ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

JÜRGEN MOSER

A rapidly convergent iteration method and non-linear partial differential equations - I

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 3e série, tome 20, nº 2 (1966), p. 265-315

<http://www.numdam.org/item?id=ASNSP_1966_3_20_2_265_0>

A RAPIDLY CONVERGENT ITERATION METHOD AND NON-LINEAR PARTIAL DIFFERENTIAL EQUATIONS - I.

JÜRGEN MOSER (*)

Introduction:

In the following lectures we shall discuss a number of problems connected with nonlinear differential equations, and the construction of their solutions. There are several methods available to cope with the difficulties encountered in the theory of nonlinear functional analysis. We mention iteration methods, the contraction principle which can be viewed as a generalization of the « regula falsi » to Banach spaces, and fixed point methods, as they were initiated by Leray and Schauder and their fixed point theorems. Schauder applied his method to the study of quasi linear hyperbolic differential equations and established the existence of the solutions « in the small » [1].

In Schauder's work careful a priori estimates for the solutions of some linear partial differential equations are basic for the applicability of the method. We shall not describe them here but mention only that these are square integral estimates which are also fundamental if one wants to establish the existence of weak solutions of hyperbolic equations [2].

For a long time problems were known which could not be attacked with these methods.

As a first example we mention the embedding problem: Given an abstract compact Riemannian manifold which possesses an infinitely differentiable structure, can one realize it as a submanifold of a finite-dimensional Euclidean space? It is understood that the metric should agree with the metric which is induced by the natural metric of the Euclidean space.

This question of « isometric embedding » has been answered by J. NASH using ingenious methods [3]. One can easily put this problem into the form of

---

Pervenuto alla Redazione il 23 Giugno 1965.

(*) Presented at the C.I.M.E. Summer Seminar in August 1964.

a system of partial differential equations, which indeed were unaccessible to the methods known before. In particular, these differential equations cannot be classified as hyperbolic, elliptic or as equations of a definite type, for which the theory has been developed to some extent. On the contrary the system is highly degenerate and the solution is not unique. The method of J. Nash was put into the form of an abstract implicit function theorem by J. Schwartz [4]. We also refer to [4′] and an application of the ideas in [4″].

A second example which we want to mention is connected with the stability problem of celestial mechanics, or more specifically, the problem of finding some almost periodic solutions for the three or n body problems. This question has been known for centuries and is related to the so called difficulty of the small divisors. The first steps towards surmounting this difficulty were made by C. L. Siegel [5], [5′] However, his results could not be adapted so as to give definite results for the differential equations of celestial mechanics. In 1954 Kolmogorov [6] [7] announced some new theorems for Hamiltonian systems of differential equations and in subsequent years V. I. Arnold supplied proofs [10] and gave striking applications of his results to the n body problem [9].

Again the relevant problem can be transformed into nonlinear partial differential equations which were not tractible previously, in spite of many serious attempts.

It would be difficult to give an account and the back-ground for either of the two problems. It turns out, however, that both results can be derived by essentially the same method (although the original approach by Nash seems to be different). Therefore we intend to present the ideas of this method removed from these particular problems and apply it rather to some simpler problems, namely the nonlinear theory of positive symmetric systems, as they were introduced by K. O. Friedrichs in the linear case [11]. It is conceivable that these equations are amenable to a different approach but we use them to illustrate our method. The result obtained will be applied to the study of invariant manifolds of vector fields as they were studied by and Bogolioubov and Mitropolsky ([1]), Diliberto [13], Kyner [14]. (*)

The last chapter contains a discussion of the results of Kolmogorov and Arnold which are relevant in celestial mechanics. The proofs are given for a simplified problem only.

([1]) A recent study by Kupka [15] on invariant surfaces in a very general con text will appear in the near future.

(*) Added in proof: Recently another application of this method was found by P. Rabinowitz in his dissertation at New York University, 1966. He established the existence of periodic solutions for second order hyperbolic differential equations which contain the highest order terms in nonlinear form.

CHAP. I. — Iteration and fast convergence.

§ 1 Approximation of functions by smoother ones
§ 2 Some lemmata and Inequalities
§ 3 Approximate solutions of linear equations
§ 4 Galerkin method
§ 5 Solution of Nonlinear problems

CHAP. II. — Positive Symmetric Systems of Partial Differential Equations.

§ 1 Linear Systems
§ 2 Existence theorem for nonlinear systems
§ 3 The analytic case
§ 4 Discussion of Invariant Surfaces for ordinary differential equations
§ 5 A priori estimates for the linear system
§ 6 A method avoiding loss of derivatives.

CHAP. III. — Conjugacy Problems.

§ 1 Siegel's theorem
§ 2 A construction for conjugacy problems
§ 3 Proof of Siegel's theorem
§ 4 A theorem by N. Levinson
§ 5 Vector fields on a torus and Kolmogorov's theorem
§ 6 Proof of theorem 1 (analytic case)
§ 7 Vector fields on a torus (differentiable case).

CHAP. I — Approximate Solutions.

In this section we intend to show how one can construct the solutions of a nonlinear problem by an iteration process where at each step an *approximate* solution of a linear equation is required. Several methods for finding such approximate solutions of the linear equations will be explained in a later section.

We want to emphasise that for the convergence of the process it is of advantage to work with approximate, rather than with exact solutions of the linearized equations. It is more advantageous to retain a high degree of smoothness of the approximation at the expense of the accuracy. In fact the natural iteration process may lead to divergence if one solves the linear equations exactly at each step.

The purpose of this section is to make these ideas precise and give definitions of approximate solutions of the linear and nonlinear problems.

Although these concepts are applicable in a much wider setting we will restrict ourselves here to vector functions on a torus and square integral norms which permit a particulary simple discussion.

§ 1. Approximation of functions by smoother ones.

a) We consider real functions $v(x)$ of n variables $x_1, \dots, x_n$ of period 2π in each of these variables. With the help of the Laplacean operator

$$\Delta = \sum_{\nu=1}^{n} \left(\frac{\partial}{\partial x_\nu}\right)^2$$

we introduce the inner product.

$$(1.1) \qquad (v, w)_\varrho = \int v \cdot (-\Delta)^\varrho w \, dx \qquad \text{for } \varrho = 0, 1, \dots, r.$$

where the integration is taken over $0 \leq x_\nu \leq 2\pi$ and dx abbreviates the volume element $dx_1 \dots dx_n$.

The norm $\| v \|_\varrho = (v, v)_\varrho^{1/2}$ vanishes for constant functions if $\varrho > 0$ but

$$(\| v \|_0^2 + \| v \|_\varrho^2)^{\frac{1}{2}}$$

represents a proper norm. The closure of all C^∞ functions (of period 2π) under this norm form a Hilbert space which we denote by V^ϱ (Sobolev space).

Using the Fourier expansion

$$v = \sum_k v_k e^{i(k, x)}, \qquad k = (k_1, \dots, k_n) \; (k_\nu \text{ integers})$$

one can introduce the spaces V^ϱ for non-integral values. We define

$$(1.2) \qquad \| v \|_\varrho^2 = 2\pi \sum_k | k |^{2\varrho} | v_k |^2$$

where $| k |^2 = k_1^2 + \dots + k_n^2$. The closure of the trigonometrical polynomicals in the norm (1.2) with $\| v \|_\varrho$ just defined for real ϱ will be called V^ϱ.

For integer $\varrho > 0$ this definition agrees with the previous one.

b) The norms $\| v \|_0$, $\| v \|_\varrho$, $\| v \|_r$ are related by several inequalities. We list the properties which will be needed later on.

It is well-known that for a given $v \in V^\varrho$, the expression $\log \| v \|_\varrho = \varphi(\varrho)$ is a convex function of ϱ in $(0, r)$ provided $0 < \| v \|_r < \infty$. This can be seen from the fact, that (1.2) defines $\varphi(\varrho)$ as an analytic function even for complex values of ϱ in the strip $0 \leq \operatorname{Re} \varrho \leq r$ and

$$\max_{\operatorname{Re} z = \varrho} | \varphi(z) | = \varphi(\varrho)$$

We assume here that v has no constant term.

Therefore, Hadamard's three line theorem ensures the convexity of $\varphi(\varrho)$ in $0 \leq \varrho \leq r$. Hence

$$(1.3) \qquad \| v \|_\varrho \leq \| v \|_{\varrho_1}^{\alpha} \| v \|_{\varrho_2}^{\beta} \qquad \text{for } \varrho = \alpha \varrho_1 + \beta \varrho_2$$

where α, β are non negative numbers with $\alpha + \beta = 1$.

In particular,

$$(1.3') \qquad \| v \|_\varrho \leq \| v \|_0^{1-\varrho/r} \| v \|_r^{\varrho/r}. \qquad \text{for } 0 < \varrho < r$$

We note that $V_\varrho \supset V_{\varrho'}$, if $0 \leq \varrho \leq \varrho' \leq r$.

The latter inequalities have been proven, if v has no constant terms.

But adding a constant to v does not affect the left and does not decrease the right hand side. Thus (1.3) and (1.3') hold in general.

c) Secondly we investigate how well one can approximate a function $v \in V^\varrho$ by functions of V^r.

Lemma 1: For $v \in V^\varrho$ $(0 < \varrho < r)$ and $Q \geq 1$ there exists a $w \in V^r$ such that

$$(1.4) \qquad \begin{cases} \| v - w \|_0 \leq K Q^{-\mu} \\ \| w \|_r \leq K Q \end{cases}$$

where $K = \| v \|_\varrho$ and

$$(1.5) \qquad \mu = \frac{\varrho}{r - \varrho} \text{ or } \frac{\varrho}{r} = \frac{\mu}{\mu + 1}.$$

Proof: We just have to choose for w the truncated Fourier series

$$w = \Sigma\, v_k\, e^{i(k, x)}$$

with an appropriate integer $N \geq 1$.

If we denote $v - w = z$ we have obviously

$$\begin{cases} \| w \|_r \leq N^{r-\varrho} \| w \|_\varrho \\ \| z \|_\varrho \geq N^\varrho \| z \|_0 \end{cases}$$

and therefore — using the orthogonality of the e^{ikx} with respect to all inner products $(v, w)_\varrho$:

$$\begin{cases} \| v - w \|_0 \leq N^{-\varrho} \; \| z \|_\varrho \leq KN^{-\varrho} \\ \qquad \| w \|_r \leq N^{r-\varrho} \| w \|_\varrho \leq KN^{r-\varrho} \end{cases} \tag{1.6}$$

Choosing $N = Q^{\frac{1}{r-\varrho}}$ and $\mu = \dfrac{\varrho}{r - \varrho}$ the lemma follows.

Conversely, we have

LEMMA 2 : If $v \in V^0$ has the property that for every $Q > 1$ there exists a $w \in V^r$ such that

$$\| v - w \|_0 \leq KQ^{-\mu} , \qquad (\mu > 0),$$

$$\| w \|_r \leq KQ$$

then $v \in V^\varrho$ for every ϱ satisfying

$$\frac{\varrho}{r} < \frac{\mu}{\mu + 1} \tag{1.7}$$

and $\| v \|_\varrho \leq cK$ if $\| v \|_0 \leq K$, where c depends on ϱ, r and n.

PROOF : Choosing $Q' = 2Q$ and denoting the corresponding approximation by w' we find

$$\| w - w' \|_0 \leq K (Q^{-\mu} + Q'^{-\mu}) \leq 2 \, KQ^{-\mu}$$

$$\| w - w' \|_r \leq K (Q + Q') \qquad \leq 3 \, KQ$$

and by (1.3') we have

$$\| w - w' \|_\varrho \leq 3 \; KQ^{-q} \text{ with } q = \mu \left(1 - \frac{\varrho}{r}\right) - \frac{\varrho}{r} .$$

The assumption (1.7) ensures that $q > 0$. Hence if we set $Q = Q_n = 2^n Q_0$ and call the corresponding approximations w_n then

$$\| w_n - w_{n+1} \|_\varrho \leq 3K\, Q_0^{-q}\, 2^{-nq} .$$

Thus w_n converges in V^ϱ to a limit w_∞.

If w_∞ is interpreted as element in V^0 then the assumption

$$\| v - w_n \|_0 \leq K Q_n^{-\mu}$$

shows that $v = w_\infty$. Thus $v \in V^\varrho$.

Moreover, we obtain an estimate for $\| v \|_\varrho$ from

$$\| v - w_0 \|_\varrho \leq \sum_{n=0}^{\infty} \| w_n - w_{n+1} \|_\varrho \leq K Q_0^{-q}\, C$$

where $C = 3 \sum 2^{-nq}$. Using

$$\| w_0 \|_r \leq K Q_0 ; \; \| w_0 \|_0 \leq \| v \|_0 + K Q^{-\mu} \leq 2K$$

we have

$$\| v \|_\varrho \leq \| w_0 \|_\varrho + Q_0^{-q}\, CK$$

$$\leq K (2 Q_0^{\varrho/r} + C Q_0^{-q})$$

For $Q_0 = 1$ we get the desired estimate with $C + 2$ as constant.

These lemmata show that the function spaces V^ϱ can nearly be characterized by their approximation properties by functions in V^r.

