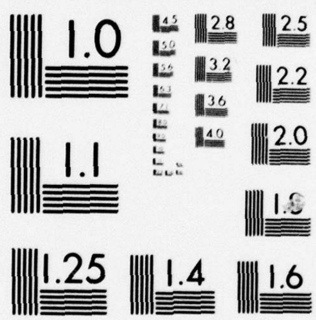
AD-A060 618 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
ON THE DEVELOPMENT OF OPTIMIZATION THEORY.(U)
APR 78 A PREKOPA DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1842 NL

| OF |
AD
A060618

END
DATE
FILMED
01-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #1842

ON THE DEVELOPMENT OF OPTIMIZATION THEORY

András Prékopa

Mathematics Research Center
University of Wisconsin–Madison
610 Walnut Street
Madison, Wisconsin 53706

April 1978

(Received February 15, 1978)

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

ON THE DEVELOPMENT OF OPTIMIZATION THEORY

András Prékopa

Technical Summary Report #1842
April 1978

ABSTRACT

The method of Lagrange for finding extrema of functions subject to equality constraints was published in 1788 in the famous book Mécanique Analytique. The works of Karush, John, Kuhn and Tucker concerning optimization subject to inequality constraints appeared more than 150 years after that. The purpose of this paper is to call attention to important papers, published as contributions to mechanics, containing fundamental ideas concerning optimization theory. The most important works in this respect were done primarily by Fourier, Cournot, Farkas and further by Gauss, Ostrogradsky and Hamel.

AMS(MOS) Subject Classification - 15-03, 15A39, 90-03, 90C30

Key Words - Fourier principle, Farkas theorem, nonlinear programming, linear inequalities

Work Unit Number 5 - Mathematical Programming and Operations Research

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

SIGNIFICANCE AND EXPLANATION

This is a historical paper. It concerns the necessary conditions of optimality for nonlinear optimization.

The solution of nonlinear programming problems is based on the necessary conditions formulated for the case of equality constraints by Lagrange in 1788 and for the case of inequality constraints by several authors between 1939 and 1950 (Karush, John, Kuhn, Tucker). This paper tries to complete the general historical picture by showing that, in the form of contributions to mechanics, the basic ideas were already discovered jointly by Fourier, Cournot and Farkas, among others.

Many referenced papers are thoroughly analyzed.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

ON THE DEVELOPMENT OF OPTIMIZATION THEORY

András Prékopa

1. Introduction

Farkas' famous paper of 1902 published in Grelle's Journal became a principal reference concerning linear inequalities after the publication of the paper of Kuhn and Tucker, "Nonlinear Programming" in 1950 [53]. In this latter paper the fundamental theorem on linear inequalities, Farkas' theorem, was used to derive necessary conditions for optimality for the nonlinear programming problem. The results obtained enabled rapid development in nonlinear optimization theory. The work of John containing similar but weaker results for optimality published in 1948, has been generally known but it was not until a few years ago that Karush's work of 1939 became widely known where essentially the same result was given as by Kuhn and Tucker in 1951.

In this paper we call attention to some important works done already in the last century and before. We show that the fundamental ideas concerning the necessary optimality conditions for nonlinear optimization subject to inequality constraints can be found in papers primarily by Fourier, Cournot, Farkas and further by Gauss, Ostrogradsky and Hamel.

To start to disclose the early development of optimization theory it is very helpful to have a glance at the first two sentences in Farkas' paper [23]:

> "Die naturgemässe und zugleich systematische Behandlung der analytischen Mechanik muss das zuerst von Fourier und dann später von Gauss formulierte Ungleichheitsprincip der virtuellen Verschiebungen zur grundlage haben.
>
> Die Möglichkeit einer solchen Behandlung erfordert aber gewisse Kenntnisse uber die homogenen linearen ganzen Ungleichungen, welche bisher so zu sagen, gänzlich gefehlt haben."

Here we see that Farkas had a definite purpose for developing the theory of linear inequalities. He was professor of theoretical physics at the University of Kolozsvár. We can very well assume that he himself had already applied his

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

inequality theorem to the problem of mechanical equilibrium and in fact he reported first "on the applications of the mechanical principle of Fourier" at the session of the Hungarian Academy held on December 17, 1894. This was published later in Hungarian [13] and in the German language [14]. We shall analyze this paper later. First we have to say a few words about those mechanical principles which played an important role in the development of optimization theory.

2. On the principles of mechanical equilibrium

The principle of virtual work was enunciated by Johann Bernoulli in 1717. It appeared first in a book by Varignon in the same year. Let us quote Lagrange [55/I.p.21,22]:

> "Le principe des vitesses virtuelles peut être rendu très général de cette manière:
>
> Si un système quelconque de tant de corps ou points que l'on veut, tirés chacun par des puissances quelconques, est en équilibre, et qu'on donne à ce système un petit mouvement quelconque, en vertu duquel chaque point parcoure un espace infiniment petit qui exprimera sa vitesse virtuelle, la somme des puissances, multipliées chacune par l'espace que le point où elle est appliquée parcourt suivant la direction de cette même puissance, sera toujours égale à zéro, en regardant comme positifs les petits espaces parcourus dans le sens des puissances, et comme négatifs les espaces parcourus dans un sens opposé.
>
> Jean Bernoulli est le premier, que je sache, qui ait apercu cette grande généralité du principe des vitesses virtuelles, et son utilité pour resoudre les problèmes de Statique. C'est ce qu'on voit dans une de ses Lettres à Varignon, datée de 1717, que ce dernier a placée à la tête de la Section neuvième de sa Nouvelle Mécanique, Section employée tout entière à montrer par différentes applications la vérité et l'usage du principe dont il s'agit."

This principle was considered by Lagrange as an axiom of mechanics. On page 27 of the same work we read:

> "C'est dans cette loi que consiste ce qu'on appelle communément le principe des vitesses virtuelles, principe reconnu depuis longtemps pour le principe fondamental de l'equilibre, ainsi que nous l'avons montré dans la Section precedente, et qu'on peut, par consequent, regarder comme une espece d'axiome de Mécanique."

For the case of a conservative system of forces i.e. when the forces are given as negative partial derivatives of a scalar function, the principle, first enunciated by Courtivron, applies. Concerning this, Lagrange writes the following [55/I., p.70]:

> "...ce qui donne cet autre principe de Statique, que, de toutes les situations que prend successivement le système, celle où il a la plus grande ou la plus petite force vive est aussi celle où il le faudrait placer d'abord pour qu'il restât en équilibre. [Voir Courtivron, les Mémoires de l'Académie des Sciences de 1748 et 1749.]"

Lagrange gave sufficient conditions that the potential takes it minimum. As Bertrand remarked, his proof was incomplete and Dirichlet later gave a correct proof [10,11] for the theorem involved.

The mechanical principle of Fourier was first published in 1798 in his paper "Mémoire sur la Statique". This concerns the case of inequality constraints. Farkas remarked [14, p.458] that the declaration of the inequality principle is not the main contribution of the paper. In fact it has the subtitle: "Contenant la démonstration du principe des vitesses virtuelles". This very general "proof" has not been accepted as a proof, however (see e.g. [75]), thus the main merit of the paper is still, contrary to Farkas' remark, the declaration of the inequality principle. We quote page 488 of Fourier's paper [34]:

"Comme il *arrive* souvent que les points du système s'appuient seulement sur les obstacles fixes, sans y être attachés, il est évident qu'il y a des déplacements possibles qui ne satisfont pas aux équations de condition : on voit encore que, par ces déplacements, le moment des résultantes est nécessairement positif, puisque la *direction* de *ces forces* doit être perpendiculaire aux surfaces resistantes. Ainsi la somme des moments des forces appliquées est positive pour tous les déplacements de cette espèce; mais il est impossible que l'on dérange un corps dur, en équilibre, de sorte que le moment total des forces appliquées soit *négatif*. Au reste, si l'on considère les résistances comme des forces, ce qui fournit, comme on le sait, le moyen d'estimer ces résistances, le corps peut être regardé comme libre, et la sommes des moments est nulle pour tous les déplacements possibles."

The moment of the forces $P,Q,R,\ldots$ acting on a mechanical system is defined as the sum of scalar products

$$P\delta p + Q\delta q + R\delta r + \ldots \tag{1}$$

where $\delta p, \delta q, \delta r$ are variations of the displacements. In a small neighborhood of the point describing the state of the system these variations - roughly speaking - do not violate the constraints. The Bernoulli principle declares (1) to be equal to zero, while the Fourier principle declares (1) to be less than or equal to zero in case of equilibrium. Fourier was using the "fluxion" (as it turns out from other parts of his paper) instead of the moment (1), these two being negatives of each other. This explains why Fourier in his principle required (1) to be nonnegative.

