REGULARIZATION OF DIFFERENCE SCHEMES*

A.A. SAMARSKII

Moscow

(*Received* 1 *April* 1966)

WE examine the possibilities of transforming or regularizing schemes in such a way that the new schemes are stable and satisfy auxiliary requirements as regards accuracy and economy.

Difference schemes are treated as operator equations in real linear normed space [1, 2]. Two- and three-layer schemes are discussed in real Hilbert space (more precisely, in unitary space, since no use is made of completeness). Stable schemes are classified. Schemes are written in a canonical form that enables stability operators, or what will be termed regularizers, to be introduced. Sufficient conditions for stability only weakly restrict the arbitrary selection of these stability operators. By varying the regularizers but remaining in the class of stable schemes, economic schemes of bounded order of accuracy can be obtained.

All the regularization schemes discussed have in common the use of energy-wise equivalent (en. eq.) operators as regularizers (similar operators are employed in linear algebra [3], see also [4, 16]).

All the schemes familiar in the literature can be treated as schemes obtained from natural explicit schemes by means of some regularization procedure.

We only need to mention the widely used implicit schemes. As a rule, the same operator is taken (as regularizer) on the upper as on the lower layer (or on a smooth part of it).

Another type of regularization is typified by the explicit three-layer rhombus scheme [5] (the Dufort - Frankel scheme) and the asymmetric two-layer scheme of V.K. Saul'ev [6] for the equation of heat conduction.

Both have the following characteristic properties: (1) the schemes are
unconditionally stable, (2) explicit formulae are used for computation,
and (3) the order of approximation is lower than for conventional ex-
plicit schemes. The rhombus scheme is obtained from the absolutely un-
stable Richardson scheme by replacing $y_i{}^j$ by $0.5(y_i{}^{j+1} + y_i{}^{j-1})$. Surprise
is expressed in Wasow and Forsythe's book [5] that such a slight modifica-
tion should transform an absolutely unstable into an unconditionally
stable scheme. It will be shown in § 3 of Section 3 below that trans-
formation of a Richardson to a rhombus scheme is equivalent to introduc-
ing an elementary type of regularizer. The type of regularizer for the
asymmetric scheme is described in § 4. Multidimensional analogues of
these schemes can be used as iterative schemes for solving elliptic equa-
tions with variable coefficients in arbitrary regions [7].

Our theory of the regularization of difference schemes starts with
sufficient conditions for stability and *a priori* estimates for two- and
three-layer schemes in real Hilbert space (Theorems 1 - 4). These
theorems will be proved in a later paper.

A general principle for regularization of two- and three-layer schemes
is laid down in Section 3, using similitude operators. Some examples of
selecting en. eq. operators for 2nd order elliptic operators are given
in Section 4. The regularizer R must be arranged in each specific case
so as to simplify the inversion of the operator to the upper layer and
retain the required order of approximation.

Factorized schemes (FS), i.e. those in which the operator is repre-
sented on the upper layer as the product of a finite number of operators
(factorized), are considered in §§ 4 and 5.

Numerous interesting papers have appeared in recent years on economic
difference methods for solving multidimensional problems of mathematical
physics, e.g. by Peaceman, Reckford, Douglas, Baker, Oliphant, V.K.
Saul'ev, N.N. Yanenko, E.G. D'yakonov, S.K. Godunov and V.B. Andreev
(the necessary references may be found in [2, 5 - 11]. In spite of differ-
ences in approach, in symbolism and terminology (method of alternating
directions, splitting method, scheme with splitting operator), many of
the difference schemes proposed by these authors have the following in
common: after reduction to the canonical form, the resulting schemes are
factorized, and by definition, approximate to the initial differential
equation. The domain of variation of the space variables $x_1, x_2, \ldots, x_p$
is a p-dimensional parallelepiped.

A large number of economic FS for the equations and systems of equa-
tions of mathematical physics have been investigated by E.G. D'yakonov

(see e.g. [9, 14], the latter with a list of references). The initial
schemes approximating the differential equations and considered by
D'yakonov were two- and three-layer FS of a special type, which he
termed splitting operator schemes. These schemes were shown to be stable
and convergent under some auxiliary restrictions (a fine mesh, smooth
coefficients, conditions beyond the natural conditions of ellipticity on
the matrix of coefficients of the space operator when mixed derivatives
are present, etc.). It was noted in [2] that these restrictions need to
be removed in the case of equations with discontinuous coefficients, and
in particular, the question of the applicability of FS (splitting
schemes) investigated. V.B. Andreev has proposed two- and three-layer FS
for parabolic equations with mixed derivatives, stable under a natural
parabolicity condition.

Paragraphs 4 and 5, which represent a development of [2, 10], give a
general method for constructing FS by regularization of an initial
scheme, whose regularizer is the sum of a finite number of operators:

$R = \sum_{\alpha=1}^{m} R_\alpha.$ Sufficient conditions for stability of the FS are given in

the form of conditions on the operators R_α, which distinguish the class
of stable FS. This class includes the family of familiar FS (including
schemes with a splitting operator).

The FS is shown to be stable on any sequence of meshes provided the
initial scheme has this property, and in addition, $\{R_\alpha\}$ is a system of
mutually commuting selfadjoint positive operators. This leads to an ex-
tremely wide class of unconditionally stable FS.

By using the en. eq. operators described in Section 4, a family of
unconditionally stable economic FS is obtained for parabolic and hyper-
bolic equations and systems of equations with variable, and even dis-
continuous, coefficients. For instance, in the case of a system of para-
bolic equations with mixed derivatives, the regularizer is a difference
elliptic operator with a diagonal matrix and constant coefficients. The
solution of a difference problem then amounts to successive application
of the formulae of one-dimensional successive substitution [9]. The re-
sulting FS are stable whatever τ and h under natural parabolicity con-
ditions, and are convergent in the class of discontinuous coefficients
(provided the coefficients have discontinuities of the 1st kind on hyper-
planes parallel to the coordinate hyperplanes). By selecting special
mesh sequences $\omega_h(k)$ [12], schemes can be obtained with the accuracy
$O(h^2 + \tau)$ and $O(h^2 + \tau^2)$ (three-layer schemes) in the class of discon-
tinuous coefficients. Only second order equations in $x_1, \ldots, x_p$ are

considered here. The theory of FS (see Section 3) is also applicable to higher order equations (this is obvious, e.g. in the case of periodic boundary conditions).

Every difference operator A can be associated with at least one en. eq. operator of simpler structure, which can be used as a regularizer for two- and three-layer schemes. Computational practice should benefit considerably from the compilation of a list of regularizers and rules for selecting the best. The same regularizer can be used for various elliptic operators, thereby enabling a standard program to be used for solving classes of problems. The program for finding the solution on a new layer remains unchanged, while the specific form of the problem operator is allowed for when computing the right-hand side.

Regularization methods can be used to construct economic universal algorithms for stationary problems, or more exactly, for the equations of the first kind

$$Au = f,$$

where A is a linear selfadjoint, positive definite operator on H. Classes of two- and three-layer iterative schemes with one or more relaxation parameters are considered here. The general stability theorems of Section 2 enable the optimal relaxation parameters and convergence rate estimates to be obtained. These latter results will be described in a separate paper.

1. Schemes in abstract spaces

1. SCHEMES AS OPERATOR EQUATIONS

Difference schemes will be considered as operator equations in abstract spaces. Let $\{H_N, \; V = 1, \; 2, \; \ldots\}$ be a sequence of linear normed spaces, representing the analogue of the spaces of mesh functions defined on a sequence of meshes in Euclidean space (finite- or infinite-dimensional). In the interval $0 \leqslant t \leqslant t_0$, introduce the mesh $\bar{\omega}_\tau = \{t_j = j\tau, \; j = 0, \; 1, \; \ldots, \; j_0, \; \tau = t_0/j_0\}$ with mesh interval τ. Let $y_N(t)$ be an abstract function of the argument t with values in H_N, so that, with fixed $t = j\tau \in \bar{\omega}_\tau$, $y_N(t)$ is a point (vector, element) of space H_N. We shall be concerned throughout with linear operators $A(t)$, $B(t)$, $R(t)$, $S(t)$ etc. (in general, unbounded), depending on the parameter $t \in \bar{\omega}_\tau$ and mapping H_N into H_N for each $t \in \bar{\omega}_\tau$. In general, all operators depend on N and τ, so that e.g. $A = A(V, \; \tau; \; t)$; the arguments V, τ will be

omitted when no confusion can arise.

Let B, S_1, S_2, ..., S_{r-1} be linear operators given in H_N. By analogy with [13], we shall describe as an r-layer (r-point) linear scheme an operator equation $By^{j+1} = S_1 y^j + S_2 y^{j-1} + \ldots + S_{r-1} y^{j-r+2} + \tau \varphi^{j+1}$, $j \geqslant r - 1$, $y^{j'} = y(t_j') \in H_N$, connecting the values of $y(t) \in H_N$ on the r layers $t = t_{j-r+1}, \ldots, t = t_{j+1}$. The initial values are given for $j = 0, 1, \ldots, r - 2$.

Only two- and three-layer schemes will be discussed.

2. CANONICAL FORM OF TWO-LAYER SCHEMES

First take the ordinary two-layer scheme

$$By^{j+1} = Cy^j + \tau \varphi^{j+1}, \qquad j = 0, 1, \ldots, \qquad \varphi = \varphi(t) \in H_N \text{ for } t \in \omega_\tau. \quad (1.1)$$

The initial condition

$$y^0 = y(0) = y_0 \in H_N.$$

is given for $j = 0$. We introduce the notation

$$y^{j+1} = \hat{y}, \qquad y^j = y(t) = y, \qquad y_t = \frac{y^{j+1} - y^j}{\tau} = \frac{\hat{y} - y}{\tau}$$

and rewrite (1.1) as

$$By = Cy + \tau \varphi, \qquad 0 \leqslant t \in \omega_\tau, \qquad y(0) = y_0.$$

Substituting $\hat{y} = y + \tau y_t$, and putting $B = C + \tau A$, the two-layer scheme is obtained in the canonical form

$$By_t + Ay = \varphi, \qquad 0 \leqslant t \in \omega_\tau, \qquad y(0) = y_0. \quad (1.2)$$

Writing the operator B as the sum

$$B = E + \tau R,$$

the second canonical form

$$y_t + \tau R y_t + Ay = \varphi, \qquad 0 \leqslant t \in \omega_\tau, \qquad y(0) = y_0. \quad (1.3)$$

is obtained.

3. CANONICAL FORMS OF THE THREE-LAYER SCHEME

Three-layer schemes connect the values of the required function $y(t) \in H$ ($t \in \bar{\omega}_\tau$) on 3 layers $t = t_{j+1}$, $t = t_j$, $t = t_{j-1}$. Consider any linear three-layer scheme

$$B_1 y^{j+1} + B_2 y^j + B_3 y^{j-1} = \tau \varphi^j, \qquad j = 1, 2, \ldots, \tag{1.4}$$

where $B_s = B_s(V, \tau; t)$ ($s = 1, 2, 3$, $t = t_j$ are linear operators given on H_N.

The following index-less notation will be used in future:

$$\hat{y} = y^{j+1}, \qquad y = y(t) = y^j, \qquad \check{y} = y(t - \tau) = y^{j-1}, \qquad y_t = \frac{1}{\tau}(\hat{y} - y),$$

$$y_{\bar{t}} = \frac{1}{\tau}(y - \check{y}),$$

$$y_t = \hat{y}_{\bar{t}}, \qquad y_{\bar{t}} = \check{y}_t, \qquad y_{\mathring{t}} = \frac{1}{2}(y_t + y_{\bar{t}}) = \frac{\hat{y} - \check{y}}{2\tau},$$

$$y_{\bar{t}t} = \frac{1}{\tau}(y_t - y_{\bar{t}}) = \frac{1}{\tau^2}(\hat{y} - 2y + \check{y}).$$

A three-layer scheme (1.4) can be written in the canonical form

$$B y_{\mathring{t}} + \tau^2 R y_{\bar{t}t} + A y = \varphi(t), \qquad \tau \leqslant t \in \omega_\tau, \tag{1.5}$$

where B, R and A are linear operators from H into H.

