THE DETERMINISTIC STOCHASTIC TRANSITION IN CONTROL PROCESSES AND THE USE OF MAXIMUM AND INTEGRAL TRANSFORMS

P. WHITTLE
UNIVERSITY OF MANCHESTER

1. Introduction

Consider the "minimum transform" $\Phi(y)$ of a function $F(x)$ defined by

$$\Phi(y) = \min_x [F(x) - xy]. \tag{1.1}$$

Then, under certain conditions on F, the essential one being that of convexity, the inverse relation is simply

$$F(x) = \max_y [\Phi(y) + xy], \tag{1.2}$$

that is, F is the "maximum transform" of Φ. We shall refer to transforms of either type generically as "maximum transforms."

In this paper we shall show that use of the transform leads to a very natural treatment of certain control problems.

The pair of relations (1.1), (1.2) is strikingly analogous in form to a Fourier integral transformation and its inversion. The analogy is more than fortuitous. Consider a pair of functions F, Φ linked via the reciprocal Fourier relations

$$\begin{aligned} \exp\left[\frac{\Phi(y)}{c}\right] &= \text{const.} \int \exp\left[\frac{F(x) - xy}{c}\right] dx \\ \exp\left[\frac{F(x)}{c}\right] &= \text{const.} \int \exp\left[\frac{\Phi(y) + xy}{c}\right] dy, \end{aligned} \tag{1.3}$$

where the integration contours are taken appropriately, one of them at least leaving the real axis. We suppose that c is a small constant. As c tends towards zero it is plausible that the integrals in (1.3) can be asymptotically evaluated by steepest descents, so that the relations (1.3) between F and Φ will reduce to relations (1.1) and (1.2).

The significance of this transition will emerge from our heuristic examination of some rather special cases. We shall consider some control problems in which there is a random disturbance whose variance is measured by c. In some cases these can be solved by use of Fourier transforms. As c tends to zero and the situation reduces to a deterministic one, we shall find that the Fourier treatment reduces to the previous treatment by use of maximum transforms.

The situation is, of course, quite similar to the familiar transition between wave optics and geometrical optics, a stationary phase approximation in the first case leading to a minimum (or more generally a stationary) time path in the limit of infinitesimally small wavelengths. This seems to be one way in which "natural extremal principles" arise in the physical world. However, in the case of the control processes, we have an imposed extremal principle: the system shall function optimally subject to given physical limitations.

The idea of the maximum transform was described by Bellman and Karush ([1], p. 354). However, it has quite a history. Karlin treats the idea very thoroughly (under the name of "conjugate functions" [5], chapter 7), attributing the work largely to Fenchel [3]. However, in a work as classic as Courant and Hilbert's *Methods of Mathematical Physics* [2] the idea is quite well developed, though not fully so (this time under the name of "involutory transformations," p. 234) and reference is made to authors as early as Friedrichs [4] and Trefftz [6]. The transformation is, of course, also related to the classic Legendre transformation.

2. Solution of linear problems by the maximum transform

Consider a linear control problem in discrete time t, with state vector x_t and control vector u_t, so that

$$x_{t+1} = Ax_t + Bu_t, \qquad t \text{ integral.} \tag{2.1}$$

Typically, one might wish to choose u so as to minimize $\sum_{j=t}^{T-1} g_j(u_j) + f(x_T)$, that is, so as to minimize some integrated measure of control effort plus a measure of deviation in x from its desired value at the terminal instant $t = T$. If

$$F(x_t, T - t) = \min_{u_t \cdots u_{T-1}} \left[\sum_{j=t}^{T-1} g_j(u_j) + f(x_T) \right], \tag{2.2}$$

then we have by the dynamic programming principle

$$\begin{aligned} &F(x, 0) = f(x) \\ &F(x, s) = \min_u \left[g_{T-s}(u) + F(Ax + Bu, s - 1) \right], \quad s = 1, 2, \cdots. \end{aligned} \tag{2.3}$$

Define now

$$\begin{aligned} \Phi(y, s) &= \min_x \left[F(x, s) - y'x \right], \\ \phi(y) &= \min_x \left[f(x) - y'x \right], \\ \psi(z, t) &= \min_u \left[g_t(u) - z'u \right]. \end{aligned} \tag{2.4}$$

One finds then readily from relations (2.3), (2.4) that

$$\Phi(A'y, s) = \Phi(y, s - 1) + \psi(-B'y, T - s). \tag{2.5}$$

The first transformation in (2.4) has eliminated the minimizing operation in relation (2.3) and left us with a purely linear relation, (2.5). If $C = (A')^{-1}$ exists, then (2.5) can be solved immediately,

$$\Phi(y, s) = \sum_{j=0}^{s-1} \psi(-B'C^{j+1}y, T - s + j) + \varphi(C^s y), \tag{2.6}$$

with, under appropriate conditions, an inversion

$$F(x, s) = \max_y [\Phi(y, s) + y's]. \tag{2.7}$$

This solution can be adapted also to some cases in which the values of x and u are constrained (to positive values, for example).