If a potential V exists, i.e. if we have

$$P = -\frac{\partial V}{\partial p}, \quad Q = -\frac{\partial V}{\partial q}, \quad R = -\frac{\partial V}{\partial r}, \ldots$$

where the derivatives on the right hand sides denote vectors, then the requirement that (1) be less than or equal to zero takes the form:

$$\frac{\partial V}{\partial p}\delta p + \frac{\partial V}{\partial q}\delta q + \frac{\partial V}{\partial r}\delta r + \ldots \geq 0. \tag{2}$$

On the left hand side we have a total differential. A correct mathematical proof - according to our present standard - cannot be found in Fourier's work. We know, on the other hand, that (2) cannot be derived as a necessary condition that V takes its minimum, without a constraint qualification. Fourier remarked that in case his principle reduces to (2), we can find the equilibrium state by minimizing (!) the function V.

In 1829 Gauss [38] again enunciated the inequality principle without mentioning Fourier. Gauss seems to have observed the necessity of a constraint qualification. He added a footnote to the enunciation of the inequality principle where he required the constraints to be of special type. We quote Gauss [38, p.234]:

> "Nach dem Princip der virtuellen Geschwindigkeiten erfordert dies Gleichgewicht, dass die Summe der Producte aus je drei Factoren, nemlich jeder der Massen m, m', m" u.s.w., den Linien cb, c'b', c"b" u.s.w., und irgend welchen auf letztere resp. projicirten, vermöge der Bedingungen des Systems möglichen Bewegungen jener Punkte, immer = 0 sei, wie man es gewöhnlich ausspricht, oder richtiger, dass jene Summe niemals positiv werden könne."

> Footnote:

> "Der gewöhnliche Ausdruck setzt stillschweigend solche Bedingungen voraus, dass die jeder möglichen Bewegung entgegengesetzte gleichfalls möglich sei, wie z. B., dass ein Punkt auf einer bestimmten Fläche zu bleiben genöthigt, dass die Entfernung zweier Punkte von einander unveränderlich sei u. dergl. Allein dies ist eine unnöthige und der Natur nicht immer angemessene Beschränkung. Die Oberfläche eines undurchdringlichen Körpers zwingt einen auf ihr befindlichen materiellen Punct nicht, auf ihr zu bleiben, sondern verwehrt ihm bloss das Austreten auf die Eine Seite; ein gespannter, nicht ausdehnbarer aber biegsamer Faden zwischen zwei Punkten macht nur die Zunahme, nicht die Abnahme der Entfernung unmöglich u. s. w. Warum wollten wir also das Gesetz der

virtuellen Geschwindigkeiten nicht lieber *gleich anfangs* so ausdrücken, dass es alle Fälle umfasst?"

This requires - according to the interpretation of the author of this paper - that the point where the equilibrium is reached should be either an interior point of the set determined by the constraints or a point at which only one of the constraints is active.

In the year 1834 at a session of the "Académie Impériale des Sciences de Saint-Pétersbourg", Ostrogradsky also enunciated the inequality principle. Voss remarks [75, p.74] that "daher heisst in Russland das Prinzip von Fourier auch wohl das von Ostrogradsky." It is interesting to read what Ostrogradsky wrote about the inequality principle:

> "Il est très surprenant de voir que dans la nouvelle édition de la *Mécanique analytique*, édition publiée à l'époque où l'on connaissait déjà toute l'étendue du principe des vitesses virtuelles, Lagrange non seulement n'a fait aucun usage de l'observation, que dans l'équilibre des forces le moment total pouvait acquérir une valeur négative, mais qu'il l'a en quelque sorte écartée quand elle s'est présentée comme d'elle même, dans la démonstration qu'il a donnée du principe des vitesses virtuelles; cependant, faute d'y avoir eu égard, ce grand géomètre a incomplètement énuméré les *déplacemens* possibles dans la plupart des questions de la première partie de la Mécanique analytique, et il est facile de reconnaître que les déplacemens qu'il a négligé de considérer, ne sont empêchés par aucune condition, en sorte que, toutes les équations qu'il a établies pour l'équilibre étant satisfaites, l'équilibre *cependant* pourrait *n*'avoir pas lieu.
>
> Nous nous proposons, dans ce mémoire, d'exposer l'analyse relative à l'emploi du principe des vitesses virtuelles considéré dans toute sa généralité et de compléter la solution de plusieurs questions traitées dans la première partie de la *Mécanique analytique*."

Lagranges's method of multipliers was published in the first volume of Mécanique Analytique (p. 77-79) as a tool to find the equilibrium of a mechanical system using Bernoulli's principle. The validity of the method was proved by Lagrange entirely on an algebraic basis. (This ingenious method was unable to attract mathematics students and teachers for a long time. The method of presentation used nowadays by many instructors is to prove first the necessary condition in case of inequality constraints, give a geometric meaning to this and then refer to the case of equality constraints. This can make the students more enthusiastic toward this theory). We shall see in Sections 4-7 what happened to Fourier's principle. Now we devote a few pages to Farkas' theorem.

3. The theorem of Farkas on linear inequalities

Gyula Farkas was born in Sárosd, Hungary, in 1847 and died in Pestszentlőrinc, Hungary, in 1930. First he studied law at the University of Budapest and music at the same time until he decided to study mathematics and physics at the Faculty of Philosophy. He received a high school teacher's diploma in 1876. He became a professor of theoretical physics at the University of Kolozsvár in 1887 and held this position until 1915. He was active in organization too, holding the positions of dean of his faculty and of rector of the university for some time. In 1898 he was elected a corresponding member and in 1914 an ordinary member of the Hungarian Academy of Sciences. Due to increasing eye disease he decided to retire from his professorship and moved to Pestszentlőrinc in 1915. As a young man he wrote several studies in numerical analysis, differential equations, elliptic functions and music theory. His most important scientific contributions are those concerning the mechanical principle of Fourier, the theory of linear inequalities and various problems of thermodynamics. Commemorating him in 1974, the János Bolyai Mathematical Society founded a Farkas prize to honor young mathematicians for their scientific contribution to applied mathematics. For further information concerning the life and work of Farkas see [32,33,65,66,73].

Farkas' 1902 paper [23] is generally referred to concerning the fundamental theorem on homogeneous linear inequalities. In what follows we shall write vectors in column form and the transpose will be denoted by a prime.

Let $g_1,\ldots g_M,g$ denote m-component vectors and form the following homogeneous linear inequalities:

(3.1) $$g_i'x \geq 0, i = 1,\ldots,M,$$

(3.2) $$g'x \geq 0.$$

If (3.2) holds for every x for which all inequalities in (3.1) hold, then we say that the inequality (3.2) is a *consequence* of the system of inequalities (3.1). Farkas' theorem is the following.

Theorem 3.1. The inequality (3.2) is a consequence of the inequalities (3.1) if and only if there exist nonnegative numbers $\lambda_1,\ldots,\lambda_M$ such that

$$g = \lambda_1 g_1 + \ldots + \lambda_M g_M.$$

Farkas published this theorem first in 1895 [13,14]. However, that proof contains a gap. He gave the first good proof in 1896 in the paper [15] published in the Hungarian language. Essentially the same paper appeared in German in 1899 [17]. Following this first proof, he gave a simple proof and published it in German [19]. This latter proof is included in his best known Crelle's Journal paper [23].

Very likely Farkas never recognized that it is not at all easy to complete his first proof. In a paper [29] which appeared in 1918 he wrote that there are six proofs for this theorem. He referred to the works of Minkowski [60] and Haar [42] to three proofs mentioned above and to a proof given in his university lecture notes (not available to the author of this paper). His last paper on linear inequalities [31] appeared in 1926 (a short paper on some geometrical aspects of his theorem). It is instructive to see what kind of error he made. We shall quote from [14].

