Базисы Гребнера. Алгоритмические и сложностные аспекты

Шокуров А.В

5 мая 2022 г.

Оглавление

Введение

Системы полиномиальных уравнений — один из наиболее универсальных математических объектов. Огромное количество задач, относящихся к различным разделам математики, могут быть переформулированы как задачи решения таких систем. В частности, почти все задачи криптографического анализа можно свести к поиску решений систем полиномиальных уравнений. Соответствующее направление исследований принято называть алгебраическим криптоанализом.

Для читателя, недостаточно знакомого с криптографией, особо подчеркнем, что алгебраический криптоанализ не сводится к одному лишь поиску методов решения систем полиномиальных уравнений. Задача построения такой системы для данной задачи криптоанализа зачастую оказывается весьма непростой.

С точки зрения вычислительной сложности, системы полиномиальных уравнений (далее, для краткости, просто — системы уравнений) охватывают весь диапазон возможных вариантов, от алгоритмической неразрешимости диофантовых уравнений до хорошо известных эффективных методов решения линейных систем. Задача распознавания разрешимости систем квадратичных уравнений NP-полна уже в булевом случае [**GJ**].

Систему уравнений над конечным полем всегда можно решить методом полного перебора. Но к системам уравнений над полем рациональных чисел этот метод неприменим. Долгое время открытым оставался вопрос об алгоритмической разрешимости задач, связанных с такими системами уравнений. Здесь следует заметить, что существуют задачи двух типов. Первый тип — задачи распознавания. Дана система уравнений. Требуется ответить на вопрос, совместна ли она. Второй тип — задачи поиска. Дана

система уравнений. Требуется найти хотя бы одно решение (другая постановка: найти все ее решения).

В действительности имеются две принципиально различные постановки задач для систем алгебраических уравнений: в зависимости от того, требуется ли определить совместность (найти хотя бы одно решение или же найти все решения) в самом поле или его алгебраическом замыкании. Например, ответ на вопрос о существовании нуля квадратичной формы от n переменных в поле рациональных чисел (основном поле) дает теорема Минковского—Хассе (см. [**Borshaf**]), утверждающая что такое решение существует тогда и только тогда, когда когда эта форма разрешима над полем вещественных чисел, а также над полями p-адических чисел для всех простых p. Другим примером задачи проверки разрешимости систем уравнений в основном поле является проблема существования алгоритмов распознавания совместности систем диофантовых уравнений. Она была поставлена Гильбертом в 1900 году (знаменитая 10 проблема Гильберта). Отрицательный ответ на поставленный вопрос был дан лишь 70 лет спустя Ю.В. Матиясевичем [**Матияс**]. Более того, можно построить однопараметрическое семейство диофантовых уравнений от 22 переменных, для которого не существует алгоритма распознавания существования решения (см. [**Borshaf**]).

Ответ на вопрос о существовании решения системы алгебраических уравнений в некотором расширении (алгебраическом замыкании) основного поля дает теорема Гильберта о нулях (1893 г.). Эта теорема, однако, не является конструктивной, поскольку требует в общем случае решения бесконечной линейной системы уравнений. В дальнейшем в 1903 г. Кёнигом была доказана возможность проверки существования решения за конечное число шагов, т.е. доказана ее алгоритмическая разрешимость. Позднее Херманн [Her25] получила явную оценку достаточного размера системы линейных уравнений, зависящую только от размеров исходной системы алгебраических уравнений (метод линеаризации). Однако, эта оценка — двойная экспонента от степени исходной системы уравнений, причем в общем случае неулучшаемая.

Новый импульс исследованиям в этой области придала работа Бухбергера [Buc06] (1965 г.), в которой определялся специальный базис идеала, порожденного системой полиномиальных уравнений. Этот базис он назвал

в честь своего учителя базисом Гребнера. Введенное понятие вызвало целый поток работ в этой области. В частности, введение базисов Гребнера позволило обобщить метод Жордана—Гаусса на системы алгебраических уравнений (см., например, [Laz83]). Метод Бухбергера приводит систему алгебраических уравнений к системе специального вида, определяемой базисом Гребнера для исходной системы уравнений и позволяющей использовать исключение переменных. В частности, базис Гребнера системы линейных уравнений — это тот базис, в котором матрица системы имеет ступенчатый вид.

Остановимся кратко на основных понятиях теории базисов Гребнера. Фундаментом определения базиса Гребнера является допустимое упорядочение на множестве термов (мономов с коэффициентом единица) от n независимых переменных. Упорядочение на множестве термов называется допустимым, если оно согласовано с умножением термов и множество термов является вполне упорядоченным относительно такого упорядочения. Примерами допустимых упорядочений на множестве термов являются лексикографическое упорядочение, лексикографическое упорядочение, согласованное с градуировкой, обратное лексикографическое упорядочение, согласованное с градуировкой. Множество допустимых упорядочений на множестве термов бесконечно и даже континуально.

При выборе допустимого упорядочения базис Гребнера идеала I системы алгебраических уравнений определяется как конечное множество многочленов, являющееся базисом идеала I, старшие термы которых делят старшие термы всех многочленов идеала I. Из определения следует, что базис Гребнера может зависеть от выбранного допустимого упорядочения. И это действительно так: имеются примеры допустимых упорядочений, для которых базисы Гребнера существенно различны.

С помощью такого базиса можно ответить на вопрос о принадлежности идеалу произвольного многочлена, свести задачу решения системы полиномиальных уравнений к задачам нахождения решений полиномиальных уравнений от одной переменной.

Алгоритм нахождения базисов Гребнера достаточно прост. Алгоритм всегда завершается за конечное число шагов, и это число зависит от выбора допустимого упорядочения. Поэтому множество вопросов связано с оценкой мощности этого базиса и сложностью его вычисления. Известно,

что в общем случае задача нахождения такого базиса лежит в классе $EXPSPAC$ (задач, разрешимых алгоритмами, использующими не более чем экспоненциальную память).

Доказано, что, зная базис Гребнера относительно лексикографического упорядочения на множестве термов для идеала I размерности ноль и имея оракул, позволяющий находить решение алгебраического уравнения от одной переменной, можно определить, совместна ли система уравнений, соответствующая этому идеалу, и найти некоторое решение в случае совместности системы. Однако это решение может не принадлежать основному полю.

В случае когда основное поле конечно, т.е. является полем $\mathbb{F}_q$, где $q = p^n$ для некоторого простого p, задача нахождения решения в основном поле решается так. Добавим элементы $x_i^q - x_i$ к базису идеала I для всех переменных x_i. Далее построим базис Гребнера полученного идеала J и найдем некоторое его решение. Несовместность полученной системы будет означать отсутствие решения в основном поле.

Вот пример, иллюстрирующий тезис о том, что метод Бухбергера решения систем алгебраических уравнений является обобщением метода Гаусса.

Рассмотрим систему из двух алгебраических уравнений от двух переменных над полем рациональных чисел

$$\begin{cases} x_1^3 + x_1^2 x_2 + x_1 x_2 + x_2^2 - 1 &= 0 \\ x_1^2 x_2 + x_2^3 + 1 &= 0 \end{cases} \tag{1}$$

Базис Гребнера относительно лексикографического порядка для идеала I этой системы уравнений состоит из многочленов x_1 и $x_2 + 1$. Чтобы убедиться в этом, достаточно проверить выполнение соотношений

$$I \subset (x_1, x_2 + 1) \tag{2}$$

и

$$I \supset (x_1, x_2 + 1). \tag{3}$$

Соотношения (2) очевидны. Соотношения (3) следуют из метода нахождения базиса.

Следовательно, (1) эквивалентна тривиальной системе уравнений

$$\begin{cases} x_1 &= 0 \\ x_2 + 1 &= 0 \end{cases} \tag{4}$$

Что будет, если нам известен базис Гребнера для некоторого допустимого порядка? Оказывается в этом случае имеется простой алгоритм (см. [Col97]), преобразующий его в базис Гребнера для любого другого допустимого порядка, в частности для лексикографического.

За последние 10–15 лет появился ряд работ, авторы которых предлагают использовать базисы Грёбнера для решения систем уравнений, возникающих в алгебраическом криптоанализе. В разделе 7 приводится весьма краткое обсуждение некоторых из этих работ. При этом не ставилась задача составления обзора исследований в данном направлении. Материал раздела 7 следует трактовать как иллюстративный.

Задача поиска секретного ключа — типичная задача криптоанализа — практически всегда может быть решена методом полного перебора. Естественным образом возникает следующий вопрос: существуют ли примеры, когда методы, основанные на базисах Грёбнера, оказываются эффективнее полного перебора. Заметим, что с научной точки зрения отсутствие таких примеров еще ничего не говорит о перспективности самого подхода.

В случае, если система алгебраических уравнений над полем имеет конечное число решений в алгебраическом замыкании этого поля (иными словами, идеал такой системы имеет размерность 0), задача распознавания алгоритмически разрешима. Более того, существует алгоритм, дающий ответ на вопрос о конечности решений в алгебраическом замыкании.

Для уравнений над конечными полями задачи распознавания совместности и поиска решения имеют, по существу, одинаковую вычислительную сложность. Это следует из существования так называемой самосводимости: задача поиска сводится к задаче распознавания.

Самосводимость легко понять на примере системы булевых уравнений. Предположим, что у нас есть оракул, который отвечает на вопросы о совместности систем, и требуется найти хотя бы одно решение. Пусть x_1 — первая из переменных. Зафиксируем одно из значений 0 или 1. Пусть это будет 0. Подставим $x_1 = 0$ во все уравнения и спросим у оракула, совместна ли система. Если да, то в решение записываем $x_1 = 0$, иначе $x_1 = 1$. При

этом исходная система сведена к системе от меньшего числа переменных.

Для систем уравнений над полем рациональных чисел вопрос о существовании самосводимости остается открытым. Поэтому распознавательная задача и задача поиска исследовались отдельно.

Задача поиска решений систем алгебраических уравнений еще десятилетия ожидала своего решения. Поскольку множество рациональных чисел бесконечно, а набор переменных в системе конечен, следовало обратиться к символьным методам. Основу символьных алгоритмов составляют преобразования решаемой системы уравнений, не зависящие от того поля, в котором ищется решение.

Подчеркнем, что для криптографии наиболее важными являются алгоритмические и сложностные аспекты построения базисов Грёбнера. Именно этой тематике и посвящены основные разделы данной главы.

Почти все известные из литературы результаты о вычислительной сложности задач, связанных с базисами Грёбнера, доказаны для меры сложности "в худшем случае". Для криптографии это всего лишь информация к размышлению. Гораздо интереснее (хотя все еще недостаточна) сложность в среднем. Эта область остается практически неисследованной. Заметим, что если мы зафиксировали множество систем уравнений и способ задания входных данных для алгоритма, то это еще не определяет вычислительную задачу для меры сложности "в среднем". Необходимо еще задать распределение вероятностей на множестве систем уравнений. Таких распределений бесконечно много, и каждое из них определяет вычислительную задачу, вообще говоря, со своей сложностью в среднем.

Остается открытым также вопрос о том, как влияет допустимый порядок на длину базиса Гребнера и вычислительную сложность его построения. В частности, можно ли за счет выбора подходящего допустимого порядка понизить трудоемкость известных алгоритмов построения базиса Гребнера. Есть также и совершенно неисследованная смежная область, связанная с поиском ответов на вопрос о существовании других типов базисов, облегчающих решение систем уравнений. Если ответ на этот вопрос положительный, то конечная цель исследований в данном направлении весьма амбициозна: описать все типы базисов идеалов, отличных от базиса Гребнера, знание которых позволяет решать системы уравнений эффективно (за полиномиальное время), и провести классификацию всех типов

базисов указанного вида по вычислительной сложности их построения.

Глава состоит из двух частей. Первая часть алгебраическая. В ней содержатся все необходимые для понимания результаты и определения из теории колец многочленов и идеалов в кольцах многочленов.

Далее приводятся различные определения базиса Гребнера идеала, доказывается его существование и описывается алгоритм Бухбергера для его нахождения.

Базис Гребнера не определяется однозначно, а зависит от допустимого порядка на множестве мономов в кольце многочленов.

Знание базиса Гребнера позволяет алгоритмически решить ряд задач. И здесь возникает вопрос о сложности нахождения этого базиса и множество вопросов о том, дает ли нам знание базиса Гребнера возможность быстро решить такие задачи: задача о принадлежности идеалу, задача о перечислении всех решений соответствующей системы уравнений и т.д.

Во второй части рассматриваются вопросы связанные со сложностью задач нахождения базисов Гребнера, а также других задач, использующих для решения такие базисы. Примером такой задачи является задача определения принадлежности многочлена идеалу. В случае идеалов в кольцах многочленов над целыми (или рациональными) числами эта задача является $EXPSPACE$-полной (Мейр и Майр).

Важной является задача нахождения решения систем уравнений в кольце многочленов специального вида. В случае систем булевых уравнений, систем, имеющих конечное число решений, и некоторых других случаях задача нахождения базиса Гребнера может быть решена алгоритмически с памятью, ограниченной полиномом от длины записи входных данных.

На самом деле, значительный интерес представляет анализ поведения известных алгоритмов построения базисов Гребнера на типичных (в некотором смысле) данных, т.е. анализ сложности в среднем. Несмотря на очевидную потребность в такого рода исследованиях, их число относительно невелико. Исследованиям в этом направлении посвящены, например, работы [Bar04; HL11; CD11].

Известно, в частности, что дважды экспоненциальные нижние оценки степеней многочленов базиса Гребнера в худшем случае для типичного случая не имеют места и для многих достаточно типичных случаев эти оценки почти линейны [Giu84].

С другой стороны, большой интерес (в частности для криптографии) представляет случай конечных полей (особенно булевых). А здесь никакие оценки степеней многочленов не могут помочь в получении нижних оценок сложности алгоритмов.

Известно, что в булевом случае можно построить базис Гребнера с памятью, ограниченной полиномом (аналог класса $PSPACE$). Однако найти решение системы булевых уравнений всегда можно перебором, т.е. задача разрешимости принадлежит классу NP.

И это типично и для других случаев: построить базис Гребнера обычно сложнее, чем просто решить систему перебором (конечно, только для случая конечных полей, однако это и есть наиболее интересный случай).

Отметим в этой связи один интересный (в первую очередь для криптографии) специальный случай систем булевых уравнений, которые заведомо имеют решение и притом единственное (ключ). В данном случае, как следует из рассуждений в разделе 6 (см. теорему 50), нахождение решения полиномиально эквивалентно нахождению базиса Гребнера (приведенного и потому единственного).

Недостаточно исследованным представляется и вопрос о зависимости сложности алгоритма построения базиса Гребнера от выбранного допустимого порядка на мономах.

Наиболее полная библиография работ по данной тематике, охватывающая сотни работ, имеется на сайте http://www.risc.jku.at/Groebner-Bases-Bibliography.

Глава 1

Упорядоченные множества

В данной главе используются стандартные обозначения $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ и $\mathbb{R}$ для множеств натуральных, целых, рациональных и вещественных чисел. Определен необходимых понятий алгебры приведены в приложении. Рассматриваемые моноиды предполагаются коммутативными.

1.1 Упорядоченные множества

Определение 1.1.1. *Множество M называется частично упорядоченным, если на M определено отношение $x \preceq y$, удовлетворяющее условиям:*

1 *Для любого элемента $x \in M$ выполняется соотношение $x \preceq x$.*

2 *Если выполняются соотношения $x \preceq y$ и $y \preceq z$, то $x \preceq z$.*

3 *Если выполняются соотношения $x \preceq y$ и $y \preceq x$, то $x = y$.*

Условие 1 называют условием рефлексивности, условие 2 — условием транзитивности, а условие 3 условием антисимметричности. Отношение $\preceq$ называется отношением частичного порядка на множестве M. Отношени $\prec$, заданное формулой

$$x \prec y \;\Leftrightarrow\; x \preceq y \text{ и } x \neq y$$

называется строгим частичным порядком, соответствующим частичному порядку $\preceq$.

Определение 1.1.2. *Подмножество N частично упорядоченного множества M называется линейно упорядоченным, если для каждой из пар $x, y \in N$ выполняется хотя бы одно из соотношений*

$$x \preceq y \quad \text{и} \quad y \preceq x.$$

В частности, если $N = M$, то отношение порядка $\preceq$ называется линейным.

Элемент b частично упорядоченного множества M называется верхней гранью подмножества $S \subset M$, если для всех $x \in S$ выполняется соотношение $x \preceq b$.

Частично упорядоченное множество M называется индуктивным, если для любого линейно упорядоченного подмножества $N \subset M$[1] существует верхняя грань в M.

Элемент a частично упорядоченного множества M называется максимальным, если из условий $x \in M$ и $a \preceq x$ следует, что $x = a$.

Следующее утверждение называется леммой Цорна.

Лемма 1. *Любое непустое индуктивное частично упорядоченное множество содержит по крайней мере один максимальный элемент.*

Определение 1.1.3. *Строгое линейное упорядочение $\prec$ на моноиде M называется допустимым, если выполняются следующие свойства:*

1. *Если $\nu_1 \prec \nu_2$, то для любого $\nu \in M$ выполняется соотношение $\nu \cdot \nu_1 \prec \nu \cdot \nu_2$.*

2. *В любом непустом подмножестве $M_0 \subset M$ всегда существует наименьший элемент $\nu_0 \in M_0$:*

$$\forall\, \nu \in M_0 \quad \nu_0 \prec \nu \ \text{ или } \ \nu_0 = \nu.$$

[1] Включение множеств $\subset$ не предполагается строгим. Для строгого включения используется обозначение $\subsetneq$.

Отношение $\nu \prec \mu$ будем записывать также как $\mu \succ \nu$.

Определение 1.1.4. *Транзитивное отношение $\preceq$ на моноиде M называется допустимым предупорядочением, а отношение $\preceq$ — допустимым предпорядком, если выполняются следующие свойства:*

1. *Для любых элементов $\nu_1, \nu_2 \in M$ выполняется хотя бы одно из соотношений: $\nu_1 \preceq \nu_2$ или $\nu_2 \preceq \nu_1$.*

2. *Если $\nu_1 \preceq \nu_2$, то для любого $\nu \in M$ выполняется соотношение $\nu \cdot \nu_1 \preceq \nu \cdot \nu_2$.*

3. *В любом непустом подмножестве $M_0 \subset M$ всегда существует наименьший элемент $\nu_0 \in M_0$:*

$$\forall\, \nu \in M_0 \quad \nu_0 \preceq \nu.$$

Отношение $\nu \preceq \mu$ будем записывать также как $\nu \succeq \mu$. Если одновременно выполняются соотношения $\nu \preceq \mu$ и $\mu \preceq \nu$, то будем писать, что $\nu \equiv_{\preceq} \mu$ или, эквивалентно, $\mu \equiv_{\preceq} \nu$. Если из условия $\nu \equiv_{\preceq} \mu$ следует, что $\nu = \mu$, то будем говорить, что допустимый предпорядок $\preceq$ соответствует допустимому порядку $\prec$.

Поэтому, если существует допустимый предпорядок, соответствующий допустимому порядку $\prec$, то он определяется так

$$\nu \preceq \mu \;\Leftrightarrow\; \nu \prec \mu \text{ или } \nu = \mu.$$

Пусть задан допустимый порядок $\succ$ на моноиде термов $T\langle X\rangle$. Тогда, в частности, линейно упорядочено само множество X. Для множества $X = \{x_1, \ldots, x_n\}$ без ограничения общности будем считать, что $x_1 \succ x_2 \succ \ldots \succ x_n$. Докажем, что единица является наименьшим элементом этого моноида.

Лемма 2. *Пусть $\succ$ — допустимый порядок на моноиде термов $T\langle X\rangle$. Тогда для всех $x \in X,\ x \neq 1$ выполняются соотношения $x \succ 1$.*

Доказательство. Пусть это не так. Тогда согласно линейности порядка $\succ$ для некоторого $x \in T\langle X\rangle$ выполняется соотношение $1 \succ x$. Умножив левую и правую части соотношения на терм x, получим, в силу свойства 1 определения 1.1.3, соотношение $x \succ x^2$. Следовательно, $1 \succ x \succ x^2$. Повторяя эту процедуру умножения, получим бесконечную цепочку неравенств

$$1 \succ x \succ x^2 \ldots \succ x^n \succ x^{n+1} \succ \ldots,$$

не имеющую минимального элемента, что противоречит свойству 2 определения 1.1.3 порядка $\succ$. □

Теорема 1. *Пусть порядок $\succ$ на моноиде термов $T\langle X\rangle$ удовлетворяет условию 1 определения 1.1.3. Тогда условие 2 этого определения эквивалентно следующему:*

$2'$. $\forall x \in X$ *выполняется неравенство* $x \succ 1$.

Доказательство. Выполнение условия $2'$ для допустимого порядка установлено в лемме 2.

Докажем теперь, что из условий 1 и $2'$ следует условие 2 определения 1.1.3.

Заметим, что из свойств 1 и $2'$ и свойства транзитивности следует, что для любого терма $\nu \in T\langle X\rangle$ выполняется соотношение $\nu \succ 1$.

Пусть N — произвольное непустое множество термов. Требуется доказать, что в этом множестве существует наименьший элемент. Согласно следствию 8 теоремы Гильберта о базисе (заметим, что для доказательства теоремы Гильберта не используется доказываемая здесь теорема 1), в множестве N существует конечное подмножество N^0 такое, что каждый элемент $x \in N$ делится на некоторый элемент множества N^0, а следовательно, согласно условию 1 определения 1.1.3 и неравенству $\nu \succ 1$ для всех термов ν, каждый элемент множества N не меньше, чем некоторый элемент его подмножества N^0. Поэтому наименьший элемент конечного множества N^0 и будет наименьшим элементом множества N. □

Формула (10.6) из приложения, сопоставляющая терму X^ω его степень ω, определяет изоморфизм между моноидом термов $T\langle X\rangle$, где $X =$

$\{x_1, \ldots, x_n\}$, и моноидом $\mathbb{Z}_+^n$ целочисленных векторов с неотрицательными координатами. Поэтому порядок на множестве термов задает порядок на соответствующем аддитивном моноиде $\mathbb{Z}_+^n$. Свойство 1 допустимого порядка определения 1.1.3 на аддитивном моноиде переформулируются так:

1. Пусть $\alpha, \beta \in \mathbb{Z}_+^n$. Если $\alpha \succ \beta$, то для любого вектора $\gamma \in \mathbb{Z}_+^n$ выполняется соотношение $\gamma + \alpha \succ \gamma + \beta$.

Примеры допустимых порядков на множестве векторов из $\mathbb{Z}_+^n$.

1. **Лексикографический порядок.**[2] Для $\alpha, \beta \in \mathbb{Z}_+^n$ будем считать, что $\alpha \succ \beta$, если $\alpha \neq \beta$ и при некотором $1 \leq i_0 \leq n$
   - $\alpha_{i_0} > \beta_{i_0}$,
   - $\alpha_i = \beta_i$ при любом $1 \leq i < i_0$.

2. **Лексикографический порядок в градуированном моноиде.** Градуируем моноид $\mathbb{Z}_+^n = \bigcup_{i=0}^{\infty} M_i$, полагая
$$M_i = \{\alpha \in \mathbb{Z}_+^n | \alpha_1 + \ldots + \alpha_n = i\}.$$
   При $i > j$ будем считать, что всякий элемент множества M_i больше любого элемента множества M_j. Порядок внутри M_i зададим как лексикографический порядок.

3. **Обратный лексикографический порядок в градуированном моноиде.** Градуируем моноид $\mathbb{Z}_+^n = \bigcup_{i=0}^{\infty} M_i$ как выше. При $i > j$ будем считать, что всякий элемент множества M_i больше любого элемента множества M_j. Порядок внутри M_i зададим как обратный лексикографический, т. е. для элементов $\alpha, \beta \in M_i$ будем считать, что $\alpha \succ \beta$, если $\alpha \neq \beta$ и при некотором $1 \leq i_0 \leq n$

[2] Данное определение применимо также к M^n для любого линейно упорядоченного множества M. В результате получаем лексикографическое линейное упорядоченние на M^n.

- $\alpha_{i_0} \prec \beta_{i_0}$,
- $\alpha_i = \beta_i$ при любом $i > i_0$.

Легко проверить, что приведенные выше отношения задают допустимые порядки на моноиде $\mathbb{Z}_+^n$.

На каждом моноиде $\mathbb{Z}_+^n$ существует бесконечно много допустимых порядков. Ниже определяется бесконечное семейство допустимых порядков на полугруппе неотрицательных целочисленных векторов.

Определение 1.1.5. *Пусть A — $\mathbb{Q}$-матрица размера $n \times k$ и ранга n, в которой первый ненулевой элемент любой строки положителен. Определим порядок $\succ_A$ на моноиде $\mathbb{Z}_+^n$. Для элементов α, β этого моноида положим*

$$\alpha \succ_A \beta \iff \alpha A \succ \beta A,$$

где $\succ$ — лексикографический порядок на множестве $\mathbb{Q}^k$.

Теорема 2. *Пусть A — $\mathbb{Q}$-матрица размера $n \times k$ и ранга n, в которой первый ненулевой элемент любой строки положителен. Тогда отношение $\succ_A$ является допустимым порядком на моноиде $\mathbb{Z}_+^n$.*

Доказательство. Проверим выполнение свойств определения 1.1.3. Свойство линейности следует из инъективности отображения

$$A : \mathbb{Z}_+^n \to \mathbb{Q}^k.$$

Докажем первое свойство допустимого порядка. Действительно, $xA \succ yA \Rightarrow \forall z\ (x+z)A \succ (y+z)A$, поскольку $xA - yA = (x+z)A - (y+z)A$. Следовательно, $x \succ_A y \Rightarrow \forall z\ x + z \succ_A y + z$.

Чтобы проверить второе свойство, согласно теореме 1 достаточно убедиться, что $e_i = (0, \ldots, 1, \ldots, 0) \succ_A (0, \ldots, 0)$. Иными словами достаточно проверить выполнение неравенства $e_i A \succ 0$. Выполним умножение

$$e_i A = (a_{i,1}, \ldots, a_{i,k}).$$

Неравенство теперь следует из положительности первого ненулевого элемента i-ой строки. □

1.2 Теорема Роббиано

Доказанная в данном разделе теорема, известная как теорема Роббиано [Rob85], переоткрывает результаты [Riq10], [Kol73], [Gio52], и [МИ53], а также обобщает и уточняет теорему 2.

Вначале сделаем несколько замечаний к обозначениям теоремы. Напомним, что поле $\mathbb{R}$ можно рассматривать как бесконечномерное векторное пространство над полем рациональных чисел. Для любого вещественного вектора $v \in \mathbb{R}^n$ и любого $\mathbb{Q}$-подпространства V в $\mathbb{Q}^n$ скалярное произведение вектора v на векторы из пространства V определяет линейное над $\mathbb{Q}$ отображение $v : V \to \mathbb{R}$. Образ V при таком отображении будем обозначать через $V \cdot v$. Поскольку векторное $\mathbb{Q}$ -пространство $\mathbb{Q}^n$ имеет конечную размерность n, то и образ его подпространства V при линейном отображении v является конечномерным векторным пространством над $\mathbb{Q}$. Вещественное замыкание векторного $\mathbb{Q}$-пространства $V \subset \mathbb{R}^n$ определяется как векторное $\mathbb{R}$-пространство, порожденное элементами $V \subset \mathbb{R}^n$. Как множество вещественное замыкание V совпадает с топологическим замыканием множества V в векторном пространстве $\mathbb{R}^n$ с топологией, определяемой, например, евклидовой метрикой.

В формулировке теоремы используется понятие модуля над кольцом целых чисел. Соответствующие определения приведены в разделе 2.1.

Определение 1.2.1. *Линейный порядок $>$ на $\mathbb{Z}$-модуле M называется согласованным, если*

- $\forall x, y, z \in M\ (x > y \Rightarrow x + z > y + z)$,
- $\forall x \in M\ (x > 0 \Rightarrow 0 > -x)$.

Для каждого натурального числа n определим подмножество $\mathcal{M}_n$ в множестве всех вещественных матриц с n строками. Вещественная $n \times k$-матрица A принадлежит множеству $\mathcal{M}_n$ тогда и только тогда, когда

1. $k \leq n$;[3]

[3] Отметим, что в определении 1.1.5 предполагается, что $n \leq k$ и рассматриваются матрицы над рациональными числами, а здесь над вещественными числами.

2. для всех $j = 1, \ldots, k$ вектор-столбец a_j матрицы A принадлежит замыканию в $\mathbb{R}^n$ подпространства

$$V_j = \{v \in \mathbb{Q}^n \mid v \cdot a_l = 0, l = 1, \ldots, j-1\}$$[4]

3. $a_j \cdot a_j = 1$ для всех $j = 1, \ldots, k$.

4. пусть $d_j = \dim_{\mathbb{Q}}(V_j \cdot a_j)$, тогда выполняется соотношение

$$d_1 + \ldots + d_k = n.$$

Пусть $\succ$ — лексикографический порядок на множестве векторов в $\mathbb{R}^k$. Сопоставим матрице $A \in \mathcal{M}_n$ размера $n \times k$ порядок $\succ_A$ на $\mathbb{Z}$-модуле $\mathbb{Q}^n$ формулой

$$\alpha \succ_A \beta \iff \alpha A \succ \beta A. \tag{1.1}$$

Прокомментируем условия 2 и 4 на элементы множества $\mathcal{M}_n$.

Поскольку $\mathbb{R}$ является $\mathbb{Q}$-модулем, а линейное отображение $A : \mathbb{Q}^n \to \mathbb{R}^k$, заданное формулой $A(v) = (v \cdot a_1, \ldots, v \cdot a_k)$, гомоморфизмом $\mathbb{Q}$-модулей[5], то условие 4 означает невырожденность $\mathbb{Q}$-линейного отображения A: $\operatorname{rank}_{\mathbb{Q}} A = n$.

Условие 2 проиллюстрируем примером. Пусть $a = \left(\sqrt{\frac{1}{3}}, \sqrt{\frac{2}{3}}\right) \in \mathbb{R}^2$. Тогда для любого $b = (\alpha, \beta) \in \mathbb{Q}^2$ всегда $a \cdot b \neq 0$, т.е. ортогональное дополнение вектора a в $\mathbb{Q}^2$ нулевое, а ортогональное дополнение в $\mathbb{R}^2$ одномерно над $\mathbb{R}$. Легко проверить, что 2×1 матрица со столбцом a^t задает согласованный порядок на $\mathbb{Q}^2$.

Из свойства 4 множества $\mathcal{M}_n$ следует, линейность порядка $\succ_A$ для любой матрицы $A \in \mathcal{M}_n$, а из формулы (1.1) его согласованность на $\mathbb{Z}$-модуле $\mathbb{Q}^n$.

Следующая теорема позволяет однозначно представлять согласованные порядки на $\mathbb{Z}$-модулях матрицами из множеств $\mathcal{M}_n$.

---

[4] $V_1 = \mathbb{Q}^n$, поскольку множество соотношений, определяющих это множество, пусто.

[5] см.раздел 2.1

Теорема 3. *Пусть $\mathcal{M}_n$ — определенное выше множество матриц, а $\mathcal{O}_n$ — множество линейных согласованных порядков на $\mathbb{Z}$-модуле $\mathbb{Q}^n$. Тогда отображение $\varphi : \mathcal{M}_n \to \mathcal{O}_n$, заданное формулой*

$$A \mapsto \succ_A,$$

определяет взаимно однозначное соответствие между множествами $\mathcal{M}_n$ и $\mathcal{O}_n$.

Для доказательства теоремы 3 потребуются следующие леммы.

Лемма 3. *Для любого согласованного линейного порядка $\prec$ на $\mathbb{Z}$-модуле $\mathbb{Q}^n$, для любого элемента $x \in \mathbb{Q}^n$ и любого элемента $\alpha \in \mathbb{Q}$ выполняются свойства*

1. $x \succ 0, \alpha > 0 \Rightarrow \alpha x \succ 0.$
2. $x \succ 0, \alpha < 0 \Rightarrow \alpha x \prec 0.$
3. $x \prec 0, \alpha > 0 \Rightarrow \alpha x \prec 0.$
4. $x \prec 0, \alpha < 0 \Rightarrow \alpha x \succ 0.$
5. $x \succ 0,\ y \succ 0 \ \Rightarrow\ x + y \succ 0.$
6. $x \prec 0,\ y \prec 0 \ \Rightarrow\ x + y \prec 0.$
7. $x \succ 0,\ y \succ 0, 0 < \alpha < 1 \Rightarrow \alpha x + (1 - \alpha)y \succ 0.$

Доказательство. Свойство 1 выполняется для всех $\alpha \in \mathbb{Z} \mid \alpha > 0$. Пусть теперь $m \in \mathbb{Z} \mid m > 0$ и для некоторого $x \succ 0$ выполнено $\frac{1}{m} \cdot x \preceq 0$. Тогда $-\frac{1}{m} \cdot x \succeq 0$ и, согласно уже доказанному, $-x = m\left(-\frac{1}{m} \cdot x\right) \succeq 0$. Полученное противоречие показывает, что $\frac{1}{m} \cdot x \succ 0$. Представляя $\alpha > 0$ в виде отношения $\frac{n}{m}$, где $n, m > 0$, получаем свойство 1.

Свойства 2—4 следуют из свойства 1 и свойства согласованности умножения на -1.

Свойства 5 и 6 следуют из транзитивности и согласованности порядка со структурой $\mathbb{Z}$-модуля.

Свойство 7 является следствием свойств 1 и 5. $\square$

Заметим, что из леммы 3 следует, что отображение ограничения $\psi : \mathcal{O}_n \to \mathcal{O}(\mathbb{Z}^n)$, где $\mathcal{O}(\mathbb{Z}^n)$ — множество линейных согласованных порядков на $\mathbb{Z}^n$, является взаимно однозначным соответствием. В дальнейшем множество всех линейных согласованных порядков на $\mathbb{Z}^n$ будем обозначать также через $\mathcal{O}_n$.

Лемма 4. *Для любого линейного порядка $\prec$ на $\mathbb{Q}^n$, согласованного со структурой $\mathbb{Z}$-модуля, и любого ненулевого $\mathbb{Q}$-подпространства $V \subset \mathbb{Q}^n$ множество*

$$V_+ = \{v \in V \mid v \succ 0\}$$

является полупространством в V. Иными словами, существует единственный вектор a длины 1, находящийся в вещественном замыкании $\mathbb{R}(V)$ пространства $V \subset \mathbb{Q}^n \subset \mathbb{R}^n$, для которого

$$V_+ \subset \{v \in V \mid v \cdot a \geq 0\}.$$

Доказательство. Определим функцию $\mathrm{sgn} : \mathbb{R}(V) \to \{-1, 0, 1\}$ формулой

$$\mathrm{sgn}(x) = \begin{cases} -1, & \text{если} \quad \exists \varepsilon > 0 \mid \forall v \in V \cap U_\varepsilon(x)\ v \prec 0 \\ 0, & \text{если} \quad \forall \varepsilon > 0 \exists v_1, v_2 \in V \cap \dot{U_\varepsilon}(x) \mid v_1 \prec 0 \wedge v_2 \succ 0 \\ 1, & \text{если} \quad \exists \varepsilon > 0 \mid \forall v \in V \cap U_\varepsilon(x)\ v \succ 0, \end{cases}$$

где $U_\varepsilon(x)$ — ε-окрестность точки x в пространстве $\mathbb{R}^n$. Положим

$$H^+ = \mathrm{sgn}^{-1}(1), \quad V^0 = \mathrm{sgn}^{-1}(0), \quad H^- = \mathrm{sgn}^{-1}(-1).$$

Тогда $\mathbb{R}(V) = H^+ \cup V^0 \cup H^-$. Докажем выпуклость множества H^+.

Пусть $x, y \in H^+$ и точка z лежит на отрезке, соединяющем x и y, т.е. при некотором $0 < \alpha < 1$ выполняется соотношение $z = \alpha x + (1 - \alpha)y$.

Согласно определению x, существует $\varepsilon > 0$, что для всех $v \in U_\varepsilon(x) \cap V$ выполняется соотношение $v \succ 0$.

Выберем $\beta \in \mathbb{Q}$, удовлетворяющее условиям

1. $0 < \beta < 1$,
2. $\left|1 - \frac{\alpha}{\beta}\right| < \frac{\varepsilon}{4|x-y|}$.

Положим $x_1 = x + \left(1 - \frac{\alpha}{\beta}\right)(y - x)$. Из условия 2 следует, что $U_{\varepsilon/4}(x_1) \subset U_\varepsilon(x)$. Поэтому

$$z = \beta x_1 + (1-\beta)y,$$

где $\beta \in \mathbb{Q}$ и $x_1 \in H^+$.

Согласно определению x_1 и y, существует такое $\delta > 0$, что для всех $v \in U_\delta(x_1) \cap V$ и всех $u \in U_\delta(y) \cap V$ выполняются соотношения $v \succ 0$ и $u \succ 0$.

Представим некоторый произвольно выбранный $w \in U_{\delta/2}(z) \cap V$ в виде суммы $w = z + v_0$. Тогда $|v_0| < \delta/2$ и

$$z + v_0 = \beta(x_1 + v_0) + (1-\beta)(y + v_0),$$

причем $x_1 + v_0 \in U_{\delta/2}(x_1)$ и $y + v_0 \in U_{\delta/2}(y)$. Для любого $v_1 \in \mathbb{R}(V)$ выполнено соотношение

$$w = z + v_0 = \beta(x_1 + v_0 + v_1) + (1-\beta)\left(y + v_0 - \frac{\beta}{1-\beta} \cdot v_1\right).$$

Положим

$$\gamma = \min\{1, (1-\beta)/\beta\} \cdot \frac{\delta}{2} \leq \frac{\delta}{2}.$$

Поэтому существует $v_1 \in U_\gamma(0)$, что $x_1 + v_0 + v_1 \in V \cap U_\delta(x_1)$. Поскольку также $w \in V$ и $\beta \in \mathbb{Q}$, то и $y + v_0 - \frac{\beta}{1-\beta} \cdot v_1 \in V$, а в силу определения γ выполняется $y + v_0 - \frac{\beta}{1-\beta} \cdot v_1 \in U_\delta(y)$. Следовательно, $y + v_0 - \frac{\beta}{1-\beta} \cdot v_1 \in V \cap U_\delta(y)$. Тогда $x_2 = x_1 + v_0 + v_1 \succ 0$ и $y_2 = y + v_0 - \frac{\beta}{1-\beta} \cdot v_1 \succ 0$ и, следовательно, по лемме 3, $w = z + v_0 = \beta x_2 + (1-\beta)y_2 \succ 0$, т.е. множество H^+ выпукло.

Для доказательства выпуклости H^- заметим, что это множество центрально симметрично H^+.

Докажем, что множество V^0 является гиперплоскостью в пространстве $\mathbb{R}(V)$, т.е. существует такой вектор $v \in \mathbb{R}(V)$, что

$$V^0 = \{x \in \mathbb{R}(V) \mid x \cdot v = 0\}.$$

Сначала докажем, что V^0 является подпространством в $\mathbb{R}(V)$.

Заметим, что $0 \in V^0$. Пусть $x_1, \ \ldots, x_k \in V^0$ — ненулевые элементы и $z = a_1x_1 + \ \ldots + a_kx_k \in \mathbb{R}(V)$. Требуется доказать, что для любого $\varepsilon > 0$ существуют $v_1, v_2 \in U_\varepsilon(z) \cap V$ такие, что $v_1 \succ 0$, а $v_2 \prec 0$.

Итак, пусть заданы $\varepsilon > 0$ и $z = a_1x_1 + \ \ldots + a_kx_k \in \mathbb{R}(V)$. Без ограничения общности можно считать, что $a_i \neq 0, \ i = 1, \ \ldots, k$. Выберем такие $\alpha_i \in \mathbb{Q} \setminus \{0\}, \ i = 1, \ \ldots, k$, что

$$|a_i - \alpha_i| < \frac{\varepsilon}{4k|x_i|}.$$

Формула корректна, поскольку $x_i \neq 0$. Определим $z' = \alpha_1x_1 + \ \ldots + \alpha_kx_k$. Согласно условию $z' \in U_{\varepsilon/4}(z)$.

Положим $\varepsilon_i = \dfrac{\varepsilon}{4k\alpha_i|x_i|}, \ i = 1, \ \ldots, k$. По условию на x_i и лемме 3 существуют такие $x_i^+, x_i^- \in U_{\varepsilon_i}(x_i) \cap V, \ i = 1, \ \ldots, k$, что

- $\alpha_i x_i^+ \succ 0$,
- $\alpha_i x_i^- \prec 0$.

Тогда

$$z^+ = \alpha_1x_1^+ + \ \ldots + \alpha_kx_k^+ \in U_\varepsilon(z) \cap V$$

и

$$z^- = \alpha_1x_1^- + \ \ldots + \alpha_kx_k^- \in U_\varepsilon(z) \cap V$$

и, следовательно, по лемме 3 выполняются соотношения $z^+ \succ 0$ и $z^- \prec 0$. Поэтому, $z \in V^0$, т.е. V^0 является линейным подпространством в $\mathbb{R}(V)$.

Выберем в пространстве V базис $v_1, \ \ldots, v_m$. Поскольку эти векторы ненулевые, то для каждого v_i выполняется ровно одно из двух условий: $v_i \succ 0$ или $v_i \prec 0$. Положим

$$w_i = \begin{cases} v_i & \text{при} \quad v_i \succ 0, \\ -v_i & \text{при} \quad v_i \prec 0. \end{cases}$$

Векторы w_i также составляют базис. Рассмотрим отрезки $l_i \subset \mathbb{R}^n$, соединяющие w_1 и $-w_i$ для всех $i = 2, \ldots, n$. Поскольку множества H^+ и H^- согласно уже доказанному выпуклы, множества

$$S_i = \{\alpha \in \mathbb{Q} \cap (0, +\infty) \mid \alpha w_1 - (1-\alpha) w_i \succ 0\}$$

являются сечениями в рациональных числах. Из соотношений $w_1 \in H^+$ и $-w_i \in H^-$ при $i = 2, \ldots, m$ следует, что сечения S_i определяют вещественные числа $a_i \in (0, 1)$. Определим независимые векторы $z_i = a_i w_1 - (1 - a_i) w_i \in V^0$ при $i = 2, \ldots, m$. Согласно уже доказанному линейная оболочка над $\mathbb{R}$ этих векторов также содержится в пространстве V^0. Если бы в V^0 содержался хотя бы еще один независимый вектор, то V^0 совпало бы со всем пространством $\mathbb{R}(V)$. В таком случае $H^+ = \emptyset$. Но это не так. Докажем, что $w = w_1 + \ldots + w_m \in H^+$.

По определению векторов w_i и согласно лемме 3 для всех элементов $h = (h_1, \ldots, h_m) \in \mathbb{Q}^m$ в шаре радиуса $1/2$ выполняются соотношения

$$(1 + h_1) w_1 + \ldots + (1 + h_m) w_m \succ 0.$$

Поэтому $\mathrm{sgn}(w) = 1$ и, следовательно, $w \in H^+$.

Поэтому V^0 является гиперплоскостью в $\mathbb{R}(V)$, а гиперплоскость однозначно с точностью до знака $\sigma \in \{1, -1\}$ определяется вектором $\sigma b \in \mathbb{R}(V)$ длины 1

$$V^0 = \{x \in \mathbb{R}(V) \mid x \cdot b = 0\}.$$

Гиперплоскость V^0 разбивает пространство $\mathbb{R}(V)$ на три части: на два полупространства

$$V^+ = \{x \in \mathbb{R}(V) \mid x \cdot b > 0\} \quad \text{и} \quad V^- = \{x \in \mathbb{R}(V) \mid x \cdot b < 0\}$$

и саму гиперплоскость V^0.

Поскольку множество H^+ выпукло и не пересекается с V^0, оно лежит в одном из двух полупространств V^+ или V^-. Пусть это полупространство V^+. В этом случае полагаем $a = b$. Из равенства $\mathbb{R}(V) = H^+ \cup V^0 \cup H^- = V^+ \cup V^0 \cup V^-$ теперь следует, что $H^+ = V^+$. Аналогично, если выполнено соотношение $H^+ \subset V^-$, то полагаем $a = -b$ и также выполняется равенство $H^+ = V^-$. В обоих случаях

$$H^+ = \{x \in \mathbb{R}(V) \mid x \cdot a > 0\} \quad \text{и} \quad V^0 = \{x \in \mathbb{R}(V) \mid x \cdot a = 0\}.$$

Поскольку выполняется соотношение $V_+ \cap H^- = \emptyset$, то

$$V_+ \subset H^+ \cup V^0 = \{v \in V \mid v \cdot a \geq 0\}.$$

□

Доказательство теоремы 3. Докажем мономорфность отображения $\varphi : \mathcal{M}_n \to \mathcal{O}_n$.

Для каждой матрицы $A \in \mathcal{M}_n$ размера $n \times k$ определим последовательность $\mathbb{Q}$-пространств

$$\mathbb{Q}^n = V_0 \supsetneq V_1 \supsetneq \ \ldots \supsetneq V_k = \{0\}$$

по индукции. Пусть $V_i \subset \mathbb{Q}^n$ определено и $i < k$. Тогда скалярное произведение на вектор a_{i+1} определяет $\mathbb{Q}$-гомоморфизм $\varphi_{i+1} : V_i \to \mathbb{R}$. Положим $V_{i+1} = V_i \cap \ker(\varphi_{i+1})$. Из свойства 2 определения 3 следует выполнение соотношений

$$\dim_{\mathbb{Q}} V_{i+1} = \dim_{\mathbb{Q}} V_i - \dim_{\mathbb{Q}} \varphi_{i+1}(V_i) \quad \text{и} \quad V_i \supsetneq V_{i+1}. \tag{1.2}$$

Из свойства 4 определения 3 и соотношения (1.2) следует, что $V_k = \{0\}$.

Пусть матрицы $A, B \in \mathcal{M}$ не равны. Рассмотрим их первые не совпадающие столбцы $a_i \neq b_i$. В силу свойства 3 определения 3 возможны два случая: либо $a_i = -b_i$, либо эти столбцы не коллинеарны. В первом случае для любого ненулевого элемента $v \in V_{i-1}$ либо $v \succ_A 0$ и $v \preceq_B 0$, либо $v \succ_B 0$ и $v \preceq_A 0$, т.е. порядки $\succ_A$ и $v \succ_B$ различаются. Во втором случае рассмотрим $(i-1)$-е члены последовательностей подпространств, соответствующих матрицам A и B. Согласно построению $E = V_{i-1} = W_{i-1}$, а согласно свойству 2 определения 3 $a_i, b_i \in \mathbb{R}(V_i)$. Рассмотрим полупространства $V_{i-1}^+ = \{v \in V_{i-1} \mid v \cdot a_i > 0\}$ и $V_{i-1}^+ = \{w \in W_{i-1} \mid w \cdot b_i > 0\}$ пространства E. Поскольку $a_i \neq b_i$ выполняется соотношение $U = V_{i-1}^+ \setminus (V_{i-1}^+ \cap W_{i-1}^+) \neq \emptyset$. Выберем элемент $v \in U$. Тогда $v \succ_A 0$ и $v \prec_B 0$, т.е. порядки $\succ_A$ и $\succ_B$ различаются. Следовательно, отображение φ мономорфно.

Теперь докажем эпиморфность отображения φ.

Для заданного порядка $\prec \in \mathcal{O}_n$ построим по индукции ортонормальную последовательность векторов-столбцов $a_1, \ldots, a_k \in \mathbb{R}^n$ и $\mathbb{Q}$-подпространств $\mathbb{Q}^n = V_0 \supsetneq V_1 \supsetneq \ \ldots \supsetneq V_k = \{0\}$, для которой

$\dim_{\mathbb{Q}} V_i - \dim_{\mathbb{Q}} \varphi_{i+1}(V_i) = \dim_{\mathbb{Q}} V_{i+1}$, где $\varphi_i : V_i \to \mathbb{R}$ определяется скалярным умножением на вектор a_i.

Начало индукции: $m = 0$. Выбираем пустую последовательность векторов-столбцов и пространство $\mathbb{Q}^n = V_0$.

Шаг индукции. Пусть построены ортонормальная последовательность векторов-столбцов $a_1, \ldots, a_m$ и последовательность $\mathbb{Q}$-подпространств $\mathbb{Q}^n \supsetneq V_1 \supsetneq \ldots \supsetneq V_m$. Если $V_m = \{0\}$, то $m = k$ и построение завершено. Если $V_m \neq \{0\}$, то продолжим построение. Определим вектор a_{m+1} длины 1 и $\mathbb{Q}$-подпространство V_{m+1}.

Рассмотрим множество $H^+_{m+1} = \{v \in V_m \mid v \succ 0\}$. Согласно лемме 4 существует единственный вектор $a_{m+1} \in \mathbb{R}(V_m)$ длины 1, для которого

$$H^+_{m+1} \subseteq \{v \in V_m \mid v \cdot a_{m+1} \geq 0\}.$$

Положим

$$V_{m+1} = \{v \in V_m \mid v \cdot a_{m+1} = 0\}.$$

Очевидно, что $\dim_{\mathbb{Q}} V_{m+1} = \dim_{\mathbb{Q}} V_m - \dim_{\mathbb{Q}} \varphi_{m+1}(V_i)$. Поскольку $\dim_{\mathbb{Q}} \mathbb{Q}^n = n$, то $m \leq n$ и процедура заканчивается не более чем за n шагов. Итак, каждому упорядочению, согласованному со структурой $\mathbb{Z}$-модуля, однозначно сопоставляется последовательность векторов $a_1, \ldots, a_k$. Из свойств векторов-столбцов $a_1, \ldots, a_k$ и соотношений $\dim_{\mathbb{Q}} V_{m+1} = \dim_{\mathbb{Q}} V_m - \dim_{\mathbb{Q}} \varphi_{m+1}(V_i)$ следует, что $n \times k$-матрица A, состоящая из этих столбцов, принадлежит множеству $\mathcal{M}_n$. Из построения также следует, что $\succ_A = \succ$. $\square$

Для каждого натурального числа n определим подмножество $\mathcal{M}^+_n$ в множестве $\mathcal{M}_n$. Матрица $A \in \mathcal{M}_n$ принадлежит множеству $\mathcal{M}^+_n$ тогда и только тогда, когда в каждой строке матрицы A имеется ненулевой элемент и первый ненулевой элемент строки положительный.

Обозначим через $\mathcal{O}^+_n$ множество всех допустимых порядков на множестве $\mathbb{Z}^n_+$. Из определения множества $\mathcal{M}^+_n$ и теоремы 2 следует, что для матрицы $A \in \mathcal{M}^+_n$ порядок $\succ_A$ всегда допустимый на множестве $\mathbb{Z}^n_+$, т.е. $\succ_A \in \mathcal{O}^+_n$. Поэтому ограничение $\varphi : \mathcal{M}_n \to \mathcal{O}_n$ на $\mathcal{M}^+_n$ задает отображение $\varphi^+ : \mathcal{M}^+_n \to \mathcal{O}^+_n$. Воспользовавшись теоремой 1, получаем

Следствие 1. *Пусть $\mathcal{M}^+_n$ — определенное выше множество матриц, а $\mathcal{O}^+_n$ — множество допустимых порядков на $\mathbb{Z}^n_+$. Тогда отображение φ^+ :*

$\mathcal{M}_n^+ \to \mathcal{O}_n^+$ определяет взаимно-однозначное соответствие между множествами $\mathcal{M}_n^+$ и $\mathcal{O}_n^+$.

Существенным недостатком следствия 1 является присутствие вещественных чисел в описании допустимых упорядочений на $\mathbb{Z}_+^n$. Следующая теорема показывает, как описание порядка может быть сделано конечным.

Теорема 4. *Для любого допустимого порядка $\succ$ на $\mathbb{Z}_+^n$ и любого $m \in \mathbb{N}$ существует такая $\mathbb{Q}$-матрица A размера $n \times n$, удовлетворяющая условиям теоремы 2, что допустимые порядки $\succ$ и $\succ_A$ совпадают на всех термах от n переменных степени не выше m.*

Доказательство. Термы будем интерпретировать здесь как элементы пространства $\mathbb{Z}_+^n$. Тогда степень терма $v \in \mathbb{Z}_+^n$ совпадает с нормой вектора v в $\mathbb{Z}^n$, заданной суммой абсолютных величин компонентов вектора v. Обозначим ее через $|v|$. Пусть допустимое упорядочение $\succ$ задается вещественной $n \times k$-матрицей

$$B = \begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,k} \\ \cdots & \cdots & \cdots \\ \alpha_{n,1} & \dots & \alpha_{n,k} \end{pmatrix} \in \mathcal{M}_n^+$$

ранга $k \leq n$. Тогда существует такое $\delta > 0$, что любая δ-аппроксимация матрицы B также имеет ранг k.

Найдем такое приближение $a_i \in \mathbb{Q}^n$ i-го столбца матрицы B, что для любых $v, w \in \mathbb{Z}_+^n \mid |v| < m,\ |w| < m$ соотношения

$$\alpha_{1,i}v_1 + \ \dots + \alpha_{n,i}v_n > \alpha_{1,i}w_1 + \ \dots + \alpha_{n,i}w_n \tag{1.3}$$

и

$$a_{1,i}v_1 + \ \dots + a_{n,i}v_n > a_{1,i}w_1 + \ \dots + a_{n,i}w_n \tag{1.4}$$

эквивалентны.

Эквивалентные соотношения (1.3) и (1.4) перепишем в виде

$$\alpha_{1,i}(v_1 - w_1) + \ \dots + \alpha_{n,i}(v_n - w_n) > 0 \tag{1.5}$$

и

$$a_{1,i}(v_1 - w_1) + \ \ldots + a_{n,i}(v_n - w_n) > 0. \tag{1.6}$$

Поскольку множество M разностей $v - w \in \mathbb{Z}^n$, для которых $|v| < m$ и $|w| < m$ и выполнено соотношение (1.5), конечно, то существует такое $\delta_i > 0$, что

$$\alpha_{1,i}(v_1 - w_1) + \ \ldots + \alpha_{n,i}(v_n - w_n) > \delta_i > 0 \tag{1.7}$$

для всех $v - w \in M$.

Чтобы неравенства (1.5) и (1.6) были эквивалентны на M достаточно, чтобы на M выполнялось неравенство

$$|(\alpha_{1,i} - a_{1,i})(v_1 - w_1) + \ \ldots + (\alpha_{n,i} - a_{n,i})(v_n - w_n)| < \delta_i/2. \tag{1.8}$$

Поскольку для всех $v - w \in M$ выполнено неравенство $|v - w| < 2m$, то

$$\begin{array}{l} |(\alpha_{1,i} - a_{1,i})(v_1 - w_1) + \ \ldots + (\alpha_{n,i} - a_{n,i})(v_n - w_n)| < \\ 2|\alpha_{1,i} - a_{1,i}|m + \ \ldots + |\alpha_{n,i} - a_{n,i}|m = \\ 2(|\alpha_{1,i} - a_{1,i}| + \ \ldots + |\alpha_{n,i} - a_{n,i}|)m. \end{array} \tag{1.9}$$

Выбирая теперь аппроксимацию только ненулевых компонент матрицы B рациональными числами $a_{j,i}$ с погрешностью, не превосходящей $\min\left\{\dfrac{\delta_i}{2mn}, \delta\right\}$, получаем, что неравенства (1.3) и (1.4) эквивалентны на M, а полученная матрица A удовлетворяет условиям теоремы 2.

Если $k < n$, то добавим к полученным столбцам $n - k$ рациональных столбцов так, чтобы получилась невырожденная матрица. Если $n = k$, выберем такую аппроксимацию, чтобы матрица A оставалась невырожденной.

Отметим, что матрица A в общем случае не принадлежит множеству $\mathcal{M}_n^+$. □

Следствие 2. *Для любого допустимого порядка $\succ$ и любого $m \in \mathbb{N}$ существует такая $\mathbb{Q}$-матрица A размера $n \times n$, состоящая из неотрицательных элементов, что допустимые порядки $\succ$ и $\succ_A$ совпадают на всех термах степени не выше m.*

Доказательство. Согласно теореме 4 существует такая невырожденная $\mathbb{Q}$-матрица A размера $n \times n$, в которой первый ненулевой элемент каждой строки положителен, а допустимые порядки $\succ$ и $\succ_A$ совпадают на всех термах степени не выше m. Теперь прибавим ко всем столбцам этой матрицы такое кратное первого столбца, состоящего из неотрицательных элементов, чтобы в полученной матрице строки, соответствующие положительным элементам первого столбца, стали бы положительными. Полученная матрица будет определять тот же порядок, а первый ненулевой элемент каждой строки останется положительным. Теперь второй столбец полученной матрицы состоит из неотрицательных элементов. Прибавим его кратные к столбцам матрицы, начиная с третьего столбца, чтобы получить положительные строки, соответствующие положительным элементам этого столбца. Повторяя эту процедуру для всех столбцов, получим матрицу из неотрицательных элементов, задающую тот же порядок, что и исходная матрица A. □

Рассмотрим теперь допустимое упорядочение $\succ_A$, заданное неотрицательной рациональной $n \times n$-матрицей A. Пусть $v = (v_1, \ \ldots, v_n) \in \mathbb{Z}^n$. Положим

$$\omega_i(v) = \sum_{k=1}^{n} v_k a_{k,i}, \ \ i = 1, \ \ldots, n.$$

Согласно определению упорядочения $\succ_A$ для любых $v, w \in \mathbb{Z}_+^n$

$$v \succ_A w \Leftrightarrow (\exists 1 \leq k \leq n \,|\, \omega_k(v) > \omega_k(w) \wedge (\forall i < k \, \omega_i(v) = \omega_i(w))). \quad (1.10)$$

Представим рациональные компоненты матрицы A в виде отношений взаимно простых неотрицательных целых чисел

$$a_{i,j} = \frac{m_{i,j}}{n_{i,j}}, \ \ i, j = 1, \ \ldots, n.$$

Положим $N = \max\{m_{i,j}, \text{НОК}(n_{i,j}) \mid i, j = 1, \ \ldots, n\}$. Для произвольного рационального $B > 1$ определим рациональную линейную форму W_B формулой

$$W_B(v) = \frac{1}{N} \sum_{i=1}^{n} N\omega_i(v) \cdot B^{n-i} = \sum_{i=1}^{n} \omega_i(v) \cdot B^{n-i}. \quad (1.11)$$

По определению N для всех $i = 1, \ldots, n$ и всех таких $v \in \mathbb{Z}^n$, что $\omega_i(v) \neq 0$, выполняется соотношение $N|\omega_i(v)| \geq 1$. Фиксируем натуральное число M. Тогда для любого $v \in \mathbb{Z}^n$, удовлетворяющего неравенству

$$|v| = |v_1| + \ldots + |v_n| < M,$$

для всех $i = 1, \ldots, n$ выполняются неравенства $|\omega_i(v)| < nNM$. Пусть теперь $B = nN^2M + 2$. Докажем, что для всех v, w, таких, что $|v| < M/2$ и $|w| < M/2$, соотношения $v \succ_A w$ и $W_B(v) > W_B(w)$ эквивалентны.

Фиксируем пару $v \succ w$, что $|v| < M/2$ и $|w| < M/2$. Выберем такое число k, что $\omega_i(v - w) = 0$ при всех $i < k$ и $\omega_k(v - w) \neq 0$. Соотношение $W_B(v - w) > 0$ в этом случае эквивалентно соотношению

$$\omega_k(v - w) > -\sum_{i=k+1}^{n} \omega_i(v - w)B^{k-i}.$$

Теперь для доказательства эквивалентности соотношений $\omega_k(v - w) > 0$ и $W_B(v - w) > 0$ согласно формуле (1.10) достаточно проверить, что при $|v| < M/2$ и $|w| < M/2$ всегда

$$|\omega_k(v - w)| > \left|\sum_{i=k+1}^{n} \omega_i(v - w)B^{k-i}\right|.$$

Ввиду неравенств $N|\omega_k(v - w)| \geq 1$ и $|\omega_i(v - w)| < nNM$, достаточно доказать, что

$$\frac{1}{N} > nNM \sum_{i=k+1}^{\infty} B^{k-i} > nNM \sum_{i=k+1}^{n} B^{k-i}.$$

Последние неравенства следуют из определения числа B.

Таким образом, получено

Следствие 3. *Для любого допустимого порядка $\succ_A$, определяемого рациональной матрицей A, существует такая рациональная форма W_B, что при $|v| < M$ и $|w| < M$ соотношения $v \succ_A w$ и $W_B(v) > W_B(w)$ эквивалентны.*

Переформулируем последнее утверждение применительно к термам. В силу изоморфизма моноида термов $T\langle X\rangle$ и моноида $\mathbb{Z}_+^n$ можно считать, что форма W_B определена на термах.

Следствие 4. *Пусть рациональная матрица A определяет допустимый порядок $\succ_A$ на множестве термов $T\langle X\rangle$ и N — максимум числителей и наименьшего общего кратного знаменателей несократимого представления рациональных коэффициентов матрицы A в виде отношения натуральных чисел. Тогда при $B = nN^2M$ соотношения $v \succ_A w$ и $W_B(v) > W_B(w)$ эквивалентны на множестве термов степени не выше M.*

1.3 Стандартные обозначения

В последующих разделах будут использованы следующие обозначения.

