TEKST NR 377 1999

The Kuhn-Tucker Theorem in Nonlinear Programming: A Multiple Discovery?

Tinne Hoff Kjeldsen

TEKSTER fra

IMFUFA ROSKILDE UNIVERSITETSCENTER
INSTITUT FOR STUDIET AF MATEMATIK OG FYSIK SAMT DERES FUNKTIONER I UNDERVISNING, FORSKNING OG ANVENDELSER

IMFUFA, Roskilde University, Postbox 260, DK-4000 Roskilde, Denmark.

The Kuhn-Tucker Theorem in Nonlinear Programming: A Multiple Discovery?

by

Tinne Hoff Kjeldsen

IMFUFA text no. 377/1999 37 pages ISSN 0106-6242

Abstract:

When Kuhn and Tucker proved the Kuhn-Tucker theorem in 1950 they launced the theory of nonlinear programming. However, in a sense this theorem had been proven already: In 1939 by W. Karush in a master's thesis, which was never published. In 1948 by F. John in a paper that was at first rejected by the Duke Mathematical Journal. And maybe earlier by Ostrogradsky and Farkas. The questions whether the Kuhn-Tucker theorem can be seen as a multiple discovery and why the different occurences of the theorem was so differently received by the mathematical communities is discussed on the basis of a contextualized historical analysis of these works. The significance of the contexts both mathematically and socially for these questions are discussed including the role played by the military in the shape of Office of Naval Research (ONR) and operations research (OR).

MSC 1991 subject classifications: 01A60, 49-03, 52-03, 90-03, 90C30.
Key Words: Nonlinear programming; The Kuhn-Tucker theorem; W. Karush; F. John; Lagranges multiplier method; The calculus of variations; The theory of convexity; Multiply discovery.

The Kuhn-Tucker Theorem in Nonlinear Programming: A Multiple Discovery?

1 Introduction

In the summer of 1950 at *the Second Berkeley Symposium on Mathematical Statistics and Probability* which were held in Berkeley, California a mathematician from Princeton, Albert W. Tucker, who was generally known as a topologist, gave a talk with the title *Nonlinear Programming.* It was based on a joint work of Tucker and a young mathematician Harold W. Kuhn, who had just finished his Ph.D. study at Princeton University. The talks was published in a conference proceeding, and for the first time the name *nonlinear programming* – the title Kuhn and Tucker chose for their paper – appeared in the mathematical literature [48]. In the paper Kuhn and Tucker introduced the nonlinear programming problem and proved the main theorem of the theory – the so-called *Kuhn-Tucker theorem.* This theorem which gives necessary conditions for the existence of an optimal solution to a nonlinear programming problem launced the mathematical theory of nonlinear programming.

The result is very famous and not long after its publication people began to talk about it as the Kuhn-Tucker theorem, but apparently Kuhn and Tucker were not the first ones to prove this theorem. In modern textbooks on nonlinear programming there will often be a footnote telling that William Karush proved the theorem in 1939 in his master's thesis from the University of Chicago, and that Fritz John derived (almost) the same result in a paper published in 1948 in an essay collection for Richard Courants 60. birthday. Now a days one will often see the theorem refered to as the *Karush-Kuhn-Tucker theorem* to acknowledge the work of Karush. But the fact is that when he handed in his master's thesis in December 1939 nothing happened: the work was not published, nobody encouraged him to publish his result, apparently is was not very interesting. Fritz John's paper came out only two years before Kuhn's and Tucker's paper, again nobody noticed it. Actually the fact is that John tried to get it published earlier in Duke Mathematics

Journal but they rejected the paper! What I find striking here is that only two years later when Kuhn and Tucker derived the result, it became famous almost instantaneously and caused the launching of a new mathematical research area.

These historical facts leads to the following interesting questions: Was it really the same result they had derived? Is it fair here to talk about a multiple discovery, and in what sense is it or is it not a multiple discovery? Why were the reactions from the mathematical community so different in the three cases? Why did nothing happen the first two times? Or maybe more interesting why did Kuhn and Tucker's work have such an enormous impact?

This paper is centered about these questions. They will be addressed and discussed on the basis of a contextualized historical analysis of the work of John, Karush, Kuhn and Tucker. Both mathematical and social contexts will be considered and the paper will end with a discussion of the role played by the military through Office of Naval Research (ONR) and operations research (OR).

1.1 Mathematical Prerequisites

Let me just very briefly explain what is to be understood by the concept a *nonlinear programming problem* and state more precisely what the Kuhn-Tucker theorem says.

Definition of a Nonlinear Programming Problem: A nonlinear programming problem is an optimization problem of the following type

$$\begin{array}{lll} \text{Minimize} & f(x) & \\ \text{subject to the constraints} & g_i(x) & \leq 0 \quad \text{for } i = 1, \ldots, m \\ & x \in X. & \end{array}$$

Here, X is a subset of $\mathbf{R}^n$, the functions $f, g_1, \ldots, g_m$ are defined on X, and x is an n-dimensional vector $(x_1, \ldots, x_n)$.[1]

Thus a nonlinear programming problem is a finite dimensional optimization problem where the variables has to fulfill some inequality constraints. A variable, $x \in \mathbf{R}^n$, who satisfy all the constraints is said to be *feasible*.

The Kuhn-Tucker Theorem: Suppose X is a nonempty open set in $\mathbf{R}^n$. Let $\bar{x}$ be feasible and the functions f, $g_1, \ldots, g_m$ differentiable at $\bar{x}$. Suppose

[1]For an exposition on the mathematical theory of nonlinear programming see for example [3], [4], [50], [58].

the gradient vectors $\nabla g_i(\bar{x})$, for the binding – or active – constraints, i.e constraints, g_i, for which $g_i(\bar{x}) = 0$, are linearly independent. Then the following will be true:

> A necessary condition for $f(\bar{x})$ to be a minimum for the nonlinear programming problem above is that there exist scalars (multipliers) $u_1, \ldots, u_m$ such that
>
> $$\nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0, \tag{1}$$
>
> $$u_i g_i(\bar{x}) = 0 \quad i = 1, \ldots, m, \tag{2}$$
>
> $$u_i \geq 0 \, , \; i = 1, \ldots, m.$$

These necessary conditions are called *The Kuhn-Tucker Conditions.*

You may recognize the first of these conditions, (1), as to say that the corresponding Lagrangian function, $\phi(x, u) = f(x) + \sum_i^m u_i g_i(x)$, has a critical point in $(\bar{x}, u)$. The second necessary condition, (2), ensures that if $g_i(\bar{x}) \neq 0$, that is, if g_i is not active in $\bar{x}$, then the corresponding multiplier, u_i, is equal to 0.

2 The Theorem of Karush: A Result in the Calculus of Variations

In December 1939 William Karush recieved a master's degree in mathematics from the University of Chicago. His master's thesis had the title "Minima of Functions of Several Variables with Inequalities as Side Conditions" [38].[2] Today we would say that such an optimization problem subject to inequality constraints belong to the subject of nonlinear programming. But since nonlinear programming didn't exist at that time we need to take a closer look at Karush's thesis in order to figure out which field of mathematics it was considered a contribution to. This student project was proposed by Karush's supervisor Lawrence M. Graves, [39], so how did it fit in with the activities in the department of mathematics at Chicago at the time? Why was this problem interesting and what was Karush trying to do?

In the introduction to his thesis Karush stated the purpose of his work, and he also gave a hint where to look for the motivation behind the proposal of the problem. He wrote:

[2] I am very gratefull to Professor W. Karush for providing me with a copy of his thesis.

The problem of determining necessary conditions and sufficient conditions for a relative minimum of a function $f(x_1, \ldots, x_n)$ in the class of points $x = (x_1, \ldots, x_n)$ satisfying the equations $g_\alpha(x) = 0$ ($\alpha = 1, \ldots, m$), where the functions f and g_α have continuous derivatives of at least the second order, has been satisfactorily treated [1]. This paper [Karush's thesis] *proposes to take up the corresponding problem in the class of points x satisfying the inequalities*

$$g_\alpha(x) \geq 0 \qquad (\alpha = 1, 2, \ldots, m),$$

where m may be less than, equal to, or greater than n. [38, 1]

The reference, *[1]*, in the above quotation from the introduction to Karush's thesis is to a paper titled "Normality and Abnormality in the Calculus of Variations" [8]. It had been published just the year before by Gilbert Ames Bliss, who was the head of the mathematics department at Chicago. The problem that Karush's supervisor proposed for the thesis originated from this paper by Bliss on the calculus of variations. So, the roots of the problem Karush set out to work on was burried in the mathematical subdiscipline called the calculus of variations, a field in mathematics that had a special connection to the mathematical department at University of Chicago.

2.1 The Chicago School in the Calculus of Variations

The mathematical department at the University of Chicago was founded with the opening of the university in 1892. The first leader of the department was Eliakim M. More (1862-1932) who in cooperation with the two Germans Oskar Bolza (1857-1936) and Heinrich Maschke (1853-1903) created a mathematical environnement that soon became the leading department of mathematics in the USA [56].

It was Bolza who introduced the calculus of variations as a major research field at the department. His own interest in the topic stemmed from Weierstrass' famous lectures in 1879 and Bolza taught the subject to graduate students at Chicago. From 1901 Bolza also turned his own research towards the calculus of variations. This indicated a shift in research direction and it was caused by a series of talks Bolza gave at the third *American Mathematical Society* (AMS) symposium. The purpose of these AMS meetings was to give an overview of selected mathematical topics for a broader audience of mathematicians and thereby give directions for new research. Bolza had been chosen as one of the main speakers for the 1901 meeting and was asked to talk about hyperelliptic functions but chose instead to give talks on the calculus of variations. Interesting unsolved problems became visible and from then on Bolza was deeply involved in research in that field [56, 394].

Bolza was very popular as a thesis advisor and he often guided his students to work in the field in which he was currently doing research himself. The result was that Bolza created a solid foundations for research in the calculus of variations at Chicago – the so-called Chicago School of the Calculus of Variations [56, 393].

In 1908 Maschke died and two years later Bolza returned to Germany. Chicago then lost two of its leading mathematicans and from 1910 on there seems to have been a decline in the reputation of the Chicago mathematics department. Acording to some Chicago people this decline was caused by a too narrow focus on the calculus of variations.[3] The "new team" at Chicago consisted of Bliss, Dickson and Wliczynski. It was Bliss who as a student of Bolza continued the calculus of variations tradition.