First he shows that we may suppose, that there are as many linearly independent inequalities in the system (of which one consequence is considered) as the number of variables. Then he represents the coefficient vector of the consequence inequality as a linear combination of the other coefficient vectors. Now, if, for the sake of simplicity, the first n inequalities are linearly independent, then we can express the multipliers belonging to these in terms of the others in the following manner

$$\begin{aligned} \lambda_1 &= I_0 + I_1\lambda_{n+1} + I_2\lambda_{n+2} + \dots \\ \lambda_2 &= K_0 + K_1\lambda_{n+1} + K_2\lambda_{n+2} + \dots \qquad (3.4) \\ \lambda_n &= \dots \end{aligned}$$

Farkas continues to write [14, p. 270]:

"Nur dann ist es unmöglich, dass alle λ-Grössen zugleich nicht-negative Werte erhalten können, wenn in (8) ((3.4) in this paper, A.P.) wenigstens eine rechte Seite- oder eine Summe von nicht-negativen vielfachen der rechten Seiten die doppelte Beschaffenheit aufweist, dass ihr erstes Glied Negativ ist, und die Coëfficienten ihrer übrigen Glieder entweder negativ sind, oder höchstens verschwinden (da durch Eliminationen positiver Glieder mittels entsprechender negativen, aus einem mit positiven λ durchaus unerfüllbarem Teile von (5) ((3.2 in this paper, A.P.) nach und nach aequivalente Teil-Systeme erhalten werden können). In diesem Falle hat aber (4) ((3.1) in this paper, A.P.) schon Auflösungen, durch welche (5) ((3.2) in this paper, A.P.) nicht befriedigt wird. Sei es nämlich, dass diese doppelte Eigenschaft der rechten Seite der ersten Gleichung in (8) ((3.4 in this paper, A.P.) zukommt, dass also

$$I_0 < 0; \quad I_1 \leqq 0, \quad I_2 \leqq 0, \dots"$$

Farkas anticipated the procedure applied in the dual simplex method. We know, however, that cycling is possible there [2]. By the application of the lexicographic dual method [57] (a short presentation is given in [70]) we can correct the proof. Let us summarize the whole corrected proof briefly.

Proof of Farkas' theorem. First we remark that if we apply a linear transformation $y = Bx$, where B is a nonsingular square matrix, in the inequalities (3.1), (3.2), then the new (3.2) inequality will be a consequence of the new (3.1) system. Furthermore, proving Farkas' theorem for the new inequalities, we see that (3.3) is satisfied with the same nonnegative multipliers $\lambda_1,\ldots,\lambda_M$.

Let us assume first that the system (3.1) contains as many linearly independent relations as the number of variables. Consider the linear programming problem:

$$\text{minimize } (0\lambda_1 + \ldots + 0\lambda_M)$$

subject to

$$\lambda_1 g_1 + \ldots + \lambda_M g_M = g,$$

$$\lambda_1 \geqq 0,\ldots,\lambda_M \geqq 0.$$

Starting from any basis and applying the lexicographic dual method, at the end we reach a system of the form (3.4) where either $I_0 \geqq 0$, $K_0 \geqq 0,\ldots$ or there exists a row, where on the right hand side the first term is negative and all variables there have positive or zero coefficients. The second case cannot occur. In fact the lexicographic dual method guarantees that the column vectors of (3.4) are obtained from the vectors $g_1,\ldots,g_M,g$ by a nonsingular linear transformation, thus the system of linear inequalities (in the variables $y_1,\ldots,y_n$):

$$\begin{array}{rl} y_1 & \geqq 0 \\ y_2 & \geqq 0 \\ \vdots & \vdots \\ y_n & \geqq 0 \\ -I_1y_1-K_1y_2- \ldots & \geqq 0 \\ -I_2y_1-K_2y_2- \ldots & \geqq 0 \\ \text{- - - - - - - -} & \end{array} \tag{3.6}$$

has the consequence

$$I_0y_1 + K_0y_2 + \ldots \geqq 0. \tag{3.7}$$

Now, if e.g. we had $I_0 < 0$, $I_1 \leq 0$, $I_2 \leq 0, \ldots,$ then the vector of components $y_1 = 1$, $y_2 = \ldots = y_n = 0$ would satisfy (3.6) and would not satisfy (3.7).

If in (3.1) we have $h(<M)$ linearly independent relations and (for the sake of simplicity) $g_1, \ldots, g_h$ are linearly independent, then we choose n-component vectors $d_1, \ldots, d_{n-h}$ so that $B = (g_1, \ldots, g_h, d_1, \ldots, d_{n-h})$ be a nonsingular matrix and apply the transformation $y = Bx$ for the inequalities in (3.1) and (3.2). Then (3.1) will depend only on $y_1, \ldots y_h$ and since the inequality in (3.2) is a consequence of the former ones, $y_{h+1}, \ldots, y_n$ will have zero coefficients also there. Now we can forget about $y_{h+1}, \ldots, y_n$ and apply the above reasoning to the system of variables $y_1, \ldots, y_h$. At the end we can reestablish $y_{h+1}, \ldots, y_n$ everywhere with zero coefficients and reach the original inequalities by setting $y = Bx$.

Thus we have proved Farkas' theorem. We have also proved

Theorem 3.2. If (3.2) is not a consequence of the system (3.1), then there is a linear transformation $y = Bx$ with nonsingular square matrix B such that for at least one i, the variable y_i has negative coefficient in the new inequality (3.2) and has nonnegative coefficients in the new system (3.1).

The proof given by Farkas in [15] for his theorem can be summarized as follows. First we observe, that if we use a transformation $\underline{v} = B\underline{x}$, where B is a nonsingular square matrix, then, proving the theorem for the obtained inequalities with the variable $\underline{v}$, we also prove it for the original inequalities (the λ factors can be taken the same in both cases). Let $\underline{x} \in R^n$ and, using induction, assume that the theorem is true for every positive integer smaller than n. For $n = 1$ the assertion holds trivially. We may assume that $g \neq \underline{0}$ (otherwise the assertion is trivial). Introduce the nonsingular transformation: $v_1 = \underline{g}'\underline{x}, v_2 = x_2, \ldots, v_n = x_n$, or in compact form: $\underline{v} = B\underline{x}$. Then $v_1 \geq 0$ is a consequence of the inequalities $\underline{h}_1'\underline{v} = \underline{g}_1'B^{-1}\underline{v} \geq 0, \ldots, \underline{h}_M'\underline{v} = \underline{g}_M'B^{-1}\underline{v} \geq 0$. In these inequalities there must be at least one, in which v_1 has positive coefficient, because otherwise $v_1 \geq 0$ would not be a consequence. If the first k of these inequalities have that property, then already the inequalities

$$(3.8)\qquad \begin{array}{c} h_{11}v_1 + \ldots + h_{1n}v_n \geq 0 \\ \text{-------------------} \\ k_{k1}v_1 + \ldots + h_{kn}v_n \geq 0 \end{array}$$

have the consequence $v_1 \geqq 0$. Consider the following inequalities in the $n - 1$ variables $v_2, \ldots, v_n$:

$$(3.9) \qquad \begin{aligned} h_{12}v_2 + \ldots + h_{1n}v_n &\geqq 0, \\ &\text{-----------} \\ h_{k2}v_2 + \ldots + h_{kn}v_n &\geqq 0.. \end{aligned}$$

At least one of them must be identically zero, because otherwise we could find such $v_2, \ldots, v_n$, for which all inequalities in (3.9) were strictly positive but then $v_1 \geqq 0$ would not be a consequence of the system of inequalities (3.8). Now we use Farkas' theorem for the case of $n - 1$. Starting from (3.9), we assume that the first relation is identically zero. Then

$$(3.10) \qquad -h_{12}v_2 - \ldots - h_{1n}v_n \geqq 0$$

is a consequence of (3.9), thus there exist nonnegative numbers $\mu_1, \ldots, \mu_k$ such that

$$\begin{aligned} (\mu_1 + 1)h_{12} + \mu_2 h_{22} + \ldots + \mu_k h_{k2} &= 0, \\ &\text{-----------} \\ (\mu_1 + 1)h_{1n} + \mu_2 h_{2n} + \ldots + \mu_k h_{kn} &= 0 . \end{aligned}$$

This implies that there exist $\lambda_1 \geqq 0, \ldots, \lambda_k \geqq 0$ such that

$$\begin{aligned} \lambda_1 h_{11} + \lambda_2 h_{21} + \ldots + \lambda_k h_{k1} &= 1, \\ \lambda_1 h_{12} + \lambda_2 h_{22} + \ldots + \lambda_k h_{k2} &= 0, \\ &\text{-----------} \\ \lambda_1 h_{1n} + \lambda_2 h_{2n} + \ldots + \lambda_k h_{kn} &= 0 . \end{aligned}$$

Setting $\lambda_{k+1} = \ldots + \lambda_M = 0$, the assertion follows for the case of n and thus the proof is complete.

4. Necessary condition for the equilibrium.

We consider a mechanical system the state of which is described by the vector $x \in R^n$ which is subject to the following constraints:

$$(4.1) \qquad g_i(x) \geq 0, \; i = 1, \ldots, m.$$

Denote by $X_1, \ldots, X_n$ the components of the forces acting on the system and suppose that at X^* equilibrium is reached. Then by Fourier's principle the inequality

$$(4.2) \qquad X_1 \delta x_1 + \ldots + X_n \delta x_n \leqq 0$$

is satisfied, where $\delta x_1, \ldots, \delta x_n$ are variations of the coordinates $x_1, \ldots, x_n$ i.e. small quantities with the property that the vector of components

$$(4.3) \qquad x_1^* + \delta x_1, \ldots, x_n^* + \delta x_n$$

satisfies approximately (4.1). In what follows we shall use general terms but there are not very restrictive mathematical conditions under which our statements can be made exact.