The loss in ϱ is unavoidable in the present set up as a simple example shows:

$$v = \sum_{|k| > 0} | k |^{-\left(\sigma + \frac{n}{2}\right)} e^{ikx}, \qquad 0 < \sigma < r,$$

admits an approximation in the sense of lemma 2 with

$$\mu = \frac{\sigma}{r - \sigma},$$

provided by the truncated Fourier series. However, v does not belong to V^σ but only to V^ϱ with $\varrho < \sigma$.

But in the following these crude estimates will be sufficient.

§ 2. Composition of functions.

a) In the preceding section we proved for the « Sobolev » spaces the inequalities (1.3′) and the approximation properties expressed by Lemma 1 and 2.

There are several other families of spaces V^ϱ $(0 \leq \varrho \leq r)$ satisfying these properties which we list again.

The V^ϱ $(0 \leq \varrho \leq r)$ are assumed to be Banach spaces in the order $V^0 \supset V^\varrho \supset V^r$, satisfying:

α) In V^ϱ a norm $(\| v \|_0^2 + \| v \|_\varrho^2)^{1/2}$ is defined where $\| v \|_\varrho$ satisfies

$$(2.1) \qquad \| v \|_\varrho \leq c \, \| v \|_0^{1 - \frac{\varrho}{r}} \, \| v \|_r^{\frac{\varrho}{r}}$$

and c depends on ϱ, r, n only. V^ϱ are assumed to be Banach spaces.

β) If $v \in V^\varrho$ and $\| v \|_\varrho \leq K$ then for $Q > 1$, there exists a $w \in V^r$ such that

$$(2.2) \qquad \left\{ \begin{aligned} & \| v - w \|_0 \leq c \, K Q^{-\mu} \\ & \| w \|_r \leq K Q \end{aligned} \right.$$

with c depending on r, ϱ, n only and $\mu = \dfrac{\varrho}{r - \varrho}$.

Conversely if $v \in V^0$ satisfies (2.2) with some $K \geq \| v \|_0$ then $v \in V^{\varrho'}$ where

$$\frac{\varrho'}{r} < \frac{\mu}{\mu + 1}$$

b) Clearly the vector valued functions $v = (v_1, \dots, v_m)$ with

$$\| v \|_\varrho^2 = \sum_{\mu=1}^{m} \| v_\mu \|_\varrho^2$$

and $\| v_\mu \|_\varrho$ defined as in (1.2) satisfy all the hypotheses.

A more interesting example is provided by the pair of norms

$$| v |_0 = \max_x \sqrt{\sum_{\mu=1}^{m} | v_\mu |^2}$$

$$\| v \|_r^2 = \int v \, (- \Delta)^r \, v \, d x$$

Let V^0 consist of all continuous functions with norm $| v |_0$ and V^r defined with the norm $\| v \|_r + | v |_0$.

How to define V^ϱ and the intermediate norms?
We shall give these only for integer ϱ in $0 < \varrho < r$.

These norms are given by the left side of the inequality

$$(2.3) \qquad \left\{ \sup \int | D_x^\varrho v_\mu |^{2 \frac{r}{\varrho}} dx \right\}^{\frac{\varrho}{2r}} \leq c \, | v |_0^{1 - \frac{\varrho}{r}} \, \| v \|_r^{\frac{\varrho}{r}}$$

where the supremum is taken over all derivatives D^ϱ of order ϱ and all components v_μ. The constant c depends on r, ϱ, n again.

This inequality — which provides the requirement (2.1) — is contained as a special case of a much more general theorem by L. Nirenberg [16], see also [17].

c) The inequality (2.3) allows the estimate for the composition of two functions: Let $\varphi = \varphi(x, y)$ be defined for $y = (y_1, \ldots, y_m)$ in $| y |^2 = \Sigma \, y_\mu^2 < 1$ and all $x = (x_1, \ldots, x_n)$, being of period 2π in the latter variables.

LEMMA: Assume that φ possesses continuous derivatives up to order r which are bounded by B. Then we have for

$$\varphi \circ v = \varphi(x, v(x))$$

the estimate

$$(2.4) \qquad \| \varphi \circ v \|_r \leq c B (\| v \|_r + 1)$$

provided $v \in V^r$ and $\max_x | v | = | v |_0 < 1$.

This estimate shows that $\| \varphi \circ v \|_r$ grows at most linearly with $\| v \|_r$ which seems not quite obvious at first.

PROOF: It suffices to prove (2.4) for scalar C^∞ functions $v(x)$.

Denoting by D^ϱ any derivative with respect to $x_1, \ldots, x_n$ of order ϱ we can write the « chain-rule » symbolically as

$$(2.5) \qquad D_x^r (\varphi \circ v) = \sum_{\varrho + \sigma \leq r} \left(D_\lambda^\sigma \frac{\partial^\varrho \varphi}{\partial y^\varrho} \right) \Sigma \, C_{\sigma \varrho \alpha} \, (Dv)^{\alpha_1} (D^2 v)^{\alpha_2} \ldots (D^r v)^{\alpha_r}$$

with constants $C_{\sigma\varrho\alpha}$ and non negative integers satisfying

$$(2.6) \qquad \begin{aligned} \alpha_1 + \ldots + \alpha_r &= \varrho \\ 1\alpha_1 + 2\alpha_2 + \ldots + r\alpha_r + \sigma &= r. \end{aligned}$$

These relations can be read off by counting the order of differentiation with respect to y and x.

To estimate the square integrals of the product in (2.5) we use Hölder's inequality for a multiple product:

Setting

$$v_0 = D_x^\sigma \frac{\partial^\varrho \varphi}{\partial y^\varrho}; \quad v_\lambda = D^\lambda v \qquad (\lambda = 1, \dots, r)$$

and

$$p_0 = \frac{r}{\sigma}; \quad \alpha_0 = 1; \quad p_\lambda = \frac{r}{\lambda \alpha_\lambda}$$

we have by (2.6)

$$\sum_{\lambda=0}^{r} \frac{1}{p_\lambda} = 1.$$

Note that $p_\lambda = \infty$ will be admitted. Hölder's inequality can now be written as

$$\int \prod_{\lambda=0}^{r} v_\lambda^{2\alpha_\lambda} dx \leq \prod_{\lambda=0}^{r} \left(\int | v_\lambda |^{2\alpha_\lambda p_\lambda} dx \right)^{\frac{1}{p_\lambda}} = \left(\int | v_0 |^{2\frac{r}{\sigma}} dx \right)^{\frac{\sigma}{r}} \left(\prod_{\lambda=1}^{r} \int | D^\lambda v |^{\frac{2r}{\lambda}} dx \right)^{\frac{\lambda\alpha_\lambda}{r}}.$$

The first factor will be estimated by B^2 and the second by (2.3) with the result:

$$\int \prod_{\lambda=0}^{r} v_\lambda^{2\alpha_\lambda} dx \leq B^2 C^r \prod_{\lambda=0}^{r} | v |_0^{\left(1 - \frac{\lambda}{r}\right) 2\alpha_\lambda} \| v \|_r^{\frac{2\lambda\alpha_\lambda}{r}}.$$

The expression simplifies because of $| v |_0 < 1$ and

$$\Sigma \frac{\lambda \alpha_\lambda}{r} = 1 - \frac{\sigma}{r} \leq 1$$

so that the right hand side is less than

$$B^2 C^r \| v \|_r^2 \qquad \text{or} \quad B^2 C^r.$$

This estimate is valid for each term in the sum (2.5) hence

$$\| \varphi \circ v \|_r^2 \leq B^2 C^r (1 + | v \|_r^2)$$

This proves (2.4) (with a different constant c).

For later purposes we list a similar estimate:

Let $\varphi(x, y, p)$ be a function of $x_1, \dots x_n$ of period 2π, of $y_1, \dots y_m$, and $p_{\mu\nu}$, $\mu = 1, \dots, m$, $\nu = 1, \dots, n$ in $| y_\mu | < 1$; $| p_{\mu\nu} | < 1$.

Moreover, let φ admit $r-1$ continuous derivatives with respect to all variables, bounded by B.

Consider the function $\varphi(x, v, v_x)$ which is obtained by substituting $y_\mu = v_\mu$; $p_{\mu\nu} = \dfrac{\partial v_\mu}{\partial x_\nu}$ where v is a function in V^r with

$$|v_\mu| < 1 ; \quad |v_{\mu x_\nu}| < 1. \tag{2.5}$$

Then

$$\|\varphi(x, v, v_x)\|_{r-1} \leq B(1 + \|v\|_r) \tag{2.6}$$

The proof of this inequality can easily be reduced to (2.4) by considering v, v_x as independent functions.

§ 3 Approximate solutions of linear equations.

a) We consider two such families of spaces satisfying conditions α) and β) in § 2 : $V^\varrho\,(0 \leq \varrho \leq r)$ with norms $\|v\|_\varrho$ for $v \in V^\varrho$ and $G^\sigma\,(0 \leq \sigma \leq s)$ with norms $\|g\|_\sigma$ for $g \in G^\sigma$.

We denote by L a linear operator mapping V^r into G^s. Usually we shall be dealing with differential operators of first order and we could identify G^s with V^{r-1} and $s = r - 1$. However, we wish to distinguish the domain V^r and range G^s by a different notation.

In the theory of elliptic differential equations, for example, the existence theory is based on the construction of spaces which are mapped one to one into each other (bijective). For the differential operator considered here this will not be the case. We give a simple example:

$$Lv = v_{x_1} + 2v_{x_2} + v$$

maps the space V^r into $V^{r-1} = G^{r-1}$. But an element $g \in G^{r-1}$ need not be the preimage of $v \in V^r$. Namely for a given function g_0 which depends on $2x_1 - x_2$ we have $v = g_0$ and $Lv = v$, hence $g_0 \in G^{r-1}$ implies $v_0 \in G^{r-1} = V^{r-1}$. In this particular case it is easy to discuss the solvability of the equation $Lv = g$ since the problem has constant coefficients.

To study such equations in a more general case we shall first construct approximate solutions: we speak of an approximate solution $w = w_Q$ of $Lv = g$ if for every $Q > 1$ there exists a $w \in V^r$ such that

$$\left\{\begin{aligned} &\|Lw - g\|_0 \leq K\eta(Q) \\ &\|w\|_r \leq KQ \end{aligned}\right. \tag{3.1}$$

if $\| g \|_0 \leq 1$, $\| g \|_s \leq K$, and $\eta(Q) \rightarrow 0$ as $Q \rightarrow \infty$. Usually we shall require

$$\eta(Q) \leq c Q^{-\mu} \tag{3.1'}$$

and call μ degree of the approximation.

For example, if L is the identity map from $v \in V^r$ to V^s with $0 < s < r$. Then the problem of solving $Lv = g$ approximately reduces to that of Lemma 1 and we can choose $\mu = \frac{s}{r-s}$.

b) We remark: If L is an operator which admits an estimate,

$$\| v \|_0 \leq \| Lv \|_0 \leq c \| v \|_\alpha$$

for $v \in V^r$ then the existence of an approximate solution for every $Q > 1$ with $\mu > \frac{\alpha}{r-\alpha}$ implies the existence of an exact solution, if $g \in G^s$.

The proof of this statement follows the same lines as Lemma 2 of § 1. Choose $Q = Q_n = 2^n$ and denote by w_n the corresponding approximate solution; from (3.1) we have

$$\| w_n - w_{n+1} \|_0 \leq \| L(w_n - w_{n+1}) \|_0 \leq c \, 2K \, Q_n^{-\mu}$$

and

$$\| w_n - w_{n+1} \|_r \leq 3 \, K Q_n .$$

Hence

$$\| w_n - w_{n+1} \|_\varrho \leq c \, K Q_n^{-q} \quad \text{with } q = \mu \left(1 - \frac{\varrho}{r}\right) - \frac{\varrho}{r} .$$

If $q > 0$ we conclude that w_n converges in V^ϱ to an element $v \in V^\varrho$. We take $\varrho = \alpha$ so that $q = \mu \frac{r-\alpha}{r} - \frac{\alpha}{r} > 0$ and Lw_n converges to Lv in G^0. Hence $Lv = g$ and v is the desired solution.

This shows that the requirement of an approximate solution for all $Q > 1$ is actually more stringent than the knowledge of an exact solution!

c) We shall describe now how one can construct approximate solutions for some linear operators which share some positivity properties with positive symmetric operators.

We shall assume that V^r are the Sobolev spaces of § 1 on which the inner product $(v, w)_r$ is defined. Similarly we use the same notation for G^s.

Let L be a linear operator which maps C^∞ into C^∞ and satisfies the estimates

$$\text{(3.2)} \qquad \begin{cases} \| v \|_0^2 \leq (Lv, v)_0 \\ \| v \|_s^2 \leq c\,((Lv, v)_s + K_1^2 \| v \|_0^2) \end{cases}$$

for $s = r - 1$ and all $v \in V^r$.

Here K_1 is a number greater than 1 which depends on L. Moreover, let

$$\text{(3.3)} \qquad \| Lv \|_s \leq c \| v \|_r .$$

For the construction of approximate solutions of $Lv - g = 0$ various methods are available. Here we reduce the problem to an elliptic problem by adding to L an « artificial viscosity term ». This trick is well known in numerical analysis (P. D. Lax) and has been used by L. Nirenberg in other connections [18]. It has to be mentioned, however, that for particular problems one can usually reduce the problem of finding approximate solutions to a *finite* one and the present approach is more complicated than necessary.