Suppose that the stopping rule, instead of being $t = T$, is a time independent one, $x \in D$. If the loss functions g are also time independent, then one can expect the same of F, so that instead of (2.3), one has

$$F(x) \text{ prescribed by } f(x), \qquad (x \in D), \tag{2.8}$$

$$F(x) = \min_u [g(u) + F(Ax + Bu)], \qquad (x \in \bar{D}). \tag{2.9}$$

If now

$$\Phi(y) = \min_{x \in \bar{D}} [F(x) - y'x], \tag{2.10}$$

then on a "negligible overshoot" assumption, we derive an approximate relation from (2.9)

$$\Phi(y) \sim \Phi(Cy) + \psi(-B'Cy). \tag{2.11}$$

By "negligible overshoot," we mean that $x_t \in \bar{D}$ implies that $x_{t+1} = Ax_t + Bu_t$ lies in or only just out of $\bar{D}$. Relation (2.11) does not in itself determine Φ or F, and must be supplemented by the boundary condition (2.8).

All these relations have obvious continuous time analogues. If

$$\dot{x} \equiv \frac{dx}{dt} = \alpha x + \beta u, \tag{2.12}$$

$$F(x_t, T - t) = \min_{\substack{u_\tau \\ t \leq \tau \leq T}} \left[\int_t^T g_\tau(u_\tau)\, d\tau + f(x_T)\right], \tag{2.13}$$

then we find that corresponding to (2.5) and (2.11) we have the relations

$$\frac{\partial \Phi(y, s)}{\partial s} + y'\alpha \frac{\partial \Phi(y, s)}{\partial y} = \psi(-\beta' y), \tag{2.14}$$

$$y'\alpha \frac{\partial \Phi(y)}{\partial y} = \psi(-\beta' y), \tag{2.15}$$

where $\partial\Phi/\partial y$ is the column vector of first order derivatives. We have written (2.15), the analogue of (2.11), as an equality rather than as an approximation because there will presumably be no overshoot into D in case (2.12). The point requires proof, of course, and raises issues such as the continuity of F at the boundary of D, the stopping region.

As is usual in dynamic programming problems, we have put all the emphasis on the determination of the loss function F. This is a sufficient preliminary, and generally seems to be a necessary one, to the immediate practical problem of determining the policy or control function $u(x)$.

3. Some examples

Suppose that the state variable x is a scalar for which

$$\dot{x} = \alpha x + u, \tag{3.1}$$

and that we wish to minimize $\int (\lambda + \mu u^2)\, dt$ under a stopping rule $x = x_0$. That is, we wish to choose u in (3.1) so as to bring x to the value x_0 in such a way as to minimize the integral of $\lambda + \mu u^2$—a criterion that gives weight both to time taken and the amount of "control energy" employed.

If the future loss at (x, t) is $F(x, t)$, then the dynamic programming equation is

$$F_t + \min_u \, [\lambda + \mu u^2 + (\alpha x + u)F_x] = 0, \tag{3.2}$$

with boundary condition

$$F(x_0, t) = 0. \tag{3.3}$$

The subscripts in (3.2) denote partial derivatives. Now, in fact, F will be independent of t, so that (3.2) reduces to

$$\lambda + \alpha x F_x - \frac{1}{4\mu} F_x^2 = 0, \tag{3.4}$$

$$u = -\frac{1}{2\mu} F_x. \tag{3.5}$$

Equation (3.4) can in fact be solved subject to condition (3.3); we find

$$F(x) = H(x) - H(x_0), \tag{3.6}$$

where

$$H(x) = \frac{\lambda}{2} \{X^2 + \operatorname{sgn}(X - X_0)[X(X^2 + 1)^{1/2} + \sinh^{-1} X]\}, \tag{3.7}$$

$$X = x\alpha(\mu/\lambda)^{1/2}. \tag{3.8}$$

However, in a situation with more variables the equation analogous to the nonlinear relation (3.4) would have been quite intractable. Even in the present case there is an indeterminacy in the solution of (3.4) (leading to the sgn $(X - X_0)$ term in the solution) which must be resolved by special arguments.

