$Ч$ — кількість чавуну, який виплавляється із шихти однієї подачі, кг. Значення $O_{шп}$ та $Ч$ визначати із наступних співвідношень:

$$O_{шп} = 0{,}222 A_г FeO + 0{,}43(A_г Fe - 0{,}778 A_г FeO) +$$
$$+0{,}43 КЖFe_ж + Ч(1{,}14[Si] + 0{,}291[Mn] + 1{,}291[P]) +$$
$$+0{,}5)0{,}85 KS_K Ч[S]) + 0{,}727 ИCO_{2кр} - 0{,}43 ПFe_п\,; \quad (35)$$

$$Ч = (A_г \times Fe + К \times Ж \times Fe_ж + Д \times Fe_д -$$
$$-П \times Fe_п) / Fe_ч\,, \quad (36)$$

де $A_г$ — маса рудної частини шихти у подачі, кг;
$К$ — маса коксу у подачі, кг;
$П$ — винос колошникового пилу, кг/подачу;
$И$ — маса вапняку у подачі, кг;
$Д$ — маса металобрухту у подачі, кг;
$Ж$ — вміст золи у коксі, частка одиниці маси;
Fe, FeO — вміст заліза та його оксиду у рудній частині шихти, частка одиниці маси;
$Fe_ч$, $Fe_д$, $Fe_п$, $Fe_ж$ — вміст заліза, відповідно, у чавуні, металобрухті, колошниковому пилу та у золі коксу, частка одиниці маси;
$CO_{2\ кр}$ — вміст вуглекислоти у вапняку, частка одиниці маси;
$[Si]$, $[S]$, $[Mn]$, $[P]$ — вміст у чавуні, відповідно, кремнію, сірки, марганцю та фосфору, частка одиниці маси;
S_K — вміст сірки у коксі, частка одиниці маси;
0,222 — частка кисню у FeO;
0,778 — частка заліза у FeO;
0,727— частка кисню у CO_2;
0,43 — відношення маси кисню до маси заліза у Fe_2O_3;
1,14 —відношення маси кисню до маси кремнію у SiO_2;
0,291 — відношення маси кисню до маси марганцю у MnO;
1,291 — відношення маси кисню до маси фосфору у P_2O_5;
0,85 — кількість сірки коксу, яка переходить у шлак та чавун, частка одиниці маси;
0,5 — частка маси газифікованого кисню від маси сірки при її ошлакуванні за реакцією $FeS + CaO = Fe + CaS + 0{,}5O_2$.

Висновки. За допомогою наведеного в статті математичного алгоритму можливо визначити основні параметри доменної печі, необхідні для автоматизованого контролю ефективності використання та оптимізації кількості пиловугільного палива.

ЛІТЕРАТУРА

1. Довгалюк Б.П. Автоматизована система керування технологічними процесами доменної плавки. Дніпродзержинськ: — ДДТУ, 2009. — 245 с.
2. Довгалюк Б.П. Вплив втрат тепла з охолоджувальною водою і колошниковим газом на тепловий режим плавки. / Б.П. Довгалюк, Р.В. Волошин // Математичне моделювання. — 2013. — № 1 (28). — С. 64—67.

пост. 30.11.2015

Вычисление максимальних абсолютных погрешностей округления чисел в IEEE-стандарте

И.И. ЖУЛЬКОВСКАЯ, О.А. ЖУЛЬКОВСКИЙ

Днепродзержинский государственный технический университет

В статье рассмотрены особенности стандартных способов округления чисел с плавающей запятой при вычислениях на компьютере, а также точность представления чисел с плавающей запятой. Получены формулы для вычисления максимальных абсолютных погрешностей округления чисел, представленных в базовых форматах стандарта *IEEE* 754.

У статті розглянуто особливості стандартних способів округлення чисел з плаваючою комою під час обчислень на комп'ютері, а також точність представлення чисел з плаваючою комою. Отримані формули для обчислення максимальних абсолютних похибок округлення чисел, представлених в базових форматах стандарту *IEEE* 754.