The inactive constraints in (4.1) i.e. those for which we have strict inequality at the point x^*, do not restrict small changes in the coordinates. Thus to see what are the conditions on small changes we only have to consider the constraints active at x^*. Let us assume that these are the first M constraints. Then, writing dx_1 instead of δx_1, the Fourier principle requires that

$$X_1 dx_1 + \dots + X_n dx_n \leqq 0 \tag{4.4}$$

and the increments are subject to the inequalities.

$$\frac{\partial g_i(x^*)}{\partial x_1} dx_1 + \dots + \frac{\partial g_i(x^*)}{\partial x_n} dx_n \geqq 0, i = 1,\dots,M. \tag{4.5}$$

Now we can forget about the order of magnitude of the quantities $dx_1,\dots,dx_n$ satisfying the above inequalities. In fact if $dx_1,\dots,dx_n$ satisfy a homogeneous linear inequality, then the same holds for $tdx_1,\dots,tdx_n$ where t is any nonnegative constant.

Let X denote the vector of components $X_1,\dots,X_n$. We shall write it in a row form and also accept the convention used in many textbooks that the gradients are row vectors.

If x^* is an equilibrium point, then, using Fourier's principle, we find that the linear inequality (4.4) is a consequence of the system of linear inequalities (4.5), hence by Farkas' theorem, there exist nonnegative numbers $\lambda_1,\dots,\lambda_M$ such that

$$X + \lambda_1 \nabla g_1(x^*) + \dots + \lambda_M \nabla g_M(x^*) = 0 \tag{4.6}$$

If the system of forces is conservative, i.e. there exists a function $V(x)$ so that at every point of the state space we have

$$X_i = -\frac{\partial V}{\partial x_i}, i = 1,\dots,n, \tag{4.7}$$

then (4.4) becomes

$$\frac{\partial V}{\partial x_1} dx_1 + \dots + \frac{\partial V}{\partial x_n} dx_n \geqq 0. \tag{4.8}$$

Here on the left hand side we have a total differential. Having a potential, we can start from (4.8), by applying Courtivron's principle, stating that if a mechanical system is in stable equilibrium, then the potential has a local minimum and observing that (except for pathological cases) the total differential is nonnegative at the minimum point of the function.

We can ask: who were the persons - if any - who inferred the equation (4.6) from the inequality principle for the mechanical equilibrium? To answer this question we can start by analyzing the papers referenced by Farkas in his Crelle's Journal paper [23], and it is also advisable to have a look at the volumes of Enzyklopädie der Mathematischen Wissenschaften published at the beginning of this century. In the four volumes written on mechanics there are two papers where theories of statics and dynamics under inequality constraints are mentioned. The authors of these two papers are Voss [75] and Stäckel [71]. By Farkas and these two about thirty books and papers are referenced. Some of them are unfortunately not available to the author of this paper, but Farkas, Voss and Stäckel together very likely give a good picture about the history of the subject. Voss writes [75, p.74]:

> "Auf diesen von Lagrange nicht berücksichtigten Fall hat unabhängig von Fourier erst wieder Gauss dann Ostrogradsky hingewiesen...daher heisst in Russland das Prinzip von Fourier auch wohl das von Ostrogradsky. In Frankreich ist Fourier's Prinzip nicht so unbeachtet geblieben; A. A. Cournot entwickelt schon in 1827 die gleichungen von Ostrogradsky;"

One paper by Ostrogradsky is also referenced by Farkas; this paper is [67]. Stäckel mentions [68] too.

Voss does not refer to Farkas but Stäckel references the papers [14,16]. We can forgive that Stäckel does not recognize the importance of Farkas' work because the subject of his paper is "Elementare Dynamik". The work of Mayer is highly esteemed by Stäckel:

"Statt holonomer oder nicht holonomer Bedingungsgleichungen können auch Bedingungsungleichheiten auftreten; auch können Bedingungen plötzlich fortfallen oder plötzlich durch andere ersetzt werden. Nachdem bereits früher solche Fälle behandelt worden waren, hat, durch E. Study angeregt, A. Mayer, vom Gausschen Prinzipe des kleinsten Zwanges ausgehend, dargelegt, wie man hier die Bewegung des Punktes in allen Fällen bestimmen kann."

The excellent papers by A. Mayer referred to [58,59] handle dynamical problems, therefore are only partially relevant. However, it is very instructive to read these papers containing interesting ideas, good examples and proofs of nonnegativity of the multipliers in special cases. Zermelo [76] gave an interesting proof for more general cases. Farkas gave the most general form of the theory in [24]. Among the papers containing solution to dynamical problems subject to inequality constraints [39] should also be mentioned.

Fourier, Cournot and Farkas seem to be the principal contributors to the present form of the necessary condition of static equilibrium. Cournot (and later Ostrogradsky) presented the equations (4.6) in the form of a conjecture (as we may say now). Farkas proved (4.6) by relying on Fourier's work concerning the first part of the theorem (considering now conservative system of forces) where - as it is obvious today but did not seem necessary for those in the last century dealing with mechanical problems - a constraint qualification is needed. We have to add to these that the work of Cournot was based to a great extent on the work of Poinsot [69].

Fourier and Farkas both contributed further important ideas to optimization theory. Let us summarize briefly their principal results in this respect.

Fourier - anticipated the formulation of the linear programming problem [36/II., p.325-328], [40];

- formulated the inequality principle for the mechanical equilibrium [34];
- initiated the parametric solution of homogeneous linear inequalities [35].

Farkas - proved the basic theorem concerning homogeneous linear inequalities; first in [15];

- gave a rigorous proof for the "dual form" of Fourier's mechanical inequality principle [14,15];
- gave an elegant parametric representation for the solutions of homogeneous linear inequalities; first in [19].

5. The work of Cournot on the problem of equilibrium

Antoine-Augustin Cournot (1801-1877) famous polyhistor, was one of the greatest mathematical economists. The work of his which we are analyzing here was done as a contribution to mechanics. It is interesting to remark that in his famous book "Recherches sur les Principes Mathématiques de la Théorie des Richesses" no application can be found of his previous investigation on the problem of mechanical equilibrium. Since the first publication of this book several editions have appeared e.g. [7] with the critiques of Walras Bertrand and Pareto, and introduction and biographical notes by Lutfalla, further [8] with the notes of Irving Fisher.

Cournot became docteur ès sciences in 1829 in Paris. His thesis [5,6] contributed to dynamics, where he applied his earlier result published in 1827[4]. One year earlier he published an elementary paper [3] on inequalities; there is no anticipation in that paper of any form of Farkas' theorem.

In the paper [4] that is at present most important for us, Cournot does not refer to the work of Fourier. He seems to have been unaware of the enunciation of the inequality principle by Fourier in 1798. Cournot rediscovered the principle but also derived the necessary conditions for the equilibrium. We quote from his short paper [4, p.166]:

> "...il arrive souvent que les liaisons du système ne peuvent être exprimées par des équations; il résulte de ces liaisons des conditions d'équilibre que l'on a toujours regardées comme ne pouvant se déduire du principe des vitesses

virtuelles et étant en dehors de ce principe (Mécanique de M. Poisson, tom. I, p. 241). Notre but est ici de faire voir que, lorsque les liaisons du système peuvent s'ex primer algébriquement par le moyen des signes d'inégalité, le principe des vitesses virtuelles, convenablement modifié, s'y applique encore, et fournit, par une méthode uniforme, les conditions de l'équilibre."

On pages 167 and 168 of the same paper we read further:

"Admettons qu'un système soit soumis à un certain nombre de liaisons semblables, exprimées par les inégalités

$$I > 0,\ I' > 0,\ \text{etc.}\ J < 0,\ J' < 0,\ \text{etc.}$$

les signes > et < étant toujours supposés ne pas exclure le cas d'égalité.