Пусть $X = \{x_1, \ldots, x_n\}$ — множество переменных, A — коммутативное кольцо с 1 и $\succ$ — допустимый порядок на $T\langle X\rangle$.

Упорядочим множество термов многочлена $p \in A[X]$ относительно порядка $\succ$

$$T_p\langle X\rangle = \{t_1, \ldots, t_k\},$$

где $t_1 \succ t_2 \ldots \succ t_k$. В этом случае будем применять запись $T_p\langle X\rangle = [t_1, \ldots, t_k]$. Тогда формулу (10.5) можно записать как

$$p = p(X) = \sum_{i=1}^{k} a_i t_i. \tag{1.12}$$

В дальнейшем будем использовать следующие обозначения и названия

- $\mathsf{HC}(p) = a_1$ — старший коэффициент многочлена p,
- $\mathsf{HT}(p) = t_1$ — старший терм многочлена p,
- $\mathsf{HM}(p) = a_1 t_1$ — старший моном многочлена p.

Если F подмножество в кольце многочленов $A[X]$, то определим множества его старших термов и старших мономов формулами

$$\mathsf{HT}(F) = \{\mathsf{HT}(f) \mid f \in F\} \quad \text{и} \quad \mathsf{HM}(F) = \{\mathsf{HM}(f) \mid f \in F\}\ .$$

Глава 2

Нетеровы кольца и модули

В дальнейшем, если не оговорено противное, будем рассматривать только коммутативные кольца с 1.

2.1 Модули. Идеалы

Определение 2.1.1. *Абелева группа M называется модулем над кольцом A, или A-модулем, если определена операция умножения $A \times M \to M$, удовлетворяющая условиям:*

- *Для любого $a \in A$ и любых $m_1, m_2 \in M$ выполняется равенство*

$$a(m_1 + m_2) = am_1 + am_2,$$

- *Для любых $a, b \in A$ и любого $m \in M$ выполняется равенство*

$$(a + b)m = am + bm,$$

- *Для любых $a, b \in A$ и любого $m \in M$ выполняется равенство*

$$(ab)m = a(bm),$$

- $1 \cdot m = m \quad \forall m \in M.$

Подмодулем модуля M называется любая его подгруппа, замкнутая относительно умножения на элементы кольца A.

Определение 2.1.2. *Подмножество N модуля M над кольцом A называется системой образующих (базисом) модуля, если любой элемент $m_0 \in M$ представим в виде суммы*

$$m_0 = \sum_{m \in N} a_m m, \;\; a_m \in A,$$

содержащей конечное число ненулевых слагаемых. Модуль M называется конечно порожденным, если существует конечная система его образующих.

Отметим, что в определении базиса модуля, в отличие от базиса векторного пространства, не требуется независимости элементов.

Примеры.

1 Векторное пространство $\mathbb{Q}^n$ размерности n является модулем над полем рациональных чисел. Систему образующих этого модуля составляют векторы $\{e_i \mid i = 1, \; \ldots, n\}$, где $e_i = (\delta_{1,i}, \; \ldots, \delta_{n,i})$, а $\delta_{i,j}$ — символ Кронекера

$$\delta_{i,j} = \begin{cases} 1 & \text{при} \quad i = j, \\ 0 & \text{при} \quad i \neq j. \end{cases}$$

2 Аддитивная группа рациональных чисел является $\mathbb{Z}$-модулем. Базис этого модуля составляет, например, множество

$$M = \left\{ \frac{1}{p^n} \;\middle|\; p \in \mathbb{P}, n \in \mathbb{N} \right\},$$

где $\mathbb{P}$ — множество простых чисел.

3 Кольцо вычетов по модулю n является $\mathbb{Z}$-модулем с единицей в качестве образующей.

4 Любая абелева группа A является $\mathbb{Z}$-модулем. Операция умножения $a \in A$ на целое число n задается формулой:

$$n \cdot a = \begin{cases} \underbrace{a + \ \ldots + a}_{n \text{ слагаемых}} & \text{при} \quad n > 0, \\ 0 & \text{при} \quad n = 0, \\ \underbrace{-a - \ \ldots - a}_{-n \text{ слагаемых}} & \text{при} \quad n < 0. \end{cases}$$

В частности, любое кольцо или поле является $\mathbb{Z}$-модулем.

Кольцо A, рассматриваемое как аддитивная группа (относительно операции сложения в этом кольце) с операцией умножения на элементы из A, представляет собой A-модуль. Такой A-модуль будем называть модулем, соответствующим кольцу A.

Определение 2.1.3. *Подмодуль модуля, соответствующего кольцу A, называется идеалом этого кольца. Идеал, имеющий одноэлементный базис, называется главным. Главный идеал I кольца A с образующим a обозначается через (a), т.е. $I = (a)$.*

Определение 2.1.4. *Ядром гомоморфизма колец*

$$\varphi : A \to B,$$

называется прообраз нуля, обозначаемый $\ker \varphi$.

Лемма 5. *Пусть $\varphi : A \to B$ — гомоморфизм колец. Тогда* $\ker \varphi$ *— идеал кольца A.*

Доказательство. Следует непосредственно из определений 2.1.4, 2.1.3 и определения гомоморфизма. □

Примеры.

1 Нулевой идеал — $I = (0)$.

2 Единичный идеал — $I = (1) = A$.

3 Пусть $n \in \mathbb{Z}$. Тогда идеал $(n) \subset \mathbb{Z}$ состоит из целых чисел, делящихся на n.

4 Пусть $a_1, \ \ldots, a_n \in A$ — конечный набор элементов кольца. Подмножество кольца A, состоящее из всех элементов вида

$$\sum_{i=1}^{n} \alpha_i a_i,$$

где α_i — произвольные элементы кольца A, является идеалом кольца A и обозначается через $(a_1, \ \ldots, a_n)$.

5 Пусть $A_0 \subset A$. Подмножество кольца A, состоящее из всех элементов вида

$$\sum_{a \in A_0} \alpha_a a,$$

где α_i — произвольные элементы кольца A и только конечное число элементов α_a не равно нулю, является идеалом кольца A и обозначается через (A_0). Элементы множества A_0 называются образующими этого идеала.

Определение 2.1.5. *Кольцо A называется кольцом главных идеалов, если любой идеал этого кольца главный.*

Теорема 5. *Кольцо целых чисел $\mathbb{Z}$ является кольцом главных идеалов.*

Доказательство. Нулевой идеал, очевидно, является главным. Пусть I ненулевой идеал кольца целых чисел. Согласно определению идеала, если $x \in I$, то и $-x = (-1)x \in I$. Если $x \neq 0$, то одно из чисел x или $-x$ положительно. Рассмотрим множество всех положительных элементов идеала I. Поскольку это множество содержится в множестве натуральных чисел, оно содержит наименьший элемент $x_0 \neq 0$. Проверим, что этот элемент является образующим идеала I.

Пусть это не так, т.е. $I \neq I_0 = \{kx_0 \mid k \in \mathbb{Z}\}$. Следовательно существует элемент $x \in I$, не делящийся на x_0. Выполним деление с остатком

$$x = qx_0 + d, \text{ где } 0 < d < x_0.$$

Тогда d — положительный элемент идеала, меньший элемента x_0, что противоречит определению элемента x_0. □

Аналогично доказывается

Теорема 6. *Кольцо многочленов $K[x]$ от одной переменной над полем K является кольцом главных идеалов.*

Пример. Идеал $(x^2, 2x)$ кольца многочленов $\mathbb{Z}[x]$ не является главным.

Поэтому условие, что коэффициенты кольца многочленов лежат в поле, в данной теореме существенно.

Теорема 7. ***(О конечно порожденных модулях)*** *Пусть A — коммутативное кольцо, а M — A-модуль. Следующие условия эквивалентны:*

1. *Всякий подмодуль в M конечно порожден.*
2. *Всякая возрастающая цепочка подмодулей в M конечна.*
3. *Всякое непустое множество подмодулей в M имеет максимальный относительно отношения включения элемент.*

Доказательство. $1 \Rightarrow 2$. Пусть

$$M_1 \subset M_2 \subset M_3 \subset \ \ldots$$

— бесконечная возрастающая цепочка подмодулей. Положим $M_0 = \bigcup\limits_{i=1}^{\infty} M_i$. Очевидно, M_0 является подмодулем модуля M. Поскольку согласно предположению о модуле M все его подмодули M_i конечно порождены, в них существуют конечные системы образующих $N_i \subset M_i$. Положим $N_0 = \bigcup\limits_{i=1}^{\infty} N_i$. Тогда N_0 — система образующих модуля M_0. Множество N_0 не более чем счетно, поэтому можно перенумеровать его элементы: $N_0 = \{\beta_1, \ \ldots\}$. Согласно предположению, в A-модуле M_0 имеется конечная система образующих $N = \{\alpha_1, \ \ldots, \alpha_k\}$. Поскольку A-модуль M_0 порожден множеством N_0, каждый элемент множества N можно представить в виде конечной суммы вида:

$$\alpha_i = \sum_{j=1}^{k_i} a_{i,j}\beta_j, \ \ a_{i,j} \in A.$$

Положим $s = \max\limits_{i=1,\ldots,k} k_i$. Тогда $\{\beta_1, \ldots, \beta_s\} \subseteq N_0$ — конечная система образующих в M_0. Поэтому при некотором $m > 0$ верно

$$\{\beta_1, \ldots, \beta_s\} \subseteq \bigcup_{i=1}^{m} N_i.$$

Следовательно, $M_0 \subset M_m$ и, значит, по определению модуля M_0, выполняется равенство $M_0 = M_m$, и, следовательно, при всех $i \geq m$ выполняются равенства $M_i = M_{i+1}$, что противоречит предположению о бесконечности возрастающей цепочки подмодулей.

$2 \Rightarrow 3$. Пусть $\mathcal{M}$ — некоторое непустое множество подмодулей в M, не имеющее максимального элемента. Поскольку $\mathcal{M}$ не пусто, в нем существует хотя бы один элемент, например, $M_1 \in \mathcal{M}$. Положим $S_1 = \{M_1\}$. Предположим, что для некоторого n построена строго возрастающая цепочка S_n

$$M_1 \subset \ \ldots \subset M_n, \quad M_i \in \mathcal{M} \quad \forall i = 1, \ \ldots n.$$

Поскольку в множестве $\mathcal{M}$ не существует максимального относительно включения элемента, существует модуль $M_{n+1} \in \mathcal{M}$, строго содержащий модуль M_n. Таким образом, определена цепочка S_{n+1}

$$M_1 \subset \ \ldots \subset M_{n+1}, \quad M_i \in \mathcal{M} \quad \forall i = 1, \ \ldots n+1,$$

причем $M_i \neq M_j$ при $i \neq j$ и т.д. Поэтому имеется бесконечная строго возрастающая цепочка подмодулей модуля M вопреки условию 2.

$3 \Rightarrow 1$. Пусть M_0 — некоторый подмодуль модуля M и $\mathcal{S}$ — множество всех его конечных подмножеств. Обозначим через M_S подмодуль, порожденный образующими $S \in \mathcal{S}$. Положим $\mathcal{M} = \{M_S \mid S \in \mathcal{S}\}$. Тогда согласно условию 3 в $\mathcal{M}$ существует максимальный элемент. Предположим, что это M_{S_0}. Тогда $M_{S_0} = M_0$. Действительно, если это не так, то существует $\alpha \in M_0 \setminus M_{S_0}$. Положим $S_1 = S_0 \cup \{\alpha\}$. Тогда M_{S_1} строго включает M_{S_0}, что противоречит максимальности модуля M_{S_0}. Поэтому $M_{S_0} = M_0$, и конечное множество S_0 является базисом подмодуля M_0. □

Определение 2.1.6. *Модуль называется нетеровым, если всякий подмодуль в нем конечно порожден. Кольцо называется нетеровым, если модуль, соответствующий этому кольцу, является нетеровым.*

Итак, кольцо A нетерово тогда и только тогда, когда каждый его идеал имеет конечную систему образующих. В частности, кольцо главных идеалов нетерово. Следовательно, кольцо целых чисел $\mathbb{Z}$ нетерово.

Определение 2.1.7. *Будем называть идеал* $\mathfrak{m}$ *делителем идеала* $\mathfrak{n}$*, если выполняется соотношение* $\mathfrak{m} \supset \mathfrak{n}$*. Если* $\mathfrak{m} \supsetneq \mathfrak{n}$*, то будем называть* $\mathfrak{m}$ *собственным делителем идеала* $\mathfrak{n}$*.*

Теорема 7 о конечно порожденных модулях может быть переформулирована так:

Теорема 8. *Пусть* A *— нетерово кольцо. Тогда всякая неубывающая бесконечная цепочка делителей*

$$\mathfrak{m}_1 \subset \mathfrak{m}_2 \subset \mathfrak{m}_3 \subset \ \ldots,$$

стабилизируется, т.е., начиная с некоторого n*, выполняются равенства*

$$\mathfrak{m}_n = \mathfrak{m}_{n+1} = \ \ldots.$$

Теорема 9. ***(Принцип индукции по делителям).*** *Пусть* E *— предикат на множестве идеалов нетерового кольца* A*, обладающий следующим свойством*

- $\forall \mathfrak{m}_0 \ (\forall \mathfrak{m} \supsetneq \mathfrak{m}_0 \ E(\mathfrak{m}) = 1) \Rightarrow E(\mathfrak{m}_0)$

Тогда $E(\mathfrak{m}) = 1$ *для всех идеалов* $\mathfrak{m}$ *кольца* A*.*

Доказательство. Для единичного идеала $E(A) = 1$, поскольку этот идеал не имеет собственных делителей.

Пусть теперь $E(\mathfrak{m}_0) = 0$ для некоторого идеала $\mathfrak{m}_0 \subsetneq A$. Тогда в непустом множестве идеалов, для которых $E(\mathfrak{m}) = 0$, имеется максимальный идеал $\mathfrak{m}_1$. В силу его максимальности равенство $E(\mathfrak{m}) = 1$ выполняется для всех его собственных делителей, а потому должно выполняться и для самого идеала $\mathfrak{m}_1$. Следовательно, $E(\mathfrak{m}) = 1$ для всех идеалов кольца. $\square$

Определение 2.1.8. *Идеал* $\mathfrak{p}$ *кольца* A *называется простым, если* $\mathfrak{p} \neq A$ *и из* $xy \in \mathfrak{p}$ *следует, что либо* $x \in \mathfrak{p}$*, либо* $y \in \mathfrak{p}$*.*

Пусть I — идеал кольца A. Определим факторкольцо A/I следующим образом. Введем отношение $\equiv$ на элементах кольца формулой

$$a \equiv b \Leftrightarrow a - b \in I.$$

Очевидно, что $\equiv$ является отношением эквивалентности. Класс эквивалентности элемента a будем обозначать через $a + I$. Тогда

$$a \equiv b \Leftrightarrow a + I = b + I.$$

Определим операции на классах эквивалентности формулами

$$(a + I) + (b + I) = (a + b) + I. \tag{2.1}$$

и

$$(a + I)(b + I) = ab + I. \tag{2.2}$$

Нетрудно убедиться, что операции сложения и умножения на классах эквивалентности определены корректно и множество классов эквивалентности с так определенными операциями является кольцом. Определим отображение

$$\pi : A \to A/I \tag{2.3}$$

формулой $\pi(a) = a + I$. Это отображение является эпиморфизмом колец.

Для любого натурального $n > 1$ идеал $I = (n) \subset \mathbb{Z}$ определяет кольцо $\mathbb{Z}/(n)$. Это кольцо называется кольцом вычетов по модулю n. Если n — простое, то это кольцо является полем и обозначается $\mathbb{F}_n$.

Пусть заданы вложение колец $A \subset B$ и B_0 — подмножество кольца B. Тогда определены множество X_{B_0} и эпиморфизм колец многочленов (см. формулу ((10.8)) раздела 10.1)

$$\varphi : A[X_{B_0}] \to A[B_0].$$

Очевидно, выполнено

Предложение 1. *Отображение*

$$i : A[X_{B_0}]/\ker\varphi \to A[B_0],$$

определенное формулой $i(p + \ker\varphi) = \varphi(p)$*, задано корректно и является изоморфизмом.*

Элемент a кольца A называется делителем нуля (кольца A), если существует такой $b \in A$, что $b \neq 0$ и $ab = 0$. Коммутативное кольцо, не имеющее делителей нуля, называется целостным. Легко проверить, что выполняется следующее

Предложение 2. *Идеал* p *целостного кольца* A *прост тогда и только тогда, когда факторкольцо* $A/$p *не имеет делителей нуля.*

Определение 2.1.9. *Идеал* I *кольца* A *называется максимальным, если* $A \neq I$ *и не существует содержащего* I *идеала, отличного от* A*.*

Предложение 3. *Для максимального идеала* I *в коммутативном кольце факторкольцо* A/I *является полем.*

Доказательство. Пусть I — максимальный идеал. Очевидно, что нулем и единицей факторкольца A/I являются различные элементы I и $1 + I$ соответственно. Поэтому достаточно проверить существование обратных элементов. Если для $a + I$ не существует обратного элемента и $a \notin I$, то определим идеал J присоединением к базису идеала I элемента a . Идеал $J \neq I$ и $I \subset J$. Поэтому $J = A$, т.е. $1 \in J$. Следовательно (согласно определению 2.1.2 базиса идеала), существуют такие $b \in A$ и $c \in I$, что $ab + c = 1$. Тогда $(a + I)(b + I) = 1 + I$ в факторкольце A/I, т.е. элемент $b + I$ является обратным для $a + I$. □

Предложение 4. *Для любого собственного идеала* I *кольца существует содержащий его максимальный идеал.*

Доказательство. Рассмотрим множество всех идеалов в A, содержащих идеал I и отличных от кольца A. Это множество не пусто, поскольку в него входит идеал $I \neq A$, и индуктивно относительно порядка $\subset$. Поэтому по лемме Цорна это множество содержит максимальный элемент. □

Предложение 5. *Главный идеал* (p) *для неприводимого многочлена* $p \in k[x]$*, где* k *— поле, является максимальным в кольце* $k[x]$*.*

Доказательство. Пусть главный идеал (p) некоторого неприводимого многочлена $p \in k[x]$ не является максимальным. Тогда существует содержащий его максимальный идеал I. Поскольку $k[x]$ — кольцо главных идеалов $I = (q) \neq (p)$ для некоторого многочлена $q \in k[x]$. Следовательно, $p = qr$ для некоторого $r \in k[x]$, что означает приводимость многочлена p.

□

Из предложений 3 и 5 получаем

Следствие 5. *Для любого неприводимого многочлена $p \in k[x]$ факторкольцо $k[x]/(p)$ является полем.*

Пусть A — целостное кольцо. Тогда множество A^* его ненулевых элементов замкнуто относительно умножения в этом кольце. Легко проверить, что отношение $\sim$ на множестве пар $A \times A^*$, заданное формулой

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc,$$

является отношением эквивалентности. Зададим операции сложения и умножения на множестве пар $A \times A^*$ формулами

- $(a, b) + (c, d) = (ad + cb, bd)$,
- $(a, b) \cdot (c, d) = (ac, bd)$.

Легко проверить, что так определенные операции инвариантны относительно отношения $\sim$. Поэтому определены операции на классах эквивалентности, причем множество классов эквивалентности является полем относительно введенных операций. Построенное поле называется полем частных целостного кольца A.

Поле частных кольца многочленов от переменных $x_1, \ldots, x_n$ с кольцом коэффициентов в поле k называется кольцом рациональных функций от переменных $x_1, \ldots, x_n$ с кольцом коэффициентов в поле k и обозначается $k(x_1, \ldots, x_n)$.

2.2 Теорема о разложении идеалов

Определение 2.2.1. *Частным идеала* $\mathfrak{m}$ *кольца* A *и множества* $\mathfrak{b} \subset A$ *называется множество*

$$\mathfrak{m} : \mathfrak{b} = \{x \in A | \; xa \in \mathfrak{m}, \text{для всех } a \in \mathfrak{b}\}.$$

Это множество, очевидно, является идеалом кольца A.

Определение 2.2.2. *Идеал, порожденный элементами объединения двух идеалов, называется суммой этих идеалов, или наибольшим общим делителем этих идеалов. Для суммы идеалов* m *и* n *используют обозначение* (m, n). *Если* $\mathsf{n} = (a)$, *то используют обозначение* (m, a).

Пересечение двух идеалов также является идеалом и называется наименьшим общим кратным этих идеалов. Для пересечения идеалов m *и* n *используют обозначение* $[\mathsf{m}, \mathsf{n}]$.

Идеал, порожденный попарными произведениями элементов идеалов m *и* n, *называется произведением идеалов и обозначается через* mn.

Определение 2.2.3. *Идеал, не представимый в виде пересечения двух собственных делителей, называется неприводимым.*

Теорема 10. *В нетеровом кольце любой идеал представим в виде пересечения неприводимых идеалов.*

Доказательство. Пусть E — утверждение о представимости идеалов в виде пересечения неприводимых идеалов.

Поскольку единичный идеал, совпадающий как множество с самим кольцом, не имеет собственных делителей, то E выполняется для единичного идеала.

Рассмотрим произвольный идеал m. Предположим, что E выполняется для всех его собственных делителей. Докажем, что в этом случае E выполняется и для идеала m.

Если идеал неприводим, то он имеет тривиальное представление требуемого вида.

Если же идеал приводим, представим его как пересечение двух собственных делителей

$$\mathsf{m} = [\mathsf{m}_1, \mathsf{m}_2].$$

Согласно сделанному предположению, собственные делители представимы как пересечения неприводимых идеалов, т.е.

$$\mathsf{m}_1 = [\mathsf{m}_{1,1}, \ \ldots, \mathsf{m}_{n,1}]$$
$$\mathsf{m}_2 = [\mathsf{m}_{1,2}, \ \ldots, \mathsf{m}_{k,2}],$$

где $\mathsf{m}_{i,j}$ — неприводимые идеалы. Тогда идеал m представим в виде пересечения неприводимых идеалов

$$\mathsf{m}_1 = [\mathsf{m}_1, \mathsf{m}_2] = [\mathsf{m}_{1,1}, \ \ldots, \mathsf{m}_{n,1}, \mathsf{m}_{1,2}, \ \ldots, \mathsf{m}_{k,2}].$$

Следовательно, согласно принципу индукции по делителям (теорема 9) каждый идеал представим в виде пересечения неприводимых идеалов. □

Определение 2.2.4. *Идеал* q *кольца* A *называется примарным, если для любых* $x, y \in A$ *из* $xy \in \mathfrak{q}$ *следует, что либо* $x \in \mathfrak{q}$*, либо* $y^\alpha \in \mathfrak{q}$ *при некотором натуральном* α*.*

Пример. Идеал $\mathfrak{p} = (2, x)$ является простым в кольце $\mathbb{Z}[x]$, а идеал $\mathfrak{q} = (4, x)$ является примарным в кольце $\mathbb{Z}[x]$. Единственным нетривиальным (отличным от единицы) простым делителем идеала q является идеал p. При этом $\mathfrak{p}^2 = \mathfrak{pp} = (4, 2x, x^2) \subsetneq \mathfrak{q}$.

Предложение 6. *Для каждого примарного идеала* $\mathfrak{q} \subset A$ *существует делящий его простой идеал* p*, определяемый формулой*

$$\mathfrak{p} = \{a \in A \mid \exists \alpha \in \mathbb{N} \,:\, a^\alpha \in \mathfrak{q}\}.$$

Доказательство. Докажем, что данное множество действительно является идеалом. Если $a \in \mathfrak{p}$, то существует такое $\alpha \in \mathbb{N}$, что $a^\alpha \in \mathfrak{q}$. Следовательно, учитывая что q является идеалом, для любого $b \in A$ выполнено $(ab)^\alpha = a^\alpha b^\alpha \in \mathfrak{q}$. Поэтому $ab \in \mathfrak{p}$.

Если $a, b \in \mathfrak{p}$, то $a^\alpha, b^\beta \in \mathfrak{q}$ при некоторых натуральных α, β и, следовательно, $(a-b)^{\alpha+\beta} \in \mathfrak{q}$, поскольку после раскрытия скобок каждое слагаемое содержит хотя бы один из сомножителей a^α или b^β. Следовательно, $a - b \in \mathfrak{p}$.

Докажем теперь, что этот идеал прост. Пусть

$$ab \in \mathfrak{p} \quad \text{и} \quad a \notin \mathfrak{p}.$$

Тогда при некотором натуральном α

$$a^\alpha b^\alpha \in \mathfrak{q} \quad \text{и} \quad a^\alpha \notin \mathfrak{q}.$$

Тогда по свойству примарности идеала q при некотором натуральном β имеет место включение

$$b^{\alpha\beta} \in \mathfrak{q}.$$

Следовательно, согласно определению идеала p, верно включение

$$b \in \mathfrak{p}.$$

И, наконец, поскольку $a^1 \in \mathfrak{q}$ для любого элемента $a \in \mathfrak{q}$, согласно определению идеала p выполняется включение $\mathfrak{q} \subset \mathfrak{p}$. □

Определение 2.2.5. *Идеал* p*, сопоставляемый в предложении 6 примарному идеалу* q*, называется простым идеалом, ассоциированным с примарным идеалом* q*.*

Условия ассоциированности простого и примарного идеалов приведены в следующем предложении.

Предложение 7. *Пусть для идеалов* p *и* q *выполняются свойства*

1. *Если* $ab \in \mathfrak{q}$ *и* $a \notin \mathfrak{q}$*, то* $b \in \mathfrak{p}$*.*
2. $\mathfrak{q} \subset \mathfrak{p}$*.*
3. *Если* $b \in \mathfrak{p}$*, то* $b^\alpha \in \mathfrak{q}$ *для некоторого натурального* α*.*

Тогда идеал q *примарный, а* p *— ассоциированный с ним простой идеал.*

Доказательство. Если $ab \in \mathfrak{q}$ и $a \notin \mathfrak{q}$, то из условий 1 и 3 следует, что $b^\alpha \in \mathfrak{q}$ при некотором натуральном α. Поэтому идеал q примарный.

Осталось проверить, что

$$\mathfrak{p} = \{a \in A | a^\alpha \in \mathfrak{q} \text{ при некотором натуральном } \alpha\}.$$

Согласно условию 3, если $b \in \mathfrak{p}$, то $b^\alpha \in \mathfrak{q}$ для некоторого натурального α. Поэтому достаточно проверить, что если $b^\alpha \in \mathfrak{q}$ для некоторого натурального α, то $b \in \mathfrak{p}$. Выберем минимальное α при котором $b^\alpha \in \mathfrak{q}$. Если $\alpha = 1$, то из условия 2 следует, что $b \in \mathfrak{p}$. Пусть теперь $\alpha > 1$. Тогда $b^{\alpha-1}b \in \mathfrak{q}$ и $b^{\alpha-1} \notin \mathfrak{q}$ и, следовательно, согласно условию 1 выполняется соотношение $b \in \mathfrak{p}$. □

Теорема 11. *Неприводимый идеал нетерова кольца является примарным.*

Доказательство. Пусть некоторый идеал m не является примарным. Докажем, что он приводим.

Поскольку идеал m не является примарным, существуют такие элементы $a, b \in A$, что

$$ab \in \mathsf{m},$$
$$a \notin \mathsf{m},$$
$$b^\alpha \notin \mathsf{m} \text{ для всех натуральных } \alpha.$$

Поскольку кольцо A нетерово, то в силу теоремы 8 цепочка делителей

$$\mathsf{m} : \{b\} \subset \mathsf{m} : \{b^2\} \subset \ \ldots$$

стабилизируется, т.е. при некотором k выполняется равенство $\mathsf{m} : \{b^k\} = \mathsf{m} : \{b^{k+1}\}$. Докажем, что выполняется равенство

$$\mathsf{m} = [(\mathsf{m}, a), (\mathsf{m}, b^k)].$$

Достаточно проверить включение $(\mathsf{m}, a) \cap (\mathsf{m}, b^k) \subset \mathsf{m}$. Пусть элемент c находится в пересечении этих идеалов. Поскольку $c \in (\mathsf{m}, a)$ и $c \in (\mathsf{m}, b^k)$, то при некоторых $m, m' \in \mathsf{m}$, и $r, r' \in A$, выполняются соотношения

$$c = m + rb^k = m' + r'a.$$

Тогда $cb = m'b + r'ab \in \mathsf{m}$, поскольку $ab \in \mathsf{m}$. Поэтому $cb = mb + rb^{k+1} \in \mathsf{m}$. Следовательно, $rb^{k+1} \in \mathsf{m}$, и, согласно определению 2.2.1, $r \in \mathsf{m} : b^{k+1} = \mathsf{m} : b^k$. Поэтому $rb^k \in \mathsf{m}$ и, следовательно, $c = m + rb^k \in \mathsf{m}$. Включение доказано. Заметим, что делители (m, a) и (m, b^k) являются собственными, поскольку $a \notin \mathsf{m}$ и $b^k \notin \mathsf{m}$. Следовательно, идеал приводим. □

Из теорем о представлении идеала в виде пересечения неприводимых идеалов и примарности неприводимого идеала (теоремы 10, 11) следует

Теорема 12. ***(О разложении идеалов)*** *Каждый идеал нетерова кольца представим в виде пересечения конечного числа примарных идеалов.*

Из такого представления

$$\mathsf{m} = [\mathsf{q}_1, \ \ldots, \mathsf{q}_r]$$

можно исключить идеалы q_i, содержащие пересечения

$$\bigcap_{k\neq i} q_k.$$

Полученное после таких исключений представление называется несократимым.

Предложение 8. *Для любого кольца выполняются следующие свойства:*

1. *Класс примарных идеалов, ассоциированных с данным простым идеалом, замкнут относительно конечных пересечений.*

2. *Несократимое пересечение примарных идеалов, не ассоциированных с одним и тем же простым идеалом, не является примарным идеалом.*

Доказательство. 1. Воспользуемся критерием ассоциированности примарного и простого идеала из предложения 7. Пусть

$$\mathsf{m} = [\mathsf{q}_1,\ \ldots, \mathsf{q}_r]$$

и все примарные идеалы q_i ассоциированы с простым идеалом p. Пусть $ab \in \mathsf{m}$ и $a \notin \mathsf{m}$. Тогда для всех i выполняются соотношения $ab \in \mathsf{q}_i$ и $a \notin \mathsf{q}_k$ при некотором k. Из ассоциированности идеалов q_k и p и соотношения $a \notin \mathsf{q}_k$ следует, что $b \in \mathsf{p}$, т.е. для идеалов m и p выполняется свойство 1 предложения 7.

Поскольку свойство 2 предложения 7 выполняется для всех идеалов q_i, оно тем более выполняется и для их пересечения, т.е. для идеала m.

И, наконец, если $b \in \mathsf{p}$, то при всех i и некоторых α_i выполняются соотношения

$$b^{\alpha_i} \in \mathsf{q}_i.$$

Поэтому при $\alpha = \max\limits_i \alpha_i$ выполняется соотношение

$$b^{\alpha} \in \mathsf{q}_i \quad \text{для всех } i$$

и, следовательно,

$$b^{\alpha} \in \mathsf{m}.$$

Утверждение 1 доказано.

Докажем теперь утверждение 2. Пусть имеется несократимое пересечение примарных идеалов

$$\mathsf{m} = [\mathsf{q}_1, \ \ldots, \mathsf{q}_r]. \tag{2.4}$$

Если заменить группы идеалов, ассоциированных с одним и тем же простым идеалом, их пересечением, то в силу уже доказанного утверждения, снова получим несократимое пересечение примарных идеалов, ассоциированных с различными простыми идеалами. Далее будем считать несократимое пересечение (2.4) таким.

Рассмотрим теперь среди конечного множества простых идеалов p_i, ассоциированных с примарными идеалами q_i из пересечения (2.4), минимальный, т.е. не содержащий ни один из остальных, идеал, например, p_1. Тогда существуют элементы

$$a_i \in \mathsf{p}_i, \ i = 2, 3, \ \ldots, r \mid \forall i \ \ a_i \notin \mathsf{p}_1.$$

Поэтому при достаточно большом α выполняются соотношения

$$a_i^{\alpha} \in \mathsf{q}_i \quad \text{при } i = 2, \ \ldots, r.$$

Если бы $\mathsf{q}_1 = \mathsf{m}$, то представление $\mathsf{m} = [\mathsf{q}_1, \ \ldots, \mathsf{q}_r]$ было бы сократимым. Поэтому существует элемент $q_1 \in \mathsf{q}_1$, не принадлежащий идеалу m. Следовательно, произведение

$$q_1(a_2 \ \ldots a_r)^{\alpha}$$

принадлежит идеалам $\mathsf{q}_1, \mathsf{q}_2, \ \ldots, \mathsf{q}_r$ и $q_1 \notin \mathsf{m}$. Если идеал m был бы примарным, то отсюда бы следовало, что

$$(a_2 \ \ldots a_r)^{\alpha\beta} \in \mathsf{m}$$

при некотором натуральном β. Поэтому из включений

$$\mathsf{m} = [\mathsf{q}_1, \ \ldots, \mathsf{q}_r] \subset \mathsf{q}_1 \subset \mathsf{p}_1$$

следует, что

$$(a_2 \ \ldots a_r)^{\alpha\beta} \in \mathsf{p}_1.$$

А ввиду простоты идеала $\mathfrak{p}_1$, при некотором $i = 2, \ldots, r$ выполняется соотношение $a_i \in \mathfrak{p}_1$, что противоречит выбору элементов a_i при $i = 2, \ldots, r$. □

Из теоремы 12 и предложения 8 получаем в качестве следствия вторую теорему Э. Ласкера о разложении примарными компонентами (см. [**Lask**]).

Теорема 13. *Каждый идеал I нетерова кольца допускает несократимое представление в виде пересечения конечного множества примарных идеалов, ассоциированных с попарно различными простыми идеалами.*

Примарные идеалы из теоремы 13 будем называть примарными компонентами идеала I. Заметим, что несократимое представление идеала не единственно.

Пример. Идеал $\mathfrak{m} = (x^2, xy) \subset K[x, y]$ представим как несократимые пересечения примарных компонент

$$\mathfrak{m} = [\mathfrak{q}_1, \mathfrak{q}_2]$$

и

$$\mathfrak{m} = [\mathfrak{q}_1, \mathfrak{q}_3],$$

где $\mathfrak{q}_1 = (x)$, $\mathfrak{q}_2 = (x^2, xy, y^2)$ и $\mathfrak{q}_3 = (x^2, y)$.

Тем не менее справедлива

Теорема 14. *Для любых двух несократимых представлений идеала* $\mathfrak{m}$ *наибольшими примарными компонентами количество компонент и наборы ассоциированных простых идеалов совпадают.*

Доказательство. Доказательство проведем индукцией по минимальному числу примарных компонент идеала $\mathfrak{m}$.

Для примарного идеала утверждение, очевидно, следует из предложения 8. Пусть теорема доказана для идеалов, минимальное несократимое представление которых имеет длину не более $l - 1$.

Пусть имеются два представления

$$\mathfrak{m} = [\mathfrak{q}_1, \ \ldots, \mathfrak{q}_{l_1}] = [\mathfrak{q}'_1, \ \ldots, \mathfrak{q}'_{l_2}], \tag{2.5}$$

где $l \leq l_i$ при $i = 1, 2$. Из всех ассоциированных простых идеалов

$$\mathfrak{p}_1, \ \ldots, \mathfrak{p}_{l_1}, \mathfrak{p}'_1, \ \ldots, \mathfrak{p}'_{l_2}$$

выберем максимальный. Без ограничения общности можно считать, что это идеал $\mathfrak{p}_1$. Докажем, что в таком случае $\mathfrak{p}_1$ совпадает с одним из простых идеалов $\mathfrak{p}'_1, \ \ldots, \mathfrak{p}'_{l_2}$. Поделим обе части равенства (2.5) на идеал $\mathfrak{q}_1$. Получим

$$[\mathfrak{q}_1 : \mathfrak{q}_1, \ \ldots, \mathfrak{q}_{l_1} : \mathfrak{q}_1] = [\mathfrak{q}'_1 : \mathfrak{q}_1, \ \ldots, \mathfrak{q}'_{l_2} : \mathfrak{q}_1].$$

Поскольку идеал $\mathfrak{p}_1$ не ассоциирован ни с одним из идеалов

$$\mathfrak{q}_2, \ \ldots, \mathfrak{q}_{l_1}, \mathfrak{q}'_1, \ \ldots, \mathfrak{q}'_{l_2},$$

выполняются равенства

$$\begin{aligned} \mathfrak{q}_i : \mathfrak{q}_1 &= \mathfrak{q}_i \quad (i = 2, \ \ldots, l_1) \\ \mathfrak{q}'_i : \mathfrak{q}_1 &= \mathfrak{q}'_i \quad (i = 1, \ \ldots, l_2) \end{aligned}$$

и $\mathfrak{q}_1 : \mathfrak{q}_1 = (1)$. Поэтому выполняется равенство

$$[\mathfrak{q}_2, \ \ldots, \mathfrak{q}_{l_1}] = [\mathfrak{q}'_1, \ \ldots, \mathfrak{q}'_{l_2}],$$

т.е. представление

$$\mathfrak{m} = [\mathfrak{q}_1, \ \ldots, \mathfrak{q}_{l_1}]$$

не является несократимым. Следовательно, идеал $\mathfrak{p}_1$ совпадает с одним из идеалов $\mathfrak{p}'_1, \ \ldots, \mathfrak{p}'_{l_2}$.

Без ограничения общности можно считать, что $l = l_1 \leq l_2$ и $\mathfrak{p}_1 = \mathfrak{p}'_1$. Разделим теперь обе части равенства (2.5) на идеал $\mathfrak{q}_1\mathfrak{q}'_1$. Получим неприводимые представления

$$[\mathfrak{q}_2, \ \ldots, \mathfrak{q}_l] = [\mathfrak{q}'_2, \ \ldots, \mathfrak{q}'_{l_2}]$$

примарными компонентами некоторого идеала, причем длина первого представления $l - 1$. Поэтому по предположению индукции для такого идеала теорема выполняется, т.е. разложения $[\mathfrak{q}_2, \ \ldots, \mathfrak{q}_l]$ и $[\mathfrak{q}'_2, \ \ldots, \mathfrak{q}'_{l_2}]$ совпадают. Следовательно, разложения $[\mathfrak{q}_1, \ \ldots, \mathfrak{q}_{l_1}]$ и $[\mathfrak{q}'_1, \ \ldots, \mathfrak{q}'_{l_2}]$ идеала $\mathfrak{m}$ также совпадают. □

Глава 3

Теоремы Гильберта

3.1 Теорема Гильберта о базисе

Следующая теорема известна как теорема Гильберта о базисе.

Теорема 15. *Кольцо многочленов $A[x]$ над нетеровым кольцом A нетерово.*

Доказательство. Пусть I — идеал в кольце многочленов $A[x]$. Найдем конечный базис в этом идеале. Поскольку кольцо A нетерово, идеал $J_\infty \subset A$, порожденный коэффициентами при старших термах многочленов из I, имеет конечный базис $\{a_1, \ldots, a_m\}$. Для каждого элемента a_i этого базиса выберем такой многочлен $q_i(x) \in I$, что коэффициент при его старшем терме равен a_i. Множество этих многочленов обозначим $Q_\infty = \{q_1(x), \ldots, q_m(x)\}$. Пусть n — максимальная из степеней многочленов $\{q_1(x), \ldots, q_m(x)\}$. Тогда любой элемент $p(x) \in I$ можно представить в виде

$$p(x) = \sum_{i=1}^{m} c_i(x) q_i(x) + p_n(x), \;\; c_i(x) \in A[x],$$

где $p_n(x) \in I$ — многочлен степени, меньшей n. Действительно, пусть $p(x) = \sum\limits_{i=0}^{k} b_i x^i$, причем b_k — ненулевой элемент кольца A. Если $k < n$,

то утверждение очевидно. Пусть $k \geq n$ и $b_k = \sum\limits_{i=1}^{m} \alpha_i a_i$, где $\alpha_i \in A$ (такое представление возможно, поскольку $b_k \in J_\infty$ лежит в идеале J_∞, а множество $\{a_1, \ldots, a_m\}$ является базисом этого идеала). В таком случае

$$p(x) = \sum_{i=1}^{m} \alpha_i q_i(x) x^{k-n_i} + p_{k-1}(x), \tag{3.1}$$

где $n_i = \deg q_i(x)$, $\deg p_{k-1}(x) < k$ и $p_{k-1}(x) \in I$. Если $\deg p_{k-1}(x) < n$, то получим искомое представление. В противном случае применим к p_{k-1} описанную выше процедуру (3.1) разложения многочлена $p(x)$. Получим p_{k-2} и либо найдем требуемое представление, либо снова повторим эту процедуру. Поскольку на каждом шаге степень многочлена снижается не менее чем на 1, то за конечное число шагов требуемое представление будет найдено.

Чтобы завершить доказательство, осталось найти конечный базис для разложения многочленов идеала I, степень которых не превосходит $n-1$.

Пусть $I_k = \{p \in I \mid \deg(p) = k\}$, а J_k — идеал кольца A, порожденный коэффициентами при старших термах этих многочленов. Поскольку кольцо A нетерово, в идеалах J_k существуют конечные базисы. Пусть $\{a_{1,k}, \ldots, a_{m_k,k}\}$ — базис идеала J_k и пусть $Q_k = \{q_{1,k}(x), \ldots, q_{m_k,k}(x)\}$ — многочлены множества I_k, старшими коэффициентами которых являются элементы множества $\{a_{1,k}, \ldots, a_{m_k,k}\}$ соответственно. Положим $Q = \bigcup\limits_{k<n} Q_k$.

Поскольку степень многочлена $p_n(x) \in I$ меньше n, этот многочлен можно представить в виде

$$p_n(x) = \sum_{i=0}^{n-1} b_i x^i.$$

Старший коэффициент b_k ($k \leq n-1$) многочлена $p_n(x)$ принадлежит идеалу J_k и, следовательно, может быть выражен через элементы $\{a_{1,k}, \ldots, a_{m_k,k}\}$:

$$b_k = \sum_{i=1}^{m_k} \beta_i a_{i,k}.$$

Тогда многочлен

$$p_n(x) - \sum_{i=1}^{m_k} \beta_i q_{i,k}(x)$$

имеет степень меньше k и также лежит в идеале I. Поэтому, используя эту процедуру не более чем n раз, получаем разложение многочлена

$$p_n(x) = \sum_{k=0}^{n-1} \sum_{i=1}^{m_k} \beta_i q_{i,k}(x).$$

Следовательно, конечное множество $Q_\infty \cup Q$ является базисом идеала I. □

Следствие 6. *Для нетерова кольца A кольцо многочленов $A[x_1, \ldots, x_n]$ нетерово. В частности, кольцо многочленов над полем нетерово.*

Обозначим $X = \{x_1, \ldots, x_n\}$.

Определение 3.1.1. *Идеал I кольца многочленов $A[x_1, \ldots, x_n]$ называется мономиальным, если для каждого многочлена*

$$p(X) = \sum_{\omega \in \mathbb{Z}_+^n} a_\omega X^\omega \in I \Rightarrow \forall \omega \in \mathbb{Z}_+^n,\ a_\omega X^\omega \in I.$$

Непосредственно из определения 3.1.1 получаем

Предложение 9. *Идеал, порожденный произвольным множеством мономов, является мономиальным.*

Следствие 7. *В мономиальном идеале кольца многочленов над нетеровым кольцом существует конечный базис из мономов.*

Доказательство. Если N — конечный базис мономиального идеала, то множество всех мономов всех многочленов этого базиса конечно и, согласно определению мономиального идеала, лежит в идеале. Следовательно, является его конечным базисом. □

В произвольном моноиде S можно определить отношение делимости. Элемент x делит элемент y (записывают как $x|y$), если существует такой z, что $xz = y$. Отношение делимости переносится на произвольные подмножества M и N моноида S.

Определение 3.1.2. *Пусть M и N подмножества моноида S. Множество M делит множество N, если для любого $n \in N$ существует такой $m \in M$, что $m|n$.*

Следствие 8. *Пусть X — конечное множество переменных. В любом бесконечном множестве термов $M \subset T\langle X\rangle$ существует конечное подмножество $M^0 \subset M$, делящее множество M.*

Это следствие известно как лемма Диксона.

Доказательство. Пусть I — мономиальный идеал кольца $\mathbb{Z}_2[X]$, порожденный элементами множества M (см. предложение 9). Согласно следствию 7 из теоремы 15 в I существует конечный базис из мономов $N = \{n_1, \ldots, n_k\}$, а учитывая, что кольцом коэффициентов является $\mathbb{Z}_2$, этот базис состоит из термов. Согласно определению идеала I множество M делит мономы идеала I, и, в частности, множество N, т.е. для каждого $1 \leq i \leq k$ существует моном m_i, делящий n_i. Поэтому множество $M^0 = \{m_1, \ldots, m_k\}$ делит множество M. □

Следствие 9. *Пусть A — поле и X — конечное множество переменных. В любом бесконечном множестве мономов $M \subset A[X]$ существует конечное подмножество $M^0 \subset M$, делящее множество M.*

Доказательство. Пусть $M_T = \{\mathrm{HT}(M)\}$. Согласно следствию 8 существует подмножество $M_T^0 \subset M_T$, делящее множество M_T. Согласно определению множества M_T, для каждого $t \in M_T$ существует моном $m(t) \in M$, термом которого является t. Тогда конечное множество $M^0 = \{m(t) \mid t \in M_T^0\}$ делит M, поскольку A — поле. □

Замечание. Если A не является полем, то утверждение следствия 9 не выполняется.

Например, пусть $A = \mathbb{Z}$, $X = \{x\}$ и

$$M = \{2x^2, 3x^3, \ldots, px^p, \ldots\},$$

где $2, 3, \ldots$ — бесконечная возрастающая последовательность простых чисел. Тогда никакое конечное множество $N \subset M$ не делит множество M. Действительно, пусть это не так и такое множество $N \subset M$ существует. Выберем в N элемент n максимальной степени, а в M — элемент m большей степени. Очевидно, что m не делится ни на один элемент множества N, т.е. множество N не делит множество M.

3.2 Теорема Гильберта о нулях

3.2.1 Расширения полей

В этом разделе приведены необходимые определения и результаты из теории полей. Подробно с теорией полей можно познакомиться, например, по книгам [Вар76] или [Лен68].

Определение 3.2.1. *Расширением поля k называется любое сохраняющее операции вложение полей $k \subset K$. Поле k в этом случае называется подполем поля k.*

Отметим, что умножение элементов поля k на элементы поля K задает на K естественную структуру векторного k-пространства.

Определение 3.2.2. *Поле k, не содержащее подполя, отличного от него, называется простым.*

Предложение 10. *Для каждого поля k существует единственное простое подполе.*

Доказательство. Возьмем пересечение всех подполей поля k. Это множество содержит элементы 0 и 1 и является согласно определению простым полем. □

Напомним, что любое поле k является $\mathbb{Z}$-модулем.

Определение 3.2.3. *Полем характеристики 0 называется поле k, в котором при любом натуральном n выполняется соотношение $n \cdot e \neq 0$, где e — единица поля k. Для остальных полей (не имеющих характеристику 0) характеристика поля определяется как наименьшее натуральное число n, для которого выполняется равенство $n \cdot e = 0$.*

Число 1, очевидно, не может быть характеристикой поля, поскольку в поле выполнено неравенство $e \neq 0$.

Предложение 11. *Характеристика любого поля либо простая, либо нулевая.*

Доказательство. Предположим, что характеристика поля равна $n \neq 0$ и не является простой. Тогда $n = pq$ для некоторых $p, q < n$. Имеем

$$0 = n \cdot e = pq \cdot e = (p \cdot e)(q \cdot e).$$

Поскольку в поле нет делителей 0, то либо $p \cdot e = 0$, либо $q \cdot e = 0$. Следовательно, характеристика поля меньше n. Поэтому предположение неверно и n простое. □

Следствие 10. *Простое подполе поля характеристики 0 изоморфно полю рациональных чисел, для поля ненулевой характеристики p — полю вычетов по модулю p.*

Доказательство. Предположим, что характеристика поля k равна 0. Очевидно, что поле рациональных чисел не содержит собственных подполей. Поэтому достаточно доказать, что k содержит в качестве подполя поле рациональных чисел. Формула $n \mapsto n \cdot e$ определяет гомоморфное вложение колец $\mathbb{Z} \subset k$, продолжающееся до гомоморфизма $\varphi : \mathbb{Q} \to k$ полей частных, являющегося вложением, т.к. гомоморфизм полей всегда инъективен.

Пусть теперь характеристика поля k ненулевая, тогда она равна некоторому простому числу p. В поле сравнений по простому модулю p нет собственных подполей. Поэтому достаточно проверить существование в k подполя, изоморфного $\mathbb{Z}_p$. Формула $n \mapsto n \cdot e$ определяет гомоморфизм колец $\varphi : \mathbb{Z} \to k$. Согласно определению характеристики поля и гомоморфизма φ выполняется равенство $\ker \varphi = (p)$. Поэтому определен инъективный гомоморфизм колец

$$\tilde{\varphi} : \mathbb{Z}/(p) \to k$$

и из равенства $\mathbb{Z}/(p) = \mathbb{F}_p$ следует, что простым полем для k является $\mathbb{F}_p$. □

Определение 3.2.4. *Пусть $k \subset K$ — расширение поля k и S — подмножество поля K. Тогда поле частных кольца многочленов $k[S]$[1] называется*

[1] См. определение 10.1.14 кольца многочленов над произвольным множеством.

расширением поля k с помощью элементов S. Это поле обозначается через $k(S)$ и называется полем рациональных функций над k на множестве переменных S. В этом случае говорят, что поле $k(S) \supset k$ получено присоединением к полю k элементов S.

Данное определение корректно, поскольку кольцо многочленов $k[S]$ является подкольцом в поле K и, следовательно, является целостным, и выполняется $k \subset k(S) \subset K$. Из определения, очевидно, следует, что поле $k(S)$ является наименьшим подполем поля K, содержащим поле k и множество S. В частности, получаем

Предложение 12. *Пусть $S = S_1 \cup S_2$. Тогда $k(S) = k(S_1)(S_2)$.*

Определение 3.2.5. *Элементы множества $S \subset K$ расширения $k \subset K$ называются алгебраически независимыми, если*

$$\forall m \geq 0 \, \forall \theta_1, \ \ldots, \theta_m \in S \, \forall p \in k[x_1, \ \ldots, x_m] \, p(\theta_1, \ \ldots, \theta_m) = 0 \Leftrightarrow p = 0.$$

Мощность максимального множества S, состоящего из алгебраически независимых элементов, называется степенью трансцендентности расширения $k \subset K$.

Из инвариантности размерности векторного пространства над произвольным полем следует

Предложение 13. *Любые два максимальные множества S, состоящие из алгебраически независимых элементов, равномощны.*

Предложение 14. *Пусть имеется последовательность расширений $k \subset K \subset L$. Тогда степень трансцендентности расширения $k \subset L$ равна сумме трансцендентностей расширений $k \subset K$ и $K \subset L$.*

Доказательство. Следует непосредственно из определения 3.2.5. □

Пусть теперь множество S состоит из одного элемента θ.

Определение 3.2.6. *Пусть $k \subset K$ — расширение поля k, и θ — элемент поля K. Тогда поле $k(\theta)$ называется простым расширением поля k.*

Рассмотрим произвольное простое расширение $k(\theta)$, где $\theta \in K$. Тогда соответствие $x \mapsto \theta$ задает эпиморфизм φ кольца многочленов $k[x]$ в кольцо $k[\theta]$, а, следовательно, и изоморфизм $\tilde{\varphi} : k[x]/\ker\varphi \xrightarrow{\approx} k[\theta]$. Поскольку кольцо многочленов $k[x]$ является кольцом главных идеалов, $\ker\varphi = (p)$ для некоторого многочлена $p \in k[x]$. Имеются две возможности: $p = 0$ или $p \neq 0$.

В случае $p = 0$ получаем изоморфизм целостных колец $\tilde{\varphi} : k[x] \xrightarrow{\approx} k[\theta]$, а, следовательно, и изоморфизм $\tilde{\varphi} : k(x) \xrightarrow{\approx} k(\theta)$ соответствующих полей частных. Полученное расширение $k(\theta)$ называется простым трансцендентным расширением, а элемент θ — трансцендентным над полем k элементом. Размерность векторного пространства $k(\theta)$ над полем k бесконечна.

Докажем, что при $p \neq 0$ многочлен p является неприводимым в кольце $k[x]$. Пусть это не так. Тогда существует разложение $p = p_1 p_2$, где $\deg p_i < p$. Тогда

$$0 = \tilde{\varphi}(p) = \tilde{\varphi}(p_1)\tilde{\varphi}(p_2) \in k[\theta] \subset k(\theta).$$

Тогда $\tilde{\varphi}(p_1) = 0$ или $\tilde{\varphi}(p_2) = 0$, поскольку $k(\theta) \subset K$. Пусть, например, $\tilde{\varphi}(p_1) = 0$. Тогда $p_1 \in \ker\varphi$ и, следовательно, $\ker\varphi \neq (p)$. Поэтому p неприводим и согласно следствию 5 факторкольцо $k[x]/(p)$ является полем. Поскольку соответствие $x \mapsto \theta$ задает изоморфизм колец

$$\tilde{\varphi} : k[x]/(p) \xrightarrow{\approx} k[\theta], \tag{3.2}$$

кольцо $k[\theta]$ является полем и выполняется соотношение $p(\theta) = 0$. Такой элемент θ называется алгебраическим над полем k, а соответствующее расширение $k(\theta)$ — простым алгебраическим расширением поля k. Если $\deg p = n$, то каждый элемент поля $k[\theta] = k(\theta)$ однозначно представим многочленом $r(\theta)$ степени, меньшей n. Поэтому как k-векторное пространство $k[\theta]$ имеет базис $1, \theta, \ldots, \theta^{n-1}$. Элементы θ^i соответствуют элементам $x^i + (p)$ (при изоморфизме $\tilde{\varphi}$), а алгебраические операции над элементами такого векторного пространства описываются операциями в $k[x]/(p)$. Следовательно, справедливо

Предложение 15. *Простое расширение $k \subset k(\theta)$ трансцендентно тогда и только тогда, когда размерность векторного k-пространства $k(\theta)$ бесконечна. Соответственно, простое расширение алгебраично*

тогда и только тогда, когда его размерность как векторного k-пространства конечна.

Определение 3.2.7. *Элемент $\theta \in K$ для расширения $k \subset K$ называется алгебраическим над k, если существует такой $p \in k[x]$, что $p(\theta) = 0$.*

Определение 3.2.8. *Расширение $k \subset K$ называется алгебраическим над k, если все элементы поля K алгебраичны над k.*

Определение 3.2.9. *Расширения $k \subset K$ и $k \subset K'$ называются эквивалентными, если существует изоморфизм полей*

$$\tilde{\varphi} : K \xrightarrow{\sim} K',$$

тождественный на k. Эквивалентность расширений K и K' будем записывать в виде соотношений $K \cong K'$.

Очевидно, что два простых трансцендентных расширения произвольного поля k эквивалентны.

Предложение 16. *Два простых алгебраических расширения $k(\alpha)$ и $k(\beta)$ эквивалентны, если α и β являются корнями одного и того же неприводимого многочлена в $k[x]$. В этом случае существует изоморфизм эквивалентности $k(\alpha) \cong k(\beta)$, преобразующий α в β.*

Доказательство. Пусть $p(x)$ — неприводимый многочлен, определяющий простые расширения $k(\alpha)$ и $k(\beta)$. Тогда согласно формуле (3.2) имеются изоморфизмы полей

$$\tilde{\varphi} : k[x]/(p) \xrightarrow{\sim} k[\alpha]$$

и

$$\tilde{\psi} : k[x]/(p) \xrightarrow{\sim} k[\beta].$$

Композиция

$$\tilde{\psi} \circ \tilde{\varphi}^{-1} : k[\alpha] \to k[\beta]$$

определяет требуемый изоморфизм полей. □

Многочлен $p \in k[x]$, неприводимый над полем k и определяющий расширение $k(\alpha)$, не обязан оставаться неприводимым над расширением $k \subset k(\alpha)$. Рассмотрим его разложение на неприводимые множители над полем $k(\alpha)$

$$p(x) = (x - \theta_1) \cdot \ \ldots \cdot (x - \theta_j) \cdot p_1(x) \cdot \ \ldots \cdot p_m(x), \ \ m \geq 0,$$

где степени многочленов p_i больше 1 при $i = 1, \ \ldots, m$. Согласно предложению 16 все расширения $k(\theta_i)$ эквивалентны

$$k(\theta_1) \cong k(\theta_2) \cong \ \ldots \cong k(\theta_j),$$

где эквивалентности $k(\theta_i) \cong k(\theta_j)$ задаются преобразованиями $\theta_i \mapsto \theta_j$.

Определение 3.2.10. *Эквивалентные конечные расширения $k(\theta_i)$, лежащие в общем поле K, а также элементы θ_i называются сопряженными относительно k.*

Таким образом, доказано обобщение предложения 16

Предложение 17. *Все корни неприводимого в $k[x]$ многочлена, принадлежащие расширению $k \subset K$, являются сопряженными относительно k.*

Очевидно, верно и обратное. Сопряженные над полем k элементы являются корнями одного и того же многочлена.

Простое расширение определяется как подполе некоторого поля K (см. определение 3.2.6). Рассмотрим теперь следующие задачи. Дано поле k. Требуется построить расширение $k \subset K$ и $\theta \in K$, чтобы элемент θ: 1) был бы трансцендентным над k; 2) являлся корнем наперед заданного неприводимого многочлена $p \in k[x]$.

Для первой задачи решением является поле рациональных функций $K = k(x)$, а в качестве элемента θ можно выбрать x. Очевидно, что это поле определено однозначно с точностью до эквивалентности полей.

Во втором случае для неприводимого многочлена $p \in k[x]$ положим $K = k[x]/(p)$. Согласно следствию 5 это факторкольцо является полем. Формула (2.3) определяет эпиморфизм колец

$$\pi : k[x] \to k[x]/(p). \tag{3.3}$$

Положим $\theta = \pi(x)$. Тогда выполняется равенство $k(\theta) = k[x]/(p)$ и, следовательно, θ удовлетворяет требуемым условиям.

Напомним, что для расширения $k \subset K$ поле K является векторным пространством над полем k.

Определение 3.2.11. *Расширение $k \subset K$ называется конечным над k, если размерность K над полем k конечна. Величина* $\dim_k K$ *в этом случае называется степенью расширения и обозначается* $(K : k)$.

Простое алгебраическое расширение конечно и его степень равна степени соответствующего неприводимого многочлена.

Очевидно, справедлива

Теорема 16. *Если $k \subset K$ и $K \subset L$ конечные расширения, то и $k \subset L$ конечно и выполняется равенство $(L : k) = (L : K)(K : k)$.*

Теорема 17. *Каждое конечное расширение K поля k алгебраично и получается присоединением конечного числа элементов.*

Доказательство. Пусть $n = (K : k)$. Тогда для любого $\alpha \in K$ элементы $1, \alpha, \ldots, \alpha^n$ линейно зависимы над k. Следовательно, для некоторых элементов $a_0, \ldots, a_n$ поля k, среди которых имеется ненулевой элемент, выполняется соотношение $\sum_{i=0}^{n} a_i \alpha^i = 0$. Поэтому α является корнем многочлена $p(x) = \sum_{i=0}^{n} a_i x^i$. Пусть $p(x) = p_1(x) \cdot \ldots \cdot p_j(x)$ — разложение многочлена на неприводимые множители. Тогда α является корнем одного из сомножителей и, следовательно, является алгебраическим над k. В качестве образующих элементов S можно взять любой k-базис в K. □

Верно и обратное утверждение.

Теорема 18. *Каждое расширение поля k, полученное присоединением конечного множества S алгебраических элементов к полю k, конечно.*

Доказательство. Пусть $S = \{\theta_1, \ldots, \theta_m\} \subset K$. Тогда имеется последовательность конечных расширений

$$k \subset k(\theta_1) \subset k(\theta_1)(\theta_2) \ldots \subset k(\theta_1)(\theta_2) \ldots (\theta_m) = L.$$

Следовательно, согласно теореме 16 расширение $k \subset L$ конечно. □

Таким образом, элементы алгебраического расширения $k(\alpha_1, \ldots, \alpha_n)$ представляются многочленами

$$k(\alpha_1, \ldots, \alpha_n) = k[\alpha_1, \ldots, \alpha_n].$$

Следствие 11. *Сумма, разность, произведение и частное алгебраических элементов являются также алгебраическими.*

Теорема 19. *Если α алгебраичен относительно K, а $k \subset K$ — алгебраическое расширение поля k, то α алгебраичен относительно k.*

Доказательство. Пусть $S = \{\theta_1, \ldots, \theta_m\}$ — множество коэффициентов многочлена $p \in K[x]$, определяющего алгебраический элемент α над полем K. Тогда простое расширение $k(S) \subset k(S)(\alpha)$ алгебраично и, следовательно, конечно. Поскольку расширение $k \subset K$ алгебраическое, то согласно теореме 18 расширение $k \subset k(S)$ конечное. Поэтому по теореме 16 расширение $k \subset k(S)(\alpha)$ также конечно. Тогда из теоремы 17 следует, что элемент α алгебраичен над k. □

Определение 3.2.12. *Полем разложения многочлена $p \in k[x]$ называется минимальное конечное расширение поля k, в котором многочлен p разлагается на линейные множители.*

Предложение 18. *Для каждого многочлена $p \in k[x]$ существует его поле разложения L, причем $L = k(S)$, где S — множество всех корней многочлена p в L.*

Доказательство. Если многочлен имеет степень 1, то утверждение тривиально: полем разложения этого многочлена является само поле k.

