УДК 535.411.01

ИСПОЛЬЗОВАНИЕ СВОЙСТВ ЦЕЛЫХ ЧИСЕЛ ДЛЯ РАСШИФРОВКИ ИНТЕРФЕРОГРАММ

Гужов В. И., Солодкин Ю. Н.

Показано, что решения систем сравнений целых чисел позволяют определять полную разность фаз периодических сигналов по значению фазы в пределах одного периода. Это позволяет устранить фазовую неоднозначность в интерферометрах с контролируемым фазовым сдвигом.

Задачей интерферометрии является получение полной разности фаз световых волн, которая определяется либо по двумерной интерференционной картине, либо по периодическому изменению яркости в заданных точках. Суть предлагаемого метода интерферометрии заключается в том, что для определения полной разности фаз световых волн нет необходимости считать интерференционные полосы [1].

Теоретическое обоснование метода

Метод основан на свойствах делимости, изучаемых в теории чисел [1]. Поскольку теория чисел представляет собой одну из самых абстрактных областей математики, не получившую широкого использования при решении инженерных задач, изложим кратко основные сведения из этой теории, необходимые для понимания предлагаемого метода.

Теория рассматривает только целые числа. Если какие-то числа при делении на число m дают один и тот же остаток r, то они называются сравнимыми по модулю m. Например, числа 7, 12, 27 сравнимы по модулю 5, так как при делении любого из них на число 5 получается остаток 2. Математически это записывается так

$$a \equiv b \pmod{m}, \tag{1}$$

что означает: a сравнимо с b по модулю m.

Числа, сравнимые по модулю m, образуют класс чисел по модулю m, причем все числа класса можно получить из выражения $mq+r$, где q принимает значения всех целых чисел. Например, для $m=5$ и $r=2$ получим ряд 2, 7, 12, 17, . . ., представляющий собой класс чисел по модулю 5.

Любое число класса называется вычетом по модулю m. Вычеты, дающие все возможные значения r, образуют полную систему вычетов по модулю m. Полная система вычетов может быть получена при заданном значении q, если r принимает допустимые значения от 0 до $m-1$. Для наглядности это можно изобразить графически. На рис. 1 показаны полные системы вычетов по модулю 5. При $q=0$ получаем наименьшие вычеты $r=0$, 1, 2, 3, 4. При $q=1$ полная система вычетов 5, 6, 7, 8, 9 и т. д. Взаимосвязь между r и $mq+r$ можно отобразить пилообразной функцией с условием, что r и $mq+r$ принимают только целые значения, отмеченные точками. Важно для дальнейшего, что полная система вычетов периодически повторяется с периодом, равным m.

Если в сравнении (1) одно из чисел, например, a, неизвестно, то, обозначив его через X, можем записать

$$X \equiv b \pmod{m}. \tag{2}$$

Выражение (2) является сравнением первой степени с одним неизвестным. Решить такое сравнение — это значит найти все удовлетворяющие ему значения X. Если сравнению (2) удовлетворяет некоторое $X=X_0$, то ему же будут удовлетворять все числа, сравнимые с X_0 по модулю m. Следовательно, весь класс чисел, определяемых сравнением

$$X \equiv X_0 \pmod{m}, \tag{3}$$

будет решением сравнения (2), и в теории чисел доказано, что это решение для сравнений вида (2) является единственным. Для одного сравнения решение находится тривиально. Например, если $b=12$, а $m=5$, то $X \equiv 12 \pmod{5}$, т. е. класс чисел 2, 7, 12, 17, . . . и есть единственное решение сравнения.

