.41) to solve the initial value problem for (3.44) we could reason as follows. Let u_o be the initial value for (3.44). Then take a bounded function ψ_o such that u_o and ψ_o satisfy (3.41) (From scattering theory it is known that such a ψ_o can only be found if λ belongs to the spectrum of the scattering problem (3.42) with u_o as potential. Note that, in this reasoning, λ is taken fixed). Then, let ψ evolve according to an equation of the form (3.43). If $\psi(t)$ has been found, u(t) is immediately found

from (3.41), and the solution of (3.44) with u_o as initial value would have been found. Of course, this method is unsuccessful because the evolution of ψ is described by a non-linear equation and is usually as difficult to solve as the original problem (3.44). For instance, for n = 1, (3.44) is the KdV equation and (3.43) reads

$$\partial_t\psi + (2\lambda + 3\frac{\psi_{xx}}{\psi})\cdot\psi_x - \psi_{xxx} = C(\lambda)\psi. \tag{3.45}$$

The much more successful method of G.G.K.M. prevents (when it is applicable) the necessity to determine explicitely the solution of equations like (3.45), at the cost of taken an infinite (continuum) number of transformations (λ becomes a parameter ranging over the spectrum of the initial value scattering problem).
If u_o is the initial value for (3.44), consider the scattering problem (3.42) with $u = u_o$ and determine a suitable set of scattering data (viz. the spectrum, the reflectioncoefficient for the generalized eigenfunctions and normalization coefficients for the eigenfunctions corresponding to the discrete part of the spectrum). From the inverse scattering theory it is known that for such a set of data, the process can be inverted: once these data are known, the potential of the problem can be determined. Therefore, if from the evolution equation for u, evolution equations for these scattering data can be found for which the initial value problem can be solved, it is possible to determine the value of these data for every $t > 0$, and then, with the inverse scattering theory, the value of the potential u(t). Now, the main observation is that the equations for the scattering data are simple ordinary differential equations which can easily be solved, if the equations for the eigenfunctions are local equations: in that case the evolution equations for the scattering data can be obtained merely from the knowledge of the asymptotic behaviour (e.g. for $x\to\infty$) of the eigenfunctions. Moreover, using the fact that $u\to 0$ for $|x|\to\infty$, this asymptotic behaviour is described by a linear equation. For instance, (3.45) can be written as

$$\partial_t\psi + (3u - \lambda)\,\psi_x - \psi_{xxx} = C(\lambda)\cdot\psi. \tag{3.46}$$

This short description may indicate why precisely equations (3.44), which led to the local equations (3.43) for ψ, can be solved with the inverse scattering theory. For more details we refer to the literature, e.g. Whitham [45] and in particular Flaschka & Newell, [42, section 3] and G.G.K.M. [44].

CHAPTER 5: WAVE PROPAGATION IN ONE-DIMENSIONAL HAMILTONIAN SYSTEMS

5.1. INTRODUCTION.

Let us start with a Hamiltonian system (h,Q,P) where Q and P are function spaces consisting of functions defined on the whole real line $R\ell$ (one space variable $x \in R\ell$). Hamilton's equations are

$$\partial_t q = \frac{\delta h}{\delta p}(q,p) \ , \quad \partial_t p = -\frac{\delta h}{\delta q}(q,p). \tag{1.1}$$

For a large class of Hamiltoniansh these equationscan often be interpreted as a set of "wave equations". Although it is very difficult to give a precise definition of this notion, an equation will be called a wave equation if its solutions (or a subset of solutions) can be interpreted as waves, i.e. if these solutions describe some propagating phenomenon. Standard phrases in connection with wave equations are: normal mode solutions (for linear equations), propagation along characteristic curves in x,t plane (for non-linear, non-dispersive equations) and steady state solutions (periodic or solitary wave solutions for non-linear, dispersive equations).

Before analyzing this point any further, we want to rewrite equations (1.1) in another form. Therefore we assume that we are dealing with a *potential system*, which means that h does not depend on the variable q itself but only on expressions of $\partial_x q$. In that case, q can be interpreted as a *potential* for the underlying system, and the system can be more simply described with the variables $u = -\partial_x q$ and p. (Note that the transformation $(q,p) \to (u = -\partial_x q, p)$

is *not* a canonical transformation). If we rewrite the Hamiltonian $h(q,p)$ in terms of u and p, and denote this transformed Hamiltonian by $\bar{h}(u,p)$, Hamilton's equations (1.1) may be rewritten as

$$(1.2) \qquad \partial_t u = -\partial_x \frac{\delta \bar{h}}{\delta p}(u,p) \ , \quad \partial_t p = -\partial_x \frac{\delta \bar{h}}{\delta u}(u,p)$$

where we have used the relation $\frac{\delta h}{\delta q}(q,p) = \partial_x \frac{\delta \bar{h}}{\delta u}(u = -\partial_x q, p)$. Unlike the variable q in a potential system, the variables u and p are required to vanish (together with their derivatives) as $|x| \to \infty$. The equations (1.2) shall be written in a more convenient way as

$$(1.3) \qquad \partial_t \operatorname{grad} \bar{m}(u,p) = -\partial_x \operatorname{grad} \bar{h}(u,p).$$

Here $\bar{m}(u,p)$ is the momentum functional

$$(1.4) \qquad \bar{m}(u,p) = \langle u,p \rangle \ ,$$

and for arbitrary differentiable functional $f(u,p)$, grad $f(u,p)$ denotes the two-component functional derivative of f with respect to u and p :

$$\operatorname{grad} f(u,p) = \begin{pmatrix} \frac{\delta f}{\delta u}(u,p) \\ \frac{\delta f}{\delta p}(u,p) \end{pmatrix} \ .$$

With the respect to the formulation (1.3) we shall make several remarks.

REMARKS 5.1.1. (i) Of course the foregoing can also be rephrased in a variational form: Introducing $u = -q_x$ and $\bar{h}(u,p)$, the canonical action principle for (1.1) is equivalent to

$$(1.5) \qquad \overline{CA}(u,p) := -\int_I dt \, [\langle p, \partial_t \partial_x^{-1} u \rangle + \bar{h}(u,p)].$$

Stationary points of this functional satisfy

$$\partial_x^{-1} \partial_t u + \frac{\delta \bar{h}}{\delta p} = 0 \quad , \quad \partial_x^{-*} \partial_t p - \frac{\delta \bar{h}}{\delta u} = 0.$$

Applying ∂_x to these equations, we obtain (1.2).

(ii) As we have seen in section 3.4 , the functional $\bar{h}$ entering in (1.3) is an invariant integral (the energy) and $\bar{m}$ is an invariant integral if $\bar{h}$ is translation invariant.

(iii) The first order Hamiltonian systems as described in section 4.2 can also be written in the form (1.3) : for instance, the equation

$$\partial_t D u = - \partial_x \frac{\delta e}{\delta u} \tag{1.6}$$

where $D = D^*$ may be written as

$$\partial_t \operatorname{grad} n(u) = -\partial_x \operatorname{grad} e(u) \tag{1.7}$$

where now

$$n(u) = \frac{1}{2} < u, Du > \tag{1.8}$$

and $\operatorname{grad} f(u) = \frac{\delta f}{\delta u}$ by definition. The functional e, the energy, is again an invariant integral for solutions of (1.7), and the momentum functional n is an invariant integral if e is translation invariant.

(iv) *Steady state solutions* of (1.3) are functions u and p which depend on x and t only through the variable x-ct where c is the constant propagation velocity. The equation for these solutions is

$$\operatorname{grad} \bar{h}(u,p) - c \operatorname{grad} \bar{m}(u,p) = 0. \tag{1.9}$$

This equation may be envisaged as the equation for the stationary points of the constrained variational problems

$$\bar{h}(u,p) = \text{Stat.} \quad \text{subject to } \bar{m}(u,p) = \text{constant}$$

or

$$\bar{m}(u,p) = \text{Stat.} \quad \text{subject to } \bar{h}(u,p) = \text{constant},$$

where c is a multiplier. This formulation may be particularly useful

when an extremum variational formulation is valid. In that case it follows e.g. that $\bar{h}$ is an extremum on the class of functions for which $\bar{m}$ has a prescribed value and, because of remark (ii) above, this property then holds for all times. Such a property may have important consequences concerning the stability of steady states as has been remarked by Benjamin [35] (cf. also Benjamin [46] and Bona [47] where this property is used to prove stability of the solitary wave solutions of the KdV and BBM equation).

(v) The *most important aspect* of equation (1.3) for the rest of this chapter is that equation (1.3) is invariant for a time-independent linear transformation of the variables u and p. To be more precise, consider a transformation $(u,p) \to (\alpha,\beta)$ of the form

$$\begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} A & -B \\ C & D \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad , \tag{1.10}$$

where A, B, C and D are pseudo differential operators with constant coefficients. The transformation is *regular* if AD+BC is invertible and under a regular transformation equation (1.3) is equivalent to

$$\partial_t \operatorname{grad} \tilde{\bar{m}}(\alpha,\beta) = -\partial_x \operatorname{grad} \tilde{\bar{h}}(\alpha,\beta) \tag{1.11}$$

where

$$\tilde{\bar{h}}(\alpha,\beta) = \bar{h}(u,p) \quad , \quad \tilde{\bar{m}}(\alpha,\beta) = \bar{m}(u,p) \text{ under (1.10)}, \tag{1.12}$$

and where now "grad" denotes the two component functional derivative with respect to α and β. The equivalence of the equations (1.3) and (1.11) follows immediately from the fact that the operators A, B, C and D commute with ∂_x and ∂_t, and from transformation properties of an arbitrary functional f : if $f(\alpha,\beta) = f(u,p)$ then

$$\begin{pmatrix} \frac{\delta f}{\delta \alpha} \\ \frac{\delta f}{\delta \beta} \end{pmatrix} = \begin{pmatrix} A^* & C^* \\ -B^* & D^* \end{pmatrix} \begin{pmatrix} \frac{\delta f}{\delta u} \\ \frac{\delta f}{\delta p} \end{pmatrix} \quad .$$

Of course, these results can also be obtained from the variational formulation (1.5): under a regular transformation (1.10), stationary

points of

$$CA(\alpha,\beta) \; : \; = \; \overline{CA}(u = A\alpha - B\beta, \quad p = C\alpha + D\beta)$$

are in a one-to-one correspondence with those of $\overline{CA}(u,p)$, and satisfy equation (1.11).

Returning to the wave character of systems described by equation (1.3), it is often found that such an equation admits solutions which can be interpreted as being composed of waves running to the right and of waves running to the left (i.e. in the direction of the positive and negative x-axis respectively). For instance, if from the set (1.3) a single equation, of second order in time, for u can be extracted (which is then actually the Euler equations for the variable q) which does *not* contain terms with $\partial_t u$, the equation is invariant for time inversion from which the absence of any preferred direction of propagation follows (if u(x,t) is a solution which can be interpreted as propagating to the right, u(x,-t) is a solution propagating to the left).

For such wave equations with solutions running in both directions it is tempting to write the solutions as superpositions of two unidirectional waves running in opposite directions and to find the equations describing each of these unidirectionally propagating waves. To that end we perform a linear transformation of the form (1.10) and take the transformation such that the transformed momentum functional $\tilde{\bar{m}}(\alpha,\beta)$ does not contain a product term with α and β. For instance, for $\bar{m}$ as given by (1.4) we perform a transformation

$$\begin{aligned} u &= D(\alpha - \beta) \\ p &= \alpha + \beta \end{aligned} \tag{1.13}$$

where $D = D^*$ and invertible, and find that

$$\tilde{\bar{m}}(\alpha,\beta) = \langle \alpha, D\alpha \rangle - \langle \beta, D\beta \rangle \; . \tag{1.14}$$

Under this transformation, $\tilde{\bar{h}}(\alpha,\beta)$ can be written as

(1.15) $$\tilde{h}(\alpha,\beta) = a(\alpha) + b(\beta) + i(\alpha,\beta),$$

where a and b are functionals depending only on α and β respectively and where the interaction functional $i(\alpha,\beta)$ consists of all those terms which contain both the α and the β variable:

$$i(\alpha,0) = 0 \quad , \quad i(0,\beta) = 0 \quad \forall_\alpha \forall_\beta .$$

With (1.14) and (1.15) the equations (1.11) may be written as

(1.16)
$$2\partial_t D\alpha = -\partial_x\left(\frac{\delta a}{\delta \alpha}(\alpha) + \frac{\delta i}{\delta \alpha}(\alpha,\beta)\right)$$
$$2\partial_t D\beta = \partial_x\left(\frac{\delta b}{\delta \beta}(\beta) + \frac{\delta i}{\delta \beta}(\alpha,\beta)\right).$$

For a large class of *linear* equations it shall be shown in the next section that it is possible to find a transformation such that $i(\alpha,\beta) = 0$. In that case the two equations (1.16) are uncoupled and the solutions of the original equation (1.3) can be written as a super position of solutions of

(1.17) $$2\partial_t D\alpha = -\partial_x \frac{\delta a}{\delta \alpha}(\alpha)$$

(1.18) $$2\partial_t D\beta = \partial_x \frac{\delta b}{\delta \beta}(\beta).$$

Each of these equations is a first order Hamiltonian system such that, if the α and β equation describe waves travelling to the right and to the left respectively, we have obtained an *exact separation* of the original Hamiltonian system (1.3) into two unidirectionally propagative first order Hamiltonian systems.

For *non linear* equations it is generally not possible to find a linear (!) transformation such that $i(\alpha,\beta)$ vanishes identically. Then $i(\alpha,\beta)$ can be interpreted as an *interaction functional* which couples the two uncoupled equations (1.17), (1.18) as in (1.16). As they stand equations (1.17) and (1.18) are two first order Hamiltonian systems, the α- and the β- "mode" respectively .
For a restricted set of solutions of (1.16) it may be possible that

the interaction between the α- and β- mode can be "neglected", such that, within some approximation, these solutions of the original Hamiltonian system can be written as a linear combination of solutions of two uncoupled first order Hamiltonian systems. For instance, consider the solution (α,β) of (1.16) corresponding to initial data (α_o, β_o) with $\beta_o = 0$. It is likely that this solution is, at least for sufficiently small times, in some sense "close" to the solution of (1.17) with initial value α_o and $\beta \equiv 0$. Moreover, it may be argued that the correspondence will be as good as possible if the interaction terms in (1.16) are as "small" as possible. This will be the case if the linear transformation is chosen to achieve an exact separation of the linearized equations (1.16). A rigorous formulation and justification of these heuristic observations seems to be impossible in this generality. For a specific system of Chapter 6 we shall be able to say somewhat more about this point. Another result which can be obtained from (1.16) is that if $\bar{h}$ is an even functional of one of its variables, it is possible to define one first order Hamiltonian system which describes a subset of solutions of the original system exactly. This first order system is then *not* translation invariant. This shall be examined for general systems in section 5.3 and for a specific system in section 6.3 Up to now we have repeatedly used expressions such as "unidirectional propagation" without specifying the meaning of this notion. It is well known that for linear systems such a notion is closely related to the concept of group velocity, but especially for non-linear equations, this point seems to have had not so much attention in literature. In an attempt to illuminate this subject somewhat, we give in section 5.4 a (physically acceptable) definition of unidirectional propagativity in conservative evolution equations. This definition and its consequences are then investigated for linear Hamiltonian systems in section 5.5, and for non-linear first order Hamiltonian systems in sections 5.6 and 5.7.