Предположим, что утверждение доказано для многочленов степеней не больших m, и пусть $\deg p = m + 1$. Если многочлен имеет хотя бы один корень $\theta \in k$, то $p = (x - \theta)q$, где $\deg q = m$. Согласно предположению индукции для q существует поле разложения L и $L = k(S_q)$, где S_q — множество корней многочлена q. Тогда это же поле является полем разложения многочлена p и $L = k(S_q \cup \{\theta\})$. Если же многочлен p не имеет корней в k, то имеется его неприводимый множитель p_1 степени большей 1 и тогда определены конечное расширение $k \subset k[x]/(p_1) = K$ и элемент $\theta = \pi(x)$,

где $\pi : k[x] \to k[x]/(p_1)$ — естественный гомоморфизм на факторкольцо, причем $K = k(\theta)$. Тогда θ — корень многочлена p и, следовательно, $p = (x - \theta)q$, где $q \in K[x]$. Согласно предположению индукции для q существует поле разложения $L \supset K$ и $L = K(S_q)$, где S_q — множество всех корней многочлена q в L. Тогда расширение $k \subset L$ является полем разложения для p и выполняется равенство $L = K(S_q) = k(\theta)(S_q) = k(S)$, где $S = \{\theta\} \cup S_q$ — множество всех корней многочлена p в поле L. □

Определение 3.2.13. *Гомоморфизмом расширений $k \subset K$ и $\bar{k} \subset \bar{K}$, содержащихся в общем поле, называется пара гомоморфизмов полей $\varphi : k \to \bar{k}$ и $\psi : K \to \bar{K}$, для которых коммутативна диаграмма*

$$\begin{array}{ccc} k & \subset & K \\ \varphi\downarrow & & \downarrow\psi \\ \bar{k} & \subset & \bar{K}. \end{array}$$

Гомоморфизм ψ в этом случае называют продолжением гомоморфизма φ. Если φ и ψ — эпиморфизмы, то они являются изоморфизмами полей и пара (φ, ψ) называется эквивалентностью или изоморфизмом расширений. Изоморфизм ψ называется в этом случае продолжением изоморфизма φ.

Замечание 3.2.1. Определение 3.2.13 обобщает определение 3.2.9, данное в случае $k = \bar{k}$ и $\varphi = id_k$.

Изоморфизм полей $\varphi : k \to \bar{k}$ определяет изоморфизм колец многочленов $\varphi^* : k[x] \to \bar{k}[x]$.

Теорема 20. *Пусть $k(\theta)$ — простое расширение поля k, $p \in k[x]$ — соответствующий неприводимый многочлен и $\varphi : k \to \bar{k}$ — некоторый изоморфизм. Тогда многочлен $\bar{p} = \varphi^*(p) \in \bar{k}[x]$ неприводим и существует изоморфизм (φ, ψ) простых расширений $k \subset k(\theta)$ и $\bar{k} \subset \bar{k}(\bar{\theta})$, причем расширение $\bar{k} \subset k(\bar{\theta})$ определяется неприводимым многочленом $\bar{p}$.*

Доказательство. Изоморфизм ψ задается композицией

$$k(\theta) = k[x]/(p) \xrightarrow{\approx} \bar{k}[x]/(\bar{p}) = \bar{k}(\bar{\theta}).$$

□

Следствие 12. *Пусть $\varphi : k \to \bar{k}$ — изоморфизм, $p \in k[x]$, $\bar{p} = \varphi^*(p) \in \bar{k}[x]$, и в полях их разложения $k(\theta_1, \ldots, \theta_n)$ и $\bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_n)$ выполняются соотношения $p = (x - \theta_1) \cdot \ldots \cdot (x - \theta_n)$ и $\bar{p} = (x - \bar{\theta}_1) \cdot \ldots \cdot (x - \bar{\theta}_n)$. Тогда расширения $k \subset k(\theta_1, \ldots, \theta_n)$ и $\bar{k} \subset \bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_n)$ эквивалентны, причем множество $\{\theta_1, \ldots, \theta_n\}$ отображается взаимно однозначно на множество $\{\bar{\theta}_1, \ldots, \bar{\theta}_n\}$.*

Доказательство. Для многочленов p степеней 1 или 0 утверждение очевидно. Предположим, что утверждение доказано для всех многочленов степени, не большей m. Докажем его для многочленов степени $m + 1$.

Пусть p и $\bar{p}$ — многочлены степени $m + 1$. Тогда эти многочлены имеют неприводимые множители p_1 и $\bar{p}_1 = \varphi^*(p_1)$. Согласно теореме 20 простые расширения $k(\theta_1)$ и $\bar{k}(\bar{\theta}_1)$ эквивалентны, и пусть ψ — их некоторая эквивалентность. В этом случае в соответствующих расширениях выполняются соотношения

$$p = (x-\theta_1)\cdot \ldots\cdot(x-\theta_s)q_1\cdot \ldots\cdot q_t \text{ и } \bar{p} = (x-\bar{\theta}_1)\cdot \ldots\cdot(x-\bar{\theta}_s)\bar{q}_1\cdot \ldots\cdot \bar{q}_t,$$

где степени неприводимых многочленов q_i и $\bar{q}_i$ больше 1. Поэтому, $k(\theta_1) = k(\theta_1, \ldots, \theta_s)$ и $\bar{k}(\bar{\theta}_1) = \bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_s)$, и эквивалентность ψ устанавливает взаимно однозначное соответствие $\{\theta_1, \ldots, \theta_s\}$ и $\{\bar{\theta}_1, \ldots, \bar{\theta}_s\}$. Положим $q = q_1 \cdot \ldots \cdot q_s$ и $\bar{q} = \bar{q}_1 \cdot \ldots \cdot \bar{q}_s$. Тогда $\bar{q} = \psi^*(q)$. По предположению индукции существует эквивалентность расширений $k(\theta_1, \ldots, \theta_s) \subset k(\theta_1, \ldots, \theta_n)$ и $\bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_s) \subset \bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_n)$ с требуемыми свойствами. Соответствующий изоморфизм полей $k(\theta_1, \ldots, \theta_n) \cong \bar{k}(\bar{\theta}_1, \ldots, \bar{\theta}_n)$ задает искомую эквивалентность расширений. □

Из предложения 18 и следствия 12 при $p = \bar{p}$ и $\varphi = \mathrm{id}_k$ получаем

Следствие 13. *Поле разложения многочлена $p \in k[x]$ определено однозначно с точностью до эквивалентности.*

Из теоремы 16 и следствия 12 получаем

Следствие 14. *Пусть $k(\theta_1, \ldots, \theta_n)$ — конечное алгебраическое расширение, лежащее в некотором поле L. Тогда множество изоморфных вложений поля $k(\theta_1, \ldots, \theta_n)$ в L конечно и его мощность не превосходит величину $(k(\theta_1, \ldots, \theta_n) : k)$.*

Определение 3.2.14. *Поле K называется алгебраически замкнутым, если каждый отличный от константы многочлен $p \in K[x]$ имеет корень в K.*

Теорема 21. *Для каждого поля k существует алгебраически замкнутое алгебраическое расширение K. С точностью до эквивалентности это расширение определено однозначно.*

Доказательство этой теоремы см. в книгах [Лен68] и [Вар76].

3.2.2 Алгебраические многообразия

В данном разделе, если не оговорено противное, фиксируем расширение $K \supset k$ поля k и множество переменных $X = \{x_1, \ldots, x_n\}$. Предполагается также, что рассматриваются только идеалы в $k[X]$.

Элемент $f \in k[X]$ задает отображение вычисления

$$f : K^n \to K,$$

определяемое подстановками в $f \in K[X]$ значений переменных $\xi = (\xi_1, \ldots, \xi_n) \in K^n$. Оно обозначается $f(\xi_1, \ldots, \xi_n)$, или $f(\xi)$. Точка $\xi = (\xi_1, \ldots, \xi_n) \in K^n$ называется корнем многочлена $f \in k[X]$, если $f(\xi) = 0$. Точка $\xi = (\xi_1, \ldots, \xi_n) \in K^n$ называется корнем идеала m, если $f(\xi) = 0$ для любого $f \in$ m.

Алгебраическим многообразием в K^n называется множество всех корней некоторого идеала m в $k[X]$. Согласно теореме Гильберта о базисе для каждого идеала кольца многочленов $k[X]$ существует конечный базис. Пусть, например, m $= (f_1, \ldots, f_m)$. Легко убедиться, что множество корней идеала m совпадает с множеством решений системы алгебраических уравнений

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ \cdots\cdots\cdots \\ f_m(x_1, \ldots, x_n) = 0 \end{cases} \tag{3.4}$$

Выбрав другой базис идеала m, например, $\mathsf{m} = (g_1, \ \ldots, g_s)$, получаем эквивалентную систему алгебраических уравнений

$$\begin{cases} g_1(x_1, \ \ldots, x_n) &= 0, \\ \cdots\cdots\cdots & \\ g_s(x_1, \ \ldots, x_n) &= 0 \end{cases}, \tag{3.5}$$

множество решений которой также совпадает с алгебраическим многообразием идеала m. Таким образом, множество корней идеала m характеризует эквивалентные системы уравнений.

Может оказаться, что разные идеалы определяют одно и то же многообразие. Например, идеалы $((x-1)^2, y-1)$ и $((x-1), (y-1)^2)$ не совпадают, но определяют одно и то же многообразие в $\mathbb{C}^2$.

Определение 3.2.15. *Пусть $M \subset K^n$ — алгебраическое многообразие некоторого идеала в $k[X]$. Идеал*

$$\mathsf{m} = \{f \in k[X] \mid \forall \xi \in M \ f(\xi) = 0\}$$

называется соответствующим многообразию M или просто идеалом многообразия M.

Очевидно, что многообразие, определяемое идеалом m, соответствующим многообразию M, совпадает с самим многообразием M. Следовательно, различным многообразиям соответствуют различные идеалы.

Любой делитель n идеала m, т.е. $\mathsf{n} \supset \mathsf{m}$, определяет подмножество N в многообразии M, называемое подмногообразием многообразия M.

Теорема 22. *В любом семействе непустых многообразий существует минимальное непустое многообразие, т.е. многообразие, не содержащее ни одного собственного подмногообразия из этого семейства.*

Доказательство. Сопоставим семейству многообразий $\mathcal{M}$ семейство соответствующих идеалов $\mathcal{I}$. Согласно теореме 7 в $\mathcal{I}$ имеется максимальный идеал m, и многообразие этого идеала минимально в $\mathcal{M}$. □

Определение 3.2.16. *Будем говорить, что многочлен f содержит многообразие M, если он обращается в нуль во всех точках этого многообразия.*

Поэтому идеал m многообразия M состоит из всех многочленов, содержащих это многообразие.

Легко проверить, что, если многочлен f содержит многообразие M, то многообразие идеала (f) содержит многообразие M.

Предложение 19. *Пересечение двух многообразий M и N является многообразием.*

Доказательство. Если многообразия M и N являются корнями идеалов m и n соответственно, то множество корней объединения идеалов (m, n) является многообразием, совпадающим с $M \cap N$. □

Предложение 20. *Объединение двух многообразий M и N является многообразием.*

Доказательство. Если многообразия M и N являются корнями идеалов m и n соответственно, то множество корней пересечения идеалов $[\mathsf{m}, \mathsf{n}]$ (или произведения идеалов $\mathsf{m} \cdot \mathsf{n}$) является многообразием, совпадающим с $M \cup N$.

Действительно, если ξ принадлежит объединению $M \cup N$, то ξ — корень всех многочленов либо из m, либо из n и, следовательно, корнем всех многочленов из $[\mathsf{m}, \mathsf{n}]$ и, в частности, из $\mathsf{m} \cdot \mathsf{n}$.

Если ξ не принадлежит объединению $M \cup N$, то существуют $f \in \mathsf{m}$ и $g \in \mathsf{n}$, не обращающиеся в нуль в точке ξ. Тогда $fg \in \mathsf{m} \cdot \mathsf{n} \subset [\mathsf{m}, \mathsf{n}]$ и, следовательно, ξ не является корнем ни одного из идеалов $[\mathsf{m}, \mathsf{n}]$ и $\mathsf{m} \cdot \mathsf{n}$. □

Далее будем рассматривать только непустые многообразия идеала в $k[X]$.

Определение 3.2.17. *Многообразие, которое можно представить как объединение двух собственных (непустых) подмногообразий называется приводимым над полем k. Многообразие, не являющееся приводимым, называется неприводимым.*

Теорема 23. *Многообразие M неприводимо над k тогда и только тогда, когда идеал* m*, соответствующий M, прост.*

Доказательство. Предположим, что M приводимо, т.е. $M = M_1 \cup M_2$, $M \neq M_1$ и $M \neq M_2$. В таком случае в идеале многообразия M_1 существует многочлен f, не содержащий M, т.е. $f \notin \mathsf{m}$. Аналогично, существует многочлен g из идеала многообразия M_2, также не содержащий M, т.е. $g \notin \mathsf{m}$. Произведение fg содержит M_1 и M_2, а следовательно, и M, т.е. $fg \in \mathsf{m}$. Поэтому идеал m не прост.

Если M неприводимо и идеал m не прост, то для некоторых $f, g \notin \mathsf{m}$ произведение fg принадлежит m. Тогда $M = M_1 \cup M_2$, где M_1 — многообразие идеала (f, m), а M_2 — многообразие идеала (g, m). Согласно построению, многообразия M_1 и M_2 являются собственными в M, следовательно, многообразие M приводимо, что противоречит сделанному предположению. □

Аналогично доказывается

Теорема 24. *Если неприводимое многообразие M содержится в объединении многообразий M_1 и M_2, то M содержится либо в M_1, либо в M_2.*

Теорема 25. *Каждое многообразие над полем k представимо как объединение конечного числа неприводимых многообразий.*

Доказательство. Пусть существует многообразие M, не представимое как объединение конечного числа неприводимых многообразий. Рассмотрим семейство $\mathcal{M}$ всех таких многообразий. Согласно теореме 22 в этом семействе имеется минимальное по включению многообразие M_0. Поскольку $M_0 \in \mathcal{M}$, это многообразие приводимо: $M_0 = M_1 \cup M_2$. Тогда, в силу минимальности M_0, многообразия M_1 и M_2 представляются как конечные объединения неприводимых многообразий. Следовательно, и многообразие M_0 также представимо как конечное объединение неприводимых многообразий, что противоречит сделанному предположению. □

Определение 3.2.18. *Представление*

$$M = M_1 \cup M_2 \cup \ \ldots \cup M_s \tag{3.6}$$

называется минимальным неприводимым представлением многообразия M, если все многообразия M_i неприводимы и из этого разложения нельзя убрать ни одного слагаемого.

Согласно теореме 25 минимальное неприводимое представление всегда существует.

Теорема 26. *Минимальное неприводимое представление многообразия единственно.*

Доказательство. Пусть существует минимальное неприводимое представление многообразия M:

$$M = N_1 \cup N_2 \cup \ \ldots \cup N_r, \tag{3.7}$$

отличное от минимального неприводимого представления (3.6). Тогда в силу теоремы 24 многообразие M_1 содержится в одном из слагаемых N_i. Без ограничения общности можно считать, что $M_1 \subset N_1$. Аналогично, при некотором k выполняется включение $N_1 \subset M_k$. Тогда

$$M_1 \subset N_1 \subset M_k.$$

Если $k \neq 1$, то разложение (3.6) не является минимальным. Следовательно, $k = 1$ и $M_1 = N_1$. Аналогично получаем $M_2 = N_2, \ \ldots, M_s = N_s$ и $s = r$. □

Определение 3.2.19. *Универсальным расширением поля k будем называть поле Ω, являющееся алгебраическим замыканием поля L, полученного из k присоединением бесконечного числа алгебраически независимых над k элементов.*

Далее всюду Ω обозначает универсальное расширение поля k.

Лемма 6. *Всякое расширение $k(\alpha_1, \ \ldots, \alpha_n)$, получающееся присоединением конечного числа элементов к полю k, можно изоморфно вложить в универсальное расширение Ω.*

Доказательство. Выберем максимальное подмножество алгебраически независимых над k элементов из множества $\{\alpha_1, \ \ldots, \alpha_n\}$. Без ограничения общности будем считать, что это $\{\alpha_1, \ \ldots, \alpha_r\}$, где $r \leq n$. Пусть $\{\alpha'_1, \ \ldots, \alpha'_r\}$ — алгебраически независимые над k элементы в Ω. Тогда соответствие $\alpha_i \mapsto \alpha'_i$ определяет изоморфное вложение

$$\varphi_r : k(\alpha_1,\ \ldots, \alpha_r) \xrightarrow{\approx} k(\alpha'_1,\ \ldots, \alpha'_r) \subset \Omega.$$

Предположим, что построено изоморфное вложение

$$\varphi_m : k(\alpha_1,\ \ldots, \alpha_m) \xrightarrow{\approx} k(\alpha'_1,\ \ldots, \alpha'_m) \subset \Omega$$

для некоторого $m \geq r$. Если $r = n$, то построение завершено. В противном случае, поскольку элементы $\alpha_1,\ \ldots, \alpha_r, \alpha_{m+1}$ алгебраически зависимы, а элементы $\alpha_1,\ \ldots, \alpha_r$ алгебраически независимы, в кольце многочленов над полем $k(\alpha_1,\ \ldots, \alpha_m)$ имеется неприводимый многочлен p, корнем которого в $k(\alpha_1,\ \ldots, \alpha_n)$ является элемент α_{m+1}. Поэтому и многочлен $\varphi^*_m(p) \in k(\alpha'_1,\ \ldots, \alpha'_m)[x]$ неприводим над $k(\alpha'_1,\ \ldots, \alpha'_m)$. Обозначим через α'_{m+1} любой его корень в Ω. Тогда согласно теореме 20 существует изоморфизм расширений

$$\varphi_{m+1} : k(\alpha_1,\ \ldots, \alpha_m)(\alpha_{m+1}) \xrightarrow{\approx} k(\alpha'_1,\ \ldots, \alpha'_m)(\alpha'_{m+1}) \subset \Omega.$$

□

Определение 3.2.20. *Точка $\xi \in \Omega^n$ называется общим корнем идеала $I \subset k[x_1,\ \ldots, x_n]$, если выполняются условия*

- *из $f \in I$ следует, что $f(\xi) = 0$,*
- *из $f \in k[x_1,\ \ldots, x_n]$ и $f(\xi) = 0$ следует, что $f \in I$.*

Теорема 27. *Пусть $K \supset k$ — произвольное расширение поля k. Тогда для любого $\xi \in K^n$ множество $\mathfrak{p}_\xi$ многочленов кольца $k[x_1,\ \ldots, x_n]$, обращающихся в ноль в точке ξ, составляют отличный от $k[x_1,\ \ldots, x_n]$ простой идеал.*

Доказательство. Если $f(\xi) = g(\xi) = 0$, то и $f(\xi) - g(\xi) = 0$ и для любого $h \in k[x_1,\ \ldots, x_n]$ выполнено равенство $h(\xi)f(\xi) = 0$. Следовательно, $\mathfrak{p}_\xi$ — идеал.

Если теперь $f(\xi)g(\xi) = 0$ и $g(\xi) \neq 0$, то $f(\xi) = 0$, т.к. поле не имеет делителей нуля. Следовательно, указанный идеал $\mathfrak{p}_\xi$ прост. Поскольку $1 \notin p_\xi$, то $p_\xi \neq k[x_1,\ \ldots, x_n]$. □

В дальнейшем идеал $\mathfrak{p}_\xi$ из теоремы 27 будем называть идеалом, соответствующим точке ξ.

Пример. Пусть

$$\xi_i = \alpha_i + \beta_i t, \quad i = 1, \ \ldots, n, \tag{3.8}$$

где $\alpha_i, \beta_i \in k$ и $t \in \Omega$ трансцендентен над k. Согласно теореме 27, идеал $\mathfrak{p}_\xi$, состоящий из всех многочленов в $k[X]$, обращающихся в ноль в точке $\xi \in \Omega^n$, прост, и точка $\xi = (\xi, \ \ldots, \xi_n)$ является общим корнем этого идеала. Точку ξ в этом случае можно интерпретировать как параметрическое уравнение прямой (3.8), а идеал $\mathfrak{p}_\xi$ — как множество многочленов, обращающихся в ноль на этой прямой.

Теорема 28. *Если $\mathfrak{p}_\xi$ — идеал из теоремы 27, соответствующий точке $\xi \in \Omega^n$, то поле $\Lambda = k(\xi_1, \ \ldots, \xi_n)$ изоморфно полю частных Π факторкольца $k[x_1, \ \ldots, x_n]/\mathfrak{p}_\xi$. Изоморфизм определяется соответствиями $x_i \mapsto \xi_i$.*

Доказательство. Рассмотрим подкольцо $k[\xi_1, \ \ldots, \xi_n] \subset \Lambda$. Тогда поле Λ является полем частных рассматриваемого кольца. Соответствия $f(\xi) \mapsto f(X)$ задают изоморфизм

$$k[\xi_1, \ \ldots, \xi_n] \to k[x_1, \ \ldots, x_n]/\mathfrak{p}_\xi.$$

Действительно, из $f(\xi) - g(\xi) = 0$ следует, что $f - g \in \mathfrak{p}_\xi$, т.е. $f = g$ в кольце $k[x_1, \ \ldots, x_n]/\mathfrak{p}_\xi$.

Потому изоморфны и соответствующие поля частных Π и Λ. □

Определение 3.2.21. *Точки $\xi, \xi' \in \Omega^n$ называются сопряженными, если соответствие $\xi_i \mapsto \xi_i'$ определяет изоморфизм колец $k[\xi] \to k[\xi']$.*

Теоремы 27 и 28 означают, что сопряженные точки являются общими корнями однозначно определенного простого идеала и все общие корни простого идеала сопряжены.

Теорема 29. *Каждый простой идеал $\mathfrak{p}$ кольца $k[x_1, \ \ldots, x_n]$, отличный от самого кольца, имеет общий корень $\xi \in \Omega^n$.*

Доказательство. Поскольку факторкольцо кольца многочленов по простому идеалу не содержит делителей нуля, определено поле частных Λ кольца $k[x_1,\ \ldots, x_n]/\mathfrak{p}$, являющееся расширением поля k, и, очевидно, выполняется равенство $\Lambda = k(\alpha_1,\ \ldots, \alpha_n)$, где $\alpha_i \equiv x_i \pmod{\mathfrak{p}}$. Согласно лемме 6 расширение $k(\alpha_1,\ \ldots, \alpha_n)$ вкладывается в универсальное расширение Ω. Обозначим образы элементов α_i через ξ_i. Тогда, ввиду изоморфизма $k(\alpha_1,\ \ldots, \alpha_n) \approx k(\xi_1,\ \ldots, \xi_n)$, выполняется равенство

$$\mathfrak{p} = \mathfrak{p}_\alpha = \mathfrak{p}_\xi$$

и, следовательно, $\xi = (\xi_1,\ \ldots, \xi_n) \in \Omega^n$ — общий корень простого идеала $\mathfrak{p}$. □

Согласно теореме 29 каждый простой идеал $\mathfrak{p} \neq k[x_1,\ \ldots, x_n]$ имеет общий корень ξ в универсальном поле Ω, а по теореме 28 этот корень определен однозначно с точностью до сопряженения. Точка ξ является корнем идеала $\mathfrak{p}$ и, следовательно, принадлежит многообразию $M \subset \Omega^n$ этого идеала. Идеал, соответствующий M, совпадает с $\mathfrak{p}$, поскольку для $f \in k[x_1,\ \ldots, x_n]$, обращающегося в ноль во всех точках многообразия M, в частности, $f(\xi) = 0$, т.е. $f \in \mathfrak{p}$. Так как идеал $\mathfrak{p}$ прост, то по теореме 23 многообразие M неприводимо. Следовательно, справедлива

Теорема 30. *Каждый простой идеал $\mathfrak{p} \neq k[x_1,\ \ldots, x_n]$ соответствует некоторому неприводимому многообразию корней и служит идеалом этого многообразия.*

Определение 3.2.22. *Общей точкой многообразия M называется любой общий корень идеала, соответствующего этому многообразию.*

Согласно теореме 23 неприводимому многообразию M соответствует простой идеал $\mathfrak{p}$. В этом случае общий корень идеала $\mathfrak{p}$ является общей точкой многообразия M над полем k. Если общие корни идеала сопряжены, то соответствующие общие точки многообразия этого идеала называются сопряженными. Следовательно, справедлива

Теорема 31. *Многообразие M обладает общей точкой над полем k тогда и только тогда, когда оно неприводимо.*

Определение 3.2.23. *Пусть ξ — общая точка над k некоторого неприводимого многообразия M. Степень трансцендентности поля $k(\xi)$ будем называть размерностью над полем k (неприводимого) многообразия M, а также размерностью соответствующего простого идеала* p *и обозначать* $\dim_k(\mathfrak{p}) = \dim_k(M)$. *Полагаем, по определению, размерность единичного идеала равной -1. Соответственно, размерность пустого многообразия полагается также равной -1.*

Очевидно, что размерность простого идеала $\mathfrak{p} \neq k[x_1, \ldots, x_n]$ может принимать значения от 0 до n. В силу теоремы 28 размерность неприводимого многообразия M не зависит от выбора общей точки ξ.

Определение 3.2.24. *Степенью трансцендентности элемента $\xi \in \Omega^n$ называется степень трансцендентности расширения $k \subset k(\xi)$.*

Пусть $(\xi_1, \ldots, \xi_n)$ общий корень простого идеала p, а $(\xi'_1, \ldots, \xi'_n)$ — произвольный корень этого же идеала. Тогда соответствие $\xi_i \mapsto \xi'_i$ при всех $i = 1, \ldots, n$ задает эпиморфизм колец

$$\varphi : k[\xi_1, \ldots, \xi_n] \to k[\xi'_1, \ldots, \xi'_n]. \tag{3.9}$$

Этот эпиморфизм является изоморфизмом тогда и только тогда, когда $(\xi'_1, \ldots, \xi'_n)$ — также общий корень идеала p.

Если размерность идеала p равна нулю, то кольцо $k[\xi_1, \ldots, \xi_n]$ является полем, а тогда эпиморфизм φ — изоморфизмом. Следовательно, все корни простого идеала нулевой размерности являются общими и эквивалентными друг другу. Согласно следствию 14 мощность множества таких эквивалентностей не превосходит величины $(k(\xi) : k)$, т.е. многообразие простого идеала размерности ноль конечно. Итак, доказана

Теорема 32. *Пусть K — алгебраическое замыкание поля k. Тогда многообразие простого идеала* $\mathfrak{p} \subset k[x]$ *размерности ноль является конечным подмножеством в K^n, а все нули этого идеала общие и, следовательно, являются сопряженными.*

В частности, если поле k алгебраически замкнуто, то простой идеал размерности 0 имеет вид $(x_1 - \xi_1, \ldots, x_n - \xi_n)$.

Теорема 33. *Различные корни простого идеала размерности r имеют степени трансцендентности не более r. Если степень трансцендентности такого корня равна r, то этот корень общий.*

Доказательство. Пусть $(\xi_1, \ldots, \xi_n)$ общий корень простого идеала p, а $(\xi'_1, \ldots, \xi'_n)$ — корень степени трансцендентности s и $\xi'_1, \ldots, \xi'_s$ алгебраически независимы. Поскольку эпиморфизм (3.9) преобразует ξ_i в ξ'_i, элементы $\xi_1, \ldots, \xi_s$ также алгебраически независимы. Поэтому $r \geq s$.

Если же $r = s$, то расширение

$$k(\xi_1, \ldots, \xi_s) \subset k(\xi_1, \ldots, \xi_n) = k(\xi_1, \ldots, \xi_s)[\xi_{s+1}, \ldots, \xi_n] \tag{3.10}$$

конечно. Для доказательства последнего утверждения теоремы достаточно проверить мономорфность φ из формулы (3.9). Пусть это не так. Тогда существует такой ненулевой $f \in k(\xi_1, \ldots, \xi_n)$, что $\varphi(f) = 0$. Ввиду конечности расширения (3.10), в поле $k(\xi_1, \ldots, \xi_n)$ выполняется равенство

$$\frac{1}{f(\xi_1, \ldots, \xi_n)} = \frac{g(\xi_1, \ldots, \xi_n)}{h(\xi_1, \ldots, \xi_s)}.$$

Следовательно,

$$h(\xi_1, \ldots, \xi_s) = f(\xi_1, \ldots, \xi_n) g(\xi_1, \ldots, \xi_n).$$

Поэтому $h(\xi'_1, \ldots, \xi'_s) = 0$, что противоречит алгебраической независимости элементов $\xi'_1, \ldots, \xi'_s$. □

Следствие 15. *Пусть простой идеал* $\mathfrak{p}'$ *делит простой идеал* $\mathfrak{p}$. *Тогда* $\dim \mathfrak{p}' \leq \dim \mathfrak{p}$. *Если* $\dim \mathfrak{p} = \dim \mathfrak{p}'$, *то* $\mathfrak{p} = \mathfrak{p}'$.

Доказательство. Первое утверждение следствия очевидно. Докажем второе.

Пусть ξ' — общий корень идеала $\mathfrak{p}'$. Тогда степень его трансцендентности равна $\dim \mathfrak{p} = \dim \mathfrak{p}'$. Тогда ξ' — также является корнем идеала $\mathfrak{p}$, степень трансцендентности которого равна $\dim \mathfrak{p}$, и, следовательно, по теореме 33 этот корень общий для идеала $\mathfrak{p}$. Поскольку идеалы $\mathfrak{p}$ и $\mathfrak{p}'$ соответствуют одной и той же общей точке, то эти идеалы совпадают. □

Согласно теоремам 25 и 26 каждое многообразие M имеет единственное минимальное представление в виде объединения неприводимых многообразий $M_1, \ldots, M_s$.

Определение 3.2.25. *Размерностью многообразия называется наибольшая из размерностей его неприводимых компонент.*

С другой стороны, по теореме 13 каждый идеал допускает несократимое представление в виде пересечения примарных идеалов, ассоциированных с попарно различными простыми идеалами. Согласно теореме 14 набор ассоциированных простых идеалов определен однозначно.

Определение 3.2.26. *Размерностью идеала называется наибольшая из размерностей ассоциированных с ним простых идеалов.*

Лемма 7. *Пусть k — поле и $I \subsetneq k[x_1, \ldots, x_n]$ — собственный идеал кольца $k[x_1, \ldots, x_n]$ и $K \supset k$ — алгебраическое замыкание поля k. Тогда существует гомоморфизм колец*

$$\varphi : k[x_1, \ldots, x_n]/I \to K,$$

продолжающий вложение $k \subset K$.

Доказательство см. в [Лен68, гл.X] .

Теорема 34. *Пусть K — алгебраическое замыкание поля k. Тогда непустое многообразие над полем k в универсальном пространстве Ω всегда имеет непустое пересечение с пространством K^n.*

Доказательство. Поскольку каждое многообразие представимо как объединение неприводимых многообразий, то достаточно доказать теорему для неприводимых многообразий.

Для неприводимых многообразий размерности ноль доказываемое утверждение вытекает из теоремы 32.

Для многообразий размерности n утверждение теоремы тривиально, поскольку в этом случае многообразие совпадает с пространством Ω^n и выполняется включение $\Omega^n \supset K^n$.

Для остальных многообразий утверждение теоремы следует из леммы 7. Действительно, если I — идеал, определяющий неприводимое многообразие M, и

$$\varphi : k[x_1, \ldots, x_n]/I \to K$$

гомоморфизм из леммы 7, то элементы $\xi_i = \varphi(x_i)$ определяют точку многообразия M, принадлежащую пространству K^n. Отметим, что в этом случае такие точки не являются общими точками многообразия M. □

Следующее утверждение представляет критерий существования решений систем алгебраических уравнений.

Теорема 35. *Пусть идеал I порожден многочленами системы уравнений*

$$\begin{cases} f_1(x_1, \ldots, x_n) &= 0, \\ \cdots\cdots\cdots & \\ f_m(x_1, \ldots, x_n) &= 0, \end{cases}$$

над полем k. Эта система уравнений имеет решение в алгебраическом замыкании K поля k тогда и только тогда, когда $1 \notin I$.

Доказательство. Пусть $1 \in I$, то $1 = r_1 f_1 + \ldots + r_m f_m$. Поэтому, если решение существует, то $1 = 0$.

Пусть $1 \notin I$. Тогда существует максимальный идеал J кольца $k[x_1, \ldots, x_n]$, содержащий I и не содержащий единицу. В этом случае по лемме 7 существует гомоморфизм ψ факторкольца $k_0 = k[x_1, \ldots, x_n]/J$ в алгебраическое замыкание K. Пусть

$$\varphi : k[x_1, \ldots, x_n] \to k_0$$

соответствующий естественный гомоморфизм. Тогда

$$(\psi(\varphi(x_1)), \ldots, \psi(\varphi(x_n))) \in K^n$$

решение исходной системы уравнений. □

Следующая теорема называется теоремой Гильберта о нулях. Пусть K — алгебраическое замыкание поля k.

Теорема 36. *Пусть многочлен $f(x_1,\ \ldots, x_n) \in k[x_1,\ \ldots, x_n]$ обращается в нуль во всех корнях $(c_1,\ \ldots, c_n) \in K^n$ системы уравнений*

$$\begin{cases} f_1(x_1,\ \ldots, x_n) &= 0, \\ \cdots\cdots\cdots & \\ f_s(x_1,\ \ldots, x_n) &= 0. \end{cases}$$

Тогда при некотором $m > 0$ многочлен $f^m(x_1,\ \ldots, x_n)$ принадлежит идеалу I, порожденному многочленами этой системы.

Доказательство. Доказательство использует прием, предложенный Рабинович [Rab30].

Введем дополнительную переменную z. Система уравнений

$$\begin{cases} f_1(x_1,\ \ldots, x_n) &= 0, \\ \cdots\cdots\cdots & \ldots\ \ \ldots \\ f_m(x_1,\ \ldots, x_n) &= 0, \\ zf(x_1,\ \ldots, x_n) &= 1 \end{cases}$$

не имеет решений. Следовательно, согласно теореме 35 о существовании решений выполняется соотношение

$$r_1(x_1,\ \ldots, x_n, z)f_1(x_1,\ \ldots, x_n) + \cdots + r_n(x_1,\ \ldots, x_n, z)f_n(x_1,\ \ldots, x_n) + r_{n+1}(x_1,\ \ldots, x_n, z)(1 - zf(x_1,\ \ldots, x_n)) = 1.$$

Подставляя в это соотношение $z = \dfrac{1}{f(x_1,\ \ldots, x_n)}$ и домножая на подходящую степень знаменателя $f(x_1,\ \ldots, x_n)$, получим, что при некотором $m > 0$ многочлен $f^m(x_1,\ \ldots, x_n)$ лежит в идеале I. □

3.3 Теорема Херманн

Следующая теорема доказана в работе [Her25].

Теорема 37. *Пусть K — поле. Если система уравнений*

$$\begin{cases} \alpha_{1,1}f_1 + \ \ldots + \alpha_{1,k}f_k = b_1, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \alpha_{m,1}f_1 + \ \ldots + \alpha_{m,k}f_k = b_m \end{cases} \tag{3.11}$$

над кольцом $K[x_1, \ldots, x_n]$ относительно неизвестных $f_1, \ldots, f_k$ имеет решение в этом кольце, а коэффициенты этой системы удовлетворяют условиям

$$\deg \alpha_{ij} < d \quad \text{и} \quad \deg b_i < B,$$

то существует решение этой системы степени не выше $B + (kd)^{2^n}$.

Доказательство. Очевидно, что из существования требуемого в теореме решения в кольце многочленов над полем, являющимся расширением поля K, следует существование такого решения в кольце $K[x_1, \ldots, x_n]$. Поэтому достаточно доказать теорему в предположении, что поле K алгебраически замкнуто.

Пусть ранг матрицы системы равен $r \leq \min\{m, k\}$. Без ограничения общности можно считать, что

$$\Delta = \begin{vmatrix} \alpha_{1,1} & \ldots & \alpha_{1,r} \\ \cdots & \cdots & \cdots \\ \alpha_{r,1} & \ldots & \alpha_{r,r} \end{vmatrix} \neq 0.$$

По условию теоремы $\deg \alpha_{ij} < d$, поэтому $\deg \Delta < rd$.

Поскольку кольцо многочленов над полем не имеет делителей нуля, последние $m - r$ уравнений можно отбросить, т.к. каждое из них является линейной комбинацией первых r уравнений с коэффициентами в поле рациональных функций (поле частных кольца многочленов) и, следовательно, выполняется, если выполняются первые r соотношений. Далее произведем такую обратимую замену переменных $x_i = t_i + \beta_i t_n$, $i = 1, \ldots, n-1$, $x_n = \beta_n t_n$, $\beta_i \in K$, при которой старший член определителя Δ относительно переменной t_n равен $t_n^{\deg \Delta}$, т.е. $\deg \Delta = \deg_{t_n} \Delta$.

Для доказательства существования такой замены рассмотрим ненулевой многочлен

$$p(x_1, \ldots, x_n) = \sum_{\omega \in \mathbb{Z}_+^n,\ |\omega| = \deg \Delta} c_\omega x^\omega,$$

представляющий члены старшей степени многочлена Δ. Тогда коэффициент при $t_n^{\deg \Delta}$ равен значению этого многочлена $p(\beta_1, \ldots, \beta_n) \in K$. Поэтому существование требуемой замены следует из разрешимости уравнения $p(\beta_1, \ldots, \beta_n) = 1$ в алгебраически замкнутом поле K.

Заметим, что степени элементов кольца многочленов при таком преобразовании не меняются. Без ограничения общности оставим имена переменных неизменными.

Сопоставим данной системе уравнений однородную систему (напомним, что ввиду сказанного выше, система содержит только r уравнений):

$$\begin{cases} \alpha_{1,1}f_1 + \ \ldots + \alpha_{1,k}f_k = 0, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \alpha_{r,1}f_1 + \ \ldots + \alpha_{r,k}f_k = 0. \end{cases} \tag{3.12}$$

Перепишем эту однородную систему в виде неоднородной системы уравнений (т.к. $r \leq k$)

$$\begin{cases} \alpha_{1,1}f_1 + \ \ldots + \alpha_{1,r}f_r = -\alpha_{1,r+1}f_{r+1} - \ \ldots - \alpha_{1,k}f_k, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \alpha_{r,1}f_1 + \ \ldots + \alpha_{r,r}f_r = -\alpha_{r,r+1}f_{r+1} - \ \ldots - \alpha_{r,k}f_k. \end{cases} \tag{3.13}$$

Поскольку $\Delta \neq 0$, выразим $f_1, \ \ldots, f_r$ через $f_{r+1}, \ \ldots, f_k$ по формулам Крамера

$$\begin{cases} \Delta f_1 = \Delta_1, \\ \ldots\ldots\ldots \\ \Delta f_r = \Delta_r, \end{cases}$$

где Δ_i определитель матрицы, полученной из матрицы системы уравнений (3.13) заменой i-го столбца на столбец из правой части этой системы. Применяя разложение определителей по i-му столбцу и группируя слагаемые с множителями f_i при $i = r+1, \ \ldots, k$, получим эквивалентную систему

$$\begin{cases} \Delta f_1 = A_{1,r+1}f_{r+1} + \ \ldots + A_{1,k}f_k, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \Delta f_r = A_{r,r+1}f_{r+1} + \ \ldots + A_{r,k}f_k, \end{cases} \tag{3.14}$$

где $A_{i,j} \in K[x_1, \ \ldots, x_n]$. Тогда из условия теоремы и формул вычисления определителя следует, что $\deg A_{i,j} < rd$. Легко проверить, что система уравнений (3.14) имеет следующие решения:

$$l_j = (A_{1,j}, \ \ldots, A_{r,j}, \overbrace{\underbrace{0, \ \ldots, 0}_{j-r-1}, \Delta, 0, \ \ldots, 0}^{k-r}), \ j = r+1, \ \ldots, k.$$

Согласно сделанным выше замечаниям о степенях многочленов Δ и $A_{i,j}$, все компоненты векторов l_j имеют степень, меньшую чем rd.

Поскольку старший член определителя Δ по переменной x_n равен $x_n^{\deg \Delta}$, можно подобрать такие многочлены $c_i \in K[x_1, \ldots, x_n]$ при $i = 1, \ldots, r$, для которых выполняются соотношения

$$\begin{gathered} \deg(\Delta c_i) \leq \deg b_i < B, \\ \deg_{x_n}(b_i - \Delta c_i) < \deg_{x_n} \Delta = \deg \Delta < rd \leq kd, \\ \deg(b_i - \Delta c_i) \leq \deg b_i. \end{gathered} \tag{3.15}$$

Согласно правилу Крамера система уравнений

$$\begin{cases} \alpha_{1,1} f_1 + \ldots + \alpha_{1,k} f_k = \Delta c_1, \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ \alpha_{r,1} f_1 + \ldots + \alpha_{r,k} f_k = \Delta c_r \end{cases}$$

имеет частное решение $m = (m_1, \ldots, m_r, 0, \ldots, 0)$, где m_i — определитель матрицы, полученной из матрицы системы заменой i-го столбца на столбец, состоящий из многочленов $c_1, \ldots, c_r$. Из определения такого решения следует, что его степень меньше, чем $B + rd \leq B + kd$.

Рассмотрим теперь систему

$$\begin{cases} \alpha_{1,1} f_1 + \ldots + \alpha_{1,k} f_k = b_1 - \Delta c_1, \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ \alpha_{r,1} f_1 + \ldots + \alpha_{r,k} f_k = b_r - \Delta c_r. \end{cases} \tag{3.16}$$

Эта система имеет решение, поскольку существует решение исходной системы уравнений (3.11). Вычитая из любого решения системы (3.16) произвольные линейные комбинации над $K[x_1, \ldots, x_n]$ решений однородной системы (3.12) (в частности, решения l_j при $j = r+1, \ldots, k$), снова получаем решение системы (3.16). Следовательно, если имеется некоторое решение исходной системы уравнений, то существует также такое решение $f_1, \ldots, f_k$ системы (3.16), для которого выполнены соотношения

$$\deg_{x_n} f_i < \deg \Delta \leq rd \leq kd \text{ при } i = r+1, \ldots, k. \tag{3.17}$$

Заметим теперь, что в силу неравенств (3.15) при $i = 1, \ldots, r$ выполняются соотношения $\deg_{x_n}(b_i - \Delta c_i) < kd$.

Перепишем теперь систему (3.16) в следующем виде:

$$\begin{cases} \alpha_{1,1}f_1 + \ \ldots + \alpha_{1,r}f_r = -\alpha_{1,r+1}f_{r+1} - \ \ldots - \alpha_{1,k}f_k + b_1 - \Delta c_1, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \alpha_{r,1}f_1 + \ \ldots + \alpha_{r,r}f_r = -\alpha_{r,r+1}f_{r+1} - \ \ldots - \alpha_{r,k}f_k + b_r - \Delta c_r. \end{cases} \tag{3.18}$$

Применяя правило Крамера, получаем эквивалентное представление системы (3.18):

$$\begin{cases} \Delta f_1 = \Delta_1, \\ \ldots\ \ldots \\ \Delta f_r = \Delta_r, \end{cases}$$

где Δ_i — определитель матрицы, полученной из матрицы системы (3.18) заменой i-го столбца на столбец из правой части этой системы. Тогда

$$\deg_{x_n} f_i = \deg_{x_n} \Delta_i - \deg_{x_n} \Delta, \ \ i = 1, \ \ldots, r.$$

Разлагая определители Δ_i по столбцам и группируя слагаемые по многочленам f_j при $j = r+1, \ \ldots, k$, а также учитывая неравенства (3.17) и (3.15), получаем соотношения

$$\deg_{x_n} \Delta_i < rd + \deg_{x_n} \Delta \leq kd + \deg_{x_n} \Delta.$$

Следовательно,

$$\deg_{x_n} f_i < rd \leq kd, \quad \text{при } i = 1, \ \ldots, r.$$

Представим теперь f_i при $i = 1, \ \ldots, k$ и $b_j - \Delta c_j$ при $j = 1, \ \ldots, r$ как многочлены от x_n степени, меньшей kd, с коэффициентами в кольце $K[x_1, \ \ldots, x_{n-1}]$ и подставим эти разложения в систему (3.16). Приравнивая коэффициенты при соответствующих степенях переменной x_n от 0 до $kd + d - 1$, получим новую систему не более чем из $(kd+d)r$ уравнений относительно $s = k^2 d$ неизвестных $g_i \in K[x_1, \ \ldots, x_{n-1}]$, являющихся коэффициентами разложения многочленов $f_1, \ \ldots, f_k$ по степеням x_n:

$$\begin{cases} \beta_{1,1}g_1 + \beta_{1,s}g_s = b'_1, \\ \ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \beta_{t,1}g_1 + \beta_{t,s}g_s = b'_t \end{cases}$$

над кольцом $K[x_1, \ldots, x_{n-1}]$. В силу определения элементов $\beta_{i,j}$ и b'_j выполняются соотношения $\deg \beta_{ij} < d$ и $\deg b'_i < B$.

Обозначим минимальную степень решения такой системы через $m(k, n, d, B)$, где k, n, d, B — параметры, описывающие исходную систему уравнений (3.11). В силу доказанной редукции исходной системы относительно переменных $x_1, \ldots, x_n$ к системе (3.16) и последующей редукции этой системы к системе уравнений относительно переменных $x_1, \ldots, x_{n-1}$ выполняется неравенство

$$m(k, n, d, B) \leq \max\{kd + m(k^2 d, n-1, d, B), kd + B\}.$$

Поэтому

$$m(k, n, d, B) \leq kd + k^2 d^2 + k^4 d^4 + \cdots + (kd)^{2^{n-1}} + B \leq B + (kd)^{2^n}.$$

□

Глава 4

Базис Гребнера

4.1 Определение базиса Гребнера

Пусть $X = \{x_1, \ldots, x_n\}$ — конечное множество. Фиксируем некоторый допустимый порядок $\prec$ на множестве термов $T\langle X\rangle$. Будем использовать обозначения раздела 1.3.

Сначала сформулируем определение базиса Гребнера для полей.

Определение 4.1.1. *Пусть A — поле. Конечное множество G называется базисом Гребнера идеала I кольца $A[X]$, если*

- *G — базис идеала I,*
- $\forall f \in I\ \exists g \in G \mid \mathsf{HT}(f) \vdots \mathsf{HT}(g)$.

Поскольку A — поле, старший моном любого элемента идеала делится на старший моном некоторого элемента базиса Гребнера этого идеала. Отметим также, что базисы Гребнера идеала для различных допустимых порядков на множестве термов $T\langle X\rangle$, вообще говоря, различны.

Обобщением этого определения для нетеровых колец A является следующее

Определение 4.1.2. *Пусть A — нетерово кольцо. Конечное множество G называется базисом Гребнера идеала I кольца $A[X]$, если*

- *G — базис идеала I,*
- *идеал, порожденный старшими мономами элементов идеала I, и идеал, порожденный старшими мономами множества G, совпадают.*

Очевидно, выполняется следующее

Предложение 21. *Если A — поле, то определения 4.1.1 и 4.1.2 базиса Гребнера G эквивалентны.*

Установим существование базиса Гребнера с помощью теоремы Гильберта о базисе.

Теорема 38. *Пусть A — нетерово кольцо, X — конечное множество. Для произвольного идеала I кольца многочленов $A[X]$ и для любого допустимого порядка на множестве термов $T\langle X\rangle$ существует (конечный) базис Гребнера.*

Доказательство. Фиксируем допустимый порядок на множестве термов. Пусть I — идеал кольца $A[X]$, а J — мономиальный идеал, порожденный старшими мономами многочленов идеала I. Тогда, по следствию 7 из теоремы Гильберта, этот идеал J имеет состоящий из мономов конечный базис H_1.

Рассмотрим множество термов НТ(H_1) (см. раздел 1.3). Для каждого $t \in$ НТ(H_1) определим идеал $I_t \subset A$, базис которого состоит из коэффициентов мономов $m \in H_1$ с термами t. Поскольку A — нетерово кольцо, в каждом идеале I_t можно выбрать конечный базис, т.е. $I_t = (\lambda_{1,t},\ \ldots, \lambda_{k_t,t})$ для некоторого $k_t \in \mathbb{N}$.

Поскольку $\lambda_{i,t}t \in J$, для каждой пары i, t, удовлетворяющей условиям $1 \le i \le k_t$ и $t \in$ НТ(H_1), существует многочлен $g_{i,t} \in I$ со старшим мономом $\lambda_{i,t}t$. Зафиксируем для каждой такой пары (i, t) такой многочлен $g_{i,t}$ и составим из всех таких многочленов множество G. (Заметим, что это множество, как и элементы $g_{i,t}$, определено, вообще говоря, неоднозначно.)

Из построения следует, что конечный набор элементов НМ$(g_{i,t}) = \lambda_{i,t}t$, по всем допустимым парам i, t, составляет базис идеала J. Следовательно, G — базис Гребнера идеала I. □

В работе [Dub90] получена оценка сверху максимальной степени базиса Гребнера идеала в кольце многочленов для любого допустимого порядка.

Теорема 39. *Пусть $I = (f_1, \ldots, f_m)$ — идеал кольца многочленов $K[X]$, где K — поле, и для некоторого d выполняются неравенства* deg $f_i \leq d$. *Тогда элементы приведенного базиса Гребнера идеала I имеют степень не более чем $2\left(\frac{d^2}{2} + d\right)^{2^{n-2}}$.*

4.2 Редукция относительно множества

Пусть $X = \{x_1, \ldots, x_n\}$ — множество переменных, A — коммутативное кольцо главных идеалов, $A[X] = A[x_1, \ldots, x_n]$ — кольцо многочленов. Выберем допустимый порядок $\prec$ на множестве термов $T\langle X\rangle$.

Пусть задан многочлен f. Рассмотрим его разложение f по ненулевым мономам (см. определение 10.1.11 и формулу (10.5))

$$f = \sum_{t \in T_f} \lambda_t t, \tag{4.1}$$

где

$$T_f = T_f\langle X\rangle = \{t \in T\langle X\rangle \mid \lambda_t \neq 0\},\ \lambda_t \in A.$$

Определение 4.2.1. *Пусть $f, h \in A[X]$ и G — конечное подмножество в $A[X]$. Будем говорить, что существует простая редукция f к h по модулю G, если существует $t \in T_f\langle X\rangle$ и для всех $g \in G$ существуют $\mu_g \in A[X]$, для которых выполняются равенства*

$$\begin{aligned} \lambda_t t &= \sum_{g \in G} \mathsf{НМ}(\mu_g g), \\ t &= \mathsf{НТ}(\mu_g g), \qquad \text{если} \quad \mathsf{НМ}(\mu_g) \neq 0, \\ f &= \sum_{g \in G} \mu_g g + h. \end{aligned} \tag{4.2}$$

Если также $\mathsf{НТ}(f) = t$, *то такую простую редукцию будем называть простой монотонной редукцией.*

Формулу (4.2) будем называть формулой простой редукции (простой монотонной редукции) и записывать как $f \to_G h$. Если существует последовательность элементов $h_1, \ldots, h_m \in A[X]$ и последовательность простых редукций (простых монотонных редукций)

$$f \to_G h_1, h_1 \to_G h_2, \ldots, h_{m-1} \to_G h_m,$$

то будем говорить, что f редуцируется (монотонно редуцируется) к форме $h = h_m$ по модулю множества G, и записывать это в виде формулы $f \to_{G^*} h$.

Пример. Пусть $G = \{x_1^2 + x_1, (t+1)x_1x_2x_3 - 1, 2x_1x_2x_3 - 1\}$ — подмножество в множестве многочленов над кольцом $\mathbb{Z}[t]$ от переменных $\{x_1, x_2, x_3\}$, $f = x_1^2 + (t+3)x_1x_2x_3 - 2$ и на множестве мономов задан лексикографический порядок. Тогда имеется редукция

$$f \to_G x_1^2, \tag{4.3}$$

не являющаяся монотонной, поскольку НТ$(f) = x_1^2$.

Если $G = \{(t+1)x_1x_2x_3 - 1, 2x_1x_2x_3 - 1\}$, то для многочлена f имеется редукция (4.3), но не существует монотонной редукции.

Приведенные далее утверждения, если не оговорено противное, справедливы как для редукций, так и для монотонных редукций. Все отличия будут специально оговорены. Определения для редукций также переносятся на монотонные редукции. Поэтому все определения и утверждения, справедливые для редукций обоих типов, формулируются только для редукций.

Определим тривиальную редукцию $f \to_{G^*} f$, заданную пустой последовательностью простых редукций. Отметим, что непустой последовательности простых редукций, приводящей к тривиальной редукции $f \to_{G^*} f$ не существует.

Определение 4.2.2. *Пусть $f \to_{G^*} h$ и из $h \to_{G^*} h_1$ следует, что $h = h_1$. В этом случае будем говорить, что редукция $f \to_{G^*} h$ является полной, f приводится к* нормальной форме *h по модулю множества G, и записывать формулой $f \to_{G^*} \underline{h}$.*

Следующая лемма вытекает непосредственно из определений базиса Гребнера идеала и монотонной редукции относительно множества.

Лемма 8. *Конечное множество $G \subset I$ является базисом Гребнера идеала I тогда и только тогда, когда для любого $f \in I$ выполняется $f \to_G \underline{0}$.*

Отметим, что нормальная форма относительно монотонной редукции может не быть нормальной формой в смысле общих редукций, поскольку может допускать дополнительные (немонотонные) редукции.

Пусть $G = \{g_1, \ \ldots, g_m\}$.

Лемма 9. *Если $f \to_{G^*} h$ — монотонная редукция, то существуют такие $f_{i,j} \in A[X]$, что*

- $f = \sum\limits_{j=1}^{k} \sum\limits_{i=1}^{m} f_{i,j} g_i + h$;
- $\mathsf{HT}\left(\sum\limits_{i=1}^{m} f_{i,j} g_i\right) = \mathsf{HT}(f_{k,j} g_k)$ *для всех таких пар k, j, что $f_{k,j} \neq 0$;*
- $\mathsf{HT}\left(\sum\limits_{i=1}^{m} f_{i,1} g_i\right) \succ \mathsf{HT}\left(\sum\limits_{i=1}^{m} f_{i,2} g_i\right) \succ \ \ldots \succ \mathsf{HT}\left(\sum\limits_{i=1}^{m} f_{i,k} g_i\right) \succ \mathsf{HT}(h)$.

Доказательство. Следует из определения монотонной редукции, как последовательности простых монотонных редукций. Число k равно количеству простых редукций, а каждая сумма вида $\sum\limits_{i=1}^{m} f_{i,j} g_i$ соответствует простой монотонной редукции. □

Определим предупорядочение[1] $\prec$ на множестве многочленов, согласованное с допустимым порядком $\prec$ на множестве термов.

Пусть имеется два многочлена f и g. Рассмотрим их разложения по ненулевым мономам

$$f = \sum_{t \in T_f} \lambda_t t \quad \text{и} \quad g = \sum_{s \in T_g} \mu_s s.$$

[1] В определении 1.1.4 предпорядка для произвольного множества нужно исключить пп.2,3.

Перепишем эти разложения по убывающим термам (см. формулу (1.12)):

$$f = \sum_{k=1}^{m(f)} \alpha_k t_k \quad \text{и} \quad g = \sum_{k=1}^{n(g)} \beta_k s_k,$$

где

$$t_1 \succ t_2 \succ \ \ldots \succ t_{m(f)} \quad \text{и} \quad s_1 \succ s_2 \succ \ \ldots \succ s_{n(g)}.$$

Будем считать что $f \prec g$ тогда и только тогда, когда $f = 0$ и $g \neq 0$ или существует такое $p \in \mathbb{N}$, что выполняются условия[2]

- $1 \leq p \leq n(g)$,
- если $p \leq m(f)$, то $t_p \prec s_p, t_1 = s_1, \ \ldots, t_{p-1} = s_{p-1}$,
- если $p > m(f)$, то $t_1 = s_1, \ \ldots, t_{m(f)} = s_{m(f)}$.

При выполнении условий $f \prec g$ и $f \neq 0$ такое число p обозначим через $p(f, g)$, а соответствующий терм t_p через $t(f, g)$.

Порядок $\prec$ индуцирует естественным образом отношение нестрогого порядка $\preceq$. А именно, $f \preceq g$ тогда и только тогда, когда $f \prec g$ или $m(f) = n(g)$ и выполняются равенства $t_1 = s_1, \ \ldots, t_{m(f)} = s_{n(g)}$. В последнем случае выполняется равенство $T_f = T_g$, определяющее отношение эквивалентности $f \equiv g$. Таким образом, $f \preceq g$ тогда и только тогда, когда $f \prec g$ или $f \equiv g$. Для любых многочленов f, g выполняется одно и только одно из соотношений $f \prec g$, $g \prec f$ или $f \equiv g$.

Согласно определениям редукции и допустимого порядка из соотношения $f \to_G h$ для простой монотонной редукции следует, что НТ$(f) \succ$ НТ(h), а для простой редукции верно соотношение $f \succeq h$ для отношения порядка $\succeq$ на множестве многочленов.

Лемма 10. *Пусть имеется бесконечная счетная невозрастающая последовательность многочленов из кольца* $A[X]$

$$f_1 \succeq f_2 \succeq \ \ldots \succeq f_n \succeq \ \ldots \ . \tag{4.4}$$

Тогда существует такое m, *что*

$$f_{m+1} \equiv f_{m+2} \equiv \ \ldots \equiv f_k \equiv \ \ldots \ .$$

[2]Данный предпорядок напоминает лексикографический.

Доказательство. Пусть утверждение леммы неверно. Тогда существует строго убывающая подпоследовательность последовательности $\{f_n\}$

$$f_{1,1} \succ f_{2,1} \succ \ \ldots \succ f_{n,1} \succ \ \ldots \ . \tag{4.5}$$

Согласно определению порядка $\succeq$ выполняются соотношения

$$\mathsf{HT}(f_{1,1}) \succeq \mathsf{HT}(f_{2,1}) \succeq \ \ldots \mathsf{HT}(f_{n,1}) \succeq \ \ldots \ .$$

Из допустимости порядка $\succ$ следует существование наименьшего терма в последовательности $\mathsf{HT}(f_{1,1}), \mathsf{HT}(f_{2,1}), \ \ldots, \mathsf{HT}(f_{n,1}), \ \ldots$. Пусть это $\mathsf{HT}(f_{m_1,1})$. Поэтому,

$$\mathsf{HT}(f_{1,1}) \succeq \mathsf{HT}(f_{2,1}) \succeq \ \ldots \succeq \mathsf{HT}(f_{m_1,1}) = \mathsf{HT}(f_{m_1+1,1}) = \ \ldots \ .$$

Определим последовательность многочленов формулой

$$f_{n,2} = \begin{cases} f_{n,1} & \text{при} \quad n \le m_1, \\ f_{n,1} - \mathsf{HM}(f_{n,1}) & \text{при} \quad n > m_1. \end{cases}$$

Заметим, что выполняется соотношение $\mathsf{HT}(f_{1,1}) \succ \mathsf{HT}(f_{m_1+1,2})$.

Согласно определению порядка $\prec$ на множестве многочленов эта последовательность строго убывает

$$f_{1,2} \succ f_{2,2} \succ \ \ldots \succ f_{k,2} \succ \ \ldots \ .$$

Для последовательности $\{f_{n,2}\}$, по тем же формулам, что и для последовательности $\{f_{n,1}\}$, строим строго убывающую последовательность $\{f_{n,3}\}$ и такое число m_2, для которого выполняется соотношение

$$\mathsf{HT}(f_{m_1+1,2}) \succ \mathsf{HT}(f_{m_2+1,3}).$$

Повторяя это построение для каждого натурального i, получим убывающие последовательности многочленов $\{f_{n,i+1}\}$ и числа m_i, для которых выполняются соотношения $\mathsf{HT}(f_{m_{i-1}+1,i}) \succ \mathsf{HT}(f_{m_i+1,i+1})$ и $m_i > m_{i-1}$. Следовательно, бесконечная цепочка термов

$$\begin{aligned} \mathsf{HT}(f_{1,1}) \ &\succ \mathsf{HT}(f_{m_1+1,2}) \succ \mathsf{HT}(f_{m_2+1,3}) \succ \ \ldots \\ &\succ \mathsf{HT}(f_{m_{i-1}+1,i}) \succ \mathsf{HT}(f_{m_i+1,i+1}) \succ \ \ldots \end{aligned}$$

строго убывает, что противоречит предположению о допустимости порядка $\prec$ на множестве термов. $\square$

Следующая лемма утверждает, что последовательность простых редукций (простых монотонных редукций) всегда приводит к нормальной форме.

