

EUR 411.e

EUROPEAN ATOMIC ENERGY COMMUNITY - EURATOM

THE GENERALIZED GRADIENT,
ITS COMPUTATION ASPECTS AND ITS RELATIONS
TO THE MAXIMUM PRINCIPLE

by

W. DE BACKER

1963



Joint Nuclear Research Center
Ispra Establishment - Italy

Scientific Data Processing Center - CETIS

Text presented at the IFAC-Congress, Basel (Switzerland)
27 August - 4 September 1963

LEGAL NOTICE

This document was prepared under the sponsorship of the Commission of the European Atomic Energy Community (EURATOM).

Neither the EURATOM Commission, its contractors nor any person acting on their behalf :

- 1° — Make any warranty or representation, express or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this document, or that the use of any information, apparatus, method, or process disclosed in this document may not infringe privately owned rights ; or
- 2° — Assume any liability with respect to the use of, or for damages resulting from the use of any information, apparatus, method or process disclosed in this document.

This report can be obtained, at the price of Belgian Francs 40, from : PRESSES ACADEMIQUES EUROPEENNES — 98, Chaussée de Charleroi, Brussels 6.

Please remit payments :

- to BANQUE DE LA SOCIETE GENERALE (Agence Ma Campagne) — Brussels — account No 964.558,
- to BELGIAN AMERICAN BANK AND TRUST COMPANY — New York — account No 121.86,
- to LLOYDS BANK (Foreign) Ltd. — 10 Moorgate — London E.C.2,

giving the reference : « EUR 411.e — The generalized gradient, its computation aspects and its relations to the maximum principle ».

This document was duplicated on the basis of the best available copy.

EUR 411.e

EUROPEAN ATOMIC ENERGY COMMUNITY - EURATOM

THE GENERALIZED GRADIENT,
ITS COMPUTATION ASPECTS AND ITS RELATIONS
TO THE MAXIMUM PRINCIPLE

by

W. DE BACKER

1963



Joint Nuclear Research Center
Ispra Establishment - Italy

Scientific Data Processing Center - CETIS

Text presented at the IFAC-Congress, Basel (Switzerland)
27 August - 4 September 1963

CONTENTS

1. INTRODUCTION
 2. STATEMENT OF THE PROBLEM
 3. THE GENERALIZED GRADIENT AS A SPECIAL CASE OF THE MAXIMUM PRINCIPLE
 - 3.1 The optimal trajectory completely lies in the interior of G
 - 3.2 The trajectory partly lies in the interior of G and partly on the boundary of G .
 - 3.3 The transversality conditions
 - 3.4 Conclusions
 4. THE STEEPEST DESCENT AS A SPECIAL CASE OF THE GENERALIZED GRADIENT
 5. COMPUTATIONAL ASPECTS OF THE GENERALIZED GRADIENT
 6. EXAMPLE OF A GENERALIZED GRADIENT APPLICATION
 7. APPLICATION FIELDS
- ACKNOWLEDGEMENTS
- REFERENCES

THE GENERALIZED GRADIENT, ITS COMPUTATIONAL ASPECTS
AND ITS RELATIONS TO THE MAXIMUM PRINCIPLE

SUMMARY

The maximum principle and the related theorems of the Pontryagin team lead to a generalization of the well-known steepest descent possible.

This generalization comprises the introduction of dynamic stresses.

For this reason the state vector no longer moves along the steepest descent, but the ψ -vector of the adjoint system does.

With respect to the simulation of the dynamic stresses and the elegant technique of implicit computing of Lagrange multipliers, the analogue computer appears to be appropriate for generalized gradient optimization.

1. INTRODUCTION

We consider the problem of minimizing a given function $F(x^1, \dots, x^i, \dots, x^n)$ where the variables x^i [†]) are subject to constraints of the form:

$$g^j(x^1, \dots, x^i, \dots, x^n) \leq 0 \quad j = 1, \dots, m$$

One possible technique for this kind of problem is based upon the construction of a set of differential equations of the steepest descent type for the variables x^i . Pyne [2] has done this for linear programming problems on the analog computer, but his method is equally valid for nonlinear objective functions and constraints. Representing the whole set of variables x^i by the vector $x = (x^1, \dots, x^i, \dots, x^n)$ the steepest descent equations can be written as

$$\frac{dx}{dt} = -k \left[\text{grad } F(x) + \sum_{j=1}^m k_j g^j(x) \text{ grad } g^j(x) \right] \quad (1)$$

