Современные вопросы вычислительной математики

А.Н.Коновалов

§ 1. Дискретные (сеточные) законы сохранения

В евклидовом пространстве R^3 зададим криволинейную систему координат y_i с локальным естественным базисом e_i (кобазисом -e^i) и фундаментальным тензором G, определяющим евклидову метрику. Пусть $D(M) = D(y_1, y_2, y_3)$ – односвязная область с достаточно гладкой границей S и вектором нормали $\boldsymbol{n} = n_m \boldsymbol{e}^m$. Если t – время, то $Q = [0 \leq t \leq T] \times D(M)$ - фазовый объем. Будем считать, что изучаемые физические процессы связаны с изменением параметров сплошной среды $u_j(M,t), (M,t) \in Q$. В качестве метода (способа) изучения изменений $u_j(M,t)$ выберем вычислительный эксперимент. Современная технология проведения вычислительного эксперимента предполагает реализацию следующих этапов. От изучаемого процесса I к непрерывной математической модели II ; далее к дискретной математической модели III, численному алгоритму IV, программе V; наконец, собственно к вычислениям на ЭВМ VI . Если имеется возможность сравнить результаты вычислений с результатами натурных экспериментов, то в нужных случаях происходит уточнение математической модели VI $\rightarrow$ II и т.д. При таком подходе реальному процессу изменения изучаемых параметров $u_j(M,t)$ соответствует контролируемый процесс смены дискретных состояний ЭВМ. Обратимся к этапу I $\rightarrow$ II. Коль скоро речь идет о сплошной среде, то непременными составляющими математической модели при описании какого-либо процесса, протекающего в фазовом пространстве Q, являются:

$$- \text{законы} \quad \text{сохранения}, \tag{1.1}$$
$$- \text{определяющие} \quad \text{соотношения}, \tag{1.2}$$
$$- \text{уравнения} \quad \text{состояния}. \tag{1.3}$$

Соотношения (1.1)–(1.3) можно задать в виде дифференциальных и конечных уравнений, к которым следует добавить начальные и краевые условия. Говоря о замкнутой математической модели II , обычно предполагают, что число определяемых параметров модели II совпадает с числом независимых соотношений (1.1)–(1.3). В идеальном варианте как замкнутая математическая модель II, так и дискретная модель III должны быть корректными по Адамару: решение существует, единственно и непрерывно зависит от входных данных. Это требование является необходимым при проведении вычислительного эксперимента на ЭВМ. Также играет существенную роль структура каждой компоненты вычислительного эксперимента при переходах II $\rightarrow$ III $\rightarrow$ IV $\rightarrow$ V. Опять-таки в идеальном случае:

–"малому изменению" математической модели должно соответствовать "малое изменение" алгоритма; (1.4)

–"малому изменению" алгоритма должно соответствовать "малое изменение" программы. (1.5)

Как представляется, именно выполнение требований (1.4)–(1.5) приводит к эффективности вычислительного эксперимента на классе задач, а не на какой-то конкретной задаче. Модульный анализ вычислительного алгоритма IV позволяет строить програмные реализиции V , удовлетворяющие требованию (1.5). Что же касается требования (1.4), то ему можно удовлетворить только на соответствующих друг другу классах структурированных моделей и алгоритмов.

Сказанного повидимому достаточно, чтобы оценить важность выявления структуры как непрерывной математической модели (1.1)–(1.3), так и соответствующей ей дискретной модели III. Именно в этой связи в дальнейшем будет рассмотрен достаточно широкий класс математических моделей механики сплошной среды, основанный на (1.1)–(1.3), который изначально обладает некоторой структурой. Пока же ограничимся описанием одного из возможных способов перехода от непрерывных законов сохранения к дискретным.

Обратимся к уравнению неразрывности (закон сохранения массы):

$$\int_{t_1}^{t_2} \int_V (\frac{\partial \rho}{\partial t} + div\, \rho v)\, dV\, dt = 0 \leftrightarrow \frac{\partial \rho}{\partial t} + div\, \rho v = 0 \tag{1.6}$$

Напоминаем, что здесь ρ – плотность, $\boldsymbol{v}$ – скорость, $V \subseteq D$ – произвольный пространственный “макрообъем”, $t_2 - t_1 \le T$. Тогда $\delta Q(M,t) = [t_1 \le t \le t_2] \times V(M)$ можно интерпретировать как произвольный фазовый “макрообъем”. Ради некоторых упрощений будем рассматривать декартову прямоугольную систему координат. В этом случае $x_0 = t, x_1, x_2, x_3$ можно принять за фазовые независимые переменные и ввести вектор

$$\boldsymbol{w}_1 = (w_{01}, w_{11}, w_{21}, w_{31}) = (\rho, \rho v_1, \rho v_2, \rho v_3).$$

По определению дивергенции вектора

$$div\, \boldsymbol{w}_1 = \frac{\partial}{\partial t}\rho + \frac{\partial}{\partial x_1}\rho v_1 + \frac{\partial}{\partial x_2}\rho v_2 + \frac{\partial}{\partial x_3}\rho v_3$$

и (1.6) перепишется следующим образом

$$\int_{\delta Q} div\, \boldsymbol{w}_1 dV\, dt \ = \ 0 \leftrightarrow div\, \boldsymbol{w}_1 \ = \ 0 \tag{1.7}$$

Если ∂V – граница пространственного макрообъема V, то $\omega = [t_1 \le t \le t_2] \times \partial V$ – граница фазового макрообъема δQ. Использование формулы Гаусса- Остроградского

$$\int_{\delta Q} div\, \boldsymbol{w}_1\, dV\, dt \ = \int_\omega \boldsymbol{w}_1 \cdot \boldsymbol{n}\, dw$$

вместе с (1.7) дает

$$\int_\omega \boldsymbol{w}_1 \cdot \boldsymbol{n}\, d\omega = \ 0 \leftrightarrow div \boldsymbol{w}_1 \ = 0 \tag{1.8}$$

В точности такие же рассуждения, связанные с использованием дифференциального закона сохранения в дивергентной форме, можно применить и для закона сохранения полной энергии. При отсутствии массовых сил и тепловых источников (стоков) получим

$$\int_\omega \boldsymbol{w}_2 \cdot \boldsymbol{n}\, d\omega \,=\, 0 \,\leftrightarrow\, div\, \boldsymbol{w}_2 \,=\, 0, \tag{1.9}$$

где

$$div\, \boldsymbol{w}_2 \,=\, \frac{\partial}{\partial t}\rho\varepsilon \,+\, div[\boldsymbol{v}(\rho\varepsilon - P) + \boldsymbol{q}],\, P = P^*.$$

Здесь $\boldsymbol{q}$ – плотность потока тепла, P – тензор истинных напряжений Коши, $\varepsilon \,=\, E + \frac{1}{2}|\boldsymbol{v}|^2$ – сумма внутренней и кинетической энергий. Ее иногда называют массовой плотностью полной энергии.

В дивергентной форме закона сохранения массы

$$\frac{\partial\rho}{\partial t} \,+\, div(\rho\boldsymbol{v}) \,=\, 0 \tag{1.10}$$

и полной энергии

$$\frac{\partial(\rho\varepsilon)}{\partial t} \,+\, div[\boldsymbol{v}(\rho\epsilon - P) + \boldsymbol{q}],\, P = P^* \tag{1.11}$$

операция div применяется к векторным полям. В дивергентной форме закона сохранения импульса

$$\frac{\partial(\rho\boldsymbol{v})}{\partial t} \,=\, div[P - \rho\boldsymbol{v}\otimes\boldsymbol{v}],\, P = P^* \tag{1.12}$$

эта же операция div применяется к тензорному полю ранга два. Для приведения (1.12) к виду (1.8), (1.9) следует в соответствии с определением дивергенции тензора ранга два перейти от (1.12) к покомпонентной записи (1.12)

$$\frac{\partial(\rho v_i)}{\partial t} \,=\, \frac{\partial}{\partial x_j}(p_{ji} - \rho v_i v_j),\, i = 1, 2, 3.$$

Далее аналогично предыдущему следует ввести векторы $\boldsymbol{w}_{2+i},\, i \,=\, 1, 2, 3$ с компонентами

$$\boldsymbol{w}_{2+i} = (\rho v_i,\, -p_{1i} + \rho v_1 v_i,\, -p_{2i} + \rho v_2 v_i,\, -p_{3i} + \rho v_3 v_i).$$

Тогда полная система законов сохранения (1.10) - (1.12) может быть записана следующим образом

$$\int_\omega \boldsymbol{w}_m \cdot \boldsymbol{n} d\omega \,=\, 0 \,\leftrightarrow\, div\boldsymbol{w}_m = 0,\, P = P^*, m = 1, 2, 3, 4, 5. \tag{1.13}$$

Итак,будем считать, что непрерывные интегральные и дифференциальные модели механики сплошной среды M_1 и M_2 основаны на формулировке (1.13). Следует отметить, что поскольку модели M_1 и M_2 пред'являют различные требования к гладкости изучаемых параметров, то M_1 и M_2, вообще говоря, не эквивалентны. Поэтому обычно предполагают, что существует непрерывно дифференцируемое тензорное поле истинных напряжений $P \,=\, P^*$ и векторное поле плотности потока тепла q, на которых интегральная M_1 и дифференциальная M_2 модели из (1.13) эквивалентны.

