

Светлой памяти А.С. Холодова посвящается

УДК 519.63

## БЕЗМАТРИЧНАЯ РЕАЛИЗАЦИЯ НЕЯВНЫХ СХЕМ МЕТОДОМ СОПРЯЖЕННЫХ ГРАДИЕНТОВ

© 2018 г. Н. Г. Бураго<sup>1</sup>, И. С. Никитин<sup>2,3</sup>

<sup>1</sup>119526 Москва, пр-т Вернадского, 101, кор. 1, ИПМех РАН;

<sup>2</sup>123056 Москва, ул. 2-я Брестская, 19/18, ИАП РАН;

<sup>3</sup>125993 Москва, Волоколамское ш., 4, МАИ)

e-mail: buragong@yandex.ru

Поступила в редакцию 05.03.2018 г.

Рассматриваются безматричные алгоритмы метода сопряженных градиентов для решения больших систем алгебраических уравнений, возникающих при реализации неявных схем численного решения задач механики сплошных сред. Дано описание алгоритмов и показана их эффективность для минимизации потребностей в оперативной памяти и в быстродействии ЭВМ, а также в отношении свойств робастности, простоты реализации и отладки алгоритмов решения. Библ. 13. Фиг. 5.

**Ключевые слова:** безматричный метод сопряженных градиентов, большие системы алгебраических уравнений, неявные схемы, механика сплошных сред, робастность.

**DOI:**

### 1. ОПИСАНИЕ МЕТОДА

Метод сопряженных градиентов был предложен в [1]. В основе метода лежит идея итерационной минимизации квадратичного функционала

$$F(\mathbf{x}) = (\mathbf{C}\mathbf{x}, \mathbf{x}) / 2 - (\mathbf{d}, \mathbf{x}),$$

имеющего минимум на решении исходной системы алгебраических уравнений

$$\mathbf{C}\mathbf{x} = \mathbf{d},$$

где  $\mathbf{C}$  – положительная симметричная  $N \times N$  матрица,  $N$  – число неизвестных,  $\mathbf{x}$  – вектор неизвестных размера  $N$  и  $\mathbf{d}$  – вектор правой части. На каждой итерации  $n$  направлением поиска  $\mathbf{s}_n$  является линейная суперпозиция направления поиска  $\mathbf{s}_{n-1}$  предыдущей итерации и направления градиента  $\mathbf{g}_n = \mathbf{C}\mathbf{x}_n - \mathbf{d}$  функционала  $F$ :

$$\alpha_n = (\mathbf{g}_n, \mathbf{s}_n) / (\mathbf{C}\mathbf{s}_n, \mathbf{s}_n), \quad \mathbf{x}_{n+1} = \mathbf{x}_n - \alpha_n \mathbf{s}_n, \quad \mathbf{g}_{n+1} = \mathbf{g}_n - \alpha_n \mathbf{C}\mathbf{s}_n, \quad (1.1)$$

$$\beta_n = (\mathbf{g}_{n+1}, \mathbf{C}\mathbf{s}_n) / (\mathbf{C}\mathbf{s}_n, \mathbf{s}_n), \quad \mathbf{s}_{n+1} = \mathbf{g}_{n+1} - \beta_n \mathbf{s}_n, \quad (1.2)$$

где  $\mathbf{s}_0 = \mathbf{g}_0 = \mathbf{C}\mathbf{x}_0 - \mathbf{d}$ . Выражения для коэффициентов  $\alpha_n$  и  $\beta_n$  получаются из условий минимума функционала вдоль направления поиска. Доказано (см. например [2]), что метод сопряженных градиентов вырабатывает  $\mathbf{C}$ -ортогональный базис в пространстве решений

$$(\mathbf{C}\mathbf{s}_i, \mathbf{s}_j) = c_i \delta_{ij},$$

где  $c_i > 0$  и  $\delta_{ij}$  – дельта Кронекера. На каждой итерации определяется проекция решения на очередной вырабатываемый базисный вектор. Для реализации процесса в начале при  $n=0$  вычисляется начальная невязка уравнений, а затем на каждой итерации требуется вычислять однородную часть невязки  $\mathbf{C}\mathbf{s}_n$ . Всего алгоритм использует 4 вектора длины  $N$ :

$$\mathbf{a}_1 = \mathbf{x}_n, \quad \mathbf{a}_2 = \mathbf{s}_n, \quad \mathbf{a}_3 = \mathbf{g}_n, \quad \mathbf{a}_4 = \mathbf{C} \cdot \mathbf{s}_n$$

В случаях когда матрица исходной системы алгебраических уравнений является несимметричной и/или нейтральной, метод сопряженных градиентов применяется к подправленной системе уравнений  $\mathbf{C}^T(\mathbf{C}\mathbf{x} - \mathbf{d}) = 0$  и принимает следующий вид:

$$\alpha_n = (\mathbf{g}_n, \mathbf{C}\mathbf{s}_n) / (\mathbf{C}\mathbf{s}_n, \mathbf{C}\mathbf{s}_n), \quad \mathbf{x}_{n+1} = \mathbf{x}_n - \alpha_n \mathbf{s}_n, \quad \mathbf{g}_{n+1} = \mathbf{g}_n - \alpha_n \mathbf{C}\mathbf{s}_n, \quad (1.3)$$

$$\beta_n = (\mathbf{C}\mathbf{g}_{n+1}, \mathbf{C}\mathbf{s}_n) / (\mathbf{C}\mathbf{s}_n, \mathbf{C}\mathbf{s}_n), \quad \mathbf{s}_{n+1} = \mathbf{g}_{n+1} - \beta_n \mathbf{s}_n, \quad (1.4)$$

где  $\mathbf{s}_0 = \mathbf{g}_0 = \mathbf{C}\mathbf{x}_0 - \mathbf{d}$ . В этом алгоритме на каждой итерации приходится выполнять не одно, а два вычисления однородной части невязки ( $\mathbf{C}\mathbf{s}_n$  и  $\mathbf{C}\mathbf{g}_{n+1}$ ). Для реализации потребуется 5 векторов длины  $N$ :  $\mathbf{a}_1 = \mathbf{x}_n$ ,  $\mathbf{a}_2 = \mathbf{s}_n$ ,  $\mathbf{a}_3 = \mathbf{g}_n$ ,  $\mathbf{a}_4 = \mathbf{C} \cdot \mathbf{s}_n$ ,  $\mathbf{a}_5 = \mathbf{C} \cdot \mathbf{g}_n$ .

Заметим, что для реализации итерационного процесса метода сопряженных градиентов формирование, запоминание и какие-либо операции с матрицей системы уравнений  $\mathbf{C}$  не требуются. Требуется только алгоритм вычисления невязки  $\mathbf{C}\mathbf{x}_n - \mathbf{d}$  и ее однородной части  $\mathbf{C}\mathbf{x}_n$ . Такой алгоритм значительно проще, чем формирование матрицы  $\mathbf{C}$ . При отсутствии ошибок округления точное решение определяется за  $N$  итераций (базис конечномерного пространства решений построен и проекции решения найдены), то есть итерационный процесс сходится за конечное число итераций, равное размерности вектора неизвестных.

