The general theory of relaxation methods applied to linear systems

By G. Temple

(*Communicated by S. Chapman, F.R.S.—Received* 11 *October* 1938)

1. Introduction and summary

During the last few years Southwell and his fellow-workers have developed a new method for the numerical solution of a very general type of problem in mathematical physics and engineering. The method was originally devised for the determination of stresses in frameworks, but it has proved to be directly applicable to any problem which is reducible to the solution of a system of non-homogeneous, linear, simultaneous algebraic equations in a finite number of unknown variables.* Southwell's "relaxation method" is one of successive approximation and, in order to complete the previous investigations of this method, it is necessary to prove that the successive approximations do actually converge towards the exact solutions. This formal proof is given in § 4.

Southwell's "relaxation" methods are not *directly* applicable to continuous systems, where the number of unknown variables is infinite, but it is shown here that simple extensions and modifications of the relaxation method render it suitable for application to either discrete or continuous systems (§ 5).

The general theory of relaxation methods is then developed in terms of the theory of linear operators (§ 7) and sufficient conditions are prescribed for the convergence of the process of approximation (§ 8).

These general methods are then applied to the solution of non-homogeneous, linear integral equations (§ 10) and to the solution of non-homogeneous, linear differential equations (§§ 11, 12).

Broadly speaking it appears that almost all linear systems of equations, algebraic, integral or differential, can be brought within the scope of relaxation methods, which seem to constitute one of the most powerful methods of computation in mathematical physics and engineering.

* In a paper to be published shortly Southwell has extended the method to deal with homogeneous equations such as occur in vibration problems involving characteristic numbers.

I. LINEAR ALGEBRAIC EQUATIONS IN n UNKNOWN VARIABLES

2. The problem of stress calculation in frameworks

Southwell's method of "Systematic Relaxation of Constraints" was originally devised for the determination of stresses in redundant frameworks, and, although the method appears to be applicable to any minimal problem, it is convenient to retain the notation and terminology of the stress problem. We therefore begin by considering an elastic framework whose configuration is completely specified by n co-ordinates,

$$u_1, u_2, \ldots, u_n,$$

which determine the displacements of the joints of the framework from the positions of equilibrium which they occupy when there are no external forces. In any configuration of the framework which is geometrically possible the generalized forces corresponding to the co-ordinate u_k are

(i) the internal force exerted by the framework, A_k,
(ii) the external force exerted by the loads, B_k, and
(iii) the residual force which would be required to produce equilibrium,

$$C_k = A_k + B_k.$$

The principal properties of an elastic framework are the following:

(i) the internal forces and displacements satisfy a generalized Hooke's Law,
(ii) the framework forms a conservative system, and
(iii) the position of equilibrium under prescribed external forces is stable.

It is upon these properties that relaxation methods depend, and we therefore proceed to express them in analytical form.

(i) In a system which obeys a generalized Hooke's Law the internal force A_k is a homogeneous, linear function of the co-ordinates $\{u_k\}$, i.e.

$$A_k = \Sigma A_{kj} u_j,$$

where the "influence coefficients", A_{kj}, are constants which depend only on the form and the material of the framework, and where the summation here and elsewhere (in I) is with respect to the repeated suffix and is taken over all the joints of the framework. The fundamental problem is to determine the co-ordinates when the framework is in equilibrium under the action of prescribed external forces. In these circumstances there are no residual forces and the co-ordinates satisfy the n simultaneous, linear, algebraic equations,

$$\Sigma A_{kj} u_j + B_k = 0. \tag{2·1}$$

In order that these equations should possess a unique solution for any system of external forces it is necessary and sufficient that the determinant A formed with the influence coefficients should not be zero. We shall denote by A^{kj} the cofactor of A_{kj} in this determinant.

(ii) Since an elastic framework forms a conservative system the influence coefficients satisfy Maxwell's reciprocal relations,

$$A_{kj} = A_{jk},$$

and hence there exists a strain energy function

$$U(u) = \tfrac{1}{2}\Sigma A_{kj} u_k u_j, \tag{2·2}$$

such that the internal forces are given by

$$A_k = \partial U/\partial u_k. \tag{2·3}$$

The potential energy of the external forces is

$$V(u) = \Sigma B_k u_k, \tag{2·4}$$

and the total energy of the framework is

$$W(u) = U(u) + V(u). \tag{2·5}$$

(iii) Finally, since the position of equilibrium is stable, the strain energy $U(u)$ is necessarily positive for all sets of co-ordinates $\{u_k\}$, except, of course, the null-set, $u_k = 0$, $(k = 1, 2, \ldots, n)$; i.e., in the terminology of the theory of operators, the quadratic form (2·2) is "positive definite"

It follows at once that

$$A_{kk} > 0 \quad (k = 1, 2, \ldots, n); \tag{2·6}$$

and a well-known argument shows that the determinant A is essentially positive. The positive character of the strain energy $U(u)$, and, in particular, the relations (2·6) also show that in the position of equilibrium the total energy $W(u)$ is an absolute minimum, $W_{\min}$.

Lastly, we note that, if by an orthogonal transformation of the type

$$u_k = \Sigma l_{kj} x_j,$$

the strain energy is brought into the form

$$U(u) = \tfrac{1}{2}\Sigma L_k x_k^2,$$

while

$$\Sigma u_k^2 = \Sigma x_k^2,$$

then we must have $L_k > 0$ for $k = 1, 2, \ldots, n$. Hence the set of coefficients, $\{L_k\}$, has a positive minimum, say m, and a finite maximum, μ. We can thus infer the important relation, required later (§ 5),

$$0 < m \leqslant 2U(u)/\Sigma u_k^2 \leqslant \mu, \tag{2·7}$$

for all displacements $\{u_k\}$.

3. The method of successive approximation

The solution of the system of simultaneous equations (2·1) which determines the position of equilibrium is excessively laborious save for the very simplest frameworks, and, according to Southwell (1935, p. 58), it is effectively impracticable when n, the number of degrees of freedom, exceeds ten or twelve. On the other hand the direct problem of calculating the residual forces $\{C_k\}$ required to give the framework specified co-ordinates $\{u_k\}$ is extremely simple, the requisite residual forces being given explicitly by the equations

$$C_k(u) = \Sigma A_{kj} u_j + B_k. \tag{3·1}$$

Now the residual forces are given by the relation

$$C_k(u) = \partial W / \partial u_k,$$

(by equations (2·2)–(2·5)), and in the position of equilibrium when W is a minimum,

$$C_k(u) = 0.$$

In any other configuration the magnitudes of the residual forces furnish an estimate of the deviation of this configuration from that of equilibrium; and, moreover, we shall now show that a knowledge of the values of the residual force enables us to prescribe displacements which will reduce the value of the total energy W.