Bliss was head of the department from 1927 to 1941 and this period in the life of the institute was characterized by intense research in the calculus of variations. In the 10-year period from 1927 til 1937 the Chicago institute produced 117 Ph.D. thesis'. Bliss supervised 35 of these and 34 fell within the caluclus of variations [51, 138]. Several mathematicians connected with Chicago later held a very critical view of Bliss' program in the calculus of variations. They seem to share the following view put forward by A. L. Duren who himself was a student of Bliss and wrote a Ph.D. in the calculus of variations:

> *The subject itself* [calculus of variations] *had come to be too narrowly defined as the study of local, interior minimum points for certain prescribed functionals given by integrals of a special form. Generalization came only at the cost of excessive notational and analytical complications. It was like defining the ordinary calculus to consist exclusively of the chapter on maxima and minima* [17, 245].

This is of course a characterization of the Chigago School in the Calculus of Variations under Bliss with hinsight, but it tells something about how extensive the research in the calculus of variations was at the department at the time and that this field was quite narrowly defined at Chicago. As a student in Chicago Karush was a product of this calculus of variations tradition and his master's thesis must be analyzed and discussed within that context.[4]

---

[3]See f.ex. [51], [10], [66], [17].

[4]For more information on the mathematical institute at Chicago under the leadership of Moore see [56], [18]. For the history of the calculus of variations see [29], [30].

2.2 Karush's Master's Thesis

The purpose of Karush's work was to determine necessary and sufficient conditions for a relative minimum of a function $f(x_1, \ldots, x_n)$ in the class of points $x = (x_1, \ldots, x_n)$ satisfying the inequalities $g_\alpha(x) \geq 0$ for $(\alpha = 1, 2, \ldots, m,)$ where the functions f and g_α are subject to various continuity and differentiability conditions. He carried out this work in 1939 at a time were the research was centred on variational calculus problems with *inequalities* as side conditions. Viewed in that context Karush's problem can be interpretated as a finite dimensional version of such a problem in the calculus of variations.

At first sight it can seem a little strange to ask Karush – who was a promising student – to work on a finite dimensional version of the problem that was the real focus of attention. Karush didn't explain the importance of his work in a broader perspective but from his introduction it is clear that he viewed his work as an extension of the work of Bliss, mentionned above, from the year before. From the mid-30.s Bliss had been interested in some properties called *normality* and *abnormality* for the minimizing arc of an *equality* constrained problem in the calculus of variations. The purpose of the paper by Bliss that Karush took as point of departure for his thesis was to

> *... analyze, more explicitly than has been done before, the meaning of normality and abnormality for the calculus of variations. To do this I have emphasized in §1 below the meaning of normality for the problem of a relative minimum of a funtion of a finite number of variables.* [8, 365]

Because as Bliss wrote

> *The significance of the notion of abnormality in the calculus of variations can be indicated by a study of the theory of the simpler* [finite dimensional] *problem.* [8, 367]

Hence, Bliss' idea was that valuable insight into the general more complex cases could be obtained through a thorough study of the finite dimensional case. In the light of this it is reasonable to presume that the same would hold true for the *inequality* constrained case – a study of the finite dimentional version of the problem could throw light on the infinite case. This shows that even though the problem proposed for Karush's thesis didn't fell directly in the main research area in the calculus of variations at the department it would still make sense to work it out.

The theorem which is interesting in relation to the Kuhn-Tucker theorem in nonlinear programming appears in the third section of the thesis. Here Karush examined the minimum problem under the condition that the functions f and g_α, that is the objective function and the contrained functions, are C^1-functions near a point x^0.

Before he proved the theorem which is now recognized as the Kuhn-Tucker theorem he showed a less restricted version, namely the following theorem:

> *Theorem 3:1. If $f(x^0)$ is a minimum then there exist multipliers l_0, l_α not all zero such that the derivatives F_{x_i} of the function*
>
> $$F(x) = l_0 f(x) + l_\alpha g_\alpha(x)$$
>
> *all vanish at x^0.* [38, 12-13][5]

Note, that there is no sign-restriction on the multipliers in these first necessary conditions. Also the multiplier l_0 associated with the objective function, f, can take the value zero, in which case x^0 is called an *abnormal* point. In order to avoid the abnormal case one need some kind of regularity conditions or *constraint qualification* as Kuhn and Tucker later called it.

The concepts Karush introduced to construct such a regularity condition was *admissible direction*, *admissible curve* and *normal point*. By an admissible direction Karush understood a nonzero vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ that solves the inequality system

$$\sum_{i=1}^{n} \frac{\partial g_\alpha}{\partial x_i}(x^0)\lambda_i \geq 0$$

[38, 11]. In other words, Karush considered a direction admissible if the directional derivatives of the constrained functions, g_α, in the direction of λ is nonnegative, which means that "you stay" in the feasible area if "you walk" from x^0 in the direction of λ. He called a regular arc, $x_i(t)$ $(i = 1, 2, \ldots, n;\ \ 0 \leq t \leq t_0)$, an admissible arc if

$$g_\alpha(x(t)) \geq 0 \quad \text{for all} \quad \alpha \quad \text{and} \quad t$$

[38, 11]. This means that a regular arc is admissible if "you stay" feasible when "you move" along the arc. Finely he called a point x^0 normal if the Jacobian matrix for g has rank m at x^0, that is, if the gradients

$$\nabla g_1(x^0),\ \nabla g_2(x^0), \ldots, \nabla g_m(x^0)$$

[5] Karush used the Einstein summation sybolisme, i. e $F(x) = l_0 f(x) + l_\alpha g_\alpha(x)$ means $F(x) = l_0 f(x) + \sum_{\alpha=1}^{m} l_\alpha g_\alpha(x)$.

are linearly independent.

Karush then formulated the later so celebrated Kuhn-Tucker theorem in the following way:

> *Theorem 3:2. Suppose that for each admissible direction λ there is an admissible arc issuing from x^0 in the direction λ. Then a first necessary condition for $f(x^0)$ to be a minimum is that there exist multipliers $l_\alpha \leq 0$ such that the derivatives F_{x_i} of the function*
>
> $$F = f + l_\alpha g_\alpha$$
>
> *all vanish at x^0* [38, 13].[6]

By a curve $x_i(t)$ $(0 \leq t \leq t_0)$, "issuing from x^0 in the direction λ" he meant that $x_i(0) = x_i{}^0$ and $x_i'(0) = \lambda_i$ [38, 13].

His idea was to use Farkas' lemma[7] to guarantee the existence of non-positive multipliers, l_α, and the assumptions in the theorem – the regularity condition – ensures precisly that Farkas' lemma can be brought into action.

2.3 The Acknowledgement of Karush's Thesis in Non-linear Programming

Karush's theorem looks indeed very much like the version of the Kuhn-Tucker theorem I showed in the introduction. The symbol $l_\alpha g_\alpha$ in Karush's formulation of the theorem means $\sum l_\alpha g_\alpha$. That is there should exist multipliers (l_α) such that the Lagrangian function, F, has a critical point at x^0, (l_α). The condition $l_\alpha g_\alpha(x^0) = 0$ is missing because Karush only considered the active constraints, e.i constraints for which $g_\alpha(x^0) = 0$.

Actually in 1975 Harold Kuhn wrote a letter to Karush saying:

> *First let me say that you have clear priority on the results known as the Kuhn-Tucker conditions (including the constraint qualification). I intend to set the record as straight as I can in my talk.* [42]

The talk Kuhn refered to is one he had been asked to give on the history of nonlinear programming at an AMS symposium. Kuhn became aware of the work of Karush through Takayama's book *Mathematical Economics* [67], [44,

[6]Karush used the Einstein summation sybolisme, i. e $F(x) = f(x) + l_\alpha g_\alpha(x)$ means $F(x) = f(x) + \sum_{\alpha=1}^{m} l_\alpha g_\alpha(x)$.

[7]For a short historical account on Farkas lemma see [9]. To consult Farkas' own work see [22].

10]. During the research for the AMS talk Kuhn took contact to Karush and offered a partial publication of the master's thesis as an appendix to Kuhn's historical paper in the AMS proceeding that was to be published after the meeting. In this paper Kuhn announced Karush's thesis as an unpublished classic in the field of nonlinear programming [44].

Just looking at Karush's result independent of the context of discovery one can only agree with Kuhn and say that Karush actually had the later so celebrated Kuhn-Tucker theorem. In the light of its later importance one is then naturally led to the questions: Why wasn't Karush's result valued at the time? Why wasn't it published?

The main interest in Chicago at the time was – as we have seen – variational calculus with inequality contraints and if Karush's work is evaluated in this context instead of the context on nonlinear programming (which didn't exist at the time) Karush's work was only a minor, finite dimentional thing. It was not very interesting or exciting just a minor corner, some 'cleaning up' in a research direction where variational calculus with inequality constraints was the main field. Neither the posed problem nor the results was something special, the interesting questions in this field were different from the questions that were important and later guided the research in nonlinear programming.

The letter from Kuhn to Karush qouted above also gives a flavor of how important the theorem is considered to be in the mathematical community of people doing nonlinear programming. Kuhn tells in the letter that Richard Cottle, who was among the organizers of the AMS symposium, made the following remark about Karush when he heard about Kuhn's intentions of "setting the record straight":

> *'you must be a saint' not to complain about the absence of recognition.* [42]

Kuhn also writes about Tucker's reaction when he learned about the result in Karush's thesis. Tucker was truly amased that Karush never had told him about his work when they met at *the Rand Cooperation* [42]. Richard Bellman wrote the following to Kuhn when he learned about Kuhn's upcoming talk:

> *I understand from Will Karush that you will try and set the record straight on the famous Kuhn-Tucker condition. I applaud your effort. Fortunately, there is enough credit for everybody. It would certainly be wonderful if you wrote it as the Kuhn-Tucker-Karush condition.*
>
> *Like many important results, it is not difficult to establish, once observed. That does not distract from the importance of the condition.* [7]

Also Phil Wolfe informed Kuhn how pleased he was that Karush's work would now be recogniced [43]. These reactions from wellknown members of the nonlinear programming society not directly involved in the development of the Kuhn-Tucker theorem shows that the theorem is considered to be very important and is something that people would like very much to be associated with.

From the letters it is clear that the mathematicians working in the field are truly amazed that Karush had not come forward to claim if not priority then at least recognition. To this Karush himself gave the following explanation:

> *That does not answer the question of why I did not point to my work in later years when nonlinear programming took hold and flourished. The thought of doing this did occur to me from time to time, but I felt rather diffident about that early work and I don't think I have a strong necessity to be 'recognized'. In any case, the master's thesis lay buried until a few years ago when Hestenes urged me to look at it again to see if it shouldn't receive its proper place in history ... . So I did look at the thesis again, and I looked again at your work with Tucker. I concluded that you two had exploited and developed the subject so much further than I, that there was no justification for my announcing to the world, 'Look what I did, first'.* [39]

From a history of mathematics point of view I think Karush is right here. He did derive a result that was comparable to the result later developed by Kuhn and Tucker, but Karush did not exploite the subject further, his work was *not* nonlinear programming, it was a work that came into being in a completely different context. The institute at Chicago had became under Bliss a place with focus on a very narrowly defined calculus of variations research programme and within this research direction nobody was interested in exploring the possibilities for important potentially applications of Karush's result. The questions that drove the research in the calculus of variations at Chicago was different from the ones that later guided the research in nonlinear programming.