For substituting

$$\hat{y} = y + \tau y_{\mathring{t}} + \frac{1}{2}\tau^2 y_{\bar{t}t}, \qquad \check{y} = y - \tau y_{\mathring{t}} + \frac{1}{2}\tau^2 y_{\bar{t}t}$$

in (1.4) and denoting

$$(B_1 - B_3)\tau = B, \qquad B_1 + B_2 + B_3 = A, \qquad \frac{1}{2}(B_1 + B_3) = R,$$

(1.5) follows.

The scheme (1.5) is written with $t = \tau$, 2τ, ... The initial conditions $y(0) = y_0$, $y(\tau) = y_1$ or $y_t(0) = \bar{y}_0$, y_0, $\bar{y}_0$, $y_1 \in H$ must be given for $t = 0$ and $t = \tau$.

Note. If a three-layer scheme is written for a first-order differential equation in t, only the initial condition $y(0)$ will be given. To find $y(\tau)$, either the initial equation with $t = 0$ is employed, or a

two-layer scheme (see [10]) of the type $By_t + \Lambda y = \varphi$ for $t = 0$. Since $y = y(0)$ is known, it can be assumed in this case that $y_t(0)$ ($y_t(0) = B^{-1}(\varphi - \Lambda y_0)$) is given.

The problem to be considered is thus

$$By_{\hat{t}} + \tau^2 R y_{\bar{t}t} + Ay = \varphi, \qquad \tau \leqslant t \in \bar{\omega}_\tau, \qquad y(0) = y_0, \qquad y_t(0) = \bar{y}_0$$
$$\text{or } y(\tau) = y_1, \qquad\qquad (1.6)$$

where $y = y(t)$, $\hat{y} = y(t + \tau)$, $\check{y} = y(t - \tau)$.

Formally replacing $\tau^2 R$ by $E + \tau^2 R$, we get the second canonical form of the three-layer scheme

$$y_{\bar{t}t} + \tau^2 R y_{\bar{t}t} + By_{\hat{t}} + Ay = \varphi, \qquad 0 < t \in \omega_\tau, \qquad y(0), \; y(\tau) \text{ given.} \quad (1.7)$$

4. SOME CONCEPTS OF THE THEORY OF DIFFERENCE SCHEMES

By analogy with [1, 13], the following concept of correctness of a difference scheme can be introduced.

The introduction of different norms into a linear set H produces different normed spaces consisting of the same elements. The solutions of problems (1.2) and (1.6) and the right-hand side φ will in general be regarded as elements of distinct spaces that differ only in their norms. Let $\|\cdot\|_{(1)}, \|\cdot\|_{(2)}, \ldots$ be norms on the linear set H.

Problem (1.2) is correctly posed (scheme (1.2) is correct) if numbers τ_0, V_0 can be indicated such that, for $\tau \leqslant \tau_0$, $V \geqslant V_0$:

(1) a solution of problem (1.2) exists for any $y_0 \in H$ and $\varphi(t) \in H$, $t \in \omega_\tau$;

(2) the solution of problem (1.2) is a continuous function of y_0 and φ uniformly with respect to V and τ; the function may be, for instance,

$$\|y(t)\|_{(1)} \leqslant M_1 \|y_0\|_{(1)} + M_2 \max_{\omega_\tau} \|\varphi(t)\|_{(2)}. \qquad (1.8)$$

Here, and throughout what follows, M_1, M_2, M, c_1, c_2, $\ldots$ are positive constants, independent of τ and V.

If the solution of problem (1.2) depends continuously (uniformly with respect to V, τ) on y_0, φ, the scheme is said to be stable with respect

to the initial data and with respect to its right-hand side. More stringent conditions may be imposed on the stability of the right-hand side: the solution of problem (1.2) depends continuously on $\varphi_t(t)$ as well as on $\varphi(t)$. The following definition of correctness will therefore be used below:

Scheme (1.2) is correct, if, for $\tau \leqslant \tau_0$ and $V \geqslant V_0$,

(1) problem (1.2) is solvable, whatever $y_0 \in I$, $\varphi(t) \in I$,

(2)

$$\|y(t)\|_{(1)} \leqslant M_1 \|y_0\|_{(1)} + M_2 \max_{\omega_\tau} \|\varphi(t)\|_{(2)} + M_2' \max_{\omega_\tau} \|\varphi_t(t)\|_{(2^*)}, \qquad (1.9)$$

where $M_2' \geqslant 0$ is independent of V and τ.

Generally speaking, $\|\cdot\|_{(2)}$ and $\|\cdot\|_{(2^*)}$ are different norms.

A similar concept of correctness can be developed for three-layer schemes.

Returning to scheme (1.2), it is solvable if the inverse operator B^{-1} exists. In this case,

$$\hat{y} = B^{-1}Sy + \tau B^{-1}\varphi, \qquad S = B - \tau A.$$

Hence

$$\|\hat{y}\| \leqslant \|B^{-1}S\| \|y\| + \tau \|B^{-1}\varphi\|. \qquad (1.10)$$

The existence of B^{-1} and the condition

$$\|B^{-1}S\| \leqslant 1 + c_1\tau, \qquad c_1 = \text{const} > 0, \qquad (1.11)$$

are obviously sufficient for stability of scheme (1.2) provided

$$\|y\|_{(1)} = \|y\|, \qquad \|\varphi\|_{(2)} = \|B^{-1}\varphi\|.$$

For, from (1.10), (1.11),

$$\|\hat{y}\| \leqslant (1 + c_1\tau) \|y\| + \tau \|B^{-1}\varphi\|,$$

$$\|y^j\| \leqslant \exp(c_1 t_j) \left[\|y_0\| + \sum_{j'=1}^{j} \tau \|B^{-1}\varphi\|^{j'} \right]. \qquad (1.12)$$

Comparison with (1.9) shows that (1.9) is satisfied, where

$$M_1 = \exp{(c_1 t_0)}. \qquad M_2 = M_1 t_0.$$

If it is required that $\|B^{-1}\|$:

$$\|B^{-1}\| \leqslant c_2, \tag{1.13}$$

(1.12) gives

$$\|y^j\| \leqslant \exp{(c_1 t_j)} \left[\|y_0\| + c_2 \sum_{j'=1}^{j} \tau \|\varphi^{j'}\| \right]. \tag{1.14}$$

The definitions of correctness and the sufficient conditions (1.13), (1.14) hold whatever the normed spaces.

The basic question is: what properties must the operators A, B have, in order for scheme (1.2) to be correct? An effective answer can be obtained by considering schemes in real Hilbert space (see Section 2).

Before considering the idea of approximation for schemes, and the abstract space H_N with the norm $\| \cdot \|_N$, consider space H_0 with norm $\| \cdot \|_0$. We assume the existence of a linear operator P_N, associating a vector $u \in H_0$ with a vector $u_N = P_N u$ of space H_N and that the norms are matched

$$\lim_{N \to \infty} \|P_N u\|_N = \|u\|_0. \tag{1.15}$$

Given the set $V \in H_0$ (say the set of solutions of the initial problem), consider the difference

$$z = y - P_N v, \qquad z(t) \in H_N \quad \text{for } t \in \omega_\tau,$$

where $v(t) \in V$, $y(t) \in H_N$ is the solution of equation (1.2). Substituting $y = z + P_N v$ in (1.2), we get $\Pi z = B z_t + A z = \psi$, where $\psi = \varphi - (B v_{t,N} + A v_N)$, $v_N = P_N v$, $\psi \in H_N$. The vector $\psi = \psi(t; v)$ of set H_N is called the error of the approximation for scheme (1.2) on the set $V \subset H_0$. With each H_N we associate a number $h_N > 0$ such that the sequence $\{h_N\}$ is convergent to zero as $N \to \infty$:

$$\lim_{N \to \infty} h_N = 0.$$

Scheme (1.2) will be said to have an approximation $O(\tau^k + h_N^m)$ in the norm $\| \cdot \|_{(2)}$ on V if, given any $v \in V$ and any $N \geqslant N_0 (h_N \leqslant h_0)$ and $\tau \leqslant \tau_0$

we have $\|\psi(t;v)\|_{(2)} = O(\tau^k + h_N{}^m)$ or $\|\psi(t,v)\|_{(2)} \leqslant M(\tau^k + h_N{}^m)$.

Consider two schemes: $\Pi_1 y = \varphi$, $\Pi_2 \tilde{y} = \tilde{\varphi}$.

Let $\psi(v)$ and $\tilde{\psi}(v)$ be the corresponding approximation errors on V.

The schemes Π_1 and Π_2 will be said to belong to the class of schemes $S(k, m)$ or to be equivalent in the sense of the approximation $O(\tau^k + h_N{}^m)$ on V in the norm $\|\cdot\|_{(3)}$, if

$$\|\widetilde{\psi}(v) - \psi(v)\|_{(3)} \leqslant M_1 \tau^k + M_2 h_N{}^m,$$

where $M_1 \geqslant 0$, $M_2 \geqslant 0$ are independent of τ and h_N; if $M_2 = 0$, Π_1, Π_2 belong to the class $S(k, \infty)$, or with $M_1 = 0$, to the class $S(\infty, m)$.

Various classes $S(k, m)$ of schemes equivalent in their order of approximation will be considered.

2. Sufficient conditions for stability

1. LINEAR OPERATORS IN HILBERT SPACE

Let $H = H_N$ be real unitary space with scalar product $(\,,\,)_N$ and the norm $\|y\|_N = \sqrt{(y, y)_N}$ (the subscript V will be omitted in future). We shall consider linear operators $A(t)$, $B(t)$, $R(t)$ etc. dependent on the parameter $t \in \bar{\omega}_\tau$ (and τ, V) and mapping H into H for each $t \in \bar{\omega}_\tau$.

We recall some elementary properties of linear operators [4]. Let v, z be vectors of H. We shall write

$$A \geqslant B, \quad \text{if} \quad (Av, v) \geqslant (Bv, v) \; *);$$
$$A = B, \quad \text{if} \quad Av = Bv \text{ for all } v \in H.$$

The operator A is selfadjoint: $A^* = A$, if $(Av, z) = (v, Az)$; positive: $A > 0$, if $(Av, v) > 0$, $v \neq 0$; positive definite: $A \geqslant \delta E$, if $(Av, v) \geqslant \delta \|v\|^2$ (E is the unit operator, δ is a positive number); semibounded from below: $A \geqslant - \delta E$, if $(Av, v) \geqslant \delta \|v\|^2$. (Note: if $(Av, v) = (Bv, v)$ for all $v \in H$, the operator form of the condition $A \geqslant B$ will not be used.)

A positive operator $A = A(t)$, dependent on $t \in \bar{\omega}_\tau$, will be termed Lipschitz continuous in t if

$$|(A_{\bar{t}} v, v)| \leqslant c_2 (\check{A} v, v), \qquad (2.1)$$

where $c_2 = \text{const} > 0$, $A_{\bar{t}} = (A - \check{A})/\tau$, $A = A(t)$, $\check{A} = A(t - \tau)$, $t = t_j \in \omega_\tau$.

The necessary and sufficient condition for the existence of the inverse operator A^{-1} is that $Av = 0$ only if $v = 0$. Hence, if A is positive $(A > 0)$, A^{-1} exists.

2. TWO-LAYER SCHEMES

Consider the set of all two-layer schemes

$$By_t + Ay = \varphi(t), \quad 0 \leqslant t \in \omega_\tau, \quad y(0) = y_0, \qquad (2.2)$$

where $B = B(t)$, $A = A(t)$ are linear operators given on H.

Suppose that, for all $t \in \bar{\omega}_\tau$:

1) $A = A(t)$ is

$$\text{selfadjoint,} \quad A^* = A; \qquad (2.3)$$

$$\text{positive,} \quad A > 0; \qquad (2.4)$$

Lipschitz continuous in t: $|(A_{\bar{t}} z, z)| \leqslant c_2 (\check{A} z, z)$, $z \in H$; $\qquad (2.5)$

2) the operator $B = B(t)$ is

$$\text{positive,} \quad B > 0. \qquad (2.6)$$

The set of two-layer schemes (2.2), whose operators $A = A(t)$ and $B = B(t)$ satisfy conditions (2.3) - (2.6) will be termed the initial family of two-layer schemes and denoted by IF2.