Consider now the use of the minimum transform of F. This obeys an equation (see (2.15))

$$\alpha y \Phi'(y) = \lambda - \frac{1}{4\mu} y^2 \tag{3.9}$$

of a simple linear form, which can be integrated immediately to give

$$\Phi(y) = \text{const.} + \frac{\lambda}{\alpha}\log y - \frac{y^2}{8\alpha\mu}. \tag{3.10}$$

Applying the inversion (1.2) to (3.9), we recover solution (3.6). In this case $F(x)$ can be evaluated explicitly by either method, but in general it is probably true that the most explicit solution for F would be just the representation (1.2), with Φ a solution of the linear equation (2.15).

A rather more interesting problem is the following one in which control is restricted by imposing a limit rather than a cost on "fuel."

Suppose that the state variables are scalars x, w with

$$\dot{x} = u, \tag{3.11}$$

$$\dot{w} = -u^2, \tag{3.12}$$

and that the only cost is a terminal cost at $t = T$ of $\nu x^2 - \theta w$. The control u is to be chosen to minimize this terminal cost subject to the restriction $w \geqq 0$ for $t \leqq T$. That is, w has the interpretation of a "fuel reserve," and one wishes to achieve $x(T) = 0$ as closely as possible, but residual fuel at $t = T$ still has a value θ. The case is rather more realistic when, instead of (3.11), we have

$$\ddot{x} = u, \tag{3.13}$$

and it can be reduced to the present one by invariance arguments.

Let $s = T - t$ denote "time to go" and let the future loss at x, w, s be $F(x, w, s)$. We have then

$$F(x, w, 0) = \nu x^2 - \theta w, \tag{3.14}$$

$$F_s = \min_u \,[uF_x - u^2 F_w], \qquad s > 0, \tag{3.15}$$

so that

$$F_s = 0, \qquad w = 0,\ s > 0, \tag{3.16}$$

$$F_s = \frac{F_x^2}{4F_w}, \qquad w > 0,\ s > 0. \tag{3.17}$$

From (3.16) we find that

$$F(x, 0, s) = \nu x^2. \tag{3.18}$$

The nonlinear equation (3.17) can in fact be solved subject to boundary conditions (3.14) and (3.18). However, use of the minimum transform again gives a much more expeditious treatment. Define

$$\Phi(y, z, s) = \min_{\substack{x \\ w \geqq 0}} \,[F(x, w, s) - sy + wz]. \tag{3.19}$$

We find, by the same derivation as that of equation (2.15), that

$$\Phi_s = M(y, z), \tag{3.20}$$

where

$$M = \begin{cases} 0 & \text{if the minimizing } w \text{ value in (3.19) is zero,} \\ \dfrac{-y^2}{4z} & \text{if the minimizing } w \text{ value in (3.19) is positive.} \end{cases} \tag{3.21}$$

The y, z values corresponding to $w = 0$ as minimizing value in (3.19) are those maximizing $\Phi + xy$ for some s, that is, those for which $\Phi_z = 0$. We thus have either

$$\Phi_s = 0, \qquad \Phi_z = 0, \tag{3.22}$$

or

$$\Phi_s + \frac{y^2}{4z} = 0, \qquad \Phi_z > 0. \tag{3.23}$$

We have further from (3.14) and (3.19),

$$\Phi(y, z, 0) = \begin{cases} \dfrac{-y^2}{4\nu}, & z \geqq \theta, \\ -\infty, & z < \theta. \end{cases} \tag{3.24}$$

For minimal Φ we choose the option (3.23) rather than (3.21) for $s > 0$, and so in general

$$\Phi = \begin{cases} \dfrac{-y^2}{4}\left(\dfrac{1}{\nu} + \dfrac{s}{z}\right), & z \geqq \theta' \\ -\infty, & z < \theta. \end{cases} \tag{3.25}$$

The fact that $\Phi = -\infty$ for $z < \theta$ implies that the marginal utility of fuel is never less than 0, as one would expect.

