УДК 004-62,51-37

Раскрытие и слияние логарифмических и степенных функций при компьютерном упрощении символьных выражений

©А.С. Епифанов, А.М. Никифоров

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

Проанализированы сложности, возникающие при почленном логарифмировании и возведении в рациональные степени символьных комплексных произведений («раскрытие») и при обратных преобразованиях («слияние»). Применительно к использованию в малых системах компьютерной математики предложены легко программно проверяемые достаточные критерии, обеспечивающие корректность выполняемых преобразований. Рассмотрен метод упрощения символьных выражений, основанный на раскрытии и последующем слиянии логарифмических и степенных функций. Обсуждаются свойства сигнатуры комплексного переменного в связи с ее использованием в рассматриваемых преобразованиях. Примеры использования предлагаемых методов с помощью программы ПифМат доступны на сайте http://[*PifMath.ru*](http://PifMath.ru).

Ключевые слова: *символьные вычисления, компьютерная математика, логарифмическая функция комплексного переменного, степенная функция, упрощение выражений.*

Задача упрощения символьных выражений весьма трудна и, главное, во многих смыслах неоднозначна. Вполне определенной можно считать только так называемую проблему «эквивалентности нулю», заключающуюся в установлении факта, что данное выражение упрощается к нулю. Д. Ричардсон [1] показал, что не существует единого алгоритма, решающего эту проблему, даже для класса выражений, составленных только из рациональных чисел, символа x, действительных констант π и $ln(2)$, синуса, экспоненты, абсолютной величины, возведения в целочисленную степень, соединяемых знаками сложения и умножения. Хотя для некоторых других случаев решения могут быть найдены (см., например, [2]), их практический смысл невелик, поскольку обычно требуется, так или иначе, именно упростить символьное выражение.

Если речь идет не об эквивалентности нулю, то сразу возникает вопрос, что значит «упростить», поскольку некоторая форма выражения может быть проще, например, по числу арифметических операций, необходимых для получения числового результата, но при этом хуже с точки зрения получающейся погрешности; немаловажными часто оказываются наглядность и чисто эстетическое восприятие. Однако существуют преобразования (вынесение общих множителей за скобку, сокращение символьных рациональных выражений и т.п.), безусловно упрощающие исходное выражение. Здесь мы под упрощениями будем иметь в виду именно такие преобразования.

Малые системы компьютерной математики (МСКМ), получающие в последнее время все более широкое распространение, предъявляют к используемым алгоритмам особые требования с точки зрения простоты – в их понимании, реализации и использовании. Это относится и к алгоритмам упрощения символьных выражений. Желательно, чтобы МСКМ предлагала упрощенное выражение по умолчанию, без использования каких-либо дополнительных операторов, поэтому использование сложных, хотя и более общих, алгоритмов, часто бывает нежелательным.

Последовательность упрощающих преобразований «*раскрытие – упрощение – сжатие»* восходит к приемам простейшей арифметики: раскрыть скобки, привести подобные члены, вынести за скобку общие сомножители или, возможно, факторизовать получившееся выражение. Аналогично упрощаются и произведения, только в роли приведения подобных членов выступает сокращение:

$$xy\sqrt{u}\cdot\sqrt{\frac{v}{x^2y^2}} \to \frac{xy\sqrt{u}\sqrt{v}}{xy} \to \sqrt{u}\sqrt{v} \to \sqrt{uv}.$$

Все входящие в этот пример символы (только здесь!) представляют положительные (стало быть, действительные) переменные или выражения.[1] При нарушении этого требования показанные преобразования, вообще говоря, *не верны.* Это связано с тем, что интересующие нас функции не однозначны. Поскольку любые упрощенные символьные выражения, в конечном счете, предназначены для численных расчетов, требующих однозначных числовых результатов, МСКМ обычно имеют дело с *главными ветвями* этих функций (иначе, с их *главными* значениями).

Определение степенной функции в общем случае дается через логарифмическую функцию, поэтому сначала коротко напомним свойства главной ветви логарифмической функции комплексного переменной

$$z = x + iy = r\cdot\exp(i\varphi). \qquad (1)$$

Здесь $x = Re(z), y = Im(z), r = |z|$ и $\varphi = \arg(z)$ – соответственно действительная, мнимая, модуль и аргумент (фаза) этого числа.[2]

Для краткости величину φ, удовлетворяющую условию

$$-\pi < \varphi \le \pi \qquad (2)$$

мы будем называть приведенной.

По определению, главная ветвь логарифма есть

$$\ln(z) = \ln(|z|) + i\cdot\arg(z), \qquad (3)$$

причем аргумент должен быть *приведен*, т.е. должен удовлетворять условию (2):

[1] Выражение называется действительным (положительным и т.п.) если оно остается таковым при любых допустимых значениях его символов.

[2] Мы зарезервируем буквы $x.y$ и p,q,r за действительными, а m и n – целыми переменными и выражениями.

$$-\pi < \arg(z) \leq \pi. \qquad (2a)$$

Так определенная функция однозначная, всегда обратна экспоненте в том смысле, что

$$\exp(\ln(z)) = z,$$

однако равенство $\ln(\exp(x + iy)) = x + iy$ справедливо, только если величина y приведена.

Достигнутая однозначность сразу приводит к некоторым неприятностям. Отрицательная часть действительной оси начинает играть особую роль (говорят, что по ней проходит *разрез* на комплексной плоскости): при пересечении ее аргументом многозначный логарифм переходил бы на следующую ветвь, а требование (2a) для сохранения главной ветви приводит к скачку логарифма – он перестает быть непрерывным на ней, так для $x < 0$

$$\ln(x) = \ln(|x|) + i\pi, \ \ln(x + i0_+) = \lim_{y \to +0} \ln(x + iy) = \ln(|x|) + i\pi ,$$

но

$$\ln(x + i0_-) = \lim_{y \to -0} \ln(x + iy) = \ln(|x|) - i\pi .$$

Первые два равенства указывают на непрерывность «против часовой стрелки», а последнее – на разрыв со стороны отрицательных значений мнимой части аргумента логарифма. Введенные здесь обозначения «нуля со знаком» [3] отражают один из возможных путей решения этой проблемы, как и осложнений из-за того, что $\ln(z^*) \neq (\ln(z))^*$ и $\ln(1/z) \neq -\ln(z)$ только на линии разреза. Однако во многих случаях нарушение этих равенств только в одномерной области комплексной плоскости либо не влияет на корректность результата,[1] либо сравнительно легко может быть проверено.

Гораздо сложнее дело обстоит с нарушением привычных соотношений

$$\ln(uv) = \ln(u) + \ln(v) \qquad (4)$$

и

$$\ln(u^p) = p \cdot \ln(u). \qquad (5)$$

Если какое-либо преобразование выводит аргумент из области $(-\pi, \pi]$, то значение логарифма переходит на другую ветвь и результат должен быть скорректирован прибавлением слагаемого $2\pi i n, n = \pm 1, \pm 2, \ldots$ так, чтобы новый аргумент был приведен.

Операцию приведения действительного выражения мы будем обозначать квадратными скобками. Заметим, что в записи комплексного числа (1) аргумент φ всегда считается приведенным, поэтому все комплексные величины *перед любой операцией* должны быть приведены. Пусть, например, вычисляется логарифм от экспоненты $\exp(ix)$. Необходимость приведения означает, что фактически будет вычисляться

$$\ln(\exp(i[x]) = i[x] \neq ix$$

согласно сказанному выше.

