Fourier's Method of Linear Programming and Its Dual

Author(s): H. P. Williams

Source: *The American Mathematical Monthly*, Nov., 1986, Vol. 93, No. 9 (Nov., 1986), pp. 681-695

Published by: Taylor & Francis, Ltd. on behalf of the Mathematical Association of America

Stable URL: https://www.jstor.org/stable/2322281

FOURIER'S METHOD OF LINEAR PROGRAMMING AND ITS DUAL

H. P. WILLIAMS

Faculty of Mathematical Studies, University of Southampton, Southampton SO9 5NH, U.K.

Introduction. There has been widespread popular interest in recent years in suggested improved methods for solving Linear Programming (LP) models. In 1977 Shor [**13**] described a new algorithm for LP. Khachian [**7**] modified this algorithm in order to prove that the number of computational steps was, in the worst case, bounded by a polynomial function of the size of the data. This method has become known as the *Ellipsoid Method*. It has in practice been disappointing in experimental computational performance. In 1984 Karmarker [**6**] produced another algorithm which was also "polynomially bounded" with spectacular practical computational claims. Controversy continues as to whether Karmarker's method will displace the Simplex Method. *The Simplex Method* was invented by Dantzig in 1948 and is well explained in Dantzig [**1**]. Although it is not polynomial in the worst case it has proved a remarkably powerful method in practice and its major extension, the *Revised Simplex Method*, is the method used in all commercial systems.

The reason for the widespread popular interest (both Khachian and Karmarker's methods received headlines in the national press) is that LP models are among the most widely used type of Mathematical Model. Applications of LP arise in Manufacturing, Distribution, Finance, Agriculture, Health, Energy and general Resource Planning. A practical discussion of application areas is contained in Williams [**16**].

In this article we show that, predating all these methods, a method discovered by Fourier in 1826 for manipulating linear inequalities can be adapted to Solving Linear Programming models. The theoretical insight given by this method is demonstrated as well as its clear geometrical interpretation. By considering the dual of a linear programming model it is shown how the method gives rise to a dual method. This dual method generates all extreme solutions (including the optimal solution) to a linear programme. Therefore if a polytope is defined in terms of its facets the dual of Fourier's method provides a method of obtaining all vertices.

An LP model consists of variables (e.g., x_1, $x_2,\ldots,$etc.) contained in a linear expression known as an *objective function*. Values are sought for the variables which *maximise* or *minimise* the objective function subject to *constraints*. These constraints are themselves linear expressions which must be either less-than-or-equal to ($\leqslant$), greater-than-or-equal to ($\geqslant$) or equal to ($=$) some specified value. For example, the following is a small LP model.

Find values for $x_1, x_2,\ldots$ among the real numbers so as to:

$$\text{Maximise} \quad -4x_1 + 5x_2 + 3x_3$$

$$\mathrm{P} \quad \begin{array}{c}\text{subject to}\\ \text{constraints}\end{array} \begin{cases} -x_1 + x_2 - x_3 \leqslant 2 \\ \;\; x_1 + x_2 + 2x_3 \leqslant 3 \\ x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$$

It is usually the case that the variables are restricted to be non-negative as in the example above.

Paul Williams was born in Cornwall, England, in 1943 and educated at Redruth Grammar School. He graduated in Mathematics from Cambridge University as well as obtaining an Athletics Blue as a hurdler. This was followed by a Ph.D. in Mathematical Logic from Leicester University. It was during this time that he first "discovered" the procedure described in this paper only to find out, some years later, about Fourier's work.

He worked for IBM for a number of years developing Mathematical Programming Software and liaising with clients. In 1976 he was appointed to the first Chair of Management Science at Edinburgh University. Then, in 1984, he moved to the Chair of Operational Research at Southampton University. He is the author of a well-known book "Model Building in Mathematical Programming". His main research interest is in Integer programming.

Professor Williams is married with three children. He is still most at home in Cornwall where he has a cottage and spends as much time as he can.

In practical applications there are sometimes thousands of variables and constraints (a mixture of $\leqslant$, $\geqslant$ and $=$). Typical objective functions represent *profit* (to be maximised) or *cost* (to be minimised).

It is not widely known that in 1826 the French mathematician Fourier [5] devised a method of manipulating Linear Inequalities. He was not concerned with optimising any expression but rather with deriving the *set of solutions* to a system of inequalities (in an analogous way to solving a set of simultaneous equations). His method has been rediscovered a number of times in different contexts. A brief account of some of these is given later.

Fourier's method can comparatively easily be adapted to solving LP models, i.e., *Optimising* an objective function subject to linear inequalities and equations. While the method results in prohibitively large storage requirements for anything but small models it is extremely illuminating and much easier to understand than the Simplex Algorithm. In addition it is a clear way of demonstrating certain theoretical properties of LP models as well as providing more information about other possible solutions.

The method can also be used in a *dual* form to provide another algorithm for solving LP models which generates all *vertex solutions*. Geometrical interpretations of both the original (known as the *primal*) method and the dual method are given later.

Fourier's method. In order to demonstrate Fourier's method we will consider an LP model in a standard form as a maximisation subject to $\leqslant$ constraints. Clearly any model can be converted into this standard form.

When we try to solve an LP one of three possibilities results.

(i) The model is *infeasible*, i.e., there are no values for the variables which satisfy all constraints simultaneously.

(ii) The model is *unbounded*, i.e., the value of the objective function can be increased without limit by choosing values for the variables.

(iii) The model is *solvable*, i.e., there exists a set of values for the variables giving a finite optimal value to the objective function.

Although case (iii) applies to our illustrative numerical example, it will be obvious in the method how cases (i) and (ii) manifest themselves.

In order to demonstrate the method we will use the model P above. Since we wish to maximise $-4x_1 + 5x_2 + 3x_3$ as well as solve the inequalities we will consider the model in the form:

$$\text{P1} \qquad \begin{array}{llrl} \text{Maximise } z & & & \\ \text{subject to:} & 4x_1 - 5x_2 - 3x_3 + z & \leqslant 0 & \text{C0} \\ & -x_1 + x_2 - x_3 & \leqslant 2 & \text{C1} \\ & x_1 + x_2 + 2x_3 & \leqslant 3 & \text{C2} \\ & -x_1 & \leqslant 0 & \text{C3} \\ & -x_2 & \leqslant 0 & \text{C4} \\ & -x_3 & \leqslant 0. & \text{C5} \end{array}$$

Constraint C0 is really a way of saying we wish to maximise z where

$$z \leqslant -4x_1 + 5x_2 + 3x.$$

By maximising z we will "drive" it up to the maximum value of the objective function. It would clearly be possible to treat C0 as an equation but for simplicity of exposition we are treating all constraints as $\leqslant$ inequalities.