[†])Upper indices, which obviously should not be considered as exponents, will be used throughout the text.

$$\begin{aligned} \text{with: } k_j &= 0 && \text{if } g^j(x) \leq 0 \\ k_j &= \text{large and positive} && \text{if } g^j(x) > 0 \end{aligned} \quad (2)$$

An additional aspect of these equations is the implicit computation of the functions $k_j g^j(x)$ as approximations of the Lagrange multipliers for $g^j(x) = 0$. This is related to the theory of the penalty functions studied by Courant and Moser [3]. Indeed, relation (1) could have been written as

$$\frac{dx}{dt} = -k \text{ grad } P(x) \quad (3)$$

where $P(x)$ is a penalty function,

$$P(x) = F(x) + \frac{1}{2} \sum_{j=1}^m k_j [g^j(x)]^2 \quad (4)$$

which has to be minimized without additional constraints. The approximation is the better the more k_j tends to infinity for $g^j(x) > 0$.

It is obvious that the set of equations (1) and (2) or (3), (4) and (2) is not the only possible one describing trajectories of $x(t)$ ending at (at least a local) minimum of $F(x)$. This means that additional constraints and even additional criteria with respect to the optimal trajectory could be satisfied before it is uniquely determined. These latent degrees of freedom make some generalizations possible. The maximum principle and other theorems of the Pontryagin team [1] will permit us to analyse a related but more general problem and to define the concept of generalized gradient.

2. STATEMENT OF THE PROBLEM

$F(x^1, \dots, x^i, \dots, x^n)$ is a given function which has to be guided to its (local) minimum at some time τ , taking account of following system of constraints.

We consider $x(t)$ to be a state vector $x = (x^1, \dots, x^i, \dots, x^n)$, belonging to a closed subset G of the n -dimensional state space and whose evolution is described by a system of ordinary differential equations

$$\frac{dx(t)}{dt} = f[x(t), u(t)] \quad (5)$$

with $f = (f^1, \dots, f^i, \dots, f^n)$. In this system the control vector $u = (u^1, \dots, u^r)$ has to be chosen as an element of the closed subset U of the r -dimensional control space. The subset G is defined by

$$g^j(x) \leq 0 \quad j = 1, \dots, m \quad (6)$$

and the subset U is defined by

$$q^l(u) \leq 0 \quad l = 1, \dots, s \quad (7)$$

We shall show that minimizing $F(x)$ by generalized gradient corresponds to minimizing the functional

$$J(x) = F[x(t+T)] - F[x(t)] = \int_t^{t+T} f^0(x, u) dt' \quad (8)$$

with

$$f^0(x, u) = \sum_{\alpha=1}^n \frac{\partial F(x)}{\partial x^\alpha} f^\alpha(x, u) \quad (9)$$

The initial condition x_0 of the optimal trajectory reduces to one point. At every instant $x_0 = x(t)$. The final condition $x_1 = x(t+T)$ belongs to the set of all possible x , reachable in a time T from t , with admissible $u \in U$. We call this set $R_T(t)$ and its boundary $\rho[x(t+T)] = 0$.

3. THE GENERALIZED GRADIENT AS A SPECIAL CASE OF THE MAXIMUM PRINCIPLE

3.1. The optimal trajectory completely lies in the interior of G

Requiring that $u(t)$ is piecewise continuous and that all f^i and all $\partial F / \partial x^i$ are defined and continuous on $G \times U$ together with their partial derivatives, we can apply the maximum principle for the problem defined in § 2. By definition we have

$$\mathcal{H} = \sum_{i=0}^n \psi_i f^i = \sum_{i=1}^n (\psi_i + \psi_0 \frac{\partial F}{\partial x^i}) f^i \quad (10)$$

The vector function $\psi = (\psi_0, \psi_1, \dots, \psi_n)$ of the adjoint system is given by the Hamiltonian system

$$\frac{d\psi_i}{dt} = -\frac{\partial \mathcal{H}}{\partial x^i} = -\psi_0 \frac{d}{dt} \left(\frac{\partial F}{\partial x^i} \right) - \sum_{\alpha=1}^n (\psi_\alpha + \psi_0 \frac{\partial F}{\partial x^\alpha}) \frac{\partial f^\alpha}{\partial x^i}$$

