

# Dual Method for the Solution of a One-Stage Stochastic Programming Problem with Random RHS Obeying a Discrete Probability Distribution

By A. Prékopa<sup>1</sup>

*Abstract:* In this paper we present a method for the solution of a one stage stochastic programming problem, where the underlying problem is an LP and some of the right hand side values are random variables. The stochastic programming problem that we formulate contains probabilistic constraint and penalty, incorporated into the objective function, used to penalize violation of the stochastic constraints. We solve this problem by a dual type algorithm. The special case where only penalty is used while the probabilistic constraint is disregarded, the simple recourse problem, was solved earlier by Wets, using a primal simplex algorithm with individual upper bounds. Our method appears to be simpler. The method has applications to nonstochastic programming problems too, e.g., it solves the constrained minimum absolute deviation problem.

*Zusammenfassung:* In dieser Arbeit wird eine Methode vorgestellt zur Lösung einstufiger stochastischer Programme, wobei das zugrundeliegende Problem ein LP mit zufälligen rechten Seiten darstellt. Das resultierende stochastische Programm enthält Wahrscheinlichkeitsrestriktionen und Strafterme, letztere innerhalb der Zielfunktion zur Bestrafung von Abweichungen in den stochastischen Restriktionen. Wir lösen dieses Problem mit einem dualen Algorithmus. Der Spezialfall, in dem ausschließlich Strafterme benutzt werden und Wahrscheinlichkeitsrestriktionen unberücksichtigt bleiben, d.h. das einfache Kompensationsmodell, wurde bereits früher von Wets mittels eines primalen Simplex-Algorithmus mit einzelnen oberen Schranken gelöst. Unsere Methode scheint einfacher zu sein. Die Methode ist auch auf nicht-stochastische Programme anwendbar, z.B. auf das Problem minimaler absoluter Abweichungen von Nebenbedingungen.

## 1 Introduction

Stochastic programming problems are formulated so that first we start from an underlying problem that would be the optimization problem if there were no random variables in it. Having observed that some of the parameters in the underlying problem are random, we formulate a new problem, by using some statistical decision principle. One underlying problem we are dealing with is the following type

---

<sup>1</sup> András Prékopa, Rutcor – Rutgers Center for Operations Research, Hill Center for the Mathematical Sciences, Rutgers's University, New Brunswick, New Jersey 08903, USA.

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b, \quad x \geq 0, \quad Tx \geq \xi, \end{aligned} \tag{1.1}$$

where  $A$  is an  $m \times n$ ,  $T$  is an  $r \times n$  matrix and the vectors  $c, x, b, \xi$  have dimensionalities consistent with the formulation (1.1). Assuming  $\xi$  to be a random vector, the expectation of which exists, we reformulate (1.1) in the following manner

$$\begin{aligned} &\text{minimize} && \left\{ c^T x + \sum_{i=1}^r q_i E[\xi_i - T_i x]^+ \right\} \\ &\text{subject to} && P(Tx \geq \xi) \geq p, \quad Ax = b, \quad x \geq 0, \end{aligned} \tag{1.2}$$

where  $\xi_1, \dots, \xi_r$  are the components of  $\xi$  and  $T_1, \dots, T_r$  are the rows of  $T$ ;  $q_1, \dots, q_r$  are positive constants, the symbol  $E$  designates the expectation,  $[z]^+ = z$  if  $z \geq 0$ , otherwise it is 0 and  $p$  is a fixed probability, given in advance. Typical values for  $p$  are 0.8, 0.9, 0.95, 0.99. Problem (1.2) is a one stage problem because decision is made in only one stage. For the expectation of a random variable  $\eta$  we will use both notations  $E(\eta)$  and  $E\eta$ . The second underlying problem we are dealing with differs from (1.1) by the last constraint and it has the form

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && Ax = b, \quad x \geq 0, \quad Tx = \xi. \end{aligned} \tag{1.3}$$

For this case the stochastic programming problem is

$$\begin{aligned} &\text{minimize} && \left\{ c^T x + \sum_{i=1}^r q_i^+ E[\xi_i - T_i x]^+ + \sum_{i=1}^r q_i^- E[T_i x - \xi_i]^+ \right\} \\ &\text{subject to} && Ax = b, \quad x \geq 0. \end{aligned} \tag{1.4}$$

Here we assume that  $q_i^+ + q_i^- > 0, i = 1, \dots, r$ . We have not included a probabilistic constraint of the type  $P(Tx = \xi) \geq p$ . The reason for this is that if  $p > 1/2$  than  $\xi$  can take only one value, if any, for which the above probabilistic inequality can be satisfied. If this possible value of  $\xi$  is designated by  $z$  then the constraint  $(Tx = \xi) \geq p$  can be replaced by  $Tx = z$  which, in turn, can be included among the deterministic constraints of the problem.

Let  $F_i(z)$  designate the probability distribution function of  $\xi_i$ , i.e.  $F_i(z) = P(\xi_i \leq z)$  for every real  $z$ . It is well known that problems (1.2) and (1.4) can be reformulated as

minimize

$$\left\{ c^T x + \sum_{i=1}^r q_i \int_{T_i x}^{\infty} [1 - F_i(z)] dz = c^T x + \sum_{i=1}^r q_i \left[ \mu_i - T_i x + \int_{-\infty}^{T_i x} F_i(z) dz \right] \right\}$$

subject to

$$P(Tx \geq \xi) \geq p, \quad Ax = b, \quad x \geq 0, \quad (1.5)$$

and

$$\text{minimize} \quad \left\{ c^T x + \sum_{i=1}^r \left[ q_i^+ (\mu_i - T_i x) + (q_i^+ + q_i^-) \int_{-\infty}^{T_i x} F_i(z) dz \right] \right\}$$

$$\text{subject to} \quad Ax = b, \quad x \geq 0, \quad (1.6)$$

respectively. The objective function of problem (1.6) reduces to that of problem (1.5) by setting  $q_i^+ = q_i$ ,  $q_i^- = 0$ ,  $i = 1, \dots, r$ .

Stochastic programming problems of the type (1.6) were formulated first by Dantzig (1955) and Beale (1955) whereas the probabilistic constrained stochastic programming model was introduced by Carnes, Cooper and Symonds (1958). The combined use of penalties and probabilistic constraint, as in problem (1.5) was proposed by the author of this paper (1973). Problem (1.6) is called by Wets (1983) the simple recourse problem. For other references regarding this model construction the reader is referred to his paper.