5.2. EXACT SEPARATION OF LINEAR SYSTEMS

In this section we consider linear equations of the form (1.3):

$$\partial_t \operatorname{grad} \bar{m}(u,p) = -\partial_x \operatorname{grad} \bar{h}(u,p). \tag{2.1}$$

More specifically we assume that the functionals $\bar{m}$ and $\bar{h}$ are given by

$$\bar{m}(u,p) = \langle u,p \rangle \tag{2.2}$$

$$\bar{h}(u,p) = \frac{1}{2}\langle u,Uu\rangle + \frac{1}{2}\langle p,Np\rangle + \langle u,Qp\rangle \ . \tag{2.3}$$

These functionals are considered on the set $\mathcal{A} \times \mathcal{A}$, wherein

$$\mathcal{A} := \{f \in L_2(R\ell) \mid \hat{f}(k) = 0 \quad \text{for } |k| \geq k_o\}, \tag{2.4}$$

with k_o some arbitrary but fixed positive number and $\hat{f}$ denoting the Fouriertransform of the function f. The operators U, N and Q are assumed to be translation invariant pseudo differential operators from $\mathcal{A}$ into $\mathcal{A}$ (this implies that the operators are bounded on $\mathcal{A}$), with U and N symmetric on $\mathcal{A}$: $U = U^*$ and $N = N^*$. Moreover, we shall assume that

$$U \textit{ and } N \textit{ are boundedly invertible on } \mathcal{A} \ . \tag{2.5}$$

The equations (2.1) can be written explicitely as

$$\begin{aligned} \partial_t u &= -\partial_x [Np + Q^* u] \\ \partial_t p &= -\partial_x [Uu + Qp] \ , \end{aligned} \tag{2.6}$$

and it is easily seen that a solution (u,p) belongs to $\mathcal{A} \times \mathcal{A}$ for every t if the initial data belongs to $\mathcal{A} \times \mathcal{A}$.
To look for a separation of these equations, we perform a regular transformation of the form (1.10) and require that mixed terms (containing both the α and the β variable) entering in the transformed functionals $\tilde{m}$ and $\tilde{h}$ vanish. This gives conditions for the operators A, B, C and D. If these conditions can be fulfilled we have obtained an exact separation of the system (2.1). Reasoning along these lines, the following theorem is a straightforward result. In the formulation of it we use the decomposition of the operator Q in its symmetric and anti-symmetric part:

$$Q = Q_s + Q_a \ , \ Q_s^* = Q_s \ , \ Q_a^* = -Q_a$$

THEOREM 5.2.1. *The linear system described by* (2.1), (2.2), (2.3) *can be exactly separated if the operator*

$$NU^{-1} + [Q_a U^{-1}]^2 \tag{2.7}$$

is positive definite. In that case we may define an operator S *by* $S = S_s + S_a$, *where the symmetric part* S_s *is defined to be the (bounded) positive definite square root of* (2.6):

$$S_s^{\,2} = NU^{-1} + [Q_a U^{-1}]^2 \tag{2.8}$$

and where the anti-symmetric part is given by

$$S_a = -Q_a.U^{-1} \ . \tag{2.9}$$

Then the operator S *is bounded and has bounded inverse*

$$S^{-1} = S^*(SS^*)^{-1} = S^*(S_s^{\,2} - S_a^{\,2})^{-1} = S^*UN^{-1} \ , \tag{2.10}$$

and the transformation

$$\mathcal{H} \times \mathcal{H} \ni (u,p) \to (\alpha,\beta) \in \mathcal{H} \times \mathcal{H} \tag{2.11}$$

is regular and well-defined by

$$\begin{cases} p = \alpha + \beta \\ u = S\alpha - S^*\beta \end{cases} \qquad \begin{cases} 2\alpha = S_s^{-1}(S^*p+u) \\ 2\beta = S_s^{-1}(Sp-u) \end{cases} \ , \tag{2.12}$$

and transforms the functionals $\bar{m}$ *and* $\bar{h}$ *into*

$$\tilde{m}(\alpha,\beta) = \langle S_s\alpha,\alpha\rangle - \langle S_s\beta,\beta\rangle$$

$$\tilde{h}(\alpha,\beta) = \langle\{US_s^{\,2} + S_sQ_s\}\alpha,\alpha\rangle + \langle\{US_s^{\,2} - S_sQ_s\}\beta,\beta\rangle,$$

such that the transformed equations are given by

$$\partial_t \alpha = -\partial_x (US_s + Q_s)\alpha$$

(2.13)

$$\partial_t \beta = \partial_x (US_s - Q_s)\beta \ .$$

We shall end this section on linear systems with some remarks.

REMARK 5.2.2. For simplicity of exposition, assume that $Q = 0$. Then $S_a = 0$ and

$$S_s^2 = S^2 = NU^{-1},$$

and the transformation (2.12) becomes

$$\text{(2.14)} \qquad \begin{cases} p = \alpha + \beta \\ u = S_\alpha - S_\beta \end{cases} \qquad \begin{cases} 2\alpha = p + S^{-1}u \\ 2\beta = p - S^{-1}u \ . \end{cases}$$

Using the identity

$$\text{(2.15)} \qquad u = \frac{1}{2} S(p + S^{-1}u) - \frac{1}{2} S(p - S^{-1}u),$$

we are able to comment on condition (2.5). Indeed, if condition (2.5) is dropped, it is easy to construct examples for which S^{-1} is a bounded operator on $\mathcal{A}$, but S is not bounded on $\mathcal{A}$.
[A well known example is the linear equation for an elastic bar, which has $N = I$ and $U = -\partial_x^2$. Then, using the symbols of these operators, we have

$$\hat{S}^2 = \hat{N}.\hat{U}^{-1} = \frac{1}{k^2} \geq \frac{1}{k_o^2} \text{ on } \mathcal{A},$$

from which it follows that S^2 is positive definite, and thus S^{-2} is bounded, but S^2 is not bounded. Consequently, α and β are elements of $\mathcal{A}$:

$$\alpha := p + S^{-1}u \in \mathcal{A}$$

$$\beta := p - S^{-1}u \in \mathcal{A}$$

but there is no need for $S\alpha$ and $S\beta$ to be elements of $\mathcal{A}$, although the difference $S\alpha - S\beta$ belongs to $\mathcal{A}$, according to (2.15)]. Although a further investigation of this matter might indicate that such a separation can be given a mathematically sound base, from a physical point of view a more satisfactory result is obtained if condition (2.5) is fulfilled.

REMARK 5.2.3. In some sense, the character of a linear equation is completely reflected in its *dispersion relation*. The dispersion relation for the set of equations (2.6) can be expressed with the symbols of the pseudo differential operators as

$$(\omega - k\,\hat{Q}_s)^2 = k^2.(\hat{N}.\hat{U}+\hat{Q}_a^{\,2}), \tag{2.16}$$

and we see that a *necessary* condition for an exact separation to be possible is that this dispersion relation admits two real-valued (for $k \in R\!\!I$) solutionbranches $\omega(k)$, which is true provided $\hat{N}\cdot\hat{U}+\hat{Q}_a^{\,2} \geq 0$ (compare this with the requirement for the operator (2.7)). Then

$$\omega_+ = k.\hat{Q}_s + k.(\hat{N}.\hat{U}+\hat{Q}_a^{\,2})^{\frac{1}{2}} = k.(\hat{U}.\hat{S}_s+\hat{Q}_s)$$

is the dispersion relation for the α-mode of (2.13) and

$$\omega_- = k.\hat{Q}_s - k.(\hat{N}.\hat{U}+\hat{Q}_a)^{\frac{1}{2}} = -k.(\hat{U}.\hat{S}_s-\hat{Q}_s)$$

is the dispersion relations for the β-mode of (2.13).

5.3. REDUCTION FROM A CLASSICAL TO A FIRST ORDER HAMILTONIAN SYSTEM THROUGH SYMMETRY.

In this section we consider a set of Hamilton's equations in the variables $u = -\partial_x q$ and p, such that (c.f. (1.4))

$$\begin{aligned} \partial_t u &= -\partial_x \frac{\delta \bar{h}}{\delta p}(u,p) \\ \partial_t p &= -\partial_x \frac{\delta \bar{h}}{\delta u}(u,p)\,. \end{aligned} \tag{3.1}$$

We shall show that if $\bar{h}$ satisfies some assumptions, the solutions (u,p) of (3.1) corresponding to a restricted class of initial data can be obtained from *one* first order Hamiltonian system (without any approximation). For the surface waves to be studied in chapter 6 the assumptions are satisfied and the restricted class of initial data has a clear physical meaning.

To derive the results, it is somewhat simpler first to perform a linear transformation such as (1.13):

$$u = D(\alpha-\beta) \quad , \quad p = \alpha+\beta \tag{3.2}$$

where D is any selfadjoint , invertible operator (cf. remark 5.3.5, where the results are derived without such a transformation). Then, with $\tilde{h}(\alpha,\beta) = \bar{h}(u,p)$, the equations for α and β are (cf. (1.16)

$$2\partial_t D\alpha = -\partial_x \frac{\delta\tilde{h}}{\delta\alpha}(\alpha,\beta) \tag{3.3}$$

$$2\partial_t D\beta = \partial_x \frac{\delta\tilde{h}}{\delta\beta}(\alpha,\beta) \, . \tag{3.4}$$

For the following it is convenient to introduce the operator J which is the inversion of the real axis with respect to the origin:

$$Ju(x) := u(-x) \, , \tag{3.5}$$

for arbitrary function $u : R\ell \to R\ell$. Note that J satisfies

$$\begin{aligned} \langle u, Jv\rangle &= \langle Ju, v\rangle \\ &\qquad\qquad \forall u,v, \\ \langle u, v\rangle &= \langle Ju, Jv\rangle \end{aligned}$$

such that

$$J^* = J \quad \text{and} \quad J^2 = Id. \tag{3.6}$$

Moreover, if A and D are pseudo differential operators with constant coefficients, then

(3.7) $\quad JD = DJ \quad \text{if} \quad D = D^*$

(3.8) $\quad JA = -AJ \quad \text{if} \quad A = -A^*$

HYPOTHESIS 5.3.1. The functional $\tilde{h}$ satisfies

(3.9) $\quad \tilde{h}(\alpha,\beta) = \tilde{h}(\beta,\alpha) \qquad \forall \alpha,\beta$.

(3.10) $\quad \tilde{h}(J\alpha, J\beta) = \tilde{h}(\alpha,\beta)$

Differentiating the identity (3.9) with respect to α(or β), there results:

(3.11) $\quad \frac{\delta\tilde{h}}{\delta\alpha}(\phi,\psi) = \frac{\delta\tilde{h}}{\delta\beta}(\psi,\phi) \qquad \forall\phi,\psi$

Differentiating (3.10) with respect to α and β there results (using $J^* = J$)

(3.12)
$$\frac{\delta\tilde{h}}{\delta\alpha}(\phi,\psi) = J\,\frac{\delta\tilde{h}}{\delta\alpha}(J\phi,J\psi) \qquad \forall\phi,\psi$$
$$\frac{\delta\tilde{h}}{\delta\beta}(\phi,\psi) = J\,\frac{\delta\tilde{h}}{\delta\beta}(J\phi,J\psi) \ .$$

Now, applying the operator J to the equations (3.3), (3.4) we find with (3.7), (3.8):

$$2\partial_t\, DJ\alpha = \partial_x\, J\,\frac{\delta\tilde{h}}{\delta\alpha}(\alpha,\beta)$$

$$2\partial_t\, DJ\beta = -\partial_x\, J\,\frac{\delta\tilde{h}}{\delta\beta}(\alpha,\beta) \ ,$$

and with (3.11) and (3.12) this may be written as

$$2\partial_t\, D\, J\alpha = \partial_x\,\frac{\delta\tilde{h}}{\delta\alpha}(J\alpha,J\beta) = \partial_x\,\frac{\delta\tilde{h}}{\delta\beta}(J\beta,J\alpha)$$

$$2\partial_t\, D\, J\beta = -\partial_x\,\frac{\delta\tilde{h}}{\delta\beta}(J\alpha,J\beta) = -\partial_x\,\frac{\delta\tilde{h}}{\delta\alpha}(J\beta,J\alpha).$$

From this result we immediately obtain

PROPOSITION:5.3.2. *If* $\tilde{h}$ *satisfies hypothesis* 5.3.1., *then we have:if* (α,β) *is a solution of* (3.3), (3.4)*then* $(J\beta,J\alpha)$ *is a solution of* (3.3), (3.4).

Stating explicitely that the initial value problem for (3.3), (3.4) is assumed to have a unique solution, we can prove the following

THEOREM 5.3.3. *If hypothesis* 5.3.1. *is satisfied, the unique solution* (α,β) *of* (3.3), (3.4) *corresponding to any initial data* (α_o,β_o) *for which*

$$\alpha_o = J\beta_o \tag{3.13}$$

satisfies

$$\alpha(t) = J\beta(t) \quad \forall_t \ . \tag{3.14}$$

Moreover, we have $\gamma(t) = \alpha(t) = J\beta(t)$, *where* γ *is the solution of the initial value problem*

$$\partial_t D\gamma = -\partial_x \frac{\delta g}{\delta \gamma}(\gamma) \tag{3.15}$$

$$\gamma(0) = \alpha_o$$

wherein the (non-translation invariant) functional g *is defined by*

$$g(\gamma) := \frac{1}{4}\tilde{h}(\gamma,J\gamma) \ . \tag{3.16}$$

PROOF. If (α,β) is the solution of (3.3), (3.4) with $(\alpha(0), \beta(0)) = (\alpha_o,\beta_o)$, then according to proposition 5.3.2.,$(J\beta,J\alpha)$ is also a solution with $(J\beta(0), J\alpha(0)) = (J\beta_o, J\alpha_o)$. Now, if α_o and β_o satisfy (3.13), then $(J\beta_o, J\alpha_o) = (\alpha_o,\beta_o)$. Hence, in that case $(\alpha(t),\beta(t))$ and $(J\beta(t), J\alpha(t))$ are both solutions of (3.3), (3.4), corresponding to the same initial value. Then $\alpha(t) = J\beta(t)$ (and $\beta(t) = J\alpha(t)$) by the uniqueness assumption. For such a solution we may insert $\beta(t) = J\alpha(t)$ in (3.3) and find that α has to satisfy the equation

$$2\partial_t D\gamma = -\partial_x \frac{\delta\tilde{h}}{\delta\gamma}(\gamma, J\gamma).$$

This is precisely equation (3.15) as follows from differentiating the functional g given by (3.16):

$$4\frac{\delta g}{\delta\gamma}(\gamma) = \frac{\delta\tilde{h}}{\delta\alpha}(\gamma, J\gamma) + J\frac{\delta\tilde{h}}{\delta\beta}(\gamma, J\gamma),$$

which can be written using (3.11) and (3.12) as

$$4\frac{\delta g}{\delta\gamma}(\gamma) = \frac{\delta\tilde{h}}{\delta\alpha}(\gamma, J\gamma) + J\frac{\delta\tilde{h}}{\delta\alpha}(J\gamma, \gamma) = 2\frac{\delta\tilde{h}}{\delta\alpha}(\gamma, J\gamma) .$$

This proves the theorem. □

These results are now easily reformulated in terms of the original equation (3.1).

THEOREM 5.3.4. *Suppose* $\bar{h}(u,p)$ *satisfies*

$$\bar{h}(u,p) = \bar{h}(-u,p) \qquad \forall u,p , \tag{3.17}$$

$$\bar{h}(Ju,Jp) = \bar{h}(u,p) \tag{3.18}$$

and assume that the initial value problem for (3.1) *has a unique solution. Then the solutions* (u,p) *of* (3.1) *which correspond to intitial data* (u_o,p_o) *for which*

$$Ju_o = -u_o , \quad Jp_o = p_o ,$$

satisfy

$$\begin{aligned} Ju(t) &= -u(t) \\ Jp(t) &= p(t) \end{aligned} \qquad \forall t \geq 0. \tag{3.19}$$

Moreover, for arbitrary, selfadjoint pseudo differential operator D, *these solutions can be represented as*

$$u = D(\gamma - J\gamma) \quad , \quad p = \gamma + J\gamma, \tag{3.20}$$

where γ *is the solution of*

$$\begin{aligned} \partial_t D\gamma &= -\partial_x \frac{\delta g}{\delta \gamma}(\gamma) \\ \gamma(0) &= \frac{1}{2}\,(D^{-1}u_o + p_o) \end{aligned} \tag{3.21}$$

with

$$g(\gamma) = \frac{1}{4}\,\bar{h}(u = D(\gamma - J\gamma),\ p = \gamma + J\gamma). \tag{3.22}$$

REMARK 5.3.5. Once it is observed that there exist solutions with the property (3.19), it is possible to define a function γ as in (3.20). Then the equation for γ can be obtained with the variational principle (1.5) for (3.1). Inserting (3.20), with $D = Id$ for simplicity, there results a functional $A(\gamma) := \overline{CA}(u,p)$:

$$A(\gamma) = -\int_I dt\ [\langle \gamma + J\gamma, \partial_t \partial_x^{-1}(\gamma - J\gamma)\rangle + 4g(\gamma)]\ . \tag{3.23}$$

Using the relations

$$\partial_x^{-1} J = J\partial_x^{-*} \quad \text{and} \quad J\partial_x^{-1} = \partial_x^{-*} J$$

(where ∂_x^{-*} is the adjoint of ∂_x^{-1}, given by 4.(2.9)), (3.23) may be simplified to

$$A(\gamma) = -4 \int dt\ [\ \frac{1}{2}\langle \gamma,\ \partial_t \partial_x^{-1}\gamma\rangle + g(\gamma)]\ , \tag{3.24}$$

where we have omitted some uninteresting terms at the endpoints of the time interval. Stationary points of (3.24) satisfy the evolution equation (3.21).

5.4. DEFINITION OF ONE-WAY PROPAGATIVITY

Consider a first-order Hamiltonian system described by

$$\partial_t \, Du = -\partial_x \frac{\delta h}{\delta u}(u). \tag{4.1}$$

It is our aim to investigate here what sense can be given to such statements as "equation (4.1) describes unidirectionally propagating waves". We shall give a definition of unidirectional propagativity, the idea of which is quite simple: intuitively it is clear that, when speaking about propagation, we mean propagation of some "property" of the system (such as wave-form, energy etc.). Let E(u) stand for such a property, depending on the considered solution u. Then it must be assured that, when considering E(u) at a fixed place x as a function of time, changes in the value of E(u) are only caused by propagation and not by any dissipative effects. Therefore, it is argued that especially *conserved* densities are able to "measure" propagation. With this observation in mind, the following definitions are proposed.