Пусть, далее, перемножаются два числа

---

[1] В системах компьютерной математики желательна опция, разрешающая игнорировать некорректность результата на линиях разреза. В ПифМате такой опцией служит директива $si+

$u = |u| \cdot \exp(i\varphi_1)$ и $v = |v| \cdot \exp(i\varphi_2)$: [1] (6)

$z = uv = |u||v| \cdot \exp(i(\varphi_1 + \varphi_2))$.

Тогда $\ln(z) = \ln(|u||v|) + i[\varphi_1 + \varphi_2]$. Произведение положительных чисел всегда, очевидно, приведено, и мы приходим к выводу, что формула (4) верна, только если

$$[\varphi_1 + \varphi_2] = \varphi_1 + \varphi_2, \tag{7}$$

т.е. если операция приведения не требуется, поскольку *сумма аргументов* не выходит из области (2). Аналогично, равенство (5) справедливо, если

$$-\pi < p \cdot \arg(z) \leq \pi.$$

Поскольку общее определение возведения в степень комплексного числа[2]

$$z^p = \exp(p \cdot \ln(z)) \tag{8}$$

содержит логарифмическую функцию, то степенная функция наследует ту же линию разреза и те же осложнения при выделении ее главной ветви. Так, равенство

$$(uv)^r = u^r v^r \tag{9}$$

справедливо, если сумма $\arg(u) + \arg(v)$ удовлетворяет условиям

$$-\pi < \arg(u) + \arg(v) \leq \pi \tag{10}$$

или, что то же, условию (7) с учетом представлений (6).

Кроме того, равенство (9) справедливо, если r – целое число, так как в этом случае приведение сводится к добавлению к аргументу экспоненты (8) числа $2\pi inr$, не изменяющему ее значение.

Преобразования (4), (5) и (9), выполняемые слева направо мы будем для краткости называть раскрытием или разложением (expansion), а в обратном направлении – слиянием или сжатием (contraction). В отсутствие других упрощений слияние приводит к упрощению выражения и, вообще говоря, должно выполняться после раскрытия.[3]

Возведение степенной функции в степень также возможно не всегда. В отличие от основного свойства экспоненты ($e^{u+v} = e^u \cdot e^v$) равенство $(e^z)^p = e^{pz}$ верно, только если $-\pi < Im(z) \leq \pi$, т.е. мнимая часть z приведена; формула возведения степени в степень

$$(z^p)^r = z^{p \cdot r} \tag{11}$$

справедлива, если

$$[p \cdot \arg(z)] = p \cdot \arg(z), \text{ т.е. } -\pi < p \cdot \arg(z) \leq \pi, \tag{12}$$

или если показатель степени r – целое число (тогда $e^{i2\pi \cdot r \cdot n} = 1$). Это условие позволяет трактовать возведение числа в произвольную рациональную степень n/m как возведение m-го корня из этого числа в целую степень n:

$$z^{n/m} = (z^{1/m})^n = (\sqrt[m]{z})^n.$$

---

[1] Мы будем считать, что аргументы операндов всегда уже приведены, поскольку операция приведения не изменяет значение экспоненты: $\exp(z + 2\pi in) = \exp(z)$.

[2] Часто возведение комплексного числа в целую степень определяют через умножение, но легко убедиться, что это одно и то же.

[3] Мы увидим, что в общем случае для упрощения выражений необходимо пытаться выполнить как раскрытие, так и слияние.

Заметим, что из определения (8) и условия (12) сразу видно, что если $|p| < 1$, то преобразования (5) и (11) справедливы, например, [1]

$\ln(\sqrt{z}) = \frac{1}{2}\ln(z)$ и $\ln(1/\sqrt{z}) = -\frac{1}{2}\ln(z)$.

Преобразование (11) обычно выполняется только «слева направо» и выполняется, когда это возможно, на этапе раскрытия.

Упрощение алгебраических выражений, содержащих логарифмические функции и рациональные степени, представляет особую трудность при работе со многими выражениями, в частности, с обратными тригонометрическими и гиперболическими функциями. Суть дела заключается в том, что формулы (4), (5), (9) и (11) чрезвычайно важны: без них часто оказывается невозможным вынести общие множители, сократить дроби (см. вступительный пример), провести факторизацию и т.п. Для того, чтобы обойти описанные выше сложности, были предложены [4] так называемые «раскручивающие числа» (“unwinding number”), как бы вбирающие в себя поправки к перечисленным выше и некоторым другим формулам. Несмотря на некоторую общность, получающиеся выражения вряд ли можно назвать упрощенными, поскольку арифметика этих чисел весьма сложна и мало наглядна.

С другой стороны, оказывается, что эти формулы справедливы во многих практически важных случаях, и задача заключается в том, чтобы сформулировать сравнительно простые, хотя бы и не исчерпывающие, достаточные условия (назовем их «*признаки* раскрываемости») выполнения соотношений (7), (10), (12), без особых сложностей проверяемые программно. Некоторые такие условия рассмотрены и сформулированы в виде теорем в гл. 3 книги [5]. Здесь мы их несколько расширим, а также и обсудим возможности их применения как по отдельности, так и в совокупности.

Мы будем называть *выражениями с определенным знаком* действительные выражения, знак которых при любых входящих в них параметров может быть однозначно определен в рамках данной системы.[2]

Рассматриваемые преобразования представимы в виде *частичного раскрытия*

$$\ln(u^p \cdot v^q \cdot \ldots) \leftrightarrow \ln(u^p) + \ln(v^q) + \ \ldots, \tag{13}$$

$$(u^p \cdot v^q \cdot \ldots)^r \leftrightarrow (u^p)^r \cdot (v^q)^r + \ \ldots \tag{14}$$

или, желательно, *полного раскрытия*

$$\ln(u^p \cdot v^q \cdot \ldots) \leftrightarrow p \cdot \ln(u) + q \cdot \ln(v) + \ \ldots \tag{15}$$

и

$$(u^p \cdot v^q \cdot \ldots)^r \rightarrow u^{pr} \cdot v^{qr} \cdot \ldots \tag{16}$$

---

[1] Мы не будем делать различия между $\sqrt{z}$ и $z^{1/2}$. Кроме того, напомним, что $u^{-p} = (u^p)^{-1}$.

[2] Вообще говорят, что выражение действительно, положительно и т.п., если это верно при любых значениях входящих в него символов. Например, $(x+y)^2 + 1 > 0$ при любых вещественных x и y. Однако фактически такие атрибуты выражений во многом определяются возможностями МСКМ.

где r – действительное,[1] u, v – комплексные, вообще говоря, выражения, а p и q – выражения с определенным знаком.

Далее будем считать выполненными соотношения

$$-1 \le p \le 1,\ p \ne 0 \text{ и } -1 \le q \le 1,\ q \ne 0. \qquad (17)$$

Если дополнительно потребовать, что ни одно из выражений p и q не обращалось в минус единицу, то из допустимости преобразований (13) и (14) следует корректность и корректность преобразований (15) и (16), см. (12).

Рассмотрим основные признаки справедливости преобразований (13) – 16).

1. Если x – действительная величина, то, очевидно, имеют место формулы (о сигнатуре *sgn(x)* см. Дополнение, в частности, формулы (29)):

$$\ln(xz) = \ln(|x|) + \ln(\operatorname{sgn}(x) \cdot z),$$
$$(xz)^r = |x|^r (\operatorname{sgn}(x) \cdot z)^r.$$

Эти формулы особенно полезны, если выражение x имеет определенный знак. При раскрытии всегда сначала необходимо вынести такие сомножители. Например (здесь $a \ge 0, b \le 0$, z – комплексное выражение),[2]

$$(4a(x+y)^2 z)^{1/2} \to 2\sqrt{a} \cdot |x+y| \cdot \sqrt{z},$$
$$\ln((3b^3(x+y)^2 z) \to \ln(3) + 3\ln(-b) + 2\ln(|x+y|) + \ln(-z).$$

Заметим, что одновременно применяются и преобразования (5) и (11) для действительных выражений, т.е. происходит полное раскрытие по типу (15) и (16). В результате задача сводится к анализу выражений вида левых частей (13) и (14) без дополнительных сомножителей с определенным знаком.