Assume that the random variables  $\xi_1, \dots, \xi_r$  have discrete probability distributions with finite possible values. Let  $z_{i1}, \dots, z_{ik_i}$  be the possible values of  $\xi_i$  arranged in increasing order. Assume furthermore that there exist two numbers  $z_{i0}$  and  $z_{ik_i+1}$ , corresponding to each  $i$  ( $1 \leq i \leq r$ ), such that for every  $x$  satisfying  $Ax = b$ ,  $x \geq 0$ , we have  $z_{i0} < T_i x \leq z_{ik_i+1}$  and

$$z_{i0} < z_{i1} < \dots < z_{ik_i} < z_{ik_i+1}.$$

Under this condition the objective functions of problems (1.5), (1.6) can be written in the forms of separable, piecewise linear, convex functions, by introducing some additional, linear constraints. This can be done by the use of two different linearization methods which are the  $\delta$ , and the  $\lambda$ -methods.

Given a continuous, piecewise linear, convex function  $f$ , defined in the interval  $[z_0, z_{k+1}]$  with breakpoints at  $z_0 < \dots < z_{k+1}$  we define

$$\delta_j = \frac{f(z_j) - f(z_{j-1})}{z_j - z_{j-1}}, \quad j = 1, \dots, k+1. \quad (1.7)$$

By the assumptions, we have  $\delta_1 < \delta_2 < \dots < \delta_{k+1}$  and any function value  $f(y)$ , corresponding to  $y \in [z_0, z_{k+1}]$ , can be represented in the form

$$\begin{aligned}
 f(y) = \min & \quad \left\{ f(z_0) + \sum_{j=1}^{k+1} \delta_j v_j \right\} \\
 \text{subject to} & \quad z_0 + \sum_{j=1}^{k+1} v_j = y \\
 & \quad 0 \leq v_j \leq z_j - z_{j-1}, \quad j = 1, \dots, k+1.
 \end{aligned} \tag{1.8}$$

This is the  $\delta$ -representation of the function value  $f(y)$ . The  $\lambda$ -representation is the following

$$\begin{aligned}
 f(y) = \min & \quad \sum_{j=0}^{k+1} f(z_j) \lambda_j \\
 \text{subject to} & \quad \sum_{j=0}^{k+1} z_j \lambda_j = y \\
 & \quad \sum_{j=0}^{k+1} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j = 0, \dots, k+1.
 \end{aligned} \tag{1.9}$$

Both representations can be applied to problems (1.5) and (1.6). We will consider problem (1.6) first. Disregarding a constant term, the objective function is the sum of  $c^T x$  and  $\sum_{i=1}^r f_i(T_i x)$ , where

$$f_i(y_i) = -q_i^+ y_i + (q_i^+ + q_i^-) \int_{z_{i0}}^{y_i} F_i(z) dz, \quad i = 1, \dots, r. \tag{1.10}$$

These are piecewise linear convex functions in the intervals  $[z_{i0}, z_{ik_i+1}]$ ,  $i = 1, \dots, r$  thus, both the  $\delta$ - and the  $\lambda$ - representations are applicable. Wets (1983) applied the  $\delta$ - representation and solved the problem by the use of a primal simplex method with individual upper bound technique. He exploits the special structure of the problem but the number of variables is large and the presentation of the method is somewhat complicated. We will apply the  $\lambda$ - representation, work with much smaller number of variables and present a simple, dual type algorithm. This is done in section 2. In section 3 we show how the proposed algorithm can be applied to solve other problems, e.g. the constrained minimum absolute deviation problem. In section 4 we solve the problem (1.5).

Let us introduce the notations

$$c_{ij} = -q_i^+ z_{ij} + (q_i^+ + q_i^-) \int_{z_{i0}}^{z_{ij}} F_i(z) dz, \quad j = 0, \dots, k_i + 1, \quad i = 1, \dots, r.$$

Using the fact that the functions (1.10) are piecewise linear and convex, we apply the  $\lambda$ -representation and reformulate problem (1.6) in the following manner:

$$\begin{aligned} \min_x \quad & \left\{ c^T x + \min_{\lambda} \sum_{i=1}^r \sum_{j=0}^{k_i+1} c_{ij} \lambda_{ij} \right\} \\ \text{subject to} \quad & \sum_{j=0}^{k_i+1} z_{ij} \lambda_{ij} = y_i \\ & \sum_{j=0}^{k_i+1} \lambda_{ij} = 1, \quad \lambda_{ij} \geq 0 \quad \text{all } i, j \\ & T_i x = y_i, \quad i = 1, \dots, r \\ & Ax = b, \quad x \geq 0. \end{aligned} \tag{1.11}$$

This, in turn, can be written in the following manner

$$\begin{aligned} \min_{x, \lambda} \quad & \left\{ c^T x + \sum_{i=1}^r \sum_{j=0}^{k_i+1} c_{ij} \lambda_{ij} \right\} \\ \text{subject to} \quad & Ax = b, \\ & T_i x - \sum_{j=0}^{k_i+1} z_{ij} \lambda_{ij} = 0, \\ & x = 0, \quad \lambda_{ij} \geq 0, \quad j = 0, \dots, k_i + 1, \quad i = 1, \dots, r. \end{aligned} \tag{1.12}$$

The matrix of the equality constraints has the following structure (on the top the objective function coefficients are listed):

$c_1 \quad \dots \quad c_n$	$c_{10} \quad \dots \quad c_{1k_1+1} \quad \dots \quad c_{r0} \quad \dots \quad c_{rk_r+1}$
$A$	
$T$	$-z_{10} \quad \dots \quad -z_{1k_1+1} \quad \dots \quad \dots \quad \dots \quad -z_{r0} \quad \dots \quad -z_{rk_r+1}$
	$1 \quad \dots \quad 1 \quad \dots \quad \dots \quad \dots \quad 1 \quad \dots \quad 1$

## 2 Dual Method for the Solution of Problem (1.12)

Problem (1.9) has a fundametal property that makes possible the development of a simple dual algorithm for the solution of problem (1.12). This is expressed by

*Lemma 2.1:* All dual feasible bases of problem (1.9) are dual non-degenerate and consist of two consecutive columns of the matrix of the equality constraints.