Нас будет интересовать решение системы сравнений вида (2)

$$X \equiv b_1 \pmod{m_1}, \quad X \equiv b_2 \pmod{m_2}, \ \ldots, \ X \equiv b_k \pmod{m_k}, \tag{4}$$

причем модули не равны между собой и являются взаимно простыми числами. Решением системы сравнений (4) считается класс чисел $X=X_0 \pmod{m}$, кото-

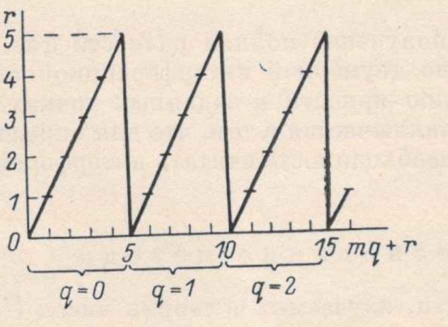

Рис. 1. Полные системы вычетов по модулю 5.

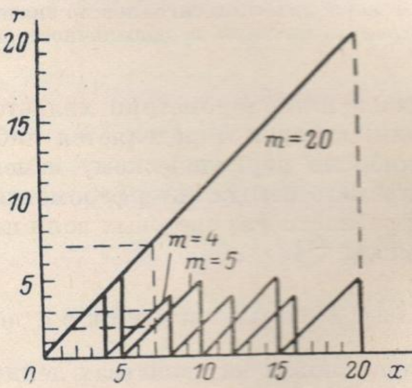

Рис. 2. Полные системы вычетов по модулям 4, 5 и 20.

рый удовлетворяет всем сравнениям одновременно. Требуется определить X_0 и m.

Именно свойства решений системы сравнений (4) лежат в основе предлагаемого метода интерферометрии.

В теории чисел доказывается теорема, по которой

$$X_0 = M_1 M_1' b_1 + M_2 M_2' b_2 + \ldots + M_k M_k' b_k, \tag{5}$$

$$m = m_1 m_2 \ldots m_k, \tag{6}$$

где M_k и M_k' — числа, определенные из условий

$$M_S m_S = m_1 m_2 \ldots m_k, \tag{7}$$

$$M_S M_S' \equiv 1 \pmod{m_S}, \tag{8}$$

в которых S последовательно принимает значения от 1 до k.

Рассмотрим простой пример. Решим систему из двух сравнений

$$X \equiv b_1 \pmod{4}, \quad X \equiv b_2 \pmod{5}. \tag{9}$$

Определим числа M_S и M_S'. В данном случае $S=1, 2$. Из (7) следует $M_1 m_1 = m_1 m_2$ и $M_2 m_2 = m_1 m_2$, откуда $M_1 = m_2 = 5$ и $M_2 = m_1 = 4$. Из (8) находим M_S'. Имеем два сравнения: $M_1 M_1' \equiv 1 \pmod{4}$ и $M_2 M_2' \equiv 1 \pmod{5}$. Подставляем найденные значения M_1 и M_2, получаем $5M_1' \equiv 1 \pmod{4}$ и $4M_2' \equiv 1 \pmod{5}$ и перебираем целые значения M_1' и M_2', начиная с единицы, пока произведения, стоящие слева, деленные на соответствующие модули, не дадут в остатке единицу. Получаем $M_1'=1$ и $M_2'=4$.

Равенства (5) и (6) дают искомые значения $X_0=5b_1+16b_2$, $m=20$, а решение системы сравнений (9) имеет вид

$$X \equiv 5b_1 + 16b_2 \pmod{20}. \tag{10}$$

Это означает, что числа, являющиеся решением, пробегают полную систему вычетов по модулю 20. Поясним полученный результат графически. На рис. 2 показаны полные системы вычетов по модулям 4 и 5 (сравнения (9)) и полная система вычетов по модулю 20 (сравнение (10)). Если b_1 и b_2 — наименьшие вычеты, то сравнение (10) дает наименьший вычет решения. Например, $b_1=3$ и и $b_2=2$. Тогда $X_0=5\cdot3+16\cdot2=47$, и наименьший вычет решения равен 7. В этом легко убедиться по графикам на рис. 2. Вертикальная штриховая линия пересекает графики при значениях r, равных 7, 3, 2. Число 7 удовлетворяет обоим сравнениям (9), так как при делении на модули 4 и 5 дает соответственно остатки 3 и 2. Таблица содержит все решения для рассматриваемого случая. Она понадобится при обсуждении вопроса точности определения полной фазы.