Fourier gives a method of *eliminating* variables from inequalities. We will eliminate $x_1, x_2, \ldots$, etc., from the inequalities $\text{C0}, \text{C1}, \ldots$, etc., until we are left with inequalities in z above. Then the maximum possible value of z will be apparent.

To eliminate a variable from a set of *inequalities*, Fourier pointed out that we must consider all *pairs of inequalities* in which the inequality has opposite sign and eliminate between each pair. To demonstrate this we will first consider the import of constraints C0 and C1 above.

C0 can be written as

$$4x_1 \leqslant (5x_2 + 3x_3 - z);$$

C1 can be written as

$$x_1 \geqslant -2 + x_2 - x_3.$$

Therefore we have

$$-2 + x_2 - x_3 \leqslant x_1 \leqslant \frac{1}{4}(5x_2 + 3x_3 - z). \tag{1}$$

Since x_1 is a real number and the real numbers form a continuum (in contrast to the natural numbers), the import of the pair of inequalities above is that

$$-2 + x_2 - x_3 \leqslant \frac{1}{4}(5x_2 + 3x_3 - z), \tag{2}$$

i.e.,

$$-x_2 - 7x_3 + z \leqslant 8. \tag{3}$$

This constraint is more easily arrived at by simply adding 4 times constraint C1 to C0 above to eliminate x_1.

We have shown that if there is a solution to the inequalities such as C0 and C1, there must be a solution to the derived inequality (3). Conversely, if there is a solution to an inequality such as (3), writing it in the form (2) demonstrates that there exists a value of x_1 satisfying (1). In order to give x_1 a value we can take the value of either the left-hand-side or the right-hand-side of the inequality (2) (or any value in between).

Since x_1 also occurs in constraints C2 and C3, we must also eliminate it between all the other pairs in which it has opposite sign, i.e., (C0, C3), (C1, C2) and (C2, C3). If we fail to consider every possible pair we are in danger of losing information and generating spurious (infeasible) solutions.

These eliminations result in the transformed model:

$$\text{P2} \qquad \begin{array}{llll} \text{Maximise } z & & & \\ \text{subject to:} & -x_2 - 7x_3 + z & \leqslant 8 & \text{C0} + 4\text{C1} \\ & -5x_2 - 3x_3 + z & \leqslant 0 & \text{C0} + 4\text{C3} \\ & 2x_2 + x_3 & \leqslant 5 & \text{C1} + \text{C2} \\ & x_2 + 2x_3 & \leqslant 3 & \text{C2} + \text{C3} \\ & -x_2 & \leqslant 0 & \text{C4} \\ & -x_3 & \leqslant 0. & \text{C5} \end{array}$$

The origins of the combined constraints are indicated. It is convenient (but not strictly necessary) always to keep the coefficient of z (the objective), where it occurs, as 1. In order to do this, for example, we add 4 times C1 to C0 in preference to $\frac{1}{4}$ times C0 to C1 or any other combination which reduces the new coefficient of x_1 to zero.

If P1 has a solution (giving values for x_2, x_3 and z), then we have shown P2 must have a solution. Conversely, if P2 has a solution, then a value of x_1 can be found, which satisfies P1, using the argument above.

It is worth contrasting this elimination procedure (for inequalities) with Gaussian elimination (for equations). If the variable to be eliminated has a nonzero coefficient in an equation, this equation (known as the pivot equation) can be used to eliminate the variable from all other equations. With inequalities our elimination procedure is clearly more complex, although should such an equation be present (together with inequalities) we can still use it in this way.

Having eliminated x_1 we now eliminate another variable. There is complete flexibility in the order in which the variables are eliminated. For convenience we will continue to eliminate the variables in consecutive order and to choose x_2. The pairs of constraints in which x_2 have opposite sign are

$$(\mathrm{C0} + 4\mathrm{C1}, \mathrm{C1} + \mathrm{C2}), (\mathrm{C0} + 4\mathrm{C3}, \mathrm{C1} + \mathrm{C2}), (\mathrm{C1} + \mathrm{C2}, \mathrm{C4}),$$
$$(\mathrm{C0} + 4\mathrm{C1}, \mathrm{C2} + \mathrm{C3}), (\mathrm{C0} + 4\mathrm{C3}, \mathrm{C2} + \mathrm{C3}) \quad \text{and} \quad (\mathrm{C2} + \mathrm{C3}, \mathrm{C4}).$$

Combining those constraints in suitable multiples in order to eliminate x_2 reduces the model to:

$$\begin{array}{lllll}
 & \text{Maximise } z & & & \\
 & \text{subject to:} & -\frac{13}{2}x_3 + z \leqslant \frac{21}{2} & & (\mathrm{C0} + 4\mathrm{C1}) + \frac{1}{2}(\mathrm{C1} + \mathrm{C2}) \\
\mathrm{P3} & & -\frac{1}{2}x_3 + z \leqslant \frac{25}{2} & & (\mathrm{C0} + 4\mathrm{C3}) + \frac{5}{2}(\mathrm{C1} + \mathrm{C2}) \\
 & & x_3 \leqslant 5 & & (\mathrm{C1} + \mathrm{C2}) + 2\mathrm{C4} \\
 & & -5x_3 + z \leqslant 11 & & (\mathrm{C0} + 4\mathrm{C1}) + (\mathrm{C2} + \mathrm{C3}) \\
 & & 7x_3 + z \leqslant 15 & & (\mathrm{C0} + 4\mathrm{C3}) + 5(\mathrm{C2} + \mathrm{C3}) \\
 & & 2x_3 \leqslant 3 & & (\mathrm{C2} + \mathrm{C3}) + \mathrm{C4} \\
 & & -x_3 \leqslant 0. & & \mathrm{C5}
\end{array}$$