*Proof:* This lemma is a special case of theorem 3.1 in the paper by Prékopa (1990). A simple direct proof is presented below.

The dual feasibility of a basis means that the (sufficient) condition of optimality is satisfied. Let  $a_i, i = 0, \dots, k + 1$  be the columns in the equality constraints of problem (1.9) and let  $B = (a_i, a_j)$ , where  $i < j$ . Let furthermore  $f_i = f(z_i), i = 0, \dots, k + 1$  and  $f_B$  be the vector of basic components of the coefficient vector of the objective function. Finally, let  $z_p = f_B^{T-1} \alpha_p, d_p = b^{-1} \alpha_p, p = 0, \dots, k + 1$ . Since we have

$$\begin{pmatrix} 1 & f_B^T \\ 0 & B \end{pmatrix}^{-1} \begin{pmatrix} f_p \\ a_p \end{pmatrix} = \begin{pmatrix} 1 - f_B^T B^{-1} \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} f_p \\ a_p \end{pmatrix} = \begin{pmatrix} f_p - z_p \\ d_p \end{pmatrix},$$

it follows that

$$\begin{pmatrix} 1 & f_B^T \\ 0 & B \end{pmatrix} \begin{pmatrix} f_p - z_p \\ d_p \end{pmatrix} = \begin{pmatrix} f_p \\ a_p \end{pmatrix}$$

and by Cramer's rule we get

$$f_p - f_B^T B^{-1} a_p = \frac{1}{|B|} \begin{vmatrix} f_p & f_i & f_j \\ a_p & a_i & a_j \end{vmatrix}, \tag{2.1}$$

where  $|B|$  designates the determinant of  $B$ . For this we have  $|B| = \begin{vmatrix} z_i & z_j \\ 1 & 1 \end{vmatrix} = z_i - z_j < 0$ . On the other hand, given  $i < j < \ell$ , we have

$$\begin{vmatrix} f_i & f_j & f_l \\ a_i & a_j & a_l \end{vmatrix} = -(z_l - z_j)(z_j - z_i) \left( \frac{f_l - f_j}{z_l - z_j} - \frac{f_j - f_i}{z_j - z_i} \right) < 0, \tag{2.2}$$

because, by the convexity of the function  $f$ , the difference  $(f_l - f_j) (z_l - z)^{-1} - (f_j - f_i) (z_i - t_i)^{-1}$  is positive. Hence, the value  $f_p - f_B^T B^{-1} a_p$  in(2.1) is always different from 0 and is positive for every nonbasic  $p$  if and only if  $j = i + 1$ . In fact, by (2.2) we get that if the column containing  $f_p, a_p$  in (2.1) is in its right place, allowing

for an increasing order of the subscripts, then the result is a positive number. This excludes the case  $j > i + 1$  because otherwise the choice  $p = i + 1$  would produce a negative value in (2.1).

The structure of the matrix (1.13) implies that every feasible basis of problem (1.12) has at least one but at most two columns from any block  $i$  ( $2 \leq i \leq r$ ). The above theorem, on the other hand, implies that in the latter case the two columns must be consecutive.

Below we present our algorithm to solve Problem (1.12). We assume that  $A \neq 0$ ,  $T \neq 0$  and designate the columns of  $A$  and  $T$  by  $a_1, \dots, a_n$  and  $t_1, \dots, t_n$ , respectively. The columns of the matrix (1.13) will be designated by  $h_1, \dots, h_n, h_{10}, \dots, h_{1k_1+1}, \dots, h_{r0}, \dots, h_{rk_r+1}$ , respectively. We will say that the matrix (1.13) is subdivided into  $r + 1$  blocks. The first block consists of the first  $n$  columns and the  $i + 1$ st block consists of the columns  $h_{i0}, \dots, h_{ik_i+1}$ .

Our solution of problem (1.12) applies to the general case, i.e., we do not assume that the coefficients  $c_{ij}$  are those, derived for the stochastic programming problem (1.6). We assume, however, that for every  $1 \leq i \leq r$  the discrete function

$$f_i(z_{ij}) = c_{ij}, \quad j = 0, \dots, k_i + 1$$

is convex, in other words, its second order divided differences are positive. Let us introduce the notations for the first and second order divided differences, respectively:

$$\begin{aligned} \frac{c_{ij+1} - c_{ij}}{z_{ij+1} - z_{ij}} &= [z_{ij}, z_{ij+1}]c_i, & j = 0, \dots, k_i \\ \frac{[z_{ij}, z_{ij+1}] - [z_{ij-1}, z_{ij}]}{z_{ij+1} - z_{ij-1}} &= [z_{ij-1}, z_{ij}, z_{ij+1}]c_i, & j = 1, \dots, k_i. \end{aligned} \quad (2.3)$$

In the case of the stochastic programming problem (1.6) we have the following equalities

$$\begin{aligned} [z_{ij}, z_{ij+1}]c_i &= -q_i^+ + \frac{q_i^+ + q_i^-}{z_{ij+1} - z_{ij}} \int_{z_{ij}}^{z_{ij+1}} F_i(z) dz \\ &= -q_i^+ + (q_i^+ + q_i^-)(p_{i0} + \dots + p_{ij}), \end{aligned} \quad (2.4)$$

$$[z_{ij-1}, z_{ij}, z_{ij+1}]c_i = \frac{1}{z_{ij+1} - z_{ij-1}} (q_i^+ + q_i^-) p_{ij}. \quad (2.5)$$

In steps 0,1 an initial dual feasible basis is constructed, whereas in the other steps we perform iterations according to the dual method of linear programming.

*Step 0:* Select two consecutive vectors out of the last  $r$  blocks of the matrix (1.13). Let  $j_1, j_1 + 1, \dots, j_r, j_r + 1$  be the subscripts of the selected vectors in the 2nd, ...,  $r + 1$ st blocks, respectively. Solve the systems of linear equations

$$\begin{aligned} -z_{1j_1} v_1 + w_1 &= c_{1j_1}, \\ -z_{1j_1+1} v_1 + w_1 &= c_{1j_1+1}, \\ &\vdots \\ -z_{rj_r} v_r + w_r &= c_{rj_r}, \\ -z_{rj_r+1} v_r + w_r &= c_{rj_r+1} \end{aligned}$$

and define the vectors  $v^T = (v_1, \dots, v_r)$ ,  $w^T = (w_1, \dots, w_r)$ .

