ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2024, том 64, № 6, с. 895–913

ОБЩИЕ ЧИСЛЕННЫЕ МЕТОДЫ

УДК 519.6

РАЦИОНАЛЬНАЯ АРИФМЕТИКА С ОКРУГЛЕНИЕМ

© 2024 г. В. П. Варин[1,*]

[1] *125047 Москва, Миусская пл., 4, Институт прикладной математики им. М. В. Келдыша РАН, Россия*

[*] *e-mail: varin@keldysh.ru*

Поступила в редакцию 16.01.2024 г.
Переработанный вариант 08.02.2024 г.
Принята к публикации 05.03.2024 г.

Вычисления на компьютере в плавающей арифметике всегда являются приближенными. Напротив, вычисления в рациональной арифметике (например, в компьютерной алгебре) всегда абсолютно точны и воспроизводимы как на других компьютерах, так и (теоретически) вручную. Поэтому такие вычисления могут быть доказательными в том смысле, что доказательство, полученное с их помощью, ничем не отличается от традиционного. Однако обычно такие вычисления в достаточно сложной задаче невозможны ввиду ограниченности ресурсов памяти и времени. Мы предлагаем механизм округления рациональных чисел при расчетах в рациональной арифметике, который решает эту проблему (ресурсов), т.е. вычисления по-прежнему могут быть доказательными, но уже не требуют неограниченных ресурсов. Приведен ряд примеров реализации стандартных численных алгоритмов в этой арифметике. Результаты имеют приложения к аналитической теории чисел. Библ. 22. Фиг. 3.

Ключевые слова: рациональная арифметика, подходящие дроби, доказательные вычисления, критерий иррациональности Бруна.

DOI: 10.31857/S0044466924060015, **EDN:** XZRESQ

1. ВВЕДЕНИЕ

Вычисления на компьютере в плавающей арифметике всегда являются приближенными, так же как и вычисления вручную с карандашом и бумагой. Однако помимо разной производительности этих вычислений между ними существует одно принципиальное отличие. Вычисления вручную контролирует сам вычислитель, т.е. человек определяет, сколько цифр сохранить в промежуточных и окончательных результатах и как при этом происходит округление. В компьютере же за эти операции отвечает, помимо прочего, тип данных, выбранных в программе, тип операционной системы, а также тип процессора, установленного на данном компьютере.

Хотя вся эта информация, в принципе, доступна пользователю, однако практическое ее использование затруднительно даже в одной элементарной операции типа $x = y + z$. Кроме того, на другом компьютере все это будет уже несколько по-другому.

Иными словами, вычисления на компьютере – это всегда обращение к “черному ящику”, “размер” которого зависит от желания и возможности пользователя изучать то, как конкретный компьютер работает на физическом уровне. В этом смысле вычисления на компьютере вполне аналогичны проведению физического эксперимента.

Вероятно, в том числе и эти соображения учитывались при создании математической дисциплины “экспериментальная математика” (см., например, [1]). Заметим также, что В. И. Арнольд полагал, что “математика – это экспериментальная наука” (см. [2]).

В то же время вычисления на компьютере могут быть абсолютно точными, в случае если это символьные вычисления в системе компьютерной алгебры (CAS). Символьные вычисления, проделанные с помощью компьютера, могут быть при этом чрезвычайно сложны и громоздки, однако они, в принципе, воспроизводимы в разных системах CAS и проверяемы вручную.

В этом смысле доказательство, полученное с помощью компьютера (computer assisted proof), ничем не отличается от традиционного и признается как таковое.

Хотя сам термин “доказательные вычисления” исторически появился значительно раньше, однако он не получил (и вряд ли получит) столь же широкое признание.

По нашему мнению, вычисления в плавающей арифметике в принципе не могут быть доказательными потому, что они не проверяемы вручную, так как человеку (как правило) неизвестен алгоритм, по которому округляются промежуточные вычисления и результат, а также потому, что на разных компьютерах это делается по-разному и дает (вообще говоря) различные результаты.

При этом с прикладной точки зрения результат может не вызывать ни малейших сомнений, как и результат физического эксперимента, однако математически это не будет доказательством.

Все это относится и к интервальной арифметике, которая одно время также претендовала на доказательность. Однако в настоящее время в контексте интервальной арифметики говорят, скорее, о достоверных вычислениях (см. [3]).

Таким образом, доказательность или бездоказательность вычислений зависят не только от качества алгоритмов и полученных решений, но и от введения в математику новых аксиом, которые будут относиться к компьютеру, и поэтому никогда не будут приняты большинством математиков.

Однако вычисления на компьютере могут быть доказательными в традиционном смысле, если они абсолютно точны, воспроизводимы на других компьютерах независимо от их архитектуры, и, главное, теоретически проверяемы вручную.

Мы привели символьные вычисления в CAS как один из примеров таких вычислений. Другой пример дают вычисления в рациональной арифметике (необязательно в CAS), которые эквивалентны вычислениям с целыми числами. Такие вычисления всегда осуществляются абсолютно точно на любом компьютере, а также всегда (теоретически) проверяемы вручную. Поэтому такие вычисления полностью укладываются в понятие “математически строгое доказательство”, если оно было получено с их помощью.

Одним из ярких примеров такого доказательства является проблема четырех красок, которая была решена с помощью вычислений на компьютере (см. [4]).

Разумеется, абсолютно точные вычисления на компьютере можно осуществлять не только в рациональной арифметике, но и, например, в квадратичных полях $\mathbb{Q}(\sqrt{n})$, $n \in \mathbb{Z}$. Такие вычисления применяются, в частности, в вычислительной геометрии (см. [5]).

При вычислениях в рациональной арифметике возникают те же проблемы с ресурсами памяти и времени, которые существуют при вычислениях в плавающей арифметике с расширенной разрядной сеткой. Теоретически вычисления в плавающей арифметике можно проводить с любым числом десятичных разрядов. Например, вполне реально вычислять с тысячами десятичных разрядов на домашнем компьютере. Однако практически ресурсы всегда ограничены.

Вычисления в рациональной арифметике, как правило, предъявляют еще бóльшие требования к ресурсам памяти и времени, чем вычисления в плавающей арифметике. К тому же, в отличие от последней, вычисления с рациональными числами значительно менее предсказуемы в плане требуемых ресурсов.

Например, итерационный алгоритм Айткена при вычислении константы Эйлера—Гомпертца уже на 10-м шаге дает приближение этой константы с погрешностью менее $1.2 \cdot 10^{-10}$. Однако суммарное количество цифр в полученном рациональном приближении равно 180980 (см. [6]).

Один из способов если не решить, то контролировать эту проблему в CAS — использование модулярной арифметики, которая широко применяется, например, в базисах Грёбнера и в алгоритмах факторизации полиномов с целыми коэффициентами.

Модулярная арифметика в таких вычислениях (с целыми числами) является аналогом вычислений в плавающей арифметике с фиксированной разрядной сеткой. Однако, в отличие от последней, вычисления в модулярной арифметике по-прежнему абсолютно точны и воспроизводимы как на других компьютерах, так и (теоретически) человеком вручную, и поэтому являются доказательными в указанном выше смысле.

Из сказанного выше следует, что сделать вычисления доказательными, просто заменив плавающую арифметику на рациональную, не получится, в силу того, что вычисления проводятся на реальном, а не на идеальном компьютере с неограниченными ресурсами.

В данной статье предлагается механизм округления рациональных чисел, который позволяет использовать рациональную арифметику в обычных вычислениях, например, для численного интегрирования ОДУ, в задачах линейной алгебры и т.п.

Вычисления в рациональной арифметике на компьютере (или вручную) эквивалентны вычислениям с целыми числами, поэтому "округленный" результат, если он получается детерминистским и предсказуемым образом, всегда (теоретически) проверяем вручную.

В качестве такого механизма округления рационального числа r (или вещественного числа x) предлагается использовать обыкновенные подходящие дроби этого числа, которые мы обозначаем через $K(r, n)$, $n \in \mathbb{N}$, или:

$$K(x, n) = a_0 + \frac{1}{a_1+}\ \frac{1}{a_2+}\ \ldots\ \frac{1}{a_n} = [a_0;\, a_1, \ldots, a_n], \quad x \in \mathbb{R}.$$

Уже сама эта запись является аналогом записи числа в виде q-адичной дроби с фиксированной разрядной сеткой. Однако здесь "цифры" a_k, $k \in \mathbb{N}$ – это натуральные числа для $k \geq 1$ и $a_0 \in \mathbb{Z}$.