3 The Theorem of F. John: A Contribution to the Theory of Convexity

Fritz John's version of the Kuhn-Tucker theorem appeared in his essay *Extremum Problems with Inequalities as Subsidiary Conditions* which was published in 1948 in the *Courant Anniversary Volume* [37].

John was a student of Richard Courant in Göttingen where he recieved a Ph.D in 1933. He had Jewish ancestors and Courant worked hard to get John a position outside of Germany. In 1934 he succeeded in getting John a research scholarship at Cambridge, England. John moved to USA a year later where he got an offer from the university in Kentucky. He worked there until 1943 and after some years of war related work at the Ballistic Research Laboratory at Aberdeen Proving Ground he returned "home" to Courant at his institute at the New York University [61, 131-132, 154-155].

Fritz John was a world class mathematician. His list of publications counts 101 mathematical texts - papers as well as monographs, and he has received a lot of prizes and fellowships. Today he is probably most recognized for his work on partial differential equations but he has also made important contributions in the fields of geometry, analysis and nonlinear elasticity. At the time where Courant's *Anniversary Volume* was published John had mostly been working within the theory of convexity – more than half of his mathematical publications from the first one that appeared in 1934 until this one published in 1948 was in the field of convexity and quite a few of these are now considered as "classics" in the theory of convexity [34]. The point here is that in the period leading up to the publication of Johns paper on extremum problems with inequalities as side conditons he was deeply rooted in a mathematical environment of the theory of convexity [53].

3.1 John's Paper

What was John's intentions in this paper? The title of the paper was *Extremum Problems with Inequalities as Subsidiary Conditions* and in the introduction he wrote:

> *This paper deals with an extension of Lagrange's multiplier rule to the case, where the subsidiary conditions are inequalites instead of equations. Only extrema of differentiable functions of a finite number of variables will be considered.* [37, 187]

Like Karush John only looked at the finite dimensional case so judging from the title and the introduction it sounds very much like John was interested in the same kind of questions as Karush. This impression gets reinforced later in the introduction where John pointed ahead to further directions of research on the problem:

> *from the point of view of applications it would seem desirable to extend the method used here to cases, where the functions involved .... do not depend on a finite number of independent variables.* [37, 187]

This extension of the problem clearly belongs to the calculus of variations but if John considered his work as a contribution to this field, it would seem unlikely that he shouldn't know the work of the Chicago School in the calculus of variations – well known at the time – who had already carried out this work for the general case.

Apparently John didn't know the Chigago work. There is no reference at all to the calculus of variations in his paper. It seems that John was not really "interested" in the calculus of variations. What was his real interest, then? In the following I will scrutinize his paper to see *what* he actually did and *how* he did it. The paper is divided into two parts where the first one is concerned with the question of necessary and sufficient conditions for the existence of a minimum and the second part is devoted to two geometrical applications of the theoretical result in part one.

John formulated the result that later got acknowledges as a version of the Kuhn-Tucker theorem in the following way:

> Let R be a set of points x in $\mathbf{R}^n$,[8] and $F(x)$ a real-valued function defined in R. We consider a subset R' of R, which is described by a system of inequalities with parameter y:
>
> $$G(x, y) \geq 0,$$
>
> where G is a function defined for all x in R and all "values" of the parameter y. ... we assume that the "values" of the parameter y vary over a set of points S in a space H. ... We are interested in conditions a point x^0 of R' has to satisfy in order that
>
> $$M = F(x^0) = \min_{x \in R'} F(x)$$
>
> [37, 187-188].

Under some further continuity and differentiability conditions John was able to prove the following theorem:

> Theorem I.
> Let x^0 be an interior point of R, and belong to the set R' of all points x of R, which satisfy the contraints $G(x, y) \geq 0$ for all $y \in S$. Let
>
> $$F(x^0) = \min_{x \in R'} F(x).$$

[8] Instead of $\mathbf{R}^n$ John wrote ... *in a space E,* but in the following he restricted himself to the case, where the space E containing the set R is the n-dimensional euclidean space, which I have called $\mathbf{R}^n$ [37, 188].

Then there exists a finite set of points $y^1, \dots, y^s$, in S and numbers $\lambda_0, \lambda_1, \dots, \lambda_s$, which do not all vanish, such that

$$G(x^0, y^r) = 0 \qquad \text{for } r = 1, \dots, s$$

$$\lambda_0 \geq 0, \ \lambda_1 > 0, \dots, \lambda_s > 0,$$

$$0 \leq s \leq n,$$

the function

$$\phi(x) = \lambda_0 F(x) - \sum_{r=1}^{s} \lambda_r G(x, y^r)$$

has a critical point at x^0 i.e. the partial derivatives are zero at x^0:

$$\phi_i(x^0) = 0 \ \text{ for } \ i = 1, \dots, n.$$

(See [37, 188-189]).

John's way of "attacking" the problem was the same one as Karush used, but where Karush invoked Farkas' lemma as his main tool John used other similar results from the theory of convexity, results that he was familiar with through works of Dines and Stokes among others [15], [65].

John's formulation of the theorem though looks a little different than Karush's but the stated conditions are the Kuhn-Tucker conditions. The difference is the appearence of the parameter, y, in the parameter set, S, and that the multiplier, λ_0, associated with the objective function, F, can become zero as in Karush's first theorem. The last thing is caused by the fact that John didn't have the constraint qualification as Kuhn and Tucker called it or the normality condition as Karush would have said.

3.2 The Two Geometrical Applications

From reading the second part of the paper which is concerned with the two geometrical applications it becomes clear why John chose this construction with the parameter, y, and a parameter set, S. It also offers an explanation to why John did not touch upon the problem of abnormality and thereby didn't consider the problem of constraint qualification.

More than half of the paper is devoted to these geometrical applications. The first one is *Application to minimum sphere containing a set* and the second is about the ellipsoid of least volume containing a set, S, in $\mathbf{R}^m$ [37, 193]. In the first one John considered the following problem:

Let S be a bounded set in $\mathbf{R}^m$. Find the sphere of least positive radius enclosing S.

John was not interested in the *existence* of such a sphere. If the assumption is made that the bounded set, S, contains at least two distinct points it is quite clear that such a sphere exist [37, 194].

To be able to use his theorem John characterized spheres in $\mathbf{R}^m$ as points in $\mathbf{R}^{m+1}$:

$$x = (x_1, \dots, x_{m+1})$$

where $(x_1, \dots, x_m)$ are the coordinates of its center and x_{m+1} the square of its radius. He could then rewrite the problem as an optimization problem subject to inequality constraints:

Minimize the function $F(x) = x_{m+1}$ subject to the constraints

$$G(x, y) = x_{m+1} - \sum_{i=1}^{m} (x_i - y_i)^2 \geq 0 \quad \text{for all} \quad y \in S.$$

The constraints ensures that the minimum is only sought among spheres containing S.

John used a similar procedure in the second application about the ellipsoid. In both cases he knew that a minimum, x^0, existed, so the necessary conditions of the theorem is fulfilled. He then used these to derive significant properties of the minimum sphere and the minimum ellipsoid. From the last application – the one with the ellipsoid – he derived several general properties of closed convex sets [37, 201-202].

3.3 The Link to the Theory of Convexity

In the application part of John's paper and especially in the last application concerning the ellipsoid it becomes clear that John's main interest was the results about closed convex sets that he developed through the applications of his theoretical result – the extension of Lagrange's multiplier method – to problems with inequality constraints. In connection with Kuhn's talk on the history of nonlinear programming Kuhn also had a small correspondence[9] with John. According to Kuhn, John should have revealed that he was led to the theorem when he was

[9] This correspondence is apparently lost.

trying to prove the theorem ... that asserts that the boundary of a compact convex set S in $\mathbf{R}^n$ lies between two homothetic ellipsoids of ratio $\leq n$, and that the outer ellipsoid can be the ellipsoid of least volume containing S. [44, 15]

Eventhough John in his title and introduction gives the impression that he is concerned with problems in the calculus of variations it is my opinion that his paper rather should be viewed as a contribution to the theory of convexity. John had at the time – as mentioned earlier – made fine contributions to convexity theory. All the references in the paper are either to the theory of convexity or to less general works – by John and others – on the two applications.[10] From reading the application part of the paper it becomes quite clear that the applications has a justification in them selves, they serve a deeper purpose than just as illustrations of the theoretical result in the first part of the paper. The conlusion must be that the guiding questions – the important issues for John – was the applications and the results he could derive from these.

3.4 The Status of The Theorem

In Karush's work the theorem was important in itself. The hole purpose of Karush's work was to derive these necessary conditions for the existence of a minimum or maximum. In John's work on the other hand the theorem had a completely different status. It was not at all the main thing it was only derived as a tool for deriving general results about convex sets. The applications guided the formulation of the theorem which explains John's contruction of the 'parameter set' which clearly is dictated by the applications. Another difference between Karush's and John's work is the 'normality' condition, as Karush called it, or the 'constraint qualification' as Kuhn and Tucker would call it, John does not touch upon that problem. This can also be explained from the applications, both of them are actually examples of the normal case, so John never ran into that problem.

In his paper on the history of nonlinear programming Kuhn wrote about John's work that it "*very nearly joined the ranks of unpublished classics in our subject* [e. i. nonlinear programming] " [44, 15]. But John himself apparently did not view this work as a contribution to what later became nonlinear programming. He never came forward with priority claims or anything like that.

[10][35], [36], [5], [6], [1].

4 The Theorem of Kuhn and Tucker: An Extension of Linear Programming

Albert W. Tucker was born in Canada at the year 1905 and he died in Princeton, US. a few years ago more precisely January 27, 1995. He got a Bachelor degree in mathematics from University of Toronto in 1928 and a year later he began a Ph.D study at Princeton University. This turned out to be the beginning of a life long connection to the mathematical department at Princeton. In 1932 he recieved the Ph.D degree on a thesis in the field of topology, and two years later he got appointed assistent professor. In 1938 he became associate professor and then full professor in 1946. He was an important figure in the establishment in the 30.s and 40.s of Princeton as a prestigious place for mathematical research and he served as head of the department for the ten year period 1953-1963. He had a tremendous influence on the students who came in contact with him, and he is often characterized as a very good teacher and leader [68], [47].

