ON THE WELL–POSEDNESS OF DIFFERENCE SCHEMES WITH GENERALIZED DATA

N. Yu. Bakaev

1. Initial-boundary value problems with non-smooth or generalized data are extensively studied in mathematical physics, and in this connection the concept of a generalized solutions is often used [1]. In the approximation of such problems by difference schemes one uses discrete elements, which may be thought of as approximations of some generalized elements in the original formulation of the problem (for example, discrete δ-functions). Singularity properties of such approximations cannot be exhibited in a fixed mesh, but in a sequence of meshes with sufficiently refined step. A characteristic sign of generalized discrete elements is a lack of uniformity with respect to the mesh size in the bounds of these elements in the original mesh norm. It is possible to obtain uniform bounds for generalized elements only in a weaker discrete norm. Because of this, in the presence of generalized elements in the difference scheme, the use of ordinary stability estimates in the original discrete norm may lead to a lowering of the order of convergence of the scheme. For the construction of maximally accurate convergence estimates for difference schemes with generalized data, Tikhonov and Samarskiĭ [2–5] suggested the use of stability estimates, comparing the solution of the discrete problem in a basic norm to the data in a weaker norm. In the following years these ideas have been further developed [6,7].

In the present work we develop a general approach to the study of the well-posedness of difference schemes with generalized data in Banach space norms. As in [2–5] the discrete solution and the data are considered in discrete spaces, with the data in a space with weaker norm. The approach presented below is based on principles, that were used in [8–12], allowing a uniform approach to the study of wide classes of difference schemes with generalized data. We obtain general results that allow us, in particular, to give conditions for the well-posedness of difference schemes for parabolic partial differential equations in a scale of Banach spaces L_{ph}, $1 \le p \le \infty$ (discrete analogues of the spaces L_p).

2. Let there be given two families of Banach spaces E_h and $\tilde{E}_h$, parametrized by a scalar or vectorial parameter h. We call the spaces E_h of the first family basic, and the spaces $\tilde{E}_h$ of the second family negative, if the spaces E_h and $\tilde{E}_h$ are connected by the imbedding

$$E_h \subset \tilde{E}_h, \qquad \forall h,$$

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65J10, 65M10.

uniformly in h in the sense that

$$\|u\|_{\tilde{E}_h} \leq C_0\|u\|_{E_h}, \qquad \forall h,\ \forall u \in E_h \tag{1}$$

with a constant C_0 independent of h and u.

In the family of pairs of basic and negative spaces $(E_h, \tilde{E}_h)$ we consider a difference scheme, written in canonical form,

$$\begin{aligned} &y(t_{k+1}) = [I - \tau\hat{A}_{\tau h}(t_k)]y(t_k) + \tau Q_{\tau h}F_{\tau h}(t_k), \qquad k = 0,\dots,T/\tau - 1,\\ &y(t_0) = \mathcal{R}_{\tau h}y_{0h}, \end{aligned} \tag{2}$$

where $t_k = k\tau \in [0,T]$, $k = 0,\dots,T/\tau$ is a discrete argument ($T > 0$ is the upper limit of its size, and τ is the discretization step), $y(t_k)$ is a discrete function of the argument t_k with values in E_h, representing the solution of the discrete problem (2), $F_{\tau h}(t_k)$ is a given function of the discerete argument t_k with values in $\tilde{E}_h$, defining the inhomogeneity of the scheme (2) and depending on τ and h, $y_{0h} \in \tilde{E}_h$ is the initial datum of the scheme, $\hat{A}_{\tau h}(t_k)$, $Q_{\tau h}(t_k)$ and $\mathcal{R}_{\tau h}$ are linear operators acting in $\tilde{E}_h$ (not necessarily defined in all of $\tilde{E}_h$), depending on t_k (except for $\mathcal{R}_{\tau h}$), τ and h. If $y(t_k)$ does not belong to the space E_h for all $k = 0,\dots,T/\tau$, then we shall say that the scheme (2) is not completely solvable, but the totality of elements $y(t_k) \in E_h$ for all corresponding k defines an incomplete solution of the problem (2) in the family of pairs $(E_h, \tilde{E}_h)$. The operator $\hat{A}_{\tau h}(t_k)$ is called the generator of the scheme [12], and the operators $Q_{\tau h}(t_k)$ and $\mathcal{R}_{\tau h}$ are naturally called smoothing operators for the right-hand side and initial data, respectively. Since the data $F_{\tau h}(t_k)$ and y_{0h} of the scheme (2) may not belong to the space E_h, we are lead to consider generalized data.