Критерием окончания итерационного процесса является условие

$$(\mathbf{g}_n, \mathbf{g}_n) \leq \varepsilon \vee (s_n, s_n) \alpha_n^2 \leq \varepsilon,$$

где  $\varepsilon$  — малое число, близкое к машинному эпсилон (т. е. к числу, добавление которого к единице не меняет ее величину), которое зависит от разрядности представления чисел в ЭВМ. Для четырехбайтового представления вещественных чисел  $\varepsilon \approx 10^{-6}$ .

Случаи  $(As_n, s_n) \leq \varepsilon$  ( $A > 0$ ) или  $(As_n, As_n) \leq \varepsilon$  указывают на некорректную постановку задачи (вырожденность матрицы системы).

Вычисления проводятся с ограниченным числом значащих цифр (разрядов) и число обусловленности матрицы  $\mathbf{C}$  может быть много больше единицы, из-за этого в итерациях накапливаются ошибки в определении базисных векторов и соответствующих проекций искомого решения. В результате при очень плохой обусловленности (при очень больших значениях числа обусловленности) сходимость итераций метода сопряженных градиентов может вообще отсутствовать. Поэтому решаемая система уравнений преобусловливается путем умножения на приближенную обратную матрицу  $\tilde{\mathbf{C}}^{-1}$ :

$$\tilde{\mathbf{C}}^{-1}(\mathbf{C}\mathbf{x} - \mathbf{d}) = 0,$$

где  $\tilde{\mathbf{C}}^{-1}\mathbf{C} \approx \mathbf{E}$ ,  $\mathbf{E}$  — единичная матрица. В идеальном случае для  $\tilde{\mathbf{C}}^{-1}\mathbf{C} = \mathbf{E}$  такое умножение точно решает задачу.

Наибольшее распространение получили способы преобусловливания, основанные на ленточных приближениях матриц  $\mathbf{C}$ , обрабатываемых по методу Холецкого (см. [3]), сохраняющего ширину ленточных матриц. Однако даже применение ленточных приближений матрицы  $\mathbf{C}$  при большом числе неизвестных ( $N > 10^5 \div 10^6$ ) связано со значительными трудностями, особенно в трехмерных задачах с неизбежно большой шириной лент ненулевых элементов.

Операции с матрицами из алгоритмов метода сопряженных градиентов можно полностью исключить, что будет показано далее. Практика применения метода сопряженных градиентов для начально-краевых задач математической физики при использовании конечно-разностных и конечно-элементных аппроксимаций показала, что сходимость метода сопряженных градиентов вполне обеспечивается применением простейшей приближенной обратной матрицы, составленной из обратных диагональных элементов исходной матрицы, что эквивалентно масштабированию неизвестных. Вычисление, запоминание и обращение диагональных элементов для матрицы  $\mathbf{C}$  не встречает никаких трудностей.

Таким образом, подобно методам исключения Гаусса метод сопряженных градиентов требует конечного числа операций для отыскания решения, но абсолютно не нуждается в каких-либо операциях с матрицей системы линейных алгебраических уравнений, включая ее формирование и запоминание. Для реализации метода сопряженных градиентов достаточно построить алгоритм определения градиента  $\mathbf{g}_n = \mathbf{C}\mathbf{x}_n - \mathbf{d}$  и его однородной части  $\mathbf{C}\mathbf{x}_n$  для приближенного решения  $\mathbf{x}_n$ . Матрицы для такого алгоритма вообще не требуются.

Потребление машинной памяти с ростом числа неизвестных растет линейно и не зависит от числа пространственных координат. Требуемый объем машинной памяти пропорционален  $4N - 5N$  в зависимости от наличия или отсутствия свойства симметрии и положительности оператора  $\mathbf{C}$ . К привлекательным качествам метода также относится унаследованное от градиентных методов свойство описывать основные (по величине) проекции искомого решения уже на первых итерациях. Благодаря этому

для получения решений с машинной точностью (т. е. с точностью до малых, добавление которых к единице не меняет ее машинное представление) достигается уже после  $\sqrt{N}$  итераций. Поскольку объем вычислений на отдельной итерации пропорционален  $N$  (как на временных шагах явных схем), то скорость роста объема вычислений, требуемых для решения, в зависимости от числа неизвестных пропорциональна  $N^{3/2}$ , т. е., например, при числе неизвестных, равном миллиону, для определения решения требуется примерно тысяча итераций. При этом, как ниже будет показано, вычисление градиентов (или невязок уравнений) на каждой итерации абсолютно идентично вычислению решений на новом временном слое для двухслойных явных схем.

Необходимо также отметить предпочтительность итерационных безматричных методов для распараллеливания вычислений на многопроцессорных ЭВМ.

Имеется много модификаций метода сопряженных градиентов (см. например [2]). Вариант (1.1), (1.2) происходит от градиентного метода наискорейшего спуска (на каждой итерации минимизируется исходный квадратичный функционал). Вариант (1.3), (1.4) происходит от градиентного метода наименьших невязок (на каждой итерации минимизируется норма невязки).

## 2. БЕЗМАТРИЧНАЯ РЕАЛИЗАЦИЯ МЕТОДА СОПРЯЖЕННЫХ ГРАДИЕНТОВ

Для реализации итераций по методу сопряженных градиентов достаточно располагать алгоритмами вычисления векторов невязок однородной и неоднородной систем уравнений, которые реализуются аналогично явным схемам. Рассмотрим эти алгоритмы для решении задач механики сплошных сред (МСС) на примере вариационных сеточных методов.

Поскольку задачи механики сплошных сред в большинстве случаев нелинейны, то для их решения как правило применяется квазилинеаризация. Причем нестационарные задачи квазилинеаризуются на каждом шаге по времени, а стационарные задачи квазилинеаризуются на каждой итерации какого-либо квазиньютоновского итерационного процесса (варианты методов Ньютона–Канторовича или дифференцирования по параметру). После этого на каждом шаге по времени или на каждой итерации по нелинейности возникают вспомогательные линейные краевые задачи, для решения которых в нашем случае применяется метод сопряженных градиентов.

Важно отметить, что квазилинеаризация вынужденно проводится на исходных интегродифференциальных уравнениях, поскольку дискретные нелинейные уравнения как правило не имеют аналитического представления и единственным способом задания нелинейных уравнений на ЭВМ являются алгоритмы вычисления невязок таких уравнений.

### 2.1. Формулировка типовой начально-краевой задачи

Постановка задач МСС содержит балансные соотношения для основных процессов, которые в общем виде при использовании произвольно подвижной системы координат записываются (см. [4]) в виде

$$\frac{\partial}{\partial t} \int_V A \delta A dV + \int_V (\mathbf{B} - A\Omega) \cdot \nabla \delta A dV = \int_V f \delta A dV + \int_S (\mathbf{B} - A\Omega) \cdot \mathbf{n} \delta A dS. \quad (2.1)$$

Здесь  $A$  – сохраняемая величина,  $\mathbf{B}$  – поток величины  $A$ , обусловленный взаимодействием и хаотическими движениями молекул (упругость, вязкость, диффузия),  $A\Omega$  – конвективный поток величины  $A$  (обусловленный упорядоченным движением материальной среды относительно координатной среды),  $f$  – источник/сток величины  $A$ ,  $V$  – область решения с границей  $S$ ,  $\Omega = \mathbf{u} - \mathbf{w}$  – конвективная скорость,  $\mathbf{u}$  и  $\mathbf{w}$  – скорости материальной и координатной сред. Временная производная  $\partial/\partial t$  вычисляется вдоль траекторий произвольно подвижных координат (вдоль траекторий узлов подвижной сетки). В частности, если в данном узле  $\mathbf{w} = \mathbf{u}$ , то имеем узел лагранжев, если  $\mathbf{w} = 0$ , то – узел эйлеров. Использование здесь и далее термина “координатная среда” оправдано тем, что на практике управление движением подвижных координат (движением узлов и ячеек) удобно реализуется с помощью представления множества расчетных узлов и ячеек в виде разновидности сплошной (“координатной”) среды.