It has been shown by Kohler [8] that after n variables have been eliminated any constraint that depends on more than $n + 1$ of the original constraints must be redundant (implied by the other constraints). In this case after eliminating 2 variables the 2nd and 4th of the above inequalities depend on more than 3 of the original inequalities (both depend on C0, C1, C2 and C3). Therefore Kohler's result allows us to ignore the 2nd and 4th inequalities giving the representation:

$$\begin{array}{llll}
 & \text{Maximise } z & & \\
 & \text{subject to:} & -\frac{13}{2}x_3 + z \leqslant \frac{21}{2} & \mathrm{C0} + \frac{9}{2}\mathrm{C1} + \frac{1}{2}\mathrm{C2} \\
\mathrm{P3}' & & x_3 \leqslant 5 & \mathrm{C1} + \mathrm{C2} + 2\mathrm{C4} \\
 & & 7x_3 + z \leqslant 15 & \mathrm{C0} + 5\mathrm{C2} + 9\mathrm{C3} \\
 & & 2x_3 \leqslant 3 & \mathrm{C2} + \mathrm{C3} + \mathrm{C4} \\
 & & -x_3 \leqslant 0. & \mathrm{C5}
\end{array}$$

Finally we eliminate x_3 between pairs of inequalities where x_3 has coefficients of opposite sign. Again Kohler's result enables us to ignore the elimination between the 1st and 4th constraint. The resultant transformed model is:

$$\begin{array}{llll}
 & \text{Maximise } z & & \\
 & \text{subject to:} & z \leqslant 43 & \mathrm{C0} + 11\mathrm{C1} + 7\mathrm{C2} + 13\mathrm{C4} \\
\mathrm{P4} & & 0 \leqslant 5 & \mathrm{C1} + \mathrm{C2} + 2\mathrm{C4} + \mathrm{C5} \\
 & & z \leqslant \frac{38}{3} & \mathrm{C0} + \frac{7}{3}\mathrm{C1} + \frac{8}{3}\mathrm{C2} + \frac{13}{3}\mathrm{C3} \\
 & & z \leqslant 15 & \mathrm{C0} + 5\mathrm{C2} + 9\mathrm{C3} + 7\mathrm{C5} \\
 & & 0 \leqslant 3. & \mathrm{C2} + \mathrm{C3} + \mathrm{C4} + 2\mathrm{C5}
\end{array}$$

Clearly the maximum value of z satisfying *all* these constraints is $38/3$. This arises as the *minimum* constant on the right-hand side of the three inequalities involving z. In order to obtain the values of the variables $x_1, x_2, \ldots$, etc., which give rise to the maximum value of z we can work backwards as follows.

The 3rd constraint in P4 is that one which gives $z = 38/3$. This arises from combining the 1st and 3rd constraints in P3′. If $z = 38/3$ (instead of $z < 38/3$), then we must have the 1st and 3rd constraints of P3′ satisfied as equations. Solving these equations gives $x_3 = 1/3$. These constraints in turn arise from the 1st, 2nd, 3rd and 4th constraints in P2 which when solved as equations give $x_2 = 2\frac{1}{3}$. Finally the origins of these constraints are C0, C1, C2, and C3 in P1 which when solved as equations give $x_1 = 0$.

Alternatively we could observe immediately that constraint $z \leqslant 38/3$ in P4 arises from C0, C1, C2 and C3. If we set $z = 38/3$, this forces us to treat these constraints as equations, which when solved simultaneously give this optimal solution.

This method gives us much more information that just the specific solution to a specified model. The coefficients (multipliers) of C0, C1, C2 and C3 in the 3rd constraint of P4 are 1, 7/3, 8/3 and 13/3, since C0 in P1 consists of the negated original objective (plus z) this points out the obvious result that

$$\begin{array}{l} \frac{7}{3}(-x_1 + x_2 - x_3 \leqslant 2) \\ +\frac{8}{3}(x_1 + x_2 + 2x_3 \leqslant 3) \\ +\frac{13}{3}(-x_1 \qquad\qquad \leqslant 0) \\ \qquad\qquad\downarrow \\ -4x_1 + 5x_2 + 3x_3 \leqslant \frac{38}{3} \end{array}$$

These multipliers therefore show 38/3 to be an *upper bound* for the maximum value of $-4x_1 + 5x_2 + 3x_3$.

Similarly the multipliers of $C_1, C_2, \ldots$, etc., in the other inequalities (the 1st and 4th) in P4 involving z give (non-strict) upper bounds of 43 and 15, respectively, for the objective.

Our method has not only provided us with multipliers for the constraints and an upper bound for the objective function. It has also provided us with a set of values for the variables for which the objective attains the least upper bound derived. This is the main import of the famous *Duality Theorem* of LP which is discussed later.

The significance of the other inequalities $0 \leqslant 5$ and $0 \leqslant 3$, not involving z, in P4 will become apparent when we describe the dual of the method above.

Should an LP model be *infeasible* the method demonstrates this. The final inequalities will contain a *contradiction*, i.e., a constraint such as

$$0 \leqslant -1.$$

If a model is *unbounded*, this will be apparent as in the final inequalities there will be no upper limit to the value of z.

Although we have solved model P for specific values of the right-hand-side coefficients of the inequalities, it should be apparent that those values were not used until we derived the maximum value of z from P4. Therefore we could, with no extra work, have found the maximum value of z as a function of the right-hand-side coefficients. Such a function is known as the *value function* of an LP. If the right-hand-side values of the two constraints (apart from the nonnegativity constraints) in P were b_1 and b_2 instead of 2 and 3, the multipliers of C0, C1, etc., in P4 would tell us that the final inequalities would be

$$\begin{aligned} z &\leqslant 11b_1 + 7b_2 \\ 0 &\leqslant b_1 + b_1 \\ z &\leqslant \frac{7}{3}b_1 + \frac{8}{3}b_2 \\ z &\leqslant 5b_1 + 9b_2 \\ 0 &\leqslant b_2. \end{aligned}$$

Therefore if $b_1 + b_2$ or b_2 is negative, the model is infeasible. Otherwise

$$z = \text{minimum}\left(11b_1 + 7b_2, \frac{7}{3}b_1 + \frac{8}{3}b_2, 5b_1 + 9b_2\right).$$

A geometrical interpretation of the method. The method can be interpreted geometrically. It is possible to represent the model P1 in 4-dimensional Euclidean space. A point represents a *feasible* solution if its coordinates give values for the variables which satisfy the constraints. The set of feasible solutions can be shown to give a *polyhedron* in 4 dimensions. For a general model with n variables we will have a *polyhedron* in n dimensions. (The polyhedron may not be bounded as in this example.) By eliminating a variable we *project* the polyhedron down into a space of one less dimension. While we cannot visualise a space of 4 dimensions, we can visualise the transformed model P2 which has been reduced to 3 variables, and therefore is represented in a space of 3 dimensions in Fig. 1(i). By maximising z we are trying to find the *highest* point in this three-dimensional polyhedron. Each of the inequalities in P2 gives rise to a 2-dimensional face of the polyhedron. These are the faces ABC, $ACED$, $ABHI$, $ADFI$, $FDEG$ and $GECBH$. In order to visualise the diagram more easily, the coordinates (x_2, x_3, z) of the 5 vertices A, B, C, D and E are marked. The lines FD, GE, HB and IA are all parallel to the z axis. In this example none of the inequalities is redundant. If there were redundant inequalities, these would give rise to 2-dimensional planes outside the polyhedron and not therefore forming boundaries.