Дискретная (сеточная) модель M_3 требуется, как правило, для численной реализации непрерывных моделей M_1, M_2. Для формальных переходов $M_1 \,\rightarrow\, M_3$ или $M_2 \,\rightarrow\, M_3$ следует перейти от области изменения непрерывного аргумента $M \,\in$

$Q, M = (x_0 = t,\, x_1, x_2, x_3) = (x_\alpha)$ к конечному множеству точек $M_h \in Q_h$, т.е. построить сетку Q_h. Искомыми параметрами дискретной модели M_3 будут $u_{jh}(M_h) \neq u_j(M_h)$. Пусть, например, фазовый об'ем Q является декартовым произведением отрезков $[a_\alpha \leq x_\alpha \leq b_\alpha]$, $\alpha = 0, 1, 2, 3$. Дискретное множество Q_h (сетку) можно определить следующим образом: $M_h \in Q_h$, если $M_h = (a_\alpha + i_\alpha h_\alpha)$, где $b_\alpha - \alpha_\alpha = N_\alpha h_\alpha$ и $0 \leq i_\alpha \leq N_\alpha$. Сетки подобного рода называются регулярными. Аналогичным способом вводится регулярная сетка Q_h и в том случае, когда для описания изучаемого процесса выбрана криволинейная система координат y_1, y_2, y_3, а исходная пространственная область D является "криволинейным параллелепипедом на гранях которого $y_\beta = const., \beta = 1, 2, 3$. Тогда в каждой точке $M_h \in Q_h$ определены также и стационарный естественный базис e_i, кобазис e^j, фундаментальный тензор G.

В целях максимального упрощения изложения в дальнейшем, как правило, используются регулярные сетки. Тем не менее следует хотя бы в самых общих чертах обозначить те проблемы, которые возникают при построении "хороших"сеток Q_h в реальных задачах.

Пусть $Q_h = [0 \leq n\tau \leq T] \times D_h$. Здесь и далее положено $a_0 = 0$, $b_0 = T$, $i_0 = n$, $h_0 = \tau$ и $n\tau = t$. Тогда основные проблемы будут связаны с построением пространственной сетки D_h. Однако независимо от конкретного способа построения D_h, переход $Q \to Q_h$ привносит в дискретную модель M_3 не только малые параметры τ, h_1, h_2, h_3, но и параметры τ/h_i, h_i/h_j, $i \neq j$. Поэтому $M_3 = M_3(\tau,\, h_i,\, \tau/h_i,\, h_i/h_j)$. Следовательно, одна и та же сеточная модель M_3 при различных законах предельного перехода

$$\tau \to 0,\; h_i = h_i(\tau) \to 0 \tag{1.14}$$

может, вообще говоря, приводить к различным непрерывным моделям $\hat{M}_1 \leftrightarrow \hat{M}_2$, которые существенно отличаются от исходных $M_1 \leftrightarrow M_2$. Достаточно содержательные примеры подобного рода будут приведены ниже.

Другая проблема, которая нас ожидает, связана с "произвольностью"исходной области D изменения пространственных переменных x_1, x_2, x_3 или y_1, y_2, y_3. Здесь возможны два случая: граничные узлы S_h сетки D_h таковы, что $S_h S$, либо $S_h \subset S$. В первом случае говорят о несогласованной с S сеткой D_h, во втором - о согласованной. В обоих случаях сеточная модель M_3 содержит плохо контролируемые источники погрешности. Действительно, пусть искомые параметры $u_j(x_\alpha, t)$ непрерывной модели удовлетворяют некоторым краевым условиям, например, $u_j(x_\alpha, t) = g_j$, $(x_\alpha) \in S$. Если $S_h S$, то источником плохо контролируемой погрешности может являться (и действительно является!) краевое условие сеточной модели $M_3 : u_{jh}(x_{\alpha h}, n\tau) = g_j$ при $x_{\alpha h} \in S_h$. Если же $S_h \subset S$, то это неминуемо приводит к неравномерной по каждой пространственной переменной x_α сетке D_h. Иными словами, если $(x_{\alpha h})$ и $(x_{\alpha h} \pm h_\alpha)$ "соседние"узлы сетки D_h, то $h_{i\alpha} = h_i(x_{\alpha h})$. Но тогда при переходе от M_3 к непрерывным моделям $\hat{M}_1 \leftrightarrow \hat{M}_2$ наряду с (1.14) следует также учитывать зависимость сеточной модели M_3 от параметров $min\, h_{i\alpha}/max\, h_{i\alpha}$. Проблема построения "хорошей"сетки Q_h может быть связана также и с тем, что само решение исходной непрерывной задачи $u_j(x_\alpha, t)$ имеет особенности в некоторых подобластях $D_m \subset D$. Обычно эти особенности связаны с "относительно большими"градиентами искомых параметров $u_j(x_\alpha, t), (x_\alpha) \in D_m$. Интуитивно понятно, что в D_{mh} следует использовать "относительно малые"параметры дискретизации $h_{i\alpha}$. Это приводит к нерегулярной сетке D_h, для которой ,как уже говорилось,$h_{i\alpha} = h_i(x_{\alpha h})$. Сами подобласти D_m больших градиентов могут зависеть от t и тогда $h_{i\alpha} = h_i(x_{\alpha h}, n\tau)$. Сетки подобного рода называются адаптивными.

Пусть однако "хорошая"сетка Q_h построена. Дальнейшее при переходе к M_3 заключается в следующем. Функции непрерывного аргумента - искомые параметры $u_j(x_\alpha, t), (x_\alpha, t) \in Q$ непрерывной модели заменяются функциями дискретного аргумента $u_{jh} = u_{jh}(x_\alpha, n\tau)$, $(x_\alpha, n\tau) \in Q_h$. Затем от соотношений (1.1)-(1.3) непрерывной модели следует перейти к конечным соотношениям, которые связывают параметры u_{jh} в узлах сетки Q_h. Этот этап является одним из основных при проведении вычислительного эксперимента и заслуживает подробного описания. В качестве основы для построения M_3 выберем интегральную модель M_1 в форме (1.13).Эта модель имеет дело с произвольным фазовым макрооб'емом δQ и ω -граница этого макрооб'ема.Произвольность ω в (1.13) предполагает, что существует $\cup\delta Q$, совпадающее с рассматриваемым фазовым пространством Q. При переходе $M_1 \to M_3$ достаточно в качестве δQ выбрать элементарный сеточный фазовый "микрооб'ем"δQ_h, соответствующий узлу сетки $(x_{\alpha h}, n\tau)$. Таковым, например, может быть $\delta Q_h = \tau_n h_1 h_2 h_3$, где $h_i = h_i(x_{\alpha h})$. Тогда в (1.13) следует положить $\omega = \omega_h$, где ω_h – граница δQ_h. Следует также предположить существование $\cup\delta Q_h$ такого, что

$$\lim_{h\to 0} mes(\cup\delta Q_h - Q) = 0 \tag{1.15}$$

Сразу же и отметим, что именно с выполнением условия (1.15) могут быть связаны основные трудности при построении "хороших"сеток D_h для "произвольной"области D.