We shall have then

$$\begin{aligned} F(x, w, s) &= \max_{y,z} \left[\Phi + xy - wz\right] \\ &= \max_{\substack{y \\ z \geqq \theta}} \left[\frac{-y^2}{4}\left(\frac{1}{\nu} + \frac{s}{z}\right) + xy - wz\right]. \end{aligned} \tag{3.26}$$

The maximizing z value must certainly be positive, and so given by

$$z = \frac{|y|}{2}\left(\frac{s}{w}\right)^{1/2}, \tag{3.27}$$

if this quantity exceeds θ. Thus,

$$F = \max_y \left[-\frac{y^2}{4\nu} - |y|(ws)^{1/2} + xy\right], \tag{3.28}$$

if the maximizing y value in (3.28) makes expression (3.27) exceed θ; otherwise, $z = \theta$ and

$$F = \max_y \left[-\frac{y^2}{4\nu} - \frac{y^2 s}{4\theta} + xy - w\theta\right]. \tag{3.29}$$

Now expression (3.28) has its maximum value at

$$y = \begin{cases} 2\nu[x - (ws)^{1/2}] & \text{if} \quad x \geqq (ws)^{1/2}, \\ 0 & \text{if} \quad |x| \leqq (ws)^{1/2}, \\ 2\nu[x + (ws)^{1/2}] & \text{if} \quad x \geqq -(ws)^{1/2}, \end{cases} \tag{3.30}$$

so that the expression is equal to $\nu[x - \operatorname{sgn}(x)(ws)^{1/2}]^2$ if $|x| > (ws)^{1/2}$, zero otherwise. However, the condition $|y| > 2\theta(s/w)^{1/2}$ implies that one of the solutions (3.30) can be accepted only if

$$\nu|x - \operatorname{sgn}(x)(ws)^{1/2}|(s/w)^{1/2} \geqq \theta \tag{3.31}$$

or

$$|x| \geqq (ws)^{1/2} + \frac{\theta}{\nu}(w/s)^{1/2}; \tag{3.32}$$

otherwise, F is determined by (3.29). We thus obtain the final solution

$$F = \begin{cases} \nu[x - \operatorname{sgn}(x)(ws)^{1/2}]^2 & \text{if} \quad |x| \geqq (ws)^{1/2} + (\theta/\nu)(w/s)^{1/2}, \\ \dfrac{\nu x^2}{1 + \nu s/\theta} - \theta w & \text{otherwise.} \end{cases} \tag{3.33}$$

These two very different functions can both be verified to be solutions of (3.17), but the approach taken here is both simpler and more automatic than direct argument from (3.17). The two regions are respectively those of "fuel shortage" and "fuel plenty"; we find that

$$u = \frac{F_x}{2F_w} = \begin{cases} -\frac{1}{2}\operatorname{sgn}(x)(w/s)^{1/2}, \\ \dfrac{\nu x}{\theta + \nu s}, \end{cases} \tag{3.34}$$

in the two cases, respectively.

A point to be noted is that, whereas the loss function F takes different analytic forms in different regions, its transform Φ does not. It is in the inversion from Φ that one discovers the existence of qualitatively different regimes.

4. A stochastic control process

Consider again the process associated with equation (3.1), but let this now be modified to

$$\dot{x} = \alpha x + u + \epsilon, \tag{4.1}$$

where ϵ is a "white noise" process with zero mean, whose integral per unit time has variance c. Such a process is a perfectly proper one when understood in the context of generalized processes.

Relations (3.2) and (3.4) now become modified to

$$F_t + \min_u [\lambda + \mu u^2 + (\alpha x + u)F_x] + \frac{c}{2} F_{xx} = 0, \tag{4.2}$$

$$\lambda + \alpha x F_x - \frac{1}{4\mu} F_x^2 + \frac{c}{2} F_{xx} = 0. \tag{4.3}$$

If we take

$$J(x) = \exp\left[-\frac{F(x)}{2\mu c}\right] \tag{4.4}$$

as independent variable instead of F, then equation (4.3) takes the linear form

$$c^2 J_{xx} + 2\alpha c x J_x - \frac{\lambda}{\mu} J = 0. \tag{4.5}$$

This is just the Hermite equation, whose general solution is a linear combination of two integrals of the form

$$\int \frac{1}{y} \exp\left\{-\frac{1}{2\mu c}[\Phi(y) + xy]\right\} dy, \tag{4.6}$$

where the contour of integration follows the imaginary y-axis, with an indentation either to the right or the left of the origin, and $\Phi(y)$ is the function defined by equation (3.10). These two integrals are to be combined in such a way that $F(x)$ is zero and minimal at $x = x_0$. Now, as c tends to zero (and the stochastic relation (4.1) degenerates to the deterministic one (3.1), it is plausible that the integrals (4.6) can be evaluated by steepest descents; and we see then that the integral relation between F and Φ determined by relations (4.4) and (4.6) will reduce to the maximum transformation (1.2). That is, the maximum transformation (and its inverse) arises as a limiting case of a Fourier transformation and its inverse, corresponding to a "minimum phase" evaluation in a transition from wave to geometric optics.