Theorem 1. *Let the operator $\hat{A}_{\tau h}(t_k)$, the generator of the scheme (2), be defined in the family of pairs of basic and negative spaces $(E_h, \tilde{E}_h)$, linear and bounded in the space E_h, and satisfying in $\tilde{E}_h$ the $(\Lambda_{2\tau})$-condition of [12], uniformly with respect to t_k (in the sense that all objects occuring in the $(\Lambda_{2\tau})$-condition do not depend on t_k), and let, in addition, the following condition hold*

$$\begin{aligned} &\|[\hat{A}_{\tau h}(t_k) - \hat{A}_{\tau h}(t_l)]u\|_{E_h} \leq C_1\sum_{j=1}^{r_0}|t_k - t_l|^{\theta_j}\|[\hat{A}_{\tau h}(t_k) + \mu_0 I]u\|_{E_h}^{\chi_j}\|u\|_{E_h}^{1-\chi_j},\\ &\forall h,\ \forall \tau \in (0,\tau_0],\ \forall t_k, t_l \in [0,T],\ \forall u \in E_h \end{aligned}$$

for some natural number $r_0 \geq 1$ and with some constants $C_1,\ \mu_0 \geq 0,\ \tau_0 > 0,\ \theta_j \in [0,1],\ \chi_j \in [0,1],\ j = 1,2,\dots,r_0$, independent of $h,\ \tau,\ u,\ t_k,\ t_l$. Let also the operators $Q_{\tau h}$ and $\mathcal{R}_{\tau h}$ be bounded and linear from $\tilde{E}_h$ into E_h, satisfying the conditions

$$\begin{aligned} &\|R(\lambda, \tau\hat{A}_{\tau h}(t_k))Q_{\tau h}(t_k)u\|_{E_h} \leq C_2|\lambda|^{\gamma-1}\tau^{-\gamma}\|u\|_{\tilde{E}_h},\\ &\forall h,\ \forall \tau \in (0,\tau_0],\ \forall t_k \in [0,T],\ \forall \lambda \notin Int\,\Lambda_\tau,\ \forall u \in \tilde{E}_h; \end{aligned} \tag{3}$$

$$\begin{aligned} &\|R(\lambda, \tau\hat{A}_{\tau h}(t_0))\mathcal{R}_{\tau h}u\|_{E_h} \leq C_3|\lambda|^{\gamma-1}\tau^{-\gamma}\|u\|_{\tilde{E}_h},\\ &\forall h,\ \forall \tau \in (0,\tau_0],\ \forall \lambda \notin Int\,\Lambda_\tau,\ \forall u \in \tilde{E}_h \end{aligned} \tag{4}$$

with constants C_2, C_3 and $\gamma \in [0,1],\ \gamma < \theta_j + 1 - \chi_j,\ j = 1,2,\dots,r_0$, where $\Lambda_\tau = \{\lambda : |\lambda| \le a\tau\} \cup \Lambda$, $a = const \ge 0$, and the fixed set Λ has a configuration of type (Λ_2) (cf. [12]).