Итак, для сеточной модели M_3 мы предполагаем, что

$$\int_{\omega_h} \boldsymbol{w}_m \cdot \boldsymbol{n}\, d\omega \,=\, 0. \tag{1.16}$$

Дальнейшее заключается в построении для (1.16) некоторой кубатурной формулы

$$\sum_\beta c_{m,\beta}\, \boldsymbol{w}_{m,\beta} = 0 \tag{1.17}$$

такой, что

$$\lim_{h\to 0}[\int_{\omega_h} \boldsymbol{w}_m \cdot \boldsymbol{n}\, d\omega \,-\, \sum_\beta c_{m,\beta}\, \boldsymbol{w}_{m,\beta}] = 0 \tag{1.18}$$

Здесь и далее мультииндекс β задает множество узлов сетки Q_h, принадлежащих ω_h, на котором определены компоненты вектора $\boldsymbol{w}_m$. Поэтому в (1.17),(1.18) вместо $\boldsymbol{w}_{m,\beta}$ можно использовать также и стандартное обозначение $\boldsymbol{w}_{m,h}$.

Пусть определено соответствие

$$\int_{\omega_h} \boldsymbol{w}_m \cdot \boldsymbol{n}\, d\omega \,\leftrightarrow\, \sum_\beta c_{m,\beta} \boldsymbol{w}_{m,h},$$

для которого выполнено условие аппроксимации (1.18). Поскольку для достаточно гладкого класса допустимых векторов $H(\boldsymbol{w}_m)$

$$\int_{\omega_h} \boldsymbol{w}_m \cdot \boldsymbol{n}\, d\omega \;=\; \int_{\delta Q_h} div\, \boldsymbol{w}_m\, dV dt,$$

то на $H(\boldsymbol{w}_m)$ определено также и некоторое соответствие

$$div\, \boldsymbol{w}_m \;\leftrightarrow\; div_h \boldsymbol{w}_{m,h},$$

для которого в силу (1.18) будем иметь

$$\lim_{h\to 0} ||div\, \boldsymbol{w}_m - div_h \boldsymbol{w}_{m,h}|| = 0.$$

Как нетрудно понять, тем самым определен и сеточный оператор div_h такой, что при соответствующем законе предельного перехода (1.14) на достаточно гладком классе $H(\boldsymbol{w}_m)$

$$\lim_{h\to 0} ||(div - div_h)\boldsymbol{w}_m|| = 0 \tag{1.19}$$

Хотя предыдущие рассуждения и не являются достаточно строгими, тем не менее можно подвести некоторые итоги. Формулировка законов сохранения в интегральной форме (1.13), переход в (1.13) от ω к ω_h и (1.17) определяют способ построения сеточного оператора $div_{h,m}$. Формулировка законов сохранения в дифференциальной форме (1.13) использует векторы $\boldsymbol{w}_m$ с тензорными компонентами различных рангов. Поэтому в (1.17) определен сеточный оператор div_h для конкретного m, т.е. $div_{h,m}$. Наконец, запись условия аппроксимации в виде (1.19) предполагает эквивалентность интегральной и дифференциальной моделей (1.13). Конкретизация $||\cdot||$ в (1.19) неизбежно связана с конкретизацией функционального пространства $H(\boldsymbol{w}_m)$.

Для конкретизации $H(\boldsymbol{w}_m)$ следует, по существу, дать определение решения $\boldsymbol{w}_m$. Поясним только что сказанное на простейшем, но достаточно характерном примере. Пусть $u(x,t)$ – гладкое решение уравнения

$$div\boldsymbol{w} \equiv \frac{\partial}{\partial x}u + \frac{\partial}{\partial x}au = 0, \quad a = const. > 0 \tag{1.20}$$

Выберем <u>любую гладкую</u> финитную функцию $f(x,t)$. Тогда

$$\int_{\delta Q} f(\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x})dxdt \;=\; 0, \quad \delta Q \subset Q$$

и после интегрирования по частям получим

$$\int_{\delta Q} u(\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x})\, dxdt \;= 0, \quad \delta Q \subset Q. \tag{1.21}$$

Верно и обратное. Если $u(x,t)$ – гладкая и для любой финитной и гладкой $f(x,t)$ выполнено (1.21),то справедливо и (1.21). Поскольку в (1.21) не нужно требовать гладкости $u(x,t)$, то (1.21) можно принять за определение <u>обобщенного решения</u> уравнения (1.20). Эквивалентное, но более конструктивное для наших дальнейших целей определение обобщенного решения уравнения (1.20) можно связать с равенством

$$\oint_\omega udx - audt = o, \tag{1.22}$$

потребовав его выполнения для любого замкнутого контура $\omega \subset Q$. Поэтому переход в (1.13) от дифференциальной модели M_2 к интегральной модели M_1 связан, вообще говоря, с переходом от рассмотрения классического решения системы дифференциальных уравнений $divw_m = 0$ к обобщенному : $\oint w_m \cdot nd\omega = 0$. Понятно, что гладкое обобщенное решение совпадает с классическим.

Для дифференциального уравнения (1.20) рассмотрим задачу Коши:

$$\frac{\partial}{\partial t}u + \frac{\partial}{\partial x}au = 0,\, |x| < \infty,\, t > 0, \quad u(x,0) = \varphi(x). \tag{1.23}$$

Пусть

$$u(x,t) = \varphi(x - at) = \varphi(\xi) \tag{1.24}$$

Если $\varphi(\xi)$ непрерывно дифференцируема по ξ,то $u(x,t)$ в (1.24) задает классическое решение задачи (1.23). В противном случае в (1.24) задано обобщенное решение задачи (1.23). Для интегрируемых $\varphi(\xi)$ таковым, например, является функция $u(x,t) = sign\xi$. Поэтому естественное желание расширить класс возможных начальных данных $\varphi(x)$ приводит к иной формулировке задачи Коши (1.23):

$$\forall\, \omega \subset Q : \oint_\omega udx - audt = 0, \quad u(x,0) = \varphi(x). \tag{1.25}$$

Гладкость классического решения задачи Коши (1.23) определяется гладкостью начальных данных $u(x,0) = \varphi(x)$. Пусть в (1.25) $\varphi(x)$ не является непрерывно дифференцируемой. Пусть также существует непрерывно дифференцируемая $\varphi_\delta(x)$ такая, что

$$\lim_{\delta \to 0} \varphi_\delta(x) \;=\; \varphi(x).$$

Классическое решение задачи (1.23), соответствующее начальным данным $\varphi_\delta(x)$ обозначим через $u_\delta(x,t)$. Тогда для обобщенного решения $u(x,t)$ задачи (1.25) имеет место

$$\lim_{\delta \to 0} u_\delta(x,t) = u(x,t). \tag{1.26}$$

Поэтому переход в (1.25) от "негладких"$\varphi(x)$ к "сглаженным"$\varphi_\delta(x)$ позволяет наряду с (1.25) иметь дело с эквивалентной классической задачей Коши (1.23), решение которой обладает "нужной по ходу изложения гладкостью".

Наряду с (1.23) будем рассматривать также смешанную задачу Коши (начально-краевую задачу). Если $Q = [0 \le t \le T] \times [0 \le x \le l]$, то речь идет о задаче

$$\begin{gathered} Lu \equiv (\frac{\partial}{\partial t} + a\frac{\partial}{\partial x})u = 0,\, a = const. > 0, \\ u(x,0) = \varphi(x),\, u(0,t) = \mu(t),\, \varphi(0) = \mu(0) \end{gathered} \tag{1.27}$$

Если в (1.27) $a = const. < 0$, то краевое условие нужно задавать при $x = l$. Тогда в (1.27) следует положить $u(x,0) = \varphi(x)$, $u(l,t) = \nu(t)$, $\varphi(l) = \nu(0)$. Если $u(x,t)$ – решение задачи (1.27), то

$$u(x,t)=\begin{cases}\varphi(x-at),\, x\geq at\\ \mu(t-\dfrac{x}{a}),\, x<at\end{cases} \tag{1.28}$$

Теперь имеется все необходимое, чтобы для непрерывной задачи (1.27) построить сеточную модель M_3. Выберем регулярную сетку $Q_h=[0\leq nt\leq T]\times[0\leq ih\leq l]$, так что узел сетки имеет координаты $(ih,n\tau)$, а индексы i,n пробегают всю последовательность значений, обеспечивающих принадлежность узла $(ih,n\tau)$ сеточной области Q_h. В качестве "элементарного фазового об'ема"δQ_h выберем такой прямоугольник $ABCD$, для которого

$$A=((i-1)h,n\tau),\quad B=(ih,n\tau),\quad C=(ih,(n+1)\tau),\quad D=((i-1)h,(n+1)\tau).$$

Для достаточно гладкого решения дифференциального уравнения из (1.27) имеем

$$div\boldsymbol{w}\equiv\frac{\partial}{\partial t}u-\frac{\partial}{\partial x}au=0\,\leftrightarrow\,\oint_\omega\boldsymbol{w}\cdot\boldsymbol{n}\,d\omega=0\rightarrow\oint_\omega udx-audt=0 \tag{1.29}$$

Напоминаем, что в (1.13),(1.29) n-вектор внешней по отношению к δQ нормали, поэтому обход контура ω в (1.29) совершается против часовой стрелки. В качестве элементарного сеточного контура ω_h выберем замкнутый контур $ABCDA$ с тем же направлением обхода. Тогда из (1.29) получим

$$\int_A^B udx-\int_D^C udx=a(\int_B^C udt-\int_A^D udt) \tag{1.30}$$

Дальнейшее будет связано с использованием в (1.30) конкретных кубатурных формул.