Harold W. Kuhn - twenty years younger than Tucker - was born and raised in California. He got a Bachelor degree in Science from California Institute of Technology in 1947 and then moved on to Princeton where he wrote a Ph.D thesis on *Subgroup Theorems for Groups Presented by Generators and Relations* in 1950 [41]. After some traveling and a seven year appointment at Bryn Mawr College Kuhn returned to Princeton as associate professor. He was connected to both the mathematical and the economic departments [45].

Together Kuhn and Tucker defined the nonlinear programming problem and proved their famous theorem *The Kuhn-Tucker Theorem* in the joint paper *Nonlinear Programming* published in 1950 [48].

4.1 The Nonlinear Programming Paper

The main point in Kuhn and Tucker's paper was to find necessary and sufficient conditions for the existence of a solution to the following *Maximum Problem* – as they called it then:

> *To find an x^0 that maximizes $g(x)$ constrained by $Fx \geqq 0$, $x \geqq 0$* [48, 483].

Here $x^0 \in \mathbf{R}^n$, $x \longrightarrow u = Fx$ is a differential mapping of nonnegative n-vectors x into m-vectors u. That is, Fx is an m-vector whose components $f_1(x), \ldots, f_m(x)$ are differentiable functions of x defined for $x \geq 0$. $g(x)$ is a differentiable real function of $x \in \mathbf{R}^n$ defined for $x \geq 0$ [48, 483].

Kuhn and Tucker handled this problem by taking the so-called *saddle value problem* as their point of departure. They defined the saddle value problem as the problem of finding nonnegative vectors $x^0 \in \mathbf{R}^n$ and $u^0 \in \mathbf{R}^m$, such that

$$\phi(x, u^0) \leqq \phi(x^0, u^0) \leqq \phi(x^0, u) \qquad \text{for all } x \geqq 0,\ u \geqq 0,$$

where $\phi(x, u)$ is a differentiable function of an n-vector x with components $x_i \geqq 0$ and an m-vector u with components $u_h \geqq 0$.

They led ${\phi^0}_x$, ${\phi^0}_u$ denote the partial derivatives, evaluated at a particular point x^0, u^0. That is ${\phi^0}_x$ is an n-vector:

$${\phi^0}_x = (\frac{\partial \phi}{\partial x_1}(x^0), \ldots, \frac{\partial \phi}{\partial x_n}(x^0)),$$

and ${\phi^0}_u$ is an m-vector:

$${\phi^0}_u = (\frac{\partial \phi}{\partial u_1}(u^0), \ldots, \frac{\partial \phi}{\partial u_m}(u^0)).$$

They used the "′"-notation to denote the transposed vector.

The first theorem Kuhn and Tucker then proved in the paper concerned the question of necessary and sufficient conditions for the existence of a solution to the saddle value problem. They proved that the conditions

$$(1) \qquad {\phi^0}_x \leqq 0, \quad {\phi^0}'_x x^0 = 0, \quad x^0 \geqq 0$$

$$(2) \qquad {\phi^0}_u \geqq 0, \quad {\phi^0}'_u u^0 = 0, \quad u^0 \geqq 0$$

are necessary that x^0, u^0 provide a solution for the saddle value problem [48, 482-483]. For the second part of the question they proved that the conditions (1), (2) together with the following two conditions

$$(3) \qquad \phi(x, u^0) \leqq \phi(x^0, u^0) + {\phi^0}'_x(x - x^0)$$

$$(4) \qquad \phi(x^0, u) \geqq \phi(x^0, u^0) + {\phi^0}'_u(u - u^0)$$

for all $x \geqq 0$, $u \geqq 0$, are *sufficient* that x^0, u^0 provide a solution for the saddle value problem [48, 483].

Equiped with these conditions for the saddle value problem Kuhn and Tucker phrased their later so celebrated theorem as:

> *Theorem 1. In order that x^0 be a solution of the maximum problem, it is necessary that x^0 and some u^0 satisfy conditions* (1) *and* (2) *for $\phi(x, u) = g(x) + u'Fx$.* [48, 484]

If the condition $x^0 \geqq 0$ is incoorporated in the constraint function, F, the first and last condition in (1) together means that the Lagrangian function $\phi(x,u)$ has a critical point in x^0, u. The second condition in (1) ensures that the multipliers associated with the non-binding components of x^0 is equal to zero. The first condition in (2) ensures that x^0 is feasible, the second that the multipliers associated with non-binding constraints are equal to zero and the last one is the sign-restriction on the multipliers.

These conditions later got named *the Kuhn-Tucker conditions* and they constitute one of the main results in the mathematical theory of nonlinear programming.

Actually the first time Kuhn and Tucker revealed this theorem was not at the Berkeley Symposium but a few month earlier at a seminar held at the RAND in Maj 1950. Among the audience was C. B. Tompkins who came up with something as unpleasent as a counter example. [44, 14]. The result – as it was – could not rule out the *abnormal* case as Karush would have called it. Kuhn and Tucker got back to work and realized the need for some regularity conditions on the constraint functions. This led them to introduce the term *constraint qualification*. The constraint qualification they used in the 1950 paper was the same as Karush's that for each x^0 of the boundary of the set determined by the constraints and for any vector differential, dx, for which the directional derivatives of the binding constraints in the direction of dx are non-negative there corresponds a differentiable arc $x = a(\theta)$, $0 \leqq \theta \leqq 1$, contained in the constrained set, with $x^0 = a(0)$, and some positive scalar λ such that $a'(x^0) = \lambda dx$ [48, 483].

It can as Kuhn and Tucker pointed out in the paper seem artificial to introduce the conditions (3) and (4) that occured in the sufficiency part of the saddle value problem but these conditions are satisfied if $\phi(x, u^0)$ is a concave function of x and $\phi(x^0, u)$ is a convex function of u [48, 483]. In order to get full equivalence between solutions of the maximum problem and the saddle value problem Kuhn and Tucker then required that the involved functions, g, $f_1, \dots, f_m$, were concave as well as differentiable for $x \geqq 0$. With these extra requirements Kuhn and Tucker showed that:

> ... x^0 *is a solution of the maximum problem if, and only if,* x^0 *and some* u^0 *give a solution of the saddle value problem for* $\phi(x,u) = g(x) + u'F(x)$. [48, 486]

4.2 The Saddle Value Problem: A Detour?

Kuhn and Tucker's formulation of the theorem is different from that of Karush's and John's neither of them considered the concept of saddle points. Why did Kuhn and Tucker choose the saddle point formulation? and why were they looking for an equivalence theorem between the maximum problem and the saddle value problem? The mathematical context the work of Kuhn and Tucker originated in can provide an answer to these questions.

The cooperation between Kuhn and Tucker had begun two years earlier, in 1948, where they had examined the relation between game theory and the linear programming model for a logistic problem in the U. S. Air-Force that had just been developed by George B. Dantzig. Kuhn was still a student at the time and together with another student, David Gale, the three of them worked out the mathematical foundations for linear programming [31]. They formulated the corresponding dual problem, proved the duality theorem and showed the relation between linear programming and game theory.[11]

Tucker's first association when he got introduced to the linear programming problem was that it reminded him of Kirchoff's laws for electrical networks [2, 342-343]. In the fall of 1949 just after Kuhn, Gale and Tucker had presented their work on linear programming and game theory at the first conference on linear programming which was held in Chicago in June 1949 Tucker went on leave to Standford. Here he dug deeper into this first association of his and discovered the underlying optimization problem of minimizing heat loss. According to Kuhn, this knowledge led Tucker to the recognition that the Lagrangian multiplier method which is normally used to solve *equality* constrained optimization problems could be adapted to optimization problems subject to *inequality* constraints [44, 12-13]. Tucker then wrote Kuhn and Gale and invited them to continue the work and extend their duality result for linear programming to quadratic programming, i.e to problems where the involved functions no longer has to be linear but can be quadratic [44, 12-13]. David Gale declined the offer but Kuhn accepted and he and Tucker developed the theory in a correspondence between Standford and Princeton.[12]

Thus, the original purpose of Kuhn and Tucker's work was to extend the duality result from linear programming to quadratic programming and the idea was to adapt the classical Langrangian multiplier method. In the introduction to the *Nonlinear programming* paper Kuhn and Tucker explained

[11]A linear programming problem is a nonlinear programming problem where all the involved functions are linear functions. To a linear programming problem one can formulate another linear programming problem on the same data called the *dual* programme. The duality theorem says that the original, *primal*, problem has a finite optimal solution if and only if the dual problem has a finite optimal solution, and the optimum values will be the same.

[12]This coorespondence is lost, I know about it from an interview with Kuhn, who also mentioned it in [44, 13].

how this would work for linear programming. From a linear programming problem

$$\text{maximize} \quad g(x) = \sum c_i x_i, \qquad c_i \in \mathbf{R},$$

where $x_1, \ldots, x_n$ are n real variables constrained by $m+n$ linear inequalities:

$$f_h(x) = b_h - \sum a_{hi} x_i \geqq 0, \quad x_i \geqq 0,$$

where $h = 1, \ldots, m$, $i = 1, \ldots, n$, a_{hi}, $b_h \in \mathbf{R}$, they formed the corresponding Lagrangian function:

$$\phi(x, u) = g(x) + \sum u_h f_h(x), \qquad u_h \in \mathbf{R}.$$

They realised that $x^0 = ({x^0}_1, \ldots, {x^0}_n)$ will maximize $g(x)$ subject to the given constraints if, and only if, there exists a vector $u^0 = ({u^0}_1, \ldots, {u^0}_m) \in \mathbf{R}^m$ with components ${u^0}_i \geq 0$ for all i, such that (x^0, u^0) is a saddle point for the Lagrangian function $\phi(x, u)$ [48, 481]. The really neat and interesting thing about this saddle point result for linear programming was, as Kuhn and Tucker phrased it:

> *The bilinear symmetry of $\phi(x, u)$ in x and u yields the characteristic duality of linear programming.* [48, 481]

The bottom line here is that a linear programming problem has a solution if, and only if, the correspronding Lagrangian function has a saddle point, this saddle point then constitute a solution not only to the linear programming problem but also to the dual programme. Considering now that Kuhn and Tucker actually were searching for a way to extend the duality theorem for linear programming to more general cases[13] it seems perfectly naturally to take the saddle point for the Lagrangian function as the starting point.[14]

In section 6 I will return to the significance of the mathematical and social contexts that Kuhn and Tucker's work originated in.