В учебниках, где рассматриваются свойства цепных дробей, насколько нам известно, никогда не отмечается отличие вычислений с рациональными числами от вычислений с вещественными числами в плавающей арифметике. Между тем, если вычисления проводятся на реальном, а не на идеальном компьютере с неограниченной разрядной сеткой, алгоритм Евклида работает принципиально отличным образом для этих арифметик.

Алгоритм Евклида вычисления обыкновенной подходящей дроби рационального числа (в отличие от числа, заданного в плавающей арифметике) всегда корректно определен и всегда дает один и тот же результат на любом компьютере независимо от его архитектуры (или вручную).

Таким образом, приближенные вычисления в рациональных числах можно рассматривать, в определенном смысле, как абсолютно точные. Поэтому такие вычисления, по нашему мнению, укладываются в концепцию доказательных вычислений.

В разд. 2 мы приводим некоторые свойства функций $K(x, n)$, $x \in \mathbb{R}$, включая их аппроксимационные возможности. Оказывается, что функции $K(x, n)$, $x \in [0, 1]$, весьма быстро стремятся к x при $n \to \infty$, причем максимальное отклонение $|x_n - K(x_n, n)|$ достигается в рациональных x_n, которые стремятся к золотому сечению.

В разд. 3 приводятся примеры численных алгоритмов, реализованных с помощью этой арифметики. Показано, что такие вычисления могут давать хорошие рациональные приближения нужных величин при использовании стандартных численных методов.

В разд. 4 рассматриваются некоторые приложения этой арифметики к аналитической теории чисел.

2. ЦЕПНАЯ ДРОБЬ $K(x, n)$ КАК ФУНКЦИЯ СВОИХ АРГУМЕНТОВ

Напомним, что функции $K(x, n)$ мы определили в [6]. Они использовались для построения рациональных аппроксимаций некоторых фундаментальных постоянных (см. [6, 7]).

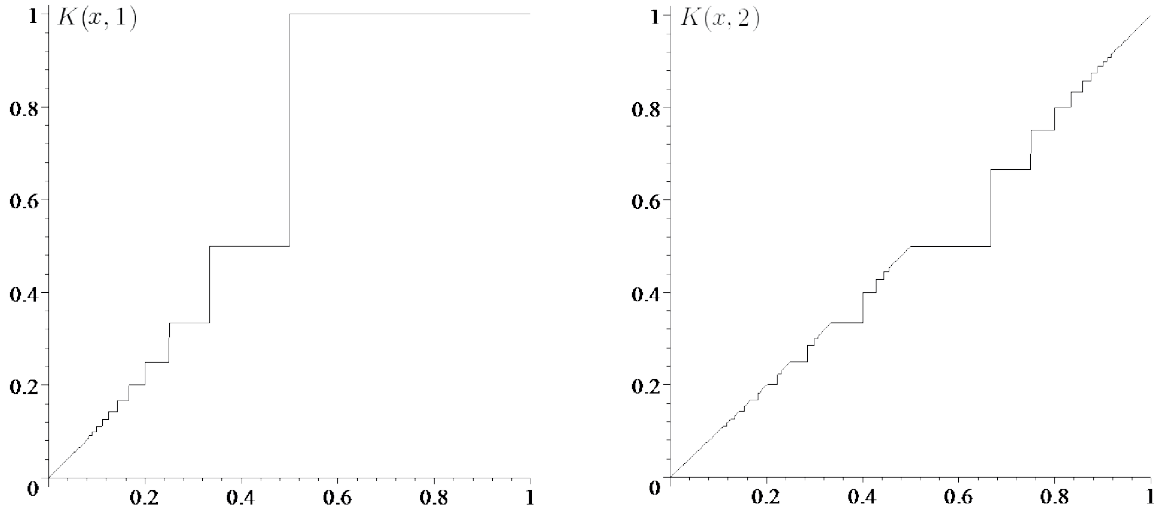
Хотя аппроксимационные свойства обыкновенных цепных дробей хорошо изучены для фиксированных x (см., например, [8]), свойства цепных дробей как функций аргумента x и порядка дроби n, насколько нам известно, нигде не изучались.

Функции $P(x, n) = K(x, n) - x$ являются, очевидно, 1-периодическими, причем

$$P(x, 2n) \leq 0, \quad P(x, 2n - 1) \geq 0,$$

по известному свойству подходящих дробей. Поэтому достаточно изучать эти функции на интервале $[0, 1]$.

На фиг. 1 приведены графики функций $K(x,1)$ и $K(x,2)$ на интервале $[0,1]$, где вертикальные отрезки – это чисто технический артефакт (удалять их было бы слишком накладно). Из фиг. 1 видно, что график функции $K(x,2)$ на отрезках $[1/2,1]$, $[1/3,1/2]$, $[1/4,1/3]$, … повторяет масштабированный график функции $K(x,1)$. Так же ведут себя графики функций $K(x,n)$, $n>2$. Таким образом, графики этих функций имеют весьма сложную квазифрактальную структуру.



Фиг. 1. Графики функций $K(x,1)$ и $K(x,2)$.

Поведение функций $K(x,n)$ на интервале $x\in[0,1]$ описывает следующее (очевидное)

Предложение 1. *Все точки разрыва функции $K(x,n)$, $n\in\mathbb{N}$, $x\in[0,1]$, даются формулой*

$$x_{k_1,\dots,k_n}=[k_1,k_2,\dots,k_n],\quad k_2,\dots,k_n\in\mathbb{N}_0,\ k_1\in\mathbb{N}. \tag{1}$$

При этом $K(x,n)=x_{k_1,\dots,k_n}$ для x, расположенных между величинами $[k_1,k_2,\dots,k_n]$ и $[k_1,k_2,\dots,k_n+t]$, $0\le t<1$, $k_1,\dots,k_n\in\mathbb{N}$.

Заметим, что выбор $k_j=0$, $j\ne 1$, в (1) означает просто, что особенности функций $K(x,m)$ наследуются функцией $K(x,n)$, $m<n$, что видно также из очевидного тождества

$$K(x,m)=K(K(x,m),n),\quad m,n\in\mathbb{N},\ m\le n.$$

Последовательности подходящих дробей $K(x,2n)$ и $K(x,2n-1)$, $n\in\mathbb{N}$, являются, соответственно, возрастающей и убывающей, т.е.

$$\dots\le K(x,2n-2)\le K(x,2n)\le x\le K(x,2n-1)\le K(x,2n-3)\le\dots,$$

причем равенства начиная с номера m здесь достигаются только для рациональных x, которые представимы в виде обыкновенной цепной дроби порядка $m\le 2n$.

Таким образом, у нас всегда имеются верхняя и нижняя оценки округляемого числа, что делает этот механизм округления схожим с интервальной арифметикой.

Другим преимуществом такого округления, т.е. замены (рационального) числа x на его подходящую дробь $K(x,n)$, является оптимальность этого представления. Как известно (см. [8]), дробь $K(x,n)$ приближает число x наилучшим образом в множестве рациональных чисел, знаменатели которых не превосходят знаменателя дроби $K(x,n)$ (а она всегда неприводима).

Пользуясь тождеством $x=1/([1/x]+\{1/x\})$, любую функцию $K(x,n)$, $0\le x\le 1$, можно выразить явно в виде композиций функций целой и дробной частей числа x. Однако это представление не имеет практического значения, так как реально вычисления проводятся по алгоритму Евклида.

Предложение 1 дает возможность записать интегралы от функций $K(x,n)$ на интервале $[0,1]$ в виде мультисумм с индексами $\{k_1,\dots,k_n\}$. Однако в явном виде такая сумма имеется только для $K(x,1)$:

$$\int_0^1 K(x,1)\,dx = \int_0^1 1/[1/x]\,dx = \frac{\pi^2}{6} - 1.$$

Согласно предложению 1, скачок, например, функции $K(x,2)$ в произвольной точке ее разрыва равен

$$\frac{1}{n+\dfrac{1}{m+1}} - \frac{1}{n+\dfrac{1}{m}} = \frac{1}{(n\,m+n+1)\,(n\,m+1)}, \quad m,n\in\mathbb{N},$$

т.е. m и n попадают в знаменатель, а числитель равен единице. Скачок любой функции $K(x,n)$, $n\in\mathbb{N}$, будет выражаться похожим образом по известному свойству подходящих дробей.