DEFINITION 5.4.1. Let E(u) be any conserved density for (4.1) and e(u) the corresponding invariant functional: $e(u) = \int_{R\ell} E(u)dx$. The *centre of gravity* of E(u), denoted by $X_E(u)$ (t) is defined by

$$\int_{R\ell} (x - X_E) \cdot E(u)dx = 0 \, . \tag{4.2}$$

Hence, X_E is a time-dependent functional of u, which can be defined for solutions u for which $e(u) \neq 0$. The *velocity* of E(u), denoted by $V_E(u)(t)$ is defined to be the velocity of the centre of gravity

$$V_E(u)(t) \, := \, \partial_t \, X_E(u)(t) \, . \tag{4.3}$$

[If T(u) is the flux density corresponding to E(u), i.e. if E and T

satisfy

$$\partial_t E(u) + \partial_x T(u) = 0 \tag{4.4}$$

for solutions of (4.1), V_E can be expressed as

$$V_E(u) = \frac{\int T(u)dx}{e(u)} \quad]. \tag{4.5}$$

We say that E is propagating to the right (to the left) for a solution u at time t if $V_E(u)(t) > 0$ ($V_E(u)(t) < 0$ respectively). Equation (4.1) is said to be *unidirectionally propagative* (in the strict sense) *with respect to the conserved density* E if $V_E(u)(t)$ can be defined for every solution (i.e. $e(u) \neq 0$ for every u) and has the same sign for *every* solution and all time.

With respect to this definition some remarks have to be made.

REMARKS 5.4.2. (i) Although the centre-of-gravity velocity of a conserved density has some physical significance, it is by no means the only possible method to describe propagation phenomena. An important practical reason to deal with the above described notion is implicitely given in the next sections: the possibility to apply this definition and to formulate, with relatively ease, general conditions on D and h under which equations of the form (4.1) are unidirectionally propagative with respect to some conserved density. Furthermore, an obvious requirement that must be imposed on any sensible definition of propagation is that, if the equation under consideration admits a solution which travels undisturbed in shape with constant velocity c, say $u(x,t) = \phi(x-ct)$, the propagation velocity to be defined must equal c for this special solution, for all time. Because of the relation $\partial_t E(\phi) = -c\partial_x E(\phi)$ for $\phi = \phi(x-ct)$, this requirement is satisfied by the centre-of-gravity velocity of every conserved density.
(ii) As a consequence of the proposed definition, with every conserved density there is associated a velocity for every solution. Suppose E_1 and E_2 are two different conserved densities (possibly with the same invariant functional!). Then, if V_1 and V_2 denote the

corresponding velocities, the velocity V_{12} of the conserved density $E_{12} = E_1 + E_2$ is easily found to be

$$V_{12} = \frac{V_1 \cdot e_1 + V_2 \cdot e_2}{e_1 + e_2}$$

for every solution u.

Furthermore, in general there is no evidence at all that if equation (4.1) is unidirectionally propagative with respect to E_1, the same is true with respect to E_2. However, for linear equations with constant coefficients it will be shown in section 5.5 that if the equation is unidirectionally propagative with respect to some definite, quadratic conserved density, the same is true for all quadratic conserved densities. [Moreover, this unidirectional propagativity is shown to be true if and only if the group velocity of the linear equation is definite. (This result gives also some confidence in the proposed definition)]. For non-linear equations no such strong relationship between the propagativity of different conserved densities has been found.

(iii) Closely related with the foregoing remark is the following observation. If E is a conserved density, and T the corresponding flux density, then E^*, defined by

$$E^*(u) = E(u) + \partial_x F(u) ,$$

where F(u) is any expression in u satisfying $F \to 0$ for $|x| \to \infty$ on the considered class of solutions, is also a conserved density with the same invariant functional

$$e(u) = \int E(u)dx = \int E^*(u)dx .$$

The flux density T^* corresponding to E^* is given by

$$T^*(u) = T(u) - \partial_t F(u) ,$$

and if X, X^* denote the centres of gravity of E and E^* respectively, with corresponding velocities V and V^* we have

$$X(u) - X^*(u) = e(u)^{-1}.\int F(u)dx$$

and

$$V(u) - V^*(u) = e(u)^{-1}.\partial_t\int F(u)dx = e(u)^{-1}.\int [T(u)-T^*(u)]dx \ .$$

From this it follows that adding a term $\partial_x F$ to the density, the corresponding velocity will change in general: only if the total flux is not altered, the velocity remains the same. This may seem a serious shortcoming of the applicability of the proposed definition. However, relying on the physics of the problem at hand, the physically most relevant densities can often be distinguished from less relevant ones. For instance, among all the densities which give the functional that can be interpreted as the total energy of the system, we take the positive definite density (if possible) as the energy density, whose centre-of-gravity velocity is then interpreted as "the energy-velocity".

(iv) *Conserved flux property.*

In general the velocity functional is not an invariant functional. However, if the conserved density E(u) has a conserved flux, i.e. if the total flux $\int T(u)dx$ itself is an invariant functional:

$$\partial_t \int T(u)dx = 0 \ , \tag{4.6}$$

then V_E as given by (4.5) is an invariant functional. In that case, it follows from (4.3) that the centre-of-gravity is a linear function of t:

$$X(u)(t) = t.V(u) + X_o(u) \ , \tag{4.7}$$

where X_o is an invariant functional (the position of the centre of gravity at t = 0). Inserting (4.7) into (4.2) gives

$$\int [x\ E(u) - t\ T(u)]dx = X_o(u).\int E(u)dx.$$

This gives rise to the following invariant functional which contains the x and t-variable explicitely:

(4.8) $$\partial_t \int [xE(u) - t\, T(u)]dx = 0 \ .$$

(v) Instead of considering all the solutions of a given equation, one is often only interested in a subset S, say, of the complete solution set. Typically, this situation is encountered if it is a-priori known that only solutions from this subset S describe the behaviour of a physical phenomenon in a required approximation. In that case, the above described definition of unidirectional propagativity of an equation is too strict, and one would like to define this notion only with respect to solutions which lie in S by requiring $V_E(u)$ to be of the same sign for all solutions from S. In section 5.7 we shall demonstrate this idea for the BBM equation, which equation is known to be a good description of "fairly long, fairly low" water waves. There it will be shown that the energy density is propagating in the same direction for any solution which can be qualified as a "fairly long, fairly low" wave, which property is lacking if the complete solution set is considered.
(vi) Finally, we note that the proposed definition makes also sense in those cases where one is dealing with more general systems than those described by (4.1): the only requirement is that there exists a local conservation law, in which case the velocity of the centre of gravity of the conserved density is again given by (4.5). An example will be given in the next section.

5.5. PROPAGATION IN LINEAR SYSTEMS.

Linear first order Hamiltonian systems are described by an equation of the form

(5.1) $$\partial_t u = -\partial_x Lu,$$

where L is some selfadjoint operator. The Hamiltonian for such equations is the quadratic functional

(5.2) $$h(u) = \frac{1}{2} \langle u, Lu \rangle .$$

In the following we shall restrict to the simplest class of operators viz. the class of pseudo-differential operators with constant coefficients. (When we speak about a "pseudo-differential operator" in the following we shall mean a pseudo-differential operator with constant coefficients).

THEOREM 5.5.1 *Any linear density* $E(u) = Pu$, *where* P *is a pseudo-differential operator, is conserved; the centre of gravity and its velocity can be defined for solutions for which* $\int Pu\ dx \neq 0$ *and we have*

$$V_E(u) = \hat{L}(0)\ , \tag{5.3}$$

where $\hat{L}$ *denotes the symbol of the operator* L. *Hence all linear densities are propagating with the same constant speed* $\hat{L}(0)$, *independent of the particular solution.*

PROOF: As P commutes with ∂_x we have

$$\partial_t P_u + \partial_x PLu = 0\ .$$

From this it follows with (5.2) and (4.5) and Fouriertransformation ($\hat{u}$ denotes the Fouriertransform of the function u) that

$$V_E(u)(t) = \frac{\int PLu\ dx}{\int Pu\ dx} = \frac{\hat{P}(0).\hat{L}(0).\hat{u}(0)}{\hat{P}(0).\hat{u}(0)} = \hat{L}(0)\ .$$

These simple observations prove the theorem. □

Quadratic conserved densities are more interesting and have been studied in great detail. It is at this point that the concept of group-velocity enters the discussion of propagation. The *dispersion relation* for equation (5.1) is

$$\omega = k.\hat{L}(k), \tag{5.4}$$

and corresponding to the group velocity $\frac{d\omega}{dk}$ we define an operator G.

DEFINITION 5.5.2. The *group velocity operator* G is defined to be the pseudo-differential operator with symbol

$$\hat{G}(k) := \frac{d}{dk}\omega(k) = \frac{d}{dk}(k.\hat{L}(k)) \quad . \tag{5.5}$$

LEMMA 5.5.3 *As* L *is self-adjoint, the operator* G *is selfadjoint.*

PROOF: As $\hat{L}(k)$ is an even function of k ($k \in R\ell$), $\hat{G}$ is an even function of k for $k \in R\ell$. Thus G is a selfadjoint operator. □

LEMMA 5.5.4. *Any quadratic density* E(u) = Pu.Qu *where* P *and* Q *are pseudo-differential operators, is a conserved density for the equation* (5.1).

PROOF: Generally, E(u) = Pu.Qu is a conserved density of (5.1) if P and Q satisfy

$$P^{*}Q\partial_x L = L\,\partial_x P^{*}Q \quad , \tag{5.6}$$

in which case the flux density corresponding to E is given by

$$T(u) = \int_{-\infty}^{x} [Pu.Q\ \partial_x Lu - P\partial_x Lu.Qu](\xi)d\xi \quad . \tag{5.7}$$

For pseudo-differential operators P and Q condition (5.6) is clearly satisfied. □

As a typical result concerning the relation between group-velocity and the propagation of conserved densities by monochromatic solutions of (5.1), we quote the following result

THEOREM 5.5.5. *Let* E *be a quadratic conserved density with corresponding flux density* T. *Then* E *is propagating with the group velocity in the following sense: for monochromatic solutions* $\Phi(x,t) = \Phi_o . \exp i(k_o x - \omega_o t)$, *where* Φ_o *is a constant and* $\omega_o = \omega(k_o)$, *the following relation holds:*

(5.8) $$\frac{T(\Phi)}{E(\Phi)} = \hat{G}(k_o) .$$

This theorem is well known and can be found e.g. in de Graaf & Broer [48]. With the proposed definition 5.4.1. it turns out to be possible to relate the group-velocity to non-periodic solutions.

THEOREM 5.5.6. *Consider a definite conserved quadratic density* $E(u) = Au.Au$, *where* A *is some pseudo-differential operator. Then the centre-of-gravity velocity of this density is an invariant functional and is given by*

(5.9) $$V_E(u) = \langle Au, Au\rangle^{-1} . \langle Au, G\, Au\rangle ,$$

where G *is the group velocity operator. Hence* $V_E(u)$ *equals the weighted group velocity with weightfunction* $|\widehat{Au}|^2$.

PROOF: Inserting equation (5.1) directly into definition (4.3) it follows that

$$V_E(u)(t) = 2.\langle Au, Au\rangle^{-1} . \langle Au, -x\partial_x L\, Au\rangle .$$

With Parseval's theorem and the expression (5.4) we find

$$V_E(u)(t) = 2\, \langle \widehat{Au}, \widehat{Au}\rangle^{-1} . \langle \widehat{Au}, \partial_k[\omega . \widehat{Au}]\rangle$$

and after some straightforward manipulations

(5.10) $$V_E(u)(t) = \langle \widehat{Au}, \widehat{Au}\rangle^{-1} . \langle \widehat{Au}, \frac{d\omega}{dk} . \widehat{Au}\rangle = \langle Au, Au\rangle^{-1} . \langle Au, G\, Au\rangle .$$

With Lemma 5.5.4. it follows that V_E is an invariant functional and the theorem is proved. □

An immediate consequence of this theorem is

COROLLARY 5.5.7. *Equation* (5.1) *is unidirectionally propagative (to the right) with respect to any definite, conserved quadratic density if and only if the group velocity is non-negative for all wave numbers:*

$$\hat{G}(k) \geq 0 \quad \textit{for all } k \in Rl \ . \tag{5.11}$$

REMARK 5.5.8. For more general conserved quadratic densities of the form $E(u) = Au.Bu$, A and B pseudo-differential operators, the corresponding centre-of-gravity velocity is found to be the invariant functional

$$V_E(u) = \langle Au, Bu \rangle^{-1} . \langle Au, G\, Bu \rangle \ . \tag{5.12}$$

To conclude this section on propagation in linear systems, we consider a classical Hamiltonian system and relate the energy propagation of the complete system to the energy propagation of the subsystems in which it may be separated. By way of example we restrict to the simple set of equations

$$\begin{aligned} \partial_t u &= -\partial_x L\, p \\ \partial_t p &= -\partial_x L\, u \end{aligned} \tag{5.13}$$

where $L = A^2$, with A some positive self-adjoint operator. The Hamiltonian for (5.13) expressed in the non-canonical variables u and p

$$h(u,p) = \frac{1}{2} \langle p, Lp \rangle + \frac{1}{2} \langle u, Lu \rangle \tag{5.14}$$

is the total energy of the system and is an invariant functional for (5.13) to which there corresponds a local conservation law of the form (4.4). Hence the velocity of the energy propagation may be defined in a meaningful way. Defining the centre of gravity $X(u,p)$ of the total energy density $\frac{1}{2}$ (Au.Au + Ap.Ap) by

$$\int (x-X)(Au.Au + Ap.Ap)dx = 0 ,$$

its velocity can be evaluated along the same lines as was done in the proof of theorem 5.5.6. with the aid of Fouriertransform techniques:

$$h(u,p).\ V(u,p) = \langle Au, xA\partial_t u\rangle + \langle Ap, xA\partial_t p\rangle$$

$$= \langle Au, -x\partial_x LAp\rangle + \langle Ap, -x\partial_x LAu\rangle$$

$$= \frac{1}{2}\langle A(u+p), -x\partial_x LA(u+p)\rangle - \frac{1}{2}\langle A(u-p), -x\partial_x LA(u-p)\rangle$$

$$= \frac{1}{4}\langle \hat{A}.(\hat{u}+\hat{p}), \frac{d\omega}{dk}.\hat{A}.(\hat{u}+\hat{p})\rangle - \frac{1}{4}\langle \hat{A}.(\hat{u}-\hat{p}), \frac{d\omega}{dk}.\hat{A}.(\hat{u}-\hat{p})\rangle ,$$

thus

$$V(u,p) = \frac{1}{4}.h(u,p)^{-1}.[\langle A(u+p), GA(u+p)\rangle - \langle A(u-p), GA(u-p)\rangle], \tag{5.15}$$

where ω is a solution branch of the dispersion relation of (5.13):

$$\omega(k) = k.\ \hat{L}(k) \tag{5.16}$$

and G is the group velocity operator with symbol $\frac{d\omega}{dk}$. Under the transformation

$$p = \alpha + \beta \tag{5.17}$$

$$u = \alpha - \beta$$

the equations for α and β become

$$\partial_t \alpha = -\partial_x L\alpha \tag{5.18}$$

$$\partial_t \beta = \partial_x L\beta ,$$

and for these separated equations the energy velocities are given by (cf. (5.9)):

(5.20) $$V(\alpha) = e(\alpha)^{-1} \cdot \langle A\alpha, GA\alpha \rangle$$

(5.21) $$V(\beta) = -e(\beta)^{-1} \cdot \langle A\beta, GA\beta \rangle$$

respectively, where

$$\alpha(\alpha) = \langle \alpha, L\alpha \rangle = \langle A\alpha, A\alpha \rangle$$

is the invariant energy functional for the α-mode (5.18) and $e(\beta)$ is the invariant energy functional for the β-mode (5.19). Note that

(5.22) $$h(u,p) \equiv e(\alpha) + e(\beta),$$

and with (5.20) and (5.21) the expression (5.15) can be written as

(5.23) $$V(u,p) = V(\alpha) \cdot \frac{e(\alpha)}{h(u,p)} + V(\beta) \cdot \frac{e(\beta)}{h(u,p)} .$$

This result clearly shows how the α- and the β-mode contribute to the energy propagation of the complete system. (Note that all the functionals entering in (5.23) are invariant functionals).

5.6. ONE WAY PROPAGATIVE FIRST ORDER HAMILTONIAN SYSTEM.

In this section we shall consider first order Hamiltonian systems as introduced in section 4.2.:

(6.1) $$\partial_t u = -\partial_x \frac{\delta h}{\delta u}(u) ,$$

where h is a translation invariant functional. As was observed before, this equation admits three invariant integrals

(6.2) $$\ell(u) = \int_{R\ell} u \, dx \qquad \text{mass-functional}$$

(6.3) $m(u) = \frac{1}{2} \langle u,u \rangle$ momentum-functional

(6.4) $h(u)$ energy functional.