Потоки  $\mathbf{B}$  связаны с градиентами сохраняемых зависимых переменных так называемыми определяющими соотношениями сплошной среды, которые в достаточно общем виде для сред дифференциального типа (среды с короткой памятью) можно представить в виде

$$\mathbf{B} = \mathbf{K} \cdot \nabla A + \mathbf{L}, \quad (2.2)$$

где тензоры  $\mathbf{K}$  и  $\mathbf{L}$  – заданные функции координат, времени и искомого решения  $A$ , точка означает скалярное умножение. Движение сеточной среды считается заданным или определенным из решения вспомогательной отдельной “сеточной задачи”. Задача (2.1), (2.2) дополняется начальными

$$t = 0 : A = A^0(\mathbf{x})$$

и граничными условиями

$$\mathbf{x} \in S_A : A = A_*(\mathbf{x}, t), \quad \mathbf{x} \in S_B = S \setminus S_A : \mathbf{B} \cdot \mathbf{n} = \mathbf{B}_*(\mathbf{x}, t).$$

Звездочками отмечены заданные функции. Для обсуждения безматричных алгоритмов метода сопряженных градиентов для неявных схем дальнейшая детализация формулировки типовой задачи МСС не требуется.

### 2.2. Аппроксимации решения и уравнений

Пусть сетка сгенерирована. Сетка содержит  $N_1$  узлов с координатами  $\mathbf{x}_i$ ,  $i = 1, \dots, N_1$ . Узлы сетки объединены в  $N_2$  ячеек, заданных информационным целочисленным массивом  $J_2(k, j)$ ,  $k = 1, \dots, N_2$ ,  $j = 1, \dots, M_2$ , где  $J_2(k, j)$  – глобальный номер локального узла  $j$  в ячейке  $k$ ,  $N_2$  – число ячеек,  $M_2$  – число узлов в ячейке. Граничная сетка задана информационным МАССИВОМ граничных ячеек  $J_3(k, j)$ ,  $k = 1, \dots, N_3$ ,  $j = 1, \dots, M_3$ , где  $J_3(k, j)$  – глобальный номер локального узла  $j$  в граничной ячейке  $k$ ,  $N_3$  – число граничных ячеек,  $M_3$  – число узлов в граничной ячейке.

В дальнейшем будем именовать ячейки сетки (в данном случае, ячейки Дирихле) *конечными элементами*. Значение искомой функции в узле  $i$  в момент времени  $t_n$  ( $n$  – номер временного слоя) обозначим через  $A_i^n$ . Используем простейшую кусочно-линейную аппроксимацию решения. Значения искомых функций  $[A]_k$  в центре каждого элемента  $k$  вычисляются по формуле

$$[A]_k = \sum_{j=1}^M \frac{1}{M} A_{J_2(k,j)}, \tag{2.3}$$

где  $M$  – число узлов в элементе,  $J_2(k, j)$  – глобальный номер узла с локальным номером  $j$  в элементе  $k$ . Формула (2.3) одинакова и для внутренних, и для граничных ячеек (элементов). В общем случае в сетках могут присутствовать ячейки с разным числом узлов, так что число узлов  $M$  может быть переменным. Но это не принципиально и номер элемента у числа  $M$  опущен, чтобы не загромождать запись.

Пространственные производные вычисляются по формуле

$$[\nabla A]_k = \sum_{j=1}^M \nabla_{kj} A_{J_2(k,j)}.$$

Выражения для коэффициентов оператора пространственного дифференцирования  $\nabla_{kj}$  имеются в учебниках и справочниках, так что выписывать их здесь необходимости нет.

Интегралы в балансных уравнениях вычислим простейшими квадратурами Гаусса (сумма интегралов по ячейкам, определяемых произведениями значений подынтегральных выражений в центре элемента на длину, площадь или объем ячейки в зависимости от числа пространственных измерений) с гауссовой точкой в центре элемента, например,

$$\int_{V_k^n} (\mathbf{B} - A\Omega) \cdot \nabla \delta A dV = ([\mathbf{B}]_k^n - [A]_k^n [\Omega]_k^n) \cdot [\nabla \delta A]_k^n V_k^n.$$

В одномерном случае это квадратурная формула прямоугольников.

Для членов уравнений с временными производными от искомой функции используются квадратурные формулы с гауссовыми точками в узлах элемента

$$\int_{V_k^n} \frac{\partial A}{\partial t} \delta A dV = \sum_{j=1}^M \frac{A_{J_2(k,j)}^{n+1} - A_{J_2(k,j)}^n}{M \Delta t_n} \delta A_{J_2(k,j)} V_k^n.$$

Такая формула в одномерном случае отвечает интегральной формуле трапеций. Здесь для расчета решения  $A_i^{n+1}$  на новом временном слое  $t_{n+1}$  производная по времени представлена в разностной форме. Вариация решения от времени не зависит, так что верхний индекс у вариации решения отсутствует. Благодаря принятому представлению интегралов от членов с временными производными для их определения не требуется обращать ленточные матрицы масс, поскольку они являются диагональными (так называемые несогласованные матрицы масс [5]).

Заметим, что раньше в литературе по методу конечных элементов было распространено мнение, что применение несогласованных матриц масс в конечно-элементных алгоритмах является источником их неустойчивости в нестационарных задачах, а согласованные матрицы масс, полученные, например,

интегрированием по формулам с гауссовыми точками в центрах элементов, наоборот, способствуют более устойчивому расчету. Но благодаря теории разностных схем стало ясно, что причина неустойчивости та же самая, что и в явных центрально-разностных схемах, а именно, наличие диффузионных членов в первых дифференциальных приближениях с отрицательными коэффициентами вязкости. Для устойчивости схем надо вводить дополнительную явную или неявную вязкость. Многочисленные пути введения таких стабилизирующих членов или аппроксимаций представляют отдельную проблему, лежащую вне рамок данного изложения (подробности см. в [6]).