Предложение 2. *Наибольший скачок функции $K(x,n)$, $n\in\mathbb{N}$, $x\in[0,1]$, достигается в точке*

$$r_n = \frac{1}{1+}\,\frac{1}{1+}\,\cdots\,\frac{1}{1} = \mathop{\mathbf{K}}_{k=1}^{n+1}\frac{1}{1}.$$

Доказательство. Как известно, числители p_n и знаменатели q_n подходящих дробей удовлетворяют рекуррентным соотношениям

$$\begin{aligned} p_n &= k_n\,p_{n-1} + p_{n-2}, \quad p_{-1}=1,\ p_0=0,\\ q_n &= k_n\,q_{n-1} + q_{n-2}, \quad q_{-1}=0,\ q_0=1. \end{aligned} \tag{2}$$

Скачок функции $K(x,n)$ можно записать в виде

$$\frac{(k_n+1)\,p_{n-1}+p_{n-2}}{(k_n+1)\,q_{n-1}+q_{n-2}} - \frac{k_n\,p_{n-1}+p_{n-2}}{k_n\,q_{n-1}+q_{n-2}} = \frac{p_{n-1}\,q_{n-2}-p_{n-2}\,q_{n-1}}{((k_n+1)\,q_{n-1}+q_{n-2})\,(k_n\,q_{n-1}+q_{n-2})}.$$

Поэтому (см. теорему 2 в [8]) числитель последней дроби всегда равен ±1 согласно тождеству

$$\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n q_{n+1}}. \tag{3}$$

Знаменатель дроби, выражающей скачок функции $K(x,n)$, очевидно, достигает минимального значения при $k_1=k_2=\cdots=k_n=1$. Что и требовалось доказать.

Таким образом, наибольший скачок функций $K(x,n)$ достигается в дробях

$$\{r_n,\ n\in\mathbb{N}\} = \left\{\frac{1}{2},\frac{2}{3},\frac{3}{5},\frac{5}{8},\frac{8}{13},\frac{13}{21},\frac{21}{34},\frac{34}{55},\dots\right\},\ r_n\to\frac{1}{2}\,(\sqrt{5}-1)\approx 0.61803,$$

где числители и знаменатели — это не что иное, как числа Фибоначчи, F_n (см. [9]), так как формула (2) дает именно эти числа при $k_n=1$. Величина скачка при этом равна

$$\delta_n = \mathop{\mathbf{K}}_{k=1}^{n}\frac{1}{1} - \mathop{\mathbf{K}}_{k=1}^{n-1}\frac{1}{1+}\,\frac{1}{2} = \mathop{\mathbf{K}}_{k=1}^{n}\frac{1}{1} - \mathop{\mathbf{K}}_{k=1}^{n+1}\frac{1}{1} = \frac{F_n}{F_{n+1}} - \frac{F_{n+1}}{F_{n+2}}, \quad n\in\mathbb{N}.$$

По известному свойству чисел Фибоначчи, которое, впрочем, следует из формулы (3), получаем

$$\left\{\delta_n = \frac{(-1)^{n+1}}{F_{n+1}\,F_{n+2}},\ n\in\mathbb{N}\right\} = \left\{\frac{1}{2},-\frac{1}{6},\frac{1}{15},-\frac{1}{40},\frac{1}{104},-\frac{1}{273},\dots\right\}.$$

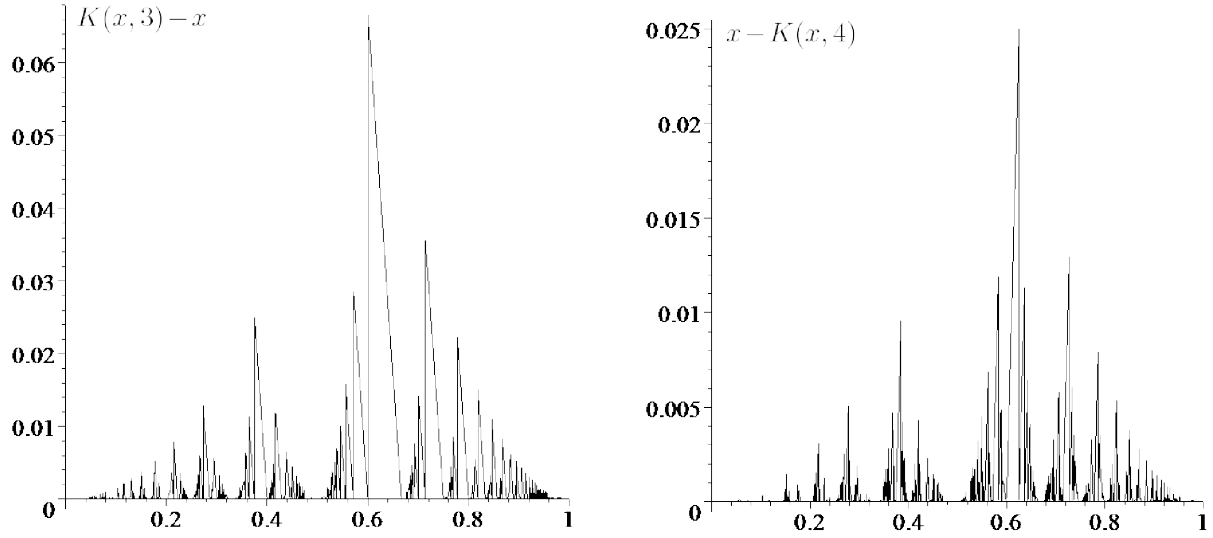
Величина $|\delta_n|$, очевидно, является максимально возможной погрешностью приближения числа $x\in[0,1]$ дробью $K(x,n)$. Например, при $n=40$ максимальная погрешность равна

$$|\delta_{40}| = \frac{1}{44361286907595736} \approx 0.225421\times10^{-16}.$$

Реальная погрешность, разумеется, может быть значительно меньше.

Таким образом, мы подтвердили известный факт: золотое сечение (или его сдвиги на целое число) приближается рациональными числами наихудшим образом.

На фиг. 2 приведены графики функций $K(x,3)-x$ и $x-K(x,4)$ на интервале $[0,1]$ (где вертикальные отрезки — это технический артефакт). Из этих рисунков (а также других для $n>4$) видно, что существуют промежутки, где вещественные числа приближаются цепными дробями в численном смысле значительно лучше, чем в других интервалах отрезка $[0,1]$. Эти "благоприятные" для аппроксимации промежутки расположены вблизи точек $1/2,\ 1/3,\ 1/4,\ldots$, а также вблизи точек $2/3,\ 3/4,\ 4/5,\ldots$.



Фиг. 2. Графики функций $K(x,3)-x$ и $x-K(x,4)$.

Этот экспериментальный факт, возможно, объясняет, почему дробно-линейное преобразование с рациональными коэффициентами иррационального числа может радикальным образом изменить скорость сходимости обыкновенной цепной дроби новой константы по сравнению с исходной (см. [7]).

В книге [8, p. 28–29] Хинчин сравнивает преимущества и недостатки представления вещественных чисел обыкновенными цепными дробями с их представлением обычными (q-адичными в современной терминологии) дробями. Согласно Хинчину, с теоретической точки зрения цепные дроби имеют явные и значительные преимущества.

С практической точки зрения цепные дроби также имеют ряд преимуществ. Хинчин отмечает лишь один, но существенный их недостаток — неудобство проведения арифметических операций с этими дробями. Однако следует вспомнить, что эта книга вышла в 1935 году, когда арифметические операции могли выполняться только вручную. При вычислениях на современных компьютерах такие ограничения во многом становятся неактуальными.

Кроме того, мы не предлагаем выполнять арифметические операции с цепными дробями. Операции выполняются исключительно с рациональными числами, т.е. с парами целых чисел. Цепные дроби нужного порядка служат только для целей округления и вычисляются только для (некоторых) промежуточных результатов. Иными словами, вычисления в приближенной рациональной арифметике не требуют неограниченных ресурсов компьютера, как это было бы при вычислениях в обычной рациональной арифметике, но по-прежнему являются абсолютно точными в указанном выше смысле.

В следующих разделах мы приведем ряд примеров реализации некоторых классических алгоритмов в этой арифметике и результаты их работы.

3. ФУНКЦИИ $K(x, n)$ И ЧИСЛЕННЫЕ МЕТОДЫ

Напомним, что символ $K(x, n)$ мы использовали как для обозначения соответствующей цепной дроби, вычисляемой по алгоритму Евклида, так и для обозначения рационального числа, выражаемого этой дробью. Далее символ $K(x, n)$ понимается только в этом последнем смысле.