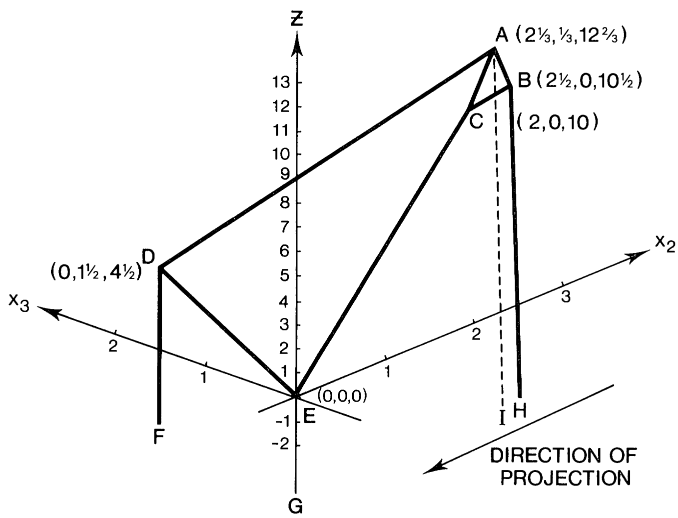

FIG. 1(i)

The elimination of variable x_2 projects this polyhedron down onto the plane (x_3, z) giving the model P3 (or P3′). In effect what we are doing by eliminating x_2 is shining rays of light parallel to the axis x_2 in the direction of the (x_3, z) plane. The *shadow* of the 3-dimensional polyhedron on the plane gives the polyhedron associated with P3′ represented in Fig. 1(ii). The inequalities in P3′ (apart from the second) respectively give rise to the lines PQ, PR, RT and QS. Although Kohler's observation allowed us to remove some redundant inequalities in P to produce $P3'$, it does not remove them all. From Fig. 1(i) it is apparent that the inequality $x_3 \leqslant 5$ is redundant (implied by the other inequalities). PQ, PR, RT and QS form the 1-dimensional faces of the 2-dimensional polyhedron. (In fact the inequality $x_3 \leqslant 5$ is the "shadow" of the line of intersection of the extended faces $ABHI$ and $FDEG$ in Fig. 1(i).)

Finally, eliminating x_3, we project the 2-dimensional polyhedron in 1(ii) down onto the z axis to give the 1-dimensional polyhedron in Fig. 1(iii). The 1st (redundant) inequality in P4 gives the point $z = 43$ which is not marked. The 3rd inequality gives the point X and the 4th inequality the point Y. Clearly X is the only 0-dimensional face of the 1-dimensional polyhedron and all inequalities apart from the 4th are redundant. For completeness we observe that the point $z = 43$ is the shadow of the intersection of lines QP and $x_3 = 5$ in Fig. 1(ii). Point Y is the shadow of the intersection of the extensions of SQ and PR. The redundant inequality $0 \leqslant 5$ is the "shadow" of the "intersection" of the parallel lines SQ and $x_3 = 5$; similarly $0 \leqslant 3$ is the "shadow" of the "intersection" of RT and SQ.

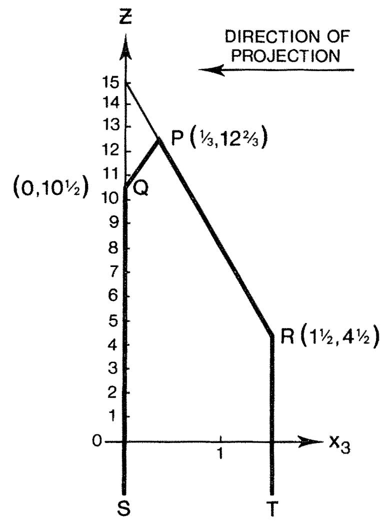

FIG. 1(ii)

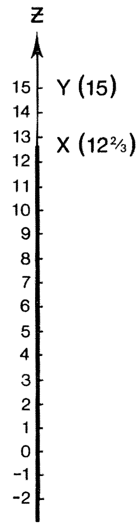

FIG. 1(iii)

Having shown that the maximum possible value of z arises from point X in 1(iii), we backtrack to the point P in 1(ii) of which X is the shadow giving $x_3 = 1/3$. P is the shadow of A in 1(i) giving $x_2 = 2\frac{1}{3}$. If it were possible to visualise 4 dimensions, A would be the shadow of a vertex of the 4-dimensional polyhedron represented by P1.

Were the original model to be *infeasible*, it would be represented by an empty polyhedron whose projections would clearly be empty. If the model were unbounded, the polyhedron would be unbounded in the z-direction which would be revealed in the projection onto the z axis.

In practice the build-up in inequalities resulting from the elimination of each variable can be explosive. If, for example, a variable to be eliminated occurs with a negative coefficient in m_1 inequalities, a positive coefficient in m_2 inequalities, and does not occur in the remaining m_3 inequalities, the result of eliminating it will be to produce $m_1 m_2 + m_3$ inequalities. Many of these resultant inequalities will be redundant. Although Kohler's observation may allow us to remove some of them, the number can still become very large even for quite modest values of m_1 and m_2. It is this, potentially explosive, growth in inequalities which makes the method computationally impractical for real life models. No efficient method has yet been devised for removing all the redundant inequalities generated.