It is always assumed in future that conditions (2.3) - (2.6) are satisfied, i.e. only schemes belonging to the initial family are considered.

Recalling the definition of correctness of scheme (1.2) given in § 4 of Section 1, the solution of problem (1.2) will be estimated in the energy norm

$$\| y \|_{(1)} = \| y \|_a = \sqrt{(A(t)y,\, y)}. \tag{2.7}$$

For the right-hand side we use the norms

$$\|\varphi\|_{(2)} = \|\varphi\| \quad \text{and} \quad \|\varphi\|_{(2)} = (A^{-1},\, \varphi,\, \varphi)^{1/2} + (\check{A}^{-1}\varphi_{\bar{t}},\, \varphi_{\bar{t}}). \tag{2.8}$$

Since $B > 0$, problem (1.2) is solvable whatever $y_0 \in \mathcal{I}$, $\varphi(t) \in \mathcal{I}$.

The sufficient conditions for two-layer schemes (2.2) of IF2 can now be formulated. Everywhere, c_1, c_2, ... denote positive numbers independent of V and τ.

Theorem 1

If

$$B \geqslant \tfrac{1}{2}\tau(1 - c_1\tau)A, \tag{2.9}$$

scheme (2.2) of IF2 is stable for sufficiently small $\tau \leqslant \tau_0$, $\tau_0 < 1/2\, c_1$, so that the solution of problem (1.2) satisfies

$$\| y(t) \|_a \leqslant M_1 \| y_0 \|_a + M_2 \max_{\omega_\tau} [(A^{-1}\varphi,\, \varphi)^{1/2} + (\check{A}^{-1}\varphi_{\bar{t}},\, \varphi_{\bar{t}})^{1/2}], \tag{2.10}$$

where M_1, M_2 are positive constants dependent only on c_1, c_2 and t_0.

Corollary

If

$$B \geqslant \tfrac{1}{2}\tau A, \tag{2.11}$$

scheme (1.2) is stable whatever $\tau > 0$ and M_1, M_2 depend only on c_2 and t_0.

Theorem 2

If

$$B \geqslant \varepsilon E + \tfrac{1}{2}\tau(1 - c_1\tau)A, \qquad \varepsilon = \text{const} > 0, \tag{2.12}$$

then the *a priori* inequality holds:

$$\| y(t) \|_a \leqslant M_1 \| y_0 \|_a + \frac{1}{\sqrt{\varepsilon}} M_2 \max_{\bar{\omega}_\tau} \| \varphi(t) \| \quad \text{for} \quad \tau < \frac{1}{2c_1}. \tag{2.13}$$

Corollary

With

$$B \geqslant {}^{1}/_{2}\tau A + \varepsilon E, \ 0 < \varepsilon \leqslant 1, \tag{2.14}$$

(2.13) holds for all $\tau > 0$.

Conditions (2.9) and (2.12) distinguish the classes of stable schemes from the initial family of two-layer schemes (2.2).

As a rule, (2.9) and (2.12) with $c_1 = 0$, i.e. (2.11) and (2.14), will be used as sufficient conditions for stability.

Putting $B = E + \tau R$, we write scheme (2.2) in the second canonical form

$$y_t + \tau R y_t + Ay = \varphi. \tag{2.15}$$

From (11),

$$E + \tau R \geqslant \frac{1}{2}\tau A \quad \text{or} \quad R \geqslant \frac{1}{2}A - \frac{1}{\tau}E.$$

This condition will be satisfied for

$$R \geqslant \sigma_0 A, \text{ where } \sigma_0 = \frac{1}{2} - \frac{1}{\tau\|A\|}. \tag{2.16}$$

(If A is an unbounded operator, we formally put $1/\|A\| = 0$. everywhere in (2.16), (2.17) etc.).

For, if $R \geqslant \sigma_0 A$, i.e. $(Ry, y) \geqslant \sigma_0(Ay, y)$, then

$$(By, y) - {}^{1}/_{2}\tau(Ay, y) = ((E + \tau(R - {}^{1}/_{2}A))y, y) = \left[\|y\|^2 - \frac{1}{\|A\|}(Ay, y)\right] +$$

$$+ \tau((R - \sigma_0 A)y, y) \geqslant \tau((R - \sigma_0 A)y, y) \geqslant 0,$$

since $(Ay, y) \leqslant \|A\|\|y\|^2$. The writing in the proof can be simplified by using the operator inequalities and recalling that $E \leqslant (1/\|A\|)A$ and $E + \tau R \geqslant (1/\|A\|)A + \tau R$. The condition $E + \tau R \geqslant {}^{1}/_{2}\tau A$ will be satisfied if $(1/\|A\|)A + \tau R \geqslant {}^{1}/_{2}\tau A$, i.e. $R \geqslant \sigma_0 A$.

Similarly, condition (2.14) can be seen to be satisfied for

$$R \geqslant \sigma_\varepsilon A, \qquad \sigma_\varepsilon = \frac{1}{2} - \frac{1 - \varepsilon}{\tau\|A\|}, \qquad 0 < \varepsilon \leqslant 1. \tag{2.17}$$

Theorem 3

If

$$B \geqslant -c_1\tau^2 A, \qquad c_1 = \text{const} > 0, \qquad (2.20)$$

$$R \geqslant \frac{1+\varepsilon}{4} A, \quad \varepsilon = \text{const}, \qquad 0 < \varepsilon \leqslant 1, \qquad (2.21)$$

scheme (2.19) is correct for sufficiently small $\tau \leqslant \tau_0$, $\tau_0 < 1/4c_1$ and the solution of problem (2.19) satisfies

$$\|y(t)\|_{(1)} \leqslant M_1 \|y(\tau)\|_{(1)} + M_2 \max_{\omega_\tau} [(A^{-1}\varphi, \varphi) + (A^{-1}\varphi_{\bar{t}}, \varphi_{\bar{t}})]^{1/2}, \quad (2.22)$$

$$\|y\|_{(1)}^2 = {}^1/_4 (A(y+\check{y}), \ y+\check{y}) + \tau^2 (R - {}^1/_4 A)\, y_{\bar{t}}, \, y_{\bar{t}}), \qquad (2.23)$$

where M_1, M_2 are positive numbers depending only on c_1, c_2, ε, and t_0.

Notice that $B + 2\tau R \geqslant ((1+\varepsilon)/2 - c_1\tau)\tau A > 0$ for $\tau < 1/2c_1$. Scheme (2.19) is therefore solvable.

Theorem 4

Let

$$B \geqslant \delta E, \quad \delta = \text{const} > 0, \quad R \geqslant {}^1/_4(1+\varepsilon) A, \quad 0 < \varepsilon \leqslant 1. \qquad (2.24)$$

Then

$$\|y(t)\|_{(1)} \leqslant M_1 \|y(\tau)\|_{(1)} + M_2 \max_{\omega_\tau} \|\varphi(t)\|, \qquad (2.25)$$

holds for (2.19) (whatever $\tau > 0$), where $M_1 = M(c_2, \ t_0, \ c_2/\varepsilon)$, $M_2 = M(1/\delta, \ c_2, \ t_0, \ c_2/\varepsilon)$.

Note 1. If the operator $R - {}^1/_4 A \geqslant 0$ is Lipschitz continuous in t, Theorems 3 and 4 hold provided the condition $R \geqslant [(1+\varepsilon)/4]A$ is replaced by

$$R \geqslant {}^1/_4 A. \qquad (2.26)$$

In particular, this is true for constant operators R and A.

Consider as an example the scheme with weights

$$y_{\bar{t}} + A(\sigma_1 \hat{y} + (1 - \sigma_1 - \sigma_2) y + \sigma_2 \check{y}) = \varphi. \qquad (2.27)$$

Writing it in the canonical form (2.19), we get

$$B = E + (\sigma_1 - \sigma_2)\tau A, \qquad R = \frac{1}{2}(\sigma_1 + \sigma_2)A. \qquad (2.28)$$

In view of Theorem 3 and Note 1, the scheme (2.27) is stable with

$$\sigma_1 - \sigma_2 \geqslant -\frac{1}{\tau\|A\|} - c_1\tau, \qquad \sigma_1 + \sigma_2 \geqslant {}^1/_2. \qquad (2.29)$$

The condition $B \geqslant \delta E$ is satisfied for

$$\sigma_1 - \sigma_2 \geqslant -\frac{1-\delta}{\tau\|A\|}, \qquad 0 < \delta \leqslant 1. \qquad (2.30)$$

In this case (2.25) holds.

The symmetric scheme ($\sigma_1 = \sigma_2 = \sigma$) is stable if

$$\sigma \geqslant {}^1/_4. \qquad (2.31)$$

Note 2. Putting $R = 0$ in (2.19), we get the explicit scheme

$$B y_{\overline{t}} + Ay = \varphi,$$

which is unstable whatever the operator $B > 0$.

We recall that the two-layer explicit scheme ($R \equiv 0$)

$$y_t + Ay = \varphi$$

is conditionally stable for $\tau \leqslant 2/\|A\|$.

While the stability of the three-layer scheme is determined by the operator R only, that of the two-layer scheme is determined by $R + E/\tau$.

3. Regularization of schemes

We now turn to some general methods of regularization.

1. BASIC REGULARIZATION PRINCIPLE

The sufficient conditions for stability

$$R \geqslant \sigma_0 A, \qquad \sigma_0 = \frac{1}{2} - \frac{1}{\tau \|A\|}, \text{ for the two-layer scheme} \qquad (3.1)$$

$$R \geqslant \frac{1+\varepsilon}{4} A, \qquad \varepsilon > 0, \text{ for the three-layer scheme} \qquad (3.2)$$

impose very weak restrictions on the selection of the operator R. In consequence, it becomes possible to transform (regularize) a scheme by modifying R in such a way that conditions (3.1) - (3.2) are still satisfied. In future, R will be termed a regularizer.

Consider, for instance, the scheme

$$(E + \tau R) y_t + A y = \varphi, \qquad 0 \leqslant t \in \omega_\tau, \qquad y(0) = y_0. \qquad (3.3)$$

Let condition (1) be satisfied, i.e. scheme (3) is stable. Any other scheme (3.3) with regularizer $\tilde{R} \geqslant R$ will now also be stable. The aim of regularization is to remain within the class of stable schemes and to select R in each concrete case so as to satisfy extra requirements such as economy (i.e. minimum amount of computation for finding the solution of the difference problem) or a given order of approximation. These requirements compete with one another, and it is far from easy to satisfy both simultaneously. Since solving problem (3.3) amounts to inverting

the operator $E + \tau R$: $\hat{y} = (E + \tau R)^{-1} (E - \tau (A - R)) y + (E + \tau R)^{-1} \varphi$, R must be selected so as to minimize the number of operations $q(R)$ required for inversion of $E + \tau R$. Concrete methods of selecting R will be found in Section 4, as well as in §§ 2 and 3 of the present section. The basic principle of regularization consists in selecting en. eq. operators as regularizers. (The case when the operator A also changes during regularization will not be considered.)

The positive operators $A = A(N, \tau; t)$ and $B = B(N, \tau; t)$ will be termed en. eq. operators if

$$\gamma_1 (By, y) \leqslant (Ay, y) \leqslant \gamma_2 (By, y) \qquad (3.4)$$

for any $y \in H_N$ and $t \in \omega_\tau$, where γ_1, γ_2 are positive numbers, in general dependent on τ and N; if they are not dependent on τ, N, we call A, B uniformly en. eq. operators (with respect to τ, N).*

---

* In [15], the difference operator B satisfying (4) is called majorant with respect to A; in [16], A and B for which (3.4) is satisfied are termed spectrally equivalent. A, B are assumed selfadjoint, and H a finite-dimensional space. Neither of these assumptions is made here, nor do we use any information about the spectra of A and B.