Функцию $K(x, n)$ мы вычисляли, в основном, по встроенному алгоритму в CAS Maple. Однако для рациональных чисел, которые мы далее используем, алгоритм Евклида легко (и столь же эффективно) реализуется как в системах CAS, так и в обычных языках программирования.

В частности, в C++ версии 11 и выше существует встроенный тип данных, предназначенный для вычислений в рациональной арифметике. Добавление операции округления, определенной выше, позволяет использовать эту арифметику для (практически) любых численных расчетов.

Приведем два примера реализации численных алгоритмов в округленной рациональной арифметике.

Рассмотрим сначала классический алгоритм Рунге—Кутты 4-го порядка интегрирования ОДУ

$$y'(x) = f(x,\ y(x)) \tag{4}$$

на конечном интервале. В современных обозначениях (см. [10, p. 134-135]) один шаг метода Рунге—Кутты имеет вид

$$y_k = y_0 + h \sum_{i=0}^{k-1} a_{k,i}\, f(x_0 + h\, c_i,\ y_i), \quad k = 1, 2, 3,$$

затем

$$y_h = y_0 + h \sum_{i=0}^{3} b_i\, f(x_0 + h\, c_i,\ y_i), \tag{5}$$

где x_0, y_0 — это данные Коши (начальные значения), а y_h — это приближенное значение решения $y(x)$ в точке $x = x_0 + h$.

Затем следует положить $x_0 := x_0 + h$, $y_0 := y_h$ и сделать следующий шаг, и т.д. В частности, для начального значения $x_0 = 0$, N шагов данного алгоритма с шагом $h = 1/N$ проинтегрируют уравнение (4) на интервале $[0, 1]$.

Коэффициенты

$$\{a_{k,i},\ b_j,\ c_j,\ i = 0, 1, 2,\ j = 0, 1, 2, 3,\ k = 1, 2, 3\}$$

выбираются таким образом, что

$$|y(x_0 + h) - y_h| \le C\, h^5, \quad C = \mathrm{const}(x_0, f),$$

т.е. тейлоровское разложение погрешности в точке $x = x_0 + h$ начинается с члена h^5.

Набор коэффициентов a, b, c может быть выбран неединственным образом. Классический выбор:

$$\begin{array}{lll} a_{1,0} = 1/2, & a_{1,1} = 0, & a_{1,2} = 0, \\ a_{2,0} = 0, & a_{2,1} = 1/2, & a_{2,2} = 0, \\ a_{3,0} = 0, & a_{3,1} = 0, & a_{3,2} = 1; \end{array}$$

$$b_0 = \frac{1}{6},\ b_1 = \frac{1}{3},\ b_2 = \frac{1}{3},\ b_3 = \frac{1}{6}; \quad c_0 = 0,\ c_1 = \frac{1}{2},\ c_2 = \frac{1}{2},\ c_3 = 1.$$

Насколько нам известно, методы Рунге—Кутты никогда не применялись в рациональной арифметике по понятной причине: эти вычисления (даже для полиномиальной правой части (4), не говоря уже о рациональной) быстро приведут к неконтролируемому росту объема промежуточных результатов.

Рассмотрим, например, задачу Коши $y(0) = 0$ для ОДУ

$$y'(x) = \frac{1}{1 + x^2}$$

и проинтегрируем его данным выше методом на интервале [0, 1] в рациональной арифметике с шагом $h = 1/10$. Получаем приближение

$$y(1) = \frac{5988585315838311774901484536676836463}{76249036426504635203016941416552830 00}, \quad y(1) - \frac{\pi}{4} \approx -1.550 \cdot 10^{-10}.$$

Ясно, что вычисления в рациональной арифметике по этому алгоритму при измельчении шага быстро становятся нереализуемыми.

Рассмотрим теперь модификацию алгоритма Рунге—Кутты, в которой последняя операция на каждом шаге (5) заменяется на операцию

$$y_h = K(y_0 + h \sum_{i=0}^{3} b_i \, f(x_0 + h\, c_i,\, y_i), M),$$

где M — это параметр алгоритма, т.е. порядок цепной дроби, который можно менять. Тогда с тем же шагом для $M = 18$ получим аппроксимацию

$$y(1) = \frac{77072475}{98131723}, \quad y(1) - \frac{\pi}{4} \approx -1.551 \cdot 10^{-10},$$

т.е. практически с той же точностью, но уже без неконтролируемого роста объема промежуточных величин.

Рассмотрим итерационный алгоритм Айткена ускорения сходимости знакопеременного ряда (или осциллирующей последовательности) (см. [11]). В отличие от классического алгоритма Айткена, который точен на геометрически сходящихся последовательностях, итерационный алгоритм является весьма эффективным ускорителем сходимости, применимым также к расходящимся рядам. Он дается рекуррентным соотношением

$$A(n, 0) = s(n),$$
$$A(n, k+1) = A(n, k) - \frac{(A(n+1, k) - A(n, k))^2}{A(n+2, k) - 2\,A(n+1, k) + A(n, k)}, \quad k \geq 0.$$

Последовательность $\{A(n, 1),\ n = 1, 2, \dots\}$ соответствует обычному методу Айткена.

Об асимптотических свойствах итерационного алгоритма Айткена известно немного (см. ссылки в [11]). Применим этот алгоритм для суммирования расходящегося ряда, дающего константу Эйлера—Гомпертца в качестве своей обобщенной суммы.

Осциллирующая последовательность $s(n)$ для этой константы имеет вид

$$s(n) = \sum_{k=0}^{n} (-1)^k\, k!,$$

т.е. последовательность быстро расходится. Вычисление (обобщенной) суммы этого ряда, $\delta = \mathbf{e}\,\mathrm{Ei}(\mathbf{1}, \mathbf{1})$ (где $\mathrm{Ei}()$ — это интегральная экспонента), является классической задачей, впервые решенной Эйлером.

Так называемая диагональная (так как подобные итерационные схемы традиционно изображают в виде ромба) последовательность $\{A(0, n),\ n = 1, 2, \dots\}$ весьма быстро сходится к δ, но только для небольших значений n. При $n > 10$ точность стабилизируется, а затем начинает убывать. При этом не вполне ясно, является ли этот эффект свойством алгоритма как такового или результатом накопления ошибок.

Численная неустойчивость итерационного алгоритма Айткена в плавающей арифметике ранее уже отмечалась в литературе (см. [11]). Например, расчеты в обычной D-арифметике (≈ 16 десятичных разрядов) дают для константы δ точность приближения $\approx 9.2 \cdot 10^{-11}$ на 10-м шаге. Далее точность падает, а затем процесс быстро расходится.

Расчеты в QD-арифметике (≈ 64 десятичных разряда) дают аналогичную точность. Различие состоит в том, что достигнутая точность стабилизируется на более длинном отрезке последовательности итераций, но затем также начинает убывать.

Вычисления в рациональной арифметике по алгоритму Айткена могли бы прояснить вопрос о влиянии погрешностей округления на свойства его сходимости (или расходимости). Однако такие вычисления, очевидно, невозможны (см. Введение). Поэтому применим рациональное округление в этом алгоритме, т.е. вместо $A(n, k)$ будем вычислять (и хранить)

$$\tilde{A}(n, k) = K(A(n, k), M(k)),$$

где $M(k) \in \mathbb{N}$ — это некоторая функция, которую можно выбирать различным образом.

Вычислительные эксперименты показали слабую зависимость точности вычисления константы δ от способа округления (т.е. от выбора функции $M(k)$). В частности, для $M(k) = k + 10$ достигнута точность приближения $\approx 5.8 \cdot 10^{-12}$ на 18-м шаге.

Таким образом, с большой вероятностью итерационный алгоритм Айткена не сходится к константе δ, а генерирует некоторый асимптотический расходящийся процесс.

4. ВЫЧИСЛЕНИЯ С ЦЕПНЫМИ ДРОБЯМИ

Отличие вычислений в рациональной арифметике от вычислений в плавающей арифметике, по-видимому, нигде не проявляется более отчетливо, чем при вычислении обыкновенных цепных дробей.

В рациональной арифметике все вычисления сводятся к манипуляциям с целыми числами и поэтому всегда выполняются абсолютно точно. Компьютер при этом является лишь инструментом, ускоряющим работу.

Напротив, вычисление цепной дроби в плавающей арифметике требует обращения к функциям вычисления целой и дробной части числа, $[x]$ и $\{x\}$, и поэтому всегда опирается на аппаратную реализацию приближенных операций плавающей арифметики.