The dual model. Another illuminating way of looking at the method is to consider the *dual model*. It has already been pointed out that the multipliers of C0, C1, C2, ..., etc., in P4 demonstrate different ways in which the constraints of P can be added together to give an upper bound for the objective function. If we look at the rows of *detached coefficients* of x_1, x_2 and x_3 in the constraints C1, C2, ..., we have

-1	1	-1	C1
1	1	2	C2
-1	0	0	C3
0	-1	0	C4
0	0	-1	C5

The multipliers of C1, C2, ..., etc., (which will always be nonnegative) in the 1st, 3rd and 4th inequalities of P4, give different ways in which these rows can be added together to give the rows of detached coefficients of the objective in P i.e.,

-4	5	3

If we let the multipliers of C1, C2, ..., C5 be $y1, y2, \ldots, y5$, we must have

$$\begin{aligned} -y_1 + y_2 - y_3 \qquad &= -4 \\ y_1 + y_2 \quad - y_4 \quad &= 5 \\ -y_1 + 2y_2 \qquad - y_5 &= 3 \end{aligned}$$

The multipliers for the 1st inequality in P4 provide a solution to this set of equations

$$y_1 = 11, \quad y_2 = 7 \quad y_3 = 0, \quad y_4 = 13, \quad y_5 = 0.$$

The multipliers for the 3rd inequality in P4 provide another solution to the equations

$$y_1 = \frac{7}{3}, \quad y_2 = \frac{8}{3}, \quad y_3 = \frac{13}{3}, \quad y_4 = y_5 = 0.$$

The multipliers for the 4th inequality in P4 provide yet another solution to the equations.

$$y_1 = 0, \quad y_2 = 5, \quad y_3 = 9, \quad y_4 = 0, \quad y_5 = 7.$$

What we are seeking are a set of nonnegative multipliers (values for the y variables) which give the *least upper bound* for the objective, in P. In order to do this we wish to

$$\text{Minimise} \quad 2y_1 + 3y_2.$$

where these coefficients 2 and 3 are the values on the right-hand sides of C1 and C2 in P1.

The problem which we have posed involving variables $y_1, y_2, \ldots$, etc., is itself an LP model. The variables y_3, y_4 and y_5 in the above three equations are sometimes known as *surplus variables*. Since they (like all the variables) cannot take negative values, the three equations above can be written as $\geqslant$ inequality constraints. If the expression $2y_1 + 3y_2$ is regarded as the new objective function, we have the new model in the form:

$$\text{D} \qquad \begin{aligned} &\text{Minimise} && 2y_1 + 3y_2 \\ &\text{subject to:} && -y_1 + y_2 \geqslant -4 \\ & && y_1 + y_2 \geqslant 5 \\ & && -y_1 + 2y_2 \geqslant 3 \\ & && y_1 \geqslant 0, \quad y_2 \geqslant 0. \end{aligned}$$

This model D is known as the *dual model* to the (*primal*) model P.

We have already, indirectly, found, by Fourier's method, a solution to D, where the value of the objective is equal to the maximum possible value of the objective of P. Since it should now be apparent that any solution to D provides an upper bound for the maximum objective of P, the solution we have obtained for D must also minimise the objective of D. This, it has already been pointed out, is an instance of a general powerful and famous result known as the *Duality Theorem* of LP. To every LP model there corresponds a dual model. If both are solvable (i.e., not infeasible or unbounded) the optimal objective values of both are the same. Fourier's method provides a clear demonstration of this.

The dual method. The fact that every LP model has a dual model allows us to convert Fourier's method into a *dual method*. Each of the steps in our original (primal) method applied to the original model can be mirrored by steps applied to the dual model. The resultant method is also intuitive and has a clear geometrical interpretation.

In our primal method we combined *rows* (constraints) together, two at a time, so as to eliminate variables (columns) from the model. Ultimately we arrived at nonnegative combinations of the original rows which gave the objective function. For the dual method we will combine *columns* together, two at a time, so as to eliminate constraints (rows) from the model. Ultimately we will arrive at nonnegative combinations of the columns which give the column of right-hand-side coefficients of the model. The multipliers in these non-negative linear combinations will constitute feasible solutions to the dual model. We seek a feasible solution which minimises the dual objective function.

Just as it was convenient to convert model P into model P1 by representing the objective by a variable z, it is convenient in the dual method, applied to model D, to represent the right-hand-side constants as coefficients of a new variable y_0 fixed at value 1. We also, in P1, explicitly included the nonnegativity conditions $-x_1 \leqslant 0$, etc. The dual correspondence to this is to include the surplus variables so making the constraints of D into equations. This gives us the form D1 of the model.

$$
\text{D1}\qquad
\begin{array}{lrcl}
\text{Minimise} & 2y_1 + 3y_2 & & \\
\text{subject to:} & 4y_0 - y_1 + y_2 - y_3 & = 0 & \text{A1} \\
 & -5y_0 + y_1 + y_2 - y_4 & = 0 & \text{A2} \\
 & -3y_0 - y_1 + 2y_2 + y_5 & = 0 & \text{A3} \\
 & y_0 & = 1 & \text{B} \\
 & y_1, y_2, y_3, y_4, y_5 \geqslant 0. & &
\end{array}
$$

The parallel between P1 and D1 should be obvious. The coefficients in the four *rows* of D1 are the same as the coefficients in the four *columns* of P1. The objective coefficients of D1 are the same as the right-hand-side coefficients of P1.

In order to eliminate constraint A1 we apply a transformation of variables. New (nonnegative) variables u_1, u_2, u_3 and u_4 are introduced which are related to the variables $y_0, y_1, \ldots,$ etc., by the equations

$$
\begin{aligned}
4u_1 + u_3 &= y_1, & 4y_2 + u_4 &= y_3, \\
4u_1 + 4u_2 &= 4y_0, & u_3 + u_4 &= y_2.
\end{aligned}
$$

When these equations are used to substitute y_0, y_1, y_2 and y_3 out of the equations in D1 it can easily be verified that the equation A1 disappears.

A graphic way of interpreting the row variables has been suggested by Dantzig and Eaves [**2**].This can best be understood through Fig. 2. In equation A1 we have a mixture of negative quantities ($-y_1$ and $-y_3$) and positive quantities ($4y_0$, y_2) which must sum to zero to satisfy the equation. $4y_0$ is split up into $4u_1$ and $4u_2$ (the coefficients of u_1 and u_2 are kept the same as y_0 so as to keep all the coefficients in the transformed equation $y_0 = 1$ as unity. This is for convenience rather than necessity). Similarly the other quantities in equation A1 are split up as indicated in Fig. 2. Fig. 2 may be interpreted as a "Transportation Problem." The Transportation problem is itself a particularly simple type of the LP model and is described in Dantzig [**1**].

