UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

(2)

AD-A214 489

DTIC ELECTE NOV 20 1989 S B D

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1b. RESTRICTIVE MARKINGS

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION: University of Washington

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research

6c. ADDRESS (City, State, and ZIP Code)
Department of Mathematics
GN-50
Seattle, WA 98195

7b. ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL (If applicable): NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0281

8c. ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
61102F	2304	A8	

11. TITLE (Include Security Classification)
PIECEWISE LINEAR-QUADRATIC PROGRAMMING AND ITS APPLICATIONS

12. PERSONAL AUTHOR(S)
R. T. Rockafellar

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 1 Jun 87 TO 30 Oct 88		

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

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

> Piecewise linear-quadratic programming problems are a fundamental class class of optimization problems in the mathematical modeling of multistage decision-making and large-scale dynamical systems, with or without the presence of uncertainty. Patterns of mathematical structure in such problems have been identified that cover wide areas of application and are conductive to the development of solution methodology. Work has gone forward on utilizing this structure in new numerical procedures, which include "finite-envelope methods" and a "double conjugate gradient method:, as well as a simplex-like algorithm for solving small scale subproblems. Preliminary tests have been made of these precedures on problems of modest size. To pave the way for experiments with larger examples, program modules for handling descrete-time dynamics have been coded in part. For decision problems involving scenarios, a progressive hedging algorithm have been devised. This provides a systematic approach to optimization in cases where uncertainty cannot be modeled in terms of standard random variables.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: DR NEAL GLASSMAN

22b. TELEPHONE (Include Area Code): (202) 767- 5026

22c. OFFICE SYMBOL: NM

PIECEWISE LINEAR-QUADRATIC PROGRAMMING AND ITS APPLICATIONS

Final Scientific Report

Grant No. AFOSR-87-0281

December 1, 1988

R. T. Rockafellar, Principal Investigator
Department of Mathematics
University of Washington, GN-50
Seattle, WA 98195

Summary. Piecewise linear-quadratic programming problems are a fundamental class of optimization problems in the mathematical modeling of multistage decision-making and large-scale dynamical systems, with or without the presence of uncertainty. Patterns of mathematical structure in such problems have been identified that cover wide areas of application and are conducive to the development of solution methodology. Work has gone forward on utilizing this structure in new numerical procedures, which include "finite-envelope methods" and a "double conjugate gradient method", as well as a simplex-like algorithm for solving small-scale subproblems. Preliminary tests have been made of these precedures on problems of modest size. To pave the way for experiments with larger examples, program modules for handling discrete-time dynamics have been coded in part. For decision problems involving scenarios, a progressive hedging algorithm has been devised. This provides a systematic approach to optimization in cases where uncertainty cannot be modeled in terms of standard random variables.

Research Directions and Motivation.

Quadratic programming problems have long been recognized as fundamental in optimization, but in the traditional formulation they suffer from serious drawbacks that make them unsuitable for large-scale modeling of dynamic or stochastic structure. While work has been done on iterative methods of solution for quadratic programming problems of high dimension (as surveyed by Y.Y. Lin and J.-S. Pang [**1**], J.-S. Pang [**2**]), the attitude has mainly been one of extending existing approaches through clever use of subproblem decomposition, sparseness of coefficient matrices, and efficient approaches to pivoting. The real difficulty, however, is that all such approaches emphasize exact treatment of constraints and aim ultimately at identifying the "active" constraints at the solution. In problems that are infinite-dimensional because of continuous time or continuous probability distributions, that kind of thinking makes little sense, so it is not surprising that quadratic programming approaches to the discretized versions of such problems, or to problems that involve many time periods and possible events even if there is no continuous analogue, have not been very successful.

Piecewise linear-quadratic programming, a newly developing subject that has been the focus of this grant, starts from a quite different point of view. Constraints that are simple variable bounds or dynamical equations are handled in a special pattern that brings out their primal-dual potential, but other constraints are not introduced as such. They are replaced by "monitoring terms", for instance penalty expressions, that allow for trade-offs between costs and other factors. This tactic has the advantage of much greater flexibility in computation, avoiding the conceptual bias toward low dimensions, but it also requires theoretical innovations and a fresh start in numerical methodology.