Разумеется, вычисления в плавающей арифметике с расширенной разрядной сеткой, необходимые для получения большого количества подходящих дробей, имеют программную, а не аппаратную реализацию. Но проблемы недетерминистского округления при этом не исчезают, а лишь отодвигаются к более поздним стадиям работы алгоритма.

Обыкновенные цепные дроби трансцендентных констант обычно вычисляются именно с использованием плавающей арифметики с расширенной разрядной сеткой (см., например, [12]). Тем более удивительным представляется факт, что для вычисления подходящих дробей алгебраических иррациональностей достаточно использовать только вычисления с целыми числами.

Алгоритм, который мы представим, не нов и принадлежит Лагранжу (см. [13, p. 560]). С тех пор он неоднократно использовался многими авторами (см., например, [14]), однако факт вычислений только с целыми числами, по-видимому, ранее не отмечался, кроме (неявно) в [15, p. 375].

Пусть $P(x)$ — это полином с целочисленными коэффициентами степени d, имеющий единственный вещественный иррациональный корень $\alpha > 0$, который и разлагается в цепную дробь. Сам корень при этом неизвестен.

На самом деле, вещественных корней у полинома $P(x)$ может быть много, но ищется положительный корень ближайший к нулю, причем следующий положительный корень должен отстоять от α более чем на единицу.

Введем комплексные величины $\{r_n,\ n = -2, -1, 0, 1, \dots\}$ (*гауссовы целые*), которые кодируют подходящие дроби p_n/q_n числа α,

$$r_{-2} = 1,\ r_{-1} = i, \quad r_n = q_n + i\,p_n.$$

Смысл данного обозначения прояснится позднее.

Целые величины $\{a_n,\ n = 0, 1, 2, \dots\}$ — это частные знаменатели разложения

$$\alpha = [a_0; a_1, a_2, \dots], \quad 0 \le a_0.$$

Дадим сначала готовую процедуру в CAS Maple, а затем необходимые пояснения:

```
cf:=proc(P,N) local d,f,n,k,j,s,sk: global a,r,x:
f:=P: d:=degree(f,x):
r[-2]:=1: r[-1]:=I:
j:=n->‘if‘(n=0,0,1):
for n from 0 to N do
s:=sign(eval(f,x=j(n))):
for k from j(n)+1 do
sk:=sign(eval(f,x=k)):
if s<>sk then a[n]:=k-1: break fi
od:
f:=expand(x^d*subs(x=a[n]+1/x,f)):
r[n]:=a[n]*r[n-1]+r[n-2]:
od
end:
```

Структура алгоритма очевидна. Для определения следующего частного знаменателя a_n вычисляются значения полинома с целочисленными коэффициентами, полученного на предыдущем шаге, в точках $k = 1, 2, \ldots$ до тех пор, пока не встретится первое отличие в знаках. Это означает, что корень этого полинома больше, чем k, т.е. $a_n = k - 1$ и т.д.

Например, для полинома $P = x^3 - 2$ вычисление 5000 подходящих дробей числа $2^{1/3}$ на нашем компьютере занимает менее 0.6 сек., а 10000 — менее 1.28 сек. Встроенная процедура Maple вычисляет 10000 частных знаменателей разложения числа $2^{1/3}$ за 8.1 сек. Выборка из первых 10000 частных знаменателей содержит следующие большие знаменатели:

$$a_{35} = 534, \quad a_{571} = 7451, \quad a_{619} = 4941, \quad a_{1990} = 12737, \quad a_{2247} = 2897, \ \ldots,$$

причем $a_{1990} = 12737$ является самым большим знаменателем из этой выборки.

Численные эксперименты показывают, что частные знаменатели подходящих дробей алгебраических иррациональностей степени больше двух неограничены. Хотя, по-видимому, не существует доказательства того, что это свойство выполняется у какой-либо конкретной иррациональности этого вида. Вообще, существует крайне мало информации о степени роста частных знаменателей конкретных алгебраических иррациональностей (см. [16]).

Необычно большие знаменатели в разложении алгебраической иррациональности в цепную дробь являются слабым местом представленного алгоритма. Это видно из строчки “`for k from j(n)+1 do`”, после которой вычисляются значения полинома на шаге n в точках $k = 1, 2, \ldots$ до смены знака. В интернете можно найти способы обойти эту проблему с использованием алгоритмов бинарного поиска, т.е., например, использовать перебор $k = 2^m$, $m = 0, 1, 2, \ldots$, и т.д.

Существует алгоритм, позволяющий вычислить следующую подходящую дробь алгебраической иррациональности по двум предыдущим, если значение приближаемого числа известно. Этот алгоритм опирается на свойства *промежуточных дробей* (см. [8]). А именно: дроби

$$\frac{p_{n-2}}{q_{n-2}}, \frac{p_{n-2} + p_{n-1}}{q_{n-2} + q_{n-1}}, \frac{p_{n-2} + 2\,p_{n-1}}{q_{n-2} + 2\,q_{n-1}}, \ldots \frac{p_{n-2} + a_n\,p_{n-1}}{q_{n-2} + a_n\,q_{n-1}} = \frac{p_n}{q_n}$$

образуют при четном n возрастающую, а при нечетном n — убывающую последовательность, расположенную по одну сторону от приближаемого числа. Поэтому если значение приближаемого числа известно, то $k = a_n$ нумерует последний член этой последовательности. Следующий будет лежать уже по другую сторону от приближаемого числа. Например, для $\alpha = 2^{1/3}$ проводится сравнение

$$(p_{n-2} + k\,p_{n-1})^3 <> 2\,(q_{n-2} + k\,q_{n-1})^3, \quad k = 1, 2, \ldots,$$

в зависимости от четности n.

Этот алгоритм (для $\alpha = 2^{1/3}$) также использует только вычисления в целых числах, но имеет тот же недостаток, что и предыдущий алгоритм, а именно: большое число проверок для больших частных знаменателей.

Независимо от того, известен корень или нет, эти проверки можно не делать, если иметь хорошее приближение следующей подходящей дроби, p_n/q_n. Для этого можно использовать метод Ньютона вычисления корня полинома. А именно: пусть $N(x)$ обозначает ньютоновскую итерацию для данного полинома, т.е.

$$N(x) = x - \frac{P(x)}{P'(x)},$$

где x — это приближенное значение корня α, $P(\alpha) = 0$.

Пусть $h = x - \alpha$, тогда

$$N(x) - \alpha = \frac{1}{2}\,\frac{P''(\alpha)}{P'(\alpha)}\,h^2 + O(h^3),$$

или

$$|\alpha - N(x)| < C\,h^2 \tag{6}$$

для некоторой константы C, которая зависит только от полинома $P(x)$ и его корня α (квадратичная сходимость метода Ньютона).

Выберем в качестве приближения α последнюю найденную подходящую дробь, $x = \frac{p_{n-1}}{q_{n-1}}$, и найдем корень k уравнения

$$\frac{p_{n-2} + k\,p_{n-1}}{q_{n-2} + k\,q_{n-1}} = N\left(\frac{p_{n-1}}{q_{n-1}}\right) \approx \alpha$$

в рациональной арифметике. Затем положим $a_n = [k]$, т.е. вычислим целую часть рационального числа k. Здесь мы используем тот факт, что p_{n-1}/q_{n-1} и p_n/q_n лежат по разные стороны от α. Поэтому целое $[k]$ и рациональное k также с большой вероятностью лежат по разные стороны от α.

На самом деле, дробь p_n/q_n почти всегда вычисляется точно этим способом, если использовать функции $K(x, n)$. Приведем сначала

Предложение 3. *Пусть дано произвольное иррациональное число $x > 0$ и его разложение в обыкновенную цепную дробь. И пусть рациональное число y находится между двумя подходящими дробями числа x, p_n/q_n и p_{n+1}/q_{n+1}, $n \in \mathbb{N}$. Тогда подходящие дроби числа y совпадают с таковыми для числа x вплоть до n-й включительно.*

Доказательство очевидно (см. [6]). Поэтому справедливо следующее

Предложение 4. *Пусть дан полином с целыми коэффициентами $P(x)$, имеющий иррациональный корень $\alpha > 0$, $P'(\alpha) \neq 0$. И пусть вычислена подходящая дробь p_{n-1}/q_{n-1} числа α с достаточно большим номером n. Тогда последующие подходящие дроби вычисляются по формуле*

$$\frac{p_n}{q_n} = K\left(N\left(\frac{p_{n-1}}{q_{n-1}}\right), n\right) \tag{7}$$

при условии, что частные знаменатели a_{n+1} или a_{n+2} не окажутся слишком большими. Например, достаточно, чтобы выполнялось условие

$$\max(a_{n+1}^3, a_{n+2}^3) < \text{const}\, q_{n-1}^2. \tag{8}$$

Доказательство. Согласно предложению 3, достаточно показать, что рациональное число $x_n = N(p_{n-1}/q_{n-1})$ лежит между числами p_n/q_n и p_{n+1}/q_{n+1}.