Лемма 11. *Фиксируем допустимый порядок $\prec$ на множестве термов. Тогда для любого конечного подмножества G в кольце многочленов $A[x_1, \ldots, x_n]$ последовательность простых редукций по G не может быть бесконечной.*

Доказательство. Пусть утверждение леммы неверно. Тогда имеется бесконечная цепочка простых редукций (простых монотонных редукций)

$$f_1 \to_G f_2 \to_G \ \ldots \to_G f_m \to_G \ \ldots \ .$$

В силу определения редукции выполняются соотношения

$$f_1 \succ f_2 \succ \ \ldots \succ f_m \succ \ \ldots \ .$$

Но в силу леммы 10 это невозможно. Поэтому цепочка редукций всегда конечна и, следовательно, всегда приводит к нормальной форме. □

Утверждение леммы справедливо и для монотонных редукций. Доказательст повторяется дословно заменой редукции на монотонную редукцию.

4.3 Определение S-разности многочленов

Будем считать, что задано множество переменных X и на термах $T\langle X\rangle$ задан допустимый порядок $\prec$. Пусть A — кольцо главных идеалов.

Для $f, g \in A[X]$ положим

$$u_{f,g} = \frac{\text{НОК}(\text{НМ}(f), \text{НМ}(g))}{\text{НМ}(f)}, \quad u_{g,f} = \frac{\text{НОК}(\text{НМ}(f), \text{НМ}(g))}{\text{НМ}(g)}. \tag{4.6}$$

Определение 4.3.1. *S-разностью многочленов f и g называется многочлен*

$$S(f, g) = u_{f,g} f - u_{g,f} g. \tag{4.7}$$

Пусть $F = \{f_1, \ldots, f_m\} \subset A[x_1, \ldots, x_n]$ и $I = (f_1, \ldots, f_m)$ — идеал кольца $A[X]$. Тогда для любого $f \in I$ и всех $i = 1, \ldots, m$ существуют такие $\alpha_i \in A[X]$, что выполняется равенство многочленов

$$f = \sum_{i=1}^{m} \alpha_i f_i. \tag{4.8}$$

Правую часть соотношения (4.8)[3] будем называть формулой разложения f по множеству F. Отметим, что многочлен f может иметь различные формулы разложения по множеству F. Например, пусть $X = \{x, y\}$, $A = \mathbb{Z}$, $F = \{2x, y, x^2\}$ и $f(x, y) = 2x^2$. Тогда имеется бесконечно много различных формул разложения f по множеству F, например две такие:

$$f(x, y) = xf_1 + 0 \cdot f_2 + 0 \cdot f_3 \quad \text{и} \quad f(x, y) = 0 \cdot f_1 + 0 \cdot f_2 + 2 \cdot f_3.$$

Равенство формул разложений означает равенство векторов коэффициентов $(\alpha_1, \ldots, \alpha_n)$ этого разложения.

Определение 4.3.2. *Пусть I — идеал, порожденный множеством $F = \{f_1, \ldots, f_m\} \subset A[x_1, \ldots, x_n]$. Главным термом формулы разложения* (4.8) *элемента $f \in I$ по множеству F называется терм*

$$D\left(\sum_{i=1}^{m} \alpha_i f_i\right) = \max_{1 \le i \le n} \{\mathsf{HT}(\alpha_i)\mathsf{HT}(f_i)\}.$$

Всегда выполняется неравенство

$$\mathsf{HT}\left(\sum_{i=1}^{m} \alpha_i f_i\right) \preceq D\left(\sum_{i=1}^{m} \alpha_i f_i\right).$$

Отметим, что в левой части неравенства выражение $\sum_{i=1}^{m} \alpha_i f_i$ представляет многочлен, а в правой — формулу разложения.

[3] Формула разложения однозначно определяется вектором $(\alpha_1, \ldots, \alpha_m)$

Лемма 12. *Пусть* $F = \{f_1, \ldots, f_m\} \subset A[x_1, \ldots, x_n]$ *— конечное множество и* A *— целостное кольцо главных идеалов. Если при некотором* $k \leq m$ *формула разложения*

$$f = \sum_{i=1}^{m} \alpha_i f_i$$

элемента $f \in (f_1, \ldots, f_m)$ *по* F *такова, что* $\alpha_i = 0$ *при* $i > k$*,*

$$d = D\left(\sum_{i=1}^{k} \alpha_i f_i\right) = \mathsf{HT}(\alpha_1)\mathsf{HT}(f_1) = \ldots = \mathsf{HT}(\alpha_k)\mathsf{HT}(f_k) \tag{4.9}$$

и

$$\sum_{i=1}^{k} \mathsf{HM}(\alpha_i f_i) = 0, \tag{4.10}$$

то существуют мономы $\beta_1, \ldots, \beta_{k-1}$ *и такие многочлены* $\gamma_1, \ldots, \gamma_k$*, что*

$$f = \sum_{i=1}^{k} \alpha_i f_i = \sum_{0<i<j\leq k} \beta_{i,j} S(f_i, f_j) + \sum_{i=1}^{k} \gamma_i f_i, \tag{4.11}$$

причем $d \succ \mathsf{HT}(\gamma_i)\mathsf{HT}(f_i)$ *для всех* $i = 1, \ldots, k$ *и* $d \succ \mathsf{HT}(\beta_{i,j})\mathsf{HT}(S(f_i, f_j))$ *для всех* $0 < i < j \leq k$*.*

Доказательство. Доказательство проведем по индукции. Пусть $k = 2$, тогда $\mathsf{HT}(\alpha_1 f_1) = \mathsf{HT}(\alpha_2 f_2) = d$ и из соотношения (4.10) следует, что

$$\alpha_1 f_1 + \alpha_2 f_2 = \beta S(f_1, f_2) + \sum_{i=1}^{2} (\alpha_i - \mathsf{HM}(\alpha_i))\, f_i = \beta S(f_1, f_2) + \sum_{i=1}^{2} \gamma_i f_i,$$

где $d = \mathsf{HT}(\alpha_i)\mathsf{HT}(f_i) \succ \mathsf{HT}(\alpha_i - \mathsf{HT}(\alpha_i))\mathsf{HT}(f_i) = \mathsf{HT}(\gamma_i)\mathsf{HT}(f_i)$ при $i = 1, 2$, и

$$\mathsf{HM}(\alpha_1) f_1 + \mathsf{HM}(\alpha_2) f_2 = \beta S(f_1, f_2)$$

и, следовательно, выполняется соотношение $d \succ \mathsf{HT}(\beta)\mathsf{HT}(S(f_1, f_2))$.

Пусть возможность разложения по формуле (4.11) доказана для всех k, не превосходящих $s < m$. Докажем существование такого разложения для $k = s + 1$.

Ведем обозначения: НС$(\alpha_i) = a_i$, НС$(f_i) = b_i$. Поскольку A — кольцо главных идеалов, для некоторого $c_0 \in A$ выполняется равенство

$$(a_1 b_1, \ \ldots, a_s b_s) = (c_0).$$

Следовательно, существуют представления

$$c_0 = \sum_{i=1}^{s} d_i a_i b_i, \ d_i \in A \ \text{ и } \ a_i b_i = e_i c_0, i = 1, \ \ldots s. \tag{4.12}$$

Тогда $\sum\limits_{i=1}^{s} d_i e_i = 1$. Учитывая соотношения (4.9) и (4.10), получаем, что для некоторого $e_{s+1} \in A$ выполняется соотношение $a_{s+1} b_{s+1} = e_{s+1} c_0$. Поэтому выполняются равенства

$$\begin{aligned}
&\sum_{i=1}^{s} \alpha_i f_i + \alpha_{s+1} f_{s+1} = \sum_{i=1}^{s} \alpha_i f_i + \left(\sum_{i=1}^{s} d_i e_i \right) \alpha_{s+1} f_{s+1} \quad = \\
&\sum_{i=1}^{s} \alpha_i f_i + \sum_{i=1}^{s} d_i e_i \alpha_{s+1} f_{s+1} \quad = \\
&\sum_{i=1}^{s} (\alpha_i f_i + d_i e_{s+1} \alpha_i f_i) + \sum_{i=1}^{s} (d_i e_i \alpha_{s+1} f_{s+1} - d_i e_{s+1} \alpha_i f_i)
\end{aligned} \tag{4.13}$$

и, учитывая соотношения $a_i b_i = c_0 e_i$,

НМ$(c_0(d_i e_i \alpha_{s+1} f_{s+1} - d_i e_{s+1} \alpha_i f_i)) =$ НМ$(c_0 d_i e_i \alpha_{s+1} f_{s+1}) -$
НМ$(c_0 d_i e_{s+1} \alpha_i f_i) =$ НМ$(d_i a_i b_i a_{s+1} b_{s+1} d) -$ НМ$(d_i a_{s+1} b_{s+1} a_i b_i d) = 0.$

Поскольку кольцо A — целостное, выполняется равенство

$$\text{НМ}(d_i e_i \alpha_{s+1} f_{s+1} - d_i e_{s+1} \alpha_i f_i) = 0.$$

Поэтому, согласно начальному шагу индукции для каждого $i = 1, \ \ldots, s$ и условию (4.9), выполняется равенство

$$d_i e_i \alpha_{s+1} f_{s+1} - d_i e_{s+1} \alpha_i f_i = \beta_i S(f_i, f_{s+1}) + \sum_{j=1}^{s} \gamma_{i,j} f_j,$$

а β_i и $\gamma_{i,j}$ удовлетворяют условиям леммы. Из равенства (4.13) следует, что

$$\sum_{i=1}^{s}(1+d_ie_{s+1})\text{НМ}(\alpha_if_i)=0.$$

Для завершения доказательства, достаточно воспользоваться предположением индукции для формулы разложения

$$g=\sum_{i=1}^{s}(1+d_ie_{s+1})\alpha_if_i.$$

□

В следующей лемме сформулирован критерий Бухбергера.

Лемма 13. *Пусть кольцо A является целостным кольцом главных идеалов и конечное множество $G\subset A[x_1,\ \ldots,x_n]$ обладает следующим свойством:*

> *для любых $g_1,g_2\in G$ существует редукция к нулю их S-разности, т.е. $S(g_1,g_2)\to_{G^*}\underline{0}$.*

Тогда для любого ненулевого элемента идеала I, порожденного множеством G, любая его нормальная по модулю G форма нулевая и, следовательно, G — базис Гребнера идеала I.

Доказательство. Пусть существуют элементы f идеала, порожденного множеством $G=\{g_1,\ \ldots,g_m\}$, имеющие ненулевые нормальные формы. Среди таких нормальных форм выберем многочлен g, имеющий наименьший старший терм НТ(g). Такой многочлен существует, поскольку порядок на множестве термов является допустимым. Выберем среди всех разложений элемента g по множеству G разложение

$$g=\sum_{i=1}^{m}\alpha_ig_i \tag{4.14}$$

с минимальным главным термом $d=D\left(\sum\limits_{i=1}^{m}\alpha_ig_i\right)$.

Выделим теперь в разложении (4.14) элемента g по базису G слагаемые, старшие термы которых равны d. Без ограничения общности

$$g = \sum_{i=1}^{k} \alpha_i g_i + \sum_{i=k+1}^{m} \alpha_i g_i, \quad k \geq 1,$$

где $d = \mathsf{HT}(\alpha_i)\mathsf{HT}(g_i)$ при $i \leq k$ и $d \succ \mathsf{HT}(\alpha_i)\mathsf{HT}(g_i)$ при $i > k$. При $d = \mathsf{HT}(g)$ простая редукция

$$g = \sum_{i=1}^{k} \alpha_i g_i + \sum_{i=k+1}^{m} \alpha_i g_i \to_{G^*} \sum_{i=k+1}^{m} \alpha_i g_i = g'$$

приводит к многочлену g', имеющему ненулевую нормальную форму по модулю G, для которого $\mathsf{HT}(g) \succ \mathsf{HT}(g')$, что противоречит определению многочлена g. Поэтому $d \succ \mathsf{HT}(g)$ и $k > 1$. Отметим также, что из условия $d \succ \mathsf{HT}(g)$ следует, что

$$\sum_{i=1}^{k} \mathsf{HM}(\alpha_i)\mathsf{HM}(g_i) = 0. \tag{4.15}$$

Из соотношения (4.15) и леммы 12 следует существование таких многочленов β_i, γ_j, для которых выполняется соотношение

$$\sum_{i=1}^{k} \alpha_i g_i = \sum_{i=1}^{k-1} \beta_i S(g_i, g_{i+1}) + \sum_{i=1}^{k} \gamma_i g_i, \tag{4.16}$$

причем $d \succ \mathsf{HT}(\gamma_i)\mathsf{HT}(g_i)$ и $d \succ \mathsf{HT}(\beta_i)\mathsf{HT}(S(g_i, g_{i+1}))$.

Согласно предположению леммы $S(g_i, g_j) \to_{G^*} \underline{0}$. Поэтому $S(g_i, g_j) = \sum_{k=1}^{m} \lambda_{k,i,j} g_k$, где $\lambda_{k,i,j}$ — многочлены, причем

$$\mathsf{HT}(S(g_i, g_j)) \succeq \mathsf{HT}(\lambda_{k,i,j})\mathsf{HT}(g_k)$$

для всех i, j, k. Тогда

$$\sum_{i=1}^{k-1} \beta_i S(g_i, g_{i+1}) = \sum_{i=1}^{k-1} \beta_i \sum_{k=1}^{m} \lambda_{k,i,i+1} g_k = \sum_{k=1}^{m} \theta_k g_k,$$

где $\theta_k = \sum\limits_{i=1}^{k-1} \beta_i \lambda_{k,i,i+1}$, причем $d \succ \text{НТ}(\theta_k g_k)$. Следовательно, имеется разложение

$$g = \sum_{k=1}^{m} \theta_k g_k + \sum_{i=1}^{k} \gamma_i g_i + \sum_{i=k+1}^{m} \alpha_i g_i = \sum_{i=1}^{m} (\theta_i + \delta_i) g_i, \tag{4.17}$$

где

$$\delta_i = \begin{cases} \gamma_i & \text{при} \quad 1 \le i \le k, \\ \alpha_i & \text{при} \quad k+1 \le i \le m, \end{cases}$$

причем $d \succ D\left(\sum\limits_{i=1}^{m} (\theta_i + \delta_i) g_i\right)$. Полученная формула (4.17) противоречит предположению о минимальности разложения (4.14). Теперь из леммы 8 следует, что G является базисом Гребнера идеала I. □

Отметим, что в действительности доказано больше, а именно, что монотонная нормальная форма любого элемента идеала равна нулю.

Определение 4.3.3. *Пусть даны множество $G = \{g_1, \ \ldots, g_m\}$, элемент f кольца многочленов $P = A[x_1, \ \ldots, x_n]$ над кольцом главных идеалов A, терм t и задано представление*

$$f = \sum_{i=1}^{m} \alpha_i g_i, \ \ \alpha_i \in P. \tag{4.18}$$

Если $\text{НТ}(\alpha_i g_i) \preceq \text{НТ}(t)$ для всех $i = 1, \ \ldots, m$, то формула (4.18) называется t-представлением многочлена f по $G = \{g_1, \ \ldots, g_m\}$.

Доказательство леммы 13 приводит нас к следствию, усиливающему критерий Бухбергера.

Следствие 16. *Пусть кольцо A является кольцом главных идеалов и для конечного множества $G \subset A[x_1, \ \ldots, x_n]$ выполнено свойство:*

- *для любых $g_1, g_2 \in G$ при некотором $t \prec \text{НОК}(\text{НТ}(g_1), \text{НТ}(g_2))$ существует t-представление их S-разности.*

Тогда для любого ненулевого элемента идеала, порожденного множеством G, существует простая редукция по G, т.е. любая нормальная форма произвольного элемента этого идеала нулевая и множество G является базисом Гребнера идеала, порожденного множеством G.

Следствие 17. *Пусть A — кольцо главных идеалов. Следующие свойства конечного подмножества G кольца многочленов над A эквивалентны:*

- *для любых $g_1, g_2 \in G$ при некотором $t \prec$ НОК(НТ(g_1), НТ(g_2)) существует t-представление их S-разности;*
- *для любых $g_1, g_2 \in G$ существует редукция*

$$S(g_1, g_2) \to_{G^*} \underline{0};$$

- *для любых $g_1, g_2 \in G$*

$$(S(g_1, g_2) \to_{G^*} \underline{h}) \Rightarrow h = 0;$$

- *множество G является базисом Гребнера идеала, порожденного множеством G.*

Для редукций, не являющихся монотонными, выполняется

Теорема 40. *Следующие свойства конечного множества G в кольце многочленов от n переменных над полем A эквивалентны:*

1. *нормальная форма любого многочлена относительно G определена однозначно, т.е.*

$$\forall f \in A[x_1, \ \ldots, x_n] \ (f \to_{G^*} \underline{h_1}, \ f \to_{G^*} \underline{h_2}) \Rightarrow h_1 = h_2.$$

2. *для любых $g_1, g_2 \in G$*

$$(S(g_1, g_2) \to_{G^*} \underline{h}) \Rightarrow h = 0.$$

Доказательство. Сначала докажем, что $1 \Rightarrow 2$.

Пусть $S(g_1, g_2) \to_{G^*} \underline{h}$. Тогда определена цепочка редукций

$$S(g_1, g_2) \to_G \ \ldots \to_G \underline{h}.$$

Согласно определению 4.3.1

$$S(g_1, g_2) = u_{g_1,g_2} g_1 - u_{g_2,g_1} g_2.$$

Поэтому

$$u_{g_1,g_2} g_1 = u_{g_2,g_1} g_2 + S(g_1, g_2),$$

т.е.

$$u_{g_1,g_2} g_1 \to_G S(g_1, g_2).$$

Соединяя две полученные цепочки редукций, получим редукцию

$$u_{g_1,g_2} g_1 \to_G S(g_1, g_2) \to_G \ \ldots \to_G \underline{h},$$

т.е.

$$u_{g_1,g_2} g_1 \to_{G^*} \underline{h}. \tag{4.19}$$

С другой стороны $u_{g_1,g_2} g_1 \to_G \underline{0}$. Следовательно $h = 0$.

Докажем теперь, что $2 \Rightarrow 1$.

Предположим, что свойство 1 не выполняется для множества G. Тогда существует такой многочлен f, что

$$f \to_{G^*} \underline{h_1},\ f \to_{G^*} \underline{h_2}\ ,\ h_1 \neq h_2 \text{ и } \mathsf{HT}(h_1) \succeq \mathsf{HT}(h_2).$$

Тогда определена цепочка простых редукций

$$f \to_G f_1 \ \ldots \to_G f_k \to_G \underline{h_1}. \tag{4.20}$$

Поскольку h_2 не редуцируем, цепочке (4.20) соответствует редукция

$$f - h_2 \to_G f_1 - h_2 \ \ldots \to_G f_k - h_2 \to_G h_1 - h_2.$$

В силу определения нормальной формы, элемент $h_1 - h_2$ — нередуцируем. Следовательно,

$$f - h_2 \to_G f_1 - h_2 \ \ldots \to_G f_k - h_2 \to_G \underline{h_1 - h_2}. \tag{4.21}$$

Согласно построению, элемент $h_1 - h_2$ принадлежит идеалу, порожденному множеством G. Поэтому из свойства 2 и леммы 13 следует редуцируемость многочлена $h_1 - h_2$, что противоречит формуле (4.21). □

Следствие 18. *Приведение к нормальной форме по базису Гребнера идеала в кольце многочленов над полем k определяет гомоморфизм векторных k-пространств*

$$\varphi : k[x_1, \ \ldots, x_n] \to k[x_1, \ \ldots, x_n].$$

Доказательство. Пусть G — базис Гребнера идеала I. Отображение φ зададим формулой $\varphi(f) = h$, где $f \to_{G^*} \underline{h}$. Проверим гомоморфность.

Пусть $f \to_{G^*} \underline{h_1}$ и $g \to_{G^*} \underline{h_2}$. Из определения полной редукции следует, что сумма h_1+h_2 нередуцируема. Пусть $(f+g) \to_{G^*} \underline{h}$. Тогда $h-(h_1+h_2) \in I$. Поэтому $h - (h_1 + h_2) \to_{G^*} \underline{0}$ и, следовательно, учитывая нередуцируемость h_1, h_2 иh получаем равенство $h - (h_1 + h_2) = 0$, что означает равенство $\varphi(f + g) = \varphi(f) + \varphi(g)$.

Пусть теперь $f \to_{G^*} \underline{h}$ и $\alpha \in k$. Тогда $\alpha f \to_{G^*} \underline{\alpha h}$ и, следовательно, $\varphi(\alpha f) = \alpha h = \alpha\varphi(f)$. □

Замечание 4.3.1. При рассмотрении идеала над кольцом главных идеалов сумма нередуцируемых многочленов, а также произведение нередуцируемого многочлена на элемент кольца может уже быть редуцируемым многочленом, и в этом случае отображение φ уже не будет гомоморфизмом A-модулей.

Для монотонных редукций выполняется следующее утверждение.

Теорема 41. *Следующие свойства конечного множества G в кольце многочленов над полем эквивалентны:*

1. *старшие мономы нормальной формы относительно монотонных редукций любого многочлена определены однозначно*[4]

$$\forall f \in A[x_1, \ \ldots, x_n] \ \ (f \to_{G^*} \underline{h_1}, \ f \to_{G^*} \underline{h_2}) \Rightarrow \mathsf{HM}(h_1) = \mathsf{HM}(h_2).$$

2. *для любых $g_1, g_2 \in G$ и монотонных редукций*

$$(S(g_1, g_2) \to_{G^*} \underline{h}) \Rightarrow h = 0.$$

[4] для немонотонных редукций недостаточно совпадения старших мономов.

Далее редукции не предполагаются монотонными.

Определение 4.3.4. *Базис Гребнера $F = \{f_1, \ldots, f_k\}$ называется приведенным, если из $f_i \to_{F_i^*} \underline{h}$, где $F_i = F \setminus \{f_i\}$, следует, что $h = f_i$.*

Теорема 42. *Приведенный базис Гребнера идеала в кольце многочленов над полем относительно допустимого порядка $\succ$ определен однозначно с точностью до множителей из поля.*

Доказательство. Будем считать, что коэффициенты при старших термах равны 1.

Существование. Выберем произвольный базис Гребнера F идеала I и рассмотрим множество T_F всех старших термов многочленов из F конечное множество. Каждому терму $t \in T_F$ сопоставим многочлен из множества F со старшим термом t и составим из всех этих многочленов множество $F_0 = \{f_1, \ldots, f_m\}$. Поскольку множество старших термов множества F_0 совпадает с T_F, то согласно определению 4.1.1 множество F_0 также является базисом Гребнера идеала I. Без ограничения общности, $\mathrm{HT}(f_1) \succ \ldots \succ \mathrm{HT}(f_m)$. Положим $F_i = F_0 \setminus \{f_i\}$ и $f_i \to_{F_i^*} \underline{g_i}$. Тогда $\mathrm{HT}(g_i) = \mathrm{HT}(f_i)$ и элементы g_i удовлетворяют условиям определения 4.3.4 и, следовательно, составляют приведенный базис Гребнера идеала I.

Единственность. Пусть $G = \{g_1, \ldots, g_k\}$ и $G' = \{g'_1, \ldots, g'_m\}$ — два приведенных базиса идеала I, элементы которых перенумерованы в порядке убывания старших термов. Если $g_k \neq g'_m$ и, например, $\mathrm{HT}(g_k) \succ \mathrm{HT}(g'_m)$, то элемент g'_m принадлежит идеалу, но его старший терм не делится ни на один старший терм из базиса G, что противоречит определению базиса Гребнера. Следовательно, $\mathrm{HT}(g_k) = \mathrm{HT}(g'_m)$. Если $g_k \neq g'_m$, то разность $g_k - g'_m$ принадлежит идеалу и ее старший терм не делится ни на один из старших термов многочленов из G и G', что противоречит определению базиса Гребнера. Следовательно, $g_k = g'_m$.

Повторяя это рассуждение для элементов g_{k-1} и g'_{m-1}, получаем их совпадение. И так до тех пор, пока не переберем все элементы одного из множеств G или G'. Теперь пусть $k > m$. Тогда из $g_1 \to_{G'^*} \underline{0}$ следует, что $g_1 \to_{G_1^*} \underline{0}$. Поэтому $g_1 = 0$ и, следовательно, $k = m$ и $G = G'$. □

4.4 Оценка степени нормальной формы многочлена

Согласно теореме 2 положительная рациональная матрица A определяет допустимое упорядочение $\succ_A$ на множестве термов $T\langle X\rangle$. Пусть N — максимум числителей и наименьшего общего кратного знаменателей неприводимых дробей, представляющих элементы матрицы A, и W_B — рациональная форма, заданная формулой

$$W_B(v) = \sum_{i=1}^{n} \omega_i(v) \cdot B^{n-i} \tag{4.22}$$

(см. формулу (1.11)). Согласно следствию 4 отношение порядка $t \succ_A s$ эквивалентно отношению $W_B(t) > W_B(s)$ для термов степени не выше M для $B = nN^2M$. Поскольку согласно теореме 39 степень элементов базиса Гребнера идеала, порожденного многочленами степени не более чем d, не выше чем $\left(\frac{d^2}{2} + d\right)^{2^{n-1}}$, при

$$B = 2nN^2 \left(\frac{d^2}{2} + d\right)^{2^{n-1}}.$$

отношение $\succ_A$ определяется формой W_B.

Теорема 43. *Пусть $I \subset K[X]$ — идеал в кольце многочленов над полем K, порожденный многочленами степени не выше d. Тогда для любого многочлена $h \in K[X]$ степень его нормальной формы относительно базиса Гребнера идеала I для упорядочения $\succ_A$ не превосходит*

$$\left(\left(2n\left(\frac{d^2}{2} + d\right)^{2^{n-1}}\right)^n N^{2n} \deg(h)\right)^{n+1}.$$

Доказательство. Отметим, что согласно следствию 4 и теореме 39 для сравнения термов многочленов, составляющих базис Гребнера G идеала I относительно упорядочения $\succ_A$ можно использовать форму W_B.

Рассмотрим монотонную редукцию многочлена $h \in K[X]$ к нормальной форме относительно G

$$h \to_G h_1 \to_G \ \ldots \to_G \underline{h_k}. \tag{4.23}$$

Согласно определению формы W_B для любого многочлена $g \in G$ и любого терма t многочлена g для всех термов $s \in T\langle X\rangle$ выполняются неравенства $W_B(s\mathsf{HT}(g)) \geq W_B(st)$, для монотонной редукции (4.23) выполняются неравенства

$$W_B(\mathsf{HT}(h)) \geq W_B(\mathsf{HT}(h_1)) \geq \ \ldots \geq W_B(\mathsf{HT}(h_k)).$$

Поэтому

$$M = \max_{t\in T_h\langle X\rangle}\{W_B(t)\} \geq \max_{t\in T_{h_1}\langle X\rangle}\{W_B(t)\} \geq \ \ldots \geq \max_{t\in T_{h_k}\langle X\rangle}\{W_B(t)\}.$$

Пусть t — терм многочлена h, для которого значение $W_B(t)$ достигает максимума M. Тогда согласно формуле (4.22) и условию $B \geq 2$

$$W_B(t) = \sum_{i=1}^{n} \omega_i(t)\cdot B^{n-i} \leq \sum_{i=0}^{n-1} B^i N \deg(h) \leq B^n N \deg(h). \tag{4.24}$$

Для любого терма $s = x_1^{v_1}\cdot \ \ldots \cdot x_n^{v_n}$

$$W_B(s) \geq \omega_n(s) = \sum_{i=1}^{n} \frac{m_{i,n}}{n_{i,n}} v_i \geq \frac{v_1 + \ \ldots + v_n}{N} = \frac{\deg(s)}{N}.$$

Поэтому из неравенства $W_B(s) \leq W_B(t)$ следует, что

$$\frac{\deg s}{N} \leq B^n N \deg(h)$$

или, эквивалентно,

$$\deg s \leq B^n N^2 \deg(h).$$

Количество термов степени не выше k не превосходит k^n, поэтому количество термов s, для которых выполняются соотношения $W_B(s) \leq M = W_B(t)$, не превосходит

$$\left(B^n N^2 \deg(h)\right)^n = \left(\left(2nN^2\left(\frac{d^2}{2}+d\right)^{2^{n-1}}\right)^n N^2 \deg(h)\right)^n.$$

Поскольку длина редукции (4.23) не превосходит числа термов s, для которых $W_B(s) \leq M$, ее длина не более

$$\left(\left(2nN^2\left(\frac{d^2}{2}+d\right)^{2^{n-1}}\right)^n N^2 \deg(h)\right)^n. \tag{4.25}$$

Поскольку на каждом шаге редукции (4.23) степень многочлена возрастает не более чем на максимальную степень многочлена из базиса Гребнера, а число таких шагов оценивается формулой (4.25), степень нормальной формы многочлена h не превосходит

$$2\left(\left(2nN^2\left(\frac{d^2}{2}+d\right)^{2^{n-1}}\right)^n N^2 \deg(h)\right)^n \cdot \left(\frac{d^2}{2}+d\right)^{2^{n-1}} + \deg(h)$$

и, следовательно, не более

$$\left(\left(2n\left(\frac{d^2}{2}+d\right)^{2^{n-1}}\right)^n N^{2n} \deg(h)\right)^{n+1}.$$

□

Глава 5

Алгоритмы нахождения базиса Гребнера

Пусть $X = \{x_1, \ \ldots, x_n\}$ — множество переменных и A — нетерово коммутативное кольцо главных идеалов. Выберем допустимый порядок $\succ$ на $T\langle X\rangle$.

5.1 Алгоритм Бухбергера

Теорема 44. ***(Критерий Бухбергера).*** *Множество многочленов $G \subset A[X]$ является базисом Гребнера идеала (G), тогда и только тогда, когда S-разность любой пары многочленов g_1, g_2 из G редуцируема к нулю, т.е.*

$$S(g_1, g_2) \to_{G^*} \underline{0}. \tag{5.1}$$

Доказательство. **Необходимость.** Пусть G — базис Гребнера идеала, порожденного самим множеством G. Как следует из леммы 11, для любой пары многочленов g_1, g_2 из множества G существует такой нередуцируемый многочлен $h \in (G)$, для которого выполняется соотношение

$$S(g_1, g_2) \to_{G^*} \underline{h}.$$

Предположим, что $h \neq 0$. Тогда его старший моном ненулевой и, согласно определению базиса Гребнера приводим, а это противоречит нередуцируемости многочлена h. Следовательно, $h = 0$.

Достаточность. Следует из леммы 13. □

5.1.1 Сизигии старших членов

Чтобы уменьшить количество проверок 5.1 рассмотрим следующую конструкцию.

Определение 5.1.1. *Пусть $F = (f_1, \ldots, f_m) \in A[X]^m$. Сизигией старших членов F называется такой $S = (h_1, \ldots, h_m) \in A[X]^m$, что*

$$\sum_{i=1}^{m} h_i \mathsf{HM}(f_i) = 0.$$

Множество всех сизигий старших членов F будем обозначать $\mathcal{S}(F)$. В дальнейшем сизигию старших членов F будем называть просто сизигией.

Согласно определению 5.1.1 выполняется включение $\mathcal{S}(F) \subset A[X]^m$. Рассматривая $A[X]^m$ как $A[X]$-модуль, получаем представление $S = \sum_{i=1}^{m} h_i e_i$, где e_i, $i = 1, \ldots, m$, — стандартный $A[X]$-базис в $A[X]^m$. Очевидно, $\mathcal{S}(F)$ является подмодулем $A[X]$-модуля $A[X]^m$. Формула (4.7) может быть записана с использованием сизигий в виде скалярного произведения

$$S(f_i, f_j) = S_{f_i,f_j} \cdot F, \tag{5.2}$$

где $S_{f_i,f_j} = u_{f_i,f_j} e_i - u_{f_j,f_i} e_j \in \mathcal{S}(F)$. Сизигии $S_{f_i,f_j} \in \mathcal{S}(F)$ называются *критическими*.

Определение 5.1.2. *Сизигия $S \in \mathcal{S}(F)$ называется однородной степени $\omega \in \mathbb{Z}_+^n$, если*

$$S = (c_1 x^{\omega_1}, \ldots, c_m x^{\omega_m}) = \sum_{i=1}^{m} c_i x^{\omega_i} e_i, \ \ \omega_i \in \mathbb{Z}_+^n,$$

где $c_i \in A$ и $\mathsf{HT}(x^{\omega_i} f_i) = x^{\omega}$ при $c_i \neq 0$. Положим также по определению $\deg e_i = \mathsf{HT}(f_i)$.

Определение 5.1.3. *Старшим термом критической сизигии* $\sigma = S_{f_i,f_j}$ *называется* $u_{f_i,f_j}e_i$, *если* $\deg e_i \succ \deg e_j$ *или* $u_{f_j,f_i}e_j$ *в противном случае. Другое слагаемое этой сизигии* S_{f_i,f_j} *называется ее младшим термом. Старший терм обозначим через* $\mathrm{HT}(\sigma)$, *а младший —* $\mathrm{LT}(\sigma)$.

В силу определения 5.1.2 для всех $i = 1, \ldots, m$ выполняется равенство $\deg S = \omega_i + \deg e_i$. В частности сизигии S_{f_i,f_j} из формулы (5.2) являются однородными и $\deg S_{f_i,f_j} = \deg u_{f_i,f_j} + \deg e_i = \deg u_{f_j,f_i} + \deg e_j$.

Лемма 14. *Сизигии* $\mathcal{S}(F)$ *допускают единственное представление в виде суммы однородных сизигий.*

Доказательство. Пусть $S = (h_1, \ldots, h_m) \in \mathcal{S}(F)$. Для каждого $\omega \in \mathbb{Z}_+^n$ определим $h_{i,\omega}$ как слагаемое многочлена h_i, для которого $\deg(h_{i,\omega}f_i) = \omega$. Тогда согласно определению сизигии $S_\omega = (h_{1,\omega}, \ldots, h_{m,\omega})$ также является сизигией и выполняется соотношение $S = \sum\limits_{\omega\in\mathbb{Z}_+^n} S_\omega$. Единственность такого представления очевидна. □

Из доказательства леммы 12 следует

Лемма 15. *Пусть* A *— нетерово кольцо главных идеалов. Тогда любая однородная сизигия* S *представима в виде суммы* $S = \sum\limits_{1\le i<j\le m} g_{i,j}S_{f_i,f_j}$, *где* $g_{i,j}$ *— однородные многочлены, причем* $g_{i,j} = 0$ *при* $\deg S \neq \deg(g_{i,j}S_{f_i,f_j})$.

Следующая лемма уточняет критерий Бухбергера из леммы 13.

Следствие 19. *Базис* $G = \{g_1, \ldots, g_m\}$ *является базисом идеала* $I = (g_1, \ldots, g_m)$ *тогда и только тогда, когда для каждого элемента* $S = (h_1, \ldots, h_m)$ *из однородного базиса пространства сизигий* $\mathcal{S}(G)$ *выполняется соотношение*

$$S \cdot G \to_{G^*} \underline{0}.$$

Доказательство. Следует из леммы 16. □

5.1.2 Алгоритм Бухбергера

В данном разделе рассматриваются не только монотонные, но и общие редукции.

Введем некоторые обозначения. Рассмотрим произвольное множество M. Обозначим через $F_S(M)$ множество всех конечных подмножеств $F \subset M$ с нумерацией, т.е.$F = \{f_1, \ldots, f_k\}$. В случае $M = A[X]$ для обозначения множества $F_S(A[X])$ будем использовать запись F_S.

Фиксируем некоторый линейный порядок $<$ на кольце многочленов $A[X]$. В множестве $F_S(F \times F)$, где $F \in F_S$ выделим подмножество множества пар

$$F_P = \{H \in F_S(F \times F) \mid (f, g) \in H \Rightarrow f < g\}.$$

Функцию упорядочения пар sym : $A[X] \times A[X] \to A[X] \times A[X]$ определим формулой

$$\mathsf{sym}(p, q) = \begin{cases} (p, q) & \text{при} \quad p < q, \\ (q, p) & \text{при} \quad q < p, \\ (p, q) & \text{при} \quad p = q. \end{cases} \tag{5.3}$$

Опишем процедуры, используемые в алгоритме Бухбергера.

1. Процедура формирования множества пар вычисляет функцию $Q : F_S \to F_P$, которая для любого конечного упорядоченного множества $H = \{h_1, \ldots, h_k\}$, $H \subset A[X]$, порядок в котором задан нумерацией его элементов, строит множество $Q(H)$, $Q(H) \subset H \times H$, по формуле
$$Q(H) = \{(h_i, h_j) \in H \times H \mid i < j\}.$$

2. Вычисление S-разности многочленов из конечного множества $Q(H) \subset A[X] \times A[X]$. Результатом этой операции является конечное множество $F \subset A[X]$, где $F = S(H)$ и $S : A[X] \times A[X] \to A[X]$ — функция S-разности многочленов. Таким образои, определено отображение $S : F_P \to F_S$.

3. Редукция конечного множества многочленов к нормальной форме относительно конечного множества многочленов
$$R : F_S \times F_S \to F_S.$$

Пусть $(F, G) \in F_S \times F_S$. Для каждого $f \in F$ выберем полную редукцию (см. определение 4.2.2)

$$f \to_{G^*} \underline{h_f}.$$

Будем полагать

$$R(F, G) = \{h_f \mid f \in F, f \to_{G^*} \underline{h_f}, h_f \neq 0\}.$$

Для вычисление базиса Гребнера применяется следующая итеративная схема.

Метод Бухбергера.

- Вход: конечное множество $F = \{f_1, \ \ldots, f_m\}$ многочленов.
- Выход: базис Гребнера G идеала $I = (f_1, \ \ldots, f_m)$.

Шаг 1. $k := 1$, $G_1 = F$, $H_1 := Q(F)$, $M_1 := S(H_1)$, $N_1 := R(M_1, G_1)$;

Шаг 2. При $N_k = \emptyset$ положить $G := G_k$ и перейти к шагу 9;

Шаг 3. $l := k$, $k := k + 1$;

Шаг 4. $H_k := (G_l \times N_l) \cup Q(N_l)$ — формирование множества пар;

Шаг 5. $G_k := G_l \cup N_l$ — добавление к G_l нормальных форм S-разностей пар многочленов из множества H_l;

Шаг 6. $M_k := S(H_k)$ — вычисление S-разности пар многочленов множества H_k;

Шаг 7. $N_k := R(M_k, G_k)$ — полная редукция множества многочленов M_k относительно множества многочленов G_k.

Шаг 8. Перейти к шагу 2.

Шаг 9. Завершение алгоритма. Результат алгоритма — множество G.

Для любого конечного множества многочленов F алгоритм завершает вычисление спустя конечное число шагов k, поскольку в противном случае строго возрастающей последовательности множеств образующих G_k соответствовала бы строго возрастающая цепочка мономиальных идеалов с базисами НТ(G_k), которая, согласно теореме 8, всегда конечна.

Замечание 5.1.1. Поскольку при поиске базиса Гребнера с помощью алгоритма Бухбергера выполняется конечное число шагов, все многочлены, участвующие в вычислении, имеют конечную степень, ограниченную некоторым числом m. Это число m зависит от выбранного порядка и заданного конечного базиса идеала I. Согласно теореме 4 допустимое упорядочение может быть задано некоторой рациональной матрицей.

Теорема 45. *Задача нахождения базиса Гребнера в кольце многочленов над нетеровым кольцом главных идеалов является алгоритмически разрешимой.*

В приведенном алгоритме имеются избыточные операции. Для оптимизации алгоритма можно уменьшить множество выбранных пар, исключив ненужные пары, не приводящие к новым элементам базиса. Для этого необходимо модифицировать схему вычислений множеств M_k и H_k. Для достижения этой цели потребуются критерии исключения лишних пар в множествах H_k.

Лемма 16. *(Первый критерий Бухбергера.) Если* $p, q \in G$ *и* НОД(НМ(p), НМ$(q)) = 1$, *то* $S(p, q) \to_{\{p,q\}^*} \underline{0}$.

Доказательство. Из условия $p, q \in G$ следует, что $u_{p,q} =$ НМ(q) и $u_{q,p} =$ НМ(p). Положим

$$\alpha = p - \text{НМ}(p) \quad \text{и} \quad \beta = q - \text{НМ}(q).$$

Заметим, что НМ$(p) \succ$ НМ(α) и НМ$(q) \succ$ НМ(β). Представим многочлены α и β в виде сумм мономов по убыванию термов

$$\alpha = a_1 + \ \ldots + a_{m_p}, \quad \beta = b_1 + \ \ldots + b_{m_q}.$$

Тогда

$$S(p,q) = u_{p,q}p - u_{q,p}q = (q-\beta)p - (p-\alpha)q = \alpha q - \beta p,$$

и при этом для всех $0 < i \leq m_p$ и $0 < j \leq m_q$ выполняется $a_i\text{НМ}(q) \neq b_j\text{НМ}(p)$. Следовательно, имеется редукция

$$S(p,q) = \alpha q - \beta p \to_{\{p,q\}^*} \underline{0},$$

заданная представлением

$$S(p,q) = a_1 q + \ \ldots + a_{m_p} q - b_1 p - \ \ldots - b_{m_q} p.$$

□

Теперь можно сформулировать второй критерий Бухбергера.

Лемма 17. *Пусть $p, q, s \in A[X]$ и выполняются соотношения*

$$\text{НМ}(r) \mid \text{НОК}(\text{НМ}(p), \text{НМ}(q)), \tag{5.4}$$

то сизигия $S_{p,q}$ представляется в виде

$$S_{p,q} = aS_{r,q} - bS_{r,p},$$

где a и b — мономы.

Доказательство. Рассмотрим произвольную тройку многочленов p, q, r, для которой выполнено условие

$$\text{НМ}(r) \mid \text{НОК}(\text{НМ}(p), \text{НМ}(q)) \tag{5.5}$$

Пусть $d = \text{НОД}(\text{НМ}(p), \text{НМ}(q))$. Тогда имеются такие разложения $\text{НМ}(p) = dt$ и $\text{НМ}(q) = ds$, что мономы s и t взаимно просты. Из условия (5.5) следует, что $\text{НМ}(r) = d_1 t_1 s_1$, где $d_1 \mid d$, $t_1 \mid t$ и $s_1 \mid s$. Тогда

$$\begin{aligned} \text{НМ}(r) &= d_1 t_1 s_1, \\ \text{НМ}(p) &= d_1 d_2 t_1 t_2, \\ \text{НМ}(q) &= d_1 d_2 s_1 s_2. \end{aligned} \tag{5.6}$$

Из соотношений (5.6) получаем

$$vS(r,p) = d_2t_2r - s_1p,$$
$$uS(r,q) = d_2s_2r - t_1q,$$

где $u = \mathsf{НОД}(t_1, d_2)$ и $v = \mathsf{НОД}(s_1, d_2)$. Поэтому

$$\begin{aligned} S(p,q) &= s_1s_2p - t_1t_2q \\ &= s_2(s_1p - d_2t_2r) + s_2d_2t_2r - t_2(t_1q - d_2s_2r) - t_2d_2s_2r \\ &= t_2uS(r,q) - s_2vS(r,p). \end{aligned}$$

□

Из леммы 17 и следствия 19 получаем, что сизигия $S_{p,q}$ может быть исключена из критерия приводимости критических пар в алгоритме Бухбергера.

Второй критерий Бухбергера можно обобщить, построив произвольный базис в $A[X]$-модуле сизигий $\mathcal{S}(F)$, где F — конечное подмножество многочленов. Варианты таких обобщений описаны в работах [Cab04] и [CB12], однако приведенные в этих работах определения и доказательства неполны.

Далее будем предполагать, что кольцо коэффициентов A является полем. Введем несколько обозначений. Пусть $F = \{f_1, \ldots, f_m\} \subset A[X]$. Положим $T(F) = \{t_1, \ldots, t_m\}$, где $t_i = \mathsf{НТ}(f_i)$. Поскольку A — поле, то без ограничения общности можно считать, что $\mathsf{НМ}(f_i) = t_i$. При $i < j$ положим $S_{f_i,f_j} = \alpha_{i,j}e_i + \beta_{i,j}e_j$. Обозначим множество критических сизигий для F через Σ. Согласно лемме 15 множество Σ является однородным базисом $A[X]$-модуля $\mathcal{S}(F)$. Введем обозначение

$$\Sigma_\omega = \{s \in \Sigma \mid \deg s \prec \omega\}, \ \ \omega \in \mathbb{Z}_+^n.$$

Определение 5.1.4. *Однородная сизигия s степени $\omega \in \mathbb{Z}_+^n$ называется разложимой, если выполняется соотношение*

$$s = \sum_{\sigma \in \Sigma_\omega} c_\sigma t_\sigma \sigma, \ \ \omega = \deg(t_\sigma\sigma), t_\sigma \in T(A[X]), \ c_\sigma \in A.$$

Непосредственно из определения 5.1.4 следует

Лемма 18. *Сумма разложимых однородных сизигий одинаковой степени разложима.*

Из определения 5.1.4 и леммы 18 следует

Лемма 19. *Пусть Σ_0^ω — подмножество множества всех разложимых критических сизигий размерности $\omega \in \mathbb{Z}^n$, $\Sigma'_\omega = \Sigma_\omega \cup \Sigma_0^\omega$, s — однородная сизигия размерности ω и*

$$s = \sum_{\sigma \in \Sigma'_\omega} c_\sigma t_\sigma \sigma, \;\; \omega = \deg(t_\sigma \sigma), t_\sigma \in T(A[X]), \; c_\sigma \in A. \tag{5.7}$$

Тогда сизигия s — разложима.

В дальнейшем количество ненулевых элементов c_σ в разложении (5.7) будем называть длиной этого разложения.

Лемма 20. *Пусть заданы число k, подмножество $L = \{l_1, \; \ldots, l_p\}$ в $\{1, \; \ldots, m\} \setminus \{k\}$, $\omega \in \mathbb{Z}^n$ и такая однородная сизигия s размерности ω*

$$s = \sum_{i \in L} a_i t_i S_{f_i, f_k}, \;\; t_i \in T(A[X]), \deg(t_i S_{f_i, f_k}) = \omega,$$

что $a_i \neq 0$ при $i \in L$ и $\sum\limits_{i \in L} a_i = 0$. Тогда имеется разложение

$$s = \sum_{i=1}^{p-1} \left(\sum_{j=1}^{i} a_{l_j} \right) u_i S_{f_{l_i}, f_{l_{i+1}}}, \;\; u_i \in T(A[X]).$$

Доказательство. Согласно определению критических сизигий имеем

$$S_{f_i, f_j} = u_{i,j} e_{f_i} - u_{j,i} e_{f_j},$$

причем $\deg S_{i,j} = \deg u_{i,j} + \deg e_{f_i} = \deg u_{j,i} + \deg e_{f_j}$, $u_{i,j}, u_{j,i} \in T(A[X])$. По условию леммы для всех $i \in L$ выполняются равенства $\deg t_i + \deg u_{i,k} + \deg e_{f_i} = \omega$ и, следовательно, для всех пар $i, j \in L, \; i \neq j$

$$\begin{array}{l} t_i S_{f_i, f_k} - t_2 S_{f_j, f_k} = t_i(u_{i,k} e_{f_i} - u_{k,i} e_{f_k}) - t_j(u_{j,k} e_{f_j} - u_{k,j} e_{f_k}) = \\ (t_i u_{i,k} e_{f_i} - t_j u_{j,k} e_{f_j}) + (t_j u_{k,j} - t_i u_{k,i}) e_{f_k} = t_i u_{i,k} e_{f_i} - t_j u_{j,k} e_{f_j} = \\ t_{i,j} S_{f_i, f_j}, \;\; \text{где } t_{i,j} = x^{\omega - \deg S_{f_i, f_j}}. \end{array} \tag{5.8}$$

Сдедовательно,

$$s = \sum_{i \in L} a_i t_i S_{f_i,f_k} = \sum_{i=1}^{s-1} \left(\sum_{j=1}^{i} a_{l_j} \right) (t_{l_i} S_{f_{l_i},f_{l_k}} - t_{l_{i+1}} S_{f_{l_{i+1}},f_{l_k}}) =$$
$$\sum_{i=1}^{s-1} \left(\sum_{j=1}^{i} a_{l_j} \right) t_{l_i,l_{i+1}} S_{f_{l_i},f_{l_{i+1}}}.$$

□

Лемма 21. *Множество неразложимых критических сизигий Σ^* получается с помощью алгоритма:*

- *Шаг 1.* $S := \Sigma$.
- *Шаг 2. Найти сизигию* $s \in S$, *представимую в виде* $s = t_u u + t_v v$, *где* $\forall w \in \{u, v\} \subset \Sigma$ *выполняется* $t_w \in T(A[X])$ *и либо* $\deg t_w \succ 1$, *либо* $\deg t_w = 1$ *и* $w \notin S$.
- *Шаг 3. Если сизигия* s *найдена, то* $S := S \setminus \{s\}$ *перейти к шагу 2.*
- *Шаг 4.* $\Sigma^* := S$.

Доказательство. Без ограничения общности можно считать, что $\deg f_1 \prec \deg f_2 \prec \ldots \prec \deg f_m$. Определим порядок $<$ на Σ формулой:

$$\sigma_1 < \sigma_2 \Leftrightarrow (\mathsf{HT}(\sigma_1) < \mathsf{HT}(\sigma_2) \vee (\mathsf{HT}(\sigma_1) = \mathsf{HT}(\sigma_2) \wedge \mathsf{LT}(\sigma_1) < \mathsf{LT}(\sigma_2))).$$

Согласно лемме 5.1.4 все исключаемые сизигии разложимы. Достаточно проверить, что будут исключены все разложимые сизигии. Пусть это не так. Выберем в множестве Σ^* разложимую сизигию S_{f_i,f_j} наименьшей степени ω, минимальную относительно порядка $<$. Из ее возможных разложений вида (5.7), для которых $\Sigma_0^\omega = \Sigma^\omega \setminus \Sigma^*$, выберем разложение наименьшей длины. В этом разложении выберем максимальное относительно порядка $<$ ненулевое слагаемое $c_{\sigma_0} t_{\sigma_0} \sigma_0$, где $\sigma_0 = S_{f_k,f_l}$ и $k > l$. Тогда $j \leq k$.

При $i < j$ положим $S_{f_i,f_j} = \alpha_{i,j} e_i + \beta_{i,j} e_j$.

Пусть $j = k > l$. Тогда из равенства (5.7) и леммы 20 следует, что $S_{f_i,f_j} = t_{\sigma_0} S_{f_l,f_j} + t_2 S_{f_i,f_l} = t_2 S_{f_i,f_l} - t_{\sigma_0} \sigma_0$. Если $t_2 \neq 1$, то сизигия S_{f_i,f_j}

должна быть удалена на шаге 3 и, следовательно, не принадлежит Σ^*. Пусть $t_2 = 1$, тогда $S_{f_i,f_l} = S_{f_i,f_j} + t_{\sigma_0}\sigma_0$. Тогда, поскольку $i < j$ и $l < j$, то $S_{f_i,f_l} < S_{f_i,f_j}$ и имеется разложение

$$S_{f_i,f_l} = (c_{\sigma_0} - 1)t_{\sigma_0}\sigma_0 + \sum_{\sigma\in\Sigma'_\omega,\sigma\neq\sigma_0} c_\sigma t_\sigma \sigma,$$

что противоречит свойству минимальности сизигии S_{f_i,f_j}.

Следовательно, $k > j > i$. Положим

$$\Sigma_\omega(\sigma_0) = \{\sigma \in \Sigma'_\omega \mid \mathsf{HT}(t_\sigma\sigma) = t_{\sigma_0}t_0e_k,\ c_\sigma \neq 0\}.$$

Из соотношения (5.7) и неравенства $k > j > i$ следует, что

$$\sum_{S_{f_l,f_k}\in\Sigma_\omega(\sigma_0)\ |k>l} c_{S_{f_l,f_k}} t_{S_{f_l,f_k}} \alpha_{l,k} e_k = 0.$$

Поэтому

$$\sum_{S_{f_l,f_k}\in\Sigma_\omega(\sigma_0)\ |k>l} c_{S_{f_l,f_k}} = 0.$$

Тогда к разложению

$$\sum_{S_{f_l,f_k}\in\Sigma'_\omega\ |k>l} c_{S_{f_l,f_k}} t_{S_{f_l,f_k}} S_{f_l,f_k} \tag{5.9}$$

применима лемма 20, позволяющая получить новое разложение

$$\sum_{S_{f_l,f_k}\in\Sigma_\omega(\sigma_0)\ |k>l} c_{S_{f_l,f_k}} t_{S_{f_l,f_k}} S_{f_l,f_k} = \sum_{S_{f_i,f_k}\in\Sigma'_\omega\ |i<k,j<k} d_{S_{f_i,f_j}} t_{S_{f_i,f_j}} S_{f_i,f_j}, \tag{5.10}$$

имеющее меньшее число ненулевых слагаемых. Поэтому, заменяя слагаемые вида (5.9) в разложении (5.7) с помощью формулы (5.10) получаем новое разложение сизигии s, имеющее меньшее число слагаемых, что противоречит выбору этой сизигии. □

Из леммы 21 следует, что $\Sigma_d = \Sigma \setminus \Sigma^*$ — множество всех разложимых критических сизигий, а из алгоритма построения Σ^* вытекает, что Σ^* является базисом пространства сизигий $\mathcal{S}(G)$. При доказательстве леммы 21 был также определен порядок $<$ на множестве критических сизигий.

Лемма 22. *Простая редукция относительно множества Σ_d и порядка $<$ на множестве критических сизигий преобразует элемент $\sigma \in \Sigma^*$ в $\sigma' \in \Sigma^*$.*

Доказательство. Пусть

$$\sigma \to_{\Sigma_d} \sigma'.$$

Тогда $\sigma' \in \Sigma$ и выполняется равенство $\sigma' = \sigma - \sigma_1$, где $\sigma_1 \in \Sigma_d$. Пусть $\sigma' \in \Sigma_d$. Тогда $\sigma = \sigma' + \sigma_1 \in \Sigma_d$, что противоречит условию $\sigma \in \Sigma^* = \Sigma \backslash \Sigma_d$. Следовательно, $\sigma' \in \Sigma^*$. □

Поэтому определена полная редукция сизигий из множества Σ^* относительно множества разложимых сизигий Σ_d, задающая отображение $\Sigma^* \to \Sigma^*$. Образ этого отображения обозначим через Σ^{**}. Относительно этого множества справедлива

Лемма 23. *Множество Σ^{**} является базисом $A[X]$-модуля сизигий $\mathcal{S}(F)$, и при выполнении соотношения*

$$\sum_{\sigma \in \Sigma^{**}(\omega)} a_\sigma \sigma = \sum_{\sigma \in \Sigma_d} p_\sigma \sigma, \text{ где } \Sigma^{**}(\omega) = \{\sigma \in \Sigma^{**} \mid \deg \sigma = \omega\}, a_\sigma \in A, \tag{5.11}$$

всегда

$$\sum_{\sigma \in \Sigma_d} p_\sigma \sigma = 0.$$

Доказательство. Пусть это не так. Среди разложений вида (5.11) выберем соотношение, с ненулевой правой частью минимальной длины. Без ограничения общности можно считать, что $\deg(p_\sigma \sigma) = \omega$ и p_σ — мономы. Высотой критической сизигии $\sigma = \alpha e_f + \beta e_g$ будем называть e_f, если $f \succ g$, или e_g в противном случае. Высоту сизигии σ обозначим через $\mathsf{V}(\sigma)$. Положим

$$L = \{\sigma \in \Sigma_d \,|\, p_\sigma \neq 0\},$$

$$e = \max_{\sigma \in L} \mathsf{V}(\sigma),$$

и

$$L_0 = \{\sigma \in L \mid \mathsf{V}(\sigma) = e\}.$$

Тогда из определения Σ^{**} следует, что

$$\forall \alpha e_f + \beta e_g \in \Sigma^{**} \ \forall \tau \in L \ e_f \neq e, \ e_g \neq e.$$

Поэтому из условия (5.11) следует, что $\sum_{\sigma \in L_0} p_\sigma \text{НТ}\sigma = 0$ и, следовательно, можно воспользоваться леммой 20 для разложения

$$s = \sum_{\sigma \in L_0} p_\sigma \sigma$$

и получить разложение

$$s = \sum_{\sigma \in L_1} q_\sigma \sigma, \text{ где } L_1 \subset \Sigma_d,$$

меньшей длины. Следовательно, имеется соотношение вида (5.11), с правой частью меньшей длины. □

Для любого $L \subset \Sigma^{**}$ определен максимальный элемент в L. Обозначим этот элемент через $M(L)$.

Лемма 24. *Пусть Σ^{***} — результат следующего алгоритма:*

- *Шаг 1. $S^{***} := \emptyset$, $T := \Sigma^{**}$.*
- *Шаг 2. Пусть σ — нормальная форма (см. определение 4.2.2) элемента $M(T)$ относительно множества $T \setminus \{M(T)\}$.*
- *Шаг 3. Если $\sigma \neq 0$, то $S^{***} := S^{***} \cup \{\sigma\}$.*
- *Шаг 4. $T := T \setminus M(T)$.*
- *Шаг 5. Если $T \neq \emptyset$ перейти к шагу 2.*
- *Шаг 6. $\Sigma^{***} := S^{***}$.*

Тогда Σ^{***} является минимальным базисом $A[X]$-модуля сизигий $\mathcal{S}(F)$.

Доказательство. Каждый раз на шаге 4 данного алгоритма множество сизигий $S^{***} \cup T$ является базисом $A[X]$-модуля сизигий $\mathcal{S}(F)$. Поэтому достаточно доказать минимальность базиса Σ^{***}.

□

Пусть $\Sigma^{**} = \{s_1, \ \ldots, s_k\}$. Достаточно доказать, что любая однородная линейная комбинация элементов из Σ^{**} неразложима. Пусть это не так. Тогда среди множества разложимых линейных комбинаций из множества Σ^{**} выберем однородные разложения, имеющие минимальную степень ω:

$$\sum_{i=1}^{k} \alpha_i s_i = \sum_{S_{k,l} \in \Sigma_\omega} h_{k,l} S_{k,l}. \tag{5.12}$$

(5.12)

Опишем процедуру дальнейшего сокращения множества $T_n \in F_P$, фигурирующего в описании алгоритма Бухбергера, т.е. преобразование

$$\mathcal{A} : F_P \to F_P. \tag{5.13}$$

- Вход: Конечные множества $G = \{g_1, \ \ldots, g_m\} \subset A[X]$ и $\Sigma = \{S_{g_i,g_j} \mid i < j\}$.
- Выход: Множество $\Sigma^{**} \subset \Sigma$ — минимальная система образующих $A[X]$ -модуля сизигий $\mathcal{S}(G)$.

  Шаг 1. $M := \Sigma$; $k := 2$;

  Шаг 2. Исключить из S все сизигии S_{g_i,g_j}, такие, что $i < j < k$ и $\mathrm{НТ}(g_k) | \mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_j))$;

  Шаг 3. Исключить из S все сизигии S_{g_j,g_k}, $j < k$, такие, что при некотором $i < k$, $\mathrm{НТ}(g_k) \nmid \mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_j))$, $\mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_k)) \neq \mathrm{НОК}(\mathrm{НТ}(g_j), \mathrm{НТ}(g_k))$ и $\mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_k)) | \mathrm{НОК}(\mathrm{НТ}(g_j), \mathrm{НТ}(g_k))$;

  Шаг 4. Исключить из S все сизигии S_{g_i,g_k}, $j < k$, такие, что при некотором $i < j < k$, $\mathrm{НТ}(g_k) \nmid \mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_j))$ и $\mathrm{НОК}(\mathrm{НТ}(g_i), \mathrm{НТ}(g_k)) = \mathrm{НОК}(\mathrm{НТ}(g_j), \mathrm{НТ}(g_k))$;

Шаг 5. $k := k + 1$;

Шаг 6. Если $k \leq m$, перейти к шагу 2.

Шаг 7. Окончание алгоритма.

Для доказательства корректности алгоритма достаточно воспользоваться следствием 19 и леммой 17. На шаге 2 алгоритма исключаются сизигии согласно второму критерию Бухбергера.

Будем использовать переменные G, P, M, N, значениями которых являются подмножества множеств $A[X]$ и $A[X] \times A[X]$. Конечное подмножество $F \subset A[X]$ задает базис идеала $I \subset A[X]$. Требуется найти базис Гребнера этого идеала. Пусть имеется оракул $\mathcal{A}$, сокращающий количество пар: $\mathcal{A} : F_P \to F_P$, $\mathcal{A}(T) \subset T$. Согласно следствию 19 достаточно, чтобы множество критических сизигий для множества пар из $\mathcal{A}(T) \subset T$ составляло базис всех сизигий для рассматриваемого базиса G_i идеала I.

Модифицированный метод Бухбергера.

- Вход: конечное множество $F = \{f_1, \ldots, f_m\}$ многочленов.
- Выход: базис Гребнера G идеала $I = (f_1, \ldots, f_m)$.