*Step 1:* Solve the linear programming problem

$$\begin{aligned} \text{minimize} \quad & \{(c_1 - v^T t_1)x_1 + \dots + (c_n - v^T t_n)x_n\} \\ \text{subject to} \quad & a_1 x_1 + \dots + a_n x_n = b, \quad x_1 \geq 0, \dots, x_n \geq 0, \end{aligned} \tag{2.6}$$

by a method which provides us with a primal-dual feasible basis. Let  $B$  be this optimal basis and let  $d$  be a dual vector corresponding to this optimal basis  $B$ , i.e., any solution of the equation  $d^T B = c_B^T - v^T T_B$ , where  $c_B$  and  $T_B$  are those parts of  $c$  and  $T$ , respectively, which correspond to the basis subscripts. If  $A$  has full rank then  $B$  is a square matrix and  $y$  is uniquely determined.

We have obtained a dual feasible basis for problem (1.12). It consists of those vectors that trace out  $B$  from  $A$  and  $T_B$  from  $T$ , in the first block, furthermore the previously selected consecutive pairs from the other blocks. If  $A$  has full rank and the optimal basis in problem (2.6) consists of the vectors  $a_1, \dots, a_m$ , then for problem (1.12) we have the dual feasible basis

$$\left( \begin{array}{cccccc} a_{11} & \cdots & a_{1m} & & & \\ & & \cdots & & & \\ a_{m1} & \cdots & a_{mm} & & & \\ t_{11} & \cdots & t_{1m} & -z_{1j_1} & -z_{1j_1+1} & \\ & & \cdots & & & \ddots \\ t_{r1} & \cdots & t_{rm} & & & -z_{rj_r} & -z_{rj_r+1} \\ 0 & \cdots & 0 & 1 & 1 & & \\ & & \cdots & & & & \ddots \\ 0 & \cdots & 0 & & & 1 & 1 \end{array} \right) \tag{2.7}$$



In the later steps of the procedure the basis structure may change so that out of the variables  $x_1, \dots, x_n$ , there are  $m + s$  in the basis and out of  $s$  of the variable pairs  $z_{1j_1}, z_{1j_1 + 1}; \dots; z_{rj_r}, z_{rj_r + 1}$ , only one is in the basis, where  $0 \leq s \leq r$ . Go to step 2.

*Step 2:* Let  $S$  designate the set of those row subscripts of  $T$ , corresponding to which only one  $-z_{ij}$  is in the basic column and let it be  $-z_{iji}$ . We designate furthermore by  $Q$  the set  $\{1, \dots, r\} - S$  and let  $-z_{iji}, -z_{iji+1}$  be those elements in row  $i \in Q$  which are in basic columns.

Determine the basic components of the basic solution corresponding to the basis. Let  $x_B$  be the vector consisting of the basic components out of  $x_1, \dots, x_n$  ( $x_B$  may have more components than the rank of  $A$ ), determined by the equations

$$\begin{aligned} A_B x_B &= b \\ T_{iB} x_B &= z_{iji}, \quad i \in S, \end{aligned} \quad (2.8)$$

where  $T_{iB}$  is that part of  $T_i$  (the row of  $T$ ) which corresponds to  $x_B$ .

As regards the basic components of  $\{\lambda_{ij}\}$  we easily find that

$$\begin{aligned} \lambda_{iji} &= 1 && \text{for } i \in S, \\ \lambda_{iji} &= \frac{z_{iji+1} - T_{iB} x_B}{z_{iji+1} - z_{iji}}, && i \in Q, \\ \lambda_{iji+1} &= \frac{T_{iB} x_B - z_{iji}}{z_{iji+1} - z_{iji}}, \end{aligned}$$

*Step 3:* Test for primal feasibility:  $x_B \geq 0$ ,  $\lambda_{iji} \geq 0$ ,  $\lambda_{iji+1} \geq 0$ ,  $\notin Q$ . If all these inequalities are satisfied then stop, the basis is optimal. If it is not the case then choose any basic component which is negative and let the corresponding vector leave the basis. Go to step 4.

*Step 4:* Update nonbasic columns, i.e., represent them as linear combinations of the basic vectors and compute the corresponding reduced costs that we designate by the symbols  $\bar{c}_p, \bar{c}_{ij}, j = 0, \dots, k_i + 1, i = 1, \dots, r$ . Not all nonbasic vectors have to be updated, just those which may enter the basis. This depends, however, on the outgoing vector, as described below.

I. Let a column from the first block leave the basis. Then either a column containing  $-z_{iji-1}$  (if  $j_i > 0$ ) or  $-z_{iji+1}$  (if  $j_i < k_i + 1$ ),  $\notin S$  or a column from the first block may enter.

Ia. To update the column containing  $-z_{iji-1}$ , where  $j_i > 0$  and  $\notin S$ , first we represent a part of it by solving the equation with respect to  $u$ :

$$\begin{aligned} A_B u &= 0 \\ T_{iB} u - z_{ij_i} &= -z_{ij_i-1}, \\ T_{hB} u &= 0, \quad h \in S, \quad h \neq i. \end{aligned}$$

Let  $u_i$  be the solution of the equation

$$\begin{aligned} A_B u_i &= 0 \\ T_{iB} u_i &= 1, \\ T_{hB} u_i &= 0, \quad h \in S, \quad h \neq i. \end{aligned}$$

Then we have  $u = u_i(z_{ij_i} - z_{ij_i-1})$ . To update the remaining part of the column of  $z_{ij_i-1}$ , we solve the equations for  $d_{h1}, d_{h2}$ :

$$\begin{aligned} T_{hB} u - d_{h1} z_{hj_h} - d_{h2} z_{hj_{h+1}} &= 0, \\ d_{h1+} \quad \quad \quad d_{h2} &= 0, \quad h \in Q \end{aligned}$$

and obtain

$$d_{h2} = -d_{h1} = \frac{T_{hB} u}{z_{hj_{h+1}} - z_{hj_h}}, \quad h \in Q. \tag{2.9}$$