The article describes the features of the standard methods of rounding floating point numbers in the calculations on the computer, precision floating-point numbers. In the article obtained the formulas for calculating the maximum absolute round off errors of numbers, presented in the basic format of the *IEEE* 754 standard.

Общая характеристика проблемы. Одним из наиболее доступных методов исследования объектов и процессов, происходящих во всех сферах человеческой деятельности, является математическое моделирование. Источниками ошибок в процессе математического моделирования могут быть погрешности исходной математической модели, неверный выбор численных методов и алгоритмов, а также вычислительные ошибки. Вычислительные ошибки возникают в связи с точностью представления действительных чисел в памяти ЭВМ при вводе данных или при проведении арифметических операций.

К особенностям машинной арифметики можно отнести конечность набора действительных чисел и ограниченность диапазона их изменения. Это, в результате, приводит к следующим последствиям: полученное решение верно лишь с машинной точностью, результат каждой промежуточной арифметической операции округляется, определенные элементы алгоритма (например, критерии останова) будут зависеть от машинной точности.

Таким образом, при проведении расчетов на ЭВМ нужно иметь представление о погрешностях машинной арифметики.

Неустранимые ошибки вычислений связаны, в частности, с ошибками округления чисел при вычислениях на компьютере. Их обязательно надо свести к минимуму или сообщать пользователю соответствующей программы о недопустимой величине этих ошибок.

Краткий аналитический обзор. Ассоциацией *IEEE* (*Institute of Electrical and Electronics Engineers*) был разработан и внедрен единый стандарт для представления в двоичном коде чисел с плавающей запятой *IEEE* 754 [1].

Из математических функций стандартом *IEEE* 754 описывается только извлечение квадратного корня. Поэтому во многих публикациях [2–4] обсуждается необходимость и подходы к стандартизации реализаций большого количества математических функций, работающих с числами с плавающей запятой в форматах *IEEE* 754. Наряду с этим, множество работ [4–6] посвящены тестированию на соответствие стандарту *IEEE* 754, вопросам погрешностей результатов при вычислении функций [7]. Ранее авторами данной работы исследован и описан алгоритм современного подхода к формированию машинного представления и хранения числовой информации в формате с плавающей запятой [8], рассмотрены особенности представления, а также вычислены граничные (максимальные и минимальные) значения субнормальных и нормализованных чисел [9, 10].

Цель работы (постановка задачи) – рассмотрение особенностей стандартных способов округления чисел с плавающей запятой при вычислениях на компьютере, а также получение формул для вычисления максимальных абсолютных погрешностей округления чисел, представленных в базовых форматах стандарта *IEEE* 754.

Результаты работы. Значение двоичного числа с плавающей запятой определяется выражением:

$$A = \pm a_0.a_1a_2a_3...a_{n-1} \times S^e \qquad (1)$$

где S – основа системы счисления;

n – число значащих разрядов мантиссы;

a_i – цифры ($0 \le a_i < S$);

e – порядок или экспонента (не путать с числом e).

Стандарт *IEEE* 754 однозначно определяет основные параметры форматов представления чисел с плавающей запятой, которые отличаются диапазоном представимых в них значений. Стандарт *IEEE* 754 определяет несколько возможных типов чисел с плавающей запятой, из которых чаще всего используются числа одинарной точности (*single precision*), числа двойной точности (*double precision*) и числа двойной расширенной точности (*double-extended precision).* Кроме того, иногда используются числа так называемой четырехкратной точности (*quadruple*). Параметры базовых форматов приведены в табл.1.

Стандарт настоятельно рекомендует к реализации формат расширенной точности, где для хранения экспоненты выделено как минимум 15 бит, а для хранения дробной части мантиссы отведено как минимум 63 бита. Некоторые архитектуры имеют поддержку такого формата, но детали реализации отличаются от производителя к производителю. Так, например, микропроцессоры *Intel* поддерживают формат расширенной точности с 1 битом знака, 15-битной экспонентой и 64-битной мантиссой, и хранят их в регистрах модуля *FPU* (*Floating Point Unit*) шириной 80 бит. Микропроцессоры *Sun SPARC* реализуют поддержку формата расширенной точности программно, а для хранения таких чисел задействуется 128 бит.