Шаг 1. $k := 1$, $G_1 = F$, $T_1 := Q(F)$, $P_1 := \mathcal{A}(T_1)$, $M_1 := S(P_1)$, $N_1 := R(M_1, G_1)$;

Шаг 2. При $N_k = \emptyset$ положить $G := G_k$ и перейти к шагу 9;

Шаг 3. $l := k$, $k := k + 1$;

Шаг 4. $P_k := \mathcal{A}((G_l \times N_l) \cup Q(N_l))$ — формирование множества пар;

Шаг 5. $G_k := G_l \cup N_l$ — добавление к G_l нормальных форм S-разностей пар многочленов из множества T_l;

Шаг 6. $M_k := S(P_k)$ — вычисление S-разности пар многочленов множества P_k;

Шаг 7. $N_k := R(M_k, G_k)$ — полная редукция множества многочленов M_k относительно множества многочленов G_k.

Шаг 8. Перейти к шагу 2.

Шаг 9. Завершение алгоритма. Результат алгоритма — множество G.

Отличие данного модифицированного алгоритма от представленного выше алгоритма Бухбергера только в выборе пар $P_i \subset T_i$. Мы не уточняем здесь, как найти P_i. Уменьшить T_i можно, используя **первый критерий Бухбергера** из леммы 16, удалив пары (p, q), удовлетворяющие условию НОД(НМ(p), НМ$(q)) = 1$.

5.2 Метод Фожера F_4

Описание метода Фожера F_4 взято из работы [Fau99] .

Введем обозначения, используемые при описании алгоритма. Пусть A — кольцо, $R = A[X]$ — кольцо многочленов на множестве переменных X, $F = \{f_1, \ldots, f_k\} \subset R$ — конечное множество многочленов и $\prec$ — допустимое упорядочение на множестве термов $T\langle X\rangle$. Положим $T_F = \bigcup_{i=1}^{k} T_{f_i}$. Упорядоченное (по убыванию относительно $\succ$)множество термов T_F будем обозначать как

$$T_{\succ}(F) = [t_1, \ldots, t_m].$$

Следовательно, m — количество элементов множества T_F и $t_1 \succ \ldots \succ t_m$. Для краткости множество $T_{\succ}(F)$ будем обозначать через $T(F)$.

Определение 5.2.1. *Пусть M — матрица размера $k \times m$, состоящая из элементов $M_{i,j} \in A$, $T_M = [t_1, \ldots, t_m]$ — упорядоченное множество термов на множестве переменных X и $e_1, \ldots, e_m$ — стандартный базис в модуле A^m. Рассмотрим линейное отображение $\varphi : V_{T_M} \to A^m$ (V_{T_M} — подмодуль A-модуля $A[X]$, порожденный элементами T_M), заданное формулой $\varphi(t_i) = e_i$. Обратное преобразование обозначим через ψ_{T_M}. Отображение ψ_{T_M} позволяет интерпретировать векторы из A^m как многочлены. В частности, строки матрицы M можно интерпретировать как многочлены. Обозначим через (M, T_M) матрицу M с такой интерпретацией.*

Определение 5.2.2. *Строкам матрицы (M, T_M) сопоставим множество многочленов*

$$r(M, T_M) = \{\psi_{T_M}(M_i) \mid i = 1, \ \ldots, s\}\backslash\{0\},$$

где M_i — i-я строка матрицы M.

Для каждого списка многочленов $F = \{f_1, \ \ldots, f_k\}$ и упорядоченного множества термов $T(F) = [t_1, \ \ldots, t_m]$ сопоставим $k \times m$-матрицу $M^{(F,T_F)}$. Элементом $M_{i,j}$ этой матрицы является коэффициент многочлена f_i при терме t_j. Такую матрицу будем изображать как

$$M(F) = \begin{array}{c} \\ f_1 \\ f_2 \\ \vdots \\ f_k \end{array} \begin{array}{c} \begin{array}{cccc} t_1 & t_2 & t_3 & \ldots \end{array} \\ \begin{pmatrix} M_{1,1} & M_{1,2} & M_{1,3} & \ldots \\ M_{2,1} & M_{2,2} & M_{2,3} & \ldots \\ \vdots & \vdots & \vdots & \ldots \\ M_{k,1} & M_{k,2} & M_{k,3} & \ldots \end{pmatrix} \end{array} \tag{5.14}$$

и обозначать как $M = M^{(l,T_l,\prec)}$

Определение 5.2.3. *Пусть M — $k \times m$-матрица и $Y = [Y_1, \ \ldots, Y_m]$ — упорядоченное множество новых переменных. Эти переменные можно рассматривать как термы степеней 1. Тогда определено конечное множество линейных уравнений $F = r(M, Y)$, для которых можно найти приведенный базис Гребнера $\tilde{F}$ относительно лексикографического порядка $Y_1 \succ \ \ldots \succ Y_m$. По этому базису определим ступенчатую матрицу $\tilde{M} = A^{(\tilde{F},Y)}$, которую будем называть ступенчатой формой матрицы M. Соответствующая матрица выглядит так*

$$\tilde{M}(F) = \begin{array}{c} \\ \tilde{f}_1 \\ \tilde{f}_2 \\ \vdots \\ \tilde{f}_l \\ \tilde{f}_{l+1} \\ \vdots \\ \tilde{f}_k \end{array} \begin{array}{c} \begin{array}{ccccccccc} t_1 & t_2 & \ldots & t_s & t_{s+1} & & \ldots & & t_m \end{array} \\ \begin{pmatrix} 1 & 0 & \ldots & 0 & \star & & \ldots & & \star \\ 0 & 1 & \ldots & 0 & \star & & \ldots & & \star \\ & & \ddots & & \star & & \ldots & & \star \\ 0 & 0 & \ldots & 1 & \star & & \ldots & & \star \\ 0 & 0 & \ldots & 0 & 0 & 0 & & & 0 \\ & & \vdots & & 0 & & \ldots & & 0 \\ 0 & 0 & \ldots & 0 & 0 & 0 & & & 0 \end{pmatrix} \end{array} \tag{5.15}$$

Теорема 46. *Пусть $M(F)$ — $k \times m$-матрица, $[Y_1, \ldots, Y_m]$ — множество новых переменных, $F = r(M, Y)$, $\tilde{M}$ — ступенчатая форма матрицы M и $\tilde{F} = r(\tilde{M}, Y)$. Определим множества*

$$\tilde{F}^+ = \left\{ g \in \tilde{F} | \mathsf{HT}(g) \notin \mathsf{HT}(F) \right\}$$
$$\tilde{F}^- = \tilde{F} \backslash \tilde{F}^+$$

Для любого подмножества F_- множества F, такого что

- $\mathsf{size}(F_-) = \mathsf{size}(\mathsf{HT})(\tilde{F})$
- $\mathsf{HT}(F_-) = \mathsf{HT}(\tilde{F})$,

множество $G = \tilde{F}^+ \cup F_-$ является приведенным базисом R-модуля V_M, порожденного элементами множества F, т.е. для любого элемента $f \in V_{T_F}$ существует $\lambda_1, \ldots, \lambda_j \in A$ и такие элементы $g_1, \ldots, g_j \in G$, что $f = \sum_{i=1}^{j} \lambda_i g_i$, $\mathsf{HT}(g_1) = \mathsf{HT}(f)$, $\mathsf{HT}(g_i) \succ \mathsf{HT}(g_{i+1})$.

Опишем теперь метод F_4, предложенный Фожером. Пусть $R = A[X]$.

Определение 5.2.4. *Определим отображение*

$$\mathcal{P} : R \times R \to T\langle X \rangle \times T\langle X \rangle \times R \times T\langle X \rangle \times R,$$

заданное формулой $\mathcal{P}(f, g) := (\mathsf{НОК}(\mathsf{HT}(f), \mathsf{HT}(g)), t(u_{f,g}), f, t(u_{g,f}), g)$, где $u_{f,g}$ и $u_{g,f}$ определены в формуле (4.6) при задании S-разности и $t(m)$ — терм монома m. Образ этого отображения называется множеством критических пар многочленов, а $\mathcal{P}(f, g)$ критической парой для пары многочленов (f, g). Степенью критической пары $\mathcal{P}(f, g)$ называется степень терма $\mathsf{НОК}(\mathsf{HT}(f), \mathsf{HT}(g))$. Определим также две проекции для критической пары

$$L(\mathcal{P}(f, g)) := (t(u_{f,g}), f) \text{ и } R(\mathcal{P}(f, g)) := (t(u_{g,f}), f).$$

Пусть $\mathcal{R}$ — множество конечных подмножеств кольца многочленов R, $\mathcal{P}(R)$ — множество конечных подмножеств множества $R \times R$ и $\Delta(R) = \{(x, x) \in R \times R\}$.

Алгоритм F_4.

- **Вход.** $\begin{cases} \text{Конечное подмножество F — базис идеала } I \subset A[X]. \\ \text{Функция Sel} : \mathcal{P}(R) \to \mathcal{P}(R), \text{ такая что Sel}(l) \neq \emptyset \text{ при } l \neq \emptyset. \end{cases}$
- **Выход.** Конечное подмножество в R.
- шаг 1. $G := F,\ \tilde{F}_0^+ := F,\ d := 0.$
- шаг 2. $P := \mathcal{P}(G \times G \setminus \Delta(G))$.
- шаг 3. Если $P = \emptyset$, то переходим к шагу 12.
- шаг 4. $d := d + 1$.
- шаг 5. $P_d :=$ Sel(P).
- шаг 6. $P := P \setminus P_d$.
- шаг 7. $L_d := L(P_d) \cup R(P_d)$.
- шаг 8. $\tilde{F}_d^+ :=$ Reduction(L_d, G).
- шаг 9. $P := P \cup \mathcal{P}(\tilde{F}_d^+ \times G)$.
- шаг 10. $G := G \cup \tilde{F}_d^+$.
- шаг 11. Перейти к шагу 3.
- шаг 12. Конец вычислений.

Опишем используемую в алгоритме процедуру приведение Reduction.
Reduction.

- **Вход.** $\begin{cases} \text{Конечное подмножество L множества } T\langle X\rangle \times R. \\ \text{Конечное подмножество } G \text{ множества } R. \end{cases}$
- **Выход.** Конечное подмножество в R (возможно пустое).
- шаг 1. $F :=$ Symbcomp(L, G).
- шаг 2. $\tilde{F} :=$ Приведение к ступенчатому виду множества F.

- шаг 3. $\tilde{F}^+ := \{f \in \tilde{F} \mid \mathsf{HT}(f) \notin \mathsf{HT}(F)\}$.
- шаг 4. Конец вычислений.

Опишем теперь процедуру Symbcomp — символьную предобработку.
Symbcomp.

- **Вход.** $\begin{cases} \text{Конечное подмножество L множества } T\langle X\rangle \times R. \\ \text{Конечное подмножество } G \text{ множества } R. \end{cases}$
- **Выход.** Конечное подмножество в R (возможно пустое).
- шаг 1. $F := \{t \cdot f \mid (t, f) \in L\}$.
- шаг 2. $D := \mathsf{HT}(F)$.
- шаг 3. Если $T(F) = D$, то перейти к шагу 10.
- шаг 4. Выберем $m \in T(F) \setminus D$.
- шаг 5. $D := D \cup \{m\}$
- шаг 6. Если не существует простой монотонной редукции терма m по множеству G, то переходим к шагу 5.
- шаг 7. Выбираем $f \in G$ и $m' \in T\rangle X\langle$, для которых $m = m' \cdot \mathsf{HT}(f)$.
- шаг 8. $F := F \cup \{m' \cdot f\}$.
- шаг 9. Вернуться к шагу 3.
- шаг 10. Конец вычислений

Выбор множества пар (функция Sel) с помощью критерия F_5 описана в следующем разделе. Соответствующий алгоритм F_4 для критерия F_5 называется методом Фожера F_5.

5.3 Критерий Фожера F_5

Доказательство критерия F_5 взято из работы [Ger10].

5.3.1 Формулировка критерия

5.3.1.1 Определение индекса, сигнатуры и атрибута многочлена

Обозначим через R кольцо многочленов $R = K[x_1, \ldots, x_n]$ над полем K. Фиксируем допустимый порядок $\preceq$ на множестве термов $T = T\langle x_1, \ldots, x_n\rangle$. Пусть также заданы многочлены $f_1, \ldots, f_m \in R$. Положим $I = (f_1, \ldots, f_m)$.

Для всех $i = 1, \ldots, m$ определим идеалы $I_i = (f_i, \ldots, f_m)$, а также $I_{m+1} = (0)$. Рассмотрим произвольный многочлен $f \in I$. Максимальное из чисел i, для которых выполняется соотношение $f \in I_i$, будем называть индексом многочлена f и обозначать через $i(f)$. В этом случае, очевидно, всегда существует такой элемент $\lambda \in R$, для которого выполняется соотношение

$$f \equiv \lambda f_{i(f)} \pmod{I_{i(f)+1}}.$$

Минимальный терм среди старших термов таких многочленов λ назовем сигнатурой многочлена f и обозначим через $s(f)$, а представитель такого терма $s(f)$ — через λ_f. Тогда выполняются соотношения

$$f \equiv \lambda_f f_{i(f)} \pmod{I_{i(f)+1}} \text{ и } s(f) = HT(\lambda_f).$$

Пару $L(f) = (i(f), s(f)) \in \mathbb{N} \times T$ назовем атрибутом многочлена f. Умножение атрибута (i, t) и терма определим с помощью формулы: для любого терма $t_1 \in T$ положим

$$t_1(i, t) = (i, t_1 t)$$

Определим порядок $\trianglelefteq$ на множестве атрибутов $\mathbb{N} \times T$:

$$L(f) \trianglelefteq L(g) \Leftrightarrow \left[\begin{array}{l} i(f) > i(g) \\ i(f) = i(g),\ s(f) \preceq s(g). \end{array} \right.$$

Заметим, что отношение $\trianglelefteq$ задает допустимый порядок на множестве $\mathbb{N} \times T$ относительно умножения на термы. Соответствующим образом задаем строгий порядок $\triangleleft$ на множестве атрибутов:

$$L(f) \triangleleft L(g) \Leftrightarrow \left[\begin{array}{l} i(f) > i(g) \\ i(f) = i(g),\ s(f) \prec s(g). \end{array} \right.$$

Далее символ $\prec$ всегда обозначает строгий порядок, соответствующий порядку $\preceq$.

5.3.1.2 Формулировка критерия

Пусть $f \in I$. Будем говорить, что $f = o_G(h)$, если существует такое представление

$$f = \sum_{g \in G} \alpha_g g,$$

что для всех $g \in G$, выполняются неравенства:

$$\mathsf{HT}(\alpha_g g) \prec \mathsf{HT}(h) \quad \text{и} \quad L(\alpha_g g) \trianglelefteq L(h).$$

Пусть M — произвольное подмножество кольца многочленов R. Обозначим через $\mathsf{HT}(M)$ множество всех старших термов многочленов из множества M. Следующая теорема представляет собой критерий Фожера F_5.

Теорема 47. *Пусть G — конечное подмножество идеала $I = (f_1, \ldots, f_m)$, содержащее многочлены $f_1, \ldots, f_m$. Если для всех $f, g \in G$, удовлетворяющих условиям*

$$u_{f,g} s(f) \notin \mathsf{HT}(I_{i(f)+1}), \quad u_{g,f} s(g) \notin \mathsf{HT}(I_{i(g)+1}), \quad u_{f,g} L(f) \neq u_{g,f} L(g),$$

выполняются следующие соотношения

$$S(f, g) = o_G(u_{f,g} f),$$

то G является базисом Гребнера идеала I.

5.3.2 Доказательство критерия

Далее при доказательстве критерия будем предполагать, что множество G удовлетворяет условиям теоремы 47.

5.3.2.1 Определение символа $m(f)$ и множеств B_i, N_i, M_i

Положим $G_i = G \cap I_i$. Тогда $(G_i \setminus G_{i+1}) \cap I_{i+1} = \emptyset$.

Выберем произвольное число $i \in \{1, \ldots, m\}$. Согласно определению сигнатуры для каждого $f \in I_i \setminus I_{i+1}$, выполняется равенство

$$s(f) = \min \max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\alpha_g s(g)),$$

где min берется по всем представлениям

$$f \equiv \sum_{g \in G_i \setminus G_{i+1}} \alpha_g g \pmod{I_{i+1}}. \tag{5.16}$$

Введем обозначение:

$$m(f) = \min \max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\alpha_g g),$$

где min берется по всем представлениям (5.16), на которых достигается значение $s(f)$.

Множества B_i, N_i и M_i определим формулами:

$$B_i = \{f \in I \mid i(f) = i, \mathsf{HT}(f) \prec m(f)\},$$

$$N_i = \{f \in B_i \mid s(f) = \min_{g \in B_i} s(g)\}, \quad M_i = \{f \in N_i \mid m(f) = \min_{g \in N_i} m(g)\}.$$

Заметим, что $B_m = \emptyset$. Отметим также, что если для некоторого индекса i множество B_i не пусто, то множество M_i также непусто.

Рассмотрим отображение $T : R \to \mathbb{N} \times T$, заданное формулой $T(f) = (i(f), \mathbf{HT}(f))$. Определим порядок $\propto$ на множестве $T(R)$:

$$(i,t) \propto (j,s) \Leftrightarrow \left[\begin{array}{l} i > j \\ i = i,\ t \prec s. \end{array} \right.$$

Множество $T(R)$ вполне упорядочено относительно $\propto$.

Лемма 25. *Если все $B_k = \emptyset$, то G — базис Гребнера идеала I.*

Доказательство. Достаточно проверить, что для любого $f \in I$ существует его простая монотонная редукция с помощью множества G. Поскольку множество $T(R)$ вполне упорядочено, доказательство можно провести с помощью трансфинитной индукции относительно $(i,t) \in T(R)$.

Пусть $T(f) = (m, 1)$. Тогда из равенства $B_m = \emptyset$ следует, что $m(1) \preceq \mathsf{HT}(f) = 1$, т.е. существует $1 \in G_m$. Тогда f редуцируется к нулевому многочлену.

Пусть доказано, что для всех многочленов $f \in I$, таких что $T(f) \propto (i,t)$, существует простая монотонная редукция.

Докажем, что тогда и для всех многочленов $f \in I$, таких что $T(f) = (i,t)$, такая редукция существует. Поскольку $B_i = \emptyset$, существует такое представление

$$f = \sum_{g \in G_i \setminus G_{i+1}} \alpha_g g \ (\mathrm{mod} I_{i+1}),$$

что

$$m(f) = \max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\alpha_g g) \preceq \mathsf{HT}(f) = t.$$

Если при некотором $g \in G_i$ выполняется равенство $\mathsf{HT}(\alpha_g g) = t$, то существование простой монотонной редукции для f доказано. В противном случае многочлен $f^1 = f - \sum_{g \in G_i \setminus G_{i+1}} \alpha_g g \in I$ имеет старшим термом t и $T(f^1) \propto (i,t)$ и, следовательно, по предположению индукции существует его простая монотонная редукция. Поэтому существует простая монотонная редукция многочлена f. □

5.3.2.2 Неприводимость $s(f)$

Лемма 26. *Для любого $f \in I$ выполняется соотношение*

$$s(f) \notin \mathsf{HT}(I_{i(f)+1}).$$

Доказательство. По определению $s(f)$ выполняется соотношение $f \equiv \lambda_f f_{i(f)} \ (\mathrm{mod} I_{i(f)+1})$, причем $\mathsf{HT}(\lambda_f) = s(f)$. Пусть $s(f) \in \mathsf{HT}(I_{i(f)+1})$, тогда существует многочлен $h \in I_{i(f)+1}$, для которого $\mathsf{HT}(\lambda_h) = s(f)$. Тогда $f \equiv (\lambda_f - h) f_{i(f)} \ (\mathrm{mod} I_{i(f)+1})$, причем $\mathsf{HT}(\lambda_f - h) \prec s(f)$, что противоречит минимальности $s(f)$. □

Из леммы 26 непосредственно выводится следующее утверждение.

Лемма 27. *Пусть $g \in G$ и*

$$\mathsf{HT}(\lambda s(g)) \in \mathsf{HT}(I_{i(g)+1}).$$

Тогда либо $\lambda g \in I_{i(g)+1}$, либо $s(\lambda g) \prec \mathsf{HT}(\lambda s(g))$.

Доказательство. Пусть $\lambda g \notin I_{i(g)+1}$. Тогда $i(\lambda g) = i(g)$. Согласно определению λ_g выполняется равенство

$$\lambda g \equiv \lambda\lambda_g f_{i(g)} \pmod{I_{i(g)+1}}$$

Поскольку $\mathsf{HT}(\lambda s(g)) \in \mathsf{HT}(I_{i(g)+1})$, существует такой многочлен $h \in I_{i(g)+1}$, что $\mathsf{HT}(\lambda\lambda_g - h) \prec \mathsf{HT}(\lambda s(g))$ и

$$\lambda g \equiv (\lambda\lambda_g - h) f_{i(g)} \pmod{I_{i(g)+1}}.$$

Следовательно, согласно определению сигнатуры, учитывая равенство $i(\lambda g) = i(g)$, получим

$$s(\lambda g) \preceq \mathsf{HT}(\lambda\lambda_g - h) \prec \mathsf{HT}(\lambda s(g)).$$

□

5.3.2.3 Представление многочленов из M_i

Лемма 28. *Пусть $f \in M_i$ для некоторого $i = 1, \ldots, m$, и пусть задано представление*

$$f \equiv \sum_{g \in G_i \setminus G_{i+1}} \alpha_g g \pmod{I_{i+1}},$$

на котором достигаются значения $s(f)$ и $m(f)$. Тогда существует такой многочлен $g \in G_i \setminus G_{i+1}$, для которого выполняются соотношения

$$\mathsf{HT}(\alpha_g s(g)) = s(f) \quad \text{и} \quad \mathsf{HT}(\alpha_g g) = m(f).$$

Доказательство. Положим

$$J = \{g \in G_i \setminus G_{i+1} \mid \mathsf{HT}(\alpha_g s(g)) = s(f)\}.$$

Пусть

$$\max_{g \in J} \mathsf{HT}(\alpha_g g) \prec m(f). \tag{5.17}$$

Определим

$$h = f - \sum_{g \in J} \alpha_g g. \tag{5.18}$$

Тогда $s(h) \prec s(f)$. Докажем, что в этом случае $h \notin B_i$.

Пусть это не так, т.е. $h \in B_i$. Заметим, что $M_i \subset N_i \subset B_i$. Поскольку $f \in M_i$, то $f \in B_i$, $f \in N_i$ и, следовательно, по определению множества N_i, для всех $g \in B_i$ выполняется соотношение $s(f) \preceq s(g)$. Поэтому из включения $h \in B_i$ следует, что $s(f) \preceq s(h)$, а это противоречит установленному выше соотношению $s(h) \prec s(f)$.

Из соотношения (5.17) следует, что $h \notin I_{i+1}$. Поскольку $h \notin B_i$, существует такое представление

$$h \equiv \sum_{g \in G_i \setminus G_{i+1}} \nu_g g \pmod{I_{i+1}}, \tag{5.19}$$

для которого выполняются соотношения

$$\max_{g \in G_i \setminus G_{i+1}} \mathrm{HT}(\nu_g s(g)) = s(h) \text{ и } \max_{g \in G_i \setminus G_{i+1}} \mathrm{HT}(\nu_g g) \preceq \mathrm{HT}(h).$$

Из соотношений (5.17) и (5.18) следует, что $\mathrm{HT}(h) \prec m(f)$, а ввиду условия $f \in B_i$ выполняется соотношение $\mathrm{HT}(f) \prec m(f)$. Поэтому справедливо неравенство

$$\max_{g \in G_i \setminus G_{i+1}} \mathrm{HT}(\nu_g g) \prec m(f). \tag{5.20}$$

Теперь из соотношений (5.18) и (5.19) следует существование представления

$$f \equiv \sum_{g \in J} (\alpha_g + \nu_g) g + \sum_{g \in (G_i \setminus G_{i+1}) \setminus J} \nu_g g \pmod{I_{i+1}},$$

которое, согласно соотношениям (5.17) и (5.20), противоречит минимальности $m(f)$. □

Лемма 29. *Пусть $f \in M_i$. Тогда существует такое представление*

$$f \equiv \sum_{g \in G_i \setminus G_{i+1}} \alpha_g g \pmod{I_{i+1}}, \tag{5.21}$$

для которого значения $s(f)$ и $m(f)$ достигаются на единственном многочлене $g \in G_i \setminus G_{i+1}$:

$$\mathrm{HT}(\alpha_g s(g)) = s(f) \text{ и } \mathrm{HT}(\alpha_g g) = m(f). \tag{5.22}$$

Доказательство. Рассмотрим произвольное представление (5.21), на котором достигаются значения $s(f)$ и $m(f)$. Согласно лемме 28 множество $G_i \setminus G_{i+1}$ содержит по крайней мере один многочлен, удовлетворяющий условиям (5.22). Пусть такой многочлен не единственный. Обозначим такие многочлены через $g_1, g_2 \in G_i \setminus G_{i+1},\ g_1 \neq g_2$,

$$\mathsf{HT}(\alpha_{g_j} s(g_j)) = s(f) \ \text{ и } \ \mathsf{HT}(\alpha_{g_j} g_j) = m(f) \ \text{ при } j = 1, 2. \tag{5.23}$$

Выберем такой элемент c поля K, для которого старшие коэффициенты многочленов $c\alpha_{g_1} s(g_1)$ и $\alpha_{g_2} s(g_2)$ совпадают.

Рассмотрим представление

$$f \equiv (1+c)\alpha_{g_1} g_1 + (\alpha_{g_2} g_2 - c\alpha_{g_1} g_1) + \sum_{g \in (G_i \setminus G_{i+1}) \setminus \{g_1, g_2\}} \alpha_g g \ (\mathsf{mod}\, I_{i+1}).$$

Положим $h = \alpha_{g_2} g_2 - c\alpha_{g_1} g_1$. Имеются две возможности: либо $h \notin I_{i+1}$, либо $h \in I_{i+1}$.

Пусть $h \notin I_{i+1}$. Согласно выбору многочленов g_1 и g_2 выполняется неравенство $\mathsf{HT}(h) \preceq \mathsf{HT}(\alpha_2 g_2) = m(f)$. По определению сигнатуры s, очевидно, что $s(h) \prec s(f)$. В этом случае (см. доказательстве леммы 28) $h \notin B_i$, т.е. либо $h \in I_{i+1}$ (вторая возможность), либо существует такое представление

$$h \equiv \sum_{g \in G_i \setminus G_{i+1}} \nu_g g \ (\mathsf{mod}\, I_{i+1}), \tag{5.24}$$

для которого достигаются значения $s(h)$ и $m(h)$ и, следовательно, выполняются соотношения

$$\max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\nu_g s(g)) = s(h) \ \text{ и } \ \max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\nu_g g) \preceq \mathsf{HT}(h) \preceq m(f).$$

Поскольку для $h \in I_{i+1}$ можно записать представление (5.24) с нулевыми слагаемыми ν_g, вне зависимости от выполнения включения $h \in I_{i+1}$, всегда можно предполагать, что существует представление (5.24), для которого выполняются условия:

$$\max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\nu_g s(g)) \prec s(f) \ \text{ и } \ \max_{g \in G_i \setminus G_{i+1}} \mathsf{HT}(\nu_g g) \preceq m(f). \tag{5.25}$$

Если для некоторого $g \in G_i \setminus G_{i+1}$ выполнено соотношение $\mathsf{HT}(\alpha_g s(g)) = s(f)$, то из первого неравенства формулы (5.25) следует, что $\mathsf{HT}(\nu_g) \prec \mathsf{HT}(\alpha_g)$ для такого g, и, таким образом,

$$\mathsf{HT}(\nu_g g) \prec \mathsf{HT}(\alpha_g g) \preceq m(f).$$

Поэтому в представлении

$$f \equiv ((1+c)\alpha_{g_1} + \nu_{g_1})g_1 + \nu_{g_2} g_2 + \sum_{g \in (G_i \setminus G_{i+1}) \setminus \{g_1, g_2\}} (\alpha_g + \nu_g) g \,(\mathrm{mod}\, I_{i+1}),$$

число элементов $g \in G_i \setminus G_{i+1}$, на которых одновременно достигаются значения $s(f)$ и $m(f)$, меньше, чем в исходном представлении на 1 или 2, в зависимости от того, равен ли нулю коэффициент $1 + c$.

Повторяя эту процедуру соответствующее число раз, получаем представление с требуемым в лемме свойством единственности. □

5.3.2.4 Представление многочленов из $I_i \setminus B_i$

Лемма 30. *Пусть $B_i = \emptyset$ для всех $i = k+1, \ldots, m$. Если $f \in I_k \setminus B_k$ и*

$$L(f) \trianglelefteq \min_{g \in B_k} L(g),$$

то существует представление

$$f = \sum_{g \in G_k} \alpha_g g, \tag{5.26}$$

такое что

$$\max_{g \in G_k}(\mathsf{HT}(\alpha_g) L(g)) = L(f) \;\text{ и }\; \max_{g \in G_k} \mathsf{HT}(\alpha_g g) \preceq \mathsf{HT}(f),$$

и, кроме того, для любого элемента $g \in G_k$, удовлетворяющего равенству $\mathsf{HT}(\alpha_g g) = \mathsf{HT}(f)$, выполняется соотношение

$$\mathsf{HT}(\alpha_g s(g)) \notin \mathsf{HT}(I_{i(g)+1}). \tag{5.27}$$

Доказательство. Существование представления (5.26), удовлетворяющего всем требованиям леммы, за исключением соотношения (5.27), доказывается также как в лемме 25. Рассмотрим такое представление.

Положим

$$J = \{g \in G_k \mid \mathsf{HT}(\alpha_g g) = \mathsf{HT}(f),\ \mathsf{HT}(\alpha_g s(g)) \in \mathsf{HT}(I_{i(g)+1})\}.$$

Если $\alpha_g g \in B_k$, то по условию $s(f) \preceq s(\alpha_g g)$, а по лемме 27 выполняется соотношение $s(\alpha_g g) \prec \mathsf{HT}(\alpha_g s(g))$. Следовательно, $s(f) \prec \mathsf{HT}(\alpha_g s(g))$, что противоречит условию

$$\max_{g \in G_k}(\mathsf{HT}(\alpha_g) L(g)) = L(f).$$

Поэтому для всех $g \in J$ выполняется включение $\alpha_g g \notin B_k$, а поскольку все множества B_i при $i = k+1,\ \ldots, m$ пустые, для каждого такого g существует представление

$$\alpha_g g = \sum_{h \in G} \nu_h h, \tag{5.28}$$

удовлетворяющее условиям

$$\max_{h \in G}(\mathsf{HT}(\nu_h) L(h)) = L(\alpha_g g) \ \text{ и } \ \max_{h \in G} \mathsf{HT}(\nu_h h) \preceq \mathsf{HT}(\alpha_g g).$$

Согласно лемме 27, для каждого $g \in J$ выполняется неравенство

$$L(\alpha_g g) \lhd \mathsf{HT}(\alpha_g) L(g).$$

Поэтому, преобразуя соотношение (5.26) по всем $g \in J$, получим в итоге требуемое представление. □

5.3.2.5 Доказательство теоремы 47

Согласно лемме 25 достаточно доказать, что $B_i = \emptyset$ при $i = 1,\ \ldots, m$.

Утверждение выполняется при $i = m$ по определению множества B_m. Предположим, что $B_i = \emptyset$ для $i = k+1,\ \ldots, m$. Докажем, что $B_k = \emptyset$.

Предположим, что множество B_k не пусто. Тогда M_k также не пусто. По лемме 29 для многочлена $f \in M_k$ существует представление

$$f \equiv \sum_{g \in G_k \setminus G_{k+1}} \gamma_g g \ (\mathsf{mod}\, I_{k+1}),$$

такое что существует единственный многочлен $g \in G_k \setminus G_{k+1}$, для которого выполняются равенства $\mathrm{HT}(\gamma_g s(g)) = s(f)$ и $\mathrm{HT}(\gamma_g g) = m(f) \succ \mathrm{HT}(f)$. Обозначим его через g_1. Согласно лемме 26 выполняется соотношение $s(f) \notin \mathrm{HT}(I_{k+1})$, поэтому $\mathrm{HT}(\gamma_g s(g)) \notin \mathrm{HT}(I_{k+1})$.

Определим множество $H = \{g \in G_k \setminus G_{k+1} \mid \gamma_g s(g) = s(f)\}$, и рассмотрим многочлен $h = f - \sum_{g \in H} \gamma_g g$. По определению выбранного представления многочлена f выполняются соотношения

$$\mathrm{HT}(h) = m(f) \text{ и } L(h) \trianglelefteq (k, \max_{g \in G_k \setminus G_{r+1} \setminus H} \gamma_g g) \triangleleft (k, s(f)) = L(f), \quad (5.29)$$

т.е. $\mathrm{HT}(h) = m(f)$ и $L(h) \triangleleft L(f)$. Поэтому из определений множеств B_k, N_k и M_k следует, что многочлен h удовлетворяет условиям леммы 30. Следовательно существует представление $h = \sum_{g \in G_k} \beta_g g$, такое что

$$\max_{g \in G_k}(\mathrm{HT}(\beta_g) L(g) = L(h) \triangleleft (k, s(f)) \text{ и } \max_{g \in G_k} \mathrm{HT}(\beta_g g) \preceq \mathrm{HT}(h) = m(f), \quad (5.30)$$

и, кроме того, для любого элемента $g \in G_k$, удовлетворяющего равенству $\mathrm{HT}(\beta_g g) = \mathrm{HT}(h)$, выполняется соотношение $\mathrm{HT}(\beta_g s(g)) \notin \mathrm{HT}(I_{i(g)+1})$.

Рассмотрим представление многочлена f в виде суммы полученных выше представлений

$$f = \left(f - \sum_{g \in H} \gamma_g g\right) + \sum_{g \in H} \gamma_g g = \sum_{g \in G_k} \beta_g g + \sum_{g \in H} \gamma_g g = \sum_{g \in G_k} \alpha_g g. \quad (5.31)$$

Из соотношений (5.29) и (5.30) следует, что только для $g = g_1$ выполняются условия $\alpha_g s(g) = s(f)$ и $\mathrm{HT}(\alpha_g g) = m(f)$. С другой стороны, поскольку $m(f) \succ \mathrm{HT}(f)$, выполняется соотношение $\mathrm{HT}(h) = m(f)$ и, следовательно, существует $g_2 \in G_k \setminus H$, такой что $\mathrm{HT}(\alpha_{g_2} g_2) = m(f) = \mathrm{HT}(h)$ и $\mathrm{HT}(\alpha_{g_2}) L(g_2)) \triangleleft (k, s(f)) = \mathrm{HT}(\alpha_{g_1}) L(g_1)$. Поэтому $g_2 \neq g_1$ и выполняется соотношение $\mathrm{HT}(\alpha_{g_2} s(g_2)) \notin \mathrm{HT}(I_{i(g_2)+1})$. Итак, доказано, что

$$\mathrm{HT}(\alpha_{g_2}) L(g_2) \triangleleft (\mathrm{HT}(\alpha_{g_1}) L(g_1), \quad (5.32)$$

и

$$\mathrm{HT}(\alpha_{g_j} s(g_j)) \notin \mathrm{HT}(I_{i(g_j)+1}), \ \ j = 1, 2. \quad (5.33)$$

Так как $\mathsf{HT}(\alpha_{g_1} g_1) = \mathsf{HT}(\alpha_{g_2} g_2) = m(f)$, существует такой моном μ, для которого $\mu u_{g_1,g_2}$ является старшим мономом многочлена α_{g_1}, а $\mu u_{g_2,g_1}$ является старшим мономом многочлена α_{g_2} т.е.

$$\mathsf{HT}(\alpha_{g_1} - \mu u_{g_1,g_2}) \prec \mathsf{HT}(\alpha_{g_1}). \tag{5.34}$$

и

$$\mathsf{HM}(\alpha_{g_1}) = \mathsf{HM}(\mu u_{g_1,g_2}) \neq 0 \ \text{ и } \ \mathsf{HM}(\alpha_{g_2}) = \mathsf{HM}(\mu u_{g_2,g_1}) \neq 0. \tag{5.35}$$

Из соотношений (5.32) и (5.33) теперь следует, что

$$u_{g_2,g_1} L(g_2) \lhd u_{g_1,g_2} L(g_1), \quad u_{g_j,g_{3-j}} s(g_j) \notin \mathsf{HT}(I_{i(g_j)+1}), \ \ j = 1, 2.$$

Поэтому согласно условию теоремы

$$S(g_1, g_2) = o_G(u_{g_1,g_2} g_1),$$

т.е. существует такое представление

$$S(g_1, g_2) = \sum_{g \in G_k} \nu_g g \tag{5.36}$$

что для всех $g \in G_k$ выполняются неравенства

$$\mathsf{HT}(\nu_g g) \prec \mathsf{HT}(u_{g_1,g_2} g_1) \ \text{ и } \ L(\nu_g g) \trianglelefteq L(u_{g_1,g_2} g_1).$$

Но ввиду того, что

$$L(u_{g_1,g_2} g_1) \trianglelefteq u_{g_1,g_2} L(g_1),$$

для всех $g \in G_k$ справедливо

$$\mathsf{HT}(\nu_g g) \prec \mathsf{HT}(u_{g_1,g_2} g_1) \ \text{ и } \ L(\nu_g g) \trianglelefteq u_{g_1,g_2} L(g_1). \tag{5.37}$$

Тогда из (5.35) и (5.37) следует, что для всех $g \in G_k$ выполняются соотношения

$$\mathsf{HT}(\mu \nu_g g) \prec \mathsf{HT}(\alpha_{g_1} g_1) = m(f) \ \text{ и } \ L(\mu \nu_g g) \trianglelefteq \mathsf{HT}(\alpha_{g_1}) L(g_1) = L(f). \tag{5.38}$$

Прибавляя многочлен $\mu \mathsf{S}(g_1, g_2)$ к правой части соотношения (5.31), а затем вычитая его, получим, используя равенство (5.36), представление

$$f = (\alpha_{g_1} - \mu u_{g_1,g_2} + \mu\nu_{g_1})g_1 + (\alpha_{g_2} + \mu u_{g_2,g_1} + \mu\nu_{g_2})g_2 + \sum_{g \in G_k \setminus \{g_1,g_2\}} (\alpha_g + \mu\nu_g)g. \tag{5.39}$$

Из соотношений (5.34), (5.35) и (5.38), следует, что значения $L(f)$ и $m(f)$ достигаются на этом представлении.

Если $\mathsf{HT}(\alpha_g g) = m(f)$ для некоторого $g \in G_k \setminus G_{k+1}$, то из первого неравенства в формуле (5.38) следует, что $\mathsf{HT}(\mu\nu_g) \prec \mathsf{HT}(\alpha_g)$ и, таким образом,

$$\mathsf{HT}(\mu\nu_g s(g)) \prec \mathsf{HT}(\alpha_g s(g)) \preceq s(f),$$

т.е. для такого g выполняются соотношения

$$s(\mu\nu_g g) \prec s(f) \text{ и } \mathsf{HT}(\mu\nu_g g) \prec m(f).$$

В этом случае элемент $\mu\nu_g g$ можно выразить как

$$\mu\nu_g g \equiv \sum_{g \in G_k \setminus G_{k+1}} \mu_g g \,(\mathrm{mod}\, I_{k+1}), \tag{5.40}$$

где

$$\max_{g \in G_k \setminus G_{k+1}} (\mathsf{HT}(\mu_g)s(g)) \prec s(f) \text{ и } \max_{g \in G_k \setminus G_{k+1}} \mathsf{HT}(\mu_g g) \prec m(f).$$

Используя теперь в равенстве (5.39) формулы (5.40) для всех $g \in G_k \setminus G_{k+1}$ с условием $\mathsf{HT}(\alpha_g g) = m(f)$, получим такое представление многочлена f, на котором достигаются значения $s(f)$ и $m(f)$, и при этом не существует элементов $g \in G_k \setminus G_{k+1}$, на которых одновременно достигаются значения $s(f)$ и $m(f)$, (см. (5.34)), что противоречит лемме 28.

Следовательно, множество B_k пусто.

Глава 6

Использование базиса Гребнера для решения систем уравнений

Пусть k — поле, а K — его алгебраическое замыкание. Многообразие идеала $I \subset k[x_1, \ldots, x_n]$ в K^n обозначим через $V(I)$.

Задача 1. Задано конечное множество F элементов в кольце многочленов над полем и базис Гребнера G идеала (F). Определить, какой из следующих случаев имеет место:

- $V((F)) = \emptyset$, или
- $V((F)) \neq \emptyset$ и конечно, или
- $V((F)) \neq \emptyset$ и бесконечно.

Задача 2. Задано конечное множество F элементов в кольце многочленов над полем k и базис Гребнера G идеала (F). Определить, какой из следующих случаев имеет место:

- $\dim_k(F) = 0$ или
- $\dim_k(F) \neq 0$.

Задача 3. Задано конечное множество F элементов в кольце многочленов $k[x_1, \ldots, x_n]$, для которого $\dim_k(F) = 0$, и базис Гребнера G идеала (F). Найти (построить) множество $V((F))$.

Задача 4. Задано конечное множество F элементов в кольце многочленов $k[x_1, \ldots, x_n]$, для которого $\dim_k(F) = 0$, и базис Гребнера G идеала (F). Найти (построить) множество решений идеала (F) в основном поле, т.е. множество $V((F)) \cap k^n$.

Для описания алгоритмов решения поставленных задач потребуются некоторые утверждения о свойствах базисов Гребнера идеалов размерности ноль.

6.1 Идеалы нулевой размерности

Лемма 31. *Многообразие $V(I)$ решений идеала $I \subset k[x_1, \ldots, x_n]$ конечно тогда и только тогда, когда величина* $\dim_k k[x_1, \ldots, x_n]/I$ *конечна.*

Доказательство. **Необходимость.** Если решений нет, то согласно теореме 35 идеал I совпадает с кольцом многочленов $k[x_1, \ldots, x_n]$, и в этом случае $\dim_k k[x_1, \ldots, x_n]/I = 0$.

Пусть теперь множество $V(I)$ непусто, конечно и $(\lambda_{i,1}, \ldots, \lambda_{i,n})$, при $i = 1, \ldots, m$, — все его элементы. Поскольку $\lambda_{i,j}$ принадлежат алгебраическому замыканию поля k, то для каждого такого $\lambda_{i,j}$ существует многочлен $p_{\lambda_{i,j}}(x) \in k[x]$ с корнем $\lambda_{i,j}$. Тогда многочлены

$$p_j(x_j) = \prod_{i=1}^{m} p_{\lambda_{i,j}}(x_j) \in k[x_1, \ldots, x_n], \quad j = 1, \ldots, n,$$

обращаются в ноль на всех решениях идеала, и, следовательно, по теореме Гильберта о нулях существуют такие $k_j \in \mathbb{N}$, что $p_j^{k_j} \in I$. Поэтому $(p_1^{k_1}, \ldots, p_n^{k_n}) \subset I$ и

$$\dim_k k[x_1, \ldots, x_n]/I \leq \dim_k k[x_1, \ldots, x_n]/(p_1^{k_1}, \ldots, p_n^{k_n}) = \prod_{j=1}^{n}(m_j \cdot k_j + 1),$$

где $m_j = \deg p_j$

Достаточность. Пусть теперь $\dim_k k[x_1, \ldots, x_n]/I$ конечна. Если эта размерность нулевая, то $I = k[x_1, \ldots, x_n]$ и, следовательно, множество решений пусто, т.е. конечно.

Рассмотрим случай когда размерность $\dim_k k[x_1, \ldots, x_n]/I$ конечна, но не равна нулю. Рассмотрим базис Гребнера G идеала I для лексикографического порядка на множестве многочленов. Рассмотрим многочлены вида $f(x_1)$. По определению лексикографического порядка, любой терм, в который входит хотя бы одна из переменных $x_2, \ldots, x_n$, старше любого терма вида x_1^k. Поэтому существует многочлен $p_1(x_1)$, принадлежащий базису Гребнера G идеала I (в противном случае $\dim_k k[x_1, \ldots, x_n]/I$ была бы бесконечной), а следовательно, и самому идеалу I. Аналогично для всех остальных переменных существуют $p_i(x_i) \in I$. Тогда все решения идеала I лежат в произведении

$$X_1 \times \ldots \times X_n \subset K^n,$$

где $X_i \subset K$ — множества всех решений уравнений $p_i(x) = 0$, т.е. в конечном множестве. □

Лемма 32. *Размерность идеала $I \subset k[x_1, \ldots, x_n]$ равна нулю тогда и только тогда, когда многообразие $V(I)$ конечно и непусто.*

Доказательство. Пусть $I = [I_1, \ldots, I_s]$ — неприводимое представление идеала I примарными идеалами и $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ — ассоциированные с этим представлением простые идеалы (см. теоремы 13 и 14). Размерность идеала I, по определению, равна максимальной из размерностей идеалов $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$.

Предположим, что $\dim_k I = 0$. Тогда для всех $i = 1, \ldots, s$ выполнено $\dim \mathfrak{p}_i = 0$. Непосредственно из определений следует, что $V(\mathfrak{p}_i) = V(I_i)$ и, следовательно, множества $V(I_i)$ конечны. Поэтому и множество решений $V(I) = V(I_1) \cup \ldots \cup V(I_s)$ конечно.

Пусть теперь множество $V(I)$ конечно. Тогда и все $V(\mathfrak{p}_i)$ конечны. Достаточно проверить, что для простого идеала $\mathfrak{p}$ размерности, большей нуля, множество $V(\mathfrak{p})$ бесконечно. Для этого, согласно лемме 31, достаточно убедиться, что $\dim_k k[x_1, \ldots, x_n]/\mathfrak{p}$ бесконечна. Пусть $(\xi_1, \ldots, \xi_n) \in \Omega^n$ — общий корень идеала $\mathfrak{p}$. Тогда определены вложения

$$k \subset k[\xi_1, \ldots, \xi_n] \subset k(\xi_1, \ldots, \xi_n) \subset \Omega^n.$$

Поскольку $\dim \mathfrak{p} > 0$, компоненты общего корня этого идеала содержат трансцендентные элементы. Без ограничения общности можно считать,

что ξ_1 трансцендентен. Тогда элементы $\xi_1, \xi_1^2, \ldots, \xi_1^m, \ldots$ линейно независимы над k. Следовательно, размерность пространства $k[\xi_1, \ldots, \xi_n]$ бесконечна, а поскольку в силу теоремы 28 выполняется равенство $k[\xi_1, \ldots, \xi_n] = k[x_1, \ldots, x_n]/\mathfrak{p}$, $\dim_k k[x_1, \ldots, x_n]/\mathfrak{p}$ бесконечна.

□

Лемма 33. *Пусть G — базис Гребнера идеала I. Отображение*

$$\pi : k[x_1, \ldots, x_n]/I \to k[x_1, \ldots, x_n],$$

заданное формулой $f + I \mapsto h$, где $f \to_{G^*} \underline{h}$, определено корректно, взаимно однозначно и является гомоморфизмом векторных k-пространств.

Доказательство. Согласно следствию 18 формула $f \to_{G^*} \underline{h}$ определяет гомоморфизм векторных k-пространств

$$\varphi : k[x_1, \ldots, x_n] \to k[x_1, \ldots, x_n],$$

ядром которого является идеал I. Следовательно, π определено корректно и является мономорфизмом. □

Следствие 20. *Пусть G — базис Гребнера идеала I. Размерность идеала I равна нулю тогда и только тогда, когда множество неприводимых относительно базиса Гребнера G термов конечно.*

Теорема 48. *Пусть G — базис Гребнера собственного идеала кольца многочленов $k[x_1, \ldots, x_n]$. Этот идеал имеет размерность 0 тогда и только тогда, когда для любого допустимого порядка при каждом $1 \leq i \leq n$ существует многочлен $g_i \in G$ со старшим термом $x_i^{\nu_i}$, где ν_i — некоторое неотрицательное целое число.*

Доказательство. Если для некоторого i не существует элемента базиса Гребнера идеала I со старшим термом $x_i^{\nu_i}$, то термы вида x_i^m неприводимы. Следовательно, множество неприводимых термов бесконечно, и поэтому, согласно следствию 20, размерность $\dim_k I \neq 0$.

Пусть базис Гребнера идеала I содержит многочлены g_i, старшими термами которых являются $x_i^{\nu_i}$. В этом случае необходимым условием неприводимости терма t является выполнение соотношений $\deg_{x_i} t < \nu_i$ для всех $i = 1, \ldots, n$. Поэтому мощность множества неприводимых термов не превосходит величину $\nu_1 \cdot \ldots \cdot \nu_n$ и, следовательно, конечна. А тогда, согласно следствию 20, размерность идеала I равна нулю. □

6.2 Решение систем уравнений

Решение задачи 1. Согласно теореме Гильберта о нулях, условие $V((F)) = \emptyset$ эквивалентно условию $1 \in (F)$, или эквивалентно $1 \in G$.

Пусть теперь $1 \notin G$. Тогда $V((F)) \neq \emptyset$. Вопрос о конечности или бесконечности многообразия $V((F))$ решается теперь теоремой 48 и леммой 32. Достаточно проверить, содержит ли базис Гребнера G многочлены со старшими термами $x_i^{\nu_i}$ при всех $i = 1, \ldots, n$.

Задача 2 полностью решается в теореме 48.

Для решения **задачи 3** достаточно решить следующую задачу.

Задача нахождения хотя бы одного решения идеала нулевой размерности, если задан приведенный базис Гребнера этого идеала относительно лексикографического порядка (Задача 3а). Предположим, что задан базис Гребнера $\{g_1, \ldots g_m\}$ идеала $I \subset k[x_1, \ldots, x_n]$ нулевой размерности. Пусть также имеется оракул $\mathcal{A}$, решающий задачу нахождения корня любого многочлена от одной переменной над K, где K — алгебраическое замыкание поля k.

Решение задачи 3а.

Пусть $I \subset k[x_1, \ldots, x_n]$ — идеал и G — приведенный базис Гребнера этого идеала относительно лексикографического порядка. Тогда пересечение $G \cap k[x_1]$ состоит в точности из одного многочлена $f(x_1) \in k[x_1]$ и является базисом Гребнера идеала $I_1 = I \cap k[x_1]$ кольца $k[x_1]$. Находим с помощью оракула $\mathcal{A}$ решение $\xi_1 \in K$ уравнения $f(x_1) = 0$. Заметим, что для любого решения $(x_1^0, \ldots, x_n^0)$ идеала I выполняется соотношение $f(x_1^0) = 0$.

Предположим теперь, что найдено решение $(\xi_1, \ldots, \xi_i)$ идеала $I_i = I \cap k[x_1, \ldots, x_i]$. Найдем решение $(\xi_1, \ldots, \xi_i, \xi_{i+1})$ идеала I_{i+1}. Для этого найдем элементы базиса Гребнера G, находящиеся в кольце $k[x_1, \ldots, x_{i+1}]$. Подставим в полученные многочлены от переменных $x_1, \ldots, x_{i+1}$ значения первых i переменных: $x_1 = \xi_1, \ldots, x_i = \xi_i$. Получим набор многочленов, зависящих только от одной переменной x_{i+1}. Вычислим их наибольший общий делитель $g(x_{i+1})$. Как будет показано ниже, полученный многочлен имеет степень не менее единицы и, следовательно, имеет непустое множество решений в алгебраическом замыкании поля k. Находим с помощью оракула $\mathcal{A}$ решение $\xi_{i+1} \in K$ уравнения $g(x_{i+1}) = 0$. Тогда вектор

$(\xi_1, \ldots, \xi_i, \xi_{i+1})$ является решением идеала I_{i+1}.

Далее повторяем описанную процедуру до тех пор, пока не найдем полный вектор решения идеала I.

Ниже приведен алгоритм для описанной процедуры нахождения решения алгебраической системы уравнений.

Алгоритм А. Дано: Базис Гребнера $G = \{g_1, \ldots g_m\}$ относительно лексикографического порядка идеала $I \subset k[x_1, \ldots, x_n]$ нулевой размерности.

Выход: Точка $(x_1^0, \ldots, x_n^0) \in K^n$, где K — алгебраическое замыкание поля k.

Шаг 1 $i := 1$.

Шаг 2 Находим пересечение $G_1 = G \cap k[x_1]$, состоящее в точности из одного многочлена $g(x_1)$.

Шаг 3 $x_i^0 := \mathcal{A}(g(x_i))$ — некоторое решение уравнения $g(x_i) = 0$.

Шаг 4 $i := i + 1$.

Шаг 5 Если $i > n$, перейти к шагу 10.

Шаг 6 Находим пересечение $G_i = G \cap k[x_1, \ldots, x_i]$.

Шаг 7 $G_i(x_1^0, \ldots, x_{i-1}^0) := \{g(x_1^0, \ldots, x_{i-1}^0, x_i) | g \in G_i\} \subset K[x_i]$.

Шаг 8 Находим $g(x_i)$ — наибольший общий делитель элементов множества $G_i(x_1^0, \ldots, x_{i-1}^0)$.

Шаг 9 Переходим к шагу 3.

Шаг 10 Выход: Точка $(x_1^0, \ldots, x_n^0) \in K^n$.

Теорема 49. *Для любого идеала размерности ноль алгоритм А находит некоторое его решение.*

Для доказательства теоремы достаточно показать, что на шаге 8 приведенного выше алгоритма всегда получаем многочлен степени не меньше единицы, или, эквивалентно, каждое решение $(\xi_1, \ldots, \xi_i) \in K^i$ идеала

$I \cap k[x_1, \ldots, x_i]$ продолжается до решения $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ идеала $I \cap k[x_1, \ldots, x_{i+1}]$. Доказательство этого утверждения потребует нескольких вспомогательных утверждений.

Для любого подмножества M кольца $k[x_1, \ldots, x_n]$ и любого $0 < i \leq n$ будем использовать следующее **обозначение**: $M_i = M \cap k[x_1, \ldots, x_i]$.

Напомним также, что Ω — алгебраическое замыкание поля рациональных функций $k(t_1, \ldots)$ от бесконечного множества алгебраически независимых переменных.

Лемма 34. *Если G — приведенный базис Гребнера относительно лексикографического порядка идеала $I \subset k[x_1, \ldots, x_n]$, то G_i — приведенный базис Гребнера идеала I_i в кольце многочленов от переменных $x_1, \ldots, x_i$ относительно лексикографического порядка на термах.*

Доказательство. Следует из определения 4.1.1 базиса Гребнера. □

Лемма 35. *Пусть $I \subset k[x_1, \ldots, x_n]$ — простой идеал. Тогда для любого $0 < i \leq n$ идеал I_i простой.*

Доказательство. Следует из определения 2.1.8 простого идеала. □

Лемма 36. *Пусть $I \subset k[x_1, \ldots, x_n]$ — примарный идеал. Тогда для любого $0 < i \leq n$ идеал I_i примарный.*

Доказательство. Следует из определения 2.2.4 примарного идеала. □

Лемма 37. *Пусть $I \subset k[x_1, \ldots, x_n]$ — примарный идеал и J — ассоциированный с ним простой идеал. Тогда множества корней идеалов I и J совпадают.*

Доказательство. Следует из определения 2.2.5 простого идеала ассоциированного с примарным идеалом. □

Лемма 38. *Пусть $I \subset k[x_1, \ldots, x_n]$ — примарный идеал и J — ассоциированный с ним простой идеал. Тогда простой идеал J_i ассоциирован с идеалом I_i.*

Доказательство. Следует из лемм 35 и 36 и определения 2.2.5 простого идеала ассоциированного с примарным идеалом. □

Лемма 39. *Если $I \subset k[x_1, \ldots, x_n]$ — идеал размерности 0, то I_m является идеалом размерности 0 в кольце многочленов $k[x_1, \ldots, x_m]$.*

Доказательство. Пусть G — приведенный базис Гребнера относительно лексикографического порядка идеала I размерности ноль. Согласно теореме 48, для всех $i = 1, \ldots, n$ существуют многочлены $g_i \in G$ со старшими термами $x_i^{\nu_i}$. Тогда $g_i \in G_i = G \cap k[x_1, \ldots, x_i]$ и для любого $i \leq m$ элементы $g_i \in I_m$ при $i = 1, \ldots, m$. Поскольку старший терм g_i равен $x_i^{\nu_i}$, то, по теореме 48, $\dim_k I_m = 0$. □

Следствие 21. *Пусть G — базис Гребнера относительно лексикографического порядка идеала $I \subset K[x_1, \ldots, x_n]$ размерности 0. Тогда G_1 состоит в точности из одного многочлена $f \in k[x_1]$ положительной степени.*

Лемма 40. *Пусть $I \subset k[x_1, \ldots, x_n]$ — простой идеал размерности 0 и $(\xi_1, \ldots, \xi_i) \in K^i$ — корень идеала I_i при некотором $1 \leq i \leq n-1$. Тогда существует $\xi_{i+1} \in K$, такой, что $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ — корень идеала I_{i+1}.*

Доказательство. Поскольку идеал I — размерности 0, он не совпадает со своим кольцом и, следовательно, по теореме Гильберта имеет некоторый корень $(\omega_1, \ldots, \omega_n)$ в K^n. В частности, $(\omega_1, \ldots, \omega_i) \in K^i$, так же как и $(\xi_1, \ldots, \xi_i)$, является корнем идеала I_i. Поскольку согласно леммам 35 и 39 идеал I_i прост и имеет нулевую размерность, из теоремы 32 следует, что эти корни сопряжены, т.е. имеется изоморфизм подполей поля K

$$\varphi : k(\omega_1, \ldots, \omega_i) \to k(\xi_1, \ldots, \xi_i),$$

заданный соответствиями $\omega_j \mapsto \xi_j$ для всех $j = 1, \ldots, i$.

Пусть G — базис Гребнера идеала I и многочлен $h \in k(\omega_1, \ldots, \omega_i)[x_{i+1}]$ является наибольшим общим делителем многочленов $f(x_{i+1}) = g(\omega_1, \ldots, \omega_i, x_{i+1})$, где g пробегает G_{i+1}. Поскольку $(\omega_1, \ldots, \omega_i, \omega_{i+1})$ является корнем идеала I_{i+1}, элемент ω_{i+1} удовлетворяет соотношению $h(\omega_{i+1}) = 0$. Верно и обратное, для любого корня ζ уравнения $h(x) = 0$, точка $(\omega_1, \ldots, \omega_i, \zeta)$ также корень идеала I_{i+1}, а поскольку размерность простого идеала I_{i+1} равна нулю, эта точка также

является общим корнем этого идеала. Поскольку число корней идеала размерности ноль конечно, число решений уравнения $h(x_{i+1}) = 0$ не пусто и конечно. Поэтому степень многочлена h положительна.

Обозначим через $\tilde{h}$ наибольший общий делитель многочленов $f(x_{i+1}) = g(\xi_1, \ldots, \xi_i, x_{i+1})$, где g пробегает G_{i+1}. Тогда согласно определению изоморфизма φ выполняется соотношение $\varphi^*(h) = \tilde{h}$[1] и степени многочленов h и $\tilde{h}$ совпадают. Следовательно, степень многочлена $\tilde{h}$ положительна. Поэтому уравнение $\tilde{h}(x_{i+1}) = 0$ всегда разрешимо в K. Пусть его решение ξ_{i+1}. Тогда $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ — корень идеала I_{i+1}. □

Лемма 41. *Пусть $(\xi_1, \ldots, \xi_n) \in \Omega^n$ — общий корень простого идеала I. Тогда $(\xi_1, \ldots, \xi_i) \in \Omega^i$ — общий корень простого идеала I_i.*

Доказательство. Обозначим через J_i — простой идеал, состоящий из всех многочленов кольца $k[x_1, \ldots, x_i]$, обращающихся в ноль в точке $(\xi_1, \ldots, \xi_i)$. Согласно определению 3.2.20 точка $(\xi_1, \ldots, \xi_i)$ является общим корнем идеала J_i. Очевидно, что $J_i \supset I_i$, а поскольку точка $(\xi_1, \ldots, \xi_n) \in \Omega^n$ является общим корнем идеала I, из определения 3.2.20 следует, что $J_i \subset I$. Следовательно, $J_i \subset I_i$. Поэтому $J_i = I_i$, и $(\xi_1, \ldots, \xi_i)$ — общий корень идеала I_i. □

Лемма 42. *Пусть $I \subset k[x_1, \ldots, x_n]$ — примарный идеал размерности 0 и $(\xi_1, \ldots, \xi_i) \in K^i$ — корень идеала I_i. Тогда существует $\xi_{i+1} \in K$, такой, что $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ — корень идеала I_{i+1}.*

Доказательство. Следует из лемм 36, 39, 37, 40. □

Напомним, что $V(I)$ — многообразие идеала $I \subset k[x_1, \ldots, x_n]$.

Лемма 43. *Если идеал I является пересечением идеалов $\mathfrak{q}_j$, где j пробегает целые числа от 1 до m, то $V(I) = \bigcup_{j=1}^{m} V(\mathfrak{q}_j)$.*

[1]Напомним, что изоморфизм полей $\varphi : k_1 \to k_2$ задает изоморфизм колец многочленов $\varphi^* : k_1[x] \to k_2[x]$

Доказательство. Пусть $\xi \in \bigcup\limits_{j=1}^{m} V(\mathfrak{q}_j)$. Тогда при некотором $1 \leq j \leq m$ выполняется $\xi \in V(\mathfrak{q}_j)$. А поскольку $I \subset \mathfrak{q}_j$, верно, что $\xi \in V(I)$. Следовательно, $V(I) \supset \bigcup\limits_{j=1}^{m} V(\mathfrak{q}_j)$.