Then there exists a solution $y(t_k) \in E_h,\ k = 0,1,\dots,T/\tau$, of the difference scheme (2) and it satisfies the following stability estimate

$$\begin{aligned}(5)\qquad &\|[\hat{A}_{\tau h}(t_k) + \bar{\mu} I]^\xi y(t_k)\|_{E_h} \le C_4(\xi)\Big\{[(k+1)\tau]^{-(\xi+\gamma)}\|y_{0h}\|_{\tilde{E}_h}\\ &+ \tau\sum_{l=1}^{k}[(k-l+1)\tau]^{-(\xi+\gamma)}\|F_{\tau h}(\tau_{l-1})\|_{\tilde{E}_h}\Big\}, \qquad k = 0,1,\dots,T/\tau,\\ &\forall h,\ \forall\tau \in (0,\tau_0],\ \forall \xi \in [0, 1-\gamma)\end{aligned}$$

with nonnegative constants $C_4(\xi),\ \bar{\mu}$, independent of h, τ, and k ($C_4(\xi)$ may depend on ξ). If in addition the operator $\hat{A}_{\tau h}(t_k)$ can be extended to be linear and bounded on $\tilde{E}_h$, and instead of (4) the following inequality holds

$$\begin{aligned}&\|[R(\lambda, \tau\hat{A}_{\tau h}(t_k)) - (\lambda - 1)^{-1} I]u\|_{E_h} \le C_5|\lambda|^{\gamma-1}\tau^{-\gamma}\|u\|_{\tilde{E}_h}, \quad C_5 = const,\\ &\forall h,\ \forall \tau \in (0,\tau_0],\ \forall t_k \in [0,T],\ \forall \lambda \notin Int\,\Lambda_\tau,\ \forall u \in \tilde{E}_h\end{aligned}$$

for the same values of γ as in (3), then for $\mathcal{R}_{\tau h} = I$ there exists an incomplete solution $y(t_k) \in E_h,\ k = 1,2,\dots,T/\tau$, of problem (2), and the estimate (5) holds for these k. In the case where the operator $\hat{A}_{\tau h}(t_k)$ does not depend on t_k the same statement holds with no upper bound for ξ (in this case we may formally consider $r_0 = 1,\ \theta_1 = 1,\ \chi_1 = 0$).

The proof is based on techniques devoloped in [8, 9, 12].

3. In the family of pairs $(E_h, \tilde{E}_h)$ of basic and negative spaces we consider a difference scheme of the form

$$\begin{aligned}(6)\qquad &y(t_{k+1}) = y(t_k) - \tau\sum_{j=1}^{\nu_1} b_j A_h(t_k + c_j\tau)Y_j + \tau\sum_{j=1}^{\nu_2}\bar{b}_j f_h(t_k + \bar{c}_j\tau),\\ &y(t_0) = \mathcal{R}_{\tau h} y_{0h}, \qquad k = 0,1,\dots,T/\tau - 1,\end{aligned}$$

where Y_j are defined by the following system of equations

$$(7)\qquad Y_j = \beta_j y(t_k) - \tau\sum_{l=1}^{\nu_1} a_{jl} A_h(t_k + c_l\tau)Y_l + \tau\sum_{l=1}^{\nu_2}\bar{a}_{jl} f_h(t_k + \bar{c}_l\tau),\ j = 1,2,\dots,\nu_1;$$

here $b_j,\ \bar{b}_l,\ \beta_j,\ a_{jr},\ \bar{a}_{jl},\ c_j,\ \bar{c}_l;\ j,r = 1,2,\dots,\nu_1,\ l = 1,2,\dots,\nu_2$, are coefficients defining the scheme, $A_h(t)$ is an operator depending on $t \in [0,T]$ and h, $f_h(t)$ is a function of $t \in [0,T]$ with values in $\tilde{E}_h$, depending on h and defining the inhomogeneity of the scheme, and the remaining notation has the same meaning as above (see no. 2). A solution (incomplete solution) $y(t_k) \in E_h$ of the difference problem (6), (7) is understood in the same sense as in the case of the difference scheme (2). The data $f_h(t) \in \tilde{E}_h$ and $y_{0h} \in \tilde{E}_h$ of the scheme (6), (7) are generalized.