However, the picture is not always this simple; for example, it does not seem possible to linearize the stochastic version of equation (3.22),

$$F_s = \frac{F_x^2}{4F_w} + \frac{c}{2} F_{xx}, \tag{4.7}$$

by a simple transformation F. In such cases the stochastic problem is presumably not soluble in terms of linear transforms, and the maximum transform, successful in the deterministic limit, must be the limit version of some more general transformation.

5. The replacement of minimization by a functional integration

Suppose one has the problem of calculating a loss function

$$F(x, t) = \min_{x(\tau)} \int_t^T L[\dot{x}(\tau), x(\tau), \tau]\, d\tau, \tag{5.1}$$

where $x = x(t)$, and the integrated loss is minimized with respect to $x(\tau)$, for $t < \tau \leqq T$. One might replace the analytically awkward minimization by a functional integration; that is, suppose that F was instead given by

(5.2) $$\exp(-F/\gamma) = \text{const.} \int \cdots \int dx(\tau) \exp\left(-\int_t^T L\, d\tau/\gamma\right).$$

Here by the right member we mean the functional integral

(5.3) $$\lim_{N\to\infty} \int dx_1 \int dx_2 \cdots \int dx_N \exp\left[-\frac{\delta}{\gamma}\sum_{j=1}^{N} L\left(\frac{x_j - x_{j-1}}{\delta}, x_j, t + j\delta\right)\right],$$

where $\delta = (T - t)/N$ and $x_0 = x$. Expression (5.3) may well involve terms in δ which lead to singularities as $\delta \to 0$. These terms are absorbed in the "constant" in (5.2); the point is that even if expression (5.3) does not tend to a positive limit as $\delta \to 0$, the quotient of two such expressions for different x may well do so, and this is all that is needed.

As γ tends to zero in (5.2), one will expect that the value of F given by relation (5.2) will become identical with that given by (5.1).

Suppose one is dealing with a stochastic dynamic programming problem, so that integral (5.1) must be averaged in some way. In the modification (5.2) we can combine this averaging with the functional integration, and so have the great advantage that both our operations, integration and averaging, are linear.

The validity of the replacement of evaluation (5.1) by (5.2) (or its limit as $\gamma \to 0$) is not clear, still less are the real life implications of replacing an averaging of $\int L\, d\tau$ by that of $\exp(-\gamma^{-1}\int L\, d\tau)$. Nevertheless, we shall find that an application of this technique to the two examples we have discussed leads to explicit and suggestive results, and this may justify some investigation of the method.

Consider the problem of section 4, the minimization of $\int (\lambda + \mu u^2)\, d\tau$ subject to (2.1) and the stopping rule $x = x_0$. If

(5.4) $$J(x) = e^{-F(x)/\gamma},$$

we have

(5.5) $$J(x) = \text{const.}\ E \int \cdots \int du(\tau) \exp\left[-\frac{1}{\gamma}\int (\lambda + \mu u^2)\, d\tau\right],$$

so that, to within $o(\delta)$ terms,

(5.6) $$J(x) = \frac{1}{\pi(2c\gamma)^{1/2}} \int du \int dx' J(x') \exp\left\{-\frac{1}{2c\delta}[x' - x - \delta(\alpha x + u)]^2 - \frac{\delta}{\gamma}(\lambda + \mu u^2)\right\}.$$

Taking Fourier transforms

(5.7) $$K(\xi) = \int e^{ix\xi} J(x)\, dx,$$

we find that (5.6) becomes

(5.8) $$K(\xi) = K(\xi - \alpha\delta\xi) \exp\left[-\frac{\lambda\delta}{\gamma} - \delta\xi^2\left(\frac{c}{2} + \frac{\gamma}{4\mu}\right)\right] + o(\delta),$$

or

(5.9)
$$\alpha\xi \frac{K'(\xi)}{K(\xi)} = -\frac{\lambda}{\gamma} - \left(\frac{c}{2} + \frac{\gamma}{4\mu}\right)\xi^2,$$

and

(5.10)
$$\log K(\xi) = \text{const.} - \frac{\lambda}{\gamma}\log \xi - \left(\frac{c}{4} + \frac{\gamma}{8\mu}\right)\xi^2.$$