For the reduced cost  $\bar{c}_{j_i-1}$  we derive

$$\begin{aligned} \bar{c}_{j_i-1} &= c_B^T u + c_{ij_i} + (z_{ij_i} - z_{ij_i-1}) \sum_{h \in Q} [z_{hj_h}, z_{hj_{h+1}}] c_h T_{hB} u_i - c_{ij_i-1} \\ &= (z_{ij_i} - z_{ij_i-1}) \\ &\quad \times \left( c_B^T u_i + [z_{ij_i-1}, z_{ij_i}] c_i + \sum_{h \in Q} [z_{hj_h}, z_{hj_{h+1}}] c_h T_{hB} u_i \right) \end{aligned} \tag{2.10}$$

Ib. To update the column of  $-z_{ij_i+1}$ , where  $j_i < k_i + 1$  and  $i \in S$ , the same reasoning can be used, the only difference is that now we define  $u = u_i(z_{ij_i} - z_{ij_i+1})$ , while  $u_i$  is the same as before. The coefficients (2.9) change accordingly. The reduced cost  $\bar{c}_{ij_i+1}$  equals

$$\begin{aligned} \bar{c}_{ij_i+1} &= -(z_{ij_i+1} - z_{ij_i}) (c_B^T u_i + [z_{ij_i}, z_{ij_i+1}] c_i \\ &\quad + \sum_{h \in Q} [z_{hj_h}, z_{hj_{h+1}}] c_h T_{hB} u_i). \end{aligned} \tag{2.11}$$

Ic. To update that column from the first block which traces out  $a_p$  from  $A$ , we solve the equations for  $d_p$ :

$$A_B d_p = a_p$$

$$T_{iB} d_p = t_{ip}, \quad i \in S,$$

furthermore, the equations for  $d_{h1}$ ,  $d_{h2}$ :

$$\begin{aligned} T_{hB} d_p - \frac{d_{h1} z_{hj_h} - d_{h2} z_{hj_{h+1}}}{d_{h1} + d_{h2}} &= t_{hp} \\ &= 0, \quad h \in Q. \end{aligned}$$

For the latters we obtain

$$\begin{aligned} d_{h1} &= \frac{t_{hp} - T_{hB} d_p}{z_{hj_{h+1}} - z_{hj_h}} \\ d_{h2} &= \frac{T_{hB} d_p - t_{hp}}{z_{hj_{h+1}} - z_{hj_h}}, \end{aligned} \quad h \in Q \quad (2.12)$$

For the reduced cost  $\bar{c}_p$  we derive

$$\bar{c}_p = c_B^T d_p + \sum_{h \in Q} [z_{hj_h}, z_{hj_{h+1}}] c_h (T_{hB} d_p - t_{hp}) - c_p. \quad (2.13)$$

II. Assume now that one of the columns  $-z_{qj_q}, -z_{qj_q+1}, q \in Q$ , leaves the basis.

Ia. To update the column of  $-z_{qj_q-1}$ , where  $j_q > 0$ , we solve the equations for  $d_{q1}, d_{q2}$ :

$$-d_{q1} z_{qj_q} - d_{q2} z_{qj_q+1} = -z_{qj_q} - 1, \quad d_{q1} + d_{q2} = 1$$

which gives

$$d_{q1} = \frac{z_{qj_q+1} - z_{qj_q-1}}{z_{qj_q+1} - z_{qj_q}}, \quad d_{q2} = \frac{z_{qj_q-1} - z_{qj_q}}{z_{qj_q+1} - z_{qj_q}}, \quad (2.14)$$

From here we derive the reduced cost

$$\begin{aligned} \bar{c}_{qj_q-1} &= c_{qj_q+1} d_{q2} + c_{qj_q} d_{q1} - c_{qj_q-1} \\ &= -(z_{qj_q+1} - z_{qj_q-1})(z_{qj_q} - z_{qj_q-1}) [z_{qj_q-1}, z_{qj_q}, z_{qj_q+1}] c_q. \end{aligned} \quad (2.15)$$

Ib. To update the column of  $-z_{qj_q+2}$ , where  $j_q < k_q$ , we solve the equations

$$\begin{aligned} -d_{q1}z_{qj_q} - d_{q2}z_{qj_q+1} &= -z_{qj_q+2}, \\ d_{q1} + d_{q2} &= 1 \end{aligned}$$

which gives

$$d_{q1} = \frac{z_{qj_q+1} - z_{qj_q+2}}{z_{qj_q+1} - z_{qj_q}}, \quad d_{q2} = \frac{z_{qj_q+2} - z_{qj_q}}{z_{qj_q+1} - z_{qj_q}}, \tag{2.16}$$

From here we derive the reduced cost

$$\begin{aligned} \bar{c}_{qj_q+2} &= c_{qj_q+1}d_{q2} + c_{qj_q}d_{q1} - c_{qj_q+2} \\ &= -(z_{qj_q+2} - z_{qj_q+1})(z_{qj_q+2} - z_{qj_q})[z_{qj_q}, z_{qj_q+1}, z_{qj_q+2}]c_q. \end{aligned} \tag{2.17}$$

IIc. The update formulas and the reduced costs concerning the columns of  $-z_{iji-1}$  (if  $j_i > 0$ ),  $-z_{ji+1}$  (if  $j_i < k_i + 1$ ),  $\notin S$  and the nonbasic columns in the first block are given in Ia, b, c.

Step 5: Determine the vector that enters the basis. The two cases handled below are the same as those mentioned in the description of step 4.

I. Let the outgoing vector be the  $l$ th nonbasic vector from the first block. Designate by  $u(l)$  and  $u_i(l)$  the  $l$ th components of the vectors  $u$  and  $u_i$ , respectively. If  $u$  is defined concerning  $-z_{iji-1}$ , then  $u(l) = u_i(l) (z_{iji} - z_{iji-1})$  and if  $u$  is defined concerning  $-z_{iji+1}$ , then  $u(l) = u_i(l) (z_{iji} - z_{iji+1})$ .