В стандарте *IEEE* 754 порядок представлен в виде беззнакового числа, называемого смещенным порядком или характеристикой (X), которое отличается от порядка на фиксированную для данного формата величину, называемую смещением (b).

Таблица 1. Параметры базовых форматов чисел с плавающей запятой

Формат	**Всего бит**	**Бит в порядке**	**Бит в мантиссе (m)**	**Смещение порядка (b)**
single	32	8	23	127
double	64	11	52	1023
extended	80	15	64	16383
quadruple	128	15	112	16383

Естественно, не каждое действительное число можно представить в формате с плавающей запятой. К представимым относятся числа, не имеющие дробной части (т.е. множество целых чисел), а также числа, которые можно представить в виде конечной двоичной дроби. Точнее, представимые в формате с плавающей запятой числа должны быть двоично-рациональными, т.е. должны иметь вид $p/2^m$, где p и m – целые числа, причем m неотрицательно. Только в этом случае число представляется в формате с плавающей запятой без округления и, соответственно, без потери точности. Остальные же числа представлены приближенно, т.е. они округляются до точно представимых в формате с плавающей запятой чисел.

Как известно, стандарт *IEEE* 754 описывает операции сложения, умножения, вычитания, деления, вычисления остатка от деления, извлечения квадратного корня и преобразований между различными типами чисел. Общий принцип всех операций заключается в том, что результат получается из точного путем приведения к представимому числу, согласно установленному режиму округления. Таким образом, исходное число модифицируют так, чтобы результат поместился в целевой формат.

Для представления результатов вычислений в виде чисел с плавающей запятой стандарт *IEEE* 754 определяет следующие четыре режима округления.

1. Округление к ближайшему (*round to nearest*). В этом режиме должно возвращаться ближайшее к результату представимое значение. В случае равноудаленности (*tie-breaking rule*) необходимо возвращать

число с нулевым младшим битом мантиссы (*lsb – least significand bit*).

2. Округление к отрицательной бесконечности. В этом случае результатом будет ближайшее представимое число, не превосходящее точного значения, т.е. ближайшее меньшее представимое число.

3. Округление к положительной бесконечности. Режим такого округления предусматривает возвращение ближайшего представимого числа, которое не меньше точного значения, т.е. ближайшее большее представимое число.

4. Округление к нулю. Такой режим округления возвращает ближайшее представимое число, не превосходящее по абсолютной величине точного значения, что означает выбор после усечения числа, ближайшего к нулю. Таким образом, для положительного точного результата округление происходит, как в режиме округление к $-\infty$, а для отрицательного – как в округлении к $+\infty$.

Три последних режима называют режимами направленного округления (*directed rounding*).

Любая реализация стандарта *IEEE* 754 обязана поддерживать режим *round to nearest* в качестве дефолтного (*default settings*).

Основным показателем качества машинной арифметики с плавающей запятой считается точность (*accuracy*), с которой арифметика округляет действительные числа. Чем больше расстояние от исходного числа до ближайшего представимого, тем с меньшей точностью оно может быть представлено.

Расстояние между соседними представимыми числами, т.е. числами с единым десятичным значением порядка и с различающимися в один бит мантиссами, называют шагом числа. Следовательно, максимальная абсолютная погрешность округления для числа в формате *IEEE* 754 равна половине шага. Следовательно, вычисленный результат отличается от точного значения не более, чем на половину единицы последнего разряда мантиссы результата (*unit in the last place*, *ulp*) [11].

Шаг чисел удваивается с увеличением экспоненты двоичного числа на единицу. То есть, чем дальше от нуля, тем шире шаг чисел в формате *IEEE* 754 по числовой оси.