Обратное преобразование – слияние – с величинами определенного знака всегда выполняется в последнюю очередь, чтобы не затруднять проверку нижеследующих признаков.

2. Утверждения этого пункта основаны на теореме 3 книги [5] (стр. 64). Здесь мы ее несколько расширим и дадим более удобные для программирования формулировки. Речь пойдет о случаях, когда выражения в левых частях (13) – (16) содержат только два сомножителя, причем выражения u и v имеют вид

$$u = x + z,\ v = y + bz, \qquad (18)$$

где x и y – действительные выражения, b – выражение с определенным знаком.

Общее условие для всех рассматриваемых ниже в этом пункте случаев формулируется так:

Если степени $\boldsymbol{p}$ и $\boldsymbol{q}$ имеют одинаковый знак, то множитель $\boldsymbol{b}$ должен быть меньше нуля, $\boldsymbol{b < 0}$, в противном случае $\boldsymbol{b > 0}$.

---

[1] Поскольку для целых значений r преобразования (14) и (16) всегда справедливы, далее будем считать, что «внешний» показатель степени r – не целый. Кроме того, имея в виду выделение целой части, можно считать, что $0 < r < 1$.

[2] Все приводимы в статье (и многие другие) примеры можно просмотреть с помощью программы PifMath.exe, расположенной на сайте PifMath.ru. Там же находятся файлы с примерами. Текущие примеры содержатся в файле \Additions2\LnsPowers\LnsPowers.mtp.

При выполнении этого условия, если мнимая часть z отлична нуля, то мнимые части выражений u^p и v^q, а с ними и их аргументы, имеют различные знаки, и, стало быть, не могут при умножении привести к нарушению соотношений (7). Кроме того, поскольку равенства $\ln(z^{-1}) = -z$ и $(z^{-1})^r = z^{-r}$ нарушаются только на линии разреза, то корректны и преобразования (15) и (16). Сформулируем этот факт как первый вариант достаточного признака раскрываемости.

а) *Если выполнено общее условие и либо разрешено игнорировать значения на линиях разреза, либо программно устанавливается, что $Im(z) \neq 0$, то преобразования (15) и (16) корректны.*

Например, равенство $((x+i)^p(x-2i)^q)^r = (x+i)^{rp}(x-2i)^{rq}$ заведомо верно, если p и q имеют одинаковый знак.

Нам осталось изучить ситуацию, когда мнимая часть выражения z обращается в нуль, т.е. когда произведение $u^p \cdot v^q$ сводится к произведению $x^p \cdot (y+ba)^q$, где все символы соответствуют действительным выражениям. Если хотя бы одно из выражений x и $(y+ba)$ положительно, то соответствующую степенную функцию можно вынести можно вынести, как в п.1, так что преобразования (13) и (14) оказываются справедливыми. Если опции пользователя разрешают игнорировать ошибки на линиях разреза или ни один из показателей степени p и q не равен -1, то можно использовать и преобразования (15) и (16).

Сначала исключим из рассмотрения случай, когда хотя бы один из показателей степеней равен минус единице, т.е. положим, что

$$-1 < p \leq 1,\ p \neq 0 \text{ и } -1 < q \leq 1,\ q \neq 0.$$ [1]

Тогда нам достаточно исследовать только случай, когда оба выражения (18) отрицательны.

б) *Если показатели степени p и q имеют противоположные знаки, то никакого дополнения к основному условию не требуется.*

В самом деле, в этом случае, очевидно, в чуть измененной форме выполняется условие (10):

$$-\pi < p \cdot \arg(u) + q \cdot \arg(v) \leq \pi.$$

Пример. Равенство

$$\ln\left((x+z)(y+az)^{-\frac{2}{3}}\right) = \ln(x+z) - \frac{2}{3}\ln(y+az)$$

верно, если $a > 0$.

в) *Если показатели степени p и q имеют одинаковые знаки, то для справедливости преобразований (13) – (16) достаточно дополнительно потребовать, чтобы*

[1] Если какой-либо из показателей степеней обращается в ноль, то все преобразование сводится к преобразованиям (5) или (11), которые, очевидно, корректны, поскольку выполнено условие (12).

либо чтобы выражение z было чисто мнимым,[1] *а хотя бы одно из выражений x, y имело определенный неотрицательный знак,*

либо чтобы выполнялось условие (см. (18))

$$-bx + y \geq 0, \tag{19}$$

В этом случае недопустимо, чтобы обе величины u и v были отрицательны, и эту возможность надо исключить, т.е. *не* допустить, чтобы одновременно выполнялись неравенства[2]

$$x + z < 0,\ \ y + bz < 0.$$

Это, очевидно, гарантирует первый вариант условия, кроме того, из этих соотношений[3] следует неравенство $-bx + y < 0$, что и доказывает достаточность условия (19).

Пример. Если $c > 0$, то

$$\ln((-c + 2z)(c - z)^{2/3}) = \ln(-c + 2z) + (2/3)\ln(c - z).$$

Обратимся к случаю, когда один из показателей степени равен минус единице, пусть для определенности $q = -1$, т.е. речь идет о выражениях типа

$$\ln((x + z)^p (y + bz)^{-1}),\ \ ((x + z)^p (y + bz)^{-1})^r.$$

Как и в двух предыдущих пунктах достаточно рассмотреть ситуацию, когда величина z действительна.

Если и показатель степени $p = -1$, то

если оба выражения (18) могут быть отрицательными, то ни одно из преобразований (13) – (16) не справедливо,[4]

если не положительно только одно из них, то возможно преобразование лишь по пункту 1, т.е. некоторое сочетание преобразований (13), (14) и (15), (16). Например, если $u > 0$, то

$$\ln(u^{-1}v^{-1}) = -\ln(u) + \ln(v^{-1}).$$

Пусть теперь $p \neq -1$. Сначала положим, что также $p \neq 1$. Если $v > 0$, то преобразования (15) и (16) справедливы по п.1. Поэтому сконцентрируем внимание на случае $v < 0$. Если $u > 0$, то справедливо только преобразование по п.1, если же $u < 0$, то необходимо рассмотреть два случая.

1) $0 < p < 1$. Тогда $\arg(-u^p) = \arg(u^p) - \pi = p\pi - \pi,$ если.

С одной стороны,

$$\ln(u^p v^{-1}) = \ln(-u^p |v|^{-1}) = \ln(|u|^p |v|^{-1}) + i(p\pi - \pi),$$

а с другой –

$$\ln(u^P) - \ln(v) = \ln(|u|^p) + ip\pi - (\ln(|v|) + i\pi),$$

---

[1] Отличие от признака 2а заключается в том, что чисто мнимое выражение z здесь может обращаться в ноль.

[2] Напомним, что в рассматриваемом случае z действительно и $b < 0$ по общему условию.

[3] На практике возможность проверить оба эти неравенства представляется очень редко.

[4] Если оба выражения заведомо отрицательны ($u < 0, v < 0$), то возможны, конечно, преобразования по п.1, например:
$$\ln(u^{-1}v^{-1}) = \ln((-u)^{-1}(-v^{-1})) = -\ln(-u) - \ln(-v)\ .$$

Поскольку правые части, очевидно, равны, то оказывается справедливым преобразование (15) и, аналогично, (16). Происходит компенсация двух ошибок: величина

$$\arg(u^p) + \arg(v^{-1}) = p\pi + \pi$$

требует приведения вычитанием 2π, котрое и происходит из-за того, что $-\ln(v) - \ln(v^{-1}) = -2\pi i$ на линии разреза.