It is possible to write down the local conservation laws corresponding to these invariant integrals in a fairly general way. From these expressions the velocities of the conserved densities are then found with (4.5). However, these general formulae are not very transparent, and for shortness we shall therefore restrict to two simple classes of equations.

DEFINITION 5.6.1. Equation (6.1) is said to be a *local equation* if the operator $\frac{\delta h}{\delta u}$ is a local operator, i.e. if $\frac{\delta h}{\delta u}(u)(x)$ depends on u and its derivatives with respect to x at the point x only. Equation (6.1) is said to be a *non-local* equation if $\frac{\delta h}{\delta u}$ is not a local operator.

EXAMPLE 5.6.2. The k.d.V equation 4.(1.1)

(6.5) $$\partial_t u = -\partial_x (u + \frac{1}{2} u^2 + \partial_x^2 u)$$

is a local equation, but the B.B.M. equation 4.(1.2)

(6.6) $$\partial_t (1 - \partial_x^2) u = -\partial_x (u + \frac{1}{2} u^2)$$

is a non-local equation.

Generally speaking, local equations are rather easy to deal with. The velocities of the conserved densities are usually readily found and sufficient conditions can be given which assure that the equation is unidirectionally propagative. For a simple, but representative, class of local equations we shall summarize some results. Non-local equations are usually more difficult. Although the expressions for the velocities of conserved densities may be written down, these expressions are functionals with non-quadratic, non-local integrands, the positivity of which is difficult to

investigate. Nevertheless, for a special class of non-local equations we shall derive some remarkable results, especially in connection with the propagation of the energy density.

A class of local equations.

We consider local equations of the form (6.1) for which the Hamiltonian h(u) is given by

$$h(u) = \int dx\, [N(u) + S(u_x)] \,, \tag{6.7}$$

where N and S are smooth (C^2-) functions of their arguments with derivatives n and s respectively:

$$n(y) := N'(y) = \frac{dN}{dy}(y) \qquad y \in Rl$$

$$s(z) := S'(z) = \frac{dS}{dz}(z) \qquad z \in Rl$$

(primes denote differentiations with respect to the arguments). We assume that

$$N(0) = S(0) = n(0) = s(0) = 0 :$$

$N(0) = S(0) = u(0)$ to assure that $h(0) = \frac{\delta h}{\delta u}(0) = 0$ and $s(0) = 0$ is no restriction. Equation(6.1) with h as in (6.7) then reads

$$\partial_t u = -\partial_x (n(u) - \partial_x s(u_x)) \,. \tag{6.8}$$

In a straightforward way the following results can be obtained

LEMMA 5.6.3. *The centre-of-gravity velocities of the conserved densities*

$$u, \qquad u.u \qquad , \qquad N(u) + S(u_x) \tag{6.9}$$

are given by

$$V(u) = [\int u\,dx]^{-1}.[\int n(u)dx] \tag{6.10}$$

$$V(u) = 2.\langle u,u\rangle^{-1}.\int dx[u.n(u) - N(u) + 2u_x.s(u_x) - S(u_x)] \tag{6.11}$$

$$V(u) = h(u)^{-1}.\int dx[\tfrac{1}{2}n^2(u) + 2u_x.s(u_x).n'(u) + \tfrac{3}{2}(\partial_x s(u_x))^2] \tag{6.12}$$

which are called the mass velocity, the momentum velocity and the energy velocity respectively.

It is now a simple matter to state conditions for N and S that assure that the momentum- and/or energy velocity are positive for all solutions:

LEMMA 5.6.4. *Equation* (6.8) *is unidirectionally propagative to the right with respect to*
(i) *the momentum density if*

$$y.n(y) - N(y) \geq 0\ ,\quad 2z.s(z) - S(z) \geq 0 \quad \forall y,z \in Rl$$

(ii) *the enenergy density (required to be positive) if*

$$N(y) \geq 0\ \&\ n'(y) \geq 0\ ,\quad S(z) \geq 0\ \&\ 2z.s(z) \geq 0 \quad \forall y,z \in Rl$$

(iii) *both the momentum-and the energy density if*

$$N(y) \geq 0\ \&\ y.n(y) - N(y) \geq 0\ \&\ n'(y) \geq 0 \quad \forall y \in Rl$$

$$S(z) \geq 0\ \&\ 2z.s(z) - S(z) \geq 0 \quad \forall z \in Rl.$$

REMARKS 5.6.5.
(i) Linearizing equation (6.8) gives

$$\partial_t u = -\partial_x(n'(0).u - s'(0).\partial_x^2 u)\ , \tag{6.13}$$

which equation has the dispersion relation

$$\omega(k) = n'(0).k + s'(0).k^3$$

and group velocity

$$\frac{d\omega}{dk}(k) = n'(0) + 3\, s'(0).k^2 .$$

The velocity functionals of the momentum density u^2 and the (linearized) energy density $\frac{1}{2} n'(0).u^2 + \frac{1}{2} s'(0).u_x^2$ of this linear equation (6.13) (as given by (5.9)) are easily seen to be the quadratic terms in a Taylor expansion of the integrands of (6.11) and (6.12) respectively.

(ii) As u and u^2 are conserved densities for (6.8), the mass velocity (6.10) is an invariant functional if

$$n(u) = \alpha u + \beta u^2, \quad \alpha,\beta \in R\!\!\!Z .$$

In that case, u is a conserved density with conserved flux and according to remark 5.4.2.(iv) there exists an invariant functional which depends on x and t explicitely; in this case

$$\partial_x \int dx\, [xu - t\, n(u)] = 0 . \tag{6.14}$$

(iii) The KdV equation (6.5) belongs to the considered class of equations with

$$N(y) = \frac{1}{2} y^2 + \frac{1}{6} y^3, \quad S(z) = -\frac{1}{2} z^2 .$$

This equation is neither unidirectionally propagative with respect to the momentum- nor with respect to its energy density, as may be confirmed from the expressions (6.11) and (6.12). The mass velocity is an invariant functional, and (6.14) reads

$$\partial_t \int dx\, [xu - t(u + \frac{1}{2} u^2)] = 0 . \tag{6.15}$$

This functional has already been met in section 4.3. in a slightly

different form as the functional $C_1(u)$.

A class of non-local equations.

Here we shall examine equations of the form

$$\partial_t Du = -\partial_x \frac{\delta e}{\delta u}(u), \tag{6.16}$$

where the functional e is given by

$$e(u) = \int N(u)dx, \tag{6.17}$$

with N a smooth (C^2-) function of its argument, $n(u) := \frac{dN}{du}(u)$, and $N(0) = n(0) = 0$, and where $D = A^2$ with A some positive selfadjoint pseudo-differential operator. From section 4.2 it follows that via a simple linear transformation equation (6.16) can be brought into an equation of the form (6.1), but except when A^{-1} is an ordinary differential operator this equation will be of non-local type. In these cases it is somewhat simpler to deal directly with the form (6.16). The three invariant functionals for (6.16), corresponding to (6.2), (6.3) and (6.4), are given by

$$\ell(u) = \int Du\, dx \qquad \text{(mass)} \tag{6.18}$$

$$m(u) = \frac{1}{2}\langle u, Du\rangle \qquad \text{(momentum)} \tag{6.19}$$

$$e(u) = \int N(u)dx \qquad \text{(energy)} \tag{6.20}$$

The mass velocity is again given by (6.10), and remark 5.6.5. (ii) applies as well: if $\int n(u)dx$ is an invariant functional, then

$$\partial_t \int dx\, [x\, Du - t\, n(u)] = 0\ . \tag{6.21}$$

For the following we define the (self-adjoint) operator G as the pseudo-differential operator with symbol $\hat{G}$ where

$$\hat{G}(k) = \partial_k[k.\hat{D}^{-1}(k)] = \hat{D}^{-1}(k).[1 - k.\hat{D}^{-1}(k).\partial_k\hat{D}(k)]\ . \tag{6.22}$$

[Note that if n has a linear term, say $n'(0) = 1$, then $\omega(k) := k.\hat{D}^{-1}(k)$ is the dispersion relation of the linearized equation (6.16), and $\hat{G}(k)$ the corresponding group velocity. However, if equation (6.16) does not admit a formal linearization, i.e. if $n'(0) = 0$, this interpretation of ω and $\hat{G}$ makes no longer sense, but the results to be derived remain valid!].

THEOREM 5.6.6. *The centre-of-gravity velocity of the energy density is given by*

$$V_E(u)(t) = \frac{1}{2}.e(u)^{-1}.\langle n(u),\ G\ n(u)\rangle\ . \tag{6.23}$$

Consequently, equation (6.16), (6.17) *is unidirectionally propagative to the right with respect to the energy density (assumed to be positive) if and only if*

$$\hat{G}(k) \geq 0 \qquad \forall k \in R\ell. \tag{6.24}$$

PROOF: The proof of this result is analogous to the proof of theorem 5.5.6: using Fourier-transform techniques and writing $\hat{n}$ for the Fourier-transform of the expression n(u) we find

$$\partial_t \int xN(u)dx = \int x.n(u).\partial_t u\ dx = \int n(u).(-x\partial_x D^{-1}\ n(u))dx =$$

$$= \frac{1}{2}\int \bar{\hat{n}}(k).\hat{G}(k).\hat{n}(k)dk.$$

Hence

$$V_E(u) = \frac{1}{2}\ e(u)^{-1}.\langle \hat{n},\hat{G}.\hat{n}\rangle = \frac{1}{2}\ e(u)^{-1}.\langle n(u),G\ n(u)\rangle, \tag{6.25}$$

and the theorem follows. □

REMARKS 5.6.7.

(i) The BBM equation (6.6) belongs to the considered class of equations with

$$D = 1 - \partial_x^{\,2}\ ,\quad N(u) = \frac{1}{2}\ u^2 + \frac{1}{6}\ u^3. \tag{6.26}$$

For this operator D, the function G is not positive for all $k \in \mathbb{R}l$:

$$\hat{G}(k) = (1+k^2)^{-2}.(1-k^2). \tag{6.27}$$

Hence the BBM equation is not unidirectionally propagative with respect to its (non-definite) energy density (See however the results of the next section for a restricted set of solutions).
(ii) More generally, if D has symbol

$$\hat{D}(k) = (1+a^2k^2)^{\sigma} \quad \text{with } a \in \mathbb{R}l, \sigma \in \mathbb{R}l, \tag{6.28}$$

then $\hat{G}$ is given by

$$\hat{G}(k) = (1+a^2k^2)^{-\sigma-1}.(1+(1-2\sigma)a^2k^2) \ . \tag{6.29}$$

Hence, for such operators, condition (6.24) is satisfied if and only if

$$\sigma \leq \frac{1}{2} \ . \tag{6.30}$$

The energy velocity as given by (6.23) is remarkably simple. However, matters are much more complicated for the velocity of the (positive) quadratic momentum density Au.Au. Using Fourier transform techniques in intermediate steps it is possible to derive the following result.

LEMMA 5.6.8. *The centre-of-gravity velocity of the positive momentum density can be written with the operator* G *as*

$$V(u)(t) = \langle Au, Au \rangle^{-1}. \int dx \ [u.n(u) - 2N(u) + n(u).DGu]. \tag{6.31}$$

Note that for linear equations, $N(u) = \frac{1}{2} u^2$, (6.31) agrees with (5.9). But in the more interesting case of non-linear equations it seems to be impossible to derive conditions on N and D such that (6.31) is positive.

5.7. ONE WAY PROPAGATIVE LONG-LOW WAVE MODELS.

In this section we shall once again examine equations of the form

$$\partial_t \, Du = -\partial_x \, n(u) \tag{7.1}$$

where D is a pseudo-differential operator and n(u) a smooth function of its argument. The energy density N(u), with $n(u) = \frac{dN}{du}(u)$, $N(0) = n(0) = 0$, is no longer required to be positive. In view of the results of the foregoing section we shall only consider the velocity of the centre of gravity of the energy density. This velocity is given by (6.23).

$$V(u) = \frac{1}{2} \, .e(u)^{-1} . \langle n(u), G \, n(u) \rangle \tag{7.2}$$

where G is the pseudo-differential operator with symbol given by (6.22). We shall suppose that

$$n'(0) = 1 \, , \tag{7.3}$$

such that G can be interpreted as the group velocity operator of the linearized problem. In the foregoing section it was shown that the BBM equation

$$(1-\partial_x^{\,2})\partial_t \, u = - \, \partial_x (u + \frac{1}{2} \, u^2) \tag{7.4}$$

is not unidirectionally propagative with respect to the energy density. However, the BBM equation (as many other equations of this type) is derived as an approximate equation for the description of "fairly long, fairly low" waves (c.f. chapter 6 for more details about this approximate chapter of the equation). Therefore it is reasonable to investigate the positivity of the functional (7.2) on the restricted class of functions which can be described as long, low waves. To make this idea more concrete, let us suppose that we

can define two functionals ε and λ whose values $\varepsilon(u)$ and $\lambda(u)$ are a measure of the height and of the "length" of the function u respectively. Then the class of long, low waves can be described as the set of functions satisfying

$$\begin{aligned} \varepsilon(u) &< \varepsilon_o \\ \lambda(u) &> \lambda_o \end{aligned} \tag{7.5}$$

where ε_o and λ_o^{-1} are small positive numbers. Now suppose that numbers ε_o and λ_o can be found such that $V_E(u)$ is of the same sign (positive say) for every function u which satisfies (7.5). Then, if u is a solution of (7.1) which satisfies (7.5) at some instant t_o, $V_E(u)(t)$ will be positive at $t = t_o$ and for times $t > t_o$ as long as u(t) satisfies condition (7.5). *A priori*, it is by no means clear that solutions corresponding to initial data which satisfy (7.5), satisfy this condition for all $t > 0$. Especially for non-linear equations this is a critical point. To demonstrate this for the long wave length condition for instance, consider the solution of (7.1) corresponding to an initial value g(x) whose Fourier transform $\hat{g}$ satisfies $\hat{g}(k) = 0$ for $|k| \geq k_o$, $k \in Rl$ (i.e. g consists of long wave components only). A Fourier transformation of equation (7.1) shows that if the equation is *linear*, then $\hat{u}(k,t) = 0$ for $|k| \geq k_o$ for all $t \geq 0$, but if the equation is *non-linear* then $\hat{u}(k,t) \neq 0$ for almost all $k \in Rl$, no matter how small $t > 0$: *initial long wave components generate short wave components instantly.*

From these remarks and observations the following definition will be acceptable.

DEFINITION 7.1. Equation (7.1) is said to be unidirectionally propagative with respect to the energy density for long waves if positive numbers ε_o and λ_o can be found such that $V_E(u)(t) \geq 0$ for all $t \geq 0$ and every solution u whose initial value satisfies

$$\begin{aligned} \varepsilon(u(x,0)) &< \varepsilon_o \\ \lambda(u(x,0)) &> \lambda_o \,. \end{aligned} \tag{7.6}$$

For the following we suppose that the symbol of the group velocity operator G can be estimated as

$$\hat{G}(k) \geq \hat{G}(0).\left[1 - \ell^2 k^2\right] \quad , \quad k \in R^{\ell} \tag{7.7}$$

wherein $\hat{G}(0)$ and ℓ are positive numbers. For long wave models such an estimate is generally possible: the long wave components propagate with the largest, positive speed (the groupvelocity has a positive maximum at k = 0). With (7.7) the velocity functional (7.2) can be estimated as

$$V(u) \geq \hat{G}(0).\frac{||n(u)||^2}{e(u)} . \left\{1 - \ell^2 . \frac{||\partial_x n(u)||^2}{||n(u)||^2}\right\} . \tag{7.8}$$

From this it immediately follows that $V(u) \geq 0$ if

$$\lambda(u) \geq \ell \, , \tag{7.9}$$

if the functional λ is defined by

$$\lambda(u)^{-1} := \frac{||\partial_x n(u)||}{||n(u)||} . \tag{7.10}$$

REMARK 5.7.2. The functional λ defined by (7.10) can indeed be interpreted as an averaged wave length : $\lambda(u)^{-2}$ is the weighted average of k^2 with weight function $|\widehat{n(u)}|^2$. Another way to interpret $\lambda(u)$ as a measure of the "length" of the function u follows from the observation

$$\lambda(u_\mu)^{-2} = \mu^2 . \lambda(v)^2 \qquad \text{for } u_\mu(x) := v(\mu x) \; ; \tag{7.11}$$

then $\lambda(u_\mu) \to \infty$ for $\mu \to 0$.

In the following we shall show that it is sometimes possible to find conditions of the form (7.6), i.e. conditions imposed on the initial data, which assure that the resulting solutions satisfy condition (7.9) for all $t \geq 0$. For simplicity we shall restrict ourselves to a specific equation, viz. the BBM equation (7.4). Note that this

equation satisfies (7.7) with

$$\hat{G}(0) = 1, \quad \ell^2 = 3. \tag{7.12}$$

For the sup norm $|\ |_\infty$ and the first Sobolev norm $||\ ||_1$, defined by

$$|u|_\infty := \sup_{x \in R\ell} |u(x)|; ||u||_1^2 := ||u||^2 + ||u_x||^2 ,$$

we shall need the estimate in the following lemma.