To establish this result we observe that if $\{v_k\}$ is a displacement to be considered in this connexion, then it will also be reasonable to consider the whole family of displacements $\{tv_k\}$, where t is an arbitrary, numerical multiplier. We shall describe any displacement of this family as being of the "type" $\{v_k\}$. Since W is a minimum in the equilibrium configuration the best displacement of type $\{v_k\}$ is that which reduces W by the maximum amount. This displacement is easily determined as follows:

In an obvious notation, the total energy in the configuration specified by the co-ordinates $\{u_k + tv_k\}$ is

$$\left.\begin{aligned} W(u+tv) &= W(u) + W_1 t + W_2 t^2, \\ \text{where} \qquad W_1 &= \Sigma A_{kj} u_k v_j + \Sigma B_k v_k = \Sigma v_k C_k(u), \\ \text{and} \qquad W_2 &= \tfrac{1}{2} \Sigma A_{kj} v_k v_j = U(v). \end{aligned}\right\} \tag{3·2}$$

Hence $\qquad W(u+tv) - W(u) = (t + \frac{1}{2} W_1/W_2)^2 W_2 - \frac{1}{4} W_1^2/W_2.$

Since W_2 is the strain energy for the co-ordinates $\{v_k\}$, it follows that W_2 is positive. Therefore the best displacement of type $\{v_k\}$ is obtained when

$$t = -\tfrac{1}{2} W_1/W_2, \tag{3·3}$$

and then $$W(u+tv) - W(u) = -\tfrac{1}{4} W_1^2/W_2. \tag{3·4}$$

Hence the displacement $\{tv_k\}$ has certainly not increased the total energy.

Moreover, by (3·1)

$$W_1 = \Sigma v_k C_k(u), \tag{3·5}$$

so that, unless the initial configuration $\{u_k\}$ is one of equilibrium (when $C_k(u) = 0$), we can always choose the type of the displacement $\{v_k\}$ so that W_1 does not vanish. Then $W(u+tv)$ will be definitely less than $W(u)$, and the configuration $\{u_k + tv_k\}$ will be nearer to the position of equilibrium than the configuration $\{u_k\}$.

It is clear that, starting from any arbitrary configuration, we can thus determine a sequence of configurations in each of which the total energy is less than in the preceding configuration. The configurations so determined will form successive approximations to the equilibrium position. It only remains to show that the sequence of configurations so determined does actually converge to the equilibrium position.

To establish this result, upon which all relaxation methods depend, we write $(u_1^{(p)}, u_2^{(p)}, \ldots, u_n^{(p)})$ for the co-ordinates in the pth configuration of the sequence, and we denote the corresponding value of the total energy by $W(u^{(p)})$. Then

$$W(u^{(p)}) \geqslant W(u^{(p+1)}) \geqslant W_{\min}.$$

Hence the sequence $\{W(u^{(p)})\}$, $(p = 1, 2, \ldots)$ converges to a unique limit. Now, by (3·4) and (3·5)

$$W(u^{(p)}) - W(u^{(p+1)}) = \tfrac{1}{4}\{\Sigma v_k^{(p)} C_k(u^{(p)})\}^2 / U(v^{(p)}),$$

$\{v_k^{(p)}\}$ being the type of displacement employed in forming the $(p+1)$th configuration of the sequence. Therefore,

$$\{\Sigma v_k^{(p)} C_k(u^{(p)})\}^2 / U(v^{(p)}) \to 0, \quad \text{as } p \to \infty. \tag{3·6}$$

We shall consider only those methods of approximations in which all the types of displacement $\{v_k^{(p)}\}$ employed are such that the relation (3·6) implies that

$$C_k(u^{(p)}) \to 0, \text{ as } p \to \infty, \quad \text{for } k = 1, 2, \ldots, n. \tag{3·7}$$

When this condition is satisfied the proof of the convergence of the sequence $\{u_k^{(p)}\}$ is easily completed.

In the notation of § 2, the solution of equations (2·1),

$$\Sigma A_{kj} u_k + B_j = 0$$

is

$$A u_k = -\Sigma A^{kj} B_j.$$

Now, in the pth configuration of the sequence, the co-ordinates $\{u_k^{(p)}\}$ satisfy the equation

$$\Sigma A_{kj} u_k^{(p)} + B_j = C_j(u^{(p)}).$$

Hence

$$A u_k^{(p)} = -\Sigma A^{kj} B_j + \Sigma A^{kj} C_j(u^{(p)}).$$

Therefore, by (3·7),

$$A u_k^{(p)} \to -\Sigma A^{kj} B_j, \text{ as } p \to \infty, \quad \text{for } k = 1, 2, \ldots, n,$$

i.e. the sequence of configurations converges to the equilibrium configuration.

We have therefore shown that the co-ordinates in the position of equilibrium, in which W is a minimum, can be determined by a method of successive approximation in which each step in the process depends upon the "direct" problem of the calculation of the residual forces in terms of prescribed co-ordinates. It is perhaps desirable to emphasize that it is the total energy W and not the strain energy U which converges to an extreme value in this process. In Southwell's first paper (1935, § 7) attention was focused on changes in the strain energy, but it is easy to see from simple examples that no *general* statement can be made about the direction of changes in the strain energy.*

The proof of the convergence of the successive approximations given by Black and Southwell (1938, § 3) is not quite complete as it assumes that the convergence of the sequence $\{W(u^{(p)})\}$ always implies the convergence of the sequence $\{u_k^{(p)}\}$. This assumption is easily verified for Southwell's method of successive relaxation, and an elementary formal proof is given in § 4.

* If $n = 1$, write $u_1 = x$, $v_1 = y$, $U(x) = \frac{1}{2}Ax^2$, $V(x) = Bx$. Then

$$W(x+ty) - W(x) = (Axy + By)\,t + \tfrac{1}{2}Ay^2t^2,$$

and the right-hand side is a minimum when $t = -(Ax+B)/(Ay)$. The change in the strain energy is

$$U(x+ty) - U(x) = \tfrac{1}{2}A(2x+ty)\,(ty) = -\tfrac{1}{2}A(x^2 - B^2/A^2),$$

and this can be either positive or negative according to the value of x.

4. The method of successive relaxation

The preceding section gives the general theory of all relaxation methods for elastic frameworks. Special methods of approximation can now be obtained by particularizing the types of displacement $\{v_k^{(p)}\}$ to be employed at each stage of the process. There appear to be two general methods—Southwell's method of successive relaxation in which the type of displacement is dependent on the co-ordinate system adopted to describe the configuration of the framework, and the method given in the next section in which the type of displacement is determined by the intrinsic properties of the framework as expressed in the total energy function $W(u)$.