The following papers, presenting accomplishments supported by this grant, address issues in this area and lay a foundation for further research now in progress or anticipated in the near future.

[A] "Computational schemes f   lving large-scale problems in extended linear-quadratic programming", submitted   ath. Programming.

[B] "A simplex-active-set algorithm for monotropic piecewise quadratic programming", submitted to Math. Programming (with J. Sun).

[C] "Generalized linear-quadratric problems of deterministic and stochastic optimal control in discrete time", accepted for publication in SIAM J. Control Opt. (with R. J-B. Wets).

[D] "Scenarios and policy aggregation in optimization under uncertainty", submitted to Math. of Operations Research.

The results in these papers will be outlined below. Besides addressing the nature of piecewise linear-quadratic programming in general, they set up tools for the mathematical modeling of multistage optimization problems subject to various kinds of uncertainties.

On the numerical side, considerable effort has gone into comptational experiments with new methods for solving piecewise linear-quadratic programming problems in the kind of format suitable for large-scale applications. Code modules have been programmed and tested for later incorporation into a "finite envelope method" package for the multistage problems being modeled. Another approach which has been opened up and looks very promising is that of primal-dual conjugate gradient schemes for such problems. Progress in these directions will be described too.

An overall aim of this research is to develop a better version of linear-quadratic modeling for use in optimization, get corresponding new solution methods set up for it, and then tackle large nonquadratic problems by this means through iterative approximation in terms of the new generalized linear-quadratic subproblems.

Piecewise Linear-Quadratic Programming.

The challenge posed by piecewise linear-quadratic programming can best be appreciated through a brief comparison with ordinary quadratic programming, as conceived up until now. The latter, in the standard description, consists of minimizing a quadratic convex function subject to a system of constraints in the form of linear equalities and inequalities. The constraints define a kind of set that is said to be *polyhedral*. To be specific, let us write this set in terms of vectors $u = (u_1, \ldots, u_k)$ in $\mathbb{R}^k$ as

$$\{u \in U \mid Du - q \leq 0\}, \tag{1}$$

where U is a relatively simple subset of $\mathbb{R}^k$ such as a box giving upper and lower bounds on the variables, $q \in \mathbb{R}^l$ and $D \in \mathbb{R}^{l \times k}$. The objective function is given by a expression

$$p{\cdot}u + \tfrac{1}{2}u{\cdot}Pu, \tag{2}$$

where p is a vector of coefficients and P is a symmetric, positive semidefinite matrix. The ordinary problem thus consists of minimizing the expression (2) over all u belonging to the set (1). When $P = 0$, one has the case of linear programming.

This traditional view of quadratic programming suffers from serious limitations. In mathematical modeling it emphasizes the introduction of exact constraints $q - Du \leq 0$, which may well be the wrong direction to take in a large-scale setting. The possibility of modeling with penalty terms is ignored or, at the least, shoved into the background. Furthermore, dualization is made inconvenient or impossible in anything but the most elementary cases. This precludes the effective use of primal-dual approaches to computation such as otherwise might be envisioned.

In piecewise linear-quadratic programming, as it has emerged through this project, a typical problem instead takes the form

$$\text{minimize } f(u) = p{\cdot}u + \tfrac{1}{2}u{\cdot}Pu + \rho(q - Du) \text{ over } u \in U,$$

where U is still a relatively simple, polyhedral set, but ρ is a *monitoring* function which may, for instance, give penalties on the difference between $q - Du$ and 0 if this difference is in an undesired direction. The monitoring function is generally just piecewise linear quadratic, i.e., given by different linear or quadratic expressions on various polyhedral cells, and the objective function f is therefore only piecewise linear-quadratic as well. This is a serious complication, but research under this grant has shown that in all of the most important cases, at least, f can be given an alternative representation in which an underlying simplicity is regained.