Difference schemes of the form (6), (7) arise from the discretization in t of a family of Cauchy problems with parameter h of the form

$$\frac{dy}{dt} + A_h(t)y = f_h(t), \quad 0 < t \leq T; \qquad y(t=0) = y_{0h}. \tag{8}$$

The discretization method, consisting in the passage from the family of problems (8) to the difference scheme (6), (7), was studied from the point of view of stability of (6), (7) in the norms of the basic spaces in [11]. This method includes both the well-known Runge-Kutta method [13], widely popular for the solution of ordinary differential equations, and the method of Vinokurov [14] developed recently.

We extend somewhat the concept of a uniformly strongly semi-positive operator defined in [11]. We say that a linear operator $A_h(t)$ acting in the Banach space E_h is uniformly (with respect to $t \in [0,T]$ and h) strongly semi-positive with an angle $\varphi \in (0, \pi/2)$, if it is closed, has a domain of definition D_h, which is independent of t and dense in E_h, and in addition

$$\|R(\lambda, A_h(t))\|_{E_h} \leq C_6 |\lambda|^{-1}, \qquad \forall h,\ \forall t \in [0,T],\ \forall \lambda : |\arg \lambda| \geq \varphi,\ |\lambda| \geq b_0$$

with some constants C_6 and b_0 independent of h, t and λ. In addition we will use below the following concepts introduced in [11]: the symbol $\alpha(z)$ of the generator of the scheme, the correcting symbols $\omega_j(z)$, $j = 1, 2, \ldots, \nu_2$, and difference scheme of type $R(\varphi)$. We also define the concept of underlying symbols, by which we understand functions $\eta_j(z)$, $j = 1, 2, \ldots, \nu_1$, coinciding with the corresponding correcting symbols $\omega_j(z)$ when the parameters ν_2, $\bar{a}_{jr}$, $\bar{b}_r$ of the scheme (6), (7) are prescribed in a special way, namely, when $\nu_2 = \nu_1$, $\bar{a}_{jr} = a_{jr}$, $\bar{b}_r = b_r$, $j, r = 1, 2, \ldots, \nu_1$. Finally, we let the range of the scheme be the complex number λ_0 defined by

$$\lambda_0 = \lim_{|z| \to \infty} \alpha(z),$$

where $\alpha(z)$ is the symbol of the generator of the scheme.

Theorem 2. *Let the difference scheme (6), (7) be given in a family of pairs $(E_h, \tilde{E}_h)$ of basic and negative spaces, and belong to the type $R(\varphi_0)$ for some angle $\varphi_0 \in (0, \pi/2)$, whereby $deg[\alpha(z)] = 0$, where $\alpha(z)$ is the symbol of the generator of the scheme (the parameter $deg[\rho(z)]$ is defined in [11] for an arbitrary rational function $\rho(z)$). Let, in addition, the following conditions hold for the scheme (6), (7)*

$$deg[g_{jl}(z)] \leq 0,\ deg\left[\sum_{k=1}^{\nu_1} g_{jk}(z)\bar{a}_{kr}\right] \leq -1,\ j, l = 1, 2, \ldots, \nu_1,\ r = 1, 2, \ldots, \nu_2; \tag{9}$$

$$deg\left[\sum_{r=1}^{\nu_1} g_{jr}(z)\beta_r\right] \leq -1,\ j = 1, 2, \ldots, \nu_1, \tag{10}$$

where $g_{jl}(z)$ *are the elements of the inverse matrix of the matrix* $(\delta_{jl}+za_{jl})$ $(\delta_{jl}$ *is the Kronecker symbol), and we have the following representations of the underlying symbols* $\eta_j(z)$ *and the correcting symbols* $\omega_l(z)$

$$\eta_j(z)=[1-\lambda_0^{-1}\alpha(z)]\bar{\eta}_j(z),\ j=1,2,\ldots,\nu_1, \tag{11}$$

$$\omega_l(z)=[1-\lambda_0^{-1}\alpha(z)]\bar{\omega}_l(z),\ j=1,2,\ldots,\nu_2, \tag{12}$$