These have to be compared with the reduced costs (2.10) and (2.11), respectively. If, on the other hand, we look at a nonbasic column in the first block, the subscript of which is  $p$ , say (i.e., it is the column intersecting  $A$  at  $a_p$ ), then the  $l$ th component of  $d_p$ , that we designate by  $d_p(l)$ , has to be compared with

$\bar{c}_p$  in (2.13). If the matrix  $\begin{pmatrix} A_B \\ T_{SB} \end{pmatrix}$  is nonsingular then

$$d_p = \begin{pmatrix} A_B \\ T_{SB} \end{pmatrix}^{-1} \begin{pmatrix} a_p \\ t_{Sp} \end{pmatrix}.$$

Thus the incoming vector is determined by taking the minimum of the following three minima (in the first two lines  $z_{iji} - z_{iji-1}$  and  $z_{iji} - z_{iji+1}$ , respectively, are already cancelled):

$$\min_{i \in S, j_i < 0, u_i(l) > 0} \left\{ \frac{1}{u_i(l)} \left( c_B^T u_i + [z_{iji-1}, z_{iji}]c_i + \sum_{h \in Q} [z_{hj_h}, z_{hj_h+1}]c_h T_{hB} u_i \right) \right\},$$

$$\min_{i \in S, j_i < k_i, u_i(l) > 0} \left\{ \frac{1}{u_i(l)} \left( c_B^T u_i + [z_{ij_i}, z_{ij_i+1}] c_i + \sum_{h \in Q} [z_{hj_h}, z_{hj_h+1}] c_h T_{hB} u_i \right) \right\},$$

$$\min_{d_p(l) < 0} \left\{ \frac{1}{d_p(l)} \left( c_B^T d_p + \sum_{h \in Q} [z_{hj_h}, z_{hj_h+1}] c_h (T_{hB} d_p - t_{hp}) - c_p \right) \right\}. \quad (2.18).$$

If the minimum is attained in the first line at  $i$ , then the column of  $-z_{ij_i-1}$  is the incoming one.

If the minimum is attained in the second line at  $i$ , then the column of  $-z_{ij_i+1}$  is the incoming one.

If the minimum is attained in the third line at  $p$ , then the column of  $a_p$  is the incoming one.

II. Let the outgoing column be either the column of  $-z_{qj_q}$  or the column of  $-z_{qj_q+1}$ , where  $q \in Q$ .

IIa. If it is the column of  $-z_{qj_q}$  then the column of  $-z_{qj_q+2}$  may enter, provided  $j_q < k_q$ . The other candidates can be subdivided into three disjoint groups. The first group is formed by the nonbasic columns of the first block. The second (third) group is formed by the columns of  $-z_{ij_i-1} (-z_{ij_i+1})$ ,  $i \in S$ . We take the minimum of the fractions of the reduced costs and the coefficients of the outgoing vector, in the representation of the candidates in terms of the basic vectors, restricting ourselves to negative coefficients, as prescribed by the dual method. The coefficient that multiplies the column of  $-z_{qj_q}$  in the representation of  $-z_{qj_q+2}$  is negative and is given by (2.16). We take the fraction of  $-\bar{c}_{qj_q+2}$  and this number.

To determine the incoming vector we have to take the minimum of the thus obtained four numbers. Since all of them contain, as factor, the difference  $-z_{qj_q+1} - z_{qj_q}$ , we can cancel it everywhere and obtain the following

$$(z_{qj_q+2} - z_{qj_q}) [z_{qj_q}, z_{qj_q+1}, z_{qj_q+2}] c_q,$$

$$\min_{T_{qB} d_p > t_{qp}} \frac{\bar{c}_p}{t_{qp} - T_{qB} d_p},$$

$$\min_{i \in S, j_i > 0, T_{qB} u_i > 0} \frac{\bar{c}_{ij_i-1}}{-(z_{ij_i} - z_{ij_i-1}) T_{qB} u_i}, \quad (2.19)$$

$$\min_{i \in S, j_i < k_i, T_{qB} u_i < 0} \frac{\bar{c}_{ij_i-1}}{(z_{ij_i+1} - z_{ij_i}) T_{qB} u_i},$$

where the reduced costs are given by (2.10), (2.11) and (2.13). Some of the lines in (2.19) may be absent. E.g., the first line is absent if  $j_q = k_q$ .

If the minimum of the four numbers in (2.19) is attained in the first line then the column of  $-z_{qj_q+2}$  comes in. If it is attained in the second line at  $p$  then the co-



We have checked that  $0 \leq T_i x \leq 80$ ,  $i = 1, 2$  whenever  $Ax = b$ ,  $x \geq 0$ . For the objective function coefficients  $c_{ij}$  we obtain  $c_{10} = c_{20} = c_{11} = c_{21} = 0$ ,  $c_{1i} = c_{2i} = 10/7 (1 + 2 + \dots + i - 1)$ ,  $i = 2, 3, 4, 5, 6, 7, 8$ .

To describe the results in the subsequent steps we number the columns of the matrix (1.13) from 2 through 25. The initially chosen pairs from blocks 2 and 3 have subscripts 11, 12, 20, 21 (Step 0). Corresponding to these we have obtained the vectors subscripted by 3 and 4, from the first block (Step 1). The subsequent dual feasible bases are

		Block 1	Block 2	Block 3
	Initial	3, 4	11, 12	20, 21
Iteration	1	3, 4	11, 12	21, 22
	2	3, 4	11, 12	22, 23
	3	0, 3, 4	11, 12	23
	4	0, 3, 4	12	23, 24
	5	0, 3, 4	12, 13	24
	6	0, 3, 4	13	24, 25
	7	0, 3, 4	13, 14	25
	8	0, 3, 4	14	25, 26
	9	0, 3, 4	14, 15	26
	10	0, 3, 4	15, 16	26
	11	0, 3, 4, 5	16	26
	12	0, 3, 5	16, 17	26

The optimal solution is

$$x_1 = 2.45386571, \quad x_4 = 2.954389113, \quad 1.398684014$$

$$\lambda_{1,8} = 0.1025123374, \quad \lambda_{1,9} = 0.8974876626, \quad \lambda_{2,9} = 1.$$

### 3 Applications for Deterministic Problems

Consider problem (1.3), where  $\xi$  is non-random now and assume that the constraints in the equality  $Tx = \xi$  are not required to be satisfied at any price. Instead, we define a cost of deviation in its  $i$ th row by taking  $f_i(T_i x - \xi_i)$  and then formulate the following problem

$$\begin{aligned} & \text{minimize } \left\{ c^T x + \sum_{i=1}^r f_i(T_i x - \xi_i) \right\} \\ & \text{subject to } Ax = b, \quad x \geq 0, \end{aligned} \quad (3.1)$$

where  $f_1, \dots, f_r$  are piecewise linear convex functions, defined in some intervals.

