КОНКРЕТНАЯ ТЕОРИЯ КОЛЕЦ

ELVEN RINGS

Николай Вавилов

Three Rings for the Elven-kings under the sky,
Seven for the Dwarflords in their halls of stone,
Nine for Mortal Men doomed to die,
One for the Dark Lord on his dark throne
In the Land of Mordor where the Shadows lie.
One Ring to rule them all, One Ring to find them,
One Ring to bring them all and in the darkness bind them.

J.R.R.Tolkien, The Lord of the Rings

A good stack of examples, **as large as possible**, is indispensable for a thorough understanding of any concept, and when I want to learn something new, I make it my first job to build one.

Paul Halmos[1]

[1]Цитируется по книге J.Gallian, Contemporary abstract algebra.

Оглавление

Глава 1: КОЛЬЦА И ГОМОМОРФИЗМЫ

§ 1. Определение кольца, простейшие следствия из аксиом

Здесь мы введем второй важнейший тип алгебраических структур, в которых определены две основные операции.

1. Кольца. До сих пор мы рассматривали структуры с одной бинарной операцией. Однако в действительности на обычных числовых множествах задано две операции: сложение и умножение.

Определение. *Непустое множество R с двумя бинарными операциями, сложением '$+$' и умножением '$\cdot$', называется* **кольцом**, *если R образует абелеву группу по сложению и умножение двусторонне дистрибутивно относительно сложения.*

Иными словами, предполагается, что выполнены следующие четыре аксиомы, относящиеся к сложению:

A1. $\forall x, y, z \in R$, $(x + y) + z = x + (y + z)$;

A2. $\exists 0 \in R$, $\forall x \in R$, $x + 0 = x = 0 + x$;

A3. $\forall x \in R$, $\exists -x \in R$, $x + (-x) = 0 = (-x) + x$;

A4. $\forall x, y \in R$, $x + y = y + x$;

Через R^+ обозначается **аддитивная группа** кольца R, которая получается, когда мы забываем о том, что на R задано умножение. Однако в действительности умножение в кольце есть и оно связано со сложением следующими двумя аксиомами дистрибутивности:

D1. $\forall x, y, z \in R$, $x(y + z) = xy + xz$ (правая дистрибутивность);

D2. $\forall x, y, z \in R$, $(x + y)z = xz + yz$ (левая дистрибутивность).

Кольца чаще всего обозначаются R от немецкого 'Ring' (или английского 'ring'). Коммутативные кольца очень часто обозначаются A, от французского 'anneau'. Во многих книгах некоммутативные кольца обозначаются буквой Λ,

2. Простейшие следствия из аксиом кольца. Вычитание и 0 обладают обычными свойствами относительно умножения, которые сразу вытекают из аксиом кольца. Проверим, например, что $0x = 0 = x0$. В самом деле, $xy = x(y + 0) = xy + x0$ и $yx = (y + 0)x = yx + 0x$ для любых $x, y \in R$. Аналогично, $x(-y) = 0 - xy = (-x)y$, $(-x)(-y) = xy$. Докажем, для примера, первое из этих равенств: $x(-y) + xy = x(-y + y) = x0 = 0$, откуда $x(-y) = -xy$. По аналогии читатель без труда докажет выполнение двух других равенств и двусторонней дистрибутивности умножения относительно вычитания: $x(y - z) = xy - xz$, $(x - y)z = xz - yz$. В дальнейшем мы будем без всяких упоминаний пользоваться подобными очевидными свойствами операций в кольце. Заметим, что все они имеют место *независимо* от ассоциативности и/или коммутативности умножения!

Задача. Показать, что в кольце с 1 не нужно требовать коммутативность сложения. Она автоматически вытекает из остальных аксиом.

Решение. Пусть $x, y \in R$. Два раза воспользовавшись дистрибутивностью и определением 1, получим $2x + 2y = 2(x + y) = (1 + 1)(x + y) = (x + y) + (x + y)$. Воспользовавшись теперь ассоциативностью сложения и сокращая на x слева и на y справа, получаем $x + y = y + x$.

3. Поучительный контр-пример. Заметим, что в случае некоммутативного умножения левая и правая дистрибутивность, вообще говоря, *независимы* и для проверки того, что множество с данными операциями образует кольцо, необходимо проверять обе! Например, рассмотрим множество многочленов $K[x]$ относительно сложения $+$ и композиции $\circ$. Тогда $K[x]$ образует абелеву группу по сложению, композиция ассоциативна и допускает нейтральный элемент $e = x$. Очевидно и выполнение *левой* дистрибутивности $(f+g)\circ h = f\circ h + g\circ h$. Тем не менее, это **не кольцо**, так как *правая* дистрибутивность *не имеет места*, т.е., вообще говоря, $f\circ(g+h) \neq f\circ g + f\circ h$ (в характеристике $p \neq 2$ возьмите $f = x^2$, $g = h = 1$). Равенство здесь имеет место только для линейных многочленов $f = ax$ и в характеристике $p > 0$ для многочленов от x^p (см. Глава ?).

4. Ассоциативные кольца. Обычно мы будем иметь дело только с кольцами, в которых умножение также удовлетворяет обычным свойствам.

Определение. *Говорят, что* R – **ассоциативное кольцо с 1**, *если* R *образует моноид по умножению, т.е. если дополнительно к A1 – A4, D1, D2 выполнены две следующие аксиомы:*

M1. $\forall x, y, z \in R$, $(xy)z = x(yz)$;

M2. $\exists 1 \in R$, $\forall x \in R$, $x1 = x = 1x$.

Если, кроме того, умножение коммутативно, т.е.

M3. $\forall x, y \in R$, $xy = yx$,

то кольцо R *будет называться* **коммутативным** *и обозначаться через* R *(от английского 'ring').*

Более общо, $e \in R$ называется **левой единицей**, если $ex = x$ для всех $x \in R$, и **правой единицей**, если $xe = x$ для всех $x \in R$. Обычная выкладка, использующая ассоциативность умножения, показывает, что если в *ассоциативном* кольце существует как левая единица e_1, так и правая единица e_2, то $e_1 = e_2$, так что R – кольцо с единицей. Однако в неассоциативном кольце это не так. Кроме того, легко привести примеры ассоциативных колец в которых существует бесконечно много левых или правых единиц.

Задача. Пусть A – произвольное ассоциативное кольцо с 1, R – множество пар A^2 с покомпонентным сложением и умножением $(a,b)(c,d) = (ac, ad)$. Убедитесь, что в кольце R нет правых единиц, но, вообще говоря, много левых единиц. Постройте пример кольца в котором не существует левых единиц, но есть правые единицы.

5. Вычисления в коммутативных кольцах. Класс коммутативных ассоциативных колец с 1 представляет собой *естественную* общность, в которой выполняются **все** обычные формулы школьной алгебры. Так, например, $(x+y)^2 = (x+y)(x+y) = x^2+xy+yx+y^2$ и, если элементы x и y коммутируют (и только в этом случае!), это равно $x^2 + 2xy + y^2$. Вообще, для любого $n \in \mathbb{N}$ имеет место формула **бинома Ньютона**

$$(x+y)^n = \sum_{i=0}^{n} C_n^i x^i y^{n-1}.$$

и обобщающая ее **мультиномиальная формула Лейбница**

$$(x_1 + \ldots + x_m)^n = \sum_{i_1+\ldots+i_m=n} \frac{n!}{i_1! \ldots i_m!} x^{i_1} \ldots x_m^{i_m},$$

сумма в которой берется по всевозможным разбиениям n на m слагаемых из $\mathbb{N}_0$. Имеет место и обычная формула для суммы геометрической прогрессии

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \ldots + y^{n-1})$$

и **все** остальные формулы, знакомые из школьной алгебры.

§ 2. Первые примеры колец

Обычные сложение и умножение на $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ превращают их в ассоциативные коммутативные кольца с 1, но, в действительности, они удовлетворяют еще одной дополнительной аксиоме, и мы рассмотрим их в § ?, посвященном полям. Приведем некоторые простейшие примеры ассоциативных колец.

• **Кольцо целых чисел.** Множество $\mathbb{Z}$ целых чисел является коммутативным кольцом (в действительности, областью целостности, см. § ?) относительно обычных операций сложения и умножения.

• **Кольцо десятичных дробей.** Следующее кольцо является основой всей школьной арифметики. Обозначим через R множество всех десятичных дробей с обычными операциями сложения и умножения[2]. Из курса арифметики для младших классов известно, что R является кольцом. С точки зрения алгебры $R = \mathbb{Z}\left[\frac{1}{2}, \frac{1}{5}\right]$ это кольцо, получающееся из $\mathbb{Z}$ обращением 2 и 5, мы подробно обсудим эту конструкцию в Главе ?. Древние обычно работали с большим кольцом $\mathbb{Z}\left[\frac{1}{2}, \frac{1}{3}, \frac{1}{5}\right]$ шестидесятиричных дробей, в котором, кроме того, обращается 3.

• **Кольца классов вычетов.** В Главе 1 мы уже упоминали о возможности корректно ввести операции на множестве $\mathbb{Z}/m\mathbb{Z}$ классов вычетов по модулю m, в § ? мы познакомимся с конструкцией фактор-колец, обобщающей эту идею. Кольца классов вычетов $\mathbb{Z}/2\mathbb{Z}$ и $\mathbb{Z}/3\mathbb{Z}$ в действительности являются полями и будут рассмотрены в § 4. Приведем в качестве иллюстрации таблицы сложения и умножения кольца $\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$ вычетов по модулю 4:

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

×	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Мы уже встречали эту таблицу сложения, когда обсуждали латинские квадраты – это таблица Кэли циклической группы порядка 4. Обратите внимание, что $2 \cdot 2 = 0$, в то время, как $2 \neq 0$, так что $\mathbb{Z}/4\mathbb{Z}$ – кольцо с делителями нуля.

• **Булево кольцо множеств.** Пусть R обозначает множество всех подмножеств некоторого множества X. Тогда R не может, разумеется, быть кольцом

[2]**конечных** десятичных дробей, **никто** не умеет складывать или умножать "бесконечные" десятичные дроби, см. § ?.

относительно объединения (как сложения) и пересечения (как умножения). В самом деле, относительно объединения R коммутативный моноид, но не группа – ни одно непустое подмножество не будет обратимым относительно объединения. Возьмем поэтому в качестве сложения симметрическую разность $\triangle$, относительно которой R является абелевой группой, а в качестве умножения – пересечение $\cap$, которое дистрибутивно относительно $\triangle$. С этими операциями R образует коммутативное ассоциативное кольцо с 1, являющееся основным примером булева кольца (не путать с булевыми алгебрами!). Ассоциативное кольцо R с 1 называется **булевым**, если $x^2 = x$ для всех $x \in R$.

Упражнение. Проверить, что булево кольцо R коммутативно.

• **Нулевое кольцо.** Если в кольце R выполнено равенство $0 = 1$, то вообще для любого элемента этого кольца $x = x1 = x0 = 0$. Таким образом, R состоит из одного элемента. Такое кольцо называется **нулевым** и обозначается через 0.

• **Кольцо с нулевым умножением.** Пусть R – абелева группа по сложению, содержащая по крайней мере 2 элемента. Определим в R умножение, положив $xy = 0$ для любых $x, y \in R$. Получившееся ассоциативное коммутативное кольцо без 1 называется **кольцом с нулевым умножением**. Любой R-модуль M, в частности, любое векторное пространство V над полем K можно рассматривать как кольцо с нулевым умножением.

2. Построение неассоциативных колец из ассоциативных. В действительности, легко придумать много способов так испортить операцию умножения в ассоциативном кольце, чтобы оно перестало быть ассоциативным, но осталось кольцом.

Задача. Пусть R – ассоциативное кольцо, $c \in R$. Определим в R новую операцию умножения, полагая $x \cdot y = cxy$. Убедитесь, что с этой новой операцией R по прежнему представляет собой кольцо.

В действительности любое кольцо является подкольцом, в каком-то из колец $R(c)$ построенным таким образом из некоторого ассоциативного кольца[3].

Задача. Пусть $R = A \oplus B$ – разложение аддитивной группы кольца R в прямую сумму двух аддитивных групп, $\pi : R \longrightarrow A$ – каноническая проекция на прямое слагаемое A. Определим в R новую операцию умножения, полагая $x \cdot y = \pi(xy)$. Убедитесь, что с этой новой операцией R по прежнему представляет собой кольцо[4].

§ 3. Некоторые основные конструкции

Перечислим некоторые основные конструкции, которые позволяют строить новые кольца по одному или нескольким известным кольцам. В дальнейшем мы подробно изучим эти конструкции, их варианты и обобщения.

• **Кольцо многочленов.** Пусть K обозначает одно из следующих множеств: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, или $\mathbb{C}$. Тогда множество всех многочленов $K[t]$ от одной независимой переменной над K образует коммутативное кольцо – в действительности снова область целостности. Мы подробно обсудим определение этого кольца с произвольным кольцом коэффициентов K в следующем параграфе и вернемся к его детальному изучению в Главах 2 и 3.

---

[3]А.И.Мальцев, Об одном представлении неассоциативных колец. – Успехи Мат. Наук, 1952, т.7, N.1, с.181–185.

[4]Л.А.Скорняков, Представление неассоциативных колец в ассоциативных. – Докл. АН СССР, 1955, т.102б N.1, с.33–35.

• **Кольцо матриц.** Пусть K обозначает одно из следующих множеств: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, или $\mathbb{C}$. В § ? мы построим кольцо матриц $M(n, K)$ (в действительности, конечно, над произвольным ассоциативным кольцом R), которое будет ассоциативным, но не коммутативным (если $n \geq 2$) кольцом с 1. При всех $n \geq 2$ это кольцо имеет делители нуля.

• **Противоположное кольцо.** Для любого кольца R через R^o (или R^{op}) обозначается кольцо, аддитивная группа которого совпадает с R^+, а умножение задается формулой $x \circ y = yx$. Кольцо R^o называется **противоположным** к R. Заметим, что буковка 'o' в обозначении противоположного кольца – это не нолик, а первая буква слова 'opposite'. Кольцо R в том и только том случае коммутативно, когда $R^o = R$. Заметим, что некоммутативное кольцо может быть *изоморфно* своему противоположному $R^0 \cong R$, при этом $R^o \neq R$.

• **Прямая сумма колец.** Пусть $R_1, \ldots, R_s$ – кольца. Обозначим через $R_1 \oplus \ldots \oplus R_s$, кольцо, которое как множество совпадает с прямым произведением $R = R_1 \times \ldots \times R_s$, с покомпонентными операциями сложения и умножения:

$$(x_1, \ldots, x_s) + (y_1, \ldots, y_s) = (x_1 + y_1, \ldots, x_s + y_s),$$

$$(x_1, \ldots, x_s)(y_1, \ldots, y_s) = (x_1 y_1, \ldots, x_s y_s).$$

Заметим, что прямая сумма конечного числа колец часто называется также их **прямым произведением** и в этом случае обозначается $R = R_1 \times \ldots \times R_s$.

Предложение. *Пусть аддитивная группа кольца R есть прямая сумма аддитивных групп его подколец $R_1, \ldots, R_s$. Тогда следующие условия эквивалентны:*

1) $R = R_1 \oplus \ldots \oplus R_s$,

2) $R_i \trianglelefteq R$ для всех i,

3) $R_i R_j = 0$ для всех $i \neq j$.

Доказательство. 1) $\Longrightarrow$ 2) Ясно, так как операции в R покомпонентные. 2) $\Longrightarrow$ 3) $R_i R_j \subseteq R_i \cap R_j = 0$. 3) $\Longrightarrow$ 1) Дистрибутивность умножения.

Предостережение. Понятия прямого произведения и прямой суммы обобщается и на бесконечные семейства колец R_i, $i \in I$, но в этом случае эти конструкции различаются точно так же, как понятия прямого произведения и прямой суммы абелевых групп. Иными словами, **прямое произведение** $\prod R_i$, $i \in I$, совпадает с прямым произведением R_i как множеств. В то же время **прямая сумма** $\bigoplus R_i$, $i \in I$, является собственным подмножеством в $\prod R_i$, $i \in I$, совпадающим с множеством таких списков $(x_i)_{i \in I}$, где $x_i \in R_i$, что $x_i = 0$ для *почти всех* i. Для бесконечных семейств колец эти конструкции приводят к кольцам, которые, как правило, не только не изоморфны, но и обладают *совершенно различными свойствами*. Например, прямое произведение колец с единицей само является кольцом с 1, в качестве которой можно взять список, i-й компонентой которого является единица кольца R_i. В то же время прямая сумма бесконечного семейства колец с единицей является кольцом без единицы!

• **Присоединение 1.** Пусть R – любое кольцо, вообще говоря, без 1. Рассмотрим кольцо R_1, которое совпадает с $\mathbb{Z} \oplus R$ как аддитивная группа, с умножением, определенным равенством

$$(m, x)(n, y) = (mn, my + nx + xy).$$

Легко видеть, что R_1 – кольцо с 1. Про него говорят, что оно получается из R присоединением 1. Эта конструкция сводит все вопросы о кольцах без 1 к соответствующим вопросам для колец с 1. Поэтому в дальнейшем мы будем, как правило, включать 1 в сигнатуру кольца (все наши гомоморфизмы и подкольца будут унитальными!)

• **кольцо Штуди.** Пусть M – R-модуль, рассматриваемый как кольцо с нулевым умножением, т.е. $uv = 0$ для любых $u, v \in M$. Тогда присоединяя 1 мы получим кольцо $R \oplus M$. Частным случаем этой конструкции является кольцо дуальных чисел.

§ 4. Подкольца

1. Подкольца. Подкольца определяются точно так же, как все остальные подобъекты, а именно, подкольцом кольца R называется подмножество, которое является кольцом относительно тех же операций.

Определение. *Непустое подмножество $S \subseteq R$ называется* **подкольцом** *кольца R, если для любых $x, y \in S$ имеем $x - y, xy \in R$. Если R – кольцо с 1, то полкольцо S называется* **унитальным**, *если $1 \in S$.*

Сейчас мы приведем несколько простейших примеров подколец.

• **Центр кольца** Элемент $x \in R$ называется **центральным**, если он коммутирует со всеми элементами кольца R, т.е. $xy = yx$ для всех $y \in R$. Множество $C(R)$ всех центральных элементов кольца R называется **центром** кольца R,

$$C(R) = \{x \in R \mid \forall y \in R, xy = yx\}.$$

Легко видеть, что центр является унитальным подкольцом.

• **Кольцо четных чисел.** Так как сумма и произведение двух четных чисел являются четным числом, то множество $2\mathbb{Z}$ четных чисел является подкольцом в $\mathbb{Z}$. Очевидно, это подкольцо без 1.

• **Подкольца в $\mathbb{Q}$.** Любое подкольцо с 1 поля $\mathbb{Q}$ имеет вид $\mathbb{Z}\left[\frac{1}{p}, p \in S\right]$ для некоторого множества простых $S \subseteq \mathbb{P}$. Таким образом, в $\mathbb{Q}$ имеется континуум подколец. Позже мы убедимся, что все эти кольца попарно неизоморфны.

2. Кольца целых алгебраических чисел. Существует множество интересных колец, родственных $\mathbb{Z}$, которые получаются из $\mathbb{Z}$ присоединением корней алгебраических уравнений. Все эти кольца являются подкольцами в кольце $\mathbb{A}$ **целых алгебраических чисел**. Перечислим несколько простейших примеров таких подколец, которые будут встречаться нам в дальнейшем:

• Кольцо $\mathbb{Z}[i]$ **целых гауссовых чисел**, состоящее из всех чисел вида $a+bi$, где $a, b \in \mathbb{Z}$, а $i^2 = -1$.

• Кольцо $\mathbb{Z}[\omega]$ **целых гауссовых чисел**, состоящее из всех чисел вида $a + b\omega$, где $a, b \in \mathbb{Z}$, а $\omega = \dfrac{-1 + i\sqrt{3}}{2}$.

• $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$.

Нам встретятся и другие примеры.

§ 5. Гомоморфизмы колец

Как обычно, рассматривая какой-то класс структур, мы должны одновременно ввести допустимый класс отображений.

1. Гомоморфизмы колец. Пусть R и S – два кольца.

Определение. *Отображение $f : R \longrightarrow S$ называется* **гомоморфизмом колец**, *если f одновременно является гомоморфизмом аддитивной и мультипликативной структур, т.е. если для любых $x, y \in G$ выполняются равенства*

$$f(x+y) = f(x) + f(y), \qquad f(xy) = f(x)f(y).$$

Так как R образует группу по сложению, то, как мы знаем, из того, что f сохраняет сложение, автоматически вытекает, что $f(0) = 0$. В то же время, из того, что f сохраняет умножение, отнюдь не следует, что $f(1_R) = 1_S$. Гомоморфизм f, для которого $f(1_R) = 1_S$, называется **унитальным**. Обычно для колец с 1 мы включаем 1 в сигнатуру и рассматриваем только унитальные гомоморфизмы.

Как обычно, гомоморфизм f называется **мономорфизмом**, если он инъективен, **эпиморфизмом**, если он сюръективен, **изоморфизмом**, если он биективен, **эндоморфизмом**, если $R = S$, и, наконец, **автоморфизмом**, если $R = S$, а f биективен.

Приведем несколько простейших примеров гомоморфизмов.

• **Каноническая проекция.** Пусть R – кольцо, $I \trianglelefteq R$, R/I – факторкольцо. Тогда по самому определению операций в R/I каноническая проекция $R \longrightarrow R/I$, $x \mapsto x + I$, является гомоморфизмом колец.

• **Характеристическая функция.** Пусть $(2^X, \triangle, \cap)$ – булево кольцо подмножеств в X. Тогда отображение $\chi : 2^X \longrightarrow \mathbb{F}_2^X$, которое сопоставляет каждому $Y \subseteq X$ его характеристическую функцию, является гомоморфизмом колец:

$$\chi_{Y \triangle Z} = \chi_Y + \chi_Z, \qquad \chi_{Y \cap Z} = \chi_Y \chi_Z.$$

Так как соответствие $Y \leftrightarrow \chi_Y$ биективно, то булево кольцо 2^X изоморфно прямому произведению $|X|$ экземпляров поля $\mathbb{F}_2$.

• **Действие перестановок на подмножествах.** Пусть $\sigma \in S_X$ – перестановка множества X. Сопоставим σ перестановку множества 2^X полагая $\sigma(Y) = \{\sigma(y) \mid y \in Y\}$. Ясно, что σ – автоморфизм булева кольца 2^X:

$$\sigma(Y \triangle Z) = \sigma(Y) \triangle \sigma(Z), \qquad \sigma(Y \cap Z) = \sigma(Y) \cap \sigma(Z).$$

Задача. Докажите, что **каждый** автоморфизм булева кольца 2^X имеет такой вид, т.е. $\mathrm{Aut}(2^X) \cong S_X$.

• **Матрица линейного отображения.** Пусть $M \cong R^n$ свободный R-модуль и $u_1, \ldots, u_n$ – какой-то базис этого модуля. Тогда отображение

$$\mathrm{End}_R(M) \longrightarrow M(n, R), \quad \phi \mapsto [\phi]_u,$$

сопоставляющее эндоморфизму модуля M матрицу в базисе $u_1, \ldots, u_n$, является изоморфизмом колец.

§ 6. Гомоморфизмы, связанные со структурой кольца

• **Внутренние автоморфизмы.** Пусть $u \in R^*$. Тогда отображение $I_u : R \longrightarrow R$, $x \mapsto uxu^{-1}$ является автоморфизмом кольца R, называемым **внутренним автоморфизмом**. В самом деле $I_u(x+y) = I_u(x) + I_u(y)$ и $I_u(xy) = I_u(x)I_u(y)$.

Задача. Убедитесь, что отображение $R^* \mapsto \mathrm{Aut}(R)$, $u \mapsto I_u$ является гомоморфизмом групп. Чему равно ядро этого гомоморфизма?

• Гомоморфизм $\mathbb{Z} \longrightarrow R$, $n \mapsto n1$, является **единственным** (унитальным) гомоморфизмом $\mathbb{Z}$ в R.

Задача. Два подкольца в $\mathbb{Q}$ тогда и только тогда изоморфны, когда они равны.

Решение. Любой изоморфизм колец должнен быть изоморфизмом их аддитивных групп, т.е. гомотетией с коэффициентом $c \in \mathbb{Q}$. Но как мы только что заметили, изоморфизм колец должен быть постоянным на $\mathbb{Z}$.

• **Проекции на слагаемые.** Если $R = R_1 \oplus \ldots \oplus R_s$ – прямая сумма колец $R_1, \ldots, R_s$, то $\mathrm{pr}_i : R \longrightarrow R_1 \oplus \ldots \oplus R_s$, $(x_1, \ldots, x_s) \mapsto x_i$, является кольцевым гомоморфизмом.

Задача. Докажите, что проекции $\mathrm{pr}_i : \mathbb{Z}^n \longrightarrow \mathbb{Z}$, $(a_1, \ldots, a_n) \mapsto a_i$ являются *единственными* кольцевыми гомоморфизмами $\mathbb{Z}^n \longrightarrow \mathbb{Z}$.

Решение. Пусть $e_1, \ldots, e_n$ – стандартный базис $\mathbb{Z}^n$. Это полный набор ортогональных идемпотентов в $\mathbb{Z}^n$, т.е. $1 = e_1 + \ldots + e_n$, $e_i^2 = e_i$, и $e_ie_j = 0$ при $i \neq j$. Таким образом, $\phi(e_1) + \ldots + \phi(e_n) = \phi(e_1 + \ldots + e_n) = 1$, все $\phi(e_i)$ равны 0 или 1, причем $\phi(e_i)\phi(e_j) = 0$ при $i \neq j$. Это возможно только если **ровно** одно из $\phi(e_i)$ равно 1.

Задача. Найти все автоморфизмы кольца $\mathbb{Z}^n$ (с поточечными операциями).

• **Пополняющий гомоморфизм.** Пусть R – коммутативное колцьо с 1, M – мультипликативный моноид. Тогда тождественный автоморфизм $\mathrm{id} : R \longrightarrow R$ и тривиальный гомоморфизм моноидов $M \longrightarrow R^*$, $g \mapsto 1$, определяют эпиморфизм колец

$$\mathrm{aug} : R[M] \longrightarrow R, \quad \sum a_g e_g \mapsto \sum a_g,$$

называемый **аугментацией** или **пополняющим гомоморфизмом**. Ядро аугментации называется **пополняющим идеалом** кольца $R[M]$

Количество гомоморфизмов $\mathbb{Z}/m\mathbb{Z}$ в $\mathbb{Z}/n\mathbb{Z}$ равно $2^{\omega(n)-\omega(n/\gcd(m,n)}$, где $\omega(n)$ обозначает количество различных простых делителей n, см.[5].

§ 2. Теорема косинусов и теорема Хуа Локена

1. Теорема Хуа Локена. Мы знаем, что любой гомоморфизм мультипликативных групп *автоматически* переводит 1 в 1 и обратный в обратный. Сейчас произойдет нечто совершенно удивительное! Оказывается, для того, чтобы аддитивный гомоморфизм тела в себя был гомоморфизмом колец, *почти* достаточно, чтобы он переводил 1 в 1 и обратный в обратный.

Теорема Хуа Локена. *Если $\sigma : T \longrightarrow T$ – отображение тела T в себя такое, что*

1) $\sigma(x+y) = \sigma(x) + \sigma(y)$;

2) $\sigma(x^{-1}) = \sigma(x)^{-1}$;

[5] J.A.Gallian, J.Van Buskirk, The number of homomorphisms from $\mathbb{Z}_m$ into $\mathbb{Z}_n$. – Amer. Math. Monthly, 1984, vol.91, p.196–197.

3) $\sigma(1) = 1$.

Тогда σ *является либо автоморфизмом, либо антиавтоморфизмом.*

Приводимое ниже доказательство является обработкой доказательства из замечательной книги Эмиля Артина, которое, в свою очередь, является обработкой первоначальной идеи Хуа Локена.

2. Сохранение антикоммутатора. Начнем с проверки несколько более слабого утверждения. А именно, в этом пункте мы покажем, что σ является автоморфизмом **йордановой алгебры** $T^{(+)}$ получающейся, если заменить умножение в T на **антикоммутирование** $x \circ y = xy + yx$. Для этого мы начнем с еще более слабого утверждения, а именно, покажем, что σ сохраняет **квадратичное умножение** $x \bullet y = xyx$. Сделать это совсем легко, так как следующее **рациональное тождество** выражает квадратичное умножение через сложение и взятие обратного элемента.

Лемма. *Пусть* $x, y \in T$, $x, y \neq 0$, $x^{-1} \neq y$. *Тогда элемент* $x^{-1} + (y^{-1} - x)^{-1} \in T^*$, *и*

$$xyx = x - (x^{-1} + (y^{-1} - x)^{-1})^{-1}.$$

Доказательство. Рассмотрим $x^{-1} + (y^{-1} - x)^{-1}$. Вынося x^{-1} слева и $(y^{-1} - x)^{-1}$ справа, получим

$$x^{-1} + (y^{-1} - x)^{-1} = x^{-1}((y^{-1} - x) + x)(y^{-1} - x)^{-1} = x^{-1}y^{-1}(y^{-1} - x)^{-1} \in T^*.$$

Таким образом,

$$(x^{-1} + (y^{-1} - x)^{-1})^{-1} = (y^{-1} - x)yx = x - xyx.$$

Лемма. *Если* $\sigma : T \longrightarrow T$ *– отображение тела* T *в себя, удовлетворяющее условиям теоремы, то* σ *автоморфизм относительно квадратичного умножения* $x \bullet y = xyx$, *иными словами,*

$$\sigma(xyx) = \sigma(x)\sigma(y)\sigma(x).$$

Доказательство. Применяя σ к обеим частям содержащегося в Лемме тождества и пользуясь тем, что σ аддитивный гомоморфизм и переводит обратный в оборатный, мы получаем такое же выражение с заменой x и y на $\sigma(x)$ и $\sigma(y)$. Таким образом, $\sigma(xyx) = \sigma(x)\sigma(y)\sigma(x)$ для всех $x, y \neq 0$ таких, что $x^{-1} \neq y$. Но в исключительных случаях это равенство очевидно верно. Таким образом, оно верно всегда.

Лемма. *Если* $\sigma : T \longrightarrow T$ *– отображение тела* T *в себя, удовлетворяющее условиям теоремы, то* σ *автоморфизм йордановой алгебры* $T^{(+)}$, *иными словами,*

$$\sigma(xy + yx) = \sigma(x)\sigma(y) + \sigma(y)\sigma(x).$$

Доказательство. В частности, при $y = 1$ получаем $\sigma(x^2) = \sigma(x)^2$ так что σ переводит квадраты в квадраты. Теперь обычное рассуждение (**теорема косинусов**) показывает, что σ переводит антикоммутаторы в антикоммутаторы. В самом деле,

$$\sigma((x + y)^2) = (\sigma(x + y))^2 = (\sigma(x) + \sigma(y))^2.$$

Вычисляя левую часть, получаем $\sigma(x)^2 + \sigma(xy + yx) + \sigma(y)^2$, в то время как вычисление правой части дает $\sigma(x)^2 + \sigma(x)\sigma(y) + \sigma(y)\sigma(x) + \sigma(y)^2$. Для доказательства леммы достаточно сравнить два эти выражения.

Заметим, что этой леммы уже достаточно, чтобы завершить доказательство теоремы Хуа Локена в случае, когда $T = K$ – поле характеристики $\neq 2$, так как в этом случае $xy = (x \circ y)/2$. Нам понадобится еще несколько строчек, чтобы убедиться в том, что для любых x, y, всегда $\sigma(xy) = \sigma(x)\sigma(y)$ или $\sigma(xy) = \sigma(y)\sigma(x)$, так что теорема верна и для полей характеристики 2. Таким образом, **вся** остальная (нетривиальная!) часть доказательства нужна исключительно для борьбы с некоммутативностью.

3. Окончание доказательства теоремы. Вот ключевое соображение.

Основная лемма. *Для любых x, y имеем*

$$\sigma(xy) = \sigma(x)\sigma(y) \quad \textit{или} \quad \sigma(xy) = \sigma(y)\sigma(x).$$

Доказательство. Так как оба эти равенства очевидно верны, если хотя бы один из элементов x или y равен 0. Если же $x, y \neq 0$, то пользуясь двумя предшествующими леммами, мы убеждаемся, что произведение

$$(\sigma(x) - \sigma(x)\sigma(y))\sigma(xy)^{-1}(\sigma(xy) - \sigma(y)\sigma(x))$$

равно 0. Тем самым, хотя бы один из сомножителей обязан равняться 0.

Таким образом, для завершения доказательства теоремы нам остается лишь показать, что предполагая существание пар x, y и u, v таких, что $\sigma(xy) = \sigma(x)\sigma(y) \neq \sigma(y)\sigma(x)$ и $\sigma(uv) = \sigma(v)\sigma(u) \neq \sigma(u)\sigma(v)$ мы неизбежно придем к противоречию. В самом деле, возьмем **любое** $z \in T$. По основной лемме имеет место (по крайней мере) одна из двух возможностей

$$\sigma(x)\sigma(y) + \sigma(xz) = \sigma(x(y+z)) = \sigma(x)\sigma(y+z) = \sigma(x)\sigma(y) + \sigma(x)\sigma(z)$$

или

$$\sigma(x)\sigma(y) + \sigma(xz) = \sigma(x(y+z)) = \sigma(y+z)\sigma(x) = \sigma(y)\sigma(x) + \sigma(z)\sigma(y).$$

В первом случае $\sigma(xz) = \sigma(x)\sigma(z)$. Во втором случае $\sigma(xz) \neq \sigma(z)\sigma(x)$ так что по основной лемме снова $\sigma(xz) = \sigma(x)\sigma(z)$. Рассматривая выражения $\sigma((x+z)y)$, $\sigma(u(v+z))$, $\sigma((u+z)v)$, точно так же получаем, что $\sigma(zy) = \sigma(z)\sigma(y)$, $\sigma(uz) = \sigma(z)\sigma(u)$, $\sigma(zv) = \sigma(v)\sigma(z)$, для всех z.

Подставляя сюда $z = v$, получим $\sigma(xv) = \sigma(x)\sigma(v)$, а подставляя $z = x$, получим $\sigma(xv) = \sigma(v)\sigma(x)$. Таким образом, $\sigma(x)\sigma(v) = \sigma(v)\sigma(x)$ и точно так же доказывается, что $\sigma(y)\sigma(u) = \sigma(u)\sigma(y)$.

Теперь у нас все готово, чтобы получить финальное противоречие. Для этого рассмотрим $\sigma(xy) + \sigma(xv) + \sigma(uy) + \sigma(uv) = \sigma((x+u)(y+v))$. По основной лемме имеет место (по крайней мере) одна из двух следующих возможностей $\sigma((x+u)(y+v)) = \sigma(x+u)\sigma(y+v)$ или $\sigma((x+u)(y+v)) = \sigma(y+v)\sigma(x+u)$. Раскрывая скобки в правой части и пользуясь тем, что, как мы только что доказали, $\sigma(x)$ коммутирует с $\sigma(v)$, а $\sigma(y)$ коммутирует с $\sigma(u)$, мы видим, что первая возможность вынуждает $\sigma(u)$ коммутировать с $\sigma(v)$, а вторая возможность вынуждает $\sigma(x)$ коммутировать с $\sigma(y)$, вопреки нашему первоначальному предположению! Теорема доказана.

Тема 2: ПОЛЯ И ТЕЛА

§ 1. Поля и тела

1. Поля и тела. Особенно большое значение имеют такие кольца, в которых на любой ненулевой элемент можно сократить.

Определение. *Ассоциативное кольцо T с $1 \neq 0$ называется* **телом** *(или* **кольцом с делением***), если все его ненулевые элементы обратимы, иными словами, если выполняется следующая аксиома:*

M4. $\forall x \in T,\ x \neq 0,\ \exists x^{-1} \in T,\ xx^{-1} = 1 = x^{-1}x$.

Таким образом, в теле выполнены аксиомы **A1 – A4, D1, D2, M1, M2, M4**. В частности, все ненулевые элементы тела образуют (не обязательно коммутативную!) группу по умножению, называемую **мультипликативной группой** этого тела и обозначаемую $T^* = T^\bullet = T \setminus \{0\}$.

Определение. *Коммутативное тело K называется* **полем**.

Иными словами, в поле K множество K всех ненулевых элементов образует абелеву группу по умножению и, таким образом, выполнены аксиомы **A1 – A4, D1, D2, M1 – M4**. Поля принято обозначать буквой K от немецкого 'Körper' – 'поле' или F – от английского 'field' – 'поле'. Этот термин был впервые использован Рихардом Дедекиндом. Некоторые ранние авторы предлагали передавать немецкое 'Körper' и французское 'corps' по-русски как 'корпус', что было чрезвычайно удачно, но, к сожалению, не получило широкого распространения. В западных языках термин 'тело' не имеет отдельного существования, по-английски тело называется skew-field (с итальянской калькой campo sghembo), по-немецки – Schiefkörper или Divisionsbereich, по-французски – corps gauche.

Упражнение. Проверьте, что центр тела является полем.

2. Первые примеры полей. Приведем некоторые очевидные примеры полей.

• **Числовые поля.** Рациональные числа $\mathbb{Q}$, вещественные числа $\mathbb{R}$ и комплексные числа $\mathbb{C}$ образуют поля. Напомним (подробнее об этом см. Главу 4), что комплексные числа можно истолковать как пары вещественных чисел (a, b), $a, b \in \mathbb{R}$, с покомпонентным сложением и умножением, имитирующим теоремы сложения для cos и sin:

$$(a, b)(c, d) = (ac - bd, ad + bc).$$

• **Конечные поля.** По определению поле K обязано содержать по крайней мере два элемента, 0 и 1. Оказывается, существует поле ровно из двух элементов. А именно, определим на множестве $\mathbb{F}_2 = \{0, 1\}$ алгебраические операции, полагая

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

Так определенные операции превращают $\mathbb{F}_2$ в поле (это частный случай булева кольца множеств, относительно операций симметрической разности и объединения, рассмотренного в примере 3 выше, для случая, когда множество X одноэлементно).

Также и трехэлементное множество $\mathbb{F}_3 = \{0, 1, -1\}$ превращается в поле введением следующих операций:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\times$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

В действительности, в обоих случаях так построенные поля совпадают с кольцами вычетов $\mathbb{Z}/2\mathbb{Z}$ и $\mathbb{Z}/3\mathbb{Z}$, о которых шла речь в примере 8 выше, и то же самое верно для любого простого числа p: кольцо вычетов $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ является полем (называемым полем из p элементов или простым полем характеристики p).

Как мы убедились в том же примере 8, кольцо $\mathbb{Z}/4\mathbb{Z}$ не может быть полем, потому что оно имеет делители 0. Тем не менее, поле из четырех элементов существует, просто оно не изоморфно $\mathbb{Z}/4\mathbb{Z}$. Чтобы убедиться в этом, рассмотрим множество $\mathbb{F}_4 = \{0, 1, u, v\}$, операции в котором вводятся следующим образом:

$+$	0	1	u	v
0	0	1	u	v
1	1	0	v	u
u	u	v	0	1
v	v	u	1	0

$\times$	0	1	u	v
0	0	0	0	0
1	0	1	u	v
u	0	u	v	1
v	0	v	1	u

Аналогично, хотя и несколько сложнее, можно построить конечное поле $\mathbb{F}_q$ любого **примарного** порядка $q = p^m$, например, поля $\mathbb{F}_8$, $\mathbb{F}_9$, $\mathbb{F}_{16}$, $\mathbb{F}_{25}$, $\mathbb{F}_{27}$ и так далее. Поле порядка $q = p^m$ единственно с точностью до изоморфизма. В то же время, если q не является примарным, то не существует поля из q элементов, например, нет полей, содержащих ровно 6, 10, 12, 14, 15 элементов. Эти результаты будут доказаны в Главе ?.

• **Поля алгебраических чисел.** Так называются поля, получающиеся из $\mathbb{Q}$ присоединением корней алгебраических уравнений, обычно в конечном числе. Например, можно рассмотреть поле $\mathbb{Q}(i)$ **гауссовых чисел**, состоящее из всех чисел вида $a + ib$, где $a, b \in \mathbb{Q}$, а i – мнимая единица, $i^2 = -1$. Аналогично, можно рассмотреть поле $\mathbb{Q}(\sqrt{2})$, состоящее из всех чисел вида $a + b\sqrt{2}$, где $a, b \in \mathbb{Q}$. Вообще, можно показать, что множество корней **всех** алгебраических уравнений вида $x^n + a_{n-1}x^{n-1} + ... + a_1x + a_0 = 0$, где $a_{n-1}, \dots, a_1, a_0 \in \mathbb{Q}$, образует поле, называемое **полем алгебраических чисел**, и обозначаемое через $\overline{\mathbb{Q}}$.

• **Поле разложения многочлена.** Пусть K какое-то поле, $f \in K[x]$ – многочлен над K. В главе ? мы свяжем с f наименьшее поле L, в котором f полностью раскладывается на линейные множители. В случае, когда f *неприводим*, построить такое поле L совсем просто, можно, например, положить $L = K[x]/(f)$. Идейно эта конструкция ('**поле корней многочлена**') была известна еще Лагранжу, но на промышленную основу ее поставили только Дедекинд и Кронекер.

• **Поле достижимых чисел.** Рассмотрим точки 0 и 1 на вещественной прямой. Теория Галуа учит, что все точки комплексной плоскости, которые

можно построить отправляясь от 0 и 1 при помощи циркуля и линейки, образуют поле, называемое полем достижимых чисел и обозначаемое $\mathbb{K}$. Это поле **квадратично замкнуто**, т.е. для любого $x \in \mathbb{K}$ существует $y \in \mathbb{K}$ такое, что $y^2 = x$. В то же время, отрицательный вопрос на классическую проблему удвоения куба означает в точности, что $\sqrt[3]{2}$ не является достижимым числом.

Рациональные числа строились как дроби, числитель и знаменатель которых суть целые числа. В действительности эта конструкция без всяких изменений проходит вообще для любой области целостности R и приводит к построению некоторого поля $Q(R)$, называемого полем частных кольца R. Вот несколько примеров этой конструкции, которые изучаются во втором семестре.

• **Поле рациональных дробей.** Применяя ту же конструкцию образования дробей к кольцу многочленов $K[x]$ над некоторым полем K, мы получим поле $K(x)$, называемое **полем рациональных функций от одной переменной**.

• **Поле рациональных дробей от нескольких переменных.** В связи с изучением многочленов от нескольких переменных мы построим и изучим поле $K(x_1, \dots, x_n)$, являющееся полем частных кольца многочленов $K[x_1, \dots, x_n]$.

• **Поле формальных рядов Лорана.** Следующее поле является полем частных кольца $K[[x]]$ формальных степенных рядов

$$K((x)) = \Big\{ \sum a_i x^i \mid a_i \in K, a_i = 0 \text{ для почти всех } i < 0 \Big\}.$$

§ 2. Поля алгебраических чисел

Одним из основных типов задач, обсуждаемых в школьной алгебре, является “избавление от иррациональности в знаменателе”. При этом, например, предлагается переписать число

$$\frac{1}{3 + 5\sqrt[3]{2} + 7\sqrt[3]{4}}$$

в виде $x + y\sqrt[3]{2} + z\sqrt[3]{4}$, где x, y, z снова рациональные числа. В настоящем пункте мы покажем, что это возможно ровно потому, что $\mathbb{Q}(\sqrt[3]{2})$ образует поле и приведем другие примеры полей алгебраических чисел.

1. Квадратичные поля. Наиболее известны квадратичные поля, которые получаются присоединением к $\mathbb{Q}$ квадратного корня из *одного* элемента $d \in \mathbb{Q}$, который не является квадратом в $\mathbb{Q}$. Ясно, что не теряя общности число d можно считать *целым* и *бесквадратным*. Второе из этих требований означает, что d не делится на квадрат простого числа. Поле

$$\mathbb{Q}(\sqrt{d}) = \mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$$

называется **квадратичным полем**. При этом d называется **дискриминантом** $\mathbb{Q}(\sqrt{d})$. Поле $\mathbb{Q}(\sqrt{d})$ называется **вещественным квадратичным**, если $d > 0$ и **мнимым квадратичным**, если $d < 0$. Проверка всех аксиом поля для этого случая очевидна, единственный чуть менее тривиальный момент – показать, что число $a + b\sqrt{d} \neq 0$ обратимо в этом поле. Для этого достаточно заметить, что $(a + b\sqrt{d})^{-1} = (a - b\sqrt{d})/(a^2 - db^2)$ (знаменатель не обращается в 0, так как d не является квадратом в $\mathbb{Q}$).

Вот несколько наиболее часто используемых примеров квадратичных полей.

• Присоединяя квадратный корень из $d=-1$, мы получаем поле **гауссовых чисел** $\mathbb{Q}(i)=\{a+bi \mid a,b\in\mathbb{Q}\}$.

• Присоединяя квадратный корень из $d=-3$, мы получаем поле **эйзенштейновых чисел** $\mathbb{Q}(i)=\{a+b\omega \mid a,b\in\mathbb{Q}\}$, где $\omega=-\frac{1+\sqrt{-3}}{2}$.

• Присоединяя квадратный корень из $d=2$, мы получаем поле $\mathbb{Q}(\sqrt{2})=\{a+b\sqrt{2} \mid a,b\in\mathbb{Q}\}$.

• Присоединяя квадратный корень из $d=5$, мы получаем поле **золотого сечения** $\mathbb{Q}(\sqrt{5})=\{a+b\sqrt{5} \mid a,b\in\mathbb{Q}\}$.