Наиболее просто и достаточно эффективно вариационное балансное уравнение (3.1) можно решить по следующей явно-неявной схеме:

$$\frac{\partial}{\partial t} \int_{V^n} \frac{A^{n+1} - A^n}{\Delta t_n} \delta AdV + \int_{V^n} (\tilde{\mathbf{B}}^{n+1} - A^n \Omega^n) \cdot \nabla^n \delta AdV = \int_{V^n} f^{n+1} \delta AdV + \int_{S_B^n} (\tilde{\mathbf{B}}_*^{n+1} - A^n \Omega^n) \mathbf{n}^n \delta AdS, \quad (2.4)$$

где для устойчивости сделаны добавки искусственной вязкости в потоковые члены  $\mathbf{B}$ :

$$\tilde{\mathbf{B}}^{n+1} = \mathbf{B}^{n+1} + v_{art}^n \mathbf{I} \cdot \nabla^n A^{n+1}.$$

Здесь коэффициент дополнительной стабилизирующей вязкости  $v_{art}^n$  вводится явно или неявно (за счет нецентральной аппроксимации производных, например) в соответствии с конкретно используемой конечно-элементной схемой. Единичный тензор второго ранга обозначен здесь и далее через  $\mathbf{I}$ . Напомним, что звездочкой отмечены заданные на границе функции.

Заметим, что в уравнении (3.4) учтены естественные граничные условия (ограничения на граничные потоки). Поскольку на части границы  $S_A^n \subset S^n$  искомая функция  $A$  задана, то ее вариации там равны нулю, поэтому граничный интеграл берется только по остальной части границы  $S_B^n = S^n \setminus S_A^n$ . Верхний индекс  $n$  использован, так как положение границ и граничные условия переменны во времени.

После подстановки пространственных аппроксимаций производных и интегралов получаем следующее приближенное представление вариационного уравнения:

$$\begin{aligned} & \sum_{k=1}^{N_2} \left[ \sum_{j=1}^{M_2} \left( \frac{A_{J_2}^{n+1}(k,j) - A_{J_2}^n(k,j)}{\Delta t_n} - f_{J_2}^n(k,j) \right) \frac{1}{M_2} \delta A_{J_2}(k,j) V_k^n \right] + \\ & + \sum_{k=1}^{N_2} \left[ \left( \tilde{K}_k^n \sum_{j=1}^{M_2} \nabla_{kj}^n A_{J_2}^{n+1}(k,j) \right) \cdot \sum_{l=1}^{M_2} \nabla_{kl}^n \delta A_{J_2}(k,j) V_k^n \right] + \\ & + \sum_{k=1}^{N_2} \left[ L_k^n \cdot \sum_{l=1}^{M_2} \nabla_{kl}^n \delta A_{J_2}(k,j) V_k^n \right] - \sum_{k=1}^{N_2} \left[ \left( \sum_{j=1}^{M_2} \frac{1}{M_2} A_{J_2}^n(k,j) \right) \cdot \sum_{l=1}^{M_2} \nabla_{kl}^n \delta A_{J_2}(k,j) V_k^n \right] - \\ & - \sum_{k=1}^{N_3} \left[ \left( B_{*k}^{n+1} - \Omega_k^n \sum_{l=1}^{M_3} \frac{A_{J_3}^n(k,l)}{M_3} \right) \sum_{j=1}^{M_3} \frac{\delta A_{J_3}(k,j)}{M_3} S_k^n \right] = 0. \end{aligned}$$

Здесь  $\tilde{K}_k^n = (K + v_{art} \mathbf{I})_k^n$ . В программах для ЭВМ суммам соответствуют циклы по ячейкам, внутри которых имеются циклы по локальным узлам в ячейках. Дискретные вариации искомых функций в узлах никаких значений не принимают (основное свойство вариаций это их произвольность), так что если собрать подобные члены при каждой дискретной вариации, то в сумме они обязаны на решении обратиться в нуль (лемма Эйлера). Такие суммы и есть уравнения, образующие линейную алгебраическую систему для данной задачи. Приведение подобных членов реализуется в процессе обхода ячеек, при этом дискретные вариации служат указателями того, с какой компонентой вектора градиента (невязки) суммировать очередной множитель при данной вариации. Алгоритм приведения подобных членов совпадает с расчетом временного шага в явных двухслойных схемах. Процесс подсчета градиентов и их однородных частей можно записать в виде формул:

$$\mathbf{g} = \mathbf{C}\mathbf{y} - \mathbf{d} = G_i^n \frac{A_i^{n+1} - A_i^n}{\Delta t_n} - P_i^n - Q_i^{n+1}, \quad \mathbf{g}_0 = \mathbf{C}\mathbf{y} = G_i^n \frac{A_i^{n+1}}{\Delta t_n} - Q_i^{n+1},$$

где вектор неизвестных  $\mathbf{y} = \{A_i^{n+1}\}_{i=1}^N$  составлен из искомых узловых значений функции  $A$ , а множитель  $G_i^n$ , а также слагаемые  $P_i^n$  и  $Q_i^{n+1}$  определяются следующими формулами:



$$\begin{aligned}
 G_i^n &= \sum_{k=1}^{N_2} V_k^n M_2^{-1} \sum_{j=1}^{M_2} \tilde{H}(i - J_2(k, j)), \\
 P_i^n &= \sum_{k=1}^{N_2} V_k^n M_2^{-1} \sum_{j=1}^{M_2} \left[ f_{J_2(k, j)}^n \tilde{H}(i - J_2(k, j)) \right] - \sum_{k=1}^{N_2} V_k^n L_k^n \sum_{l=1}^{M_2} \left[ \nabla_{kj}^n \tilde{H}(i - J_2(k, j)) \right] + \\
 &\quad + \sum_{k=1}^{N_2} V_k^n \frac{1}{M_2} \left( \sum_{l=1}^{M_2} A_{J_2(k, l)}^n \right) \sum_{j=1}^{M_2} \left[ \nabla_{kj}^n \tilde{H}(i - J_2(k, j)) \right] + \\
 &\quad + \sum_{k=1}^{N_3} S_k^n \left( B_{*k}^{n+1} - \Omega_k^n \sum_{l=1}^{M_3} \frac{A_{J_3(k, l)}^n}{M_3} \right) \frac{1}{M_3} \sum_{j=1}^{M_3} \left[ \tilde{H}(i - J_3(k, j)) \right], \\
 Q_i^n &= - \sum_{k=1}^{N_2} V_k^n (K + \text{var} I)_k^n \left( \sum_{l=1}^{M_2} \nabla_{kl}^n A_{J(k, l)}^{n+1} \right) \sum_{j=1}^{M_2} \left[ \nabla_{kj}^n \tilde{H}(i - J_2(k, j)) \right],
 \end{aligned}$$

где функция  $\tilde{H}$  равна единице, если аргумент равен нулю и равна нулю в противном случае.

Главные граничные условия, т. е. задание искомой функции в граничных узлах, учитываются при задании начального приближения, а затем в конце каждой итерации. Для этого достаточно в узлах с главными граничными условиями обнулить невязки, т. е. обнулить поправки к заданным граничным значениям искомой функции. Естественные граничные условия (ограничения на граничные потоки) учитываются в смысле обобщенного решения в самом вариационном уравнении. Так что по сравнению со многими другими численными методами учет граничных условий реализуется предельно просто.

В традиционных матричных методах реализации алгоритмов для вычисления коэффициентов матрицы системы уравнений  $S$  подобные члены пришлось бы собирать при общих множителях – произведениях  $A_i^{n+1} \delta A_j$ , да и память для запоминания коэффициентов матриц заняла бы значительный объем. Причем этот объем растет нелинейно в зависимости от числа пространственных переменных и количества дискретных неизвестных.