The representation involves a Lagrangian function $L(u,v)$ on a product set $U \times V$, where V is a polyhedral set in a space $\mathbb{R}^l$. One has $L(u,v)$ quadratic in (u,v), convex in u and concave in v; Thus in general

$$L(u,v) = p{\cdot}u + \tfrac{1}{2}u{\cdot}Pu + q{\cdot}v - \tfrac{1}{2}v{\cdot}Qv - v{\cdot}Du \tag{3}$$

for symmetric, positive semidefinite matrices P and Q. One has

$$\max_{v\in V} L(u,v) = p{\cdot}u + \tfrac{1}{2}u{\cdot}Pu + \rho_{V,Q}(q - Du) = f(u) \tag{4}$$

for a certain monitoring function $\rho_{V,Q}$ determined from the specification of V and Q. Correspondingly, one has

$$\min_{u\in U} L(u,v) = q{\cdot}v - \tfrac{1}{2}v{\cdot}Qv - \rho_{U,P}(D^T v - p) = g(v) \tag{5}$$

for a certain monitoring function $\rho_{U,P}$. The primal problem thus takes the form

$(\mathcal{P})$ minimize over $u \in U$ the function f in (4),

where f is piecewise linear-quadratic and convex, while the dual problem takes the form

$(\mathcal{Q})$ maximize over $v \in V$ the function g in (5),

where g is piecewise linear-quadratic and concave. Solutions these problems correspond to saddle points $(\bar{u},\bar{v})$ of L relative to $U \times V$.

The formulas for the monitoring functions $\rho_{V,Q}$ and $\rho_{U,P}$ in $(\mathcal{P})$ and $(\mathcal{Q})$ can be made quite explicit in the context of penalty models for constraints, although we shall forgo this here. The main point is that the simple algebraic character of this class of problems is not to be seen in terms of $(\mathcal{P})$ and $(\mathcal{Q})$ themselves, but in the associated *saddle point* problem. A shift toward solution methods directed at finding saddle points is thereby suggested, in contrast to the usual thinking devoted exclusively to primal descent or dual ascent. It is principally in this direction that one needs to turn, because the functions $f(u)$ and $g(v)$ are generally ***not*** "smooth" (continuously differentiable to whatever degree is desired), and yet the kinds of procedures in existence for nonsmooth optimization, such as in the book of Kiwiel [3], are essentially conceived for low-dimensional spaces only.

Especially useful is the feature that in being able to calculate $f(u)$ for any $u \in U$ and $g(v)$ for any $v \in V$, such as points constructed in the course of some iterative method of solution, one can obtain the difference $\varepsilon = f(u) - g(v)$. From basic duality theory it is then known that the given u comes within ε of achieving the minimum in $(\mathcal{P})$ while the given v comes within ε of achieving the maximum in $(\mathcal{Q})$. Thus progress can carefully be measured, and a stopping criterion for suboptimality is always conveniently at hand.

other than $v \in V$.

Double decomposability assumption: For each $u \in U$ the maximum in $v \in V$ can be calculated in (4). Likewise, for each $v \in V$ the minimum in $u \in U$ can be calculated in (5).

The finite monitoring assumption means that all constraints that might have been forseen for the vector $q - Du$ in $(\mathcal{P})$ or the vector $D^T v - p$ in $(\mathcal{Q})$ have been modeled instead with finite penalties. It is satisfied if the sets U and V are bounded, for instance, or if the matrices P and Q are positive definite. This assumption is therefore harmless; other cases are readily approximated by it.

The double decomposability assumption is satisfied in practice through the fact that $L(u, v)$, as defined in (3), can be taken to be separable in u for fixed v and also separable in v for fixed u. It is not obvious that this is generally possible, and the demonstration that it is so is one of the chief contributions of paper [C]. The upshot is that the minimization of $L(u, v)$ in u for fixed v reduces to a large number of separate subproblems in the individual components of u. These subproblems can be solved *in parallel*; indeed they often are one-dimensional and have closed-form solutions that just require plugging into a formula. This holds as well for the maximization of $L(u, v)$ in v for fixed u.

The challenge is how to make good use of these advantageous properties in solving $(\mathcal{P})$ and $(\mathcal{Q})$. The type of thinking that is involved breaks with the past: no one has previously hit upon this pattern and realized its fundamental importance for large-scale applications.