On ne peut pas, en général, différentier une inégalité comme une équation, ni en déduire une relation entre les incrémens des variables; mais lorsqu'il s'agit de la recherche des conditions d'équilibre, deux cas seulement peuvent se présenter. Ou la situation du système est telle que l'on pourrait supprimer les liaisons, sans changer son état, comme il arrive, si les fils qui joignent certains de ses points ne sont pas tendus, ou si les points auxquels certaines surfaces opposent des obstacles impénétrables, ne sont point contigus à ces surfaces: dans ce cas l'équilibre doit avoir lieu, indépendamment des conditions de liaison, et il est superflu d'en tenir compte; ou bien le système est placé de manière à ce que les liaisons produisent leur effet, et puissent diminuer le nombre des conditions nécessaires pour l'équilibre: alors on a pour les valeurs actuelles des coordonnées $I = 0, J = 0$, etc., et si l'on fait varier ces valeurs, leurs incrémens ne seront pas entièrement arbitraires; ils devront, en vertu des conditions primitives, satisfaire aux relations:

(a) $$\delta I > 0, \delta I' > 0,\ \text{etc.}\ \delta J < 0, \delta J' < 0,\ \text{etc.}$$

D'ailleurs, quand le système est soumis à des liaisons de l'espèce de celles que nous considérons ici, il est clair qu'on pourrait les supprimer, pourvu qu'on lui appliquât certaines forces, capables d'en tenir lieu; ainsi la résistance d'une surface peut être remplacée par l'application d'une force normale à cette surface et dirigée dans le sens convenable, celle d'un fil tendu, par l'application d'une force dirigée dans le sens de la tension du fil, etc. Soient F,F',....les forces directement appliquées au système, suivant les directions f,f',....et P,P',...les forces auxiliaires qui tiendraient lieu de l'existence des obstacles, et seraient dirigées suivant les lignes p,p'....; on aura, en vertu du principe des vitesses virtuelles, l'équation fondamentale

$$F\delta f + F'\delta f' + \text{etc.} + P\delta p + P'\delta p' + \text{etc.} = 0.$$

Or il est aisé de voir que tous les mouvemens virtuels, pour lesquels P et δp,P' et $\delta p'$, etc., seraient de signes contraires, tendraient à surmonter les obstacles que les surfaces, les fils, etc., opposent aux differens points de système, et par conséquent sont incompatibles avec les liaisons du système, exprimées par les inégalités (a). Il n'y a de compatibles avec ces liaisons que les mouvemens pour lesquels P et δp,P' et $\delta p'$ sont de même signe, et pour lesquels par conséquent les quantités $P\delta p$,$P'\delta p'$, etc. sont essentiellement positives."

The "proof" given by Cournot for his equations is only a short reference to our imagination. When dealing with the special case of a system of points positioned on planes parallel to the xy-plane, Cournot refers to the work of Poinsot [69], first published in 1803. In fact, for this special case Poinsot already obtained the conditions of equilibrium. Assume for the sake of simplicity that one body is supported at the points A,B,C,D,..., by a plane parallel to the xy-plane. On page 125 of [69] we read:

"...toutes les forces appliquées au système doivent se réduire à une seule, perpendiculaire au plan fixe, et dont la valeur ne soit pas positive.

En second lieu, je dis que sa direction doit rencontrer ce plan dans l'intérieur du polygone formé par les points d'appui A,B,C,D,...."

The papers [3,4] of Cournot seem to be less well known then his other works. They are not listed in the bibliography of his works published in [7].

6. The work of Ostrogradsky on the problem of equilibrium

Mikhail Vasilevich Ostrogradsky (1801-1862) contributed important results (among others) to mechanics. He was a student in Paris and attended the courses of Fourier, Poisson, Cauchy and other well-known French mathematicians. He returned to Russia (St. Petersburg) in 1828. In 1830 he became an extraordinary member and in 1832 an ordinary member of the Academy of St. Petersburg.

Two papers of Ostrogradsky should be mentioned concerning the problem of mechanical equilibrium. Farkas refers only to [67] (presented in 1834 before the Academy) but the other [68] is an improved version of his theory. In the earlier paper four applications are mentioned: a.) a point that is allowed to move in one part of the space subdivided by a surface; b.) the funicular polygon; c.) the flexible file; d.) the incompressible liquid. Further problems are mentioned concerning dynamics.

In [68] Ostrogradsky refers to (what we call now) Farkas' theorem as an obvious algebraic fact and derives the equation of equilibrium (4.6). We reproduce here part of the pages 589 and 590 of [68]:

"Supposons que les quantités $\delta s, \delta s', \delta s'', \delta s''', \ldots$ appartiennent non seulement à ceux des déplacemens du système, dont les forces perdues sont capables, mais encore, à tous les autres déplacemens tant possibles que non, ou plutôt considérons $\delta s, \delta s', \delta s'', \delta s''', \ldots$ comme tout-à-fait arbitraires. Nous devons exprimer que les forces perdues $R, R', R'', R''', \ldots$ sont incapables de produire aucun déplacement des systèmes satisfaisant aux conditions

$$(15)\quad \begin{cases} \delta L > 0 \\ \delta L_1 > 0 \\ \delta L_2 > 0 \\ \delta L_3 > 0 \\ \ldots\ldots \end{cases}$$

le signe $>$ n'exclut point celui de l'égalité.

Or on sait qu'un système des forces est capable de tout déplacement qui fournit, pour son moment total, une valeur positive et aucun de ceux qui correspondent aux valeurs négatives ou zéro du moment total. Ainsi, pour que les forces perdues soient incapables de produire aucun des déplacemens satisfaisans aux conditions (15), il faut que leur moment soit négatif ou zéro pour ces déplacemens, c'est-à-dire il faut que la fonction

$$R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots..$$

dans laquelle $\varphi, \varphi', \varphi'', \varphi''', \ldots..$ désignent respectivement les angles $\widehat{R\delta s}, \widehat{R'\delta s'}, \widehat{R''\delta s''}, \widehat{R'''\delta s'''}, \ldots.$ et qui par conséquent représente le moment des forces $R, R', R'', R''', \ldots..$ soit negative ou zéro toutes les fois que $\delta s, \delta s', \delta s'', \delta s''', \ldots.$ remplissent les conditions (15).

La solution de la question qui consiste à rendre la fonction

$$R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R''' ds'''\cos\psi''' + \ldots..$$

négative ou zéro toutes les fois que les fonctions de même nature, $\delta L, \delta L_1, \delta L_2, \delta L_3, \ldots.$ sont positives ou zéro, appartient à l'algèbre la plus élémentaire. Il est nécessaire, et il suffit que $R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots..$ puisse se réduire à une fonction linéaire de $\delta L, \delta L_1, \delta L_2, \delta L_3, \ldots.$ avec des coëfficiens négatifs. Ainsi il n'y a qu'à faire, quels que soient $\delta s, \delta s', \delta s'', \delta s''', \ldots..,$

$$R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots$$
$$= \lambda\delta L + \lambda_1\delta L_1 + \lambda_2\delta L_2 + \lambda_3\delta L_3 + \ldots.$$

et à y ajouter la condition que les λ sont tous négatifs. Ou bien, si l'on veut éviter de considérer les λ comme négatifs, on peut faire

$$R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots$$
$$= -(\lambda\delta L + \lambda_1\delta L_1 + \lambda_2\delta L_2 + \lambda_3\delta L_3 + \ldots)$$

alors tous les λ seront positifs. Il est évident, par la dernière équation, comme par celle qui la précède, que le moment $R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots.$ sera négatif ou zéro toutes les fois que les fonctions $\delta L, \delta L_1, \delta L_2, \delta L_3, \ldots.$ seront positives ou zéro.

En transportant tous les termes d'un même côté, l'équation de l'équilibre des forces perdues deviendra

$$(16)\qquad R\delta s\cos\psi + R'\delta s'\cos\psi' + R''\delta s''\cos\psi'' + R'''\delta s'''\cos\psi''' + \ldots\ldots + \lambda\delta L + \lambda_1\delta L_1 + \lambda_2\delta L_2 + \lambda_3\delta L_3 \ldots\ldots = 0$$

elle doit avoir lieu quelles que soient $\delta s, \delta s', \delta s'', \delta s''', \ldots\ldots$ tant en grandeur que pour la direction. Mais il ne faut pas oublier d'ajouter à l'équation (16) les inégalites

$$(17)\quad \begin{cases} \lambda > 0 \\ \lambda_1 > 0 \\ \lambda_2 > 0 \\ \lambda_3 > 0 \\ \ldots\ldots \end{cases}$$ "

Ostrogradsky made two errors. One consists in the declaration without proof of the assertion of Farkas' theorem. (In this respect Ostrogradsky's authority was so strong, that many authors took it for granted [58,59,64,75]). The second error was discovered by Study as communicated by Mayer [58, p.225]. On page 583 of [59] Ostrogradsky infers that some of his inequalities should be equal to zero at the equilibrium. This would have made it possible to determine the multipliers in the equation of equilibrium. He thought that those inequalities should be equal to zero which would turn out to be negative if we did not have the other constraints. Hamel gave a counterexample for this in [47, p.41-42].