В рассмотренном варианте матрица  $S$  для определения невязок вообще не нужна и алгоритм расчета стал проще, нагляднее и экономнее как по памяти, так и по быстродействию. Работа по написанию и отладке программ значительно упрощается, поскольку текст программы по сути является записью исходных уравнений в дискретной форме. Алгоритмы становятся очень легко модифицируемыми, легко адаптируемыми под совершенно различные задачи механики сплошных сред.

Применение рассмотренных итерационных процессов экономит значительно не только машинные ресурсы, но, что особенно ценно, сокращает затраты рабочего времени вычислителей на разработку алгоритмов, на рутинные поиски ошибок программирования, неизбежных для сложных алгоритмов, и позволяет сосредоточиться на более интересной работе по исследованию решений широкого диапазона задач МСС.

### 2.3. Учет дополнительных условий (ограничений)

Дополнительные условия задач (ограничения), например, контактные условия, условия разрушения/ консолидации, условия фазовых переходов, условия несжимаемости и другие в рассматриваемом вариационном подходе без затруднений учитываются в сквозном счете (см. обзор [7]) путем модификации вариационных уравнений с применением методов множителей Лагранжа или штрафных функций (в разных вариантах). Такие модификации не связаны с необходимостью применения именно метода сопряженных градиентов, они хорошо известны, разработаны и широко применяются для вариационных формулировок задач МСС.

### 2.4. Расщепление по физическим процессам

Отметим, что на каждом шаге по времени система уравнений механики сплошных сред, как правило, расщепляется по физическим процессам и полное решение на новом слое получается в результате решения вспомогательных краевых задач по очереди для каждого балансного уравнения задачи (например, для уравнений движения, уравнения неразрывности, уравнения энергии, баланса примесей и так далее).

Нередко от решения вспомогательной краевой задачи не требуется большой точности. В этих случаях вместо метода сопряженных градиентов в неявных схемах можно успешно использовать еще более

простые итерационные алгоритмы метода установления путем искусственного превращения уравнений эллиптического типа в параболические, интегрируемые по явным схемам. Это достигается добавлением в исходное эллиптическое уравнение нестационарного члена с производными по фиктивному времени. Примерами таких вспомогательных задач могут служить, например, краевые задачи для уравнений управления подвижными адаптивными сетками (см. обзор [8]). Методом установления приближенное установившееся решение получается после заранее заданного (подбором) фиксированного числа итераций.

Корректность такого упрощения итерационных алгоритмов проверяется расчетами. Если повышение точности решений вспомогательных задач практически не влияет на решение основной задачи, то применение упрощенных алгоритмов оправдано.

### 2.5. Консервативность алгоритмов

Консервативность алгоритмов обеспечивается представленной выше проинтегрированной по частям вариационной формой Галеркина–Петрова применяемых законов сохранения и последующей конечно-элементной аппроксимацией решения. Локальная консервативность, т. е. баланс потоков между приузловыми объемами обосновывается элементарно и является следствием вариационной записи и факта равенства нулю производных от константы. Глобальная консервативность определяется корректным заданием граничных условий.

### 2.6. Задание искусственной вязкости

Применяемые схемы расчета по времени (физическому или фиктивному в итерациях) полностью аналогичны центрально-разностным двухслойным схемам. Поскольку конвекция рассчитывается по явной центрально-разностной схеме, то для устойчивости необходима искусственная вязкость. Среди множества рецептов введения искусственной вязкости (см. [6]) в настоящей работе после огромного множества тестов предпочтение было отдано способу стабилизированной конечно-элементной схемы Петрова–Галеркина с разностями против потока (см. [9], [10]). Однако схема эта используется в упрощенном варианте без введения направленных разностей. Искусственная вязкость определяется так, чтобы на границах между приузловыми объемами уравновесить по норме вклад от диффузионных членов и вклад от прочих членов однородной части уравнения. Поясним, что “на границах между приузловыми объемами” означает “в гауссовых точках используемых квадратур интегрирования вариационного уравнения”, расположенных в центрах конечных элементов. Такую упрощенную разновидность стабилизированной схемы Галеркина–Петрова называем *схемой уравновешивающей вязкости*. Формулы для вычисления коэффициентов искусственной вязкости  $v_{art}$  имеют вид (см. [11])

$$[v_{art}]_k = \alpha_k \frac{d_{1k}}{d_{2k}},$$

где  $k$  – номер конечного элемента, индекс временного слоя у всех переменных опущен, чтобы не загромождать записи. Параметр  $\alpha_k = 0.5$  при  $|\nabla \cdot \mathbf{u}|_k \Delta t < 0.1$ , иначе  $\alpha_k = 1.0$ . Шаг по времени  $\Delta t$  определяется далее условием устойчивости. Коэффициенты  $d_{1k}^{(j)}$  и  $d_{2k}^{(j)}$ ,  $j = 1, 2, 3$ , определяются в виде

$$(d_{1k})^2 = (|\nabla \cdot (A\Omega)|_k)^2, \quad (d_{2k})^2 = \sum_{l=1}^M (|\nabla A|_k \cdot \nabla_{kl})^2.$$

Свойство коэффициентов пространственного дифференцирования  $\nabla_{kl} \left( \sum_{l=1}^M \nabla_{kl} = 0 \right)$  определяет равенство нулю производной от константы и обеспечивает консервативность обмена сохраняемой величины между приузловыми объемами. В зонах сильного разрежения, вызванного ускорением потока при  $|\nabla \cdot \mathbf{u}|_k \Delta t \geq 0.1$ , коэффициент уравновешивающей вязкости специально увеличивается вдвое, что необходимо для устойчивости расчета на больших временах.

### 2.7. Корректировка физической вязкости

Для повышения точности решения в пограничных слоях физическая вязкость  $k$  корректируется (уменьшается с ростом искусственной вязкости) по методу экспоненциальной подгонки, предложенному в 60-х гг. А.А. Самарским и развитому в [12]. Простейшая корректировка имеет вид

$$\tilde{k} = k^2 / (k + v_{art}).$$

### 2.8. Устранение мелкомасштабных паразитных возмущений

Поскольку все элементы алгоритма обосновываются только на уровне простейших модельных задач и не строго, то в решении задач по полным уравнениям в области решения местами возникают паразитные возмущения решения с длиной волны, пропорциональной шагу пространственной сетки. Такие нефизические возмущения (“пила”) немедленно отлавливаются по смене знака второй производной по проверяемому координатному направлению в соседних узлах, принадлежащих ребру сетки. Взаимная ориентация ребра и координатного направления роли не играет. Такие возмущения хотя и не нарушают устойчивость метода (не растут), но их присутствие портит графики решения (дает негладкие картины изолиний). Для устранения этих возмущений применяется простейшее консервативное покоординатное сглаживание. Реально эта процедура проводится в крайне малом количестве узлов в конце каждого шага.