Подходящие дроби p_n/q_n и p_{n-1}/q_{n-1} всегда лежат по разные стороны от числа α. Всего существует четыре случая: n четно или нечетно; и функция $P(x)$ локально выпукла или вогнута в окрестности α.

В зависимости от случая достаточно, чтобы

$$|\alpha - x_n| < \left|\alpha - \frac{p_n}{q_n}\right| \quad \text{или} \quad |\alpha - x_n| < \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right|. \tag{9}$$

Но второе из этих неравенств влечет первое, поэтому рассмотрим второе.

Пусть

$$h = \left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| < \frac{1}{q_{n-1}\,q_n}. \tag{10}$$

Данное неравенство — это теорема 9 в [8]. Согласно теореме 13 в [8], имеем

$$\frac{1}{q_{n+1}\,(q_{n+1} + q_{n+2})} < \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right|.$$

Поэтому достаточно показать, что

$$|\alpha - x_n| < \frac{1}{q_{n+1}\,(q_{n+1} + q_{n+2})}.$$

Используя оценки (6) и (10), получаем неравенство

$$C < \frac{q_{n-1}^2\,q_n^2}{q_{n+1}\,(q_{n+1} + q_{n+2})},$$

достаточное для выполнения (7). Заменяя q_m по их рекуррентным формулам и загрубляя неравенство, получим (8). Что требовалось доказать.

Заметим, что для выполнения первого из неравенств (9) достаточно выполнения неравенства

$$a_{n+1} < \operatorname{const} q_{n-1}^2. \tag{11}$$

Нарушение любого из неравенств (8) или (11) означало бы появление катастрофически большого частного знаменателя у алгебраической иррациональности. Это маловероятно даже в единичных случаях, а нарушение этих условий бесконечное число раз возможно, по-видимому, только у трансцендентных чисел.

Об этом же свидетельствует теорема Рота (см. [17]), которая ограничивает скорость стремления подходящих дробей алгебраической иррациональности к своему пределу.

Например (см. [6]), положим $x_0 = 1$ и

$$x_{n+1} = K\left(x_n - \frac{x_n^3 - 2}{3\,x_n^2}, n+1\right), \quad n \in \mathbb{N}_0.$$

Тогда последовательность $\{x_n\}$ — это все подходящие дроби числа $\sqrt[3]{2}$. Во всяком случае, это справедливо для $n \le 10000$.

Алгоритм вычисления подходящих дробей алгебраических иррациональностей, изложенный в предложении 4, представляется заведомо менее оптимальным, чем алгоритм

Лагранжа, изложенный выше. Однако это, вообще говоря, не так. Рассмотрим уравнение (см. [18])

$$P(x) = x^3 - 8x - 10, \quad P(\alpha) = 0, \quad \alpha \approx 3.318628217750,$$

корень которого имеет разложение в обыкновенную цепную дробь с экстремально большими частными знаменателями, которые встречаются среди первых 200 членов разложения α в цепную дробь. Например,

$$a_{17} = 22986, \quad a_{33} = 1501790, \quad a_{121} = 16467250, \quad a_{161} = 325927.$$

В [18] показано, что это явление связано с тем, что данное уравнение имеет отношение к квадратичному полю $\mathbb{Q}(\sqrt{-163})$ и что порядок класса этого поля равен единице. Такой пример, разумеется, не единственный (см. [18]).

Алгоритм Лагранжа считает на нашем компьютере первые 100 подходящих дробей за 4.72 сек, а первые 200 — за 72.14 сек. В то время как наш алгоритм дает 200 подходящих дробей за 0.1 сек.

Правда, при нулевом приближении $x_0 = 3$ первые три подходящие дроби неправильны, но они легко восстанавливаются по полученным правильным с помощью функций $K(x, n)$ (см. ниже).

Классический критерий Дирихле иррациональности числа α имеет вид

$$|\alpha q - p| \to 0, \quad \frac{p}{q} \to \alpha, \quad p, q \in \mathbb{N}.$$

Подходящие дроби всех иррациональных чисел обладают этим свойством.

Как оказалось, для всех случаев, когда иррациональность числа удается установить по критерию Дирихле с помощью быстро сходящейся последовательности рациональных чисел (например, для π, $\ln 2$, $\zeta(2)$, $\zeta(3)$), эти последовательности стремятся к своему пределу быстрее в численном смысле, чем подходящие дроби этого предела (разумеется, это экспериментальный факт).

Это, вероятно, связано с тем, что подходящие дроби являются наилучшими диофантовыми приближениями своего предела. Поэтому другие дроби, удовлетворяющие критерию Дирихле, должны как-то компенсировать свою неоптимальность, например, за счет более быстрой сходимости. Однако данное рассуждение хотя и правдоподобно, но, вообще говоря, неверно (см. контрпример ниже).

Применение функций $K(x, n)$ позволяет генерировать подходящие дроби числа α из любой последовательности рациональных чисел p/q, если она стремится к α быстрее в обычном численном смысле, чем последовательность его подходящих дробей. А именно, будем говорить, что последовательность $s_n \to \alpha$ является *быстро сходящейся*, если выполнено условие

$$|\alpha - s_n| < \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right|, \quad \text{const} < n \in \mathbb{N}. \tag{12}$$

Тогда (см. предложение 3)

$$\frac{p_n}{q_n} = K(s_n, n), \tag{13}$$

где p_n/q_n — это обыкновенные подходящие дроби числа α. Например (см. [6]), мы сгенерировали таким образом подходящие дроби (до 100-й включительно) констант π, $\ln 2$, $\zeta(n)$, $n = 2, 3, \ldots, 9$, а также константы γ.

Возникает вопрос, как определить, является ли данная последовательность рациональных чисел быстро сходящейся, если последовательность подходящих дробей предела α неизвестна. На этот вопрос дает ответ следующее

Предложение 5. *Пусть дана сходящаяся неприводимая последовательность рациональных чисел* $t_n = p_n/q_n$. *Тогда она является* (*начиная с некоторого индекса* n) *последовательностью подходящих дробей своего предела* α *тогда и только тогда, когда*

$$t_n = K(t_n, n) = K(t_{n+2}, n), \quad \text{const} < n. \tag{14}$$

Доказательство. В одну сторону утверждение очевидно, т.е. все последовательности подходящих дробей иррациональных чисел обладают свойством (14).

Для однозначного представления рационального числа p_{n+2}/q_{n+2} в виде обыкновенной цепной дроби необходимо, чтобы $a_{n+2} \neq 1$ (см. [8]). Но в любом случае,

$$t_{n+2} = K(t_{n+2}, n+2) = [a_0; a_1, \ldots, a_n, a_{n+1}, a_{n+2}] = [a_0; a_1, \ldots, a_n + \frac{1}{a_{n+1} + \frac{1}{a_{n+2}}}].$$

Поэтому $t_n = K(t_n, n) = [a_0; a_1, \ldots, a_n]$ находится однозначно. Что требовалось доказать.

Недостающие подходящие дроби последовательности $\{t_n\}$ в начале списка однозначно восстанавливаются с помощью функций $K(x, n)$, т.е. надо положить

$$t_n = K(t_{n+2}, n), \quad n < \text{const},$$

заменив неправильные t_n в начале списка.

Наконец, частные знаменатели разложения числа α в обыкновенную цепную дробь находятся по формуле

$$\frac{r_n - r_{n-2}}{r_{n-1}} = a_n \in \mathbb{N}, \tag{15}$$

где $r_n = q_n + i\,p_n$. Иными словами, цепная дробь Эйлера, построенная по алгоритму Д. Бернулли из последовательности $\{t_n\}$ (см. [6]), совпадает с обыкновенной цепной дробью предела α, если начальный отрезок последовательности $\{t_n\}$ пересчитать с помощью функций $K(x, n)$.