LEMMA 5.7.3. *For arbitrary* $u \in H^1(R\ell)$ *we have* $u \in C^0(R\ell)$ *and*

$$|u|_\infty^2 \leq \frac{1}{2} ||u||_1^2 \tag{7.13}$$

PROOF: This result has already been recalled in section 0.2.3. (the Embedding theorem). To give an independent proof of the relation (7.13), note that the first Sobolev space $H^1(R\ell)$ $(=H_o^1(R\ell))$ can be defined with Fouriertransform techniques as ($\hat{u}$ denotes the Fourier transform of a function u):

$$H^1(R\ell) = \{u \in L_2(R\ell) |\ \hat{u}.(1+k^2)^{\frac{1}{2}} \in L_2(R\ell)\}.$$

Then for $u \in H^1(R\ell)$:

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{u}(k).e^{ikx}\, dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{u}(k).(1+k^2)^{\frac{1}{2}}.(1+k^2)^{-\frac{1}{2}} dk$$

and thus

$$|u|_\infty^2 \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} (1+k^2).|\hat{u}(k)|^2 dk. \int_{-\infty}^{\infty} (1+k^2)^{-1}\, dk = \frac{1}{2} ||u||_1^2 ,$$

which proves the estimate (7.13). □

Concerning the existence of a classical solution of the initial value problem for the BBM equation, we quote the following result:

THEOREM 5.7.4. *Let* $u_o \in C^2(R) \cap H^1(R)$. *Then there exists a unique (classical) solution* u *of equation* (7.4) *with*

$$u(x,0) = u_o(x)$$

and $u(.,t), \partial_t u(.,t) \in C^2(R\ell) \cap H^1(R\ell)$ *for all* $t \geq 0$.

Consequently, the momentum and energy functional are neatly defined and are invariant:

$$\begin{aligned} m(u) &= m(u_o) \\ e(u) &= e(u_o) \end{aligned} \quad \forall t \geq 0.$$

PROOF: The proof of this theorem can be found in Benjamin et al [39]□

We are now in a position to formulate the main result.

THEOREM 5.7.5. *The BBM equation* (7.4) *is unidirectionally propagative to the right with respect to the energy density for the class of long, low waves, which class is characterized as the solutions whose initial value satisfy*

$$\begin{aligned} & u(x,0) \in C^2(R\ell) \cap H^1(R\ell) \\ & \lambda(u(x,0)) > \lambda_o \\ & \varepsilon(u(x,0)) < \varepsilon_o \end{aligned} \tag{7.14}$$

for sufficiently small positive numbers ε_o *and* λ_o^{-1}. *Here,* λ *is the functional defined by* (7.10) *and*

$$\varepsilon(u) := ||u||_1 .$$

PROOF: In view of the estimate (7.8) and result (7.12) we have to show that λ_o and ε_o can be found such that

$$\lambda(u)^{-2} \leq \frac{1}{3} \tag{7.15}$$

for every $t \geq 0$ and every solution with initial data satisfying (7.14). Let u_o denote the initial value and define $\delta > 0$ by

$$||u_o||_1^2 = \frac{1}{2}\delta^2.$$

As

$$m(u) = \frac{1}{2} < u,(1 - \partial_x^2)u > = \frac{1}{2}||u||_1^2 \tag{7.16}$$

is an invariant functional, it follows that

$$||u||_1^2 = \frac{1}{2}\delta^2 \qquad \forall t \geq 0$$

and with (7.13) that

$$|u|_\infty < \delta \qquad \forall t \geq 0.$$

Then we can derive the following useful estimates for the functionals e and λ :

$$\frac{1}{2}(1 - \frac{1}{3}\delta).||u||^2 \leq e(u) \leq \frac{1}{2}(1 + \frac{1}{3}\delta).||u||^2 \qquad \forall t \geq 0 \tag{7.18}$$

and, provided $\delta<2$,

$$\left(\frac{1-\delta}{1+\frac{1}{2}\delta}\right)^2.\left(\frac{||u||_1^2}{||u||^2} - 1\right) \leq \lambda(u)^{-2} \leq \left(\frac{1+\delta}{1-\frac{1}{2}\delta}\right)^2.\left(\frac{||u||_1^2}{||u||^2} - 1\right) \tag{7.19}$$

$$\forall t \geq 0.$$

As e is an invariant functional, it follows from (7.18) that

$$(1-\frac{1}{3}\delta).||u_o||^2 \leq (1+\frac{1}{3}\delta).||u||^2 \qquad \forall t \geq 0. \tag{7.20}$$

Writing $\lambda_o = \lambda(u(x,0))$ it follows from (7.19) that,

$$\left(\frac{1-\delta}{1+\frac{1}{2}\delta}\right)^2.\left(\frac{||u_o||_1^2}{||u_o||^2} - 1\right) \leq \lambda_o^{-2},$$

and, as $||u||_1$ is invariant, we obtain, provided $\delta<1$:

$$\frac{||u||_1^2}{||u_o||^2} \leq 1 + \lambda_o^{-2}.\left(\frac{1+\frac{1}{2}\delta}{1-\delta}\right)^2 \qquad \forall t \geq 0 \tag{7.21}$$

With (7.20) and (7.21) we can majorize the right hand side of (7.19) and obtain

$$(7.22)\quad \lambda(u)^{-2} \le \left(\frac{1+\delta}{1-\frac{1}{2}\delta}\right)^2 \cdot \left[\left(\frac{1+\frac{1}{3}\delta}{1-\frac{1}{3}\delta}\right) \cdot \left(1+\lambda_o^{-2} \cdot \left(\frac{1+\frac{1}{2}\delta}{1-\frac{1}{2}\delta}\right)^2 \right) - 1 \right] \quad \forall t \ge 0$$

This result shows that $\lambda(u)^{-2}$ can be majorized uniformely with respect to t in terms of initial value δ and λ_o. Moreover, it is easily seen that the right hand side of (7.22) can be bounded above by $\frac{1}{3}$ if δ and λ_o^{-1} are taken sufficiently small. This shows that condition (7.15) is satisfied for δ(and hence ε_o) and λ_o sufficiently small. With the extra observation that e(u) is positive if $\delta<3$, as follows from (7.18), this proves the theorem. □

REMARKS 5.7.6. (i) From a physical point of view the foregoing theorem is satisfactory because the requirements define the functions to be low waves, as follows from the estimate (7.13), and to be long waves in the sense of remark 5.7.2. However, it is possible to show that the velocity functional (7.2) is positive on a larger class of functions. Therefore, define the functional Λ by

$$(7.23)\qquad \Lambda(u) := \frac{m(u)}{e(u)} .$$

Then it can be shown that

$$(7.24)\qquad V(u) > 0 \quad \text{for every } u \in S_\gamma ,$$

where S_γ is the set of functions for which

$$(7.25)\qquad \begin{aligned} \varepsilon(u) &:= ||u||_1 < \gamma \\ \Lambda(u) &\le \Gamma(\gamma) \end{aligned}$$

for some γ, $0<\gamma<2$, where the function $\Gamma(\gamma)$ is given by

$$\Gamma(\gamma) = (1+\tfrac{1}{3}\gamma)^{-1} \cdot \left\{ 1+\tfrac{1}{3}\,\frac{1-\frac{1}{2}\gamma}{1+\gamma} \right\} .$$

[Note that $\Gamma(0) = \frac{4}{3}$, $\Gamma(2) = \frac{3}{5}$ and

(7.26) $\Gamma(\gamma) > 1$ for $0 < \gamma < \gamma_0 = \frac{1}{2}\left(\sqrt{\frac{11}{2}} - \frac{3}{2}\right)$].

As the functionals ε and Λ are invariant functionals for the BBM equation, it follows that $V(u)(t) > 0$ for all $t \geq 0$ for every solution whose initial value satisfies condition (7.25). Although the functional Λ has the advantage of being an invariant functional, its relevance as a measure of the "length" of a function is less clear. Nevertheless, for functions

$$u_{\delta,\mu}(x) := \delta\, v(\mu x)$$

we have

$$\Lambda(u_{\delta,\mu}) = \left[||v||^2 + \mu^2||\partial_x v||^2\right] \cdot \left[||v||^2 + \frac{1}{6}\delta\int_{-\infty}^{\infty} dy\; v^3(y)\right]^{-1},$$

such that

$$\Lambda(u_{\delta,\mu}) \to 1 \quad \text{for} \quad \delta,\mu \to 0 .$$

From this it follows that for the class of long, low waves $\Lambda \approx 1$, and hence, because of (7.26), this class is included in the set S_γ for $\gamma < \gamma_0$. This shows that the result stated above includes the contents of theorem 5.7.5.

(ii) It is illustrative to apply the above described method to more general equations of the form (7.1) where D is given by (6.28). The essential tools used above to derive the results are (a) the existence theorem, (b) the estimate between $|\ |_\infty$ and the invariant functional m:

(7.27) $$|u|_\infty^2 \leq \text{const. } m(u).$$

(Smoothness of the function n is assumed; an estimate as (7.7) is possible for the considered class of operators D.). In fact, reviewing the proof of theorem 5.7.4., it may be seen that an estimate of the form (7.27) is essential to prove the existence of a

solution for *all* time: With the contracting mapping principle, the existence of a regular solution can be proved over a limited range of time $[0,T_o]$, where T_o depends, for given n, only on $|u(t=0)|_\infty$ and $||u(t=0)||_\sigma$, where $||\ ||_\sigma$ is the norm of the Sobolev space H^σ ($\sigma>0$):

$$H^\sigma := \{u \in L_2 \mid \hat{u}.(1+k^2)^{\frac{\sigma}{2}} \in L_2\},$$

$$||u||_\sigma^2 = \int_{-\infty}^{\infty} (1+k^2)^\sigma . |\hat{u}|^2 \, dk \ .$$

As

$$m(u) = \tfrac{1}{2}||u||_\sigma^2 \ , \tag{7.28}$$

if an estimate (7.27) is available, T_o depends only on $m_o = m(u(t=0))$. Taking $u(T_o)$ as initial value, the process can be repeated to prove the existence over the time interval $[T_o, T_1]$. Then $T_1 - T_o$ depends only on $m(u(T_o))$, and as $m(u(T_o)) = m_o$, $T_1 - T_o = T_o$, such that $T_1 = 2T_o$. Repeating this process, the existence can be proved for all $t \geq 0$. However, as H^σ is continuously embedded in $C^o(R\ell)$ only if $\sigma > \frac{1}{2}$, an estimate of the form (7.27) (and an existence theorem) is available only if

$$\sigma > \frac{1}{2} \ . \tag{7.29}$$

Therefore we can conclude that qualitatively the same results as obtained for the BBM equation above, can be obtained for equations of the form (7.1) with D given by (6.28), provided (7.29) is satisfied. (It is intriguing that this requirement conflicts the condition $\sigma \leq \frac{1}{2}$ for positivity of the group velocity c.f.(6.30)).

CHAPTER 6: THEORY OF SURFACE WAVES.

6.1. INTRODUCTION.

In this last chapter we shall demonstrate the ideas developed in the foregoing chapters of part II to a specific system, viz. the irrotational motion of a two-dimensional inviscid layer of fluid over a horizontal bottom under influence of gravity. The fluid is assumed to be incompressible and its constant density ρ_o, the constant of gravity g and the undisturbed fluid height are all normalized to 1.
The motion of the fluid can be described in terms of a velocity potential ϕ:

$$\underline{v} = \text{grad } \phi = (\phi_x, \phi_y)$$

and the elevation η of the free surface measured from the equilibrium configuration. The complete system is then described by (c.f. section 3.5)

$$\phi_{xx} + \phi_{yy} = 0 \qquad 0 < y < 1 + \eta(x,t) \tag{1.1}$$

$$\phi_y = 0 \qquad y = 0 \tag{1.2}$$

$$\left.\begin{aligned} &\partial_t \eta + \phi_x \eta_x - \phi_y = 0 \qquad (1.3)\\ &\partial_t \phi + \tfrac{1}{2}(\phi_x^2 + \phi_y^2) + \eta = 0 \qquad (1.4)\end{aligned}\right\} \quad y = 1 + \eta(x,t)$$

Notwithstanding its long history, there are hardly any achievements concerning concrete results for the free surface problem (1.1)-(1.4). From Benjamin [35], p.7., we quote: "In respect of the initial value problem for the system (1.1) - (1.4) , virtually nothing in the way of a rigorous theory is available. Moreover, some degree of mathematical intractability seems inevitable here. We recognize the probability that the general initial-value problem cannot be correctly posed (well set), because we know that in practice water waves may break - that is the motion may become turbulent and so lose continuous dependence on initial data. This aspect of the subject still remains largely mysterious, and reservations regarding it are needed to put any theoretical work on water waves into a properly scientific perspective".
From the few exact results which are available we mention the existence of periodic (Levi-Cevita, Krasovskii) and solitary wave (Friedrichs & Hyers) solutions, which solutions represent progressing waves.

For these reasons it is clear that one has studied approximations of (1.1) - (1.4). We shall deal with some of these approximations, and the appraisal of each of them for specific situations, in the next sections. But it will be clear from the onset that knowledge of various characteristic features of the exact problem is necessary to compare approximate equations with the exact equations. It is at this point that the Hamiltonian character of the exact equations can be given an important place. For this reason we shall study that aspect of the exact equations and some of its consequences in the rest of this section.

In section 3.5 it was shown that upon introducing

$$\psi(x,t) := \phi(x, y = 1 + \eta(x,t), t) \ , \tag{1.5}$$

the exact problem can be formulated as a Hamiltonian system in the variables η and ψ. The Hamiltonian is given by

$$h(\psi,\eta) = \int dx \, \frac{1}{2} \eta^2(x,t) + k(\psi,\eta) \tag{1.6}$$

wherein the *kinetic energy functional* is implicitely given in terms of ψ and η by

$$(1.7) \qquad k(\psi,\eta) := \int dx \int_0^{1+\eta(x,t)} dy \, \frac{1}{2}(\phi_x^2 + \phi_y^2) ,$$

where ϕ is the solution of the *linear* potential problem

$$(1.8) \qquad \begin{aligned} \phi_{xx} + \phi_{yy} &= 0 && 0 < y < 1 + \eta(x,t) \\ \phi_y &= 0 && y = 0 \\ \phi &= \psi && y = 1 + \eta(x,t) . \end{aligned}$$

From simple properties of the boundary value problem (1.8) one arrives at the following

OBSERVATION 6.1.1 *The exact hamiltonian* (1.6) *satisfies*

(i) $h(\psi,\eta) \geq 0$ *for every* ψ *and every* η *for which* $\eta(x) > -1 \quad \forall x \in R\ell$.

(ii) $h(\psi,\eta) = 0 \iff \eta \equiv 0$ and $\psi_x \equiv 0$.

(iii) $h(-\psi,\eta) = h(\psi,\eta)$ *and* $h(J\psi,J\eta) = h(\psi,\eta) \quad \forall\psi \; \forall\eta$, *where* J *is defined by* 5.(3.5).

(iv) $h(\psi+c,\eta) = h(\psi,\eta) \quad \forall\psi \; \forall\eta$ *for arbitrary constant* $c \in R\ell$.

Especially from this last property it follows that we are dealing with a potential problem. Therefore we define a new variable

$$(1.9) \qquad u(x,t) := \partial_x \psi(x,t).$$

With (1.5) it follows that

$$u(x,t) = \phi_x + \phi_y \cdot \eta_x = \underline{v} \cdot \underline{\tau} \qquad \text{at } y = 1 + \eta(x,t),$$

where $\underline{\tau} = (1, \eta_x)$ and $\underline{\tau} \cdot |\underline{\tau}|^{-1}$ is the unit tangent at the surface. This shows that $u \cdot |\underline{\tau}|^{-1}$ is the component of the *velocity tangent*

to the surface. Defining $\bar{h}(u,\eta) := h(\psi,\eta)$ we write Hamilton's equations in the variables u and η in the same way as in section 5.1. as

$$\partial_t \operatorname{grad} \bar{m}(u,\eta) = -\partial_x \operatorname{grad} \bar{h}(u,\eta) \tag{1.10}$$

where

$$\bar{m}(u,\eta) = \langle u,\eta \rangle . \tag{1.11}$$

Thus

$$\partial_t u = -\partial_x \frac{\delta \bar{h}}{\delta \eta} \tag{1.12}$$

$$\partial_t \eta = -\partial_x \frac{\delta \bar{h}}{\delta u} . \tag{1.13}$$

The functional $\bar{m}$ can be related to the *total horizontal momentum of the motion* M:

$$M := \int dx \int_0^{1+\eta} dy\ \phi_x = \int dx \left[\partial_x \int_0^{1+\eta} dy\ \phi - \eta_x . \psi\right] \tag{1.14}$$

$$= \langle \eta, \psi_x \rangle + \left[\int_0^{1+\eta} dy\ \phi - \eta . \psi \right]_{x=-\infty}^{x=\infty}.$$

As we shall restrict to motions which vanish for $|x| \to \infty$ i.e.

$$\begin{cases} \eta \to 0 \quad \text{for } |x| \to \infty \\ \phi(x,y;t) \to \psi(\pm\infty,t) \text{ for every } y \in (0,1), \text{ for } x \to \pm\infty \text{ (i.e. } \underline{v} \to 0 \text{ for } |x| \to \infty) \end{cases} \tag{1.15}$$

we find

$$M = \bar{m}(u,\eta) + \psi(\infty,t) - \psi(-\infty,t) = \bar{m}(u,\eta) + \int dx\ u(x,t) . \tag{1.16}$$

In terms of $\bar{h}$, the observations for h may be reformulated as follows

PROPOSITIONS 6.1.2. *The Hamiltonian* $\bar{h}$ *can be written as*

$$\bar{h}(u,\eta) = \frac{1}{2} \langle \eta, \eta \rangle + \bar{k}\ (u,\eta)$$

and satisfies

(i) $\bar{k}(u,\eta) \geq 0$ *and thus* $\bar{h}(u,\eta) \geq 0$ *for every* u *and every* η *for which* $\eta > -1$.