2. Поле $\mathbb{Q}(\sqrt[3]{2})$. Рассмотрим теперь множество чисел

$$\mathbb{Q}(\sqrt[3]{2})=\left\{a+b\sqrt[3]{2}+c\sqrt[3]{4} \mid a,b,c\in\mathbb{Q}\right\}.$$

Совершенно ясно, что все такие числа образуют коммутативное ассоциативное кольцо с 1. Доказать, что это поле чуть сложнее. Для этого, прежде всего, полезно знать, что 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$ линейно независимы над $\mathbb{Q}$, т.е. что если $a+b\sqrt[3]{2}+c\sqrt[3]{4}=0$ для некоторых $a,b,c\in\mathbb{Q}$, то $a=b=c=0$. Это весьма частный случай теоремы Безиковича, которую мы обсуждаем в части, посвященной линейной алгебре. Конечно, для этого случая этот результат легко проверить и методами школьной алгебры,

Задача. Докажите, что 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$ линейно независимы над $\mathbb{Q}$.

Указание. Если $c\neq 0$, то на равенство $a+b\sqrt[3]{2}+c\sqrt[3]{4}=0$ можно смотреть как на *квадратное* уравнение относительно $\sqrt[3]{2}$. Выражая его корень через a,b,c мы видим, что $\sqrt[3]{2}$ является *квадратичной* иррациональностью, т.е. принадлежит квадратичному полю $\mathbb{Q}(\sqrt{d})$, где $d=b^2-4ac$. Возводя представление $\sqrt[3]{2}=u+v\sqrt{d}$ в куб, мы легко придем к противоречию.

Итак, пусть $\alpha=a+b\sqrt[3]{2}+c\sqrt[3]{4}\neq 0$. Будем искать $\beta=\alpha^{-1}$ в виде $\beta=x+y\sqrt[3]{2}+z\sqrt[3]{4}$, где $x,y,z\in\mathbb{Q}$. Вычисляя произведение $\alpha\beta$, получаем

$$(ax+2cy+2bz)+(bx+ay+2cz)\sqrt[3]{2}+(cx+by+az)\sqrt[3]{4}=1.$$

В силу линейной независимости 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$ это равенство эквивалентно следующей системе линейных уравнений относительно x,y,z

$$\begin{cases} ax+2cy+2bz=1\\ bx+ay+2cz=0\\ cx+by+az=0\end{cases}$$

Матрица этой системы

$$g=\begin{pmatrix} a & 2c & 2b\\ b & a & 2c\\ c & b & a\end{pmatrix}$$

является квазициркулянтом и ее определитель равен

$$\det(g)=a^3+2b^3+4c^3-6abc.$$

Почему $\det(g)\neq 0$, для всех $\alpha\neq 0$? Это можно доказать, например, следующим образом[6]. Воспользуемся тождествами

$$u^3+v^3+w^3-3uvw=(u+v+w)(u^2+v^2+w^2-uv-uw-vw),$$
$$u^2+v^2+w^2-uv-uw-vw=\frac{1}{2}((u-v)^2+(u-w)^2+(v-w)^2).$$

[6]W.Więslaw, Algebra geometryczna, Wroclaw, 1974, p.1–405, стр.60–61.

Подставляя в первое из этих тождеств $u = a$, $v = b\sqrt[3]{2}$, $w = c\sqrt[3]{4}$, получим

$$a^3 + 2b^3 + 4c^3 - 6abc = \frac{\alpha}{2}((a - b\sqrt[3]{2})^2 + (a - c\sqrt[3]{4})^2 + (b\sqrt[3]{2} - c\sqrt[3]{4})^2).$$

Если левая часть равна 0, то, так как $\alpha \neq 0$, то

$$(a - b\sqrt[3]{2})^2 + (a - c\sqrt[3]{4})^2 + (b\sqrt[3]{2} - c\sqrt[3]{4})^2 = 0,$$

что невозможно так как тогда все слагаемые должны были бы равняться 0. Но ведь, если $a - b\sqrt[3]{2} = a - c\sqrt[3]{4} = 0$, то в силу линейной независимости 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$, получим $a = b = c = 0$.

Таким образом, определитель матрицы g всегда $\neq 0$ и, значит, система уравнений относительно x, y, z имеет единственное решение. Теперь уже совсем несложно найти это решение:

$$\beta = \frac{(a^2 - 2bc) + (2c^2 - ab)\sqrt[3]{2} + (b^2 - ac)\sqrt[3]{4}}{a^3 + 2b^3 + 4c^3 - 6abc}.$$

В дальнейшем мы будем обычно опускать такого рода тривиальные выкладки необходимые для проверки того, что то или иное множество чисел образует поле.

Задача. Избавьтесь от иррациональности в знаменателе выражения, приведенного в самом начале этого параграфа.

Задача. Докажите, что поля $\mathbb{Q}(\sqrt[3]{2})$ и $\mathbb{Q}(\omega\sqrt[3]{2})$ изоморфны.

Указание. Проверьте, что $a+b\sqrt[3]{2}+c\sqrt[3]{4} \mapsto a+b\omega\sqrt[3]{2}+c\omega^2\sqrt[3]{4}$ задает требуемый изоморфизм.

§ 3. Поля функций

• **Поле рациональных функций на алгебраической кривой.** Сейчас мы обсудим широкое обобщение поля $K(x)$. Пусть X – **плоская алгебраическая кривая**, заданная уравнением $f(x, y) = 0$, где $f \in K[x, y]$. Если многочлен f неприводим, то с f можно связать поле $K(X)$, называемое полем рациональных функций на кривой X. Для этого рассмотрим подкольцо $R = K[x, y]_{(f)}$ кольца $K[x, y]$, состоящее из дробей **определенных на** X, т.е. всех дробей вида g/h, $g, h \in K[x, y]$ таких, что $f \not| h$. Две дроби, g_1/h_1 и g_2/h_2 определенные на X называются **равными на** X, если $f|g_1h_2 - h_1g_2$. Тогда $K(X)$ – это множество классов дробей определенных на X относительно равенства на X. Например, если X – окружность, заданная уравнением $x^2 + y^2 = 1$, то $(1 + y)/x = x/(1 - y)$ в $K(X)$. Операции в $K(X)$ – это обычные операции над дробями (нужно еще, конечно, доказать, что они корректно определены!)

• **Поле мероморфных функций.** Пусть $U \subseteq \mathbb{C}$ – открытое связное подмножество. **Мероморфной функцией** на U называется функция вида $U \mapsto \overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, $z \mapsto f(z)/g(z)$, где $f, g : U \longrightarrow \mathbb{C}$ – две голоморфные функции, причем $g \neq 0$. Мероморфные функции на U образуют поле K_U относительно сложения и умножения функций. В самом деле, если $f/g \neq 0$, то $(f/g)^{-1} = g/f$. Это поле является полем частных кольца $\mathcal{O}_U$ голоморфных функций на U.

• **Поле алгеброидных функций.** По определению это алгебраическое замыкание поля мероморфных функций. В конкретных терминах это поле можно описать следующим образом. Рассмотрим, прежде всего, поле формальных рядов вида

$$\sum_{j=0}^{\infty} a_j z^{(m+j)/n}, \quad a_j \in \mathbb{C}, a_0 \neq 0, m \in \mathbb{Z}, n \in \mathbb{N}.$$

Это алгебраически замкнутое поле. Выделим теперь такие ряды, которые сходятся в какой-то окрестности 0 и которые можно продолжить до конечнозначной аналитической функции на плотном открытом подмножестве в $\mathbb{C}$, не имеющей никаких особенностей, кроме алгебраических. Такие ряды тоже образуют алгебраически замкнутое поле, которое, как раз и

является алгебраическим замыканием поля $K_{\mathbb{C}}$ мероморфных функций. Обратите внимание, что по нашему определению, алгеброидная функция это именно ряд, так что, скажем, $z^{1/2}$ и $-z^{1/2}$ являются различными функциями!

• **Поле эллиптических функций.** Зафиксируем *мнимое* число $\tau \in \mathbb{C} \setminus \mathbb{R}$ и рассмотрим решетку $L = \{m + n\tau \mid m, n \in \mathbb{Z}\}$ в комплексной плоскости. Мероморфная на $\mathbb{C}$ функция f называется **эллиптической** относительно решетки L, если $f(z + \omega) = f(z)$ для всех $z \in \mathbb{C}$ и всех $\omega \in L$. В анализе такие функции часто называются **двоякопериодическими** с периодами 1 и τ, но мы будем придерживаться терминологии, принятой в геометрии и теории чисел. Легко видеть, что эллиптические функции относительно данной решетки L образуют поле. Классически известно[7], что это поле порождается $\wp$-функцией Вейерштрасса

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in L \setminus \{0\}} \left(\frac{q}{(z-\omega)^2} - \frac{1}{\omega^2} \right)$$

и ее производной

$$\wp'(z) = -2 \sum_{\omega \in L} \frac{1}{(z-\omega)^3},$$

которые связаны между собой соотношением

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3,$$

где

$$g_2 = 60 \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^4}, \qquad g_3 = 140 \sum_{\omega \in L \setminus \{0\}} \frac{1}{\omega^6}.$$

Таким образом, поле эллиптических функций является полем рациональных функций на плоской алгебраической кривой X с уравнением $y^2 = 4x^3 - g_2 x - g_3$. Такая кривая X называется **эллиптической кривой**.

§ 4. Алгебры кватернионов

1. Тело кватернионов. Первый пример некоммутативного тела был построен ирландским математиком Гамильтоном. Его конструкция обобщает обычную конструкцию комплексных чисел как пар вещественных чисел, только в теперь вместо пар рассматриваются *четверки*, с чем и связано название этих новых **гиперкомплексных** чисел: **кватернионы**.

Рассмотрим четыре **кватернионных единицы** $1, i, j, k$, линейно независимых над $\mathbb{R}$ и обозначим через $\mathbb{H}$ множество их линейных комбинаций с вещественными коэффициентами

$$\mathbb{H} = \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\},$$

где вместо $a \cdot 1$ мы пишем просто a. Линейная независимость $1, i, j, k$ означает, что два кватерниона $z_1 = a_1 + b_1 i + c_1 j + d_1 k$ и $z_2 = a_2 + b_2 i + c_2 j + d_2 k$ в том и только том случае равны, когда их соответствующие коэффициенты совпадают, т.е. $a_1 = a_2$, $b_1 = b_2$, $c_1 = c_2$, и $d_1 = d_2$. Складываются кватернионы почленно, т.е.

$$(a_1 + b_1 i + c_1 j + d_1 k) + (a_2 + b_2 i + c_2 j + d_2 k) = \\ (a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k.$$

[7]Курант, Гурвиц, Теория функций, ч.II, гл.I, §§ 8,9.

а их умножение продолжает по линейности уже известное нам из примера § ? умножение кватернионных единиц:

$$(a_1 + b_1 i + c_1 j + d_1 k) \cdot (a_2 + b_2 i + c_2 j + d_2 k) =$$
$$(a_1 a_2 - b_1 b_2 - c_1 c_2 - d_1 d_2) + (a_1 b_2 + b_1 a_2 + c_1 d_2 - d_1 c_2) i +$$
$$(a_1 c_2 + c_1 a_2 + d_1 b_2 - b_1 d_2) j + (a_1 d_2 + d_1 a_2 + b_1 c_2 - c_1 b_2) k.$$

Сами эти *формулы* были известны уже Эйлеру, но только Гамильтон осознал, что их можно рассматривать как определение произведения в $\mathbb{R}^4$. Все аксиомы кольца, кроме ассоциативности умножения, очевидны. В ассоциативности умножения тоже несложно убедиться непосредственно, но в Главе 4 мы предложим **гораздо** более интеллигентный способ проверки ассоциативности, основанный на реализации кватернионов как 2×2 комплексных матриц (или, что почти то же самое, как 4×4 вещественных матриц). Однако это умножение некоммутативно так как, например, $ij = k$, но $ji = -k$. Остается проверить еще, что все ненулевые элементы элементы получившегося кольца обратимы. В самом деле, пусть $z = a + bi + cj + dk \neq 0$. По определению это означает, что хотя бы один из коэффициентов a, b, c, d отличен от 0 а так как все они являются вещественными, то $a^2 + b^2 + c^2 + d^2 \neq 0$. Теперь легко убедиться в том, что

$$(a + bi + cj + dk)^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}(a - bi - cj - dk).$$

2. Обобщенные алгебры кватернионов. Пусть K – поле, A – четырехмерное векторное пространство над K с базисом $1, i, j, k$ и умножением, определенным условиями

$$i^2 = a, \qquad j^2 = b, \qquad ij = -ji = k.$$

Пользуясь ассоциативностью легко восстановить остальную часть таблицы умножения:

$$k^2 = -ab, \qquad ik = -ki = aj, \qquad jk = -kj = -bi.$$

Получившаяся алгебра называется **обобщенной алгеброй кватернионов** над K и обозначается $\left(\frac{a,b}{K}\right)$. В частном случае $K = \mathbb{R}$, $a = b = -1$, получается алгебра **гамильтоновых кватернионов** $\mathbb{H} = \left(\frac{-1,-1}{\mathbb{R}}\right)$. Впрочем, часто $\left(\frac{a,b}{K}\right)$ называется просто **алгеброй кватернионов**, в этом случае об $\mathbb{H}$ говорят как о **классических кватернионах**.

Теорема. *При любых $a, b \in K^*$ алгебра $\left(\frac{a,b}{K}\right)$ является простой центральной алгеброй над K.*

Теорема. *Алгебра кватернионов $\left(\frac{a,b}{K}\right)$ в том и только том случае является телом, когда уравнение $ax^2 + by^2 = z^2$ не имеет в K нетривиальных решений.*

$x = y = z = 0$

Доказательство. см. Пирс, с.29.

§ 5. Циклические алгебры

Конструкция Диксона (1906 год). Пусть L/K – циклическое расширение Галуа, иными словами, конечное нормальное сепарабельное расширение такое, что группа Галуа $\mathrm{Gal}(L/K) = \langle\sigma\rangle$ циклическая. При этом $o(\sigma) = n = \dim_K L$. Зафиксируем $a \in K^*$ и символ x. Положим

$$D = L \cdot 1 \oplus L \cdot x \oplus \ldots \oplus L \cdot x^{n-1}$$

и будем умножать элементы D используя дистрибутивность и следующие правила $x^s = a$, $x \cdot b = b^\sigma x$. Легко видеть, что $K \leq C(D)$, так что размерность K-алгебры D равна n^2. Эта алгебра обозначается через $(L/K, \sigma, a)$ и называется циклической алгеброй, ассоциированной с $(L/K, \sigma)$ и $a \in K^*$.

Тело $\mathbb{H}$ Гамильтоновых кватернионов является частным случаем этой конструкции, получающимся при $K = \mathbb{R}$, $L = \mathbb{C}$, σ – комплексное сопряжение, $a = -1$. Иными словами, $\mathbb{H} = (\mathbb{C}/Re, , -1)$. В этом случае $x = j$, так что

$$\mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}ij = \mathbb{C} \oplus \mathbb{C}j$$

и $j(a + bi) = (a - bi)j = \sigma(a + bi)j$.

Еще один способ построить $(L/K, \sigma, a)$ – использовать кольцо косых многочленов $R = L[t, \sigma]$. Тогда $(L/K, \sigma, a) = L[t, \sigma]/(t^n - a)$ Заметим, что это фактор-кольцо по идеалу, порожденному *центральным* многочленом $x^n - a$. Пока мы не знаем, тело это или нет!

§ 6. Тело Гильберта

Циклические алгебры конечномерны над центром. Сейчас мы построим пример бесконечномерного тела.

Гильбертов твист. Пусть R – коммутативное кольцо, $\sigma : R \longrightarrow R$, $a \mapsto a^\sigma$, – эндоморфизм R. Тогда, как мы знаем из ? можно определить скрученные кольца многочленов и формальных степенных рядов $R[x, \sigma]$, $R[[x, \sigma]]$, в которых переменная x коммутирует с коэффициентами по следующей формуле $xa = a^\sigma x$. Если предполагать дополнительно, что σ является не просто эндоморфизмом, а автоморфизмом кольца R, то $x^{-1}b^\sigma = (b^\sigma)^{\sigma^{-1}} x^{-1} = bx^{-1}$. Таким образом, в этом случае можно определить скрученные кольца многочленов Лорана и формальных рядов Лорана $R[x, x^{-1}, \sigma]$, $R((x, \sigma))$. В этих кольцах умножение задается формулой

$$\left(\sum a_i x^i\right)\left(\sum b_j x^j\right) = \sum a_i b_j^{\sigma^i} x^{i+j}$$

Ограничимся теперь случаем, когда $R = K$ – поле.

Теорема. *Если $\sigma \in \mathrm{Aut}(K)$ – автоморфизм поля K, то $D = K((x, \sigma))$ – тело. Если обозначить через $K_0 = K^\sigma$ поле неподвижных элементов K, то*

$$C(D) = \begin{cases} K_0, & \text{если } o(\sigma) = \infty, \\ K_0((x^n)), & \text{если } o(\sigma) = n. \end{cases}$$

Доказательство. Возьмем $f \in K((x, \sigma))$, $f \neq 0$. Подберем x^i так, чтобы $fx^i = a_0 + a_1 x + a_2 x^2 + \ldots$, где $a_0 \neq 0$. blablabla

В частности, тело $D = K((x, \sigma))$ в том и только том случае конечномерно над центром, когда $o(\sigma) < \infty$. В оригинальном примере Гильберта $K = \mathbb{Q}((t))$, а σ это $\mathbb{Q}$-автоморфизм поля $\mathbb{Q}((t))$, отображающий t и $2t$, так что $xa(t) = a(2t)x$. Это был первый пример тела бесконечномерного над центром.

Почему в элементарных учебниках нет примеров тел? Единственный пример тела, фигурирующий в элементарных текстах – это тело Гамильтоновых кватернионов $\mathbb{H}$. В то же время имеется большое количество классических конструкций тел, принадлежащих Гильберту, Диксону–Веддербарну, Нетер, Кебе, Оре, Хану–Мальцеву–Магнусу, ..., Кону, Платонову–Дракслу, Скофилду и другим. Почему же эти примеры не попали в учебники алгебры? Дело в том, что в некоммутативном случае понятие **поля** расщепляется на несколько понятий:

- **тело**,
- **простое артиново кольцо**,
- **простое кольцо**.

В коммутативном случае все эти понятия совпадают. Вспомним деление алгебры на High-school algebra (многочлены, дроби, матрицы), College algebra (группы, кольца, модули), University algebra (категории, пучки, когомологии). При помощи College algebra довольно легко построить большое число примеров *простых артиновых колец*. Однако вопрос о том, будут ли эти кольца телами, как правило, не допускает простого решения в терминах College algebra. Это арифметический вопрос, требующий для своего решения либо методов University algebra, либо довольно искусственного и громоздкого перевода рассуждений относящихся к University algebra на язык College algebram, что обычно приводит к *чудовищным* текстам. Что касается простых (неартиновых) колец, то 'их сложность не поддается анализу' ('is beyond analysis' – а также, конечно, 'beyond algebra').

Тема 3: СПЕЦИАЛЬНЫЕ ЭЛЕМЕНТЫ КОЛЕЦ

§ 1. Мультипликативная группа

1. Обратимые элементы кольца. Пусть R кольцо с 1. Рассмотрим его **мультипликативный моноид** $R^\times$, т.е. алгебраическую систему R в которой мы **забыли** (alias **стерли** или проигнорировали) все элементы структуры, кроме умножения и 1. Заметим, впрочем, что для колец без делителей 0 мультипликативным моноидом обычно называется множество $R^\bullet = R \setminus \{0\}$.

Говоря об обратимых элементах кольца R имеют в виду обратимые элементы его мультипликативного моноида.

Определение. *Говорят, что элемент $x \in R$* **обратим слева**, *если существует такое $y \in R$, что $yx = 1$. В этом случае y называется* **левым обратным** *для x. Говорят, что элемент $x \in R$* **обратим справа**, *если существует такое $z \in R$, что $xz = 1$. В этом случае z называется* **левым обратным** *для x. Элемент $x \in R$ называется* **обратимым**, *если он обратим как слева, так и справа, т.е. если найдется такой $y \in X$, что $yx = 1 = xy$. В этом случае y называется* **обратным** *к x и обозначается x^{-1}.*

Задача. Докажите, что если $x, y, x + y \in R^*$, то $x(x + y)^{-1}y = y(x + y)^{-1}x$.

Обратимые слева **или** справа элементы называются **односторонне обратимыми**. Как мы знаем, в *ассоциативном* кольце каждый **двусторонне обратимый** элемент x, т.е. элемент, у которого существует как левый обратный y, так и правый обратный z, обратим: $z = (yx)z = y(xz) = y$. Тем не менее вскоре мы приведем примеры таких элементов x, у которых существует левый обратный, но не правый обратный или, наоборот, правый обратный, но не левый обратный. Следующие две задачи показывают, что двусторонние обратные это *в точности* единственные односторонние обратные.

Задача. Докажите, что если у элемента x есть два различных правых обратных, то он имеет бесконечно много правых обратных.

Решение. Обозначим через X множество всех правых обратных к элементу x, зафиксируем какой-то элемент $y \in X$ и рассмотрим отображение $\phi : X \longrightarrow X$, $z \mapsto y + zx - 1$. Так как у x нет левых обратных, то $y \notin \phi(X)$. С другой стороны, так как $\phi(u) = \phi(v)$ влечет $ux = vx$ влечет $u = v$, то отображение ϕ инъективно. Тем самым, мы построили инъективное, но не сюръективное отображение множества X в себя. Это возможно только если множество X бесконечно.

В действительности, использованное в этой задаче рассуждение можно прочесть и в обратную сторону.

Задача. Докажите, что если у элемента x существует единственный правый обратный, то x обратим.

Решение. В самом деле, пусть $xy = 1$. Если $yx \neq 1$, то $y + yx - 1 \neq y$ – второй правый обратный для x.

2. Примеры мультипликативных групп.

- Если $R = T$ – тело, то $T^* = T^\bullet$.
- $\mathbb{Z}^* = \{\pm 1\}$.
- Если R – область целостности (например, поле), то $R[x]^* = R^*$.

- $R[[x]]^* = \{a_0 + a_1x + \ldots \in R[[x]] \mid a_0 \in R^*\}$.
- Если $R = R_1 \oplus \ldots \oplus R_s$, то $R^* = R_1^* \times \ldots \times R_s^*$.
- Если A – абелева группа, то $\mathrm{End}(A)^* = \mathrm{Aut}(A)$.
- $M(n, R)^* = \mathrm{GL}(n, R)$.
- Если $S \leq R$ – подкольцо в R, то $S^* \leq R^*$.

Задача. Постройте пример подкольца, для которого $S^* \neq S \cap R^*$. Докажите, что $C(R)^* = C(R) \cap R^*$.

§ 1. Конечность по Дедекинду

1. Слабая конечность. Кольцо R называется **слабо 1-конечным** (weakly 1-finite) или **конечным по Дедекинду** (Dedekind finite), если односторонняя обратимость в нем совпадает с двусторонней обратимостью. Иными словами, если $xy = 1$ для каких-то $x, y \in R$, то $yx = 1$. Большинство колец, с которыми сталкивается начинающий, слабо 1-конечны, вот несколько примеров:

- тела;
- коммутативные кольца;
- кольца матриц над коммутативными кольцами, это вытекает из теории определителей: если $xy = e$, то $\det(xy) = \det(x)\det(y) = 1$;
- почти коммутативные кольца. Напомним, что кольцо R называется **почти коммутативным**, если оно конечно порожденно как модуль над коммутативным кольцом;
- Кольцо без делителей 0. В самом деле, пусть $xy = 1$. Тогда $x(yx - 1) = 0$, так что $yx = 1$;
- Нетерово кольцо.

В то же время, кольца операторов в бесконечномерных пространствах как правило не обладают этим свойством. Вот очевидный пример в кольце бесконечных конечно столбцовых матриц:

$$x = \begin{pmatrix} 0 & 1 & 0 & \ldots \\ 0 & 0 & 1 & \ldots \\ 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & \ldots \\ 0 & 1 & 0 & \ldots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

§ 2. Рациональные тождества

> A theorem may be hard to discover, even though, once discovered, it is easy to prove. The point of this note is to emphasize that completely nonrigorous (some may say nonsensical) reasoning is perfectly acceptable in the discovery stage, and that may furnish clues that enable one to make a good guess. Proofs can come later.
>
> Walter Rudin

1. Рациональные тождества. Для характеризации радикала Джекобсона нам понадобится следующая лемма, которая чрезвычайно важна и сама по себе.

Лемма. *Если* $1+yx \in R^*$, *то* $1+xy \in R^*$.

Доказательство. Мы утверждаем, что

$$(1+xy)^{-1} = 1 - x(1+yx)^{-1}y.$$

Проверим, для примера, что это действительно правый обратный для $1+xy$. В самом деле, легко видеть, что $x(1+yx) = (1+xy)x$ и $y(1+xy) = (1+yx)y$ и, таким образом,

$$(1+xy)(1-x(1+yx)^{-1}y) = 1+xy-(1+xy)x(1+yx)^{-1}y =$$

$$1+xy1-x(1+yx)(1+yx)^{-1}y = 1+xy-xy = 1.$$

Проверка с другой стороны аналогична.

2. Страшная тайна. Начинающего, вероятно, волнует вопрос, как можно *догадаться* до таких формул как $(1+xy)^{-1} = 1-x(1+yx)^{-1}y$? Понятно, что после того, как такая формула угадана, *доказать* ее ничего не стоит. Разумеется, мало кто, кроме специалистов по некоммутативным кольцам *помнит* подобные формулы, называемые **рациональными тождествами** (rational identities). Однако искусство быстро находить их составляет часть ремесла математика. Обычно такие формулы ищутся при помощи вычислений с матрицами или формальными рядами от некоммутирующих переменных. Например, в нашем случае сравнение рядов

$$(1+xy)^{-1} = 1-xy+(xy)^2-\dots, \qquad (1+yx)^{-1} = 1-yx+(yx)^2-\dots,$$

сразу подсказывает ответ. После того, как ответ угадан, обычно не представляет никакого труда вывести его из аксиом кольца. Вычисление с рядами представляет собой страшную тайну, которая известна, по всей видимости, всем профессионалам. Я придумал для себя этот метод сам, но потом обратил внимание на статьи[8,9] – то, что могут независимо придумать *три* математика, может придумать свинья.

§ 3. Делители 0 и регулярные элементы

1. Делители 0 и регулярные элементы. Пусть теперь R ассоциативное кольцо с 1. Элемент $x \in R$ называется **левым делителем нуля**, если существует такой $y \in R$, $y \neq 0$, что $xy = 0$. Если такого y не существует, т.е. если $xy = 0$ влечет $y = 0$, то x называется **регулярным справа** (sic!). На регулярный *справа* элемент можно сокращать *слева*: $xy = xz$ влечет $y = z$. В самом деле, равенство $xy = xz$ можно переписать в виде $x(y-z) = 0$, и, так как x регулярен справа, $y - z = 0$.

Понятия **правого делителя нуля** и **регулярного слева** элемента определяются аналогично. Элемент $x \in R$ называется **правым делителем нуля**, если существует такой $y \in R$, $y \neq 0$, что $yx = 0$, в противном случае x называется **регулярным слева**. Элемент называется **регулярным**, если он регулярен как слева, так и справа. Множество всех регулярных элементов кольца R будет обозначаться через $\mathrm{Reg}(R)$. Большинство колец, с которыми мы будем работать на первых порах, коммутативны, и там можно говорить просто о **делителях нуля** и **регулярных элементах**.

[8] P.R.Halmos, Does mathematics have elements, Math. Intelligencer, 1981, v.3, p.147–153.

[9] W.Rudin, Unique inverses are two-sided, Amer. Math. Monthly, 1985, August-September, p.489–490.

Лемма. *Множество левых/правых регулярных элементов замкнуто относительно умножения*

Доказательство. Нам нужно показать, что если x, y регулярны слева, то xy тоже регулярен слева. В самом деле, пусть $z \neq 0$. Тогда $z(xy) = (zx)y$. Так как x регулярен слева, то $zx \neq 0$, а так как y регулярен слева, то $(zx)y \neq 0$. Но это и значит, что $z(xy) \neq 0$ для любого $z \neq 0$. Доказательство для регулярных справа элементов аналогично.

Так как, кроме того, 1 является, а 0 не является регулярным элементом, эта лемма означает, что множество регулярных элементов коммутативного кольца образует **мультипликативную систему**.

Задача (принцип Дирихле). Докажите, что в конечном кольце каждый регулярный элемент обратим.

Указание. Иными словами, каждый необратимый элемент является делителем 0.

Задача. Пусть $x, y \in R$, причем x регулярен. Покажите, что если $x^m = y^m$ и $x^n = y^n$, то $x^d = y^d$ для $d = \gcd(m, n)$.

Задача. Пусть R коммутативное кольцо. Тогда матрица $x \in M(n, R)$ тогда и только тогда является левым делителем нуля в $M(n, R)$, когда она является правым делителем 0 в $M(n, R)$, когда $\det(x)$ является делителем 0 в R.

2. Кольца без делителей 0.

Определение. *Ассоциативное кольцо R называется* **кольцом без делителей** *0, alias* **целостным кольцом**, *если в нем нет ненулевых делителей 0. Ненулевое кольцо без делителей 0 называется* **областью**. *Ненулевое коммутативное целостное кольцо называется* **областью целостности**.

nullteilerfrei, Bereich, Integritätsbereich, integral, domain, integral domain
В кольце без делителей 0 $\mathrm{Reg}(R) = R^\bullet$.
Примеры!!!

Теорема Титчмарша. Сейчас мы приведем *нетривиальный* пример кольца без делителей 0, возникающий в анализе. Обозначим через R множество непрерывных функций на $[0, \infty)$ со значениями в $\mathbb{R}$. Это множество образует абелеву группу относительно обычного (поточеченого) сложения функций. Как мы знаем, оно является кольцом относительно умножения функций.

Задача. Приведите примеры нетривиальных делителей 0 в кольце непрерывных функций.

В действительности, обычно в качестве умножения в R рассматривается не обычное умножение функций, а свертка $f * g$, определяемая в этом случае как

$$(f * g)(t) = \int_0^t f(t - s)g(s)ds.$$

Следующий исключительно важный результат был получен Титчмаршем в 1924 году.

Теорема Титчмарша. *Кольцо R непрерывных функций на $[0, \infty)$ относительно свертки является областью целостности.*

Эта теорема лежит в основе 'операционного исчисления'. Как мы узнаем в дальнейшем, любая область целостности (даже если в ней нет 1) вкладывается в некоторое поле. В данном случае это и будет поле 'операторов' Хэвисайда. На этом пути Микусиньский[10] дал чисто алгебраическое обоснование операционного исчисления, вне всякой связи с преобразованием Лапласа.

§ 4. Нильпотенты и унипотенты

1. Нильпотенты. Кольцо дуальных чисел содержит нильпотенты.

Определение. *Элемент кольца $x \in R$ называется* **нильпотентом**, *если $x^n = 0$ для некоторого натурального n.*

Очевидно, 0 нильпотентен, нильпотенты же, отличные от 0, называются **нетривиальными**. Кольцо R называется **приведенным**, если в нем нет нетривиальных идемпотентов. Нам уже встречались кольца, с нетривиальными нильпотентами. Так в кольце $\mathbb{Z}/4\mathbb{Z}$ имеем $2 \neq 0$, но $2^2 = 4 = 0$. Как мы вскоре увидим, нетривиальные нильпотенты есть и в кольце матриц $M(n, K)$ при любом $n \geq 2$.

2. Приведенные кольца. Во многих вопросах условие отсутствия в кольце делителей нуля является черезчур ограничительным. Например, кольцо $C(X)$ непрерывных функций на топологическом пространстве X как правило имеет делители 0. Сейчас мы определим важнейшее ослабление условия целостности.

Определение. *Кольцо R называется* **приведенным**, *если в нем нет нетривиальных нильпотентных элементов. Иными словами, для $x \in R$ из того, что $x^n = 0$ вытекает, что $x = 0$.*

Если кольцо не является приведенным, то в нем есть делители 0. Однако, как показывает следующий пример, и приведенное кольцо может иметь делители 0. Скажем, в кольце $\mathbb{Z}/6\mathbb{Z}$ делители 0 есть, так как в нем $2 \cdot 3 = 0$, но нетривиальных нильпотентов там нет.

Приведем несколько примеров колец, не являющихся приведенными.

1) При любых $p \in \mathbb{P}$, $m \geq 2$ элемент p кольца $\mathbb{Z}/p^m\mathbb{Z}$ нильпотентен.

2) При $m \geq 2$ элемент x кольца $K[x]/(x^m)$ усеченных многочленов нильпотентен. В частности, это относится к кольцу дуальных чисел $K[x]/(x^2)$.

3) При любом $n \geq 2$ кольцо матриц $M(n, K)$ содержит громадное количество нильпотентов. Например, нильпотентна любая стандартная матричная единица e_{ij}, $i \neq j$.

3. Основные свойства нильпотентов.

Лемма. *Пусть $x, y \in R$ – коммутирующие элементы кольца R.*

1) Если x нильпотентен, то xy нильпотентен.

2) Если x и y нильпотентны, то $x - y$ нильпотентен.

3) Если x нильпотентен, а $y \in R^*$, то $y - x \in R^*$.

Доказательство. 1) Если $x^m = 0$, то $(xy)^m = x^m y^m = 0$.

2) Если $x^m = y^n = 0$, то по биному Ньютона $(x - y)^{m+n} = 0$.

[10] Я.Микусиньский, Операторное исчисление. – М., ИЛ. 1959.

3) Так как $y - x = y(1 - y^{-1}x)$, то достаточно показать, что $1 - y^{-1}x \in R^*$. По пункту 1) элемент $z = y^{-1}x$ нильпотентен, скажем, $z^n = 0$. Поэтому $(1 - z)(1 + z + \ldots + z^{n-1}) = 1 - z^n = 1$.

Определение. *Элемент $u \in R$ такой, что $u - 1$ нильпотентен, называется* **унипотентом**.

Обозначим множество унипотентных элементов кольца R через $U(R)$. По определению $U(R) = 1 + N(R)$. По лемме любой унипотентный элемент обратим, т.е. $U(R) \subseteq R^*$. Вообще говоря, произведение двух унипотентов уже совсем не обязательно является унипотентом. Например, если

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad v = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

то их произведение uv не будет унипотентом (см. пример в § ?, связанный с числами Фибоначчи).

Предложение. *1) Если $u \in U(R)$, то $u^{-1} \in U(R)$.*

2) Если $u, v \in U(R)$ – коммутирующие унипотенты, то $uv \in U(R)$.

Доказательство. 1) Пусть $u = e - x$, $x^m = 0$. Тогда $u^{-1} = e + x + \ldots + x^{m-1}$, причем $(x + \ldots + x^{m-1})^m = 0$.

2) Пусть $u = e - x$, $v = e - y$, где $x^m = y^n = 0$. Тогда $uv = e - x - y + xy$, причем так как x и y коммутируют, то $(x + y - xy)^{m+n} = 0$.

Следствие. *Если кольцо R коммутативно, то $U(R) \leq R^*$.*

§ 5. Анализ по Ферма: дуальные числа

> Немного найдется строк в истории человеческой мысли, столь знаменитых как короткое сочинение, посланное Ферма Декарту на следующий день после опубликования "Геометрии" – "Метод отыскания наибольших и наименьших значений ".
>
> L.Brunschwieg, "Этапы математической философии"

> But so great is the average person fear of the infinite that to this day calculus is being taught as a study of *limit processes*, instead of what it really is: *infinitesimal analysis*.
>
> Rudy Rucker, "Infinity and the Mind"

Следующая замечательная конструкция была предложена Пьером де Ферма в 1638 году[11] и переоткрыта Клиффордом и Штуди в XIX веке. Текст Ферма не дает никаких возможностей для истолкования его в инфинитезимальном смысле, в духе Лейбница, или в смысле теории пределов. Напротив, это чисто алгебраический текст, притом абсолютно строгий[12].

---

[11] Р.Декарт, Геометрия, с приложением избранных работ П.Ферма и переписки Декарта, М.–Л., 1938, с.154–155

[12] Часто высказывается точка зрения, что Ферма не может считаться основателем анализа потому, что у него нет 'бесконечных процессов'. При этом происходит подмена понятий: в самом инфинитезимальном исчислении нет никаких 'предельных переходов', пределы – это лишь один из способов **обоснования** инфинитезимального исчисления, притом далеко не самый красивый или удобный.

1. Кольцо дуальных чисел. Пусть K – произвольное поле, например, $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, а d – не принадлежащий K элемент такой, что $d^2 = 0$. Рассмотрим множество $K[d]$, элементами которого являются формальные линейные комбинации 1 и d с коэффициентами из K (мы пишем просто $x + dy$ вместо $x1 + yd$, причем $x + dy = u + dv$ в том и только том случае, когда $x = u$ и $y = v$), сложение покомпонентное, а умножение продолжает умножение 1 и d по линейности, иными словами, для $x, y, u, v \in K$ имеем $(x + dy)(u + dv) = xy + d(xv + yu)$. Построенное так кольцо называется **кольцом дуальных чисел** над K.

2. Нильпотенты. В частности, в $K[d]$ имеем $(dy)^2 = 0$ для любого $y \in K$. Элементы $dy \in K[d]$ являются **актуально бесконечно малыми первого порядка**, рассматриваемыми с точностью до бесконечно малых второго порядка (понятия, относящиеся к актуально бесконечно малым, называются обычно **инфинитезимальными** – 'infinitesimal'). Это значит, что кольцо $K[d]$ уже достаточно для определения **первых производных**, причем в нем производная находится непосредственно как **отношение бесконечно малых**, а не как предел отношений конечных величин.

Порядок на дуальных числах над $\mathbb{R}$

3. Производные. Функция $f : K \mapsto K$ называется **дифференцируемой**, если ее можно продолжить на $K[d]$ так, чтобы отношение

$$f'(x) = \frac{f(x + dy) - f(x)}{dy},$$

называемое **производной** f, не зависело от выбора $y \in K$. Для этого необходимо (но, вообще говоря, не достаточно), чтобы $f(x + dy) - f(x)$ было бесконечно малой. Разумеется, для дифференцируемой функции мы можем определить производную так, как это делается в анализе, заменив dy на dx:

$$f'(x) = \frac{f(x + dx) - f(x)}{dx},$$

замечательно, что нам не приходится при этом писать знака предела!

4. Производные многочленов. Вычислим, например, первую производную функции $x \mapsto x^n$ в K. Для этого придадим x бесконечно малое приращение dy (традиционный аналист написал бы в этом месте dx) и вычислим отвечающее ему приращение функции $x \mapsto x^n$. Вычисляя $(x + dy)^n$ по биному Ньютона, и пользуясь тем, что бесконечно малые второго и более высоких порядков **равны нулю**, получаем $(x+dy)^n - x^n = nx^{n-1}dy$, так что производная функции $x \mapsto x^n$ равна

$$\frac{(x + dy)^n - x^n}{dy} = \frac{nx^{n-1}dy}{dy} = nx^{n-1}.$$

5. Производные экспоненты и логарифма. Продифференцируем теперь экспоненту и логарифм. Экспонента должна оставаться гомоморфизмом аддитивной структуры в мультипликативную. С другой стороны, в бесконечно малой окрестности 0 очень просто построить гомоморфизм, переводящий сложение в умножение: $\exp : dx \mapsto 1 + dx$. Таким образом, экспоненту естественно определить формулой

$$e^{x+dy} = e^x e^{dy} = e^x(1 + dy).$$

Тем самым, логарифм в инфинитезимальной окрестности 1 должен иметь вид $\ln : 1 + dx \mapsto dx$. Таким образом,

$$\frac{e^{x+dy} - e^x}{dy} = \frac{e^x(1+dy) - e^x}{dy} = e^x.$$

и, аналогично,

$$\frac{\ln(x+dy) - \ln(x)}{dy} = \frac{\ln(1+dy/x)}{dy} = \frac{1}{x}.$$

6. Производные тригонометрических функций. В качестве еще одного примера вычислим производную функции $x \mapsto \sin(x)$. Снова придадим x бесконечно малое приращение dy и вычислим отвечающее ему приращение функции $x \mapsto \sin(x)$. Вычисляя $\sin(x+dy)$ по формуле сложения для тригонометрических функций и, воспользовавшись тем, что $\cos(dy) = 1$, а $\sin(dy) = dy$ (бесконечно малые высших порядков равны нулю!), получаем

$$\frac{\sin(x+dy) - \sin(x)}{dy} = \frac{\sin(x)\cos(dy) + \cos(x)\sin(dy) - \sin(x)}{dy} = \cos(x).$$

Комментарий 1. Адресованное Декарту письмо Ферма, содержащее метод нахождения максимумов и минимумов на основе дуальных чисел, вне всякого сомнения представляет собой один из **самых** поразительных документов в истории человечества. Это письмо опередило свое время более, чем на три века. Получающийся на этом пути подход к анализу гораздо современнее не только XIX века, но и первой половины XX века. Точка зрения Ферма начала возрождаться в работах итальянских алгебраических геометров, но ее триумфальное возвращение произошло только в 1960-х годах, когда Гротендик установил, что **правильным** подходом к инфинитезимальному исчислению конечного порядка – **единственным** возможным подходом в случае положительной характеристики – является переход от полей к кольцам, содержащим нильпотенты произвольных порядков[13]. Но именно об этом и говорит Ферма всем, кто умеет читать: “и не одно сокровище, быть может, минуя внуков к правнукам уйдет, и снова скальд чужую песню сложит, и как свою ее произнесет”.

Д.Мамфорд, Лекции о кривых на алгебраической поверхности, М., Мир, 1968, 236с.

Комментарий 2. Эта идея легко переносится и на производные высших порядков, при этом вместо кольца двойных чисел нужно рассматривать его обобщение – кольцо **усеченных многочленов** $K[t]/(t^n)$, которое мы введем в § ?. Разумеется, единственная сложность здесь – доопределить функцию, априори заданную на K, на всем $K[t]/(t^n)$. Конечно, для функций, заданных рядами (а **только** такие функции и рассматривались классиками!) этого вопроса не возникает.

Комментарий 3. Построить чисто алгебраическую версию дифференциального исчисления **бесконечного** порядка *не в пример сложнее*. Наиболее известный подход предложен А.Робинсоном в его ‘нестандартном анализе’, где

[13]Как замечает по этому поводу Мамфорд [Mu], с.25, еще в 1950-х годах возможность существования не равных 0 функций, все значения которых равны 0, представлялась многим ‘скандальной’.

он вводит нестандартную модель ${}^*\mathbb{R}$ поля $\mathbb{R}$ вещественных чисел, содержащую актуально бесконечно малые и бесконечно большие произвольных порядков, причем так, что все свойства поля $\mathbb{R}$, *выразимые на языке первого порядка*, сохраняются в ${}^*\mathbb{R}$ (в частности, ${}^*\mathbb{R}$ является полем, а не просто кольцом!). Имеются и другие чисто алгебраические обоснования анализа, основанные на теоретико-кольцевых характеризациях алгебры дифференцируемых функций.

§ 6. Идемпотенты и инволюции

1. Идемпотенты. Сейчас мы определим еще один важнейший тип элементов, которые отвечают, в частности, за разложение колец в прямую сумму.

Определение. *Элемент кольца $e \in R$ называется* **идемпотентом**, *если $e^2 = e$.*

В этом случае $e^n = e$ для любого натурального n. В каждом кольце 0 и 1 являются идемпотентами – это **тривиальные** идемпотенты. Ясно, что любой нетривиальный идемпотент является делителем 0: $e(1-e) = 0 = (1-e)e$.

Лемма. *Пусть $e \neq 0, 1$ – нетривиальный идемпотент кольца R. Тогда $1-e$ тоже является идемпотентом.*

Доказательство. В самом деле, $(1-e)^2 = 1 - 2e + e^2 = 1 - e$.

Идемпотенты e и f называются **ортогональными**, если $ef = 0 = fe$. Таким образом, e и $f = 1 - e$ – ортогональные **дополнительные** идемпотенты, $1 = e + f$.

Задача. Проверьте, что если e и f коммутирующие идемпотенты, то $(e-f)^2$ тоже идемпотент. Обобщите.

Задача. Докажите, что для любого идемпотента e в кольце R множество $eRe = \{exe, x \in R\}$ является подкольцом в R с единицей e.

Подкольцо eRe вообще говоря, неунитально, если $e \neq 1$.

Теперь мы можем определить булево кольцо, как кольцо, **все** элементы которого являются идемпотентами.

Задача. Докажите, что булево кольцо коммутативно.

Решение. В булевом кольце $2 = 2^2 = 4$, таким образом, $2 = 0$ или, что то же самое, $-1 = 1$. С другой стороны, из равенства $x + y = (x+y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y$ вытекает, что $xy + yx = 0$. Тем самым, $xy = -yx = yx$.

2. Теорема Джекобсона. Продолжим вариации на тему доказанной в последней задаче коммутативности булевых колец.

Задача. Докажите, что если в кольце R для любого x имеет место равенство $x^3 = x$, то R коммутативно.