Приближая полученный из теории чисел результат к решаемой в интерферометрии задаче, можно интерпретировать его следующим образом. Если линейные периодические функции заданы своими целыми значениями и периоды их являются взаимно простыми числами, то им в однозначное соответствие может быть поставлена также линейная периодическая функция с периодом, равным произведению периодов заданных функций.

Решения системы (9)

b_1	b_2				
	0	1	2	3	4
0	0	16	12	8	4
1	5	1	17	13	9
2	10	6	2	18	14
3	15	11	7	3	19

Описание метода и его сравнительная характеристика

Разность фаз световых волн в интерферометрах определяется измеряемой величиной, например смещением $\mathbf{u}$, и геометрическими параметрами интерферометра — единичными векторами освещения $\mathbf{r}_0$ и наблюдения $\mathbf{r}_н$. При этом полная разность фаз вычисляется, как

$$\varphi = \frac{2\pi}{\lambda}\,\mathbf{u}\,(\mathbf{r}_0 + \mathbf{r}_н), \tag{11}$$

где λ — длина световой волны. Ту же разность фаз можно определить через число интерференционных полос N

$$\varphi = 2\pi N. \tag{12}$$

Приравнивая (11) и (12), имеем

$$\mathbf{u}\,(\mathbf{r}_0 + \mathbf{r}_н) = \lambda N. \tag{13}$$

В частности, если $\mathbf{u}$ совпадает по направлению с $\mathbf{r}_0$ и $\mathbf{r}_н$, то

$$|\mathbf{u}| = \frac{\lambda}{2} N. \tag{14}$$

Равенство (14) справедливо для лазерных измерителей перемещений, выполненных на основе интерферометра Майкельсона, а выражение (13) используется для определения смещений и деформаций в голографических интерферометрах. Непосредственно в интерферометрах определяется число полос N.

Если интерференционный эксперимент проводится в реальном времени, то подсчитывается число периодов изменения яркости в исследуемой точке. Для определения полей смещений этот метод обычно не используется из-за трудоемкости реализации.

Расшифровка неподвижной интерференционной картины, содержащей информацию о поле измеряемых величин, имеет существенные отличия. Полосы в этом случае образованы точками объекта, имеющими равные смещения, и двумерная функция яркости может иметь сложный вид. Поэтому количественная информация о разности фаз может быть получена только вдоль линий максимумов и минимумов интерференционных полос. При определении разности фаз в других точках возникают ошибки интерполяции, особенно значительные при небольшом числе полос.

Автоматизация процесса выделения экстремальных значений яркости интерференционной картины и определения полной разности фаз оказывается довольно сложной. Алгоритм включает в себя поиск непрерывных зон, в которых фаза плавно меняется в пределах периода, определение знака перехода по изменению фазы внутри зоны, устранение разрывов интерференционных полос, имеющих одинаковые номера, и определение полной разности фаз.

Таким образом, определение полной разности фаз по числу интерференционных полос связано по крайней мере с двумя принципиальными трудностями — сложностью автоматизации и ошибками при отсчете долей интерференционных полос в пределах одного периода.

Точное значение фазы световой волны в исследуемой точке можно определить по значению яркости в этой точке, однако при этом необходимо знать амплитуду и начальную фазу световой волны

$$A = A_0 \cos \varphi. \tag{15}$$

Имея A_0 (точнее, $I_0=A_0^2$) и измерив $I=A^2$, можно из (15) получить значение φ. Очевидно, что это значение можно определить только в пределах одного периода.