In Southwell's original method (1935) the displacement $\{v_k^{(p)}\}$ affects only one co-ordinate, namely that for which the absolute value of the residual force is a maximum; and the magnitude of the change in this co-ordinate is adjusted so as to reduce the corresponding residual force to zero. The method therefore proceeds by a successive relaxation of the constraints represented by the residual forces. In symbols, if

$$|C_k(u^{(p)})| \leqslant |C_m(u^{(p)})| \quad \text{for } k = 1, 2, \ldots, n, \tag{4·1}$$

then

$$v_k^{(p)} = 0 \quad \text{if } k \neq m,$$
$$\text{or} \quad 1 \quad \text{if } k = m.$$

It is easily verified that the magnitude of the displacement prescribed by Southwell's method is the best displacement of this type, i.e. that which reduces the total energy by the greatest amount. In the notation of equation (3·2) we have for Southwell's method,

$$\begin{aligned} W_1 &= \Sigma A_{kj} u_k^{(p)} v_j^{(p)} + \Sigma B_k v_k^{(p)} \\ &= \Sigma A_{km} u_k^{(p)} + B_m \\ &= C_m(u^{(p)}), \end{aligned}$$

and

$$W_2 = \tfrac{1}{2}\Sigma A_{kj} v_k^{(p)} v_j^{(p)} = \tfrac{1}{2}A_{mm}.$$

Hence, by (3·3),

$$t = -\tfrac{1}{2}W_1/W_2 = -C_m(u^{(p)})/A_{mm}, \tag{4·2}$$

and the residual force corresponding to the mth co-ordinate is thus reduced to

$$\begin{aligned} C_m(u^{(p)} + tv^{(p)}) &= C_m(u^{(p)}) + t\Sigma A_{mk} v_k^{(p)} \\ &= C_m(u^{(p)}) + tA_{mm} = 0, \quad \text{by (4·2)}, \end{aligned}$$

in agreement with the value of the displacement prescribed by Southwell.

It only remains to verify that this type of displacement satisfies the condition that the convergence of the sequence $\{W(u^{(p)})\}$ implies the convergence of the sequence $\{u_k^{(p)}\}$. We have that

$$W(u^{(p)}) - W(u^{(p+1)}) = \tfrac{1}{4}\{\Sigma v_k^{(p)} C_k(u^{(p)})\}^2/U(v^{(p)})$$
$$= \tfrac{1}{2}\{C_m(u^{(p)})\}^2/A_{mm}. \tag{4·3}$$

Now $A_{mm} > 0$ for $m = 1, 2, \ldots, n$, by (2·6), and

$$|C_k(u^{(p)})| \leqslant |C_m(u^{(p)})|, \quad \text{for } k = 1, 2, \ldots, n, \text{ by (4·1).}$$

Hence it follows from (4·3) that

$$C_m(u^{(p)}) \to 0, \quad \text{as } p \to \infty,$$

whence $$C_k(u^{(p)}) \to 0, \text{ as } p \to \infty, \quad \text{for } k = 1, 2, \ldots, n.$$

The condition (3·7) is therefore satisfied, and it follows that the sequence of approximation $\{u_k^{(p)}\}$ actually converges to the co-ordinates in the equilibrium configuration.

5. The method of steepest descents

The second general method of successive approximation is most easily visualized by considering the simple example in which there are only two co-ordinates, say x and y, while

$$U = \tfrac{1}{2}(ax^2 + 2hxy + by^2),$$

and $$V = gx + fy, \quad \text{with} \quad |f| < |g|.$$

Then the components of the residual force are

$$C_1(x, y) = ax + hy + g, \quad C_2(x, y) = hx + by + f.$$

If we regard the total energy, $U + V$, as a third co-ordinate z, the expression for the total energy,

$$z = U + V = \tfrac{1}{2}(ax^2 + 2hxy + by^2 + 2gx + 2fy), \tag{5·1}$$

can be regarded as the equation to a surface in ordinary space. The position of equilibrium, given by $C_1 = 0$, $C_2 = 0$, is stable by hypothesis, and therefore corresponds to the bottom of the valley represented by (5·1).

In Southwell's method of approximation the successive displacements which converge to the valley bottom are alternately in the direction of the axes of x and of y, i.e. their direction is determined by the co-ordinate system employed. In the method to be developed in this section each displacement is in the direction of the line of steepest descent at the point reached by the previous displacement. In symbols, Southwell's type of displacement from the position $(0, 0)$ is

$$x = 1, \ y = 0 \quad (\text{since } |f| < |g|),$$

while the type of displacement in the method of steepest descents is

$$x = C_1(0,0) = g, \quad y = C_2(0,0) = f.$$

A few numerical calculations with this example are sufficient to show that Southwell's method is undoubtedly simpler in practical computation, but the method of steepest descents is theoretically better, and, as we shall see later (II), it is easily extended to continuous systems with an infinite number of degrees of freedom to which Southwell's method cannot be applied *directly*.*

For the general framework with n degrees of freedom the direction of "steepest descent" will be given by a displacement of the type

$$v_k^{(p)} = \partial W(u^{(p)})/\partial u_k^{(p)} = C_k(u^{(p)}). \tag{5·2}$$

The direction of this displacement is that in which the total energy W changes most rapidly. It is independent of the co-ordinate system employed, and is organically connected with the intrinsic properties of the framework as expressed in the form of W. In Southwell's method the displacement affects only one co-ordinate, viz. that for which the magnitude of the residual force is a maximum. *In the method of steepest descents the displacement affects all co-ordinates and affects them in the ratio of their residual forces.*

In the notation of § 3, (3·2),

$$W_1 = \Sigma [C_k(u^{(p)})]^2,$$

and

$$W_2 = \tfrac{1}{2} \Sigma A_{kj} C_k(u^{(p)}) \, C_j(u^{(p)}).$$

Hence, by (3·3),

$$t = -\tfrac{1}{2} W_1/W_2 = \frac{-\Sigma [C_k(u^{(p)})]^2}{\Sigma A_{kj} C_k(u^{(p)}) \, C_j(u^{(p)})},$$

and, by (3·4),

$$W(u^{(p+1)}) - W(u^{(p)}) = \frac{\frac{1}{2}\{\Sigma [C_k(u^{(p)})]^2\}^2}{\Sigma A_{kj} C_k(u^{(p)}) \, C_j(u^{(p)})}$$

$$= \tfrac{1}{4}\{\Sigma [c_k^{(p)}]^2\}^2 / U(c^{(p)}),$$

on writing $c_k^{(p)}$ for $C_k(u^{(p)})$.

* It should be added that Southwell has shown that his method can be adapted to continuous systems if these are first replaced (approximately) by appropriate discrete systems. Thus, for example, differential equations must be replaced by difference equations (Bradfield and Southwell 1937; Atkinson, Bradfield and Southwell 1937).

To verify that this type of displacement satisfies the general condition postulated in § 3, we note that by the inequalities (2·7),

$$0 < m \leqslant \frac{2U(c^{(p)})}{\Sigma [c_k^{(p)}]^2} \leqslant \mu.$$

Hence
$$W(u^{(p+1)}) - W(u^{(p)}) \geqslant \frac{1}{2\mu} \Sigma [c_k^{(p)}]^2.$$

Now the sequence $\{W(u^{(p)})\}$, $(p = 1, 2, \ldots)$ converges to a limit. Therefore

$$\Sigma [c_k^{(p)}]^2 \to 0 \quad \text{as} \quad p \to \infty$$

and therefore,

$$C_k(u^{(p)}) = c_k^{(p)} \to 0 \text{ as } p \to \infty, \quad \text{for } k = 1, 2, \ldots, n.$$

Thus the successive approximations determined by the method of steepest descents converge to the equilibrium co-ordinates.

In the terminology introduced by Richards (1937), the type of displacement in the method of steepest descents is a "group displacement", and it is clear that, in theory, it is the best possible displacement, although in practice Southwell's methods will probably be quicker for numerical computation.