Положим

$$\begin{aligned}&\int_A^B udx=hu(B)=hu_i^n,\quad \int_D^C udx=hu(C)=hu_i^{n+1},\\ &\int_B^C udt=\tau u(B)=\tau u_i^n,\quad \int_A^D udt=\tau u(A)=\tau u_{i-1}^n\end{aligned} \tag{1.31}$$

Вместе с (1.30) это дает

$$\frac{u_i^{n+1}-u_i^n}{\tau}+a\frac{u_i^n-u_{i-1}^n}{h}=0,\quad u_i^{n+1}=ru_{i-1}^n+(1-r)u_i^n,\quad r=\frac{a\tau}{h}>0, \tag{1.32}$$

Если же в (1.31) положить

$$\int_B^C udt=\tau u(C)=\tau u_i^{n+1},\quad \int_A^D udt=\tau u(D)=\tau u_{i-1}^{n+1},$$

то вместо (1.32) будем иметь

$$\frac{u_i^{n+1}-u_i^n}{\tau}+a\frac{u_i^{n+1}-u_{i-1}^{n+1}}{h}=0,\quad u_i^{n+1}=\frac{r}{1+r}u_{i-1}^{n+1}+\frac{1}{1+r}u_i i^n. \tag{1.33}$$

И, наконец,если в (1.31)

$$\int_A^B udx = hu(A) = hu_{i-1}^n, \quad \int_D^C udx = hu(D) = hu_{i-1}^{n+1},$$

то (1.30) перепишется следующим образом

$$\frac{u_{i-1}^{n+1} - u_{i-1}^n}{\tau} + a\frac{u_i^n - u_{i-1}^n}{h} = 0$$

или , после смены индекса $(i-1) \to i$,

$$\frac{u_i^{n+1} - u_i^n}{\tau} + au_{i+1}^n - u_i^n h = 0, u_i^{n+1} = -ru_{i+1}^n + (1+r)u_i^n. \tag{1.34}$$

Далее во избежание разночтений мы сохраним обозначение $u_i^n = u(ih, n\tau)$ для функции непрерывного аргумента $u(x,t)$ в узле сетки $(ih, n\tau)$ и будем использовать обозначение y_i^n для функции дискретного аргумента $y(ih, n\tau)$ в том же узле.

Непрерывной задаче (1.27) поставим в соответствие сеточную (дискретную) задачу – разностную схему:

$$\begin{aligned} &(L_h y)_i^n \equiv L_h y_i^n \equiv (\frac{T_{+0} - E}{\tau} + a\frac{E - T_{-1}}{h})y_i^n = 0, \\ &n = 0, 1, ..., N_0;\ i = 0, 1, ..., N_1;\ N_0\tau = T,\ N_1 h = l; \\ &y_i^0 = \varphi(ih),\ y_0^n = \mu(n\tau),\ r = a\tau/h = const. \text{ при } \tau \to 0, h \to 0. \end{aligned} \tag{1.35}$$

В (1.35) и далее $T_{\pm 0}, T_{\pm 1}$-операторы сдвига,E-тождественный оператор , так что

$$T_{\pm 0} f(x,t) = f(x, t \pm \tau), \quad T_{\pm 1} f(x,t) = f(x \pm h, t), \quad Ef(x,t) = f(x,t) \tag{1.36}$$

Введенные в (1.36) операторы сдвига определены как для функций непрерывного аргумента $(x,t) \in Q$, так и для функций дискретного аргумента $(ih, n\tau) \in Q_h$. Для нерегулярной сетки Q_h сдвиги аргументов в (1.36) могут зависеть от координат $(x = ih, t = n\tau)$ конкретного узла (x,t) сетки Q_h.

Искомыми величинами в (1.35) являются y_i^n. Нетрудно подсчитать, что число линейных уравнений в (1.35) для определения неизвестных y_i^n равно числу неизвестных. Разностное уравнение из (1.35) перепишем в таком виде (сравни с (1.32):

$$y_i^{n+1} = ry_{i-1}^n + (1-r)y_i^n, \quad (ih, n\tau) \in Q_h \tag{1.37}$$

Своеобразие системы линейных алгебраических уравнений (1.37) заключается в том, что в силу начальных данных и краевых условий разностной схемы (1.35) возможно последовательное определение неизвестных $y_1^{n+1}, ..., y_{N_1}^{n+1}$ сначала для $n = 0$, затем для $n = 1$ и т.д.Этот вывод остается в силе, если в (1.35) использовать операторы L_h, соответствующие (1.33) или (1.34). Тогда

$$\begin{aligned} &L_h y_i^n \equiv (\frac{T_{+0} - E}{\tau} + aT_{+0}\frac{E - T_{-1}}{h})y_i^n = 0, \\ &y_i^{n+1} = \frac{r}{1+r}y_{i-1}^{n+1} + \frac{1}{1+r}y_i^n,\ (x,t) \in Q_h, \end{aligned} \tag{1.38}$$

если L_h соответствует (1.33) и

$$
\begin{gathered}
L_h y_i^n \equiv (\frac{T_{+0}-E}{\tau} + a\frac{T_{+1}-E}{h})y_i^n = 0, \\
y_i^{n+1} = -r y_{i+1}^n + (1+r)y_i^n,\ (x,t) \in Q_h,
\end{gathered}
\tag{1.39}
$$

когда L_h соответствует (1.34). Если говорить о (1.38),(1.39) как о разностных схемах для непрерывной задачи (1.27), то следует дополнительно указать закон предельного перехода $h = h(\tau)$ при $\tau \to 0$, $h \to 0$, а также задать предельные условия y_i^0, y_0^n. Будем считать, в (1.38),(1.39) закон предельного перехода, а также y_i^0, y_0^n те же , что и в (1.35).

Сеточное уравнение (1.37) для каждого $0 < i < N_1$ содержит только одно неизвестное: y_i^{n+1}. Такого рода уравнения принято называть явными. Этот же термин обычно используется и для разностных схем, в основе которых лежат сеточные уравнения подобного типа. Разностные схемы (1.35),(1.39) являются явными в только что указанном смысле. Число искомых величин y_{i-1}^{n+1}, y_i^{n+1} в сеточном уравнении разностной схемы (1.38) больше единицы. Разностные схемы, построенные на сеточных уравнениях такого типа принято называть неявными.

Пусть y^n- вектор, компонентами которого являются величины y_i^n. Тогда рассматриваемые двухслойные разностные схемы (1.35),(1.38),(1.39) можно записать в операторном виде

$$
\hat{B}y^{n+1} = \hat{C}y^n, \quad y^0 = \varphi. \tag{1.40}
$$

Здесь $\varphi = (\varphi_0, \varphi_1, ..., \varphi_{N_1})$, $\varphi_i = \varphi(ih)$. К областям определения $R(\hat{B}), R(\hat{C})$ сеточных операторов $\hat{B}$ и $\hat{C}$ следует отнести те и только те y^n, у которых $y_0^n = \mu(n\tau)$. Для явных разностных схем (1.35),(1.39) $\hat{B} = E$, для неявной разностной схемы (1.38)

$$
(E - \frac{r}{1+r}T_{-1})y^{n+1} = \frac{1}{1+r}y^n, \quad y^o = \varphi \tag{1.41}
$$

и $\hat{B} \neq E$. Для двухслойных разностных схем (1.40)

$$
y^{n+1} = \hat{B}^{-1}\hat{C}y^n = Cy^n, \quad y^0 = \varphi \tag{1.42}
$$

Определенный в (1.42) оператор C обычно называют оператором перехода разностной схемы (1.40).