Knowing that ψ_0 is negative and constant this relation can be reduced to

$$\frac{d}{dt} (\psi_i + \psi_0 \frac{\partial F}{\partial x^i}) = - \sum_{\alpha=1}^n (\psi_\alpha + \psi_0 \frac{\partial F}{\partial x^\alpha}) \frac{\partial f^\alpha}{\partial x^i}$$

Calling $\psi_i + \psi_0 \frac{\partial F}{\partial x^i} = \psi_i'$ we can write the Hamiltonian system as

follows

$$\mathcal{H} = \sum_{i=1}^n \psi_i' f^i$$

$$\frac{dx^i}{dt} = \frac{\partial \mathcal{H}}{\partial \psi_i'} = f^i$$

$$\frac{d\psi_i'}{dt} = -\frac{\partial \mathcal{H}}{\partial x^i} = - \sum_{\alpha=1}^n \psi_\alpha' \frac{\partial f^\alpha}{\partial x^i}$$

This formulation obviously corresponds to another problem, which has the same $f^i(x, u)$ as the original one of § 2, but with a different adjoint system ($\psi_i \neq \psi_i'$) and with $f'^0 = 0$. Such a problem has a trivial solution. All admissible $u \in U$ satisfying the boundary conditions are optimal. This is only possible if $\mathcal{H} \equiv 0$, which means $\psi_i' = 0$, independantly of U , initial or final conditions. For this reason we have

$$\psi_i = \frac{\partial F}{\partial x^i} \quad \psi_0 = -1 \quad \mathcal{H} \equiv 0 \quad (11)$$

In contrast to most other applications, the control variable is not determined by the maximum principle.

3.2. The trajectory partly lies in the interior of G and partly on the boundary of G.

Introducing the Lagrange multipliers for the boundary of G and the jump conditions for Ψ at the junction points where the optimal trajectory reaches or leaves the boundary of G, we apply the maximum principle for restricted state coordinates. A reasoning similar to that of § 3.1. leads to the relations (12) and (13). For simplicity we took $m = 1$ in relation (6).

$$\Psi_i = \frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \quad \Psi_0 = -1 \quad \mathcal{H} \equiv 0 \quad (12)$$

$$\lambda = 0 \quad \text{for } g(x) < 0 \quad (13 \text{ a})$$

$$\lambda = - \frac{\sum_{i=1}^n \frac{\partial F}{\partial x^i} \cdot \frac{\partial g}{\partial x^i}}{\sum_{i=1}^m \left(\frac{\partial g}{\partial x^i} \right)^2} \quad \text{for } g(x) = 0 \quad (13 \text{ b})$$

Relation (13 b) satisfies the boundary condition $(\Psi, \text{grad } g) = 0$ for $g(x) = 0$ +)

3.3. The transversality conditions

Since $\mathcal{H} \equiv 0$ it finally turns out that the optimal control u has to be determined by the transversality conditions of the theorems of Pontryagin, stating that $\Psi(t_1)$, which corresponds in our problem to $\Psi(t + T)$, has to be orthogonal to the set $R_T(t)$ of the final events $x_1 = x(t_1) = x(t + T)$. Two possibilities have to be considered. They are illustrated in Fig. 1.

a) $x(t + T)$ lies in the interior of $R_T(t)$. In this case the orthogonality condition can only be satisfied by

$$\Psi_i(t + T) = \left(\frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \right) x(t + T) = 0 \quad (14)$$

+) By (a, b) we mean the scalar product of the vectors a and b.

This is true whatever be the value of T .

b) $x(t + T)$ lies on the boundary $\rho[x(t + T)] = 0$ of $R_T(t)$.