5 The Aspect of Muliple Discovery

The reason why a question of multiple discovery pops up in connection with a historical study of the Kuhn-Tucker theorem in nonlinear programming is

[13] Somewhere along the proces they shiftet the focus from the quadratic case to the general nonlinear case.

[14] It is striking then that Kuhn and Tucker did not mention duality for nonlinear programming in the paper. The first duality result for nonlinear programming was derived by Werner Fenchel in 1951, published 1953 [23].

that the result today in textbooks and in papers on the history of mathematics is ascribed to all of them - Karush, John and Kuhn and Tucker.[15]

One can also see the result ascribed to the Russian physicist Mikhail Ostrogradsky (1801-1862) and the Hungarian physicist Julius Farkas (1847-1930). In three papers O. I. Franksen discusses Fourier's extension of the Principle of Virtual Work in mechanics and how it sheds new light on the development of the Second Law of thermodynamics and mathematical programming [26], [27], [28]. He concludes that the Kuhn-Tucker theorem is an independent rediscovery, by Kuhn and Tucker, of a theorem derived by Ostrogradsky in a paper which was read for the French Academy in 1834 and published four years later, in 1838 [28, 337-338, 353, 355]. A. Prékopa gives an account on the development of optimization theory in a paper from 1980. He had searched for the first appearence of the Kuhn-Tucker conditions in the literature and he found it at Ostrogradsky and Farkas [59, 528].

Before I return to the question whether Karush's, John's, Kuhn's and Tucker's work can be said to count as a multiple discovery I will briefly deal with these older sources which discuss questions belonging to the field of analytical mechanics – questions that came out of Fourier's extension of the Principle of Virtual Work.

5.1 The Kuhn-Tucker Theorem in Analytical Mecanics

John as well as Kuhn and Tucker mentioned explicitly that their work in one way or another was connected with the Lagrangian multiplier method. John wrote directly in his introduction that the purpose of his work was to extend this method to problems with inequality constraints. Tucker associated the network nature of linear programming with Kirchoff's laws for electrical networks and got the idea that maybe the Lagrangian multiplier method could be adapted to inequality constraint cases.

Lagrange developed his multiplier method in *Mécanique analitique* from 1788 as a method for finding an equilibrium for a mechanical system [49]. He founded his theory of equilibrium on what is now called the Principle of Virtual Work which he took as an axiom. In modern terms the principle states that in order for an equilibrium to take place the virtual work of the applied forces acting on the system is equal to zero. This principle was stated in terms of reversible displacements which means that if a virtual displacement, δr, is alloved then the opposite displacement, $-\delta r$, is also possible without breaking the constraints on the system. This means that the mecanical sy-

[15] [4, 149], [58, 169]. For an account on the prehistory of linear and nonlinear programming see [32], [33]. For an account on the history of nonlinear programming see [44], [40].

stem is subject only to constraints that can be formulated as equations [26, 137].

The inequalities entered the picture in 1798 where Fourier extended the Principle of the Virtual Work to irreversible displacements, that is to mechanical systems subject to inequality constraints [25]. Based on arguments concerning "le moment de la force" [25, 479] he formulated the conditions for equilibrium for such systems as an inequality condition, as he realized that such a system is in equilibrium if, and only if, the virtual work of the applied forces is nonpositive [25, 494]. This inequality is often called the Fourier Inequality.

Ostrogradsky derived in 1834, and published 1838, the conditions for equilibrium for such a system [55]. He denoted the applied forces acting on a system by P, Q, R, ... and the equilibrium condition, the Fourier Inequality, then stated that the total work

$$Pdp + Qdq + Rdr + \ldots$$

has to be nonpositive for every feasible displacement. The constraints was named L, M, ... and because these constraints are given by inequalities Ostrogradsky argued that dl, dM, ... can only change sign in cases where one moves from feasible to infeasible displacements [55, 131].

The 'trick' he used was to change the coordinates by introducing so-called *generalized* coordinates. So instead of considering dp, dq, dr ... Ostrogradsky introduced some other variations $d\xi$, $d\eta$, $d\psi$, ... which are functions of dp, dq, dr, ... and in number equals the number of the original variables. Since dL, dM, ..., are also functions of dp, dq, dr, ... Ostrogradsky took these to be the first of the new generalized coordinates.[16] He then reformulated the hole thing with these new coordinates and ended up with the following equilibrium condition:

$$\lambda dL + \mu dM + \ldots + Ad\xi + Bd\eta + Cd\zeta + \ldots \leq 0$$

for every feasible displacement [55, 131]. Using arguments about the impossibility of changing signs for dL, dM, ... and the possibility of sign changing for $d\xi$, $d\eta$, $d\psi$, ... Ostrogradsky concluded that $A = B = C = \ldots = 0$. This meant that the total work, $Pdp + Qdq + Rdr + \ldots$, equals $\lambda dL + \mu dM + \ldots$, i.e

$$Pdp + Qdq + Rdr + \ldots = \lambda dL + \mu dM + \ldots$$

[16]This means that Ostrogradsky's method can be used only when the number of constraints does not exceed the number of variables.

for all feasible displacements. Since dL, $dM, \ldots$ cannot change sign the equilibrium condition can only take place, he concluded, if the multipliers λ, $\mu, \ldots$ has the opposite sign as the cooresponding constraint, dL, $dM, \ldots$ [55, 132].

Ostrogradsky then ended up by concluding that:

> ... *les conditions de l'équilibre d'un système quelconcque seront exprimées*
>
> 1^{mo} *par l'équation*
>
> $$0 = Pdp + Qdq + Rdr + \ldots + \lambda dL + \mu dM + \ldots$$
>
> *qui doit avoir lieu pour tous les déplacemens imaginables,*
>
> 2^{do} *par la condition que les quantités λ, $\mu, \ldots$ aient respectivement les mêmes signes que les différentielles dL, $dM, \ldots$ pour les déplacemens possibles.* [55, 132-133]

Today, in a situation were a potential, V, exists, that is if $P = -\frac{\partial V}{\partial p}$, $Q = -\frac{\partial V}{\partial q}$, $R = -\frac{\partial V}{\partial r}, \ldots$, one can 'translate' the question of finding an equilibrium into a problem about minimizing the potential energy. So, the conclusion of Franksen and Prékopa that Ostrogradsky here formulated as well as argued for what we call the Kuhn-Tucker theorem in nonlinear programming can only be understood – I think – with this interpretation and 'translation' of Ostrogradsky's work. My opinion – as an historian of mathematics – is that in ascribing the Kuhn-Tucker theorem to Ostrogradsky too much has been read into these sources. In the next section – The Significance of the Context – I will provide further reasons for this response.

The mathematical foundations for the extension of Lagrange's multiplier method to equilibrium for mechanical systems subject to irreversible displacements was treated by Farkas. The main mathematical result that came out of this is Farkas' lemma about linear inequality systems [22]. Farkas developed it in some earlier papers [19], [20], [21] whose main focus was

> ... *zu erweisen, dass mit einer passenden Modifikation die Methode der Multiplikatoren von Lagrange auch auf das Fourier'sche Princip übertragen werden kann.* [19, 266]

There is a remarkable resemblance with the goal stated in John's introduction but here in a context of analytical mechanics.

Farkas knew the work of Ostrogradsky and he made a remark about the limitations of the method used by Ostrogradsky to situations were the number of constraints does not exeed the number of variables [19]. Farkas wanted to come up with a method that could be used in any problem no matter what the relationship is bewteen the number of constraints and the number of variables [19, 266]. He was very much concerned with the mathematical foundations of the method and he had a clear insight that homogeneous linear inequalities could provide a satisfactory foundation so he began his 1895 paper with such a theory:

> *§I. enthält eine algebraische Einleitung über die homogenen linearen Ungleichheiten als mathematische Grundlage der weiteren Betrachtungen.* [19, 266]

This "algebraische Einleitung" consist of a proof of what we now call Farkas lemma. With the help of that Farkas was able to reach the same conclusion as Ostrogradsky but this time for the general problem were there are no restriction on the relation between the number of variables and the number of constraints on the system. Again, if a potenial, V, exists Farkas' results can be translated and interpreted as the Kuhn-Tucker conditions, but the same conclusion as for Ostrogradsky also holds for Farkas.

The work of Ostrogradsky and Farkas had no direct influence on the development of nonlinear programming. It is true that Farkas' lemma functions as an important tool in both the work of Karush and that of Kuhn and Tucker but this is in a version were Farkas lemma is completely removed from analytical mechanics and questions about equilibrium conditions – a separation that was initiated by Farkas in his 1901 paper, *Theorie der einfachen Ungleichungen*, where both the title and the content solely refer to abstract theory of inequalities [22].

5.2 Theories on Multiple Discoveries

The mathematical community does not ascribe the Kuhn-Tucker theorem to Ostrogradsky and Farkas, but they do consider the work of Karush and John as papers belonging to the field of nonlinear programming and both names now appear in textbooks on nonlinear programming. The Kuhn-Tucker theorem is now often renamed the Karush-Kuhn-Tucker theorem and there is also a Fritz John theorem [4]. Also my analysis of Karush's, John's, Kuhn's and Tucker's work seems to indicate that we may actually have a multiple discovery. What I find particulary interesting is the fact that three occurences of a result – which the scientific community later viewed as the samme – developed within a time span of only 11 years was received so differently. In order to examine and understand this phenomenon I turned to the theories on multiple discoveries.

A central figur in the literature on multiple discoveries in science is Robert K. Merton. His main criteria for talking about a multiple discovery is independent discovery of the same scientific result and his theory is that multiple discoveries in science is not something special on the contrary it is the discoveries that on the surface appear to be single that deserves special attention. It is Merton's hypothesis that a thorough investigation will show that these singletons will turn out to be if not multiple then at least potentially multiple. According to Merton "all scientific discoveries are in principle multiples," [52, 356]. He has 10 different arguments for this hypothesis: First of all he points to the huge class of singletons which later turns out to be rediscoveries of results found in earlier work - unpublished or published "obscure" places. Then he has six arguments that all are concerned with the problem of "being anticipated". Merton here describes situations were the scientist for some reason suddently realize that some one else already has developed the result he or she is working on. If the scientist then let go of the project the discovery is an example of a singleton which in reality were a potential multiple discovery. If the scientist goes ahead an publishes anyway there will typically be a footnote saying that this or that person arrived at this conclusion in this or that paper. The last three of Merton's arguments deals with the behavior of the scientist which in Merton's view reveal that they themselves believe that all scientific discoveries are potentially multiple. Here Merton refer to all the different things scientist do in order to secure that they will not be anticipated by another scientist: they carefully date their notes, they "leak" informations about their ideas and circulate incomplete versions of their work [52, 358-361]. The reason for this behavior, Merton points out, is based on a whish to ensure priority which is very important in the scientific world:

> *the culture of science puts a premium not only on originality but on chronological firsts in discovery, this awareness of multiples understandably activates a rush to ensure priority.* [52, 361]

Evaluated according to Merton's theory for multiple discoveries the Kuhn-Tucker theroem is a triple discovery. Some of the circumstances Merton points out can be found in the work of Kuhn and Tucker: Tucker presented their work at a meeting before they had the theory thoroughly worked out. Kuhn told me that he felt that the *Berkely Symposium on Mathematical Statistics and Probability* was an odd place to present their work but explained it by arguing that is provided an oppertunity to get the result published fast [46]. Another of Merton's points also holds for Kuhn and Tucker. They do not have a footnote in the paper saying that Fritz John had worked on the same problem but they do have a reference to his paper. In an interview Kuhn told me that the reference to John was made in the proof reading stage were some one told them about his work [46].