From (5.7), (5.10),

(5.11)
$$\begin{aligned} J(x) = e^{-F(x)/\gamma} &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} e^{-ix\xi}K(\xi)\, d\xi \\ &= \text{const.} \int \exp\{-\gamma^{-1}[\psi(y) + xy]\}\, dy/y, \end{aligned}$$

where $y = i\gamma\xi$ and

(5.12)
$$\Psi(y) = \frac{\lambda}{\alpha}\log y - \left(\frac{1}{\mu} + \frac{2c}{\gamma}\right)\frac{y^2}{8\alpha}.$$

Note the close resemblance between $\Psi(y)$ and the minimum transform $\Phi(y)$ of (3.10), also between (5.11) and (4.6). In the deterministic case $c = 0$, we can go to the limit $\gamma \to 0$ and obtain the correct evaluation of $F(x)$ as the maximum transform of $\Psi(y)$. The apparent effect of a stochastic element ϵ is to decrease μ to $\mu\gamma/(\gamma + 2\mu c)$; it is plain that the value of γ must be related to that of c if sensible results are to be obtained.

Consider the stochastic version of the second problem associated with equations (3.11) and (3.12)

(5.13)
$$\begin{aligned} \dot{x} &= u + \epsilon, \\ \dot{w} &= -u^2, \end{aligned}$$

the only cost being the terminal one $\nu x^2 - \theta w$ at $t = T$, or $s \equiv (T - t) = 0$. Defining

(5.14)
$$J(x, w, s) = e^{-F(x,w,s)/\gamma},$$

(5.15)
$$K(\xi, \eta, s) = \int_{-\infty}^{\infty} dx \int_0^{\infty} dw\, e^{ix\xi - \eta w} J(x, w, s),$$

we find, using the same functional integration technique as before, that

(5.16)
$$K(\xi, \eta, s) = \text{const.}\ (\eta - \theta/\gamma)^{-1}\left\{\exp - \frac{\xi^2}{2}\left[\frac{\gamma}{2\nu} + \left(c + \frac{1}{2\eta}\right)s\right]\right\},$$

so that

(5.17)
$$\begin{aligned} J = e^{-F/\gamma} &= \text{const.} \iint e^{-ix\xi + \eta w} K\, d\xi\, d\eta \\ &= \text{const.} \iint (z - \theta)^{-1} \exp\left\{-\frac{1}{\gamma}\left[-\frac{y^2}{4}\left(\frac{1}{\gamma} + \frac{s}{z} + \frac{2cs}{\gamma}\right)\right.\right. \\ &\qquad\qquad \left.\left. + xy - wz\right]\right\} dz\, dw, \end{aligned}$$

where $y = i\gamma\xi$, $z = \gamma\eta$.

The square bracketed expression in the exponent is just the bracketed expression in relation (3.26), at least in the deterministic case $c = 0$. The fact that maximization is restricted to $z \geqq \theta$ in (3.26) now corresponds to the fact that in the asymptotic evaluation ($\gamma \to 0$) of integral (5.17) one must consider not only the saddle points of the exponent, but also the pole at $z = \theta$.

REFERENCES

[1] R. Bellman and S. Dreyfus, *Applied Dynamic Programming*, Princeton, Princeton University Press, 1962.

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, New York, Interscience, 1953 (English edition).

[3] W. Fenchel, *Convex Cones, Sets and Functions*, lecture notes Department of Mathematics, Princeton University, 1953.

[4] K. O. Friedrichs, "Ein Verfahren der Variationsrechnung," *Nachr. Akad. Wiss. Göttingen, Math.-Phys. K1. II*, No. 1 (1929), pp. 13–20.

[5] S. Karlin, *Mathematical Methods and Theory in Games, Programming, and Economics*, Vol. 1, London, Pergamon, 1959.

[6] E. Trefftz, "Ein Gegenstück zum Ritzschen Verfahren," *Verhandlungen des zweiten Internationalen Kongresses für Technische Mechanik*, Zürich, 1927, p. 131.