Uniformly en. eq. operators will always be considered, except in § 2. Conditions (3.4) are written as the operator inequalities

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B. \tag{3.5}$$

Theorem 5

Let $A = A(t)$ be positive, selfadjoint, and let it satisfy the Lipschitz continuity condition

$$|(A_{\bar{t}}(t)y, y)| \leqslant c_3(\check{A}^{(0)}y, y), \qquad \check{A}^{(0)} = A^{(0)}(t - \tau), \tag{3.6}$$

where $A^{(0)}(t)$ is uniformly en. eq. with $A(t)$, so that

$$\gamma_1 A^{(0)} \leqslant A \leqslant \gamma_2 A^{(0)}. \tag{3.7}$$

Scheme (3.3) with regularizer $R = \sigma A^{(0)}$ is stable, and Theorems 1, 2 hold for it, with $\tau > 0$ arbitrary and

$$\sigma \geqslant {}^1/_2\gamma_2. \tag{3.8}$$

Proof. From (3.6) and (3.7),

$$|(A_{\bar{t}}y, y)| \leqslant c_3(\check{A}^{(0)}y, y) \leqslant \frac{c_3}{\gamma_1}(\check{A}y, y) = c_2(\check{A}y, y), \qquad c_2 = \frac{c_3}{\gamma_1},$$

i.e. (3.3) belongs to IF2. The conditions of Theorems 1, 2 are satisfied, since $B = E + \tau R \geqslant E + {}^1/_2\tau\gamma_2 A^{(0)} \geqslant E + {}^1/_2\tau A$. The theorem is proved.

Note 1. Theorem 5 remains in force in any of the following cases: (1) $\sigma = \sigma_0\gamma_2$, $\sigma_0 = {}^1/_2 - 1 / (\tau\|A\|)$; (2) A is independent of t, and instead of (3.7), $0 < A \leqslant \gamma_2 A^{(0)}$.

Note 2. The operator $A^{(0)}$ may be non-selfadjoint.

Now consider the three-layer scheme

$$By_{\mathring{t}} + \tau^2 R y_{\bar{t}t} + Ay = \varphi, \qquad 0 < t \in \omega_\tau, \qquad y(0),\ y(\tau) \text{ given} \tag{3.9}$$

Theorem 6

Let (3.6) and (3.7) be satisfied, $A(t) > 0$ and $A^*(t) = A(t)$. If $A^{(0)}(t)$ is selfadjoint and Lipschitz continuous in t, i.e.

$$|(A_{\bar{t}}^{(0)}(t)y, y)| \leqslant c_4(\check{A}^{(0)}y, y), \tag{3.10}$$

then the *a priori* estimates (2.22) and (2.25) for

$$\sigma \geqslant \frac{1+\varepsilon}{4}\gamma_2, \qquad 0 < \varepsilon \leqslant 1. \tag{3.11}$$

hold for scheme (3.9) with regularizer

$$R = \sigma A^{(0)}$$

This theorem is proved simply by showing that scheme (3.9) belongs to IF3. We just observe that condition (3.2) is satisfied, since

$$R \geqslant \frac{1}{4}(1+\varepsilon)\gamma_2 A^{(0)} \geqslant \frac{1}{4}(1+\varepsilon)A.$$

Note 3. The constants M_1, M_2 in the *a priori* estimates depend on γ_1 but not on γ_2. The requirement that γ_1 be independent of τ and V is therefore natural. If A, $A^{(0)}$ are constant operators, (3.7) can be replaced by the weaker condition

$$0 < A \leqslant \gamma_2 A^{(0)}.$$

Theorems 5 and 6 also remain in force when γ_2 depends on τ and V. If γ_2 depends on τ and V, only the error of the approximation of the scheme is affected. The case when γ_2 depends on $\|A\|$, and hence on V, τ, will be considered in § 2.

The choice of a suitable en. eq. operator is of basic importance in constructing stable schemes for concrete problems. Several en. eq. operators $A^{(0)}$ will be listed in Section 4 for the case when $A = -\Lambda$, Λ being the elliptic difference operator.

Two elementary examples of difference scheme regularization will be · considered here, and each will be shown to correspond to a specific method of selecting the en. eq. operator (by Theorems 5 and 6, the regularizer R is then taken as $R = \sigma A^{(0)}$).

2. EXPLICIT THREE-LAYER SCHEMES

The simplest en. eq. operator is the unit operator

$$A^{(0)} = E.$$

Since $A \leqslant \|A\|E$, the condition $A \leqslant \gamma_2 A^{(0)}$ will be satisfied with $\gamma_2 = \|A\|$. We shall assume for simplicity that A is a constant operator

(independent of t). Consider the three-layer scheme. For this, $R = \sigma 4^{(0)}$, $\sigma \geqslant \frac{1}{4} \gamma_2$. Recalling that $\gamma_2 = \| 4 \|$, we get $R = \sigma E$, where $\sigma \geqslant \frac{1}{4} \| 4 \|$. The corresponding three-layer scheme is

$$By_{\stackrel{\circ}{t}} + \sigma \tau^2 y_{\bar{t}t} + Ay = \varphi.$$

This regularization is meaningful if R is either the unit operator or is such that inversion of $B + 2\sigma\tau E$ can be performed with a small number of operations. With $B = E$, we get the explicit three-layer scheme

$$y_{\stackrel{\circ}{t}} + \sigma \tau^2 y_{\bar{t}t} + Ay = \varphi, \tag{3.12}$$

which is unconditionally stable with

$$\sigma \geqslant \frac{1}{4} \| A \|. \tag{3.13}$$

In the introduction we mentioned the rhombus scheme for the equation of heat conduction

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t), \qquad 0 < x < 1, \qquad 0 \leqslant t \leqslant t_0. \tag{3.14}$$

We specify boundary conditions of the 1st kind with $x = 0$ and $x = 1$. Let $\bar{\omega}_h = \{x_i = ih, i = 0, 1, \ldots, N, h = 1/N\}$ be a mesh on $0 \leqslant x \leqslant 1$, $\bar{\omega}_\tau = \{t_j = j\tau, j = 0, 1, \ldots, N_0, \tau = t_0/N_0\}$ be a mesh on $0 \leqslant t \leqslant t_0$, $\bar{\omega}_{h\tau} = \bar{\omega}_h \times \bar{\omega}_\tau = \{(x_i, t_j)\}$ be a mesh on

$$(0 \leqslant x \leqslant 1, \ 0 \leqslant t \leqslant t_0),$$

H be the space of mesh functions specified on $\bar{\omega}_h$ and vanishing for $i = 0$, N,

$$(y, v) = \sum_{i=1}^{N-1} y_i v_i h, \qquad \|y\| = \sqrt{(y, y)}$$

are the scalar product and norm in H.

The rhombus scheme is

$$\frac{y_i^{j+1} - y_i^{j-1}}{2\tau} = \frac{y_{i-1}^j - (y_i^{j+1} + y_i^{j-1}) + y_{i+1}^j}{h^2} + \varphi_i{}^j.$$

To reduce it to the canonical form, we recall that

$$y_i^{j+1} + y_i^{j-1} = \hat{y} + \check{y} = (\hat{y} - 2y + \check{y}) + 2y = 2y + \tau^2 y_{\bar{t}t},$$

$$y_{i-1} - 2y_i + y_{i+1} = h^2 y_{\bar{x}x, i},$$

which gives

$$y_{\hat{t}} + \frac{1}{h^2}\,\tau^2 y_{\hat{t}t} = \Lambda y + \varphi, \qquad \Lambda y = y_{\bar{x}x}. \tag{3.15}$$

Comparison of (3.15) and (3.9) shows that $\Lambda y = - Ay$, $Ry = y/h^2$. Since $\| A \| = (4/h^2)\cos^2(\pi h/2) < 4/h^2$, then $R = E/h^2 > A/4$, i.e. the stability condition is satisfied. The *a priori* estimates (2.22) and (2.25) hold for the rhombus scheme whatever τ and h.

Next, comparison of (3.15) and (3.12) shows that the rhombus scheme belongs to the class of three-layer schemes with regularizer $R = \sigma E$, where σ satisfies (3.13).

The analogue of scheme (3.15) is easily written for the equation of heat conduction with variable coefficients when one or more space variables x_1, x_2, ..., x_p are involved. In (3.15) Λy has to be replaced by a difference operator approximating the elliptic operator, and σ has to satisfy $\sigma \geqslant \frac{1}{4} \| A \|$. It only remains to evaluate

$$\|A\| = \|\Lambda\| \leqslant c_2 \sum_{\alpha=1}^{p} 4 h_\alpha^{-2},$$

where c_2 is the maximum of the thermal conductivity, and to put

$\sigma = c_2 \sum\limits_{\alpha=1}^{p} h_\alpha^{-2}$ (it is assumed for simplicity that the region of vari-

ation of x_1, ..., x_p is a p-dimensional parallelepiped and the mesh ω_h is uniform with respect to each x_α; these restrictions can easily be lifted).

The explicit three-layer scheme is well known to involve an approximation error $O(\tau^2 h^{-2} + h^2)$ and to converge at the rate $O(\tau + h^2)$ provided $\tau h^{-2} \leqslant c_0 = $ const. It can be shown to be uniformly convergent for $p \leqslant 3$.

3. ASYMMETRIC SCHEMES

R does not need to be selfadjoint for a two-layer scheme. Consider the case when $A^{(0)}$ is expressible as the sum of two mutually adjoint operators

$$A^{(0)} = A_1 + A_2, \qquad A_2 = A_1^*.$$

Then

$$(A_1 y, y) = (A_2 y, y) = 0.5 (A^{(0)} y, y).$$

The regularizer R can be

$$R_1 = \gamma_2 A_1 \quad \text{or} \quad R_2 = \gamma_2 A_2, \qquad R_2 = R_1^*. \qquad (3.16)$$

Since $(A_1 y, y) = (A_2 y, y) = 0.5 (A^{(0)} y, y), \quad \gamma_2 (A^{(0)} y, y) \geqslant (Ay, y),$
$(R_1 y, y) = (R_2 y, y) = 0.5 \gamma_2 (A^{(0)} y, y) \geqslant 0.5 (Ay, y),$ the resulting schemes
are stable and satisfy Theorem 5.

In particular, with $A^{(0)} y = Ay$ $(\gamma_1 = \gamma_2 = 1)$ we have

$$R_1 y = A_1 y, \qquad R_2 y = A_2 y, \qquad R_2 = R_1^*. \qquad (3.17)$$

Schemes with regularizers (3.16) will be termed asymmetric or tri-
angular.