Merton's hypothesis has not survieved undisputed. It has been criticized by Don Patinkin who points out especially two issues which he finds have not received proper attention: First, what is it actually that has been discovered and second to what degree does the discovery form part of the central message of the scientist [57, 306]. Patinkin claimes that a lot of so-called multiple discoveries will turn out to be singletons if they are subject to an analysis that takes these two issues seriously. Patinkin's own *central message* is that a scientist cannot be considered as having made a discovery unless this discovery form part of the central message of the scientist. The question now is of course how to identify the central message. Patinkin sets up the following criteries:

> *... the central message of a scientic work is announced by its presentation early in the work (and frequently in its title) and by repetition, either verbatism or modified in accordance with the circumstances.* [57, 314]

Patinkin's reason for the importance of the central message is first the scientific reward system. In order for this system to be "fair" Patinkin finds that it is important that:

> *... its rewards must go to the true discoverers: to those who brought about a cognitive change.* [57, 316]

Second, in Patinkin's view the function of scientific discoveries is to

> *stimulate a new research program on the part of colleagues in his field of inquiry, for only in that way can the full scientific potential of the discovery be efficiently exploited.* [57, 316]

Using Patinkin's criteria for multiplicity the picture gets a little more subtle. Using his method for uncovering the central message of the scientist and taking John's introduction for face value it must be said that the Kuhn-Tucker theorem is indeed part of the central message in all three papers. The titles of both Karush's and John's paper indicate that the subject of their paper is optimization constrained by inequality conditions. The title of Kuhn and Tucker's paper is simply *Nonlinear Programming* but at the time linear programming was well known in the circles Kuhn and Tucker appeared in so in 1950 this word could not refer to anything but finite dimensional optimization subject to inequality constraints. So using only this criteria we must once again conclude that the Kuhn-Tucker theorem is a triple discovery.

This however is not very satisfactory and if one is also considering Patinkin's reasons for putting such a high empasis on the central message namely that the purpose of scientific discoveries is to stimulate further research in the field it becomes clear that only Kuhn and Tucker can be said to be the true discoverer of the Kuhn-Tucker theroem in nonlinear programming. Neither Karush's nor John's work stimulated any further research. Their work had no influence what so ever on the development of nonlinear programming or any other discipline, only the work of Kuhn and Tucker can fulfill this requirement.

This however does not shed light on why the three different versions of the result were so differently recieved in the scientific community. I think that Patinkin's second essential point – what is it exactly that has been discovered – analyzed with respect to the different contexts the three papers originated in is a more fruitful approach to understanding this phenomenon.

6 The Significance of the Context

In the following I shall distinguish between a mathematical and a sociological context. I shall make a further division of the mathematical context into what I call the context of "pure mathematical content" which refer to analysis of mathematical results without taking into account the context of discovery or the mathematical environment in which it is presented, and the context of mathematical subdisciplines such as the theory of convexity or the calculus of variations.

Today mathematicians concieve Karush's and Kuhn and Tucker's result as the same result – as the Kuhn-Tucker theroem and John's result as the Kuhn-Tucker theorem without the constraint qualification. The reason for this is an analysis of the results in relation to "pure mathematical content" that is an analysis based on the theoretical knowledge of today without taking into account the mathematical subdisciplines, that is the internal mathematical contexts the results were derived in. In such an analysis mathematicians disregard the differences and focus solely on the similarities between the three results. They look at the theorems independently of the context they were developed in.

An analysis which instead focuses on the *differences* in the three formulations of the theorem and take the internal mathematical context i.e. the context of the subdisciplines into account can provide an explanation for the different influences on the mathematical development and the different reception in the mathematical community at the time of the three occurences of the result.

The reason why the works of Karush and John were 'overlooked' was not because their result didn't form part of the central message of their work but rather because their work wasn't central in relation to the internal mathematical – and maybe also sociological – context it appeared in.

Karush's work was part of the calculus of variations. It was work on a finite dimensional version of the research problems the group at the mathematical department at University of Chicago was deeply involved in at the time, namely variational calculus with inequalities as side conditions. Interpreted in this subdiscipline Karush's work was just a minor corner, some 'cleaning up', it was not very exciting, it did not touch on the main questions that guided the research in this area. This explains why nobody at Chicago encouraged him to publish his results, they simply were not very interesting in this narrowly defined mathematical context of variational calculus as it was viewed in Chicago at that time.

John's work originated in a mathematical context of the theory of convexity. His real interest was not finite dimensional optimization with inequality contraints but rather to develop a tool useful for deriving some general theorems in the theory of convexity. The important results, the results he was looking for emerged from his geometrical applications of the theorem. It is also obvious that his hole formulation of the theorem and the 'missing' regularity condition, the 'contraint qualification', is due to the geometrical applications.

In contrast to this Kuhn and Tucker's work was a continuation of their work on linear programming. They derived their theorem in an attempt to extent the duality theorem for linear programming to the nonlinear case, which explain their reformulation of the problem in terms of saddle points. Their work took place within linear programming – a newly established research field – which it enlarged. The Kuhn-Tucker theorem answered a central question in this discipline and thereby stimulated further research in nonlinear programming. That is, their work was defined in an internal mathematical context which was just beginning and was rapidly developing into a major discipline. Regarded with this knowledge it is not surprising that Kuhn and Tucker's paper could launch a new research area.

6.1 The Significance of the military: ONR and OR

Until now I have almost exclusively considered the *mathematical* contexts of the different versions of the Kuhn-Tucker theorem. In the case of Kuhn and Tucker there is also a very important *social* context – namely the military represented by *Office of Naval Research* (ONR) that encouraged mathematicians to do research in linear programming and thereby played a major role in the development of nonlinear programming into a succesful and growing research area.[17]

In the section on the work of Kuhn and Tucker I mentioned that the coorporation of the two began in 1948. The direct reason for this was the establishment of a university based project with the purpose of exploring the relationship between linear programming and game theory and do research in the underlying mathematical structure of linear programming. This project was not only financed by ONR they also initiated it.

The background for the project was a model that George B. Dantzig had developed for a logistic problem within the Air Force.[18] The military became very enthusiastic when they learned about this model and it even caused ONR to set up a special logistic branch within its mathematics program.[19] Mina Rees who was the head of the Mathematics Division of ONR described it in [62, 111]:

> ... *when, in the late 1940's the staff of our office became aware that some mathematical results obtained by George Dantzig, who was then working for the Air Force, could be used by the Navy to reduce the burdensome costs of their logistics operations, the possibilities were pointed out to the Deputy Chief of Naval Operations for Logistics. His enthusiasm for the possibilities presented by these results was so great that he called together all those senior officers who had anything to do with logistics, as well as their civilian counterparts, to hear what we always referred to as a "presentation". The outcome of this meeting was the establishment in the Office of Naval Research of a separate Logistics Branch with a separate research program. This has proved to be a most successful activity of the Mathematics Division of ONR, both in its usefulness to the Navy, and in its impact on industry and the universities.*

ONR was established by the Navy in 1946. It grew out of the mobilization of scientists in the US during World War II. When the war was about to end there was a common concern that the scientists would just go back to their university duties after the war. There also was a strong belief that the US had be to strong scientifically in order to be strong military. A lot of people

[17]For historical accounts on ONR see f.ex [63], [64], [54].

[18]This was the beginning of linear programmering. For historical accounts on linear programmering see [12], [13], [14], [16].

[19]The simultaneous development of the computer also had a major influence in this hole development and financing of linear and nonlinear programming.

were concerned with the further financing of science, military related science as well as basic science. The first four years of its existence ONR was the head sponsor for government supported research in the US. It was organized after the same model Vannevar Bush created for the war effort of civilian scientists.[20] The scientists continued to work in universities and industries and their relationship with ONR were based on contracts. Every project had a principal investigator and the financial support from ONR covered salaries during the summers, salaries for research assistants working on the projects, conferences, invitations of guests etc.

In the spring of 1948 Dantzig went to Princeton on behalf of ONR to meet with John von Neumann in order to discuss the possibilities for a university based project financed by ONR on linear programming and its relations to game theory [2, 342-343]. During this visit Dantzig got introduced to Tucker who gave him a ride to the train station. During this short car trip Dantzig gave Tucker a brief introduction to the linear programming problem. Tucker made a remark about a possible connection to Kirchoff-Maxwell's law of electric network and because of this remark Tucker was contacted by the ONR a few days later and asked if he would set up such a mathematics project [2, 342-343].

Until this moment Tucker had been absorbed in research in topology. He agreed in becomming principal investigator and that completely changed his research direction. The same happened for Kuhn, who at the time was finishing up a Ph.D project on group theory. Kuhn went to Tucker to ask for summer employment in the summer of 1948, because he needed the additional income. Tucker hired him together with Gale, to work with him on the ONR project [46]. The three of them presented the results of their work on the project at the first conference on linear programming which took place in Chicago in June 1949 [31]. The most prominent among their results was the duality theorem for linear programming. After that Kuhn and Tucker in a way got "stuck" in the project, the duality theorem caught their interest – it was an interesting result from a mathematical point of view. From there on Kuhn and Tucker, proceeding according to the "inner" rules for research in pure mathematics, tried to extend this result to more general cases. This work resulted in the *nonlinear programming* paper and the Kuhn-Tucker theorem. This work was also sponsored by ONR who continued to support Tucker's project until 1972 where the National Science Foundation took over.

Another social factor also related to the military was the development of operations reserach (OR) during the war and the following establishment

[20]See [69].

after the war of OR as a scientific discipline at the universities.[21] ONR also played a major role in this proces. Fred Rigby the head of the logistic program of ONR later described the significance of ONR:

> *We did indeed influence the introduction of operations research into business schools. The subdiscipline called management science is our invention, in quite a real sense. That is, we and our contract researchers recognized its potentials, planned its early growth, and, as it turned out, set the dominant pattern in which it has developed* (quoted in [62, 111]).