Now we have

$$\psi_i(t + T) = \left(\frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \right) x(t + T) = \mu \left(\frac{\partial \rho}{\partial x^i} \right) x(t + T) \quad (15)$$

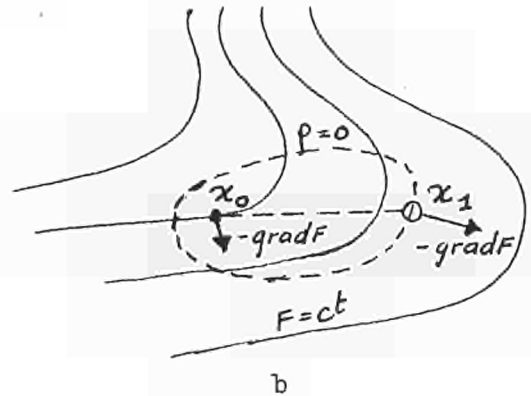
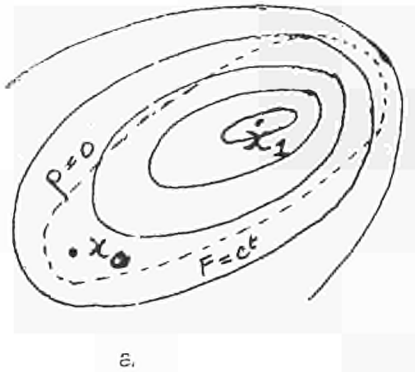


Fig. 1

3.4. Conclusions

1. With the generalized gradient it is no longer true that $dx/dt = f(x, u)$ moves along the steepest descent, but the vector ψ of the adjoint system takes over the same function.
2. Since the Hamiltonian is identically zero, the optimal control can only be determined by the transversality conditions and not necessarily in a unique way.
3. If the optimal trajectory $x(t)$ stops somewhere ($dx/dt = 0$) at a time τ in a point $x(\tau)$, it necessarily stops inside $R_T(\tau)$. This means (cfr (14)) that this point is a (local) optimum of $F(x)$. The fact that the trajectory stops implies $u = \text{constant}$, but not necessarily $u = 0$.
4. The optimal trajectory does not necessarily stop anywhere, even at a (local) optimum of $F(x)$. Indeed, the imposed system of constraints (cfr. (5), (6), (7)) may very well be of such a kind that it is completely impossible to keep the system at the optimum of $F(x)$. This is not at all in contradiction with the optimality of the trajectory. Fig. 2 illustrates this phenomenon.

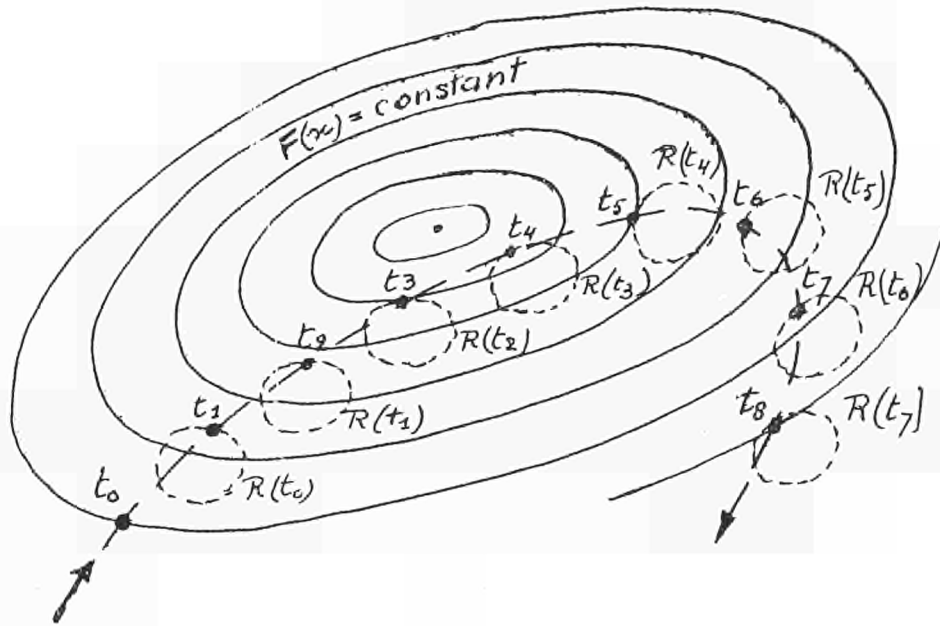


Fig. 2

5. Whenever $x(t + T)$ lies on the boundary $\rho = 0$ of $R_T(t)$, the trajectory can be determined by (15), the definition of $\rho = 0$ and the equations (7) giving the boundary of U . Indeed these relations determine $u(t)$ and μ . Examples will be given in § 4 and § 6.
6. Whenever $x(t + T)$ lies in the interior of $R_T(t)$ the trajectory is not uniquely determined unless $T = \Delta t$ is arbitrarily small. Indeed, the trajectory is composed of a sequence of initial conditions $x(t)$ of an optimization problem for $t \leq t' \leq t + T$ of which only the final condition is given by our optimization criterion (8). In this case we propose a new optimization criterion.

$$J = \int_t^{t+T} F(x) dt' \quad (16)$$

The problem will be subject of further study. Meanwhile the difficulty will be by-passed by taking T arbitrarily small, which eliminates the optimization between t and $t + T$.