6. Gyrostatic systems

The relaxation methods described above are essentially methods of successive approximation for the solution of a system of simultaneous, linear, algebraic equations

$$\Sigma A_{kj} u_j + B_k = 0,$$

in the canonical form in which

$$A_{kj} = A_{jk}, \tag{6·1}$$

and
$$\Sigma A_{jk} u_j u_k > 0, \tag{6·2}$$

for all $\{u_k\}$ except $u_k = 0$. It has been shown by Black and Southwell (1938, §§ 16–20) that the relaxation method is easily extended to "gyrostatic" systems of equations in which neither the relation (6·1) nor (6·2) is fulfilled.

In fact *any* system of equations

$$\Sigma P_{ik} u_k + Q_i = 0, \tag{6·3}$$

such that
$$\det P_{ik} \neq 0,$$

is exactly equivalent to a system in the canonical form. For, let

$$A_{jk} = \Sigma P_{ij} P_{ik},$$

and
$$B_j = \Sigma P_{ij} Q_i. \tag{6·4}$$

Then, it follows from (6·3) that

$$\Sigma A_{jk} u_k + B_j = 0, \tag{6·5}$$

and, conversely, since $\det P_{ik} \neq 0$, the original equations (6·3) can be deduced from (6·5).

Now, it is clear from (6·4) that

$$A_{jk} = A_{kj},$$

and that

$$\sum_{j,k} A_{jk} u_j u_k = \sum_{i,j,k} P_{ij} P_{ik} u_j u_k$$

$$= \sum_i \{\sum_j P_{ij} u_j\}^2 \geqslant 0.$$

The sign of equality holds only if

$$\sum_j P_{ij} u_j = 0 \quad \text{for } i = 1, 2, \ldots, n,$$

and these equations are consistent only if $u_j = 0$, $(j = 1, 2, \ldots, n)$. Hence equations (6·5) are in the canonical form. They are, in fact, the familiar conditions that

$$\sum_i \{\sum_j P_{ij} u_j + Q_i\}^2$$

should be a minimum.

II. LINEAR OPERATIONAL EQUATIONS

7. The problem of linear operational equations

So far we have applied relaxation methods only to systems with a finite number of degrees of freedom, but the extension to continuous systems with an infinite number of degrees of freedom is immediate, and the necessary generalizations are discovered at once when the problem is stated in terms of the theory of linear operators. In this part of the paper we shall therefore study the general operational problem and in Parts III and IV apply our results to linear integral equations and linear differential equations.

In Part I we dealt with two classes of entities—sets of co-ordinates $\{u_k\}$ and sets of influence coefficients $\{A_{jk}\}$. In this part we replace these by the more general concepts of vectors $\alpha, \ldots$ in a Hilbert space $\mathfrak{h}$, and linear operators $A, \ldots$ acting in this space. These concepts are defined by the following axioms:

The Hilbert space (an abstract, complex, Euclidean space) is linear, metric and complete:

(i) It is *linear* in the sense that if α, β are vectors belonging to $\mathfrak{h}$, so also is $x\alpha + y\beta$, where x, y are any complex numbers.

(ii) It is a *metric* space in the sense that there exists a function (β, α) for any two vectors, which is a complex number such that

$$(\gamma, x\alpha + y\beta) = x(\gamma, \alpha) + y(\gamma, \beta);$$

(β, α) and (α, β) are conjugate complex numbers; and $(\alpha, \alpha) \geqslant 0$.*

(iii) It is a *complete* space in the sense that, if $\{\alpha_p\}$ is an infinite sequence of vectors with the property that

$$|\alpha_m - \alpha_n| \to 0 \quad \text{as} \quad m, n \to \infty,$$

then there exists a vector α such that

$$|\alpha - \alpha_p| \to 0 \quad \text{as} \quad p \to \infty.$$

$\mathfrak{h}$ may have an *infinite* number of dimensions in the sense that for *any* positive integer n there exists n linearly independent vectors, $\alpha_1, \alpha_2, \ldots, \alpha_n$.

For any operator A and any vector α in $\mathfrak{h}$, there exists a "product" $A\alpha$, which is also a vector in $\mathfrak{h}$. The operators are linear in the sense that

$$A(x\alpha + y\beta) = xA\alpha + yA\beta,$$

for all complex numbers x, y and all vectors α, β in $\mathfrak{h}$. The quadratic forms $\Sigma A_{jk} u_j v_k$ are now replaced by expressions of the form $(\beta, A\alpha)$. The most important classes of these forms are defined as follows:

If $(\beta, A\alpha) = (A\beta, \alpha)$ for all α, β in $\mathfrak{h}$, A is said to be "self-adjoint" (symmetric or Hermitian).

If $|(\alpha, A\alpha)| \leqslant k(\alpha, \alpha)$ for all α in $\mathfrak{h}$, A is said to be "bounded".

If $(\alpha, A\alpha) \geqslant 0$ for all α in $\mathfrak{h}$, the sign of equality holding only if $|\alpha| = 0$, A is said to be "positive definite".

The set of linear equations in canonical form,

$$\Sigma A_{kj} u_j + B_k = 0,$$

is now replaced by the operational equation

$$A\alpha + \beta = 0, \tag{7·1}$$

* (β, α) is called the "scalar product" of α by β, and $|\alpha|$ the positive value of $(\alpha, \alpha)^{\frac{1}{2}}$, is called the "length" of α. We make considerable use of the inequality $(\alpha, \beta)(\beta, \alpha) \leqslant (\alpha, \alpha)(\beta, \beta)$, which is derived from the principle that $(\alpha + t\beta, \alpha + t\beta)$ is non-negative for all complex numbers t. The "distance function" for the two vectors α and β is $|\alpha - \beta|$. It is assumed that there is a unique null-vector o, such that $|o| = 0$.

where β is a prescribed vector and α is the unknown vector. The generalization of Maxwell's reciprocal relations is provided by the equation

$$(\beta, A\alpha) = (A\beta, \alpha) \tag{7·2}$$

which implies that the "influence" operator A is self-adjoint. The condition of stability, viz.

$$\Sigma A_{jk} u_j u_k \geqslant 0$$

now becomes $(\alpha, A\alpha) > 0$ unless $|\alpha| = 0$, so that A is a positive definite operator.

There is, however, an important difference between the operational problem and the problem of a finite number of linear, algebraic equations. In the latter case we showed (equation (2·7)) that the positive definite character of the strain energy $U(u)$ implied that

$$0 < m \leqslant 2U(u)/\Sigma u_k^2 \leqslant \mu.$$

But it is not true that, because A is positive definite, then

$$0 < m \leqslant \frac{(\alpha, A\alpha)}{(\alpha, \alpha)} \leqslant \mu, \tag{7·3}$$

where m is positive and μ is finite. In fact m may be zero and μ may be infinite. We shall, however, discuss only the case in which the inequalities (7·3) are satisfied. Then A is bounded and may be described as *strictly positive*.*

8. The method of steepest descents

Since vectors in a Hilbert space have infinitely many components it is no longer possible to apply the method to successive relaxation, in which only one component of the vector approximating to α is changed at each stage of the process. The method of steepest descents can, however, be extended at once to the operational problem.