Linear programming was imediately incoorporated in the toolbox of OR which meant that OR stood ready to provide a 'home' also for nonlinear programming as soon as it was developed. In this way it can be seen that the Office of Naval Research had an enormous influence in creating a scientific community of people doing linear programming, and in this community it was almost inevitable that the nonlinear programming paper of Kuhn and Tucker would give rise to the new research field of nonlinear programming.

7 Closing Remarks

The Kuhn-Tucker theorem exemplifices that it is not always the mathematical theorem in itself, that is its "pure mathematical content" that decide whether it will stimulate further research or not. Whether a theorem is going to be famous or not to launch new research areas or not is not independent of the mathematical and social contexts. Eventhough the three results today are viewed as the same theorem they were in praxis very different. The significance of a result and its potential for stimulating further research in its area is determined by the mathematical – and sometimes also the social – context it was developed in. The Kuhn-Tucker theorem was an important result in the mathematical discipline Kuhn and Tucker was working in but this was not at all the case in the discplines the papers of Karush and John appeared in.

The fact that both Karush, John and Kuhn and Tucker gets credit for the theorem in the scientific community of nonlinear programming is due to the influence of *third parties* – a notion introduced by S. Cozzens. In her book *Social Control and Multiple Discoveries in Science: The Opiate Receptor Case* she focused on how discoveries later gets established as multiple discoveries [11]. She points out that it is often due to an *after-the-fact process* where the

[21] For historical accounts on OR in the USA see f.ex [60], [24].

case is settled by influence from third parties, that is members of the scientific community who are not directly involved in the discovery. Through later references and acknowledgement the third parties establishes the discoveries as multiple. The quotes I showed in the section on Karush shows that this also was the case for the establishment of the Kuhn-Tucker theorem as a multiple discovery, even though Kuhn himself, that is one of the involved scientist, here played a major role in the establishment.

References

[1] O. B. Ader, An Affine Invariant of Convex Regions, *Duke Mathematics Journal*, **4** (1938), 291-299.

[2] J. D. Albers and G. L. Alexanderson, *Mathematical People, Profiles and Interviews*, Boston: Birkhäuser, 1985.

[3] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, New York: John Wiley & Sons, Inc., 1979.

[4] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms*, New York: John Wiley & Sons, Inc., 1993, (2. ed.).

[5] Felix Behrend, Über einige Affininvarianten konvexer Bereiche, *Mathematische Annalen*, **113** (1937), 713-747.

[6] Felix Behrend, Über die kleinste umbeschriebene und die grosste einbeschriebene Ellipse eines konvexen Bereichs, *Mathematische Annalen*, **115** (1938), 379-411.

[7] Richard Bellman, Letter to Harold W. Kuhn, dated 11. Februar, 1975. (Unpubliced, in possession of the author).

[8] Gilbert A. Bliss, Normality and Abnormality in the Calculus of Variations, *Transactions of the American Mathematical Society*, **43** (1938), 365-376.

[9] Sonja Brentjes, Bemerkungen zum Beitrag von Julius Farkas zur Theorie der linearen Optimierung, *NTM-Schriftenreihe zur geschicte der Naturwissenschaften, Technik und Medizin*, **13** (1976), 21-23.

[10] Felix E. Browder, The Stone Age of Mathematics on the Midway, in *A Century of Mathematics in America, part II*, ed. Peter Duren, American Mathematical Society, History of Mathematics, 2, 1989, pp. 191-193.

[11] Susan E. Cozzens, *Social Control and Multiple Discovery in Science: The Opiate Receptor Case*, Albany: State University of New York Press, 1989.

[12] George B. Dantzig, Reminiscences about the Origins of Linear Programming, *Operations Research Letters*, **1** (1982), 43-48.

[13] George B. Dantzig, Impact of Linear Programming on Computer Development, *OR/MS Today*, (1988), 12-17.

[14] George B. Dantzig, Linear Programming, in *History of Mathematical Programming, A Collection of Personal Reminiscences*, eds. Jan Karel Lenstra, Alexander H. G. Rinnooy Kan and Alexander Schrijver, Amsterdam: Noth-Holland, 1991, pp. 19-31.

[15] L. L. Dines, Convex Extension and Linear Inequalities, *Bulletin of the American Mathematical Society*, **42** (1936), 353-365.

[16] Robert Dorfman, The Discovery of Linear Programming, *Annals of the History of Computing*, **6** (1984), 283-295.

[17] A. L. Duren, Graduate Student at Chicago in the Twenties, *The American Mathematical Monthly*, **83** (1976), 243-248.

[18] Peter Duren, A Century of Mathematics in America, part II, *American Mathematical Society, History of Mathematics*, 2, 1989.

[19] Gyula (Julius) Farkas, Über die Anwendungen des mechanischen Princips von Fourier, *Mathematische und Naturwissenschaftliche Berichte aus Ungarn*, **12** (1895), 263-281.

[20] Gyula (Julius) Farkas, Die algebraischen Grundlagen der Anwendung des Fourier'sschen Principes in der mechanik, *Mathematische und Naturwissenschaftliche Berichte aus Ungarn*, **15** (1897), 25-40.

[21] Gyula (Julius) Farkas, Die algebraische Grundlage der Anwendungen des mechanischen Princips von Fourier, *Mathematische und Naturwissenschaftliche Berichte aus Ungarn*, **16** (1899), 154-157.

[22] Gyula (Julius) Farkas, Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik*, **124** (1901), 1-27.

[23] Werner Fenchel, *Convex Cones, Sets, and Functions*, Lecture Notes: Department of Mathematics, Princeton University, 1953.

[24] M. Fortun and Sam S. Schweber, Scientists and the Legacy of World War II: The Case of Operations Research (OR), *Social Studies of Science*, **23** (1993), 595-642.

[25] Jean B. J. Fourier, Mémoire sur la statique, contenant la démonstration du principe de vitesses virtuelles, et la théorie des momens, in *Oeuvres*, ed. Gaston Darboux, 2 vols, Académie des Sciences, Paris: Gauthier-Villars, 1890, 2: 475-521.

[26] Ole I. Franksen, Irreversibility by Inequality Constraints I: On Fourier's Inequality, *Syst. Anal. Model. Simul. 2*, **2** (1985), 137-149.

[27] Ole I. Franksen, Irreversibility by Inequality Constraints Part II: The Second Law of Thermodynamics, *Syst. Anal. Model. Simul. 2*, **3** (1985), 251-273.

[28] Ole I. Franksen, Irreversibility by Inequality Constraints Part III: Towards Mathematical Programming, *Syst. Anal. Model. Simul. 2*, **4** (1985), 337-359.

[29] Craig G. Fraser, Isoperimetric Problems in the Variational Calculus of Euler and Lagrange, *Historia Mathematica*, **19** (1992), 4-23.

[30] Craig G. Fraser, The Origin of Euler's Variational Calculus, *Archive for History of Exact Sciences*, **47** (1994), 103-141.

[31] David Gale, Harold W. Kuhn and Albert W. Tucker, Linear Programming and the Theory of Games, in *Activity Analysis of Production and Allocation*, ed. Tjalling C. Koopmans, Cowles Commission Monograph, 13, New York: Wiley, 1951, pp. 317-329.

[32] Ivor Grattan-Guinness, Joseph Fourier's Anticipation of Linear Programming, *Operational Research Quaterly*, **21** (1970), 361-364.

[33] Ivor Grattan-Guinness, "A New Type of Question" : On the Prehistory of Linear and Non-Linear Programming, 1770-1940, in *The History of Modern Mathematics. III*, eds. Eberhard Knobloch and David E. Rowe, Boston: Academic Press, 1994, pp. 43-89.

[34] Lars Gårding, 'Foreword', in *Fritz John, Collected Papers, Volume 1*, ed. J. Moser, Boston: Birkhäuser, 1985.

[35] Fritz John, Moments of Inertia of Convex Regions, *Duke Mathematics Journal*, **2** (1936), 447-452.

[36] Fritz John, An inequality for convex bodies, *University of Kentucky Research Club Bulletin*, **8** (1942), 8-11.

[37] Fritz John, Extremum Problems with Inequalities as Subsidiary Conditions, in *Studies and Essays, Presented to R. Courant on his 60th Birthday January 8, 1948*, New York: Interscience, 1948, pp. 187-204.

[38] William Karush, *Minima of Functions of Several Variables with Inequalities as Side Conditions*, Dissertation, Department of Mathematics, University of Chicago, 1939.

[39] William Karush, Letter to Harold W. Kuhn, dated 10. Februar, 1975. (Unpubliced, in possession of the author).

[40] Tinne H. Kjeldsen, *A Contextualised Mathematico-Historical Analysis of Nonlinear Programming: Development and Multiple Discovery*, (in Danish), IMFUFA, Text **372**, Roskilde University, 1999.

[41] Harold W. Kuhn, Subgroup Theorems for Groups Presented by Generators and Relations, *Annals of Mathematics*, **56** (1952), 22-46.

[42] Harold W. Kuhn, Letter to William Karush, dated 4. Februar, 1975. (Unpubliced, in possession of the author).

[43] Harold W. Kuhn, Letter to William Karush, dated 21. Februar, 1975. (Unpubliced, in possession of the author).

[44] Harold W. Kuhn, Nonlinear Programming: A Historical View, *SIAM-AMS Proceedings*, **9** (1976), 1-26.

[45] Harold W. Kuhn, *Curriculum Vitae*, 1986. (Unpubliced, in possession of the author).

[46] Harold W. Kuhn, Personal interview, Princeton University, Princeton, New Jersey, 23. april, 1998.

[47] Harold W. Kuhn, Tucker, A. W.: Some Reminiscence, Prepared with the assistance of Alan and Tom Tucker. Based on an oration given by Harold W. Kuhn. *Notices of the AMS*, October 1995.

[48] Harold W. Kuhn and Albert W. Tucker, Nonlinear Programming, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman, Berkeley, 1950, pp. 481-492.

[49] J. -L. Lagrange, *Mécanique analitique*, 1788; reprint ed., Paris: Albert Blanchard, 1965.

[50] David G. Luenberger, *Linear and Nonlinear Programming*, Reading, Massachusetts: Addison-Wesley publishing company, 1973.

[51] Saunders Mac Lane, Mathematics at the University of Chicago A Brief History, in *A Century of Mathematics in America, part II*, ed. Peter Duren, American Mathematical Society, History of Mathematics, 2, 1989, pp. 127-154.