The device is the following: In order to solve the equation $Lv = g$ approximately and retaining more smoothness we solve the modified equation

$$\text{(3.5)} \qquad L_h w = (h^{2\alpha} (-\Delta)^\alpha + L)\, w = g$$

exactly where h is a small parameter in $0 < h < 1$ and $2\alpha \leq s$. This equation is elliptic and satisfies the same inequalities like L. The question of existence and uniqueness for this elliptic equation is standard by the projection method, the argument of Lax-Milgram and others. If g and the coefficients of L are in C^∞ so is the solution of (3.5). As $h \to 0$ one may expect the solution to converge to the exact solution of $Lv = g$.

Yet we shall not set $h = 0$ but rather keep the parameter not too small in order to hold the size of the higher derivatives down.

We shall show now that the solution of (3.5) yields an approximate solution of $Lw = g$ with a degree of approximation

$$\text{(3.6)} \qquad \mu = 2\alpha$$

provided

$$\text{(3.7)} \qquad 1 \leq \alpha \leq s/2 .$$

For this purpose we make use of the estimates (3.2) to derive

$$h^{2\alpha} \| w \|_\alpha^2 + \| w \|_0^2 \leq (L_h w, w)_0 = (g, w)_0 \leq \frac{1}{2} (\| g \|_0^2 + \| w \|_0^2)$$

which yields

$$\| w \|_0 \leq c_0 \| g \|_0 .$$

Similarly, we find for higher derivatives

$$h^{2\alpha} \| w \|^2_{\alpha+s} + \| w \|^2_s \leq c \{(L_h w, w)_s + K_1^2 \| w \|^2_0\}$$

$$\leq \frac{1}{2} (\| w \|^2_s + c^2 \| g \|^2_s) + cK_1^2 \| w \|^2_0$$

hence

$$h^{2\alpha} \| w \|^2_{\alpha+s} + \frac{1}{2} \| w \|^2_s \leq c^2 (K^2 + K_1^2) \leq 2c^2 K^2 \text{ if } \| g \|_s \leq K$$

for $K \geq K_1$. Thus we have

$$\| w \|_{\alpha+s} \leq c' h^{-\alpha} K ; \quad \| w \|_s \leq c' K$$

and, since $r = s + 1 \leq s + \alpha$ we have from (1.3)

$$\| w \|_r \leq c_1 h^{-1} K . \tag{3.8}$$

By (3.7)

$$\| (-\Delta)^\alpha w \|_0 = \| w \|_{2\alpha} \leq c_2 K.$$

Hence

$$\| Lw - g \|_0 = h^{2\alpha} \| (-\Delta)^\alpha w \|_0 \leq c_2 K h^{2\alpha} . \tag{3.9}$$

The relations (3.8), (3.9) verify that w is an approximate solution in the sense of (3.1): With $Q = \frac{c_1}{h}$ one has

$$\eta (Q) = c_2 c_1^{2\alpha} Q^{-2\alpha} .$$

This proves (3.6).

If one chooses s as an even number the strongest approximation is obtained for $\alpha = s/2$ in which case we have $\mu = s$. For odd s we have $\mu = s - 1$ for $2\alpha = s - 1$.

d) For later applications we shall discuss the situation when the additional term $K_1^2 \| v \|^2_0$ in (3.2) is replaced by $K_1^2 \, | v |^2_0$ (where $| v |_0 = \sup_x | v |$), i. e. we replace (3.2) by

$$\begin{cases} \| v \|^2_0 \leq (Lv, v)_0 \\ \| v \|^2_s \leq c \, ((Lv, v)_s + K_1^2 \, | v |^2_0). \end{cases} \tag{3.10}$$

We shall show that also in this case we can produce an approximate solution of the linear equation $Lv = g$ satisfying the same inequalities (3.1) provided that

$$\| g \|_0 \leq K^{-\frac{n}{2s-n}}; \quad \| g \|_s \leq K; \quad s > \frac{n}{2}. \tag{3.11}$$

To prove this remark we construct w again as solution of

$$L_h w = g.$$

It remains to be shown that w satisfies (3.1). This follows precisely as before if we can show that the additional term $K_1^2 \, | w |_0^2$ can be estimated by K^2. Therefore we shall prove now that indeed

$$K_1 \, | w |_0 \leq K \text{ for } K > c \, K_1 \tag{3.12}$$

holds as a consequence of (3.11).

For this propose we derive the a priori estimates

$$\| w \|_0 \leq c_0 \, \| g \|_0$$

$$\| w \|_s^2 \leq c \left((L_h w, w)_s + K_1^2 \, | w |_0^2 \right)$$

$$\leq \frac{1}{2} \left(\| w \|_s^2 + c^2 \, \| g \|_s^2 \right) + c K_1^2 \, | w |_0^2 .$$

Hence

$$\| w \|_s \leq c_1 \left(\| g \|_s + K_1 \, | w |_0 \right). \tag{3.13}$$

We combine these estimates with the general Sobolev inequality

$$| w |_0 \leq c_2 \, \| w \|_0^{1 - \frac{n}{2s}} \, \| w \|_s^{\frac{n}{2s}}, \quad \text{for} \quad s > \frac{n}{2} .$$

Therefore we have

$$| w |_0 \leq c_2 \left(c_0 \, \| g \|_0 \right)^{1 - \frac{n}{2s}} \, \| w \|_s^{\frac{n}{2s}}$$

and with (3.11), (3.13)

$$| w |_0 \leq c_2 \, c_0^{1 - \frac{n}{2s}} \left(K^{-1} \, \| w \|_s \right)^{\frac{n}{2s}} \leq \left\{ c_3 \left(1 + \frac{K_1}{K} \, | w |_0 \right) \right\}^{\frac{n}{2}}$$

with another constant $c_3 > 1$.

Assuming that, (3.12) would not hold we would have

$$| w |_0 < \left(2c_3 \frac{K_1}{K} | w |_0\right)^{\frac{n}{2s}} < \left(| w |_0\right)^{\frac{n}{2s}} \quad \text{for } K > 2c_3 K_1$$

which is a contradiction and (3.12) is established with $c = 2c_3$.

§ 4 Galerkin Method.

a) Here we want to describe a second method for construction of approximate solutions, the so called Galerkin method. It has the advantage to reduce the problem to a finite dimensional one. In fact, it amounts to solving the linear equation $Lu = f$ projected into a finite dimensional space.

To introduce these finite dimensional spaces we assume that the set

$$\| v \|_r \leq 1$$

is compact with respect to the norm $\| v \|_0$. (This assumption is certainly fulfilled for functions on a torus with the norms introduced). We can represent $(v, w)_0$ in the form

$$(v, w)_0 = (v, Rw)_r$$

where R is a symmetric and compact operator [2].

For any number $N \geq 1$ let H_N be the eigen space of R corresponding to the part of the spectrum where $| \lambda | > N^{-r}$. Then H_N is a finite dimensional space in which

$$\| v \|_0 \geq N^{-r} \| v \|_r \quad \text{for } v \in H_N$$

and in the orthogonal complement:

$$\| v \|_0 \leq N^{-r} \| v \|_r .$$

If P_N denotes the projection of V^0 into H_N we have then for $P = P_N$

$$\text{(4.1)} \qquad \begin{cases} \| Pv \|_r \leq N^r \| Pv \|_0 \\ \| (I - P) v \|_0 \leq N^{-r} \| (I - P) v \|_r \leq N^{-r} \| v \|_r . \end{cases}$$

---

(2) In order to avoid any difficulty with the elements for which $\| v \|_r = 0$ (as the constants on the torus) we restrict ourselves to the orthogonal complement of these elements.

More generally, since the P_N commute with differentiation

$$(4.1') \qquad \| Pv \|_r \leq N^{r-s} \| Pv \|_s$$

$$\| (I - P) v \|_0 \leq N^{-s} \| (I - P) v \|_s$$

and the P_N are a family of commuting self adjoint projections satisfying $P_N P_{N'} = P_N$ for $N' > N$.

In the case discussed in the previous section the P_N projects functions into trigonometrical polynomials. Clearly, one can introduce a similar projection of functions on a sphere into spherical harmonics, or any closed manifold into the eigen-functions of the Beltrami Laplace equations.

b) We come to the construction of an approximate solution for a linear equation $Lv = g$. Let us note that for the identity of $L = I$ the approximate solution of $v = g$ can be given by

$$v = P_N g,$$

as we saw in the previous section.

Similarly, we construct an approximate solution of $Lv = g$, assuming that

$$(4.2) \qquad \begin{cases} \| v \|_0^2 \leq (v, Lv)_0 \\ \| v \|_s^2 \leq c \{(v, Lv)_s + K_1^2 \| v \|_0^2\} \end{cases}$$

and

$$(4.3) \qquad \| Lv \|_s \leq c \| v \|_r$$

Let $\| g \|_s \leq K$, $\| g \|_0 \leq 1$ and find $v = v_N$ as a solution of the linear equations

$$(4.4) \qquad P_N (Lv - g) = 0\,; \; P_N v = v.$$

Since the range of P_N is finite dimensional, (4.4) constitutes finitely many equations in equally many unknowns. The solvability is guarenteed if the determinant does not vanish, i. e. it suffices to prove uniqueness of the solution of homogeneous equation. We prove more by the following estimate:

From (4.2) we have for $v = v_N$

$$\| v \|_0^2 \leq (v, Lv)_0 = (Pv, Lv)_0 = (v, g)_0$$

and hence

$$(4.5) \qquad \| v \|_0 \leq \| g \|_0 .$$

To estimate the degree of approximation we find from (4.1′) and (4.2)

$$\| v \|_r^2 \leq (N^{r-s} \| v \|_s)^2 \leq N^{2(r-s)} c ((v, g)_s + K_1^2 \| v \|_0^2)$$

Estimating the right hand side by the Schwarz inequality, we arrive at

$$\| v \|_r \leq c_1 N^{r-s} K \text{ for } K \geq K_1 . \tag{4.6}$$

It remains to estimate $\| Lv - g \|_0$.

For this purpose we note that $Lv - g$ is orthogonal to H_N and hence by (4.1′) admits the estimate

$$\| Lv - g \|_0 \leq N^{-s} \| Lv - g \|_s \leq N^{-s} (\| Lv \|_s + \| g \|_s).$$

Using (4.3) and (4.6) we have

$$\| Lv - g \|_0 \leq N^{-s} (c c_1 N^{r-s} K + K) \leq c_2 K N^{r-2s} \tag{4.7}$$

if $s < r < 2s$ and $N > 1$. Thus the degree of approximation, which is implied by (4.7), is

$$\mu = \frac{2s - r}{r - s} \text{ if } s < r < 2s. \tag{4.8}$$

For example, for $r = s + 1$ we have $\mu = s - 1$.

The result shows that the construction of an approximate solution leads to a finite dimensional problem, whenever $\| v \|_r \leq 1$ is a compact set in V^0, for example on closed manifolds.

§ 5 The Nonlinear Case.

a) We shall show now how the concept of an approximate solution can be used for the construction of exact salutions of non linear problems.

We consider a functional $\mathcal{F}(u)$ which is defined in a neighborhood of an element u_0.

The result will be applied mainly to partial differential equations of positive type, but we shall formulate the results more generally.

Our result is of the type of the inverse function theorem, and establishes the existence of a solution u near u_0 of $\mathcal{F}(u) = f$ if f is close to $\mathcal{F}(u_0) = f_0$. For this purpose we shall require that the linearized equations

$$\lim_{t \to 0} \frac{\mathcal{F}(u + tv) - \mathcal{F}(u)}{t} = \mathcal{F}'(u) v = g$$

admits approximate solutions. This is required not only for the linearization at $u = u_0$ but for *every u near* u_0.

We formulate the conditions more precisely:

Let V^r, V^0 denote the function spaces defined in Section 1 and let $u \in V^r$. We shall denote the domain

$$\| u - u_0 \|_0 < 1 \text{ (}^3\text{)} \tag{5.1}$$

by $\mathcal{U}$. In this domain $\mathcal{F}(u)$ is defined, mapping u into an element f in

$$\begin{cases} \|f - f_0\|_0 < M; & f_0 = \mathcal{F}(u_0) \in G^s \\ \|f - f_0\|_s < \infty & (0 < s < r): \end{cases} \tag{5.2}$$

and for every $K > 1$ and $u \in \mathcal{U}$

$$\| \mathcal{F}(u) \|_s \leq MK \qquad \text{if } \| u \|_r < K,\ u \in \mathcal{U}. \tag{5.3}$$

The derivative operator $\mathcal{F}'(u)$ defined by

$$\mathcal{F}'(u)\, v = \lim_{\to 0} \frac{1}{t} (\mathcal{F}(u + tv) - \mathcal{F}(u))$$

is assumed to exist with values in G^s for $u \in \mathcal{U}$ and $v \in V^r$. Moreover, the linearized equation $\mathcal{F}'(u)\, v = g$ is supposed to admit an approximate solution $v \in V^r$ in the following sense:

If $g \in G^s$ and $\| g \|_0 \leq 1$ ([4]), $\| g \|_s \leq K$ and $\| u \|_r < K$ we require for every $Q > 1$ the existence of a $v = v_Q \in V^r$ satisfying

$$\begin{cases} \| \mathcal{F}'(u)\, v - g \|_0 \leq K Q^{-\mu} \\ \| v \|_r \leq KQ \end{cases} \tag{5.4}$$

([3]) In later applications we will replace (5.1) by $\| u - u_0 \|_0 + \| u - u_0 \|_\varrho < 1$ with some $\varrho > 0$.