Пусть теперь $\xi \notin \bigcup\limits_{j=1}^{m} V(\mathfrak{q}_j)$. Тогда для всех $j = 1, \ldots, m$ существуют $p_j \in \mathfrak{q}_j$, для которых $p_j(\xi) \neq 0$. Поскольку все $\mathfrak{q}_j$ являются идеалами, произведение $p = \prod\limits_{s=1}^{m} p_s$ является их общим элементом и, следовательно, принадлежит идеалу I. Элемент ξ не принадлежит многообразию $V(I)$, поскольку выполняется $p(\xi) = \prod\limits_{s=1}^{m} p_s(\xi) \neq 0$. Следовательно, $V(I) \subset \bigcup\limits_{j=1}^{m} V(\mathfrak{q}_j)$. □

Лемма 44. *Пусть $I \subset k[x_1, \ldots, x_n]$ — идеал нулевой размерности и $(\xi_1, \ldots, \xi_i) \in K^i$ — корень идеала I_i. Тогда существует $\xi_{i+1} \in K$, такой, что $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ — корень идеала I_{i+1}.*

Доказательство. Согласно теореме 13 существует представление идеала I в виде пересечения конечного множества примарных компонент

$$I = \bigcap_{j=1}^{m} \mathfrak{q}_j.$$

Поскольку I — идеал размерности ноль, то и все идеалы $\mathfrak{q}_j$ — также нулевой размерности (см. определение 3.2.25). Очевидно, выполняется равенство

$$I_i = \bigcap_{j=1}^{m} \mathfrak{q}_{j,i}, \tag{6.1}$$

где $\mathfrak{q}_{j,i} = \mathfrak{q}_j \cap k[x_1, \ldots, x_i]$ — примарные идеалы размерности ноль. Поэтому, согласно лемме 43, для всех $i = 1, \ldots, n$ имеет место разложение

$$V(I_i) = \bigcup_{j=1}^{m} V(\mathfrak{q}_{j,i}). \tag{6.2}$$

Поскольку $(\xi_1, \ldots, \xi_i) \in K^i$ — корень идеала I_i, то из представления (6.2) следует, что при некотором j элемент $(\xi_1, \ldots, \xi_i) \in K^i$ является корнем идеала $\mathsf{q}_{j,i}$. Поэтому по лемме 42, существует $\xi_{i+1} \in K$ такой, что $(\xi_1, \ldots, \xi_{i+1}) \in K^{i+1}$ — корень идеала $\mathsf{q}_{j,i+1}$, а следовательно, согласно формулам (6.1) и (6.2), является корнем идеала I_{i+1}. □

Определение 6.2.1. *Размерностью системы алгебраических уравнений называется размерность соответствующего идеала этой системы. Систем алгебраических уравнений размерности ноль будем называть полными.*

Лемма 45. *Пусть $p(x) \in \mathbb{Q}[x]$. Тогда уравнение $p(x) = 0$ алгоритмически разрешимо в поле рациональных чисел.*

Доказательство. Рассмотрим уравнение

$$x^n + b_{n-1}x^{n-1} + \ \ldots + b_0 = 0.$$

Выполняется очевидное ограничение на его корни:

$$|x| \leq 1 + \mathsf{max}\{|b_{n-1}, \ \ldots, |b_0|\} = 1 + \lambda.$$

Представим ненулевые рациональные коэффициенты уравнения в виде отношения взаимно простых чисел: $b_i = \frac{p_i}{q_i}$, $i = 0, \ \ldots, n-1$. Тогда для рациональных решений $x = \frac{p}{q}$ выполняются соотношения $q_n \vdots q$, где

$$q_n = \mathsf{НОК}(q_0, \ \ldots, q_{n-1}).$$

Следовательно, все рациональные решения могут быть получены процедурой перебора из конечного множества размера не более $2(1+\lambda)q_n$. □

Следствие 22. *Задача нахождения рационального решения полной системы алгебраических уравнений над полем рациональных чисел является алгоритмически разрешимой.*

Доказательство. Поскольку задача нахождения базиса Гребнера идеала I относительно лексикографического порядка по теореме 45 является алгоритмически разрешимой, а по теореме 48 для идеала размерности

ноль для каждой переменной x_i существуют многочлены $f_i(x_i)$, зависящие только от этой переменной и принадлежащие идеалу I, то задача построения таких многочленов алгоритмически разрешима. Пусть X_i — множество рациональных решений уравнения $f_i(x_i) = 0$. Согласно лемме 45 конечные множества X_i могут быть построены алгоритмически. Поскольку все рациональные решения системы идеала I принадлежат произведению $X_1 \times \ \ldots \times X_n$, то, используя процедуру перебора, можно найти все решения идеала размерности ноль, а следовательно, полной системы алгебраических уравнений над полем рациональных чисел. □

6.3 Нахождение решений в основном поле

Используя алгоритм А, если дан приведенный базис Гребнера относительно лекисикографического порядка, можно найти все решения системы алгебраических уравнений размерности ноль, а затем из этого конечного множества найти решение, принадлежащее основному полю, либо доказать, что решения нет. Однако, такой подход может оказаться весьма трудоемким. Проиллюстрируем это следующим примером.

Рассмотрим систему алгебраических уравнений

$$\left\{\begin{array}{lcl} x_1^2 - x_1 & = & 0, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots & & \ldots \\ x_{n-2}^2 - x_{n-2} & = & 0, \\ (n-2)x_{n-1}^2 - x_{n-1}^2 x_{n-2} - \ \ldots - x_{n-1}^2 x_1 + x_{n-1} + 1 & = & 0, \\ x_n + \ \ldots + x_1 - n + 2 & = & 0. \end{array}\right. \tag{6.3}$$

Эта система уравнений имеет единственное вещественное решение

$$\left\{\begin{array}{lcl} x_1 & = & 1, \\ \ldots & & \ldots \\ x_{n-2} & = & 1, \\ x_{n-1} & = & -1, \\ x_n & = & 1, \end{array}\right. \tag{6.4}$$

Это решение будет получено алгоритмом А, только в том случае, если для

первых $(n-2)$ уравнений оракул укажет в качестве решения

$$\begin{cases} x_1 & = & 1, \\ \dots & \dots & \\ x_{n-2} & = & 1. \end{cases} \tag{6.5}$$

Отметим, что общее число различных решений системы (6.3) в $\mathbb{C}$ равно $2^{n-1}-1$. Базис Гребнера относительно лексикографического порядка совпадает с левыми частями частями уравнений, т.е. имеет ту же сложность, что и описание идеала.

Для конечных полей задача нахождения решения идеала I в поле $\mathbb{F}_q$, где $q = p^n$ и p — некоторое простое число, сводится к нахождению решения идеала J, полученного добавлением многочленов $x_i^q - x_i$ для всех $i = 1, \dots, n$. В этом случае все решения идеала J принадлежат полю $\mathbb{F}_q$, а также все решения идеала I, принадлежащие полю $\mathbb{F}_q$, являются решениями идеала J.

Решение задачи 4 для конечного поля $k = \mathbb{F}_q$ можно получить с помощью следующего алгоритма.

Алгоритм В. Нахождение решения в основном поле.

Вход: Базис M идеала I в кольце $\mathbb{F}_q[x_1, \dots, x_n]$.

Выход: Решение $(\xi_1, \dots, \xi_n)$ в основном поле $\mathbb{F}_q$ или доказательство несуществования такого решения.

Шаг 1 Строим базис идеала $J = I \cup (x_1^q - x_1, \dots, x_n^q - x_n)$: $F \cup \{x_1^q - x_1, \dots, x_n^q - x_n\}$.

Шаг 2 Строим базис Гребнера G идеала J относительно лексикографического порядка на термах.

Шаг 3 Если $1 \in G$, то решений в основном поле нет и переходим к шагу 5.

Шаг 4 Применяем алгоритм А нахождения решения.[2]

Шаг 5 Завершение алгоритма.

[2]Поскольку идеал J содержит многочлены $x_i^q - x_i$, решения идеала лежат в основном поле F_q.

Определение 6.3.1. *Идеал I в кольце многочленов $k[x_1, \ldots, x_n]$ над полем k называется булевым, если $x_i^2 - x_i \in I$ для любого $0 < i \leq n$. Система уравнений, определяющая такой идеал, называется булевой.*

Алгоритм B может быть применен также для булевых идеалов для колец многочленов над любым полем.

Теорема 50. *Если булева система уравнений над полем k имеет единственное решение, то приведенный базис Гребнера идеала этой системы с точностью до умножения на ненулевые множители не зависит от выбора допустимого порядка на множестве мономов.*

Доказательство. В силу соотношений $x_i^2 = x_i$ все x_i принимают значения в множестве $\{0, 1\}$, и , следовательно, решения, принадлежащие алгебраическому замыканию поля k, лежат в поле k. Пусть единственным решением булевой системы уравнений является набор $x_i = a_i$, $i = 1, \ldots, n$. Докажем вначале, что идеал

$$J = (x_1 - a_1, \ldots, x_n - a_n)$$

совпадает в этом случае с идеалом I, заданным исходной булевой системой уравнений. Действительно, поскольку $x_i - a_i$ обращается в ноль во всех решениях исходной системы (а оно одно), согласно теореме Гильберта о нулях идеала при некотором натуральном m многочлен $(x_i - a_i)^m \in I$, а тогда и сам многочлен $x_i - a_i$ принадлежит этому идеалу, поскольку $x_i^2 - x_i \in I$. Следовательно, идеал J лежит в идеале I, заданном булевыми уравнениями исходной системы уравнений.

Положим $F = \{x_1 - a_1, \ldots, x_n - a_n\}$. Очевидно, что для любого многочлена $f \in I$, независимо от выбора допустимого порядка, выполняется соотношение $f \to_{F^*} \underline{a}$, где $a \in k$. Если этот элемент ненулевой, то система не имеет решения. Если этот элемент нулевой, то это означает, что $f \in J$. Поэтому $I = J$.

Пусть F_0 — приведенный базис Гребнера идеала I относительно некоторого допустимого порядка. Без ограничения общности можно считать, что x_n — наименьший моном положительной степени относительно этого порядка. Поскольку $x_n - a_n \in I$, выполняется соотношение $f \to_{F_0^*} \underline{0}$. Поэтому существует многочлен $h \in F_0$, старший моном которого делит

моном x_n. Поскольку моном x_n — наименьший, этот моном делится на старший моном многочлена h. Следовательно, $x_n - a_n \in F_0$ или $h = 1$ и $F_0 = \{1\}$. Переменная x_n не входит ни в один из оставшихся многочленов приведенного базиса Гребнера F_0. Рассуждая как выше, получаем, что в приведенный базис Гребнера входит также $x_{n-1} - a_{n-1}$, где x_{n-1} — следующая по старшинству переменная и т.д., пока не переберем все переменные. □

Глава 7

Криптоанализ

7.1 Базисы Грёбнера в криптографии

Данный раздел предполагает знакомство читателя с криптографией. Читателю, не владеющему такими знаниями, вполне доступно математическое содержание материала, но не значимость обсуждаемых исследований для криптографии.

В таких ситуациях нередко предпринимаются попытки написать для математиков краткое введение в криптографию. Но это не более осмысленно, чем пытаться создать для неспециалистов краткие изложения состояния исследований в различных математических дисциплинах.

Тем не менее заметим, что введение в криптографию можно найти в одноименной книге [**1**], а криптографические термины и определения на сайте `cryptography.ru`.

7.2 Оценка параметров

Линейный метод криптоанализа, предложенный Мацуи [**3**], предназначен для блоковых шифров типа DES. Описание шифра DES можно найти, например, в книге [**5**]. Для нас важно только, что эта криптосистема построена как композиция 16 однотипных преобразований, называемых раундовыми. Линейный криптоанализ реализует угрозу полного раскрытия (опре-

деления секретного ключа) на основе атаки с известным открытым текстом.

Основная идея метода состоит в нахождении линейных функций, приближающих раундовые преобразования шифра с вероятностью, отделенной от $\frac{1}{2}$. Одним из параметров линейного криптоанализа является требуемый объем материала, т.е. количество пар (открытый текст, шифртекст). Для шифра DES этот параметр имеет порядок 2^{43}. В работе [**11**] этот параметр оценивается для модифицированного метода. Модификация состоит в замене линейных приближений квадратичными и работает, в общих чертах, следующим образом. Основу каждого раундового преобразования составляют так называемые S-боксы. Математически каждый S-бокс может быть задан системой четырех булевых функций от шести переменных:

$$S_i: \quad \begin{array}{l} y_1 = f_1^{(i)}(x_1, x_2, \ldots, x_6) \\ y_2 = f_2^{(i)}(x_1, x_2, \ldots, x_6) \\ y_3 = f_3^{(i)}(x_1, x_2, \ldots, x_6) \\ y_4 = f_4^{(i)}(x_1, x_2, \ldots, x_6) \end{array}, \quad i = 1, 2, \ldots, 8.$$

Для каждой системы S_i ($1 \le i \le 8$) авторы рассматривают идеал

$$J_i = \left(y_1 \oplus f_1^{(i)}(x_1, \ldots, x_6),\ \ldots, y_4 \oplus f_4^{(i)}(x_1, \ldots, x_6) \right)$$

в $\mathbb{Z}_2[y,\ \ldots, y_4, x_1,\ \ldots, x_6]$.

Для идеалов $J_1, J_2, \ldots, J_8$ вычисляются приведенные базисы Грёбнера. Используя эти базисы, находятся квадратичные функции в этих идеалах. Оказалось, что квадратичные функции есть в J_1 (одна функция), J_4 (пять функций) и J_5 (одна функция). Далее в [**2**] используют квадратичную функцию из J_5. Необходимо отметить, что квадратичное соотношение для раундовой функции, построенное на основе квадратичной функции из J_5, выполняется с вероятностью 1. Авторы, комбинируя полученное квадратичное соотношение для 2-го раунда с линейными соотношениями, получают оценку количества необходимых пар (открытый текст, шифртекст) для модифицированного метода, примерно равную $\frac{25}{34} \cdot 2^{43}$, что несколько меньше, чем в методе М. Мацуи.

Еще один пример применения базисов Грёбнера относится к оценке такого параметра, как нелинейность булевой функции.

Для произвольной булевой функции f от n переменных ее нелинейностью $nl(f)$ называется расстояние (по Хэммингу) до множества $\mathcal{A}_n$ аффинных булевых функций

$$nl(f) = \mathsf{dist}(f, \mathcal{A}_n) = \min_{g \in \mathcal{A}_n} \mathsf{dist}(g, f).$$

Для вычисления этого параметра известен метод, основанный на быстром преобразовании Фурье (преобразовании Адамара) с трудоемкостью $O(n2^n)$ (см. [**6**]).

В работе [**7**] для вычисления нелинейности булевой функции предлагается использовать аппарат базисов Грёбнера.

Пусть $\mathbb{Z}_2^n = \{v_1, \ldots, v_{2^n}\}$ — множество всех наборов над $\mathbb{Z}_2$ длины n, выписанных в определенном порядке, $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_{2^n}\}$ и $A = \{a_0, a_1, \ldots, a_n\}$ — булевы переменные. Обозначим через $M_{2^n,t}$, $1 \le t \le 2^n$, множество мономов из $\mathbb{Z}_2[y_1, \ldots, y_{2^n}]$, каждый из которых содержит t переменных. Для произвольной булевой функции f из $\mathbb{Z}_2[X]$ и функции $g_n = a_0 \oplus \bigoplus_{j=1}^n a_j x_j$ из $\mathbb{Z}_2[A, X]$ определим идеал в $\mathbb{Z}_2[A]$ вида

$$J_t^n(f) = \langle \{ m(g_n(A, v_1) \oplus f(v_1), \ldots, g_n(A, v_{2^n}) \oplus f(v_{2^n})) \mid m \in M_{2^n,t} \} \cup E(A) \rangle, \tag{7.1}$$

где $E(A) = \{a_0^2 \oplus a_0, a_1^2 \oplus a_1, \ldots, a_n^2 \oplus a_n\}$.

Основой для построения алгоритма, вычисляющего $nl(f)$, является следующее утверждение из работы [**7**].

Лемма 46. *Пусть $f \in \mathbb{Z}_2[x]$ и $1 \le t \le 2^n$. Тогда следующие условия эквивалентны:*

1. $\mathcal{V}(J_t^n(f)) = \{u \in \mathbb{Z}_2^n \mid h(u) = 0 \, \forall h \in J_t^n(f)\} \ne \emptyset$;
2. *существует функция $g = a_0 \oplus \bigoplus_{j=1}^n a_j x_j$ из $\mathcal{A}_n$ такая, что* $\mathsf{dist}(f, g) \le t - 1$.

Утверждение леммы почти очевидно. Действительно, выполнение условия 1 означает, что некоторая фиксация переменных $a_0, \ldots, a_n$ определяет аффинную функцию g, для которой вес вектора

$$\big(g(v_1) \oplus f(v_1), \ldots, g(v_{2^n}) \oplus f(v_{2^n})\big)$$

не превосходит $t-1$ (условие 2). Обратное утверждение очевидно.

Из леммы 46 непосредственно следует

Теорема 51. *Для произвольной функции $f \in \mathbb{Z}_2[X]$ нелинейность $nl(f)$ совпадает с минимальным t, таким, что*

$$\mathcal{V}(J_t^n(f)) \neq \emptyset.$$

Предлагаемый автором алгоритм по существу сводится к вычислению базисов Грёбнера идеалов J_t^n, $t = 1, 2, \ldots$.

7.3 Криптоанализ

Системы полиномиальных уравнений — весьма универсальный объект. Почти все вычислительные задачи, возникающие в криптографии, могут быть сформулированы как задачи решения таких систем. Однако во многих случаях оказывается, что просто выписать требуемую систему уже не просто.

Авторам работы [**9**] удалось построить систему полиномиальных уравнений для шифра AES-128 над полем $\mathbb{F}_{2^8}$. Система состоит из 200 уравнений алгебраической степени 254 и 152 линейных уравнений. Имеющиеся в настоящее время вычислительные возможности не позволяют построить базис Грёбнера для идеала, порожденного уравнениями данной системы, и использовать его для нахождения решения системы (в том числе секретного ключа AES-128).

Построенная для AES-128 относительно той же пары атака-угроза система булевых уравнений (см. [**10**]) также имеет параметры, которые не позволяют с использованием современной вычислительной техники найти соответствующий базис Грёбнера.

Важно отметить, что в настоящее время отсутствуют какие-либо теоретические результаты, позволяющие хотя бы приблизительно оценивать трудоемкость построения базиса Грёбнера исходя из параметров исследуемого идеала.

Приведем пример использования базисов Грёбнера в криптоанализе потокового шифра.

Регистром сдвига длины L с линейными обратными связями (linear feedback shift register) над полем $\mathbb{F}_2$ называют конечный автомат, вырабатывающий линейную рекуррентную последовательность порядка L. Множество внутренних состояний совпадает с $\mathbb{F}_2^L$. Начальному состоянию $U_0 = (u_0 u_1 \ldots u_{L-1}) \in \mathbb{F}_2^L$ соответствует последовательность состояний $\mathbb{U} = (U_l)_{l=0}^{\infty}$, где $U_l = (u_l u_{l+1} \ldots u_{l+L-1})$. Функция переходов регистра соответствует невырожденному линейному преобразованию:

$$A : (u_l \ldots u_{l+L-1}) \mapsto (u_{l+1} \ldots u_{l+L-1} \bigoplus_{j=1}^{L} p_j u_{l+L-j}),$$

где $p(\lambda) = 1 \oplus p_1\lambda \oplus p_2\lambda^2 \oplus \cdots \oplus p_{L-1}\lambda^{L-1} \oplus p_L\lambda^L$ — примитивный над $\mathbb{F}_2$ полином степени L.

Пусть k — фиксированное натуральное число, $k < L$, и пусть $\gamma = (\gamma_j)_{j=1}^{k}$ — набор чисел, таких, что $0 \le \gamma_1 < \gamma_2 < \ldots < \gamma_k \le L-1$.

Пусть f — фильтрующая функция, то есть булева функция $f(x_0, x_1, \ldots, x_{L-1})$, существенно зависящая только от переменных $x_{\gamma_1}, x_{\gamma_2}, \ldots, x_{\gamma_k}$.

Совокупность регистра сдвига и фильтрующей функции называют фильтрующим генератором. Выходная (шифрующая) последовательность фильтрующего генератора имеет вид

$$z_0 = f(U_0),\ z_1 = f(AU_0),\ \ldots,\ z_{N-1} = f(A^{N-1}U_0) = f(U_{N-1}),\ \ldots$$

Шифрование открытого текста $a = (a_0 a_1 \ldots a_{N-1})$ осуществляется с помощью преобразования

$$c_0 = a_0 \oplus z_0,\ c_1 = a_1 \oplus z_1,\ \ldots,\ c_{N-1} = a_{N-1} \oplus z_{N-1},$$

где $c = (c_0 c_1 \ldots c_{N-1})$ — шифрованный текст.

Ключом данного потокового шифра является начальное состояние регистра U_0. Принимающая сторона, обладая ключом U_0 и выработав последовательность $z = (z_0 z_1 \ldots z_{N-1})$, дешифрует полученное сообщение с помощью преобразования

$$a_0 = c_0 \oplus z_0,\ a_1 = c_1 \oplus z_1,\ \ldots,\ a_{N-1} = c_{N-1} \oplus z_{N-1}.$$

Рассмотрим угрозу полного раскрытия шифра (определения ключа U_0) на основе атаки с известным открытым текстом. В данном случае угрозу можно осуществить, если решить систему булевых уравнений

$$f(x) = z_0,\ f(Ax) = z_1,\ \ldots,\ f(A^{N-1}x) = z_{N-1}. \tag{7.2}$$

Будем предполагать, что данная система имеет единственное решение $x = U_0$.

Рассмотрим в кольце

$$\mathcal{R} = \mathbb{F}[x_0, x_1,\ \ldots, x_{L-1}]/(x_0^2 \oplus x_0, x_1^2 \oplus x_1,\ \ldots, x_{L-1}^2 \oplus x_{L-1})$$

идеал, ассоциированный с системой (7.2),

$$I_N = \langle z_0 \oplus f(x_0x_1 \ldots x_{L-1}),\ \ldots, z_{N-1} \oplus f(A^{N-1}(x_0x_1 \ldots x_{L-1}))\rangle.$$

Как отмечалось ранее (см. стр.) для любого упорядочивания на множестве мономов базис Грёбнера идеала I_N в нашем случае имеет вид $x_0 \oplus u_0, x_1 \oplus u_1,\ \ldots, x_{L-1} \oplus u_{L-1}$, где $U_0 = (u_0u_1 \ldots u_{L-1})$ — решение системы (7.2).

В работе [**11**] авторы используют обратный лексикографический порядок для нахождения базисов Грёбнера идеалов, ассоциированных с системами вида (7.2). В этой работе приводятся результаты компьютерных экспериментов с использованием алгоритмов F_4 и F_5. При этом успешное построение базисов Грёбнера для идеалов I_N проводится для значений параметра L (длина регистра сдвига), равных нескольким десяткам.

Кроме того, в этой же работе получены интересные теоретические результаты, связанные с трудоёмкостью решения этого класса систем булевых уравнений.

Пусть $M(L,d) = \sum\limits_{i=0}^{d} \binom{L}{i}$. Справедливо следующее утверждение.

Теорема 52. *Пусть $N \geq M(L, \lfloor \frac{k+1}{2} \rfloor)$. Тогда трудоёмкость построения базиса Грёбнера для идеала I_N равна*

$$O\left(\left(\sum_{i=0}^{\lfloor \frac{k+1}{2} \rfloor} \binom{L}{i}\right)^{\omega}\right) = O\big(L^{\lfloor \frac{k+1}{2} \rfloor \omega}\big),$$

где ω — константа линейной алгебры (показатель степени полинома в оценке трудоемкости наиболее эффективного метода решения системы линейных уравнений).

Замечание 7.3.1. Необходимо отметить, что рассматриваемые задачи криптоанализа можно решать и другими методами, например методами, основанными на линейной алгебре, математической статистике или теории кодирования. Метод, использующий базисы Грёбнера, выгодно отличается от них тем, что обладает надежностью, равной 1.
В частности, метод линеаризации системы (7.2) при $N \geq M(L, \lfloor \frac{k+1}{2} \rfloor)$ рассматривает каждый моном уравнений как новую переменную, при этом множество решений линеаризованной системы существенно зависит от ее ранга. Трудоёмкость этого метода по порядку такая же, как в теореме 52.
Единственное решение получается только в том случае, если линеаризованная система имеет полный ранг. Хорошо известно, что для случайных линейных систем вероятность получить полный ранг высока. Но система (7.2) порождается применением функции f к орбите начального состояния, и поэтому гипотеза о том, что полученная из неё линеаризованная система близка к случайной, выглядит неправдоподобной.

7.4 Криптосинтез

Можно ли использовать базисы Грёбнера не для взлома криптографических схем, а для их построения? Имеется очень небольшое количество публикаций, в которых обсуждается данная техника (см., например, [**12**; **14**; **15**; **16**]). В таких работах предлагаются конструкции криптосистем с открытым ключом.

И это своего рода «поцелуй Иуды». Действительно, авторы работ по криптоанализу (см. подраздел 7.3) считают, что с помощью аппарата базисов Грёбнера можно снизить трудоемкость криптоаналитических алгоритмов. А тут оказывается, что вычислительная сложность задачи построе-

ния базиса Грёбнера настолько высока, что может быть основой стойкости криптографических систем.

Еще в 1994 году была опубликована статья Б. Барки и др. [**13**] под красноречивым заголовком «Почему нет никакой надежды использовать базисы Грёбнера в криптографии с открытым ключом». Следует отметить уникальность этой работы. Трудно определить ее жанр — научная публикация или пасквиль. К примеру, второй из авторов — на самом деле фамилия D. Naccache после применения инверсии. А сама статья написана в форме открытого письма и начинается с обращения «Уважаемый Заблудший Автор». Отметим, что тезис, вынесенный в заголовок работы [**13**], остается необоснованным в его общности. В частности, авторы работы [**16**] полемизируют с работой [**13**]. Обсуждение этой тематики можно найти также в книге [**17**], являющейся довольно объемным обзором приложений аппарата базисов Грёбнера в кодировании и криптографии.

Следуя работе [**13**], опишем кратко идею построения криптосистемы с открытым ключом.

Предварительно напомним некоторые понятия и обозначения. Пусть I идеал кольца $k[x_1, \ldots, x_n]$ (k — конечное поле). Обозначим через $\mathbb{T}$ упорядоченное множество всех мономов в $k[x_1, \ldots, x_n]$. Тогда любой полином f из $k[x_1, \ldots, x_n]$ можно представить единственным образом в виде $f = \sum_{i=1}^{r} c_i t_i$, где $c_i \in k \setminus \{0\}$, $t_i \in \mathbb{T}$ и $t_1 > t_2 > \ldots > t_r$. Полиному $f \in k[x_1, \ldots, x_n]$ сопоставим $LT(f) = t_1$ — старший моном — и $LC(f) = c_1$ — старший коэффициент. Для идеала I рассмотрим два множества

$$T(I) = \{LT(f) \mid f \in I \setminus \{0\}\} \subset \mathbb{T},$$
$$O(I) = \mathbb{T} \setminus T(I).$$

Из теорем 40 и 42 следует следующее утверждение.

Предложение 22. *1.* $k[x_1, \ldots, x_n] = I \oplus Span_k(O(I))$.

2. Существует k-изоморфизм векторных пространств $k[x_1, \ldots, x_n]/I$ и $Span_k(O(I))$.

3. *Для любого* $f \in k[x_1, \ldots, x_n]$ *существует единственный* $g = Can(f, I) \in Span_k(O(I))$*, такой, что* $f - g \in I$*. Кроме того:*

   (a) $Can(f, I) = Can(g, I)$ *тогда и только тогда, когда* $f - g \in I$.

   (b) $Can(f, I) = 0$ *тогда и только тогда, когда* $f \in I$.

Пусть имеется подходящее бесконечное семейство идеалов (в кольцах полиномов над конечным полем), для которого задача построения базисов Грёбнера имеет высокую вычислительную сложность. Пусть I — идеал из этого семейства.

Предположим, что Арчибальду каким-то образом становится известен базис Грёбнера идеала I. Это его секретный ключ. Открытый ключ Арчибальда состоит из множества мономов $T \subset O(I)$ и набора полиномов $\{g_1, \ldots, g_l\} \subset I$ невысокой степени. Этот набор является базисом либо идеала I, либо некоторого идеала, вложенного в I.

Предположим также, что открытый текст M, который Балтазар хочет послать Арчибальду, представлен в виде $M = \sum_{t_i \in T} c_i t_i$. Для шифрования Балтазар выбирает случайные полиномы $p_1, \ldots, p_l$ и вычисляет шифртекст $C = M + \sum_{i=1}^{l} p_i g_i$.

Поскольку Арчибальд знает базис Грёбнера идеала I, он может вычислить каноническую форму полинома C. Но, так как $C - M \in I$, то $Can(C, I) = Can(M, I) = M$. Таким образом выполняется дешифрование.

Их (не)друг (в оригинале — f(r)iend) Фантомас знает $T, g_1, \ldots, g_l$ и C. Но Фантомас знает также, что M — каноническая форма полинома C.

Основной тезис работы состоит в том, что Фантомасу, на самом деле, требуется вычислить не базис Грёбнера, а каноническую форму полинома C. А последняя задача, с вычислительной точки зрения, может оказаться существенно проще. Дается краткое пояснение почему.

Суть проблемы в следующем. Рассмотрим функцию F, которая отображает $(T, g_1, \ldots, g_l, M, p_1 \ldots, p_l)$ в $(T, g_1, \ldots, g_l, C)$. Здесь $(T, g_1, \ldots, g_l)$ — параметры, выбирающие функцию из семейства функций. Какие у нас имеются основания считать функцию F односторонней? Подчеркнем, что вычислительной трудности в худшем случае задачи инвертирования функ-

ции F недостаточно. Также как и трудности в среднем. Требуется именно односторонняя функция.

Если удастся построить, быть может, модифицированную функцию F для которой гипотеза об односторонности будет обоснована математически, то возникнет вторая проблема. Требуется найти эффективный метод генерации параметров функции F с уже известным базисом Грёбнера.

В случае решения обеих обозначенных проблем будет получен второй, после функции Рабина, пример семейства функций с секретом.

Глава 8

О вычислительной сложности задач построения базисов Гребнера

8.1 Постановка задачи

Данный раздел предполагает знакомство читателя с основными понятиями теории сложности вычислений. Необходимую информацию можно найти в монографиях по этой теории [**GJ**], [**Kuz1**], либо в приложении 10.2.

Задача построения базиса Грёбнера может быть сформулирована в следующем стиле. Даны набор многочленов, определяющих идеал, и допустимый порядок. Найти базис Грёбнера этого идеала. Это — математическая задача, для которой вычислительная сложность не может быть определена.

В отличие от математических задач для вычислительных задач необходима точная спецификация форматов входных и выходных данных. Очевидно, что математической задаче может соответствовать бесконечное множество вычислительных задач. Последние могут иметь различную вычислительную сложность.

Всюду далее рассматривается следующая постановка вычислительной задачи построения базиса Грёбнера.

В случае многочленов над полем рациональных чисел данные представляются следующим образом. Множество переменных в задаче — $X = \{x_1, x_2, \ldots, x_n, \ldots\}$. Многочлен будем описывать (кодировать) конечным словом в алфавите A, состоящем из четырех символов $0, 1, \$, \#$. Предполагается, что все переменные занумерованы натуральными числами. Кодом $K(x_i)$ каждой переменной x_i объявляется кратчайшая запись натурального числа i в двоичной системе счисления. Кодом $K(r)$ каждого рационального числа $r = p/q$ считается слово $\sigma(r)\$K(p)\$K(q)$, где

- $\sigma(r) = 1$, если $r \geq 0$, и $\sigma(r) = 0$, если $r < 0$,
- $K(p)$ и $K(q)$ — кратчайшие записи натуральных чисел $|p|$ и $|q|$ в двоичной системе счисления.

Кодом $K(t)$ одночлена $t = rx_{i_1}^{k_1} \ldots x_{i_n}^{k_n}$ является слово

$$K(r)\$K(x_{i_1})\$K(k_1)\$ \ldots \$K(x_{i_n})\$K(k_n) .$$

Кодом $K(f)$ многочлена $f = t_1 + t_2 + \cdots + t_N$ является слово

$$K(t_1)\#K(t_2)\# \ldots \#K(t_N) .$$

Очевидно, что два разных многочлена имеют разные коды, и для каждого слова в алфавите A можно легко проверить, является ли это слово кодом какого-либо многочлена.

Таким же способом можно закодировать многочлены над любым полем, которое порождается некоторым конечным подмножеством его элементов.

Для заданной системы многочленов допустимый порядок на множестве термов определяется на основании теорем 4 и 43 квадратными матрицами над полем рациональных чисел. Матрицы $\{a_{ij}\}_{i,j=1}^{n}$ такого вида также могут быть закодированы словами

$$K(a_{11})K(a_{12}) \ldots K(a_{1n})K(a_{21}) \ldots K(a_{nn})$$

в алфавите A.

Таким образом, входными данными вычислительной задачи построения базиса Гребнера являются конечные наборы конечных слов в конечном алфавите A, которые представляют собой коды многочленов, определяющих некоторый идеал I в кольце многочленов, и код соответствующей им матрицы рациональных чисел, задающей допустимый порядок $\prec$ на множестве термов.

Алгоритм построения базиса Гребнера, получив входные данные в указанном выше формате, должен завершить вычисление и выдать на выходе конечный набор слов в алфавите A, которые являются кодами многочленов, образующих базис Гребнера идеала I относительно допустимого порядка $\prec$.

При разработке алгоритмов построения базисов Гребнера и оценке вычислительной сложности этой задачи в качестве модели вычислений будет использоваться последовательная машина с произвольным доступом к памяти (см. [GJ]). Эта модель очень удобна для описания алгоритмов решения вычислительных задач. Наряду с машинами Тьюринга эта модель широко используется в теории сложности вычислений.

Заметим, что ни в одной из работ, посвященных вычислительной сложности задачи построения базисов Грёбнера, нет постановки вычислительной задачи, т. е. нет форматов входных и выходных данных. А они весьма существенны. Например, многочлен может быть задан набором всех коэффициентов (включая нулевые) вплоть до старшего терма. Такая спецификация выглядит искусственной. Но лишь на первый взгляд. Рассмотрим семейство систем уравнений. Каждая система содержит единственное уравнение вида $x_n = 0$, где n — целочисленный растущий параметр. Ставится задача найти хотя бы одно решение системы. Если для постановки вычислительной задачи использовать принятую нами спецификацию кодирования многчленов, то такая задача будет иметь экспоненциальную сложность.

Поскольку базисы Гребнера определяются относительно некоторого фиксированного допустимого порядка на множестве термов, то возможны два варианта постановки задачи:

- найти базис Гребнера заданного идеала. Входными данными являются многочлены, задающие идеал. Результатом является описа-

ние некоторого допустимого порядка на множестве термов и базис Гребнера относительно этого порядка.

- Найти базис Гребнера заданного идеала относительно заданного допустимого порядка на множестве термов. Входными данными являются многочлены, задающие идеал, и описание допустимого порядка на множестве термов. Результатом является базис Гребнера данного идеала относительно этого порядка.

Нам неизвестны работы, в которых проводилось исследование первой из указанных задач. Мы ограничимся рассмотрением второй задачи.

Хорошо известно, что уже в булевом случае задача распознавания разрешимости систем квадратичных уравнений NP-полна [**GJ**]. Но даже для булева случая неясно, принадлежит ли задача построения базиса Грёбнера классу FNP (функциональный аналог класса NP). Известно лишь, что эта задача принадлежит классу FPSPACE, см. подраздел 8.3).

В подразделе 8.2 построен пример системы полиномиальных уравнений, для которой задача нахождения решений тривиальна, а приведенный базис Гребнера имеет экспоненциальную длину.

Основной результат раздела (теорема 58 Майра и Мейера) доказывается в подразделе 8.8: задача построения базиса Грёбнера EXPSPACE-трудна. В подразделе 8.11 показано, что базис Грёбнера может быть построен на экспоненциальной памяти.

Последующие подразделы посвящены доказательству вспомогательных утверждений.

Прежде чем перейти к формулировкам и доказательствам результатов, на наш взгляд уместно сделать следующие замечания.

1. В определениях классов языков с ограничениями на память (пример — класс PSPACE), безразлично, какая память учитывается — вся или только рабочая. Но для функциональных аналогов таких классов это может иметь существенное значение. Всюду в данном разделе рассматриваются ограничения только на рабочую память.

2. Хорошо известно, что результаты об NP-полноте или NP-трудности являются на данный момент лишь аргументами в пользу высокой вычислительной сложности задач. Следует подчеркнуть, что результаты, скажем, об EXPSPACE-трудности абсолютны. Как бы в дальнейшем ни были

решены открытые вопросы теории сложности вычислений, любой алгоритм решения такой задачи требует экспоненциальной памяти.

3. С точки зрения криптографических приложений, важнейшим является случай системы полиномиальных булевых уравнений, которая заведомо имеет решение и притом единственное. Сложностной статус таких задач неясен. Во-первых, они принадлежат классу FUP (определение см. в приложении). Во-вторых, это так называемые задачи с обязательством (promise problem). Входные данные удовлетворяют дополнительному требованию (обязательству): решение существует.

8.2 Простая система уравнений со сложным базисом Гребнера

Основным результатом данного раздела является доказательство следующей теоремы.

Теорема 53. *Задача построения базиса Гребнера булева идеала не принадлежит классу* $FPSPACE$.

Рассмотрим кольцо многочленов $K[X]$ от переменных $X = \{x_1, \ldots, x_n\}$ над произвольным полем K. Фиксируем целое число $0 < s \leq n$. Для целого $0 < i \leq s$ обозначим через $\sigma_i^{(s)}(x_1, \ldots, x_s)$ i-й симметрический многочлен. Далее будем предполагать, что характеристика поля K либо 0, либо больше s.

Фиксируем неотрицательное целое число $k \leq n$, где n — число переменных в кольце многочленов. Рассмотрим идеал I, порожденный многочленами

$$\begin{aligned} f_0(x_1, \ldots, x_n) &= x_1 + \ldots + x_n - k, \\ f_i(x_1, \ldots, x_n) &= x_i^2 - x_i, \ i = 1, \ldots, n. \end{aligned} \tag{8.1}$$

Рассмотрим множество многочленов $F = \{f_i \mid i = 1, \ldots, n\}$.

Лемма 47. *Пусть базисом идеала I являются многочлены из формулы ((8.1)) и заданы два целых числа $0 < i \leq s < n$. Тогда существует такой многочлен $P_{i,s}$ степени не выше i от переменных $x_{s+1}, \ldots, x_n$, что*

$$\sigma_i^{(s)}(x_1, \ldots, x_s) + P_{i,s}(x_{s+1}, \ldots, x_n) \in I.$$

Доказательство. Воспользуемся индукцией по i.

При $i = 1$ определим многочлен $P_{i,s}(x_{s+1},\ \ldots, x_n)$ формулой

$$P_{1,s}(x_{s+1},\ \ldots, x_n) = x_{s+1} + \ \ldots + x_n - k = \sigma_1^{(n-s)}(x_{s+1},\ \ldots, x_n) - k.$$

Тогда

$$I \ni f_0(x_1,\ \ldots, x_n) = \sigma_1^{(s)}(x_1,\ \ldots, x_s) + P_{1,s}(x_{s+1},\ \ldots, x_n).$$

Предположим, что существование многочленов $P_{i,s}$ доказано для всех $i < j \leq s$. Докажем существование многочлена $P_{j,s}$. Заметим, что для любых a и b в силу равенства

$$a^j - b^j = (a-b)(a^{j-1} + \ \ldots + b^{j-1})$$

и определения идеала из соотношения $a - b \in I$ следует, что $a^j - b^j \in I$.

Положим

$$\begin{array}{lcl} a = x_1 + \ \ldots + x_s & = & \sigma_1^{(s)}(x_1,\ \ldots, x_s), \\ b = -x_{s+1} - \ \ldots - x_n + k & = & -\sigma_1^{(n-s)}(x_{s+1},\ \ldots, x_n) + k. \end{array}$$

Легко видеть, что

$$a^j = j!\cdot\sigma_j^{(s)}(x_1,\ \ldots, x_s) + L\left(\sigma_1^{(s)}(x_1,\ \ldots, x_s),\ \ldots, \sigma_{j-1}^{(s)}(x_1,\ \ldots, x_s)\right) \ (\text{ mod } F),$$

где L — линейная форма. Тогда при некотором $h \in I$ выполняется соотношение

$$\begin{array}{rcl} a^j & = & j!\cdot\sigma_j^{(s)}(x_1,\ \ldots, x_s) + L\left(\sigma_1^{(s)}(x_1,\ \ldots, x_s),\ \ldots, \sigma_{j-1}^{(s)}(x_1,\ \ldots, x_s)\right) \\ & + & h(x_1,\ \ldots, x_n). \end{array}$$

Согласно предположению индукции для всех $i = 1,\ \ldots, j-1$ существуют $h_i \in I$ для которых выполняются соотношения

$$\sigma_i^{(s)}(x_1,\ \ldots, x_s) = h_i(x_1,\ \ldots, x_n) - P_{i,s}(x_{s+1},\ \ldots, x_n).$$

Поскольку $a^j - b^j \in I$, и $b^j = Q_{j,s}(x_{s+1},\ \ldots, x_n)$ при некотором $Q_{j,s}$ степени не выше j, то существует $g \in I$, такой, что

$$\begin{aligned} a^j - b^j &= j! \cdot \sigma_j^{(s)}(x_1,\ \ldots, x_s) \\ &\quad -L\left(P_{1,s}(x_{s+1},\ \ldots, x_n),\ \ldots, P_{j-1,s}(x_{s+1},\ \ldots, x_n)\right) \\ &\quad -Q_{j,s}(x_{s+1},\ \ldots, x_n) + g(x_1,\ \ldots, x_n) \end{aligned}$$

и, следовательно, можно определить многочлен $P_{j,s}$ формулой

$$\begin{aligned} -P_{j,s}(x_{s+1},\ \ldots, x_n) &= \frac{1}{j!}\left(Q_{j,s}(x_{s+1},\ \ldots, x_n)\right. \\ &+ \left. L\left(P_{1,s}(x_{s+1},\ \ldots, x_n),\ \ldots, P_{j-1,s}(x_{s+1},\ \ldots, x_n)\right)\right). \end{aligned}$$

□

Введем обозначение

$$B^s = \underbrace{\{0,1\} \times\ \ldots \times \{0,1\}}_{s \text{ сомножителей}}.$$

Для $\omega = (i_1,\ \ldots, i_s) \in B^s$ положим

$$|\omega| = i_1 +\ \ldots + i_s.$$

Для любого целого $0 < k < n$ определим разбиение множества X на два множества $Y = \{x_1,\ \ldots, x_{k-1}\}$ (в случае $k = 1$ это множество пустое) и множество $Z = \{x_k,\ \ldots, x_n\}$. Напомним, что для любого $\omega \in B^{k-1}$ определен терм

$$Y^\omega = x_1^{i_1}\ \ldots x_{k-1}^{i_{k-1}},$$

где $\omega = (i_1,\ \ldots, i_{k-1})$, и для любого $\eta \in B^{n-k+1}$ определен терм

$$Z^\eta = x_k^{j_1}\ \ldots x_n^{j_{n-k+1}},$$

где $\eta = (j_1,\ \ldots, j_{n-k+1})$.

Лемма 48. *Пусть $0 < k < 2k < n$ и характеристика поля K больше n или равна 0. Тогда в кольце многочленов $K[x_k, \ldots, x_n]$ существует единственный, с точностью до умножения на ненулевую константу, ненулевой многочлен вида*

$$\sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| \leq k}} \lambda_\omega Z^\omega,$$

принадлежащий идеалу I.

Доказательство. Вначале докажем существование указанного многочлена, воспользовавшись леммой 47.

Положим $s = k - 1$. Согласно лемме 47 для всех $i < k$ определены многочлены $P_{i,k-1}$, удовлетворяющие условиям

$$\sigma_i^{(k-1)}(x_1, \ldots, x_{k-1}) + P_{i,k-1}(x_k, \ldots, x_n) \in I.$$

Поскольку $\deg P_{i,k-1} \leq i$, а идеал I — булев (см. определение 6.3.1), то существуют такие многочлены $h_{i,k} \in I$ и

$$Q_{i,k}(x_k, \ldots, x_n) = \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| \leq i}} \lambda_{\omega,k} Z^\omega \in K[Z],$$

что

$$P_{i,k-1}(x_k, \ldots, x_n) = Q_{i,k}(x_k, \ldots, x_n) + h_{i,k}(x_k, \ldots, x_n).$$

Тогда

$$\sigma_i^{(k-1)}(x_1, \ldots, x_{k-1}) + Q_{i,k}(x_k, \ldots, x_n) \in I \tag{8.2}$$

для всех пар i, k, удовлетворяющих условию $0 < i \leq k - 1 < n$.

Заметим, что в силу равенства

$$a^k - b^k = (a - b)(a^{k-1} + \ldots + b^{k-1})$$

и определения идеала из соотношения $a - b \in I$ следует, что $a^k - b^k \in I$.

Положим

$$a = x_1 + \ldots + x_{k-1}$$
$$b = -x_k - \ldots - x_n + k.$$

Раскрывая скобки и используя соотношения $x_i^2 = x_i$ при всех $i = 1, \ldots, n$ получаем равенства

$$a^k = \sum_{i=0}^{k-1} \lambda_i \sigma_i^{(k-1)}(x_1, \ldots, x_{k-1}), \pmod I$$
$$b^k = \sum_{i=0}^{k-1} \mu_i \sigma_i^{(n-k+1)}(x_k, \ldots, x_n) + k! \cdot \sigma_k^{(n-k+1)}(x_k, \ldots, x_n) \pmod I.$$

Тогда в силу соотношений ((8.2)) выполняется равенство

$$a^k = \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| < k}} \nu_\omega Z^\omega \pmod I,$$

и, учитывая, что $a^k - b^k \in I$, получаем

$$k! \cdot \sigma_k^{(n-k+1)}(x_k, \ldots, x_n) + \sum_{i=0}^{k-1} \mu_i \sigma_i^{(n-k+1)}(x_k, \ldots, x_n) - \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| < k}} \nu_\omega Z^\omega \in I.$$

Поскольку характеристика поля K больше k, многочлен

$$k! \cdot \sigma_k^{(n-k+1)}(x_k, \ldots, x_n) + \sum_{i=0}^{k-1} \mu_i \sigma_i^{(n-k+1)}(x_k, \ldots, x_n) - \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| < k}} \nu_\omega Z^\omega$$

не равен 0 и удовлетворяет требованиям леммы.

Далее покажем, что указанный многочлен единствен. Выберем любой многочлен

$$p(x_k, \ldots, x_n) = \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| \le k}} \lambda_\omega Z^\omega,$$

удовлетворяющий условиям леммы 48. Необходимым условием принадлежности многочлена p идеалу I является равенство этого многочлена нулю во всех корнях идеала I.

Разобьем множество M всех корней идеала I на непересекающиеся подмножества M_i, где $i = 0, 1, \ldots, k-1$. Для этого определим множество M_i как множество корней $(\alpha_1, \ldots, \alpha_n)$ идеала, удовлетворяющих

условию $\alpha_1 + \ \ldots + \alpha_{k-1} = i$. Множества M_i, очевидно, не пересекаются. Поскольку идеал I булев, то компоненты корней могут быть равными только 0 или 1. Поэтому $\alpha_1 + \ \ldots + \alpha_{k-1} < k$. Следовательно, множество корней M идеала I представимо в виде разбиения

$$M = \bigcup_{i=0}^{k-1} M_i.$$

Докажем, что

$$\lambda_\omega = (-1)^{|\omega|}\lambda_{\omega_0}, \tag{8.3}$$

где

$$\omega_0 = (0, \ \ldots, 0).$$

Доказательство проведем индукцией по $|\omega|$.

Пусть $|\omega| = 1$. Рассмотрим множество корней $M_{k-1} \subset M$. Поскольку для любого корня $\alpha = (\alpha_1, \ \ldots, \alpha_n)$ из этого множества выполняется равенство $\alpha_1 + \ \ldots + \alpha_{k-1} = k-1$, то только одна из координат $(\alpha_k, \ \ldots, \alpha_n)$ равна 1, а остальные нулевые. Все эти корни находятся во взаимно однозначном соответствии с такими векторами $\omega \in B^{n-k+1}$, что $|\omega| = 1$. Поскольку многочлен p обращается в ноль на таких векторах $(\alpha_k, \ \ldots, \alpha_n)$, для всех ω, $|\omega| = 1$, выполняется равенство $\lambda_\omega = -\lambda_{\omega_0}$.

Пусть равенство ((8.3)) доказано для всехω, $|\omega| < m \leq k$. Докажем равенство ((8.3)) для ω, $|\omega| = m$.

Рассмотрим множество корней $M_{k-m} \subset M$. Пусть $(\alpha_1, \ \ldots, \alpha_n) \in M_{k-m}$ Поскольку $\alpha_1 + \ \ldots + \alpha_{k-1} = k - m$, из координат $(\alpha_k, \ \ldots, \alpha_n)$ в точности m равны 1, а остальные нулевые. Все такие решения находятся во взаимно однозначном соответствии с такими $\omega \in B^{n-k+1}$, что $|\omega| = m$. Заметим, что согласно определению множества M_{k-m} и предположению индукции для таких решений выполняется равенство

$$p(\alpha_k, \ \ldots, \alpha_n) = \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| \leq m}} \lambda_\omega A^\omega = \sum_{\substack{\omega \in B^{n-k+1} \\ |\omega| < m}} (-1)^{|\omega|}\lambda_{\omega_0} A^\omega + \sum_{\substack{\eta \in B^{n-k+1} \\ |\eta| = m}} \lambda_\eta A^\eta,$$

где $A = (\alpha_k, \ \ldots, \alpha_n)$. Согласно определению множества M_{k-m} и ввиду неравенства $2k < n$ при $0 \leq i < m$ на решениях из M_{k-m} справедливо

равенство

$$\sum_{\substack{\omega\in B^{n-k+1}\\|\omega|=i}}(-1)^{|\omega|}\lambda_{\omega_0}A^{\omega}=\sum_{\substack{\omega\in B^{n-k+1}\\|\omega|=i}}(-1)^{i}\lambda_{\omega_0}A^{\omega}=(-1)^{i}\cdot\binom{m}{i}\cdot\lambda_{\omega_0}.$$

Также выполнено равенство

$$\lambda_{\eta_0}=\sum_{\substack{\eta\in B^{n-k+1}\\|\eta|=m}}\lambda_{\eta}A^{\eta},$$

где $\eta_0=(\alpha_k,\ \ldots,\alpha_n)$. Заметим, что справедливы равенства

$$\sum_{i=0}^{m}\binom{m}{i}(-1)^{i}=(1-1)^{m}=0. \tag{8.4}$$

Поскольку многочлен p равен нулю на таких векторах $(\alpha_k,\ \ldots,\alpha_n)$, выполняются равенства

$$p(\alpha_k,\ \ldots,\alpha_n)=\sum_{i=0}^{m-1}(-1)^{i}\cdot\binom{m}{i}\cdot\lambda_{\omega_0}+\lambda_{\eta_0}=0.$$

Учитывая теперь соотношение ((8.4)), получаем равенство $\lambda_{\eta_0}=(-1)^{m}\cdot\lambda_{\omega_0}$. Следовательно, для всех $|\omega|=m$ выполняется равенство $\lambda_{\omega}=(-1)^{|\omega|}\cdot\lambda_{\omega_0}$.

Поэтому

$$p(\alpha_k,\ \ldots,\alpha_n)=\lambda_{\omega_0}\cdot\left(\sum_{i=0}^{k}(-1)^{i}\sigma_i^{(n-k+1)}(x_k,\ \ldots,x_n)\right).$$

□

Следствие 23. *Пусть выполняются предположения леммы 48. Тогда многочлен*

$$p(\alpha_k,\ \ldots,\alpha_n)=\sum_{i=0}^{k}(-1)^{i}\sigma_i^{(n-k+1)}(x_k,\ \ldots,x_n). \tag{8.5}$$

является неприводимым элементом базиса Гребнера идеала I для лексикографического упорядочения термов на множестве переменных $X=\{x_1,\ \ldots,x_n\}$.

Доказательство. Предположим, что многочлен p приводим. Тогда существует элемент g неприводимого базиса Гребнера идеала I, старший терм которого делит один из термов многочлена p. Следовательно, этот старший терм многочлена g содержит только переменные из множества $Z = \{x_k, \ldots, x_n\}$. Поскольку термы многочлена g упорядочены лексикографически, все его остальные термы также зависят только от переменных из множества Z. Поэтому многочлен g удовлетворяет условиям леммы 48, и многочлен $p \in K[x_k, \ldots, x_n]$ отличается от g ненулевым множителем, т.е. g неприводим, что противоречит выбору g. □

Следующее утверждение завершает доказательство теоремы 53.

Следствие 24. *При $k = \left[\frac{n}{2}\right]$ размер базиса Гребнера идеала I экспоненциален относительно длины описания идеала I формулами ((8.1)).*

Доказательство. Очевидно, длина описания идеала I равна n. Многочлен p из следствия 23 является неприводимым элементом базиса Гребнера идеала I и содержит $\binom{n}{k+1} > 2^{n/2}$ термов с коэффициентами ± 1. □

8.3 Уравнения в булевой алгебре

Задача нахождения базиса Гребнера для булева идеала может быть решена алгоритмически, с использованием не более чем экспоненциальной памяти (лежит в классе $FPSPACE$).

Теорема 54. *Для заданных базиса булева идеала $I \subseteq \mathbb{Q}[x_1, \ldots, x_n]$ и допустимого порядка на термах $<$ единственный приведенный базис Гребнера для I относительно порядка $<$ может быть вычислен алгоритмом, объем используемой памяти которого ограничен сверху полиномом от длины входных данных.*

8.4 EXPSPACE-трудность задачи проверки принадлежности базису Гребнера

В данном разделе под биномом понимается разность двух непустых термов (т.е. $t \neq x_1^0 \cdots x_n^0 = 1$). Идеал $I \subset K[x_1, \ldots, x_n]$ называется биномиальным, если существует его базис, состоящий из биномов.

Лемма 49. *Приведенный базис Гребнера биномиального идеала состоит из биномов.*

Доказательство. Заметим, что при применении алгоритма Бухбергера для вычисления базиса Гребнера биномиального идеала все получаемые многочлены являются биномами:

- S-многочленом биномов также является бином.
- Результатом операции приведения бинома относительно множества биномов также является бином.

□

Обозначим через A_G алгоритм нахождения базиса Гребнера относительно лексикографического порядка в градуированном моноиде термов для произвольного идеала в кольце многочленов с рациональными коэффициентами. Обозначим $SPACE(A_G(I))$ необходимую для выполнения алгоритма A_G память, если на вход алгоритма A_G поступает описание идеала I.

Через A_{IM} обозначим алгоритм проверки принадлежности многочлена идеалу. На вход алгоритма A_{IM} поступает описание идеала I и многочлена $f \in K[x_1, \ldots, x_n]$, относительно которого решается задача принадлежности. Обозначим через $SPACE(A_{IM}(I))$ необходимую для выполнения алгоритма A_{IM} память в наихудшем случае, т.е.

$$SPACE(A_{IM}(f)) = \max_f SPACE(A_{IM}(f, I)).$$

$$cn \log d + SPACE(A_G(I)) \geq SPACE(A_{IP}(I)).$$

Теорема 55. *Задача проверки равенства слов PSPACE-сводима по Тьюрингу к задаче проверки принадлежности многочлена базису Гребнера идеала.*

Доказательство. Предположим, имеется алгоритм A_G проверки принадлежности элемента базису Гребнера относительно лексикографического порядка в градуированном моноиде термов для произвольного идеала в кольце многочленов с рациональными коэффициентами. Ограничимся биномиальными идеалами. Терм, предшествующий терму α, будем обозначать α'. Бином будем записывать в виде $x^{\omega_1} - x^{\omega_2}$, причем x^{ω_1} — старший терм. Положим $d = \deg(x^{\omega_1} - x^{\omega_2})$. Решим задачу проверки принадлежности бинома $x^{\omega_1} - x^{\omega_2}$ биномиальному идеалу с помощью следующего алгоритма, использующего алгоритм A_G в качестве вспомогательной процедуры.

Шаг 0 Выполняем присваивания $b := x^{\omega_1} - x^{\omega_2}$ и $\alpha := x^{\omega_1}$.

Шаг 1 Выполняем присваивания $\eta := \alpha$, $\nu := \alpha'$.

Шаг 2 Если $\nu = 1$, то переходим к Шагу 7.

Шаг 3 Выполняем присваивание $\beta := \eta - \nu$.

Шаг 4 Выполняем проверку $\beta \in G$ с помощью алгоритма A_G. Если $\beta \in G$, переходим к шагу 11.

Шаг 5 Выполняем присваивание $\nu := \nu'$.

Шаг 6 Если $\nu \neq 1$, то переходим к шагу 3.

Шаг 7 Выполняем присваивание $\mu := \eta'$.

Шаг 8 Если $\mu = 1$, то переходим к Шагу 13.

Шаг 9 Если $\mu \not| \alpha$, то выполняем присваивание $\eta = \mu$ и переходим к Шагу 7.

Шаг 10 Выполняем присваивания $\eta := \mu$ и $\nu := \eta'$. Переходим к Шагу 2.

Шаг 11 Приводим бином b с помощью элемента β. Получаем новый бином $x^{\gamma_1} - x^{\gamma_2}$. Выполняем присваивания $b := x^{\gamma_1} - x^{\gamma_2}$ и $\alpha := x^{\gamma_1}$.

Шаг 12 Если $b \neq 0$, то переходим к Шагу 1.

Шаг 13 Если $b = 0$, то $x^{\omega_1} - x^{\omega_2}$ принадлежит идеалу, в противном случае $x^{\omega_1} - x^{\omega_2}$ не принадлежит идеалу.

Шаг 14 Стоп.

Очевидно, что приведенный алгоритм позволяет решать задачу проверки принадлежности бинома идеалу, а следовательно, решать задачу равенства слов. Используя описание алгоритма, нетрудно проверить, что для его выполнения достаточно памяти, не превышающей памяти, требуемой для проверки принадлежности биномов базису Гребнера, степени которых не превышают степени сравниваемых слов, на величину не более чем $cn \log d$. Иными словами, имеется PSPACE-сводимость по Тьюрингу задачи проверки равенства слов к задаче проверки принадлежности элемента базису Гребнера. □

8.5 Построение базиса Гребнера алгортимом с экспоненциальным объемом памяти

В данном разделе доказывается

Теорема 56. *Существует алгоритм, который для любого идеала I, порожденного многочленами $f_1, \ldots, f_m \in \mathbb{Q}[x_1, \ldots, x_n]$ степени не выше d, и допустимого порядка $<$ на множестве термов, строит приведенный базис Гребнера этого идеала, используя экспоненциальный объем памяти.*

Хотя в формулировке этой теоремы рассматривается кольцо многочленов над полем рациональных чисел, ее очевидным образом можно обобщить на случай кольца многочленов над произвольным конечным полем. Такое обобщение требует описаний представления элементов поля и алгебраических операций в нем.

Доказательство. Доказательство этой теоремы использует результат теорем 60 и 39.

Согласно теореме 39 для построения базиса Гребнера достаточно провести полный перебор термов, степень которых не превосходит величины

$$D = 2\left(\frac{d^2}{2} + d\right)^{2^{n-2}} \tag{8.6}$$

и проанализировать их нормальные формы.

На первом этапе рассмотрим множество термов, степень которых не превосходит D, и для каждого терма t из этого множества построим его нормальную форму $\mathsf{NF}(t)$, руководствуясь теоремой 60. Если старший моном нормальной формы отличен от исходного терма t, то этому терму сопоставляется элемент базиса Гребнера, равный $t - \mathsf{NF}(t)$. (Согласно определению нормальной формы этот элемент принадлежит идеалу I.) В противном случае терм исключается из рассмотрения. Очевидно, что в результате получается базис Гребнера. Однако этот базис не является приведенным, поскольку содержит элементы, старшие термы которых могут быть в отношении делимости. Чтобы построить приведенный базис Гребнера G, необходимо исключить лишние элементы так, чтобы для любых $g_1, g_2 \in G$ не выполнялось ни одно из соотношений

$$\mathsf{HT}(g_1)|\mathsf{HT}(g_2) \text{ и } \mathsf{HT}(g_2)|\mathsf{HT}(g_1).$$

С этой целью на втором этапе будем перебирать все термы степени не выше D, используя градуированный лексикографический порядок $\leq_{g-lex}$ на множестве термов. Отметим, что нормальная форма определяется относительно, вообще говоря, другого допустимого порядка. Нормальную форму многочлена h будем обозначать через $\mathsf{NF}(h)$.