При использовании неявной разностной схемы (1.38) переход $y^n \to y^{n+1}$ формально связан с вычислением обратного оператора $\hat{B}^{-1}$. Однако в данном конкретном случае существует такая вычислительная реализация перехода $y^n \to y^{n+1}$, для которой ,как и в явных разностных схемах, обращение $\hat{B}$ сводится к тривиальному обращению оператора E. Действительно, будем вычислять компоненты y_i^{n+1} вектора y^{n+1} в порядке возрастания индекса $i : 1 \to 2 \to ... \to N_1$. Учтем также краевое условие $y_0^{n+1} = \mu((n+1)\tau)$. Тогда для $i = 1$ имеем

$$
y_1^{n+1} = \frac{r}{1+r}\mu^{n+1} + \frac{1}{1+r}y_1^n.
$$

Поэтому при вычислении y_2^{n+1} можно считать заданными величины y_1^{n+1}, y_2^n. Затем переходим к $i = 3$ и т.д. Именно такой алгоритм бегущего счета позволяет переписать операторное уравнение (1.41) следующим образом (сравни с (1.38)):

$$y^{n+1} = \frac{r}{1+r}T_{-1}y^{n+1} + \frac{1}{1+r}y^n \tag{1.43}$$

и считать правую часть в (1.43) известной.

Для функций (вектор-функций) $y(x,t)$ дискретного аргумента $(x,t) \in Q_h$ будем в дальнейшем использовать следующие обозначения

$$\begin{aligned} &y_t^n = \frac{T_{+0}-E}{\tau}y = \Lambda_0^+ y, \quad y_{\bar t}^n = \frac{E-T_{-0}}{\tau}y = \Lambda_0^- y, \quad y_{\bar t}^{n+1} = y_t^n, \\ &y_{x,i} = \frac{T_{+1}-E}{h}y = \Lambda_1^+ y, \quad y_{\bar x,i} + \frac{E-T_{-1}}{h}y = \Lambda_1^-, \quad y_{\bar x,i+1} = y_{x,i} \end{aligned} \tag{1.44}$$

Для упрощения записи в (1.44) и далее, как правило, явно указаны те индексы, отсутствие которых может привести к недоразумениям. Поэтому, например, $y = y_i^n$ и

$$y_t + ay_x = (y_i^n)_t + a(y_i^n)_x = y_t + ay_{x,i} = y_t + ay_{\bar x,i+1}.$$

В обозначениях (1.44) рассматриваемые разностные схемы (1.35), (1.38),(1.39) для смешанной задачи Коши (1.27) можно записать следующим образом:

$$y_t + ay_{\bar x} \equiv y_t + a\Lambda_1^- y = 0, \quad y^0 = \varphi(x), \quad (x,t) \in Q_h, \tag{1.45}$$

$$(E + \tau a\Lambda_1^-)y_t + a\Lambda_1^- = 0, \quad y^0 = \varphi(x), \quad (x,t) \in Q_h, \tag{1.46}$$

$$y_t + ay_x \equiv y_t + a\Lambda_1^+ y = 0, \quad y^0 = \varphi(x), \quad (x,t) \in Q_h. \tag{1.47}$$

В (1.45)-(1.47) использована возможность представления двухслойной разностной схемы (1.40) в ином <u>каноническом</u> виде

$$\hat B y^{n+1} - \hat C y^n = \hat B y^{n+1} - \hat B y^n + (\hat B - \hat C)y^n = \tau \hat B y_t^n + Ay^n = By_t + Ay = 0 \tag{1.48}$$

Выпишем также и цепочку преобразований при переходе от (1.38) к (1.46) и далее к (1.48)

$$\begin{aligned} &y_t + ay_{\bar x}^{n+1} = y_t + av^{n+1} = y_t + a(v^n + \tau\frac{v^{n+1}-v^n}{\tau}) = \\ &= y_t + ay_{\bar x} + a\tau\Lambda_1^- y_t = (E + a\tau\Lambda_1^-)y_t + ay_{\bar x} = By_t + Ay. \end{aligned}$$

Отметим, что в каноническом представлении (1.48) неявной разностной схемы (1.38) $B \neq E$. Кроме того еще раз напомним, что краевое условие смешанной задачи Коши (1.26) учитывается в разностных схемах (1.45)–(1.47) при задании областей определения $R(\Lambda_0^\pm), R(\Lambda_1^\pm)$ сеточных операторов $\Lambda_0^\pm, \Lambda_1^\pm$. Не вдаваясь пока в детали, все же отметим, что с точностью до конкретизации $R(\Lambda_0^\pm), R(\Lambda_1^\pm)$ разностные схемы (1.45)–(1.47) можно рассматривать также и в качестве сеточных аналогов непрерывной задачи Коши(1.23).

§ 2.Аппроксимация. Устойчивость.

$$Lu \equiv (\frac{\partial}{\partial t} + a\frac{\partial}{\partial x})u \equiv (D_0 + aD_1)u = 0, \quad u(x,0) = \varphi(x), (x,t) \in Q \tag{2.1}$$

разностную схему

$$L_h y = 0, \quad y^0 = \varphi(x), \quad (x,t) \in Q_h \tag{2.2}$$

мы надеемся, что если параметры дискретизации τ и h "достаточно малы то y из (2.2) дает "достаточно хорошее"приближение для u из (2.1). Если это так, то мы также надеемся и на то, что чем меньше τ и h , тем "более хорошим"становится это приближение. А каковы, собственно говоря, основания для подобного рода надежд ?

Пусть Q_1 - прямоугольный треугольник, сторонами которого являются отрезки прямых $x = at, t = 0, x = l$. Для $(x,t) \in Q_1$ имеем $x - at \leq 0$ и в соответствии с (1.28) решение задачи (2.1) дается формулой

$$u(x,t) = \varphi(x - at), \quad (x,t) \in Q_1 \tag{2.3}$$

Зададимся параметрами дискретизации τ, h и через Q_{1h} обозначим множество узлов сетки Q_h, для которых $ih \geq n\tau$. Зададим также закон предельного перехода:

$$\frac{\tau}{h} = const. \quad \text{при} \quad \tau \to 0, \quad h \to 0 \tag{2.4}$$

Зафиксируем узел M_* сетки Q_{1h} : $M_* = (x_*, t_*) = (ih, n\tau)$. Предельный переход $\tau \to 0, h \to 0$ определяет некоторую последовательность сеток $Q_{h,m}$. В силу (2.4) можно считать, что $\forall m, M_* \in Q_{1h,m}$. Характеристика

$$x - x_* = a(t - t_*) \tag{2.5}$$

уравнения $Lu = 0$, проходящая через узел сетки M_*, пересекается с прямой $t = 0$ в точке $\hat{M} = (\hat{x}, 0)$, где $\hat{x} = ih - an\tau = (i - rn)h, r = a\tau/h$. В этих обозначениях (2.3) можно переписать следующим образом

$$u(M_*) = \varphi(\hat{M}). \tag{2.6}$$

Особо отметим, что при $r = 1$ точка $\hat{M} \in Q_{1h}$. При $r = 1$ узлами сетки Q_{1h} являются также и точки с координатами $x = (i-k)h$, $t = (n-k)\tau; k = 1, ..., n$. Эти узлы лежат на характеристике (2.5) и, следовательно, если $u(x,t)$- решение задачи (2.1), то при $r = 1$ (2.3) допускает и такую запись

$$u(ih, n\tau) = u_i^n = u_{i-1}^{n-1} = u((i-1)h, (n-1)\tau). \tag{2.7}$$

Для численного решения исходной непрерывной задачи (2.1) выберем разностную схему (1.39). Тогда

$$y(M_*) = y_i^n = -r y_{i+1}^{n-1} + (1+r) y_i^{n-1}, \quad y_i^0 = \varphi_i. \tag{2.8}$$

Поэтому

$$y(M_*) = \sum_{m=0}^{n} \alpha_m(r) \varphi_{i+m}. \tag{2.9}$$

С другой стороны, пусть $r = k/n$, где $k > 0$- целое и $k \leq n$. По определению параметров дискретизации: $\tau = T/N_0$, $h = l/N_1$ и в силу (2.4) такой выбор r является допустимым. Тогда $\hat{x} = (i-k)h$, точка $\hat{M} \in Q_{1h}$, а (2.6) дает

$$u(M_*) = \varphi_{i-k} \tag{2.10}$$

Из (2.9),(2.10) заключаем, что изменение начальных данных в узле $\hat{M}$ приводит к изменению решения задачи (2.1) в узле M_* и не приводит к изменению решения разностной схемы (1.39) в этом же узле. Следовательно, с помощью разностной схемы (1.39) нельзя получить как "хорошее так и "более хорошее"приближение для u_i^n из (2.1). Вывод: разностная схема (1.39) непригодна для численного решения смешанной задачи Коши (2.1). Сразу же заметим, что этот вывод относительно разностной схемы (1.39) остается в силе, если за исходную задачу (2.1) принять задачу Коши (1.23).