Решение. (З.И.Боревич, спецкурс 'Коммутативные кольца', 1986/87 учебный год) Сравнивая равенства $(x+x)^3 = x + x$ и $(x^2 - x)^3 = x^2 - x$, мы видим, что $6x = 0$ и $3x^2 = 3x$ для всех $x \in R$. Легко видеть, что множество $S = \{3x \mid x \in R\}$ является подкольцом, причем для любого $y \in S$ выполняется $y^2 = y$ (почему?) Таким образом, S булево кольцо и, значит, коммутативно. Отсюда и из того, что $6xy = 0 = 6yx$ сразу следует, что $3xy = 3yx$ для всех $x, y \in R$. Теперь, сравнивая $(x+y)^3 = x + y$ и $(x-y)^3 = x - y$, мы видим, что $2xy^2 + 2yxy + 2y^2x = 0$. Умножая это равенство на y слева и справа и вычитая получившиеся равенства, мы видим, что $2xy = 2yx$ для всех x, y. Те, кто не забыли сказанное выше, понимают, что это ровно то, что требовалось.

Если Вы уловили идею этого рассуждения, то Вам уже совсем просто будет решить следующую задачу.

Задача. Докажите, что если в кольце R для любого x имеет место равенство $x^4 = x$, то R коммутативно.

В действительности все такие утверждения допускают следующее замечательное обобщение.

Теорема Джекобсона. *Пусть R – кольцо, в котором для любого элемента x найдется такое натуральное число $n > 1$, что $x^n = x$. Тогда R коммутативно.*

Эта теорема служит, *в частности*, и очень широким обобщением малой теоремы Веддербарна.

3. Инволюции. Элемент $g \in R$ называется **инволюцией**, если $g^2 = e$.

Задача. Убедитесь, что если e – идемпотент, то $1 - 2e$ – инволюция. Если $2 \in \mathrm{Reg}(R)$, то $e \mapsto 1 - 2e$ задает инъективное отображение из множества идемпотентов кольца R в множество инволюций в этом кольце. Если $2 \in R^*$, то это соответствие является биективным.

Решение. В самом деле, $(1-2e)(1-2e) = 1-4e+4e^2 = 1$. Если $1-2e = 1-2f$, то $2e = 2f$ и, если 2 регулярна, то $e = f$. Если $2 \in R^*$, то сопоставление $g \mapsto (1-g)/2$ задает обратное отображение из инволюций в идемпотенты.

Например, при этом соответствии идемпотенту $e_{11} + \ldots + e_{rr}$ в кольце матриц $M(n, K)$ отвечает инволюция $-e_{11} - \ldots - e_{rr} + e_{r+1,r+1} + \ldots + e_{nn}$. Многие традиционные матричные вычисления основаны на использовании инволюций, а не идемпотентов [Artin], [Dieudonne], [O'Meara]. Как мы только что убедились, в случае $2 \in R^*$ это одно и то же. Тем не менее, традиционные пристрастия очень сильны: специалисты по теории колец предпочитают работать с идемпотентами, а специалисты по теории групп – с инволюциями.

§ 7. Центральные идемпотенты

1. Центральные идемпотенты. Идемпотент e называется **центральным**, если $e \in \mathrm{Cent}(R)$. Если в R существует нетривиальный центральный идемпотент e, то оно допускает нетривиальное разложение в прямую сумму $R = Re \oplus R(1-e)$.

В кольце матриц $M(n, K)$ при $n \geq 2$ есть нетривиальные идемпотенты $e_{11}, \ldots, e_{nn}$, но, тем не менее, оно не раскладывается в прямую сумму, поскольку они не центральные.

В кольце $R = \mathbb{Z}/6\mathbb{Z}$ можно взять дополнительные идемпотенты 3 и 4, $3^2 \equiv 3 \pmod 6$, $4^2 \equiv 4 \pmod 6$, $3+4 \equiv 1 \pmod 6$. Они определяют разложение кольца R в прямую сумму $R = \{0,3\} \oplus \{0,2,4\}$ двух подколец, изоморфных $\mathbb{F}_2$ и $\mathbb{F}_3$, соответственно.

Задача. Если e – идемпотент в R такой, что $eR = Re$, то $e \in \mathrm{Cent}(R)$.

Решение. В самом деле, умножая равенство $eR = Re$ на e слева, мы видим, что $eR = eRe$. Это значит, что умножение элементов из $eR = Re$ на e как справа, так и слева является тождественным преобразованием этого множества. Тем самым для любого $x \in R$ имеем $ex = exe = xe$.

Задача. Покажите, что в приведенном кольце R все идемпотенты центральны.

Решение. Пусть $e \in R$ – идемпотент, а $x \in R$ – произвольный элемент. Тогда $(exe - ex)^2 = (exe - xe)^2 = 0$. Поэтому $ex = exe = xe$.

Как разложимые, так и неприведенные кольца имеют делители 0. Однако неразложимость и приведенность представляют собой независимые условия. Убедимся в этом на примере колец $\mathbb{Z}/m\mathbb{Z}$. Пусть $p, q \in \mathbb{P}$. Тогда кольцо $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ является приведенным и неразложимым. Кольцо $\mathbb{Z}/p^2\mathbb{Z}$ неприведенное и неразложимое. Кольцо $\mathbb{Z}/pq\mathbb{Z}$ приведенное и разложимое. Наконец, кольцо $\mathbb{Z}/p^2q\mathbb{Z}$ неприведенное и разложимое.

Пусть R – коммутативное кольцо. Идемпотент $e \in R$ называется **неразложимым**, если не допускает нетривиального представления в виде суммы ортогональных идемпотентов. Иными словами, если $e = f + g$, где $f^2 = f$, $g^2 = g$, $fg = 0$, то $f = 0$ или $g = 0$.

Задача. Покажите, что два различных неразложимых идемпотента ортогональны.

Решение. Пусть $e^2 = e$, $f^2 = f$, – неразложимые идемпотенты. Предположим, что $ef \neq 0$. Тогда $e = ef + e(1-f)$ – разложение e в сумму двух ортогональных идемпотентов. Так как e неразложим, а $ef \neq 0$, то $e(1-f) = 0$, т.е. $e = ef$. Точно так же из неразложимости f следует, что $f = ef$. Поэтому $e = f$.

§ 8. Радикальные элементы

1. Радикал Джекобсона. Элемент $x \in R$ называется **радикальным**, если $1 + yx \in R^*$ для всех $y \in R$. Множество $J(R) = \mathrm{Rad}(R)$ всех радикальных элементов в R называется радикалом **Джекобсона кольца** R:

$$J(R) = \{x \in R \mid \forall y \in R, 1 + xy \in R^*\}.$$

В силу леммы предыдущего пункта радикал Джекобсона можно еще охарактеризоватиь как

$$J(R) = \{x \in R \mid \forall y \in R, 1 + yx \in R^*\}$$

или как

$$J(R) = \{x \in R \mid \forall y, z \in R, 1 + yxz \in R^*\}$$

(в самом деле $1 + yxz \in R$ в том и только том случае, когда $1 + zyx \in R$). В дальнейшем мы дадим другие характеризации радикала Джекобсона, но бесхитростная характеризация, которую мы только что привели, является с большим отрывом наиболее полезной.

Теорема. *Радикал Джекобсона $J(R)$ является идеалом в R.*

Доказательство. Покажем, прежде всего, что $J(R)$ – аддитивная подгруппа. Если $x, y \in J(R)$, то для любого $z \in R$ имеем $1 + z(x + y) = 1 + zx + zy$. Так как $1 + zx \in R^*$, то этот элемент можно переписать в виде $1 + zx + zy = (1 + zx)(1 + (1 + zx)^{-1}zy)$, где оба множителя справа принадлежат $J(R)$ по определению радикала Джекобсона. Ясно, что $J(R)$ устойчиво относительно умножения на элементы кольца R слева: если $x \in J(R)$, $z \in R$, то для любого $y \in R$ имеем $1 + y(zx) = 1 + (yz)x \in R^*$. Таким образом, радикал Джекобсона является *левым* идеалом. Так как определение радикала Джекобсона право-лево симметрично, то он является и *правым* идеалом.

Задача. Если R коммутативно, то $\mathrm{Nil}(R) \subseteq J(R)$.

Глава ?: КОЛЬЦА ФУНКЦИЙ И КОЛЬЦА ОПЕРАТОРОВ

Много вопросов, разбираемых обычно в анализе бесконечно малых, я здесь разрешил при помощи правил элементарной алгебры, чтобы лучше выявилась сущность того или иного метода.

Леонард Эйлер 'Введение в анализ бесконечно малых', т.I

Betrogene Betrüger, eure Ringe sind alle drei nicht echt. Der echte Ring vermutlich geht verloren.

Gotthold Efraim Lessing, 'Nathan der Weise'

Чем больше я размышляю над принципами теории функций, а делаю я это непрестанно, тем тверже становится мое убеждение, что ее следует воздвигать на фундаменте алгебраических истин и что потому неверен способ, при котором, наоборот, для обоснования более простых и фундаментальных алгебраических положений принимается, говоря кратко, "трансцендентное".

Карл Вейерштрасс[14]

Математический анализ слишком труден для аналистов.

Давид Гильберт

A theory is worth studying if it has at least three distinct good hard examples.

Adrian Albert

Сейчас мы рассмотрим три различные возможности определить структуру кольца на множестве отображений $X \longrightarrow Y$. Три ключевых слова, которыми описываются эти структуры, это **композиция, умножение и свертка**. Во всех случаях, которые мы будем рассматривать, множество $Y = A$ будет абелевой группой. Тем самым у нас определено поточечное сложение функций. Если $f, g : X \longrightarrow A$, то $f+g$ определяется посредством $(f+g)(x) = f(x)+g(x)$. Однако при этом мы будем рассматривать три совершенно различных типа произведений:

• композиция '$\circ$': $(f \circ g)(x) = f(g(x))$;

• поточечное произведение функций '$\cdot$': $(f \cdot g)(x) = f(x)g(x)$;

• свертка функций '$*$', когда $(f * g)(x)$ определяется как сумма произведений $f(y)g(z)$ по некоторым парам (y, z), подробнее см. § ?.

§ 1. Кольцо операторов

Кольцо отображений $A \longrightarrow A$ относительно *композиции* называется кольцом операторов.

1. Кольцо эндоморфизмов/операторов. Конструкция кольца матриц является частным случаем следующей весьма общей конструкции колец. Пусть A – произвольная абелева группа. Рассмотрим множество $R = \mathrm{End}(A)$ всех эндоморфизмов группы A со следующими операциями: обычным поточечным сложением функций и композицией отображений как умножением.

[14]Из письма Вейерштрасса Г.А.Шварцу, цитируется по книге Ф.Клейн, 'История математики в XIX веке'.

Теорема. *Для любой абелевой группы A множество* $\mathrm{End}(A)$ *с так определенными операциями представляет собой ассоциативное кольцо с 1.*

Доказательство. Заметим, прежде всего, что $\mathrm{End}(A)$ является абелевой группой по сложению,

$$(x+y)(a+b) = x(a+b)+y(a+b) = (x(a)+x(b))+(y(a)+y(b)) = \\ (x(a)+y(a))+(x(b)+y(b)) = (x+y)(a)+(x+y)(b),$$

так что сумма $x+y$ двух эндоморфизмов является эндоморфизмом (чем мы пользовались в этом вычислении?) С другой стороны, композиция отображений ассоциативна, причем тождественное отображение id выступает в роли нейтрального элемента. Осталось лишь проверить дистрибутивность умножения относительно сложения. Правая дистрибутивность очевидна, так как вообще для любых $x, y, z \in \mathrm{Map}(A, A)$ и любого $a \in A$ имеем

$$(x+y)z(a) = (x+y)(z(a)) = x(z(a))+y(z(a)) = xz(a)+yz(a).$$

С другой стороны, если, кроме того, $x \in \mathrm{End}(A)$, то

$$x(y+z)(a) = x((y+z)(a)) = x(y(a)+z(a)) = x(y(a))+x(z(a)) = zy(a)+xz(a)$$

(где мы воспользовались тем, что x эндоморфизм?), так что для этого случая имеет место и дистрибутивность слева. В частности, в $\mathrm{End}(A)$ умножение двусторонне дистрибутивно относительно сложения, так что $\mathrm{End}(A)$ действительно является кольцом.

Алгебраисты обычно называют $\mathrm{End}(A)$ **кольцом эндоморфизмов**, в то время как большинство остальных математиков предпочитают в этом случае говорить о **кольцах операторов**. Заметим, что **каждое** кольцо есть подкольцо кольца эндоморфизмов некоторой абелевой группы, например, $R \leq \mathrm{End}(A^+)$.

Задача. Пусть A – кольцо. Композиция двух кольцевых эндоморфизмов A является кольцевым эндоморфизмом. Почему, тем не менее, никто не говорит о кольце $\mathrm{End}(A)$ эндоморфизмов кольца A?

2. Кольцо R-эндоморфизмов. В различных вопросах возникают подкольца кольца $\mathrm{End}(A)$. Например, если A является R-модулем, то обычно рассматриваются только **R-эндоморфизмы** модуля A, т.е. такие эндоморфизмы аддитивной группы A, которые R-однородны, т.е. $x(\lambda a) = \lambda x(a)$ для любых $\lambda \in R$, $a \in A$. Здесь мы предполагаем, что A левый R-модуль, для правого R-модуля нужно предполагать $x(a\lambda) = x(a)\lambda$. Обозначим множество всех R-эндоморфизмов модуля A через $\mathrm{End}_R(A)$. По определению $\mathrm{End}_R(A) \leq \mathrm{End}(A)$. При этом $\mathrm{End}(A)$ можно истолковать как $\mathrm{End}_{\mathbb{Z}}(A)$.

Для свободного правого R-модуля $A = R^n$, алгебра R-эндоморфизмов естественно истолковывается как алгебра матриц $\mathrm{End}_R(R^n) = M(n, R)$.

§ 2. Кольцо непрерывных операторов

Пусть U – банахово (например, гильбертово) пространство с нормой $\| \ \|$. Пусть $L(U,U) = \mathrm{End}(U)$ – кольцо **всех** линейных операторов в U. В конечномерном случае любой линейный оператор автоматически непрерывен, но в

бесконечномерном случае это не так. Сейчас мы построим некоторые важнейшие подкольца в $\mathrm{End}(U)$, определяемые в терминах непрерывности.

1. Ограниченные операторы. Линейный оператор $x : U \longrightarrow U$ называется **непрерывным** alias **ограниченным**, если

$$\|x\| = \sup_{\|u\|\leq 1} \|xu\| = \sup_{\|u\|=1} \|xu\| < \infty.$$

В силу линейности $\|xu\| \leq \|x\|\|u\|$, причем $\|x\|$ является наименьшим из всех $c \geq 0$ таких, что $\|xu\| \leq c\|u\|$. Вещественное число $\|x\|$ называется **нормой** ограниченного оператора x (это так называемая **операторная норма**). Легко видеть, что

$$\|x+y\| \leq \|x\| + \|y\|, \qquad \|xy\| \leq \|x\|\|y\|.$$

Докажем, для примера, второе из этих неравенств. В самом деле, $\|xyu\| \leq \|x\|\|yu\| \leq \|x\|\|y\|\|u\|$, так что, действительно, $\|xy\| \leq \|x\|\|y\|$ в силу минимальности $\|xy\|$. Тем самым, множество $B(U,U)$ всех ограниченных линейных операторов образует кольцо.

Отступление. На самом деле $B(U,U)$ не просто кольцо или алгебра, а удовлетворяет сильным дополнительным условиям, выражаемым в терминах операторной нормы. А именно, $B(U,U)$ является **банаховой алгеброй** относительно нормы $\|\ \|$. Это значит, что $B(U,U)$ полна в топологии, определенной этой нормой, а $\|xy\| \leq \|x\|\|y\|$. Если U – гильбертово пространство, то U естественно изоморфно своему сопряженному U^*, так что сопряженный x^* к оператору x можно в этом случае снова рассматривать как оператор из U в U. При этом сопряженный к ограниченному оператору ограничен и $\|x^*\| = \|x\|$, так что $B(U,U)$ является **$*$-банаховой алгеброй**. Более того, из $(xu, xu) = (u, x^*x, u)$ следует, что $\|x^*x\| = \|x\|^2$, так что в действительности для гильбертовых пространств $B(U,U)$ представляет собой **C^*-алгебру**. В случае гильбертовых пространств многие важные классы операторов *автоматически* оказываются непрерывными. Например, один из самых ранних результатов, которые можно отнести собственно к функциональному анализу, теорема Хеллингера–Теплица (1910 год), утверждает, что если оператор x в гильбертовом пространстве самосопряжен, $x = x^*$, то он непрерывен.

2. Обратимые ограниченные операторы. Следующий результат весьма удивителен и достаточно небанален (не пробуйте доказать это в качестве упражнения!) Он утверждает, что обратный к обратимому ограниченному оператору автоматически ограничен[15].

Теорема Банаха об обратном операторе. *Пусть U – банахово пространство, тогда* $\mathrm{GL}(U) \cap B(U,U) = B(U,U)^*$.

3. Вполне непрерывные операторы. Линейный оператор $x : U \longrightarrow U$ называется **вполне непрерывным** alias **компактным**, если замыкание множества $\{xu \mid \|u\| \leq 1\}$ компактно. Обозначим через $C(U,U)$ множество всех компактных операторов. Ясно, что каждый компактный оператор ограничен, так что $C(U,U) \leq B(U,U)$. Ясно, что сумма двух компактных операторов снова является компактным оператором. Легко доказать, что произведение компактного оператора на любой ограниченный оператор снова является компактным. Поэтому $C(U,U) \trianglelefteq B(U,U)$.

Если U – гильбертово пространство, то x в том и том случае компактен, когда x^* компактен.

[15] Колмогоров и Фомин, ibid. стр.225.

4. Конечномерные операторы. Оператор x называется **конечномерным**, если его образ конечномерен. Легко видеть, что множество $F(U,U)$ всех конечномерных операторов образует подкольцо в $B(U,U)$. Очевидно, что любой конечномерный оператор компактен, так что $F(U,U) \leq B(U,U)$. Оказывается, если оператор x компактен, а его образ $x(U)$ замкнут, то оператор x конечномерен[16].

§ 3. Кольцо функций

Основным примером колец в математике являются кольца функций относительно *умножения*.

1. Кольца функций. Рассмотрим отображения произвольного множества X в какое-то коммутативное ассоциативное кольцо R с 1, скажем, в $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, или $\mathbb{C}$. Множество R^X всех таких отображений с поточечными операциями сложения и умножения

$$(f+g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x),$$

называется **кольцом R-значных функций** на X. Из свойств операций в кольце R сразу вытекает, что R^X действительно будет коммутативным ассоциативным кольцом с 1.

Теперь мы готовы дать основное определение этой главы. Любое подкольцо в R^X называется **кольцом функций**. Обычно рассматривают только **алгебры функций**, т.е. такие подкольца в R^X, которые содержат все постоянные функции $x \mapsto c$, где c – фиксированный элемент кольца R.

Эта конструкция колец является одной из центральных в анализе, топологии, дифференциальной и алгебраической геометрии. Много примеров колец функций получаются при помощи следующей руководящей идеи: X и R несут дополнительную структуру и, соответственно, рассматриваются не все функции из X в R, а только функции, сохраняющие эту дополнительную структуру, например, непрерывные, дифференцируемые, голоморфные, регулярные, рациональные, и т.д. Именно таким образом получаются кольцо $C(X)$ непрерывных вещественнозначных функций на топологическом пространстве, кольцо $C^r(X)$ r-кратно дифференцируемых функций на дифференцируемом многообразии, кольцо $C^\infty(X)$ гладких функций, кольцо $C^\omega(X)$ аналитических функций на вещественном или комплексном аналитическом многообразии и т.д.[17]. В следующих параграфах мы обсудим конструкции и важнейшие свойства этих колец.

Задача. Докажите, что $\left(R^X\right)^* = \left(R^*\right)^X$. Иными словами, функция f тогда и только тогда обратима в R^X, когда все ее значения обратимы в R. В частности, если $R = K$ поле, для обратимости функции в K^X достаточно, чтобы она не принимала нулевых значений.

Задача. Докажите, что единственными идемпотентами кольца K^X являются характеристические функции χ_Y подмножеств $Y \subseteq X$.

2. Замена переменной. Пусть R – коммутативное кольцо с 1, а X и Y – два произвольных множества. Тогда определены кольца функций R^X и R^Y. Пусть

[16]У.Рудин, Функциональный анализ. Мир, М., 1975, с.1–443. Теорема 4.18.

[17]Н.Бурбаки, Дифференцируемые и аналитические многообразия", 1975

теперь $\phi : X \longrightarrow Y$ – произвольное отображение. Тогда сопоставление функции $f \in R^Y$ функции $f \circ \phi \in R^X$ задает гомоморфизм колец $\phi^* : R^Y \longrightarrow R^X$, называемый **заменой переменной**. В самом деле, нужно убедиться в том, что для любых двух функций $f, g \in R^Y$ имеют место равенства $(f+g)\circ\phi = f\circ\phi + g\circ\phi$ и $(f\cdot g)\circ\phi = f\circ\phi\cdot g\circ\phi$. Но эти равенства сразу вытекают из определения действий над функциями (проверьте!). То, что при этом 1 переходит в 1 сразу следует из того, что постоянная функция $y \mapsto 1$ является левым нулем относительно композиции функций. Обратите внимание, что отображению из X в Y соответствует гомоморфизм колец в обратном направлении, из R^Y в R^X, а вовсе не из R^X в R^Y. С точки зрения теории категорий это означает, что функтор $X \mapsto R^X$ из категории множеств в категорию колец **контравариантен**.

Частные случаи замены переменной:

• **Ограничение.** В частном случае, когда $X \subseteq Y$, замена переменных, отвечающая вложению $X \hookrightarrow Y$, называется ограничением (restriction) $\mathrm{res}^Y_X : R^Y \longrightarrow R^X$, $f \mapsto f|_X$.

• **Значение в точке.** По определению операций над функциями $\mathrm{ev}_x : R^X \longrightarrow R$, $f \mapsto f(x)$, является гомоморфизмом колец (в действительности, R-алгебр). Этот гомоморфизм обычно называется **эвалюацией** (evaluation) в точке x.

3. Идеалы в кольце функций. Пусть теперь X – произвольное множество, а R – произвольное коммутативное кольцо. Рассмотрим кольцо функций R^X с обычными поточечными операциями сложения и умножения функций. Для подмножества $Y \subseteq X$ обозначим через I_Y множество функций, обращающихся в 0 на Y:

$$I_Y = \{f \in R^X \mid \forall y \in Y, f(y) = 0\}.$$

Легко видеть, что I_Y – идеал в R^X. В самом деле, если $f, g \in I_Y$, то $(f-g)(y) = f(y) - g(y) = 0$ для любого $y \in Y$, так что $f - g \in I_Y$. С другой стороны, для любой функции $h \in R^X$ и любого $y \in Y$ имеем $(hf)(y) = h(y)f(y) = h(y)0 = 0$, так что $hf \in I_Y$.

Рассмотрим теперь наиболее важный для приложений случай, когда $R = K$ – поле. Пусть $Y = \{x\}$ – одноточечное множество. Легко видеть, что в этом случае идеал $I_Y = \mathfrak{m}_x = \{f \in K^X \mid f(x) = 0\}$ максимален в R^X. В действительности, мы покажем, что в некоторых важнейших кольцах функций $A \leq K^X$ нет никаких максимальных идеалов, кроме $A \cap \mathfrak{m}_x$, $x \in X$. В тех случаях, когда результат такого типа удается получить, он оказывается центральным фактом всей теории, который позволяет свести изучение геометрии пространства X к чисто алгебраическим свойствам кольца A (теорема Гильберта о нулях, преобразование Гельфанда и т.д.).

4. Фактор-кольца. Что такое R^X/I_Y??

5. Разложение кольца функций. Разложению $X = X_1 \sqcup \ldots \sqcup X_n$ множества X в свободное объединение отвечает разложение кольца функций на нем в прямую сумму

$$R^{X_1 \sqcup \ldots \sqcup X_n} \cong R^{X_1} \oplus \ldots \oplus R^{X_n}.$$

Изоморфизм между этими кольцами устанавливается посредством

$$f \mapsto (f|_{X_1}, \ldots, f|_{X_n}).$$

В частности, если множество $X = \{x_1, \ldots, x_n\}$ конечно, описанная конструкция устанавливает изоморфизм $R^X \longrightarrow R^n = R \oplus \ldots \oplus R$, $f \mapsto (f(x_1), \ldots, f(x_n))$.

§ 4. Кольцо непрерывных функций

В действительности, как правило рассматриваются не произвольные функции, а только функции некоторых классов. Наиболее общий и важный пример получается, если рассматривать непрерывные функции.

1. Кольцо непрерывных функций. Пусть X топологическое пространство, а K – нормированное поле, например, $\mathbb{R}$, $\mathbb{C}$ или $\mathbb{Q}_p$. Тогда имеет смысл говорить о непрерывных функциях $X \longrightarrow K$. Чаще всего это понятие используется для случая, когда X – локально компактное хаусдорфово топологическое пространство. Начинающщий должен представлять себе, что речь здесь идет об $X = \mathbb{R}, \mathbb{R}^n, [0,1]$ или $\mathbb{Z}$.

• **Кольцо функций, имеющих предел в каждой точке** $A(X)$. Мы уже знаем, что

$$\lim_x(f+g) = \lim_x(f) + \lim_x(g), \qquad \lim_x(fg) = \lim_x(f)\lim_x(g)$$

для всех $x \in X$. Поэтому $A(X)$ – кольцо.

• **Кольцо ограниченных функций** $B(X)$. Очевидно, что сумма и произведение ограниченных функций ограничены.

• **Кольцо непрерывных функций** $C(X)$. Хорошо известно, что сумма и произведение непрерывных функций непрерывны.

• **Кольцо непрерывных ограниченных функций** $BC(X) = B(X) \cap C(X)$. Сразу вытекает из двух предыдущих примеров.

• **Кольцо непрерывных функций с компактным носителем** $C_c(X)$. Функции, для которых $\mathrm{Supp}(f)$ компактен. Как обычно,

$$\mathrm{Supp}(f+g) \subseteq \mathrm{Supp}(f) \cup \mathrm{Supp}(g), \qquad \mathrm{Supp}(fg) \subseteq \mathrm{Supp}(f) \cap \mathrm{Supp}(g).$$

Отсюда следует, что $C_c(X)$ кольцо.

• **Кольцо непрерывных функций обращающихся в 0 на бесконечности** $C_0(X)$. Будем говорить, что f **обращается в 0 на бесконечности**, если для любого m существует компактное подмножество Y такое, что $|f(x)| < \frac{1}{2^m}$ для всех $x \in X \setminus Y$. Легко проверить, что сумма и произведение двух функций, обращающихся в 0 на бесконечности, обращается в 0 на бесконечности.

Задача. Докажите, что χ_Y в том и только том случае непрерывна, когда оба множества $Y, X \setminus Y$ замкнуты в X. В частности, кольцо $C(X)$ в том и только том случае разлагается в прямую сумму, когда пространство X несвязно.

2. Непрерывная замена переменной. Пусть теперь $f : X \longrightarrow Y$ непрерывное отображение топологических пространств. Тогда замена неременной

§ 5. Напоминания из общей топологии

В следующем параграфе мы докажем замечательный классический результат[18], утверждающий, что **все** свойства компактного хаусдорфова пространства X выражаются в терминах кольца $C(X)$ непрерывных функций на нем

[18] И.М.Гельфанд, А.Н.Колмогоров, О кольцах непрерывных функций на топологических пространствах. – Докл. АН СССР, 1939, т.22, с.11–15.

и, таким образом, топология компактов не имеет никакого отношения к топологии, а является просто разделом коммутативной алгебры. Однако, чтобы установить эту связь, мы должны напомнить несколько определений и два классических результата общей топологии.

Топологическое пространство X называется **компактным**, если из любого его покрытия открытыми множествами можно выделить конечное подпокрытие. Иными словами, если $X = \bigcup U_\alpha$, где объединение берется по произвольному семейству U_α, $\alpha \in \Omega$, открытых множеств, то можно выбрать конечное подсемейство $U_{\alpha_1}, \dots, U_{\alpha_n}$ этого семейства такое, что $X = U_{\alpha_1} \cup \dots \cup U_{\alpha_n}$.

Комментарий. Чаще всего компактность используется в сочетании с хаусдорфовостью. Свойства *хаусдорфовых* компактных пространств, называемых **компактами**, настолько замечательны, что многие авторы даже включают хаусдорфовость в определение компактности. В этом случае, то, что мы называем компактностью, называется квазикомпактностью. Однако в алгебре и алгебраической геометрии очень часто возникают и компактные пространства, не являющиеся хаусдорфовыми.

По определению, **хаусдорфовы** пространства удовлетворяют аксиоме отделимости T_2:

T_2. *Для любых точек $x, y \in X$, $x \neq y$, найдутся такие открытые окрестности $x \in U$ и $y \in V$, что $U \cap V = \emptyset$.*

Хаусдорфовы пространства названы так в честь Феликса Хаусдорфа, который первым явно выделил это условие в 1914 году в своей книге 'Grundzüge der Mengenlehre'. Буква 'T' в названии аксиом отделимости – это первая буква немецкого слова `Trennbarkeit`, по английски эти аксиомы называются separation axioms. Оказывается *компактное* хаусдорфово пространство *автоматически* удовлетворяет значительно более сильной аксиоме отделимости T_4:

T_4. *Для любых замкнутых подмножеств $Y, Z \in X$ таких, что $Y \cap Z = \emptyset$ найдутся такие открытые окрестности $U \supseteq Y$ и $V \supseteq Z$, что $U \cap V = \emptyset$.*

Хаусдорфовы пространства, удовлетворяющие аксиоме T_4, называются **нормальными** (это условие было введено Л.Вьеторисом в 1921 году и изучено Г.Титце, П.С.Александровым и П.С.Урысоном в 1923–1924 годах). Сформулируем явно только что упомянутой ключевое свойство компактов[19,20].

Лемма. *Компактное хаусдорфово пространство нормально.*

Доказательство этого результата совсем просто и заинтересованный читатель может минут за 5 провести его в качестве упражнения. При этом в качестве промежуточного шага полезно показать вначале, что компактное хаусдорфово пространство **регулярно**, т.е. удовлетворяет следующей аксиоме отделимости T_3, введенной в 1921 году Л.Вьеторисом.

T_3. *Для любой точки $x \in X$ и любого замкнутого подмножества $Y \in X$ таких, что $x \notin Y$ найдутся такие открытые окрестности $x \in U$ и $V \supseteq Z$, что $U \cap V = \emptyset$.*

---

[19] Н.Бурбаки, Общая топология. Основные структуры. – Наука, М., 1968, с.1–272, Предложение 2 на стр.127.

[20] А.Н.Колмогоров, С.В.Фомин, Теорема 4 на стр.97.

Следующий теорема является одним из самых фундаментальных и часто используемых результатов всей классической общей топологии. По историческим причинам она обычно называется *леммой* Урысона[21,22].

Лемма Урысона. *Пусть X – нормальное пространство, $Y, Z \subseteq X$, $Y \cap Z = \emptyset$. Тогда существует функция $f \in C(X)$ такая, что $f(y) = 0$ для всех $y \in Y$ и $f(z) = 1$ для всех $z \in Z$.*

Ясно, что отсюда сразу вытекает такое

Следствие. *Пусть X – компактное хаусдорфово пространство. Тогда алгебра $C(X)$ разделяет точки, т.е. для любых $x, y \in X$, $x \neq y$ существует непрерывная функция $f \in C(X)$ такая, что $f(x) \neq f(y)$.*

Комментарий. В действительности, свойство разделения точек это почти в точности определение **тихоновских** пространств. Точнее, аксиома $T_{3\frac{1}{2}}$, требует существования непрерывной функции $f \in C(X)$ такой, что $f(x) = 0$ и $f(y) = 1$ для любого $y \in Y$, где $Y \subseteq X$ – замкнутое подмножество в X, не содержащее x. Однако доказательство того, что $T_{3\frac{1}{2}}$, промежуточна по силе между T_3 и T_4 в любом случае отнюдь не очевидно!

§ 6. Кольцо непрерывных функций на компакте

1. Кольцо $C(X)$ на компакте X.

гомоморфизмы $C(X) \longrightarrow \mathbb{C}$ и максимальные идеалы в $C(X)$. Два гомоморфизма в том и только том случае равны, когда равны их ядра.

Теорема. *Каждый гомоморфизм $\mathbb{C}$-алгебр $C(X) \longrightarrow \mathbb{C}$, есть гомоморфизм вида ev_x для единственного $x \in X$.*

Доказательство. Из леммы Урысона следует, что $C(X)$ разделяет точки компакта X, так что $\mathrm{ev}_x \neq \mathrm{ev}_y$ при $x \neq y$. С другой стороны, пусть ϕ – произвольный гомоморфизм $C(X) \longrightarrow \mathbb{C}$. Если $\phi \neq \mathrm{ev}_x$ ни для одной точки $x \in X$, то для каждого $x \in X$ найдется такая функция $f_x \in C(X)$, что $\phi(f_x) = 0$, в то время как $f_x(x) = \mathrm{ev}_x(f) \neq 0$. Так как f_x непрерывна, то найдется такая окрестность U_x точки x в которой f_x не обращается в 0. Выберем теперь из покрытия $X = \bigcup U_x$, $x \in X$, конечное подпокрытие. Пусть $X = U_1 \cup \ldots \cup U_n$, где $U_i = U_{x_i}$ – окрестности точек $x_1, \ldots, x_n$, а $f_i = f_{x_i}$ – соответствующие функции. Тогда функция $f = f_1\overline{f}_1 + \ldots + f_n\overline{f}_n \in C(X)$ не обращается в 0 ни в одной точке x и, таким образом, обратима в $C(X)$. С другой стороны, по самому определению функций f_x имеем $\phi(f_x) = 0$, так что $\phi(f) = 0$, противоречие.

$X \cong Y \Longleftrightarrow C(X) \cong C(Y)$.

§ 7. Пределы

• **Предел в точке.** Пусть X – компактное хаусдорфово пространство, $A(X) = A_{\mathbb{C}}(X)$ – кольцо комплекснозначных функций на X, имеющих предел в каждой точке. Напомним, что **предел** $\lim_x f$ функции f в точке $x \in X$ определяется следующим образом. Обозначим через X' множество неизолированных точек. Если $x \in X \setminus X'$ изолированная, то $\lim_x f = f(x)$. Если

[21] Н.Бурбаки, Общая топология. Использование вещественных чисел в общей топологии. Функциональные пространства. – Наука, М., 1975, с.1–408, Теорема 1 на стр.87.

[22] Р.Энгелькинг, Общая топология. – Мир, М., 1986, с.1–751, Теорема 1.5.10 на стр.75.

$x \in X'$ не изолированная, то $c \in \mathbb{C}$ называется **пределом** f в x, если для любой окрестности $U \subseteq \mathbb{C}$ существует **пунктированнная окрестность** (alias **окрестность с выколотой точкой**, punctured neighborhood) $V \subseteq X$ точки x такая, что $f(V) \subseteq U$. Традиционно предел в точке x обозначается $\lim_{y\to x} f(y)$, но связанные переменные здесь можно убрать.

Легко видеть, что предел в точке обладает следующими свойствами:

$$\lim_x(f+g) = \lim_x f + \lim_x g, \qquad \lim_x(fg) = \lim_x f \lim_x g,$$

для любых $f, g \in A(X)$ и, таким образом, определяет гомоморфизм колец $\lim_x : A(X) \longrightarrow \mathbb{C}$.

Легко видеть, что, кроме того, $\lim_x(cf) = c\lim_x f$ для любого $c \in \mathbb{C}$ и любой $f \in A(X)$, так что в действительности это гомоморфизм $\mathbb{C}$-алгебр. Несложная проверка[23] показывает, что $\lim_x$ это в точности **единственный** гомоморфизм $\mathbb{C}$-алгебр такой, что $\lim_x(\delta_y) = 0$ для всех $y \in X$. Предел можно охарактеризовать также как **единственный** гомоморфизм $\mathbb{C}$-алгебр такой, что $\lim_x(h) = 0$, если $h = 0$ в некоторой пунктированной окрестности V точки x.

• **Предел в бесконечности.** Пусть теперь X – локально компактное хаусдорфово пространство, например, $X = \mathbb{N}, \mathbb{R}$ или $\mathbb{C}$. Тогда можно рассмотреть одноточечную компактификацию $\overline{X} = X \cup \{\infty\}$, например, $\overline{\mathbb{R}}$ и $\overline{\mathbb{C}}$ – это, соответственно, вещественная и комплексная проективные прямые. С точки зрения $\overline{X}$ точка ∞ ничем не отличается от всех остальных точек, так что для функций $f \in A(\overline{X})$ определен предел $\lim_\infty f$. Обозначим через $A_\infty(X) = \{f|_X | f \in A(\overline{X})\}$ кольцо функций на X, имеющих предел в каждой точке $\overline{X}$. Единственными гомоморфизмами $\mathbb{C}$-алгебр $A_\infty(X)$ в $\mathbb{C}$ являются $\lim_x$ и ev_x для $x \in X$ и $\lim_\infty$.

Комментарий. В элементарном анализе существует прискорбная[24] традиция рассматривать другую компактификацию $\mathbb{R}$, а именно, расширенной вещественной прямой принято называть $\widetilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. При этом можно определить пределы $\lim_{-\infty} f$ и $\lim_{-\infty} f$. Функция f имеет предел в ∞ в нашем смысле, если оба эти предела существуют и равны, $\lim_{-\infty} f = \lim_{+\infty} f$.

• **Устранение разрывов.** Пусть X – компактное хаусдорфово пространство, $A(X)$ – кольцо функций на X, имеющих предел в каждой точке, и $C(X)$ – кольцо непрерывных функций. Тогда отображение $A(X) \longrightarrow C(X)$, $f \mapsto \widehat{f}$, где $\widehat{f}(x) = \lim_x f$, является гомоморфизмом, т.е. $\widehat{f+g} = \widehat{f}\widehat{g}$ и $\widehat{fg} = \widehat{f}\widehat{g}$. Этот гомоморфизм называется **устранением разрывов**.

§ 8. Кольцо дифференцируемых функций

Пусть X дифференцируемое или аналитическое многообразие. В этом случае можно говорить не только о непрерывных, но и о дифференцируемых функциях. Опять же начинающий может считать, что $X = \mathbb{R}, \mathbb{R}^n$ или $[0,1]$.

• **Кольцо дифференцируемых функций** $D[a,b]$. Хорошо известно, что сумма и произведение дифференцируемых функций дифференцируемы. Ясно, что $D(X) \leq C(X)$.

---

[23] J.D.Gray, An algebraic characterisation of limits. – Amer. Math. Monthly, 1975, vol.82, N.8, p.825–827.

[24] Используется как перевод английского deplorable: 'presense of wit, but deplorable absense of etiquette'.

• **Кольцо непрерывно дифференцируемых функций** $C^1[a,b]$. Функция называется непрерывно дифференцируемой, если у него существует производная и эта производная непрерывна. Ясно, что $C^1(X) \le D(X)$.

• **Кольцо r раз непрерывно дифференцируемых функций** $C^r(X)$. Функция называется r раз непрерывно дифференцируемой, если у нее существует r-я производная и она непрерывна. Ясно, что $C^r(X) \le C^{r-1}(X)$.

• **Кольцо гладких функций** $C^\infty(X)$. Гладкая (alias бесконечно дифференцируемая) функция – это функция, у которой существуют производные *всех* конечных порядков. Ясно, что $C^\infty(X) \le C^r(X)$ для всех r.

• **Кольцо гладких функций с компактным носителем** $C^\infty_{\rm c}(X)$. По определению $C^\infty_{\rm c}(X) = C^\infty(X) \cap C_{\rm c}(X)$.

• **Кольцо аналитических функций** $C^\omega(X)$. Функция называется аналитической, если в окрестности любой точки она раскладывается в ряд blablabla Ясно, что $C^\omega(X) \le C^\infty(X)$.

Комментарий. Кольца $C^\infty(X)$ и $C^\omega(X)$ кажутся очень близкими, но в действительности с алгебраической точки зрения между ними пролегает пропасть равная различию между дифференциальной и аналитической/алгебраической геометрией. Дело в том, что кольцо $C^\infty[a,b]$ является кольцом с делителями 0. Вот, например, знаменитый пример бесконечно дифференцируемой функции $f : \mathbb{R} \longrightarrow \mathbb{R}$ с носителем $[0,1]$. Положим

$$f(x) = e^{-1/x^2(1-x)^2}, \qquad 0 < x < 1,$$

и $f(x) = 0$ в противном случае[25]. Ясно, что носитель функции $g(x) = f(x-2)$ не пересекается с носителем функции f, так что $fg = 0$. Это явление играет основную роль в теории обобщенных функций. В то же время в кольце аналитических функций такое невозможно.

ДИФФЕРЕНЦИРУЕМАЯ замена переменной

§ 9. Формула Лейбница и гомоморфизм Тэйлора

1. Многочлен Тэйлора. Пусть $V = (a,b)$, $a, b \in \mathbb{R}$, $a < b$, – интервал в $\mathbb{R}$, $c \in (a,b)$ – какая-то точка этого интервала, а $C^k(V)$ – кольцо k раз дифференцируемых функций на V. Для любого целого $m \le k$ положим

$$T^m_c(f) = \sum_{i=0}^{m} \frac{1}{i!}\frac{d^i f}{dx^i}(\mathrm{c})(x-\mathrm{c})^i.$$

Многочлен $T^m_{\rm c}(f)$ называется **многочленом Тэйлора** m-го порядка функции f в c. Рассматривая вместо самой функции f композицию трансляции $x \mapsto x-c$ с этой функцией, мы можем (и будем) полагать, что $c = 0$ и писать просто $T^m(f)$ вместо $T^m_0(f)$. В элементарных учебниках анализа $T^m(f)$ рассматривается как элемент кольца многочленов $\mathbb{R}[x]$. Однако в действительности, инвариантный смысл имеет не многочлен $T^m(f)$ как таковой, а *его класс по модулю более высоких степеней* x. Иными словами, $T^m(f)$ нужно рассматривать не как элемент $\mathbb{R}[x]$, а как элемент кольца $\mathbb{R}[x]/(x^{m+1})$ усеченных многочленов.

Замечательное наблюдение Ферма (в случае $m = 1$) и Лейбница (в общем случае), которое лежит в основе дифференциального исчисления, состоит в том, что отображение $f \mapsto T^m(f)$ является **эпиморфизмом** кольца $C^k(V)$ на $\mathbb{R}[x]/(x^{m+1})$. В самом деле, $T^m(f+g) = T^m(f) + T^m(g)$ очевидно, а равенство

[25] Б.Гелбаум, Дж.Олмстед, ibid., стр.54.

$T^m(fg) = T^m(f)T^m(g)$ называется **формулой Лейбница**. В элементарных учебниках анализа формула Лейбница обычно записывается как

$$\frac{d^l(fg)}{dx^l} = \sum_{i+j=l} \binom{l}{i} \frac{d^i(f)}{dx^l} \frac{d^j(g)}{dx^l}.$$

Чтобы перейти отсюда к форме $T^m(fg) = T^m(f)T^m(g)$, достаточно разделить обе части этого равенства на $l!$ и взять значения в 0.

В частности, функции, у которых значения всех производных до m-го порядка включительно в точке 0 равны 0, образуют идеал кольца $C^k(V)$ и сопоставление $f \mapsto T^m(f)$ является в точности канонической проекцией кольца $C^k(V)$ в фактор-кольцо по этому идеалу.

2. Ряд Тэйлора. Изложенная в предыдущем пункте конструкция без изменений переносится на следующую ситуацию. Пусть $C^\infty(V)$ – кольцо бесконечно дифференцируемых функций. Сопоставим каждой функции $f \in C^\infty(V)$ формальный степенной ряд

$$T(f) = \sum_{i=0}^{\infty} \frac{1}{i!} \frac{d^i f}{dx^i}(0) x^i,$$

называемый **рядом Тэйлора** функции f. Снова в силу формулы Лейбница отображение $f \mapsto T(f)$ является гомоморфизмом $C^\infty(V)$ в кольцо $\mathbb{R}[[x]]$ формальных степенных рядов с вещественными коэффициентами. Этот гомоморфизм, называемый в дальнейшем **гомоморфизмом Тэйлора**, вообще говоря, не является мономорфизмом: как хорошо известно, существуют бесконечно дифференцируемые функции $f \neq 0$, у которых значения **всех** производных в точке 0 равны 0. Наиболее известным примером таких функций являются

$$f_n(x) = \begin{cases} x^n e^{-1/x^2} & \text{при } x \neq 0 \\ 0 & \text{при } x = 0 \end{cases}.$$

Еще эффектнее выглядят следующая незначительная модификация этого примера

$$g_n(x) = \begin{cases} x^n e^{-1/x^2} & \text{при } x < 0 \\ 0 & \text{при } x \geq 0 \end{cases}.$$

Функции f_n и g_n лежат в ядре гомоморфизма Тэйлора.

В высшей степени поучительна следующая теорема, показывающая насколько монструозным образованием является кольцо $C^\infty(V)$. Доказательство этой теоремы совсем несложно, см., например[26]

Теорема Бореля. *Гомоморфизм Тэйлора T является эпиморфизмом $C^\infty(V)$ на $\mathbb{R}[[x]]$.*

Таким образом, снова сопоставление бесконечно дифференцируемой функции ее ряда Тэйлора является канонической проекцией по модулю идеала, состоящего из функций, все производные которых в точке 0 обращаются в 0.