Пусть t — произвольный терм степени, не выше D. Необходимым и достаточным условием принадлежности многочлена $t - \mathsf{NF}(t)$ приведенному базису Гребнера идеала I является отсутствие такого терма t', который делит терм t и удовлетворяет соотношению $t' - \mathsf{NF}(t') \neq 0$. Чтобы проверить это условие достаточно перебрать все термы t' степени не выше D, которые являются делителями рассматриваемого терма t. Если таких термов нет, оставляем элемент $t - \mathsf{NF}(t)$ как элемент приведенного базиса Гребнера. В противном случае исключаем этот элемент.

Алгоритм, осуществляющий построение базиса Гребнера, приведен ниже.

1 Присваиваем $t := 1$, $G := \emptyset$. Вычисляем NF(t).

2 Если NF$(t) = 0$, то присваиваем $G := \{1\} \cup G$ и переходим к шагу 14.

3 Выбираем следующий за термом t терм t' относительно порядка $\leq_{g-lex}$ на множестве термов.

4 Если deg$(t') > D$, то переходим к шагу 14.

5 Присваиваем $t := t'$.

6 Вычисляем NF(t).

7 Если NF$(t) = t$, то переходим к шагу 3.

8 Вычисляем элемент $g(t)$ по формуле $g(t) = t -$ NF(t). Выполняем присваивание $s := t$.

9 Среди собственных делителей терма t выбираем максимальный относительно порядка $\leq_{g-lex}$ терм s_1, такой что $s_1 \leq_{g-lex} s$. Вычисляем $g(s_1) = s_1 -$ NF(s_1).

10 Если $g(s_1) \neq 0$, то переходим к шагу 3. (В этом случае элемент $g(t)$ не добавляется в базис Гребнера G.)

11 Выполняем присваивание $s := s_1$.

12 Если $s \neq 1$, то переходим к шагу 9.

13 Добавляем элемент $g(t)$ в базис Гребнера: $G := \{g(t)\} \cup G$. Переходим к шагу 3.

14 Заканчиваем вычисление: базис Гребнера G найден.

Как видно из описания алгоритма вся используемая им память расходуется для построения нормальной формы и для хранения очередного терма. Как следует из теоремы 60 и ограничения степени терма величиной D, для выполнения алгоритма достаточно экспоненциального объема памяти от размера описания идеала I. □

8.6 Метод линеаризации

Теорема Херманн 37 позволяет свести обозначенную ниже через IM задачу определения принадлежности многочлена идеалу к задаче решения системы линейных уравнений.

Пусть заданы многочлены $f_0, f_1, \ldots, f_k \in K[x_1, \ldots, x_n]$. Требуется определить, верно ли, что

$$f_0 \in (f_1, \ldots, f_k),$$

т.е. существуют ли многочлены $g_1, \ldots g_k \in K[x_1, \ldots, x_n]$, для которых выполняется равенство

$$g_1 f_1 + \ldots + g_k f_k = f_0.$$

Предположим, что deg $f_i < d$ при $i = 1, \ldots, k$ и deg $f_0 < B$. Допустим, что выполняется соотношение $f_0 \in (f_1, \ldots, f_k)$. Тогда согласно теореме Херманн соотношение $f_0 \in (f_1, \ldots, f_k)$ выполняется тогда и только тогда, когда существуют такие многочлены $g_1, \ldots, g_k \in K[x_1, \ldots, x_n]$, степеней, не выше $B + (kd)^{2^n}$, для которых выполняется равенство

$$f_1 g_1 + \ldots + f_k g_k = f_0.$$

Следовательно, задача проверки принадлежности идеалу эквивалентна задаче проверки разрешимости уравнения

$$g_1 f_1 + \ldots + g_k f_k = f_0 \tag{8.7}$$

относительно неизвестных $g_1, \ldots, g_k$ в подмножестве многочленов, степени, не выше $B + (kd)^{2^n}$.

Рассмотрим последнюю из упомянутых задач. Поскольку deg $f_0 < B$, deg $f_i < d$ и deg $g_i \leq B + (kd)^{2^n}$, имеют место равенства

$$\begin{aligned} f_0 &= \sum_{|\omega|<B} a_{0,\omega} x^\omega, \\ f_i &= \sum_{|\omega|<d} a_{i,\omega} x^\omega, \qquad i = 1, \ldots, k, \\ g_i &= \sum_{|\omega|\leq B+(kd)^{2^n}} \beta_{i,\omega} x^\omega, \quad i = 1, \ldots, k, \end{aligned}$$

где $\beta_{i,\omega}$ — неизвестные.

Тогда

$$g_1 f_1 + \ \ldots + g_k f_k \ = \ \sum_{i=1}^{k} \left(\sum_{|\omega| \le B+(kd)^{2^n}} \beta_{i,\omega} x^\omega \right) \left(\sum_{|\omega|<d} a_{i,\omega} x^\omega \right)$$

$$= \sum_{|\omega| \le B+(kd)^{2^n}+d} \left(\sum_{\substack{\omega_1+\omega_2=\omega \\ |\omega_1| \le B+(kd)^{2^n} \\ |\omega_2|<d}} \sum_{i=1}^{k} \beta_{i,\omega_1} a_{i,\omega_2} \right) x^\omega$$

Воспользовавшись полученным равенством и полагая $a_{0\omega} = 0$ при $\omega \ge B$, сводим задачу разрешимости уравнения (8.7) к системе линейных уравнений по всем ω, таким, что $|\omega| \le B + (kd)^{2^n} + d$,

$$\sum_{\substack{\omega_1+\omega_2=\omega \\ |\omega_1| \le B+(kd)^{2^n} \\ |\omega_2|<d}} \sum_{i=1}^{k} \beta_{i,\omega_1} a_{i,\omega_2} = a_{0,\omega}, \tag{8.8}$$

относительно неизвестных β_{i,ω_1}, где $|\omega_1| \le B + (kd)^{2^n}$, эквивалентную уравнению ((8.7)). Число уравнений системы ((8.8)) не превосходит мощности множества векторов $\omega \in \mathbb{Z}_+^n$ таких, что $|\omega| < B + (kd)^{2^n} + d$, т.е. величины $(B + (kd)^{2^n} + d)^n$.

Теорема 57. *Система уравнений*

$$\begin{cases} \alpha_{1,1} f_1 + \alpha_{1,k} f_k = b_1, \\ \ldots \ \ldots \ \ldots \\ \alpha_{m,1} f_1 + \alpha_{m,k} f_k = b_m, \end{cases}$$

над кольцом $K[x_1, \ \ldots, x_n]$ относительно неизвестных $f_1, \ \ldots, f_k$, коэффициенты которой удовлетворяют условиям

$$\deg \alpha_{ij} < d \quad \text{и} \quad \deg b_i < B,$$

имеет решение в этом кольце тогда и только тогда, когда разрешима система ((8.8)) из $m(B+(kd)^{2^n}+d)^n$ линейных уравнений относительно $k(B + (kd)^{2^n})^n$ неизвестных.

8.7 Метод линеаризация для булевых идеалов

Для булевых идеалов задача проверки принадлежности многочлена идеалу также сводится к задаче решения системы линейных уравнений над $\mathbb{Z}_2$, но без использования теоремы Херманн.

Действительно, задача проверки принадлежности булеву идеалу сводится к задаче проверки принадлежности идеалу в кольце

$$\mathbb{Z}_2[x_1,\ \ldots, x_n]/(x_1^2 - x_1,\ \ldots, x_n^2 - x_n).$$

Любой элемент этого кольца представим в виде

$$p(x_1,\ \ldots, x_n) = \sum_{\omega \in (\mathbb{Z}_2)^n} a_\omega x^\omega.$$

В частности, степень любого многочлена не превосходит n.

Задача принадлежности булеву идеалу эквивалентна разрешимости уравнения

$$g_1 f_1 + \ \ldots + g_k f_k = f_0 \tag{8.9}$$

относительно неизвестных $g_1,\ \ldots, g_k$ в подмножестве многочленов $\mathbb{Z}_2[x_1,\ \ldots, x_n]/(x_1^2 - x_1,\ \ldots, x_n^2 - x_n)$.

Согласно условию имеем

$$\begin{aligned} f_0 &= \sum_{|\omega|<n} a_{0,\omega} x^\omega, \\ f_i &= \sum_{|\omega|<n} a_{i,\omega} x^\omega, \quad i = 1,\ \ldots, k, \\ g_i &= \sum_{|\omega|<n} \beta_{i,\omega} x^\omega, \quad i = 1,\ \ldots, k, \end{aligned}$$

где $\beta_{i,\omega}$ — неизвестные и $\omega \in (\mathbb{Z}/(2))^n$. Поскольку $x^{\omega_1} x^{\omega_2} = x^{\omega_1 \vee \omega_2}$, где $\omega_1 \vee \omega_2$ — покомпонентная дизъюнкция векторов ω_1 и ω_2, выполняются равенства

$$
\begin{aligned}
g_1 f_1 + \ \ldots + g_k f_k &= \sum_{i=1}^{k} \left(\sum_{|\omega|<n} \beta_{i,\omega} x^{\omega} \right) \left(\sum_{|\omega|<n} a_{i,\omega} x^{\omega} \right) \\
&= \sum_{|\omega|<n} \left(\sum_{\substack{\omega_1 \vee \omega_2 = \omega \\ |\omega_1|<n \\ |\omega_2|<n}} \sum_{i=1}^{k} \beta_{i,\omega_1} a_{i,\omega_2} \right) x^{\omega} \\
&= \sum_{|\omega|<n} \sum_{i=1}^{k} \sum_{|\omega_1|<n} \left(\sum_{\omega_2 \,:\, \omega_1 \vee \omega_2 = \omega} a_{i,\omega_2} \right) \beta_{i,\omega_1} x^{\omega}.
\end{aligned}
$$

Используя полученное равенство, получим что для нахождения коэффициентов g_i, для которых выполнено соотношение (8.9), достаточно найти решение системы линейных уравнений

$$
\sum_{i=1}^{k} \sum_{|\omega_1|<n} \left(\sum_{\omega_1 \vee \omega_2 = \omega} a_{i,\omega_2} \right) \beta_{i,\omega_1} = a_{0,\omega}, \tag{8.10}
$$

где $|\omega| < n$, относительно неизвестных β_{i,ω_1}, где $|\omega_1| < n$. Число уравнений системы ((8.10)) не превосходит мощности множества векторов $\omega \in (\mathbb{Z}/(2))^n$ таких, что $|\omega| < n$.

8.8 $EXPSPACE$-полнота задачи принадлежности многочлена идеалу

Задача проверки принадлежности многочлена данному идеалу (задача IM). Пусть задано поле K и многочлены

$$
f_0, \ \ldots, f_m \in K[x_1, \ \ldots, x_n].
$$

Требуется узнать, существуют ли такие многочлены $q_1, \ \ldots, q_m$, для которых выполняется равенство

$$
f_0 = \sum_{i=1}^{m} q_i f_i.
$$

Метод решения задачи IM, включая описание структуры идеала I, приведено в теоремах Д. Гильберта, Э. Ласкера и Г. Херманн (теоремы 36, 13, 14, 37).

Понятие базиса Гребнера идеала и алгоритм Бухбергера позволили установить алгоритмическую разрешимость задачи IM.

В этом разделе нас интересует вычислительная сложность задачи IM. В качестве основного поля K будем рассматривать поле рациональных чисел. Все результаты данного подраздела справедливы и для любого другого поля K, в котором операции сложения и умножения эффективно вычислимы.

Сформулируем основной результат данного подраздела.

Теорема 58. *(Майр* [**Mayr**]*) Задача IM является* $EXPSPACE$*-трудной.*

Доказательство теоремы проводится в два этапа. Сначала доказывается EXPSPACE-трудность задачи распознавания равенства слов в коммутативных полугруппах (задача CWEP). Второй этап заключается в демонстрации сводимости задачи CWEP к задаче IM.

Начнем с постановки задачи CWEP.

Поскольку CWEP рассматривается как вычислительная задача, необходимо указать, как задаются входные данные и, в частности, как задается коммутативная полугруппа $\mathcal{G}$. Эта полугруппа должна иметь описание конечной длины. Поэтому на полугруппу накладываются два ограничения: конечность множества ее образующих и конечная порожденность множества соотношений в полугруппе.

Конечное множество образующих полугруппы обозначим $U = \{u_1, \ldots, u_n\} \subset \mathcal{G}$. $T_1\langle U\rangle = T\langle U\rangle \backslash \{1\}$ является свободной коммутативной полугруппой с тем же множеством образующих, поэтому определен эпиморфизм (“гомоморфизм на”) полугрупп

$$\varphi : T_1\langle U\rangle \to \mathcal{G},$$

определяющий задание полугруппы $\mathcal{G}$ образующими из множества U. Тогда для некоторого отношения эквивалентности $\equiv$ выполняется равенство $\mathcal{G} = T_1\langle U\rangle / \equiv$.

Имеется, например, тривиальное описание такого отношения

$$t \equiv s \Leftrightarrow \varphi(t) = \varphi(s).$$

Такое описание в общем случае является бесконечным и, конечно, неприемлемо для наших целей. Чтобы дать конечное описание конечно порожденной полугруппы, введем следующие понятия.

Определение 8.8.1. *Пусть $\mathcal{P} = \{\alpha_1 \equiv \beta_1, \ \ldots, \alpha_k \equiv \beta_k\}$ — такой конечный набор соотношений эквивалентности между элементами полугруппы $T_1\langle U\rangle$, что из $\alpha \equiv \beta \in \mathcal{P}$ следует, что $\beta \equiv \alpha \in \mathcal{P}$ и $\equiv$ — минимальное мультипликативно замкнутое отношение эквивалентности в полугруппе термов $T_1\langle U\rangle$, содержащее множество $\mathcal{P}$. Простым выводом $\alpha \to \beta$ mod $\mathcal{P}$ называется эквивалентность термов $\alpha \equiv \beta$ вида*

$$\alpha = \gamma\alpha_i \equiv \gamma\beta_i = \beta,$$

полученная умножением соотношения эквивалентности $\alpha_i \equiv \beta_i$ из множества $\mathcal{P}$ на некоторый терм $\gamma \in T\langle U\rangle$. Выводом $\alpha \to \beta$ mod $\mathcal{P}$ эквивалентности $\alpha \equiv \beta$ называется последовательность простых выводов эквивалентностей

$$\gamma_1 \to \gamma_2 \to \ \ldots \to \gamma_n \text{ mod } \mathcal{P},$$

где $\alpha = \gamma_1$ и $\beta = \gamma_n$.

Определение 8.8.2. *Пусть задана полугруппа $\mathcal{G}$, конечное множество $U = \{u_1, \ \ldots, u_n\}$, эпиморфизм полугрупп*

$$\varphi : T_1\langle U\rangle \to \mathcal{G},$$

и конечный набор соотношений эквивалентности

$$\mathcal{P} = \{\alpha_1 \equiv \beta_1, \ \ldots, \alpha_k \equiv \beta_k\},$$

между термами $T_1\langle U\rangle$. Набор $(\mathcal{G}, U, \varphi, \mathcal{P})$ называется конечным заданием полугруппы $\mathcal{G}$, если, равенство $\varphi(\alpha) = \varphi(\beta)$ эквивалентно существованию вывода эквивалентности $\alpha \equiv \beta$ mod $\mathcal{P}$. Множество U называется множеством образующих группы $\mathcal{G}$, а множество $\mathcal{P}$ множеством соотношений конечного задания полугруппы.

Отметим, что не любая конечно порожденная полугруппа может быть описана таким способом. Однако в дальнейшем под конечно порожденной полугруппой будем подразумевать только те полугруппы, которые имеют указанное описание, т.е. конечно порожденные полугруппы, заданные конечным набором соотношений на множестве термов.

Задача $CWEP$ проверки эквивалентности слов в коммутативной полугруппе. Пусть имеется конечное задание коммутативной полугруппы $\mathcal{G}$ множеством образующих $U = \{u_1, \ \ldots, u_n\}$ и множеством соотношений $\mathcal{P} = \{\alpha_1 \equiv \beta_1, \ \ldots, \alpha_k \equiv \beta_k\}$ и пара термов (слов) $a, b \in T_1\langle U\rangle$. Требуется выяснить, выполняется ли $a \equiv b$.

Покажем, что задача $CWEP$ сводится к задаче IM.

Задание U, $\mathcal{P}$ определяет идеал

$$I_{\mathbb{Q}}(\mathcal{P}) = (\alpha_1 - \beta_1, \ \ldots, \alpha_k - \beta_k).$$

в кольце многочленов $\mathbb{Q}[U]$.

Лемма 50. *Если $\alpha \equiv \beta$, то $\alpha - \beta \in I_{\mathbb{Q}}(\mathcal{P})$.*

Доказательство. Согласно условию леммы существует вывод эквивалентности с помощью задания $\mathcal{P}$:

$$\alpha = \gamma_0 \to \gamma_1 \to \ \ldots \to \gamma_s = \beta \ \text{ mod } \mathcal{P}.$$

Без ограничения общности $s \geq 1$. Поэтому для любого $1 \leq m \leq s$ существуют $\delta_m \in T\langle U\rangle$ и $1 \leq i_m \leq n$, такие, что

$$\gamma_{m-1} = \delta_m \alpha_{i_m} \quad \text{и} \quad \gamma_m = \delta_m \beta_{i_m}.$$

Следовательно,

$$\beta - \alpha = \sum_{m=1}^{n} \delta_m(\beta_{i_m} - \alpha_{i_m}) \in I_{\mathbb{Q}}(\mathcal{P}).$$

□

Лемма 51. *Пусть α, β — термы на U. Если для некоторых ненулевых $g_i \in \mathbb{Q}[U]$, где $1 \leq i \leq k$, выполняется равенство*

$$\beta - \alpha = \sum_{i=1}^{k} (\beta_i - \alpha_i) g_i, \tag{8.11}$$

то существует такой вывод эквивалентности $\alpha \equiv \beta$ с помощью набора эквивалентностей $\mathcal{P}$

$$\alpha = \gamma_0 \to \gamma_1 \to \ \ldots \to \gamma_s = \beta \bmod \mathcal{P},$$

что для всех $j \in \mathcal{I}_s = \{0, \ \ldots, s\}$ выполняются неравенства

$$\deg \gamma_j \leq \max_{1 \leq i \leq k} \deg(\beta_i g_i).$$

Доказательство. Без ограничения общности можно считать, что $\alpha \neq \beta$.

Домножим обе части формулы (8.11) на общий знаменатель d всех рациональных коэффициентов многочленов g_i, где $1 \leq i \leq k$. Тогда для некоторого $s \geq 1$ задание соотношение можно записать в виде

$$d\beta - d\alpha = \sum_{m=1}^{s} (\beta_{i_m} - \alpha_{i_m}) t_m, \tag{8.12}$$

где $t_m \in T\langle U \rangle, \ 1 \leq i_m \leq k$ и $\deg t_m \leq \deg g_{i_m}$. Здесь мы воспользовались тем, что из $\alpha_i \equiv \beta_i \in \mathcal{P}$ следует, что и $\beta_i \equiv \alpha_i \in \mathcal{P}$.

Поскольку в левую часть формулы (8.12) входит терм α с ненулевым коэффициентом, при некотором $1 \leq r \leq s$ выполнено соотношение $\alpha = \alpha_{i_r} t_r$. Поэтому

1) $d\beta - (d-1)\alpha - \beta_{i_r} t_r = \sum\limits_{m \in \mathcal{I}_s \setminus \{r\}} (\beta_{i_m} - \alpha_{i_m}) t_m$,

2) $\alpha = \alpha_{i_r} t_r \to \beta_{i_r} t_r \bmod \mathcal{P}$.

Если $\beta_{i_r} t_r = \beta$, то вывод искомой эквивалентности получен. В противном случае существует $r' \in \mathcal{I}_s \setminus \{r\}$, такой, что $\beta_{i_r} t_r = \alpha_{i_{r'}} t_{r'}$, и поэтому выполнены соотношения

1) $d\beta - (d-1)\alpha - \beta_{i_{r'}} t_{r'} = \sum\limits_{m \in \mathcal{I}_s \setminus \{r,r'\}} (\beta_{i_m} - \alpha_{i_m}) t_m$,

2) $\alpha \to \beta_{i_r} t_r = \alpha_{i_{r'}} t_{r'} \to \beta_{i_{r'}} t_{r'} \bmod \mathcal{P}$.

Если $\beta_{i_{r'}} t_{r'} = \beta$, то требуемый вывод построен. Иначе продолжим построение. Поскольку множество $\mathcal{I}_s$ содержит s элементов, не более чем за s шагов будет получен вывод

$$\alpha \to \gamma_1 \to \ \ldots \to \gamma_{s'} = \beta \bmod \mathcal{P}$$

или вывод

$$\alpha \to \gamma_1 \to \ \ldots \to \gamma_s = \beta_{i_0} t_{r_0} \bmod \mathcal{P}$$

и соотношение

$$d\beta - (d-1)\alpha - \beta_{i_0} t_{r_0} = 0.$$

из которого следует, что $\beta = \alpha$, что противоречит исходному предположению $\alpha \neq \beta$. □

Согласно леммам 50 и 51 задача $CWEP$ является специальным случаем задачи IM — принадлежности идеалу разности двух термов.

Ниже приводится доказательство $EXPSPACE$-трудности задачи $CWEP$, из которого вытекает $EXPSPACE$-трудность задачи IM.

Схема доказательства $EXPSPACE$-трудности задачи $CWEP$ следующая:

- Вначале мы обратимся к модели вычислений машин Минского [Мин71] и докажем $EXPSPACE$-трудность задачи ESC (Exponential State Complexity) проверки завершаемости вычислений 3-счетчиковых машин Минского с ограниченной высотой счетчиков (Лемма 52).

- Далее докажем, что задача проверки завершаемости вычислений 3-счетчиковой машины Минского C сводится к задаче проверки равенства слов в некоторой коммутативной полугруппе $\mathcal{G}'_C$, которая соответствует машине C.

- Однако построенная полугруппа $\mathcal{G}'_C$ имеет слишком сложное устройство, и сведение задачи ESC к задаче проверки равенства слов в этой полугруппе требует больших вычислительных затрат. Чтобы избежать этого, построим сравнительно коммутативную полугруппу $\mathcal{G}_C$ с более простой системой определяющих соотношений, в которую вложима полугруппа $\mathcal{G}'_C$. Вложение устроено так, что задача ESCсводится к задаче равенства слов в полугруппе $\mathcal{G}_C$, и это сведение требует небольшого объема вычислительных ресурсов. Отсюда следует, что задача $CWEP$ является $EXPSPACE$-трудной.

Для того чтобы формально определить задачу ESC, нам нужно обратиться к модели вычислений счетчиковых машин. Счетчиковые машины разработал и исследовал американский математик М. Минский в 1958 г. Счетчиковые машины (машины Минского) — это одна из наиболее простых алгоритмически полных моделей вычисления. Машина с n счетчиками проводит вычисления на n регистрах памяти, в каждом из которых может быть записано произвольное целое неотрицательное число. На каждом шаге вычисления машина Минского может либо изменить содержимое одного из регистров на величину -1 или 1, либо проверить, содержит ли некоторый регистр значение 0, и в зависимости от этого осуществить переход в то или иное состояние управления. В монографии [Мин71] показано, что любая частично рекурсивная функция может быть вычислена некоторой 2-счетчиковой машиной Минского.

Формальное описание машины Минского таково. Счетчиковая машина представляет собой вычислительное устройство, которое состоит из программы, имеющей конечное число команд, и конечного набора счетчиков. В каждом счетчике может храниться произвольное целое число. Программа счетчиковой машины — это автомат с конечным числом состояний; каждому состоянию q этой программы приписана команда одного из следующих трех типов:

1. увеличить на 1 число m_i, содержащееся в i-ом счетчике, и перейти в состояние q';

2. уменьшить на 1 число m_i, содержащееся в i-ом счетчике, и перейти в состояние q';

3. проверить, содержится ли в i-ом счетчике число 0, и в зависимости от результата проверки перейти либо в состояние q', либо в состояние q''.

Далее рассматриваются детерминированные машины Минского с тремя счетчиками (3-счетчиковые машины), устройство и функционирование которых определяется более формально следующим образом. *3-счетчиковая машина* — это система $C = \langle Q, q_a, q_1, \Pi \rangle$, состоящая из

- конечного множества *состояний* $Q = \{q_a, q_1, q_2, \ldots, q_N\}$,
- *начального состояния* q_1,
- *допускающего состояния* q_a,
- конечного множества *счетчиковых команд* $\Pi = \{K_1, \ldots, K_N\}$.

Каждому состоянию q_i, $0 \leq i \leq N$, машины C, отличному от допускающего состояния q_a, приписана счетчиковая команда K_i одного из следующих двух видов:

1. $K_i = \langle q_i, x, \sigma, q_j \rangle$,
2. $K_i = \langle q_i, \sigma, q_j, q_k \rangle$,

где $x \in \{-1, 1\}$, $\sigma \in \{1, 2, 3\}$, а q_j и q_k — это некоторые состояния из множества Q. В начале функционирования счетчиковая машина C устанавливается в начальное состояние q_0. На каждом шаге функционирования счетчиковая машина C пребывает в одном из состояний q_i множества Q и выполняет приписанную этому состоянию команду K_i. Счетчиковая команда первого вида прибавляет число x к содержимому счетчика с номером σ и переводит машину C в состояние q_j. Счетчиковая команда второго вида проверяет, равно ли нулю число, содержащееся в счетчике с номером σ. Если это число равно 0, то машина C переходит в состояние q_j, иначе она переходит в состояние q_k. Функционирование завершается, как только счетчиковая машина достигает допускающего состояния q_a.

Более строго вычисление 3-счетчиковой машины $C = \langle Q, q_a, q_1, \Pi \rangle$ определяется так. *Счетчиковой конфигурацией* называется четверка

(q, m_1, m_2, m_3), где q — состояние из множества Q, а m_1, m_2, m_3 — целые числа. Здесь целое число m_i является содержимым счетчика с номером i. Если $q = q_0$, то счетчиковая конфигурация называется *начальной*, если $q = q_a$, то счетчиковая конфигурация называется *заключительной*. *Высотой* счетчиковой конфигурации (q, m_1, m_2, m_3) будем называть число $n = \max(m_1, m_2, m_3)$. Счетчиковая конфигурация (q, m_1, m_2, m_3) называется *правильной*, если $0 \leq m_1,\ 0 \leq m_2,\ 0 \leq m_3$, т. е. в каждом счетчике содержится целое неотрицательное число.

На множестве всех незаключительных счетчиковых конфигураций определяется *функция переходов* δ_C. Пусть (q, m_1, m_2, m_3) — произвольная счетчиковая конфигурация, $q \neq q_0$, и K — счетчиковая команда из множества Π, приписанная состоянию q. Тогда значение функции переходов $\delta_C(q, m_1, m_2, m_3)$ определяется следующим образом:

1. если $K = \langle q, x, \sigma, q' \rangle$, то $\delta_C(q, m_1, m_2, m_3) = (q', n_1, n_2, n_3)$, где $n_i = m_i + x$ в случае $i = \sigma$, и $n_i = m_i$ в случае $i \neq \sigma$;

2. если $K = \langle q, \sigma, q', q'' \rangle$, то $\delta_C(q, m_1, m_2, m_3) = (q', m_1, m_2, m_3)$ в том случае, если $m_\sigma = 0$, и $\delta_C(q, m_1, m_2, m_3) = (q'', m_1, m_2, m_3)$ в том случае, если $m_\sigma \neq 0$.

Вычислением 3-счетчиковой машины C на начальной счетчиковой конфигурации (q_1, m_1, m_2, m_3) называется последовательность (конечная или бесконечная) счетчиковых конфигураций

$$\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_i, \ldots\ , \tag{8.13}$$

удовлетворяющая следующим трем условиям:

1. $\alpha_1 = (q_1, m_1, m_2, m_3)$,

2. для любого члена последовательности $\alpha_i,\ i > 1$, выполняется равенство $\delta_C(\alpha_{i-1}) = \alpha_i$,

3. последовательность завершается счетчиковой конфигурацией α_N тогда и только тогда, когда α_N — заключительная счетчиковая конфигурация.

Вычисление (8.13) называется *правильным*, если все счетчиковые конфигурации α_i последовательности (8.13) являются правильными. Счетчиковая машина C *допускает начальную счетчиковую конфигурацию* $\alpha = (q_1, m_1, m_2, m_3)$, если вычисление машины C на счетчиковой конфигурации α завершается заключительной счетчиковой конфигурацией $(q_a, 0, 0, 0)$.

Счетчиковая машина C *допускает натуральное число* n, если C допускает $(q_1, n, 0, 0)$ и соответствующее вычисление правильное. Обозначим записью $L(C)$ множество всех натуральных чисел, допускаемых 3-счетчиковой машиной C. Одна из теорем Минского (см. [Мин71]) гласит: для любого рекурсивно перечислимого множества натуральных чисел R существует такая 3-счетчиковая машина C, что $R = L(C)$.

Машины Минского можно также использовать для описания некоторых классов сложности.

Для произвольного натурального числа n будем говорить, что вычисление (8.13) 3-счетчиковой машины C *ограничено высотой* n, если вычисление (8.13) является правильным, и высота каждой счетчиковой конфигурации $\alpha_i = (q, m_1, m_2, m_3)$ этого вычисления не превосходит n, т. е. на каждом шаге вычисления машины C содержимое любого ее счетчика является неотрицательным целым числом, не превосходящим n. Размером $size(C)$ счетчиковой машины $C = \langle Q, q_a, q_1, \Pi\rangle$ будем называть количество команд в множестве $\Pi \cup \{q_1\}$.

Лемма 52. *Язык*

$$ESC = \{C : C \text{ — 3-счетчиковая машина, вычисление которой на начальной счетчиковой конфигурации } (q_0, 0, 0, 0) \text{ правильно, конечно и ограничено высотой } 2^{2^{size(C)}}\}$$

является $EXPSPACE$*-полным.*

Доказательство. 1. Покажем, что $ESC \in EXPSPACE$. Очевидно, всякая 3-счетчиковая машина C с множеством состояний Q имеет $(|Q| - 1)(n + 1)^3$ различных правильных незаключительных счетчиковых конфигураций, высота которых не превосходит числа n. Поскольку в завершающемся вычислении счетчиковые конфигурации не должны повторяться, длина всякого завершающегося вычисления, ограниченного высотой

n, не превосходит величины $(|Q|-1)(n+1)^3+1$. Поэтому для любой 3-счетчиковой машины C из множества ESC, имеющей k состояний, ее вычисление на счетчиковой конфигурации $(q_0,0,0,0)$ завершается спустя не более $k\cdot 2^{3\cdot 2^k}$ шагов. При этом описание машины C (т. е. список ее команд) может быть задано двоичным словом длины $O(k\log k)$. Поэтому вычисление 3-счетчиковой машины C может быть воспроизведено машиной Тьюринга (МТ) M_C, которая наряду с входной лентой, содержащей описание 3-счетчиковой машины C, использует 5 рабочих лент. На первой ленте всегда записывается состояние машины C. На трех других рабочих лентах в двоичной системе счисления записывается содержимое счетчиков машины C. Пятая лента отводится для таймера. В начале вычисления МТ на первой рабочей ленте записан код начального состояния машины C, на рабочих лентах, моделирующих счетчики, записывается число 0, а на ленте таймера записывается число $k\cdot 2^{3\cdot 2^k}$ (в двоичной системе счисления). На каждом шаге работы МТ M_C моделирует выполнение очередной команды счетчиковой машины C, изменяя соответствующим образом содержимое четырех рабочих лент, а также уменьшает на 1 показание таймера. МТ M_C допускает описание 3-счетчиковой машины C, если в процессе вычисления на первой рабочей ленте оказывается записан код допускающего состояний q_0 машины C. МТ M_C отвергает описание 3-счетчиковой машины C, если в процессе вычисления либо значение одного из счетчиков выходит за пределы диапазона $[0:2^{2^k}]$, либо показания таймера оказываются равными 0. Таким образом, МТ M_C допускает описание 3-счетчиковой машины C в том и только в том случае, когда $C\in ESC$. При этом объем рабочей памяти, используемой МТ M_C, не превосходит величины $O(2^k)$. Следовательно, $ESC\in EXPSPACE$.

2. Покажем, что для любой одноленточной МТ M и любого входного слова $w=x_0x_1\ldots x_n$ существует такая 3-счетчиковая машина $C_{M,w}$, что соотношение $C_{M,w}\in ESC$ выполняется в том и только том случае, когда МТ M допускает слово w, используя при этом рабочую зону размера $2^{c(size(w)+size(M))}$, где c — некоторая константа, не зависящая от M и w.

Без ограничения общности мы будем полагать, что слова w, поступающие на вход МТ M, являются кодами двоичных слов в следующем кодировании. Каждое двоичное слово $\sigma_1\sigma_2\ldots\sigma_s$ вначале окаймляется слева и справа символами # (ограничителями рабочей зоны), т. е. оно представля-

ется в виде #$\sigma_1\sigma_2\ldots\sigma_s$#, и после этого символ # замещается комбинацией 11, символ 1 — комбинацией 01, а символ 0 — комбинацией 00. МТ M в процессе работы поддерживает указанное кодирование, перемещая соответствующим образом коды правого и левого ограничителя # в том случае, когда рабочая зона расширяется.

Устройство и принцип функционирования машины $C_{M,w}$ таковы. Каждой ленточной конфигурации $b_k\ldots b_1b_0qa_0a_1\ldots a_m$ МТ M соответствует счетчиковая конфигурация (q, N_1, N_2, a_0) 3-счетчиковой машины $C_{M,w}$, где натуральные числа

$$N_1 = \sum_{i=0}^{k} a_i 2^i \text{ и } N_2 = \sum_{j=0}^{m} b_i 2^i \ ,$$

будучи представленными в двоичной системе счисления, являются словами $b_0b_1\ldots b_k$ и $a_0a_1\ldots a_m$, записанными соответственно слева и справа от обозреваемой ячейки на ленте МТ (обозреваемая ячейка a_0 включается в правое слово). Вычисление счетчиковой машины $C_{M,w}$ состоит из трех этапов.

На первом этапе вычисления в первый счетчик записывается натуральное число, двоичным представлением которого является слово $w = x_n, \ldots, x_1, x_0$, а в третий счетчик записывается число x_0. Для каждого из битов $x_n, \ldots, x_1, x_0$ в машине $C_{M,w}$ задействованы не более 8 счетчиковых команд, которые позволяют при поддержке третьего счетчика преобразовать счетчиковую конфигурацию вида $(q, z, 0, 0)$ в счетчиковую конфигурацию вида $(q', 2z + x_i, 0, 0)$. Это осуществляется следующим образом. Вначале содержимое первого счетчика уменьшается до нуля, и при этом, всякий раз когда из первого счетчика вычитается 1, к содержимому третьего счетчика прибавляется 2. Для осуществления этой процедуры достаточно 4 команд счетчиковой машины. Затем содержимое третьего счетчика переносится в первый счетчик. Для этого достаточно 3 команд. В заключение к содержимому первого счетчика прибавляется число x_i. Для этого, возможно, нужна еще одна команда счетчиковой машины. В результате на первом этапе начальная счетчиковая конфигурация $(q_1, 0, 0, 0)$ преобразуется в счетчиковую конфигурацию $(q, z_w, 0, x_0)$, где $z_w = \sum_{i=0}^{n} x_i 2^i$. Эта счет-

чиковая конфигурация соответствует начальной ленточной конфигурации МТ M.

На втором этапе 3-счетчиковая машина $C_{M,w}$ моделирует работу МТ M. Каждой команде МТ M соответствует несколько (не более 27) отдельных команд счетчиковой машины $C_{M,w}$, которые позволяют воспроизвести эффект выполнения одной команды МТ. Например, выполнение команды $q10q'R$ (проверить, содержит ли обозреваемая ячейка ленты символ 1, и если это так, то записать в эту ячейку символ 0, сдвинуть считывающую головку на одну ячейку вправо и перейти в состояние q'), предусматривающей преобразование ленточной конфигурации $b_k \dots b_1 b_0 q 1 a_1 \dots a_m$ в ленточную конфигурацию $b_k \dots b_1 b_0 0 q' a_1 \dots a_m$, соответствует преобразованию счетчиковой конфигурации $(q, z_1, z_2, 1)$ в счетчиковую конфигурацию $(q', [\frac{z_1}{2}], 2z_2, a_1)$. Для осуществления этого преобразования содержимое третьего счетчика вначале сравнивается с 0. Если оно отлично от 0, то указанная команда МТ может быть выполнена. В противном случае счетчиковая машина переходит к моделированию альтернативной команды МТ M. Чтобы промоделировать выполнение команды $q10q'R$ счетчиковая машина $C_{M,w}$ выполняет следующие действия:

1. Обнуляет третий счетчик. В результате счетчиковая конфигурация $(q, z_1, z_2, 1)$ преобразуется в конфигурацию $(q_1, z_1, z_2, 0)$ с использованием одной счетчиковой команды.

2. Обнуляет первый счетчик, и при этом вместо каждой пары единиц, изъятой из первого счетчика, к содержимому третьего счетчика прибавляется 1. Поскольку число z_1, хранившееся в первом счетчику нечетно (это следует из того, что число z_1 кодирует двоичное слово $1a_1 \dots a_m$), при изъятии из первого счетчика последней непарной единицы содержимое третьего счетчика остается неизменным. В результате счетчиковая конфигурация $(q_1, z_1, z_2, 0)$ преобразуется в конфигурацию $(q_2, 0, z_2, [\frac{z_1}{2}])$ с использованием 5 счетчиковых команд.

3. Далее содержимое третьего счетчика переносится в первый счетчик. В результате счетчиковая конфигурация $(q_2, 0, z_2, [\frac{z_1}{2}])$ преобразуется

в конфигурацию $(q_3, [\frac{z_1}{2}], z_2, 0)$. Для этого необходимо еще 3 счетчиковых команды.

4. Далее содержимое второго счетчика удваивается и переносится в третий счетчик. Для этого необходимо обнулить второй счетчик, и при этом вместо каждой единицы, изъятой из второго счетчика, к содержимому третьего счетчика прибавляется 2. Это можно осуществить при помощи 4 счетчиковых команд. В результате счетчиковая конфигурация $(q_3, [\frac{z_1}{2}], z_2, 0)$ преобразуется в конфигурацию $(q_4, [\frac{z_1}{2}], 0, 2z_2)$.

5. Далее содержимое третьего счетчика переносится во второй счетчик. В результате счетчиковая конфигурация $(q_4, [\frac{z_1}{2}], 0, 2z_2)$ преобразуется в конфигурацию $(q_5, [\frac{z_1}{2}], 2z_2, 0)$. Для этого необходимо еще 3 счетчиковых команды.

6. Далее проверяется четность содержимого первого счетчика (это необходимо для того, чтобы знать самый левый символ слова $a_1 \dots a_m$, который должен обозреваться МТ после сдвига считывающей головки вправо). Для этого вновь обнуляется первый счетчик, но на этот раз вместо каждой пары единиц, изъятой из первого счетчика, к содержимому третьего счетчика также прибавляется пара единиц. Если последняя единица изъятая из первого счетчика оказывается непарной, то 3-счетчиковая машина $C_{M,w}$ прибавляет 1 к содержимому третьего счетчика и переходит в состояние $q_{6,1}$. Если же после изъятия очередной пары единиц из первого счетчика его содержимое оказалось равно 0, то 3-счетчиковая машина $C_{M,w}$ переходит в состояние $q_{6,0}$. В результате счетчиковая конфигурация $(q_5, [\frac{z_1}{2}], 2z_2, 0)$ преобразуется в конфигурацию $(q_{6,a_1}, 0, 2z_2, [\frac{z_1}{2}])$. Для этого необходимы еще 7 счетчиковых команд.

7. В заключение содержимое третьего счетчика переносится в первый счетчик, а затем в третий счетчик (в зависимости от состояния q_{6,a_1}) записывается число a_1. В результате счетчиковая конфигурация $(q_{6,a_1}, 0, 2z_2, [\frac{z_1}{2}])$ преобразуется в конфигурацию $(q_7, [\frac{z_1}{2}], 2z_2, a_1)$. Для этого необходимо использовать не более 7 счетчиковых команд.

Все остальные разновидности команд МТ M моделируются счетчиковой машиной $C_{M,w}$ аналогичным образом. Таким образом 3-счетчиковая машина $C_{M,w}$ воспроизводит все шаги вычисления МТ M на начальной конфигурации $q_1x_0x_1\dots x_n$.

Если M завершает вычисление, допуская входное слово w, то $C_{M,w}$ переходит к третьему этапу вычисления. На этом этапе стирается содержимое всех счетчиков, и после этого $C_{M,w}$ переходит в допускающее состояние. Для выполнения третьего этапа задействованы 6 счетчиковых команд.

Нетрудно убедиться, что построенная таким образом 3-счетчиковая машина $C_{M,w}$ обладает следующими двумя свойствами:

1) размер $C_{M,w}$ ограничен величиной $c(size(M) + size(w))$, где c — некоторая константа, не зависящая от M и w;

2) $C_{M,w}$ имеет завершающееся вычисление, длина которого ограничена числом $2^{2^{size(C)}}$, тогда и только тогда, когда МТ M допускает слово w, используя рабочую зону, размер которой не превосходит $2^{c(size(M)+size(w))}$, где константа c не зависит от M и входного слова w.

Таким образом, для всякого языка $L(M)$, распознаваемого машиной Тьюринга M и принадлежащего классу сложности $EXPSPACE$, включение $w \in L(M)$ имеет место тогда и только тогда, когда верно включение $C_{M,w} \in ESC$. Значит, язык ESC является $EXPSPACE$-полным. □

Не составляет труда убедиться в том, что описанная трансляция 3-счетчиковой машины C в МТ M_C, равно как и трансляция пары, состоящей из МТ M и входного слова w в 3-счетчиковую машину $C_{M,w}$, осуществляется за время, линейно зависящее от размеров исходных данных, и с использованием вспомогательной памяти, объем которой пропорционален величине, равной логарифму от размера исходных данных.

Теперь перейдем к следующему этапу доказательства основной теоремы и сопоставим произвольной 3-счетчиковой машине $C = \langle Q, q_a, q_1, \Pi\rangle$ размера n коммутативную полугруппу $\mathcal{G}'_C$, которая в определенном смысле описывает поведение этой машины. Образующими этой полугруппы является объединение двух непересекающихся множеств

$$\bar{Q} = Q \cup \{g_1, h_1, g_2, h_2, g_3, h_3\}.$$

Положим $e_n \stackrel{def}{=} 2^{2^n}$. Каждой счетчиковой конфигурации $\alpha = (q, m_1, m_2, m_3)$ сопоставим элемент описываемой полугруппы, который задается словом

$$w(q, m_1, m_2, m_3) = q g_1^{e_n - m_1} h_1^{m_1} g_2^{e_n - m_2} h_2^{m_2} g_3^{e_n - m_3} h_3^{m_3} .$$

Определяющие соотношения (тождества) $\mathcal{P}'_C$ полугруппы $\mathcal{G}'_C$ таковы.

Для каждой команды вида $K = \langle q, x, \sigma, q' \rangle$, где $q, q' \in Q$, $x \in \{-1, 1\}$, $\sigma \in \{1, 2, 3\}$, вводятся следующие тождества:

$$q g_\sigma \equiv q' h_\sigma \text{ при } x = 1 \qquad (A)$$

и

$$q h_\sigma \equiv q' g_\sigma \text{ при } x = -1. \qquad (B)$$

Для каждой команды вида $K = \langle q, \sigma, q', q'' \rangle$, где $q, q', q'' \in Q$, $\sigma \in \{1, 2, 3\}$, вводятся следующие тождества:

$$q h_\sigma \equiv q'' h_\sigma \qquad (C)$$

и

$$q g_\sigma^{e_n} \equiv q' g_\sigma^{e_n}. \qquad (D)$$

Положим также

$$W \stackrel{def}{=} \{w(q, z_1, z_2, z_3) : q \in Q, 0 \le z_1, z_2, z_3 \le e_n\}.$$

Пусть $\Phi(\alpha, s)$ — число вхождений символа s в слово α. Тогда для любого слова α из множества W выполняется соотношение $\Phi(\alpha, g_k) + \Phi(\alpha, h_k) = e_n$ при всех $k \in \mathcal{I}_3$. Тождества $(A) - (D)$ соответствуют командам счетчиковой машины, поэтому нетрудно показать, применяя индукцию по длине вывода, что для любой пары слов α и β, если $\beta \equiv \alpha \pmod{\mathcal{P}'_C}$ и $\alpha \in W$, то $\beta \in W$.

Отметим, что из принадлежности счетчиковой машины C множеству ESC следует, что $w(q_1, 0, 0, 0) \equiv w(q_a, 0, 0, 0) \pmod{\mathcal{P}'_C}$, поскольку вычислению $c^0, c^1, \ldots$ машины C соответствует последовательный вывод

$w(c^0) \to w(c^1) \to \ldots$, в котором используются тождества $(A) - (D)$. Поэтому справедлива

Лемма 53. $w(q_1, 0, 0, 0) \equiv w(q_a, 0, 0, 0)$ (mod $\mathcal{P}'_C$) $\Leftrightarrow C \in ESC$.

Таким образом, задача проверки принадлежности машины C множеству ESC сводима к проблеме тождества в построенной полугруппе $\mathcal{G}'_C$.

Теперь построим коммутативную полугруппу $\mathcal{G}_n$ с системой образующих мощности $O(n)$, содержащую такие образующие S, F, B, что FB^{e_n} — единственное слово, содержащее F, которое равно S в $\mathcal{G}_n$. Полугруппа $\mathcal{G}_n$ и ее образующие U_n строятся индукцией по числу n, используя соотношение $e_{n+1} = (e_n)^2$. Положим

$$U_0 \stackrel{def}{=} \{s, f, c_1, c_2, c_3, c_4, b_1, b_2, b_3, b_4\} \text{ и } \mathcal{P}_0 \stackrel{def}{=} \{sc_i \equiv fc_ib_i^2;\ i \in \mathcal{I}_4\}.$$

Для каждого $m > 0$ выберем множество различных символов

$$\{S, Q_1, Q_2, Q_3, Q_4, F, C_1, C_2, C_3, C_4, B_1, B_2, B_3, B_4\},$$

не содержащихся в U_{m-1}, и положим

$$U_m \stackrel{def}{=} U_{m-1} \bigcup \{S, Q_1, Q_2, Q_3, Q_4, F, C_1, C_2, C_3, C_4, B_1, B_2, B_3, B_4\}.$$

Элементы множества U_0 будем называть образующими нулевого уровня, а при $n > 0$ элементы множества $U_n \setminus U_{n-1}$ — образующими уровня n. При рассмотрении полугруппы $\mathcal{G}_n$ для удобства обозначений прописными буквами $S, \ldots, B_4$ будем обозначать образующие уровня n, а строчными буквами $s, \ldots, b_4$ — соответствующие им образующие $(n-1)$-го уровня. Тогда множество определяющих соотношений $\mathcal{P}_n$ полугруппы $\mathcal{G}_n$ состоит из всех определяющих соотношений $\mathcal{P}_{n-1}$ и следующих соотношений:

$$\begin{array}{rcll} S & \equiv & Q_1sc_1, & (a) \\ Q_1fc_1b_1 & \equiv & Q_2sc_2, & (b) \\ Q_2fc_2 & \equiv & Q_3fc_3, & (c) \\ Q_3sc_3b_1 & \equiv & Q_2sc_2b_4, & (d) \\ Q_3sc_3 & \equiv & Q_4fc_4b_4, & (e) \\ Q_4sc_4 & \equiv & F & (f) \end{array}$$

и при $i \in \mathcal{I}_4$

$$Q_2C_ifb_2 \equiv Q_2C_iB_ifb_3 \quad (g) - (j).$$

Лемма 54. *Пусть S, F, C_i, B_i при $i \in \mathcal{I}_4$ — образующие уровня n. Тогда*

$$SC_i \equiv FC_iB_i^{e_n} \pmod{\mathcal{P}_n} \text{ при } i \in \mathcal{I}_4.$$

Доказательство. Проводится индукцией по n.

При $n = 0$ полугруппа $\mathcal{G}_n$ имеет только одно определяющее соотношение, которое совпадает с доказываемым тождеством.

При $n > 0$ для всех $i \in \mathcal{I}_4$ получаем следующую цепочку тождеств

$SC_i \equiv C_iQ_1sc_1$	1) по эквивалентности (a),
$\equiv C_iQ_1fc_1b_1^{e_{n-1}}$	2) по предположению индукции,
$\equiv C_ib_1^{e_{n-1}-1}Q_2sc_2$	3) по эквивалентности (b),
$\equiv C_ib_1^{e_{n-1}-1}Q_2fc_2b_2^{e_{n-1}}$	4) по предположению индукции,
$\equiv C_ib_1^{e_{n-1}-1}Q_2fc_2b_3^{e_{n-1}}B_i^{e_{n-1}}$	5) по эквивалентностям (g)-(j),
$\equiv C_iB_i^{e_{n-1}}b_1^{e_{n-1}-1}Q_3fc_3b_3^{e_{n-1}}$	6) по эквивалентности (c),
$\equiv C_iB_i^{e_{n-1}}b_1^{e_{n-1}-1}Q_3sc_3$	7) по предположению индукции,
$\equiv C_iB_i^{e_{n-1}}b_1^{e_{n-1}-2}b_4Q_2sc_2$	8) по эквивалентности (d),
$\equiv \cdots \equiv C_iB_i^{e_{n-1}e_{n-1}}b_4^{e_{n-1}-1}Q_3sc_3$	9) повтор строк (4-8), кроме строки (8), $(e_{n-1}-1)$ раз, а строки ((8), $(e_{n-1}-2)$ раза,
$\equiv C_iB_i^{e_n}Q_4fc_4b_4^{e_{n-1}}$	10) по эквивалентности (e),
$\equiv C_iB_i^{e_n}Q_4sc_4$	11) по предположению индукции,
$\equiv FC_iB_i^{e_n}$	12) по эквивалентности (f).

□

Назовем цепочку выводов приведенной, если в ней не содержатся цепочки вида

$$\alpha \to \beta \to \alpha \pmod{\mathcal{P}}.$$

Далее будет доказано, что приведенные выводы из слов SC_i в $\mathcal{G}_n$ в слова, содержащие символ F уровня n, построенные в доказательстве леммы 54, единственно возможные. Для доказательства этого потребуется определить высоту слова и описать некоторые простые свойства этого понятия.

Пусть S, C_1 — символы уровня n, и пусть $\alpha \in U_n^*$ — такое слово, что $\alpha \equiv SC_1 \pmod{\mathcal{P}_n}$. Определим высоту такого слова α формулой

$$h(\alpha) \stackrel{def}{=} \min\{m \in \mathbb{N} \mid \Phi(\alpha, c_i) > 0 \text{ для некоторого } c_i \text{ уровня } m\}.$$

Напомним, что $\Phi(\alpha, s)$ — число вхождений символа s в слово α. Следующая лемма выводится непосредственно из соответствующих определений.

Лемма 55. *Пусть $SC_1 = \gamma_0 \to \gamma_1 \to \cdots \to \gamma_r = \alpha \pmod{\mathcal{P}_n}$ — приведенный вывод слова α из слова SC_1 в полугруппе $\mathcal{G}_n$. Тогда*

(i) $\sum_{i=1}^{4} \Phi(\alpha, c_i) = 1$, *если c_1, c_2, c_3, c_4 — символы уровня m, где $h(\alpha) \le m \le n$, и эта сумма равна 0 в остальных случаях.*

(ii) $\sum_{i=1}^{4} \Phi(\alpha, q_i) = 1$, *если q_1, q_2, q_3, q_4 — символы уровня m, где $h(\alpha) < m \le n$, и эта сумма равна 0 в остальных случаях.*

(iii) $\Phi(\alpha, s) + \Phi(\alpha, f) = 1$, *если s, f — символы уровня $h(\alpha)$, и эта сумма равна 0 в остальных случаях.*

(iv) $|h(\gamma_i) - h(\gamma_{i-1})| \le 1$ *для всех $i \in I_r$.*

(v) *К слову α можно применять определяющие соотношения только из $\mathcal{P}_{h(\alpha)+1} \setminus \mathcal{P}_{h(\alpha)-1}$. По определению считаем $\mathcal{P}_{-1} = \emptyset$. Высота убывает тогда и только тогда, когда применяются определяющие соотношения из $\mathcal{P}_{h(\alpha)}$.*

Лемма 56. *Пусть S, F, C_i, B_i, где $i \in \mathcal{I}_4$, — символы алфавита уровня n и $\alpha \in U_n^*$. Если $SC_i \equiv \alpha (\mathrm{mod} \mathcal{P}_n)$ и в слово α входит либо символ S, либо символ F, то $\alpha = SC_i$ или $\alpha = FC_iB_i^{e_n}$.*

Доказательство. При $n = 0$ утверждение леммы очевидно. Пусть $n > 0$ и имеется приведенный вывод

$$SC_1 = \gamma_0 \to \gamma_1 \to \cdots \to \gamma_r = \alpha \bmod \mathcal{P}_n. \tag{8.14}$$

Докажем индукцией по n, что этот вывод совпадает с выводом в лемме 54.

Заметим, что поскольку слово α содержит S или F, то из утверждения (iii) леммы 55 следует, что $h(\alpha) = n$. Далее, заметим, что кроме SC_1 и α в последовательности (8.14) нет других слов высоты n. Действительно, если

$0 < i < r$ — минимальное значение, при котором $h(\gamma_i) = n$, то согласно утверждению (iii) леммы 55 в слово γ_i входит либо S, либо F. Поскольку символы S и F встречаются только в определяющих соотношениях (a) и (f), а также ввиду утверждения (ii) леммы 55 получаем, что единственно возможный вывод из γ_i совпадает с обращением вывода из γ_{i-1} в γ_i, что противоречит приведенности вывода (8.14).

Единственно возможными определяющими соотношениями, применимыми к γ_0 и γ_1, является эквивалентность (a). Поэтому

$$\gamma_1 = C_1 Q_1 s c_1, \quad h(\gamma_1) = n-1, \quad h(\gamma_2) = n-2.$$

Согласно утверждению (iv) леммы 55 существует самое первое слово γ_{i_1} высоты $n-1$ после слова γ_1 в последовательности (8.14). Тогда, согласно утверждению (v) леммы 55, в цепочке выводов

$$\gamma_1 \to \cdots \to \gamma_{i_1} \text{ mod } \mathcal{P}_n$$

применяются только определяющие соотношения полугруппы $\mathcal{G}_{n-1}$. Поскольку в эти соотношения не входят символы уровня n, можно представить каждое слово γ_i из данной цепочки как $Q_1 C_1 \gamma_i'$, где $\gamma_i' \in U_{n-1}^*$. Поскольку $\gamma_1' = sc_1$, из утверждения (iii) леммы 55 следует, что слово γ_i' содержит либо s, либо f. Поэтому имеется приведенная цепочка выводов

$$sc_1 = \gamma_1' \to \cdots \to \gamma_{i_1}' \text{ mod } \mathcal{P}_{n-1}.$$

Следовательно, согласно предположению индукции для $n-1$ имеем

$$\gamma_{i_1}' = f c_1 b_1^{e_{n-1}} \quad \text{и} \quad \gamma_{i_1} = Q_1 C_1 f c_1 b_1^{e_{n-1}}.$$

Поскольку цепочка выводов (8.14) приведенная, к γ_{i_1} можно применить только определяющее соотношение (b) уровня n. Поэтому

$$\gamma_{i_1+1} = Q_2 C_1 b_1^{e_{n-1}-1} s c_2, \quad h(\gamma_{i_1+1}) = n-1, \quad h(\gamma_{i_1+2}) = n-2.$$

Пусть снова γ_{i_2} — первое слово после γ_{i_1} в цепочке выводов (8.14), имеющее высоту $n-1$. Повторяя ранее приведенные рассуждения, получаем, что в цепочке выводов (8.14) между словами γ_{i_1+1} и γ_{i_2} используются

определяющие соотношения только из полугруппы $\mathcal{G}_{n-1}$. Более того, во всех словах γ_i при $i_1 < i \leq i_2$ содержится символ c_2, поскольку только определяющие соотношения уровня n могут преобразовать c_2 в другие c_k. Поэтому согласно утверждению (i) леммы 55 определяющие соотношения (g) уровня $n-1$ неприменимы к словам γ_i при $i_1 < i \leq i_2$. Поэтому каждое слово γ_i при $i_1 < i \leq i_2$ можно представить в виде $Q_2 C_1 b_1^{e_{n-1}-1} \gamma_i'$, где $\gamma_i' \in U_{n-1}^*$, и имеется цепочка выводов

$$sc_2 = \gamma_{i_1+1}' \to \cdots \to \gamma_{i_2}' \bmod \mathcal{P}_{n-1}.$$

Согласно утверждению (iii) леммы 55 слово γ_{i_2}' содержит либо s, либо f и, следовательно, согласно предположению индукции

$$\gamma_{i_2} = Q_2 C_1 b_1^{e_{n-1}-1} f c_2 b_2^{e_{n-1}}.$$

Далее единственно возможным является применение k раз определяющего соотношения (g) уровня n, где $0 \leq k \leq e_n$, а затем однократное применение определяющего соотношения (c). В результате получится

$$\gamma_{i_3} = Q_3 C_1 B_1^k b_2^{e_{n-1}-k} f c_3 b_3^k.$$

Учитывая неприведенность вывода, к полученному слову можно применить только эквивалентность (f) уровня $n-1$. Поэтому $h(\gamma_{i_3+1}) = n-2$. Согласно утверждению (iv) леммы 55 существует наименьшее $i_4 > i_3$ такое, что $h(\gamma_{i_4}) = n-1$. Заметим также, что символ c_3 присутствует во всех словах от γ_{i_3} до γ_{i_4}. Поэтому определяющие соотношения (g) и (h) уровня $n-1$ неприменимы, и при $i_3 < i \leq i_4$ каждое слово γ_i можно представить как произведение $Q_3 C_1 B_1^k b_1^{e_{n-1}-1} b_2^{e_{n-1}-k} \gamma_i'$, где $\gamma_i' \in U_{n-1}^*$. Тогда

$$f c_3 b_3^k = \gamma_{i_3}' \to \cdots \to \gamma_{i_4}' \bmod \mathcal{P}_{n-1},$$

причем в слове γ_{i_4}' содержится один из символов s или f.

Предположим, что $\gamma_{i_4}' = f c_3 \eta$, где $\eta \neq b_3^k$ и $\eta \in U_{n-1}^*$. Тогда по лемме 54 существует цепочка выводов

$$sc_3 \to \cdots \to f c_3 b_3^k b_3^{e_{n-1}-k} \to \cdots \to f c_3 \eta b_3^{e_{n-1}-k} \bmod \mathcal{P}_{n-1}.$$

Поскольку $\eta b_3^{e_{n-1}-k} \neq b_3^{e_{n-1}}$, существование такой цепочки выводов противоречит предположению индукции.

Пусть теперь $\gamma'_{i_4} = sc_3\eta$, где $\eta \in U^*_{n-1}$. Тогда из леммы 54 следует существование цепочки выводов

$$sc_3 \to \cdots \to fc_3b_3^kb_3^{e_{n-1}-k} \to \cdots \to sc_3\eta b_3^{e_{n-1}-k} \text{ mod } \mathcal{P}_{n-1}.$$

Тогда по предположению индукции для $n-1$ получаем равенство $sc_3 = sc_3\eta b_3^{e_{n-1}-k}$. Следовательно, $\eta = 1$[1] и $k = e_{n-1}$. Поэтому

$$\gamma_{i_4} = Q_3C_1B_1^{e_{n-1}}b_1^{e_{n-1}-1}sc_3.$$

К полученному слову теперь применимы только определяющие соотношения (d) и (e). Но после использования соотношения (e) из этого слова невозможно получить слово высоты $n-1$. Поэтому

$$\gamma_{i_4+1} = Q_2C_1B_1^{e_{n-1}}b_1^{e_{n-1}-2}sc_2.$$

Если теперь применить эквивалентность (b), то получим цепочку, в которой не будет слов высоты $n-1$. Поэтому к полученному слову должна применяться эквивалентность (a) уровня $n-1$. Следовательно, $h(\gamma_{i_4+2}) = n-2$. Теперь можно повторить $e_{n-1}-1$ раз рассуждение, использованное для исследования подпоследовательности $\gamma_{i_1} \to \cdots \to \gamma_{i_1}$ mod $\mathcal{P}_n$. В результате получим при некотором $i_5 > i_4 + 2$

$$\gamma_{i_5} = Q_3C_1B_1^{e_{n-1}e_{n-1}}sc_3b_4^{e_{n-1}-1}.$$

Теперь применима только эквивалентность (e) и

$$\gamma_{i_5+1} = Q_4C_1B_1^{e_n}fc_4b_4^{e_{n-1}}.$$

Поскольку теперь можно применить эквивалентность (f) уровня $n-1$, получаем $h(\gamma_{i_5+2}) = n-2$. Тогда по утверждению (iv) леммы 55 существует наименьшее $i_6 > i_5+2$, такое, что $h(\gamma_{i_6}) = n-1$. Тогда из предположения индукции следует, что

$$\gamma_{i_6} = Q_4C_1B_1^{e_n}sc_4,$$

[1]Пустому слову сопоставим элемент 1.