Take as an example the asymmetric scheme of [6] for the equation of
heat conduction (3.14) with boundary conditions of the 1st kind. Here,
$Ay = -\Lambda y, \quad \Lambda y = y_{\bar{x}x}$ (see § 2). Noting that $(y, y_x) = 0.5 h \|y_{\bar{x}}\|^2,$
$-(y, y_x) = 0.5 h \|y_x\|^2 = (y, y_{\bar{x}}), \quad Ay = -y_{\bar{x}x} = -y_x / h + y_{\bar{x}} / h,$ we can
write A as $A = A_1 + A_2$, where $A_1 y = y_{\bar{x}} / h, \quad A_2 y = -y_x / h$. Putting
$R = A_1$ and $R = A_2$ next, two asymmetric schemes are obtained (written in
the canonical form, [6]:

$$y_t + \frac{\tau}{h} y_{\bar{x}t} = y_{\bar{x}x} + \varphi, \qquad R_1 y = \frac{1}{h} y_{\bar{x}}, \qquad (3.18)$$

$$y_t - \frac{\tau}{h} y_{xt} = y_{\bar{x}x} + \varphi, \qquad R_2 y = -\frac{1}{h} y_x. \qquad (3.19)$$

Each of these is unconditionally stable, since $(R_1 y, y) = (R_2 y, y) = 0.5 (Ay, y)$. The stability is achieved at the expense of poor approxima-
tion: schemes (3.18) and (3.19) give an approximation $O(\tau/h + h^2)$,
whereas the explicit scheme $y_t = y_{\bar{x}x} + \varphi$ gives $O(\tau + h^2)$. Alternation of
schemes (3.18) and (3.19) from layer to layer was proposed in [6] for
improving the accuracy. The resulting mixed scheme has the accuracy [17]
$O(\tau^2 h^{-2} + h^2) = O(h^2)$ for $\tau h^{-2} \leqslant c_0 = \text{const.}$

Asymmetric schemes can easily be written for the equation of heat
conduction with variable coefficients by means of Theorem 5:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(k \frac{\partial u}{\partial x} \right) + f, \qquad 0 < c_1 \leqslant k \leqslant c_2. \qquad (3.20)$$

Here $Ay = -(ay_{\bar{x}})_x$, say, $a(x, t) = k(x - 0.5 h, t)$,

$$\gamma_1 = c_1, \qquad \gamma_2 = c_2, \qquad A^{(0)}y = -y_{\bar{x}x}, \qquad R_1y = \frac{c_2}{h}y_{\bar{x}}, \qquad R_2y = -\frac{c_2}{h}y_x.$$

More asymmetric schemes (4 for $p = 2$ and 8 for $p = 3$) can be written for the multidimensional equation of heat conduction

$$\frac{\partial u}{\partial t} = \sum_{\alpha=1}^{p} \frac{\partial^2 u}{\partial x_\alpha^2} + f(x, t), \qquad p > 1, \qquad x = (x_1, \ldots, x_p) \in G,$$

$$0 \leqslant t \leqslant t_0, \qquad u|_\Gamma = \mu, \tag{3.21}$$

Let $p = 2$. Possible regularizers are now

$$R_1y = \frac{1}{h_1}y_{\bar{x}_1} + \frac{1}{h_2}y_{\bar{x}_2}, \qquad R_2y = -\frac{1}{h_1}y_{x_1} - \frac{1}{h_2}y_{x_2}, \qquad R_2 = R_1^*, \tag{3.22}$$

$$R_3y = \frac{1}{h_1}y_{\bar{x}_1} - \frac{1}{h_2}y_{x_2}, \qquad R_4y = \frac{1}{h_2}y_{\bar{x}_2} - \frac{1}{h_1}y_{x_1}, \qquad R_4 = R_3^*, \tag{3.23}$$

where h_1, h_2 are the intervals of the mesh $\omega_h(G)$ relative to x_1, x_2. We shall assume for simplicity that $G = G_0$ is a p-dimensional parallele-piped. If ω_h is non-uniform, h_α in (3.22) and (3.23) must be replaced by $\hbar_\alpha = 0.5(h_\alpha + h_\alpha^{(+1\alpha)})$. The operators R_1 and R_2, R_3 and R_4 are mutually adjoint. All 4 schemes are unconditionally stable. By alternating R_1 and R_2, R_3 and R_4 from layer to layer, alternating schemes are obtained (explicit schemes with alternating directions) with the accuracy $O(\tau^2 h^{-2} + h^2)$, where $h = \min(h_1, h_2)$. Such schemes were proposed in [6]. The same regularizers, multiplied by c_2, where c_2 is the maximum of the thermal conductivity, can be used for the equation of heat conduction when the coefficient is variable. Extension to any number of dimensions involves no difficulties. We can take

$$R_1y = \sum_{\alpha=1}^{p} \frac{1}{h_\alpha}y_{\bar{x}_\alpha}, \qquad R_2y = -\sum_{\alpha=1}^{p} \frac{1}{h_\alpha}y_{x_\alpha}$$

and so on. The resulting schemes are economic, since explicit formulae are used to compute $\hat{y}$ at a new layer.

4. METHOD OF FACTORIZATION. FACTORIZED TWO-LAYER SCHEMES

Many economic schemes have been proposed (see Introduction) for the numerical solution of the multidimensional problems of mathematical

physics. After reduction to the canonical form (in the case of two-layer
schemes), most give schemes (3.3), in which B is the product of several
operators. The terminology for these schemes has not yet been entirely
settled (method of alternating directions, splitting schemes, splitting
operator schemes, method of fractional steps, etc.). Without pretending
that there is anything final about it, we shall use the term factorized
schemes (FS) to describe those in which the operator on the upper layer,
B in the case of two-layer, and $B + 2\tau R$ in the case of three-layer
schemes, is the product of a finite number of operators, e.g.

$$B = \prod_{s=1}^{m} B_s = B_1 \ldots B_s \ldots B_m, \qquad (3.24)$$

This term does have the advantage of indicating the particular nature of
the regularization procedure, i.e. factorization of the operator on the
upper layer.

The formal method of constructing FS [10, 12] is as follows.*

Given a scheme having a definite order of approximation; it it is un-
stable, we transform it, in accordance with § 1, to a stable scheme by
means of a suitable choice of R. Generally speaking, this involves a
change in the order of approximation (better or worse).

Take the case when R is

$$R = \sum_{\alpha=1}^{m} R_\alpha \qquad (3.25)$$

and satisfies the stability condition

$$R \geqslant \sigma_0 A, \qquad \sigma_* = \frac{1}{2} - \frac{1}{\tau \|A\|}, \qquad (3.26)$$

so that the initial scheme is

$$\left(E + \tau \sum_{\alpha=1}^{m} R_\alpha \right) y_t + A y = \varphi. \qquad (3.27)$$

---

* FS were considered as initial schemes in [8]. It was noted in [20]
that an earlier [21] economic algorithm for solving the multidimen-
sional equation of heat conduction (with constant coefficients in
the parallelepiped) can be obtained by a formal method of "approxi-
mate" factorization of the operator on the upper layer, as a result
of which the natural multidimensional scheme is replaced by a factor-
ized scheme (see also [11]).

Consider the scheme

$$y_t + \tau \breve{R} y_t + A y = \varphi, \tag{3.28}$$

which belongs to the same class of stable schemes $(\breve{R} \geqslant \sigma_0 A)$. Choose $\breve{R}$ so that $\breve{B} = E + \tau \breve{R}$ can be factorized, e.g.

$$E + \tau \breve{R} = \prod_{\alpha=1}^{m} (E + \tau R_\alpha \tag{3.29}$$

Scheme (28) now becomes

$$\prod_{\alpha=1}^{m} (E + \tau R_\alpha) y_t + A y = \tilde{\varphi}, \quad y(0) = y_0. \tag{3.30}$$

Comparison of (3.28) and (3.30) shows that

$$\breve{R} = R + \tau Q, \quad \breve{B} = E + \tau \breve{R} = E + \tau R + \tau^2 Q \tag{3.31}$$

$$Q = \sum_{\alpha < \beta} R_\alpha R_\beta + \tau \sum_{\alpha < \beta < \gamma} R_\alpha R_\beta R_\gamma + \ldots + \tau^{m-2} \prod_{\alpha=1}^{m} R_\alpha. \tag{3.32}$$

Turning to the stability conditions for FS (3.30), since $\tilde{B} = E + \tau R + \tau^2 Q$, the sufficient condition for stability

$$\breve{B} \geqslant 0.5 \, \tau (1 - 2 c_1 \tau) \, A \tag{3.33}$$

will be satisfied if

$$Q \geqslant -c_1 A,$$

where c_1 is an arbitrary positive constant independent of τ and N. We have thus proved

Theorem 7

Let the initial scheme (3.27) be stable, i.e. $R \geqslant \sigma_0 A$. Then the FS (3.30) will be stable in the norm $\|y\|_a = \sqrt{(Ay, y)}$ for small $\tau \leqslant \tau_0$, $\tau_0 < 1/4 \, c_1$, provided (3.33) is satisfied. The estimate (2.10) holds for the solution of problem (3.30).

Note 1. If

$$Q \geqslant -c_1 R, \quad R \geqslant 0.5 A, \tag{3.34}$$

the FS (3.30) is stable for $\tau \leqslant 1/c_1$ and (2.10), (2.12) hold for it
(with $\varepsilon = 1$).

For, $\tilde{B} = E + \tau R + \tau^2 Q \geqslant E + \tau(1 - c_1\tau)R \geqslant 0.5\tau(1 - c_1\tau)A + \quad E$, i.e.
$\tilde{B} \geqslant E$, since $R \geqslant 0.5\,A$, $1 - c_1\tau \geqslant 0$.

Convergence of a scheme only requires that it approximate a differ-
ential equation and be asymptotically stable, i.e. on a finite enough
mesh. For a scheme to be realizable, it must be stable on a practically
acceptable mesh (i.e. a "coarse" mesh). This requirement becomes speci-
ally important in the case of variable coefficients. The condition for
the mesh intervals to be small, $\tau \leqslant \tau_0$ and $h \leqslant h_0$, where τ_0, h_0 depend
on the coefficients of the differential equation, is too onerous and
can lead to the scheme being unrealizable in practice.

Let us distinguish the class of FS which is free from this defect,
i.e. FS which are stable whatever τ.

Theorem 9

Let $R_1, \ldots, R_m$ be selfadjoint $(R_s^* = R_s)$, non-negative $(R_s \geqslant 0)$ and
commutable $(R_s R_k = R_k R_s$ for all s, $k = 1, \ldots, m)$. Then scheme (3.30) is
stable whatever τ, provided

$$R = \sum_{s=1}^{m} R_s \geqslant \sigma_0 A. \tag{3.35}$$

Proof. By the Theorems of [22, 23], we have $R_s R_k \geqslant 0$, if $R_s \geqslant 0$,
$R_k \geqslant 0$, i.e. $Q \geqslant 0$. Since $R \geqslant \sigma_0 A (E + \delta R \geqslant 0.5\tau A)$, where $\tilde{R} = R +$
$\tau Q \geqslant R \geqslant \sigma_0 A$ whatever τ. Scheme (3.30) is unconditionally stable.

Note 2. When speaking of unconditional stability, all the conditions
are understood to be satisfied for the operators A (see Section 2, § 2)
and R_s whatever τ and $V(h_N)$.

Note 3. If $m = 2$, the theorem is still true if, instead of being
selfadjoint, R_1 and R_2 are adjoint to one another $(R_2 = R_1^*$, see § 4)
and $R = R_1 + R_2 \geqslant \sigma_0 A$. For, $\tilde{B} = E + \tau R + \tau^2 R_1 R_2 > E + \tau R \geqslant 0.5\tau A$,
since $(R_1 R_2 y, y) = (R_2^* R_2 y, y) = \|R_2 y\|^2 > 0$.

Two conditions need to be observed when factorizing:

(1) FS (3.30) belongs to the same class of stable schemes as the

initial scheme (3.27).

(2) FS (3.30) and the initial scheme (3.27) are equivalent in order of approximation on a set of smooth functions $V \subset H_0$ in the sense of the definition of Section 1, § 4.

The right-hand side sometimes has to be changed in order to retain the order of approximation. This applies, in particular, when the boundary conditions are inhomogeneous [8]. In the operator form of the difference equation, the boundary conditions are taken care of by varying the right-hand side at the mesh base-points immediately adjacent to the boundary. In an auxiliary operator $\tau^2 Q$ is introduced, φ has to be changed to $\tilde{\varphi}$ at these base-points [2].

The following problem is obtained for the error $z = y - P_N v$:

$$(E + \tau R + \tau^2 Q) z_t + Az = \tilde{\psi}, \qquad 0 \leqslant t \in \omega_\tau, \qquad z(0) = 0, \qquad (3.36)$$

where $\tilde{\psi}$ is the approximation error for scheme (3.30). Let $\varphi = \tilde{\varphi}$. Then we can write

$$\tilde{\psi} = \overset{\bullet}{\psi} + \psi, \qquad \overset{\bullet}{\psi} = Q\eta, \qquad \eta = \tau^2 P_N v_t,$$

where ψ is the approximation error for the initial scheme (3.27). To judge whether (3.27) and (3.30) are equivalent in order of approximation, we have to estimate $\overset{\bullet}{\psi} = Q\eta$ in some norm $\|\cdot\|_{(3)}$. This norm is

$$\|\overset{\bullet}{\psi}\|_{(3)} = \|\eta\|_q = \sqrt{(Q\eta, \eta)},$$

if Q is a selfadjoint positive operator. We write Q as

$$Q\eta = \sum_{s=2}^{m} \tau^s Q_s \bar{\eta}, \qquad \eta = \tau^2 \bar{\eta}, \qquad \bar{\eta} = P_N v_t,$$

and introduce the notation

$$\|\bar{\eta}\|_{q_s} = \sqrt{(Q_s \bar{\eta}_x, \bar{\eta})}$$

(assuming the Q_s to be positive selfadjoint operators).