7. It should be noticed that the evolution of the system only takes account of the direction of the vector $(\text{grad } F + \lambda \text{ grad } g)$ and not of its magnitude. The speed of evolution is more closely related to the system of dynamical constraints (5) and the boundary of U (7).

8. While optimizing we have to take account of the values of $F(x)$ only for x in the interior of $R_T(t)$. Outside this region $F(x)$ is completely ignored. It looks as if the optimizing system has a limited "horizon" of information around the moving point $x(t)$ and which is defined by $R_T(t)$. If the trajectory stops at a local optimum, the optimizer never finds a possibly higher optimum if this lies outside its horizon. The probability of staying at a local optimum obviously grows with decreasing T and becomes a certainty for T arbitrarily small. The concept of the horizon of an optimizing system seems to be realistic. Indeed, the ability of predicting and interpreting all possible events within some time period T is linked to a degree of complexity which is limited for most technologically realizable optimizing systems.

4. THE STEEPEST DESCENT AS A SPECIAL CASE OF THE GENERALIZED GRADIENT

We consider the problem of § 2 with the following restrictions:

$$T = \Delta t = \text{arbitrarily small}$$

$$f^i = u^i \quad i = 1, \dots, n \quad (17)$$

$$q(u) = \sum_{i=1}^n (u^i)^2 - 1 \leq 0 \quad (18)$$

By (18) U is represented by a unit hypersphere in the n -dimensional state-space. It follows immediately that $R_{\Delta t}(t)$ too is a hypersphere with radius Δt and that for $x(t + \Delta t)$ on the boundary $\rho = 0$ we have

$$\begin{aligned} \left(\frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \right)_{x(t + \Delta t)} &= \mu \left(\frac{\partial \rho}{\partial x^i} \right)_{x(t + \Delta t)} = 2\mu \left[x^i(t + \Delta t) - x^i(t) \right] \\ &= (2\mu \Delta t) u^i \end{aligned}$$

Knowing that for $x(t + \Delta t)$ on $\rho = 0$, u is on $q(u) = 0$ we can calculate

$$2\mu \Delta t = \sum_{i=1}^n u^i \left(\frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \right)$$

Since μ is negative for a minimizing problem we have the following final solution:

$$\frac{dx}{dt} = 0 \quad \text{for} \quad \text{grad } F + \lambda \text{ grad } g = 0 \quad (19)$$

$$\frac{dx}{dt} = -k (\text{grad } F + \lambda \text{ grad } g) \quad (20)$$

with $\lambda = (13 \text{ a})$ and (13 b)

$$k = \frac{1}{2\mu \Delta t} = \left[\sum_{i=1}^m \left(\frac{\partial F}{\partial x^i} + \lambda \frac{\partial g}{\partial x^i} \right)^2 \right]^{-1/2} \quad (21)$$

Relations (19) and (20) obviously are equivalent to (1) for $m = 1$. The only points meriting some comment are k and λ . As stated already in the introduction $k_j g^j(x)$ ($j = 1$) is nothing but an approximation of the Lagrange multiplier λ , generated by implicit computing. As far as k is concerned, in relation (20) it has been taken for the highest possible speed of x up to the endpoint. At the endpoint however, (19) is necessary since (20) becomes undetermined. In relation (1) k has not yet been specified. The only important restriction on k is that it is a scalar, which is the case in (1) as well as in (20).

To resume, it has been shown that for the steepest descent not only γ_i but also dx^i/dt is directed along $(\text{grad } F + \lambda \text{ grad } g)$. This is true because $T = \Delta t$ and $\rho [x(t + \Delta t)] = 0$ is a hypersphere.