([4]) This condition may be relaxed to $\| g \|_0 < K^{-\lambda}$ with a positive λ introduced below.

and

$$\| \mathcal{F}'(u)\, v \|_0 \geq \| v \|_0 \,.\text{[5] [6]} \tag{5.5}$$

Finally we require that the quadratic part

$$Q(u, v) = \mathcal{F}(u + v) - \mathcal{F}(u) - \mathcal{F}'(u)\, v$$

admits the estimate

$$\| Q(u, v) \|_0 \leq M \| v \|_0^{2-\beta} \| v \|_r^{\beta} \qquad (0 \leq \beta < 1) \tag{5.6}$$

for $u \in \mathcal{U}$ and $v \in V^r$, if $\| v \|_0 < \| v \|_r$.

b) Under the above conditions we propose to solve $\mathcal{F}(u) = f$ for u if f is close to $f_0 = \mathcal{F}(u_0)$. Like in the linear case we shall speak of an approximate solution to this problem if for every $K > 1$ there is a $u = u_K \in \mathcal{U}$ satisfying.

$$\| \mathcal{F}(u) - f \|_0 < K^{-\lambda} ; \; \| u \|_r < K \tag{5.7}$$

and call λ the degree of approximation.

The purpose of the following theorem is to show that the construction of an approximate solution of degree μ can be used to construct approximations to the nonlinear equations of degree λ. We shall have to assume here that

$$0 < \lambda + 1 < \frac{1}{2}(\mu + 1) \tag{5.8}$$

and

$$0 < \beta < \frac{\lambda}{\lambda + 1} \frac{\mu}{\mu + 1} \left(1 - 2\frac{\lambda + 1}{\mu + 1}\right) \tag{5.9}$$

where β is defined in (5.6).

---

(5) This does not imply the uniqueness of a solution of $\mathcal{F}'(u)\, v = g$ since this estimate is only required for the approximate solution constructed which may lie in a smaller subspace. We normalized the coefficient on the right hand side to 1 since this can always be achieved.

(6) If for some reason the linearizing operator $\mathcal{F}'(\widehat{u}) = L_\infty$ on the unknown solution $\widehat{u}$ is known, it suffices to impose the above requirements on L_∞ only.

THEOREM: We assume that $\mathcal{F}(u)$ has the properties listed in (5.1) to (5.6). Then we claim that there exists a constant $K_0(M, \beta, \mu, \lambda) > 1$ such that if u_0 and $\mathcal{F}(u_0) = f_0$ satisfy.

$$\|f - f_0\|_0 < K_0^{-\lambda}; \ \|u_0\|_r < K_0 \text{ and } \|f\|_s \leq MK_0 \tag{5.10}$$

then we shall construct a sequence of approximations $u_n \in \mathcal{U}$ such that

$$\|\mathcal{F}(u_n) - f\|_0 < K_n^{-\lambda}; \ \|u_n\|_r < K_n \tag{5.11}$$

where $K_n \to \infty$. u_n converges to a solution $\widehat{u}$ in the norm

$$\|\widehat{u} - u_n\|_{\varrho'} \to 0 \quad \text{for} \quad n \to \infty$$

if

$$\frac{\varrho'}{r} < \frac{\lambda}{\lambda + 2}. \tag{5.11'}$$

Assuming that $\mathcal{F}(u)$ maps $V^{\varrho'}$ continuously into G^0 it follows that $\mathcal{F}(\widehat{u}) = f$.

REMARK: The convergence of the constructed sequence is faster than linear, since we shall prove the above inequalities with

$$K_{n+1} = K_n^{\varkappa}$$

where

$$\varkappa > \left(1 - \frac{\lambda + 1}{\mu + 1}\right)^{-1} > 1. \tag{5.12}$$

PROOF: We proceed by induction and shall construct a sequence u_n verifying the inequalities

$$\begin{cases} \|\mathcal{F}(u_n) - f\|_0 < K_n^{-\lambda} \\ \|u_n - u_{n-1}\|_0 < 2K_{n-1}^{-\lambda} \\ \|u_n - u_{n-1}\|_r < \frac{1}{2} K_n \end{cases} \tag{5.13}$$

where $K_n = K_{n-1}^{\varkappa}$ with some exponent $\varkappa$ in $1 < \varkappa < 2$.

The first of these inequalities is satisfied for $n = 0$ while the other two are empty since u_{-1} is not defined.

Assuming that the inequalities (5.13) have been proven for $u_0, u_1, \dots, u_n$ we shall establish them for u_{n+1}.

From the third inequality it follows that

$$\| u_n \|_r = \| u_0 \|_r + \sum_{\nu=1}^{n} \| u_\nu - u_{\nu-1} \|_r = K_0 + \frac{1}{2} \sum_{\nu=0}^{n} K_\nu$$

if K_0 is chosen large enough [8].

Therefore by (5.3)

$$\| \mathcal{F}(u_n) \|_s < MK_n .$$

The next approximation $u_{n+1} = u_n + v$ will be chosen by solving the linearized equation

$$\mathcal{F}'(u_n)\, v + \mathcal{F}(u_n) = f$$

approximately. According to (5.4) (applied to $g = f - \mathcal{F}(u_n)$, $u = u_n$ there exists a v satisfying [7],

$$\| \mathcal{F}'(u_n)\, v + \mathcal{F}(u_n) - f \|_0 \leq 2MK_n\, Q^{-\mu} \tag{5.14}$$

$$\| v \|_r \leq 2MK_n\, Q \tag{5.15}$$

and by (5.5)

$$\| v \|_0 \leq \| \mathcal{F}'(u_n)\, v \|_0 .$$

The first and third of these relations give

$$\| v \|_0 \leq \| \mathcal{F}(u_n) - f \|_0 + 2MK_n\, Q^{-\mu} \leq K_n^{-\lambda} + 2MK_n\, Q^{-\mu} .$$

---

(7) Here we used that $g = f - \mathcal{F}(u_n)$ satisfies

$$\| g \|_0 \leq K_n^{-\lambda} \leq 1, \qquad \| g \|_s \leq \| \mathcal{F}(u_n) \|_s + \| f \|_s \leq M(K_n + K_0) \leq 2MK_n .$$

by (5.3) and (5.10).

(8) The choice of K_0 depends on $\varkappa > 1$; $\varkappa$ will be chosen as a function of μ, λ, β later on.

We shall choose Q so large that

$$2MK_n Q^{-\mu} < K_n^{-\lambda} \tag{5.16}$$

and therefore

$$\| u_{n+1} - u_n \|_0 = \| v \|_0 < 2 K_n^{-\lambda}$$

which proves the second of the inequalities (5.13). Also the last of those 3 inequalities follows immediately from (5.15) provided Q is chosen such that

$$2MK_n Q < \frac{1}{2} K_{n+1} . \tag{5.17}$$

Finally we verify the first inequality of (5.13): By (5.14) and (5.6)

$$\| \mathcal{F}(u_{n+1}) - f \|_0 = \| \mathcal{F}(u_n + v) - f \|_0 = \| \mathcal{F}(u_n) + \mathcal{F}'(u_n) v + Q(u_n, v) - f \|_0$$

$$\leq c \{ K_n Q^{-\mu} + K_n^{-\lambda(2-\beta)} (K_n Q)^\beta \} .$$

where $c > M$ depends on M and the exponents only.

The theorem will be proven if we succeed to choose Q in such a way that (5.16), (5.17) and

$$c \{ K_n Q^{-\mu} + K_n^{-\lambda(2-\beta)} (K_n Q)^\beta \} < K_{n+1}^{-\lambda} .$$

Since $K_{n+1} > K_n$ one sees that the last inequality implies (5.16). It remains to find Q such that the 3 inequalities

$$\begin{aligned} & c K_n Q < K_{n+1} \\ & c (K_n Q)^\beta K_n^{-\lambda(2-\beta)} < K_{n+1}^{-\lambda} \\ & c K_n Q^{-\mu} < K_{n+1}^{-\lambda} , \end{aligned} \tag{5.18}$$

hold. The first two inequalities are upper estimates for Q while the last yields a lower bound for Q.

We use $K_{n+1} = K_n^{\varkappa}$ and express the inequalities in terms of powers of $K = K_n$, which will be chosen sufficiently large. Comparing the exponents one finds that Q can be found if

$$(5.19) \qquad \begin{cases} \dfrac{2-\varkappa}{\varkappa} > \dfrac{\lambda+1}{\lambda} \dfrac{\mu+1}{\mu} \beta \\ \dfrac{\varkappa}{\varkappa - 1} < \dfrac{\mu+1}{\lambda+1} \ (9) \end{cases}$$

or

$$\frac{\lambda+1}{\lambda} \cdot \frac{\mu+1}{\mu} \beta < \frac{2}{\varkappa} - 1 < 2\left(1 - \frac{\lambda+1}{\mu+1}\right) - 1 = 1 - 2\,\frac{\lambda+1}{\mu+1}.$$

This shows that under the assumptions (5.8), (5.9) one can find a number $\varkappa$ satisfying (5.19)

$$1 < \left(1 - \frac{\lambda+1}{\mu+1}\right)^{-1} < \varkappa < 2.$$

Hence (5.18) are compatible if K_0 is chosen large enough; K_0 depends on M, β, μ, λ.

---

(9) Computation leading to the inequalities (5.19): Comparison of the exponents in (5.18) gives first the inequalities

A) $\varkappa\lambda + 1 < \mu(\varkappa - 1)$

B) $\varkappa\lambda + 1 < \mu \left\{ -1 + \lambda \dfrac{2-\beta}{\beta} - \varkappa \dfrac{\lambda}{\beta} \right\}$.

Adding $\varkappa - 1$ to both sides of A) gives the second inequality of (5.19). We rewrite B) as

$$\varkappa\lambda + 1 < \mu \left\{ -(1+\lambda) + (2-\varkappa)\frac{\lambda}{\beta} \right\}$$

Since $\varkappa > 1$ it is a stronges requirement if we satisfy

$$\varkappa(\lambda+1) < \mu \left\{ -\varkappa(\lambda+1) + (2-\varkappa)\frac{\lambda}{\beta} \right\}$$

which is indeed equivalent to the first relation in (5.19).

This completes the proof of (5.13). From it we derive that

$$\| u_{n+1} \|_r \leq \| u_0 \|_r + \sum_{\nu=1}^{n+1} \| u_\nu - u_{\nu-1} \|_r \leq K_0 + \sum_{\nu=1}^{n+1} \frac{1}{2} K_\nu < K_{n+1}$$

if K_0 large enough, proving the second half of (5.11).

Finally from

$$\| v \|_{\varrho'} \leq \| v \|_0^{1-\frac{\varrho'}{r}} \| v \|_r^{\frac{\varrho'}{r}}$$

we have

$$\| u_n - u_{n-1} \|_{\varrho'} \leq c K_n^{-\left(1-\frac{\varrho'}{r}\right)\lambda + \frac{\varrho'}{r}\varkappa}$$

which has a negative exponent if

$$\frac{\varrho'}{r} < \frac{\lambda}{\lambda + \varkappa}$$

Therefore u_n converges to an element $u \in V^{\varrho'}$. (10)

If K_0 is chosen large enough clearly

$$\| u - u_0 \|_0 \leq \sum_{m=1}^{\infty} \| u_m - u_{m-1} \|_0 < 1$$

so that $\mathcal{F}(u)$ is defined. Since $\mathcal{F}(u)$ is continuous as a mapping from $V^{\varrho'}$ to G^0 we conclude from (5.13) $\mathcal{F}(u) = f$.

c) Finally we discuss again the construction of approximate solutions satisfying (5.4). A sufficient condition was described in Section 3. Here we want to recapitulate that statement for the operator L defined by

$$Lv = \mathcal{F}'(u)\, v.$$

Clearly, L depends on the choice of u and so will the estimates (3.2). We

---

(10) Therefore all approximations remain in

$$\| u - u_0 \|_{\varrho'} \leq \sum_{m=1}^{\infty} \| u_m - u_{m-1} \|_{\varrho'} < 1$$

if K_0 large enough, and we could have restricted $\mathcal{U}$ in (5.1) to $\| u - u_0 \|_{\varrho'} < 1$; $\| u - u_0 \|_r < \infty$.

shall require therefore that for

$$u \in \mathcal{U} \text{ and } \| u \|_r \leq K$$

the estimates:

$$(5.20) \qquad \begin{aligned} \| v \|_0^2 &\leq (Lv, v)_0 \\ \| v \|_s^2 &\leq c \{(Lv, v)_s + K_1^2 \| v \|_0^2\}\,. \end{aligned}$$

hold, where c is independent of u. Then the argument of Section 3 ensures the existence of approximate solutions satisfying (5.4).

The above conditions (5.20) will be verified for some partial differential operators. Thus the condition (5.4) can be reduced to a priori estimates again.