Заметим, что для кольца $C^\omega(V)$ аналитических функций ситуация **абсолютно** иная. Гомоморфизм Тэйлора $T : C^\omega(V) \longrightarrow \mathbb{R}[[x]]$ перестает быть

[26] R.Narasimhan, Анализ на действительных и комплексных многообразиях, М., Мир, 1971, 232с. Глава I, § 1.5

сюръективным (ряд Тэйлора аналитической функции обязан сходиться в некоторой окрестности 0), но становится инъективным. Это значит, что кольцо формальных степенных рядов **много больше** кольца аналитических функций, но **много меньше** кольца бесконечно дифференцируемых функций.

Разумеется, оба эти примера (включая теорему Бореля) обобщаются на функции нескольких вещественных переменных, или, более общо, на функции определенные в окрестности какой-то точки x гладкого многообразия M.

§ 8. Идеалы дифференцируемых функций, струи

27,28

§ 10. Кольца алгебраических, экспоненциальных и тригонометрических многочленов

В этом параграфе мы рассмотрим несколько важнейших колец функций $\mathbb{R} \longrightarrow \mathbb{R}$ и $\mathbb{C} \longrightarrow \mathbb{C}$, которые рассматриваются в элементарной математике, теории дифференциальных уравнений, гармоническом анализе, теории аппроксимации и т.д.

1. Кольцо алгебраических многочленов. В анализе и теории аппроксимации **алгебраическим многочленом** называется конечная линейная комбинация степенных функций $x \mapsto x^n$, с целым показателем $n \in \mathbb{N}_0$, т.е. функцию вида $\widetilde{f} : \mathbb{R} \longrightarrow \mathbb{R}$, $x \mapsto a_n x^n + \ldots + a_1 x + a_0$. В школьной математике такая функция называется просто многочленом, в алгебре – полиномиальной функцией, а в алгебраической геометрии – регулярной функцией или целой алгебраической функцией. Как мы убедимся в главе 4, из леммы Дедекинда-Артина сразу вытекает, что одночлены $e^{\lambda x}$ линейно независимы над $\mathbb{R}$ и, значит образуют базис пространства алгебраических многочленов. Поэтому кольцо вещественных алгебраических многочленов на самом деле изоморфно кольцу $\mathbb{R}[x]$, причем изоморфизм устанавливается самым естественным образом, а именно, изображенная выше функция $\widetilde{f}$ является образом многочлена $f = a_n x^n + \ldots + a_1 x + a_0$. Кольцо комплексных алгебраических многочленов определяется аналогично.

2. Кольцо экспоненциальных многочленов. Сделаем первую попытку определить кольцо экспоненциальных многочленов. Назовем вещественным **экспоненциальным многочленом** конечную линейную комбинацию экспонент $x \mapsto e^{\lambda_n x}$, $\lambda \in \mathbb{R}$, т.е. функцию вида $f : \mathbb{R} \longrightarrow \mathbb{R}$,

$$x \mapsto a_1 e^{\lambda_1 x} + \ldots + a_n e^{\lambda_n x}, \quad a_i, \lambda_i \in \mathbb{R}.$$

Как мы убедимся в главе 4, из деммы Дедекинда-Артина сразу вытекает, что одночлены $e^{\lambda x}$ линейно независимы над $\mathbb{R}$ и, таким образом, образуют базис пространства $\mathrm{Exp}_{\mathbb{R}}$. При этом $e^{\lambda x} e^{\mu x} = e^{(\lambda+\mu)x}$, так что $\mathrm{Exp}_{\mathbb{R}}$ кольцо с мультипликативным базисом.

В действительности, $\mathrm{Exp}_{\mathbb{R}}$ представляет собой групповую алгебру $\mathbb{R}[\mathbb{R}^+]$ аддитивной группы вещественных чисел, с вещественными же коэффициентами. Разумеется, здесь речь идет о том, что называют групповой алгеброй *алгебраисты*, т.е. групповой алгебре $\mathbb{R}^+$ как *абстрактной*, а не топологической группы! Мы могли бы определить алгебру $\mathrm{Exp}_{\mathbb{C}}$ функций $\mathbb{C} \longrightarrow \mathbb{C}$, являющихся

[27] Б.Мальгранж, Идеалы дифференцируемых функций. – Мир, М., 1965, с.1–131.

[28] Джет Неструев, Гладкие многообразия и наблюдаемые. – МЦНМО, М., 2000, с.1–299.

линейными комбинациями $z \mapsto e^{\lambda z}$, где $\lambda \in \mathbb{C}$, с комплексными же коэффициентами.

Задача. Какие классы функций реализуют групповые алгебры $\mathbb{R}[\mathbb{R}^*]$ и $\mathbb{C}[\mathbb{R}^+]$? Верно ли, что эти алгебры изоморфны?

3. Кольцо экспоненциальных многочленов, другим манером. Однако в действительности под экспоненциальными многочленами обычно понимают не $\mathrm{Exp}_{\mathbb{R}}$, а более широкое кольцо $\mathrm{Expo}_{\mathbb{R}}$, включающее кольцо алгебраичекских многочленов. Иными словами, $\mathrm{Expo}_{\mathbb{R}}$ состоит из конечных линейных комбинаций функций $x \mapsto x^m e^{\lambda x}$. Таким образом, каждый экспоненциальный многочлен имеет вид

$$x \mapsto a_1 x^{m_1} e^{\lambda_1 x} + \ldots + a_n x^{m_n} e^{\lambda_n x}, \quad a_i, \lambda_i \in \mathbb{R},\ m_i \in \mathbb{N}_0.$$

Как и выше, произведение двух базисных элементов снова является базисным элементом:

$$x^m e^{\lambda x} \cdot x^n e^{\mu x} = x^{m+n} e^{(\lambda+\mu)x}.$$

Это кольцо и его комплексный аналог $\mathrm{Expo}_{\mathbb{C}}$, в котором $a_i, \lambda_i \in \mathbb{C}$, играют центральную роль в теории линейных дифференциальных уравнений и ее многочисленных приложениях.

4. Кольцо тригонометрических многочленов. Основную роль в теории колебаний, в частности, в математическом анализе музыки играют вещественные тригонометрические многочлены, называемые еще многочленами Фурье. Вещественным тригонометрическим многочленом называется функция $\mathbb{R} \longrightarrow \mathbb{R}$, являющаяся конечной линейной комбинацией постоянной функции $x \mapsto 1$, и тригонометрических функций $x \mapsto \cos(nx)$ и $x \mapsto \sin(nx)$. Иными словами, любой вещественный тригонометрический многочлен имеет вид

$$x \mapsto a_0 + \sum_{k=1}^{n} (a_m \cos(mx) + b_m \sin(mx)), \quad a_m, b_m \in \mathbb{R}.$$

Обычно множество всех вещественных тригонометрических многочленов обозначается через $\mathrm{Trig}_{\mathbb{R}}$. Пользуясь формулами, выражающими произведения $\cos(x)\cos(y)$, $\cos(x)\sin(y)$, $\sin(x)\sin(y)$ через $\cos(x \pm y)$, $\sin(x \pm y)$, мы видим, что произведение двух элементов базиса пространства $\mathrm{Trig}_{\mathbb{R}}$ снова является линейной комбинацией двух базисных элементов. Таким образом, в действительности $\mathrm{Trig}_{\mathbb{R}}$ является коммутативным кольцом. Этот пример снова возникнет у нас в главе 'Арифметика коммутативных колец'. Допуская комплексные коэффициенты мы можем переписать тригонометрический многочлен в виде

$$z \mapsto \sum_{k=-n}^{n} c_k e^{ikz}, \quad c_m \in \mathbb{C}.$$

Комплексные тригонометрические многочлены образуют кольцо $\mathrm{Trig}_{\mathbb{C}}$.

§ 11. Кольца функций вещественного анализа. 1st installment: ограничения на рост

В настоящем и следующем параграфах мы рассматриваем некоторые классы функций, рассматриваемых в вещественном анализе, которые образуют кольцо относительно обычного сложения и умножения функций.

• **Кольцо функций ограниченной вариации.** Пусть $f : [a,b] \longrightarrow \mathbb{R}$. Определим полную вариацию $V_a^b(f)$ функции f как точную верхнюю грань сумм вида

$$|f(x_1) - f(x_0)| + \ldots + |f(x_n) - f(x_{n-1})|,$$

по всем $n \in \mathbb{N}$, и всем наборам точек $a = x_0 < x_1 < \ldots < x_n = b$. Если $V_a^b(f) < \infty$, то говорят, что f функция ограниченной вариации на отрезке $[a,b]$. Класс $V[a,b]$ функций ограниченной вариации на $[a,b]$ образует кольцо относительно обычного сложения и умножения функций, см., например[29]. Ясно, что $V[a,b] \leq B[a,b]$, иными словами, каждая функция ограниченной вариации ограничена.

Комментарий. Монотонные функции не образуют кольца: сумма или произведение двух возрастающих функций не обязано быть возрастающей функцией. Жордан показал, что функции ограниченной вариации это *в точности* наименьшее подкольцо в $\mathbb{R}^{[a,b]}$, содержащее возрастающие функции. А именно, любая функция $f \in V[a,b]$, допускает **разложение Жордана**: существуют возрастающие на $[a,b]$ функции f^+ и f^-, с закрепленным концом $f^+(0) = f^-(0) = 0$ такие, что $f = f(a) + f^+ - f^-$. Более того, можно выбрать функции f^+ и f^- так, чтобы $V_a^x(f) = f^+(x) + f^-(x)$, в этом случае f^+ и f^- естественно истолковать как функции, определяющие **положительную вариацию** и **отрицательную вариацию** функции f.

• **Кольцо абсолютно непрерывных функций.** Функция $f : [a,b] \longrightarrow \mathbb{R}$ называется **абсолютно непрерывной** на отрезке $[a,b]$, если для любого $\epsilon > 0$ найдется такое $\delta > 0$, что для любого $n \in \mathbb{N}$ и любых

$$a < a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n < b$$

таких, что

$$|b_1 - a_1| + \ldots + |b_n - a_n| < \delta$$

выполняется неравенство

$$|f(b_1) - f(a_1)| + \ldots + |f(b_n) - f(a_n)| < \epsilon.$$

(см., например[30]). Класс $AC[a,b]$ абсолютно непрерывных функций на $[a,b]$ образует кольцо. $AC[a,b] \subseteq V[a,b]$.

Комментарий. Условие абсолютной непрерывности является промежуточным между непрерывностью и липшицевостью: всякая абсолютно непрерывная функция непрерывна, а всякая липшицева функция абсолютно непрерывна (если в определении абсолютно непрерывной функции не требовать $b_i < a_{i+1}$, то как раз и получится условие Липшица). Знаменитый пример непрерывной, но не абсолютно непрерывной функции, это функция $f : [0,1] \longrightarrow \mathbb{R}$, определенная как $f(x) = x\sin(1/x)$ для $0 < x \leq 1$ и $f(0) = 0$. Абсолютно непрерывные функции более, чем естественно возникают в теории интеграла Лебега. Теорема Лебега утверждает, что класс абсолютно непрерывных функций – это в точности класс первообразных в смысле Лебега от суммируемых функций[31]. Это показывает, что по сути условие абсолютной непрерывности **значительно** ближе к дифференцируемости, чем к обычной непрерывности. В частности, абсолютно непрерывная функция имеет конечную производную почти в каждой точке $x \in [a,b]$ – в действительности это верно уже для функций ограниченной вариации[32].

• **Кольцо липшицевых функций.**

---

[29] У.Рудин, Основы математического анализа. 2-е изд., М., Мир, 1976, 319с. теорема 6.24.

[30] [Колмогоров–Фомин], с.335

[31] Колмогоров–Фомин, ibid., стр.337–338.

[32] Колмогоров–Фомин, ibid. стр.330.

§ 12. Кольца функций вещественного анализа. 2nd installment: измеримость и интегрируемость

• **Кольцо интегрируемых функций.** Класс $R[a,b]$ **интегрируемых по Риману** $[a,b]$ образует кольцо относительно обычного сложения и умножения функций, см., например[33]. $C[a,b] \leq R[a,b] \leq B[a,b]$.

• **Кольцо измеримых функций.** Функция $f : [a,b] \longrightarrow \mathbb{R}$ называется **измеримой**, если для любого $c \in \mathbb{R}$ множество $\{x \in [a,b] \mid f(x) \geq c\}$ измеримо. ПО ЛЕБЕГУ?? Класс $M[a,b]$ измеримых функций на $[a,b]$ образует кольцо относительно обычного сложения и умножения функций, см., например. [Rudin], теорема 10.18.

Комментарий. Измеримость функции не означает еще возможности определить для нее (конечный) интеграл. Множество $\mathcal{L}[a,b]$ **суммируемых** alias **интегрируемых** по Лебегу функций, т.е. таких функций, для которых $\int_a^b |f| d\mu < \infty$, кольца не образует, это векторное пространство, но не кольцо относительно умножения функций. Например, (неотрицательная) функция $f : [0,1] \longrightarrow \mathbb{R}$, определенная посредством $f(x) = 1/\sqrt{x}$ при $0 < x \leq 1$ и $f(0) = 0$ интегрируема по Лебегу, но f^2 – нет[34]. В действительности, это пространство является модулем над кольцом ограниченных измеримых функций: если $f \in B[a,b] \cap M[a,b]$, а $g \in \mathcal{L}[a,b]$, то $fg \in \mathcal{L}[a,b]$.

Задача. Функция называется **борелевской**, если для любого $c \in \mathbb{R}$ множество $\{x \in [a,b] \mid f(x) \geq c\}$ борелевское. Образуют ли борелевские функции кольцо?

§ 13. Кольца функций вещественного анализа. 3rd installment: Кольца функций кусочно $\mathcal{C}$

• **Кольцо ступенчатых функций** $\mathrm{Step}(\mathbb{R})$. В теории интеграла Римана рассматривается класс ступенчатых функций. Функция $f : \mathbb{R} \longrightarrow X$ называется **ступенчатой**, если существует *конечная* возрастающая последовательность точек $c_0, c_1, \ldots, c_n$ такая, что на каждом из *открытых* интервалов (∞, c_0), (c_0, c_1), $\ldots$, (c_{n-1}, c_n), (c_n, ∞), функция f постоянна (никаких предположений относительно значений f в $c_0, c_1, \ldots, c_n$ не делается).

Задача. Пусть f, g – ступенчатые функции, Δ_f, Δ_g – соответствующие разбиения. Докажите, что $f+g$, fg тоже ступенчатые. Что можно взять в качетсве разбиения для этих функций.

Задача. Докажите, что ступенчатые функции с компактным носителем образуют идеал в $\mathrm{Step}(\mathbb{R})$.

• **Кольцо конечнозначных функций.** Функция $f : X \longrightarrow \mathbb{R}$ называется функцией с конечным образом, если $|\operatorname{Im}(f)| < \infty$. Ясно, что функции с конечным образом образуют кольцо. В самом деле,

$$\operatorname{Im}(f+g) \leq \operatorname{Im}(f) + \operatorname{Im}(g), \qquad \operatorname{Im}(fg) \leq \operatorname{Im}(f)\operatorname{Im}(g)$$

Легко видеть, что это в точности подалгебра в $\mathbb{R}^X$, порожденная характеристическими функциями χ_Y подмножеств $Y \subseteq X$.

Замечание. Во многих книгах по анализу такие такие функции называются **простыми**, ... см., например[35]. В[36], простыми называются *измеримые* функции, принимающие не более,

[33] Rudin, ibid. теоремы 6.10, 6.12.

[34] Б.Гелбаум, Дж.Олмстед, ibid., стр.222.

[35] У.Рудин, ibid. стр.284

[36] Колмогоров–Фомин, ibid. стр.280.

чем счетное множество значений!!! Однако термин 'конечнозначные' представляется нам гораздо более точным и суггестивным.

• **Кольцо этажных функций.** В теории интеграла Лебега рассматривается класс этажных функций. Функция $f : \mathbb{R} \longrightarrow \mathbb{R}$ называется **этажной**, относительно меры μ, если для каждого $z \in \mathbb{R}$ множество $f^{-1}(z)$ измеримо. В частности, если $X = \mathbb{R}$ с мерой Лебега, то всякая ступенчатая функция является этажной, но обратное, конечно, неверно, как показывает пример функции Дирихле $\chi_{\mathbb{Q}}$.

?. Кусочно $\mathcal{C}$. Вообще, говорят, что функция $f : \mathbb{R} \longrightarrow \mathbb{R}$ кусочно $\mathcal{C}$, где $\mathcal{C}$ – некоторый класс функций, если существует конечная последовательность точек $c_0 < c_1 < \ldots < c_n$ такая, что ограничение f на каждый из открытых интервалов (∞, c_0), (c_0, c_1), $\ldots$, (c_{n-1}, c_n), (c_n, ∞), принадлежит классу $\mathcal{C}$. Например, можно говорить о

• кусочно непрерывных функциях. (Функции с конечным множеством разрывов)

• кусочно линейных функциях.

• кусочно полиномиальных функциях.

Задача. Какие из перечисленных классов функций образуют кольцо?

Задача. Образуют ли кольцо функции с конечным или счетным множеством разрывов?

• **Дискретизация.** В инженерной практике, а также финансовой и актуарной математике широко используется отображение **дискретизации**

$$\lceil f : C(\mathbb{R}) \longrightarrow \mathrm{Step}(\mathbb{R}), \qquad f \mapsto \lceil f(x) = f(\lfloor x \rfloor),$$

сопоставляющее непрерывной функции f ступенчатую функцию $\lceil f(x)$. Это гомоморфизм колец:

$$\lceil (f + g) = \lceil f + \lceil g, \qquad \lceil (fg) = \lceil f \cdot \lceil g.$$

• **Сплайны**. Непрерывная кусочно полиномиальная функция называется сплайном.

Комментарий. Иногда требуют непрерывности не только самого сплайна, но еще согласования производных до определенного порядка.

§ 14. Кольца функций вещественного анализа. 4th installment: периодические функции

• **Периодические функции.** Напомним, что функция $f : \mathbb{R} \longrightarrow \mathbb{R}$ называется **периодической** с периодом ω, если $T_\omega f = f$, иными словами, если $f(x + \omega) = f(x)$ для всех $x \in \mathbb{R}$.

Задача. Доказать, что периодические функции и непрерывные периодические функции с **данным** периодом ω образуют алгебру.

Задача. Построить пример, показывающий, что если f и g – периодические функции с несоизмеримыми периодами ω_1, ω_2, то $f+g$ и fg могут не быть периодическими, так что множество периодических функций не образует алгебры (или, хотя бы, векторного пространства).

Ответ. Ну, хотя бы функции $\sin(x)$ и $\sin(\alpha x)$, где $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ иррационально, см.[37].

Комментарий. Однако, это легко исправить. Сумма и произведение периодических функций всегда будут **почти периодическими** функциями в смысле Бора, см., например[38]. Если мы хотим варьировать период, то правильным объектом изучения являются не периодические, а именно почти периодические функции – **они** образуют алгебру.

§ 14. Кольца функций вещественного анализа. 5th installment: пространство производных

1. Пространство производных. Рассмотрим множество V функций, которые являются производными. Обратите внимание, не *имеют* производную, а *являются* производными ('f *is*, rather than *has* a derivative'[39]). Иными словами, $f \in V$ в том и только том случае, когда существует дифференцируемая функция g такая, что $g' = f$. Ясно, что множество V является векторным пространством, содержащим все непрерывные функции. Обратное, конечно, неверно: не каждая производная непрерывна. Тем не менее, каждая производная удовлетворяет **теореме Дарбу**: если она принимает значения $a < b$, то она принимает и все промежуточные значения. Эта теорема позволяет легко строить функции, *не являющиеся* производными.

Естественно возникает вопрос, является пространство производных кольцом, т.е. будет ли произведение двух производных снова производной? В 1921 году Вилкош[40] обнаружил, что это не так. А именно, он заметил, что уже квадрат функции

$$f(x) = \begin{cases} \cos\left(\frac{1}{x}\right), & x > 0, \\ 0, & x \leq 0. \end{cases}$$

не является производной.

Задача. Докажите, что $f(x)$ является производной.

Указание. Продифференцируйте

$$g(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Что можно сказать про функцию $g' - f$?

По определению **первый класс Бэра** B_1 состоит из пределов последовательностей непрерывных функций в смысле поточечной сходимости[41,42]. Легко видеть, что B_1 образует алгебру. Так как

В 1982 году Дэвид Прайс получил полное решение старого вопроса о том,

---

[37] Б.Гелбаум, Дж.Олмстед, Контрпримеры в анализе, М., Мир, 1967, 251с., стр.219–220.

[38] Maurin, Методы гильбертова пространства. –

[39] A.M.Bruckner, J.Mařik, C.E.Weil, Some aspects of products of derivatives. – Amer. Math. Monthly, 1992, February, p.134–145.

[40] W.Wilcosz, Some properties of derivative functions. – Fund. Math., 1921, vol.2, p.145–154.

[41] Р.Бэр, Теория разрывных функций. – М.–Л., 1932.

[42] Многие аналисты называют множество непрерывных функций **нулевым классом Бэра**, при этом **первым классом Бэра** называется множество *разрывных* функций из B_1. Таким образом, наш класс B_1 является объединением нулевого и первого классов Бэра. Мне, однако, кажется, что используемая нами терминология гораздо удобнее

Теорема Прайса. *Кольцо, порожденное производными, совпадает с* B_1.

В действительности, в работе[43] доказан следующий значительно более точный результат: каждая функция класса Бэра 1 представима в виде $fg+h$, где f, g, h – производные.

§ 15. Кольцо голоморфных функций

По поводу содержания этого параграфа[44,45,46,47,48,49,50,51]

1. Голоморфные функции на открытом множестве. Пусть U – открытое подмножество в $\mathbb{C}^n$. Функция $f : U \longrightarrow \mathbb{C}$ называется **голоморфной** на U, если для каждой точки $x = (x_1, \dots, x_n) \in U$ существует содержащаяся в U открытая окрестность V точки x и степенной ряд

$$\sum a_{i_1 \dots i_n} (z_1 - x_1)^{i_1} \dots (z_n - x_n)^{i_n}, \qquad i_1, \dots, i_n \in \mathbb{N}_0,$$

абсолютно сходящийся к $f(z)$ в каждой точке $z = (z_1, \dots, z_n) \in V$. Множество голоморфных на U функций обозначается через $\mathcal{O}_U$. Функция голоморфная на всем пространстве $\mathbb{C}^n$, называется **целой**.

Композиция голоморфных функций голоморфна. Мы не будем пытаться сформулировать этот принцип в самом общем виде, так как для всех наших целей достаточно следующего частного случая: если $f_1, \dots, f_m$ – голоморфные функции на открытом подмножестве $U \subseteq \mathbb{C}^n$, а $g : \mathbb{C}^m \longrightarrow \mathbb{C}$ – целая функция, то функция $x \mapsto g(f_1(x), \dots, f_n(x))$ голоморфна на U. (Это вытекает, например, из леммы Абеля и композиции нормально сходящихся степенных рядов – ССЫЛКА ИЛИ КАК??)

В частности, так как сложение и умножение являются целыми функциями $\mathbb{C}^2 \mapsto \mathbb{C}$, то для любых голоморфных функций $f, g : U \longrightarrow \mathbb{C}$ их сумма и произведение тоже голоморфны. Тем самым, $\mathcal{O}_U$ образует кольцо относительно обычных операций над функциями. Важнейшим особенностью голоморфных функций по сравнению с вещественными бесконечно дифференциируемыми функциями, является следующий замечательный результат, называемый еще **теоремой единственности** для голоморфных функций.

Принцип аналитического продолжения. *Пусть* $f, g : U \longrightarrow \mathbb{C}$ *– голоморфные функции на открытом связном подмножестве* $U \subseteq \mathbb{C}^n$. *Тогда если* $f|_V = g|_V$ *для некоторого непустого открытого подмножества* $V \subseteq U$, *то* $f = g$ *всюду на* U.

---

[43] D.Preiss, Algebra denerated by derivatives. – Real Anal. Exch., 1982–83, vol.8, p.208–216.

[44] С.Бохнер, У.Т.Мартин, Функции многих комплексных переменных. – ИЛ, 1951.

[45] Р.Ганнинг, Ч.Росси, Аналитические функции многих комплексных переменных. – Мир, М., 1969, с.1–395.

[46] Г.Грауэрт, Р.Реммерт, Аналитические локальные алгебры. – Наука, М., 1988, с.1–303.

[47] А.Картан, Элементарная теория аналитических функций одного и нескольких комплексных переменных. – ИИЛ, М., 1963, с.1–296.

[48] Л.Хермандер, Введение в теорию функции нескольких комплексных переменных. – Мир, М., 1968, с.1–279.

[49] Е.М.Чирка, Комплексные аналитические множества. – Наука, М., 1985, с.1–272.

[50] Б.В.Шабат, Введение в комплексный анализ. – Наука, М., 1969, с.1–576.

[51] М.Эрве, Функции многих комплексных переменных. Локальная теория. – Мир, М., 1965, с.1–165.

Этот результат объясняет, в частности, известный всем специалистам факт, что комплексный анализ по духу гораздо ближе к алгебре, чем к вещественному анализу. Вот простая иллюстрация этого принципа.

Теорема. *Для любой открытой связной области $U \subseteq \mathbb{C}^n$ кольцо $\mathcal{O}_U$ является областью целостности.*

Доказательство. Пусть $fg = 0$ для каких-то $f, g \in \mathcal{O}_U$. Если $f \neq 0$, то множество V точек $z \in U$, в которых $f(z) \neq 0$ открыто и непусто. Поскольку $g(z) = 0$ для всех $z \in V$, то по принципу аналитического продолжения $g(z) = 0$ для всех $z \in U$, но это и значит, что $g = 0 \in \mathcal{O}_U$.

2. Ростки голоморфных функций. Пусть теперь X – любое подмножество в $\mathbb{C}^n$. Рассмотрим множество пар (U, f), где U, $X \subseteq U \subseteq \mathbb{C}^n$ – открытая окрестность X, а f – голоморфная функция на U. Скажем, что $(U, f) \sim (V, g)$, если существует такая окрестность W, $X \subseteq W \subseteq U \cap V$, множества X, что $f|_W = g|_W$. Класс этой эквивалентности называется **ростком голоморфных функций** на X.

Легко проверить (проделайте это!), что отношение $\sim$ согласовано со сложением и умножением функций, так что операции $(U, f) + (V, g) = (U \cap V, f + g)$, $(U, f)(V, g) = (U \cap V, fg)$, превращают множество ростков в коммутативное ассоциативное кольцо с 1, называемое **кольцом ростков голоморфных функций** на X. Это кольцо будет по прежнему обозначаться через $\mathcal{O}_X$. В случае, когда множество X само открыто, у каждого ростка есть единственный представитель вида (X, f), т.е. голоморфная функция на X, так что никакого конфликта с предшествующим не возникает!

Особенно важен случай $X = \{x\}$, кольцо $\mathcal{O}_x$ называется **кольцом ростков голоморфных функций в точке** x. Это кольцо называется еще **локальным кольцом точки** $x \in \mathbb{C}^n$. Оно изоморфно кольцу $\mathbb{C}\{x_1, \ldots, x_n\}$ сходящихся степенных рядов. Заметим, что если $(U, f) \sim (V, g)$, то $f(x) = g(x)$, так что имеет смысл говорить о значении ростка из $\mathcal{O}_x$ в точке x

Задача. Докажите, что $\mathcal{O}_x^* = \{f \in \mathcal{O}_x \mid f(x) \neq 0\}$.

Решение. Голоморфная функция непрерывна.

Задача. Докажите, что $\mathfrak{m} = \{f \in \mathcal{O}_x \mid f(x) = 0\}$ максимальный идеал в $\mathcal{O}_x$.

Решение. Очевидно, что $\mathfrak{m}$ идеал, его максимальность вытекает из изоморфизма $\mathcal{O}_x/\mathfrak{m} \equiv \mathbb{C}$.

Таким образом, все необратимые элементы кольца $\mathcal{O}_x$ образуют идеал, который в этом случае является *единственным* максимальным идеалом. Коммутативные кольца с этим свойством называются **локальными кольцами**.

Глава ?: СВЕРТКА

Эта глава посвящена детальному обсуждению **важнейшей** алгебраической операции: свертки. Точнее, мы строим и изучаем несколько типов колец функций со свертакой в качестве умножения. Основными примерами свертки, которые постоянно используются в дальнейшем, служат:

• Умножение многочленов и степенных рядов (свертка Абеля);

• Умножение матриц;

• Свертка Дирихле в алгебре арифметических функций (умножение рядов Дирихле);

• Умножение в полугрупповой/групповой алгебре;

• Свертка в алгебре $L^1(G)$ функций интегрируемых по Лебегу (в частности, свертка последовательностей и функций на $\mathbb{R}$).

Свертка является трудоемкой и дорогой операцией и значительная часть усилий математиков на протяжении последних двух веков состояла в разработке методов сводить вычисление свертки к умножению функций. Такое сведение называется **гармоническим анализом**, а любой явный изоморфизм между алгеброй функций со сверткой и алгеброй функций с умножением – **преобразованием Фурье**. Преобразования Фурье известны в десятках вариантов: ряды Фурье, интегралы Фурье, дискретное преобразование Фурье (DFT), преобразование Лапласа, ...

§ 6. Изменение порядка суммирования

Доказательство ассоциативности свертки связано с важным приемом изменения порядка суммирования, на котором стоит остановиться подробнее.

1. Обозначения, связанные с суммами. В алгебре сумма $\sum_{i\in I} a_i$ определена если в ней лишь конечное число ненулевых слагаемых, т.е. если $|\{i \in I \mid a_i \neq 0\}| < \infty$. Это всегда так, например, если уже само множество I конечно. В этом случае часто используется **натуральная индексниця**, т.е. множество I отождествляется с начальным отрезком $\underline{n}$, $n = |I|$, натурального ряда я сумма записывается в виде $\sum_{i=1}^{n} a_i$. Однако во вмногих случаях даже для конечных множеств удобнее использовать в качестве индексов не натуральные числа, а элементы других множеств, скажем, брать суммы по подмножествам $\underline{n}$, перестановкам, элементам конечной группы или поля.

Комментарий. В математическом анализе в различных ситуациях придается смысл *бесконечным* суммам: как *счетным* суммам вида $\sum_{i=1}^{\infty} a_i$, называемым там **рядами**, так и **интегралам** $\int_a^b f(x)dx$, которые естественнее всего рассматривать как *континуальные* суммы инфинитезимальных слагаемых (собственно, именно так они и определяются в нестандартном анализе). Все наши конструкции обобщаются на эти случаи и представляют в аналитическом контексте не меньший интерес, чем в алгебраическом. Однако при этом возникают различные, иногда весьма небанальные, **вопросы сходимости** – при каких условиях и *в каком смысле* существует сумма бесконечного ряда или значение интеграла? Так как эти вопросы не относятся собственно к алгебре, *в основном тексте* мы ограничиваемся только такими классами функций, для которых все возникающие суммы конечны. Для того, кто овладел необходимой алгебраической техникой на примере конечных сумм и знаком с основами анализа, перевод на язык интегралов в большинстве случаев **не представляет никакого труда**. Во многих случаях и собственно в алгебре (скажем, в кольцах формальных степенных рядов или бесконечных матриц над **произвольным** кольцом) определяются суммируемые последовательности или сходящиеся произведения. Все эти конструкции, даже

если они маскируются под чисто алгебраические, в действительности основаны на скрытом использовании топологии.

2. Обобщенная дистрибутивность. Пусть вначале R – произвольное кольцо (не обязательно коммутативное или даже ассоциативное). Рассмотрим произвольные элементы $a_1, \dots a_m, b_1, \dots b_n \in R$. Рассмотрим произведение

$$(a_1 + \ldots + a_m)(b_1 + \ldots + b_m) = \sum_{i=1}^{m} a_i \sum_{j=1}^{n} b_j = \sum_{i \in \underline{m}} a_i \sum_{j \in \underline{n}} b_j$$

и преобразуем его двумя различными способами. Вначале раскрывая *левую* сумму, и только потом правую, получаем

$$\sum_{i \in \underline{m}} a_i \sum_{j \in \underline{n}} b_j = \sum_{i \in \underline{m}} \big(a_i \sum_{j \in \underline{n}} b_j\big) = \sum_{i \in \underline{m}} \sum_{j \in \underline{n}} a_i b_j$$

С другой стороны, вначале раскрывая *правую* сумму и только потом левую, получаем

$$\sum_{i \in \underline{m}} a_i \sum_{j \in \underline{n}} b_j = \sum_{j \in \underline{n}} \big(a_i \sum_{i \in \underline{m}}\big) b_j = \sum_{j \in \underline{n}} \sum_{i \in \underline{m}} a_i b_j$$

Сравнивая два эти выражения, и полагая $I = \underline{m}$, $J = \underline{N}$, мы видим, что

$$\sum_{i \in I} \sum_{j \in J} a_i b_j = \sum_{j \in J} \sum_{i \in I} a_i b_j.$$

3. Изменение порядка суммирования. Предположим, что мы хотим сложить элементы a_{ij}, зависящие от двух индексов i и j, где, скажем, $1 \le i \le m$, а $1 \le j \le n$. Есть два очевидных способа сделать это. Один из них состоит в том, чтобы сначала просуммировать все элементы с фиксированным i, т.е. образовать суммы $b_i = \sum a_{ij}$, $1 \le j \le n$, а уже потом образовать сумму $\sum b_i$, $1 \le i \le m$. При втором же способе вначале суммируются все элементы с фиксированным j, т.е. образуются суммы $c_j = \sum a_{ij}$, $1 \le i \le m$, а потом образуется сумма $\sum c_j$, $1 \le j \le n$. В результате обеих процедур должна получаться одна и та же сумма $\sum a_{ij}$, $1 \le i \le m$, $1 \le j \le n$, иными словами,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij},$$

где i и j меняются в тех же пределах, что и выше. Именно эта формула и называется ‘изменением порядка суммирования’.

4. Момент истины (test of truth). С изменением порядка суммирования приходится часто встречаться на практике. Например, при составлении плана учебной нагрузки кафедры на год учебные часы обычно распределяются неким образом по месяцам и между преподавателями. Подсчитать общее количество часов можно двумя способами. Первый из них состоит в том, чтобы вначале найти полную нагрузку каждого преподавателя за весь учебный год, а потом сложить нагрузку всех преподавателей. При втором же способе вначале находится нагрузка всех преподавателей кафедры в каждом месяце, а потом складывается суммарная нагрузка кафедры по месяцам. Согласно только что

сказанному, оба способа должны приводить к одному и тому же ответу. Если в действительности этого обычно не происходит, то не потому, что прием изменения порядка суммирования неправильный, а потому, что, *как все знают из личного опыта*, в выполняемых вручную арифметических действиях неизбежно возникают ошибки.

5. Примеры изменения порядка суммирования. Вот несколько специальных случаев и вариантов изменения порядка суммирования.

• **Общие индексные множества.** При изменении порядка суммирования не обязательно пользоваться натуральной индексацией:

$$\sum_{i\in I}\sum_{j\in J} c_{ij} = \sum_{j\in J}\sum_{i\in I} c_{ij}.$$

В действительности обе суммы равны $\displaystyle\sum_{(i,j)\in I\times J} c_{ij}$.

• **Изменения порядка умножения.** Если операция в X записывается мультипликативно, то правило изменения порядка суммирования принимает вид

$$\prod_{i\in I}\prod_{j\in J} c_{ij} = \prod_{j\in J}\prod_{i\in I} c_{ij}.$$

Разумеется, для справедливости этой формулы необходимо, чтобы операция в X была ассоциативной и коммутативной.

• **Правило подсчета двумя способами.** Если R – отношение между элементами множеств X и Y, т.е. подмножество их декартова произведения $X \times Y$, то

$$\sum_{x\in X} |\{y \in Y \mid (x,y) \in R\}| = \sum_{y\in Y} |\{x \in X \mid (x,y) \in R\}|.$$

В самом деле, обе суммы равны $|R|$. Мы будем самым существенным образом использовать это правило в нескольких главах, посвященных теории групп.

• **Разбиение суммы.** Следующую формулу можно рассматривать как специальный случай изменения порядка суммирования, когда $|J| = 2$:

$$\sum_{i\in I}(a_i + b_i) = \sum_{i\in I} a_i + \sum_{i\in I} b_i.$$

Аналогичное утверждение имеет место для разбиения на три или большее количество слагаемых.

• **Изменение пределов суммирования.** Изменение порядка суммирования можно применять и в тех случаях, когда суммирование производится не по всем индексам $1 \le i \le m$, $1 \le j \le n$, а только по части пар (i,j), например, только по тем парам, для которых $j \le i$. Однако при этом нужно следить за тем, как изменяются пределы суммирования:

$$\sum_{i=0}^{n}\sum_{j=0}^{i} c_{ij} = \sum_{j=0}^{n}\sum_{i=j}^{n} c_{ij}.$$

Изменение пределов интегрирования всегда вызывает трудности у студентов при вычислении кратных интегралов. Это потому, что им перед этим никто не пытался объяснить, что то же самое бывает и у конечных сумм!

§ 2. Свертка, моноидная алгебра

Трудно переоценить важность операции свертки во многих областях математики.

Уиллиам Феллер[52]

В этом параграфе мы построим **важнейший** пример алгебраической операции – свертку.

1. Определение свертки. Пусть M – полугруппа с операцией $\circ$, а R – кольцо с операциями $+$ и $\times$. В обычном умножении функций из $R^M = \mathrm{Map}(M, R)$ задействовано только умножение элементов кольца R, а операция в M игнорируется. Сейчас мы введем на множестве функций R^M и/или на его подмножестве $R[M]$, состоящем из функций с конечным носителем, новую операцию, гораздо более хитрую и дорогую, чем обычное умножение функций, в определении которой будут участвовать *все три операции* $\circ$, $+$ и $\times$. В большинстве примеров роль $\circ$ в свою очередь будет играть сложение или умножение, но пока мы будем тщательно различать все три операции на уровне обозначений, чтобы четко отследить роль каждой их них.

Итак, для двух функций $f, g \in R^M$ мы хотим определить третью функцию $f * g \in R^M$, для этого достаточно задать значения $(f * g)(x) \in R$ этой функции на всех $x \in M$. Вначале мы напишем *формулу*, определяющую свертку, а потом опишем две важнейших ситуации, когда эта формула *имеет смысл*. Разумеется, следующее 'определение' само по себе, без каких-то дополнительных предположений, *ничего не определяет*.

'Определение'. Сверткой *функций f и g называется функция $f*g$ значение которой в $x \in M$ задается формулой*

$$(f * g)(x) = \sum_{y \circ z = x} f(y)g(z).$$

Разумеется, чтобы это определение имело смысл с точки зрения алгебры, все суммы в правой части должны быть *конечными*. Это безусловно так, если полугруппа M конечна. Однако это условие будет для нас слишком ограничительным, так как оно не позволяет определить ни формальные степенные ряды, ни даже многочлены. Сейчас мы сформулируем два других, гораздо более интересных условия.

Комментарий. В математическом анализе свертка определяется при помощи *бесконечных* сумм или интегралов. При этом, как всегда, возникают **вопросы сходимости**, т.е. для функций каких классов определены эти суммы и интегралы? Кроме того, если мы хотим, как и в алгебраическом случае, использовать свертку для определения структуры кольца на некоторых множествах функций, свертка двух функций некоторого класса должна быть не только определена, но и снова принадлежать тому же классу. Многие интегральные преобразования являются свертками с функциями специального вида ('ядрами'), однако в анализе XIX века ядро и преобразуемая функция не рассматривались как равноправные операнды. По мнению Н.Бурбаки, первым упоминанием об алгебраических свойствах свертки функций была статья Чебышева "О двух теоремах теории вероятностей", где доказано, что функцией распределения суммы независимых случайных величин является свертка их функций распределения.

[52]У.Феллер, Введение в теорию вероятностей и ее приложения, т.2, М., 1967.

2. Условия, гарантирующие существование свертки. Для того, чтобы формула, которой мы определили свертку, имела смысл, мы должны наложить ограничение либо на полугруппу M, либо на класс рассматриваемых функций.

• **Условие на полугруппу.** Уравнение $y \circ z = x$ имеет в полугруппе M **конечное число решений**. Иными словами, для каждого $x \in M$ существует конечное число пар $(y, z) \in M \times M$ таких, что $y \circ z = x$.

Назовем **носителем** функции $f \in R^M$ множество тех $x \in M$, для которых $f(x) \neq 0$:

$$\mathrm{Supp}(f) = \{x \in M \mid f(x) \neq 0\}.$$

Это *алгебраическое* определение носителя, в анализе носителем обычно называют *замыкание* того, что называем носителем мы. Впрочем, если M рассматривается с *дискретной* топологией, эти определения совпадают.

• **Условие на функции.** Рассматриваются только функции с **конечным носителем**:

$$R[M] = \{f \in R^M \mid |\mathrm{Supp}(f)| < \infty\}.$$

Иными словами, рассматриваются только функции, такие, что $f(x) = 0$ для **почти всех** x, где выражение 'почти всех' снова понимается в алгебраическом смысле, как **всех, кроме конечного числа**. Таким функции называются еще **формальными линейными комбинациями** элементов M с коэффициентами из R. Если M конечна, то $R[M] = R^M$.

Предостережение. Бурбаки использует для множества формальных линейных комбинаций обозначение $R^{(M)}$, а не $R[M]$!

Ясно, что каждое из этих условий гарантирует конечность всех сумм в определении свертки. Кроме того, если $f, g \in R[M]$, то

$$\mathrm{Supp}(f + g) \subseteq \mathrm{Supp}(f) \cup \mathrm{Supp}(g), \qquad \mathrm{Supp}(f * g) \subseteq \mathrm{Supp}(f) \circ \mathrm{Supp}(g),$$

так что $f + g, f * g \in R[M]$. В двух следующих пунктах мы докажем, что если R – ассоциативное кольцо с 1, то R^M и $R[M]$ тоже являются ассоциативными кольцами с 1.

Комментарий. Эти условия – это в точности условия, фигурирующие книгах Бурбаки 'Алгебра'[53] и 'Интегрирование'[54]. С содержательной точки зрения и в том и в другом случае речь идет о **компактности**. Тополог выразил бы наше условие на полугруппу M сказав, что операция в ней задает в дискретной топологии **совершенное** отображение $M \times M \longrightarrow M$, т.е. такое отображение, что прообраз компактного множества компактен[55]. Наиболее простой и важный класс функций, непосредственно обобщающий функции с конечным носителем, для которого **все** вопросы сходимости тривиальны, это **непрерывные функции с компактным носителем**. Свертка двух таких функций всегда существует и снова является непрерывной функцией с компактным носителем. Впрочем, этот пример тоже скорее относится к топологической алгебре, чем собственно к анализу.

• **Смешанные условия.** классы бесконечных матриц

3. Свертки на группе. В случае, когда $M = G$ является группой, для каждых x и y уравнение $yz = x$ имеет *единственное* решение $z = y^{-1}x$. Точно так же, если фиксированы x и z, то с необходимостью $y = xz^{-1}$. Поэтому

[53]Н.Бурбаки, Алгебра I, гл. II; пункты 9,10 стр. 294–298.

[54]Н.Бурбаки Интегрирование, Гл. VI–VIII, М., 1970; примеры 1 и 2 на стр. 222–223, – разумеется, в этой книге изложение ведется в терминах *мер*, а не функций

[55]Н.Бурбаки, Общая топология, Гл. I – II, М., 1968; см. Глава I, § 10

суммирование в формуле, выражающей свертку, достаточно вести не по двум неизвестным y или z а лишь по одной из них. Тем самым, эта формула приобретает следующий вид, более привычный для тех, кто сталкивался с ней в анализе или теории вероятностей:

$$(f * g)(x) = \sum_{y \in G} f(y)g(y^{-1}x) = \sum_{z \in G} f(xz^{-1})g(z).$$

Если G абелева и операция в ней записывается аддитивно, эти формулы переписываются в виде

$$(f * g)(x) = \sum_{y \in G} f(y)g(x - y) = \sum_{z \in G} f(x - z)g(z).$$

§ 3. Расширенная полугрупповая алгебра

Сейчас мы проверим, что при выполнении условия 2.1 множество функций R^M образует кольцо относительно сложения функций и свертки как умножения. Основной момент здесь проверка ассоциативности свертки, но с учетом ассоциативности R и M это в точности изменение порядка суммирования, которое мы рассматривали в предыдущем параграфе.

Теорема. *Пусть M – полугруппа такая, что для каждого $x \in M$ уравнение $y \circ z = x$ имеет лишь конечное число решений. Тогда*

1) *R^M – кольцо относительно операций $+$, $*$;*
2) *Если R ассоциативно, то R^M тоже ассоциативно;*
3) *Если M и R коммутативны, то R^M коммутативно;*
4) *Если M – моноид, а R – кольцо с 1, то R^M – кольцо с 1.*

Доказательство. 1) Свойства R^M по сложению нам известны, поэтому остается лишь проверить дистрибутивность свертки относительно сложения. В самом деле, пусть $f, g, h \in R^M$. Проверим, например, левую дистрибутивность. Вычислим значение функции $(f + g) * h$ в элементе $x \in M$:

$$\begin{aligned}((f+g)*h)(x) = \sum_{y\circ z=x} (f+g)(y)h(z) = \sum_{y\circ z=x} (f(y)+g(y))h(z) = \\ \sum_{y\circ z=x} f(y)h(z) + g(y)h(z) = \sum_{y\circ z=x} f(y)h(z) + \sum_{y\circ z=x} g(y)h(z) = \\ (f*h)(x) + (g*h)(x).\end{aligned}$$

Ясно, что правая дистрибутивность проверяется аналогично.