2) $-1 < p < 0$. В этом случае никакой компенсации не происходит, и преобразования (15) и (16) не верны ни для положительных, ни для отрицательных значений u.

Все рассмотренные варианты для $q = -1$ и $p \neq 1$, как мы видим, требуют тщательного анализа знака действительных частей выражений u и v при условии обращения в ноль мнимой части z и поэтому малопригодны для реализации в малых программах. В тоже время, ситуации, когда $q = -1$ и $p = 1$, часто встречается и допускает достаточное условие, которое мы сейчас и сформулируем.

г) *Если $p = 1$ и $q = -1$, то для справедливости преобразований (15) и (16) достаточно дополнительно к общему условию $b > 0$ потребовать, чтобы выполнялось неравенство* (см. (18))

$$-bx + y \geq 0.$$

Это условие совпадает с (19), но напомним, что величина b здесь должна быть *больше* нуля.

Выпишем явно формулы преобразований (15) и (16), которые нас интересуют:

$$\ln\left((x+z)\,(y+bz)^{-1}\right) \leftrightarrow \ln(x+z) - \ln(y+bz), \tag{20}$$

$$\left((x+z)\,(y+bz)^{-1}\right)^r \leftrightarrow (x+z)^r (y+bz)^{-r}. \tag{21}$$

Как и раньше, достаточно рассмотреть случаи, когда выражения u и v (18) оказываются действительными. Если $v = y + bz > 0$, то написанные формулы, очевидно, верны по п.1. Если $v < 0$, и $u = x + z < 0$, то происходит компенсация ошибок, с которой мы уже встречались. В самом деле, в левых частях равенств (20) и (21) оказываются действительные числа, а в правых – фазовые множители взаимно уничтожаются $(i\pi - i\pi)$ для логарифма и сокращаются $(\exp(ir\pi)\cdot\exp(-ir\pi))$ для степенных функций. Однако частичные преобразования (13) и (14), конечно, не верны. Наконец, если $v < 0$, и $u = x + z > 0$, то ситуация оказывается противоположной: частичные преобразования (13) и (14) верны, а полные (15) и (16) – нет. Мы видим, что для достижения единообразия и, имея в виду, что полные преобразования обычно гораздо полезнее для дальнейших упрощений, желательно исключить последний случай. Из соотношений

$$x + z > 0,\ \ y + bz < 0$$

следует неравенство $-bx + y < 0$, что и доказывает сформулированное достаточное условие.

Пример. При $a > 0$ справедливо равенство

$$\ln\left((x+z)(2x+a+2z)^{-1}\right) = \ln(x+z) - \ln((2x+a+2z).$$

Многие произведения двух сумм умножением их на -1 или i, $-i$ приводятся к виду, удовлетворяющему признакам а), в) или г). Например,

$$\ln((a - iz)(-2a + iz)) = \ln((-a + iz) + \ln(2a - iz).$$

Здесь проверка по п. в) (см. (19)) дает отрицательную сумму и сразу показывает, что знаки в суммах надо изменить на противоположные. Еще один пример, связанный с умножением на $\pm i$:

$$((ia + z)(2ia + z))^{1/2} = (-a + iz)^{1/2}(2a - iz)^{1/2}.$$

С последней возможностью связана и некоторая неприятность – возникающая порой неоднозначность раскрытия. Например, равенство

$$\ln((x + i)(x - 2i) = \ln(x + i) + \ln(x - 2i)$$

верно по п.2-а), а равенство

$$\ln\big((1 - ix)(2 + ix)\big) = \ln(1 - ix) + \ln(2 + ix)$$

– по п.2-в), но их левые части, как легко проверить, одинаковы. Эта неоднозначность может повлиять на дальнейшие упрощения, например, сумма $\ln((x + i)(x - 2i)) - \ln(x + i) + \ldots$ упростится только при выборе первого равенства в качестве формулы раскрытия.

Существует несколько возможностей преодолеть это затруднение. Во-первых, можно сделать доступ процедуры раскрытия логарифма или степени к окружающему их выражению с тем, чтобы процедура, упрощающая его, сама сделала выбор. Во-вторых, процедура раскрытия может возвращать дополнительный вариант результата. Это наиболее общий способ, поскольку выбор может осуществлять не только процедура, обрабатывающая окружающее выражение (сумму для логарифмической и произведение для степенной функций), но и ее предшественники – ведь могут быть возможны различные варианты сокращения. Наконец, как уже говорилось, в конце расширенного упрощения желательно производить слияние, после которого возможно сокращение без каких-либо неоднозначностей. Для достижения лучших результатов все суммы должны быть нормированы, например, так, чтобы выделяемые в качестве x и y (см. (18)) выражения имели, если это возможно, определенные знаки. Все эти способы не лишены недостатков: первые два из-за дополнительных сложностей, а последний – потому, что слияние может и не получиться.

Обрисуем последовательность операций для раскрытия выражений

$\ln(u^p \cdot v^q)$ и $(u^p \cdot v^q)^r$,

где выражения u и v – символьные суммы.

Процедуры раскрытия логарифмической функции или произведения вызывают общую процедуру анализа, передавая ей по ссылке (параметры-переменные) два выражения u и v, по значению степени p и q и в случае положительного ответа осуществляет разложение. Здесь мы остановимся только на процедуре анализа. Эта процедура возвращает либо положительный (тогда выражения u и v могут быть изменены), либо отрицательный результат с неизменными u и v.[1]

Чтобы расширить область применения рассматриваемой процедуры на случай, когда произведение имеет общий числовой множитель, дополнительно в нее желательно передавать этот множитель (по ссылке, потому что он может быть изменен). На всех последующих шагах можно либо его использовать, либо игнорировать, пытаясь достичь желаемого результата – выполнения одного из перечисленных выше признаков.[2]

1) Во избежание дополнительных вычислений проверяем, что оба выражения u и v имеют одинаковый комплексный тип.[3] Если этот тип – неопределенный, то вероятность успеха мала. Убеждаемся, что значения показателей степеней удовлетворяют неравенствам (17) одному из условий а) – г). Запоминаем подходящий пункт (в частности, знаки p и q).

2) Суммы представляем в форме $u = x + z_1$ и $v = y + z_2$. Здесь x и y – неполные действительные части u и v, выделяемые по следующим правилам (например, для суммы u).

Если сумма u действительна, то выражение x должно содержать все слагаемые с определенным знаком. Это позволяет раскрыть выражения типа

$\ln(1 + 2y)(1 - y))$.

В противном случае в выражение x включаются все действительные слагаемые u. Действительная часть комплексных переменных, функций и т.п. не выделяется.

3) Находим отношение $b = z_2/z_1$, символьную дробь упрощаем, в частности сокращением.

4) Проверяем выполнение общего условия, исходя из результатов шага 1). Если получается определенный, но противоположный по знаку результат, то в ситуации в) с $p = q = 1$ умножаем u и v соответственно на i и $-i$ (или наоборот, если не угадаем, то скорректируем позже, на шаге 6, хотя иногда можно сразу сделать правильный выбор), в остальных случаях – неудача.

5) В ситуациях а) и б) возвращаем положительный результат.

6) В ситуациях в) и г) дополнительно проверяем условие (19). При этом если это условие не выполнено, но знак выражения $-bx + y$ определен как отрицательный, то в ситуациях в) с $p = q = 1$ и г) возвращаем выражения u и v, умноженные на -1.