Theorem 9

Let

$$R = \sum_{s=1}^{m} R_s, \qquad R_s^* = R_s > 0, \qquad R_s R_k = R_k R_s$$

for all s, $k = 1, 2, \ldots, m$, $R \geqslant \sigma_0 A$.

Then FS (3.36) with right-hand side

$$\widetilde{\psi} = Q\eta + \psi \tag{3.37}$$

satisfies for any $\tau > 0$ the *a priori* estimates

$$\|z(t)\|_a \leqslant M \max_{\omega_\tau} [\tau \|\overline{\eta}\|_q + \|\psi\|], \tag{3.38}$$

$$\|z(t)\|_a \leqslant M \max_{\omega} \left[\sum_{s=3}^{m} \tau^{(s+1)/2} \|\overline{\eta}(t)\|_{q_s} + \|\psi(t) + \tau^2 Q\overline{\eta}\| \right], \tag{3.39}$$

where $\|z(t)\|_a = \sqrt{(A(t)z(t), z(t))}$.

Note 4. Let

$$Q \geqslant \varepsilon_0 \overline{Q} - c_1 A, \qquad \varepsilon_0 > 0, \qquad c_1 > 0, \tag{3.40}$$

$$|(Q_s y, v)| \leqslant \varepsilon_0 (\overline{Q}_s v, v) + \frac{1}{\varepsilon_0} M' (\overline{Q}_s y, y),, \tag{3.41}$$

where

$$\overline{Q} = \prod_{\alpha=1}^{m} (E + \tau \overline{R}_\alpha) - (E + \tau \overline{R}) = \sum_{s=2}^{m} \tau^s \overline{Q}_s, \tag{3.42}$$

$\overline{R}_s$, $s = 1, 2, \ldots, m$ are operators satisfying the conditions of Theorem 9. Now, if we replace $\|\overline{\eta}\|_q$ by $\|\overline{\eta}\|_{\overline{q}}$ and $\|\overline{\eta}\|_{q_s}$ by $\|\overline{\eta}\|_{\overline{q_s}}$ in (3.38) and (3.39), these latter hold for small $\tau \leqslant \tau_0$. Conditions (3.40) and (3.41) are satisfied for a number of FS's approximating a parabolic equation in the parallelepiped $(0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2, \ldots, p)$. Estimates similar to the above were obtained for them in [9].

There is not space here to dwell in detail on more complex types of estimate. For instance, if the conditions of Theorem 9 are satisfied and ψ has the form

$$\widetilde{\psi} = Q\eta + \psi^* + \sum_{\alpha=1}^{m} \bar{\psi}_\alpha,$$

the solution of problem (3.36) satisfies

$$\| z(t) \|_a \leqslant M \max_{\omega_\tau} \| \widetilde{\psi}(t) \|_{(2)}, \tag{3.43}$$

where

$$\| \psi \|_{(2)}^2 = \| \psi^* + \tau^2 Q_2 \bar{\eta} \|^2 + \sum_{s=1}^{m} \tau^{s+1} \| \bar{\eta} \|_{q_s}^2 + \tag{3.44}$$

$$+ \sum_{\alpha=1}^{m} [((A_\alpha^{(0)})^{-1} \bar{\psi}_\alpha, \bar{\psi}_\alpha) + ((A_\alpha^{(0)})^{-1} \bar{\psi}_{\alpha\bar{t}}, \bar{\psi}_{\alpha\bar{t}})],$$

$A_\alpha^{(0)}$, $\alpha = 1, 2, \ldots, m$, are selfadjoint positive operators, and

$A^{(0)} = \sum_{\alpha=1}^{m} A_\alpha^{(0)}$ is en. eq. to A operators $(\gamma_1 A^{(0)} \leqslant A \leqslant \gamma_2 A^{(0)})$.

5. THREE-LAYER FACTORIZED SCHEMES

Three-layer schemes can be factorized by several methods. Three methods will be considered here for the scheme

$$y_{\hat{t}} + \tau^2 R y_{\bar{t}t} + Ay = \varphi, \qquad 0 < t \in \omega_\tau, \quad y(0), \, y_t(0) \text{ given}, \tag{3.45}$$

belonging to IF3, assuming that

$$R = \sum_{\alpha=1}^{m} R_\alpha.$$

Using the relations

$$y_{\hat{t}} = {}^1/_2 (y_t + y_{\bar{t}}), \qquad \tau y_{\bar{t}t} = y_t - y_{\bar{t}},$$

(3.45) can be rewritten as

$$(E + 2\tau R) y_t = -F_1, \qquad F_1 = (E - 2\tau R) y_{\bar{t}} + 2Ay - 2\varphi. \tag{3.46}$$

Substitution of $\tau y_{\bar{t}t} = 2y_{\hat{t}} - 2y_{\bar{t}}$ in (3.45) gives

$$(E + 2\tau R) y_{\check{t}} = -F_2, \qquad F_2 = Ay - 2\tau R y_{\bar{t}} - \varphi. \qquad (3.47)$$

Since $y_{\check{t}} \cdot y_{\bar{t}} + \frac{1}{2}\tau y_{\bar{t}t}$, it follows from (3.45) that

$$(E + 2\tau R) y_{\bar{t}t} = -F_3, \qquad F_3 = \frac{2}{\tau}(y_{\bar{t}} + Ay - \varphi). \qquad (3.48)$$

The right-hand sides F_1, F_2 and F_3 are known. Determination of $\hat{y}$ amounts to inversion of the operator $E + 2\tau R$. Replacing it by $E + 2\tau R$ and putting

$$E + 2\tau R = \prod_{\alpha=1}^{m} (E + 2\tau R_\alpha) = E + 2\tau R + 4\tau^2 Q, \qquad (3.49)$$

$$Q = \sum_{\alpha<\beta} R_\alpha R_\beta + 2\tau \sum_{\alpha<\beta<\gamma} R_\alpha R_\beta R_\gamma + \ldots + (2\tau)^{m-2} \prod_{\alpha=1}^{m} R_\alpha, \qquad (3.50)$$

three FS's are obtained, which can be written after simple manipulations as

$$(E + 4\tau^2 Q) y_{\check{t}} + \tau^2 (R + 2\tau Q) y_{\bar{t}t} + Ay = \widetilde{\varphi}, \qquad (3.51)$$

$$(E + 4\tau^2 Q) y_{\check{t}} + \tau^2 R y_{\bar{t}t} + Ay = \widetilde{\varphi}, \qquad (3.52)$$

$$y_{\check{t}} + \tau^2 (R + 2\tau Q) y_{\bar{t}t} + Ay = \widetilde{\varphi}. \qquad (3.53)$$

Some sufficient conditions for the stability of these schemes will be stated.

Theorem 10

Let

$$R_s^* = R_s, \qquad R_s R_k = R_k R_s \quad \text{for all} \quad s, k = 1, 2, \ldots, m, \qquad (3.54)$$

$$R = \sum_{s=1}^{m} R_s \geqslant \frac{1+\varepsilon}{4} A, \qquad (3.55)$$

where R_s are independent of t. Then each of schemes (3.51) - (3.53) is stable whatever $\tau > 0$, and their solutions satisfy (2.22) and (2.25), in which we have to put

$$\|\hat{y}\|_{(1)}^2 = \frac{1}{4}\Big(A(\hat{y} + y), \hat{y} + y\Big) + \tau^2 \left(\Big(R + 2\tau Q - \frac{1}{4}A\Big)y_t, y_t\right) \qquad (3.56)$$

for schemes (3.51) and (3.53), and

$$\|\hat{y}\|_{(1)}^{2} = \frac{1}{4}\left(A\left(\hat{y} + y\right), \hat{y} + y\right) + \tau^{2}\left(\left(R - \frac{1}{4}A\right)y_{t}, y_{t}\right) \qquad (3.57)$$

for scheme (3.52).

This is proved simply by showing that the conditions of Theorems 3 and 4 are satisfied. Notice that the present conditions (3.54) are of a constructional type. For instance, there is no difficulty in selecting the operators R_{α} for systems of parabolic equations. It is easily shown that Theorems 3 and 4 imply

Theorem 11

If

$$R = \sum_{\alpha=1}^{m} R_{\alpha} \geqslant \frac{1 + \varepsilon}{4}A, \qquad Q \geqslant -c_{1}R, \qquad (3.58)$$

then (2.22) and (2.25) hold for schemes (3.51) - (3.53) when τ is small: $\tau \leqslant \tau_{0}(c_{1})$.

We write $\tau^{2}Q$ as $\displaystyle\sum_{s=2}^{m} \tau^{2}Q_{s}$.

The order of accuracy of schemes (3.51) - (3.53) can be estimated from

Theorem 12

The solution of any one of equations (3.51) - (3.53) with right-hand side

$$\tilde{\varphi} = \sum_{s} \tau^{s}Q_{s}\eta \qquad (3.59)$$

and homogeneous initial conditions $y(0) = y_{t}(0) = 0$ satisfies

$$\|y(t)\|_{(1)} \leqslant M \max_{\omega_{\tau}} \sum_{s} \tau^{(s+1)/2}\|\eta\|_{q_{s}}, \qquad \|\eta\|_{q_{s}} = \overline{\gamma(Q_{s}\eta, \eta)}, \qquad (3.60)$$

provided the conditions of Theorem 10 are satisfied, $\|y\|_{(1)}$ being defined by (3.56) and (3.57).

When the order of accuracy of an FS is estimated, the approximation error is generally written as the sum of several terms: ($\psi = \psi_1 + \ldots + \psi_k$), each of which is estimated in its own norm. Combining Theorems 11 and 12, the estimate

$$\|z(t)\|_{(1)} \leqslant M \max_{\bar{\omega}_\tau} \sum_{s=1}^{k} \|\psi_s\|_{(2_s)}$$

is obtained for the error $z = y - {}^{P_N}u$, where u is the solution of the initial problem (§ 4). Various algorithms can be suggested for solving each equation. They all amount to successive inversion of the operators $E + 2\tau R_\alpha$. For instance, we rewrite scheme (3.51) as

$$\overset{m}{\underset{\alpha=1}{}}(E + 2\tau R_\alpha)y_t = -F_1. \tag{3.61}$$

Denoting $w = y_t$, we find $\hat{y}$ by solving successively the equations [10]

$$(E + 2\tau R_1)w_1 = -F_1, \qquad (E + 2\tau R_\alpha)w_\alpha = w_{\alpha-1}, \qquad 1 < \alpha \leqslant m, \tag{3.62}$$

$$\hat{y} = y + \tau w_m.$$

Schemes (3.52) and (3.53) are written as

$$\prod_{\alpha=1}^{m}(E + 2\tau R_\alpha)y_{\hat{t}} = -F_2, \tag{3.63}$$

$$\prod_{\alpha=1}^{m}(E + 2\tau R_\alpha)y_{\bar{t}t} = -F_3. \tag{3.64}$$

If we then put $w = y_{\hat{t}}$, $w = y_{\bar{t}t}$, we get

$$(E + 2\tau R_1)w_1 = -F_2, \qquad (E + 2\tau R_\alpha)w_\alpha = w_{\alpha-1}, \qquad 1 < \alpha \leqslant m,$$

$$\hat{y} = \check{y} + 2\tau w_m, \tag{3.65}$$

$$(E + 2\tau R_1)w_1 = -F_3, \qquad (E + 2\tau R_\alpha)w_\alpha = w_{\alpha-1}, \qquad 1 < \alpha \leqslant m,$$

$$\hat{y} = 2y - \check{y} + \tau^2 w_m. \tag{3.66}$$

To realize algorithms (3.65) and (3.66), storage of three vectors is required, while for algorithm (3.62), two vectors (y, $w_{\alpha-1}$) must be stored. But the total volume of computation is the same for all three algorithms. They were used in [10, 14] to solve the multidimensional equation of heat conduction in a parallelepiped.