Таким образом, пока выполняется условие (14), выполняется и (13), т.е. подходящие дроби, полученные с помощью функций $K(x, n)$, являются также подходящими дробями предела последовательности.

Однако доказать, что последовательность s_n, полученная, например, по алгоритму ускорения сходимости (см. [6]), обладает свойством (12) для всех $n > \text{const}$, по-видимому, невозможно. Единственный пример (квадратичные иррациональности не в счет), когда это можно установить достоверно, дает константа **e** (см. [7]).

В то же время результаты Хинчина (см. [8]) и Леви (см. [19]) дают усредненную оценку сверху роста частных знаменателей для почти всех вещественных чисел. Поэтому “в среднем” подходящие дроби иррационального числа не могут стремиться к своему пределу слишком быстро.

Это означает, что вполне реалистично ожидать, что последовательность рациональных чисел s_n, сходящаяся к пределу α быстрее, чем последовательность его подходящих дробей, может быть построена, если не требовать выполнения условия (12) всюду, а допустить его нарушение в некоторых случаях. Или даже в бесконечном числе случаев. Тогда не все подходящие дроби числа α могут быть получены по формуле (13).

С другой стороны, мы узнаем об этом из предложения 5, т.е. сгенерированные подходящие дроби $t_n = K(s_n, n)$ не будут удовлетворять свойству (14).

Однако кодирование иррационального числа его подходящими дробями обладает весьма большой избыточностью. Например, справедливо

Предложение 6. *Пусть известна последовательность четных подходящих дробей иррационального числа* $\alpha > 0$,

$$\frac{p_0}{q_0}, \frac{p_2}{q_2}, \ldots, \frac{p_{2n}}{q_{2n}}, \ldots.$$

Тогда все остальные элементы разложения числа $\alpha > 0$ *в обыкновенную цепную дробь восстанавливаются однозначно.*

Доказательство. Необходимо определить все нечетные подходящие дроби p_{2n-1}/q_{2n-1}, а также все частные знаменатели a_n числа α.

Заметим сперва, что $p_{-2} = 0$, $q_{-2} = 1$ и $p_{-1} = 1$, $q_{-1} = 0$ по определению. Поэтому $p_0 = a_0$ и $q_0 = 1$. Рассмотрим последовательность

$$p_n = a_n p_{n-1} + p_{n-2},\ q_n = a_n q_{n-1} + q_{n-2}, \quad n = 1, 2, \ldots, 2N,$$

как систему $4N$ линейных уравнений относительно $4N$ неизвестных

$$a_1, a_2, \ldots, a_{2N}, \quad p_1, p_3, \ldots, p_{2N-1},\ q_1, q_3, \ldots, q_{2N-1}.$$

Тот факт, что система линейна, можно увидеть, если разбить уравнения на группы $\{1, 2, 3, 4\}$, $\{5, 6, 7, 8\}$ и т.д., а затем исключить неизвестные a_n.

Нетрудно доказать по индукции, что эта система всегда разрешима. Например, для $N = 1$ получим решение

$$a_1 = \frac{q_2 - q_0}{p_2 q_0 - p_0 q_2},\ a_2 = \frac{p_2 q_0 - p_0 q_2}{q_0},\ p_1 = \frac{q_0 (p_2 - p_0)}{p_2 q_0 - p_0 q_2},\ q_1 = \frac{q_0 (q_2 - q_0)}{p_2 q_0 - p_0 q_2}.$$

Далее формулы становятся более громоздкими, но, вспоминая, что $q_0 = 1$, и вычисляя сперва величины с четными номерами, общее решение можно записать в виде

$$a_n = \begin{cases} p_n q_{n-2} - p_{n-2} q_n, & n \text{ четно}, \\ \dfrac{p_{n+1} q_{n-3} - p_{n-3} q_{n+1} - a_{n+1} - a_{n-1}}{a_{n+1} a_{n-1}}, & n \text{ нечетно}, \end{cases}$$

где

$$p_n = \begin{cases} p_n, & \\ \dfrac{p_{n+1} - p_{n-1}}{a_{n+1}}, & \end{cases} \qquad q_n = \begin{cases} q_n, & n \text{ четно}, \\ \dfrac{q_{n+1} - q_{n-1}}{a_{n+1}}, & n \text{ нечетно}, \end{cases}$$

что требовалось доказать.

Аналогичным образом можно восстановить четные подходящие дроби разложения числа в обыкновенную цепную дробь, если знать все нечетные подходящие дроби. Отметим, что эти операции не являются обратными операциям четного или нечетного сжатия обыкновенной цепной дроби, так как частные знаменатели здесь неизвестны.

Предложение 6, впрочем, становится почти очевидным, если использовать предложение 3 и функции $K(x, n)$: достаточно вычислить последовательность подходящих дробей

$$\frac{p_n}{q_n} = K\left(\frac{p_{2N}}{q_{2N}}, n\right), \quad n = 1, 2, \ldots, 2N - 2,$$

а затем вычислить цепную дробь Эйлера, которая даст $a_n, n = 1, \ldots, 2N - 2$ (а также и a_{2N-1}, если $a_{2N} \neq 1$).

Таким образом, один из возможных сценариев доказательства иррациональности числа α — это предъявление последовательности рациональных чисел, сходящейся к α, где условие (12) выполняется если не всюду, то достаточно часто. Однако и это ослабленное требование к скорости сходимости может оказаться слишком ограничительным. К счастью, оно не является необходимым.

Напомним, что два иррациональных числа α и β называются *эквивалентными*, если существует преобразование

$$\alpha = \frac{a\beta + b}{c\beta + d}, \quad a, b, c, d \in \mathbb{Z}, \quad a d - b c = \pm 1.$$

Тогда частные знаменатели a_n числа α и b_n числа β совпадают начиная с некоторого номера $n > N$ и $n + m$, $m \in \mathbb{Z}$, т.е. $a_n = b_{n+m}$ (см., например, [20]).

Однако, как мы уже отмечали, дробно-линейное преобразование с целыми коэффициентами $a, b, c, d \in \mathbb{Z}$, не принадлежащее модулярной группе, может радикальным образом изменить скорость сходимости подходящих дробей числа α, по сравнению с β. Например, (см. [7]),

$$\mathbf{e} - K(\mathbf{e}, 10) \approx 1.102 \cdot 10^{-7}, \quad \frac{\mathbf{e}-1}{\mathbf{e}+1} - K\left(\frac{\mathbf{e}-1}{\mathbf{e}+1}, 10\right) \approx 4.216 \cdot 10^{-26}.$$

Но всегда существует цепная дробь для α с целыми коэффициентами, которая продолжается как обыкновенная цепная дробь для β и наоборот. Например, $(\mathbf{e}-1)/(\mathbf{e}+1) = 1/(1+2/(\mathbf{e}-1))$, поэтому

$$\frac{\mathbf{e}-1}{\mathbf{e}+1} = \cfrac{1}{2+\cfrac{1}{6+\cfrac{1}{10+\cfrac{1}{14+\cfrac{1}{18+\cfrac{1}{22+\ldots}}}}}} = \cfrac{1}{1+\cfrac{2}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{4+\ldots}}}}}}},$$

т.е. очень быстро сходящаяся последовательность подходящих дробей числа $(\mathbf{e}-1)/(\mathbf{e}+1)$ может быть заменена последовательностью

$$\left\{1, \frac{1}{3}, \frac{1}{2}, \frac{5}{11}, \frac{7}{15}, \frac{6}{13}, \frac{55}{119}, \frac{67}{145}, \frac{61}{132}, \frac{799}{1729}, \ldots\right\}, \tag{16}$$

которая сходится как последовательность подходящих дробей числа $\mathbf{e}$, но не является таковой (условие (14) не выполнено).

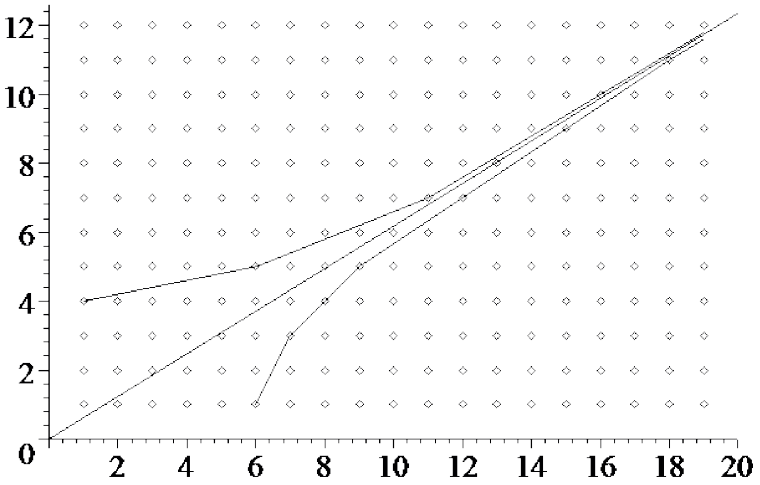
Тем не менее, последовательность (16) обладает почти всеми свойствами последовательностей обыкновенных подходящих дробей и, в частности, подразумевает иррациональность своего предела.