[1] Поэтому производить изменения, не приводящие заведомо к успеху, нужно с независимыми их копиями.

[2] Мы описали выше умножение на множители $-1, i, -1$, не меняющие выражение, но можно это делать и компенсируя изменения произведения сумм, изменяя передаваемый множитель.

[3] Комплексные типы: *неопределенный* (выражение содержит переменные с неопределенным типом), безусловно *действительный*, безусловно *мнимый*, *комплексный* – в остальных случаях. Уже на этом этапе можно учитывать дополнительный числовой множитель, позволяющий модифицировать тип одного из сомножителей.

В некоторых случаях на шаге 4 в ситуации в) с $p = q = 1$ знак b может быть не определенным (из-за комплексности этого выражения), но умножение сумм на i и $-i$, спасает ситуацию, например $(a > 0)$,

$$\ln\big((x + ia + z)(x + 2ia + z)\big) = \ln(-a + ix + iz)(2a - ix - iz)),$$

и последний логарифм раскрывается по п. в). Поэтому в этой конкретной ситуации разумно делать вторую попытку, начиная с шага 2, умножив предварительно u и v на на i и $-i$ соответственно (или используя переданный дополнительный числовой множитель).

3. Иногда оказывается возможным сделать преобразования (13) и (14) или (15) и (16), используя лишь оценку предельных значений аргументов выражений, возводимых в степень (или стоящих под знаком логарифма). Укажем некоторые простые оценки φ_{min} и φ_{max} этих значений (при этом необходимо указывать, достигается ли нижняя грань φ_{min}, чтобы проверить выполнение соотношений (2)).

Для чисел, констант и т.п.: $\varphi_{min} = \varphi_{max} = \arg(\text{этого числа})$, φ_{min} достигается.

Для степеней z^p $(-1 \le p \le 1)$:

Пусть $\varphi_{min,base}$ и $\varphi_{min,base}$ – предельные значения аргумента выражения z, и $\varphi_{base} = \max(|\varphi_{min,base}|, |\varphi_{max,base}|)$. Тогда $\varphi_{min} = -|p|\varphi_{base}$, $\varphi_{max} = |p|\varphi_{base}$, причем φ_{min} достигается при $p \le 0$ и $\varphi_{base} = \pi$.

Для произведений $\prod u_k$: $\varphi_{min} = \sum \varphi_{min,k}$, $\varphi_{max} = \sum \varphi_{max,k}$; φ_{min} не достигается, если не достигается хотя бы одно из значений $\varphi_{min,k} \ne 0$.

Для сумм $\sum u_k$:

Сначала заметим, что в рассматриваемых преобразованиях суммы, для которых надо оценивать предельные аргументы появляются в степенных функциях с показателем степени, по модулю большим единице (см. (11), (12)), например

$$\ln(z^2) \text{ или } \sqrt{z^2}, \tag{22}$$

где $z = \sum u_k$.

В приведенных примерах раскрытие возможно, если сумма принадлежит правой комплексной полуплоскости, т.е. ее аргумент находится в пределах

$$-\pi/2 < \arg(z) \le \pi/2. \tag{23}$$

Мы ограничимся случаем, когда эти неравенства выполнены,[1] причем не только для всей суммы, но и для каждого слагаемого. Тогда можно считать, что $\varphi_{max} = \max(\varphi_{max,k})$ и $\varphi_{min} = \min(\varphi_{min,k})$, и φ_{min} достигается, если он достигается хотя бы одним слагаемым.

Примеры.

$$\ln(e^{\frac{i\pi}{4}}\sqrt{z}) = \frac{i\pi}{4} + \frac{1}{2}\ln(z),\; \left(\sqrt{z} \cdot \sqrt{w}\right)^x = z^{x/2}w^{x/2},$$

[1] Если показатель степени по модулю больше двух, то это действительно так, в противном случае наши требования оказываются слишком жесткими, но никакой простой оценки для φ_{min} и φ_{max} получить не удается.

$$\ln\left(\left(1+\sqrt{w}+\sqrt{z}\right)^2\right) = 2\ln(1+\sqrt{w}+\sqrt{z}).$$

Здесь использованы формулы (4), (9) и (11), и тот факт, что $-\pi/2 < \arg(\sqrt{z}) \le \pi/2$, так что все условия оказываются выполненными.

4. При упрощении многих выражений, в частности, логарифмических представлений обратных тригонометрических и гиперболических функций, особую роль играют квадратные корни из комплексных величин. Конечно, при этом применимы все описанные выше подходы, но полезными оказываются и некоторые дополнительные формулы. Если выражение z принадлежит правой комплексной полуплоскости (23), то, конечно, справедливы простые равенства для (22)

$$\ln(z^2) = 2\ln(z),\ \sqrt{z^2} = z.$$

Наиболее последовательный способ реализации второго из этих преобразований использует *комплексную сигнатуру* $csgn(z)$ (см. [5], стр. 67-72 и Дополнение), по определению равную $+1$, если z принадлежит правой комплексной полуплоскости (23), и -1 в противном случае. В разных ситуациях ее по-разному доопределяют при $z = 0$. Имеет место основная формула[1]

$$\sqrt{z^2} = z \cdot csgn(z). \qquad (24)$$

Любая система должна поддерживать определение предельных аргументов выражения, хотя бы в простых случаях, в частности устанавливать выполнение неравенств (23). Если же система поддерживает комплексную сигнатуру, то установление неравенств (23) сводится к упрощению (часто даже автоматическому) этой функции.

Для логарифмической функции имеет место менее красивая формула

$$\ln(z^2) = 2\ln(iz) - i\pi \cdot csgn(z). \qquad (25)$$

Если неравенства (23) установить не удается, то наряду с равенствами (24) и (25) часто оказываются полезными формулы

$$\sqrt{z^2} = \sqrt{iz}\sqrt{-iz},\quad \sqrt{-z^2} = \sqrt{z}\sqrt{-z},$$

$$\ln(z^2) = \ln(iz) + \ln(-iz),\ \ln(-z^2) = \ln(z) + \ln(-z). \qquad (26)$$

По существу, эти формулы суть следствия п. 2-в), но оказывается удобнее проверять возможность их использования отдельно, поскольку форма записи сильно отличается от (18). Отметим еще простое преобразование

$$\ln(-z^2) \to \ln((\pm iz)^2) \to 2ln(\pm iz),$$

если выражение $\pm iz$ принадлежит правой комплексной полуплоскости.

5. Так же, как в п. 2, анализируются преобразования

$$\ln((x+z)^p(y-bz^*)^q) \leftrightarrow \mathrm{p}\cdot\ln(x+z) + q\cdot\ln(y-bz^*),$$

$$((x+z)^p(y-bz^*)^q)^r = (x+z)^{pr}(y-bz^*)^{qr}$$

---

[1] О сигнатуре комплексного переменной см. Дополнение. При использовании этой формулы обычно можно положить $csgn(0) = csgn_+(0) = 1$ но при сокращении $\sqrt{z^2/z} = csgn(z)$ нужно считать значение $csgn(0)$ неопределенным. Дополнительные примеры см. в файле \Additions2\LnsPowers\SqRootsC.mtg.

с показателями степени p и q, удовлетворяющими условиям (17).

Мы поставили знак «минус» перед множителем b с тем, чтобы общее условие выглядело бы в точности так же, как в п.2. Это условие достаточно, если мнимая часть z не может обращаться в ноль (см. п.2-а) или показатели степени p и q имеют различные знаки, не обращаясь в минус единицу, (см. п.2б). В остальных случаях программное использование аналога неравенства (19)

$$-bx + y - 2b \cdot Re(z) \geq 0$$

затруднено необходимостью выделять действительную часть выражения z/

Опишем кратко *алгоритм раскрытия логарифмов и степенных функций от произведений* (см. (13) – (16)) с применением вышеперечисленных признаков **1** – **5**.