We relax (5.21) to the a priori estimate

$$(5.21) \qquad \begin{cases} \| v \|_0^2 \leq (Lv, v)_0 \\ \| v \|_s^2 \leq c \{(Lv, v)_s + K_1^2 \sup v^2\} \end{cases}$$

and show that the previous derivations remain valid provided that

$$(5.22) \qquad \lambda > \frac{n}{2s - n}\,; \qquad s > \frac{n}{2}\,.$$

Indeed, we showed at the end of § 3, that an approximate solution satisfying (5.4) can be found if (3.11) holds, i. e. if

$$\| g \|_0 \leq K^{-\frac{n}{2s-n}}$$

holds. For $g = -\mathcal{F}(u_n) + f$ this relation is a consequence of (5.13) if (5.22) is satisfied. This remark will be useful in the application of the next chapter where only the a priori estimates (5.21) are available.

CHAP. II. Positive symmetric systems.

§ 1. Linear Systems

We shall show how these methods can be used in the theory of partial differential equations to study a class of so called positive symmetric systems which were introduced by Friedrichs [11] in order to handle problems in which the « type » of the equation changes, i. e. the systems may be of elliptic type in some region and of hyperbolic in another, as it occurs in transonic flow problems, for example. On the other hand the equations admit certain estimates which allow one to prove the existence of solutions.

In this Chapter we wish to investigate such systems in the non linear case applying the methods of the previous section. It will be the main point to circumvent the preliminary concept of a weak solution but rather to obtain a construction for the solution by approximations which converge pointwise with several derivatives.

In the linear case such positive symmetric systems are defined as follows: Let $x = (x_1, \dots, x_n)$ be in a domain and $u = (u_1, \dots, u_m)$ denote a vector and $a^{(\nu)}(x)$, $b(x)$ both m by m matrices. The general first order system has the form

$$Lu \equiv \sum_{\nu=1}^{n} a^{(\nu)} \frac{\partial}{\partial x_\nu} u + bu = f(x).$$

Such a system will be called positive symmetric, if all the matrices $a^{(\nu)}(x)$ are symmetric and

$$b + b^T - \sum_{v=1}^{n} a^{(\nu)}_{x_\nu}, \tag{1.1}$$

is positive definite. Actually one has to supply boundary conditions but we shall study a simple problem in which we require that $a(x)$, $b(x)$, $f(x)$ and the desired solution $u(x)$ have period, say 2π, in all variable $x_1, \dots, x_n$. That is to say, we consider the problem on the torus, which will be motivated later on.

To discuss the meaning of the positive symmetric character we note that the above assumptions imply that the quadratic form

$$(u, Lu) = \int_\Omega \left\{ u \left(\sum_{v=1}^{n} a^{(i)} \frac{\partial}{\partial x_v} u \right) + ubu \right\} dx$$

is positive definite if the integration is taken over the torus $\Omega:\ 0 \leq x_\nu \leq 2\pi$. Namely, an integration by parts yields

$$(u, Lu) = \int_\Omega u \left(b - \frac{1}{2} \Sigma\, a^{(\nu)}_{x_\nu} \right) u \, dx$$

which by (1.1) is positive definite.

Introducing the matrix

$$b_0 = b - \frac{1}{2} \sum_{\nu=1}^{n} a^{(\nu)}_{x_\nu} \tag{1.2}$$

one can write the system with Friedrichs in the form

$$\frac{1}{2} \sum_{\nu=1}^{n} \left(a^{(\nu)} \frac{\partial}{\partial x_\nu} u + \frac{\partial}{\partial x_\nu} a^{(\nu)} u \right) + b_0 u = f$$

which shows that the first term is antisymmetric and therefore does not contribute to the quadratic form (u, Lu). Since the type of the equations is governed just by the matrices $a^{(\nu)}$ the positivity of the system i. e. of b_0 depends heavily on b and not only on the first order derivatives of $a^{(\nu)}$.

For the following we shall derive some a priori estimates for the higher derivatives, that means lower estimates for

$$(Lu, u)_l$$

where l is a large integer. Special care will be taken on how these estimates depend on the high derivatives of the coefficients.

LEMMA: If for $\xi = (\xi_1, \ldots, \xi_n)$; $\eta = (\eta_1, \ldots, \eta_m)$ on $|\xi| = 1$, $|\eta| = 1$ the inequality

$$\langle (l \sum a^{(\nu)}_{x_\mu} \xi_\nu \xi_\mu + b_0 |\xi|^2) \eta, \eta \rangle > 2\gamma > 0 \, . \tag{1.3}$$

holds then one has for $v \in V^l$.

$$\| v \|_l^2 \leq \gamma (Lv, v)_l + c (1 + \| a \|_l + \| b \|_l)^2$$

with a constant c which depends on γ and on an upper bound for $(|a|_0 + |a|_2 + |b|_0 + |b|_1)$ (these are maximum norms of derivatives up to order 2 or 1 respectively) and $|v|_0$.

We shall postpone the proof to later (see § 5). Here we note that the condition (1.3) agrees with (1.1) for $l = 0$.

We shall abbreviate it in the symbolic form

$$l \langle a_x \rangle + \langle b_0 \rangle > 2\gamma \tag{1.4}$$

We note that for given matrices a, b on the torus this conditions sets an upper limit to l, except in the trivial case of constant coefficients $a^{(\nu)}$.

Namely, otherwise

$$\langle \sum_{\nu} a^{(\nu)}_{x_\mu} \xi_\nu \xi_\mu \eta, \eta \rangle$$

takes on negative values. Integrating this expression (for fixed ξ, η) over the torus gives zero, hence for some x this form takes negative values, unless it is identically zero.

This remark shows that in the a priori estimates for v one can only admit finitely many derivatives. This is not only a short coming of the estimates but corresponds to the phenomenon that even for analytic coefficients of the system the solution may admit only finitely many derivatives.

A trivial example of this sort can be given for $n = 1$, $m = 1$ i. e. an ordinary differential equation on the circle :

$$-\sin x\, u_x + bu = (\sin x)^b$$

for which the unique periodic solution is given explicitly by

$$u = -\left(\tan \frac{x}{2}\right)^b \int_{\pi/2}^{x/2} \frac{(\cos t)^{2b-1}}{\sin t}\, dt, \text{ for } 0 < x < 2\pi$$

This function behaves like

$$-\, c x^b \log x$$

at $x = 0$ and so the derivative of order b is unbounded. The derivative of order $\left(b + \frac{1}{2}\right)$ is not square integrable.

The above condition (1.4) requires

$$-\, l \cos x + \left(b + \frac{1}{2} \cos x\right) > 0 \text{ for all } x,$$

which amounts to

$$l < b + \frac{1}{2}.$$

This agrees with the expected number of square integrable derivatives.

§ 2. Nonlinear systems

In this Section we formulate an existence theorem concerning systems which are the generalizations to the nonlinear of positive symmetric systems.

We restrict ourselves to differential equations on the torus. This means that the difficulties of boundary behavior disappear and are ignored.

The systems under consideration are of the form

$$F_k(x, u, u_x) = 0 \quad \text{for} \quad k = 1, 2, \dots, m \tag{2.1}$$

where F_k are of period 2π in $x, \dots, x_n$ and admit sufficiently many derivatives in

$$|y| + |q| \leq 1;$$

where p has nm many components $p_{k\nu}$ corresponding to $\frac{\partial u_k}{\partial x_\nu}$.

We introduce the matrices

$$a_{kl}^{(\nu)}(x, y, p) = \frac{\partial F_k}{\partial p_{l\nu}} = \frac{\partial F_l}{\partial p_{k\nu}} \quad \text{for} \quad \begin{matrix} k, l = 1, \dots, m \\ \nu = 1, \dots, n \end{matrix}$$

$$b_{kl}(x, y, p) = \frac{\partial F_k}{\partial u_l}$$

where we require that the $a_{kl}^{(\nu)}$ are symmetric matrices.

We assume that an approximate solution, say $u = 0$, is known and ask for conditions which ensure that the given system has a solution. This is a perturbation problem and we can take

$$\max_x |F(x, 0, 0)|$$

as the smallness parameter.

Our result can be formulated as follows:

With some $l = l(n)$ we assume that all derivatives up to order l of F are bounded by a constant C for $|y| + |p| < 1$ and all x.

THEOREM: Assume that with the number l above the condition

$$l \langle a_x \rangle + \langle b_0 \rangle > \gamma > 0$$

(see (1.3)) is satisfied for $y = p = 0$. Then there exists a constant $\varepsilon = \varepsilon(n, C, \gamma)$ such that for

$$\sup_x |F(x, 0, 0)| < \varepsilon$$

there exists a periodic solution $u(x)$ of (2.1) which is twice continuously differentiable. The integer $l(n)$ can be chosen as any integer $> \frac{3n}{2} + 6, 15.$

The proof of this result is an application of the general theorem of the previous Section.

We shall verify the main requirements of that theorem, which are listed in Chap. I. Section 5; (5.1) to (5.6).

First we shall show that the solution u of (2.1) can be found in a prescribed C' neighbourhood of 0 if only ε is chosen small enough. For this reason the condition has to be imposed at $y = p = 0$ only. By a continuity argument one see that this condition still holds in a neighbourhood which contains the solution.

We shall show that the solution can even be found in a prescribed C'' neighbourhood. By a general inequality we have

$$| u_n - u_{n-1} |_2 \leq c \, \| u_n - u_{n-1} \|_0^{1-\alpha} \, \| u_n - u_{n-1} \|_r^{\alpha}$$

where

$$\alpha = \frac{n+4}{2r} \text{ if } r > \frac{n}{2} + 2$$

and according to (5.13)

$$| u_n - u_{n-1} |_2 \leq c K_n^{-(1-\alpha)\lambda + \alpha}$$

Hence for

$$\lambda > \frac{\alpha}{1-\alpha} = \frac{n+4}{2r-n-4} \tag{2.3}$$

the exponent is negative and one can ensure that

$$| u_n |_2 \leq \sum_{\nu=1}^{n} | u_\nu - u_{\nu-1} |_2 \leq c' \, K_0^{-(1-\alpha)\lambda+\alpha} = \delta$$

which can be made arbitrarily small. The same holds for $| u |_0$, $| u |_1$. We have to verify therefore the conditions only in this neighbourhood:

$$| u |_0 + | u |_1 + | u |_2 < \delta < 1.$$

The next condition is (5.4), namely that the linearized equation can be solved approximately. Here we make use of the construction described in Chap. I, Section 3, which required the a priori estimates (3.10), Chap I.

For this purpose we have the Lemma of Chap. II, Section 1 available. The condition (1.3) is satisfied for u_n with $| u_n |_0 + | u_n |_1$ sufficiently small. If, moreover.

$$\| u_n \|_r < K_n = K$$

then

$$\| a \|_l + \| b \|_l \leq cK \text{ for } l = r - 1$$

by Chap. I. § 2 (2.6). This means the condition (3.10) is satisfied with $s = l = r - 1$ and therefore an approximate solution with degree of approximation

(2.4) $$\mu = r - 1$$

can be constructed.

Finally we estimate the expression of (5.6): the expression

$$Q = F(u + v, p + q) - F(u, p) - F_u v - F_p v_x$$

can be estimated with the mean value theorem by

$$| Q | \leq c(| v | + | v_x |)^2$$

where c depends on C. Note that $| u |$, $| u_x | < \delta < 1$. Therefore the square integral can be estimated by

$$\| Q \|_0 \leq c(| v |_0 + | v |_1)(\| v \|_0 + \| v \|_1).$$

Using Sobolev's inequalities

$$\| v \|_1 \leq c \| v \|_0^{1 - \frac{1}{r}} \| v \|_r^{\frac{1}{r}}$$

$$| v |_1 \leq c \| v \|_0^{1-\alpha'} \| v \|_r^{\alpha'}, \qquad \alpha' = \frac{n+2}{2r}$$

we find

$$\| Q \|_0 \leq c \| v \|_0^{2-\beta} \| v \|_r^{\beta}.$$

with

(2.5) $$\beta = \frac{1}{r} + \frac{n+2}{2r} = \frac{n+4}{2r}.$$

In remains to investigate whether the condition (2.3) — (2.5) are compatible with Chap. I: (5.8), (5.9).

For $r > \frac{3}{2} n + 6$ we have $\beta < \frac{1}{3}$ by (2.5). Moreover, (2.3) is certainly verified with $\lambda = 1$. Therefore to check (5.9) we set $\lambda = 1$; $\mu = r - 1$ and have

$$0 < \beta < \frac{1}{3} \leq \frac{\lambda}{\lambda+1} \frac{\mu}{\mu+1} \left(1 - 2\frac{\lambda+1}{\mu+1}\right) = \frac{1}{2}\left(1 - \frac{1}{r}\right)\left(1 - \frac{4}{r}\right)$$

which is valid for $r \geq 15$. Therefore we assume

$$r > \frac{3}{2} n + 6, \ 15.$$

We note that with the choice of λ and $r = s + 1$ also (5.22) holds.

The second derivatives converge uniformly in the sequence on account of (2.3), which leads to continuous second derivatives of the solution u.

§ 3 The analytic Case.

We mentioned above that in general the solution of the above equation can only be expected to have finitely many derivatives. In our criterion this corresponds to the fact that the form $\langle a_x \rangle$ can never be positive definite on a closed manifolds.