Consider two cases when the scheme

$$By_{\overset{\circ}{t}} + \tau^2 Ry_{\overline{t}t} + Ay = \varphi, \qquad 0 < t \in \omega_\tau, \qquad y(0), \ y(\tau) \text{given}, \qquad (3.67)$$

can be factorized, without assuming that B is the unit operator.

1. Let B, R_1, R_2, ..., R_m be selfadjoint, constant, positive and commutable, and

$$R = \sum_{\alpha=1}^{m} R_\alpha \geqslant \frac{1+\varepsilon}{4} A.$$

Replacing $B + 2\tau R$ by

$$B \prod_{\alpha=1}^{m} (E + 2\tau B^{-1} R_\alpha) = B + 2\tau R + 4\tau^2 Q,$$

a factorized scheme similar to (3.51) is obtained:

$$(B + 4\tau^2 Q) y_{\overset{\circ}{t}} + \tau^2 (R + 2\tau Q) y_{\overline{t}t} + Ay = \varphi, \qquad y(0) = y_0, \qquad y(\tau) = y_1 \ (3.68)$$

which is stable, since $Q > 0$, $Q^* = Q$. Problem (3.68) can be solved by using the algorithm

$$(B + 2\tau R_1) w_1 = -F_1, \qquad (B + 2\tau R_\alpha) w_\alpha = B w_{\alpha-1}, \qquad \alpha > 1, \quad (3.69)$$

$$\hat{y} = y + \tau w_m.$$

2. Let $A^{(0)} = \sum_{\alpha=1}^{m} A_\alpha$, $B^{(0)} = \sum_{\alpha=1}^{m} B_\alpha$ be constant operators en. eq. to A and B, so that

$$0 < A \leqslant \gamma_1 A^{(0)}, \qquad 0 < B \leqslant \gamma_2 B^{(0)}. \qquad (3.70)$$

We put

$$R = \frac{1}{\tau} \left(E + \sum_{\alpha=1}^{m} R_\alpha \right) - \frac{1}{2\tau} B, \qquad R_\alpha = R_\alpha^{(a)} + R_\alpha^{(b)},$$

$$R_\alpha^{(a)} = \frac{1+\varepsilon}{4} \gamma_1 A_\alpha, \qquad R_\alpha^{(b)} = \frac{1}{2} \gamma_2 B_\alpha, \qquad \varepsilon > 0.$$

Recalling that $y_{t\circ} - \frac{1}{2}\tau y_{\bar{t}t} = y_{\dot{t}}$, (3.67) can be rewritten as

$$\tau \left(E + \sum_{\alpha=1}^{m} R_\alpha \right) y_{\bar{t}t} + B y_{\dot{t}} + A y = \varphi, \qquad y(0) = y_0, \qquad y(\tau) = y_1. \qquad (3.71)$$

This scheme can be factorized by replacing $E + \sum_{\alpha=1}^{m} R_\alpha$ by the product $\prod_{\alpha=1}^{m} (E + R_\alpha)$:

$$\tau \prod_{\alpha=1}^{m} (E + R_\alpha) y_{\bar{t}t} + B y_{\dot{t}} + A y = \varphi. \qquad (3.72)$$

We reduce FS (3.72) to the canonical form

$$B y_{\dot{t}} + \tau^2 \left(R + \frac{1}{\tau} Q \right) y_{\bar{t}t} + A y = \varphi, \quad 0 < t \in \omega_\tau, \quad y(0) = y_0, \quad y(\tau) = y_1. \qquad (3.73)$$

The conditions under which this scheme is stable are: (1) $B^* = B > 0$, $A_\alpha^* = A_\alpha > 0$, $B_\alpha^* = B_\alpha > 0$; (2) $A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_m$ are mutually commutable. The condition $\tilde{R} = R + Q/\tau \geqslant (1+\varepsilon)A/4$ is satisfied, since $Q > 0$, $R \geqslant E/\tau + \frac{1}{4}(1+\varepsilon)A$ by construction. Scheme (3.73) belongs to IF3, and Theorems 3 and 4 hold for it, with

$$\|\hat{y}\|_{(1)}^2 = \frac{1}{4}(A(\hat{y} + y), \hat{y} + y) + \tau^2 \left(\left(R + \frac{1}{\tau} Q - \frac{1}{4} A \right) y_t, y_t \right). \qquad (3.74)$$

The solution of (3.72) is found by successive inversion of the operators $E + R_\alpha$, $\alpha = 1, 2, \ldots, m$ (cf. (3.66)). FS (3.73) can be used for solving problems of mathematical physics that lead to the abstract Cauchy problem

$$\mathscr{B}(t) \frac{du}{dt} + \mathscr{A}(t) u = f(t) \qquad (3.75)$$

in real Hilbert space. The relevant FS (3.73) has the first order of accuracy in τ. By considering schemes with a larger number of layers, economic schemes of the second order of accuracy in τ can be constructed.

Three-layer schemes of the type

$$y_{\bar{t}t} + \tau^2 R y_{\mathring{t}} + Ay = \varphi, \ 0 < t \in \omega_\tau, \ y(0) = y_0, \ y_t(0) = \bar{y}_0, \qquad R = \sum_{\alpha=1}^{m} R_\alpha,$$
$$(3.76)$$

are factorized similarly to scheme (3.45); $E + 2\tau R$ in (3.46) - (3.48) has to be replaced by $E + \tau^2 R$. The resulting factorized schemes are

$$(E + \tau^2 R + {}^1/_2 \tau^4 Q) y_{\bar{t}t} + \tau^3 Q y_{\mathring{t}} + Ay = \varphi, \tag{3.77}$$

$$(E + \tau^2 R) y_{\bar{t}t} + \tau^3 Q y_{\mathring{t}} + Ay = \varphi, \tag{3.78}$$

$$(E + \tau^2 R + \tau^4 Q) y_{\bar{t}t} + Ay = \varphi, \tag{3.79}$$

similar to schemes (3.51 - (3.53). Here,

$$Q = \sum_{\alpha < \beta} R_\alpha R_\beta + \tau^2 \sum_{\alpha < \beta < \gamma} R_\alpha R_\beta R_\gamma + \ldots + \tau^{2m-2} \prod_{\alpha=1}^{m} R_\alpha. \tag{3.80}$$

The sufficient conditions for stability are similar to the conditions for (3.51) - (3.53); their statement must be omitted for lack of space.

4. Examples

1. The regularization methods described in Section 3 will be illustrated by examples for parabolic equations

$$\frac{\partial u}{\partial t} = Lu + f(x, t), \qquad u = u(x, t), \qquad x = (x_1, \ldots, x_p) \in G, \tag{4.1}$$

and hyperbolic equations

$$\frac{\partial^2 u}{\partial t^2} = Lu + f(x, t), \tag{4.2}$$

where Lu is a 2nd order elliptic operator.

Let G be a region of p-dimensional Euclidean space, with boundary Γ. Only boundary conditions of the 1st kind, specified on Γ, will be considered.

To apply the theory of Section 3, difference operators Λ, approximating the differential operators L, must first be obtained. Let $\bar{\omega}_h(\bar{G})$ be a rectangular mesh in the region $G + \Gamma = \bar{G}$. We consider the set of mesh functions specified at interior base-points ω_h of the mesh $\bar{\omega}_h$ and

introduce into it the scalar product [24]

$$(y, v) = \sum_{\omega_h} y(x) v(x) H, \qquad H = \prod_{\alpha=1}^{p} h_\alpha, \qquad h_\alpha = {}^1/_2(h_\alpha + h_\alpha^{(+1\alpha)}),$$

where h_α is the mesh interval relative to x_α (the mesh may be non-uni-form), and the norm $\|y\| = \sqrt{(y, y)}$. The result is a space H. Let A be an operator mapping H into H. The values of Ay are the same as the values of $-\Lambda y$ on the set $\bar{\Omega}_0$ of mesh functions which vanish on the boundary γ_h of the mesh. Using Green's difference formulae, the operator $A^{(0)}$ en. eq. with A is easily found. As a rule, the Laplace difference operator

$$\Lambda^{(0)}y = \sum_{\alpha=1}^{p} y_{\bar{x}_\alpha \hat{x}_\alpha}, \qquad A^{(0)}y = -\Lambda^{(0)}y \qquad (4.3)$$

or the operator without mixed derivatives

$$\Lambda^{(0)}y = \sum_{\alpha=1}^{p} (\bar{a}_\alpha(x_\alpha) y_{\bar{x}_\alpha})_{\hat{x}_\alpha}, \qquad A^{(0)}y = -\Lambda^{(0)}y. \qquad (4.4)$$

is taken as the similitude operator. Once $A^{(0)}$ is found, Theorems 5 and 6 can be employed.

2. We start with the elementary operator

$$\Lambda y = (a(x) y_{\bar{x}})_x, \qquad 0 < c_1 \leqslant a \leqslant c_2,$$

on the mesh $\overline{\omega}_h = \{x_i = ih, \ i = 0, 1, \ldots, N, \ h = 1/N\}$, corresponding to the differential operator $Lu = (k(x)u')'$. In this case (using the notation of [12])

$$(y, v) = \sum_{i=1}^{N-1} y_i v_i h, \qquad \|y\| = \sqrt{(y, y)}. \qquad (4.5)$$

The difference formula gives

$$(-\Lambda y, y) = (a, y_{\bar{x}}^2] \leqslant c_2(1, y_{\bar{x}}^2) = c_2(-\Lambda^{(0)}y, y), \qquad \Lambda^{(0)}y = y_{\bar{x}x},$$

whence

$$c_1(A^{(0)}y, y) \leqslant (Ay, y) \leqslant c_2(A^{(0)}y, y), \qquad Ay = -(ay_{\bar{x}})_x, \qquad (4.6)$$

$$A^{(0)}y = -y_{\bar{x}x}. \qquad (4.7)$$

3. Now consider the selfadjoint elliptic operator

$$\Lambda y = \sum_{\alpha=1}^{p} (a_\alpha(x) y_{\bar{x}_\alpha})_{\hat{x}_\alpha}, \qquad 0 < c_1 \leqslant a_\alpha(x) \leqslant c_2, \qquad (4.8)$$

corresponding to the differential operator

$$Lu = \sum_{\alpha=1}^{p} \frac{\partial}{\partial x_\alpha} \left(k_\alpha(x) \frac{\partial u}{\partial x_\alpha} \right). \qquad (4.9)$$

By using Green's difference formula, it can be seen that

$$c_1(-\Lambda^{(0)}y, y) \leqslant (-\Lambda y, y) \leqslant c_2(-\Lambda^{(0)}y, y), \qquad y \in \overline{\Omega}_0,$$

where

$$\Lambda^{(0)}y = \sum_{\alpha=1}^{p} y_{\bar{x}_\alpha \hat{x}_\alpha}.$$

In this case, therefore,

$$Ay = -\Lambda y, \qquad A^{(0)}y = -\Lambda^{(0)}y = -\sum_{\alpha=1}^{p} y_{\bar{x}_\alpha \hat{x}_\alpha}, \qquad (4.10)$$

$$A^{(0)}y = \sum_{\alpha=1}^{p} A_\alpha y, \qquad A_\alpha y = -y_{\bar{x}_\alpha \hat{x}_\alpha}.$$

If $a_\alpha(x)$ satisfies

$$c_1 \bar{a}_\alpha(x_\alpha) \leqslant a_\alpha(x) \leqslant c_2 \bar{a}_\alpha(x_\alpha), \qquad \alpha = 1, \ldots, p, \qquad (4.11)$$

the operator en. eq. with A can be the operator with separable variables

$$A^{(0)}y = \sum_{\alpha=1}^{p} A_\alpha y, \qquad A_\alpha y = -(\bar{a}_\alpha(x_\alpha) y_{\bar{x}_\alpha})_{\hat{x}_\alpha}. \qquad (4.12)$$

4. Consider the elliptic difference operator with mixed derivatives

$$\Lambda y = \frac{1}{2} \sum_{\alpha, \beta=1}^{p} [(k_{\alpha\beta}(x) y_{\bar{x}_\beta})_{x_\alpha} + (k_{\alpha\beta}(x) y_{x_\beta})_{\bar{x}_\alpha}], \qquad (4.13)$$

corresponding to the differential operator

$$Lu = \sum_{\alpha, \beta=1}^{p} \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\beta}(x) \frac{\partial u}{\partial x_\beta} \right). \qquad (4.14)$$

Suppose: (1) that

$$c_1 \sum_{\alpha=1}^{p} \xi_\alpha^2 \leqslant \sum_{\alpha,\,\beta=1}^{p} k_{\alpha\beta}(x)\xi_\alpha\xi_\beta \leqslant c_2 \sum_{\alpha=1}^{p} \xi_\alpha^2;$$

(2) that $G = G_0$ is the p-dimensional parallelepiped ($0 \leqslant x_\alpha \leqslant l_\alpha$, $\alpha = 1, 2, \ldots, p$); (3) the mesh $\bar\omega_h = \{(i_1 h_1, \ldots, i_p h_p), i_\alpha = 0, 1, \ldots, N_\alpha, h_\alpha = l_\alpha / V_\alpha, \alpha = 1, \ldots, p\}$ is uniform relative to each of the x_α.