where λ_0 *is the range of the scheme, and* $\bar{\eta}_j(z),\ j=1,2,\ldots,\nu_1,\ \bar{\omega}_l(z),\ j=1,2,\ldots,\nu_2,$ *are rational functions such that* $\deg[\bar{\eta}_j(z)]\le 0,\ j=1,2,\ldots,\nu_1,\ \deg[\bar{\omega}_l(z)]\le 0,\ j=1,2,\ldots,\nu_2,$ *and the poles of these functions are outside the closed sector* $\{z:|\arg z|\le\varphi_0\}$. *We assume also that the operator* $A_h(t)$ *is uniformly strongly semi-positive with angle* φ_0 *in the space* E_h, *admits an extension to an operator in the space* $\tilde{E}_h$ *which is uniformly strongly semi-positive with angle* φ_0 *(the value of* φ_0 *is the same as in the previous conditions of the theorem), and satisfies the conditions*

$$\begin{aligned}
&\|[A_h(t)-A_h(s)]u\|_{E_h}\le C_7\sum_{j=1}^{q_0}|t-s|^{\sigma_j}\|[A_h(t)+\mu_1 I]u\|_{E_h}^{\kappa_j}\|u\|_{E_h}^{1-\kappa_j},\\
&\forall h,\ \forall t,s\in[0,T],\ \forall u\in D_h,\\
&\|[A_h(t)-A_h(s)]u\|_{\tilde{E}_h}\le C_8\{|t-s|^{\delta_1}\|[A_h(t)+\mu_1 I]u\|_{\tilde{E}_h}+\|[A_h(t)+\mu_1 I]u\|_{\tilde{E}_h}^{\delta_2}\|u\|_{E_h}^{1-\delta_2}\},\\
&\forall h,\ \forall t,s\in[0,T],\ \forall u\in\tilde{D}_h,
\end{aligned}$$

for some fixed natural number $q_0\ge 1$ *and with some constants* $C_7,\ C_8,\ \mu_1\ge 0,\ \sigma_j\in[0,1],\ \kappa_j\in[0,1],\ j=1,2,\ldots,q_0,\ \delta_1\in(0,1],\ \delta_2\in[0,1)$, *independent of* h,t,s,u, *and the bound*

$$\begin{aligned}
&\|R(\lambda,A_h(t))u\|_{E_h}\le C_9|\lambda|^{\gamma_1-1}\|u\|_{\tilde{E}_h}, \qquad \forall h,\ \forall t\in[0,T],\ \forall u\in\tilde{E}_h,\\
&\forall\lambda:|\arg\lambda|\ge\varphi_0,\ |\lambda|\ge b_1
\end{aligned}$$

holds with constants $C_9,\ b_1$ *and* $\gamma_1\in[0,1),\ \gamma_1<1+\sigma_j-\kappa_j,\ j=1,2,\ldots,q_0$, *independent of* h,t,λ,u.

Then the difference scheme (6), (7) admits an equivalent representation in the canonical form (2) with some linear operators $\hat{A}_{\tau h}(t_k),\ Q_{\tau h}(t_k)$ *and inhomogeneity* $F_{\tau h}(t_k)$, *and*

$$Q_{\tau h}(t_k)=I-\lambda_0^{-1}\alpha(\tau A_h(t_k)),\ F_{\tau h}(t_k)=\sum_{r=1}^{\nu_2}\Omega_{\tau hr}(t_k)f_h(t_k+\bar{c}\tau)$$

(functions of operators are understood in the sense of the functional calculus of closed operators [15]*) and the operator* $\hat{A}_h(t_k)$ *is bounded both in* E_h *and* $\tilde{E}_h$; $Q_{\tau h}(t_k)$ *is bounded as an operator from* $\tilde{E}_h$ *to* E_h, *and the linear operators* $\Omega_{\tau hr}(t_k),\ r=1,2,\ldots,\nu_2,$ *are*

bounded in $\tilde{E}_h$. *If we take* $\mathcal{R}_{\tau h} = Q_{\tau h}(t_0)$, *then there exists a solution* $y(t_k) \in E_h$, $k = 0, 1, \dots, T/\tau$, *of the discrete problem (6), (7), satisfying the stability estimates*