In [A] a general foundation is built for working with such problems, and a class of methods, called *finite envelope methods*, is introduced and analyzed. These methods proceed as follows in terms of low-dimensional subproblems. At each iteration ν (where $\nu = 1, 2, \ldots$) a finite subset U^ν of U is designated along with a finite subset V^ν of V. Instead of finding a saddle point of $L(u, v)$ over $U \times V$, one does it over $\operatorname{co} U^\nu \times \operatorname{co} V^\nu$, thereby obtaining a vector pair (u^ν, v^ν). This subproblem is *small-scale* when the number of elements in U^ν and V^ν is small, and it can therefore be solved by small-scale methods of piecewise linear-quadratic programming like those discussed above.

At first this assertion may seem surprising, because $\operatorname{co} U^\nu$ and $\operatorname{co} V^\nu$ nevertheless lie in spaces of putatively high dimension, but the reduction is made possible by the quadratic structure of the Lagrangian L. Elements of $\operatorname{co} U^\nu$ and $\operatorname{co} V^\nu$ can be expressed as convex combinations of the elements of U^ν and V^ν, the coefficients being nonnegative and adding up to 1, and when these expressions are substituted into the formula for $L(u, v)$ they result in a low-dimensional quadratic form in the coefficients in question. The subproblem thus reduces to one of finding a saddle point of a low-dimensional "projected" Lagrangian function over a product of two simplexes, and this does belong to the realm of small-scale piecewise linear-quadratic programming.

Results presented in [A] give guidelines for generating the sets U^ν and V^ν out of decomposed maximization in (4) or minimzation in (5) in such a way that the sequence of pairs (u^ν, v^ν) calculated from the subproblems, or another easily derived from it, will converge to a saddle point $(\bar{u}, \bar{v})$ of $L(u, v)$ on $U \times V$ and thus provide an optimal solution $\bar{u}$ to $(\mathcal{P})$ and an optimal solution $\bar{v}$ to $(\mathcal{Q})$. Remarkably, it is *not* necessary for the number of elements in U^ν and V^ν to grow indefinitely in order to ensure this convergence. As long as the matrices P and Q are positive definite, convergence to optimality can be guaranteed with these sets kept really small. Of course this is a theoretical fact, which must be supplemented by numerical experience as to the rules that seem to work best. The positive definiteness supporting the convergence can always be introduced if necessary by an outer algorithm of the proximal point variety, first developed by the P.I. in [7], [8].

In the numerical experiments on finite envelope methods that have been carried out so far, the dimensions of the u and v vectors have been kept at about 50 each. This size, which does not correspond to ultimate large-scale aims, has been dictated by a number of factors, the most important being that of getting ahead with initial tests of ideas while the way was being prepared on a separate track for utilizing more details of structure, such as are sure to be present in dynamic or stochastic versions of $(\mathcal{P})$ and $(\mathcal{Q})$. To a good job of experimenting with large problems, it is not appropriate merely to generate bigger and bigger matrices randomly. Steve Wright, a Ph.D. student in the Mathematics Department of the University of Washington who has received support as an R.A. under this grant, has been developing code to make large, but sensible and well chosen problems available for numerical testing. This will be described below in connection with the modeling efforts.

Anyway, with the 50×50 problems the method has worked well and much has been learned about different strategies in implementation. This experience has been invaluable in setting the stage for the next level of computation. It has also been possible to draw on experience with an earlier version of finite envelope methods in the case of two-stage stochastic programming as developed in Rockafellar and Wets [9], [10], and implemented by King [11] and Wagner [12]. That version concerned cases where only the v dimension is high, so the subproblems consist of finding saddle points over sets $U \times \operatorname{co} V^\nu$, i.e., one can work directly at all times with the whole set U. Incidentally, the proximal point outer algorithm mentioned above was used in this case to great effect.

Another line of attack on large-scale piecewise linear-quadratic programming problems under the finite monitoring and double decomposibility assumptions has been a new "double" form of the conjugate gradient method. The conjugate gradient method is well known in numerical optimization as a means of minimizing an unconstrained quadratic convex function in finitely many steps and, in various generalizations, providing a good iterative approach to minimizing a smooth (twice continuously differentiable) convex function on a space of high dimensions. For problems with constraints, however, the classical method has been hampered by development only in terms of active constraint strategies which, as already pointed out, are unsuitable for large-scale applications.