Анализ произведения требует, вообще говоря, двух проходов. На первом проходе происходит выделение сомножителей с определенным знаком (п.1) и суммируются оценки максимальных аргументов согласно п.3. Если текущий сомножитель (отличный от числовой константы[1]) имеет определенный знак, то он убираются из начального произведения (и, конечно, запоминается в новом произведении). В противном случае сначала с помощью процедуры анализа, описанной выше, для него подбирается пара из оставшихся сомножителей так, чтобы удовлетворялось одно из условий п.2. (в качестве числового множителя передается множитель произведения). Если удается это сделать, то для оценки максимально аргумента используется больший аргумент из этой пары, иначе – аргумент текущего сомножителя. Если оценка суммарного аргумента выходит за допустимые пределы (2), то первый проход прерывается.

Если по завершении первого прохода оценка максимального аргумента не вышла за пределы (2), то сначала проверяется, не портит ли ситуацию (возможно, измененный) числовой множитель, если все нормально, то на втором проходе формируется полностью раскрытое выражение с учетом, возможно, модифицированных пар. В противном случае проверяет выполнение условий п.4 для оставшейся части произведения, позволяющее все еще произвести полное раскрытие. При неудаче к выделенным сомножителям присоединяется оставшаяся часть произведения целиком, и в лучшем случае раскрытие получается неполным.

Следующий признак основывается на теореме 6 книги [5].

6. Пусть $\varphi = \arg(uv)$. Равенства (4) и (9) верны, если произведение uv имеет определенный квадрант на комплексной плоскости и существует комплексное выражение[2] w с определенным квадрантом при выполнении следующих условий.

а) Если $0 < \varphi \leq \pi/2$, то $-\pi/2 \leq \arg(w) \leq 0$ и либо сумма $wu + v$, либо сумма $u + wv$ принадлежит правой комплексной полуплоскости:

---

[1] Предполагается, что произведение автонормировано, поэтому числовая константа в произведении может быть только одна, обычно на первой позиции. Произведение в ПифМате обязательно имеет числовой множитель, даже если он равен единице.

[2] Обычно приводимые здесь условия легко проверяются, только если как произведение uv, так и выражение w суть комплексные числа, однако именно эта ситуация имеет большое практическое значение.

$$-\pi/2 < \arg(wu + v) \le \pi/2 \text{ или } -\pi/2 < \arg(u + wv) \le \pi/2. \qquad (27)$$

В этом случае (произведение uv находится в первом квадранте) условие (7) может нарушаться только, если оба сомножителя принадлежат третьему квадранту. Умножение любого из них на множитель w с аргументом в указанном диапазоне оставляет произведение в левой полуплоскости, как и сложение двух величин из левой полуплоскости, что и доказывает это признак.

б) Если $-\pi/2 < \varphi \le 0$, то $0 \le \arg(w) \le \pi/2$ и выполняется условие (27). Доказательство аналогично, но условие (7) может нарушаться только, если оба сомножителя принадлежат второму квадранту, поэтому изменяются требования к множителю w.

в) Если $-\pi < \varphi \le -\pi/2$, то множитель w действителен и больше нуля, $w > 0$, и либо сумма $w(iu) + iv$, либо сумма $iu + w(iv)$ принадлежит к правой комплексной полуплоскости.

«Плохие» сомножители принадлежать верхней полуплоскости, и умножение на мнимую единицу переводит их в левую комплексную полуплоскость.

г) Если $\pi/2 < \varphi \le \pi$, то $w > 0$, и либо сумма $w(-iu) - iv$, либо сумма $-iu + w(-iv)$ принадлежит к правой комплексной полуплоскости.

Умножением на $-i$ «плохие» сомножители из нижней также переводятся в левую полуплоскость.

Примеры применения этих признаков будут даны в статье, посвященной упрощению выражений, содержащих обратные тригонометрические и гиперболические функции.

7. В некоторых случаях преобразования, рассмотренные в признаке 2, могут быть корректны и при нарушении общего условия. Здесь мы ограничимся двумя важными случаями.

а) Для того, чтобы выполнялись равенства

$$\ln\big((z+a)(z+b)\big) = \ln(z+a) + \ln(z+b),$$

$$((z+a)(z+b))^r = (z+a)^r(z+b)^r,$$

где a, b и r – действительные выражения (r не целое), достаточно, чтобы выражение z принадлежало правой комплексной полуплоскости и чтобы сумма $a + b$ имела определенный неотрицательный знак, $a + b \ge 0$.

Пусть $z = x + iy$, причем (по условию правой полуплоскости) либо $x > 0$, либо $x = 0$, но при этом $y \ge 0$. Тогда

$$Im\big((z+a)(z+b)\big) = 2xy + (a+b)y.$$

Видно, что в сформулированных условиях мнимая часть произведения имеет тот же знак, что и мнимая часть y выражения z, поэтому пересечение линии разреза возможно только в том случае, если перемножением двух сомножителей из нижней комплексной полуплоскости получается отрицательная величина. При этом, очевидно, $y = 0$,[1] но тогда хотя бы одно из чисел $x + a$ и $x + b$ должно быть положительным и, стало быть, выполняется признак 1.

Следствие.

Если величина z принадлежит правой комплексной полуплоскости, то для любого действительного a справедливы равенства

$$\ln(z^2 - a^2) = \ln(z + a) + \ln(z - a),\ \sqrt{z^2 - a^2} = \sqrt{z + a}\sqrt{z - a}.$$

Пример. $\sqrt{\sqrt{z} + \sqrt{a}} \cdot \sqrt{\sqrt{z} - \sqrt{a}} = \sqrt{z - a},\ a \geq 0.$

б) Для того, чтобы были справедливы равенства

$$\ln\big((z + a)(-z + b)^{-1}\big) = \ln(z + a) - \ln(-z + b),$$

$$((z + a)(-z + b)^{-1})^r = (z + a)^r \cdot (-z + b)^{-r},$$

достаточно, чтобы сумма $a + b$ имела определенный положительный знак, $a + b > 0$ [2] и чтобы мнимая часть z не обращалась в ноль, $Im(z) \neq 0$.

Вычисляя мнимую часть произведения, убеждаемся, что его знак определяется знаком выражения $Im(z)(a + b)$, так что при выполнении указанных условий не может быть ни пересечения линии разреза, ни выхода на нее.

Замечание. Если $Im(z) = 0$, то произведение действительно и для установления справедливости рассматриваемых равенств необходимо определение знака действительной части выражения z.

8. При слиянии логарифмов нужно использовать некоторые дополнительные возможности. Здесь мы укажем лишь одну формулу

$$\ln(z) - \ln(-z) = -i\pi \cdot csgn_+(iz), \tag{28}$$

легко проверяемую непосредственным сравнением мнимых частей левой и правой части, Заметим, что из равенств (28) и (26) получаются формула (25).

9. Следующие формулы слияния квадратных радикалов

$$(-z)^{1/2}(z)^{-1/2} = i \cdot csgn(iz),$$

$$z^{1/2}(-z)^{-1/2} = -i \cdot csgn(iz)$$

также проверяются непосредственно.