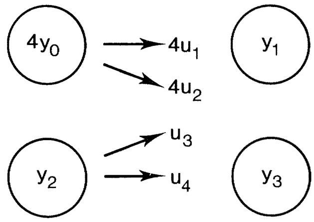

FIG. 2

While visualising the transformed variables in this way gives an interpretation to the new variables, it is not necessary for the execution of the dual method. This can be carried out mechanically by analogy with the primal method, as will become clear through the example. Performing the above substitutions transforms D1 to the model below.

$$
\text{D2}\qquad
\begin{array}{llr}
\text{Minimise} & 8u_1 \qquad + 5u_3 + 3u_4 & \\
\text{subject to:} & -u_1 - 5u_2 + 2u_3 + u_4 - y_4 \qquad = 0 & \text{A2} \\
 & 7u_1 - 3u_2 + u_3 + 2u_4 \qquad - y_5 = 0 & \text{A3} \\
 & u_1 + u_2 \qquad\qquad\qquad = 1 & \text{B} \\
 & u_1, u_2, u_3, u_4, y_4, y_5 \geqslant 0. &
\end{array}
$$

It is clear that D2 is the dual of model P2. In this dual method we have removed *constraint* A1 whereas in the primal method we removed *variable* x_1.

Rather than think in terms of transformed variables we can perform the method computationally by combining columns in pairs. The column of coefficients for u_1 in D2 arises as the column for y_0 in D1 added to 4 times the column for y_1. These two columns are combined in these multiples in order to eliminate the coefficient of the new variable u_1 in D2. Similarly the columns for the pairs of variables (y_0, y_3), (y_1, y_2) and (y_2, y_3), each having opposite signs in A1, are combined in suitable multiples. The correspondence with the elimination of x_1 in P1 in the primal method should be apparent. It is convenient to remember the origins of each column. This may conveniently be done by Table 1 below.

TABLE 1

	u_1	u_2	u_3	u_4	y_4	y_5
y_0	1	1				
y_1	4		1			
y_2			1	1		
y_3		4		1		
y_4					1	
y_5						1

The elimination of A2 from D2 can be performed similarly by combining columns for the pairs (u_1, u_3), (u_2, u_3), (u_3, y_4), (u_1, u_4), (u_2, u_4) and (u_4, y_4) in suitable multiples so that the resultant coefficients in A2 are all zero. As with the primal method some of these combinations can be ignored. If after n constraints have been eliminated a column depends upon more than $n + 1$ of the original columns, it can be shown that it may be ignored. This is the obvious dual of Kohler's observation in the primal method. The reason why it is possible to ignore such columns is pointed out below. For our example here it means that we need not combine the pairs (u_2, u_3)

and (u_1, u_4). Both would result in columns depending on y_0, y_1, y_2 and y_3. The result of eliminating constraint A2 is to produce D3′ (the dual of P3′).

$$\text{D3}' \qquad \begin{array}{lll} \text{Minimise} & \frac{21}{2}v_1 + 5v_2 + 15v_3 + 3v_4 & \\ \text{subject to:} & -\frac{13}{2}v_1 + v_2 + 7v_3 + 2v_4 - y_5 = 0 & \text{A3} \\ & v_1 + v_3 = 0 & \text{B} \\ & v_1, v_2, v_3, v_4, y_5 \geqslant 0. & \end{array}$$

The origins of the columns for the variables are given in Table 2.

TABLE 2

	v_1	v_2	v_3	v_4	y_5
y_0	1		1		
y_1	$\frac{9}{2}$	1			
y_2	$\frac{1}{2}$	1	5	1	
y_3			9	1	
y_4		2		1	
y_5					1

Table 2 can be constructed by combining the columns of Table 1 in the same multiples as the columns of D2. For example, the column for v_1 in $D3'$ arises from the column for u_1 in D2 added to $\frac{1}{2}$ times the column for u_3 in D2. Similarly, the column for v_1 in Table 2 is the column for u_1 in Table 1 added to $\frac{1}{2}$ times the column for u_3 in Table 1. Multiples of columns are chosen so as to keep the nonzero coefficients of y_0 unity in the tables of originating variables.

Finally, eliminating A3 from D3′ produces model D4 and Table 3.

$$\text{D4} \qquad \begin{array}{lll} \text{Minimise} & 43w_1 + 5w_2 + \frac{38}{3}w_3 + 15w_4 + 3w_5 & \\ \text{subject to:} & w_1 + w_3 + w_4 = 1 & \text{B} \\ & w_1, w_2, w_3, w_4, w_5 \geqslant 0. & \end{array}$$

TABLE 3

	w_1	w_2	w_3	w_4	w_5
y_0	1		1	1	
y_1	11	1	$\frac{7}{3}$		
y_2	7	1	$\frac{8}{3}$	5	1
y_3			$\frac{13}{3}$	9	1
y_4	13	2			1
y_5		1		7	2

The solution of D4 is obvious. We choose that variable from among w_1, w_3 and w_4 which has the smallest objective coefficient and set it to 1. Clearly this gives $w_3 = 1$. From Table 3 we see the multiples of the original columns of the model D1 which give rise to the column for w_3 in D4. Therefore the optimal solution to the original model is

$$y_1 = \frac{7}{3}, \quad y_2 = \frac{8}{3}, \quad y_3 = \frac{13}{3}, \quad y_4 = y_5 = 0$$

given an objective value of 38/3. The coefficient 1 for row y_0 in Table 3 indicates that we must take 1 times the negated column of right-hand-side coefficients for 0 in making up the optimal solution.

It should be obvious that again D4 is the dual model to P4 and contains the same coefficients but these are transposed. The coefficients in Table 3 are the same as the multipliers of the original constraints of P given in P4. In the same way that the primal method can provide the optimal solution for any right-hand-side coefficient, this dual method gives the optimal solution for *any objective function*. If model D were to have another objective function, then the final transformed model D4 would be the same apart from its objective coefficients. In fact, if the objective coefficients in D were b_1 and b_2 instead of 2 and 3, the transformed model D4 would be:

$$\begin{aligned} &\text{Minimise} \quad (11b_1 + 7b_2)w_1 + (b_1 + b_2)w_2 + \left(\frac{7}{3}b_1 + \frac{8}{3}b_2\right)w_3 + (5b_1 + 9b_2)w_4 + b_2 w_5 \\ &\text{subject to:} \quad w_1 + w_3 + w_4 = 1 \\ &\qquad w_1, w_2, w_3, w_4, w_5 \geqslant 0. \end{aligned}$$

If $b_1 + b_2$ or b_2 is negative, the objective can be made as small as we like and the model is said to be *unbounded* (the primal model was infeasible in these cases); otherwise the minimum value of the objective is

$$\text{minimum}\left(11b_1 + 7b_2, \frac{7}{3}b_1 + \frac{8}{3}b_2, 5b_1 + 9b_2\right).$$