The new version of the conjugate gradient method now under development, the work on it having been started in the grant period covered by this report, makes use of the assumed ability to calculate vectors giving the maximum in (4) or the minimum in (5) and interprets this operation as a sort of gradient projection onto the sets U and V. Certain formulas that are well known classically can then be emulated despite the constraints $u \in U$ and $v \in V$ to get a procedure resembling the conjugate gradient method for minimizing $f(u)$ over U in the primal problem $(\mathcal{P})$ as well as one for maximizing $g(v)$ in the dual problem $(\mathcal{Q})$. The truly novel feature—with surprising consequences—is that both the primal and dual methods are executed *simultaneously* and with a kind of *feedback*.

The feedback comes from the circumstance that each of the conjugate gradient algorithms, although applied in concept to only one of the two problems, inevitably produces information about the other problem. By running them simultaneously and letting them "listen to each other", it is possible to take advantage of this information to keep either algorithm from getting into a rut. Numerous restarts occur, namely when one of the algorithms latches on to a better point that happens to have been produced by the companion algorithm, and such events usually result in drastic improvements in the move toward optimality.

A major share of the effort on this conjugate gradient phase of the project has been put in by Ciyou Zhu, a Ph.D. student in the Applied Mathematics Department of the University of Washington. He was not able to receive R.A. support from this grant, however, due to a lack of funds. (He was supported from other sources.)

The double conjugate gradient method may supersede the finite envelope class of methods, as it has given better performance in tests so far. It has even been used tried with success on problems of size 500×500. The real tests, though, will come when the very large, structured problems come on line. Those problems may balk at the line search routine required in the conjugate gradient approach and may need more accumulation of information, such as can be accomplished through the many ways of generating envelopes in the choice of the sets U^ν and V^ν.

Large-Scale Modeling.

A strong motivation for studying problems in piecewise linear-quadratic programming has been the fact that such problems provide an apparently effective, but heretofore unrecognized, way to capture the large-scale optimization structure in a wide class of applications. This has been brought out quite strikingly in paper [C], written jointly with R. Wets of the Department of Mathematics of the University of California at Davis. Earlier work with Wets in [9] had shown that two-stage stochastic recourse problems with underlying linear-quadratic structure could advantageously be placed in this format, thereby opening new avenues of insight and computation. The P.I. had also demonstrated more recently in [13] that optimal control problems in continuous time could be viewed in terms of a sort of infinite-dimensional piecewise linear-quadratic programming, and in this manner it is possible to look numerically at *constrained* forms of linear-quadratic control, including some with nonquadratic classical penalties, that previously had been beyond the pale. In [C], however, the problems can be multistage stochastic as well as deterministic. The context is one of discrete time and discrete probability. Substantial results are obtained about the characterization of optimality, and entirely new territory is entered for setting up problem models advantageously.

It is difficult here to do justice to the full scope and potential in [C], but a particular case will convey the key ideas. For now, the description will be deterministic only, but the stochastic version will come up later.

The primal problem to be discussed concerns a dynamical system with states $x_t \in \mathbb{R}^n_t$

and controls $u_t \in \mathbb{R}^k$ for $t = 0, 1, \ldots, \tau$, the states being determined from the controls by

$$\begin{aligned} x_t &= A_t x_{t-1} + B_t u_t + b_t \text{ for } t = 1, \ldots, \tau, \\ x_0 &= B_0 u_0 + b_0, \quad \text{with } u_t \in U_t, \end{aligned} \tag{6}$$

where the constraint set U_t is polyhedral. As explained in [C], the introduction of the state vectors is to some extent an artificial construct, depending on the particular situation, but it pays handsomely in bringing out the temporal structure without incurring any loss of generality. The optimization problem for this system is

$$(\mathcal{P}_0) \qquad \begin{aligned} \text{minimize } &\sum_{t=1}^{\tau} \left[p_t{\cdot}u_t + \tfrac{1}{2}u_t{\cdot}P_t u_t + \rho_{V_t,Q_t}(q_t - C_t x_{t-1} - D_t u_t) - c_{t+1}{\cdot}x_t\right] \\ &+ \left[p_0{\cdot}u_0 + \tfrac{1}{2}u_0{\cdot}P_0 u_0\right] + \left[\rho_{V_{\tau+1},U_{\tau+1}}(q_{\tau+1} - C_{\tau+1}x_\tau) - c_{\tau+1}{\cdot}x_\tau\right] \\ &\text{subject to (6).} \end{aligned}$$