2) Снова достаточно сравнить значения функций $(f * g) * h$ и $f * (g * h)$ в каждой точке $x \in M$.

$$\begin{aligned}((f*g)*h)(x) = \sum_{y\circ z=x} (f*g)(y)h(z) = \\ \sum_{y\circ z=x} \left(\sum_{u\circ v=y} f(u)g(v) \right) h(z) = \sum_{u\circ v\circ z=x} f(u)g(v)h(z).\end{aligned}$$

Здесь мы воспользовались ассоциативностью умножения в R, ассоциативностью $\circ$ и дистрибутивностью умножения относительно сложения (в виде изменения порядка суммирования). Так как в получившееся выражение f, g и h входят симметрично, уже ясно, что ассоциативность имеет место. Для полноты доведем вычисление до конца. Полагая $v \circ z = w$, получим

$$\sum_{u\circ v\circ z=x} f(u)g(v)h(z) = \sum_{u\circ w=x} f(u)\left(\sum_{v\circ z=w} g(v))h(z)\right) = \sum_{u\circ w=x} (f)(u)(g*h)(w) = (f*(g*h))(x),$$

что и утверждалось.

3) Обычное вычисление, использующее коммутативность M и R, показывает

$$(f*g)(x) = \sum_{y\circ z=x} f(y)g(z) = \sum_{z\circ y=x} g(z)f(y) = (g*f)(x).$$

4) В качестве 1 кольца R^M выступает δ-функция δ_e, принимающая значение 1 в $x = e$ и 0 во всех $x \neq e$. В самом деле, проверим, например, что δ_e является левой единицей:

$$(\delta_e * f)(x) = \sum_{y\circ z=x} \delta_e(y)f(z) = f(x).$$

Проверка с другой стороны осуществляется совершенно аналогично.

В случае, когда R является коммутативным ассоциативным кольцом с 1, построенное нами в этой теореме кольцо называется **расширенной полугрупповой алгеброй** полугруппы M над кольцом R. Заметим, что кольцо R^M не обязательно коммутативно, даже если R коммутативно, для этого полугруппа M **тоже** должна быть коммутативной. Важнейшими примерами расширенных полугрупповых алгебр являются кольца формальных степенных рядов от одной и нескольких переменных.

5. Кольцо Дирихле и кольцо Абеля. Рассмотрим два первых примера расширенных полугрупповых алгебр.

• **Кольцо Дирихле.** В этом случае $M = (\mathbb{N}, \cdot)$ – моноид натуральных чисел по умножению, а свертка в $\mathbb{C}^{\mathbb{N}}$ – это обычная свертка Дирихле

$$(f*g)(n) = \sum_{lm=n} f(l)g(m).$$

• **Кольцо Абеля.** В этом случае $M = (\mathbb{N}_0, +)$ – моноид неотрицательных целых чисел по сложению, а свертка в $\mathbb{C}^{\mathbb{N}_0}$ – это **свертка Абеля**

$$(f*g)(n) = \sum_{l+m=n} f(l)g(m).$$

§ 4. Полугрупповая алгебра

Для функций с конечным носителем, т.е. при выполнении условия 3.2, аналог теоремы 4 справедлив уже для совершенно произвольной полугруппы.

Теорема. 1) $R[M]$ – *кольцо относительно операций* $+$, $*$;

2) *Если* R *ассоциативно, то* $R[M]$ *тоже ассоциативно;*

3) *Если* M *и* R *коммутативны, то* $R[M]$ *коммутативно;*

4) *Если* M – *моноид, а* R – *кольцо с 1, то* $R[M]$ – *кольцо с 1.*

Доказательство. Все пункты этой теоремы проверяются теми же вычислениями, что в теореме пункта 4, в которых использовалась лишь конечность всех возникающих сумм.

В случае коммутативного ассоциативного кольца с 1 получающееся так кольцо обычно называется **полугрупповой алгеброй** полугруппы M над кольцом R. Важнейшими примерами полугрупповых алгебр являются кольца многочленов и многочленов Лорана от одной и нескольких переменных.

В терминах Главы 3 полугрупповая алгебра допускает следующее конкретное описание. Сопоставим каждому элементу $y \in M$ соответствующую δ-функцию δ_y:

$$\delta_y(x) = \begin{cases} 1, & \text{если } x = y, \\ 0, & \text{если } x \neq y. \end{cases}$$

По определению $\mathrm{Supp}(\delta_x) = \{x\}$ так что $\delta_x \in R[M]$. Легко видеть, что δ-функции δ_x, $\in M$, образуют базис $R[M]$ как свободного R-модуля. Иными словами, каждая функция $f \in R[M]$ допускает представление в виде линейной комбинации $f = \sum f(x)\delta_x$, причем такое представление единственно. Фигурирующая здесь бесконечная сумма имеет смысл так как в действительности лишь конечное число слагаемых в ней отлично от 0 и она сводится к *конечной* сумме $f = \sum f(x)\delta_x$, $x \in \mathrm{Supp}(f)$. Ясно, что δ-функции перемножаются точно так же, как соответствующие элементы M, а именно, $\delta_x * \delta_y = \delta_{x\circ y}$. Заметим, что в классических книгах δ-функция δ_x обычно обозначается e_x, от немецкого Einheit. Свертка произвольных функций с конечным носителем есть не что иное, как продолжение этого умножения базисных функций *по линейности*.

Функториальность. Пусть $\phi : A \longrightarrow B$ – гомоморфизм колец, а $\psi : M \longrightarrow N$ – гомоморфизм моноидов. Тогда $A[M] \longrightarrow B[N]$, $\sum a_g e_g \mapsto \sum \phi(a_g) e_{\psi(g)}$, определяет гомоморфизм колец. В частности, при $\psi = \mathrm{id}$ мы получаем гомоморфизм $A[M] \longrightarrow B[M]$, называемый **заменой скаляров**, а при $\phi = \mathrm{id}$ – гомоморфизм $A[M] \longrightarrow A[N]$. Вот несколько простых иллюстраций функториальности.

Задача. Докажите, что

(1) $R[M \times N] \cong (R[M])[N]$.

(2) $R[M]^o \cong R^o[M^o]$.

(3) Если $I \trianglelefteq R$, то $R[M]/IR[M] \cong (R/I)[M]$.

(4) Если S – мультипликативная система в R, то $S^{-1}(R[M]) \cong (S^{-1}R)[M]$.

Групповые алгебры. ПАРА СЛОВ!!

Задача. Докажите, что если $R \neq 0$, а группа G содержит нетривиальный элемент конечного порядка, то в $R[G]$ есть делители 0.

Решение. Пусть $g \neq e$ элемент порядка n. Тогда $(e+g+\ldots+g^{n-1})(e-g) = 0 \in R[G]$.

Следующая задача опирается на структурную теорему для конечно порожденных абелевых групп.

Задача. Если R – кольцо без делителей 0, а G – абелева группа без кручения, то $R[G]$ – кольцо без делителей 0.

Решение. Пусть $x, y \in R[G]$. Тогда объединение носителей x и y конечно (вот это и есть ключевая идея!) Это значит, что x и y лежат уже в кольце $R[H]$, где $H \leq G$ – **конечно порожденная** абелева группа без кручения. По структурной теореме H свободна и, значит, $R[H] \cong R[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ есть кольцо многочленов Лорана.

Задача. Пусть R – коммутативное кольцо и G – конечная группа с классами сопряженных элементов $C_1, \ldots, C_n$. Докажите, что $\mathrm{Cent}(R[G])$ свободно порождается элементами $e_C = \sum e_g$, $g \in C$, где $C = C_1, \ldots, C_n$.

Решение. Прежде всего заметим, что для любого $h \in G$ и любого классо сопряженных элементов $hC = Ch$ (почему?) Тем самым, e_C действительно лежит в центре $R[G]$. Обратно, если $x = \sum a_g e_g$ лежит в центре $R[G]$, то для любого $h \in G$ имеем

$$x = e_h x e_{h^{-1}} = \sum_{g \in G} a_g e_{hgh^{-1}} = \sum_{g \in G} a_{h^{-1}gh} e_g.$$

Тем самым, $a_{h^{-1}gh} = a_g$ для всех $g, h \in G$, так что x действительно является линейной комбинацией элементов e_C. Осталось заметить, что так как носители e_C попарно дизъюнктны, то любая линейная зависимость между e_C дает линейную зависимость между e_g, с теми же коэффициентами.

§ 5. Сжатая полугрупповая алгебра

§ 6. Скрученная полугрупповая алгебра

Определение полугрупповой алгебры можно обобщить, введя **подкрутку**. А именно, определим таблицу умножения равенством $e_x e_y = a_{xy} e_{xy}$, где $a_{xy} \in R$. Обозначим получающееся при этом кольцо через $R^a[M]$. Выясним, какому условию должна удовлетворять матрица коэффициентов (a_{xy}), $x, y \in M$, чтобы получающееся так умножение было ассоциативным и имело 1.

Теорема. 1) *$R^a[M]$ – кольцо относительно операций $+$, $*$;*

2) *Если R ассоциативно, а $a = (a_{xy})$ удволетворяет следующему условию*

$$a_{xy} a_{xy,z} = a_{x,yz} a_{yz}, \qquad x, y, z \in M,$$

то $R[M]$ тоже ассоциативно;

3) *Если M и R коммутативны, а $a = (a_{xy})$* **симметрична**, *$a_{xy} = a_{yx}$, то $R[M]$ коммутативно;*

4) *Если M – моноид, R – кольцо с 1, а $a = (a_{xy})$* **нормирована**, *т.е. $a + x1 = 1 = a_{x1}$, то $R[M]$ – кольцо с 1.*

Доказательство. Единственный слегка новый момент в этом утверждении это пункт 2, проверим его. В самом деле,

$$(e_xe_y)e_z = a_{xy}e_{xy}e_z = a_{xy}a_{xy,z}e_{(xy)z},$$

$$e_x(e_ye_z) = e_x(a_{yz}e_{yz}) = a_{x,yz}a_{yz}e_{x(yz)}.$$

Сравнивая два эти выражения, мы и получаем сформулированное в пункте 2 условие.

Определение. *Функция $f : M \times M \mapsto R$ называется* **2-коциклом** *или, классически,* **системой факторов***, если для любых $x, y, z \in M$ имеет место равенство*

$$f(x, y)f(xy, z) = f(x, yz)f(y, z).$$

Приведем очевидный пример 2-коцикла, отвечающий **масштабированию** базиса (scaling) в $R[M]$, но не меняющий алгебру. Заменим базис e_x, $x \in M$, на базис $e'_x = u_xe_x$, где $u_x \in R^*$. Тогда в новом базисе таблица умножения примет вид

$$e'_xe'_y = u_xu_ye_xe_y = u_xu_ye_xy = u_xu_yu_{xy}^{-1}e'_{xy}.$$

Легко видеть, что если R коммутативно, то функция

$$M \times M \longrightarrow R, \qquad (x, y) \mapsto a_{xy} = u_xu_yu_{xy}^{-1},$$

удовлетворяет соотношению, определяющему 2-коцикл (проверьте это!) 2-коцикл вида $a_{xy} = u_xu_yu_{xy}^{-1}$ называется **2-кограницей** или коциклом **гомологичным** 0. Два коцикла называются **гомологичными**, если они отличаются на кограницу, $a_{xy} \sim b_{xy} = u_xu_yu_{xy}^{-1}a_{xy}$.

Задача. Убедитесь, что отвечающие гомологичным коциклам скрученные полугрупповые алгебры изоморфны.

§ 7. Дискретное преобразование Фурье

1. Дискретное преобразование Фурье. Преобразованием Фурье в алгебре называется любой гомоморфизм из кольца функций относительно свертки в кольцо функций относительно умножения. Опишем самую простую, но **чрезвычайно** важную для приложений в теории передачи информации конструкцию. **Циклическая свертка** векторов длины n – это свертка функций на аддитивной группе $\mathbb{Z}/n\mathbb{Z}$. При этом функцию $f : \mathbb{Z}/n\mathbb{Z} \longrightarrow K$ принято записывать как **столбец** $f = (f_0, \dots, f_{n-1})^t$, состоящий из ее значений на классах $0, \dots, n-1 \in \mathbb{Z}/n\mathbb{Z}$. Если f и g – две такие функции, то их сверткой является функция $h = f * g$, значение которой на $m \in \mathbb{Z}/n\mathbb{Z}$ определяется формулой $h_m = \sum f_ig_j$, где сумма берется по всем $i, j \in \mathbb{Z}/n\mathbb{Z}$ таким, что $m = i + j$. Например,

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_0b_0 + a_1b_2 + a_2b_1 \\ a_0b_1 + a_1b_0 + a_2b_2 \\ a_0b_2 + a_1b_1 + a_2b_0 \end{pmatrix}.$$

Кольцо $K^{\mathbb{Z}/n\mathbb{Z}}$ относительно циклической свертки можно истолковать еще как фактор-кольцо кольца многочленов по модулю $x^n - 1$.

Задача. Постройте изоморфизм $R^{\mathbb{Z}/n\mathbb{Z}} \cong K[x]/(x^n-1)$.

Для того, чтобы вычислить циклическую свертку двух векторов длины n, необходимо произвести n^2 умножений. Оказывается, если характеристика K не делит n и в поле K есть первообразный корень степени n из 1, это кольцо изоморфно кольцу $K^n = K \oplus \ldots \oplus K$ с покомпонентными операциями, в котором умножение двух векторов требует всего лишь n умножений в поле K. В самом деле, пусть ζ – первообразный корень степени n из 1. Образуем матрицу Вандермонда

$$V = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \zeta & \zeta^2 & \ldots & z^{n-1} \\ 1 & \zeta^2 & \zeta^4 & \ldots & \zeta^{2(n-1)} \\ & & & \ldots & \\ 1 & \zeta^{n-1} & \zeta^{2(n-1)} & \ldots & \zeta^{(n-1)^2} \end{pmatrix},$$

называемую обычно **матрицей дискретного преобразования Фурье** или, сокращенно, **матрицей DFT**. Оказывается, умножение на F и реализует требуемый гомоморфизм.

Дискретная теорема Планшереля. *Пусть n не делится на характеристику поля K и ζ – первообразный корень степени n в K. Тогда отображение*

$$F : K^{\mathbb{Z}/n\mathbb{Z}} \longrightarrow \underbrace{K \oplus \ldots \oplus K}_{n \text{ слагаемых}}, \qquad f \mapsto F(f) = Vf,$$

является изоморфизмом колец.

§ 8. Преобразование Фурье

Преобразование Лапласа. Пусть $f : [0,\infty) \longrightarrow \mathbb{C}$ – функция, интегрируемая по Лебегу на каждом конечном интервале $[0,a]$. Сопоставим ей функцию $L(f)$ комплексного переменного, полагая

$$(L(f))(x) = \int_0^\infty f(t)e^{-zt}dt.$$

Напомним, что свертка функций на $[0,\infty)$ определяется равенством

$$(f * g)(t) = \int_o^t f(t-s)g(s)ds.$$

Теорема Бореля утверждает, что преобразование Лапласа является гомоморфизмом кольца функций в вещественной области относительно свертки в кольцо функций в комплексной области относительно поточечного умножения $L(f * g) = L(f)L(g)$. УТОЧНИТЬ КЛАСС ФУНКЦИЙ!! У Бореля там еще произносятся слова, что если $L(f), L(g)$ определены в полуплоскости $\mathrm{re}(x) > r$, то $L(f * g)$ определено там же и т.д.

§ 9. Кольцо Дирихле арифметических функций

Наиболее часто используемым умножением в кольцах собственно в алгебре является *свертка*.

1. Кольцо Дирихле. Снабдим множество R всех функций $\mathbb{C}^{\mathbb{N}} = \mathrm{Map}(\mathbb{N}, \mathbb{C})$ обычным сложением, а в качестве умножения возьмем **свертку Дирихле**, определенную формулой

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right),$$

где сумма берется по всем натуральным делителям d числа n.

Теорема. *Эти операции превращают $R = \mathbb{C}^{\mathbb{N}}$ в ассоциативное коммутативное кольцо с 1.*

Мы не будем доказывать этот результат, так как он сразу вытекает из гораздо более общих результатов § ?. Заметим лишь, что в качестве 1 этого кольца служит δ-функция $e = \delta_{1n}$, принимающая значение 1 в 1, и значение 0 во всех остальных натуральных числах. Получившееся кольцо называется **кольцом Дирихле арифметических функций**. Оно вносит порядок в вековой хаос аналитической теории чисел. Многие классические результаты теории чисел и комбинаторики получают *буквальное* толкование и естественный смысл в терминах кольца Дирихле: так, скажем, формулы обращения становятся просто умножением на обратный элемент в этом кольце (см. Главу ?).

Комментарий. Более искушенный читатель наверняка слышал о рядах Дирихле. А именно, каждой арифметической функции $f : \mathbb{N} \longrightarrow \mathbb{C}$ сопоставляется **ряд Дирихле**

$$L(s, f) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s},$$

где $s \in \mathbb{C}$. Это сопоставление задает **гомоморфизм** кольца Дирихле в кольцо рядов Дирихле:

$$L(s, f + g) = L(s, f) + L(s, g), \qquad L(s, f * g) = L(s, f)L(s, g).$$

Задача. Доказать, что кольцо R арифметических функций является областью целостности. Доказать, что функция f тогда и только тогда обратима в R, когда $f(1) \neq 0$.
Указание. Предположим, что $f * g = 0$, а $f, g \neq 0$. Рассмотрите наименьшие натуральные m и n такие, $f(m) \neq 0$ и $g(n) \neq 0$ и вычислите $(f * g)(mn)$. Формула для обратного в случае $f(1) = 1$ приведена в Задаче ?.?, модифицируйте ее.

Задача. Доказать, что кольцо Дирихле не является нетеровым.
Указание. Как обычно, проще доказать значительно более сильный факт, а именно, что в действительности кольцо Дирихле изоморфно кольцу формальных степенных рядов от счетного числа переменных над $\mathbb{C}$.

Задача. Доказать, что если рассматривать в R в качестве произведения не свертку Дирихле, а унитарную свертку мы по-прежнему получим область целостности с той же самой единицей e.

В следующем параграфе мы узнаем, что обратным элементом к функции I, тождественно равной 1, в кольце Дирихле является функция Мебиуса μ.

Задача. Найдите функцию, обратную к функции I, в нашем кольце арифметических функций с унитарной сверткой.

Задача. Доказать, что множество R функций $\mathrm{Map}(\mathbb{N}_0, \mathbb{C})$ с обычным сложением и умножением Абеля в качестве свертки представляет собой коммутативное кольцо с 1. В качестве 1 этого кольца служит функция δ, принимающая значение 1 в 0 и 0 во всех натуральных числах. Показать, что это кольцо без делителей 0 и функция f в том и только том случае обратима в R, когда $f(0) \neq 0$.

Будем рассматривать ln как арифметическую функцию, значение которой в n равно $\ln(n)$.

Задача. Доажите, что умножение на ln является дифференцированием кольца Дирихле, иными словами, $\ln(f * g) = (\ln \cdot f) * g + f * (\ln \cdot g)$.

§ 10. Формула обращения Мебиуса

1. Формула обращения Мебиуса. Пусть, как и выше, $R = \mathbb{C}^{\mathbb{N}}$ – кольцо Дирихле арифметических функций. Рассмотрим постоянную функцию $E(n) = 1$. Для любой арифметической функции $f \in R$ ее свертка $E * f$ с E называется **сумматорной функцией** для f:

$$(E * f)(n) = \sum_{d|n} f(d).$$

Так как $E(1) = 1$, то функция E обратима в R. Обратная к ней функция $\mu = E^{-1}$ называется **функцией Мебиуса**. Пусть $n = p_1^{m_1} \dots p_s^{m_s}$ – каноническое разложение числа n. Положим $n_{\mathrm{red}} = p_1 \dots p_s$. Число называется **бесквадратным**, если $n = n_{\mathrm{red}}$.

Лемма. *Функция Мебиуса выражается следующим образом:*

$$\mu(n) = \begin{cases} (-1)^s, & \text{если } n = p_1 \dots p_s, \\ 0, & \text{в противном случае.} \end{cases}$$

Доказательство. По определению функции μ выполняются следующие соотношения: $\mu(1) = 1$ и $\sum_{d|n} \mu(d) = 0$ для любого $n \geq 2$. Если $n = n_{\mathrm{red}} = p_1 \dots p_s$ бесквадратное, то рассуждаем индукцией по s. База индукции: случай $s = 0$. Шаг индукции: предположим, что для всех бесквадратных чисел, являющихся произведением менее, чем $s \geq 1$ простых, формула уже доказана. Тогда определение $\mu(n)$ превращается в

$$1 - C_s^1 + C_s^2 - \dots + (-1)^{s-1} C_s^{s-1} + \mu(n) = 0.$$

Сравнивая это равенство с равенством $(1-1)^s = 0$, мы видим, что $\mu(s) = (-1)^s$.

Если $n \neq n_{\mathrm{red}}$ не является бесквадратным, то будем вести индукцию по числу простых делителей n. Ясно, что $\mu(p^2) = 0$. Предположим теперь, что для всех m, не являющихся бесквадратными и таких, у которых простых делителей меньше, чем у n, уже известно, что $\mu(m) = 0$. Тогда

$$\mu(n) = -\sum_{d|n, d \neq n} \mu(d) = -\sum_{d|n_{\mathrm{red}}} \mu(d) - \sum_{\substack{d|n, d \neq n \\ d \nmid n_{\mathrm{red}}}} \mu(d),$$

где первое слагаемое равно 0 по определению μ, а второе по индукционному предположению.

2. Формула обращения Мебиуса. Формула $f = \mu * (E * f)$ называется **формулой обращения Мебиуса**. С научной точки зрения это не что иное, как ассоциативность умножения в кольце Дирихле. Формулу Мебиуса можно переписать в виде

$$f(n) = \sum_{d|n} \mu(d)(E * f)(n/d) = \sum_{d|n} \mu(n/d)(E * f)(d)$$

Иными словами, знания сумматорной функции $E * f$ достаточно, чтобы восстановить f. Сформулируем ее еще раз в форме обычной для руководств по комбинаторике и теории чисел.

Теорема (формула обращения Мебиуса). *Если* $g(n) = \sum\limits_{d|n} f(d)$, *то*

$$f(n) = \sum_{d|n} \mu(n/d)g(d).$$

Эту формулу часто используют и в мультипликативной записи.

Теорема (мультипликативная формула обращения Мебиуса). *Если* $g(n) = \prod\limits_{d|n} f(d)$, *то*

$$f(n) = \prod_{d|n} g(d)^{\mu(n/d)}.$$

3. Функция Мангольдта. Определим **функцию Мангольдта**

$$\Lambda(n) = \begin{cases} \ln(p), & \text{если } n = p^m,\ p \in \mathbb{P},\ m \geq 1, \\ 0, & \text{в противном случае.} \end{cases}$$

Эта функция была введена Х.Мангольдтом в 1894 году.

Задача. Убедитесь, что $\Lambda(n) = \sum\limits_{d|n} \mu(d) \ln\left(\dfrac{n}{d}\right)$.

Решение. В самом деле, по самому определению функции Мангольдта

$$\sum_{d|n} \Lambda(d) = \ln(n),$$

так что это просто формула обращения Мебиуса.

4. `Mathematica` арифметических функций. Функции Эйлера $\varphi(n)$ и Мебиуса $\mu(n)$ определены в ядре программы и вызываются посредством `EulerPhi[n]` и `MoebiusMu[n]`. Кроме того, в ядре имплементированы еще несколько арифметических функций, в том числе функция `DivisorSigma[m,n]`, которая вычисляет $\sigma_m(n)$ – сумму m-х степеней, в частности, `DivisorSigma[0,n]` – количество всех различных делителей n.

Функция Мангольдта не имплементирована но, конечно, для *небольших* целых чисел ее легко вычислить при помощи функции `FactorInteger`, например, следующим образом:

```
mangoldt[n_]:=Block[{x},x=FactorInteger[n];
                If[Length[x]==1,Ln[First[First[x]]],0]]
```

Несомненно, это *очень* примитивный подход, так как уже для случайного числа n, $10^{35} \leq n \leq 10^{40}$, вычисление, использующее `FactorInteger`, часто требует времени порядка нескольких минут работы CPU.

§ 11. Алгебра L^1

Обозначим через l^1 множество последовательностей $a = (a_n)_{-\infty<n<\infty}$ таких, что $\sum_n |a_n| < \infty$. Если $b = (b_n)_{-\infty<n<\infty}$ – вторая такая последовательность, то определим их свертку формулой

$$(a * b) = \left(\sum_{i=-\infty}^{\infty} a_i b_{n-i} \right)_{-\infty<n<\infty}.$$

Несложно убедиться, что

$$\sum_{n=-\infty}^{\infty} \left| \sum_{i=-\infty}^{\infty} a_i b_{n-i} \right| = \sum_{i=-\infty}^{\infty} |a_i| \sum_{i=-\infty}^{\infty} |b_i|$$

Таким образом, для $a, b \in l^1$ их свертка $a*b$ определена и снова принадлежит l^1. Нейтральным элементом свертки является последовательность e, для которой $e_0 = 1$ и $e_i = 0$ при $i \neq 0$.

Пространство $L^1(-\infty, \infty)$ является кольцом по отношению к свертке

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy.$$

Для проверки этого нужно, прежде всего, убедиться в том, что свертка двух функций из L^1 снова попадает в L^1. В самом деле, по теореме Фубини (или Фубини-Тонелли??)

$$\int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} f(x-y)g(y)dy \right| dx \leq$$

$$\leq \int_{-\infty}^{\infty} |f(x-y)|dx \int_{-\infty}^{\infty} |g(y)|dy = \int_{-\infty}^{\infty} |f(x)|dx \cdot \int_{-\infty}^{\infty} |g(y)|dy.$$

Таким образом, действительно, если $f, g \in L^1(-\infty, \infty)$, то $f * g \in L^1(-\infty, \infty)$, причем $\|f * g\| \leq \|f\|\|g\|$. К сожалению, это кольцо без 1!!! НУЖНО ФОРМАЛЬНО ПРИСОЕДИНИТЬ!

Тема ?. МНОГОЧЛЕНЫ И ИХ РОДСТВЕННИКИ

Сейчас мы начнем изучать одну из основных конструкций алгебры – кольцо многочленов $R[x]$. Кроме того, мы введем несколько колец родственных $R[x]$, а именно, кольцо $R[x, x^{-1}]$ многочленов Лорана, кольцо $R[[x]]$ формальных степенных рядов, кольцо $R((x))$ формальных рядов Лорана, кольца скрученных многочленов $K[x, \delta]$, алгебра разделенных степеней и т.д.

§ 1. Кольцо многочленов

Сейчас мы конкретизируем конструкцию моноидной алгебры для случая аддитивного моноида $\mathbb{N}_0$. Получающаяся при этом алгебра называется кольцом многочленов.

1. Почему неправильно школьное определение многочленов? Пусть R – произвольное коммутативное кольцо с единицей. Мы построим некоторое новое кольцо $R[\mathrm{x}]$, называемое кольцом многочленов от одной переменной над кольцом R. Неформально, многочлен от x с коэффициентами из R – это *выражение* вида $f = a_n x^n + \ldots + a_1 x + a_0$, где $a_i \in R$. В школе многочлен обычно *отождествляется* с определяемой им **полиномиальной функцией** $\widetilde{f}$, которая сопоставляет каждому элементу $c \in R$ значение f в точке c, т.е. элемент $f(c)$, получающийся при подстановке c вместо x в выражение для f:

$$f(c) = a_n c^n + \ldots + a_1 c + a_0 \in R.$$

Действительно, функция $\widetilde{f}$ определяется многочленом f однозначно, однако, *даже в случае, когда $R = K$ является полем*, обратное, вообще говоря, совершенно **неверно**, т.е. различные многочлены могут определять одну и ту же полиномиальную функцию $\widetilde{f}$. В самом деле, пусть $K = \mathbb{F}_2 = \{0, 1\}$ – поле из двух элементов. Тогда, как легко видеть, многочлены x и x^2 определяют одну и ту же полиномиальную функцию $\mathrm{id} : K \longrightarrow K$. Если следовать школьному определению многочлена, то мы должны объявить, что $x = x^2$. Ясно, что такое решение нас совершенно не устраивает и, поэтому, мы должны признать, что чем бы ни был многочлен, он **не является** функцией $R \longrightarrow R$ (хотя и **определяет** такую функцию!) В действительности, мы определим многочлен как функцию, но функцию из $\mathbb{N}_0$ в R.

На самом деле, x должно быть **трансцендентным** над R (именно этот смысл вкладывается в выражение **независимая переменная**). Это значит, что все степени x должны быть *линейно независимы* над R, или, что то же самое, два многочлена $f = a_n x^n + \ldots + a_1 x + a_0$, и $g = b_n x^n + \ldots + b_1 x + b_0$, равны, если и только если все их коэффициенты при одинаковых степенях x совпадают: $a_i = b_i$ для всех $i \in \mathbb{N}_0$. В элементарной математике эту мысль выражают, говоря, что многочлен является *формальной* суммой вида $a_n x^n + \ldots + a_1 x + a_0$. Ясно, что при этом x не может быть элементом R и, таким образом, a priori совершенно непонятно, какой же смысл следует вкладывать в сложение и умножение в выражении $a_n x^n + \ldots + a_1 x + a_0$. Следующая конструкция может показаться излишне усложненной, но она как раз и служит для того, чтобы придать строгий формальный смысл этим операциям.

2. Кольцо многочленов. Построим, отправляясь от R, новое кольцо $R[\mathrm{x}]$ следующим образом. Как множество $R[\mathrm{x}]$ состоит из формально бесконечных последовательностей $(a_0, a_1, a_2, \ldots)$ с компонентами из R. Это означает, что $a_i \in R$, причем $a_i = 0$ для **почти всех** индексов i, т.е. для всех значений i, кроме конечного числа. Определим в множестве $R[\mathrm{x}]$ операции сложения и умножения следующим образом.

★ Сложение в $R[\mathrm{x}]$ *покомпонентное*:

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots),$$

ясно, что среди сумм $a_i + b_i$ лишь конечное число отлично от 0, так что последовательность в правой части действительно принадлежит $R[\mathrm{x}]$ вместе с исходными последовательностями.

★ Умножение в $R[x]$ определяется посредством

$$(a_0, a_1, a_2, \dots)(b_0, b_1, b_2, \dots) = (c_0, c_1, c_2, \dots)$$

где для всех $h \in \mathbb{N}_0$ компонента c_h равна

$$c_h = \sum a_i b_j, \quad i, j \in \mathbb{N}_0,\ i + j = h.$$

Таким образом, $c_0 = a_0b_0$, $c_1 = a_0b_1 + a_1b_0$, $c_2 = a_0b_2 + a_1b_1 + a_2b_0$, $c_3 = a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0$ и так далее. Ясно, что, во-первых, каждая из таких сумм конечна (входящая в определение c_h сумма состоит из $h + 1$ слагаемого), и, во-вторых, лишь конечное число среди этих сумм отлично от 0 (если h достаточно велико, то хотя бы одно из слагаемых i или j тоже должно быть достаточно велико, так что, соответственно, $a_i = 0$ или $b_j = 0$). Это значит, что мы действительно ввели произведение двух последовательностей из $R[x]$, которое снова принадлежит $R[x]$.

Чуть пристальнее всмотревшись в это определение, Вы увидите, что это в точности определение свертки функций $R^{\mathbb{N}_0}$, где в качестве операции $\circ$ на $\mathbb{N}_0$ берется сложение. Так что, по существу, мы просто еще раз определили в этом конкретном случае моноидную алгебру $R[\mathbb{N}_0]$. Однако, эту алгебру принято обозначать $R[x]$, где $x = \delta_1$.

Теорема. *Множество $R[x]$ с так определенными операциями является коммутативным ассоциативным кольцом с 1.*

Мы уже доказывали эту теорему в Главе ?, однако для удобства пешеходов повторим это доказательство.

Доказательство. Проверка большинства аксиом кольца очевидна. Так, например, то, что $R[x]$ образует абелеву группу по сложению, сразу вытекает из того, что сложение покомпонентно и соответствующие свойства выполнены для всех компонент. Например,

• нейтральным элементом по сложению будет последовательность $(0, 0, 0, \dots)$,

• элементом, противоположным к последовательности $(a_0, a_1, a_2, \dots)$, является последовательность $(-a_0, -a_1, -a_2, \dots)$.

В силу того, что a и b входят в определение c симметричным образом, а умножение в кольце R коммутативно, умножение в R[x] также коммутативно. При этом

• нейтральным элементом по умножению является последовательность $(1, 0, 0, \dots)$.

Непосредственно из определений действий проверяется и дистрибутивность.

Единственным чуть менее тривиальным моментом является проверка ассоциативности умножения в $R[x]$. Как мы знаем, она получается изменением порядка суммирования в выражении коэффициента в произведении трех множителей через коэффициенты сомножителей. А именно, если $(a_0, a_1, a_2, \dots)$, $(b_0, b_1, b_2, \dots)$ и $(c_0, c_1, c_2, \dots)$ суть три последовательности из $R[x]$, то изменением порядка суммирования легко убедиться, что при любой из двух расстановок скобок l-я компонента произведения этих последовательностей равна $\sum a_i b_j c_h$, где сумма берется по всем $i, j, h \in \mathbb{N}$, таким, что $i + j + h = l$. Теорема полностью доказана.

3. Одночлены. Последовательности вида $(a, 0, 0, \dots)$ складываются и умножаются так же, как элементы кольца R, поэтому в дальнейшем мы будем отождествлять такую последовательность с элементом $a \in R$, иными словами, мы полагаем $(a, 0, 0, \dots) = a$. Тем самым, R будет рассматриваться как подкольцо в R[x].

Независимую переменную x можно истолковать следующим образом. Положим $x = (0, 1, 0, 0, \dots)$. Применяя к x данное выше определение умножения,

легко убедиться, что x^n есть последовательность, у которой n-я компонента равна 1, в то время как все остальные компоненты нулевые:

$$\begin{aligned} x^0 &= (1,0,0,0,0,...),\\ x^1 &= (0,1,0,0,0,...),\\ x^2 &= (0,0,1,0,0,...),\\ x^3 &= (0,0,0,1,0,...),\\ &\dots \end{aligned}$$

и так далее. Элементы вида x^n называются **нормированными одночленами**. Ясно, что нормированные одночлены $1, x, x^2, \dots$ образуют базис свободного модуля $R[x]$. Этот базис обычно называется **стандартным базисом** (в дальнейшем нам будет встречаться много других базисов, более удобных для тех или иных типов задач.)

Вообще, любая последовательность вида ax^n, у которой только одна компонента отлична от нуля, а все остальные компоненты равны 0, называется **мономом** (или **одночленом**).

4. Стандартная запись многочлена. Таким образом, если a_n – последняя ненулевая компонента в последовательности $f = (a_0, a_1, \dots, a_n, 0, \dots)$, то f может быть записана как сумма мономов $a_i x^i$, $0 \le i \le n$:

$$(a_0, a_1, \dots, a_n, 0, \dots) = a_0 + a_1 x + \dots + a_n x^n,$$

при этом обычно (но не всегда) применяется запись по убывающим степеням x), т.е. в действительности пишут

$$(a_0, a_1, \dots, a_n, 0, \dots) = a_n x^n + \dots + a_1 x + a_0.$$

При этом представление элемента $f \in R[\mathrm{x}]$ в таком виде однозначно, так как коэффициенты a_i в правой части суть в точности компоненты последовательности f, а две последовательности $(a_0, a_1, a_2, \dots)$ и $(b_0, b_1, b_2, \dots)$ в том и только том случае равны, когда $a_i = b_i$ для всех $i \in \mathbb{Z}$. Это значит, что нам удалось придать строгий смысл неформальному понятию многочлена, которое изучалось в школе. Это мотивирует следующее определение.

Определение. *Построенное выше кольцо $R[x]$ называется* **кольцом многочленов** (*или* **полиномов**) *от одной* **переменной** (*или* **неизвестной**) *над кольцом R. Элементы этого кольца называются* **многочленами** (*или* **полиномами**).

§ 2. Степень многочлена

Основным инструментом при изучении многочленов является индукция по степени. Сейчас мы введем функцию степени и рассмотрим ее первые свойства.

1. Старший член и старший коэффициент. Пусть $R[x]$ – кольцо многочленов от одной переменной x над R. Мы договорились записывать элементы $f \in R[x]$ в виде $f = a_n x^n + \dots + a_1 x + a_0$, где $a_i \in R$. Если $a_n \neq 0$, то моном $\mathrm{lt}(f) = a_n x^n$ называется **старшим членом** (leading term) многочлена f, а соответствующий коэффициент $\mathrm{lc}(f) = a_n$ – **старшим коэффициентом**

(leading coefficient)[56]. Иными словами, a_n называется старшим коэффициентом многочлена f, если $a_n \neq 0$, но $a_m = 0$ для всех $m > n$. По определению у нулевого многочлена **нет** старшего члена, но, чтобы не делать оговорок в формулах, мы будем считать, что $\mathrm{lt}(f) = 0$ и $\mathrm{lc}(0) = 0$. Слагаемое $f(0) = a_0$ называется **свободным членом** (constant term), его можно мыслить себе и как элемент кольца R и как постоянный многочлен. Значительная часть элементарных свойств многочленов основана на том, что отображения, сопоставляющие многочлену его старший член/коэффициент являются *мультипликативными* гомоморфизмами $R[x] \longrightarrow R[x]$ и $R[x] \longrightarrow R$, соответственно:

$$\mathrm{lt}(fg) = \mathrm{lt}(f)\,\mathrm{lt}(g), \qquad \mathrm{lc}(fg) = \mathrm{lc}(f)\,\mathrm{lc}(g)$$

Сформулируем такое непосредственное следствие этого утверждения.

Лемма. *Если* $\mathrm{lc}(f) \in \mathrm{Reg}(R)$, *то* $f \in \mathrm{Reg}(R[x])$.

Сопоставление многочлену его свободного члена является даже *кольцевым* гомоморфизмом $R[x] \longrightarrow R$, т.е.

$$(f+g)(0) = f(0) + g(0), \qquad (fg)(0) = f(0)g(0),$$

– в действительности то же самое верно для любой точки $c \in R$.

2. Степень многочлена. Сейчас мы введем важнейший численный инвариант многочлена.

Определение. *Пусть* $= a_n x^n + \ldots + a_1 x + a_0 \in R[x]$, *где* $a_i \in R$, $a_n \neq 0$. *Номер* n *старшего коэффициента* $\mathrm{lc}(f) = a_n$ *многочлена* f *называется* **степенью** *многочлена* f *и обозначается* $\deg(f)$.

Сокращение deg происходит от французского degré и/или английского degree. У нулевого многочлена $f = 0$ нет старщего члена и его степень полагается равной $-\infty$. Таким образом, степень многочлена задает функцию $\deg : R[x] \longrightarrow \mathbb{N}_0 \cup \{-\infty\}$ на кольце многочленов со значениями в множестве, полученном присоединением $-\infty$ к множеству $\mathbb{N}_0$ неотрицательных целых чисел. Заметим, что на множество $\mathbb{N}_0 \cup \{-\infty\}$ можно распространить отношение порядка на множестве целых чисел и операцию сложения, положив $-\infty < n$, для любого $n \in \mathbb{Z}$; $-\infty + n = -\infty$ и $-\infty + (-\infty) = -\infty$.

Многочлены степени 0 – это ненулевые константы, $f = a \in R$, $a \neq 0$. Многочлены $f = ax + b$, $a, b \in R$, $a \neq 0$, степени 1, называются **линейными**; многочлены $f = zx^2 + bx + c$, $a, b, c \in R$, $a \neq 0$, степени 2 – **квадратичными**; а многочлены $f = ax^3 + bx^2 + cx + d$, $a, b, c, d \in R$, $a \neq 0$, степени 3 – **кубическими**. Степень многочлена $x^{17} + x^{11} + x$ равна 17.

3. Основные свойства степени. Следующая теорема резюмирует поведение степени по отношению к основным алгебраическим операциям.

[56] Для многочленов от одной переменной понятия старшего члена (highest term) и ведущего члена (leading term) совпадают и мы пользуемся более привычными словосочетаниями ‘старший член’, ‘старший коэффициент’ в качестве перевода английских ‘leading term’, ‘leading coefficient’. В случае многочленов от нескольких переменных нам придется тщательно различать эти понятия.

Теорема. 1) *Для любых* $f, g \in R$[x] *имеем*

$$\deg(f+g) \leq \max(\deg(f), \deg(g)),$$

причем при $\deg(f) \neq \deg(g)$ *здесь имеет место равенство.*

2) *Для любых* $f, g \in R[x]$ *имеем*

$$\deg(fg) \leq \deg(f) + \deg(g).$$

3) *Если* R *– область целостности, то для любых* $f, g \in R[x]$ *имеем*

$$\deg(fg) = \deg(f) + \deg(g).$$

4) *Если* R *– область целостности, то для любых* $f \in R$[x], $g \in R[x]^\bullet$, *имеем*

$$\deg(f \circ g) = \deg(f)\deg(g).$$

Доказательство. 1) В самом деле, i-й коэффициент многочлена $f+g$ равен сумме i-го коэффициента f и i-го коэффициента g, поэтому он заведомо равен 0, если $i > \deg(f), \deg(g)$. (Заметим, что он может быть равным 0 и если $i = \deg(f) = \deg(g)$, в том случае, когда старшие коэффициенты f и g сокращаются. Если же $\deg(f) \neq \deg(g)$, то в приведенной выше формуле неравенство заменяется на равенство).

2) и 3) В самом деле, пусть $\deg(f) = m$, а $\deg(g) = n$. Тогда l-й коэффициент произведения задается формулой $c_l = \sum a_i b_j$, где $i+j = l$, a_i – i-й коэффициент f, b_j – j-й коэффициент g. Если $l > m+n$, то в этом представлении либо $a_i = 0$, либо $b_j = 0$. Таким образом, все слагаемые, входящие в c_l равны 0. Заметим, что в общем случае неверно, что $\deg(fg)$ совпадает с $\deg(f)+\deg(g)$, так как возможно, что $a_m b_n = 0$, хотя $a_m \neq 0$ и $b_n \neq 0$. Это, однако, не может произойти, если R область целостности.

Отсюда сразу вытекают такие следствия.

Следствие 1. *Если* R *– область целостности, то и* $R[x]$ *также область целостности.*

Следствие 2. *Если* R *– область целостности, то* $R[x]^* = R^*$.

Доказательство. Если $\deg(f) > 0$, то для любого $g \in R[x]$, $g \neq 0$, имеем $\deg(fg) \geq \deg(f) > 0$, так что f не может быть обратимым.

Следствие 3. *Все многочлены степени 1 над областью целостности неприводимы.*

Следствие 4. *Над областью целостности только многочлены степени 1 обратимы относительно композиции.*

§ 3. Значение многочлена, эвалюация

Смысл определения кольца многочленов состоит в том, что это кольцо обладает универсальным свойством.

1. Гомоморфизм эвалюации. Пусть вначале R – произвольное коммутативное кольцо с 1, а A – произвольная, не обязательно коммутативная R-алгебра с 1.

Определение. *Для многочлена* $f = a_nx^n + \ldots + a_1x + a_0 \in R[x]$ *и элемента* $c \in A$ *результат подстановки* c *вместо* x *в* f *называется* **значением** *многочлена* f *в точке* c *и обозначается через* $f(c) = a_nc^n + \ldots + a_1c + a_01 \in A$.

Таким образом, каждый многочлен $f \in R[x]$ определяет функцию $\widetilde{f} : A \longrightarrow A$, $c \mapsto f(c)$, сопоставляющую каждому элементу $c \in A$ значение многочлена f в c. С другой стороны, каждый элемент $c \in A$ определяет отображение $\mathrm{ev}_c : R[x] \longrightarrow A$, $f \mapsto f(c)$, называемое **эвалюацией** в точке c. Сформулируем теперь универсальное свойство кольца многочленов.

Теорема. *Для любого элемента* $c \in A$ *эвалюация* $\mathrm{ev}_c : R[x] \mapsto A$ *является гомоморфизмом* R*-алгебр. Обратно, любой гомоморфизм* R*-алгебр* $\phi : R[x] \mapsto A$ *имеет вид* $\phi = \mathrm{ev}_{\phi(x)}$.

Иными словами, теорема утверждает, что для любого $c \in A$ эвалюация ev_c является **единственным** гомоморфизмом ϕ таким, что $\phi(x) = c$.

2. Первые примеры. Упомянем несколько важнейших частных случаев, которые будут постоянно встречаться нам в дальнейшем.

• Если $A = R$ есть само кольцо R, то $f \in R[x]$ определяет функцию $\widetilde{f} : R \longrightarrow R$, сопоставляющую каждому $c \in R$ значение многочлена f в c. Функции вида $\widetilde{f}$ обычно называются **полиномиальными функциями**[57]. В школьной программе многочлен обычно *отождествляется* с соответствующей функцией $\widetilde{f} : R \longrightarrow R$. Такое отождествление еще имеет какие-то основания для многочленов с коэффициентами из классических числовых областей, но как мы скоро увидим, в общем случае полиномиальная функция $\widetilde{f}$ не определяет многочлен f.

• Если $A = S$ есть некоторое расширение кольца R (т.е. кольцо, содержащее R в качестве подкольца), то можно говорить о значениях многочленов из $R[x]$ в S. Так, часто говорят о комплексных значениях вещественных многочленов, и т.д.

• Пусть $A = K[x]$ есть кольцо многочленов. Тогда $f(x) = f$, т.е. многочлен f является своим значением на независимой переменной x. В классических учебниках вместо f обычно пишут $f(x)$. Как мы только что убедились, это ничему не противоречит, это просто **совершенно бессмысленно**, так как лишний символ x ничему не служит и лишь загромождает обозначения. Отметим, что этот пример играет ключевую роль при построении поля разложения многочлена.