The analogue of the total energy $W(u)$ in the framework problem is now

$$W(\alpha) = \tfrac{1}{2}(\alpha, A\alpha) + \tfrac{1}{2}(\alpha, \beta) + \tfrac{1}{2}(\beta, \alpha). \tag{8·1}$$

From the axioms of the last section it follows that $W(\alpha)$ is a real function of α, and we shall now prove that it has a finite lower bound for all vectors

* Since we have assumed that $A\alpha$ exists for all α in $\mathfrak{h}$, the inequalities

$$(\alpha, A\alpha) \leqslant \mu(\alpha, \alpha) \quad \text{and} \quad |A\alpha| \leqslant \mu|\alpha|$$

will always hold if $\mathfrak{h}$ is *separable*.

α when β is prescribed. Since (α,β) and (β,α) are conjugate complex numbers, and A is positive definite, it follows that

$$\{(\alpha, A\alpha)+(\alpha,\beta)\}\{(\alpha, A\alpha)+(\beta,\alpha)\} \geqslant 0,$$

i.e.
$$2(\alpha, A\alpha)\; W(\alpha)+(\alpha,\beta)(\beta,\alpha)\geqslant 0.$$

Now
$$0 \leqslant (\alpha,\beta)(\beta,\alpha) \leqslant (\alpha,\alpha)(\beta,\beta),$$

and
$$(\alpha, A\alpha) \geqslant m(\alpha,\alpha).$$

Hence
$$W(\alpha) \geqslant -(\beta,\beta)/m. \tag{8·2}$$

The right-hand side of this inequality is independent of α, so that $W(\alpha)$ has a finite lower bound.

Next we consider the problem of improving any approximation α to the solution of the equation $A\alpha+\beta=0$, by replacing α by $\alpha+t\phi$, where ϕ is a prescribed vector and t an arbitrary real number. We have that

$$W(\alpha+t\phi) = W(\alpha)+W_1t+W_2t^2,$$

where
$$\begin{aligned} W_1 &= \tfrac{1}{2}(\alpha, A\phi)+\tfrac{1}{2}(\phi, A\alpha)+\tfrac{1}{2}(\beta,\phi)+\tfrac{1}{2}(\phi,\beta) \\ &= \tfrac{1}{2}(\gamma,\phi)+\tfrac{1}{2}(\phi,\gamma), \end{aligned}$$

$$\gamma = A\alpha+\beta,$$

and
$$W_2 = \tfrac{1}{2}(\phi, A\phi).$$

As before the best value of t, giving the greatest decrease in W, is

$$t = -\tfrac{1}{2}W_1/W_2,$$

and then
$$W(\alpha)-W(\alpha+t\phi) = \tfrac{1}{4}W_1^2/W_2.$$

As in the framework problem we shall take the vector ϕ to be determined by γ which now corresponds to the residual forces. In this case, $(\phi=\gamma)$,

$$W_1 = (\gamma,\gamma),$$

and
$$W_2 = \tfrac{1}{2}(\gamma, A\gamma).$$

Hence the next approximation derived from α is

$$\alpha+t\phi = \alpha-\frac{(\gamma,\gamma)}{(\gamma,A\gamma)}\gamma,$$

and the decrease in the total energy is

$$W(\alpha)-W(\alpha+t\phi) = \frac{(\gamma,\gamma)^2}{2(\gamma,A\gamma)}.$$

Starting with any arbitrary vector α_1, we therefore construct a sequence of vectors, $\alpha_1, \alpha_2, \ldots$, such that

$$\alpha_{p+1} = \alpha_p - \frac{(\gamma_p, \gamma_p)}{(\gamma_p, A\gamma_p)}\gamma_p, \tag{8·3}$$

where $$\gamma_p = A\alpha_p + \beta. \tag{8·4}$$

Then $$W(\alpha_p) - W(\alpha_{p+1}) = \frac{(\gamma_p, \gamma_p)^2}{2(\gamma_p, A\gamma_p)}. \tag{8·5}$$

It follows that the sequence $\{W(\alpha_p)\}$ is monotonic and decreasing. Now by (8·2) it has a finite lower bound. Hence this sequence must converge to a unique limit. Therefore

$$W(\alpha_p) - W(\alpha_{p+1}) \to 0 \quad \text{as} \quad p \to \infty,$$

i.e. by (8·5) $$\tfrac{1}{2}(\gamma_p, \gamma_p)^2/(\gamma_p, A\gamma_p) \to 0 \quad \text{as} \quad p \to \infty.$$

Now $$(\gamma_p, A\gamma_p) \leqslant \mu(\gamma_p, \gamma_p),$$

whence $$(\gamma_p, \gamma_p) \leqslant \mu(\gamma_p, \gamma_p)^2/(\gamma_p, A\gamma_p).$$

Therefore $$(\gamma_p, \gamma_p) \to 0 \quad \text{as} \quad p \to \infty. \tag{8·6}$$

The sequence $\{\gamma_p\}$ thus converges to the null-vector. To establish the convergence of the sequence $\{\alpha_p\}$ we note that

$$|\gamma_p - \gamma_q| \to 0 \quad \text{as} \quad p, q \to \infty.$$

Now, by (8·4), $$\gamma_p - \gamma_q = A(\alpha_p - \alpha_q). \tag{8·7}$$

Hence $$|A(\alpha_p - \alpha_q)| \to 0 \quad \text{as} \quad p, q \to \infty. \tag{8·8}$$

But $$m^2(\phi, \phi)^2 \leqslant (\phi, A\phi)^2$$
$$\leqslant (\phi, \phi)(A\phi, A\phi).$$

Therefore $$m|\phi| \leqslant |A\phi|$$

and $$m|\alpha_p - \alpha_q| \leqslant |A(\alpha_p - \alpha_q)|.$$

Hence, by (8·7), $$|\alpha_p - \alpha_q| \to 0 \quad \text{as} \quad p, q \to \infty.$$

Therefore the sequence $\{\alpha_p\}$ converges to a limit vector α.

It is not necessarily true that the sequence $\{A\alpha_p\}$ should converge to $\{A\alpha\}$, but a sufficient condition for the truth of this statement is clearly that

$$|A\alpha| \leqslant \mu|\alpha|, \tag{8·9}$$

—a more stringent form of the condition $(\alpha, A\alpha) \leqslant \mu(\alpha, \alpha)$, assumed in (7·3).

We shall assume that the relation (8·9) is verified. Then

$$|A\alpha+\beta| \leqslant |A\alpha_p+\beta| + |A(\alpha-\alpha_p)|$$
$$\leqslant |\gamma_p| + \mu|\alpha-\alpha_p|.$$

Hence
$$A\alpha+\beta = 0,$$

i.e. the limit vector α does satisfy the original equation (7·1).

The significance of the restrictions (7·3) and (8·9) upon the operator A can be exhibited by introducing the proper vectors ψ_k and the proper values λ_k of this operator. Then

$$A\psi_k = \lambda_k\psi_k,$$

whence
$$(\psi_k, A\psi_k)/(\psi_k, \psi_k) = \lambda_k,$$

and
$$|A\psi_k|/|\psi_k| = \lambda_k.$$

The conditions therefore imply that all the proper values λ_k lie in a finite range

$$0 < m \leqslant \lambda_k \leqslant \mu < \infty,$$

and, conversely, it can be shown that this restriction on the proper values is sufficient to ensure that the conditions imposed on A are fulfilled.