Монотонно возрастающие или убывающие последовательности рациональных чисел играют особую роль в установлении иррациональности своего предела. Так, четные и нечетные подходящие дроби обладают этим свойством.

Однако помимо монотонности, подходящие дроби обладают еще одним свойством, которое позволяет сразу сделать вывод об иррациональности их предела. А именно: последовательности гауссовых целых r_{2n} и r_{2n-1} всегда ведут себя так, как показано на фиг. 3.

На фиг. 3 показаны комплексные точки $r_n = q_n + i\,p_n$ для четных (снизу) и нечетных (сверху) подходящих дробей некоторого иррационального числа α, а также множество точек $\{x + i\alpha x,\ x > 0\}$, т.е. график функции $y = \alpha x$.

Монотонно возрастающая (убывающая) последовательность рациональных чисел p_n/q_n называется *выпуклой* (*вогнутой*), если последовательность $(p_{n+1} - p_n)/(q_{n+1} - q_n)$ монотонно убывает (возрастает), причем $q_{n+1} > q_n$.



Фиг. 3. Выпуклая и вогнутая последовательности.

Очевидно, это определение выпуклости (вогнутости) является дискретным аналогом обычной выпуклости (вогнутости) гладкой кривой, а также является критерием иррациональности Бруна (см. [6, 21]).

Выпуклость (вогнутость) этих ломаных влечет иррациональность α независимо от критерия Бруна (или является его доказательством), так как если бы α было рациональным, то существовала бы некоторая полоса вдоль прямой $y = \alpha x$ (или $y = \alpha x + \beta$, $\beta \in \mathbb{R}$), т.е. асимптоты этих ломаных, где все целочисленные точки лежали бы только на этой прямой либо вообще отсутствовали. В то время как существование таких выпуклых или вогнутых ломаных, имеющих целочисленные вершины, противоречит существованию такой полосы.

Как показано в [6], все последовательности четных (нечетных) подходящих дробей иррациональных чисел удовлетворяют критериям Бруна. Однако фиг. 3 указывает на существование бесконечного числа таких последовательностей для любой иррациональности. Первый такой пример дал сам Брун для числа **e**. Свойство неприводимости дробей при этом не только не является необходимым, но может препятствовать построению таких последовательностей (как для числа **e**, см. [6]).

Таким образом, имея достаточно быстро сходящуюся последовательность рациональных чисел, теоретически можно выбрать из нее (или сконструировать) выпуклую или вогнутую последовательность, что автоматически даст иррациональность предела.

Насколько нам известно, единственная работа в этом направлении — это построение выборки, удовлетворяющей критерию Бруна, из последовательности Апери для числа $\zeta(3)$ (см. [22]).

5. ЗАКЛЮЧЕНИЕ

Отметим, что вычисления в рациональной арифметике, к сожалению, пока не являются разделом численного анализа, а скорее принадлежат компьютерной алгебре.

Между тем, (помимо примеров в этой статье) существуют приложения вычислений в рациональной арифметике, относящиеся непосредственно к численному анализу (см. [7]). Так, коэффициенты разложений некоторых функций по полиномам Чебышёва или Лежандра являются рациональными числами. Квадратуры Гаусса на интервале [0, 1] от любой рациональной функции с рациональными коэффициентами также являются рациональными числами (либо неопределены). Эти и другие факты невозможно увидеть при расчетах в плавающей арифметике.

СПИСОК ЛИТЕРАТУРЫ

1. *Borwein J., Bailey D., Gigensohn R.* Experimentation in Mathematics: computational paths to discovery. A K Peters, Natick. 2004.
2. *Арнольд В. И.* Экспериментальная математика. М.: ФАЗИС, 2005.
3. *Hammer R., et.al.* Numerical Toolbox for Verified Computing I. Springer, 1993.
4. *Appel K., Haken W.* Solution of the Four Color Map Problem // Scientific American. 1977. V. 237. N 4. P. 108–121.
5. *Рухович Ф. Д.* Внешние биллиарды вне правильных многоугольников: ручной случай // Изв. РАН. Сер. матем. 2022. Т. 86. No 3. С. 105–160.
6. *Варин В. П.* Преобразование последовательностей в доказательствах иррациональности некоторых фундаментальных констант // Ж. вычисл. матем. и матем. физ. 2022. Т. 62. N 10. С. 1587–1614.
7. *Варин В. П.* Аппроксимация дифференциальных операторов с учетом граничных условий // Ж. вычисл. матем. и матем. физ. 2023. Т. 63. N 8. С. 1251–1271.
8. *Хинчин А. Я.* Цепные дроби. М.: Гостехиздат, 1961.
9. *Воробьев Н. Н.* Числа Фибоначчи. М.: Наука, 1984.
10. *Hairer, et. al.* Solving Ordinary Differential Equations I. Nonstiff Problems. 2nd ed. Berlin, Springer. 1993.
11. *Weniger E. J.* Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series // Comput. Phys. Rep. North-Holland. Amsterdam, 1989. V. 10. P. 189–371.
12. *Brent R. P.* Computation of the Regular Continued Fraction for Euler's Constant // Math. of Comput. 1977. V. 31. No 139. P. 771–777.
13. *Serret M. J.-A.* Oeuvres de Lagrange. Vol. 2. Gauthier-Villars. Paris. (M DCCC LXVIII).
14. *Брюно А. Д.* Разложение алгебраических чисел в цепные дроби // Ж. вычисл. матем. и матем. физ. 1964. Т. 4. No 2. С. 211–221.
15. *Knuth D. E.* The Art of Computer Programming. 3rd ed. Vol. 2. 1998. Addison Wesley Longman.
16. *Haas A.* The relative growth rate for partial quotients // New York J. Math. 2008. V. 14. P. 139–143.
17. *Roth K. F.* Rational approximations to algebraic numbers // Mathematika. 1955. V. 2. Part 1. No 3. P. 1–20.
18. *Stark H. M.* An explanation of some exotic continued fractions found by Brillhart // in: A.O.L. Atkin, B.J. Birch (ed.), Computers in Number Theory. Science Research Council Atlas Symposium N 2. Oxford, Academic Press, 1971.
19. *Lévy P.* Sur le lois de probabilité dont dépendent les qoutients complets et incomplets d'une fraction continue // Bull. Scc. Math. 1929. V. 57. P. 178–194.
20. *Borwein J., et. al.* Neverending Fractions. An Introduction to Continued Fractions. Australian Mathematical Society Lecture Series: 23. 2014.
21. *Brun V.* Ein Satz über Irrationalität // Arkiv for Mathematik og Naturvidenskab (Kristiania), 1910. V. 31. No 3. P. 3–6.
22. *Butler L. A.* A useful application of Brun's irrationality criterion // Expo. Math. 2015. V. 33. P. 121–134.

RATIONAL ARITHMETIC WITH A ROUND-OFF

V. P. Varin*

Keldysh Institute of Applied Mathematics RAS, Miusskaya sq. 4., Moscow, 125047 Russia
**e-mail: varin@keldysh.ru*

Received 16 January, 2024
Revised 08 February, 2024
Accepted 05 March, 2024

Abstract. Computer calculations in floating-point arithmetic are always approximate. In contrast, calculations in rational arithmetic (for example, in computer algebra) are always absolutely precise and reproducible both on other computers and (theoretically) manually. Therefore, such calculations can be demonstrative in the sense that the proof obtained with their help is no different from the traditional one. However, such calculations are usually impossible in a sufficiently complex problem due to limited memory and time resources. We propose a mechanism for rounding off rational numbers in calculations in rational arithmetic, which solves this problem (of resources), i.e. the calculations can still be demonstrative, but no longer require unlimited resources. A number of examples of the implementation of standard numerical algorithms in this arithmetic are given. The results have applications to analytical number theory.

Keywords: rational arithmetic, convergent fractions, proof calculations, Brun's criterion of irrationality.