5. COMPUTATIONAL ASPECTS OF THE GENERALIZED GRADIENT

On condition that $\text{grad } F(x)$ and $\text{grad } g(x)$ have sufficiently simple analytical expressions (or some simple analytical approximation, preferably linear, quadratic or third order polynomials) the equations (1) can be programmed without difficulty for analog and digital computers. In most cases the same is true for the generalized gradient if we take $T = \Delta t$. Especially when the set of equations (5) constitutes a complicated, high order, non-linear dynamical system the choice of an analog computer is indicated.

We draw special attention to the implicit computing of the Lagrange multipliers by $k_j g^j(x)$ or the equivalent technique of the penalty functions. This technique is extremely simple and well adapted for programming on analog computers, because of the continuous representation of the variables on the computer. But equally for digital computers the technique seems to have considerable advantages with respect to the direct computation of the Lagrange multipliers by relation (13 a, b). The accuracy of the approximation technique can be studied in terms of constraint "violations", corresponding to the maximum positive values of $g^j(x)$. These violations are roughly proportional to λ_j/k_j , a ratio which tends to zero with growing k_j . For large k_j however, difficult convergence problems in digital computers may arise. Even with the analog computer, in spite of its continuous representation, stability may become an embarrassing problem, since the stability is more strongly related to the structure of equations (5) and (7). Introducing additional damping forces for $g^j(x) > 0$ or nonlinear functions instead of constant k_j 's is very useful, but their influence upon the accuracy of the optimal trajectory still has to be examined carefully.

The simulation of U on a digital computer generally poses no difficult problems. On the analog computer it can be done in several ways. If possible the use of limiters is the most practical one. This is especially true for hyperparallelepipeds. A more general technique is once more the introduction of Lagrange multipliers γ^i for $q^i(u) = 0$, which have to be computed again in the same way as the λ 's. Since $u(t)$ can move in U without friction and without inertia, new stability problems add to the already discussed ones. They very often can be solved by a more realistic interpretation of the underlying physical phenomena.

Whenever $\text{grad } F(x)$ and $\text{grad } g(x)$ have no simple analytical expression all hope should not be lost. Modern perturbation techniques and sensitivity analysis are often very useful and easily programmable tools for estimating complicated gradients. This is especially the case for iteration procedures of the gradient type for solving two point boundary problems and parameter optimizations

[4].

Time-dependant optimizing functions and constraints, making the whole system non-autonomous, do not affect the generality of the proposed techniques. As Pontryagin states it is sufficient to consider time, wherever it appears explicitly, as a new state variable $x^{n+1} = t$ and to add a new equation of type (5): $dx^{n+1}/dt = 1$.

6. EXAMPLE OF A GENERALIZED GRADIENT APPLICATION

In terms of the statement of the problem of § 2 we take

$$dx^i/dt = u^i x^i \quad i = 1, \dots, n \quad (22)$$

$$g(x) \leq 0$$

$$q^i(u) = -u^i - A^i \leq 0$$

$$q^{i+n}(u) = +u^i - B^i \leq 0$$

$$T = \Delta t$$

The boundary of the set of possible events $\rho[x(t + \Delta t)] = 0$ takes the form of

$$\begin{aligned} x^i(t + \Delta t) &= (-A^i \Delta t) x^i(t) \\ x^i(t + \Delta t) &= (+B^i \Delta t) x^i(t) \quad \text{for } i = 1, \dots, n \end{aligned} \quad (23)$$

Fig. 3 clearly shows the structure of the solution, which can very well be approximated by the equations (24), taking full advantage of the implicit computing technique.