An important special case of problem (3.1) is the following

$$\begin{aligned} & \text{minimize } \left\{ c^T x + \sum_{i=1}^r \left| \sum_{j=1}^r t_{ij} x_j - g_i \right| \right\} \\ & \text{subject to } Ax = b, \quad x \geq 0. \end{aligned} \quad (3.2)$$

Suppose that there exist real numbers  $z_{i0}, z_{i2}$  such that  $z_{i0} < g_i < z_{i2}$  and every  $x$  which satisfies  $Ax = b, x \geq 0$ , automatically satisfies

$$z_{i0} \leq T_i x \leq z_{i2}, \quad i = 1, \dots, r. \quad (3.3)$$

Defining the functions  $f_i(x)$  so that (see figure 1)

$$f_i(x) = \begin{cases} x - g_i & \text{if } g_i \leq x \leq z_{i2}, \\ g_i - x & \text{if } z_{i0} \leq x \leq g_i, \end{cases}$$

for  $i = 1, \dots, r$ , we see that problem (3.2) is in fact a special case of problem (3.1). Let us introduce the notations  $z_{i1} = g_i, i = 1, \dots, r$ . Then problem (3.2) is equivalent to problem (1.12), where  $k_1 = \dots = k_r = 1$  and  $c_{i0} = f_i(z_{i0}), c_{i1} = 0, c_{i2} = f_i(z_{i2}), i = 1, \dots, r$ . In other words, problem (3.2) is the stochastic programming problem (1.4), where the random variable  $\xi_i$  has only one possible value  $g_i$  and  $q_i^+ = q_i^- = 1, i = 1, \dots, r$  and it is assumed that (3.3) holds for every  $x$  satisfying  $Ax = b, x \geq 0$ .

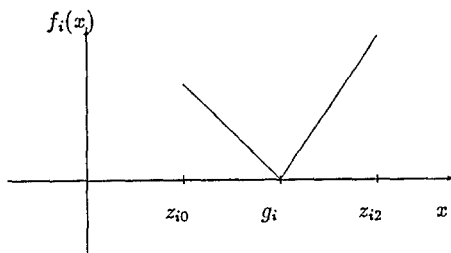


Fig. 1. Graph of the function  $f_i(x)$



#### 4 Simultaneous Use of Penalties and Probabilistic Constraint

In this section we outline an algorithm for the solution of problem (1.2), where we assume that each random variable  $x_i$  has a finite number of possible values which are  $z_{i1} < \dots < z_{ik_i}$ ,  $1 \leq i \leq r$ . We also assume that there exist numbers  $z_{i0}$ ,  $z_{ik_i+1}$  such that  $z_{i0} < z_{i1}$ ,  $z_{ik_i} < z_{ik_i+1}$  and we have

$$z_{i0} \leq T_i x \leq z_{ik_i+1}, \quad i = 1, \dots, r$$

for every  $x$  satisfying  $Ax = b$ ,  $x \geq 0$ . Let  $F_1, \dots, F_r, F$  designate the probability distribution functions of  $\xi_1, \dots, \xi_r, \xi$ , respectively, i.e.

$$\begin{aligned} F_i(z) &= P(\xi_i \leq z), & z \in \mathbf{R}^1, \quad i = 1, \dots, r, \\ F(z) &= P(\xi \leq z), & z \in \mathbf{R}^r. \end{aligned}$$

The vectors  $(z_{1j_1}, \dots, z_{rj_r})$ , where  $1 \leq j_i \leq k_i$ ,  $i = 1, \dots, r$  will be considered the set of possible values of the random vector  $\xi$ . Due to stochastic dependency, some of these may have probability 0. We will briefly designate one possible value of  $\xi$  by  $z^{(j)}$ .

We say that  $z^{(j)}$  is a  $p$  level efficient point (PLEP) of the probability distribution of  $\xi$  if  $F(z^{(j)}) \geq p$  if there is no possible value  $z^{(l)}$  of  $\xi$  such that

$$z^{(l)} \leq z^{(j)}, \quad z^{(l)} \neq z^{(j)}, \quad F(z^{(l)}) \geq p.$$

Let  $z^{(j)}$ ,  $j \in E$  be the set of PLEP's. Then the problem (1.2) is equivalent to problem (1.12), where, in addition to the constraints, we have also the constraint

$$Tx \geq z^{(j)} \quad \text{holds for at least one } j \in E. \quad (4.1)$$

In fact, problem (1.2) is equivalent to problem (1.12) supplement by the additional constraint

$$Tx \geq z^{(j)} \quad \text{holds for at least one } j \text{ such that } F(z^{(j)}) \geq p. \quad (4.2)$$

However, among all possible values satisfying  $F(z^{(j)}) \geq p$ , it is enough to take into account only those which are PLEP's because the set of feasible solutions of the problem is the same, no matter if (4.1) or (4.2) is used as the additional constraint.

Having all PLEP's, we reformulate the constraint (4.1) so that

$$Tx \in H, \quad (4.3)$$

where

$$H = \bigcup_{l \in E} H_l, \quad H_l = \{y | y \geq z^{(l)}\}$$

and solve subsequently problems of the type (1.12), supplement by the constraints

$$Tx \in H_{l_j} \setminus \bigcup_{i=1}^{j-1} H_{l_i}, \quad j \geq 1. \quad (4.4)$$

This way all possible elements in  $H$  will be allowed for  $Tx$  and an optimal solution to problem (1.2) will be obtained. The algorithm can be summarized in the following manner.

*Step 0:* Enumerate all PLEP's. This is very easy to do if the code is written in APL language which handles multidimensional arrays. In fact, applying the iterated  $+\setminus$  addition for the multidimensional array containing the probabilities of the possible values of  $\xi$ , we obtain the probability distribution function  $F$  of  $\xi$ . Then the operation  $+\setminus \dots +\setminus F \geq p$ , where there are as many  $+\setminus$  additions as the dimensionality of the array, produces an array where exactly those positions contain 1's which correspond to PLEP's. We only have to find the corresponding possible values of  $\xi$  and the enumeration is done.

Initialize  $E^{(c)}$ ,  $H^{(c)}$  and  $x^{(c)}$  as  $E^{(c)} = E$ ,  $H^{(c)} = \{y | y \geq z^{(l)}\}$  and  $x^{(c)} = 0$ , where  $l$  is arbitrarily chosen and the letter  $c$  refers to the word "current". Assuming  $x = 0$  is not a feasible solution of problem(1.12), we assign to this vector, following a generally accepted convention, the objective function value  $+\infty$ .