9. Gyrostatic systems

Similar methods of solution can be applied to any operational equation,

$$P\alpha+\beta = 0,$$

even if P is not self-adjoint, provided that it satisfies the conditions

$$0 < m(\phi,\phi) \leqslant \tfrac{1}{2}|(\phi, P\phi)+(P\phi,\phi)|$$

and
$$|P\phi| \leqslant \mu|\phi|, \tag{9·1}$$

for all vectors ϕ, m and μ being two positive numbers. We write

$$W(\alpha) = \tfrac{1}{2}|P\alpha+\beta|^2 = \tfrac{1}{2}(P\alpha, P\alpha)+\tfrac{1}{2}(P\alpha,\beta)+\tfrac{1}{2}(\beta,P\alpha)+\tfrac{1}{2}(\beta,\beta), \tag{9·2}$$

and define a sequence of approximations by the relations

$$\alpha_{p+1} = \alpha_p - \frac{\tfrac{1}{2}(\gamma_p, P\gamma_p)+\tfrac{1}{2}(P\gamma_p,\gamma_p)}{(P\gamma_p, P\gamma_p)}\gamma_p, \tag{9·3}$$

where
$$\gamma_p = P\alpha_p+\beta.$$

Then
$$W(\alpha_p)-W(\alpha_{p+1}) = \frac{[\tfrac{1}{2}(\gamma_p, P\gamma_p)+\tfrac{1}{2}(P\gamma_p,\gamma_p)]^2}{2(P\gamma_p, P\gamma_p)}. \tag{9·4}$$

It follows from (9·2) that $W(\alpha_p) \geqslant 0$, and from (9·4) that

$$W(\alpha_p) \geqslant W(\alpha_{p+1}).$$

Hence the sequence $\{W(\alpha_p)\}$ converges to a unique limit as $p \to \infty$. Hence

$$W(\alpha_p) - W(\alpha_{p+1}) \to 0 \quad \text{as} \quad p \to \infty.$$

Then, by (9·4) $$\frac{[(\gamma_p, P\gamma_p) + (P\gamma_p, \gamma_p)]^2}{(P\gamma_p, P\gamma_p)} \to 0 \quad \text{as} \quad p \to \infty.$$

Now, by (9·1)

$$\frac{m^2}{\mu^2} | \gamma_p |^2 = \frac{m^2(\gamma_p, \gamma_p)^2}{\mu^2 | \gamma_p |^2} \leqslant \frac{| (\gamma_p, P\gamma_p) + (P\gamma_p, \gamma_p) |^2}{(P\gamma_p, P\gamma_p)}.$$

Hence $$| \gamma_p | \to 0 \quad \text{as} \quad p \to \infty.$$

To establish the convergence of the sequence $\{\alpha_p\}$ we note that, since

$$| \alpha + t\beta |^2 \equiv (\alpha, \alpha) + t(\alpha, \beta) + t(\beta, \alpha) + t^2(\beta, \beta)$$

is non-negative for all real t, therefore

$$[(\alpha, \beta) + (\beta, \alpha)]^2 \leqslant 4(\alpha, \alpha)(\beta, \beta).$$

By using (9·1), we therefore find that

$$\begin{aligned} | \phi |^4 &\leqslant | (\phi, P\phi) + (P\phi, \phi) |^2/(4m^2) \\ &\leqslant | \phi |^2 | P\phi |^2/m^2. \end{aligned}$$

Hence $$| \phi | \leqslant | P\phi |/m.$$

Therefore $$\begin{aligned} m | \alpha_p - \alpha_q | &\leqslant | P\alpha_p - P\alpha_q | \\ &= | \gamma_p - \gamma_q |, \quad \text{by (9·3)}. \end{aligned}$$

But the sequence $\{\gamma_p\}$ is convergent. Hence the sequence $\{\alpha_p\}$ is also convergent.

If α is the limit of the sequence $\{\alpha_p\}$ then, as in § 8, $P\alpha$ is the limit of the sequence $\{P\alpha_p\}$, and

$$\begin{aligned} | P\alpha + \beta | &\leqslant | P\alpha_p + \beta | + | P\alpha - P\alpha_p | \\ &\leqslant | \gamma_p | + \mu | \alpha - \alpha_p |. \end{aligned}$$

Hence $$P\alpha + \beta = 0,$$

i.e. α satisfies the equation originally proposed.

III. LINEAR INTEGRAL EQUATIONS

10. Fredholm's equation of the first kind

One of the simplest examples of a Hilbert space $\mathfrak{h}$ is furnished by the set of all real, measurable functions $f(x)$ of a variable x, defined in an interval $a \leqslant x \leqslant b$, and such that the integral

$$\int_a^b f^2(x)\,dx$$

exists as a Lebesgue integral. A function of this set can be regarded as a vector with a continuously infinite number of components, and two such functions must be regarded as equivalent if they are equal almost everywhere in (a, b). The "scalar product" of two functions $\alpha(x)$ and $\beta(x)$ is given by

$$(\beta, \alpha) = (\alpha, \beta) = \int_a^b \alpha(x)\,\beta(x)\,dx.$$

The space of these functions is "complete", for, by a well-known theorem due to F. Riesz, the condition

$$\int_a^b [\alpha_p(x) - \alpha_q(x)]^2\,dx \to 0, \quad \text{as} \quad p, q \to \infty$$

is sufficient to ensure the existence of a limit $\alpha(x)$ such that

$$\int_a^b [\alpha(x) - \alpha_p(x)]^2\,dx \to 0 \quad \text{as} \quad p \to \infty,$$

whence $\alpha_p(x) \to \alpha(x)$ almost everywhere in (a, b).

In this realization of the abstract Hilbert space $\mathfrak{h}$ a linear operator A can be defined by the equation

$$A\alpha(x) = \int_a^b A(x, s)\,\alpha(s)\,ds,$$

where the nucleus $A(x, s)$ is a measurable function defined in the region $a \leqslant x$, $s \leqslant b$. The operational equation (7·1) now takes the form

$$\int_a^b A(x, s)\,\alpha(s)\,ds + \beta(x) = 0, \tag{10·1}$$

where $\beta(x)$ and $A(x, s)$ are prescribed functions and $\alpha(x)$ is unknown. This is an integral equation of Fredholm's type and of the first kind.

The conditions to which we have subjected the operator A, equations (7·3) and (8·9), now imply that, for all functions $\alpha(x)$ considered,

$$\left.\begin{aligned} &\int_a^b \alpha(x)\,dx \int_a^b A(x,s)\,\alpha(s)\,ds \geqslant m \int_a^b \alpha^2(x)\,dx, \\ \text{and} \qquad &\int_a^b \left\{\int_a^b A(x,s)\,\alpha(s)\,ds\right\}^2 dx \leqslant \mu^2 \int_a^b \alpha^2(x)\,dx, \end{aligned}\right\} \tag{10·2}$$

where m is a positive number and μ is finite.