Важные замечания. При слиянии логарифмов одновременно необходимо выполнять оба преобразования (4) и (5) (справа налево). Если символьная сумма содержит несколько логарифмов, то надо выделять пары для слияния в несколько проходов. Если рассмотрение всех вариантов нежелательно из-за больших затрат времени, то следует придерживаться следующих рекомендаций. На первом проходе игнорируются логарифмы с поло-

---

[1] Если $x = 0$, то величина $z = iy$ при $y < 0$ не принадлежит правой полуплоскости. Однако в этом случае можно разрешить отрицательные значения y, если потребовать выполнение строгого неравенства $a + b > 0$, чтобы произведение не могло быть отрицательным числом.

[2] Если $a + b = 0$ то произведение в левых частях рассматриваемых равно минус единице, поэтому преобразование «слева направо» встретиться не может, а преобразование «справа налево» см. ниже пп. 8 и 9.

жительными аргументами и пары с различными по знаку коэффициентами (p – см. (5)). Продолжать выделение пар надо до тех пор, пока число логарифмов не станет меньше двух или последний (но не первый) проход слияний не выявил. После слияния возникающая под логарифмом дробь может сокращаться, например,

$$\ln(1+x^2) - \ln(1-ix) \to \ln((1+x^2)/(1-ix)) \to \ln(1+ix).$$

В заключение продемонстрируем упрощение последовательностью «раскрытие – упрощение – сжатие» следующим простым примером. Рассмотрим сумму ($a \geq 0$)

$$S = \ln((a+ix)(3a+2ix)^{-1}(z/i)^{1/3}) + \ln((3a+2ix)(2a+ix)^{-1}(iz)^{-1/3}) +$$
$$+(i\pi/3)\sqrt{z^2}/z.$$

Проверить сразу возможность слияния логарифмов не удается, поэтому раскрываем логарифмы по описанному выше алгоритму и квадратный корень по формуле (24):

$$= \ln(a+ix) - \ln(2a+ix) + (1/3)(\ln(-iz) - \ln(iz)) + (i\pi/3)csgn(z).$$

Осталось, в частности используя формулу (28), слить логарифмы, чтобы получить упрощенный результат:

$$S = \ln((a+ix)(2a+ix)^{-1}).$$

Дополнение. Комплексная сигнатура (или, лучше, сигнатура комплексной переменной) – довольно эффективный инструмент для упрощения символьных выражений, содержащих квадратные корни из комплексных величин.[1]

Использование сигнатур сводится к двум основным моментам. Во-первых, к упрощению аргументов сигнатур и их вычислению для возможно более широкого класса символьных выражений. Для действительных функций это сводится к определению знака символьных аргументов, а для комплексных – к оценке предельных значений их фазы. Во-вторых, к реализации арифметики самих сигнатур.

Поскольку работа с комплексными сигнатурами основывается на использовании сигнатур $sgn(x)$ действительного переменного, сначала напомним некоторые свойства этой функции. Она принимает значение $+1$, если $x > 0$ и значение -1, если $x < 0$. Классическое определение устанавливает $sgn(0) = 0$. Однако в гл.2 [5] было показано, что часто бывает разумно сохранить одностороннюю непрерывность сигнатуры в нулевой точке, и были введены две дополнительные функции[2] $sgn_+(0) = 1$ и $sgn_-(0) = -1$.[3]

Это позволяет, например, записать корни квадратного уравнения с действительными коэффициентами p и q

$$x^2 + 2px + q = 0$$

в форме, гарантирующей минимальную погрешность, когда $q > 0,\ q \ll p^2$:

---

[1] Примеры использования сигнатур, например, в разложениях см. в файле /Additions2/Sgns etc/Expansions.

[2] В некоторых системах значение сигнатуры в нуле задается специальными директивами, что, однако, неудобно, если в разных местах одного выражения нужны различные значения sgn(0).

[3] Это значения устанавливаются по умолчанию, однако сохраняются пределы

$$\lim_{x\to -0} sgn_+(x) = -1 \text{ и } \lim_{x\to +0} sgn_-(x) = +1.$$

$$x_1 = -\left(p + sgn_+(p) \cdot \sqrt{p^2 - q}\right),\ x_2 = -q/\left(p + sgn_+(p) \cdot \sqrt{p^2 - q}\right).$$

Главное значение квадратного корня из комплексного числа $z = x + iy$ (x и y действительны и имеют произвольные знаки) записывается так:

$$\sqrt{x + iy} = \sqrt{(|z| + x)/2} + i \cdot sgn_+(y) \cdot \sqrt{(|z| - x)/2}\ .$$

Легко убедиться, что при использовании обычной сигнатуры эти формулы не верны.

В следующих формулах индекс … (многоточие) в левой части равенства означает *любую* из сигнатур, индекс $\pm$ – любую из сигнатур sgn_+ и sgn_-, а в правой части – *ту же* функцию.

Для любых сигнатур имеют место равенства

$$|x| = x \cdot sgn(x),\ x = |x| \cdot sgn(x), \tag{29}$$

справедливы «теорема сложения»

$$sgn_+(x) + sgn_-(x) = 2sgn(x)$$

и «теорема умножения»

$$sgn_\pm(x) \cdot sgn(x) = |sgn(x)|. \tag{30}$$

Кроме того, очень часто используются соотношения симметрии

$$sgn(-x) = -sgn(x),\ sgn_+(-x) = -sgn_-(x),\ sgn_-(-x) = -sgn_+(x) \tag{31}$$

и формулы возведения в целую степень

$$sgn^{2n}(x) = |sgn(x)|\ (n \geq 0);\ \ sgn_+^{2n}(x) = 1 \text{ и } sgn_-^{2n}(x) = 1, \tag{32}$$

$$sgn\ ...^{2n+1}\ (x) = sgn\ ...\ (x).$$

Отметим еще равенства

$$sgn_{...}(sgn_\pm(x)) = sgn_\pm(x),\ sgn_{...}(sgn(x)) = sgn_{...}(x). \tag{33}$$

и «проникновение» сигнатур в некоторые функции, например,

$$sgn(y) \cdot sin(xy) = sin(x \cdot |y|),\ \ sgn\left(x^3 \cdot arsh\left(x \cdot artg(x)\right)\right) = sgn(x).$$

Для упрощения выражений с модулями и сигнатурами важно также следующее «дистрибутивное» свойство

$$sgn(x \cdot y) = sgn(x) \cdot sgn(y), \tag{34}$$

которое не выполняется для остальных сигнатур только при обращении аргументов в ноль. Однако в большинстве случаев *при упрощении выражений* значения сигнатуры в нуле оказываются не существенными (см., например, (29)). Разрешение игнорировать эти значения должно даваться пользователем.[1] При этом резко упрощается арифметика сигнатур, прежде всего дистрибутивное свойство (33) можно использовать без ограничений, правая часть равенства (30) превращается в единицу и т.д., а, например, упрощение

$$sgn_-\left(\sqrt{x^2 + y^2} - |x| - |y|\right) = -1$$

---

[1] В ПифМате для этого используется опция $\$pm\ +$, выставленная по умолчанию. Комбинируя эту опцию с различными сигнатурами можно определить (конечно, в рамках возможностей системы) не только знак символьного выражения, но и может ли оно обращаться в ноль.

становится возможным для всех сигнатур.[1]
По поводу дальнейшей арифметики сигнатур см. гл.2 [5].

Комплексная сигнатура $csgn_+(z)$ определяется так (гл.3 [5]):

$$csgn_+(z) = \begin{cases} sgn(Re(z)), & \text{если } Re(z) \neq 0 \\ sgn_+(Im(z)), & \text{если } Re(z) = 0. \end{cases}$$

Равенство $csgn_+(z) = 1$ задает *правую комплексную полуплоскость* с включенной в нее точкой $z = 0$. Аналогично определяется и функция $csgn\_(z)$ с очевидной заменой $sgn_+(Im(z))$ на $sgn\_(Im(z))$. Что касается третьей сигнатуры, то чаще всего, надо считать ее значение в нуле неопределенным. Вообще же, все сказанное выше о возможности игнорировать значение сигнатуры в нуле еще в большей степени относится к комплексной сигнатуре.