Таким образом, максимальная абсолютная погрешность округления определяется по формуле

$$\Delta^{\max} = 0{,}5ulp\,. \qquad (2)$$

Следовательно, используя формулы (1) и (2), можно получить общую формулу для вычисления максимальной абсолютной погрешности округления чисел в формате *IEEE* 754:

$$\Delta^{\max} = 0{,}5 * 2^{X-b-m} = 2^{X-b-m-1}\,. \qquad (3)$$

Подставив в формулу (3) параметры базовых форматов из табл.1, получим формулы для вычисления максимальной абсолютной погрешности округления чисел, представленных в различных форматах стандарта. Полученные формулы сведены в табл.2.

Для повышения точности вычислений при работе с маленькими числами в стандарте предусмотрена возможность использования так называемых субнормальных (*subnormal numbers*), т.е. ненормализованных чисел.

Описание машинного представления и особенностей использования субнормальных чисел в стандарте *IEEE* 754, а также получение формул для вычисления и сам расчет граничных значений (максимальных и минимальных) субнормальных чисел в десятичной системе счисления для различных форматов авторами данной статьи рассмотрены в работе [9].

Т.к. смещенный порядок для всех форматов любого субнормального числа одинаков и равняется единице в двоичном коде, то формула для вычисления максимальной абсолютной погрешности округления субнормальных чисел будет иметь следующий вид:

$$\Delta^{\max}_{\text{subnormal}} = 0{,}5 * 2^{1-b-m} = 2^{-b-m}\,. \qquad (4)$$

Также, подставив в формулу (4) параметры базовых форматов из табл.1, получим формулы для вычисления максимальной абсолютной погрешности округления субнормальных чисел, представленных в различных форматах стандарта. Полученные формулы сведены в табл.2.

Таблица 2. Формулы вычисления максимальных абсолютных погрешностей округления чисел, представленных в базовых форматах стандарта *IEEE*

Формат	Субнормальные числа	Нормализованные числа
single	2^{-150}	2^{X-151}
double	2^{-1075}	2^{X-1076}
extended	2^{-16447}	$2^{X-16448}$
quadruple	2^{-16495}	$2^{X-16496}$

Вывод. Очень важной проблемой существующей машинной арифметики является округление чисел с плавающей запятой, поскольку числа приходится округлять как при вводе исходных значений, так и после каждой арифметической операции. Стандарт *IEEE* 754 вводит следующие обязательные к реализации округляемые операции с плавающей запятой: сложение, умножение, вычитание, деление, вычисления остатка от деления, извлечение квадратного корня, преобразование форматов. Рассмотренные особенности стандартных способов округления чисел с плавающей запятой при вычислениях на компьютере и полученные формулы для вычисления максимальных абсолютных погрешностей округления чисел, представленных в базовых форматах стандарта *IEEE* 754, позволяют провести анализ вычислительных погрешностей.

Основным средством выполнения расчетов, критичных к округлениям, является повышение точности арифметических операций за счет увеличения разрядности чисел. Современные процессоры общего назначения и компиляторы языков высокого уровня поддерживают, как правило, лишь форматы одинарной и двойной точности, а также не определенный стандартом формат двойной расширенной точности, в котором мантисса состоит из 64 бит. Арифметика повышенной точности требует значительных затрат и поэтому реализуется в программном обеспечении.

Постоянное присутствие ошибок округления при работе с машинной арифметикой предъявляет особые требования к вычислительным алгоритмам и требует дополнительного анализа решаемой задачи. Т.к. практически все числа представляются в памяти ЭВМ с погрешностью, то необходимо знать насколько решение чувствительно к изменениям параметров задачи.