• Пусть по-прежнему $A = K[x]$. Вообще, значение $f(g)$ многочлена f на многочлене g называется **композицией** многочленов g и f (sic!) и обозначается через $f \circ g$. Легко видеть, что композиция многочленов некоммутативна, т.е., вообще говоря, $f \circ g \neq g \circ f$. Например, если $f = x^2$, а $g = x+1$, то $f \circ g = f(g) = (x+1)^2$, в то время как $g \circ f = g(f) = x^2+1$. Таким образом, эти два многочлена коммутируют только если $2 = 0$, т.е. только над полем характеристики 2.

Задача. Докажите, что композиция многочленов ассоциативна, т.е. для любых трех многочленов $f, g, h \in R[x]$ имеет место равенство $(f \circ g) \circ h = f \circ (g \circ h)$.

---

[57] Впрочем, профессиональные алгебраисты часто называют такие функции **регулярными**, а профессиональные аналисты – **целыми рациональными**, в элементарных же учебниках такие функции обычно называются **алгебраическими многочленами**.

Задача. Если R – область целостности, $f, g \in R[x]$ и $f \circ g = 0$, то либо $f = 0$, либо $g \in R$.

• Важно подчеркнуть, что алгебра A не предполагается коммутативной. Так, в курсе линейной алгебры громадную роль играют 'многочлены от матриц' и 'многочлены от операторов'. Например, если $A = M(n, R)$, то значение $f(c)$ многочлена $f \in R[x]$ на матрице $c \in M(n, R)$ называется в традиционных учебниках 'многочленом от матрицы'. Это элемент $M(n, R)$, и не следует путать его с 'матричным многочленом', который является элементом $M(n, R)[x] = M(n, R[x])$.

3. Почему коммутативные коэффициенты? Почему обычно кольцо многочленов $R[x]$ рассматривается **только** в случае, когда R коммутативно? Ведь *формально* определение $R[x]$ годится для произвольного кольца коэффициентов R. Дело в том, что для некоммутативных коэффициентов обычные кольца многочленов абсолютно бесполезны! Это связано с тем, что эвалюация многочлена на нецентральном элементе $c \in R$ **не является гомоморфизмом**. В самом деле, пусть $a, b \in R$ – два некоммутирующих элемента. Рассмотрим многочлен $(x-a)(x-b) = x^2 - (a+b)x + ab \in R[x]$. Теперь вычислим значение левой и правой части этого равенства в a. Ясно, что $(a-a)(a-b) = 0$. В то же время $a^2-(a+b)a+ab = ab-ba \neq 0$. Таким образом, ev_a не является гомоморфизмом.

Разумеется, понятно, где здесь сделана ошибка. В действительности, чтобы вычислять значения в нецентральных точках, нужно рассматривать кольцо многочленов от переменной, которая **не коммутирует с коэффициентами**! В таком кольце $(x-a)(x-b) = x^2 - ax - xb + ab$. Мономами степени m теперь будут выражения вида $a_1 x a_2 \dots a_m x a_{m+1}$, в которые x входит m раз. Но тогда возникают совершенно небанальные вопросы о линейных зависимостях между такими мономами! Поэтому обычно рассматривают не общие многочлены, а многочлены, в которых переменная хотя и не коммутирует с коэффициентами, но переставляется с ними при помощи простых **коммутационных соотношений**, обычно при помощи автоморфизма и/или дифференцирования основного кольца.

• **Подстановка.** Следующий пример тоже по существу является заменой переменной, с тем, что здесь мы действуем *формально* (многочлены и степенные ряды с коэффициентами из R для нас не являются функциями из R в R). Начнем с композиции многочленов, которую мы уже обсуждали в § ?. Для фиксированного $h \in R[x]$ отображение $R[x] \longrightarrow R[x]$, $f \mapsto f \circ g$, является гомоморфизмом кольца многочленов в себя. В самом деле, $(f+g) \circ h = f \circ h + g \circ h$ и $(fg) \circ h = (f \circ h)(g \circ h)$. Такие гомоморфизмы кольца $R[x]$ на себя обычно называются **подстановками** (substitutions). В этом примере мы подставляем h вместо x. Подстановки в кольце $R[[x]]$ степенных рядов определяются аналогично, с тем, что для возможности подстановки h в любой формальный ряд $f \in R[[x]]$ свободный член ряда h должен равняться 0.

§ 4. Обратимые и регулярные многочлены

Сейчас мы посмотрим, как нужно модифицировать результаты предыдущего параграфа так, чтобы они продолжали оставаться верными для колец с делителями 0 и посмотрим, как выглядят основные типы элементов в кольце многочленов.

1. Мультипликативная группа $R[x]$. Следствие 2 можно несколько усилить. Оказывается, для того, чтобы $R[x]^* = R^*$, не обязательно требовать, чтобы R было областью целостности. Достаточно, чтобы кольцо R было приведенным. Напомним, что через $\mathrm{Nil}(R)$ обозначается множество всех нильпотентных элементов в R.

Задача. Докажите, что

$$R[x]^* = \{f = a_n x^n + \dots + a_0 \in R[x] \mid a_0 \in R^*, a_1, \dots, a_n \in \mathrm{Nil}(R)\}.$$

Решение. Пусть $g = b_m x^m + \dots + b_0$ – обратный к f многочлен. Так как $fg = 1$, то $a_0 b_0 = 1$ так что $a_0, b_0 \in R^*$. Покажем, что a_n нильпотентен. В самом деле, $a_n b_m = a_n b_{m-1} + a_{n-1} b_m = 0$. Умножая второе из этих равенств на a_n, мы видим, что $a_n^2 b_{m-1} = 0$. Продолжим действовать в таком же духе, на m-м шаге мы получим равенство $a_n b_0 + \dots +$

$a_0b_n = 0$, умножая его на a_n^m и применяя уже доказанные равенства $a_n^i b_{n-i+1} = 0$, мы видим, что $a_n^{m+1}b_0 = 0$ и, так как $b_0 \in R^*$, то $a_n \in \mathrm{Nil}(R)$. Закончите доказательство по индукции.

2. Делители 0 в $R[x]$. В действительности, даже над кольцом с делителями нуля легко построить большие классы многочленов, не являющихся делителями нуля.

Задача. Многочлен $f \in R[x]$ называется **сильно примитивным** (strongly primitive), если его коэффициенты порождают единичный идеал кольца R. Покажите, что сильно примитивный многочлен не является делителем нуля. Проверьте, что произведение сильно примитивных многочленов сильно примитивно.

В терминологии, которую мы введем в § ?, сильно примитивные многочлены образуют **мультипликативную систему**.

Теорема Маккоя. *Многочлен $f = a_nx^n + \ldots + a_0 \in R[x]$ в том и только том случае является делителем 0 в R, когда найдется такое $c \in R$, что $cf = 0$.*

Доказательство. Достаточность очевидна. Возьмем многочлен $g = b_mx^m + \ldots + b_0$ *наименьшей степени* такой, что $fg = 0$. Если $a_ig = 0$ для всех $i = 0, \ldots, n$, то доказательство закончено, так как в этом случае мы можем взять $c = b_m$. Если нет, то возьмем наибольшее j такое, что $a_jg \neq 0$. Тогда $\deg(a_jg) < \deg(g)$. В самом деле,

$$fg = (a_nx^n + \ldots + a_jx^j + \ldots + a_0)g = a_jx^jg + \ldots + a_0g = 0,$$

причем $a_jb_mx^{j+m}$ – единственное слагаемое степени $j + m$ в этом произведении. Таким образом, $fa_jg = 0$, где $\deg(a_jg) < \deg(g)$. Противоречие с определением g показывает, что такого j не существует.

Задача. Покажите, что $\mathrm{Idem}(R[x]) = \mathrm{Idem}(R)$.

3. Радикал Джекобсона кольца многочленов. В этом пункте мы вычислим радикал Джекобсона кольца $R[x]$ над произвольным кольцом коэффициентов R. Содержание следующих двух задач в совокупности составляет **теорему Снаппера** (E.Snapper). Вычислим вначале нильрадикал.

Задача. Докажите, что $\mathrm{Nil}(R[x]) = \mathrm{Nil}(R)[x]$. Иными словами, многочлен f в том и только том случае нильпотентен, когда нильпотентны *все* его коэффициенты.

Решение. Кольцо $R/\mathrm{Nil}(R)$ приведено, значит кольцо $(R/\mathrm{Nil}(R))[x]$ тоже приведено. Но ведь $(R/\mathrm{Nil}(R))[x] = R[x]/\mathrm{Nil}(R)[x]$.

Отсюда и из характеризации $R[x]^*$ сразу вытекает, что для кольца многочленов радикал Джекобсона совпадает с нильрадикалом.

Задача. Докажите, что $\mathrm{Rad}(R[x]) = \mathrm{Nil}(R[x])$.

Решение. Пусть $f \in \mathrm{Rad}(R[x])$. Тогда $1 + xf \in R^*$, и теперь из характеризации $R[x]^*$ вытекает, что все коэффициенты f нильпотентны. Но тогда по предыдущей задаче $f \in \mathrm{Nil}(R[x])$.

4. Автоморфизмы кольца многочленов. Оказывается, у кольца $K[x]$ автоморфизмов совсем немного. – ГДЕ ИСПОЛЬЗУЕТСЯ, ЧТО ЭТО ПОЛЕ?

Задача. Докажите, что любой автоморфизм кольца $K[x]$ постоянный на K является линейной заменой переменной $f \mapsto f(ax + b)$, $a \in K^*$, $b \in K$.

Таким образом, автоморфизмы $K[x]$ над K образуют группу, изоморфную группе матриц

$$\left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}, a \in K^*, b \in K \right\}.$$

Задача. Докажите, что любой автомоморфизм кольца $K[x]$ сохраняет поле K.

Из двух последних задач вытекает, что любой автоморфизм кольца $K[x]$ является композицией **полевого автоморфизма** $\phi \in \mathrm{Aut}(K)$, действующего на коэффициенты многочлена, и линейной замены переменных.

§ 5. Полиномиальные базисы: сдвинутый стандартный базис

Базис, состоящий из стандартных мономов $1, x, x^2, \ldots$ является самым известным, но далеко не единственным широко используемым R-базисом кольца $R[x]$. В действительности, в некоторых областях, таких как полиномиальная интерполяция, и во многих чисто алгебраических и комбинаторных задачах, гораздо удобнее пользоваться другими базисами. В этом и четырех следующих параграфах мы совсем коротко опишем некоторые из них.

Сдвинутый стандартный базис. Замена x^n на $(x-c)^n$ позволяет сводить вопросы, о значениях многочлена или его производных в произвольном элементе кольца к соответствующим вопросам для $c = 0$.

Как перейти от базиса к базису? Например, как перейти от базиса x^n к сдвинутому базису $(x-c)^n$? В этом простейшем случае ответ дается биномиальной теоремой/формулой Тейлора, в других случаях, которые нам встретятся, ответ будет не столь очевиден.

Задача. Убедитесь, что $x^n = \sum_{m=0}^{n} \binom{n}{m}(x-1)^m$.

Формула биномиального обращения

§ 6. Полиномиальные базисы: базис разделенных степеней

Базис разделенных степеней. $x^{(n)} = x^n/n!$

таблица умножения $x^{(m)}x^{(n)} = \binom{n}{m}x^{(m+n)}$

алгебра разделенных степеней (divided powers), с базисом $x^{(m)}$ – над полем характеристики 0 изоморфна кольцу многочленов, но над полем положительной характеристики нет и часто результаты, которые в характеристике 0 относятся к многочленам, в положительной характеристике формулируются в терминах разделенных степеней, а не многочленов!

Например, с помощью разделенных степеней можно определить экспоненту посредством $\exp(x) = 1 + x^{(1)} + x^{(2)} + \ldots$.

§ 7. Полиномиальные базисы: факториальные базисы

По поводу двух следующих базисов см. блистательную книгу Мартина Айгнера[58].

Базис возрастающих факториалов. $[x]^n = x(x+1)\ldots(x+n-1)$

Базис убывающих факториалов. $[x]_n = x(x-1)\ldots(x-n+1)$

Ясно, что $[-x]_n = (-1)^n[x]^n$

$x^n = S_{n0}[x]_0 + S_{n1}[x]_1 + \ldots + S_{nn}[x]_n$, где S_{nm} – **числа Стирлинга второго рода** (по определению $S_{00} = 1$ и $S_{n0} = 0$ для всех $n > 0$).

Числами Стирлинга первого рода называются числа s_{ni} задаваемые тождеством $[x]_n = s_{n0}x^0 + s_{n1}x^1 + \ldots + s_{nn}x^n$,

$[x]^n = |s_{n0}|x^0 + |s_{n1}|x^1 + \ldots + |s_{nn}|x^n$,

формула обращения Стирлинга

Задача. Докажите биномиальные теоремы для возрастающих и убывающих факториалов

$$[x+y]^n = \sum_{m=0}^{n} \binom{n}{m}[x]^m[y]^{n-m}, \qquad [x+y]_n = \sum_{m=0}^{n} \binom{n}{m}[x]_m[y]_{n-m}.$$

[58] М.Айгнер, Комбинаторная теория, – М., Мир, 1982.

§ 8. Полиномиальные базисы: многочлены Бернштейна

Многочлены Бернштейна $\binom{n}{m}x^m(1-x)^{n-m}$ образуют базис в модуле многочленов степени $\leq n$.

Полиномиальная последовательность (или система): (Polynomfolge) $f_n = x^n+$члены низших степеней.

в частности,

Ортогональные многочлены, которые мы изучим в Главе ?, такие как многочлены Лежандра, Эрмита, Якоби, Чебышева, Лагерра, Гегенбауэра, ... Их часто нормируют не по старшему коэффициенту, а, например, так, чтобы какой-то определенный интеграл принимал значение 1.

§ 9. Целозначные многочлены

Целозначный базис $\binom{x}{n} = [x]_n/n!$

В настоящем параграфе мы рассматриваем многочлены над $\mathbb{Q}$. Многочлен $f \in \mathbb{Q}[x]$ называется **целочисленным** (ganzzahlig, integral), если все его коэффициенты целые, т.е. если в действительности он принадлежит уже $\mathbb{Z}[x]$. Многочлен f называется **целозначным** (ganzwertig, integer-valued), если $f(\mathbb{Z}) \subseteq \mathbb{Z}$, т.е. иными словами, если для любого $n \in \mathbb{Z}$ *значение* $f(n)$ целое.

Очевидно, что любой целочисленный многочлен является целозначным. Как показывает пример многочлена $x(x-1)/2$, обратное неверно. Этот пример является специальным случаем основного примера целозначных многочленов, а именно, **биномиальных коэффициентов**:

$$\binom{x}{n} = \frac{1}{n!}x(x-1)\dots(x-n+1)$$

Классически известно, что биномиальные коэффициенты целозначны. Сейчас мы покажем, что любой целозначный многочлен является целочисленной линейной комбинацией биномиальных коэффициентов. В действительности, мы докажем чуть более сильный результат, который, собственно, и нужен в большинстве приложений (см., например, [Ha], предложение 7.3 на стр.75).

Теорема. *Пусть $f \in \mathbb{Q}[x]$ такой многочлен, что $f(n) \in \mathbb{Z}$ для всех $n \in \mathbb{Z}$, $n \gg 0$. Тогда существуют такие целые числа $a_0, \dots, a_m \in \mathbb{Z}$, что*

$$f = a_m\binom{x}{m} + \dots + a_1x + a_0.$$

Доказательство. Индукция по степени f. База индукции: $\deg(f) = 0$ – очевидно. Шаг индукции: так как $\binom{x}{m} = x^m/m!+$ члены меньшей степени, то любой многочлен степени m из $\mathbb{Q}[x]$ можно представить в указанном виде с рациональными коэффициентами $a_m, \dots, a_0 \in \mathbb{Q}$. Нужно показать, что эти коэффициенты целые. Применим к f правый разностный оператор $\Delta f(x) = f(x+1) - f(x)$. Это снова целозначный многочлен. Так как $\Delta\binom{x}{m} = \binom{x}{m-1}$ (треугольное рекуррентное соотношение для биномиальных коэффициентов), то

$$\Delta f = a_m\binom{x}{m-1} + \dots + a_2x + a_1.$$

Go индукционному предположению $a_m, \dots, a_1 \in \mathbb{Z}$. Но тогда a_0 тоже обязан быть целым так как

$$c_0 = f - a_m\binom{x}{m} - \dots - a_1x$$

есть разность двух целозначных многочленов.

§ 10. Кольцо многочленов Лорана

Первая очень важная вариация на тему конструкции кольца многочленов состоит в том, что мы рассматриваем групповое кольцо аддитивной группы $\mathbb{Z}$. В терминах предыдущего параграфа это значит, что мы допускаем не только неотрицательные, но и отрицательные степени неизвестной x. Получающееся при этом кольцо $R[x, x^{-1}]$ называется кольцом Лорановских многочленов от переменной x над кольцом R.

Многочлен Лорана изображается как

$$a_{-m}x^{-m} + \ldots + a_{-1}x^{-1} + a_0 + a_1x + \ldots + a_nx^n = \sum_{i=-m}^{n} a_ix^i,$$

где $a_i \in R$.

порядок и степень

С точки зрения теории изложенной в Главе ? кольцо многочленов Лорана является главной локализацией кольца многочленов, а именно, главной локализацией относительно мультипликативной системы стандартных одночленов $S = \{1, x, x^2, \ldots\}$. Так как эта система порождается элементом x, то можно представлять себе, что $R[x, x^{-1}]$ получается из $R[x]$ обращением *одного* элемента, а именно, x. Таким образом, каждый Лорановский многочлен имеет вид $x^m f$, где $m \in \mathbb{Z}$, $f \in R[x]$.

§ 11. Кольцо формальных степенных рядов

> What is a power series? Take the definition of a polynomial
>
> $$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$$
>
> and from it delete the last term. This gives
>
> $$a_0 + a_1X + a_2X^2 + \ldots$$
>
> as the definition of a power series; likewise for several variables. So a power series is simpler than a polynomial.
>
> Shreeram S.Abhyankar[59]

Пусть, по-прежнему, R произвольное коммутативное кольцо с 1. Определение кольца формальных степенных рядов $R[[x]]$ полностью параллельно определению кольца многочленов $R[x]$, но *проще*, так как теперь мы не предполагаем, что почти все компоненты рассматриваемых нами последовательностей нулевые.

1. Кольцо формальных степенных рядов. Как множество кольцо $R[[\mathrm{x}]]$ состоит из **всех** бесконечных последовательностей $(a_0, a_1, a_2, \ldots)$ с компонентами из R, причем операции сложения и умножения в множестве $R[[\mathrm{x}]]$ определяются *точно так же*, как в кольце многочленов $R[x]$. А именно, сложение в $R[[\mathrm{x}]]$ покомпонентное:

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots),$$

[59] S.S.Abhyankar, Historical ramblings in algebraic geometry and related algebra. – Amer. Math. Monthly, 1976, June–July, p.409–448.

а умножение определяется посредством

$$(a_0, a_1, a_2, \dots)(b_0, b_1, b_2, \dots) = (c_0, c_1, c_2, \dots)$$

где для всех $h \in \mathbb{N}_0$ компонента c_h равна $c_h = \sum a_i b_j$, причем сумма берется по всем $i, j \in \mathbb{N}_0$ таким, что $i + j = h$. Точно так же, как в параграфе 1 доказывается следующий результат.

Теорема. *Множество $R[[x]]$ с этими операциями является коммутативным ассоциативным кольцом с 1.*

Последовательность $f = (a_0, a_1, \dots, a_n, \dots)$ может быть записана как **бесконечная** сумма мономов $a_i x^i$, $0 \le i \le n$:

$$(a_0, a_1, \dots, a_n, \dots) = a_0 + a_1 x + \dots + a_n x^n + \dots = \sum_{i=0}^{\infty} a_n x^n,$$

при этом в отличие от многочленов как правило применяется запись по *возрастающим* степеням x. В этой записи сложение и умножение степенных рядов принимают вид

$$\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n,$$

$$\sum_{n=0}^{\infty} a_n x^n \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} \Big(\sum_{i+j=n} a_i b_j \Big) x^n.$$

Определение. *Построенное выше кольцо $R[[x]]$ называется* **кольцом формальных степенных рядов** *от одной* **переменной** (*или* **неизвестной**) *над кольцом R. Элементы этого кольца называются* **формальными степенными рядами** (*или просто* **степенными рядами**).

Порядок формального степенного ряда

Порядок формального ряда играет такую же роль во всех рассуждениях, касающихся степенных рядов, степень для многочленов.

$$I_m = \{f \in R[[x]] \mid \operatorname{ord}(f) \ge m\}.$$

Неравенства для порядка показывают, что $I_m \trianglelefteq R[[x]]$. Идеалы I_m образуют фундаментальную систему окрестностей 0 в некоторой топологии кольца $R[[x]]$.

Эта топология превращает $R[[x]]$ в полное метризуемое топологическое пространство (Бурбаки, Алгебра, т.II, Гл.IV, с.77-78). Кольцо многочленов $R[x]$ всюду плотно в $R[[x]]$. Понятие суммируемого семейства совпадает с обычным понятием суммируемости, определенным сходимостью в метрическом пространстве. Удобство нашего подхода состоит в том, что он чисто алгебраичексий и позволяет вообще избежать ссылки на топологию. Такого рода алгебраические субституты топологических соображений очень часто используются в алгебре, например, в теории колец, при изучении бесконечных матриц и т.д.

Задача (Ньютон–Гензель). Найти $y \in \mathbb{Q}[[x]]$ такое, что $y^2 = 1 + x$.

Решение. Ответ, конечно, дается формулой Ньютона для разложения $\sqrt{1+x}$ в ряд Тэйлора

$$y = 1 + \frac{1}{2}x - \frac{1}{2 \cdot 4}x^2 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^3 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}x^4 + \dots$$

§ 12. Обращение формальных рядов

1. Обращение рядов относительно умножения. Ясно, что 1 является нейтральным элементом относительно умножения формальных степенных рядов.

Теорема. *Пусть R – произвольное коммутативное кольцо. Тогда*

$$R[[x]]^* = \{f = \sum f_i x^i \in R[[x]] \mid f_0 \in R^*\}.$$

Доказательство. То, что левая часть содержится в правой, вытекает из того, что если $fg = 1$ для некоторого $g \in R[[x]]$, то $f_0 g_0 = 1$, так что $f_0 \in R^*$. Обратно, предположим, что $f_0 \in R^*$, построим обратный к f ряд. Для этого прежде всего представим f в виде $f = a_0(1-h)$, где h – ряд без свободного члена. Тогда $(1-h)(1+h+h^2+\dots) = 1 =$ и, таким образом, будучи произведением двух обратимых элементов, f обратим. (Где мы использовали, что h ряд без свободного члена?)

2. Разложение рациональный дроби в ряд. Таким образом, если $R = K$ – поле, то все многочлены с ненулевым свободным членом обратимы в $K[[x]]$. Это значит, что каждую рациональная дробь f/g, порядок которой в $0 \geq 0$, можно представить формальным рядом $h \in K[[x]]$, который называется **разложением** этой рациональной дроби. Пусть $h = \sum a_n x^n \in K[[x]]$. Как узнать, является ли h *разложением* рациональной дроби. В этом случае часто говорят, что h *является* рациональной дробью. Следующая задача является прекрасной иллюстрацией идеи линейной зависимости.

Задача. Докажите, что для того, чтобы $h \in K[[x]]$ было рациональной дробью, необходимо и достаточно, чтобы существовали такие $l, m \in \mathbb{N}$, что строки

$$\begin{array}{lll} (a_{m-l+1}, & a_{m-l+2}, & \dots\ ), \\ (a_{m-l+2}, & a_{m-l+3}, & \dots\ ), \\ & \dots & \\ (a_{m+1}, & a_{m+2}, & \dots\ ), \end{array}$$

линейно зависимы.

Решение. В самом деле, это значит, что существует такой многочлен $g \in K[x]^\bullet$, что $f = gh$ тоже многочлен. Пусть $g = b_0 + \dots + b_l x^l$ и $f = c_0 + \dots + c_m x^m$. Вычисляя коэффициенты произведения gh при x^n, $n > m$, мы видим, что $b_0 a_n + b_1 a_{n-1} + \dots + b_l a_{n-l} = 0$. Подставляя сюда поочередно $n = m+1, m+2$, получим

$$\begin{array}{l} b_0 a_{m+1} + b_1 a_m + \dots + b_l a_{m-l+1} = 0 \\ b_0 a_{m+2} + b_1 a_{m+1} + \dots + b_l a_{m-l+2} = 0 \\ \dots \end{array}$$

как и утверждалось.

Фигурирующее в этой задаче условие естественнее всего выражается в терминах следующего **определителя Ганкеля**:

$$H_n^{(h)} = \det \begin{pmatrix} a_n & a_{n+1} & \dots & a_{n+h-1} \\ a_{n+1} & a_{n+2} & \dots & a_{n+h} \\ & & \dots & \\ a_{n+h-1} & a_{n+h} & \dots & a_{n+2h-2} \end{pmatrix}$$

Задача. Доказать тождество Льюиса Кэррола(??)

$$H_n^{(h)}H_{n+2}^{(h)} - H_n^{(h+1)}H_{n+2}^{(h-1)} = \left(H_{n+1}^{(h)}\right)^2$$

§ 13. Обращение формальных рядов относительно композиции

Композиция формальных степенных рядов

1. Обращение рядов относительно композиции. Нейтральным элементом относительно композиции в $R[[x]]$ является ряд x: $f \circ x = f = x \circ f$. Так как композиция в $R[[x]]$ является частично определенной операцией, мы будем рассматривать подмножество $R[[x]]_0$ рядов без свободного члена, на котором композиция уже всюду определена. Следующий результат описывает обратимые ряды относительно композиции[60].

Теорема. *Пусть R – произвольное коммутативное кольцо. Тогда*

$$R[[x]]_0^* = \{f = \sum f_i x^i \in R[[x]]_0, f_1 \in R^*\}.$$

Доказательство. То, что левая часть содержится в правой, вытекает из того, что если $fg = 1$ для некоторого $g \in R[[x]]_0$, то $f_1 g_1 = 1$, так что $f_1 \in R^*$. Обратно, нам нужно показать, что если $f_0 = 0$, а $f_1 \in R^*$, то ряд f обратим в смысле композиции. В качестве первого шага, построим *правый* обратный к f ряд g. Положим $g_0 = 0$ и $g_1 = f_1^{-1}$ При всех $n \geq 2$ коэффициент при x^n в $f \circ g$ равен 0. С другой стороны, он равен коэффициенту при x^n в *многочлене* $f_1 g + f_2 g^2 + \dots f_n g^n$, являющемся начальным отрезком ряда $f \circ g$ (все остальные члены ряда $f \circ g$ начинаются с больших степеней x!). Таким образом, мы получаем рекуррентное соотношение

$$f_1 g_n + r_n(f_2, \dots, f_n, g_1, \dots, g_{n-1}) = 0,$$

где $r_n \in \mathbb{Z}[x_2, \dots, x_n, y_1, \dots, y_{n-1}]$ – некоторый многочлен с целыми коэффициентами от $2(n-1)$ переменной, линейный по переменным $x_2, \dots, x_n$. Так как $f_1 \in R^*$, то это соотношение позволяет вычислить g_n, если $g_1, \dots, g_{n-1}$ уже известны. Продолжая действовать по индукции, мы получим ряд $g = \sum g_i x^i$, такой, что $f \circ g = x$. Так как $g_0 = 0$, $g_1 \in R^*$, то к ряду g в свою очередь можно применить ту же процедуру и построить ряд $h = \sum h_i x^i$ такой, что $h \circ g = x$. Обычная выкладка, использующая ассоциативность композиции (на $R[[x]]_0$ композиция всюду определена и ассоциативна!), показывает, что $h = f$, так что g действительно является двусторонним обратным к f.

В `Mathematica` имплементрировано несколько операций с формальными степенными рядами

• `ComposeSeries[f,g]` вычисляет композицию формальных степенных рядов f и g, при этом ряд g должен быть рядом без свободного члена.

• `InverseSeries[f,x]` – обращает формальный степенной ряд f по отношению к переменной x.

[60] Анри Картан, Элементарная теория аналитических функций одного и нескольких комплексных переменных, ИИЛ, М., 1963, 296с.

§ 14. Поле формальных степенных рядов

Пусть K – произвольное п. Нашей основной целью является построение поля $\mathrm{K}((x))$, находящегося в таком же отношении к полю $K(x)$ рациональных дробей, как кольцо $K[[x]]$ формальных степенных рядов к кольцу многочленов $K[x]$. Однако, как всегда, мы вначале проведем эту конструкцию для произвольного кольца коэффициентов.

1. Кольцо формальных рядов Лорана. Пусть R – произвольное коммутативное кольцо с 1. **Формальный ряд Лорана** изображается как

$$a_{-m}x^{-m} + \ldots + a_{-1}x^{-1} + a_0 + a_1x + \ldots + a_nx^n + \ldots = \sum_{n\geq -m}^{\infty} a_nx^n,$$

где $a_i \in R$. Обозначим множество всех формальных рядов Лорана над R через $R((x))$. Ясно, что $R((x))$ является кольцом относительно обычных операций над рядами.

Для ряда Лорана $f = \sum a_ix^i$ многочлен Лорана

$$a_{-m}x^{-m} + \ldots + a_{-1}x^{-1} \in K[x^{-1}]$$

называется **главной частью** этого ряда. Основную роль в теории рядов, как с точки зрения алгебры, так и анализа, играет коэффициент $a_{-1} = \mathrm{res}(f)$, который называется **вычетом** ряда f. Как всегда, обозначим через $\mathrm{ord}(f)$ наименьший номер m такой, что $a_m \neq 0$. Если $\mathrm{ord}(f) < 0$, то $-\mathrm{ord}(f)$ называется **порядком полюса** ряда f в 0.

2. Поле формальных степенных рядов. В случае, когда $R = K$ поле, кольцо $K((x))$ допускает особенно простое описание. Дело в том, что в этом случае каждый формальный ряд из $K[[x]]^\bullet$ допускает (единственное) представление в виде $f = x^mg$, где g – формальный ряд порядка 0, а $m = \mathrm{ord}(f) \in \mathbb{N}_0$. Так как все ряды порядка 0 обратимы уже в $K[[x]]$, то для того, чтобы сделать обратимыми **все** ненулевые элементы кольца $K[[x]]$, достаточно обратить x. Таким образом, поле частных $Q(K[[x]])$ кольца формальных степенных рядов $K[[x]]$ это в точности главная локализация этого кольца в x:

$$Q(K[[x]]) = K[[x]]_x = K((x)).$$

Иными словами, каждый элемент $Q(K[[x]])$ имеет вид $f = x^mg$, где, по-прежнему, $\mathrm{ord}(g) = 0$, но теперь $m \in \mathbb{Z}$. В поле $K((x))$ можно отметить еще одно соотношение для порядка, $\mathrm{ord}(f^{-1}) = -\mathrm{ord}(f)$.

В частности, так как теперь любой ненулевой многочлен $f \in K[x]$ обратим (а не только многочлены порядка 0), то $K(x) \leq K((x))$. Каждой дроби f/g сопоставляется ряд Лорана $fg^{-1} \in K((x))$ называется ее разложением.

Предостережение. Эти результаты **не распространяются** на формальные ряды от нескольких переменных!

Задача[61]**.** Докажите, что $(x+y)^{-1} \notin K[[x,y]]_{x,y}$. Иными словами, если обратить x и y, это еще не значит, что и $x+y$ тоже станет обратимым!

Указание. Действуйте бесхитростно, умножьте $x+y$ на ряд $a_{00} + a_{10}x + a_{01}y + \ldots$ и убедитесь, что в таком произведении всегда остается по крайней мере два слагаемых.

[61] Bourbaki?, Алгебра II, Гл.IV, задача 7 на стр.80.

§ 15. Кольцо косых многочленов

$R[x; \phi, \delta]$,
$xa = \phi(a)x + \delta(a)$.
при этом $\phi \in \mathrm{Aut}(R)$, $\delta \in \mathrm{Der}(R)$.

§ 16. Кольцо многочленов от некоммутирующих переменных

$R\langle x, y\rangle$ – кольцо многочленов от некоммутирующих переменных x, y, alias свободная алгебра с двумя образующими.

§ 17. Дальнейшие варианты кольца многочленов

Квантовые многочлены $xy = qyx$.

§ 18. Лингвистические размышлизмы на тему многочленов и малочленов

Болгарин, знай свой народ и язык!

1. Многочлен versus полиномиальный. Здесь уместно сделать небольшое лингвистическое отступление. Я думаю, что большинство пишущих по-русски профессиональных математиков предпочитает пользоваться термином **многочлен** (хотя термин **полином** также вполне возможен и достаточно употребим). В то же время, единственным прилагательным, которое соотносится в математике с термином **многочлен**, является **полиномиальный**. Так, говорят о полиномиальных уравнениях, полиномиальных функциях, etc.; мне трудно даже представить себе, что могло бы означать выражение **многочленная функция**.

Еще сложнее обстоит дело с мономами, биномами и триномами. Термины **моном** и **одночлен** полностью равноправны, и большинство алгебраистов, видимо, употребляют термин **моном**, даже когда они говорят о **многочленах** (за исключением, разумеется, таких выражений, как **старший член**, **свободный член**, и т.д.). Разумеется, это тем более верно для прилагательного **мономиальный**, которое является единственно возможным в таких выражениях, как **мономиальная матрица**, **мономиальная группа**, etc. Напротив, термин **бином** настолько жестко связан с выражением **бином Ньютона**, что практически невозможно назвать **биномом** сумму двух мономов. Эпитет **биномиальный** в таких выражениях, как **биномиальный коэффициент**, **биномиальное распределение**, и т.д. также относится именно к биному Ньютона. Наконец, слово **трином**, насколько мне известно, вообще никогда не используется по-русски, будучи вытеснено знаменитым **квадратным трехчленом**. Кстати, слово **малочлен** действительно существует в русском языке, см., например, книгу А.Хованского **Малочлены**, 1997.

Это лишь один из весьма многочисленных случаев, когда русский математический узус отдает явное предпочтение калькированному существительному и заимствованному прилагательному. Приведем несколько очевидных примеров:

⋆ Сложение – **аддитивный**;

⋆ Умножение – **мультипликативный** (ну не **умножительный** же, в самом деле);

⋆ Отношение – **реляционный** (**реляционная алгебра**);

⋆ Высказывание – **пропозициональный** (в выражениях **пропозициональная связка** и т.д.; **высказывательный** просто невозможно);

⋆ Перестановка – **пермутационный** (слово **перестановочный** тоже возможно, но гораздо чаще употребляется в смысле **коммутирующий**: **перестановочные матрицы** versus **матрица перестановки** и **пермутационное представление**);

⋆ Уравнение – **эквациональный** (впрочем, этот эпитет широко употребляется лишь в логике и находящихся под ее влиянием разделах общей алгебры, большинство же математиков предпочтет парафраз **заданный уравнениями**);

⋆ Деление круга – **циклотомический** (**циклотомические полиномы** – впрочем, также и **многочлены деления круга**);

⋆ Бесконечно малая – **инфинитезимальный**;

⋆ Центр тяжести – **барицентрический**;

⋆ Последовательность – **секвенциальный** (слово **последовательный** тоже существует, но используется как синоним слова **консекутивный**: **последовательные натуральные числа**, а вовсе не для описания чего-либо, определяемого при помощи последовательностей);

⋆ Многогранник – **полиэдральный** (конечно, **многогранный** тоже употребимо в таких выражениях, как **многогранный угол** и, с другой стороны, многогранники в многомерных пространствах часто называются **полиэдрами** или **политопами** – впрочем, и в трехмерном пространстве обычно говорят о **тетраэдрах**, **октаэдрах**, ..., а не о **четырехгранниках**, **восьмигранниках**, ...; впрочем по неизвестной причине правильный **гексаэдр** называется **кубом**);

⋆ Шестиугольник – **гексагональный** (**шестиугольный** относится к форме гайки, решетки же и упаковки шаров **гексагональны**);

⋆ Определитель – **детерминантальный** (например, **детерминантальные многообразия**, конечно, и сам **детерминант** встречается довольно часто, но все же много реже **определителя**);

⋆ Слово – **вербальный** (**вербальная подгруппа**);

⋆ Сокращение – **редуцированный** (часто также **приведенный**, но, конечно, вовсе не **сокращенный**; впрочем слово **редуцированный** – не путать с **редуктивный** – поддерживается и весьма употребительным термином **редукция**);

⋆ Сдвиг – **трансляционный** (**трансляционная симметрия**);

etc., etc., etc.

Заметим, кстати, что то же явление - иногда в гораздо более радикальных формах – свойственно языку профессионалов других областей. В математике существительное **часть** дает все же прилагательное **частный** (**частные производные**), в то время как в физике – **парциальный** (**парциальные давления**). Обычными прилагательными, образованными от **спины**, **бока** и **живота** в зоологии, являются **дорсальный**, **латеральный** и **вентральный**, соответственно. Несложно привести множество подобных примеров в лингвистике, психологии и других областях знания.

2. Попытка объяснения. Можно предполагать, что отмеченное обстоятельство связано со следующими основными причинами:

• Стремлением к точности. Видимо психологически значительно легче использовать терминологически существительное, имеющее уже известное бытовое значение, чем сделать то же для прилагательных.

• Тем, что центральная часть – или, как теперь принято говорить, **твердое ядро** – употребимой сегодня русской математической терминологии сформировалась под определяющим влиянием **немецкого** языка (очевидно, что русские **множество** и **отображение** являются кальками с **Menge** и **Abbildung**, а вовсе не с **set** и **mapping** или **ensemble** и **application** – кстати, в последнем случае они бы, вероятно, так и назывались по-русски **ансамбль** и **аппликация**), а немецкой научной терминологии также свойственно использование автохтонных существительных – в противовес прилагательным, построенным на основе латинских и греческих корней.

• Трудностью образования прилагательных от выражений (хотя a priori непонятно, чем **кругоделительный** или **центротяжестный**, хуже *тысяч* других подобных искусственных образований, наподобие прилагательных **центростремительный**, **громогласный** или **златокудрый** – сотни подобных слов были **придуманы** в XVIII – XIX веках для замены заимствований из классических языков и не вызывают сегодня больше никакого протеста: **равносильный** вместо **эквивалентный**, **равноудаленный** вместо **эквидистантный**, **прямоугольный** вместо **ортогональный**, **шестиугольный** вместо **гексагональный** и т.д.[62].

• Колоссальный отпечаток наложило использование тех или иных понятий в школьной программе – и время их появления в русскоязычной литературе. Например, поскольку о **возвратных местоимениях** и **переходных глаголах** шла речь в школе, **возвратность** и **переходность** фигурируют и в серьезной лингвистической литературе. В то же время в математике эти понятия в школе традиционно не вводились и поэтому студент сразу встречается с **рефлексивностью** и **транзитивностью**. Напротив, с XVIII века в школьном преподавании речь шла о

[62]Заметим, впрочем, что *ни в одном* из этих случаев русский новодел не смог вытеснить иноязычного слова, более того, за исключением первого дублета, русские переводы вообще не воспринимаются как синонимы тех слов, которые они призваны были заменить.

переместительном, **сочетательном** и **распределительном** законах – и поэтому студенту математику приходится переучиваться на правильные термины **коммутативность**, **ассоциативность** и **дистрибутивность**.

3. Трудность русской научной речи. К счастью, русский язык не знал эпох радикального пуризма, с чем и связан тот факт, что это один из трех языков мира с *зафиксированным* словарным запасом более миллиона слов. За счет гибкости словообразования и громадного количества заимствованных слов в русском научном тексте часто используются различные термины в тех контекстах, где западные языки используют одно слово. Например, **generator** переводится, в зависимости от контекста, как **образующая**, **порождающая** или **генератор**; **relation** – как **отношение**, **соотношение**, **связь**, или даже **родственник**, **factor** – как **множитель**, **сомножитель** или **фактор**, **divisor** – как **делитель** или **дивизор**, **inversion** – как **обращение** и **инверсия**, **decomposition** – как **разложение** и **декомпозиция** и так далее. Есть и обратные примеры, скажем, английские **manifold** и **variety** оба переводятся на русский одним и тем же словом **многообразие** (калька с немецкого **Mannigfaltigkeit**, впрочем, в немецком присутствует и **Varietät**), или **unit** и **identity**, которые оба переводятся словом **единица** (снова немецкое влияние!) но таких случаев сравнительно немного. И уж совсем неизвестна другим языкам территориальная дифференциация научной речи, когда, скажем, в Петербурге говорится о **вещественных** числах, а в Москве о **действительных** (и при этом и там и там используется символ $\mathbb{R}$ подразумевающий, что эти числа называются **реальными**). Приятное разнообразие добавлялось в последние десятилетия обилием переводов *с иностранного* и отсебятиной – иногда удачной, а иногда бездарной – в передаче английских слов и конструкций (некоторых раздражает обилие непереваренных заимствований, **шейпов**, **стеков**, **пуллбэков** и **пушаутов** – но с моей точки зрения, это как раз далеко не самое худшее; опаснее всего именно отсебятина, которая не позволяет увидеть, что именно стояло в оригинале).

Поэтому читать (и в особенности писать) научные тексты по-русски значительно сложнее, чем, скажем, по-английски или по-французски. Овладение правильным и точным использованием языка представляет собой существенную часть того, что называется ‘математической культурой’. Единственный совет, который можно здесь дать – руководствоваться принципом дзен-буддийской педагогики: ‘делай как я’. Начинающий должен сознательно и бессознательно подражать тому, как используют язык мастера.

Глава 6. КОЛЬЦА МАТРИЦ

Теперь мы начнем рассматривать еще одну из важнейших алгебраических конструкций, которая позволит нам, с одной стороны, построить много новых примеров колец, а с другой, наглядно описывать линейные отображения.

§ 1. Основные определения, связанные с матрицами

В этом параграфе мы определим множество матриц фиксированного типа.

1. Матрицы. Формально, матрица — это просто семейство, индексированное двумя аргументами[63].

Определение. *Пусть I и J суть два множества (называемые в дальнейшем множеством* **строчных индексов** *и множеством* **столбцовых индексов**, *соответственно), а X — произвольное множество. Тогда* **матрицей** *типа $I \times J$ с компонентами из X называется произвольное семейство $x : I \times J \longrightarrow X$. Значение x на паре $(i, j) \in I \times J$ называется* **компонентой** (*или* **коэффициентом** *или* **матричным элементом**) *матрицы x* **в позиции** (i, j) (*иногда* **на месте** (i, j)).

Обозначим через $M(I, J, X)$ множество всех матриц типа $I \times J$ с коэффициентами из X. Матрица x обычно задается индексированной совокупностью своих матричных элементов, которые обозначаются $x(i, j) = x_{i,j}$. Обычно запятую здесь опускают и пишут просто x_{ij}. Сама матрица в этом случае обычно записывается следующим образом: $x = (x_{ij})$, $i \in I$, $j \in J$. Напомним, что понятие равенства матриц слегка отличается от понятия равенства отображений. А именно, область значений X не входит в определение равенства матриц. По определению две матрицы x и y равны, если равны соответствующие множества индексов I и J, и для любой позиции $(i, j) \in I \times J$ матричные элементы x и y в этой позиции совпадают, т.е. $x_{ij} = y_{ij}$. Отметим, что при этом матричные коэффициенты должны принадлежать пересечению областей значений X и Y отображений x и y, но сами области значений двух равных матриц могут быть различны.

Комментарий. Русский термин ‘матричный элемент’ является неудачной калькой немецкого ‘Matrixelement’. Конечно, матричный элемент является в действительности не *элементом* матрицы, а одним из ее *значений*. В английском языке для матричных элементов существует специальный термин ‘entry’, означающий, примерно, ‘запись’, ‘вхождение’, ‘словарная статья’.

2. Конечные матрицы. Часто — но далеко не всегда — множества индексов матрицы **конечны** и состоят из последовательных натуральных чисел. В случае, когда $I = \underline{m} = \{1, \dots, m\}$ и $J = \underline{n} = \{1, \dots, n\}$ говорят, что x — матрица **размера** (или **типа**) $m \times n$ (читается ‘m **на** n’). Натурально индексированная конечная матрица **изображается** прямоугольной таблицей следующего вида:

$$x = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ & \dots & \\ x_{m1} & \dots & x_{mn} \end{pmatrix}$$

[63] ‘Все, что не строчка и не столбец, есть прямоугольная матрица’ — ‘Все, что не Стренд и Пикадили, есть Белгрейвская площадь’ — Г.Джеффрис, Б.Свирлс, Методы математической физики. т.I. — М., Мир, 1969, с.1–423; стр.193.

При этом m называется **числом строк**, а n — **числом столбцов** матрицы x. Коротко матрица x записывается в виде $x = (x_{ij})$, $1 \le i \le m$, $1 \le j \le n$.

Особенно важен случай, когда множество строчных индексов совпадает с множеством столбцовых индексов, $I = J$. Такие матрицы называются **квадратными**. Множество всех квадратных матриц типа $I \times I$ с коэффициентами из X обозначается через $M(I, X) = M(I, I, X)$. В случае конечной квадратной матрицы x число $m = n$ называется **порядком** (или **степенью**) матрицы x.

Предостережение. Для того, чтобы матрица была квадратной, недостаточно, чтобы $|I| = |J|$. Необходимо, чтобы $I = J$!