Далее в качестве (2.2) рассмотрим разностную схему (1.35). В этом случае

$$y_i^n = ry_{i-1}^{n-1} + (1-r)y_i^{n-1}, \quad y_i^0 = \varphi_i \tag{2.11}$$

и тогда

$$y(M_*) = \sum_{m=0}^{n} \beta_m(r)\varphi_{i-m} \tag{2.12}$$

Пусть $r > 1, i > rn$, следовательно, $\hat{x} = (i - rn)h < (i - n)h$. Поэтому из (2.12) вытекает, что изменение начальных данных в точке $\hat{M}$ не приводит к изменению решения разностной схемы (2.11) в узле M_*. Этот факт находится в противоречии с (2.6). Вывод: при $r > 1$ разностная схема (1.35) непригодна для численного решения смешанной задачи Коши (2.1).

Пусть теперь в (2.11) $r = 1$. Вместе с (2.7) это дает

$$u_i^n = u_{i-1}^{n-1},\, u_i^0 = \varphi_i; \quad y_i^n = y_{i-1}^{n-1},\, y_i^0 = \varphi_i. \tag{2.13}$$

Если $(x,t) \in Q_{1h}$, то в силу (2.13) нет никаких оснований для того , чтобы различать u_i^n из (2.1) и y_i^n из (1.35). Практически также очевидно, что этот вывод остается справедливым и для $(x,t) \in Q_h$. Для об'яснения причины совпадения u_i^n и y_i^n на Q_h при $r = 1$ нам потребуется ряд новых понятий, одним из которых является аппроксимация.

Пусть $u(x)$ -скалярная функция векторного аргумента $(x_\alpha) \equiv (x) \in Q$, аналитическая по переменной x_α. Для введенных в (1.36) операторов сдвига $T_{\pm\alpha}$ имеем

$$T_{\pm\alpha}u(x_\alpha) = u(x_\alpha \pm h_\alpha) = u(x_\alpha) \pm h_\alpha u'(x_\alpha) + \frac{h_\alpha^2}{2!}u''(x_\alpha) \pm ... =$$
$$= [E \pm h_\alpha D_\alpha + \frac{h_\alpha^2}{2!}D_\alpha^2 \pm \frac{h_\alpha^3}{3!}D_\alpha^3 + ...]u(x_\alpha) = u(x_\alpha)exp(\pm h_\alpha D_\alpha).$$

Поэтому

$$T_{\pm\alpha} = e^{\pm h_\alpha D_\alpha}, \quad D_\alpha = \frac{\partial}{\partial x_\alpha} \tag{2.14}$$

Формула (2.14) устанавливает связь между оператором сдвига $T_{\pm\alpha}$ и оператором дифференцирования D_α на классе аналитических функций. На классе $H(u)$ функций, имеющих лишь p_α непрерывных производных по x_α, формула (2.14) является приближенной:$T_{\pm\alpha} \simeq exp(\pm h_\alpha D_\alpha)$. В зависимости от допустимого класса $H(u)$ приближенная формула для $T_{\pm\alpha}$ понимается по разному. Например, формальная запись $T_{\pm\alpha} \simeq exp(\pm h_\alpha D_\alpha)$ может означать, что

$$u(x_\alpha \pm h_\alpha) = (E \pm h_\alpha D_\alpha) u(x_\alpha) + \frac{h_\alpha^2}{2!} D_\alpha^2(\xi_\alpha), \quad \xi_\alpha \in [x_\alpha, x_\alpha \pm h_\alpha]$$

или

$$u(x_\alpha \pm h_\alpha) = (E \pm h_\alpha D_\alpha + \frac{h_\alpha^2}{2} D_\alpha^2) u(x_\alpha) \pm \frac{h_\alpha^3}{3!} D_\alpha^3(\eta_\alpha), \quad \eta_\alpha \in [x_\alpha, x_\alpha \pm h_\alpha].$$

Здесь под $H(u)$ следует понимать $H_{p_\alpha}(u)$, где в первом случае $p_\alpha = 2$, а во втором $p_\alpha = 3$. В дальнейшем вместо $H_{p_\alpha}(u(x_\alpha)) \equiv H_{p_\alpha}(u)$ используется обозначение $H(u)$. При этом предполагается, что класс $H(u)$ является допустимым, если $u(x)$ имеет нужное по ходу изложения число непрерывных производных по фазовым переменным x_α для $(x) \in Q$.

Для введенных в (1.44) операторов $\Lambda_\alpha^\pm$ использование (2.14) дает

$$\begin{aligned} \Lambda_\alpha^+ &= \frac{T_{+\alpha} - E}{h_\alpha} = \frac{e^{h_\alpha D_\alpha} - E}{h_\alpha} = D_\alpha + \frac{h_\alpha}{2} D_\alpha^2 + ... \\ \Lambda_\alpha^- &= \frac{E - T_{-\alpha}}{h_\alpha} = \frac{E - e^{-h_\alpha} D_\alpha}{h_\alpha} = D_\alpha - \frac{h_\alpha}{2} D_\alpha^2 + ... \end{aligned} \qquad (2.15)$$

Операторы $\Lambda_\alpha^\pm$ имеют смысл и для функций непрерывного аргумента, поэтому,если $H(u)$ содержит функции, имеющие непрерывные производные по x_α вплоть до второго порядка, то в силу (2.15)

$$\begin{aligned} u(x)_{x_\alpha} - D_\alpha u(x) = (\Lambda_\alpha^+ - D_\alpha) u(x) = \frac{h_\alpha}{2} D_\alpha^2 u(\xi_\alpha) = O(h_\alpha), \xi_\alpha \in [x_\alpha, x_\alpha + h_\alpha] \\ u(x)_{\bar{x}_\alpha} - D_u(x) = (\Lambda_\alpha^- - D_\alpha) u(x) = -\frac{h_\alpha}{2} D_\alpha^2 u(\eta_\alpha) = O(h_\alpha), \eta_\alpha \in [x_\alpha - h_\alpha, x_\alpha] \end{aligned} \qquad (2.16)$$

Если же допустимым классом $H(u)$ являются функции, имеющие непрерывные производные по x_α вплоть до третьего порядка, то

$$(\Lambda_\alpha^{0.5} - D_\alpha) u(x) \equiv [0.5(\Lambda_\alpha^+ + \Lambda_\alpha^-) - D_\alpha] u(x) = O(h_\alpha^2). \qquad (2.17)$$

С помощью операторов $\Lambda_\alpha^\pm, \Lambda_\alpha^{0.5}$ можно строить операторы разделенных разностей более высокого порядка, чем первый. Например,

$$\Lambda_{\alpha\alpha} = \Lambda_\alpha^+ \Lambda_\alpha^- = \frac{1}{h_\alpha} (\Lambda_\alpha^+ - \Lambda_\alpha^-) = \frac{T_{+\alpha} - 2E + T_{-\alpha}}{h_\alpha^2}, \qquad (2.18)$$

$$\Lambda_{\alpha\beta}^{0.5} = \Lambda_\alpha^{0.5} \Lambda_\beta^{0.5} = \frac{(T_{+\alpha} - T_{-\alpha})(T_{+\beta} - T_{-\beta})}{4 h_\alpha h_\beta}. \qquad (2.19)$$

Для (2.18),(2.19) можно указать такие $H(u)$, что имеет место

$$u(x)_{\bar{x}_\alpha x_\alpha} - D_\alpha^2 u(x) = (\Lambda_{\alpha\alpha} - D_\alpha^2) u(x) = O(h_\alpha^2), \qquad (2.20)$$

$$\begin{aligned} 0.25(u_{x_\alpha x_\beta} - u_{x_\alpha \bar{x}_\beta} - u_{\bar{x}_\alpha x_\beta} + u_{\bar{x}_\alpha \bar{x}_\beta}) - D_\alpha D_\beta u = \\ = (\Lambda_{\alpha\beta}^{0.5} - D_\alpha D_\beta)) u(x) = O(h_\alpha^2 + h_\beta^2). \end{aligned} \qquad (2.21)$$