Так как A_0 изменяется от точки к точке, такой способ определения фазы практически неприемлем. Существенным усовершенствованием данного способа является введение в интерферометры контролируемого фазового сдвига, предложенное в ряде работ. Смысл заключается в том, что выражения (15) получают для трех заданных значений фазы: $A_i=A_0 \cos (\varphi+\varphi_i)$, $i=1, 2, 3$. При этом измеряют соответственно три значения яркости I_1, I_2, I_3. Решая систему из трех уравнений вида (15) относительно φ, имеем

$$\varphi = \operatorname{arc\,tg} \frac{(I_3-I_2)\cos\varphi_1+(I_1-I_3)\cos\varphi_2+(I_2-I_1)\cos\varphi_3}{(I_3-I_2)\sin\varphi_1+(I_1-I_3)\sin\varphi_2+(I_2-I_1)\sin\varphi_3}. \tag{16}$$

Из (16) следует, что φ не зависит от значений A_0. Подчеркнем также, что для определения фазы не требуется отсчитывать интерференционные полосы.

Метод интерферометрии с контролируемым фазовым сдвигом позволяет создать полностью автоматизированные интерференционные измерительные системы с высокой точностью измерения, определяемой сотыми долями длины световой волны.

Перечисленные достоинства метода привлекли к нему серьезное внимание, однако неоднозначность определения фазы (так называемая 2π-неоднозначность) существенно ограничивала возможности метода. Устранение же 2π-неоднозначности подсчетом целого числа интерференционных полос (числа переходов через 2π) сводило на нет большинство преимуществ интерферометров с контролируемым фазовым сдвигом.

Для расширения диапазона измеряемых величин при допустимом изменении фазы в пределах одного периода могут быть использованы различные способы увеличения цены интерференционной полосы. Один из способов заключается в получении двух интерференционных картин с различными пространственными частотами полос, например одной с ценой полосы $\lambda_1/2$, а другой — с ценой полосы $\lambda_2/2$ в соответствии с (14). Тогда разность фаз, определенных по двум интерферограммам, будет

$$\varphi_1 - \varphi_2 = 2\pi \left(\frac{1}{\lambda_1/2} - \frac{1}{\lambda_2/2} \right) |\mathbf{u}|. \tag{17}$$

Период изменения функции (17) равен $0.5\lambda_1\lambda_2/(\lambda_2-\lambda_1)$, и предел измерения интерферометров с контролируемым фазовым сдвигом увеличивается в $\lambda_1/(\lambda_2-\lambda_1)$ или в $\lambda_2/(\lambda_2-\lambda_1)$ раз. Этот результат можно интерпретировать как устранение 2π-однозначности по отношению к исходной интерференционной картине с ценой полосы $\lambda_1/2$ в диапазоне до $0.5\lambda_1\lambda_2/(\lambda_2-\lambda_1)$.

Способ увеличения цены полосы широко используется для получения топографических карт поверхностей в голографических интерферометрах. Он же лежит в основе интерферометров для измерения длин концевых мер по совпадению дробных частей интерференционных полос. Измерение по совпадению дробных частей по крайней мере внешне похоже на измерение по предлагаемому в данной работе методу.

Предполагается, что длина концевой меры l заранее известна с точностью до $\Delta l \leqslant 0.5\lambda_1\lambda_2/(\lambda_2-\lambda_1)$, и именно эту добавку измеряют в интерферометре. Сначала находят дробную часть интерференционной полосы b_1 при длине волны λ_1, затем аналогичное измерение производят при длине волны λ_2 и получают b_2. Так как разность фаз (17) имеет период повторения $0.5\lambda_1\lambda_2/(\lambda_2-\lambda_1)$, каждой паре чисел b_1, b_2 однозначно соответствует значение Δl, и длина концевой меры определяется как $l=l_0+\Delta l$.

Основная трудность при измерении по совпадению дробных частей возникает при выборе значений λ_1 и λ_2. С одной стороны, чем меньше $\lambda_2-\lambda_1$, тем больше цена полосы и, следовательно, выше предел измерения. С другой стороны, при близких значениях λ_1 и λ_2 числа b_1 и b_2 мало отличаются друг от друга, и возрастает ошибка измерения. Компромиссное решение приводит к тому, что предел измерения не превышает нескольких микрометров. Различные модификации способа измерения (использование большого числа различных длин волн или источников с непрерывной перестройкой длины волны) существенно не расширяют его возможности. Измерения трудоемки, производятся в одной точке, их сложно автоматизировать.