Множество всех матриц размера $m \times n$ с коэффициентами из X (говорят еще **над** X) обозначается $M(m, n, X)$, либо ${}^nX^m$, либо просто $X^{m\times n}$. При этом множество всех квадратных матриц порядка n над X обозначается просто $M(n, X) = M(n, n, X)$.

Комментарий. Впрочем и в алгебре и в анализе часто рассматриваются бесконечные матрицы, индексные множества которых совпадают с $\mathbb{N}$ или $\mathbb{Z}$. При этом получается теория, которая интересна и сама по себе и как источник примеров и контр-примеров в алгебре и в связи с многочисленными приложениями в теории суммирования[64]. В анализе широко используются и континуальные матрицы $K(x, y)$, $x, y \in \mathbb{R}$, называемые там **ядрами**.

§ 2. 'Прямоугольные таблицы чисел'

В этом параграфе мы расскажем, чем **не** является матрица.

1. 'Прямоугольные таблицы чисел'. Следует иметь в виду, что

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

— это **не** единичная матрица, а один из возможных *способов задания* или *изображения* этой матрицы. В `Mathematica` **та же** матрица может задана, тысячью других способов, например, как `IdentityMatrix[3]`, как `{{1,0,0},{0,1,0},{0,0,1}}` или как

```
Table[If[i==j,1,0],{i,1,3},{j,1,3}].
```

В большинстве элементарных учебников линейной алгебры произносятся **бессмысленные заклинания**, наподобие следующего: "матрицей называется прямоугольная таблица чисел"[65]. Как отмечают Семенов и Шмидт[66], в этом определении верно все, кроме трех слов: "прямоугольная", "таблица", "чисел". Прямоугольные таблицы являются одним из способов *изображения* матриц, но отнюдь не самими матрицами. Дело в том, что обычно рассматриваются *конечные* матрицы, строки и столбцы которых *индексированы последовательными натуральными числами*. Это вводит в заблуждение.

Пусть, например, строки и столбцы матрицы индексированы элементами множества $X = \{\$, £, ¥, \mathrm{Euro}\}$, а элемент a_{xy} в позиции (x, y), $x, y \in X$ — это обменный курс из валюты x в валюту y в данном банке. Получающаяся квадратная матрица (предположительно с диагональными коэффициентами $a_{xx} = 1$ и произведениями $a_{xy}a_{yx} < 1$ при $x \ne y$) меняется от банка к банку (и ото дня ко дню). Рассмотрим теперь другую ситуацию. Пусть $X = \{\mathrm{Euro}, \mathrm{DM}, \mathrm{FF}, \mathrm{Lit}\}$. Решение EMU зафиксировало матрицу обменных курсов для всех стран еврозоны. Тем не менее в 1999–2001 годах в банках Германии, Италии и Франции эта (одна и та же!) матрица *изображалась* по разному.

---

[64] Р.Кук, Бесконечные матрицы и пространства последовательностей. — ГИФМЛ, М., 1960, с.1–471.

[65] Вариант: 'прямоугольная таблица **из** чисел' — В.А.Ильин, Э.Г.Позняк, Линейная алгебра. — Наука, М., 1974, с.1–296. стр.12.

[66] А.А.Семенов, Р.А.Шмидт, Начала алгебры. Часть II. — СПбГУ, М., 2002.

2. Опасности натуральной индексации. В большинстве элементарных учебников строки и столбцы матриц индексированы последовательными натуральными числами. Это создает у начинающих, большинства неспециалистов и даже некоторых специалистов опасную иллюзию, что строки и столбцы матриц естественным образом линейной упорядочены. Эта иллюзия поддерживается тем, что из всех алгебраических групп большинство математиков видело только полную линейную группу $\mathrm{GL}(n,K)$, притом только в векторном и ковекторном представлениях.

Однако в действительности для многих вопросов теории представлений в высшей степени существенно, что индексы, которыми нумеруются строки и столбцы матриц, являются не линейно упорядоченными, а лишь **частично** упорядоченными. Дело в том, что, кроме упомянутых выше представлений имеется еще только три типа представлений (векторное представление симплектической $\mathrm{Sp}(2l,K)$ и нечетной ортогональной $\mathrm{SO}(2l+1,K)$ групп и 7-мерное представление группы типа G_2), для которых строки и столбцы допускают естественный линейный порядок. Уже для векторного представления четной ортогональной группы $\mathrm{SO}(2l,K)$ и бивекторного представления $\mathrm{GL}(n,K)$ это совершенно не так. Подробности и много дальнейших ссылок можно найти в[67,68,69].

Разумеется, это различие становится *несравненно* более драматическим для бесконечных матриц. Дело в том, что, в то время как имеется по существу единственный способ линейно упорядочить конечное множество, для бесконечного множества это уже совершенно не так. Нет никакого *естественного* способа линейно упорядочить бесконечное множество. Если $|X|=n$, то конкретный выбор линейного порядка на X влияет, конечно, на то, какие матрицы будут верхними треугольными. Вот как выглядит кольцо $B(n,K)$ верхних треугольных матриц для $n=4$, относительно порядков $1<2<3<4$ и $1>2>3>4$:

$$\begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}, \qquad \begin{pmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & * & * & 0 \\ * & * & * & * \end{pmatrix}.$$

Однако легко видеть, что класс изоморфизма получающихся колец не зависит от выбора порядка. Но что такое верхние треугольные матрицы в $M(X,R)$, где X — счетное множество? Биекции $X\leftrightarrow\mathbb{N}$, $X\leftrightarrow\mathbb{Z}$, $X\leftrightarrow\mathbb{Q}$, где множества $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, рассматриваются с естественными порядками, определяют совершенно различные множества верхних треугольных матриц. Множества $B(\mathbb{N},K)$ и $B(\mathbb{Z},K)$ образуют различные (не изоморфные!) кольца, а множество $B(\mathbb{Q},K)$ вообще не является кольцом (потому что произведение двух матриц не определено!)

§ 3. Первые примеры матриц

Матрицы встречаются нам всюду. Здесь мы начинаем иллюстрировать это понятие.

1. Первые примеры. Вот несколько простейших примеров, в следующих параграфах описано много дальнейших примеров.

• Имеется ровно одна **пустая матрица**, не зависящая от X. Она получается, если хотя бы одно из множеств I или J пусто.

• Любую строку

$$v=(v_1,\dots,v_n)\in {}^nX$$

длины n с компонентами из X можно рассматривать как матрицу размера $1\times n$, у которой множество строчных индексов состоит из одного элемента. Обычно

[67] N.A.Vavilov, Structure of Chevalley groups over commutative rings. — Proc. Conf. Nonassociative algebras and related topics (Hiroshima — 1990), World Sci. Publ., London et al., 1991, p.219–335.

[68] E.B.Plotkin, A.A.Semenov, N.A.Vavilov, Visual basic representations: an atlas. — Int. J. Algebra and Computations, 1998, vol.8, N.1, p.61–97.

[69] N.A.Vavilov, A third look at weight diagrams. — Rend. Sem. Mat. Univ. Padova, 2000, vol.204, N.1, p.201–250.

мы будем *отождествлять* эту строку с соответствующей матрицей и писать $M(1,n,X) = {}^nX$.

• Точно так же мы *отождествим* столбец

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in X^n$$

высоты n с компонентами из X с матрицей размера $n \times 1$, у которой множество столбцовых индексов состоит из одного элемента: $M(n,1,R) = R^n$.

Часто для экономии места мы будем записывать столбец u в виде $u = (u_1, \dots, u_n)^T$, где ^{T} изображает **формальное транспонирование**.

Предостережение. Следует, однако, иметь в виду, что формальное транспонирование является не математической, а чисто **типографской** операцией, единственная цель которой сэкономить место на странице!!! За исключением случая, когда $X = R$ есть *коммутативное* кольцо, эта операция не совпадает с настоящим транспонированием, которое мы изучаем в § ?.

• Ниже приведены несколько матриц элементы которых принадлежат двухэлементному множеству $X = \{0, *\}$:

$$\begin{pmatrix} 0 & * & * \\ * & 0 & * \end{pmatrix}, \qquad \begin{pmatrix} * & * \\ 0 & * \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} * & 0 & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}.$$

Первая из них имеет размер 2×3, вторая — размер 3×2, а третья является квадратной матрицей порядка 3.

• Ниже изображена матрица, размера 3×3

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix},$$

согласно нашему определению равенства матриц, это одна и та же матрица, независимо от того, рассматриваем мы ее как матрицу с натуральными, целыми, рациональными, вещественными или комплексными элементами — или же как матрицу с элементами из множества $\{1,2,3\}$.

2. Матрицы в быту. В действительности матрицы встречаются нам всюду, где возникает нумерация объектов двумя параметрами.

• **Турнирная матрица**. На диагональ в турнирной матрице записывается 0, а в недиагональной позиции (i,j) записывается 1,1/2 или 0, в зависимости от того, как закончилась партия x и y, а именно, выиграл x у y, закончилась партия вничью или x проиграл y.

• Расстановка фигур на шахматной доске задается матрицей, индексное множество строк которой равно $\{a,b,c,d,e,f,g,h\}$, индексное множество столбцов — $\{1,2,3,4,5,6,7,8\}$, а элементы принадлежат множеству

$$\{\texttt{White,Red}\} \times \{\texttt{King,Queen,Rook,Bishop,Knight,Pawn}\} \cup \{0\}$$

Разумеется, фактически на легальную расстановку, которая может возникнуть в ходе шахматной партии, накладывается много дополнительных ограничений, вытекающих из правил шахматной игры.

• **Пиксели**. Картинка на экране монитора задается матрицей размера 640×480, 800×600, 1024×768, 1152×864, 1280×960, 1280×1024, 1600×1200, с элементами из множества `RBG`, описанного в Главе 3 Книги I. Позиции этой матрицы называются пикселями. Ясно, что это множество конечно, но довольно велико. Постараемся оценить, *насколько* оно велико, если, скажем, количество пикселей равно $1152\cdot864=995328$ (обратите внимание, что это меньше одного мегапикселя, для бытовых камер уже сегодня характерно разрешение 4-5 мегапикселей). В режиме `TrueColor` количество возможных цветов равно $|\texttt{RBG}|=256^3=16777216$. Таким образом общее число возможных картинок на экране монитора в этом режиме равно 16777216^{995328}. Это *большое* число. Чтобы понять, *насколько* оно большое, можно заметить, что уже в числе 2^{995328}, которое соответствует случаю, когда у нас всего 2 цвета, белый и черный, 299624 цифр, так что для его записи нужно 150 страниц текста обычного формата. Но ведь в действительности у нас 2^{24} цветов! Я воздержусь от того, чтобы приводить точное значение числа 16777216^{995328}, так как непосредственное вычисление `Length[IntegerDigits[16777216^995328]]` показывает, что в нем 7190967 цифр, для записи которых нужно 3596 страниц обычного формата!

§ 4. Фрагменты матриц: строки и столбцы, подматрицы, диагонали

Сейчас мы объясняем несколько способов выделять части матриц.

1. Строки и столбцы матриц. В дальнейшем мы обычно считаем, что $I=\underline{m}=1,\dots,m$ и $J=\underline{n}=\{1,\dots,n\}$ хотя все определения легко обобщаются и на общий случай.

Определение. *Пусть x — матрица порядка $m\times n$ над X. Тогда для любого индекса $1\le i\le m$ говорят о строке*

$$x_{i*}=(x_{i1},\dots,x_{in})\in {}^nX$$

*как об i-**й строке** матрицы x. Аналогично, для любого $1\le j\le n$ столбец*

$$x_{*j}=(x_{1j},\dots,x_{mj})^T\in X^n$$

*называется j-**м столбцом** матрицы x.*

Введенные нами обозначения для строк и столбцов весьма удобны и мы будем часто пользоваться ими для сокращения формул.

2. Подматрицы. Строки и столбцы являются частным случаем подматриц. Вообще для любых подмножеств $I'\subseteq I$ и $J'\subseteq J$ можно рассмотреть ограничение семейства x с $I\times J$ на $I'\times J'$. Любое такое ограничение называется **подматрицей** (Untermatrix, submatrix, субматрицей) матрицы x, со строками из I' и столбцами из J'. Например, если $1\le i_1<\dots<i_s\le m$ и $1\le j_1<\dots<j_t\le n$, где $1\le s\le m$ и $1\le t\le n$, то соответствующая подмножествам $\{i_1,\dots,i_s\}\subseteq\{1,\dots,m\}$ и $\{j_1,\dots,j_t\}\subseteq\{1,\dots,n\}$ подматрица изображается таблицей

$$\begin{pmatrix} x_{i_1j_1} & \dots & x_{i_1j_t}\\ & \dots & \\ x_{i_sj_1} & \dots & x_{i_sj_t}\end{pmatrix}$$

размера $s\times t$. На эту подматрицу можно смотреть двояко. С одной стороны, можно сказать, что она получается *выбором* строк с номерами из I' и столбцов с номерами из J'. С другой стороны, она получается *вычеркиванием* строк с номерами из $I\setminus I'$ и столбцов с номерами из $J\setminus J'$.

Рассмотрим теперь множество $M(I,X)$ квадратных матриц. Любая подматрица матрицы x, получающаяся выбором строк и столбцов из *одного и того же* множества J называется **главной подматрицей**.

3. Диагонали. Во многих вопросах играют роль позиции, стоящие на пересечении строк и столбцов с одним и тем же номером. А именно, **главной диагональю** (principal diagonal, Hauptdiagonale) матрицы x называется список

$$(x_{11}, x_{22}, \dots, x_{rr}),$$

где $r = \min(m,n)$. Позиции (i,j) в матрице, расположенные на главной диагонали, т.е. те, для которых $i = j$, называются **диагональными**. Про те позиции, для которых $i < j$, говорят, что они расположены **выше главной диагонали**, а про те, для которых $i > j$, — что они расположены **ниже главной диагонали**.

Для квадратных матриц порядка n часто говорят о **побочной диагонали** (Nebendiagonale) — это список $(x_{1n}, x_{2,n-1}, \dots, x_{n1})$. Иногда побочная диагональ называется также **косой диагональю** (skew diagonal) или **второй диагональю** (second diagonal)

Несколько реже, но тоже достаточно часто используются также наддиагонали и поддиагонали. А именно, **наддиагональю** квадратной матрицы $x = (x_{ij})$ порядка n называется список $(x_{12}, \dots, x_{n-1,n})$, а **поддиагональю** — список $(x_{21}, \dots, x_{n,n-1})$. Вообще, $(x_{1,1+r}, \dots, x_{n-r,n})$, называется **$r$-й наддиагональю**, она состоит из элементов матрицы x в позициях (i,j), где $j - i = r$. Аналогично, $(x_{1+r,1}, \dots, x_{n,n-r})$ называется **r-й поддиагональю**, она состоит из всех элементов матрицы x в позициях (i,j), где $i - j = r$.

§ 5. Некоторые индивидуальные матрицы

Сейчас мы введем несколько индивидуальных матриц, которые настолько часто возникают в дальнейшем, что мы зафиксируем их стандартные обозначения

• **Нулевая и единичная матрица.** Наиболее известными матрицами являются нулевая и единичная матрицы

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix} \qquad e = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Эти матрицы действительно являются нейтральными элементами относительно сложения и умножения матриц. В позиции (i,j) матрицы e стоит δ_{ij}. Для общего элемента единичной матрицы никогда используется запись e_{ij}, так как это обозначение используется в совершенно другом смысле.

• **Стандартные матричные единицы.** Стандартной матричной единицей называется матрица e_{ij} у которой в позиции (i,j) стоит 1, а все остальные матричные элементы равны 0. Иными словами, $(e_{ij})_{hk} = \delta_{ih}\delta_{jk}$. Изобразим, для примера, все стандартные матричные единицы степени 2:

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Стандартные матричные единицы замечательны тем, что они образуют базис кольца матриц $M(n, R)$ как свободного R-модуля.

• **Перъдиничная матрица.** Матрица с общим элементом $f_{ij} = \delta_{i,n+1-j}$ замечательна тем, что умножение на нее слева осуществляет перестановку `Reverse` на строках, а справа — на столбцах.

$$f = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots \\ 1 & \dots & 0 & 0 \end{pmatrix}$$

Перъединичная матрица часто возникает также в геометрической алгебре и теории классических групп как матрица Грама скалярных произведений по отношению к базису Витта.

• **Пробная матрица.** Матрица с общим элементом $\mathtt{Test}_{ij} = 1$ возникает в различных вопросах математической статистики, но интересна и с чисто алгебраической точки зрения

$$\mathtt{Test} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Она замечательна, например, тем, что является нейтральным элементом относительно умножения матриц *по Адамару*.

• **Шахматная доска.** Матрица с общим элементом $\mathtt{Chess}_{ij} = (-1)^{i+j}$ возникает во многих вопросах комбинаторики. Она известна как chessboard matrix или Schachbrettmatrix. Изобразим для примера эту матрицу порядка 4:

$$\mathtt{Chess} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$$

• **Сдвиг назад и сдвиг вперед.** Громадную роль во многих вычислениях играют поддиагональная матрица `Back` с общим элементом $\mathtt{Back}_{ij} = \delta_{i+1,j}$,

$$\mathtt{Back} = e_{21} + \ldots + e_{n,n-1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

и транспонированная к ней наддиагональная матрица `Forward` с общим элементом $\mathtt{Forward}_{ij} = \delta_{i,j+1}$,

$$\mathtt{Forward} = e_{12} + \ldots + e_{n-1,n} = \begin{pmatrix} 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Дело в том, что умножение строки справа на матрицу `Back` реализует преобразование этой строки известное в Computer Science как `ShiftLeft`, состоящее в том, что все элементы этой строки сдвигаются на одну позицию влево, при этом первый элемент выбрасывается, а на последнее место ставится 0. Аналогично, умножение на `Forward` осуществляет преобразование `ShiftRight`, состоящее в сдвиге всех элементов строки на одну позицию вправо.

• **Матрицы Коксетера.** Во многих вычислениях будут возникать матрица перестановки

$$\texttt{Cox} = e_{12} + \ldots + e_{n-1,n} + e_{n1} = \begin{pmatrix} 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

и транспонированная к ней матрица

$$\texttt{Cox}^{-1} = e_{1n} + e_{21} + \ldots + e_{n,n-1} = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix},$$

Для матриц перестановки транспонированная совпадает с обратной, что и отмечено в обозначении. Умножение строки справа на матрицу `Cox` реализует преобразование `RotateLeft`, состоящее в том, что все элементы этой строки *циклически* сдвигаются на одну позицию влево, иными словами, при этом первый элемент ставится на последнее место ставится 0. Аналогично, умножение на $\texttt{Cox}^{-1}$ осуществляет преобразование `RotateRight`, состоящее в циклическом сдвиге всех элементов строки на одну позицию вправо, при этом последний элемент ставится на первое место.

• **Матрица Вандермонда.** Следующая матрица десятки раз возникнет у нас в курсе по самым разным поводам

$$V(x_1, \ldots, x_n) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix}$$

§ 6. `Mathematica` матриц
1st installment: задание матриц

В системе `Mathematica` имплементировано довольно много стандартных функций, связанных с матрицами.

1. Задание матриц. Опишем, прежде всего, как порождаются матрицы. Основная внутренняя форма представления матрицы это перечисление списка, состоящего из *строк* матрицы, каждый из которых тоже трактуется как список. Так, например, запись `x={{a,b},{c,d}}` задает матрицу $x = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Таким образом, никакой симметрии между строками и столбцами матрицы в `Mathematica` нет!

• `MatrixForm[x]` позволяет фактически увидеть матрицу x в традиционной математической форме, как таблицу.

• `TableForm[x]` тоже дает обыную матричную форму матрицы x, но без круглых скобок.

Пояснение. Иными словами, различие между `MatrixForm` и `TableForm` такое же, как между TEX'овскими командами `\pmatrix`, которая ставит вокруг изображения матрицы круглые скобки (parenthesis) и `\matrix`, которая таких скобок не ставит (в действительности, есть и более тонкие отличия в трактовке ширины столбцов, связанные с форматированием вывода, которые относятся скорее к типографскому делу, чем к математике).

• `Table[f,{i,m},{j,n}]`, где f некоторая функция от i, j, генерирует матрицу $x = (x_{ij})$ размера $m \times n$ такую, что $x_{ij} = f(i, j)$.

• `Array[f,{m,n}]` где f *имя* функции, генерирует матрицу размера $m \times n$ такую, что $x_{ij} = f(i, j)$.

Пояснение. В чем различие между `Table` и `Array`? При использовании `Table` предполагается, что f действительно является функцией от i, j и ее значение фактически вычисляется для каждой пары (i, j). Так, например, `Table[f,{i,3},{j,3}]` даст `{{f,f,f},{f,f,f},{f,f,f}}`. Если i или j не входит в выражение для f, то вместо `Table[f,{i,m},{j,n}]` можно написать `Table[f,{m},{n}]`. В то же время `Array` генерирует таблицу значений, воспринимая как f как *имя* функции, так что `Array[f,{3,3}]` даст

```
{{f[1,1],f[1,2],f[1,3]},{f[2,1],f[2,2],f[2,3]},{f[3,1],f[3,2],f[3,3]}}.
```

Тем самым, `Array[f,{m,n}]` равнозначно `Table[f[i,j],{i,m},{j,m}]`.

Так, например, `Table[0,{m},{n}]` задает нулевую матрицу размера $m \times n$. Конструкция `Table[If[i<j,1,0],{i,4},{j,4}]` дает нам матрицу

```
{{0,1,1,1},{0,0,1,1},{0,0,0,1},{0,0,0,0}}.
```

Задача. В `Mathematica` есть функция `Random[ ]`, порождающая (псевдо)случайное число в отрезке $[0, 1]$. Какой из следующих конструкций Вы воспользуетесь для генерации (псевдо)случайной 3×3 матрицы: `Table[Random[ ],{3},{3}]` или `Array[Random[ ],{3,3}]`?

Решение. Первая из этих конструкций вычисляет `Random[ ]` *девять* раз, по одному для каждой пары (i, j). Это даст нам девять случайных чисел. Вторая вычислит *одно* случайное число и будет трактовать его как *имя* функции, вряд ли это то, что нам хотелось!

• `DiagonalMatrix[x]`, где $x = (x_1, \ldots, x_n)$ генерирует диагональную матрицу с диагональными коэффициентами $x_1, \ldots, x_n$, Тем самым `DiagonalMatrix[{x,y,z}]` даст нам матрицу $\mathrm{diag}(x, y, z)$.

2. Выделение фрагментов матриц. Выделение фрагментов матриц производится при помощи следующих стандартных конструкций.

• `x[[i,j]]` выделяет элемент матрицы x в позиции (i, j);

• `x[[{`$i_1,\ldots,i_r$`},{`$j_1,\ldots,j_s$`}]]` выделяет подматрицу в x, состоящую из элементов, стоящих на пересечении строк с номерами $i_1, \ldots, i_r$ и столбцов с номерами из $j_1, \ldots, j_s$;

• `x[[i]]` производит выделение i-й строки матрицы x;

• `x[[All,j]]` производит выделение i-й строки матрицы x;

• `Tr[x,List]` выделяет диагональ матрицы x.

§ 7. Сложение матриц и умножение на скаляр

В настоящем пункте мы начнем изучать алгебраические операции над матрицами. Для этого мы будем предполагать, что элементы матрицы берутся из некоторой аддитивно записываемой абелевой группы, а в дальнейшем - из некоторого модуля над ассоциативным кольцом R.

1. Сложение матриц. Пусть $X = A$ — некоторый аддитивно записанный моноид. Сейчас мы введем сумму матриц.

Определение. *Пусть x и y две матрицы одного и того же типа $I \times J$ с элементами из некоторого моноида A. Тогда* **суммой** *матриц x и y называется их сумма как отображений, т.е. такая матрица $x+y$ типа $I \times J$ с элементами из A, у которой в позиции $(i,j) \in I \times J$ стоит сумма соответствующих элементов матриц x и y: $(x+y)_{ij} = x_{ij} + y_{ij}$.*

Предложение. *Относительно сложения множество $M = M(m,n,A)$ образует абелев моноид, иными словами, сложение матриц удовлетворяет следующим аксиомам:*

A1 *Ассоциативность: $\forall x,y,z \in M$, $(x+y)+z = x+(y+z)$;*

A2 *Существование нуля: $\exists 0 \in M$, $\forall x \in M$, $0+x = x = x+0$;*

A4 *Коммутативность: $\forall x,y \in M$. $x+y = y+x$;*

Если A является группой, то и M образует абелеву группу, т.е. дополнительно удовлетворяет следующей аксиоме

A3 *Существование противоположной матрицы: $\forall x \in M$, $\exists -x \in M$, $x+(-x) = 0 = (-x)+x$.*

Доказательство. Аксиомы **A1** и **A4** выполнены автоматически, поскольку сложение в A удовлетворяет этим аксиомам, а сложение в M определяется покомпонентно.

Нейтральным элементом сложения матриц является **нулевая матрица** 0, у которой все матричные элементы равны 0, а матрицей **противоположной** к матрице $x = (x_{ij})$ является матрица, получающаяся из x заменой всех ее матричных элементов на противоположные: $(-x)_{ij} = -x_{ij}$.

2. Умножение на скаляр. Пусть теперь R — ассоциативное кольцо с 1, а $X = A$ — левый или правый R-модуль.

Определение. *Если A — левый R-модуль, то* **произведением матрицы** *x типа $I \times J$* **на скаляр** *$\lambda \in R$* **слева** *называется матрица λx типа $I \times J$ с элементами из A, у которой в позиции $(i,j) \in I \times J$ стоит произведение соответствующего элемента матрицы x на скаляр λ: $(\lambda x)_{ij} = \lambda x_{ij}$. Аналогично определяется* **произведение матрицы** *x* **на скаляр** *λ* **справа**, *в случае когда A является правым R-модулем: $(x\lambda)_{ij} = x_{ij}\lambda$.*

Если кольцо $R = R$ коммутативно, то умножение справа и слева можно не различать.

Предложение 2. *Пусть A — левый R-модуль. Тогда введенные выше операции сложения матриц и умножения на скаляр слева превращают $M = M(m,n,R)$ в левый R-модуль, иными словами, в дополнение к сформулированным выше аксиомам* (A1) — (A4) *выполняются еще следующие 4 аксиомы:*

V1 *Внешняя ассоциативность: $\forall \lambda,\mu \in R$, $\forall x \in M$, $(\lambda\mu)x = \lambda(\mu x)$;*

V2 *Дистрибутивность относительно сложения скаляров: $\forall \lambda,\mu \in R$, $\forall x \in M$, $(\lambda+\mu)x = \lambda x + \mu x$;*

V3 *Дистрибутивность относительно сложения матриц: $\forall \lambda \in R$, $\forall x,y \in M$, $\lambda(x+y) = \lambda x + \lambda y$;*

V4 *Унитальность: $\forall x \in M$, $1x = x$.*

Аналогичное утверждение выполняется с заменой левых модулей на правые, при этом левое и правое умножение связаны между собой

V5 *Двусторонняя ассоциативность:* $\forall \lambda, \mu \in R$, $\forall x \in M$, $(\lambda x)\mu = \lambda(x\mu)$.

Доказательство. Все аксиомы V1 — V4 моментально вытекают из соотвествующих аксиом для самого модуля A.

В действительности мы будем, как правило, использовать это утверждение для частного случая, когда $A = {}_R R$ или $A = R_R$. Для случая, когда кольцо R коммутативно, левые и правые умножения можно не различать и здесь говорят просто об умножении матрицы на скаляр. В частности, если K — поле, то $M(m, n, K)$ образует векторное пространство над полем K. В действительности, как мы вскоре увидим, *как векторное пространство* оно изоморфно пространству столбцов высоты mn, но имеются глубокие причины не отождествлять $M(m, n, K)$ с K^{mn}.

3. Стандартные матричные единицы как базис $M(m, n, R)$. Легко видеть, что каждая матрица $x \in M(m, n, R)$ представляется как линейная комбинация стандартных матричных единиц, причем коэффициентами этой линейной комбинации являются в точности матричные элементы матрицы x:

$$x = \sum x_{ij} e_{ij}, \qquad 1 \leq i, \leq m,\ 1 \leq j \leq n.$$

Так как e_{ij} — это единственная среди матриц e_{hk}, у которой в позиции (i, j) стоит элемент $\neq 0$, то коэффициенты этой комбинации единственны. Таким образом, стандартные матричные единицы образуют базис в $M(m, n, K)$, который называется **стандартным базисом**. Это представление особенно удобно, если матрица x имеет мало ненулевых элементов или незначительно отличается от какой-то известной матрицы, скажем, от e. В этом случае вычисления с матричными единицами становятся гораздо более удобными, чем любые другие, так как в них учитываются только те позиции, которые фактически имеют значение.

§ 8. Умножение матриц

> In my paper the fact that XY was not equal to YX was very disagreable to me. I felt this was the only point of difficulty in the whole scheme.
>
> Werner von Heisenberg
>
> Now when Heisenberg noticed that, he was **really** scared.
>
> Paul Dirac

Пусть R — ассоциативное кольцо с 1. Сейчас мы введем еще одну важнейшую операцию — умножение матриц. В отличие от сложения и умножения на скаляр, она не определяется покомпонентно, посредством формулы $(xy)_{ij} = x_{ij} y_{ij}$ — такое умножение матриц действительно рассматривается, оно называется **произведением по Адамару** и обозначается $x \circ y$. Однако в алгебре обычно рассматривается другое умножение матриц, типа свертки, которое мы сейчас и определим.

1. Произведение строки на столбец. Начнем с произведения строки на столбец.

Определение. *Пусть $u \in {}^nR$ — строка длины n, а $v \in R^n$ — столбец высоты n с компонентами из R. Тогда* **произведением** *u на v называется скаляр*

$$uv = (u_1, \dots, u_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u_1v_1 + \dots u_nv_n,$$

В действительности, здесь, конечно, не обязательно требовать, чтобы компоненты строки и столбца принадлежали некоторому кольцу, они могут быть и обьектами более общей природы, важно лишь, чтобы были определены все входящие в определение uv произведения и суммы.

2. Определение произведения матриц. Итак, произведение строки и столбца определено если и только если длина строки равна высоте столбца. Это значит, что если мы хотим определить произведение матриц в терминах произведения строк и столбцов, то длина строки первой матрицы должна равняться высоте столбца второй матрицы. Следующее определение было фактически известно уже Эйлеру (в терминах последовательного выполнения двух линейных замен переменных), но формально было введено лишь в 1842 году Артуром Кэли.

Определение. *Пусть $x \in M(m, n, R)$, $y \in M(n, p, R)$ суть матрицы над кольцом R. Тогда их* **произведением** *называется матрица $xy \in M(m, p, R)$ у которой элемент в позиции (i, j), $1 \le i \le m$, $1 \le j \le p$, равен произведению i-й строки матрицы x на j-й столбец матрицы y, иными словами, выполняется формула $(xy)_{ij} = x_{i*}y_{*j}$.*

Таким образом, произведение матриц определено, если число столбцов первого сомножителя равно числу строк второго сомножителя. При этом число строк произведения равно числу строк первого сомножителя, а число столбцов произведения — числу столбцов второго сомножителя. Подставляя в эту формулу выражение для произведения строки на столбец, приведенное в предшествующем определении, получаем явную формулу для матричного элемента $(xy)_{ij}$ в терминах матричных элементов x и y:

$$(xy)_{ij} = x_{i*}y_{*j} = x_{i1}y_{1j} + \dots + x_{in}y_{nj} = \sum x_{ih}y_{hj}, \qquad 1 \le h \le n.$$

В действительности, большинство профессиональных алгебраистов почти никогда не пользуются определением умножения матриц в таком виде, а использует одну из трех интерпретаций, изложенных в следующих параграфах.

§ 9. Столбцы и строки произведения, 1st installment

Излагаемая в настоящем параграфе интерпретация произведения матриц является исторически первой и, по существу, восходит к Эйлеру. Однако и сегодня она используется всеми, кому *реально* приходится умножать матрицы не слишком маленьких порядков, как для вычислительных, так и для теоретических целей. Именно так умножают матрицы специалисты по теории представлений, теории алгебраических групп, теории инвариантов и т.д.

1. Почему столбцы и строки? Дело в том, что в большинстве фактически возникающих матричных вычислений достаточно уметь вычислять отдельные строки и столбцы матрицы

xy, $x \in M(l,m,R)$, $y \in M(m,n,R)$. Часто нахождение всего произведения двух матриц не только не нужно, но и практически трудно осуществимо.

Например, порядок группы FG — самой большой среди спорадических конечных простых групп, первоначально известной как **большой монстр** F_1, позже получившей название **дружественный гигант** (Friendly Giant) или **группа Фишера-Грайса** (Fischer—Griess) — равен

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 =$$
$$808017424794512875886459904961710757005754368000000000.$$

Минимальная степень точного представления этой группы комплексными матрицами равна 196883. Всего несколько лет назад умножение двух матриц размера 196883×196883 на хорошей рабочей станции занимало около года вычислений. В то же время, умножение столбца высоты 196883 на такую матрицу занимает около одной минуты.

Приведу еще один пример из собственной практики. Для вычислений в самой большой исключительной группе Шевалле типа E_8 над кольцами нужно умножать матрицы размера 248×248, однако коэффициентами этих матриц являются не числа, а многочлены из $\mathbb{Z}[x_1, \dots, x_n]$. Хотя вычисление произведения двух таких матриц на бытовом компьютере на тот момент уже было возможно, оно требовало огорчительно длительного времени. Мною и Е.Б.Плоткиным[70] были развиты методы вычислений в исключительных группах Шевалле, в которых фигурируют лишь произведения таких матриц на столбцы и строки (и еще один тип вычислений, так называемые 'элементарные вычисления', которые классически известны и легко реализуются).

2. Столбцы и строки произведения. Ключом к к эффективному вычислению столбцов и строк произведения являются следующие наблюдения.

• j-й столбец матрицы xy зависит только от j-го столбца матрицы y, а именно,

$$(xy)_{*j} = xy_{*j}.$$

• i-я строка матрицы xy зависит только от i-й строки матрицы x, а именно,

$$(xy)_{i*} = x_{i*}y.$$

Конечно, эти формулы вытекают непосредственно из определения произведения матриц, но проще всего интрепретировать их следующим образом. Пусть

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

— стандартный базис модуля R^n столбцов высоты n, а

$$f_1 = (1, 0, \dots, 0), \dots, f_m = (0, \dots, 0, 1).$$

— стандартный базис модуля ${}^mR$ строк длины m. Тогда i-ю строку и j-й столбец матрицы $x \in M(m,n,R)$ проще всего интерпретировать как произведение f_i на x и как произведение x на e_i, соответственно:

$$x_{i*} = f_i x, \qquad x_{*j} = xe_j.$$

Теперь отмеченные выше формулы представляют собой просто чуть иначе записанную ассоциативность умножения:

$$f_i(xy) = (f_i x)y, \qquad (xy)e_j = x(ye_j).$$

---

[70] Описание этого проекта можно найти в N.A.Vavilov, E.B.Plotkin, Chevalley groups over commutative rings. I. Elementary calculations. — Acta Applicandae Math., 1996, vol.45, p.73–115.

§ 10. Столбцы и строки произведения, 2nd installment

1. Умножение матрицы на столбец или строку. Таким образом, нам нужно понять, как действует умножение на фиксированную матрицу x на пространствах столбцов или строк. Сменим точку зрения и будем считать, что мы хотим научиться умножать матрицу на строку или столбец.

• произведение xu матрицы x на столбец $u \in R^m$ является линейной комбинацией столбцов матрицы x, коэффициентами которой служат координаты столбца u:

$$xu = x_{*1}u_1 + \ldots + x_{*m}u_m.$$

• произведение vy строки $v \in {}^nR$ на матрицу y является линейной комбинацией строк матрицы y, коэффициентами которой служат координаты строки v:

$$vy = v_1y_{1*} + \ldots + v_mx_{m*}.$$

Таким образом,

Теорема. *Уравнение $ax = b$ в том и только том случае разрешимо, когда все столбцы матрицы b являются линейными комбинациями столбцов матрицы a.*

Понятно, что имеет место двойственное утверждение

Теорема. *Уравнение $xa = b$ в том и только том случае разрешимо, когда все строки матрицы b являются линейными комбинациями строк матрицы a.*

Во многих случаях пользуясь этой интерпретацией произведение совсем легко вычислить.

Задача. Убедитесь, что

$$\begin{aligned}
&x \cdot \mathtt{cox} = (x_{*2}, \ldots, x_{*n}, x_{*1}),\\
&x \cdot \mathtt{cox}^{-1} = (x_{*n}, x_{*1}, \ldots, x_{*,n-1}),\\
&\mathtt{cox} \cdot x = (x_{n*}, x_{1*}, \ldots, x_{n-1,*}),\\
&\mathtt{cox}^{-1} \cdot x = (x_{2*}, \ldots, x_{n*}, x_{1*}).
\end{aligned}$$

Задача. Убедитесь, что

$$\begin{aligned}
&u(e|0) = \mathtt{PadRight}[u],\\
&u(0|e) = \mathtt{PadLeft}[u],\\
&u\left(\frac{e}{0}\right) = \mathtt{DropRight}[u],\\
&u\left(\frac{0}{e}\right) = \mathtt{DropLeft}[u].
\end{aligned}$$

Разумеется, при умножении на эти матрицы столбцов, а не строчек, операции `Pad` и `Drop` меняются местами.

§ 11. Произведение матриц и матричные единицы

Специалисты по теории колец умножают матрицы иначе, в терминах матричных единиц.

1. Умножение матриц как свертка. Пусть $I = \underline{l}$, $J = \underline{m}$, $K\underline{n}$ — три множества индексов, порядков l,m и n, соответственно. Определим операцию

$$\circ : (I \times J) \times (J \times K) \longrightarrow I \times K \sqcup \{0\}$$

как **связку**, т.е. $(i,j) \circ (h,k) = \delta_{jh}(i,k)$. Вспомним теперь, что множество матриц $M(l,m,R)$ интерпретируется как множество функций $J \times J \longrightarrow R$. В этой интерпретации матричные единицы e_{ij}, $i \in I$, $j \in J$, — это в точности δ-функции, принимающие значение 1 в (i,j) и значение 0 во всех остальных парах. Множество $M(m,n,R)$ интерпретируется как $R^{J\times K}$, а $M(l,n,R)$ — как $R^{I\times K}$. Мы можем определить свертку двух функций $x : I \times J \longrightarrow R$ и $y : J \times K \longrightarrow R$ обычной формулой

$$(x * y)(i,k) = \sum_{(i,j)\circ(h,k)=(i,k)} x(i,j)y(h,k) = \sum_{j=1}^{m} x(i,j)y(j,k).$$

В частном случае $I = J = K$ это определение совпадает со определением свертки, в том виде, как она определялась в § ?.

2. Умножение матричных единиц. Обычно, вместо того, чтобы говорить об умножении матриц как свертке, этот результат переформулируют в терминах матричных единиц. При этом на языке матричных единиц связочное умножение записывается как

$$e_{ij}e_{hk} = \delta_{jh}e_{ik},$$

а определение свертки — это в точности утверждение о том, что такое умножение матричных единиц продолжается на все матрицы **по линейности**, т.е. так, чтобы выполнялись аксиомы дистрибутивности D1 и D2 и аксиома алгебры V5.

3. Выделение строки или столбца матрицы. Ясно, что умножение на стандартную матричную единицу вида e_{jh} производит выделение j-го столбца. А именно,

$$xe_{jh} = (0, \dots, 0, x_{*j}, 0, \dots, 0),$$

где x_{*j} стоит на h-м месте. Аналогично производится выделение i-й строки

$$e_{hi}x = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_{i*} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

где x_{i*} стоит на h-м месте.

Задача (умножение на пробные матрицы). Проверьте, что $x \cdot \mathtt{test} = d \cdot \mathtt{test}$, где

$$d = \mathrm{diag}(x_{11} + \ldots + x_{1n}, \ldots, x_{n1} + \ldots + x_{nn}).$$

Аналогично, $\mathtt{test} \cdot x = \mathtt{test} \cdot d$, где

$$d = \mathrm{diag}(x_{11} + \ldots + x_{n1}, \ldots, x_{1n} + \ldots + x_{nn}).$$

§ 12. Произведение столбца на строку: тензорный ранг

1. Матрицы ранга 1. Применим формулу произведения матриц к частному случаю произведения столбца на строку.

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (u_1, \ldots, u_n) = \begin{pmatrix} v_1u_1 & \ldots & v_1u_n \\ & \ldots & \\ v_nu_1 & \ldots & v_nu_n \end{pmatrix}.$$

Такая матрица называется[71] **матрицей ранга 1**. Упомянем несколько очень полезных частных случаев этой формулы:

• Тем самым, каждая стандартная матричная единица является матрицей ранга 1: $e_{ij} = e_if_j$.

• Более общо, если $u \in R^m$ — произвольный столбец высоты m, то uf_j — это матрица, у которой j-й столбец равен u, а все остальные элементы равны 0.

• Аналогично, если $v \in {}^nR$ — произвольная строка длины n, то e_i есть матрица, у которой i-я строка равна v, а все остальные элементы равны 0.

2. Умножение матриц в терминах матриц ранга 1. Если $x \in M(l, m, R)$, $y \in M(m, n, R)$, то записывая xy в виде

$$xy = xey = x(e_{11} + \ldots + e_{mm})y = xe_{11}y + \ldots + xe_{mm}y = xe_1f_1y + \ldots + xe_mf_my,$$

мы получаем еще одно очень полезное выражение для произведения двух матриц:

$$xy = x_{*1}y_{1*} + \ldots + x_{*m}y_{m*}$$

В частности, отсюда следует, что любая матрица вида xy является суммой не более, чем m матриц ранга 1.

3. Тензорный ранг. Будем говорить, что **тензорный ранг** матрицы x равен r, и писать $\mathrm{rk}(x) = m$, если матрицу x можно представить в виде суммы r матриц ранга 1 и нельзя представить в виде суммы меньшего количества таких матриц.

Комментарий. Имеется несколько конкурирующих определений ранга, строчный ранг, столбцовый ранг, ранг по минорам и т.д., и мы вернемся к этим понятиям в главе ?. В большинстве элементарных книг рангом по умолчанию называется самое бесполезное из всех определений ранга — ранг по минорам. В более продвинутых учебниках рассматривается

[71] По аналогии с тензорами, поливекторами и т.д. такую матрицу было бы естественно называть разложимой, если бы термин разложимая матрица не использовался в совершенно другом смысле, см., например, Маркус—Минк, стр.165.

столбцовый или строчный ранг, т.е. размерность пространства столбцов или строк. Однако в случае поля все эти ранги совпадают с тензорным рангом, а в случае кольца, даже коммутативного, все они не имеют большого смысла. С точки зрения обобщений на кольца тензорный ранг представляет собой *единственное* реально полезное понятие ранга.

Следующие свойства ранга очевидны.

1) Если $x \in M(m,n,R)$, то $\mathrm{rk}(x) \leq \min(m,n)$

2) Если $x,y \in M(m,n,R)$, то $\mathrm{rk}(x+y) \leq \mathrm{rk}(x)+\mathrm{rk}(y)$

3) Если $x \in M(l,m,R)$, $y \in M(m,n,R)$, то $\mathrm{rk}(xy) \leq \min(\mathrm{rk}(x),\mathrm{rk}(y))$. — ДОКАЗАТЬ!!

§ 13. `Mathematica` матриц
2nd installment: операции над матрицами

- `x.y` или `Dot[x,y]` — произведение матриц x и y.
- `Transpose[x]` — транспонированная к x матрица.
- `Inverse[x]` — обратная к x матрица.
- `Det[x]` — определитель матрицы x.
- `Tr[x]` — след матрицы x.
- `MatrixPower[x,n]` — n-я степень матрицы x.
- `MatrixExp[x]` — экспонента матрицы x.

§ 14. Элементарные матрицы и элементарная группа

Элементарные трансвекции. Матрицы вида $t_{ij}(\xi) = e + \xi e_{ii}$, $\xi \in R$, $i \neq j$, называются **элементарными трансвекциями**. Часто эпитет 'элементарные' опускают и говорят просто о трансвекциях. Ясно, что элементарные трансвекции с фиксированными (i,j) обладают следующим свойством аддитивности:

$$t_{ij}(\xi+\zeta) = t_{ij}(\xi)t_{ij}(\zeta).$$

В этом проще всего убедиться при помощи определения. В самом деле,

$$t_{ij}(\xi)t_{ij}(\zeta) = (e+\xi e_{ij})(e+\zeta e_{ij}) = e + \xi e_{ij} + \zeta e_{ij} + \xi\zeta e_{ij}e_{ij}.$$

Так как $i \neq j$, то $e_{ij}^2 = 0$, и мы получаем требуемую формулу. В частности, мы видим, что все элементарные трансвекции обратимы, $t_{ij}(\xi) = t_{ij}(-\xi)$. Таким образом, отображение $t_{ij} : R^+ \longrightarrow \mathrm{GL}(n,R)$, $\xi \mapsto t_{ij}(\xi)$, представляет собой гомоморфизм. Образ этого гомоморфизма обозначается

$$X_{ij} = \{t_{ij}(\xi),\ \xi \in R\}$$

и называется **(элементарной) корневой подгруппой** в $\mathrm{GL}(n,R)$.