Будем говорить, что сеточный оператор L_h аппроксимирует дифференциальный оператор L на $H(u)$, если для $u(x) \in H(u)$

$$\lim_{h_\alpha \to 0} ||(L - L_h)u|| = 0 \quad x \in Q_h. \tag{2.22}$$

Проверка условия (2.22) обычно связана с приближенной формулой (2.14): $T_\pm \simeq exp(\pm h_\alpha D_\alpha)$, что равносильно разложению $u(x)$ в ряд Тейлора и использованию остаточного члена ряда в форме Лагранжа (см.,например,(2.16)). При этом предполагается, что входящие в эти разложения дифференциальные операторы типа $D^m_\alpha, D^n_\beta, D^m_\alpha D^n_\beta, \alpha \neq \beta$ не связаны между собой. Однако такое сугубо формальное предположение не отвечает существу дела. В самом деле, при переходе $M_2 \to M_3$ совершается также переход от $u(x), x \in Q$ к $y(x), x \in Q_h$, где

$$Lu = f, \quad x \in Q; \quad L_h y = f_h, \quad x \in Q_h. \tag{2.23}$$

Поэтому, если L в (2.23) содержит дифференциальные операторы $D_\alpha, D_\beta, \beta \neq \alpha$, то эти операторы не являются независимыми, а связаны между собой в силу первого соотношения (2.23). Это позволяет под $H(u)$ в (2.22) понимать класс достаточно гладких решений исходной дифференциальной задачи.

Теперь снова обратимся к (2.1) и (1.35)

$$Lu \equiv (D_0 + aD_1)u = 0, \quad u(x,0) = \varphi(x); \quad L_h y \equiv (\Lambda^+_0 + a\Lambda^-_1)y^n_i = 0,,, \; y^0_i = \varphi_i \tag{2.24}$$

Краевое условие смешанной задачи Коши мы договорились учитывать при задании областей определения операторов L, L_h. Поэтому в (2.24)

$$u^n_0 = \mu^n, \quad \mu^0 = \varphi_0; \quad \Lambda^-_1 y^n_i = \begin{cases} \dfrac{y^n_i - y^n_{i-1}}{n}, & i = 2, ..., N_1 \\ \dfrac{y^n_1 - \mu^n}{h}, & i = 1; \quad \mu^0 = \varphi_0 \end{cases} \tag{2.25}$$

Далее воспользуемся (2.14).Тогда

$$L_h = L + \frac{1}{2}(\tau D^2_0 - ahD^2_1) + \frac{1}{3!}(\tau^2 D^3_0 + ah^2 D^3_1) + ... \tag{2.26}$$

Если теперь определить $H(u)$ как класс функций, имеющих непрерывные вторые производные по x и по t, то для $u \in H(u) \equiv C^{2,2}(x,t)$

$$\psi_h = (L_h - L)u = O(\tau + h) \tag{2.27}$$

В (1.35) y^n_0, y^0_i являются проекциями $u(0,t), u(x,0)$ на сетку Q_h, т.е.

$$y^n_0 - u(0, n\tau) = 0, \quad y^0_i - u(ih, 0) = 0. \tag{2.28}$$

Поэтому в соответствии с (2.22) соотношенние (2.27) означает, что <u>разностная схема (1.35)</u> <u>аппроксимирует дифференциальную задачу (2.1)</u>. К только что сказанному следует добавить: с первым порядком по τ и h. Повышение гладкости функций $u \in H(u)$ в данном случае не изменяет порядка аппроксимации. Ситуация может кардинально измениться, если в качестве $H(u)$ выбрать класс решений дифференциальной задачи (2.1). На этом классе

$$(D_0 + aD_1)u = 0 \;\rightarrow\; D_0 = -aD_1.$$

и поэтому для L_h из (2.26) справедливо любое из приводимых ниже представлений

$$L_h = \begin{cases} L + \dfrac{a}{2}(a\tau - h)D_1^2 - \dfrac{a^2}{3!}(a^2\tau^2 - h^2)D_1^3 + ... \\ L + \dfrac{1}{2}(\tau - \dfrac{h}{a})D_0^2 + \dfrac{1}{3!}(\tau^2 - \dfrac{h^2}{a^2})D_0^3 + ... \end{cases} \tag{2.29}$$

Как это следует из (1.28), гладкость решения задачи (2.1) полностью определяется гладкостью предельных условий $\varphi(x)$ и $\mu(t)$. Если $\varphi(x) \in C^\infty(x), \mu(t) \in C^\infty(t)$ и $\varphi(0) = \mu(0)$, то в (2.1)

$$u(x,t) \in C^\infty(x,t), \quad (x,t) \in Q \tag{2.30}$$

Закон предельного перехода в (1.35) зададим следующим образом

$$r = \frac{a\tau}{h} = 1, \quad \tau \rightarrow 0, \quad h \rightarrow 0. \tag{2.31}$$

В качестве следствия из (2.29)-(2.31) получаем, что для L из (2.1), L_h из (1.35) имеет место

$$L_h = L \tag{2.32}$$

Остается учесть (2.28), чтобы утверждать: при выполнении условий (2.30),(2.31) <u>разностная схема (1.35)</u> <u>аппроксимирует</u> <u>исходную задачу (1.27)</u> <u>с бесконечным порядком точности</u>. Уникальность этого результата заключается в том, что совпадение всех соотношений непрерывной задачи (1.26) и дискретной задачи (1.35) – "бесконечный порядок точности" достигается при конечных τ и h. При этом проекция решения непрерывной задачи (2.27)на сетку Q_h (2.31) совпадает с решением сеточной задачи (1.35).

Итак, при $r > 1$ разностная схема (1.35) непригодна для численного решения смешанной задачи Коши (1.27); при $r = 1$: $u(x,t) = y(x,t)$, $(x,t) \in Q_h$. Остается рассмотреть случай $r < 1$. Это условие вместе с (2.11) приводит к оценке

$$|y_i^n| \leq r|y_{i-1}^{n-1}| + (1-r)|y_i^{n-1}|. \tag{2.33}$$

Отсюда заключаем, что при $r < 1$ для разностной схемы (1.35) выполняется <u>принцип максимума</u>: $max|y_i^n|$ для $(ih, n\tau) \in Q_h$ достигается либо при $i = 0$, либо при $n = 0$. Из этого принципа вытекает очевидное, но очень полезное Следствие 1. Если в (1.27) $|\varphi(x)| < M_1 = const.$, $|\mu(t)| < M_1$, $(x,t) \in Q$, то в (1.35)

$$|y_i^n| < M_1, \quad (ih, n\tau) \in Q_h. \tag{2.34}$$

Если не оговорено противное, то используемые здесь и далее константы $M_\alpha > 0$ не зависят от τ, h.

Далее введем в рассмотрение вектор погрешности $z^n = y^n - u^n$, $(x,t) \in Q_h$. Тогда в силу (1.27),(1.35) и (2.24)

$$L_h z^n = \psi^n, \quad \psi^n = (L - L_h)u^n, \quad z_i^0 = 0, \quad z_0^n = 0 \tag{2.35}$$