$$\begin{aligned}
(13) \qquad & \|[\hat{A}_{\tau h}(t_k) + \bar{\mu} I]^{\xi} y(t_k)\|_{E_h} + \|[\tau^{-1}\alpha(\tau A_h(t_k)) + \bar{\mu} I]^{\xi} y(t_k)\|_{E_h} \\
& \le C_{10}(\xi)\Big\{[(k+1)\tau]^{-(\xi+\gamma_1)}\|y_{0h}\|_{\tilde{E}_h} \\
& + \tau \sum_{l=1}^{k} [(k-l+1)\tau]^{-(\xi+\gamma_1)} \max_{r=1,2,\dots,\nu_2} \|f_h(t_k + \bar{c}\tau)\|_{\tilde{E}_h}\Big\}, \\
& k = 0, 1, \dots, T/\tau, \ \forall h, \ \forall \tau \in (0, \tau_1], \ \forall \xi \in [0, 1-\gamma_1)
\end{aligned}$$

with some constants $C_{10}(\xi)$, $\bar{\mu} \ge 0$, $\tau_1 > 0$, *independent of* k, h, τ ($C_{10}(\xi)$ *may depend on* ξ). *If furthermore* $\lambda_0 = 1$ (λ_0 *is the range of the scheme), then for* $\mathcal{R}_{\tau h} = I$ *there exists an incomplete solution* $y(t_k), k = 1, 2, \dots, T/\tau$, *of the problem (6), (7) and for this the bound (13) is true. In the case where the operator* $A_h(t)$ *does not depend on* t, *then the statement holds with no upper bound on* ξ *(in this case we may formally assume* $q_0 = 1$, $\sigma_1 = 1$, $\kappa_1 = 0$).

The proof is based on transforming the scheme (6), (7) into the canonical form (2) and using the result of Theorem 1.

Applications of the theory developed here to the study of the well-posedness of difference schemes for parabolic partial differential equations will appear in a separate work.

References

1. O. A. Ladyzhenskaya, *The Boundary Value Problems of Mathematical Physics*, Nauka, 1973 (Russian); translation, Springer-Verlag.
2. A. N. Tikhonov and A. A. Samarskiĭ, Zh. Vychisl. Mat. i Mat. Fiz. **1** (1961), 5–63.
3. A. A. Samarskiĭ, Zh. Vychisl. Mat. i Mat. Fiz. **1** (1961), 441–460.
4. A. A. Samarskiĭ, Zh. Vychisl. Mat. i Mat. Fiz. **1** (1961), 972–1000.
5. A. A. Samarskiĭ, Zh. Vychisl. Mat. i Mat. Fiz. **2** (1962), 603–634.
6. A. A. Samarskiĭ and A. V. Gulin, *Stability of Difference Schemes*, Nauka, 1973.
7. A. A. Samarskiĭ, Different'sialnye Uravneniya **16** (1980), 1155–1171.
8. N. Yu. Bakaev, Dokl. Akad. Nauk SSSR **297** (1987), 275–279.
9. N. Yu. Bakaev, *Problems in Atomic Science and Engineering*, Methods and Programs for Numerical Solution of Problems in Mathematical Physics, issue 1, 1987, pp. 69–75.
10. N. Yu. Bakaev, Dokl. Akad. Nauk SSSR **305** (1989), 18–21.
11. N. Yu. Bakaev, Dokl. Akad. Nauk SSSR **309** (1989), 11–15.
12. N. Yu. Bakaev, *Partial Differential Equations*, Novosibirsk, 1989, pp. 3–14.
13. H. Stetter, *Analysis of Discretization Methods for Ordinary Differential Equations*, Springer-Verlag, 1973.
14. V. A. Vinokurov and N. V. Yuvchenko, Dokl. Akad. Nauk SSSR **284** (1985), 272–277.
15. *Functional Analysis* (S. G. Kreĭn, ed.), Nauka, 1973.

Translated by S. LARSSON