### 2.9. Определение вторых производных

Определение вторых производных  $A_{xx}$  от некоторой компоненты решения  $w$  по координате  $x$  в условиях кусочно-линейных аппроксимаций решения реализуется так. Очевидное вариационное уравнение для вторых производных

$$\int_V (A_{xx} - \partial^2 A / \partial x^2) \delta A_{xx} dV = 0$$

интегрированием по частям приводится к виду

$$\int_V A_{xx} \delta A_{xx} dV + \int_V \partial A / \partial x \partial A_{xx} / \partial x dV = \int_S \partial A / \partial x \delta A_{xx} n_x dV,$$

где область решения  $V$  имеет границу  $S$ . Правая часть уравнения для простоты полагается равной нулю, т. е. на границе либо первая, либо вторая производная по координате  $x$  считается равной нулю. Если для первого интеграла левой части использовать квадратуры Гаусса с точками интегрирования в узлах сетки, то вторые производные определяются вариационным уравнением явно, для их определения никаких недиагональных матриц обращать не требуется. Таким способом вторые производные быстро вычисляются по кусочно-линейным функциям приближенного решения. Вторые производные по координатам  $y$  и  $z$  вычисляются аналогично.

### 2.10. Условия устойчивости

Для явно- неявных схем шаг по времени  $\Delta t$  определяется из условия устойчивости, которое имеет обычный вид (напомним, что явная аппроксимация по времени использована только для конвекции со скоростью  $\Omega_k$ )

$$\Delta t \leq \beta \min_k \left( \frac{h_k}{|\Omega_k| + 10^{-6}}, \frac{10^{-1} \epsilon_S}{\|\mathbf{e}_k\| + 10^{-6}} \right),$$

где  $h_k$  – размер  $k$ -го конечного элемента,  $\beta$  – коэффициент запаса (обычно равный 0.3–0.6),  $\mathbf{e}$  – тензор скорости деформации,  $\epsilon_S$  – деформация, соответствующая пределу текучести или прочности. Ограничение по конвекции (слагаемое  $|\Omega_k|$  в знаменателе условия устойчивости) обусловлено явной аппроксимацией конвективных потоков. Ограничение по точности (второе ограничение) необходимо в задачах механики деформируемых твердых тел и выражает требование малости нормы приращения деформаций на шаге по времени.

Важно отметить, что явно-неявные схемы можно успешно применять к задачам о быстрых процессах, для которых ограничение шага по времени по условиям точности практически совпадает с ограничениями на шаг по времени по условиям устойчивости в явных схемах

$$\Delta t \leq \beta \min_k \left( \frac{h_k}{|\Omega_k| + c + 10^{-6}} \right),$$

где  $c$  – скорость звука (скорость распространения малых возмущений). В таких задачах за счет того, что решение на предыдущем временном слое является очень хорошим начальным приближением к решению на новом слое, число итераций метода сопряженных градиентов резко падает до двух-трех. То есть неявные схемы с методом сопряженных градиентов автоматически начинают работать так же быстро, как явные схемы. Малые числа, добавленные в знаменатели условий устойчивости, предохраняют от деления на ноль.



### 2.11. О проблемах, которые исчезают при использовании безматричных алгоритмов

Отметим дополнительные преимущества применения безматричных итераций: 1) не надо помнить ненулевые элементы матрицы системы уравнений и их расположение; 2) не надо вообще вычислять элементы матрицы системы уравнений; 3) не надо искать оптимальную нумерацию узлов сетки, обеспечивающую минимальность ширины ленты матрицы системы уравнений. Указанные преимущества присущи любым безматричным итерационным методам (методу простой итерации, методу Гаусса–Зейделя, методам релаксации и так далее), но только методы сопряженных градиентов гарантируют конечность числа итераций для определения решения.

## 3. РЕЗУЛЬТАТЫ

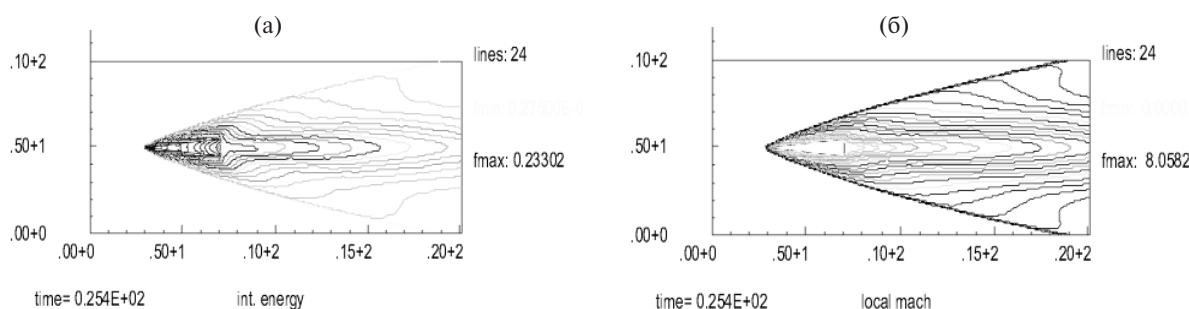
Ниже приводятся решения типичных задач, полученные с использованием явно-неявных алгоритмов итерациями по методу сопряженных градиентов.

### 3.1. Сверхзвуковые течения идеального газа с ударными волнами с помощью подвижных адаптивных сеток

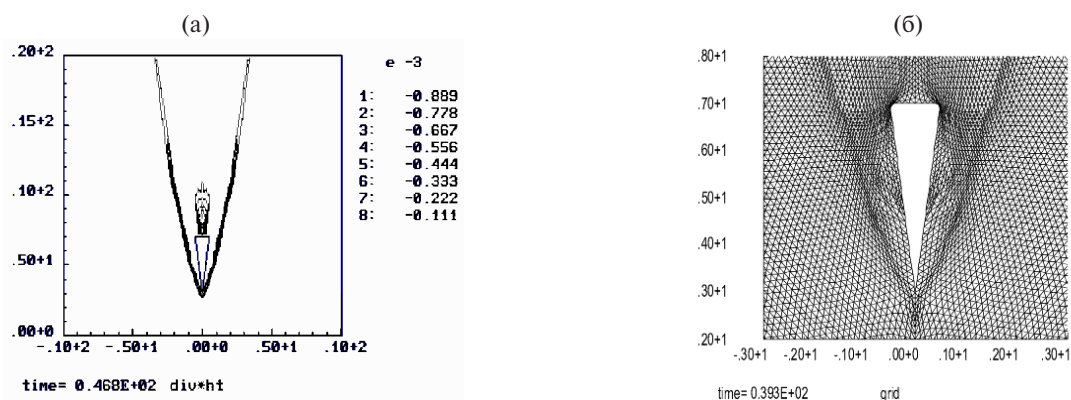
На фиг. 1 показаны результаты для сверхзвукового обтекания клина (плоская задача) с показателем адиабаты 1.4, числом Маха в набегающем потоке  $M = 8$ . На фиг. 2 даны результаты для сверхзвукового обтекания конуса (осесимметричная задача) при  $M = 134$ , а именно показана управляющая адаптацией сетки мониторинговая функция  $\tilde{T} = -\Delta \Gamma \nabla \cdot \mathbf{v}$  и фрагмент адаптивной сетки около конуса.

Конечно, в диапазоне гиперзвуковых течений ( $M > 8$ ) постановка задач должна быть существенно расширена для учета множества физико-химических эффектов высокотемпературной плазмы. Здесь показана только потенциальная пригодность алгоритма к задачам такого типа, а число  $M = 134$  является случайным. Чтобы повысить число Маха, температура в набегающем потоке была задана взятым “с потолка” маленьким числом.