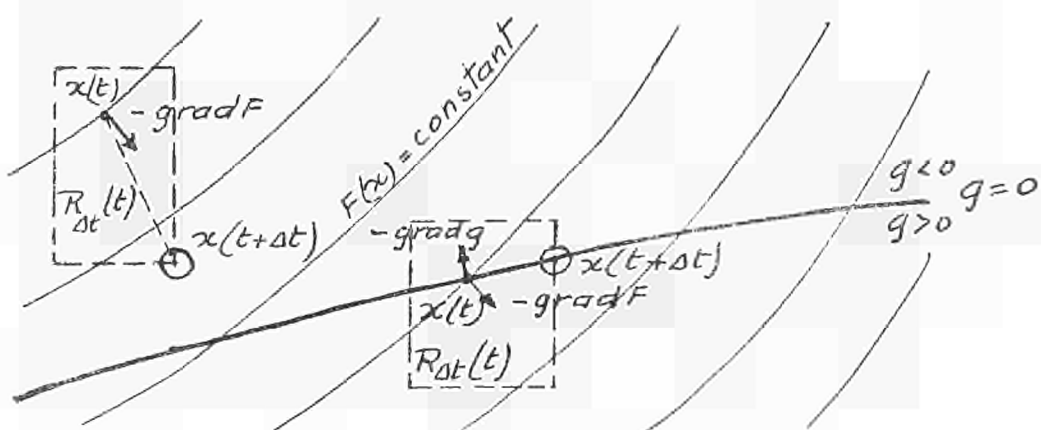


Fig. 3

$$\frac{dx}{dt} = -k x \left[\text{grad } F(x) + k_1 g(x) \text{grad } g(x) \right] \quad (24)$$

with: $-Ax \leq dx/dt \leq Bx$

k and k_1 large and positive, but $k_1 = 0$ for $g(x) \leq 0$.

A and B are constant vectors with coordinates A^i and B^i , representing the edges of the hyperparallelepiped U .

The larger k and k_1 the better the accuracy, but the more critical the stability of dx/dt whenever it lies between $-Ax$ and Bx . This particularly happens on the boundary $g(x) = 0$ (see fig. 3).

The analog computing diagram takes the form of Fig. 4.