(ii) $\bar{h}(u,\eta) = 0 \Longleftrightarrow \eta = u \equiv 0$.

(iii) $h(-u,\eta) = \bar{h}(u,\eta)$ *and* $\bar{h}(Ju, J\eta) = \bar{h}(u,\eta) \qquad \forall u\ \forall \eta$.

Furthermore we shall prove the following formal (c.f. remark 6.1.4) results

LEMMA 6.1.3. *For the system* (1.10) *we have the following invariant integrals*

(i) $\bar{h}(u,\eta)$

(ii) $\bar{m}(u,\eta) = \langle u, \eta \rangle$

(iii) $\int dx.u(x,t)$

(iv) $\int dx.\eta(x,t)$.

Moreover, we have

(v) η *is a density which has conserved flux density* $\frac{\delta \bar{h}}{\delta u}$ *and*

$$M = \int dx\ \frac{\delta \bar{h}}{\delta u}\ .$$

PROOF: (i) holds because $\bar{h}$ does not depend explicitely on t and (ii) holds because $\bar{h}$ is translation invariant. Property (iii) follows from

$$\partial_t \psi = [\partial_t \phi + \phi_y . \eta_t\ ]_{y=1+\eta(x,t)}\ ,$$

together with (1.3), (1.4) and the definition of u. Noticing that (1.3) may be written as

$$\partial_t \eta = \nabla \phi . \underline{\eta} \qquad \text{at } y = 1 + \eta \tag{1.17}$$

where $\underline{n} = (-\eta_x, 1)$ is the normal to the surface, and applying Gauss' theorem

$$\int_{-\infty}^{\infty} dx \int_{0}^{1+\eta} dy\ \Delta\phi = \int_{-\infty}^{\infty} dx\ \nabla\phi.\underline{n}\Big]_{y=1+\eta} - \int_{-\infty}^{\infty} dx\ \phi_y\Big]_{y=0} + \int_{0}^{1} dy\ \phi_x\Big]_{x=-\infty}^{x=\infty},$$

it follows with (1.2) and (1.15) that property (iv) holds. Property (v) follows from the fact that

$$\int \frac{\delta \bar{h}}{\delta u}\, dx = M = \int dx \int_{0}^{1+\eta} \phi_x\, dy. \tag{1.18}$$

which is an invariant integral because of (1.16) and properties (ii) and (iii) just proved. Relation (1.18) can be obtained as follows:

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \{ \bar{h}(u+\varepsilon,\eta) - \bar{h}(u,\eta) \} = \int dx\ \frac{\delta \bar{h}}{\delta u}(u,\eta),$$

and

$$\bar{h}(u+\varepsilon,\eta) = \bar{h}(u,\eta) + \varepsilon \int dx \int dy\ \nabla\phi.\nabla x + O(\varepsilon^2),$$

because $\phi + \varepsilon x$ is the solution of (1.8) which corresponds to $u + \varepsilon$ if ϕ is the solution which corresponds to u. This proves the lemma. □

REMARK 6.1.4. As no existence and regularity results are available for the initial value problem for the system under consideration, phrases such as: "the functional $g(u,\eta)$ is an invariant integral" have to be understood in the following sense. Let (u_o, η_o) be an initial value for which $g(u_o,\eta_o)$ is defined. Then, if there exists a solution $(u(t),\eta(t))$ of the underlying problem for $0 \leq t \leq t_o$ for which $u(0) = u_o$, $\eta(0) = \eta_o$ and for which $g(u(t), \eta(t))$ is defined in a meaningful way, then $g(u(t),\eta(t)) = g(u_o,\eta_o)$ for every t, $0 \leq t \leq t_o$.

REMARK 6.1.5. *Stability of the equilibrium solution* $u = \eta \equiv 0$. Proposition 6.1.2. (i), (ii) imply that $u = \eta \equiv 0$ is an absolute

minimum of the functional $\bar{h}$ for the class of functions (u,η) for which $\eta > -1$. Loosely speaking, as $\bar{h}$ is an invariant integral, this implies that the equilibrium solution $u = \eta \equiv 0$ of (1.10) is stable. However, because of the possible non-existence and non-regularity of solutions, such a statement can hardly be given a sensible meaning. Therefore, let us briefly indicate in what sense the equilibrium solution can be called stable. In general terms, consider a first order evolution equation $\underline{u}_t = E(\underline{u})$, where $\underline{u}$ may be a vector valued state variable, and suppose $E(\underline{0}) = \underline{0}$, such that $\underline{u} \equiv \underline{0}$ is an equilibrium solution. Then $\underline{u} \equiv \underline{0}$ is said to be *stable* (in the sense of Lyapunov/Movchan) with respect to (initial norm) $||\ ||_i$ and (evolution norm) $||\ ||_e$ if, given arbitrary small $\varepsilon > 0$, there exists $\delta > 0$ such that for arbitrary initial value $\underline{u}_o$, with $||\underline{u}_o||_i < \delta$, there exists a solution $\underline{u}(t)$ for every $t > 0$ with $\underline{u}(0) = \underline{u}_o$, such that $||\underline{u}(t)||_e < \varepsilon$ for every $t > 0$:

$$(1.19) \qquad \forall \varepsilon>0 \exists \delta>0 \forall \underline{u}_o \qquad ||\underline{u}_o||_i < \delta \Rightarrow ||\underline{u}(t)||_e < \varepsilon,\ \forall t \geq 0.$$

If S is some set of initial data, we shall say that $\underline{u} = \underline{0}$ is stable in the *restricted* sense with respect to S and the norm $||\ ||_i$ and $||\ ||_e$ if property (1.19) holds for every $u_o \in S$:

$$(1.20) \qquad \forall \varepsilon>0 \exists \delta>0 \forall \underline{u}_o \in S \quad ||\underline{u}_o||_i < \delta \Rightarrow ||\underline{u}(t)||_e < \varepsilon \quad \forall t > 0.$$

It must be remarked that in general this restricted definition is a very severe weakening of the usual definition of stability. But for the problem under consideration it seems to be impossible to avoid this weakening if one wants to formulate the intuitive idea of stability. For the water wave problem we define the following sets of initial data

$$(1.21) \quad S := \{(u_o,\eta_o) \mid \bar{h}(u_o,\eta_o) \text{ is defined; there exists a solution } (u(t),\eta(t)) \text{ for all } t \geq 0 \text{ with } (u(0),\eta(0)) = (u_o,\eta_o) \text{ and } \bar{h}(u(t),\eta(t)) = \bar{h}(u_o,\eta_o)\ \forall t \geq 0\},$$

and for $\delta_+ > 0$ and $0 < \delta_- < 1$:

(1.22) $S_{\delta_\pm} = \{(u_o,\eta_o) \in S \mid \eta(t) \in C^o(R\ell) \text{ and } -\delta_- < \eta(t) < \delta_+ \ , \forall t>o\}$.

Now observe that

(1.23) $$k(u,\delta_-) \leq k(u,\eta) \leq k(u,\delta_+) \qquad \forall (u,\eta) \in S_{\delta_\pm}$$

and that $k(u,\delta_\pm)$ can be explicitely found to be

(1.24) $$k(u,\delta) = \frac{1}{2} < u, R_\delta u > ,$$

where R_δ is a selfadjoint, positive pseudo-differential operator with symbol

(1.25) $$\hat{R}_\delta(k) = \frac{\tanh k\bar{\delta}}{k\bar{\delta}} \quad \text{where} \quad \begin{cases} \bar{\delta} = 1 + \delta_+ & \text{if } \delta = \delta_+ \\ \bar{\delta} = 1 - \delta_- & \text{if } \delta = \delta_- \end{cases}$$

(For $\delta = 1$ this will be proved in the next section).

With (1.23) it follows that

(1.26) $$\frac{1}{2}||\eta||^2 + \frac{1}{2} <u,R_{\delta_-} u> \leq \bar{h}(u,\eta) \leq \frac{1}{2}||\eta||^2 + \frac{1}{2} <u,R_{\delta_+} u>.$$

From these results it is easily seen that we may formulate the intuitive idea of stability of the equilibrium solution in the following way :
For arbitrary $\delta_+ > 0$, $0 < \delta_- < 1$, the equilibrium solution $u = \eta \equiv 0$ is stable with respect to the set $S_{\delta_\pm}$ and the norms $|| \ ||_i$ and $|| \ ||_e$, where

(1.27) $$||(u,\eta)||_i^2 := ||\eta||^2 + <u,R_{\delta_+} u>$$

(1.28) $$||(u,\eta)||_e^2 := ||\eta||^2 + <u,R_{\delta_-} u> .$$

REMARK 6.1.6. Proposition 6.1.2. (iii) shows that $\bar{h}$ satisfies

conditions 5.(3.17), (3.18) of theorem 5.3.4. Therefore, if (u_o,η_o) is an initial value for which

$$Ju_o = -u_o, \quad J\eta_o = \eta_o ,$$

the corresponding solution $(u(t), \eta(t))$ satisfies

$$Ju(t) = -u(t) , \quad J\eta(t) = \eta(t) \tag{1.29}$$

as long as (i.e. for the possible restricted set of t for which) this initial value problem has a unique solution. That such solutions exist is physically plausible because in the underlying system there is no preferred direction of propagation. For every solution which satisfies (1.29) we have

$$\int u_o \, dx = \int u(t)dx = 0$$

$$\bar{m}(u,\eta) = \bar{m}(u_o,\eta_o) = 0,$$

and consequently (c.f. (1.16)) for these solutions the total horizontal momentum of the motion M is identically zero:

$$M = 0.$$

REMARK 6.1.7. With remark 5.4.2. (iv) it follows from property 6.1.3. (v) that the functional

$$C(u,\eta) := \int \{x\eta - t \frac{\delta h}{\delta u}(u,\eta)\} dx = \int dx . x\eta - tM$$

is an invariant integral. If $\int \eta_o dx = \int \eta(t)dx \neq 0$, the centre of gravity of the free surface displacement η is propagated with constant velocity V:

$$V(u,\eta) = \frac{M}{\int dx \, \eta} \quad , \quad \partial_t V(u,\eta) = 0$$

(In particular, for the symmetric free surface displacements which satisfy (1.29) this velocity is identically zero). The same result was

derived by Benjamin & Mahony [49] for more general motion of the fluid (not necessarily irrotional, three-dimensional fluid motion).

6.2. APPROXIMATE MODELS.

In this section we shall consider some approximations for the exact set of equations (1.1) - (1.4). The exact model is difficult in that it combines two important aspects: non-linearity in the equations at the free surface and an essentially two-dimensional behaviour in the interior of the fluid (Variations of the field variables in the y-direction has a dispersive effect in the x-direction). Introducing two dimension free parameters to measure these aspects, let ε be a measure for the height of the waves : $\eta = O(\varepsilon)$ and let $\frac{1}{\lambda} = O(\mu)$, wherein λ is a characteristic length of the wave phenomenon in x-direction. (Note that we have taken the undisturbed height to be normalized to 1). *Stokes' number* is defined as

$$St := \frac{1}{\eta} \cdot \frac{1}{\lambda^2} \quad ,$$

such that $St = O(\mu^2.\varepsilon^{-1})$ is a measure for the relative importance of the two aspects: for infinitely long waves, for which variations in the y-direction can be ignored (tidal waves), there is no dispersion and $St = 0$; for infinitesimal low waves the elevation from the horizontal surface is ignored: $St = \infty$ and the equations are linear. We shall describe these two limiting cases presently.

A more interesting model, studied already by Boussinesq around 1870, accounts for both the non-linearity and dispersion in the same order of magnitude, i.e. $O(\varepsilon) = O(\mu^2)$ such that $St = O(1)$. We shall call this model, which is meant to describe what may be called the class of "fairly long, fairly low" waves, the *Boussinesq model*. In literature, several approximate equations for this model are known. Usually , these equations are obtained by expanding the exact equations in terms of ε (and $\mu^2 = O(\varepsilon)$) after which higher order terms are struck (c.f. e.g. Whitham [45, p. 464-466] as a convenient reference for a typical example of this procedure). Such a process then leads to a set of equations which approximate the

exact equations to the desired order of ε. In fact, several approximations can be written down, all of which have the same formal status of being a good approximation in the sense that they approximate the exact equations correctly to the desired order of ε. However, the distinct equations will have in general rather different mathematical properties. As the underlying problem is too difficult to admit rigorous mathematical statements about the validity of each of these approximations, it is not possible to prefer one approximation above the others on these grounds. [One would like to have a theory which gives a meaning to and proves such statements as: for every solution of the exact equations from a certain subset (the class of fairly long, fairly low waves), there exists a solution of the approximate equations such that the "difference" of these solutions is "small"]. Therefore, the best one can do is to construct approximations which have the same characteristic features as the exact model. In that respect, Broer [33] emphasized the Hamiltonian character of the exact model by looking for approximations which are also Hamiltonian systems. Then one looks for approximations of the exact Hamiltonian: the corresponding Hamilton equations then approximate the exact set of equations and resemble these in their Hamiltonian aspect. We shall briefly outline these ideas for the Boussinesq model, but we first investigate the two limiting models.

Linearized theory.

In the linearized theory the exact hamiltonian $\bar{h}(u,\eta)$ is approximated to

$$\bar{h}(u,\eta) := \frac{1}{2} <\eta,\eta> + \bar{k}(u,\eta \equiv 0).$$

In that case $\bar{k}$ can be explicitely found:

$$k(\psi,\eta\equiv 0) = \int_{-\infty}^{\infty} dx \int_{0}^{1} dy \ \frac{1}{2}(\nabla\phi)^2 = \frac{1}{2} \int_{-\infty}^{\infty} dx \ \psi(x).\ \phi_y(x,y=1,t)$$

where ϕ is the solution of

$$\begin{aligned} \Delta\phi &= 0 \qquad & 0<y<1 \\ \phi_y &= 0 \qquad & y = 0 \\ \phi &= \psi \qquad & y = 1 \end{aligned}$$

With Fouriertransform technique (writing κ for the transformed variable to distinguish from the kinetic energy functional) it is easily found that

$$\hat{\phi}(\kappa,y) = \hat{\psi}(\kappa).\ \frac{\cosh\ \kappa y}{\cosh\ \kappa}\ ,$$

such that $\hat{\phi}_y(\kappa,1) = \hat{\psi}(\kappa).\kappa.\tanh\ \kappa$ and thus

$$k(\psi,\eta\equiv 0) = \frac{1}{2}\int_{-\infty}^{\infty} d\kappa\,|\hat{\psi}(\kappa)|^2.\kappa.\tanh\ \kappa$$

which may be written as

$$k(\psi,\eta\equiv 0) = \frac{1}{2}\,\langle\psi_x,\ R\ \psi_x\rangle$$

wherein R is a pseudo-differential operator with symbol

$$\hat{R}(\kappa) = \frac{\tanh\ \kappa}{\kappa} \tag{2.1}$$

(Note that R is a selfadjoint, positive operator). Hence we find

$$\bar{h}(u,\eta) = \frac{1}{2}\,\langle\eta,\eta\rangle + \frac{1}{2}\,\langle u,Ru\rangle\ , \tag{2.2}$$

and the corresponding linear equations are

$$\begin{aligned} \partial_t u &= -\partial_x \eta \\ \partial_t \eta &= -\partial_x Ru \end{aligned} \tag{2.3}$$

Note that the equilibrium solution $u = \eta = 0$ of (2.3) is stable, as the dispersion relation is

$$\omega^2 = \kappa.\tanh\ \kappa \geq 0 \qquad \forall\kappa\in\mathbb{R}\ . \tag{2.4}$$

This reflects that the approximate Hamiltonian (2.2) is positive (as R is a positive operator).

Tidal waves.