В задаче об обтекании клина сам клин задавался методом наложенных сеток (см. [11]). Расчет проводился на основной (окаймляющей) регулярной сетке, состоящей из первоначально прямоугольных ячеек (элементов) в прямоугольной области течения. Клин имел свою (наложенную) сетку. В узлах основной сетки, накрытых сеткой клина, скорости занулялись. При адаптации сетки накрытые узлы основной



Фиг. 1. Обтекание клина  $M = 8$ : а – внутренняя энергия, б – локальное число Маха.



Фиг. 2. Обтекание конуса при  $M = 134$ : а – мониторинговая функция, б – фрагмент сетки.

сетки не двигались (клин неподвижен). Вариационная постановка задачи и конечно-элементная аппроксимация решения в сквозном счете автоматически аппроксимируют необходимые условия на границах наложенной области – равенство нормальных к границе производных от давления и температуры.

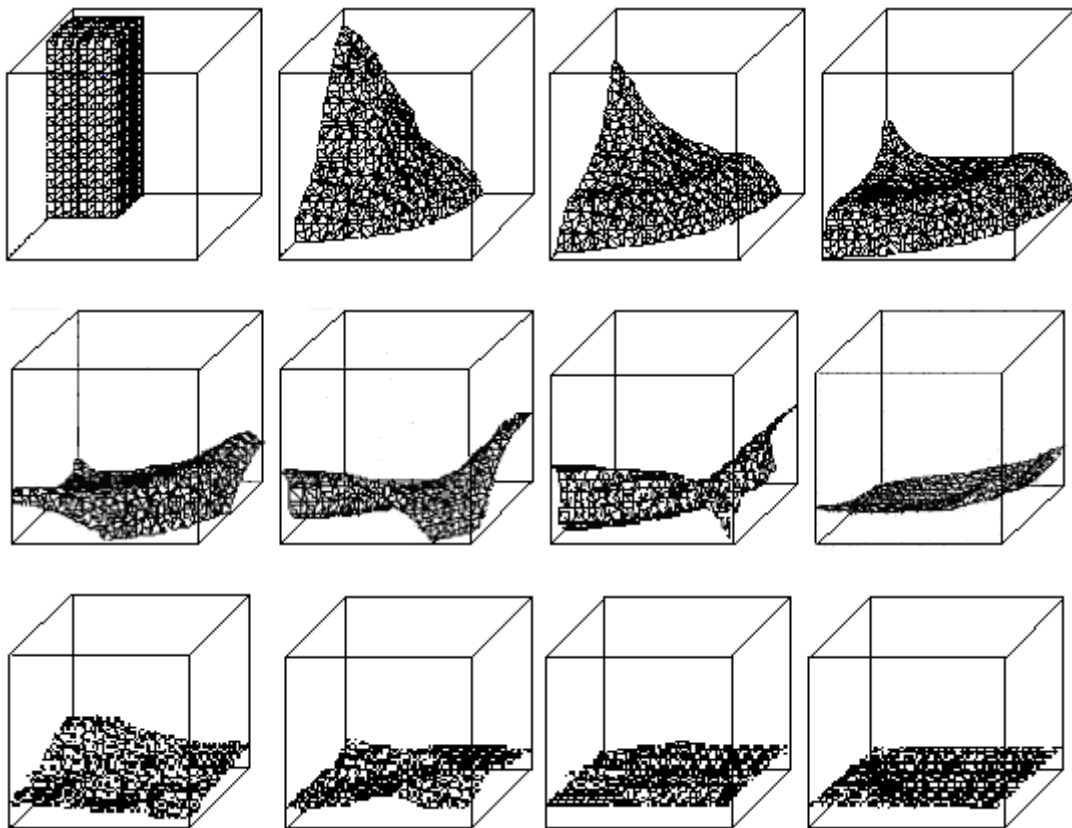
Если требуется задание граничных условий другого вида, например, задание самих функций или смешанных граничных условий, то такие условия легко учитываются как дополнительные ограничения модификацией вариационных балансных уравнений по методу штрафных функций или множителей Лагранжа.

### 3.2. Обрушение колонны тяжелой вязкой жидкости в замкнутом бассейне

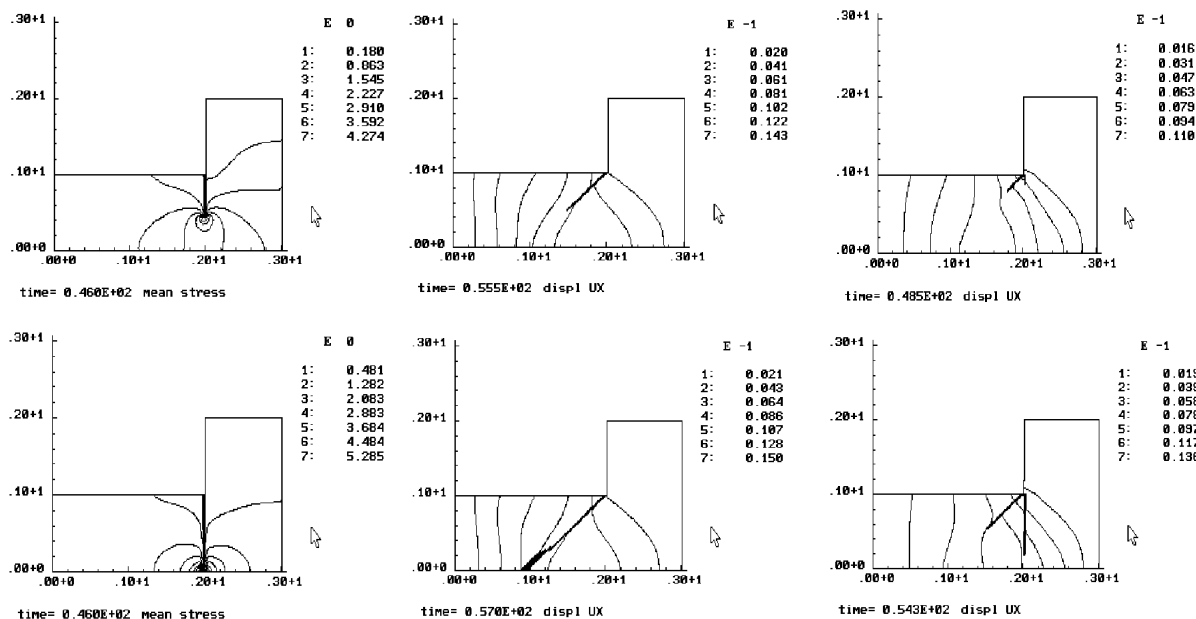
В начальный момент времени колонна тяжелой вязкой жидкости в замкнутом кубическом бассейне пребывает в состоянии покоя. Под действием силы тяжести, направленной вниз, вода приходит в движение и колонна разрушается (см. фиг. 3). Колебания воды в бассейне под действием сил вязкого трения затухают и жидкость занимает новое положение равновесия. Расчет проводился сквозным счетом с использованием непрерывной маркер-функции, принимающей значения 1 в ячейках, заполненных жидкостью, и 0 в пустых ячейках. Маркер-функция рассчитывалась по уравнению переноса. Расчет проводился в переменных скорость-давление, причем и скорость, и давление определялись в узлах сетки, и применялась кусочно-линейная аппроксимация решения так, что LBB-условие не соблюдалось, что, однако, не помешало решению задачи. Консервативность по массе не нарушена. При использовании непрерывных маркер-функций для этого необходима монотонность метода расчета и анти-диффузия маркер-функции в зоне свободной границы.