Предлагаемый метод устранения 2π-неоднозначности, основанный на свойствах целых чисел, принципиально отличен от изложенных и позволяет значительно увеличить предел измерения интерферометров с контролируемым фазовым сдвигом.

Вернемся к рис. 2. Пилообразным функциям, отражающим полные системы вычетов по модулям 4 и 5, можно придать физический смысл линейного изменения фазы гармонических сигналов с периодами $m_1=4$ и $m_2=5$, т. е. эти функции применительно к интерферометрии описывают две системы интерференционных полос. При этом по оси $mq+r$ откладывается пространственная координата, а по оси r — значение фазы. Таким образом, наименьшие вычеты b_k (4) — фазы сигналов, которые определены в заданной точке при различных значениях цены полосы (за счет изменения длины волны, коэффициента преломления среды или смещения источника).

Значения b_k получают, например, в интерферометре с контролируемым фазовым сдвигом с точностью до целого числа в пределах одного периода. Если числа m_k взаимно простые, то может быть найдена функция (5), имеющая период, равный $m_1m_2 \ldots m_k$. Подставляя в эту функцию совокупность значения b_k, получают значение X_0, имеющее смысл полной разности фаз. Максимальное число интерференционных полос, если полосой считать период m_1, составляет $m_2m_3 \ldots m_k$. Во столько же раз возрастает предел измерения интерферометра с контролируемым фазовым сдвигом. Не считая числа полос, по измерению яркости в точке можно определить полную разность фаз с точностью до $m_2m_3 \ldots m_k$ периодов.

Для сравнения предлагаемого метода и наиболее близкого к нему метода измерения по совпадению дробных частей полос рассмотрим одинаковую ситуацию, когда измерения производят при двух значениях периода интерференционных полос 0.529 и 0.633 мкм. Отметим, что выбор различных периодов не вызывает затруднений. Он особенно прост при использовании смещенного источника. Предел измерения по методу совпадения равен $0.529\cdot 0.633/(0.633-0.529)\approx$ ≈ 3.22 мкм, что составляет примерно 6 периодов по 0.529 мкм.

В предлагаемом методе используются безразмерные целые числа, поэтому задание двух указанных периодов означает, что им в соответствии ставят целые числа 529 и 633. По условию эти числа должны быть взаимно простыми, что и имеет место в данном случае. Тогда предел измерения определяется также целым безразмерным числом $529\cdot 633=334\ 857$, что составляет 633 периода по 529.

Сравнение показывает, что предел измерения по предлагаемому методу оказывается выше более чем на два порядка. Если использовать другие значения периодов или учитывать бо́льшее число десятизначных знаков, или производить измерения при трех и более различных значениях периодов, то выигрыш может быть еще значительнее.

Итак, в работе предложен новый метод интерферометрии, который позволяет по измерению яркости в точке при нескольких значениях контролируемого фазового сдвига определить не только фазу в пределах одного периода изменения яркости, а полную разность фаз световых волн, не используя при этом подсчета чисел интерференционных полос, начиная с нулевой полосы, и даже не наблюдая интерференционную картину.

Дано теоретическое обоснование метода, связанное со свойствами делимости целых чисел. Рассмотрены решения систем сравнений, исследуемых в теории чисел, и показано, что если линейные периодические функции, характеризующие изменения фазы, заданы своими целыми значениями и периоды этих функций являются взаимно простыми числами, то им в однозначное соответствие может быть поставлена также линейная периодическая функция с периодом, равным произведению периодов заданных функций, по которой и определяется полная разность фаз световых волн.

Дана сравнительная оценка предлагаемого метода с известными методами интерферометрии и отмечено, что его использование в интерферометрах с контролируемым фазовым сдвигом позволяет повысить точность, расширить динамический диапазон и полностью автоматизировать процесс измерения.

Литература

[1] Виноградов И. М. Теория чисел. М., 1978, с. 190.

Поступило в Редакцию 20 марта 1987 г.
В окончательной редакции 15 февраля 1988 г.