Вычисление сигнатуры от комплексного числа сводится в худшем случае к определению знаков двух действительных чисел, поэтому, например, формула (24) дает безусловное упрощение.

Здесь мы перечислим только основные свойства комплексной сигнатуры, используемые при упрощении символьных выражений, а также некоторые формулы, не упомянутые в [5]. Многие приводимые ниже формул практически очевидны.

Можно явно выразить комплексную сигнатуру через действительные функции:[2]

$$csgn_{...}(x + iy) = sgn(x) + \frac{1}{2} sgn_{...}(y)(sgn_+(x) - sgn\_(x)).$$

Основой для использования комплексной сигнатуры служит равенство (24) и его обобщение

$$(z^2)^{n+1/2} = z^{2n} \cdot (z^2)^{\frac{1}{2}} = z^{2n+1} \cdot csgn(z).$$

Непосредственно из определения комплексной сигнатуры следует, что

$$csgn(ix) = \mathrm{sgn}(x). \tag{35}$$

Формулы (31) – (33) имеют свои очевидные аналоги и для комплексных сигнатур. Дистрибутивное свойство справедливо только, если один из сомножителей действителен:[3]

$$csgn(xz) = sgn(x)csgn(z). \tag{36}$$

В самом деле, если множитель $x > 0$, то его присутствие ничего не меняет, а если $x < 0$, то равенство (36) следует из соотношения симметрии. Так получаются упрощения

$$csgn\left(x^3\sqrt{z}\right) = sgn(x), csgn\left(e^{\frac{2\pi}{3}}x\right) = -sgn(x).$$

На этом же свойстве (36) основываются следующие преобразования. Пусть, как и выше, x – действительное выражение, а k, l, m и n – целые числа. Рассмотрим сигнату-

[1] Другие примеры можно найти в файлах \Additions2\Sgns etc\SgnPolyAbs, \Additions2\Sgns etc\CsgnSqr.
[2] Отметим, что получится $csgn(0) = 0$, хотя часто это значение, как уже говорилось, следует считать неопределенным.
[3] Далее использование функции *csgn(z)* без индексов будет означать, что разрешено игнорировать значение сигнатуры в нуле. Более точные примеры см. в файле \Additions2\Sgns etc\CsgnSqr.mtp.

ру $csgn_{\pm}(x^{m/n})$. Представим показатель степени в виде $m/n = 2k + l/n$, где $l = -n + 1, -n+2, \dots, 0, \dots, n-1$. Это можно сделать, определив целочисленное деление по методу округления: $m\ div\ n = round(m/n)$. Тогда

$$l = m\ mod\ (2n).$$

Используя (36), приведем рассматриваемую сигнатуру к виду $csgn_{\pm}(x^{l/n})$. Если $x > 0$, то она всегда равна единице, а при $x < 0$ – равна единице, если $\arg(x^{l/n}) = l\pi/n$ соответствует правой половине комплексной полуплоскости, и минус единице в противном случае. Таким образом,[1]

$$csgn_{\pm}(x^{m/n}) = \begin{cases} 1, \text{ если } -n < 2l \leq n \\ sgn(x) \text{ в противном случае} \end{cases} \tag{37}$$

Множитель $\sqrt{x}$ при $x > 0$ не меняет направление вектора комплексной плоскости, а при $x < 0$ вращает его на $\pi/2$ против часовой стрелки. Поэтому $csgn\left(e^{i\varphi}\sqrt{x}\right)$ оказывается равным $sgn(x)$, -1, $-sgn(x)$ или 1, когда φ соответственно принадлежит первой, второй, третьей или четвертой четверти, например, $csgn\left(e^{-i\pi/4}\sqrt{x}\right) = 1, csgn\left(i\sqrt{x}\right) = sgn(x)$. Полезными также оказываются аналогичные соотношения для множителя $\sqrt{-x}$. Приведем, наконец, еще одну формулу, проверяемую непосредственно и позволяющую иногда упростить аргумент сигнатуры:

$$csgn\left(i\sqrt{z}\right) = csgn(iz). \tag{38}$$

ЛИТЕРАТУРА

[1] Richardson D. Some undecidable problems involving elementary functions of a real variable. *Journal of Symbolic Logic*, 1968, №33 (4), pp. 511 – 520.

[2] Shackell J. Zero-Equivalence in function fields defined by algebraic differential equations. *Transactions of the American Mathematical Society, 1993, vol. 336, №1. URL: http://www.ams.org/journals/tran/1993-336-01/S0002-9947-1993-1088022-2/S0002-9947-1993-1088022-2.pdf (дата обращения 21.01.2015).*

[3] Kahan W. Branch cuts for complex elementary functions. *The State of Art in Numerical Analysis, Oxford, Clarendon Press, 1987, pp. 165 – 211.*

[4] Corless R.M., Jeffry D.J. The unwinding number. *SIGSAM Bulletin*, 1996, №30 (2), pp. 28 – 35. URL: http://www.apmaths.uwo.ca/~djeffrey/Offprints/editors.pdf *(дата обращения 21.01.2015).*

[5] Епифанов А.С. *Элементарные методы символьных вычислений.* Москва, Маска, 2012, 316 с.
URL: http://PifMath.ru/1.4.2.1/PifMath_Book.pdf (дата обращения 04.04.2015).

[1] Более простой, хотя и более затратный по времени, способ получить (37) – это вычислить $csgn((-1)^{m/n})$. Если получится +1, то верна первая строка (37), иначе – вторая.

Епифанов А.С. (род. 1947 г.) – доцент кафедры «Физики» МГТУ им. Н.Э. Баумана, кандидат физ.-мат. наук; области научных интересов: теория взаимодействия лазерного излучения с конденсированными средами, компьютерная математика (в том числе непосредственно программирование); e-mail: apifpk@gmail.com.

Никифоров Александр Михайлович (род. 1985г.) - доцент кафедры «Физики» МГТУ им. Н.Э. Баумана, кандидат физ.-мат. наук; области научных интересов: квантовая теория поля, взаимодействие лазерного излучения с конденсированными средами, компьютерная математика; e-mail: nikahv@yandex.ru.

Expansions and Contraction of Logarithmic and Power functions in Computer Symbolic Expression Simplification

@A.S. Epifanov, A.M. Nikiforov

Difficulties encountered while expanding and contracting logarithmic and power functions of complex variables are analyzed. In the context of small computer mathematics systems criteria for correct merging (addition for logarithmic and multiplication for power functions of complex expressions) and inverse transformations (expansion) are proposed, a method for simplification of symbolic expressions based on these manipulations being outlined. Properties of the complex variable sign function as used in the simplifications are discussed. Examples for the method are available at http://PifMath.ru.

Key words: *symbolic computations, computer mathematics, logarithmic function of complex variable, power function, simplifying expressions.*

Epifanov A.S. (b. 1947) - Assoc. Professor of the Physics Department at Bauman Moscow State Technical University, Ph.D. Main scientific interests: Theory of the laser beam interaction with solids, computer mathematics (including programming itself). E-mail: apifpk@gmail.com.

Nikiforov A.M. (b. 1985) - Assoc. Professor of the Physics Department at Bauman Moscow State Technical University, Ph.D. Main scientific interests: quantum field theory, laser beam interaction with solids, computer mathematics. E-mail: nikahv@yandex.ru.