### 3.3. Разрушение упругопластической пластинки при нагреве и растяжении

На фиг. 4 показана четвертая часть плоского упругопластического образца. Левая и нижняя границы являются осями симметрии. Вся верхняя граница свободна от внешних нагрузок. Правая граница медленно смещается вправо, обеспечивая растяжение образца. Угол излома верхней границы создает очаг концентрации напряжений и деформаций и разрушение начинается от этого угла. Показана эволюция распространения трещин в зависимости от учета только упругости (левые рисунки), упругости и пласт-



Фиг. 3. Обрушение колонны тяжелой вязкой жидкости в замкнутом бассейне.

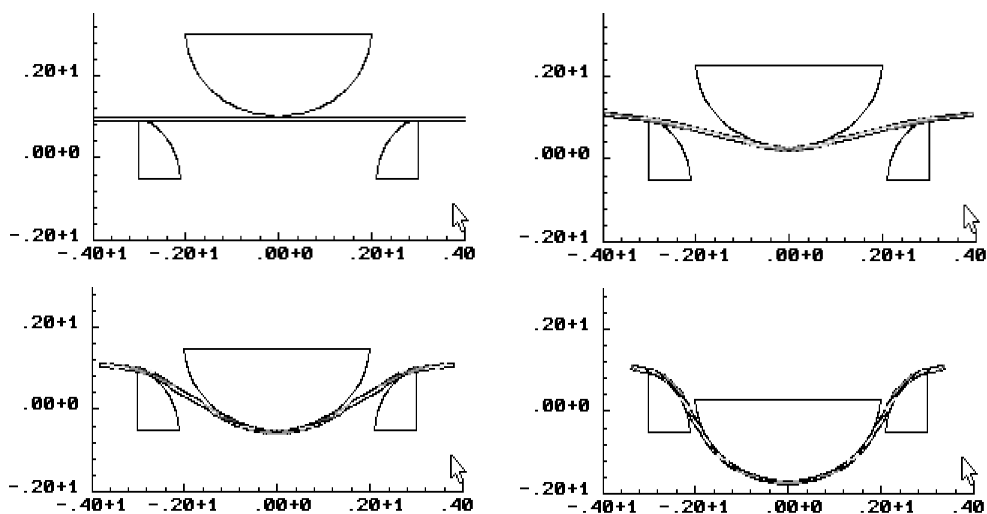


Фиг. 4. Разрушение упругопластической пластины при растяжении и нагреве.

точности (средние рисунки) и при дополнительном учете локального нагрева вертикальной зоны от излома верхней границы (правые рисунки). Картины разрушения сохраняются при измельчении сетки. Результаты показывают совершенно разные картины разрушения в зависимости от свойств материала и воздействия внешнего нагрева. Использовался метод сквозного счета, т. е. и целый, и разрушенный материал рассчитывались по единому алгоритму. Модули упругости материала и предел текучести зависели от поврежденности материала. Поврежденность определялась по кинетическому уравнению при выполнении условия разрушения (см. [13]), которое включалось по достижении максимальной растягивающей деформацией определенного предела.

### 3.4. Штамповка алюминиевой чашки

На фиг. 5 показаны результаты расчета штамповки алюминиевой чашки. Нижний абсолютно жесткий штамп неподвижен. Верхний штамп движется вниз, превращая круглую алюминиевую пластину в чашку. Время процесса многократно превышает время пробега упругих волн по радиусу пластины так,



Фиг. 5. Штамповка алюминиевой чашки.

что процесс является квазистатическим и рассчитывается по неявной схеме с использованием безматричного метода сопряженных градиентов. Расчет контактного взаимодействия выполнялся методом множителей Лагранжа, роль которых отводилась нормальным контактными нагрузкам. Трение на контакте считалось пренебрежимо малым.

#### 4. ВЫВОДЫ

Дан обзор и подробное описание простого и экономичного по памяти и быстродействию способа реализации явно-неявных схем с помощью безматричного метода сопряженных градиентов. Представлены типичные примеры применения метода в неявных схемах для задач газовой динамики (расчет адаптивной подвижной сетки по неявной схеме), механики вязкой несжимаемой жидкости (неявная схема для вязких членов и учета несжимаемости) и механики твердых деформируемых тел (неявная схема расчета контактного взаимодействия и процесса разрушения упругопластических тел). Используемые в расчетах методы известны. Новым элементом в алгоритмах является их безматричная реализация, приводящая к значительному упрощению реализации.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Hestenes M.R., Stiefel E.* Method of Conjugate Gradients for Solving Linear Systems // NBS J. 1952. Res. 49. P. 409–436.
2. *Полак Э.* Численные методы оптимизации. Единый подход. М.: Мир, 1974.
3. *Уилкинсон, Райнш.* Алгоритмы линейной алгебры на языке Фортран. М.: Мир, 1976.
4. *Бураго Н.Г.* Формулировка основных уравнений механики сплошной среды в подвижных адаптивных координатах // В книге Численные методы в механике твердого деформируемого тела // Под ред. Г.И. Пшеничнова. М.: ВЦ АН СССР, 1984. С. 32–49.
5. *Стренг Г., Фикс Дж.* Теория метода конечных элементов. М.: Мир, 1977.
6. *Куликовский А.Г., Погорелов Н.В., Семенов А.Ю.* Математические вопросы численного решения гиперболических систем уравнений. М.: ФИЗМАТЛИТ, 2001.
7. *Бураго Н.Г., Кукуджанов В.Н.* Обзор контактных алгоритмов // Изв. РАН. Механ. твердого тела. 2005. № 1. С. 44–85.
8. *Лисейкин В.Д.* Обзор методов построения структурных адаптивных сеток // Ж. вычисл. матем. и матем. физ. 2003. Т. 36. № 1. С. 3–41.
9. *Brooks A.N., Hughes T.J.R.* Streamline Upwind Petrov–Galerkin formulation for convection dominated flows // Comput. Meth. in Appl. Mechan. and Eng. 1982. V. 32. P. 199–259.
10. *Le Beau G.J., Ray S. E., Aliabadi S.K., Tezduyar T.-E.* SUPG finite element computation of compressible flows with the entropy and conservation variables formulations // Comput. Meth. in Appl. Mechan. and Eng. 1993. V. 104. P. 397–422.
11. *Бураго Н.Г., Никитин И.С., Якушев В.Л.* Гибридный численный метод решения нестационарных задач механики сплошной среды с применением адаптивных наложенных сеток // Ж. вычисл. матем. и матем. физ. 2016. Т. 56. № 6. С. 1082–1092.
12. *Дулан Э., Миллер Дж., Шилдерс У.* Равномерные численные методы решения задач с пограничным слоем. М.: Мир, 1983.
13. *Бураго Н.Г., Глушко А.И., Ковшов А.Н.* Термодинамический метод вывода определяющих соотношений для моделей сплошных сред // Изв. РАН. Механ. твердого тела. 2000. № 6. С. 4–15.