The satisfaction of the duality theorem should again be obvious. The corresponding values of the variables $y_1, y_2, \ldots,$ etc., are given in the corresponding column of Table 3. Therefore apart from the case of the model D being unbounded there are three possible optimal solutions. They are:

$$\begin{aligned} &\text{corresponding to } w_1 = 1: y_1 = 11, \quad y_2 = 7, \quad y_3 = 0, \quad y_4 = 13, \quad y_5 = 0, \\ &\text{corresponding to } w_3 = 1: y_1 = \frac{7}{3}, \quad y_2 = \frac{8}{3}, \quad y_3 = \frac{13}{3}, \quad y_4 = y_5 = 0, \\ &\text{corresponding to } w_4 = 1: y_1 = 0, \quad y_2 = 5, \quad y_3 = 9, \quad y_4 = 0, \quad y_5 = 7. \end{aligned}$$

These three solutions are obviously the three sets of multipliers for the constraints on the final form of the primal model P4. In the dual model D they are three *vertex solutions*. Model D is represented in Fig. 3. The three constraints of D are represented by the faces CE, AB and BC, respectively. AD represents the nonnegativity constraint on y_1. The nonnegativity constraint on y_2 is clearly redundant. Different objective functions will give either an unbounded solution or one of the three vertex solutions at A, B or C. For example, different values of the objective function $2y_1 + 3y_2$ give lines parallel to PQ. By minimising this objective function we move to the lowest such line which still intersects the feasible region, in this case at vertex B, giving the solution $y_1 = 7/3, y_2 = 8/3$, objective $= 38/3$ already obtained. The lines AD and CE are known as *extreme rays*. Their existence is demonstrated algebraically by the columns for w_2 and w_5 in Table 3 which have entries of 0 in row y_0. For example, we can let w_2 take any nonnegative value without violating constraint B of D4. This corresponds to keeping y_1 and y_2 in the ratio 1:1 (coefficients in Table 3) and fixing y_3 at 0 (the constraint represented by CE is therefore binding). Clearly the column for w_2 in Table 3 corresponds to the extreme ray CE. Similarly the column for w_5 corresponds to the extreme ray AD.

We have therefore demonstrated that the dual of Fourier's method generates *all vertices and extreme rays* for the feasible polyhedron of an LP model. This in itself sometimes has practical application.

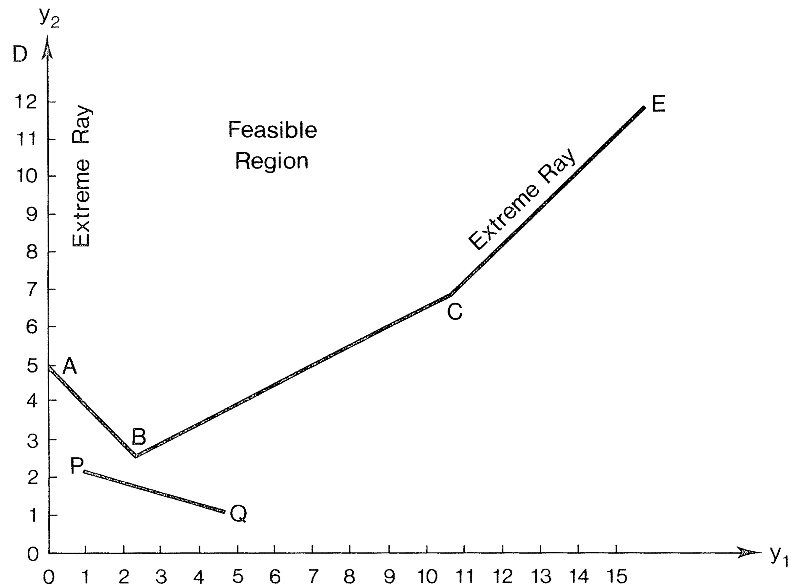

FIG. 3

It is well known that, for an LP model with m constraints, we can restrict our search for an optimal solution to solutions in which at most m variables (including slack and surplus variables) are nonzero. This is an algebraic realisation of the geometric observation that the optimal solution to an LP (if it exists) lies on the boundary of the polyhedron defined by the feasible region. If the optimal solution is unique, it will lie at a vertex, otherwise, in the case of alternate optimal solutions, there will still be among these alternatives vertex solutions which are optimal. The Simplex Algorithm restricts attention to so-called *basic solutions* which correspond to vertex solutions. This property allows us to justify Kohler's observation when applied to the dual method. When we have eliminated n constraints from our original model (D, say) we have in effect solved an LP model consisting of the first n constraints. In the optimal solution to such a model no more than n of the original variables will be nonzero. Therefore, including our right-hand-side column y_0 as a variable, no more than $n + 1$ of the original variables will go to make up a vertex solution. Hence any derived column depending on more than $n + 1$ of the original variables will correspond to a variable which can be taken as 0 in an optimal solution. Hence such a derived column may be ignored. Because of the one-to-one correspondence between derived columns in the dual method and derived rows in the primal method, this is a sufficient justification for our ignoring certain derived constraints. We did this when they depended on more than $n + 1$ of the original constraints when n constraints had been eliminated (Kohler's observation).

An outline of the history of Fourier's method and its extensions. Fourier's method was published 1826. It has been rediscovered a number of times by different authors. Motzkin [**12**] derived a method of solving 2-person zero sum games. Since any LP can be formulated as such a game (and vice versa), Motzkin's method gives rise to a method of solving LP models which in fact turns out to be Fourier's method. Hence the name Fourier-Motzkin elimination is often used for the method. Dantzig [**1**] refers to the method briefly under this name. Dines [**3**] also rediscovered the method. Langford [**10**] derived a method of solving a particular problem in Mathematical Logic. He showed, by a constructive method, that the Theory of Dense Linear Order is decidable. Williams [**14**] showed that any LP model can be posed within this restricted form of arithmetic and that hence the achievability, or otherwise, of a particular objective value can be decided. This application of Langford's method turns out to be the same as Fourier's method. Another account of Fourier's method, together with additional references, can be found in Duffin [**4**]. There is also a related article by Kuhn [**9**].