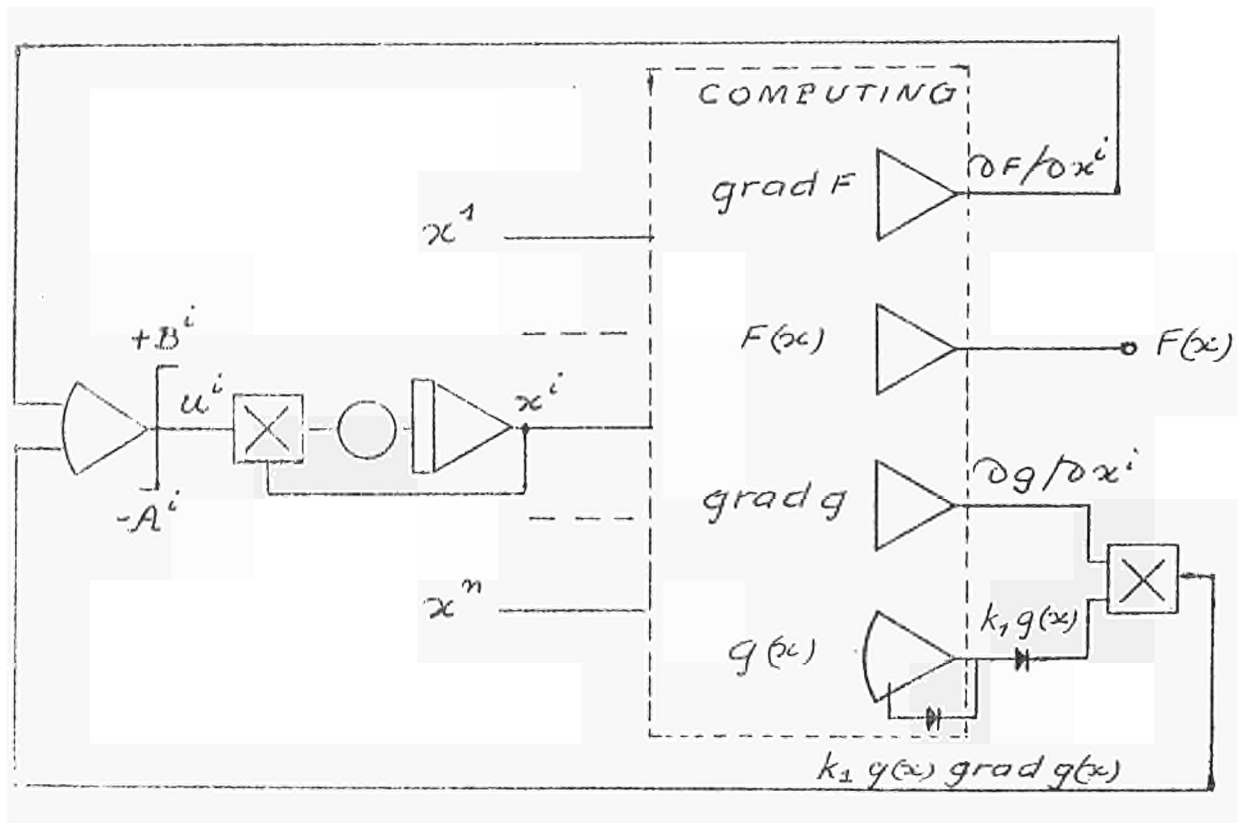


Fig. 4

7. APPLICATION FIELDS

The most important class of engineering applications probably is the optimization of industrial production processes for which the slow variation of the properties of the process constitutes a system of dynamical constraints of type (5), and for which a certain profit function corresponding to the quality of the final product, the production costs or some other criterion can be fixed [6].

Another interesting field would be the study of macro-economic structures. Indeed, the generalized gradient opens the possibility of integrating the laws of economic growth and economic optimization in one model.

ACKNOWLEDGEMENTS

This research owes much to P. Braffort, A. Gazzano and A. Riotte and I wish to express my gratitude for their support. Thanks are due to A. Riotte and C. Green for revising the paper and to many colleagues of CETIS for stimulating discussion.

REFERENCES

- [1] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko: "The Mathematical Theory of Optimal Processes", John Wiley & Sons, New York, 1962.
- [2] L.B. Pyne: "Linear Programming on an Electronic Analogue Computer", Communication and Electronics, 1956, 24: pp. 139 - 143.
- [3] H.J. Kelley: "Method of Gradients" pp. 213 - 215, in "Optimization Techniques" (G. Leitmann, ed), Academic Press, New York, 1962.
- [4] W. Brunner: "An Iteration Procedure for Parametric Model Building and Boundary Value Problems", Proc. Western Joint Computer Conference, Los Angeles, Part 12, pp. 519 - 533.
- [5] L.W. Neustadt: "Applications of linear and non-linear Programming Techniques", Proc. 3rd Int. Analogue Computation Meetings, 1961, Presses Acad. Eur., Brussels, Sect. B 1, pp. 197 - 200.
- [6] P.M.E.M. van der Grinten: "Multivariable Optimising Control by an Analogue Computer", Trans. Inst. Chem. Eng., vol. 40, no 6, 1962.

EUR 411.e

THE GENERALIZED GRADIENT, ITS COMPUTATIONAL ASPECTS AND ITS RELATIONS TO THE MAXIMUM PRINCIPLE by W. DE BACKER.

European Atomic Energy Community — EURATOM

Joint Nuclear Research Center

Ispra Establishment — Italy

Scientific Data Processing Center — CETIS

Text distributed at the IFAC-Congress, Basel (Switzerland) —

27 August - 4 September 1963

Brussels, February 1964 — pages 14 — figures 4

The maximum principle and the related theorems of the Pontryagin team make some generalizations of the well-known steepest descent possible. This generalization comprises the introduction of dynamical constraints. For this reason the state vector no longer moves along the steepest descent, but the τ -vector of the adjoint system does. With respect to the simulation of the dynamical constraints and the elegant technique of implicit computing of Lagrange multipliers, the analog computer seems the indicated one for generalized gradient optimization.

EUR 411.e

THE GENERALIZED GRADIENT, ITS COMPUTATIONAL ASPECTS AND ITS RELATIONS TO THE MAXIMUM PRINCIPLE by W. DE BACKER.

European Atomic Energy Community — EURATOM

Joint Nuclear Research Center

Ispra Establishment — Italy

Scientific Data Processing Center — CETIS

Text distributed at the IFAC-Congress, Basel (Switzerland) —

27 August - 4 September 1963

Brussels, February 1964 — pages 14 — figures 4

The maximum principle and the related theorems of the Pontryagin team make some generalizations of the well-known steepest descent possible. This generalization comprises the introduction of dynamical constraints. For this reason the state vector no longer moves along the steepest descent, but the τ -vector of the adjoint system does. With respect to the simulation of the dynamical constraints and the elegant technique of implicit computing of Lagrange multipliers, the analog computer seems the indicated one for generalized gradient optimization.



CDNA00411ENC