7. The work of Farkas on the problem of equilibrium

The two papers [13,14] being essentially the same, we choose the German version [14] rather than the Hungarian for our quotation. In the introductory part of the paper we read the following:

"In diesem besteht das Fourier'sche Princip. Ich werde dasselbe auch Ungleichheits-Princip, hingegen das gewöhnliche das Gleichheits-Princip, nennen. Das letztere ist in dem ersteren, als allgemeinerem enthalten, denn in dem Falle, dass der Zwang durch blosse Gleichungen ausgedrückt werden kann, reducirt sich das Ungleichheits-Princip auf das Gleichheits-Princip; nämlich in dem Gleichheits-Zwange entspricht einer jeden virtuellen Verrückung auch die entgegensetzte, dann kann aber die virtuelle Arbeit nur auf die Weise keinen positiven Wert annehmen, dass sie verschwindet.

So wie das Gleichheits-Princip, ist auch das Ungleichheits-Princip unabhängig von den Eigenheiten der verschiedenen Punkt-Koppelungen, und eben deshalb ist GALILEI als der Erfinder des Gleichheits-Princips anzusehen, weil dieses Princip vor ihm nur für eine gewisse Zusammensetzung und für solche, die leicht auf diese zurückgeführt werden können, aufgestellt wurde, nämlich von ARISTOTELES, und nachdem es durch eine sehr lange Zeit in Vergessenheit gerathen war, von UBALDI für hebelartige Zusammensetzungen. Die Verwertung des Princips erfordert aber selbstverständlich die Berücksichtigung der Art und Weise der gegebenen Punkt-Koppelungen wenigstens dermaassen, dass auf dem Grunde der zugehörigen Kinematik die entsprechenden Zwangs-Ausdrücke aufgestellt werden können. Ueber diese Forderung hinaus ist die Anwendung des Gleichheits-Princips seit LAGRANGE bloss eine Sache der reinen Analysis und insbesondere ein Problem der Auflösung von Gleichungen, welche jedesmal nach bereits aufliegenden Verfahrungsmethoden dargestellt werden können.

Mit dem Ungleichheits-Princip ist es nicht so weit gekommen. Von seinem Erfinder, FOURIER, wurde nur sein genetischer Zusammenhang mit älteren Principien erörtert. Die genze diesbezügliche Abhandlung von FOURIER ist unbeachtet geblieben. Dreissig Jahre später wurde das Princip von GAUSS wieder aufgestellt, oder vielmehr einfach ausgesprochen in dem kurzen Aufsatze über das Princip des kleinsten Zwanges. Ein paar Jahre später befasste sich noch der russische Mathematiker OSTROGRADSKY mit dem Princip; er hatte dessen Anwendungen in Angriff genommen, er hat aber das Princip nicht in jener ganzen Allgemeinheit aufgefasst, welche demselben beigelegt werden kann, da er nur eine gewisse Klasse der Zwangs-Verhältnisse vor Augen hielt, jene, in welcher die Anzahl der Zwangs-Ausdrücke diejenige der virtuellen Verrückungs-Componenten nicht übertrifft; auf diese Weise haben sich seine Betrachtungen auf einen verhältnissmässig sehr kleinen Gultigkeits-Bereich beschränkt. Wie es scheint hat sich sonst Niemand mit der Anwendung des Princips abgegeben, und es scheint sogar auch die Meinung vor zu walten, dass das Princip nicht nutzbar sei. Dasselbe hat sozusagen nur einen leeren Ruf beibehalten."

"Der Hauptzweck vorligender Arbeit ist zu erweisen, dass mit einer passenden Modification die Methode der Multiplicatoren von LAGRANGE auch auf das FOURIER'sche Princip übertragen werden kann."

In the first section of the paper he deals with his inequality theorem. That proof is incomplete, however, as we pointed out. In the second section the necessary condition of equilibrium is derived:

"...sollen die Zwangs-Ausdrücke in die folgenden übergehen:

$$\begin{aligned} &\Sigma F\delta q = 0, \quad \Sigma G\delta q = 0,.. \\ &\Sigma S\delta q \geq 0, \quad \Sigma T\delta q \geq 0,... \end{aligned} \qquad 11)$$

und die principielle Ungleichheit soll heissen

$$\Sigma Q\delta q \leq 0 \quad \text{oder} \quad -\Sigma Q\delta q \geq 0. \qquad 12)$$

Die Zwangs-Gleichungen sollen auch in der Form von Ungleichheiten aüsgedruckt werden, so dass das System der Zwangs-Aüsdrucke erscheint wie folgt:

$$\begin{aligned} &\Sigma F\delta q \geq 0, \quad \Sigma G\delta q \geq 0,..., \\ &-\Sigma F\delta q \geq 0, \quad -\Sigma G\delta q \geq 0,..., \\ &\Sigma S\delta q \geq 0, \quad \Sigma T\delta q \geq 0,... \end{aligned} \qquad 13)$$

Das FOURIER'sche Princip verlangt, dass die Ungleichheit (12) von allen Wertsystemen der δq befriedigt wird, welche (13) befriedigen. Dies geschieht aber nur, wie erwiesen, wenn es positive Multiplicatoren giebt, vermöge deren die Coëfficienten Q als lineare homogene Functionen der Coëfficienten $F,G,\ldots,-F,-G,\ldots,S,T,\ldots$ dargestellt werden können. Diese positiven Multiplicatoren sollen mit $\varphi',\psi',\ldots,\varphi'',\psi'',\ldots,\lambda,\mu,\ldots$ bezeichnet werden. Man muss haben

$$-Q = (\varphi' - \varphi'')F + (\psi' - \psi'')G + \ldots + \lambda S + \mu T + \ldots$$

Nun Können aber die Differenzen $\varphi' - \varphi'', \varphi' - \varphi'', \ldots$ auch negative Werte annehmen, folglich wird das FOURIER'sche Princip von jenen Q-Werten erfüllt, welche durch Gleichungen wie

$$Q + \varphi F + \psi G + \ldots + \lambda S + \mu T + \ldots = 0, \qquad 14)$$

sich bestimmen lassen, wo $\varphi,\psi,\ldots,\lambda,\mu,\ldots$ in dem Ausdrucke einer jeden Q-Grösse dieselben Werte haben und dabei sind $\varphi,\psi,\ldots$ an sich ganz willkürlich, $\lambda,\mu,\ldots$ aber willkürliche nichtnegative Quantitäten. Umgekehrt, aus dem Systeme (14) folgt mittels des Systems (11) die principielle Ungleichheit (12) durch ein leicht zu erkennendes Verfahren."

In a further section of the paper "the two main types of application" are presented: a.) the equilibrium equation for tangential solid bodies; b.) the equilibrium equation for non-solid bodies.

8. <u>On the constraint qualification</u>

The "constraint qualification" is of fundamental importance not only from the point of view of nonlinear optimization but from the point of view of mechanics too. Soon we shall see why.

In the beginning of this century the axiomatic foundation of the mathematical and physical sciences became an important activity. In his famous paper [49] Hilbert urges mathematicians to axiomatize two "physical disciplines": probability theory and mechanics. The paper of Hamel [46] of 1909 is the first serious attempt in the direction of the axiomatic foundation of the classical mechanics. The Fourier inequality principle is unfortunately not included. An improved version of his axiomatics is contained in his paper [47] which appeared in 1927. There the inequality principle of Fourier is already mentioned as one of the axioms: Axiom II 2k on page 17. On page 33 concerning "Das Energieprinzip", Axiom II 5cβ declares denoting by U the potential, that $\delta U \geq 0$ is a necessary condition of the equilibrium. Hamel had to establish consistency between the two axioms and this could only be done by a "constraint qualification;" in fact we find it in Axiom II 5cγ, on page 33. It requires that to every mass particle there correspond a scalar function u such that the following equality holds:

$$\delta U = \Sigma dm \nabla u \delta r, \tag{6.1}$$

where dm denotes the mass and r the state of a particle. If instead of (6.1) we require

$$\delta U = \nabla U \delta r, \tag{6.2}$$

where r denotes the state of the whole system of particles, then we obtain *essentially the Kuhn-Tucker constraint qualification.* (6.2) implies (6.1) because we can generate the u functions for the purpose of satisfying (6.1) in such a way that in U we subsequently fix all variables except for those belonging to one particle.

Unfortunately Hamel was unaware of the existence of Farkas' theorem. Even in his Theoretische Mechanik (first published in 1949) this theorem is not referred to, though part of the pages 69-70, 517-518 are devoted to the Fourier inequality principle. There we also see that his "constraint qualification" is not exactly derived and his reasoning concerning the Fourier principle is more heuristic.

9. Miscellaneous remarks concerning linear inequalities

Farkas' most important results concerning the Fourier principle and the theory of linear inequalities are summarized in his papers [13,14,23].