*Step 1:* Solve problem (1.12) so that we prescribe the additional constraint

$$Tx \in H^{(c)} \quad (4.5)$$

and designate by  $x_{opt}$  any optimal solution. When we first execute Step 1, then (4.5) means the constraint  $Tx \geq z^{(l)}$ . In later applications of step 1,  $H^{(c)}$  is the union of a finite number of rectangular sets, by (4.6). To solve Problem (1.12) with the additional constraint (4.5) means that we solve as many linear programming problems as the number of rectangular sets and the LP corresponding to a rectangular set is obtained so that in problem (1.12) only those  $z_{ij}$  values are allowed which are elements of the rectangular set. The  $c_{ij}$  coefficients remain unchanged but only those are used which correspond to non-deleted  $z_{ij}$ . The vector  $x_{opt}$  is defined as that optimal solution which produces the smallest optimum value among all optimum values of the above-mentioned LP's.

Step 2: Check if the objective function value of  $x_{opt}$  is smaller than that of  $x^{(c)}$ . If yes, then choose any  $l \in E^{(c)}$ , update  $H^{(c)}, x^{(c)}$  as follows

$$\begin{aligned}
 E^{(c)} &:= E^{(c)} \setminus \{l\}, \\
 H^{(c)} &:= H_l \setminus H^{(c)}, \\
 x^{(c)} &:= x_{opt}
 \end{aligned}
 \tag{4.6}$$

and go to step 1. Otherwise update only  $E^{(c)}$  and  $H^{(c)}$ , following the rule in (4.6) and go to step 1.

If cycling is somehow excluded (e.g., by applying Bland’s rule) in each LP, then the algorithm terminates in a finite number of steps, by reaching an optimal solution. This happens when the updated  $E^{(c)}$  in (4.6) becomes empty and steps 1 and 2 are executed for the last time.

The algorithm can considerably be simplified at the expense of some superfluous computation if instead of (4.4) we simply write  $Tx \in H_{lj}, j \geq 1$ . In this case step 1 and step 2 can be combined into one step where we solve problem (1.12) with the additional constraint  $Tx \in H_{lj}$ . This is iterated until all sets  $H_{lj}$  have been investigated.

This variant is supported by the fact that if the probability level  $p$  in the probabilistic constraint is relatively large, e.g.,  $p = 0.8$ , then in many cases there will be a few PLEP’s only. To see an example, let  $r = 4$  and assume that the components of  $x_j$  are independent, each can take the possible values 1, 2, 3, 4, 5, 6, 7, 8 with the same probability 1/8. If  $p = 0.8$  then there are 4 PLEP’s which are the following

7	8	8	8
8.	7	8	8
8	8	7	8
8	8	8	7

Thus if we run four times the algorithm presented in section 2, the optimal solution is obtained.

*Computational experience:* A code in APL language has been prepared for the solution of problem (1.12) and test problems were run on an IBM PC AT and a VAX 8650 mainframe. Problem sizes ranged up to  $m + r = 60, n = 200$  and  $k_i = 1,000$ . In this first variant of the code the inverse of the “working basis” (as Wets has called it)

$$\begin{pmatrix} A_B \\ T_{SB} \end{pmatrix}$$

is computed by the APL matrix inversion device at any iteration. The matrices  $A$  and  $T$  were randomly chosen. The running times depended very much on the choices of the initial consecutive column pairs in the second, ...,  $r + 1$ st blocks. In the case of  $m = 5$ ,  $r = 4$ ,  $n = 100$ ,  $k_1 = k_2 = k_3 = k_4 = 1,000$ , the solutions have been obtained instantly or in at most 35 minutes on the AT, while in case of  $m = 50$ ,  $r = 10$ ,  $n = 200$ ,  $k_1 = \dots = k_{10} = 10$ , the solution times varied between 5 and 40 minutes, on the mainframe. The sizes of the LP's in these two cases are  $13 \times 4108$  and  $70 \times 320$ , respectively. These wide ranges of the solution times suggest that the problem should be solved in an interactive way so that by observing the changes of the numbers  $j_i$ , in the course of the solution, we stop and restart the run by choosing larger or smaller initial  $j_i$  values, in agreement with the directions of their changes. Using this, the solution times remained in the lower sections of the above ranges.

The speed of the execution was improved also by the insertion of "primal steps" as follows. At the end of the execution of step 2, assuming the optimum has not been reached yet, new initial consecutive column pairs are defined (the columns of  $-z_{ij_i}$ ,  $-z_{ij_i+1}$ ,  $i = 1, \dots, r$ ) so that the equitations for the  $\lambda$ 's:

$$\begin{aligned} T_i x - \lambda_{ij_i} z_{ij_i} - \lambda_{ij_i+1} z_{ij_i+1} &= 0, \\ \lambda_{ij_i} + \lambda_{ij_i+1} &= 1 \end{aligned}$$

produce nonnegative solutions, where  $x$  is the optimal solution of problem (2.6), obtained by the execution of step 1. We then restart the solution with these  $j_1, \dots, j_r$ .

## References

- E.M.L. Beale (1955) On Minimizing a Convex Function Subject to Linear Inequalities, *Journal of the Royal Statistical Society Ser. B*, 17, 173–184
- R.G. Bland (1977) New Finite Pivoting Rules for the Simplex Method, *Mathematics of Operations Research* 2, 103–107
- A. Charnes, W.W. Cooper and G.H. Symonds (1985) Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil Production, *Management Science* 4, 236–263
- G.B. Dantzig (1955) Linear Programming under Uncertainty, *Management Science* 1, 197–206
- C.E. Lemke (1954) The Dual Method for Solving the Linear Programming Problem, *Naval Research Logistic Quarterly* 1, 36–47.
- K. Murty (1968) Linear Programming under Uncertainty: A Basic Property of the Optimal Solution, *Zeitschrift für Wahrscheinlichkeitstheorie ver. Geb.* 10, 284–288
- A. Prékopa (1973) Contributions to the Theory of Stochastic Programming, *Mathematical Programming* 4, 202–221

A. Prékopa (1990) The Discrete Moment Problem and Linear Programming. To appear in  
Discrete Applied Mathematics

R.J.B. Wets (1983) Solving Stochastic Programs with Simple Recourse, *Stochastics* 10, 219–242

Received July 1988

Revised version received September 1989