If the vertical fluid motion is completely ignored, we have

$$k(\psi,\eta) = \frac{1}{2}\int_{-\infty}^{\infty} dx\ (1+\eta).\ \psi_x^{\,2}\ ,$$

and thus

$$\bar{h}(u,\eta) = \int dx\ \{\frac{1}{2}\eta^2 + \frac{1}{2}u^2 + \frac{1}{2}\eta u^2\}\ , \tag{2.5}$$

with corresponding equations

$$\begin{aligned} \partial_t\, u &= -\partial_x\ (\eta + \tfrac{1}{2}u^2) \\ \partial_t\, \eta &= -\partial_x\ ((1+\eta)u)\ . \end{aligned} \tag{2.6}$$

This is a hyperbolic system of non-linear equations for which the standard theory may be applied (c.f. Whitham [45, p. 456]). In particular, most solutions of (2.6) will "break" and to prevent multivaluedness, discontinuities (shocks) have to be introduced.

Boussinesq model.

For the Boussinesq model, $\mu^2 = O(\varepsilon)$, Broer[50] derived an approximate Hamiltonian whose density is correct up to and including third order of ε. Although we shall essentially use approximations of this order, we start with an approximation which is correct up to and including fourth order, a result derived by Timmers [51] (c.f. also Broer et al [34])

$$\bar{h}(u,\eta) = \int_{-\infty}^{\infty} dx\ \{\frac{1}{2}u.Ru + \frac{1}{2}\eta^2 + \frac{1}{2}\eta u^2 - \frac{1}{2}\eta u_x^{\,2} + O(\varepsilon^5)\}\ , \tag{2.7}$$

wherein R is the pseudo-differential operator given by (2.1).
In the following we shall restrict ourselves to (equivalent forms of) the approximate Hamiltonian

$$\bar{h}(u,\eta) = \int dx\ \{\frac{1}{2}u\cdot Ru + \frac{1}{2}\eta^2 + \frac{1}{2}\eta u^2\}\ . \tag{I}$$

The equations corresponding to this approximation are given by

$$\partial_t \eta = -\partial_x (Ru + \eta u)$$
$$\partial_t u = -\partial_x (\eta + \frac{1}{2} u^2) \quad . \tag{2.8}$$

Because of the approximate character of (I) it is allowed to take for R any pseudo-differential operator with symbol which agrees with (2.1) up to and including κ^2 for $\kappa \to 0$. As

$$\hat{R}(\kappa) = 1 - \frac{1}{3} \kappa^2 + 0(\kappa^4) \quad ,$$

a first choice would be to take

$$\text{(a)} \qquad R = I + \frac{1}{3} \partial_x^2 \quad .$$

However, this operator is not a positive one, and the leading order terms in the Hamiltonian (I) are not positive (hence, the linearized equations of (2.8) have as dispersion relation:

$$\omega^2 = \kappa^2(1 - \kappa^2) \quad ,$$

and the equilibrium solution $u = \eta \equiv 0$ is not stable).
Therefore it is better to approximate R by a positive selfadjoint operator. In that case one may write

$$\text{(b)} \qquad R = D^{-2} \quad ,$$

wherein D is a positive selfadjoint operator. For instance, one could take

$$\text{(b1)} \qquad \hat{R}(\kappa) = (1 + \frac{1}{3} \kappa^2)^{-1} \quad , \quad D = (I - \frac{1}{3} \partial_x^2)^{\frac{1}{2}}$$
$$\text{(b2)} \qquad \hat{R}(\kappa) = (1 + \frac{1}{6} \kappa^2)^{-2} \quad , \quad D = I - \frac{1}{6} \partial_x^2 \quad .$$

In these cases it is often convenient to apply the transformation

$$v := D^{-1} u \quad , \tag{2.9}$$

such that with

$$\text{(2.10)} \qquad \bar{\bar{m}}(v,\eta) = \langle Dv,\eta\rangle \quad ; \quad \bar{\bar{h}}(v,\eta) := \bar{h}(u,\eta)$$

the equations are given by

$$\text{(2.11)} \qquad \partial_t \operatorname{grad} \bar{\bar{m}}(v,\eta) = -\partial_x \operatorname{grad} \bar{\bar{h}}(v,\eta) \ .$$

For (b1,2) the transformed Hamiltonian (I) is given by

$$\text{(Ib).} \qquad \bar{\bar{h}}(v,\eta) = \int dx \ \{ \tfrac{1}{2} v^2 + \tfrac{1}{2} \eta^2 + \tfrac{1}{2} \eta (Dv)^2 \}.$$

Expressed in terms of η and v we shall give some alternative approximate Hamiltonians which all agree with (I) in the desired order of (i.e. differences are of order ε^4). After that we briefly comment on their specific properties and differences between them.

$$\text{(II)} \quad \bar{\bar{h}}(v,\eta) = \int dx \{ \tfrac{1}{2} v^2 + \tfrac{1}{2} \eta^2 + \tfrac{1}{2}\eta \ (Dv)^2 + c.(Dv)^4 \} \ , \text{ with } c > \tfrac{1}{8}$$

$$\text{(IIIa)} \quad \bar{\bar{h}}(v,\eta) = \int dx \{ \tfrac{1}{2} v^2 + \tfrac{1}{2}\eta^2 + \tfrac{1}{2} \eta v^2 \}$$

$$\text{(IIIb)} \quad \bar{\bar{h}}(v,\eta) = \int dx \{ \tfrac{1}{2} v^2 + \tfrac{1}{2} \eta^2 + \tfrac{1}{2} \eta v^2 + c v^4 \} \qquad \text{, with } c > \tfrac{1}{8}$$

$$\text{(IVa)} \quad \bar{\bar{h}}(v,\eta) = \int dx \{ \tfrac{1}{2} v^2 + \tfrac{1}{2} \eta^2 + \tfrac{1}{2} D^{-1}\eta . v^2 \}$$

$$\text{(IVb)} \quad \bar{\bar{h}}(v,\eta) = \int dx \{ \tfrac{1}{2} v^2 + \tfrac{1}{2} \eta^2 + \tfrac{1}{2} D^{-1}\eta . v^2 + c.\big(D^{-1}(v^2)\big)^2 \} .$$

EXISTENCE AND REGULARITY.

For the equations(2.11), with $\bar{\bar{h}}$ any approximation given above, it is possible, using a contraction mapping principle, to prove the existence of a solution (v,η) corresponding to arbitrary initial data (v_o,η_o) over a *limited range of time*. The corresponding solution is in some sense as regular as the initial data, and the time interval over which the solution can be proved to exist depends only on certain norms of the initial data. However, in general this process cannot be continued to prove the existence of the solution for all time. This is possible only if some *a-priori* estimates are available. The Hamiltonians given by (I), (II) and (III), which are invariant

integrals as long as the solution is known to exist, are not suited to serve as *a-priori* estimates. However, as was shown by Bona & Smith [52], with $D = (I - \frac{1}{3}\partial_x^2)^{\frac{1}{2}}$ the Hamiltonian (IVa) can, for a restricted set of initial data, be used as such an estimate. Their results can be formulated as follows (in their paper they used a variable $\bar{\eta}$ which is related to η by $\eta = D\bar{\eta}$, such that $\bar{\eta}$ does *not* represent the wave height): If $v_o, \eta_o \in C^2(R^2) \cap L_2(R^2)$ such that

$$\bar{\bar{h}}(v_o,\eta_o) \leq \frac{2}{\sqrt{3}} \quad , \quad \text{and } D^{-1}\eta_o > -1 \ ,$$

then there exists a unique solution (v,η) which has (v_o,η_o) as initial data and which satisfies for every $t > 0$:

$$\partial_t^j v(t),\ \partial_t^j \eta(t) \in C^2(R^2) \cap L_2(R^2) \quad \text{for every } j = 0, 1, 2, \ldots$$

$$\bar{\bar{h}}(v(t),\eta(t)) = \bar{\bar{h}}(v_o,\eta_o) \quad , \quad D^{-1}\eta > -1 \ .$$

It can be shown that the same qualitative results hold for (IVb). For $c = \frac{1}{8}$, the equations for the steady states of this approximation are particularly simple and some interaction problems for these particular solutions were studied by Valkering [53].

INVARIANT INTEGRALS AND CONSERVED FLUX PROPERTY.

The equations corresponding to the approximate Hamiltonians given above all have the same four invariant integrals as given in lemma 6.1.3., of course with the approximate Hamiltonian replacing the exact one. However, it may be verified that only (I) leads to equations for which η has a conserved flux.

POSITIVITY AND STABILITY RESULTS.

For the approximation (Ib) it is not possible to give positivity statements (valid for arbitrary function u) if solutions for which η may be negative are allowed. This functional can be supplied with a higher order term (!) to give (II), which functional is non-negative. Moreover, as for arbitrary constants c and b we have

$$\bar{\bar{h}}(v,\eta) = \tfrac{1}{2}||v||^2 + \tfrac{1}{2}(1-b)||\eta||^2 + \tfrac{1}{2}b||\eta + \tfrac{1}{2b}(Dv)^2||^2 + (a - \tfrac{1}{8b})||(Dv)^2||^2,$$

if c satisfies $c > \frac{1}{8}$, a value of b can be found such that $\frac{1}{8a} < b < 1$. From this it follows that $v = \eta \equiv 0$ is a stable equilibrium solution with respect to the set S (c.f. (1.21)) and norms

$$||(v,\eta)||_{i,e}^2 := ||v||^2 + ||\eta||^2 + ||(Dv)^2||^2.$$

The functional $\bar{\bar{h}}$ given by (IIIa) satisfies proposition 6.1.2. Moreover, the solution $v = \eta \equiv 0$ is stable with respect to sets $S_{\delta\pm}$ with $0 < \delta_- < 1$ (c.f. (1.22)) and norms

$$||(v,\eta)||_{i,e}^2 := ||v||^2 + ||\eta||^2 .$$

For (IIIb), $\bar{\bar{h}} \geq 0$ for arbitrary solutions and $v = \eta \equiv 0$ is stable with respect to S with norms

$$||(v,\eta)||_{i,e}^2 := ||v||^2 + ||\eta||^2 + ||v^2||^2.$$

For (IVa) we have $\bar{\bar{h}} \geq 0$ as long as $D^{-1}\eta > -1$. Because of the available existence and regularity result, for this approximation the equilibrium solution $v = \eta \equiv 0$ is stable (in the usual sense) with norms

$$||(v,\eta)||_{i,e}^2 = ||v||^2 + ||\eta||^2.$$

For (IVb) we have $\bar{\bar{h}} \geq 0$ for arbitrary solution if $c > \frac{1}{8}$, and the same stability result as for (IVa).

REMARK 6.2.1. If it is desired to have the number of derivatives appearing in the resulting equations to be as small as possible, another transformation than (2.9) may be applied. Therefore, assume

$$R^{-1} = T^* T \text{ with } T \text{ invertible},$$

and define a new variable

$$w := T^{-1}u.$$

For the operator R given in (b_1) one may take

$$T = I + \frac{1}{\sqrt{3}}\partial_x .$$

Then

$$\tilde{\tilde{m}}(w,\eta) = \langle Tw,\eta\rangle,$$

and a suitable Hamiltonian (akin to (IIIa) but in the new variable w) would be:

$$\tilde{\tilde{h}}(w,\eta) = \int dx \quad [\frac{1}{2} w^2 + \frac{1}{2} \eta^2 + \frac{1}{2} \eta w^2].$$

The resulting equations are then found to be

$$\partial_t Tw = -\partial_x(\eta + \frac{1}{2} w^2)$$

$$\partial_t T^*\eta = -\partial_x(w + \eta w) \ .$$

6.3. FIRST ORDER EQUATIONS.

In the foregoing section we gave several approximate Hamiltonians for the description of surface waves. Here we shall show how we can apply the ideas described in section 5.1. to obtain the first order equations for unidirectionally propagative waves. In the first instance we shall formally ignore the interaction terms to obtain a complete separation. Afterwards, some remarks about the validity of the separation are made.

Consider the Hamiltonian (Ia) and perform the transformation

(3.1.)
$$\eta = \alpha + \beta$$
$$u = \alpha - \beta$$

Then the equations for α and β are given by (c.f. 5.(1.11))

(3.2) $$\partial_t \text{ grad } \tilde{m}(\alpha,\beta) = -\partial_x \text{ grad } \tilde{h}(\alpha,\beta) \ ,$$

where

(3.3) $$\tilde{m}(\alpha,\beta) = \langle\alpha,\alpha\rangle - \langle\beta,\beta\rangle \ ,$$

(3.4) $$\tilde{h}(\alpha,\beta) = e(\alpha) + e(\beta) + i(\alpha,\beta)\ ,$$

with

(3.5) $$e(\alpha) = \int dx\ [\tfrac{1}{2}\,\alpha.(I + R)\alpha + \tfrac{1}{2}\,\alpha^3],$$

(3.6) $$i(\alpha,\beta) = \int dx\ [\alpha.(I - R)\beta - \tfrac{1}{2}\,\alpha\beta(\alpha+\beta)].$$

Explicitely:

(3.7) $$\partial_t\alpha = -\partial_x(\alpha + \tfrac{1}{6}\alpha_{xx} + \tfrac{3}{4}\alpha^2 - \tfrac{1}{6}\beta_{xx} - \tfrac{1}{2}\alpha\beta - \tfrac{1}{4}\beta^2),$$

(3.8) $$\partial_t\beta = \partial_x(\beta + \tfrac{1}{6}\beta_{xx} + \tfrac{3}{4}\beta^2 - \tfrac{1}{6}\alpha_{xx} - \tfrac{1}{2}\alpha\beta - \tfrac{1}{4}\alpha^2).$$

Ignoring the interaction functional $i(\alpha,\beta)$ there result two uncoupled equations:

(3.9) $$\partial_t\alpha = -\partial_x(\alpha + \tfrac{1}{6}\,\alpha_{xx} + \tfrac{3}{4}\,\alpha^2),$$

(3.10) $$\partial_t\beta = \partial_x(\beta + \tfrac{1}{6}\,\beta_{xx} + \tfrac{3}{4}\,\beta^2).$$

each of which is a KdV equation (apart from some scaling factors).

Another interesting set of equations is obtained from the Hamiltonian (IIIa) if we perform a somewhat different transformation, viz.

(3.11) $$D^{-1}\begin{matrix} u \\ v \end{matrix} \begin{matrix} = \eta = \alpha + \beta \\ = \alpha - \beta \end{matrix} \qquad \text{where } D = I - \tfrac{1}{6}\,\partial_x^{\,2}.$$

Then the equations for α and β are given by (3.2), where now

(3.12) $$\tilde{m}(\alpha,\beta) = \langle D\alpha,\alpha\rangle - \langle D\beta,\beta\rangle\ ,$$

and $\tilde{h}$ as in (3.4) with

(3.13) $$e(\alpha) = \int dx\ [\alpha^2 + \tfrac{1}{2}\,\alpha^3]\ ,$$

(3.14) $$i(\alpha,\beta) = \int dx \ [-\frac{1}{2}\alpha\beta(\alpha+\beta)] .$$

The resulting equations are

(3.15) $$D\partial_t\alpha = -\partial_x(\alpha + \frac{3}{4}\alpha^2 - \frac{1}{2}\alpha\beta - \frac{1}{4}\beta^2)$$

(3.16) $$D\partial_t\beta = \partial_x(\beta + \frac{3}{4}\beta^2 - \frac{1}{2}\alpha\beta - \frac{1}{4}\alpha^2),$$

and, ignoring the interactionfunctional:

(3.17) $$D\partial_t\alpha = -\partial_x(\alpha + \frac{3}{4}\alpha^2)$$

(3.18) $$D\partial_t\beta = \partial_x(\beta + \frac{3}{4}\beta^2) ,$$

which are BBM equations for α and β.

For the class of solutions which is described by the Boussinesq model, the interaction functional has a density which is of order ε^3. Therefore, in general the set of coupled equations such as (3.15), (3.16) is not equivalent (in the desired approximation) to the set of uncoupled equations (3.17), (3.18). This will only be the case for those solutions for which this interaction density is of order ε^4. However it is easily seen that there are no solutions for which this is true for *all* values of time. For instance, consider initial data

(3.19) $$(\alpha_o,\beta_o) \quad \text{with} \quad \beta_o \equiv 0$$

(which correspond to initial data $v_o = \eta_o$ in the original vairables). Then

$$i(\alpha_o,\beta_o) = 0.$$

But, due to the terms $-\partial_x \frac{1}{4}\alpha^2$ in the right hand side of (3.16) (or $-\partial_x(\frac{1}{6}\alpha_{xx} + \frac{1}{4}\alpha^2)$ in (3.8)), the initial α-mode will generate a β-mode (there exists no non-trivial solution with $\beta = 0$ for all t), such that $i(\alpha,\beta)$ becomes non-zero for $t > 0$. Nevertheless, heuristically speaking, the interaction density will be of order ε^4 for such a solution for a finite, but sufficiently small, interval of time.