We shall first consider the case when A is self-adjoint. It is easily seen that this implies that the nucleus $A(x,s)$ is symmetric in x and s. In these circumstances the analogue of the total energy is given by

$$W(\alpha) = \tfrac{1}{2}\int_a^b \int_a^b A(x,s)\,\alpha(x)\,\alpha(s)\,dx\,ds + \int_a^b \alpha(x)\,\beta(x)\,dx, \tag{10·3}$$

and the sequence of approximations by the methods of steepest descents is given by the formulae—

$$\left.\begin{aligned} \gamma_p(x) &= \int_a^b A(x,s)\,\alpha_p(s)\,ds + \beta(x), \\ \text{and} \qquad \alpha_{p+1}(x) &= \alpha_p(x) - \frac{\gamma_p(x)\displaystyle\int_a^b \gamma_p^2(s)\,ds}{\displaystyle\int_a^b \int_a^b A(x,s)\,\gamma_p(x)\,\gamma_p(s)\,dx\,ds}. \end{aligned}\right\} \tag{10·4}$$

The general argument of § 8 then shows that the sequence of functions $\{\alpha_p(x)\}$ converges almost everywhere to a function $\alpha(x)$ which satisfies the equation (10·1) almost everywhere.

If A is not self-adjoint the formulae are slightly more complicated. Using the method of § 9, and defining $\gamma_p(x)$ as in (10·4), we now define the sequence of approximations by the formula—

$$\frac{\alpha_{p+1}(x) - \alpha_p(x)}{\gamma_p(x)} = \frac{\displaystyle\int_a^b \int_a^b A(x,s)\,\gamma_p(x)\,\gamma_p(s)\,dx\,ds}{\displaystyle\int_a^b \left\{\int_a^b A(x,s)\,\gamma_p(s)\,ds\right\}^2 dx}. \tag{10·5}$$

The sequence $\{\alpha_p(x)\}$ then converges in the same way as before.

IV. LINEAR DIFFERENTIAL EQUATIONS

11. The reduction of a linear differential system to an infinite system of linear algebraic equations

Although the application of the general operational theory of Part II to linear integral equations in Part III is so simple and direct, there are unavoidable difficulties in the application to linear differential equations. These difficulties arise from the fact that, if A is a "formal linear differential operator" (Stone 1932, p. 112) the class of functions $\mathfrak{F}$ to which A can be applied is necessarily restricted by conditions of continuity and differentiability, and, moreover, even if the function α belongs to the class $\mathfrak{F}$, this is by no means necessarily true of the function $A\alpha$. To avoid the difficulties of the direct application of the methods of Part II we shall therefore begin by replacing the proposed differential equation by a system of linear algebraic equations in an infinite number of unknown variables.

We consider a formal linear differential operator A of the form

$$A = \sum_{k=0}^{n} p_{n-k}(x)\frac{d^k}{dx^k}, \quad (a \leqslant x \leqslant b),$$

with which we associate a set of linear boundary operators of the form

$$M_k(f) = \sum_{j=1}^{n} \{a_{kj}f^{(j-1)}(a) + b_{kj}f^{(j-1)}(b)\}, \quad k = 1, 2, \ldots, n,$$

where $f^j(x) = df^j(x)/dx$, and a_{kj}, b_{kj} are numerical constants. The fundamental problem is then to solve the differential equation

$$\left.\begin{array}{c} A\alpha(x) + \beta(x) = 0, \\ \text{subject to the homogeneous boundary conditions} \\ M_k(\alpha) = 0, \end{array}\right\} \tag{11·1}$$

$\beta(x)$ being a prescribed function of x.

We shall assume that the differential system (11·1) satisfies conditions analogous to (7·3), which are most conveniently taken in the form

$$0 < m\int_a^b f^2(x)\,dx \leqslant \int_a^b \{Af(x)\}^2\,dx \leqslant \mu\int_a^b f^2(x)\,dx, \tag{11·2}$$

for all functions $f(x)$ satisfying the boundary conditions

$$M_k(f) = 0,$$

m being a positive number and μ being finite.

To define a space $\mathfrak{F}$ for the operator A we introduce an infinite set of functions $\{\phi_n(x)\}$ with the following properties:

(i) Each function satisfies the boundary conditions

$$M_k\{\phi_n\} = 0, \quad k = 1, 2, \ldots, n;$$

(ii) The functions are orthogonal and normal for the range (a, b),

$$\int_a^b \phi_j(x)\,\phi_k(x)\,dx = \begin{matrix} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k; \end{matrix}$$

(iii) The set of functions is complete, i.e. if

$$\int_a^b f(x)\,\phi_n(x)\,dx = 0 \quad \text{for } n = 1, 2, \ldots,$$

then $f(x)$ is zero almost everywhere in (a, b);

(iv) For each function $A\phi_n$ exists;

(v) The series $\sum_{n=1}^{\infty} \int_a^b (A\phi_n)^2\,dx$ is convergent.

We then define the space $\mathfrak{F}$ to be the set of functions $f(x)$ which can be represented by series of the form

$$\sum_{n=1}^{\infty} c_n \phi_n(x) \tag{11·3}$$

in such a way that

$$\int_a^b \left\{ f(x) - \sum_{n=1}^{p} c_n \phi_n(x) \right\}^2 dx \to 0 \quad \text{as } p \to \infty,$$

and we write

$$f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n(x).$$

The series which represents $f(x)$ will converge to $f(x)$ almost everywhere in (a, b).

As in Part III we shall take the scalar product of any two functions $f(x)$, $g(x)$ to be

$$(f, g) = (g, f) = \int_a^b f(x)\,g(x)\,dx,$$

and we shall write $|f|$ for the positive square root of (f, f).

We require three properties of the representation of functions by series of the form (11·3).* Let the sum of the first p terms in the series representing $f(x)$ be denoted by $f_p(x)$, so that

$$f_p(x) = \sum_{n=1}^{p} c_n \phi_n(x),$$

* Hobson (1926, 2, 245–49). The series considered in this section have "strong convergence" with exponent 2.

and $$|f-f_p|\to 0, \quad \text{as } p\to\infty.$$

Then (i) $$|f_p|\to|f|, \quad \text{as } p\to\infty; \tag{11·4}$$

(ii) $$(g,f_p)\to(g,f), \quad \text{as } p\to\infty, \tag{11·5}$$

if g belongs to $\mathfrak{F}$; and

(iii) $$Af(x)\sim\Sigma c_n A\phi_n(x). \tag{11·6}$$

The first and second results are well known. To establish the third result we note that, by (11·2),

$$|Af-Af_p|\leqslant\mu^{\frac{1}{2}}|f-f_p|,$$

whence $$|Af-Af_p|\to 0, \quad \text{as } p\to\infty,$$

i.e. $$\int_a^b\left\{Af(x)-\sum_{n=1}^{p}c_n A\phi_n(x)\right\}^2 dx\to 0, \quad \text{as } p\to\infty.$$

We now look for solutions of the differential system (11·1) of the form

$$\alpha(x)\sim\sum_{n=1}^{\infty}a_n\phi_n(x).$$

Since $$A\alpha(x)+\beta(x)\sim\Sigma a_n A\phi_n(x)+\beta(x),$$

it follows from (11·1) that

$$\int_a^b\left\{\sum_{n=1}^{p}a_n A\phi_n(x)+\beta(x)\right\}^2 dx\to 0 \quad \text{as } p\to\infty. \tag{11·7}$$