But we ask the question whether the solution may be analytic for analytic differential equation if $\langle a_x \rangle$ is positive on a domain with boundary. Let D be a domain in the real $(x_1, \dots, x_n)$ space with a smooth boundary. In fact we shall assume at least 2 continuous derivatives for the bounding surface. Let the form

$$\sum_{k, l, \nu, \mu} a^{(\nu)}_{klx_\mu} \xi_\nu \xi_\mu \eta_k \eta_l \geq \gamma_0 |\xi|^2 |\eta|^2 \tag{3.1}$$

with a positive constant γ_0. Moreover, we shall assume that the exterior normal $(N_1, \dots, N_n)$ satisfies at each boundary point

$$(\sum_\nu a^{(\nu)} N_\nu \eta, \eta) \geq 0. \tag{3.2}$$

Then the differential equation possesses a real analytic solution in a subdomain of D provided the functions $F_k(x, u, p)$ are real analytic (**).

The surprising fact is that no boundary conditions are imposed and the solution is unique — but we shall not prove this here.

The reason for this strange phenomenon is that usually the conditions (3.1), (3.2) imply the presence of a singularity and a solution which remains smooth at the singularity is unique.

We shall discuss an example of this type in the following section.

For the proof of this statement we shall establish an a priori estimate in a complex neighborhood of D. Let D_ϱ denote the set of all complex $z = (z_1, \dots, z_n)$ for which there is a $z_0 \in D$ with

$$|z - z_0| < \varrho.$$

(**) Added in proof.: The paper by J. J. Kohn and L. Nirenberg, Noncoercive boundary value problems, Comm. Pure Appl. Math., **18**, pp. 443-492, 1965, Section 9, contains a discussion of this problem in the linear case. Their approach differs from the present one and succeeds also if the coefficients are only C^∞ in which case the solution are also in C^∞.

If ϱ is smaller than the radius of curvature of ∂D every boundary point of D_ϱ has an unique representation

$$z = z_0 + \varrho N \tag{3.3}$$

where N is the (complex) normal.

We assume that all coefficients of L are real analytic and therefore can extend L to functions $u(z)$ which are analytic for $z = x + iy \in D_\varrho$.

As inner product we introduce,

$$(u, v)_0 = \iint_{D_\varrho} \bar{u}\, v\, d\tau$$

where $d\tau = dx_1\, dy_1, \dots, dx_n\, dy_n$ is real volume element in the $2n$ dimensional domain D_ϱ.

We shall prove the following estimate for L.

LEMMA: If the coefficients $a^{(\nu)}$ satisfy (3.1), (3.2) and if

$$(\eta, b_0\, \eta) \geq 2\, \gamma \eta^2 \text{ for } x \in D$$

holds then for ϱ sufficiently small we have

$$\operatorname{Re}\,(u, Lu)_0 \geq \gamma\, (u, u)_0$$

PROOF: A difficulty in the proofs comes from the fact, that while $a^{(\nu)}$ are symmetric for real z, they need not be selfadjoint in the complex. Therefore we shall relate $a^{(\nu)}(z)$ to a symmetric matrix $a_0^{(\nu)}$ in the following manner:

If ϱ is sufficiently small, every $z \in D_\varrho$ can be written in an unique way in the form

$$z = z_0 + rN \text{ where } 0 \leq r < \varrho \text{ and } z_0 \in D_\varrho .$$

Here N is a complex normal vector. For $r = 0$ one obtains points of D_0 and for $r = \varrho$ the boundary points of D_ϱ. We define the matrices

$$a_0^{(\nu)}(z) = a^{(\nu)}(z_0)$$

which are symmetric since $z_0 \in D_0$, hence real.

Moreover, by Taylor's theorem

$$a^{(\nu)}(z) = a_0^{(\nu)}(z) + r \sum_\mu A^{\nu\mu} N_\mu + 0\,(r^2) \tag{3.4}$$

where

$$A^{\nu\mu} = a^{(\nu)}_{x_\mu}(z_0)$$

and the estimate for $0(r^2)$ depends on the second derivatives of a.

To derive our estimate we use the complex Green's formula

$$\int\limits_{D_\varrho} \frac{\partial}{\partial z_\nu} \langle \bar{u}, a^{(\nu)} u \rangle \, d\tau = \frac{1}{2} \sum_\nu \int\limits_{\partial D_\varrho} \bar{N}_\nu \langle \bar{u}, a^{(\nu)} u \rangle \, d\sigma \tag{3.5}$$

where $d\sigma$ is the $2n-1$ dimensional surface element on ∂D_ϱ. The factor on the right is justified by the formula

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y}\right).$$

If we note that u is an analytic function, i. e. satisfies

$$\frac{\partial u}{\partial \bar{z}_\nu} = 0$$

we find

$$\frac{\partial}{\partial z_\nu} \langle \bar{u}, a^{(\nu)} u \rangle = \langle \bar{u}, a^{(\nu)} u_{z_\nu} \rangle + \langle \bar{u}, a^{(\nu)}_{z_\nu} u \rangle$$

and therefore the equation (3.5) gives

$$(u, Lu)_0 - (u, (b - \Sigma\, a^{(\nu)}_{z_\nu})\, u)_0 = \frac{1}{2} \sum_\nu \int\limits_{\partial D_\varrho} \bar{N}_\nu \langle \bar{u}, a^{(\nu)} u \rangle \, d\sigma$$

Taking the real part of this identity and noting that $(\bar{u}, a^{(\nu)}_0 u)$ is real we have

$$\mathrm{Re}\,\{(u, Lu)_0 - (u, (b - \Sigma\, a^{(\nu)}_{z_\nu})\, u)_0 =$$

$$\frac{1}{2} \int\limits_{\partial D_\varrho} \mathrm{Re}\, N_\nu \langle \bar{u}, a^{(\nu)}_0 u \rangle \, d\sigma + \frac{\varrho}{2} \int\limits_{\partial D_\varrho} \bar{N}_\nu \langle \bar{u}, A^{\nu\mu} N_\mu u \rangle \, d\sigma + 0(\varrho^2).$$

Our assumptions guarantee that the two integrals on the right hand side are positive and, dominate the error term, hence

$$\mathrm{Re}\,(u, Lu)_0 \geq \mathrm{Re}\,(u, (b - \Sigma\, a^{(\nu)}_{z_\nu})\, u)_0 .$$

Finally, since the right hand side contains

$$\frac{1}{2}\left(a_{z_\nu}^{(\nu)} + \overline{a_{z_\nu}^{(\nu)*}}\right) = \frac{1}{2}\, a_{0x_\nu}^{(\nu)} + 0\,(\varrho)$$

we can replace, with a small error, $b - \Sigma\, a_{z_\nu}^{(\nu)}$ by b_0 (defined in (1.2) and have

$$\operatorname{Re}(u, Lu)_0 \geq \operatorname{Re}(u, b_0\, u)_0 - c\varrho\,(u, u)_0$$

$$\geq \gamma\,(u, u)_0$$

which proves the Lemma.

As a consequence of this Lemma we have

$$\| Lu \|_0 \geq \gamma\, \| u \|_0$$

and it is again standard to construct weak solutions for the equation

$$Lu = f$$

if f is complex analytic in D_ϱ. However, weak solutions in this case are simply complex analytic functions and therefore classical solutions, in D_ϱ. In this sense this problem is much simpler than the previous one.

We show briefly how to establish the existence of analytic solutions for the nonlinear problem, which in this case is easier than in the previous cases:

Let

$$\mathcal{F}(u) = F(x, u, u_x)$$

and construct iteratively the solution of the linearized equation

$$\mathcal{F}'(u_s)\, v + \mathcal{F}(u_s) = 0, \qquad u_{s+1} = u_s + v$$

according to Newton's method. Assume that u_s is analytic in a complex neighborhood of radius ϱ_s where

$$\varrho_{s+1} = \varrho_s - \delta_s$$

and δ_s will be chosen presently so that

$$\Sigma \delta_s < \varrho/2.$$

Assume that

$$\| \mathcal{F}(u_s) \|_0 < \varepsilon_s .$$

Using the a priori estimates for the linearized differential equation

$$\sum_{\nu} a^{(\nu)} v_{z_\nu} + bv = - F(x, u, u_x)$$

we find

$$\| u_{s+1} - u_s \|_0 = \| v \|_0 \leq \gamma^{-1} \varepsilon_s .$$

This allows one to estimate

$$\| \mathcal{F}(u_{s+1}) \|_0 = \| \mathcal{F}(u_s) + \mathcal{F}'(u_s) v + Q(u_s, v) \|_0 = \| Q(u_s, v) \|_0 .$$

However, the quadratic term involves derivatives of v. Estimating these analogously as before one gets

$$\| Q(u_s, v) \|_0 \leq c (\| v_x \|_0 + \| v \|_0) \max (|v_x| + |v|)$$

$$\leq \gamma^{-1} c \delta_s^{-\left(\frac{n}{2} - 2\right)} \varepsilon_s^2$$

if z is restricted to the complex neighborhood of radius $\varrho_s - \delta_s$. Thus we find

$$\| \mathcal{F}(u_{s+1}) \|_0 \leq \varepsilon_{s+1}$$

if

$$\varepsilon_{s+1} \geq \gamma^{-1} c \delta_s^{-\left(\frac{n}{2} - 2\right)} \varepsilon_s^2 .$$

Setting, for example,

$$\delta_s = s^{-2} \frac{\varrho}{4}$$

we have

$$\sum_{s=1}^{\infty} \delta_s \leq \frac{\varrho}{2}$$

and

$$\varepsilon_{s+1} = c_1 s^{-n-4} \varepsilon_s^2$$

which — for sufficiently small ε_0 — converges to zero.

This proves the convergence of the procedure, in particular, of

$$\| u_{s+p} - u_s \|_0 \leq \gamma^{-1} \sum_{\nu=s}^{s+p-1} \varepsilon_\nu$$

for

$$|z - D| < \varrho/2 < \varrho - \Sigma\delta_s .$$

Thus the solution is analytic.

We summarize the result:

THEOREM: If the functions $F_k(x, u, p)$ are real analytic in a complex neighborhood of $|u| + |p| \leq 1$ and if the matrices

$$a^{(\nu)}(x) = F_{p_\nu}(x, o, o)\,;\; b = F_u(x, o, o)$$

satisfy the conditions. (3.1) and (3.2), then for sufficiently small $\sup\limits_x |F(x, o, o)|$ there exists a real *analytic* solution of the equation of $F(x, u, u_x) = 0$.

§ 4 Invariant Surfaces for Ordinary Differential Equations.

a) We shall illustrate the above results with the perturbation of invariant manifolds for ordinary differential equations.

Let

$$\dot{z} = \varphi(z)$$

be a vector field and let a closed manifold σ be called invariant, if the vector field is tangent at every point of σ. For example, a periodic solution is a one dimensional invariant manifold. We shall, however, be interested in such a manifold for higher dimensions mainly.

This concept of an invariant manifold occurs naturally for slightly coupled oscillations, i. e. systems of the form

$$\text{(4.1)} \qquad \ddot{x}_\nu = f_\nu(x_\nu, \dot{x}_\nu) + \mu g_\nu(x, \dot{x}) \qquad (\nu = 1, 2, \dots, n).$$

For $\mu = 0$ these differential equations are decoupled and represent n second order differential equations. Assume that they each possess a periodic solution which we write in the form,

$$\text{(4.2)} \qquad x_\nu = p_\nu(s_\nu)\,; \quad \dot{x}_\nu = q_\nu(s_\nu)$$

where

$$\dot{s}_\nu = 1.$$

This way it is clear that the s_ν contain n arbitrary initial values, the phases. In the $2n$ dimensional phase space (4.2) represents an n dimensional torus which is invariant under (4.1) for $\mu = 0$.

The problem arises whether such an invariant torus exists for the perturbed equation (4.1) if μ is sufficiently small. This perturbation problem of invariant surfaces has been studied by Diliberto [13], [13′], Bogoliubov and Mitropolski [12], Kyner [14], Hale [14′] and others.

We shall show how our results on positive symmetric systems apply to this situation and give new results beyond the previous ones at the expense of high smoothness requirements.

b) We start with a known n dimensional invariant torus σ_0 of a vector field, given by the unperturbed differential equations.

Introducing the variables $x_1, \dots, x_n$ (mod 2π) in the torus and considering $y_1, y_2, \dots, y_m$ as normal coordinates the differential equations can be written in the form

$$\dot{x} = a_0(x, y)$$

$$\dot{y} = -\, b_0(x, y)\, y$$

where we factored out y since $y = 0$ is assumed to be an invariant surface.

A small perturbation of this differential equation gives rise to differential equation of the form

$$\dot{x} = a(x, y)$$
$$\dot{y} = -\, b(x, y)\, y + c(x, y) \tag{4.3}$$

where $a - a_0$, $b - b_0$, c are small.

We seek an invariant torus σ in the form

$$y = u(x)$$

where u is a vector function of period 2π in x. In order that vector the field is tangential on this torus we require

$$\dot{y} = u_x\, a = -\, b(x, u)\, u + c$$

or,

$$\sum_{\nu=1}^{n} a^{(\nu)} \frac{\partial}{\partial x_\nu} u + b(x, u)\, u = c(x). \tag{4.4}$$

If we compare this system with those considered in the previous section one notices two simplifications:

The matrices $a^{(\nu)}$ are scalar multiples of the identity matrix. This reflects that the characteristic directions at each point is uniquely determined by (4.3). Thus the $a^{(\nu)}$ are trivially symmetric. Secondly the equations are quasi-linear. Both these facts allow for simpler proofs, less stringent smoothness assumption in the exposition of the proof.