ЛИТЕРАТУРА

1. IEEE Standard for Floating-Point Arithmetic. – New York, 2008.– 70p.
2. Some Notes for a Proposal for Elementary Function Implementation in Floating-Point Arithmetic / G. Hanrot, V. Lefevre, J.-M.Muller, N. Revol and other // Proc. of Workshop IEEE 754 and Arithmetic Standardization, in ARITH-15.– 2001.
3. Proposal for a standardization of mathematical function implementation in floating-point arithmetic / D. Defour, G. Hanrot, V. Lefevre, J.-M.Muller and other // Numerical Algorithms.– 2004.– №37(1–4).– P.367–375.
4. Кулямин В.В. Формальные подходы к тестированию математических функций / В.В. Кулямин // Труды ИСП РАН.– 2006.– №10.– С.69-114.
5. A Test Generation Framework for Datapath Floating-Point Verification / M. Aharoni, S. Asaf, L. Fournier, A. Koifman and other // IEEE International High Level Design Validation and Test Workshop.– 2003.– P.17–22.
6. Stehle D. Searching Worst Cases of a One-Variable Function Using Lattice Reduction / D.Stehle, V.Lefevre, P. Zimmermann // IEEE Transactions on Computers.– 2005.– №54(3).– P.340–346.
7. Ніконов О.Я. Оцінка точності обчислень спеціальних функцій при розробці комп'ютерних програм математичного моделювання / О.Я. Ніконов, О.В. Мнушка, В.М. Сав-ченко // Вісник НТУ «ХПІ». Тематичний випуск: Інформатика і моделювання.– Харків: НТУ.– 2011.– №17.– С.115-121.
8. Жульковская И.И. Алгоритм формирования машинного представления числовых данных в формате с плавающей запятой / И.И. Жульковская, О.А. Жульковский // Математичне моделювання.– Дніпродзержинськ: ДДТУ.– 2013.– №2 (29).– С.69-72.
9. Жульковская И.И. Вычисление граничных значений субнормальных чисел в IEEE-стандарте / И.И. Жульковская, О.А. Жульковский, Р.Г. Шаганенко // Математичне моделювання. – Дніпродзержинськ: ДДТУ. – 2015.– №1 (32).– С.41-44.
10. Жульковская И.И. Вычисление граничных значений действительных числовых данных в IEEE-стандарте / И.И. Жульковская, О.А. Жульковский, Ю.В. Николаенко // Зб. наук. праць ДДТУ (технічні науки).– Дніпродзержинськ, ДДТУ.– 2015.–№.1 (26).– С.240-245.
11. Goldberg D. What Every Computer Scientist Should Know about Floating-Point Arithmetic / D. Goldberg // ACM Computing Surveys.– 1991.– №.23(1).– P.5-48.

пост. 07.12.2015

Математическая модель тепло- и массопереноса в первом периоде сушки плоского слоя дисперсного материала

А.Ф. РЫЖОВ, Н.С. МИЛАШЕНКО

Днепродзержинский государственный технический университет

Представлена аналитическая методика расчета температурного поля и кинетики испарения влаги при сушке термически “массивного” плоского слоя влажного дисперсного материала.

Представлена аналітична методика розрахунку температурного поля і кінетики випаровування вологи при сушінні термічно "масивного" плоского шару вологого дисперсного матеріалу.

The analytical method for calculating the temperature field and the kinetics of evaporation of moisture during drying heat "massive" flat layer of moist particulate material.

Введение. Выбор рациональной технологии сушки материалов требует знаний его температурного поля, особенно в первом периоде сушки, в котором закладываются основы качества готовой продукции.

При нагреве “массивного” слоя влажного материала ($\mathrm{Bi} > 0{,}1$), когда существует значительный перепад температур по его толщине, происходят сложные процессы тепло - и массообмена, которые взаимно влияют на энтальпию нагреваемого материала и интенсивность испарения влаги.

Применяемые термографические методы для исследования процессов тепломассопереноса при сушке материалов [1] дают возможность получения большой информации из одного опыта (термограмма, кривая массы испаряемой влаги, кривая скорости изменения массы влаги). Вместе с тем, отсутствие разработанной теории этих методов не позволяет использовать весь объем информации, который содержат кинетические кривые, регистрируемые в опыте.

Целью работы является получение удобных для инженерной практики аналитических зависимостей, определяющих температурное поле влажного материала и интенсивность испарения влаги в первом периоде сушки, используя данные термографического метода исследования.

Постановка задачи. Рассматривается процесс радиационно-конвективного нагрева плоского слоя влажного дисперсного материала толщиной R при