[52] Robert K. Merton, Singletons and Multiples in Science, in *The Sociology of Science, Theoretical and Empirical Investigations*, ed. Robert K. Merton, Chicago: University of Chicago Press, 1973, pp. 343-382.

[53] J. Moser, *Fritz John, Collected Papers, Volume 1*, Boston: Birkhäuser, 1985.

[54] B. S. Old, The Evolution of the Office of Naval Research, *Physics Today*, **14** (1961), 30-35.

[55] Mikhail V. Ostrogradsky, Considérations génèrales sur les momens des forces, *Mémoires de l'Académie impériale des sciences de ST.-Pétersbourg*, sixieme série, **I** (1838), (read 1834), 129-150.

[56] Karen H. Parshall and David E. Rowe, *The Emergence of the American Mathematical Research Community, 1876-1900: J. J. Sylvester, Felix Klein, and E. H. Moore*, History of Mathematics, 8, American Mathematical Society, 1994.

[57] Don Patinkin, Multiple Discoveries and the Central Message, *American Journal of Sociology*, **89** (1983), 306-323.

[58] A. L. Peressini, F. E. Sullivan and J. J. Uhl, *The Mathematics of Nonlinear Programming*, New York: Springer-Verlag, 1988.

[59] A. Prékopa, On The Development of Optimization Theory, *American Mathematical Monthly*, **87** (1980), 527-542.

[60] Erik Rau, The Adoption of Operations Research in the United States during Worl War II, in *Systems, Experts, and Computers*, eds. A. C. Hughes and Tom P. Hughes, accepted for publication in Dibner Series, ed. Jed Z. Buchwald, MIT, Cambridge, Massachusetts.

[61] Constance Reid, *Courant in Göttingen and New York. The Story of an Improbable Mathematician*, New York: Springer-Verlag, 1976.

[62] Mina S. Rees, Mathematics and the Government: The Post-War Years as Augury of the Future, in *The Bicentennial Tribute to American Mathematics, 1776-1976*, ed. D. Tarwater, The American Association of America, Buffalo, NY, 1977, pp. 101-116.

[63] H. M. Sapolsky, Academic Science and the Military: The Years Since the Second World War, in *The Sciences in the American Context: New Perspectives*, ed. N. Reingold, Washington, D. C.: Smithsonian Institution Press, 1979, pp. 379-399.

[64] Sam S. Schweber, The Mutual Embrace of Science and the Military: ONR and the Growth of Physics in the United States after World War II, in *Science, Technology and the Military*, eds. E. Mendelsohn, M. R. Smith and P. Weingart, Dordrecht, The Netherlands: Kluwer Academic Publishers, 1988, pp. 3-45.

[65] R. W. Stokes, A Geometric Theory of Linear Inequalities, *Transactions of the American Mathematical Society*, **33** (1931), 782-805.

[66] Marshall H. Stone, Reminiscences of Mathematics at Chicago, in *A Century of Mathematics in America, part II*, ed. Peter Duren, American Mathematical Society, History of Mathematics, 2, 1989, pp. 183-190.

[67] A. Takyama, *Mathematical Economics*, Hinsdale, Illinois: The Dryden Press, 1974.

[68] Albert W. Tucker, *Curriculum Viate*, 1980. (Unpubliced, in possession of the author).

[69] P. G. Zachary, *Endless Frontier: Vannevar Bush, Engineer of the American Century*, New York: The Free Press, 1997.

352/98 The Herman-Swiatec Theorem with applications

by: Carsten Lunde Petersen

353/98 Problemløsning og modellering i en almendannende matematikundervisning

Specialerapport af: Per Gregersen og Tomas Højgaard Jensen

Vejleder: Morten Blomhøj

354/98 A GLOBAL RENEWABLE ENERGY SCENARIO

by: Bent Sørensen and Peter Meibom

355/98 Convergence of rational rays in parameter spaces

by: Carsten Lunde Petersen and Gustav Ryd

356/98 Terrænmodellering

Analyse af en matematisk model til konstruktion af terrænmodeller

Modelprojekt af: Thomas Frommelt, Hans Ravnkjær Larsen og Arnold Skimminge

Vejleder: Johnny Ottesen

357/98 Cayleys Problem

En historisk analyse af arbejdet med Cayley problem fra 1870 til 1918

Et matematisk videnskabsfagsprojekt af: Rikke Degn. Bo Jakobsen. Bjarke K.W. Hansen, Jesper S. Hansen. Jesper Udesen, Peter C. Wulff

Vejleder: Jesper Larsen

358/98 Modeling of Feedback Mechanisms which Control the Heart Function in a View to an Implementation in Cardiovascular Models

Ph.D. Thesis by: Michael Danielsen

359/99 Long-Term Scenarios for Global Energy Demand and Supply Four Global Greenhouse Mitigation Scenarios

by: Bent Sørensen

*360/99 **SYMMETRI I FYSIK***

En Meta-projektrapport af: Martin Niss. Bo Jakobsen & Tune Bjarke Bonné

Vejleder: Peder Voetmann Christiansen

361/99 Symplectic Functional Analysis and Spectral Invariants

by: Bernhelm Booss-Bavnbek. Kenro Furutani

362/99 Er matematik en naturvidenskab? - en udspænding af diskussionen

En videnskabsfagsprojekt-rapport af Martin Niss

Vejleder: Mogens Nørgaard Olesen

363/99 EMERGENCE AND DOWNWARD CAUSATION

by: Donald T. Campbell. MArk H. Bickhard and Peder V. Christiansen

364/99 Illustrationens kraft

Visuel formidling af fysik

Integreret speciale i fysik og kommunikation

af: Sebastian Horst

Vejledere: Karin Beyer. Søren Kjørup

365/99 To know - or not to know - mathematics. that is a question of context

by: Tine Wedege

366/99 LATEX FOR FORFATTERE

En introduktion til LATEX og IMFUFA-LATEX

af: Jørgen Larsen

367/99 Boundary Reduction of Spectral Invariants and Unique Continuation Property

by Bernhelm Booss-Bavnbek

368/99 Kvartalsrapport for projektet SCENARIER FOR SAMLET UDNYTTELSE AF BRINT SOM ENERGIBÆRER I DANMARKS FREMTIDIGE ENERGISTYSTEM

Projektleder: Bent Sorensen

369/99 DYNAMICS OF Complex Quadratic Correspondences

by: Jacob Jalving

370/99 OPGAVESAMLING

Bredde-Kursus i Fysik 1976 - 1999

(erstatter tekst nr. 350/98

371/99 Bevisets stilling

- beviser og bevisførelse i en gymnasial matematikundervisning

Matematikspeciale af: Maria Hermannsson

Vejleder: Mogens Niss

372/99 En kontekstualiseret matematikhistorisk analyse af ikke-lineær programmering: udviklingshistorie og multipel opdagelse

Ph.d.-afhandling af Tinne Hoff Kjeldsen

373/99 Criss-Cross Reduction of the Maslov Index and a Proof of the Yoshida-Nicolaescu Theorem

by: Bernhelm Booss-Bavnbek, Kenro Furutani and Nobukazu Otsuki

374/99 Det hydrauliske spring

Et eksperimentelt studie af polygoner og hastighedsprofiler

Specialeafhandling af Anders Marcussen

Vejledere: Tomas Bohr, Clive Ellegaard, Bent C. Jørgensen

Liste over tidligere udsendte tekster kan rekvireres på IMFUFA's sekretariat, tlf. 4674 2263 eller e-mail: bs@ruc.dk

332/97 ANOMAL SWELLING AF LIPIDE DOBBELTLAG
Specialerapport af:
Stine Sofia Korremann
Vejleder: Dorthe Posselt

333/97 Biodiversity Matters
an extension of methods found in the literature on monetisation of biodiversity
by: Bernd Kuemmel

334/97 LIFE-CYCLE ANALYSIS OF THE TOTAL DANISH ENERGY SYSTEM
by: Bernd Kuemmel and Bent Sørensen

335/97 Dynamics of Amorphous Solids and Viscous Liquids
by: Jeppe C. Dyre

336/97 PROBLEM-ORIENTATED GROUP PROJECT WORK AT ROSKILDE UNIVERSITY
by: Kathrine Legge

337/97 Verdensbankens globale befolkningsprognose
- et projekt om matematisk modellering
af: Jørn Chr. Bendtsen, Kurt Jensen, Per Pauli Petersen
Vejleder: Jørgen Larsen

338/97 Kvantisering af nanolederes elektriske ledningsevne
Første modul fysikprojekt
af: Søren Dam, Esben Danielsen, Martin Niss, Esben Friis Pedersen, Frederik Resen Steenstrup
Vejleder: Tage Christensen

339/97 Defining Discipline
by: Wolfgang Coy

340/97 Prime ends revisited - a geometric point of view -
by: Carsten Lunde Petersen

341/97 Two chapters on the teaching, learning and assessment of geometry
by Mogens Niss

342/97 LONG-TERM SCENARIOS FOR GLOBAL ENERGY DEMAND AND SUPPLY
A global clean fossil scenario discussion paper
prepared by Bernd Kuemmel
Project leader: Bent Sorensen

343/97 IMPORT/EKSPORT-POLITIK SOM REDSKAB TIL OPTIMERET UDNYTTELSE AF EL PRODUCERET PÅ VE-ANLÆG
af: Peter Meibom, Torben Svendsen, Bent Sorensen

344/97 Puzzles and Siegel disks
by Carsten Lunde Petersen

345/98 Modeling the Arterial System with Reference to an Anestesia Simulator
Ph.D. Thesis
by: Mette Sofie Olufsen

346/98 Klyngedannelse i en hulkatode-forstøvningsproces
af: Sebastian Horst
Vejledere: Jørn Borggren, NBI, Niels Boye Olsen

347/98 Verificering af Matematiske Modeller
- en analyse af Den Danske Eulerske Model
af: Jonas Blomqvist, Tom Pedersen, Karen Timmermann, Lisbet Øhlenschlæger
Vejleder: Bernhelm Booss-Bavnbek

348/98 Case study of the environmental permission procedure and the environmental impact assessment for power plants in Denmark
by: Stefan Krüger Nielsen
Project leader: Bent Sørensen

349/98 Tre rapporter fra FAGMAT - et projekt om tal og faglig matematik i arbejdsmarkedsuddannelserne
af: Lena Lindenskov og Tine Wedege

350/98 OPGAVESAMLING - Bredde-Kursus i Fysik 1976 - 1998
Erstatter teksterne 3/78, 261/93 og 322/96

351/98 Aspects of the Nature and State of Research in Mathematics Education
by: Mogens Niss