We shall write $$A_{jk}=\int_a^b A\phi_j(x)\,A\phi_k(x)\,dx \tag{11·8}$$

and $$b_k=\int_a^b\beta(x)\,A\phi_k(x)\,dx.$$

Then (11·7) becomes

$$\sum_{j,k=1}^{p}A_{jk}a_j a_k+2\sum_{k=1}^{p}b_k a_k+\int_a^b\beta^2(x)\,dx\to 0, \quad \text{as } p\to\infty.$$

Thus if we write $$W(\alpha)=\tfrac{1}{2}\sum_{j,k=1}^{\infty}A_{jk}a_j a_k+\sum_{k=1}^{\infty}b_k a_k, \tag{11·9}$$

it follows from (11·4), (11·5) and (11·6) that the quadratic form $W(\alpha)$ attains its minimum value, viz.

$$W(\alpha)_{\min}=-\tfrac{1}{2}\int_a^b\beta^2(x)\,dx,$$

when the constants $a_1, a_2, \ldots$ have the values appropriate to the representation of $\alpha(x)$ in the form $\alpha(x)\sim\sum_{n=1}^{\infty}a_n\phi_n(x)$.

The conditions that $W(\alpha)$ should be a minimum are clearly

$$\sum_{j=1}^{\infty} A_{jk} a_j + b_k = 0. \tag{11·10}$$

The original differential system (11·1) has therefore been replaced by an infinite set of linear algebraic equations in the variables $a_1, a_2, \ldots$.

Finally we show that this set of equations is in canonical form. It is obvious from (11·8) that

$$A_{jk} = A_{kj}.$$

Moreover,

$$\sum_{j,k=1}^{p} A_{jk} a_j a_k = \int_a^b \left\{ \sum_{n=1}^{p} a_n A\phi_n(x) \right\}^2 dx$$

$$= \int_a^b \{Af_p(x)\}^2 dx, \quad \text{say.}$$

Now

$$\int_a^b \{Af_p(x)\}^2 dx \geqslant m \int_a^b f_p^2(x)\, dx \quad \text{by (11·2)},$$

$$= m \int_a^b \left\{ \sum_{n=1}^{p} a_n \phi_n(x) \right\}^2 dx$$

$$= m \sum_{n=1}^{p} a_n^2,$$

since the functions $\{\phi_n(x)\}$ form an orthonormal system. Hence

$$m \sum_{n=1}^{\infty} a_n^2 \leqslant \sum_{n=1}^{\infty} A_{jk} a_j a_k. \tag{11·11}$$

Also,

$$\sum_{n=1}^{p} A_{jk} a_j a_k = \int_a^b \{Af_p(x)\}^2 dx$$

$$\leqslant \mu \int_a^b f_p(x)\, dx \quad \text{by (11·2)},$$

$$= \mu \sum_{n=1}^{p} a_n^2, \quad \text{as before.}$$

Therefore

$$\sum_{n=1}^{\infty} A_{jk} a_j a_k \leqslant \mu \sum_{n=1}^{\infty} a_n^2. \tag{11·12}$$

12. The solution of the infinite system of equations by relaxation methods

Since the infinite system of equations (11·10) is in canonical form we can apply the method of steepest descents as given in § 8. The vector space $\mathfrak{h}$ now consists of all sets of numbers,

$$\alpha = (a_1, a_2, \ldots, a_n, \ldots)$$

such that the series $\sum_{n=1}^{\infty} a_n^2$ is convergent. The scalar product of the vector α by the vector

$$\beta = (b_1, b_2, \ldots, b_n, \ldots)$$

is now $$(\beta, \alpha) = (\alpha, \beta) = \sum_{n=1}^{\infty} a_n b_n = \int_a^b \alpha(x)\,\beta(x)\,dx,$$

since the functions $\{\phi_k(x)\}$ form an orthonormal system. The operator A is specified by the infinite set of numbers A_{jk} as defined in (11·7).

Hence to solve the equations (11·10) we start with any arbitrary vector

$$\alpha_1 = (a_1^{(1)}, a_2^{(1)}, \ldots, a_n^{(1)}, \ldots)$$

and form the sequence of vectors defined by the recurrence relation

$$\alpha_{p+1} - \alpha_p = -\frac{(\gamma_p, \gamma_p)}{(\gamma_p, A\gamma_p)}\gamma_p,$$

where $$\gamma_p = A\alpha_p + \beta.$$

By reasoning based on the quadratic form $W(\alpha)$ of equation (11·9) we can show, as in § 8, that, as $p \to \infty$,

(i) $\{W(\alpha_p)\}$ converges to a unique limit,

(ii) $|\gamma_p| \to 0$,

(iii) α_p converges to a limit vector α.

It is not necessary to employ the condition (8·9) utilized before, for we can now proceed as follows. Let

$$\alpha(x) \sim \sum_{n=1}^{\infty} a_n \phi_n(x),$$

and $$\alpha_p(x) \sim \sum_{n=1}^{\infty} a_n^{(p)} \phi_n(x).$$

Now $$|A\alpha - A\alpha_p| \leqslant \mu^{\frac{1}{2}} |\alpha - \alpha_p|, \quad \text{by (11·2).}$$

Therefore, $$|A\alpha - A\alpha_p| \to 0 \quad \text{as } p \to \infty,$$

and $A\alpha_p$ converges to $A\alpha$ almost everywhere in (a, b). It now follows as in § 8 that the limit function $\alpha(x)$ must satisfy the equation $A\alpha(x) + \beta(x) = 0$ almost everywhere. Finally since $\alpha(x)$ is represented by a series of which each term, $a_k\phi_k(x)$, satisfies the boundary conditions (11·1), so also must $\alpha(x)$ itself. We have thus obtained a complete solution of the problem.

13. Conclusion

We have given a general account of the application of relaxation methods to the solution of linear equations. Southwell's original method for linear algebraic equations in a finite number of unknown variables has been the inspiration of the whole investigation. By replacing his method of successive relaxation by our method of steepest descents we have been able to attack linear operational equations, linear integral equations and linear differential equations. In each case our methods give successive approximations which converge to the accurate solution of the problem. Two outstanding problems must be left for future investigation—the numerical solution of particular problems of physics and engineering, and the extension of the methods to non-linear equations.

References

Atkinson, R. J., Bradfield, K. N. E. and Southwell, R. V. 1937 *Rep. Memor. Aero. Res. Comm., Lond.*, p. 1822.
Black, A. N. and Southwell, R. V. 1938 *Proc. Roy. Soc.* A, **164**, 447–67.
Bradfield, K. N. E. and Southwell, R. V. 1937 *Proc. Roy. Soc.* A, **161**, 155–81.
Hobson, E. W. 1926 "Theory of Functions of a Real Variable."
Richards, J. C. 1937 *J. Inst. Civ. Eng.*, **5**, 196–216.
Southwell, R. V. 1935 *Proc. Roy. Soc.* A, **151**, 56–95.
Stone, M. H. 1932 "Linear Transformations in Hilbert Space."