The monitoring terms in this problem $(\mathcal{P}_0)$ offer rich possibilities for specialization, which cannot be gone into here, but the resemblence with the earlier model problem $(\mathcal{P})$ is evident.

What is far from evident is that $(\mathcal{P}_0)$ is truly a problem of type $(\mathcal{P})$ for a certain choice of data elements. It is the primal problem corresponding to a Lagrangian of the form

$$\begin{aligned} L_0(u, v) = &\sum_{t=1}^{\tau} \left[p_t{\cdot}u_t + \tfrac{1}{2}u_t{\cdot}P_t u_t + q_t{\cdot}v_t - \tfrac{1}{2}v_t{\cdot}Q_t v_t - v_t{\cdot}D_t u_t\right] + l(u, v) \\ &\text{over } U \times V = (U_0 \times \cdots \times U_\tau) \times (V_1 \times \cdots \times V_{\tau+1}), \end{aligned} \tag{7}$$

where $u = (u_0, \ldots, u_\tau)$ and $v = (v_1, \ldots, v_{\tau+1})$ and the function $l(u, v)$ is a simple bilinear form expressing the dynamics. The corresponding dual problem $(\mathcal{Q}_0)$ turns out to have the same character as $(\mathcal{P}_0)$ but in terms of the vectors v_t as controls as well as certain vectors y_t as states that are determined by a dynamical system that goes backward in time from $\tau + 1$.

The crucial observation is that the Lagrangian $L_0(u, v)$ has a *natural double decomposability* with respect to time. If v is fixed, the minimization of $L_0(u, v)$ in u breaks up into separate problems for each t, which in turn could reduce smaller problems, even one-dimensional problems, if the matrices P_t and Q_t are diagonal and the polyhedral sets U_t and V_t are boxes. Similarly, the maximization of $L_0(u, v)$ in v decomposes relative to t. This fundamental property had not been recognized prior to paper [C] and the continuous-time version in [13].

The details of these problems must be omitted, but the connection with the preceding discussion of computational projects is that it makes great sense to aim in this direction for the testing out of algorithms. Accordingly, the task was formulated of developing code for inputing the data that defines $(\mathcal{P}_0)$ and setting up program modules for carrying out the operations of minimizing or maximizing $L_0(u, v)$ in its separate arguments and integrating the primal and dual state trajectories x and y from the primal and dual controls.

This project was taken on by Steve Wright, already mentioned as the Ph.D. student who received the budgeted 6 months of R.A. support under the grant. (Another of the P.I.'s Ph.D students, Maijan Qian in the Mathematics Department, got one month for her assistance when it turned out toward the end of the grant period that the funds could be put together from remnants in other categories of the expiring budget.) Wright has programmed a fairly complete package based on discretizing a continuous-time problem; it has the feature that simply by setting a single parameter one can generate piecewise linear-quadratic optimization problems of arbitrarily large dimension which still have a natural structure and are therefore good for numerical testing of finite envelope methods and the double conjugate gradient method. As of the end of the grant period, the manual for this package was being written and the code, though in principle complete, was still being checked out.

Problems Involving Uncertainty.

The dynamical problem ($\mathcal{P}_0$) extends readily to the treatment of stochastic elements, i.e., random variables, to which the decisions being made can respond in time as more is learned. The theory for this is developed in paper [C]. A fundamental model for *multistage* stochastic programming is thereby obtained, where until now two-stage problems are virtually the only ones that have been approached computationally. Trying to solve such problems would be premature at present, because the multistage deterministic framework needs to be worked with first, and that is just coming on line. But an ultimate goal obviously is to combine the deterministic methodology with that already known from two-stage stochastic programming (for instance in [9] and [10]). This will forge a very strong instrument for progress in large-scale optimization in a multitude of applications, from optimal control to decision-making under uncertainty.