The main feature, however, is that the existence of σ will follow from the positive definiteness of b. This has a simple interpretation for (4.3). If $(\eta, b\eta) > 2\gamma \mid \eta \mid^2$ then, for the unperturbed equation the y component decays exponentially, like $c^{-2\gamma t}$. Thus γ measures how fast the surface is approached (along the normal). We shall speak of an asymptotically stable invariant manifold.

On the other hand the functions $a_0^{(\nu)}(x)$ describe the vector field in the torus σ_0. In fact, our condition requires that

$$\langle (r\, \Sigma a^{(\nu)}_{x_\mu} \xi_\nu \xi_\mu + B \mid \xi \mid^2) \eta, \eta \rangle > 2\gamma \mid \xi \mid^2 \mid \eta \mid^2$$

where

$$B = b - \frac{1}{2} \Sigma a^{(\nu)}_{x_\nu} .$$

In order that this condition is verified for *all* r we need that

$$a^{(\nu)}_{x_\mu} = 0$$

i. e. that the $a^{(\nu)}$ are constants. This, indeed, is the case that has been predominantly discussed by the previous authors, except for the works by W. T. Kyner [19] [20].

We see that the number of square integrable derivatives which can be guaranted depends on the biggest eigenvalue α of

$$-\frac{1}{2}(a^{(\nu)}_{x_\mu} + a^{(\mu)}_{x_\nu}) \tag{4.5}$$

and the lowest eigenvalue β_0 of $\frac{1}{2}(B + B^T)$ where

$$B = b - \frac{1}{2} \sum_{v=1}^{n} a^{(v)}_{x_\mu} .$$

If β_0 satisfies the condition

$$r < \frac{\beta_0}{\alpha} \tag{4.6}$$

then the a priori estimate for $(Lu, u)_r$ can be established. Our theorem required that $r > r_0 = \frac{3n}{2} + 14$. This means that β_0 has to be sufficiently large compared to α in order that a twice continuously differentiable solution can be guaranteed ([11]).

We want to interpret the quantity α. If on $y = 0$ the flow is given by

$$\dot{x} = a(x)$$

then we find for the length element $(ds)^2 = \Sigma\, dx_\nu^2$

$$\frac{d}{dt}(ds)^2 = (dx, (a^{(\nu)}_{x_\mu} + a^{(\mu)}_{x_\nu})\, dx) \geq -\alpha\, (ds)^2 \tag{4.7}$$

i. e. α measures how fast characterics approach each other (***).

c) We discuss some specific situations which illustrate also our results in the analytic case.

Let us assume that the differential equations are real analytic and that the invariant manifold is a 2 dimensional torus which is asymptotically stable. For the unperturbed torus the flow

$$\dot{x}_1 = a^{(1)}(x_1, x_2)$$

$$\dot{x}_2 = a^{(2)}(x_1, x_2)$$

can be characterized by a rotation number,

$$\lim_{t \to \infty} \frac{x_2}{x_1} = \omega \qquad (\text{if } a^{(1)} > 0)$$

(introduced by Poincaré, see Coddington, Levinson, Theory of Ordinary Differential Equations, McGraw Hill). If ω is irrational every orbit is dense on the torus. But for rational $\omega = \frac{p}{q}$ there exist closed orbits for which x_2 increases by $2\pi p$ as x_1 increases by $2\pi q$.

([11]) Actually the differentiability requirements can be improved considerably if one uses in place of the L_2 norm the maximum norm which is more appropriate for scalar $a^{(\nu)}$. This has been done in a doctoral dissertation of R. Sacker, [21] NYU, 1964.

(***) Added in proof: We note that the quantities α, β_0 are *not* invariant under coordinate transformations but depend on the choice of a «metric». Recently R. Sacker and the author found conditions which are invariant by choosing the metric in an optimal way. These conditions depend on the flow near the limit sets of the caracteristics only. (See lecture by R. Sacker at the International Symposium on Differential Equations, held at Puerto Rico, Dec. 1965).

Let us consider a situation where $\omega = \frac{2}{3}$ and the torus contains one asymptotically stable and unstable orbit (see figure 1).

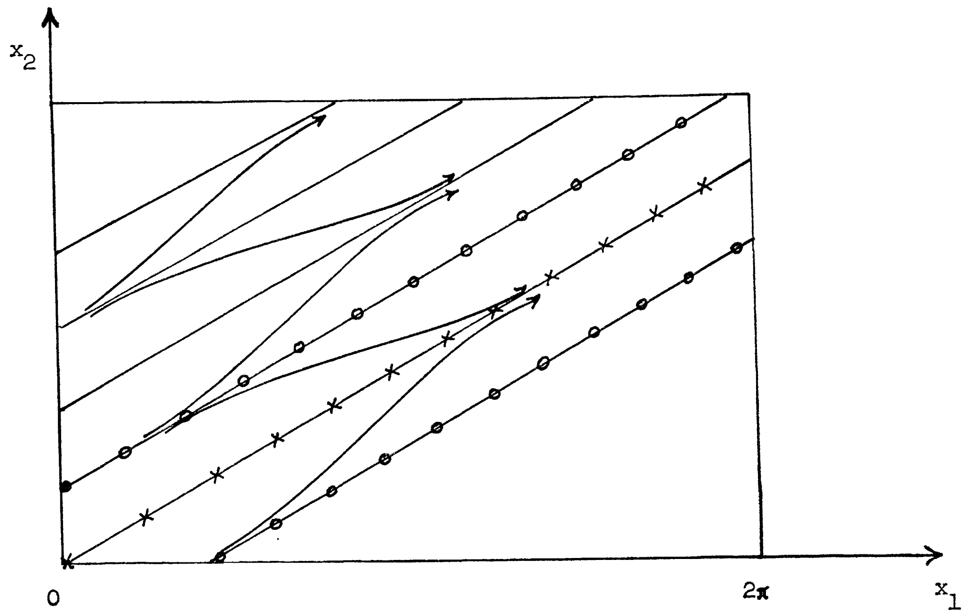

Fig. 1

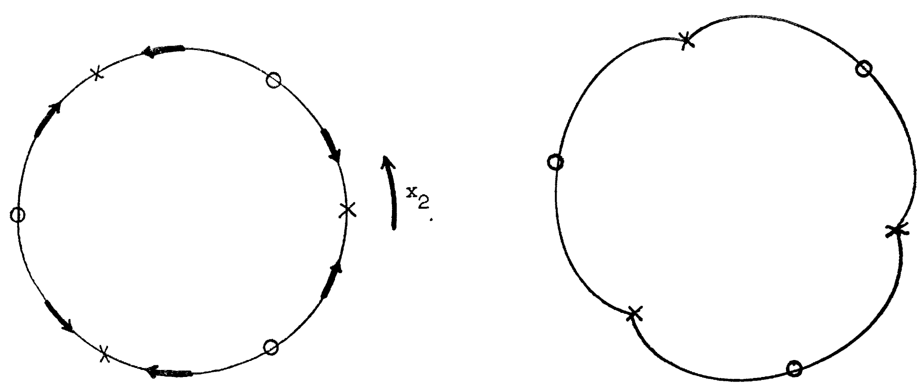

Fig. 2

Then clearly the flow is *spreading* near the unstable orbit (∘) and the matrix

$$a^{(\nu)}_{x_\mu} + a^{(\mu)}_{x_\nu}$$

will be positive definite there. Considering a neighbourhood of the unstable orbit where the characteristics exit at the boundary we can conclude that the invariant torus remains analytic near the unstable (∘) periodic solutions — after a small analytic perturbation is applied. By continuation it follows that the perturbed invariant torus is analytic except at the stable (×) periodic solutions. In fig. 2 we have drawn a cross section of the torus for $x_1 = 0$ before and after perturbation. Thus the perturbed torus consists of different pieces which are analytic and the *discontinuities* of the derivatives occur only at the periodic solutions at which the *characteristics converge*. This is indeed a similar phenomenon as that of shock wave formation when the characteristics form envelopes, except we speak only of discontinuous of higher derivatives, and not of the function u itself. Thus our result allows us to localize the possible position of the discontinuities of the derivatives. They will usually occur at asymptotically stable invariant submanifolds. If one remembers that under parameter change the rotation number will change, in general, and take on rational and irrational numbers one sees the complexity of the phenomenon. However, if one is just interested in the invariant surface and not the smoothness of it, one sees that it will be continuously dependent on a parameter as long as β is large enough compared to α, i. e. the normal approach sufficiently strong compared to the tangential flow.

Another situation of interest is that of an invariant sphere where the flow streams from the north pole to the south pole. After a small pertur-

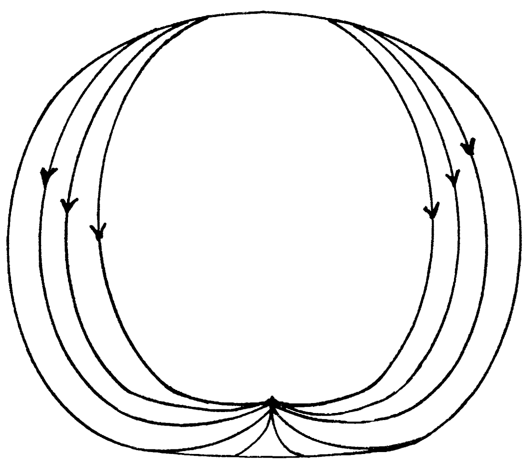

Fig. 3

bation there may develop a discontinuity of the higher derivatives at the south pole. This phenomenon reflects again that these problems in some sense are not well posed. Note that the initial values of the characteristics are not prescribed but it is required that they remain on a manifold. In this case the characteristics issuing at the north pole are well determined. Continuing these orbits it is by no means clear whether they shall fit at the south pole, and indeed the higher derivatives need not fit together down there.

d) Finally we remark that one can also discuss equally well a perturbation theory for such invariant serfaces, as long as the eigenvalues of $\frac{1}{2}(b + b^T)$ remain in

$$| \operatorname{Re} \lambda | > \beta$$

for a sufficiently large β. Such a result was communicated to the author by Kupka.

We indicate how such a result could be derived with the methods presented here. We shall assume that there exists a matrix $\mathcal{T}(x)$ such that

$$(\eta, \mathcal{T}(x)\, b\eta) > \beta\, (\eta, \eta).$$

We assume that $\mathcal{T}(x)$ varies smoothly over the surface.

Then one can derive a priori estimates for

$$(v, \mathcal{T}Lv)_0 \text{ and } (v, \mathcal{T}Lv)_s$$

For example

$$\| v \|_0 \leq c\, (v, \mathcal{T}\, Lv)_0$$

implies

$$\| v \|_0 \leq c\, \| \mathcal{T}\, Lv \|_0 \leq c_1 \| Lv \|_0$$

and the previous arguments hold again. The same remark holds for symmetric systems for which $\mathcal{T}b + b^T \mathcal{T}^T$ is positive definite for some $\mathcal{T}$.

§ 5 A priori estimates for the linear equations.

a) We supply the proof of the estimates given in the Lemma of § 1. For $l = 0$ this inequality has been derived before and here we concentrate on the estimates for higher derivatives. We shall assume that the assumption (1.3) holds and that

$$(5.1) \qquad | a |_0 + | a |_2 + | b |_0 + | b |_1 < c_0 \text{ and } | v |_0 < c_0 \,.$$

The dependence of the estimates on $\| a \|_l$, $\| b \|_l$ we shall make explicit, however. The details are somewhat lengthy but standard. Similar ideas are explained and used in the book by L. Hörmander on Linear Partial Differential Operators, Springer 1963.

We shall restrict ourselves to even $l = 2k$ and write $(-\Delta)^k = P$, $(-\Delta)^k v = w$. The quantity to be estimated is

$$(v, Lv)_l = (v, (-\Delta)^l Lv)_0 = (Pv, PLv)_0$$

We form the divergence expression

$$\frac{1}{2} \sum_\nu \frac{\partial}{\partial x_\nu} \langle Pv, a^{(\nu)} Pv \rangle = \frac{1}{2} \langle Pv, \Sigma\, a^{(\nu)}_{x_\nu} Pv \rangle + \sum_\nu \left\langle Pv, a^{(\nu)} \frac{\partial}{\partial x_\nu} Pv \right\rangle.$$

The integral of this expression vanishes. Therefore with $b_0 = b - \frac{1}{2} \sum_{\nu=1}^{n} a^{(\nu)}_{x_\nu}$ and $w = Pv$

$$\begin{cases} (v, Lv)_l = \sum_\nu \left(w, P\left(a^{(\nu)} \dfrac{\partial}{\partial x_\nu} + b\right) v\right)_0 - \sum_\nu (w, a^{(\nu)} Pv_{x_\nu})_0 - \dfrac{1}{2} \Sigma\, (w, a^{(\nu)}_{x_\nu} Pv)_0 \\ \qquad = \sum_\nu (w, (Pa^{(\nu)} - a^{(\nu)} P)\, v_{x_\nu})_0 + (w, (Pb - bP)\, v)_0 + (w, b_0\, w)_0 . \end{cases}$$

To extract the principal terms we compute the terms of order l in $(Pa^{(\nu)} - a^{(\nu)} P)\, v_{x_\nu}$. One finds

$$a^{(\nu)}_{x_\mu} (-\Delta)^{k-1} \frac{-\partial^2}{\partial x_\nu\, \partial x_\mu} v + \dots$$

where the terms not written contain derivatives of order $< l$ in v. The term $(Pb - bP)\, v$ consists only of such terms and we have

$$(v, Lv)_l = E + (w, \Phi)_0 \tag{5.2}$$

where

$$E = -\left(w, b_0\, \Delta + \sum_{\nu,\mu} a^{(\nu)}_{x_\mu} \frac{\partial^2}{\partial x_\nu\, \partial x_\mu}\right) (-\Delta)^{k-1} v)_0 . \tag{5.3}$$

the expression Φ contains only derivatives of order $< l$ of v.

b) Therefore we can estimate $\|\Phi\|_0$ by $\|v\|_{l-1}$ but the constant will depend on the size of the high derivatives of a and b. To make this dependence explicit we study the terms in Φ more carefully. We assumed (5.1) and $|v|_0 < c_0$ for some c_0 and the following constants will depend on c_0.