и возможно только применение эквивалентности (f), что приводит к равенству

$$\gamma_{i_6+1} = \gamma_r = FC_1B^{e_n}.$$

Утверждение леммы для C_2, C_3 и C_4 доказывается аналогично. □

Сопоставим 3-счетчиковой машине $C = (Q, q_a, q_1, \Pi)$ с n состояниями коммутативную полугруппу $\mathcal{G}_C$, заданную образующими и соотношениями.

Образующими полугруппы $\mathcal{G}_C$ является несвязное объединение образующих $\overline{Q}$ полугруппы $\mathcal{G}'_C$ и образующих U_n полугруппы $\mathcal{G}_n$. Определим определяющие эквивалентности $\mathcal{P}_C$ как объединение эквивалентностей $\mathcal{P}_n$, описанных выше эквивалентностей $(A)--(C)$ для $\mathcal{P}'_C$, эквивалентности $g_k = B_k$ для $k \in \mathcal{I}_3$, а также следующих эквивалентностей.

Для каждого состояния $q \in Q$ вводим символы $q_r, q_e \notin \bar{Q} \cup \mathcal{G}_n$, а также символы

$$q_{11}, q_{12}, q_{13}, q_{14}, q_{a1}, q_{a2}, q_{a3}, q_{a4}$$

для начального и конечного состояний. Добавим соотношения

$$\begin{array}{rcll} q & \equiv & q_rFC_k, & (k) \\ q_rSC_k & \equiv & q_eSC_k, & (l) \\ q_eFC_k & \equiv & q' & (m) \end{array}$$

и

$$\begin{array}{rcll} q_{11} & \equiv & q_{12}SC_1, & (n) \\ q_{12}FC_1 & \equiv & q_{13}SC_2, & (o) \\ q_{13}FC_2 & \equiv & q_{14}SC_3, & (p) \\ q_{14}FC_3 & \equiv & q_1, & (q) \\ q_a & \equiv & q_{a4}FC_3, & (r) \\ q_{a4}SC_3 & \equiv & q_{a3}FC_2, & (s) \\ q_{a3}SC_2 & \equiv & q_{a2}FC_1, & (t) \\ q_{a2}SC_1 & \equiv & q_{a1} & (u) \end{array}$$

в представление $\mathcal{P}_C$. Эти соотношения завершают описание представления $\mathcal{P}_C$.

Напомним, что $W = \{w(q, z_1, z_2, z_3); q \in Q, 0 \leq z_1, z_2, z_3 \leq e_n\}$. Определим теперь $\mathcal{W}$ как подмножество элементов коммутативной полугруппы $\mathcal{P}'_C$, порожденное словами из множества W.

Лемма 57. *Существует гомоморфизм полугруппы $\mathcal{P}'_C$ в полугруппу $\mathcal{P}_C$, инъективный на $\mathcal{W}$.*

Доказательство. Пусть гомоморфизм ι отображает g_k в B_k при $k \in I_3$ и является тождественным отображением на $\bar{Q}$ в остальных случаях. Докажем, что ι мономорфно на $\mathcal{W}$, т.е. выполняются соотношения

$$w \equiv w' \ (\text{mod } \mathcal{P}'_C) \Leftrightarrow \iota(w) \equiv \iota(w') \text{ mod } \mathcal{P}_C.$$

Каждому определяющему соотношению $qg_k^{e_n} = q'g_k^{e_n}$ типа (D) в полугруппе $\mathcal{P}'_C$ соответствует аналогичное соотношение в $\mathcal{P}_C$, поскольку

$$\begin{array}{rcll} \iota(qg_k^{e_n}) = qB_k^{e_n} & \equiv & q_rFC_kB_k^{e_n} & \text{согласно } (k), \\ & \equiv & q_rSC_k & \text{согласно лемме 54}, \\ & \equiv & q_eSC_k & \text{согласно } (l), \\ & \equiv & q_eFC_kB_k^{e_n} & \text{согласно лемме 54}, \\ & \equiv & q'B_k^{e_n} = \iota(q'g_k^{e_n}) & \text{согласно } (m). \end{array}$$

Поскольку определяющие соотношения (A)–(D) из представления полугруппы $\mathcal{P}'_C$ также выполняются для полугруппы $\mathcal{P}_C$, для всех слов $w, w' \in \bar{Q}^*$

$$w \equiv w' \ (\text{mod}\mathcal{P}'_C) \ \Rightarrow \iota(w) \equiv \iota(w') \text{ mod } \mathcal{P}_C.$$

Чтобы доказать обратное утверждение, положим, что

$$\alpha = \iota(w(q, z_1, z_2, z_3)) \equiv \iota(w(q', z'_1, z'_2, z'_3)) = \beta \text{ mod } \mathcal{P}_C$$

и пусть $\alpha = \gamma_0 \to \gamma_1 \to \cdots \to \gamma_r = \beta$ mod $\mathcal{P}_C$ — приведенный вывод. Пусть $\gamma_m \to \gamma_{m+1}$ — самый первый шаг вывода из предыдущей цепочки, использующий определяющие соотношения (k)–(u). Тогда, поскольку W замкнуто относительно соотношений из $\mathcal{P}'_C$, выполняется соотношение $\iota^{-1}(\gamma_m) \equiv \iota^{-1}(\alpha)\ (\mathcal{P}'_C)$. Поэтому возможны следующие четыре варианта.

(i) Если к γ_m применяется соотношение (q), то из леммы 56 следует, что дальнейшие преобразования должны включать использование только соотношений $(p), (o), (n)$ из $\mathcal{P}_C$, т. е. единственно возможные слова, выведенные из γ_{m+1}, будут получены при замене

$$\begin{array}{llll} q_{04}FC_3B_3^{e_n} \text{ на } & q_{04}SC_3, & \text{далее на } & q_{03}FC_2, \text{ затем} \\ q_{03}FC_2B_2^{e_n} \text{ на } & q_{03}SC_2, & \text{далее на } & q_{02}FC_1, \text{ затем} \\ q_{02}FC_1B_1^{e_n} \text{ на } & q_{02}SC_1, & \text{далее на } & q_{01}. \end{array}$$

Поэтому достичь β в этом случае невозможно.

(ii) Те же рассуждения справедливы при замене соотношения (q) соотношением (r).

(iii) Если к γ_m применяется соотношение (k), то в дальнейшем должны использоваться определяющие соотношения из $\mathcal{P}_n$. Из леммы 56 следует, что дальнейшие два слова высоты n должны получаться при замене $qB_k^{e_n}$ на q_rSC_k и далее на q_eSC_k, где число k определяется по $\delta(q)$. Следующее слово высоты n получается при замене $q_eFC_kB_k^{e_n}$ на q_eSC_k. Теперь возможно использование только соотношения (m). Поэтому при некотором $m' > m$ выполняются соотношения

$$\gamma_{m'} \in W \text{ и } \iota^{-1}(\gamma_{m'}) \equiv \iota^{-1}(\gamma_m) \bmod \mathcal{P}'_C.$$

(iv) При использовании соотношения (m) все рассуждения п. (iii) сохраняются.

По индукции относительно длины вывода над $\mathcal{P}_C$ получаем, что $\iota^{-1}(\gamma_{m'}) \equiv \iota^{-1}(\beta) \bmod \mathcal{P}'_C$. Поэтому, из существования приведенного вывода $\alpha \to \beta \bmod \mathcal{P}_C$ следует, что

$$\iota^{-1}(\alpha) \equiv \iota^{-1}(\beta) \bmod \mathcal{P}'_C.$$

□

Лемма 58. *Пусть C — 3-счетчиковая машина и $\mathcal{P}_C$ — соответствующая ей полугруппа, которая описана выше. Тогда*

$$C \in ESC \Leftrightarrow q_{01} \equiv q_{a1} \bmod \mathcal{P}_C.$$

Доказательство. Согласно леммам 53 и 57

$$C \in ESC \Leftrightarrow \iota(w(q_0,0,0,0) \equiv \iota(w(q_a,0,0,0) \bmod \mathcal{P}_C.$$

Используя рассуждение из пп. (i)–(ii) леммы 57, получим

$$q_{01} \equiv q_{a1} \bmod \mathcal{P}_C \Leftrightarrow \iota(w(q_0,0,0,0)) \equiv \iota(w(q_a,0,0,0)) \bmod \mathcal{P}_C.$$

□

Таким образом, на основании доказанных лемм справедлива следующая

Теорема 59. *[MM82] Проблема равенства слов в коммутативной полугруппе является $EXPSPACE$-трудной относительно log-lin-сводимости.*

Доказательство. Лемма 52 показывает, что задача ESC проверки завершаемости вычислений 3-счетчиковых машин Минского с ограниченной емкостью счетчиков является $EXPSPACE$-трудной. Лемма 58 утверждает, что задача ESC сводима к задаче проверки тождеств специального вида в коммутативной полугруппе $\mathcal{P}_C$. Из определения полугруппы $\mathcal{P}_C$ видно, что размер описания определяющих ее тождеств пропорционален размеру программы соответствующей счетчиковой машины C. И, наконец, из определения полугруппы $\mathcal{P}_C$ и ее элементов q_{01}, q_{a1} видно, что для любой счетчиковой машины C построение определяющих соотношений этой полугруппы, а также слов, задающих элементы q_{01}, q_{a1}, можно осуществить эффективно, используя объем вспомогательной памяти, пропорциональный логарифму от размера программы машины C. □

Поскольку проблема тождеств в конечнопорожденных коммутативных полугруппах $log-lin$-сводима к проблеме проверки принадлежности многочлена идеалу, на основании доказанной теоремы и теоремы 61 мы приходим к следующему заключению.

Следствие 25. *[May89] Проблема проверки принадлежности многочлена идеалу является $EXPSPACE$-полной относительно $log - lin$-сводимости.*

8.9 Построение нормальной формы многочлена для заданных идеала и допустимого порядка на множестве термов

Задача 2. Для заданного базиса идеала $I \subseteq \mathbb{Q}[x_1, \ldots, x_n]$, допустимого порядка $<$ на термах, описанного n линейными формами с целыми ко-

эффициентами многочлена p найти единственную нормальную форму p относительно $(I, <)$.

Согласно теореме 3 допустимый порядок на термах может быть задан последовательностью векторов $a_1, \ldots, a_k$ из пространства $\mathbb{R}^n$, где $k \leq n$. Поскольку базис Гребнера относительно любого порядка конечен и при его вычислении используется только конечное множество термов, согласно замечанию 5.1.1 для описания допустимого порядка (только на этом конечном множестве термов) достаточно аппроксимировать вещественные векторы $a_1, \ldots, a_k$ рациональными. Умножив компоненты векторов на наименьшие общие кратные знаменателей каждого вектора, получаем, что при вычислении базиса Гребнера идеала $I = (f_1, \ldots, f_k) \subset \mathbb{Q}[x_1, \ldots, x_n]$ относительно некоторого порядка достаточно описывать упорядочение с помощью n линейных форм с целыми коэффициентами (см. теорему 2). Мы будем предполагать далее именно такое задание порядка на термах.

Теорема 60. *Для заданного базиса идеала I, допустимого порядка $<$ на термах и многочлена p единственная нормальная форма p относительно $(I, <)$ (задача 2) может быть вычислена с использованием не более чем экспоненциальной памяти.*

Для доказательства теоремы 60 используется результат Херманн (теорема 37) о максимальной степени D многочлена, являющегося решением системы полиномиальных уравнений, а также оценка степени нормальной формы из теоремы 43. В обоих случаях указанная степень D оценивается сверху двойной экспонентой от максимальной степени исходных многочленов.

При этом основным методом, используемым в доказательстве этой теоремы, является метод линеаризации (см. раздел 8.6), который сводит проблему нахождения соответствующего объекта к решению системы линейных уравнений порядка D.

Далее используется результат о том, что решение системы линейных уравнений размера $N \times N$ осуществимо алгоритмом, объем памяти которого ограничен величиной $poly(\log N)$, а величина N ограничена двойной экспонентой, от размера входных данных. Это лишь набросок плана доказательства, а некоторые подробности приведены в разделе 8.11.

8.10 Сложность по объему памяти решения систем линейных уравнений

В теории сложности алгоритмов установлено, что задачу проверки разрешимости системы линейных уравнений можно решить алгоритмом, объем рабочей памяти которого не превосходит величины $poly(\log N)$, где $poly$ — это некоторый полином, а N — размер описания системы линейных уравнений. Основные этапы обоснования этого результата таковы.

1. Для того чтобы проверить разрешимость системы линейных уравнений $A\mathbf{x} = \mathbf{b}$, достаточно показать, что $rk(A) = rk(A|\mathbf{b})$.

2. Для вычисления ранга матрицы известен параллельный алгоритм с полилогарифмическим временем работы и полиномиальным числом процессоров (в модели PRAM) [Mul87].

3. Согласно известной теореме (см. [**Golds**]) о параллельных вычислениях для PRAM параллельные вычисления с временем d могут быть промоделированы на машинах Тьюринга, объем рабочей памяти которых ограничен величиной $poly(d)$. Таким образом, разрешимость системы линейных уравнений может быть проверена алгоритмом с объемом рабочей памяти, не превосходящим некоторого полинома от логарифма длины входа, т.е. указанная задача принадлежит классу сложности POLYLOGSPACE.

Соответственно если размер системы линейных уравнений экспоненциален, то задача проверки разрешимости такой системы принадлежит классу сложности PSPACE, а если дважды экспоненциален, то классу EXPSPACE.

8.11 Теорема Майра

Теорема 61. $IM \in EXPSPACE$.

Доказательство. Вначале покажем, что задача IM сводима к задаче построения нормального базиса при помощи алгоритма использующего линейный объем памяти, затем докажем, что задача построения нормальной формы может быть решена с использованием экспоненциального объема памяти.

Задачу вычисления ранга матриц достаточно исследовать лишь для квадратных симметрических матриц, поскольку матрица

$$\begin{pmatrix} 0\, A \\ A^t 0 \end{pmatrix}$$

всегда является квадратной и симметрической, а ее ранг равен удвоенному рангу A.

Заметим, что ранг матрицы не меняется при расширении исходного поля. Расширим основное поле добавлением одного трансцендентного элемента x. Получим поле $G = F(x)$ рациональных функций одной переменной. Рассмотрим матрицу $C = XA$, где X — диагональная матрица, ненулевые элементы которой имеют вид $X_{ii} = x^{i-1}$. Поскольку матрица X невырождена, $rk(C) = rk(A)$. Более того, для матрицы C справедливо равенство

$$rk(CC) = rk(C). \tag{8.15}$$

Если матрицу C рассматривать как матрицу линейного эндоморфизма $V = G^n$, то из равенства (8.15) следует, что $\ker(C) \cap C(V) = \{0\}$ (где $C(V) = \{C(v) | v \in V\}$). Значит, $V = \ker(C) \oplus C(V)$, и ограничение эндоморфизма C на множество $C(V)$ является автоморфизмом $C(V)$. Последнее означает, что $\ker(C^k) = \ker(C)$ для любого $k \geq 1$. Отсюда вытекает, что $m = \dim(\ker(C))$, где m — такое максимальное целое такое, что t^m делит характеристический многочлен $Q_C(t)$.

Согласно определению характеристический многочлен произвольной матрицы M размера $n \times n$ определяется формулой:

$$Q_M(t) = \det(M - tI) = (-1)^n t^n + \sum_{i=1}^{n} a_i t^{n-i}.$$

Как показано в статьях [Pan87], [BCP83], характеристический многочлен матрицы размера $n \times n$ в модели вычисления PRAM строится за время $O(\log^2 n)$.

Поэтому, для вычисления ранга симметричной матрицы A можно воспользоваться следующим параллельным алгоритмом.

1. Построить матрицу $C = XA$, где X — матрица, указанного выше вида.

2. Вычислить характеристический многочлен $Q_C(t) = \det(C - tI)$.

3. Выдать в качестве ранга число $n - m$, где m — такая максимальная степень, что t^m делит $Q(t)$.

Как показано в статье [KM96], описанная в разделе 8.6 техника линеаризации в сочетании с рассмотренным методом вычисления ранга матрицы позволяет решать задачи нахождения минимального неприводимого многочлена (нормальной формы многочлена) и построения базиса Гребнера. При этом объем рабочей памяти разрешающего алгоритма может быть ограничен некоторой экспонентой, зависящей от длины входных данных.

Опишем метод построения нормальной формы многочлена относительно допустимого порядка.

Напомним, что порядок $\succ_A$ на множестве термов задается рациональной формой, максимальные значения числителей и знаменателей которой ограничены числом N. Пусть $I \subset K[X]$ — идеал в кольце многочленов, порожденный многочленами $f_1, \ldots, f_m$ степени не выше d. Согласно теореме 43 степень нормальной формы любого многочлена h относительно порядка $\succ_A$ не превосходит величины

$$D_1 = \left(\left(2n \left(\frac{d^2}{2} + d \right)^{2^{n-1}} \right)^n N^{2n} \deg(h) \right)^{n+1}. \qquad (8.16)$$

Обозначим нормальную форму многочлена h через $\mathsf{NF}(h)$. Поскольку $h - \mathsf{NF}(h) \in I$, согласно теореме Херманн (теорема 37)

$$h - \mathsf{NF}(h) = \sum_{i=1}^{m} f_i c_i,$$

для некоторых многочленов $c_i \in K[X]$, степени которых не превосходят величины

$$D_2 = \deg(h - \mathsf{NF}(h)) + (md)^{2^n} \leq D_1 + (md)^{2^n}.$$

Тогда справедлива следующая цепочка равенств

$$\begin{aligned} h &= \mathsf{NF}(h) + \sum_{i=1}^{m} f_i c_i \\ &= \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le D_1}} y_t t + \sum_{i=1}^{m} \left(\sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le d}} f_{i,t} t \right) \left(\sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le D_2}} c_{i,t} t \right) \\ &= \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le D_1}} y_t t + \sum_{i=1}^{m} \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le d+D_2}} \left(\sum_{\substack{u,v\in T\langle X\rangle \\ uv=t}} f_{i,u} c_{i,v} \right) t \\ &= \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le D_1}} y_t t + \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le d+D_2}} \left(\sum_{i=1}^{m} \sum_{\substack{u,v\in T\langle X\rangle \\ uv=t}} f_{i,u} c_{i,v} \right) t. \end{aligned} \tag{8.17}$$

Таким образом, для коэффициентов многочлена $h = \sum_{\substack{t\in T\langle X\rangle \\ \deg(t)\le \deg(h)}} h_t t$ выполняются равенства

$$h_t = y_t + \sum_{i=1}^{m} \sum_{\substack{u,v\in T\langle X\rangle \\ uv=t}} f_{i,v} c_{i,v}.$$

Далее нас будут интересовать лишь соотношения для коэффициентов h_t тех термов t, степени которых не превосходят величины $\max\{D_1, d+D_2\}$. Все остальные коэффициенты h_t равны 0.

Максимальная степень термов многочлена h не превосходит величины $\max\{D_1, d+D_2\}$, поэтому число уравнений не превышает величины $\max\{D_1, d+D_2\}^n$.

Полученную систему уравнений можно представить в матричном виде

$$H = \mathcal{F}C, \tag{8.18}$$

где H — вектор коэффициентов многочлена h, C — вектор неизвестных y_t и $c_{i,v}$, а $\mathcal{F}$ — матрица системы.

Из соотношений (8.17) следует, что общее количество неизвестных равно $D_1^n + mD_2^n$, а поскольку число уравнений ограничено величиной

$\max\{D_1, d + D_2\}^n$, размер матрицы $\mathcal{F}$ (максимум высоты и ширины) не превосходит величины $M = D_1^n + m(d + D_2)^m$.

Матрица $\mathcal{F}$ имеет экспоненциальные размеры, однако, не составляет труда вычислять отдельные ее элементы по мере необходимости (на лету).

Так, в строке, соответствующей терму t, и столбце, соответствующем переменной y_u, стоит 1, если $u = t$, и 0 в противном случае.

В строке, соответствующей терму t, и столбце, соответствующем переменной $c_{i,u}$, стоит $f_{i,\frac{t}{u}}$, если терм t делится на терм u, и 0 в противном случае.

Поскольку степени термов ограничены величиной $\max\{D_1, d + D_2\}$, объем памяти используемый для их записи не превосходит величины $2n \log \max(D_1, d + D_2) + 1)$.

Построим решение $y = \sum_t y_t t$ уравнения (8.18), минимальное относительно допустимого порядка на множестве термов. Такое решение и будет представлять нормальную форму многочлена h.

В качестве промежуточного шага на этом пути, требуется найти максимальный невырожденный минор $\mathcal{F}'$ матрицы $\mathcal{F}$, для которого соответствующее (единственное) решение матричного уравнения

$$H' = \mathcal{F}'C'$$

представляет нормальную форму. Для нахождения такого минора, потребуется удалить те строки и столбцы матрицы $\mathcal{F}$, которые могут быть выражены через оставшиеся. В действительности реального удаления строки или столбца не происходит, а лишь определяется принадлежит ли строка или столбец невырожденному минору $\mathcal{F}'$.

Чтобы определить, принадлежит ли k-я строка (или k-й столбец) максимальному невырожденному минору, достаточно сравнить ранги соответствующих матриц, содержащих первые $k - 1$ строк и первые k строк (первые $k - 1$ столбцов и первые k столбцов). Если ранги совпадают, то k-ю строку (k-й столбец) не добавляют, если не совпадают, то добавляют.

Матрица $\mathcal{F}'$ зависит от порядка расположения столбцов и строк в матрице $\mathcal{F}$. Для получения минимального относительно заданного допустимого порядка решения необходимо, чтобы порядок строк удовлетворял двум условиям

- Столбцы, соответствующие неизвестным y_t, следуют за столбцами, соответствующим неизвестным $c_{i,u}$.

- Столбцы, соответствующие переменным y_t, упорядочены согласно выбранному допустимому порядку на множестве термов $t \in T\langle X\rangle$.

□

Теорема 62. *Пусть в кольце многочленов $\mathbb{Q}[x_1, \ldots, x_n]$ идеал I задан m образующими, степени которых не превосходят d. Тогда нормальная форма h заданного многочлена f может быть вычислена алгоритмом, объем рабочей памяти которого ограничен величиной $O(\log^4((D_1 + m(d + D_1 + (md)^{2^n})^n) \log M))$, где D_1 определяется формулой* (8.16)*, а M — максимум числителей и знаменателей коэффициентов многочленов.*

Глава 9

Использование базиса Гребнера для решения систем алгебраических уравнений

Задача 3. Пусть задано множество многочленов $f_1, \ldots, f_s$ над полем K. Требуется решить систему уравнений

$$\begin{cases} f_1 = 0, \\ \ldots \quad \ldots \quad \ldots \\ f_s = 0. \end{cases}$$

Связь между задачей 1 (построением базиса Гребнера) и задачей 3 неоднозначна. С одной стороны, базисы Гребнера могут помочь в решении задачи 1. Действительно, в главе 6 был описан метод нахождения корней системы алгебраических уравнений, имеющей конечное число решений. Нахождение всех корней для идеала размерности 0, соответствующего системе уравнений (см. раздел 3.2), выполняется с помощью приведенного ниже алгоритма.

Рассмотрим следующий допустимый порядок на множестве мономов. Пусть $x_1 < \ldots < x_n$ и пусть для любого $k > 0$ и $0 < i < n$ выполняются неравенства $x_i^k < x_{i+1}$. Найдем базис Гребнера для выбранного порядка. Докажем, что минимальный приведенный многочлен этого базиса зависит только от переменной x_1. Действительно, пусть $a_1, \ldots, a_s$ —

всевозможные допустимые значения переменной x_1 на решениях. Тогда многочлен $f(x_1) = (x_1 - a_1) \cdot \ldots \cdot (x_1 - a_s)$ обращается в нуль во всех корнях системы уравнений. По теореме Гильберта о нулях существует некоторое $k > 0$, такое что $f^k(x_1)$ лежит в идеале I системы уравнений. Поскольку $x_1^j < x_i$ при $i > 1$, это возможно лишь в том случае, когда минимальный приведенный многочлен зависит только от переменной x_1. Более того, этот многочлен не имеет корней, отличных от a_i, где $i = 1, \ldots, s$.

Далее, подставляя любое $x_1 = a_i$ в качестве значений x_1, получаем новую систему относительно $n - 1$ неизвестных. Размерность идеала полученной системы также равна нулю. Повторяем приведенную выше процедуру для переменной x_2, находим многочлен $f(x_2)$ в идеале $I(a_i) \subset K[x_2, \ldots, x_n]$ и так до тех пор, пока не найдем все значения неизвестных.

Из описания алгоритма следует, что для нахождения одного решения системы алгебраических уравнений достаточно n раз (n — количество переменных в кольце многочленов) найти базис Гребнера и n раз решить степенное уравнение от одной переменной степени не выше, чем количество всех решений.

С другой стороны, в разделе 8.2 описан случай системы алгебраических уравнений, решение которой найти легко, а соответствующий этой системе базис Гребнера имеет экспоненциальный размер. Более того, использование метода линеаризации позволяет показать, что в общем случае рассчитывать на построение базиса Гребнера с использованием объема памяти, существенно меньшего экспоненциального (от длины входа), не приходится, поскольку в этом случае мы могли бы с тем же объемом памяти решить задачу IM, а она является $EXPSPACE$-трудной.

Приведем теперь несколько специальных случаев, когда сложность задачи решения систем алгебраических уравнений существенно меньше экспоненциальной. Всякий раз, когда удается понизить оценки сложности решения, используются специальные оценки степени полиномов в методе линеаризации, и затем применяются те же соображения о возможности решения линейных систем с полилогарифмическим от размера системы объемом памяти.

Случай систем булевых уравнений

Как мы уже видели в разделе 8.6 размер получаемой линейной систе-

мы оценивается экспонентой от длины входа. Это означает, что разрешимость таких систем может быть проверена с памятью, объем которой не превышает полинома от размера входных данных. Базисы Гребнера для идеалов, порождаемых булевыми полиномами также могут быть найдены алгоритмами, использующими полиномиальный объем памяти.

Случай систем с конечным числом решений

Для этого случая также известны экспоненциальные верхние оценки размера линейной системы в методе линеаризации [Pan87], [May97]. Поэтому для решения таких систем существуют алгоритмы с полиномиальной памятью.

Случай проверки неразрешимости систем алгебраических уравнений.

Пусть I — идеал, порожденный многочленами $f_1, \ldots, f_s$ с рациональными коэффициентами.

В [Bro87] показано, что если $m = \min\{s, n\}$, $d = \max\{\deg(f_1), \ldots, \deg(f_s)\}$ и f_i не имеют общего нуля в $\mathbb{C}^n$, то найдутся многочлены $q_1, \ldots, q_s$ с такими рациональными коэффициентами, что $1 = q_1 f_1 + \ldots + q_s f_s$ и $\deg(q_i) \leq mnd^m + md$, при $i = 1, \ldots, s$.

Это означает, что получаемая в методе линеаризации система линейных уравнений имеет размер, ограниченный экспонентой от длины входа. Как и прежде, это влечет возможность решения таких систем с полиномиальной памятью.

Случай идеала, порожденного однородными многочленами

В случае, когда все многочлены с рациональными коэффициентами, порождающие данный идеал, являются однородными, задача принадлежности многочлена данному идеалу является $PSPACE$-полной [May97].

Глава 10

Приложения

10.1 Элементы алгебры

10.1.1 Моноиды

Пусть M — произвольное множество. Отображение

$$M \times M \to M \tag{10.1}$$

будем называть операцией.

Если для всех пар x, y из M выполняются соотношения $xy = yx$, то операция называется коммутативной.

Для операции могут использоваться мультипликативная и аддитивная запись — xy и $x + y$ соответственно. Аддитивная запись обычно используется только для коммутативной операции, т.е. когда $x + y = y + x$ для всех пар (x, y).

Операция называется ассоциативной, если для всех троек (x, y, z) выполняются соотношения $(xy)z = x(yz)$.

Элемент $e \in M$ называется единичным элементом, если для всех $x \in M$ выполняются соотношения $ex = x = xe$. В случае аддитивной записи операции единичный элемент обозначается через 0 и называется нулевым элементом. Единичный элемент единствен, если он существует. Действительно, пусть e' — другой единичный элемент. Тогда выполняются

соотношения

$$e = ee' = e'.$$

Определение 10.1.1. *Множество M с ассоциативной операцией называется полугруппой. Полугруппа, содержащая единицу, называется моноидом.*

Из определения следует, что моноид всегда не пуст. В мультипликативном моноиде M единичный элемент обычно будем обозначать символом 1_M, или просто 1, если из контекста ясно, о каком моноиде идет речь.

Определение 10.1.2. *Гомоморфизмом полугрупп M и N называется отображение $f : M \to N$, для которого $f(xy) = f(x)f(y)$. Гомоморфизмом моноидов M и N называется отображение $f : M \to N$, для которого*

- $f(xy) = f(x)f(y)$;
- $f(1_M) = 1_N$.

Взаимно однозначный гомоморфизм называется мономорфизмом. Гомоморфизм, являющийся взаимно однозначным соответствием, называется изоморфизмом.

Отметим, что в правых и левых частях соотношений определения 10.1.2 правильнее было бы использовать разные обозначения для единиц и операций: в левой части единица и операция относятся к моноиду M, а справа — к N. В дальнейшем, если это не приведет к путанице, мы не будем использовать различные записи для единиц и операций.

Пример. Пусть $N = \{1, x\}$ — множество, состоящее из двух элементов — единицы 1 и элемента x. Операцию умножения зададим таблицей

$$1 \cdot 1 = 1, \quad 1 \cdot x = x \cdot 1 = x, \quad xx = x.$$

Легко проверить, что N является моноидом относительной такой операции. Теперь в качестве M возьмем тривиальный моноид из одного элемента. Тогда отображение $f : M \to N$, заданное формулой $f(1) = x$,

является гомоморфизмом полугрупп, но не является гомоморфизмом моноидов, поскольку образом единицы моноида M не является единица моноида N.

Пусть M — моноид и $x_1, \ldots, x_n$ — его элементы (где $n > 0$ — целое число). Определим их произведение по индукции. Если $n = 1$, то положим по определению

$$\prod_{i=1}^{1} x_i = x_1.$$

При $n > 1$, положим

$$\prod_{i=1}^{n} x_i = x_1 \ldots x_n = (x_1 \ldots x_{n-1})x_n.$$

Легко доказать индукцией по k, что при $k + m \leq n$ выполняется следующее правило.

Правило умножения.

$$\prod_{i=1}^{m} x_i \cdot \prod_{j=1}^{k} x_{m+j} = \prod_{i=1}^{m+k} x_i.$$

Данное правило означает, что в произведении элементов моноида скобки можно расставлять произвольным образом.

Если определенная на моноиде M операция является коммутативной, то такой моноид будем называть коммутативным.

Легко доказать по индукции следующее правило умножения в коммутативном моноиде.

Правило умножения в коммутативном моноиде. Пусть

$$\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

произвольная перестановка (взаимнооднозначное соответствие). Тогда для любых элементов $x_1, \ldots, x_n$ коммутативного моноида выполняется соотношение

$$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} x_{\sigma(i)}.$$

Неформально это правило означает, что результат умножения произвольного конечного набора элементов коммутативного моноида не зависит ни от порядка, ни от группировки сомножителей (расстановки скобок).

Из правила умножения в коммутативном моноиде следует, что для любого конечного непустого набора S элементов коммутативного моноида M корректно определено произведение

$$\prod_{x\in S} x.$$

Для пустого набора S положим по определению

$$\prod_{x\in S=\emptyset} x = 1.$$

Тогда для любых конечных непересекающихся наборов S и T элементов коммутативного моноида M

$$\prod_{x\in S\cup T} x = \prod_{x\in S} x \cdot \prod_{x\in T} x.$$

Для любого множества X и коммутативного аддитивного моноида M обозначим через $M(X)$ множество функций $f : X \to M$, обращающихся в нуль почти всюду. Сумму элементов f и g в этом множестве зададим формулой

$$(f + g)(x) = f(x) + g(x) \qquad \forall x \in X.$$

Легко проверить, что такая операция сложения задает на $M(X)$ структуру коммутативного моноида, нулевым элементом которого является нулевая функция 0, определяемая формулой $0(x) = 0$ для всех $x \in X$.

Введем мультипликативную запись для определенной выше операции в моноиде $M(X)$. Для этого каждой функции $\nu \in M(X)$ сопоставим символ X^ν. Составленное из этих символов множество обозначим через X^M_{fin}. (Отметим, что в случае конечного множества X выполняется равенство $X^M_{fin} = X^M$). Зададим в множестве X^M_{fin} умножение формулой $X^\mu \cdot X^\nu \stackrel{\text{def}}{=} X^{\mu+\nu}$. Тогда соответствие $\nu \mapsto X^\nu$ задает изоморфизм моноидов $M(X)$ и X^M_{fin}. Поскольку X^0 — единица моноида X^M, то в соответствии с принятым соглашением $X^0 = 1$.

Для любой функции $\nu \in M(X)$ и любого элемента $x \in X$ определим функцию ν_x формулой

$$\nu_x(y) = \begin{cases} \nu(x) & \text{при} \quad x = y \\ 0 & \text{при} \quad x \neq y \end{cases}$$

Для любой функции $\nu \in M(X)$ положим по определению

$$x^{\nu(x)} \stackrel{\text{def}}{=} X^{\nu_x}.$$

Подставим в эту формулу $\nu_x = 0$. Получим,что для любого $x \in X$ выполняется равенство $x^0 = X^0 = 1$.

Для любого $\nu \in M(X)$ сумма

$$\nu = \sum_{x \in X} \nu_x,$$

содержит только конечное число ненулевых слагаемых, по правилу умножения в коммутативном моноиде выполняется соотношение

$$X^{\nu} = X^{\sum\limits_{x \in X} \nu_x} = \prod_{x \in X} X^{\nu_x} = \prod_{x \in X} x^{\nu(x)}, \tag{10.2}$$

причем произведение содержит только конечное число сомножителей не равных единице.

Рассмотрим теперь аддитивный моноид $\mathbb{Z}_+$, элементами которого являются неотрицательные целые числа. Для любого элемента $x \in X$ определим функцию $\delta_x \in X_{fin}^{\mathbb{Z}_+}$ формулой

$$\delta_x(y) = \begin{cases} 1 & \text{при} \quad x = y \\ 0 & \text{при} \quad x \neq y \end{cases}.$$

Тогда $X^{\delta_x} = x^{\delta_x(x)} = x^1$. Положим по определению $x^1 = x$. Таким образом, $x^{\delta_x(x)} = x$.

Для любой функции $\nu \in X_{fin}^{\mathbb{Z}_+}$ и любого $x \in X$ выполняется соотношение

$$\nu_x = \underbrace{\delta_x + \ \ldots + \delta_x}_{\nu(x) \text{ слагаемых}}.$$

Поэтому

$$x^{\nu(x)} = X^{\nu_x} = X^{\overbrace{\delta_x + \ \ldots + \delta_x}^{v(x) \text{ слагаемых}}} = \underbrace{X^{\delta_x} \cdot \ \ldots \cdot X^{\delta_x}}_{v(x) \text{ сомножителей}} = \underbrace{x \cdot \ \ldots \cdot x}_{v(x) \text{ сомножителей}} .$$

Иными словами, для любого $n \in \mathbb{Z}_+$ выполняется равенство $x^n = \underbrace{x \cdot \ \ldots \cdot x}_{n \text{ сомножителей}}$. Для $n = 0$ это равенство означает, что произведение пустого множества сомножителей равно 1.

10.1.2 Группы. Кольца. Поля

Определение 10.1.3. *Группой называется моноид G, в котором для любого $x \in G$ существует $y \in G$, для которого $xy = yx = 1$ (или $x + y = y + x = 0$ в аддитивной записи). Элемент y называется обратным к x и обозначается через x^{-1} (через $-x$ в аддитивной записи). Коммутативный моноид, являющийся группой называется коммутативной или абелевой группой.*

Легко видеть, что обратный элемент для x единствен. Действительно, пусть существует другой обратный, скажем y'. Тогда

$$y = y \cdot 1 = y(xy') = (yx)y' = 1 \cdot y' = y'.$$

Определение 10.1.4. *Подгруппой группы G называется подмножество G, являющееся группой относительно той же групповой операции.*

Примеры.

1 Моноид $\mathbb{N}$ натуральных чисел с операцией сложения не является группой. Множество $\mathbb{Z}$ целых чисел с операцией сложения является абелевой группой. Эта группа называется аддитивной группой целых чисел.

2 Моноид ненулевых целых чисел с операцией умножения не является группой. Множество ненулевых рациональных чисел с операцией умножения является группой.

Определение 10.1.5. *Кольцо*[1] *— это множество* A *с двумя операциями, называемыми сложением и умножением, обладающее следующими свойствами:*

1 *относительно сложения* A *— абелева группа.*

2 *относительно умножения* A *— полугруппа.*

3 *для любых* $x, y, z \in A$ *выполняются два свойства дистрибутивности*

$$z(x+y) = zx + zy \quad \text{и} \quad (x+y)z = xz + yz.$$

Если операция умножения в кольце коммутативна, то такое кольцо называется коммутативным.

Определение 10.1.6. *Коммутативное кольцо* K*, в котором множество ненулевых элементов является группой относительно умножения, называется полем.*

Примеры.

1 Множество целых чисел $\mathbb{Z}$ с операциями сложения и умножения является коммутативным кольцом. Это кольцо называется кольцом целых чисел.

2 Коммутативное кольцо целых чисел $\mathbb{Z}$ не является полем.

3 Кольцо рациональных чисел $\mathbb{Q}$ является полем и называется полем рациональных чисел.

Определение 10.1.7. *Отображение колец*

$$\varphi : A \to B,$$

являющееся моноидным гомоморфизмом для мультипликативных структур на A *и* B*, а также моноидным гомоморфизмом для аддитивных структур на* A *и* B *называется кольцевым гомоморфизмом.*

[1] Если умножение не является ассоциативным, то такие кольца называют неассоциативными.

Взаимно однозначный гомоморфизм колец называется мономорфизмом колец, а гомоморфизм, осуществляющий отображение на, называется эпиморфизмом. Гомоморфизм колец, являющийся одновременно мономорфизмом и эпиморфизмом, называется изоморфизмом. Изоморфизм колец A и B будем записывать формулой $\varphi : A \xrightarrow{\approx} B$.

Из определения следует, что гомоморфизм колец преобразует нулевой элемент кольца в нулевой, а единичный в единицу. Поэтому, гомоморфизм полей всегда является мономорфизмом.

10.1.3 Термы. Многочлены

Пусть X — конечное множество.

Определение 10.1.8. *Термом на множестве X будем называть произвольную функцию $\nu : X \to \mathbb{Z}_+$, равную нулю почти всюду.*

Множество термов на X будем обозначать через $T\langle X\rangle$. Согласно обозначениям раздела 10.1.1 выполняется равенство

$$T\langle X\rangle = X^{\mathbb{Z}_+}_{fin}.$$

Там же было показано, что это множество является коммутативным моноидом. Каждый терм ν на множестве X согласно формуле (10.2) можно отождествить с выражением

$$\prod_{x\in X} x^{\nu(x)}.$$

Поэтому терм ν можно записывать также как X^ν.

Определение 10.1.9. *Степенью терма ν называется*

$$\deg(\nu) = \deg(X^\nu) = \sum_{x\in X} \nu(x).$$

Степенью терма ν по переменной $x \in X$ называется

$$\deg_x(\nu) = \deg_x(X^\nu) = \nu(x).$$

Из определения 10.1.9 следует, что $\deg(\nu) = \sum_{x \in X} \deg_x(\nu)$.

Определение 10.1.10. *Пусть K — коммутативное кольцо. Отображение*

$$p : T\langle X\rangle \to K,$$

равное нулю всюду, кроме конечного числа термов, называется многочленом от переменных из множества X с коэффициентами в кольце K. Если отбросить ограничение о равенстве отображения нулю почти всюду, то получим формальный ряд от переменных из множества X с коэффициентами в кольце K. Многочлен, соответствующий нулевому отображению, называется нулевым многочленом. Множество многочленов с коэффициентами в кольце K будем обозначать через $K[X]$ и называть множеством многочленов от переменных из множество X или множеством многочленов от n переменных, если множество X конечно и состоит из n элементов. Если $X = \{x\}$, то множество многочленов $K[X]$ обозначим также через $K[x]$ и будем называть множеством многочленов от одной переменной.

Каждый многочлен $p \in K[X]$ отождествим с выражением

$$p \stackrel{\text{def}}{=} p(X) \stackrel{\text{def}}{=} \sum_{\nu \in T\langle X\rangle} a_\nu X^\nu, \text{ где } a_\nu = p(\nu) \in K. \tag{10.3}$$

Многочлен вида aX^ν для $a \in K$ и $\nu \in T\langle X\rangle$ называется мономом от переменных X или, кратко, мономом. Если $a = 0$, то такой моном назовем нулевым.

Определение 10.1.11. *Слагаемые $a_\nu X^\nu$ в формуле (10.3), для которых $a_\nu \neq 0$, будем называть ненулевыми мономами многочлена p, а соответствующие термы X^ν — термами многочлена p.*

Согласно определению 10.1.10 многочлен, все слагаемые которого нулевые, является нулевым многочленом.

Множество термов многочлена $p \in K[X]$, очевидно, определяется формулой

$$T_p\langle X\rangle = \{\nu \in T\langle X\rangle \mid p(\nu) \neq 0\}. \tag{10.4}$$

По определению 10.1.10 для нулевого многочлена это множество пустое $T_0\langle X\rangle = \emptyset$, а для ненулевого многочлена — конечное и непустое.

Формулу (10.3) представления многочлена как суммы мономов можно переписать в виде суммы ненулевых мономов

$$p = p(X) = \sum_{\nu \in T_p\langle X\rangle} p(\nu)X^\nu. \tag{10.5}$$

Максимум степеней термов ненулевого многочлена p определен корректно, называется его степенью и обозначается через $\deg p$. Для нулевого многочлена $\deg p = -\infty$. Примем соглашение, что

- $-\infty + k = -\infty$ для всех целых неотрицательных чисел k;
- $-\infty + -\infty = -\infty$.

Также для ненулевого многочлена p и любой его переменной $x \in X$ корректно определена величина $\max\limits_{\nu \in T_p(X)} \deg_x X^\nu$. Эта величина называется степенью многочлена относительно переменной x и обозначается через $\deg_x p$. Для нулевого многочлена принимаем по определению, что $\deg_x p = -\infty$.

Определение 10.1.12. *На множестве многочленов $K[X]$ определены две операции. Суммой многочленов p и q называется многочлен r, определяемый формулой*

$$r(\nu) = p(\nu) + q(\nu) \quad \forall \nu \in T\langle X\rangle .$$

Произведением многочленов p и q называется многочлен s, определяемый формулой

$$s(\nu) = \sum_{\mu,\lambda \in T\langle X\rangle | \mu+\lambda=\nu} p(\mu)q(\lambda) \quad \forall \nu \in T\langle X\rangle .$$

Следующие утверждения выводятся непосредственно из определения 10.1.12.

Предложение 23. *Введенные выше операции сложения и умножения на множестве многочленов $K[X]$ задают на нем структуру кольца.*

Предложение 24. *Если кольцо K не имеет делителей нуля, то для любых многочленов $p, q \in K[X]$ выполняются соотношения*

- $\deg(pq) = \deg p + \deg q$;
- $\deg(p+q) \le \max\{\deg p, \deg q\}$.

Предложение 25. *Если кольцо K не имеет делителей нуля, то для любых многочленов $p, q \in K[X]$ и любой переменной $x \in X$ выполняются соотношения*

- $\deg_x(pq) = \deg_x p + \deg_x q$;
- $\deg_x(p+q) \le \max\{\deg_x p, \deg_x q\}$.

Напомним, что множество $\mathbb{Z}_+^n$ обозначает множество неотрицательных целочисленных векторов длины n. На этом множестве имеется естественная структура моноида, определяемая операцией сложения векторов.

Пусть $X = \{x_1, \ldots, x_n\}$. Определим отображение $f : X^{\mathbb{Z}_+} \to \mathbb{Z}_+^n$ формулой

$$f(X^\nu) = (\nu(x_1), \ldots, \nu(x_n)).$$

Легко видеть, что это отображение является изоморфизмом моноидов. Следовательно, мультипликативный моноид $X^{\mathbb{Z}_+}$ отождествляется с аддитивным моноидом целочисленных векторов $\mathbb{Z}_+^n$. Следующая формула позволяет отождествить множество термов на X с множеством целочисленных неотрицательных векторов длины n

$$X^\omega \stackrel{\text{def}}{=} x_1^{k_1} \cdot \ldots \cdot x_n^{k_n}, \quad \omega = (k_1, \ldots, k_n) \in \mathbb{Z}_+^n. \tag{10.6}$$

Поэтому многочлен $p \in K[X]$ от n независимых переменных $X = \{x_1, \ldots, x_n\}$ можно представить в виде выражения

$$p(X) = p(x_1, \ldots, x_n) = \sum_{\omega \in \mathbb{Z}_+^n} a_\omega X^\omega,$$

содержащего конечное число ненулевых слагаемых.

В силу определения терма многочлена (см. определение 10.1.11), выражение $x_1^{k_1} \cdot \ \ldots \cdot x_n^{k_n}$ является термом многочлена

$$p(x_1,\ \ldots, x_n) = \sum_{(k_1,\ldots,k_n)\in\mathbb{Z}_+^n} a_{k_1,\ldots,k_n} x_1^{k_1} \cdot \ \ldots \cdot x_n^{k_n},$$

только если $a_{k_1,\ldots,k_n} \neq 0$, а соответствующее слагаемое $a_{k_1,\ldots,k_n} x_1^{k_1} \cdot \ \ldots \cdot x_n^{k_n}$ в этом случае будет мономом этого многочлена. Для степени многочлена $p(x_1,\ \ldots, x_n)$ по переменной x_i будем обозначать также обозначение $\deg_i p$, т.е. $\deg_i p = \deg_{x_i} p$.

Предложение 26. *Пусть $f : A \to B$ — гомоморфизм колец. Тогда отображение колец многочленов $f^* : A[x] \to B[x]$, заданное формулой*

$$f^*\left(\sum_{i=0}^{\infty} a_i x^i\right) = \sum_{i=0}^{\infty} f(a_i) x^i, \tag{10.7}$$

является гомоморфизмом. Если f является изоморфизмом, то этот гомоморфизм является изоморфизмом.

Доказательство. Следует непосредственно из определений. □

Определение 10.1.13. *Пусть B_0 — произвольное подмножество кольца B и n — неотрицательное целое число. Подмножество элементов из B вида*

$$\prod_{i=1}^{n} b_i$$

где $b_i \in B_0$, обозначим через $T_n\langle B_0\rangle$ и будем называть множеством термов степени n от (зависимых) переменных B_0 (из кольца B). Множество всех термов от переменных B_0 из кольца B определим формулой

$$T\langle B_0\rangle = \bigcup_{i=0}^{\infty} T_n\langle B_0\rangle .$$

Определение 10.1.14. *Пусть задано вложение колец $A \subset B$ и подмножество B_0 в B. Подмножество элементов из B вида*

$$\sum_{b \in T\langle B_0 \rangle} \alpha_b b,$$

где α_b — произвольные элементы кольца A и только конечное число сомножителей α_a не равно нулю, является кольцом. Это кольцо называется кольцом многочленов от (зависимых) переменных B_0 (из кольца B) с кольцом коэффициентов A и обозначается через $A[B_0]$.

Каждому элементу $b \in B_0$ сопоставим символ x_b, так что различным b сопоставляются различающиеся символы. Составим из всех этих символов множество X_{B_0}. Тогда определено кольцо многочленов от переменных из множества X_{B_0}. Легко проверить, что формула $\varphi(x_b) = b$ определяет эпиморфизм колец

$$\varphi : k[X_{B_0}] \to k[B_0]. \tag{10.8}$$

Следовательно, имеется изоморфизм

$$k[X_{B_0}]/\ker\varphi \approx k[B_0].$$

10.2 Немного о сложности вычислений

Ниже будут приведены некоторые сведения из теории сложности вычислений.

Классические модели вычислений – машины Тьюринга, машины Минского, RAM (машины с произвольным доступом к памяти) – предполагаются известными. В этом ряду машины Минского несколько менее популярны (и в то же время активно используются в приводимых доказательствах), поэтому мы кратко напоминаем их определение.

Счетчиковые машины (машины Минского) — это одна из наиболее простых алгоритмически полных моделей вычисления. Счетчиковая машина представляет собой вычислительное устройство, которое состоит из

программы, имеющей конечное число команд, и конечного набора счетчиков. В каждом счетчике может храниться произвольное целое число. Программа счетчиковой машины — это автомат с конечным числом состояний; каждому состоянию q этой программы приписана команда одного из следующих трех типов:

1. увеличить на 1 число m_i, содержащееся в i-ом счетчике, и перейти в состояние q';

2. уменьшить на 1 число m_i, содержащееся в i-ом счетчике, и перейти в состояние q';

3. проверить, содержится ли в i-ом счетчике число 0, и в зависимости от результата проверки перейти в либо состояние q', либо в состояние q''.

Счетчиковые машины разработал и исследовал американский математик М. Минский в 1958 г. В монографии [Мин71] он показал, что всякая частично-рекурсивная функция вычислима машиной, снабженной всего лишь двумя счетчиками. Более подробно эти машины описаны в соответствующем разделе при доказательстве $EXPSPACE$-полноты задачи $CWEP$.

В методе линеаризации решения алгебраических уравнений, используемом для получения верхних оценок сложности решения алгебраических уравнений, существенно используется модель параллельных вычислений PRAM и так называемый тезис о параллельных вычислениях. Приведем кратко некоторые сведения об этих понятиях.

PRAM (параллельная машина с произвольным доступом к памяти) — состоит из p идентичных RAM (называемых процессорами), имеющих общую память. Каждый процессор работает по своей программе (обычный набор операций RAM).

В PRAM имеется одна общая входная лента, причем каждый процессор может считывать содержимое любой ячейки этой ленты. Память является общей для всех процессоров и состоит из потенциально бесконечного числа ячеек, в каждой из которых может быть записано произвольное целое число. Время вычисления на PRAM равно максимуму из времен вычислений на каждом процессоре. Время вычисления на одном процессоре – это

либо суммарное число исполненных команд (равномерная мера), либо сумма логарифмов величин максимальных операндов для каждой команды (логарифмическая мера).

Формулировка тезиса о параллельных вычислениях: Параллельное время полиномиально эквивалентно размеру памяти (рабочей зоны) машины Тьюринга. В частности, всякая задача может быть решена на PRAM за время $T(n)$ тогда и только тогда, когда она может быть решена машиной Тьюринга, использующей объем памяти (рабочую зону) $poly(T(n))$.

В несколько ином виде этот тезис впервые сформулирован в [FW78].

Подобно известному тезису Черча тезис о параллельных вычислениях не может быть формально доказан, поскольку интуитивное понятие параллельное время не формализовано. Однако, для многих параллельных моделей вычислений этот тезис оказывается справедливым как строгое математическое доказательство (в частности, для модели PRAM).

Сложностные классы P и NP, а также понятие полиномиальной (по Карпу) сводимости предполагаются известными читателю.

Сложностные классы по объему памяти: PSPACE, EXPSPACE и по времени EXPTIME определяются так.

Задача разрешения (языков) принадлежит классу EXPSPACE (PSPACE), если существует машина Тьюринга с одной входной лентой и одной рабочей лентой решающая эту задачу с использованием объема памяти (количества ячеек рабочей ленты) не превышающего экспоненты (полинома) от длины записи исходных данных на входной ленте.

Задача разрешения (языков) принадлежит классу EXPTIME, если существует машина Тьюринга с одной входной лентой и одной рабочей лентой, решающая эту задачу за время, не превышающее экспоненты от длины записи исходных данных на входной ленте.

Приведем и формальные определения этих классов. Для этого нам понадобятся несколько вспомогательных определений.

Определение. k-ленточная машина Тьюринга M имеет **сложность по объему памяти (space complexity)** $s(n)$, если для любого входного слова длины n машина M просматривает не более $s(n)$ ячеек на всех рабочих лентах (исключая входную ленту). Сложность по объему памяти машины M будем обозначать через $space_M(n)$.

Определение. k-ленточная машина Тьюринга M имеет **временную**

сложность $t(n)$, если для любого входного слова длины n машина М заканчивает работу не более, чем за $t(n)$ шагов. Временную сложность машины Тьюринга М будем обозначать через $time_M(n)$.

Определение. Язык $L \subset \Sigma^*$ принадлежит классу $\mathcal{DSPACE}(s(n))$, если существует машина Тьюринга М, разрешающая данный язык, и $space_M(n) \leq s(n)$ для почти всех n ($\forall n \geq n_0$).

Определение. Язык $L \subset \Sigma^*$ принадлежит классу $\mathcal{DTIME}(t(n))$, если существует машина Тьюринга М, разрешающая данный язык, и $time_M(n) \leq s(n)$ для почти всех n ($\forall n \geq n_0$).

$$\mathcal{EXPTIME} = \cup_{k \geq 0} \mathcal{DTIME}(2^{n^k}).$$

$$\mathcal{EXPSPACE} = \cup_{k \geq 0} \mathcal{DSPACE}(2^{n^k}).$$

$$\mathcal{P} = \cup_{k \geq 0} \mathcal{DTIME}(n^k).$$

$$\mathcal{PSPACE} = \cup_{k \geq 0} \mathcal{DSPACE}(n^k).$$

Известно соотношение между сложностными классами.

$$P \subseteq NP \subseteq PSPACE \subseteq EXPTIME \subseteq EXPSPACE$$

Неизвестно ни про одно из включений, является ли оно собственным. С другой стороны, из теорем о иерархии (по времени и по объему памяти) вытекает, что $P \neq EXPTIME$, и $PSPACE \neq EXPSPACE$.

Для сложностного класса C и сводимости S говорят, что задача f является C-трудной относительно S-сводимости, если любая задача из C S-сводится к f.

Если к тому же $f \in C$, то такая задача называется C-полной относительно сводимости S.

Наиболее популярной в теории сложности является полиномиальная сводимость по Карпу. В нашем изложении использовалась также менее популярная $log - lin$-сводимость, которая определяется аналогично сводимости по Карпу, но с более жестким ограничением: это отображение

входов одной задачи во входы другой, вычислимое с логарифмическим объемом памяти и за линейное время.

Список иллюстраций

Список алгоритмов

Список литературы

[Bar04] Bardet, M. and Faugere, J.C and Salvy, B. “On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations”. В: *International Conference on Polynomial System Solving - ICPSS*. Paris, France, нояб. 2004, с. 71—75. URL: http://www-salsa.lip6.fr/~jcf/Papers/43BF.pdf.

[BCP83] Allan Borodin, Stephen A. Cook и Nicholas Pippenger. “Parallel Computation for Well-Endowed Rings and Space-Bounded Probabilistic Machines”. В: *Information and Control* (1983), с. 113—136.

[Bro87] D. Brownawell. “Bounds for the degrees in the Nullstellensatz”. В: *Annals of Math. Second Series* 126.3 (1987), с. 577—591.

[Buc06] B. Buchberger. “An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal”. English. Дис. … док. RISC (Research Institute for Symbolic Computation), март 2006, с. 475—511.

[Cab04] Robbiano L. Cabauer R. Kreuzer M. “Efficiently computing minimal sets of critical pairs”. В: *Journal of Symbolic computation* 38.4 (2004), с. 1169—1190.

[CD11] Daniel Cabarcas и Jintai Ding. “Linear algebra to compute syzygies and Gröbner bases”. В: *Proceedings of the 36th international symposium on Symbolic and algebraic computation*. ISSAC ’11. San Jose, California, USA: ACM, 2011, с. 67—74. ISBN: 978-1-4503-0675-1. DOI: http://doi.acm.org/10.1145/1993886.1993902. URL: http://doi.acm.org/10.1145/1993886.1993902.

[Col97] Mall D. Collart S. Kalkbrener M. “Converting bases with the Grobner walk”. В: *J. Symbolic Computation* 24 (1997), с. 465—469.

[Dub90] T.W. Dube. “The structure of polynomial ideals and Grobner bases”. В: *SIAM Jornal of Computing* 19 (1990), с. 750—773.

[Fau99] J.-C. Faugere. “A new efficient algorithm for computing Grцbner bases (F4).” В: *Journal of Pure and Applied Algebra* 139.1–3 (июнь 1999), с. 61—88. ISSN: 0022-4049. DOI: 10.1016/S0022-4049(99)00005-5. URL: http://www-salsa.lip6.fr/~jcf/Papers/F99a.pdf.

[FW78] Steven Fortune и James Wyllie. “Parallelism in Random Access Machines”. В: *STOC*. ACM, 1978, с. 114—118.

[Ger10] O. German. “Proof of the Faugere criterion for the F5 algorithm”. В: *Mathematical Notes* 88 (3 2010). 10.1134/S0001434610090191, с. 479—486. ISSN: 0001-4346. URL: http://dx.doi.org/10.1134/S0001434610090191.

[Gio52] Trevisan Giorgio. “Classificazione dei semplici ordinamenti di un gruppo libero commutativo con n generatori”. В: т. 22. CEDAM, 1952, с. 143—156. URL: http://books.google.com/books?id=OypAAQAAIAAJ.

[Giu84] Marc Giusti. “Some Effectivity Problems in Polynomial Ideal Theory”. В: *Proceedings of the International Symposium on Symbolic and Algebraic Computation*. EUROSAM ’84. London, UK: Springer-Verlag, 1984, с. 159—171. ISBN: 3-540-13350-X. URL: http://dl.acm.org/citation.cfm?id=646671.699162.

[Her25] G. Hermann. *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale: Unt. Benutzung nachgelassener Satze v. Kurt Hentzelt*. Springer, 1925. URL: http://books.google.com/books?id=DHhAHQAACAAJ.

[HL11] Amir Hashemi и Daniel Lazard. “Sharper Complexity Bounds for Zero-Dimensional Gröbner Bases and Polynomial System Solving”. В: *IJAC* 21.5 (2011), с. 703—713.

[KM96] Klaus Kühnle и Ernst W. Mayr. “Exponential space computation of Gröbner bases”. В: *Proceedings of the 1996 international symposium on Symbolic and algebraic computation*. ISSAC ’96. Zurich, Switzerland: ACM, 1996, с. 63—71. ISBN: 0-89791-796-0. DOI: http://doi.acm.org/10.1145/236869.236900. URL: http://doi.acm.org/10.1145/236869.236900.

[Kol73] E.R. Kolchin. *Differential algebra and algebraic groups*. Pure and applied mathematics т. 54. Academic Press, 1973. ISBN: 9780124176508. URL: http://books.google.com.au/books?id=yDCfhIjka-8C.

[Laz83] Daniel Lazard. “Gröbner-Bases, Gaussian elimination and resolution of systems of algebraic equations”. В: *Proceedings of the European Computer Algebra Conference on Computer Algebra*. London, UK: Springer-Verlag, 1983, с. 146—156. ISBN: 3-540-12868-9. URL: http://dl.acm.org/citation.cfm?id=646657.700393.

[May89] Ernst W. Mayr. “Membership in Plynomial Ideals over Q Is Exponential Space Complete”. В: *STACS*. Под ред. Burkhard Monien и Robert Cori. Т. 349. Lecture Notes in Computer Science. Springer, 1989, с. 400—406. ISBN: 3-540-50840-6.

[May97] Ernst W. Mayr. “Some Complexity Results for Polynomial Ideals”. В: *J. Complexity* 13.3 (1997), с. 303—325.

[MM82] E. Mayr и A. Meyer. “The complexity of the word problems for commutative semigroups and polynomial ideals”. English. В: *Adv. Math., Beijing* 46.3 (дек. 1982), с. 305—329.

[Mul87] Ketan Mulmuley. “A fast parallel algorithm to compute the rank of a matrix over an arbitrary field”. В: *Combinatorica* 7.1 (1987), с. 101—104.

[Pan87] Victor Y. Pan. “Complexity of Parallel Matrix Computations”. В: *Theor. Comput. Sci.* 54 (1987), с. 65—85.

[Rab30] J. L. Rabinowitsch. “Zum Hilbertschen Nullstellensatz”. В: *Mathematische Annalen* 102 (1 1930). 10.1007/BF01782361, с. 520—520. ISSN: 0025-5831. URL: http://dx.doi.org/10.1007/BF01782361.

[Riq10] C. Riquier. *Les systèmes d’équations aux dérivées partielles*. Cornell University Library historical math monographs. Gauthier-Villars, 1910. URL: http://books.google.fr/books?id=6jDOAAAAMAAJ.

[Rob85] Lorenzo Robbiano. *Term orderings on the polynomial ring.* Computer algebra, EUROCAL ’85, Proc. Eur. Conf., Linz/Austria 1985, Vol. 2, Lect. Notes Comput. Sci. 204, 513-517 (1985). 1985.

[Вар76] Б.Л.ван дер Варден. *Алгебра*. М.: Наука, 1976.

[Лен68] Серж Ленг. *Алгебра*. М.: Мир, 1968.

[МИ53] Зайцева М.И. “О совокупности упорядочений абелевой группы”. В: *Успехи математических наук* 8 (1953), с. 135—137.

[Мин71] Марвин Минский. *Вычисления и автоматы*. М.: Мир, 1971.

[СВ12] Агиевич С.В. “Усовершенствованный алгоритм Бухбергера”. В: *Тр. Ин-та матем. Белорусской академии наук* 20.1 (2012), с. 3—13.