The parametric representation of the solutions of linear inequalities was initiated by Fourier [35]. Minkowski gave the representation of all solutions using extremal rays [60]. In the parametric representation given by Farkas [19,23], the generation of the rays the convex combinations of which constitute all solutions of the linear inequalities is particularly simple. Writing $z_1,\dots,z_n$ instead of $dx_1,\dots,dx_n$ in (4.5) and (4.6), we can represent all solutions of (4.5) in a parametric form, insert it into (4.4) and if the number of variables and the active constraints in (4.5) is not large, we can solve the equilibrium problem in some cases. For this we have to know which are the active constraints in (4.5). This method was advised by Farkas [19]. The application of the Farkas parametric representation technique for linear programming is perhaps best done as in the paper by Uzawa [74] He completes one inequality to equality (if necessary), eliminates the right hand side non-zero constants from the other constraints, inserts the parametric form of the solutions of these into the remaining equality and the objective function and finds the optimal solution.

The papers [21,22] published in Hungarian and German, respectively, have the same content. Farkas gave further mechanical application of his theorem on linear inequalities. This concern the decomposition of the forces of reaction in a mechanical system into shocks and others satisfying the negatives of the constraints given for the displacements (Voss mentions in [75, p.75] that this kind of distinction between the forces is attributed to Painlevé). The excerpted mathematical results of the paper is reproduced in [23]. The book [20] published in Hungarian does not contain new results as compared to his earlier papers and to the paper [23].

Between 1902 and 1917 Farkas did not publish on linear inequalities. In 1917, after Haar generalized Farkas' theorem for the inhomogeneous case, Farkas returned to the area and published further papers [28,29,31] on systems of linear inequalities. The papers [25,26,27,30] are also partially relevant; [30] is the German version of [25].

Haar wrote three papers on linear inequalities [42,43,44]. The first two are essentially the Hungarian and German versions of the same paper. He gave the following generalization for Farkas' theorem (we shall use the notation of Section 3).

Theorem 6.1. If the linear inequality

(6.1) $$g'x \geq b$$

is a consequence of the linear inequalities

(6.2) $$g_i'x \geq b_i, i = 1,\ldots,M$$

i.e. every x satisfying (6.2) also satisfies (6.1), then there exist nonnegative constants $\lambda, \lambda_1, \ldots, \lambda_M$ such that for every $x \in R^n$ we have

(6.3) $$g'x - b = \sum_{i=1}^{M} \lambda_i (g_i'x - b_i) + \lambda.$$

This is the same statement referred to by Kuhn and Tucker in a footnote supplied to von Neumann's manuscript [61]. This concerns the duality theorem of linear programming proved by Gale, Kuhn and Tucker [37].

Haar remarked [42,44], that the theory of linear inequalities was developed by Farkas and Minkowski. The famous book of Minkowski [60] contains results also on linear inequalities (pages 39-45 in both editions). He considered a finite system of homogeneous linear inequalities from two points of view: to represent them in a parametric form and to discover those which are superfluous i.e. can be omitted without changing the set of solutions. This latter problem led him to prove a theorem closely related to Farkas' theorem mentioned in Section 3. For the sake of completeness we present here the exact statement of Minkowski's theorem.

We assume that in the system (3.1) the number of linearly independent relations equals n, the number of components of x. This does not restrict the generality, as Minkowski remarked. An x is said to be an extremal ray of the cone (3.1) if x is not the sum of two nonzero vectors which satisfy (3.1) and neither of them is a constant multiple of the other.

Theorem 6.2. Out of the linear forms $g_i'x \geq 0, i = 1,\ldots,M$ those which are $m - 1$ linearly independent extremal solutions have the property that each of them is essential and all other forms can be expressed as their linear combinations with nonnegative weights.

For further references to early papers on linear inequalities see [9].

REFERENCES

Remarks: - In the list given below the name of Farkas appears with J or GY initials. Up to the beginning of this century it was customary to translate the given name to the language of the paper. The letter J refers to the German name Julius, the double letter GY refers to the Hungarian name Gyula and these two names were supposed to be the same.

- Volumes of papers presented at an academy have frequently double year-assignment. One indicates the year(s) when the papers in the volume were presented before the academy whereas the other indicates the year when the publisher published the volume. In the reference list below always the latter date is given.

[1] K. J. Arrow, L. Hurwicz and H. Uzawa (editors), Studies in Linear and Nonlinear Programming (Stanford University Press, Stanford, California, 1958).

[2] E.M.L. Beale, "Cycling in the dual simplex algorithm", Naval Research Logistics Quarterly 2(1955)269-276.

[3] A. Cournot, "Sur le calcul des conditions d'inégalité, annoncé par M. Fourier", Bulletin des Sciences Mathématiques (Première Section du Bulletin Universel des Sciences et de l'Industrie, publié sous la dir. de Férussac)6(1826)1-8.

[4] A Cournot, "Extension du principe des vitesses virtuelles au cas où les conditions de liaison du système sont exprimées par des inégalités", Bulletin des Sciences Mathématiques (Première Section du Bulletin Universel des Sciences et de l'Industrie, publié sous la dir. de Ferussac)8(1827)165-170.

[5] A. A. Cournot, "Mémoire sur le mouvement d'un corps rigide, soutenu par un plan fixe", Journal für die reine und angewandte Mathematik 5(1830)133-162, 223-249.

[6] A. A. Cournot, "Du mouvement d'un corps sur un plan fixe, quand on a égard à la résistance du frottement", Journal fûr die reine und angewandte Mathematik 8(1832) 1-12.

[7] A. Cournot, Recherches sur les Principes Mathématiques de la Théorie des Richesses (Nouvelle Édition, Paris, Marcel Rivière, 1938).

[8] A. Cournot, Researches into the Mathematical Principles of the Theory of Wealth (R. D. Irwin, Homewood, Ill., 1963).

[9] L. L. Dines and N. H. McCoy, "On linear inequalities", Proceedings and Transactions of The Royal Society of Canada, Third Series, Section III. Mathematical, Physical and Chemical Sciences 27(1933)37-70.

[10] L. Dirichlet, "Sur un moyen général de vérifier l'expression du potentiel relatif à une masse quelconque, homogène ou hétérogène", Journal für die Reine und Angewandte Mathematik 32(1846)80-84

[11] L. Dirichlet, "Über die Stabilität des gleichgewichts", Journal für die Reine und Angewandte Mathematik 32(1846)85-88.

[12] Encyklopädie der Mathematischen Wissenschaften (Leipzig, Teubner, 1901-1908) Vierter Band, Mechanik, Erster Teilband.

[13] GY. Farkas, "A Fourier-féle mechanikai elv alkalmazásai" (Applications of the mechanical principle of Fourier), Mathematikai es Természettudományi Értesitő 12(1895)457-472.

[14] J. Farkas, "Über die Anwendungen des mechanischen Princips von Fourier", Mathematische und Naturwissenschaftliche Berichte aus Ungarn 12(1895)263-281.

[15] GY. Farkas, "A Fourier-féle mechanikai elv alkalmazásainak algebrai alapjáról" (On the algebraic basis of the applications of the mechanical principle of Fourier), Mathematikai és Physikai Lapok 5(1896)49-54.

[16] J. Farkas, "Die algebraischen grundlagen der Anwendungen des Fourierschen Principes in der Mechanik", Mathematische und naturwissenschaftliche Berichte aus Ungarn 15(1899)25-40.

[17] GY. Farkas, "A Fourier-féle mechanikai elv alkalmazásának algebrai alapja" (The algebraic basis of the application of the mechanical principle of Fourier), Mathematikai es Természettudományi Értesitő 16(1898)361-364.

[18] J. Farkas, "Die algebraische grundlage der Anwendungen des mechanischen Princips von Fourier", Mathematische und naturwissenschaftliche Berichte aus Ungarn 16(1899)154-157.

[19] GY. Farkas, "Paraméteres módszer Fourier mechanikai elvéhez" (Method of parameters to the mechanical principle of Fourier), Mathematikai és Physikai Lapok 7(1898)63-71.

[20] GY. Farkas, Vectortan és az egyszerü inaequatiok tana (Theory of vectors and simple inequalities) (Kolozsvár, 1900).

[21] GY. Farkas, "Általános mechanikai elvek az aether számára" (General mechanical principles for the ether), Mathematikai és Természettudományi Értesitö 19(1901)99-122.

[22] J. Farkas, "Allgemeine Principien für die Mechanik des Aethers", Archives Néderlandaises, Série II Vol. 4 (Livre jubilaire dédié à H. A. Lorentz).

[23] J. Farkas, "Theorie der einfachen Ungleichungen", Journal für die reine und angewandte Mathematik 124(1902)1-27.

[24] J. Farkas, "Beiträge zu den grundlagen der analytischen Mechanik", Journal für die reine und angewandte Mathematik 131(1906)165-201.