Consider the terms containing b in Φ. They have the form

$$D^{l-\lambda}\, b\, D^{\lambda} v = D^{\mu}(Db)(D^{\lambda} v)\,; \qquad (\lambda = 0, 1, \dots, l-1)$$

with $\lambda + \mu = l - 1$ where D stands for any first order differential operator.

We use Hölder's estimate to get

$$\| D^{\mu}(Db)(D^{\lambda} v)\|_0 \leq \left(\int | D^{\mu}(Db)|^{\frac{2(l-1)}{\mu}} dx\right)^{\frac{\mu}{2(l-1)}} \left(\int | D^{\lambda} v|^{\frac{2(l-1)}{\lambda}} dx\right)^{\frac{\lambda}{2l-2}}.$$

Using (2.3) from Chap. I and $|b|_1 \leq c_0$ we find

$$\| D^{\mu}(Db)\, D^{\lambda} v\|_0 \leq c_1 (\| Db\|_{l-1}^{\frac{\mu}{l-1}} \| v\|_{l-1}^{\frac{\lambda}{l-1}} + 1)\ (^{12})$$

$$\leq c_1 (\| Db\|_{l-1} + \| v\|_{l-1} + 1)$$

$$\leq c_1 (\| b\|_l + \| v\|_{l-1} + 1)$$

where c_1 depends on c_0. Similary the terms

$$D^{l-\lambda} a D^{\lambda} v_x = D^{\mu}(D^2 a)\, D^{\lambda} v_x\,, \qquad \lambda = 0, 1, \dots, l-2$$

with $\lambda + \mu = l - 2$ can be estimated by

$$\| D^{\mu}(D^2 a)(D^{\lambda} v_x)\|_0 \leq c_2 (\| D^2 a\|_{l-2} + \| v_x\|_{l-2} + 1)$$

$$\leq c_2 (\| a\|_l + \| v\|_{l-1} + 1)$$

where we used that $|a|_2 < c_0$.

Thus we find

$$\| \Phi\|_0 \leq c_3 \{\| a\|_l + \| b\|_l + 1 + \| v\|_{l-1}\}$$

which makes the dependence of $\| \Phi\|_0$ on the high derivatives explicit. Notice the *linear* dependence on $\| a\|_l$, $\| b\|_l$. Of course, $\| v\|_{l-1}$ can be replaced by

$$\| v\|_{l-1} \leq c_4 (1 + \| v\|_l^{1-\frac{1}{l}})$$

---

(12) In (2.3) of Chap. I the right hand side can be estimated by $c\| v\|_r^{\varrho/r}$ if $| v|_0 < 1$.

which gives an exponent < 1. With

$$K = \| a \|_l + \| b \|_l + 1$$

we have the inequality

$$(5.4) \qquad (v, Lv)_l \geq E - c_5 (K + \| v \|_l^{1-\frac{1}{l}}) \| v \|_l .$$

c) It remains to estimate E. If E had constant coefficients one could obtain the estimate

$$E \geq 2\gamma \| v \|_l^2$$

by Fourier transformation using the assumption (1.3).

Using a wellknown trick of Garding [13] to apply the above inequality to ζv where ζ is a partition of unity one obtains

$$(5.5) \qquad E \geq \frac{3}{2} \gamma \| v \|_l^2 - c_6$$

since a, b are continuously differentiable.

The constant c_6 depends on c_0 and γ. Combining (5.4), (5.5) we have

$$(v, Lv)_l \geq \frac{3}{2} \gamma \| v \|_l^2 - c_6 - c_5 \| v \|_l \left(K + \| v \|_l^{1-\frac{1}{l}} \right)$$

If $\| v \|_l$ is very large, the first term dominates the last. If $\| v \|_l$ is small then the last terms can be estimated by a constant c_1. More precisely, one finds in both cases

$$(v, Lv)_l \geq \gamma \| v \|_l^2 - c_7 K^2$$

which proves the Lemma of § 1.

§ 6 Quasilinear Differential equations.

If the system of differential equations is quasi linear, i. e. $F_k(x, u, p)$ are linear functions of p then one can devise an iteration method where no loss of derivatives occurs, i. e. where the approximation remains in a

---

(13) See, for example, HÖRMANDER's book, p. 190.

fixed sphere

$$\| u \|_l < K$$

while in the previous construction the high derivatives could tend to infinity. This construction which we shall describe is abstracted from Schauder's paper on hyperbolic differential equations [1].

Actually he uses ultimately the celebrated fixed point theorem of Leray-Schauder where the above sphere is mapped into itself. However, his approach can be turned into an iteration procedure which we shall describe. Schauder applied his method to quasi linear differential equations and noted that the general nonlinear case was not accessible to his method. It seems that in this nonlinear case the results described here are new. It would be desirable to apply these methods also to the hyperbolic equations.

We assume then that

$$F(x, u, p) = a(x, u)\, p + b(x, u) \tag{6.1}$$

is linear in p. Let $u_0 = 0, u_1, \ldots, u_s = u$ be known approximations and we construct a better approximation

$$u_{s+1} = U = u + v$$

by solving the equation

$$F_p(x, u, p)\, v_x + F_u(x, o, o)\, v + F(x, u, p) = o \tag{6.2}$$

The first term corresponds to linearization at $u = u_s$ and the second to a linearization at $u_0 = 0$. It is a mixture of Picard's and Newton's method. In order to show that this construction avoids a loss of a derivatives we express (6.2) in terms of $U = u + v$. We use that in the quasilinear case the expression

$$F(x, u, p) - F_p(x, u, p)\, u_x - F_u(x, 0, 0)\, u$$

$$= b(x, u) - b_u(x, 0, 0)\, u = g(x, u)$$

is independent of u_x. Adding this identity to (6.2) we find

$$F_p(x, u, p)\, U_x + F_u(x, 0, 0)\, U + g(x, u) = 0. \tag{6.3}$$

From this equation one reads off: if u has l square integrable derivatives, the a priori estimates which we postulated will give l square integrable derivatives for U. In fact, if K is chosen large enough one can establish from (6.3) for $U = u_{s+1}$ an inequality

$$\| U \|_l < K$$

The process (6.2) leads to linear convergence since the term $F_u(x, 0, 0)\, v$ contains a linear error in v. It is remarkable, that if one increases the accuracy and replaces this term by $F_u(x, u, p)\, v$ then the corresponding equation for $U = u + v$ is

$$F_p(x, u, p)\, U_x + F_u(x, u, p)\, U + g(x, u) - a_u\, u_x\, u = 0$$

and so leads to a loss of derivatives, since the right hand side contains u_x.

Therefore one can consider the construction (6.2) as a less accurate one which, however, preserves smoothness. In the method described in Chap. I this smoothness is provided more systematically so that even fast convergence can be assured. I am not aware of a construction analogous to (6.2) which avoids loss of derivatives and is applicable to the *general nonlinear* case.

(*To continue*)

REFERENCES

1. J. SCHAUDER, *Das Anfangswertproblem einer quasilinearen hyperbolischen Differentialgleichung zweiter Order in beliebiger Anzahl von unabhängigen Veränderlichen.* Fund. Math. **24**, 1935, pp. 213-246.

2. K. O. FRIEDRICHS, *Symmetric Hyperbolic Linear Differential Equations.* Comm. Pure Appl. Math. **7**, 1954, pp. 345-392.

3. J. NASH, *The Imbedding of Riemannian Manifolds,* Ann. Math. **63**, 1956, pp. 20-63.

4. J. SCHWARTZ, *On Nash's Implicit Function Theorem,* Comm Pure Appl. Math. **13**, 1960, pp. 509-530.

4′. J. MOSER, *A New Technique for the Construction of Solutions of Nonlinear Differential Equations.* Proc. Nat. Acad. Sciences, Vol. **47**, No. 11, pp. 1824-1831, 1961.

4″. J. MOSER, *On Invariant Curves of Area-Preserving Mappings of an Annulus.* Nachr. Akad. Wiss. Göttingen Math. Phys. Kl. IIa, No. 1, 1962, pp. 1-20.

5. C. L. SIEGEL, *Iteration of Analytic Functions.* Ann. of Math. **43**, 1942, pp. 607-612.

5′. C. L. SIEGEL, *Über die Normalform Analytischer Differentialgleichungen in der Nähe einer Gleichgewichtslösung,* Nachr. Akad. Wiss., Göttingen, Math. Phys. Kl, IIa, 1952, pp. 21-30.

6. A. N. KOLMOGOROV, Doklad. Nauk, *USSR* **98**, 1954, p. 527-530.

7. A. N. KOLMOGOROV, *General Theory of Dynamical Systems and Classical Mechanics.* Proc. of Intl. Congress of Math., Amsterdam, 1954, (Amsterdam: Erven P. Nordhoff, 1957), Vol. **1**, pp. 315-333.

8. V. I. ARNOLD, *«Small divisors» I, On mappings of a circle onto itself,* Isvest. Akad. Nauk, ser. mat., **25**, No. 1, 1961, pp. 21-86.

9. V. I. ARNOLD, *Small Divisor and Stability Problems in Classical ana Celestial Mechanics.* Uspekhi Mat. Nauk USSR, Vol. **18**, Ser. 6 (119), 1963, pp. 81-192.

10. V. I. ARNOLD, *Proof of A N. Kolmogorov's Theorem on the Preservation of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian.* Uspekhi Mat. N. USSR, Vol. **18**, Ser. 5 (113), 1963, pp 13-40.

11. K. O. FRIEDRICHS, *Symmetric Positive Linear Differential Equations,* Comm. Pure Appl. Math. **11**, 1458, pp. 333-418.

12. N. N. BOGOLIOUBOV, and Y. A. MITROPOLSKI, *The Method of Integral Manifolds in Nonlinear Mechanics,* Contrib. to Differential Equations, Vol. II, 1963, pp. 123-196.

13. S. P. DILIBERTO, *Perturbation Theorems for Periodic Surfaces.* I, Rend. del Circ. Mat. Palermo (2), **9**, 1961, pp. 265-299 II, Rend. del Circ. Mat. Palermo (2), **10**, 1962, pp. 111-161.

13′. S. P. DILIBERTO, *Pertubation Theory of Invariant Surfaces I-IV.* Mimeographed ONR Reports, Berkeley, 1956/57.

14. KYNER, W. T., *Invariant manifolds,* Rend. del Circolo Matematico Palermo, Ser. II, Vol. **9**, 1961, pp. 98-110.

14′. J. K. HALE, *Integral Manifolds of Perturbed Differential Systems,* Annals of Math., **73**, 1961, pp. 496-531.

15. I. KUPKA, *Stabilité des varietés invariants d'un champ de vecteurs pour les petites perturbations.* Compt. Rend. Acad. Sc. Paris, **258**, Group 1, 1964, pp. 4197-4200.

16. L. Nirenberg, *On elliptic partial differential equations.* Ann. Scuola Norm. Sup. Pisa, Ser. 3, **13** (1959), pp. 116-162.

17. Gagliardo, E., *Ulteriori proprietà di alcune classi di funzioni in più variabili,* Ricerche Mat., **8**, 1959, pp. 24-51.

18. J. J. Kohn and L. Nirenberg, *Non-coercive boundary value problems,* to appear in Comm. Pure Appl. Math. (See under « elliptic regularization »).

19. W. T. Kyner, *A fixed point theorem. Contributions to the Theory on Nonlinear Oscillations,* Vol. **3**, Princeton University Press, Annals of Math. Studies 36, 1956, pp. 197-205.

20. W. T. Kyner, *Small Periodic Perturbations of an Autonomous System of Vector Equations. Contributions to the Theory of Nonlinear Oscillations* Vol. **4**, Princeton, 1958.

21. R. Sacker, *On Invariant Surfaces and Bifurcation of Periodic Solutions of Ordinary Differential Equations,* NYU Report, IMM-NYU 333, Oct. 1964, to be published in Comm. Pure Appl. Math.

22. C. L. Siegel, *Vorlesungen über Himmelsmechanik,* Springer, 1956.

23. H. Cremer, *Über die Häufigkeit der Nichtzentren,* Math. Ann. **115**, pp. 607-611, 1942.

24. N. Levinson, *Transformation of an analytic function of several variables to a canonical form.* Duke Math. Journal **28**, 1961, pp. 345-353. (See Theorem 2).

25. N. I. Achieser, *Vorlesungen über Approximationstheorie,* Akd. Verlag. Berlin 1953, translated from Russian, Chapter V.