Paper [D], written under this grant with the collaboration of R. Wets, operates from a somewhat different angle. It too treats optimization problems involving uncertainty, but the focus is on the kind of uncertainty that is difficult to model probabilistically, i.e., in terms of random variables with specified distributions. Such problems are extremely common but have long stymied theoreticians. Typically they have been approached by means of *scenarios*. In this vein one makes assumptions about what might happen in the future and identifies the optimal decisions that would be taken if one were able to act validly on the basis of such assumptions. Each set of assumptions generates one scenario, and typically some dozens are worked out in a given case. The challenge then is to make use of what has been learned from the separate scenarios and come up with a mixed decision that properly hedges against the eventualities.

This "mixing" has actually been no more than intuitive guesswork in practice, completely devoid of theoretical basis. In [D], however, a more solid procedure is offered. Obviously the topic does not lend itself to really precise study, but the supposition can be made that the decision-maker does have some notion of weights for the different scenarios. One can search for a sensible, systematic approach to using such weights—from which it might also be revealed how sensitive the resulting mixed decision is to the particular weights, thereby providing feedback to the decision-maker and directing his attention to potential troublespots. This is the motivation in [D], which presents an iterative method for utilizing such weights in *progressive hedging*.

A highly attractive feature of the method is that it builds on whatever method may already be available in a given case for solving an individual scenario subproblem. Modified forms of such scenario subproblems are solved repeatedly and *in parallel*. The outputs are combined according to a scheme that at any stage does give at least an *implementable* policy, with the assurance that as the iterations go on, this policy will come closer and closer to satisfying all other requirements and being optimal (relative to the specified weights).

There is much yet to be done in this ongoing research. A number of prospects are emerging for integrating the scenario ideas with the multistage optimization models discussed earlier to obtain a highly flexible methodology.

References.

[1] Y.Y. Lin and J.-S. Pang, "Iterative methods for large convex quadratic programs: a survey", SIAM J. Control Opt. **25** (1987), 383–411.

[2] J.-S. Pang, "Methods for quadratic programming: a survey", Computers and Chem. Engineering **7** (1983), 583–594.

[3] K.C. Kiwiel, *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Math. 1133, Springer-Verlag, Berlin, 1985.

[4] R.T. Rockafellar, *Network Flows and Monotropic Optimization*, Wiley, 1984.

[5] D.P. Bertsekas, P.A. Hosein, and P. Tseng, "Relaxation methods for network flow problems with convex arc costs", SIAM J. Control Opt. **25** (1987), 1219–1243.

[6] P. Tseng and D.P. Bertsekas, "Relaxation methods for problems with strictly convex costs and linear inequality constraints," technical report LIDS-P-1717, Laboratory for Information and Decision Systems, M. I. T. (1987).

[7] R.T. Rockafellar, "Monotone operators and the proximal point algorithm", SIAM J. Control Opt. **14** (1976), 877–898.

[8] R.T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming", Math. of Op. Research **1** (1976), 97–285.

[9] R.T. Rockafellar and R. J-B Wets, "A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming", Math. Programming Studies **28** (1986), 63–93.

[10] R.T. Rockafellar and R. J-B Wets, "Linear-quadratic problems with stochastic penalties: the finite generation altorithm", in *Stochastic Optimization*, V.I. Arkin et al. (eds.), Springer-Verlag Lecture Notes in Control and Information Sciences No. 81, 1987, 545–560.

[11] A. King, "An implementation of the Lagrangian finite generation method", in *Numerical Techniques for Stochastic Programming Problems*, Y. Ermoliev et al. (eds.), Springer-Verlag, 1988, 295–312.

[12] J.M. Wagner, *Stochastic Programming with Recourse Applied to Groundwater Quality Management*, doctoral dissertation, M.I.T, 1988.

[13] R.T.Rockafellar, "Linear-quadratic programming and optimal control", SIAM J. Control and Optimization **25** (1987), 781–814.