The operator en. eq. with $4y = -\Lambda y$ is

$$A^{(0)}y = \sum_{\alpha=1}^{p} A_\alpha y, \qquad A_\alpha y = -y_{\bar x_\alpha x_\alpha}, \qquad y \in \bar\Omega_0, \tag{4.15}$$

so that $c_1 A^{(0)} \leqslant A \leqslant c_2 A^{(0)}$. Comparison of (4.15) and (4.10) shows that operator (4.15) is en. eq. with the operators (4.8) and (4.13).

5. Consider the operator of the system of elliptic equations

$$\Lambda y^i = -\frac{1}{2} \sum_{\alpha,\,\beta=1}^{p} \sum_{j=1}^{n} [(k_{\alpha\beta}^{ij}(x)\, y_{x_\beta}^j)_{x_\alpha} + (k_{\alpha\beta}^{ij}(x)\, y_{x_\beta}^j)_{\bar x_\alpha}], \quad i = 1, 2, \ldots, n, \tag{4.16}$$

$$c_1 \sum_{i=1}^{n} \sum_{\alpha=1}^{p} (\xi_\alpha^i)^2 \leqslant \sum_{i,\,j=1}^{n} \sum_{\alpha,\,\beta=1}^{p} k_{\alpha\beta}^{ij}\xi_\beta^j\xi_\alpha^i \leqslant c_2 \sum_{i=1}^{n} \sum_{\alpha=1}^{p} (\xi_\alpha^i)^2. \tag{4.17}$$

Denoting by $y = (y_1, \ldots, y_n)$ the vector, and by $k_{\alpha\beta} = (k_{\alpha\beta}^{ij})$ the $n \times n$ matrix, (4.16) can be written as

$$\Lambda y = -\frac{1}{2} \sum_{\alpha,\,\beta=1}^{p} [(k_{\alpha\beta}(x)\, y_{\bar x_\beta})_{x_\alpha} + (k_{\alpha\beta}(x)\, y_{x_\beta})_{\bar x_\alpha}]. \tag{4.18}$$

$G = G_0$ is a parallelepiped, while the mesh ω_h is uniform relative to x_α

Recalling (4.17), an operator $4^{(0)}$ with a diagonal matrix of coefficients can be taken as en. eq. with $4y = -\Lambda y$, so that

$$A^{(0)}y = \sum_{\alpha=1}^{p} A_\alpha y, \qquad (A_\alpha y)^i = -y_{\bar x_\alpha x_\alpha}^i, \tag{4.19}$$

$$c_1(A^{(0)}y,\, y) \leqslant (Ay,\, y) \leqslant c_2(A^{(0)}y,\, y).$$

The operator without mixed derivatives

$$\Lambda y^i = \sum_{\alpha=1}^{p} \sum_{j=1}^{n} (a_\alpha^{ij} y_{\bar x_\alpha}^j)_{\hat x_\alpha}, \tag{4.20}$$

$$c_1 \sum_{i=1}^{n} (\xi_\alpha{}^i)^2 \leqslant \sum_{i,\, j=1}^{n} a_\alpha{}^{ij} \xi_\alpha{}^i \xi_\alpha{}^j \leqslant c_2 \sum_{i=1}^{n} (\xi_\alpha{}^i)^2$$

can be considered on a non-uniform mesh in an arbitrary region $\overline{G}$. For this,

$$A^{(0)}y = \sum_{\alpha=1}^{p} A_\alpha y, \qquad (A_\alpha y)^i = -y_{\overline{x}_\alpha \hat{x}_\alpha}^{\, i}. \tag{4.21}$$

6. After finding the en. eq. operators $A^{(0)}$, the regularizers R for the two- and three-layer schemes can be obtained from Theorems 5 and 6:

$$R = \sigma A^{(0)},$$

where $\sigma \geqslant \frac{1}{2} c_2$ for a two-layer scheme, and $\sigma \geqslant ((1 + \varepsilon)/4)c_2$ for a three-layer scheme. All the operators A and $A^{(0)}$ written above are self-adjoint and positive definite.

Let $\overline{G} = \overline{G}_0 = (0 \leqslant x_\alpha \leqslant l_\alpha, \ \alpha = 1, 2, \ldots, p)$ be a p-dimensional parallelepiped. Then the operators A_α, $\alpha = 1, 2, \ldots, p$ commute with one another. The parabolic equations (4.1) and systems of equations ($u = (u_1, \ldots, u_n)$ is a vector) can then be solved by means of the factorized schemes of Section 3, §§ 3 and 4, after putting $R_\alpha = \sigma A_\alpha$. As a result, economic schemes are obtained for solving the first boundary value problem; they are stable whatever $\tau > 0$ on any mesh ω_h under a natural parabolicity condition, and have the accuracy $O(|h|^2 + \tau)$ in the case of two-layer schemes, or $O(|h|^2 + \tau^2$ in the case of three-layer schemes. Independently of whether a single equation without mixed derivatives is considered, or a system of equations with mixed derivatives, the transition from layer to layer is always realized by inversion of three-point difference operators of the 2nd order in accordance with the usual successive substitution formulae.

The resulting factorized schemes are convergent when the coefficients of the multidimensional parabolic equation have a finite number of discontinuities of the 1st kind on hyperplanes parallel to the coordinate hyperplanes. The approximation error of an FS is investigated in the neighbourhood of a discontinuity in the same way as for ordinary multi-dimensional schemes [18, 24]. For instance, in the case of a two-layer scheme, the main role is played by the *a priori* estimate for the problem

$$z_t + \tau(R + \tau Q)z_t + Az = \psi, \qquad z|_{t=0} = 0,$$

with right-hand side $\psi = -\tau R \eta$, $\eta = u_t$. The estimate will not be

written down, in view of its analogy with the estimates of [18, 19].
Analogues of Theorems 9 and 12 can be used for weakening the smoothness
conditions for the coefficients and solution, even in the case of con-
tinuous $k_\alpha(x, t)$.

The results of the general theory can readily be used for construct-
ing economic FS in the case of boundary conditions of the 3rd kind.

Since the same regularizers are used for schemes approximating equa-
tions of the hyperbolic type (4.2) as for a three-layer scheme corre-
sponding to equation (4.1), there is no need for a separate description
of the FS for (4.2); they follow easily from the general theory of
Section 3.

Acknowledgements. The author thanks A.N. Tikhonov for discussion and
B.V. Andreev for comments at the editing stage.

REFERENCES

1. SAMARSKII, A.A. On a theory of difference schemes, *Dokl. Akad. Nauk
    SSSR*, **165**, 5, 1007 - 1011, 1965.

2. SAMARSKII, A.A. On difference schemes for multidimensional differ-
    ential equations of mathematical physics, *Aplik. mat.* 10, 2, 146 -
    164, 1965.

3. FADDEEV, D.K. and FADDEEVA, V.N. *Computational methods of linear
    algebra* (Vychislitel'nye metody lineinoi algebry), Fizmatiz,
    Moscow, 1960.

4. KANTOROVICH, L.V. Functional analysis and applied mathematics, *Usp.
    mat. Nauk*, 3, 6, 89 - 185, 1948.

5. WASOW, W.R. and FORSYTHE, G.E. *Difference methods for partial differ-
    ential equations*, IL, Moscow, 1963.

6. SAUL'EV, V.K. *Integration of parabolic equations by the mesh method*
    (Integrirovanie uravnenii parabolicheskogo tipa metodom setok),
    Fizmatgiz, Moscow, 1960.

7. SAMARSKII, A.A. Some aspects of the theory of difference schemes,
    Zh. vychisl. Mat. mat. Fiz. 6, 4, 665 - 686, 1966.

8. D'YAKONOV, E.G. Difference schemes with a splitting operator for multidimensional non-stationary problems, *Zh. vўchisl. Mat. mat. Fis.* **2**, 4, 549 - 568, 1962.

9. D'YAKONOV, E.G. Application of difference schemes with a splitting operator to some systems of parabolic and hyperbolic equations, *Sib. mat. Zh.* **4**, 3, 509 - 515, 1965.

10. SAMARSKII, A.A. Schemes of improved order of accuracy for the multi-dimensional equation of heat conduction, *Zh. vўchisl. Mat. mat. Fiz.* **3**, 5, 812 - 840.

11. HANENKO, N.N. *A fractional step method for solving multidimensional problems of mathematical physics* (Metod drobnykh shagov resheniya mnogomernykh zadach matematicheskoi fiziki), SO AN SSSR, Novosibirsk, 1966.

12. SAMARSKII, A.A. Uniform difference schemes on non-uniform meshes for parabolic equations. *Zh. vўchisl. Mat. mat. Fiz.* **3**, 2, 266 - 288, 1963.

13. RYABEN'KII, V.S. and FILIPPOV, A.F. *The stability of difference equations* (Ob ustoichivosti raznostnykh uravnenii), Gostekh izdat, Moscow, 1956.

14. D'YAKONOV, E.G. Difference schemes with a splitting operator for multidimensional parabolic equations with variable coefficients, In Symp. *Computer methods and programming* (Vychisl. metody i programmirovanie), No. 3, 163 - 190, MGU, Moscow, 1965.

15. D'YAKONOV, E.G. A majorant operator method for solving some strongly elliptic systems, *Usp. mat. Nauk*, **19**, 5, 197 - 198, 1964.

16. D'YAKONOV, E.G. Construction of iterative methods by using spectrally equivalent operators, *Zh. vўchisl. Mat. mat. Fiz.* **6**, 1, 12 - 34, 1966.

17. SAMARSKII, A.A. An economic algorithm for numerical solution of systems of differential and algebraic equations, *Zh. vўchisl. Mat. mat. Fiz.* **4**, 3, 580 - 585, 1964.

18. SAMARSKII, A.A. Convergence and accuracy of uniform difference schemes for one- and multidimensional parabolic equations, *Zh. vўchisl. Mat. mat. Fiz.* **2**, 4, 603 - 633, 1962.

19. SAMARSKII, A.A. *A priori* estimates for difference equations, *Zh. vychisl. Mat. mat. Fiz.* **1**, 6, 972 - 1000, 1961.

20. YANENKO, N.N. On implicit difference methods for solving the multi-dimensional equation of heat conduction, *Izv. vyssh. ucheb Zaved. (Series Mat.)*, 4 (23), 148 - 157, 1961.

21. YANENKO, N.N. A difference method for solving the multidimensional equation of heat conduction, *Dokl. Akad. Nauk SSSR*, **125**, 6, 1207 - 1210, 1959.

22. LYUSTERNIK, L.A. and SOBOLEV, V.I. *Elements of functional analysis* (Elementy funktsional'nogo analiza), Moscow, Nauka, 1965.

23. KANTOROVICH, L.V. and AKILOV, G.P. *Functional analysis in normed spaces* (Funktsional'nyi analiz v normirovannykh prostranstvakh), Fizmatgiz, Moscow, 1959.

24. SAMARSKII, A.A. Locally one-dimensional schemes on non-uniform meshes, *Zh. vychisl. Mat. mat. Fiz.* **3**, 1, 5 - 63, 1963.