[25] GY. Farkas, "Biztos egyensúly potenciál nélkül", Mathematikai és Természettudományi Értesitö 32(1915)339-354.

[26] GY. Farkas, "Nemvonalas egyenlötlenségek vonalassà tétele" (Conversion of non-linear inequalities into linear ones), Mathematikai és Természettudományi Értesitö 35(1917)41-50.

[27] GY. Farkas, "Multiplicatoros módszer négyzetes alakokhoz" (Method of multipliers for quadratic forms), Mathematikai és Természettudományi Értesitö 35(1917)51-53.

[28] GY. Farkas, "Egyenlötlenségek alkalmazásának új módjai" (New ways for the application of inequalities), Mathematikai és Természettudományi Értesitö 36(1918)297-308.

[29] GY. Farkas, "A lineáris egyenlötlenségek következményei" (Consequences of linear inequalities), Mathematikai és Természettudományi Értesitö 36(1918)397-408.

[30] J. Farkas, "Stabiles gleichgewicht ohne Potential", Mathematische und Naturwissenschaftliche Berichte aus Ungarn 32(1922)43-50.

[31] GY. Farkas, "Alapvetés az egyszerű egyenlőtlenségek vektorelméletéhez" (Foundation of the vectorial theory of simple inequalities), Matematikai és Természettudományi Értesitő 43(1926)1-3.

[32] I. Fényes, "Megjegyzések és kiegészitések a mechanika elveinek Farkas Gyula-féle tárgyalásmódjához" (Remarks and complements to the treatment of the mechanical principles by Farkas), Fizikai Szemle 4(1954)99-102.

[33] L. Filep, "Farkas Gyula (1847-1930)", Természet Világa 1976/9, 410-411.

[34] J. Fourier, "Mémoire sur la statique", Journal de l'Ecole Polytechnique 5(1798), Oeuvres Vol. II. 478-520.

[35] J. Fourier, "Solution d'une question particulière du calcul des inégalités", Nouveau Bulletin des Sciences par la Société Philomatique de Paris (1826), in [36] Volume 2, 317-319.

[36] J. Fourier, Oeuvres (Gauthier-Villars, Paris, 1888).

[37] D. Gale, H. W. Kuhn and A. W. Tucker, "Linear programming and the theory of games", in [51]317-329.

[38] C. F. Gauss, "Über ein neues allgemeines grundgesetz der Mechanik", Journal für die reine und angewandte Mathematik 4(1829)232-235.

[39] J. W. Gibbs, "On the fundamental formulae of dynamics", American Journal of Mathematics 2(1879)49-64.

[40] I. Grattan-Guinness, "Joseph Fourier's anticipation of linear programming", Operational Research Quarterly 21(1976)361-364.

[41] A. T. Grigorian, Ocherki Istorii Mechaniki v Rossii (Akad. Nauk SSSR, Moskva, 1961).

[42] A. Haar, "A lineáris egyenlőtlenségekről" (On linear inequalities), Mathematikai és Természettudományi Értesitő 36(1918)279-296.

[43] A. Haar, "Die Minkowskische Geometrie und die Annäherung an stetige Funktionen", Mathematische Annalen 78(1918)294-311.

[44] A. Haar, "Über lineare Ungleichungen", Acta Scientiarum Mathematicarum 2(1924)1-14.

[45] A. Haar, Gesammelte Arbeiten (Akadémiai Kiadó, Budapest, 1959).

[46] G. Hamel, "Über die grundlagen der Mechanik", Mathematische Annaln 66(1909)350-397.

[47] G. Hamel, "Die Axiome der Mechanik", in [48] 1-42.

[48] Handbuch der Physik (H. Geiger and K. Scheel editors) Band V., Grundlagen der Mechanik, Mechanik der Punkte und Starren Körper (Springer, Berlin, 1927).

[49] D. Hilbert, "Mathematische Probleme (Vortrag gehalten auf dem internationalen Mathematiker-Kongress zu Paris 1900)", Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse (1900)253-297.

[50] F. John, "Extremum problems with inequalities as subsidiary conditions", in [71] 187-204.

[51] W. Karush, Minima of Functions of Several Variables with Inequalities as Side Conditions, Master's Thesis (Department of Mathematics, University of Chicago, 1939).

[52] T. C. Koopmans (editor), Activity Analysis of Production and Allocation (Wiley, New York, 1951).

[53] H. W. Kuhn and A. W. Tucker, "Nonlinear programming" in [61] 481-492.

[54] H. W. Kuhn, "Nonlinear programming: a historical view", in [62] 1-26.

[55] J. L. Lagrange, Mécanique Analytique I-II (Paris, 1788).

[56] J. L. Lagrange, Oeuvres (Gauthier-Villars, Paris, 1888).

[57] C. E. Lemke, "The dual method for solving the linear programming problem", Naval Research Logistics Quarterly 1(1954)36-47.

[58] A. Mayer, "Ueber die Aufstellung der Differentialgleichungen der Bewegung für reibungslose Punktsysteme, die Bedingungsungleichungen unterworfen sind", Berichte über die Verhandlungen der königlich sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Classe 50(1898)224-244.

[59] A. Mayer, "Zur Regulierung der Stosse in reibungslosen Punktsystemen, die dem Zwange von Bedingungsungleichungen unterliegen", Berichte über die Verhandlungen der königlich sächsischen gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Classe 50(1898)245-264.

[60] H. Minkowski, Geometrie der Zahlen (Teubner, Leipzig und Berlin, first edition 1896, second edition 1910).

[61] J. von Neumann, "Discussion of a maximum principle, Collected Works, Vol. VI., 89-95.

[62] J. Neyman (editor), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, California, 1950).

[63] Nonlinear Programming, SIAM-AMS Proceedings Vol. IX. (American Mathematical Society, Providence, Rhode Island, 1976).

[64] L. Nordheim, "Die Prinzipe der Dynamik", in [47] 43-90.

[65] R. Ortvay, "Farkas Gyula tudományos müködése" (The scientific work of Gyula Farkas), Matematikai és Fizikai Lapok 34(1927)5-25.

[66] R. Ortvay, "Farkas Gyula emlékezete" (The memory of Gyula Farkas), Az MTA Tagjai Felett Tartott Emlékbeszédek 21(1933)15 Füzet.

[67] M. Ostrogradsky, "Considérations génerales sur les momens des forces", Mémoires de l'Académie Impériale des Sciences de Saint-Pétersbourg, Sixieme Série 1(1838)129-150.

[68] M. Ostrogradsky, "Mémoire sur les déplacemens instantanés des systèmes assujettis a des conditions variables", Mémoires de l'Académie Impériale des Sciences de Saint-Pétersbourg, Sixième Série 1(1838)565-600.

[69] L. Poinsot, Éléments de Statique (Douzième Édition, Paris, Gauthier-Villars, 1877), first published in 1803.

[70] A. Prékopa, Lineáris Programozás (Bolyai János Matematikai Tarsulat, Budapest, 1968).

[71] P. Stäckel. "Elementare Dynamik der Punktsysteme und starren körper", in [13] 435-691.

[72] Studies and Essays, presented to Courant on his 60th birthday (Interscience, New York, 1948).

[73] B. Szénássy, A Magyarországi Matematika Torténete (History of Mathematics in Hungary), (Akadémiai Kiadó, Budapest 1970).

[74] H. Uzawa, "An elementary method for linear programming", in [1] 179-188.

[75] A. Voss, "Die Prinzipien der rationellen Mechanik" in [13] 3-121.

[76] E. Zermelo, "Ueber die Bewegung eines Punktsystemes bei Bedingungsungleichungen", Nachrichten von der königl. Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse (1898) 306-310.

AP/ed

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 1842
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): (6) ON THE DEVELOPMENT OF OPTIMIZATION THEORY
5. TYPE OF REPORT & PERIOD COVERED: Summary Report - no specific reporting period
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): (10) Andras Prekopa
8. CONTRACT OR GRANT NUMBER(s): (15) DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Work Unit Number 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709
12. REPORT DATE: (11) Apr[illegible]78
13. NUMBER OF PAGES: 35
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(9) Technical summary rept.,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(14) MRC-TSR-1842 (12) 39p.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Fourier principle, Farkas theorem, nonlinear programming, linear inequalities

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The method of Lagrange for finding extrema of functions subject to equality constraints was published in 1788 in the famous book Mecanique Analytique. The works of Karush, John, Kuhn and Tucker concerning optimization subject to inequality constraints appeared more than 150 years after that. The purpose of this paper is to call attention to important papers, published as contributions to mechanics, containing fundamental ideas concerning optimization theory. The most important works in this respect were done primarily by Fourier, Cournot, Farkas and further by Gauss, Ostrogradsky and Hamel.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

221 200