Fourier's method (and its dual) is computationally impractical for anything but small models. This is because of the large build-up in inequalities (or variables) as variables (or constraints) are eliminated. It is, however, possible that the methods could be applied in a restricted form. When all variables (apart from the objective variable) have been eliminated, one will only be interested in one of the derived inequalities. For the dual method one will only be interested in one of the final columns. Unfortunately, it is not clear how to eliminate most of the redundant inequalities (or variables) until the end. Williams [**18**] suggests applying a restricted form of the dual method as a "Crashing Procedure" prior to the Simplex Algorithm. Geometrically the Simplex Algorithm moves from vertex solution to vertex solution until it reaches the optimal vertex solution. Initially (Phase 1 of the Simplex Algorithm) it is necessary to obtain a feasible vertex solution. In practice this usually takes as much time as the second phase. For model D represented in Fig. 3 the Simplex Algorithm would start at the origin 0 and systematically move to a vertex (such as A) before proceeding to the optimal vertex at B. By applying a restricted form of the dual method one would hope to obtain a good vertex solution as a starting point.

Computational implementations of the methods using efficient data structures are possible. It is sensible to take account of the sparseness of most LP models (most coefficients in a model are usually zero) in both storing and manipulating the matrices. The transformations which eliminate variables or constraints can be represented by elementary matrices which probably gives a sparser representation than explicitly transforming the whole model. Such considerations are, however, beyond the scope of this paper.

There is a lot of interest, in view of its wide applicability, in an extension of LP known as Integer Programming (IP). Here some, or all, of the variables in a model are restricted to take integer values. Such models are much more difficult to solve than LPs. It has been shown by Lee [**11**] and Williams [**15**] how Fourier's method can be extended to allow us to eliminate integer variables. In order to do this it is necessary to introduce disjunctions of inequalities as well as congruence relations into the transformed model. The dual method can also be extended to deal with IP models by introducing congruence relations as is done by Williams [**17**].

Acknowledgement. The author would like to acknowledge the help of Mr. Robin E. Day of Edinburgh University who wrote a sophisticated computer program to perform the dual method.

References

1. G. B. Dantzig, Linear Programming and Extensions, Princeton, 1963.

2. G. B. Dantzig and B. C. Eaves, Fourier-Motzkin elimination and its dual, J. Combin. Theory Ser. A, 14 (1973) 288–297.

3. L. L. Dines, Systems of linear inequalities, Ann. of Math., 20 (1919) 191–199.

4. R. J. Duffin, On Fourier's Analysis of Linear Inequality Systems, Mathematical Programming Study I, 1974, pp. 71–95.

5. J. B. J. Fourier, Solution d'une Question Particulière du Calcul des Inégalités, Oeuvres II, Paris, 1826, pp. 317–328.

6. N. Karmarkar, A New Polynomial-Time Algorithm for Linear Programming, Proceedings of the 16th annual ACM Symposium of the Theory of Computing, Washington, DC, April 30–May 2, 1984 pp. 302–310.

7. L. G. Khachian, A polynomial algorithm in linear programming, Dokl. Akad. Nauk SSSR, 244 (1979) 1093–96 (translated as Soviet Mathematics Doklady, 20, 191–194).

8. D. A. Kohler, Projections of Convex Polyhedral Sets, Operational Research Center Report, ORC 67–29, University of California, Berkeley, 1967.

9. H. W. Kuhn, Solvability and consistency for linear equations and inequalities, this MONTHLY, 43(1956) 217–232.

10. C. H. Langford, Some theorems on deducibility, Ann. of Math. I, 28 (1927) 16–40.

11. R. D. Lee, An application of mathematical logic to the integer linear programming problem, Notre Dame J. Formal Logic, 23 (1972) 279–282.

12. T. S. Motzkin, Beitrage zur Theorie der Linearen Ungleichungen, Dissertation, University of Basel, Jerusalem, 1936.

13. N. Z. Shor, Cut-off method with space extension in convex programming problems, Kibernetika, 13 (1977) 94–96.

14. H. P. Williams, An Algorithm for the Solution of Linear Programming Problems, IBM Working Paper, 1970.

15. ______, Fourier-Motzkin elimination extension to integer programming, J. Combin. Theory Ser. A, 21 (1976) 118–123.

16. ______, Model Building in Mathematical Programming, Wiley, New York, 1985.

17. ______, A characterization of all feasible solutions to an integer programme, Discrete Applied Math., 5 (1983) 147–155.

18. ______, Restricted vertex generation applied as a crashing procedure for linear programming, Computers and Operations Research, 11 (1984) 401–407.

MUSICAL SCALES AND THE GENERALIZED CIRCLE OF FIFTHS

JOHN CLOUGH

Department of Music, SUNY at Buffalo, Amherst, NY 14260

GERALD MYERSON

Department of Mathematics, The University of Texas at Austin, Austin, TX 78712

This paper deals with the way the diatonic set (the white keys on the piano) is embedded in the chromatic scale (all the keys on the piano). To illustrate the problem, consider the chords CDF and EFA (the reader who happens to be temporarily without piano may find Fig. 1 helpful). If we ignore the black keys, these chords have the same structure; the second note is one key higher than the first, and the third note is two keys higher than the second. When actually played on the piano, the chords sound quite different, due to the embedding of the diatonic in the chromatic. From C to D is two semitones (a semitone is the distance between adjacent notes in the chromatic scale), and from D to F is three, whereas E to F is one and F to A is four. The problem

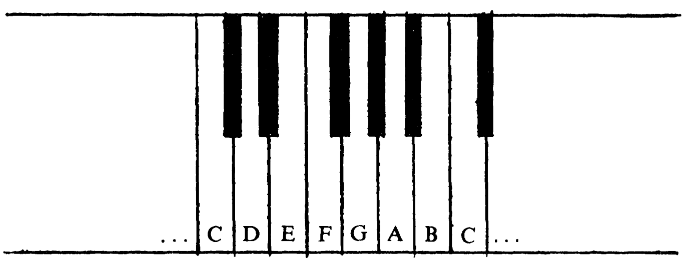

FIG. 1. Piano keyboard.

John Clough: Before coming to SUNY at Buffalo, I taught at the Oberlin College Conservatory of Music and in the School of Music at the University of Michigan. At all three places I have enjoyed the colleagueship of mathematicians who were willing to help me work through various problems in the application of mathematics to music: Edward Wong and Samuel Goldberg at Oberlin, Bernard Galler at Michigan, John Myhill and Gerald Myerson at Buffalo. Though trained only as a musician, in occasional flights of fancy I consider a second career in my first love—mathematics.

Gerald Myerson: I received my Ph.D. in Mathematics under the direction of Don Lewis at the University of Michigan in 1977. I have been on the faculty at the University of Buffalo, the University of British Columbia, and the University of Texas. I play two musical instruments: the phonograph and the cassette deck.