It is possible to investigate this matter more precisely for the equations corresponding to the Hamiltonian (IIIa). In fact, considering the interaction terms in (3.15) as source terms for the BBM equation (3.17), and using the results and ideas from BBM [39], one can prove the following

LEMMA 6.3.1. *There exist positive constants* c *and* d *such that* (i) *for arbitrary* $\alpha_o \in C^2(R\ell) \cap H^1(R\ell)$ *there exists a unique solution* (α,β) *of* (3.15) (3.16) *over the finite time interval* $[0,T]$ *where*

$$T < \frac{c}{a_o} \quad , \quad \textit{with } a_o := ||\alpha_o||_1 \ ,$$

with $\alpha(0) = \alpha_o \ , \quad \beta(0) = 0,$

and $\partial_t^j \alpha(t), \ \partial_t^j \beta(t) \in C^2(R\ell) \cap H^1(R\ell)$ *for every* $j = 0,1,2,..$

and all $t \in (0,T)$;

(ii) *for such a solution we have*

$$(3.20) \quad \sup_{0\le t\le T} \left| \ i(\alpha(t),\beta(t))\right| \le d.a_o^4.T, \ \textit{for} \quad T < \frac{c}{a_o} \ .$$

It is also possible to express these results in a somewhat different way. Therefore, let $\bar{\alpha},\bar{\beta}$ be solutionsof (3.17) and (3.18) and put

$$(3.21) \qquad \bar{\eta} := \bar{\alpha} + \bar{\beta}, \qquad \bar{v} := \bar{\alpha} - \bar{\beta} \ .$$

Then one can study the differences

$$(3.22) \qquad \eta - \bar{\eta} = (\alpha - \bar{\alpha}) + (\beta - \bar{\beta}) \ , \quad v - \bar{v} = (\alpha - \bar{\alpha}) - (\beta - \bar{\beta})$$

for solutions with the same initial values.

LEMMA 6.3.2. *For solutions corresponding to initial data* (3.19), *we have* $\bar{\eta} = \bar{v} = \bar{\alpha}$ *and*

(3.23) $$\sup_{0\leq t\leq T} [||\eta(t) - \bar{\eta}(t)||_1 + ||v(t) - \bar{v}(t)||_1] \leq \text{const.} a_o^2 T$$

for $T \leq \text{const.} \frac{1}{a_o}$ *where* $a_o = ||\alpha_o||_1$.

Unfortunately, these rigorous results are of very little practical importance. In fact, it would be far more interesting to derive results concerning the relation between solutions of (3.15), (3.16) and those of the uncoupled set of equations (3.17), (3.18) *in the limit for* $t \to \infty$. However, results in this direction seem to be impossible to obtain in an analytical way.

To conclude we note that according to the theory of section 5.3., the γ-equation, with the aid of which "symmetrical" initial value problems can be evaluated, is given for the approximate Hamiltonian (IIIa) by

(3.24) $$D\partial_t \gamma = -\partial_x(\gamma + \frac{3}{4}\gamma^2 - \frac{1}{2}\gamma J\gamma - \frac{1}{4}(J\gamma)^2).$$

The behaviour of solutions of this equation for large values of t is as difficult to investigate as for the set (3.15), (3.16). For several classes of initial data γ_o, it is possible to derive some approximations for the resulting solutions for small values of t. For instance, if $\gamma_o = J\gamma_o$, which represents a symmetrical elevation η_o of the water level without initial velocity : $v_o \equiv 0$, the resulting solution may be compared for small t with the solution of the *linear* equation

$$D\partial_t \gamma = -\partial_x \gamma.$$

REFERENCES.

1. RUDIN, W., *Functional Analysis* (Mc. Graw-Hill, New York, 1973).
2. BROWN, A.L. & A. PAGE, *Elements of Functional Analysis* (van Nostrand, London, 1970)
3. LJUSTERNIK, L.A. & W.I. SOBOLEW, *Elemente der Functionalanalysis* (Akademie Verlag, Berlin, 1968)
4. DUNFORD, N. & J.T. SCHWARTZ, *Linear Operators*, part I,(Interscience, New York, 1967)
5. TREVES, F., *Basic Linear Partial Differential Equations* (Academic Press, New York, 1975)
6. SOBOLEW, S.L., *Einige Anwendungen der Funktionalanalysis auf Gleichungen der Mathemathischen Physik* (Akademie Verlag, Berlin, 1964)
7. VAINBERG, M.M., *Variational Method and Method of Monotone Operators* (Wiley, New York, 1973)
8. DIEUDONNÉ, J., *Foundations of Modern Analysis*, (Academic Press, New York, 1960)
9. VAINBERG, M.M., *Variational Methods for the Study of Nonlinear Operators* (Holden-Day Inc., San Francisco, 1964).
10. ROCKAFELLAR, R.T., *Convex Analysis* (Princeton University Press, Princeton, 1970)
11. EKELAND, I. & R. TEMAM, *Convex Analysis and Variational Problems* (North Holland, Amsterdam, 1976)
12. LUSTERNIK, L.A., *On conditional extrema of functionals;* Math. Sb. 41 no. 3 (1934) 390-401 (in Russian)
13. MAURIN, K., *Calculus of Variations and Field Theory*, part I, Aarhus Universitet Lecture Note Series, no. 34, 1972.
14. CURTAIN, R.F. & A.J. PRITCHARD, *Functional analysis in modern applied mathematics* (Academic Press, New York, 1977)
15. KLÖTZLER, R., *Mehrdimensionale Variationsrechnung* (Birkhäuser Verlag, Basel, 1970)
16. ZUBOV, V.I., *Methods of A.M. Lyapunov and their application* (Noordhoff, Groningen, 1964)

17. DIEKMANN, O. & N.M. TEMME, (eds.) *Nonlinear Diffusion Problems*, MC Syllabus 28 (Mathematisch Centrum, Amsterdam 1976)
18. GELFAND, I.M. & S.V. FOMIN, *Calculus of Variations* (Prentice-Hall Inc., Englewood Cliffs, 1963)
19. STAKGOLD, I., *Branching of solutions of nonlinear equations*, SIAM-review 13 (1971) 289.
20. REISS, E.L., *Column Buckling-An elementary example of bifurcation*, in: *Bifurcation Theory and Nonlinear Eigenvalue Problems*, Keller, J.B. & S. Antman eds. (Benjamin, New York, 1969)
21. VAN DER VARST, P., *De elastische lijn*, Afstudeerscriptie T.H.E., vakgroep Theoretische Natuurkunde, 1977.
22. CHAFEE, N. & E. INFANTE, *A bifurcation problem for a nonlinear parabolic equation*, Appl. Anal. 4 (1974)
23. HENRY, D., *Geometric Theory of Semilinear Parabolic Equations*, to be published.
24. AMBROSETTI, A. & P.H. RABINOWITZ, *Dual variational methods in critical point theory and applications*, J. Funct.Anal., 14 (1973) 349.
25. LIN, C.C., in: *Liquid Helium*, Proc. Int. School of Physics, Course XXI, (Academic Press, New York, 1963)
26. VAN GROESEN, E.W.C., L.J.F. BROER & J. LODEWIJK , *A unified theory of variational principles in Continuum mechanics*, to be published.
27. BROER, L.J.F. & J.A. KOBUSSEN, *Conversion from material to local coordinates as a canonical transformation*, Appl.Sci.Res. 29 (1974) 419.
28. KOBUSSEN, J.A., *Some methods of classical mechanics applied to continuous systems*, thesis Eindhoven University of Technology, 1973.
29. SERRIN, J., *Mathematical principles of classical fluid mechanics*, in: *Encyclopedia of Physics*, S. Flügge ed., vol. VIII/I (Springer Verlag, Berlin, 1959)
30. BATEMAN, H., Proc.Roy.Soc. London A 125 (1929) 598.
31. BATEMAN, H., *Partial Differential Equations* (Cambridge University Press, Cambridge, 1944)
32. LUKE, J.C., *A. variational principle for a fluid with a free surface*, J. Fluid Mech. 27 (1967) 395.

33. BROER, L.J.F., *On the Hamiltonian theory of surface waves*, Appl. Sci.Res. 30 (1974) 430.

34. BROER, L.J.F., E.W.C. VAN GROESEN & J.M.W. TIMMERS, *Stable model equations for long water waves*, Appl.Sci.Res. 32 (1976) 619.

35. BENJAMIN, T.B., *Lectures on Nonlinear Wave Motion*, in: *Nonlinear Wave Motion*, Lectures in Applied Mathematics, vol.15 (A.M.S., Providence, 1974)

36. MILES, J.W., *On Hamilton's principle for surface waves*, J. Fluid Mech. 83 (1977) 153.

37. MILDER, D.M., *A note regarding "On Hamilton's principle for surface waves"*, J. Fluid Mech. 83 (1977) 159.

38. KORTEWEG, D.J. & G. DE VRIES, *On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary waves*, Phil. Mag. 39 (1895) 422.

39. BENJAMIN, T.B., J.L. BONA & J.J. MAHONY, *On model equations for long waves in nonlinear dispersive systems*, Phil.Trans.Roy. Soc. London A 272 (1972) 47.

40. GARDNER, C.S., *Korteweg- de Vries equation and generalizations IV. The Korteweg-de Vries equation as a Hamiltonian System*, J. Math. Phys. 12 (1971) 1548.

41. BROER, L.J.F., *Hidden Hamiltonians of first order equations*, Physica, 79 A (1975) 583.

42. FLASCHKA, H. & A.C. NEWELL, *Integrable systems of nonlinear evolution equations*, in: *Dynamical Systems, Theory and Applications*, J. Moser ed., Lecture Notes in Physics, vol. 38 (Springer Verlag, Berlin, 1975)

43. MIURA, R.M., *Korteweg-de Vries equation and generalizations* I. *A remarkable explicit nonlinear transformation*, J. Math.Phys. 9 (1968) 1202.

44. GARDNER, C.S., J.M. GREENE, M.D. KRUSKAL & R.M. MIURA: *Korteweg-de Vries equation and generalizations* VI. *Methods for exact solution*, Comm. Pure Appl. Math. 29 (1974) 97.

45. WHITHAM, G.B., *Linear and Nonlinear Waves*, (Wiley, New York, 1974)

46. BENJAMIN, T.B., *The Stability of Solitary Waves*, Proc. Roy. Soc. London A 328 (1972) 153.

47. BONA, J., *On the stability theory of Solitary waves*, Proc. Roy. Soc. London A 344 (1975) 363.

48. DE GRAAF, J., & L.J.F. BROER, *Plane waves in linear homogeneous media* II, Reports on Math. Phys., Nadbitka 3 (1972) 109.

49. BENJAMIN, T.B. & J.J. MAHONY, *On an invariant property of water waves*, J. Fluid Mech. 49 (1971) 385

50. BROER, L.J.F., *Approximate equations for long water waves*, Appl. Sci.Res. 31 (1975) 377.

51. TIMMERS, J.M.W., *Vergelijkingen voor lange, lage watergolven*, Afstudeerscriptie T.H.E., vakgroep Theoretische Natuurkunde 1976.

52. BONA,J.L. & R. SMITH: *A model for the two way propagation of water waves in a channel*, Math. Proc. Camb. Phil. Soc. 79 (1976) 167.

53. VALKERING, T.P., *An equation for long, low water waves. Interaction of permanent waves*. Phys. Comm., Twente University of Technology, vol. 8, no.1, 1976.

NAWOORD

Van diegenen die direct hebben bijgedragen aan het tot stand komen van dit proefschrift wil ik graag de volgenden noemen:

- *de medewerkers van de vakgroep Theoretische Natuurkunde;*
- *de vele studenten waarmee ik de afgelopen 4 jaar heb samengewerkt: door hen is veel van het in deel II beschreven onderzoek beinvloed;*
- *de leden van de werkgroep Niet Lineaire Analyse, Amsterdam, vanwege hun belangstelling voor en stimulering van het in deel I beschreven werk;*
- *dr.ir. J. de Graaf en prof.dr.ir. L.A. Peletier, vanwege het kritisch lezen van het manuscript van dit proefschrift; vooral de opmerkingen van eerstgenoemde over onderdelen van hoofdstuk 1 en 3 hebben tot aanzienlijke verbeteringen geleid;*
- *Deetje Bidlot en Marja Rooyakkers vanwege de nette verzorging van het type-werk;*
- *de medewerkers van de studiebibliotheken, in het bijzonder die van de afdeling Technische Natuurkunde, vanwege de vele plezierige contacten en hun behulpzaamheid.*

Al deze mensen wil ik hierbij van harte bedanken.

LEVENSLOOP

29 december 1949	*geboren te Breda*
juni 1967	*eindexamen H.B.S.-b, Newman H.B.S. te Breda*
september 1967	*aanvang wiskunde studie aan T.H.E.*
januari 1973	*doctoraal examen wiskundig ingenieur T.H.E.*
februari 1973 tot september 1974	*wetenschappelijk medewerker Gemeente Universiteit van Amsterdam, Instituut Toepassingen van de Wiskunde*
september 1974 tot september 1978	*wetenschappelijk medewerker T.H.E., vakgroep Theoretische Natuurkunde*

STELLINGEN

bij het proefschrift

VARIATIONAL METHODS IN MATHEMATICAL PHYSICS

van

E.W.C. van Groesen

8 december 1978.

I

Het bestaan van 2π-periodieke oplossingen van de vergelijking van DUFFING zonder demping:

$$\ddot{x}(t) + x(t) + \alpha x^3(t) = a \sin t$$

met (i) $\alpha = -1$, $a \in R\!l$ willekeurig

(ii) $\alpha = 1$, $a \in R\!l$, $|a|$ voldoende klein,

kan bewezen worden met variatiemethoden.
Het bewijs voor 't geval (i) (zie BERGER, §6.1B) is volkomen verschillend van dat voor 't geval (ii).

M.S. BERGER: *Nonlinearity and Functional Analysis*, Academic Press, New York, 1977.

II

Als $E(\rho)$ de inwendige energie per massaeenheid van een ideale vloeistof als functie van de massadichtheid ρ voorstelt, dan is de hydrostatische druk de Legendre getransformeerde van de functie $\rho \to \rho\, E(\rho)$, als deze functie convex wordt verondersteld.
Deze eenvoudige opmerking ligt ten grondslag aan het bestaan van een "principe van stationaire druk" waarmee de beweging van zo'n vloeistof beschreven kan worden.

R.L. SELIGER & G.B. WHITHAM, *Variational principles in continuum mechanis*, Proc.Roy.Soc. A 305 (1968) 1-25.

III

Periodieke oppervlaktegolven op een laag water kunnen beschreven worden met behulp van een "constrained extremum principle" van het soort zoals beschreven is in hoofdstuk 2 van dit proefschrift.
Deze formulering laat een duidelijke fysische interpretatie toe en kan gebruikt worden om de existentie van deze klasse bewegingen te bewijzen.

IV

In het recente boek van LEVINE wordt het begrip "unidirectional" uit de titel op geen enkele plaats gedefinieerd.
Bovendien zal de toevoeging van het woord *linear* op een juiste plaats in de titel het aantal gebruikers van het boek niet verkleinen.

H. LEVINE, *Unidirectional Wave Motions*, North Holland, Amsterdam, 1978

V

In [1, § 2.3.8] worden voorwaarden voor de lineaire operator A en de niet-lineaire operator B geformuleerd waaronder de met de abstracte evolutievergelijking

$$A\partial_t u = B(u)$$

overeenkomende operator $u \to A\partial_t u - B(u)$ een potentiaal operator is met betrekking tot één bepaalde klasse van functies u. VANDERBAUWHEDE [2] heeft onlangs deze resultaten op succesvolle manier gegeneraliseerd om een grotere klasse van functies u te kunnen behandelen.

[1] E.W.C. van GROESEN, *Variational methods for nonlinear operator equations*, in: Nonlinear Analysis, vol II, N.M. Temme (ed.) MC Syllabus 26.2, Mathematisch Centrum, Amsterdam, 1976, p. 100-191.

[2] A.L. VANDERBAUWHEDE, *Potential operators and variational principles: a generalization*, preprint Rijksuniversiteit Gent, august 1978.

VI

De stationaire toestanden van het systeem bestaande uit een onrekbare, flexibele stroomvoerende draad, geplaatst in een constant homogeen magneetveld, worden bij geschikte keuze van de randvoorwaarden beschreven door een stelsel vergelijkingen dat aanleiding geeft tot een karakteristiek bifurcatieprobleem.
Een bifurcatieprobleem met twee parameters ontstaat als de draad een eindige buigingsstijfheid heeft.

P.v.d. VARST: *De elastische lijn* (1977)
E.W.C. van GROESEN: *Stroomvoerende draad in magneetveld: stabiliteit van de stationaire toestanden* (1973) Afstudeerverslagen THE, vakgroep Theoretische Natuurkunde.

VII

De maatschappelijke waardering voor en beloning van onderwijsgevenden is omgekeerd evenredig met de invloed die zij hebben op de persoonlijkheidsvorming van hun leerlingen.

VIII

Het geven van onderwijs dient een wezenlijk onderdeel uit te (blijven) maken van de taken die behoren bij het bezetten van een promotieplaats.

IX

Gezien de gelaatsuitdrukking van de meeste automobilisten moet 't woordje "blij" in de kreet "Blij dat ik rij" een andere betekenis hebben dan daaraan meestal wordt toegekend.