Переход в (2.35) к покомпонентной записи дает

$$z_i^n = r z_{i-1}^{n-1} + (1-r) z_i^{n-1} + \tau \psi_i^n \tag{2.36}$$

Теперь следует учесть условие $r < 1$, чтобы из (2.36) получить

$$|z_i^n| \le max(|z_{i-1}^{n-1}|, |z_i^{n-1}|) + \tau |\psi_i^n|$$

Для сеточного вектора v^n положим

$$||v^n||_1 = \max_i |v_i^n| \tag{2.37}$$

В соответствии с этим определением предыдущее неравенство можно представить следующим образом

$$||z^n||_1 \le ||z^0||_1 + \tau \sum_{j=1}^{n} ||\psi^j||_1. \tag{2.38}$$

Если теперь

$$||\psi^n||_* = \max_{j \le n} ||\psi^j||_1,$$

то

$$||z^n||_1 \le ||z^0||_1 + n\tau ||\psi^n||_* \tag{2.39}$$

Остается учесть (2.27) и (2.35), чтобы получить

$$||z^n||_1 \le n\tau M_2(\tau + h) \equiv n\tau M_2 |h|. \tag{2.40}$$

Зафиксируем $n = n\tau$ и устремим $|y| \to 0$ так, чтобы $a\tau < h$. Тогда

$$\lim_{|h| \to 0} ||z(t)||_1 = 0 \tag{2.41}$$

Тем самым доказана Теорема 1. Пусть $u(x,t)$-решение задачи (1.27), такое, что $u(x,t) \in C^{2,2}(x,t), (x,t) \in Q$, а $y(ih, n\tau)$-решение разностной схемы (1.35) при $r < 1$. Если $ih \to x, n\tau \to t$ и $(x,t) \in Q$, то

$$\lim_{|h| \to 0} y(ih, n\tau) = u(x,t) \tag{2.42}$$

Для достаточно гладких $u \in H(u)$ справедливость теоремы 1 лимитируется условием $r < 1$. Поэтому говорят об условном характере сходимости разностной схемы (1.35). Неявная разностная схема (1.38) в этом смысле является безусловно сходящейся. Действительно, в этом случае

$$y_i^n = \frac{r}{1+r} y_{i-1}^{n-1} + \frac{1}{1+r} y_i^{n-1}$$

и для разностной схемы (1.38) при любом $r > 0$ имеет место как принцип максимума, так и вытекающее из этого принципа неравенство(2.38). Но именно на этом неравенстве и условии аппроксимации $(L - L_h)u = O(\tau + h) = O(|h|)$ основано доказательство теоремы 1. В этой связи представим достаточные условия, следствиями которых являются (2.27),(2.38) и (2.42) в некоторой абстрактной форме.В такой же абстрактной форме представим и соответствующие друг другу непрерывную и дискретные задачи. Решение исходной дифференциальной задачи $u(Q) = u(x_\alpha, t)$ будем рассматривать в качестве элемента линейного нормированного пространства H, тогда под H_h будем понимать линейное нормированное пространство элементов y, определенных на сетке Q_h. Пусть $v \in H$, а v_h-проекция v на Q_h. Норму $|| \cdot ||_{H_h}$ введем таким образом, чтобы $\forall v \in H$:

$$\lim_{|h|\to 0} ||v_h||_{H_h} = ||v||_H \tag{2.43}$$

Замечание 1.Для нестационарных задач вместо $u(Q) \in H(Q)$ можно рассмативать однопараметическое семейство u(t) такое, что

$$\forall t \in [0,T]: u(t) \in H(D), \quad (x_\alpha) \in D. \tag{2.44}$$

В зависимости от конкретной непрерывной задачи под $H(Q), H(D)$ обычно понимаются конкретные функциональные простанства Банаха.

Замечание 2. В соответствии с (2.44) решение дифференциальной задачи (1.27) опеделяется как однопараметрическое семейство $u(t) \in H(D)$ такое, что $\forall t \in [0,T]$: $u(t) \in R(\frac{\partial}{\partial x})$, далее, $u(x,0) = \varphi(x) \in H(D)$; наконец, равномерно по t

$$\lim_{\Delta t\to 0} ||\frac{u(t+\Delta t) - u(t)}{\Delta t} + a\frac{\partial u(t)}{\partial x}||_{H(D)} = 0 \tag{2.45}$$

Пусть теперь $u(Q) \in H$- решение дифференциальной задачи (2.1) и u_h- проекция $u(Q)$ на сетку Q_h. Пусть $y \in H_h$- решение разностной схемы (2.2). Говоря о разностной схеме (2.2), мы имеем ввиду семейство соотношений (2.2), которое зависит от параметров дискретизации τ, h.Задание закона предельного перехода $\tau = \tau(h)$ позволяет считать семейство (2.2) зависящем только от малого параметра дискретизации h. Пусть, наконец, $z = (y - u_h) \in H_h$. Тогда

$$L_h z^n = \psi_1^n = (L - L_h)u_h^n, \quad z^0 = y^0 - u_h^0 = \psi_2 \tag{2.46}$$

Будем говорить, что разностная схема (2.2)аппроксимирует дифференциальную задачу(2.1) с порядком m на решении $u(Q)$, если

$$||\psi_1||_{\Phi_h} = O(|h|^m), \quad ||\psi_2||_{H_h(D_h)} = O(|h|^m). \tag{2.47}$$

Линейное нормированное пространство Φ_h в (2.47) определяется следующим образом

$$\forall\, v \in H_h: \quad L_h\, v = \psi \in \Phi_h, \qquad L_h : H_h \to \Phi_h \tag{2.48}$$

Только что приведенное определение аппроксимации исходной дифференциальной задачи разностной схемы является одним из возможных. Тем не менее (2.47) реализует одно из необходимых условий при переходе от непрерывной модели к дискретной ($M_2 \to M_3$): аппроксимации подлежат все соотношения непрерывной модели. Поэтому, например, если в (2.1) и в (2.2): $Lu = f, L_h y = \tilde{f}_h$ и $\psi_3 = f - \tilde{f}_h$, то (2.47) следует дополнить соотношением $||\psi_3||_{\Phi_h} = O(|h|^m)$. Если обратиться к конкретным задачам (2.1) и (2.2), то символ $O(|h|^m)$ следует заменить на $O(h^m)$. Тогда весьма существенным в определении аппроксимации становится закон предельного перехода:$\tau = \tau(h)$. Здесь мы ограничимся простым примером. Пусть в (2.2)

$$\hat{L}_h y = [\frac{T_{+0} - 0.5(T_{+1} + T_{-1})}{\tau} + a\frac{T_{+1} - T_{-1}}{2h}]y. \tag{2.49}$$

В соответствии с (2.15),(2.18) сеточный оператор $\hat{L}_h$ можно представить следующим образом

$$\hat{L}_h = \Lambda_0^+ + 0.5a(\Lambda_1^+ + \Lambda_1^-) - \frac{h^2}{2\tau}\Lambda_{11}.$$

Поэтому на достаточно гладком допустимом классе $H(u)$ будем иметь

$$L_h u_h = Lu - \frac{h^2}{2\tau}\frac{\partial^2 u}{\partial x^2} + \frac{\tau}{2}\frac{\partial^2 u}{\partial t^2} + O(\tau^2, h^2, \frac{h^4}{\tau}).$$

Пусть $\tau \to 0, h \to 0$ таким образом, что $a\tau = rh, r = const$. Тогда

$$\hat{L}_h u_h = Lu + O(h)$$

и в соответствии с (2.22) оператор $\hat{L}_h$ аппроксимирует оператор L с первым порядком точности по h. Если же $\tau \to 0, h \to 0$ и при этом $2\tau\nu = h^2, \nu = const.$, то

$$\hat{L}_h u_h = \hat{L}u + O(h^2), \quad \hat{L}u = Lu - \nu\frac{\partial^2 u}{\partial x^2}.$$

В этом случае оператор $\hat{L}_h$ аппроксимирует оператор $\hat{L}$ со вторым порядком точности по h. Итак,аппроксимация оператора L из (2.2) оператором (2.49) имеет условный характер, ибо связана с выбираемым законом предельного перехода: $\tau = \tau(h)$. Так как нетрудно привести и другие примеры подобного рода, то закон предельного перехода следует включить в само понятие:"разностная схема". Линейную разностную схему

$$L_h y = f_h, \quad y^0 = \varphi_h, \quad (x,t) \in Q_h \tag{2.50}$$

будем называть устойчивой(по начальным данным φ_h и правой части f_h), если существуют константы M_1, M_2, не зависящие от параметров дискретизации, такие что

$$||y||_{(1)} \leq M_1||f_h||_{(2)} + M_2||\varphi_h||_{(3)}. \tag{2.51}$$

В (2.51) $||\cdot||_{(\alpha)}$- некоторая норма в пространстве сеточных функций, удовлетворяющих условию (2.43). Из (2.51) вытекает, что задача (2.50) однозначно разрешима при любых f_h, φ_h. Действительно, если положить в (2.50) $f_h = 0, \varphi_h = 0$, то в силу (2.51) $y = 0$, следовательно однородная задача (2.50) имеет только тривиальное решение. Остается заметить, что (2.50) является системой линейных алгебраических уравнений, число неизвестных y_i^n в которой равно числу уравнений.