2. Коммутационная формула Шевалле. Самое важное свойство элементарных трансвекций состоит в следующем.

$$[t_{ij}(\xi), t_{hk}(\zeta)] = \begin{cases} e & \text{если } i \neq k, j \neq h; \\ t_{ik}(\xi\zeta) & \text{если } i \neq k, j = h; \\ t_{hj}(-\zeta\xi) & \text{если } i = k, j \neq h. \end{cases}$$

В справедливости этой формулы проще всего убедиться произведя вычисления непосредственно по определению элементарных трансвекций в терминах стандартных матричных единиц. А именно, по определению

$$[t_{ij}(\xi), t_{hk}(\zeta)] = (e + \xi e_{ij})(e + \zeta e_{hk})(e - \xi e_{ij})(e - \zeta e_{hk})$$

3. Элементарная группа. Группа $E(n, R)$, порожденная всеми **элементарными трансвекциями** называется элементарной группой

$$E(n, R) = \langle t_{ij}(\xi) \mid \xi \in R, 1 \leq i \neq j \leq \rangle.$$

4. Элементарные псевдоотражения. Матрицы вида $d_i(\varepsilon) = e + (\varepsilon - 1)e_{ii}$, где $\varepsilon \in R^*$, называются **элементарными псевдоотражениями**.

$$d_i(\varepsilon) = \mathrm{diag}(1, \dots, 1, \varepsilon, 1 \dots, 1),$$

где ε стоит на i-м месте.

Группа $\mathrm{GE}(n, R)$, порожденная всеми элементарными матрицами.

§ 15. Элементарные преобразования матриц

Элементарным преобразованием матрицы называется умножение этой матрицы справа или слева на одну из приведенных в предыдущем параграфе элементарных матриц.

§ 16. Тождества для произведения

1. Ассоциативность и дистрибутивность. Умножение матриц удовлетворяет двум важнейшим тождествам, оно ассоциативно и дистрибутивно относительно сложения.

Теорема. *Умножение матриц ассоциативно в том смысле, что если определено одно из произведений $(xy)z$ и $x(yz)$, то определено и второе и они равны между собой: $(xy)z = x(yz)$.*

Доказательство. Чтобы было определено первое произведение, необходимо, чтобы было определено произведение xy, и произведение $(xy)z$, для этого число столбцов x должно равняться числу строк y, а число столбцов xy (равное числу столбцов y) должно равняться числу строк z. Тем самым, для некоторых m, n, p, q имеем $x \in M(m, n, R)$, $y \in M(n, p, R)$, $z \in M(p, q, R)$. Ясно, что при этом произведение $x(yz)$ также будет определено и — так же как и произведение $(xy)z$ — имеет размер $m \times q$. Поэтому по определению равенства матриц нам осталось лишь убедиться в том, что матричные элементы двух этих произведений на соответствующих местах совпадают.

В самом деле, фиксируем какие-то индексы $1 \leq i \leq m$, $1 \leq j \leq q$, и рассмотрим элемент $((xy)z)_{ij}$. Непосредственное вычисление, использущее изменение порядка суммирования (которое мы подробно обсуждали в связи с ассоциативностью умножения многочленов!), показывает, что

$$\begin{aligned}((xy)z)_{ij} = \sum_{k=1}^{p}(xy)_{ik}z_{kj} = \sum_{k=1}^{p}\sum_{h=1}^{n} x_{ih}y_{hk}z_{kj} = \\ \sum_{h=1}^{n}\sum_{k=1}^{p} x_{ih}y_{hk}z_{kj} = \sum_{h=1}^{n} x_{ih}(yz)_{hj} = (x(yz))_{ij}.\end{aligned}$$

что и требовалось доказать.

Не представляет труда доказать и следующее утверждение, устанавливающее дистрибутивность умножения матриц относительно сложения, в случае, когда соответствующие суммы и произведения определены.

Предложение. *Предположим, что* $x, y \in M(m, n, R)$, *а* $z, w \in M(n, p, R)$. *Тогда* $(x + y)z = xz + yz$ *и* $x(z + w) = xz + xw$.

Доказательство. Стандартное вычисление, использующее свойства операций в R.

2. Некоммутативность умножения матриц. Сразу отметим, что умножение матриц некоммутативно (это одно из первых некоммутативных умножений, появившихся в математике, оно было определено Кэли через два года после открытия кватернионов Гамильтоном). Оно некоммутативно по крайнем мере по трем следующим причинам:

• Во-первых, если произведение матриц в одном порядке определено, отсюда не следует, что и произведение в обратном порядке тоже определено. Например, если $x \in M(m, n, R)$, $y \in M(n, p, R)$, то произведение xy определено, но произведение yx определено только в случае, когда $m = p$.

• Во-вторых, даже если оба произведения определены, они могут иметь разный размер: если $x \in M(m, n, R)$, $y \in M(n, m, R)$, то как xy, так и yx определены, но при этом $xy \in M(m, R)$, в то время как $yx \in M(n, R)$. Например, как мы уже знаем, произведение строки длины n на столбец высоты n есть скаляр, но произведение того же столбца на ту же строку есть квадратная матрица порядка n.

• В третьих, даже если оба произведение xy и yx определены и имеют одинаковый порядок, то они совершенно не обязательно равны даже если умножение в самом кольце R коммутативно. В этом легко убедиться на простейших примерах, рассматривая матрицы порядка 2 над целыми числами. Скажем, если

$$e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

стандартные матричные единицы, то $e_{11}e_{12} = e_{12}$, в то время как $e_{12}e_{11} = 0$.

Замечание. Последний пример показывает еще одно необычное свойство умножения матриц: произведение двух ненулевых матриц может быть нулевым. Поэтому, когда мы введем в множестве квадратных матриц $M(n, R)$, $n \geq 2$, структуру кольца, это будет кольцо с делителями нуля.

Комментарий. Некоммутативность умножения матриц должна учитываться во всех вычислениях. Например, часто при операциях с матрицами начинающие используют обычные формулы сокращенного умножения, типа $x^2 - y^2 = (x - y)(x + y)$. Однако, как показывают простейшие примеры, делать этого нельзя! Достаточно взять здесь $x = e_{12}$, $y = e_{21}$. Это относится ко всем обычным формулам. Скажем, в математическом анализе известна формула $(1/f)' = f'/f^2$, но для матриц эта формула принимает другой — гораздо более естественный!!! — вид. А именно, если $f = (f_{ij})$, то $(f^{-1})' = f^{-1}f'f^{-1}$, где $g' = (f'_{ij})$.

§ 17. Вычислительная перспектива: алгоритм Штрассена

1. Алгоритм Штрассена для матриц степени 2. Рассмотрим умножение 2×2 матриц над ассоциативным кольцом R. Для дальнейшего весьма существенно, что кольцо R не предполагается коммутативным! Пусть $a, b \in M(n, R)$ и $c = ab$:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}.$$

где $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j}$, что требует **восемь** умножений. На первый взгляд представляется, что все эти умножения необходимы и, тем самым, число 8 здесь невозможно уменьшить. Сейчас произойдет нечто совершенно феерическое[72,73]. Ю.И.Манин заметил, что следующая лемма является одним из 10 самых замечательных математических *изобретений* XX века.

Лемма Штрассена. *Произведение двух* 2×2 *матриц можно вычислить произведя* **семь** *умножений и 18 сложений в кольце* R.

Доказательство. Вычислим

$$\begin{aligned} d_1 &= (a_{12} - a_{22})(b_{21} + b_{22}), \\ d_2 &= (a_{11} + a_{22})(b_{11} + b_{22}), \\ d_3 &= (a_{11} - a_{21})(b_{11} + b_{12}), \\ d_4 &= (a_{11} + a_{12})b_{22}, \\ d_5 &= a_{11}(b_{12} - b_{22}), \\ d_6 &= a_{22}(b_{21} - b_{11}), \\ d_7 &= (a_{21} + a_{22})b_{11}, \end{aligned}$$

Теперь матрица $c = ab$ вычисляется как

$$c = \begin{pmatrix} d_1 + d_2 - d_4 + d_6 & d_4 + d_5 \\ d_6 + d_7 & d_2 - d_3 + d_5 - d_7 \end{pmatrix}.$$

Из этой леммы легко вывести, что умножение двух матриц требует не более $O(n^{\log_2(7)})$ операций.

Теорема Штрассена. *Чтобы вычислить произведение двух* $2^n \times 2^n$ *матриц, достаточно произвести не более* 7^n *умножений в кольце* R.

Доказательство. Индукция по n. Для $n = 1$ утверждение уже доказаною Предположим, что уже известно, что умножение двух $2^n \times 2^n$ матриц требует 7^n умножений, покажем, что тогда перемножение двух $2^{n+1} \times 2^{n+1}$ матриц требует 7^{n+1} умножений. В самом деле, воспользуемся изоморфизмом $M(2^{n+1}, R) \cong M(2, M(2^n, R))$ и рассмотрим матрицы степени 2^{n+1} как блочные 2×2 матрицы с блоками размера $2^n \times 2^n$. Тогда по лемме умножение таких матриц требует 7 умножений блоков $2^n \times 2^n$, а по индукционному предположению каждое умножение таких блоков можно осуществить за 7^n умножений в кольце R. Таким образом, всего мы произвели $7 \cdot 7^n = 7^{n+1}$ умножение в R.

По порядку оценка Штрассена дает $n^{2,807}$. Около 10 лет эта оценка оставалась непревзойденной, но потом были получены многочисленные улучшения.

---

[72]Ф.Штрассен, Алгоритм Гаусса не оптимален. — Киберн. сб., нов. серия, вып. 7. — Мир., М., 1970, с.67–70.

[73]См., также А.Ахо, Дж.Хопкрофт, Дж,Ульман, Построение и анализ вычислительных алгоритмов. — Мир, М., 1979, с.1–536.

§ 18. Вычислительная перспектива: порядок умножений

Ассоциативно ли умножение матриц над ассоциативным кольцом? Смешной вопрос — мы же ДОКАЗАЛИ, что умножение матриц ассоциативно!! Да, но мы доказали это с точки зрения *математика*. Сейчас мы ДОКАЖЕМ, что с точки зрения *вычислителя* умножение (неквадратных) матриц настолько неассоциативно, насколько это можно себе представить[74,75,76].

Предположим, что для умножения матрицы $x \in M(m,n,R)$ на матрицу $y \in M(n,p,R)$ требуется mnp умножений. В действительности, как мы знаем, эта оценка заведомо является завышенной, но это не влияет на наш окончательный результат. Рассмотрим произведение трех матриц xyz, где $z \in M(p,q,R)$, и оценим количество умножений, требуемое для вычисления этого произведения в одном и в другом порядке.

∘ Вначале вычислим $(xy)z$. Вычисление произведения xy требует mnp умножений в кольце R и дает матрицу размера mp. Теперь для вычисления $(xy)z$ нужно проделать еще mpq умножений. Итого, для вычисления $(xy)z$ нам понадобилось $mnp + mpq = mp(n+q)$ умножений.

∘ С другой стороны, вычисление $x(yz)$ начинается с вычисления yz, которое требует npq умножений в кольце R и дает матрицу размера nq. Теперь для вычисления $x(yz)$ нужно проделать еще mnq умножений. Таким образом, для вычисления $x(yz)$ нам понадобилось $npq + mnq = (m+p)nq$ умножений.

Ясно, что, вообще говоря, $mp(n+q) \neq (m+p)nq$. Легко видеть, что правильная стратегия для минимизации количества умножений состоит в том, чтобы на каждом шаге получать матрицу самого маленького возможного размера. Например, вычисление произведения случайных матриц размеров $x \in M(1000,1000,\mathbb{R})$, $y \in M(1000,10,\mathbb{R})$, $z \in M(10,1000,\mathbb{R})$ в порядке $x(yz)$ требует почти в 50 раз больше умножений, чем вычисление произведения $(xy)z$, а именно $1.01 \cdot 10^9$ против $2 \cdot 10^7$. Фактически компьютерный эксперимент показывает, что вычисление произведения $x(yz)$ занимает в среднем примерно в 20 раз больше времени, чем вычисление произведения $(xy)z$, видимо потому, что при вычислении произведения матриц проделываются не только умножения, но и другие операции. Понятно, что для матриц большего размера эта разница становится еще более заметной. Для нескольких матриц, некоторые из размеров которых имеют порядок 10^6, а некоторые другие малы, вычисление произведения в одном порядке производится за секунды, а в другом требует нескольких суток. Тем самым, с вычислительной точки зрения расстановка скобок в произведении матриц исключительно важна!!!

Неассоциативность произведения с точки зрения сложности вычислений хорошо видна уже при изучении действия *квадратных* матриц на столбцах. В самом деле, пусть $h, g \in M(n,R)$, а $u \in R^n$. Тогда вычисление произведения gu требует n^2 умножений (в данном случае оценка неулучшаема, при умножении столбца на матрицу над коммутативным кольцом R никак не обойтись меньшим количеством умножений в R!). Таким образом, вычисление $h(gu)$ требует $2n^2$ умножений. А вот для наивного вычисления $(hg)u$ нужно $n^3 + n^2$

[74] S.S.Godbole, On efficient calculation of matrix chain products. — IEEE Trans. Comput., Ser. C, 1973, vol.22, N.9, p.864–966.

[75] Y.Muraoka, D.J.Kuck, On the time required for a sequence of matrix products. — Comm. Assoc. Comp. Mach., 1973, vol.16, N.1, p.22–26

[76] А.Ахо, Дж.Хопкрофт, Дж.Ульман, ibid., с.83–85.

умножений. Как мы знаем, чуть более изощренный способ действия позволяет понизить оценку n^3, но ведь не до n^2. Это соображение очень полезно, если нам нужно вычислить один элемент произведения m матриц порядка n в случае, когда число $m << n$. Мы только что убедились, что для этого достаточно $O(n^2)$ умножений, а вовсе не $O(n^3)$ и даже не $O(n^{2,701})$.

§ 18. Алгебра квадратных матриц

Резюмируем свойства операций над матрицами, о которых шла речь выше, в частном случае квадратных $n \times n$-матриц над ассоциативным кольцом с 1.

1. Кольцо квадратных матриц. Отметим, прежде всего, что две квадратных матрицы одинакового размера всегда можно перемножить.

M1 Ассоциативность: $\forall x, y, z \in M(n, R)$, $(x + y) + z = x + (y + z)$;

M2 Существование единицы: $\exists e \in M(n, R)$, $\forall x \in M$, $ex = x = xe$;

А именно, единицей относительно умножения, как раз и является единичная матрица порядка n. Объединяя эти свойства с аксиомами A1–A4, D1, D2, мы можем резюмировать.

Теорема. *Для ассоциативного кольца R с 1 множество $M(n, R)$ квадратных матриц является ассоциативным кольцом с 1 относительно операций сложения и умножения матриц.*

В то же время, как мы уже заметили, при любом $n \geq 2$ это кольцо некоммутативно и имеет делители нуля. Более того, это кольцо имеет нильпотенты, скажем, для любых $i \neq j$ имеем $e_{ij}^2 = 0$.

Если кольцо R коммутативно, то выполняется еще следующая аксиома V5, показывающая, что $M(n, R)$ в действительности является алгеброй над R.

V5 $\forall x, y \in M(n, R)$, $\forall \alpha \in R$, $(\alpha x)y = \alpha(xy) = x(\alpha y)$.

Пусть $\phi : R \longrightarrow S$ — гомоморфизм колец. Тогда $\phi : M(n, R) \longrightarrow M(n, S)$, $(x_{ij}) \mapsto (\phi(x_{ij}))$.

§ 19. Обратимые матрицы и полная линейная группа

Мы уже видели формулу для обращения матриц 2×2 над коммутативным кольцом R в главе 1, напомним ее:

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix}^{-1} = \frac{1}{xw - yz} \begin{pmatrix} w & -y \\ -z & x \end{pmatrix}.$$

Эта формула обобщается и на некоммутативный случай, но, конечно, ответ получается несколько менее наглядным. Следующая задача взята из книги[77]. Естественно, там рассматривается только случай, когда $R = M(n, A)$ для некоторого коммутативного кольца A, но, конечно, ничего не меняется и в общем случае.

Задача. Предположим, что $x, y, z, w, x - yw^{-1}z \in R^*$. Покажите, что тогда $y - xz^{-1}w, z - wy^{-1}x, w - zx^{-1}y \in R^*$ и

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix}^{-1} = \begin{pmatrix} (x - yw^{-1}z)^{-1} & (z - wy^{-1}x)^{-1} \\ (y - xz^{-1}w)^{-1} & (w - zx^{-1}y)^{-1} \end{pmatrix}$$

[77] Р.Беллман, Введение в теорию матриц. — Наука, М., 1976, с.1–351, стр.114. Заметим, что у Беллмана пропущено условие $x - yw^{-1}z \in \mathrm{GL}(n, R)$.

В действительности, приведенную в этой задаче формулу можно усовершенствовать, так как здесь вовсе не обязательно требовать, чтобы элементы y, z, w тоже были обратимыми. Правда при этом придется пожертвовать симметрией между x, y, z, w. Положим $w - zx^{-1}y$. Тогда полученную в этой задаче формулу можно переписать в виде

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix}^{-1} = \begin{pmatrix} x^{-1} + x^{-1}y(w - zx^{-1}y)^{-1}zx^{-1} & -x^{-1}y(w - zx^{-1}y)^{-1} \\ -(w - zx^{-1}y)^{-1}zx^{-1} & (w - zx^{-1}y)^{-1} \end{pmatrix}$$

Эта формула называется **формулой Фробениуса**[78].

Задача. Запишите аналоги формулы Фробениуса для случая, когда $y \in R^*$, $z \in R^*$ или $w \in R^*$.

А вот формула для обращения матрицы $x = (x_{ij})$ размера 3×3 над коммутативным кольцом:

$$x^{-1} = \frac{1}{\det(x)} \begin{pmatrix} x_{22}x_{33} - x_{23}x_{32} & -x_{12}x_{33} + x_{13}x_{32} & x_{12}x_{23} - x_{13}x_{22} \\ -x_{21}x_{33} + x_{23}x_{31} & x_{11}x_{33} - x_{13}x_{31} & -x_{11}x_{23} + x_{13}x_{21} \\ x_{21}x_{32} - x_{22}x_{31} & -x_{11}x_{32} + x_{12}x_{31} & x_{11}x_{22} - x_{12}x_{21} \end{pmatrix}$$

2. Полная линейная группа. Группа $M(n, R)^*$ обратимых элементов кольца R называется **полной линейной группой** степени n над кольцом R и обозначается $\mathrm{GL}(n, R)$.

§ 20. Транспонирование и контраградиент

В настоящем параграфе мы рассмотрим унарную операцию на матрицах, по существу состоящую в том, что мы переворачиваем матрицу относительно главной диагонали. Однако (за исключением случая, когда кольцо R коммутативно!) такая операция не обладает хорошими алгебраическими свойствами, так что нам придется слегка извернуться. Два обычных способа сделать эту операцию антиизоморфизмом относительно умножения матриц — это транспонирование и рассматриваемое в следующем параграфе эрмитово сопряжение.

1. Транспонирование. Пусть R^o — кольцо, противоположное к R, а $\alpha \mapsto \alpha^o$ — канонический антиизоморфизм $R \longrightarrow R^o$. Отображение $\mathrm{trans} : M(m, n, R) \longrightarrow M(n, m, R^o)$ сопоставляющее матрице $x \in M(m, n, R)$ матрицу x^t, у которой в позиции (i, j), где $1 \leq i \leq n$, $1 \leq j \leq m$ стоит $(x^t)_{ij} = (x_{ji})^o$, называется **транспонированием**. Следующие свойства сразу вытекают из определения.

Лемма. *Транспонирование*

1) инволютивно: $(x^t)^t = x$;

2) аддитивно: $(x + y)^t = x^t + y^t$;

3) полуоднородно относительно умножения на скаляры: $(\alpha x)^t = x^t\alpha^o$ *и* $(x\alpha)^t = \alpha^o x^t$.

Проверить аналогичное свойство для умножения матриц несколько сложнее.

Лемма. *Пусть* $x \in M(l, m, R)$, $y \in M(m, n, R)$. *Тогда* $(xy)^t = y^t x^t$.

Доказательство. Заметим, что $y^t \in M(n, m, R^o)$, $x^t \in M(m, l, R^o)$, так что произведение в правой части определено и принадлежит $M(n, l, R^o)$ вместе с

[78] Ф.Р.Гантмахер, Теория матриц. Наука, М., 1967, с.1–575. стр.59–62.

$(xy)^o$. Таким образом, чтобы проверить совпадение этих матриц, нам нужно лишь убедиться в том, что все их соответствующим матричные элементы совпадают. В самом деле,

$$((xy)^t)_{ij} = ((xy)_{ji})^o = \sum (x_{jh}y_{hi})^o = \sum y^o_{hi}x^o_{jh} = \sum (y^t)_{ih}(x^t)_{hj} = (y^t x^t)_{ij},$$

где $1 \le i \le l$, $1 \le j \le n$, а все суммы берутся по $1 \le h \le m$.

Теорема. *Транспонирование $x \mapsto x^t$ является антиизоморфизмом кольца $M(n, R)$ на $M(n, R^o)$.*

Следствие 1. *Для любого кольца $M(n, R)^o \cong M(n, R^o)$.*

Доказательство. В самом деле, по теореме $M(n, R) \cong M(n, R^o)^o$, осталось еще раз перейти к противоположным кольцам.

Пусть теперь R коммутативно. В этом случае мы можем отождествить R^o с R и рассматривать транспонирование как отображение из $M(m, n, R)$ в $M(n, m, R)$.

Следствие 2. *Для коммутативного кольца R транспонирование $x \mapsto x^t$ является антиавтоморфизмом кольца $M(n, R)$. В частности, в этом случае $M(n, R)^o \cong M(n, R)$.*

2. Контраградиент. Посмотрим, что транспонирование означает с точки зрения обратимых матриц.

Следствие 3. *Если $x \in \mathrm{GL}(n, R)$ имеем $(x^{-1})^t = (x^t)^{-1}$. В частности, для любого кольца R имеет место изоморфизм $\mathrm{GL}(n, R^o) \cong \mathrm{GL}(n, R)$.*

Доказательство. Первое утверждение сразу вытекает из теоремы (так как антиавтоморфизм перевожит e в e, он автоматически переводит обратные в обратные). Второе утверждение получается теперь из того, что каждая группа антиизоморфна себе посредством антиавтоморфизма $\mathrm{inv} : x \mapsto x^{-1}$, композиция двух антиизоморфизмов является изоморфизмом.

Получающийся при этом изоморфизм

$$\mathrm{trans} \circ \mathrm{inv} : \mathrm{GL}(n, R) \longrightarrow \mathrm{GL}(n, R)^o \longrightarrow \mathrm{GL}(n, R^o)$$

называется **контраградиентом**. Образ $x \in \mathrm{GL}(n, R)$ относительно контраградиента обозначается $x^* = (x^{-1})^t = (x^t)^{-1}$ и называется матрицей, **контраградиентной** к x. Как было отмечено в доказательстве последнего следствия, $(xy)^* = x^*y^*$. В случае коммутативного кольца R контраградиент $x \mapsto x^*$ является автоморфизмом $\mathrm{GL}(n, R)$.

§ 21. Эрмитово сопряжение

В этом параграфе мы определим антиавтоморфизм $M(n, R)$ впредположении, что

1. Антиавтоморфизмы кольца R. Пусть теперь $\overline{\ } : R \longrightarrow R$ — антиавтоморфизм кольца R. Иными словами, предполагается, что $\overline{\alpha + \beta} = \overline{\alpha} + \overline{\beta}$ и $\overline{\alpha\beta} = \overline{\beta}\,\overline{\alpha}$. Как правило (но не всегда!) этот антиавтоморфизм предполагается инволюцией, т.е. $\overline{\overline{\alpha}} = \alpha$. Вот два типичных примера:

• $R = K[G]$ — групповая алгебра группы G, а $^- : \sum a_g g \mapsto \sum a_g g^{-1}$;

• A — коммутативное кольцо, $R = M(n, A)$, а $^- : x \longrightarrow x^t$ — транспонирование.

В приложениях за пределами алгебры чаще всего возникает следующая ситуация

• $R = \mathbb{C}$ — поле комплексных чисел, $z \longrightarrow \overline{z}$ — комплексное сопряжение.

2. Эрмитово сопряжение. Теперь мы можем определить **эрмитово сопряжение** $M(m, n, R) \longrightarrow M(n, m, R)$, сопоставляющее x матрицу $x^\dagger$, для которой $(x^\dagger)_{ij} = (\overline{x})_{ji}$.

Задача. Проверьте следующие свойства эрмитова сопряжения:

1) аддитивность: $(x + y)^\dagger = x^\dagger + y^\dagger$,

2) полуоднородность относительно умножения на скаляры $(\alpha x)^\dagger = x^\dagger \overline{\alpha}$ и $(x\alpha)^\dagger = \overline{\alpha} x^\dagger$.

3) Если $^-$ является инволюцией, то $(x^\dagger)^\dagger = x$.

Задача. Докажите, что $(xy)^\dagger = y^\dagger x^\dagger$.

Решение. В самом деле,

$$(xy)^\dagger_{ij} = \overline{(xy)_{ji}} = \sum \overline{x_{jh} y_{hi}} = \sum \overline{y}_{hi}\, \overline{x}_{jh} = \sum y^\dagger_{ih} x^\dagger_{hj} = (y^\dagger x^\dagger)_{ij}.$$

Резюмируя эти свойства, мы получаем следующий результат.

Теорема. *Эрмитово сопряжение $x \mapsto x^\dagger$ есть антиавтоморфизм кольца $M(n, R)$.*

§ 22. Вычисления с матрицами

— Ты меня удивляешь, Лев, — объявляет он. — И все вы меня удивляете. Неужели вам здесь не надоело?
— Мы работаем, — возражаю я лениво.
— Зачем работать без всякого смысла?
— Почему же — без смысла? Ты же видишь, сколько мы узнали всего за один день.
— Вот я и спрашиваю: зачем вам узнавать то, что не имеет смысла? Что вы будете с этим делать? Вы все узнаёте и узнаёте и ничего не делаете с тем, что узнаёте.

Аркадий Стругацкий, Борис Стругацкий, 'Жук в муравейнике'

Сейчас мы произведем несколько типичных вычислений в кольце $M(n, R)$. Мы сделаем это двумя способами: вначале наивным, основанным на вычислениях с матрицами, так, как подобные вещи считает новичок или математик-неспециалист, а потом бесхитростным, так как такие вещи считает алгебраист ('Алгебра — это искусство избегать вычисления').

1. Числа Фибоначчи. Следующая задача часто возникает в самых разных областях математики.

Задача. Найти последовательные степени матрицы

$$x = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Наивное решение. Вычислив несколько первых степеней

$$x^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad x^3 = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}, \qquad x^4 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix},$$

легко догадаться, что элементами матриц x^n будут числа Фибоначчи:

$$x^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

В чем легко убедиться посмотрев на рекуррентное соотношение.

Бесхитростное решение. Легко видеть, что матрица x удовлетворяет уравнению $x^2 - x - e = 0$.

Задача. Вычислите последовательные степени матрицы $\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}^n$.

Задача. Вычислите последовательные степени матрицы $\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}^n$.

2. Биномиальные коэффициенты. А вот еще одна матрица, степени которой дадут нам знакомые числа.

Задача. Найти последовательные степени матрицы

$$x = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Обобщить этот пример.

Наивное решение. Пусть $x = e + e_{12} + \ldots + e_{n-1,n}$. Умножение на x справа состоит в прибавлении каждого из столбцов с номерами $1, \ldots, n-1$ к следующему. Таким образом, $x_{ij}^m = x_{ij}^{m-1} + x_{i,j-1}^{m-1}$. Но это в точности треугольное рекуррентное соотношение для биномиальных коэффициентов. Таким образом, первая строка матрицы x^m имеет вид $(1, C_m^1, C_m^2, \ldots, C_m^{n-1})$, а остальные получаются из нее последовательным применением `ShiftRight`

Бесхитростное решение. А чем же еще быть коэффициентам бинома, как не биномиальными коэффициентами? Степени матрицы x коммутируют, поэтому по сути речь идет о вычислениях в *коммутативном* кольце $R[x]$, порожденном e и степенями матрицы x. Матрица $y = x - e = e_{12} + \ldots + e_{n-1,n}$ это то, что называется `ShiftRight`, сдвиг вправо на одну позицию. Ясно, что y^m это сдвиг вправо на m позиций, т.е. теплицева матрица с единицами на m-й наддиагонали и нулями во всех остальных местах. В частности, $y^n = 0$, так что $R[x] = R[y] \cong R[t]/(t^n)$ есть кольцо усеченных многочленов. Это значит, что

$$x^m = (e+y)^m = e + C_m^1 y + C_m^2 y^2 + \ldots + C_m^{n-1} y^{n-1}.$$

Задача. Теперь, когда понятно, как *на самом деле* проводятся такие вычисления, посчитайте m-ю степень матрицы

$$z = e + y + y^2 + \ldots + y^{n-1} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & 1 & \ldots & 1 \\ & & \ldots & \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Ответ. $z^m = e + C_m^1 y + C_{m+1}^2 y^2 + \ldots + C_{m+n-2}^{n-1} y^{n-1}$.

§ 23. Простейшие примеры подколец в $M(n, R)$

В этом параграфе мы определим наиболее известные классы матриц и образующие подкольца в кольце матриц $M(n, R)$. Через R здесь обозначается произвольное ассоциативное кольцо с 1.

• **Диагональные матрицы.** Квадратная матрица $x = (x_{ij}) \in M(n, R)$ называется **диагональной**, если все ее элементы вне главной диагонали равны 0, т.е. $x_{ij} = 0$ для всех $i \neq j$. Диагональная матрица с диагональными элементами $\lambda_1, \dots, \lambda_n \in R$ обозначается через

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}.$$

Например, в этих обозначениях $e = \operatorname{diag}(1, \dots, 1)$.

Очевидно, что множество $\Delta(n, R)$ всех диагональных матриц образует подкольцо в $M(n, R)$, изоморфное прямой сумме n экземпляров основного кольца $R \oplus \dots \oplus R$, причем этот изоморфизм как раз и устанавливается отображением diag:

$$\operatorname{diag}(\lambda_1 + \mu_1, \dots, \lambda_n + \mu_n) = \operatorname{diag}(\lambda_1, \dots, \lambda_n) + \operatorname{diag}(\mu_1, \dots, \mu_n),$$
$$\operatorname{diag}(\lambda_1\mu_1, \dots, \lambda_n\mu_n) = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \operatorname{diag}(\mu_1, \dots, \mu_n).$$

• **Скалярные матрицы.** Диагональная матрица вида называется **скалярной**, если все ее диагональные элементы равны между собой. Иными словами, скалярная матрица имеет вид $\lambda e = \operatorname{diag}(\lambda, \dots, \lambda)$. Скалярные матрицы образуют подкольцо в $M(n, R)$ изоморфное кольцу R. Часто термин скалярная матрица резервируют только для таких матриц λe, у которых $\lambda \in \operatorname{Cent}(R)$. При таком определении множество всех скалярных матриц в точности совпадает с центром кольца $M(n, R)$, изоморфным центру кольца R.

• **Треугольные матрицы.** Квадратная матрица $x = (x_{ij}) \in M(n, R)$ называется **верхней треугольной**, если все ее элементы *ниже* главной диагонали равны 0, т.е. $x_{ij} = 0$ для всех $i > j$. Аналогично, x называется **нижней треугольной**, если все ее элементы *выше* главной диагонали равны 0, т.е. если $x_{ij} = 0$ для всех $i < j$. Как верхние, так и нижние треугольные матрицы образуют подкольца в $M(n, R)$, которые будут обозначаться через $\mathfrak{b}(n, R)$ и $\mathfrak{b}^-(n, R)$, соответственно.

• **Строго треугольные матрицы.** Квадратная матрица $x = (x_{ij}) \in M(n, R)$ называется **строго верхней треугольной**, если все ее элементы ниже главной диагонали *и на ней* равны 0, т.е. $a_{ij} = 0$ для всех $i \geq j$. Аналогично, x называется **строго нижней треугольной**, если все ее элементы выше главной диагонали *и на ней* равны 0, т.е. если $x_{ij} = 0$ для всех $i \leq j$. Как верхние, так и нижние строго треугольные матрицы образуют подкольца *без единицы* в $M(n, R)$. Обозначим множество всех верхних строго треугольных матриц в $M(n, R)$ через $\mathfrak{n}(n, K)$, а множество всех нижних строго треугольных матриц — через $\mathfrak{n}^-(n, K)$.

Задача. Докажите, что $\mathfrak{n}(n, R)$ — идеал в $\mathfrak{b}(n, K)$ и $\mathfrak{b}(n, R)/\mathfrak{n}(n, R) \cong \Delta(n, R)$.

• Более общо, пусть $\mathfrak{n}_r(n, R)$, $0 \leq r \leq n - 1$, — множество всех матриц, у которых все элементы под главной диагональю, на главной диагонали и еще на

$r-1$ наддиагоналях равны 0, в то время как оставшиеся элементы произвольные. В частности, $\mathfrak{n}_0(n,R)=\mathfrak{b}(n,R)$, $\mathfrak{n}_1(n,R)=\mathfrak{n}(n,R)$, так что этот пример является обобщением двух предыдущих примеров. Все эти множества, а также симметричные им множества $\mathfrak{n}_r^-(n,R)$, получающиеся заменой наддиагоналей на поддиагонали, образуют подкольца *без единицы* в $M(n,R)$.

Задача. Пусть R — кольцо с 1. Докажите, что $\mathfrak{n}(n,R)^r=\mathfrak{n}_r(n,R)$.

§ 24. Матрицы типа циркулянтов

Во многих задачах, как в самой алгебре, так и за ее пределами, естественно возникают матрицы, у которых x_{ij} зависит только от вычета $i-j$ по модулю n и их обобщения. Например, такие матрицы естественно появляются при изучении куммеровых расширений по отношению к степенному базису, в теории конченых групп, и т.д.

• **Циркулянты.** Матрица $x=(x_{ij})\in M(n,R)$ вида

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \dots & \dots & \dots & \dots \\ a_2 & a_3 & \dots & a_1 \end{pmatrix}$$

называется циркулянтом. Иными словами, x_{ij} зависит не от самих индексов i,j, а только от вычета их разности $i-j$ по модулю n. Легко проверить, что циркулянты образуют *коммутативное* подкольцо в $M(n,R)$, а обратимые циркулянты — подгруппу в $\mathrm{GL}(n,R)$. Циркулянты являются частным случаем теплицевых матриц, которые мы рассмотрим в следующем параграфе.

• **Антициркулянты.** Матрица $x=(x_{ij})\in M(n,R)$ вида

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ -a_n & a_1 & \dots & a_{n-1} \\ \dots & \dots & \dots & \dots \\ -a_2 & -a_3 & \dots & a_1 \end{pmatrix}$$

называется антициркулянтом. Иными словами, $(-1)^{\mathrm{sign}(j-i)}x_{ij}$ зависит не от самих индексов i,j, а только от вычета их разности $i-j$ по модулю n. Антициркулянты также образуют подкольцо в $M(n,R)$, а обратимые антициркулянты — подгруппу в $\mathrm{GL}(n,R)$.

• **Квазициркулянты.** Вообще, пусть d — произвольный элемент кольца R. Матрица $x=(x_{ij})\in M(n,R)$ вида

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ da_n & a_1 & \dots & a_{n-1} \\ \dots & \dots & \dots & \dots \\ da_2 & da_3 & \dots & a_1 \end{pmatrix}$$

называется квазициркулянтом. Как циркулянты, так и антициркулянты являются частными случаями квазициркулянтов, получающимися при $d=1$ и $d=-1$, соответственно.

§ 25. Теплицевы и ганкелевы матрицы

Сейчас мы введем два важных класса матриц, которые изучаются с начала XX века[79,80,81].

• **Перъдиагональные матрицы.** Матрица $x = (x_{ij})$ называется **перъдиагональной**, если $x_{ij} = 0$ для всех пар i, j, кроме $i + j = n + 1$.

Следующие классы матриц естественно возникают в различных вопросах анализа, в частности, в проблеме моментов.

• **Теплицевы матрицы.** Матрица $x = (x_{ij})$ называется **теплицевой** матрицей, если ее элемент x_{ij} в позиции (i, j) зависит не от самих i и j, а только от их разности $i-j$. Иными словами, для любого $h = -(n-1), \dots, -1, 0, 1, \dots, n-1$ определен элемент $x_h \in K$ и $x_{ij} = x_{i-j}$. Теплицевы матрицы образуют подмодуль в $M(n, R)$ содержащийся в модуле персимметричных матриц.

Задача. Будет ли множество теплицевых матриц подкольцом в $M(n, R)$?

Задача. Докажите, что пересечения множества теплицевых матриц с кольцами верхних/нижних треугольных матриц являются коммутативными подкольцами в $M(n, R)$.

• **Ганкелевы матрицы.** Матрица $x = (x_{ij})$ называется **ганкелевой** матрицей, если ее элемент x_{ij} в позиции (i, j) зависит не от самих i и j, а только от их разности $i + j$. Иными словами, для любого $h = 1, \dots, 2n$ определен элемент $x_h \in K$ и $x_{ij} = x_{i+j}$. Ганкелевы матрицы образуют подмодуль в $M(n, R)$ содержащийся в модуле симметрических матриц.

• **Ленточные матрицы.** Матрица $x = (x_{ij})$ называется **ленточной**, если $x_{ij} = 0$ для всех (i, j) таких, что $-l \leq i - j \leq m$ для некоторых $l, m \in \mathbb{N}_0$.

Фактически в приложениях особенно часто встречается следующий частный случай.

• **Трехдиагональные матрицы.** Матрица $x = (x_{ij})$ называется **трехдиагональной**, если $x_{ij} = 0$ для всех (i, j) таких, что $|i - j| \geq 2$.

• **Хессенберговы матрицы.** Матрица $x = (x_{ij})$ называется **хессенберговой**, если $x_{ij} = 0$ для всех (i, j) таких, что $i > j + 1$.

§ 26. Симметричные матрицы

• **Симметричные матрицы.** Матрица $x = (x_{ij})$ называется **симметричной**, если $x_{ij} = x_{ji}$ для всех i, j. В случае коммутативного кольца матрица x в том и только том случае симметрична, когда $x^t = x$. Обозначим множество всех симметрических матриц через $S(n, R)$.

Задача. Докажите, что произведение двух симметрических матриц $x, y \in M(n, R)$ в том и только том случае симметрично, когда эти матрицы коммутируют.

Задача. Верно ли, что обратная к симметрической матрице тоже симметрическая?

---

[79] У.Гренандер, Г.Сеге, Теплицевы матрицы и их приложения. — ИЛ, М., 1961, с.1–308.

[80] И.С.Иохвидов, Ганкелевы и теплицевы матрицы и формы. — Наука, М., 1974.

[81] Л.Д.Пустыльников, Теплицевы и ганкелевы матрицы и их применения. — Успехи Мат. Наук, 1984, т.39, N.1, с.53–84.

Задача. Докажите, что если x — симметрическая матрица, y — произвольная матрица, то y^txy тоже симметрическая.

• **Персимметричные матрицы.** Матрица $x = (x_{ij})$ называется **персимметричной**, если она симметрична относительно побочной диагонали, т.е. $x_{ij} = x + n + 1 - j, n + 1 - i$ для всех i, j.

Задача. Докажите, что в случае коммутативного кольца матрица x в том и только том случае персимметрична, когда $x^t = fxf$, где f — перъединичная матрица. Тем самым, множество персимметрических матриц равно $fS(n, R)$.

• **Кососимметричные матрицы.** Матрица $x = (x_{ij})$ называется **антисимметричной**, если $x_{ij} = -x_{ji}$ для всех i, j. В случае коммутативного кольца матрица x в том и только том случае кососимметрична, когда $x^t = x$.

• **Антисимметричные матрицы.** Кососомметричная матрица $x = (x_{ij})$ называется **антисимметричной**, если кроме того, $x_{ii} = 0$ для всех i. Обозначим множество всех симметрических матриц через $A(n, R)$. В случае, когда 2 является регулярным элементом, классы кососимметричных и антисимметричных матриц совпадают, но, вообще говоря, это совершенно не так: про диагональные элементы кососимметрической матрицы можно утверждать лишь, что $2x_{ii} = 0$, в то время как для антисимметрической матрицы $x_{ii} = 0$.

Задача. Пусть $2 \in R^*$. Докажите, что

$$M(n, R) = S(n, R) \oplus A(n, R).$$

Продолжает ли это утверждение оставаться верным без условия $2 \in R^*$?

Решение. Докажем вначале, что $M(n, R) = S(n, R) + A(n, R)$. Пусть $z \in M(n, R)$. Положим

$$x_{ij} = \frac{1}{2}(z_{ij} + z_{ji}), \qquad y_{ij} = \frac{1}{2}(z_{ij} - z_{ji}).$$

Тогда $x = (x_{ij}) \in S(n, R)$, а $y = (y_{ij}) \in A(n, R)$. Если кольцо R коммутативно, то x и y можно выразить в виде

$$x = \frac{1}{2}(z + z^t), \qquad y = \frac{1}{2}(z - z^t).$$

Для доказательства единственности такого представления, достаточно убедиться в том, что $S(n, R) \cap A(n, R) = 0$. В самом деле, пусть $z \in S(n, R) \cap A(n, R)$. Тогда по самому определению $A(n, R)$ имеем $z_{ii} = 0$. С другой стороны, для $i \neq j$ имеем $z_{ij} = z_{ji} = -z_{ij}$ так что $2z_{ij} = 0$ и, так как 2 обратима, то $z_{ij} = 0$. Нет, например, если $2 = 0$, то $A(n, R) \leq S(n, R)$.

§ 27. Эрмитовски симметричные матрицы

Предположим, что R — кольцо с инволюцией $\alpha \mapsto \overline{\alpha}$. В этом случае можно определить отображение $M(m, n, R) \longrightarrow M(m, n, R)$, переводящее $x = (x_{ij})$ в $\overline{x} = (\overline{x}_{ij})$. Это отображение аддитивно $\overline{x + y} = \overline{x} + \overline{y}$. Следующие два класса матриц особенно часто рассматриваются в случае, когда $R = \mathbb{C}$, а инволюцией является комплексное сопряжение.

• **Эрмитовски симметричные матрицы.** Матрица $x = (x_{ij})$ называется **эрмитовски симметричной** или просто **эрмитовской**, если $x_{ij} = \overline{x}_{ji}$ для всех i, j. В случае коммутативного кольца матрица x в том и только том случае эрмитовски симметрична, когда $x^t = \overline{x}$. Обозначим множество всех эрмитовски симметричных матриц через $H(n, R)$.

Задача. При каком условии произведение двух эрмитовски симметричных матриц $x, y \in M(n, R)$ эрмитовски симметрично?

Задача. Пусть R — коммутативное кольцо, $x \in M(n,R)$. Убедитесь, что матрицы x^tx и xx^t симметричные. Пусть теперь R некоммутативно, а $^- : R \longrightarrow R$ инволюция кольца R. Всегда ли матрицы $x^\dagger x$ и $xx^\dagger$ будут эрмитовски симметричными?

Задача. Пусть $z = x + iy \in M(n,\mathbb{C})$ — эрмитовски симметричная матрица, $x, y \in M(n,\mathbb{R})$ — ее вещественная и мнимая части, соответственно. Докажите, что тогда x симметрична, а y — антисимметрична.

• **Эрмитовски антисимметричные матрицы.** Матрица $x = (x_{ij})$ называется **эрмитовски антисимметричной** или просто **антиэрмитовой**, если $x_{ij} = -\overline{x}_{ji}$ для всех i, j и, кроме того, $x_{ii} = 0$. В случае коммутативного кольца первое условие на матрицу x эквивалентно тому, что $x^t = -\overline{x}$.

Задача. Убедитесь, что матрица $x \in M(n,\mathbb{C})$ в том и только том случае эрмитовски антисимметричная, когда ix эрмитовски симметричная.

Задача. Что можно сказать про вещественную и мнимую части эрмитовски антисимметричной матрицы $x \in M(n,\mathbb{C})$.

Рассмотрим теперь $z \in M(n,\mathbb{C})$. Ясно, что имеет место разложение

$$M(n,\mathbb{C}) = M(n,\mathbb{R}) \oplus iM(n,\mathbb{R}).$$

Для этого, как и выше достаточно заметить, что $z = x+iy$, где $x = \dfrac{1}{2}(z+\overline{z})$ есть вещественная, а $y = \dfrac{1}{2}(z-\overline{z})$ — мнимая часть z. Однако в действительности для многих задач $\mathrm{re}(z)$ и $\mathrm{im}(z)$ естественно понимать иначе.

Задача. Покажите, что

$$M(n,\mathbb{C}) = H(n,\mathbb{C}) \oplus iH(n,\mathbb{C}).$$

Решение. Утверждается, что каждая матрица $z \in M(n,\mathbb{C})$ однозначно однозначно представляется в виде $z = x + iy$, где $x, y \in H(n,\mathbb{C})$. Полагая

$$x = \frac{1}{2}(z + z^\dagger), \qquad y = \frac{1}{2}(z - z^\dagger)$$

мы видим, что такое представление существует. С другой стороны, если $z = x + iy$ есть *какое-то* представление z в таком виде, то $x^\dagger = x^\dagger + iy^\dagger = x - iy$. Это доказывает единственность.

§ 28. Простейшие примеры подгрупп в $\mathrm{GL}(n,R)$

• **Диагональные матрицы.** Все обратимые диагональные матрицы образуют подгруппу $D = D(n,R)$ в $\mathrm{GL}(n,R)$, называемую **группой диагональных матриц**.

• **Треугольные матрицы.** Все обратимые веръние треугольные матрицы, обратные к которым тоже верхние треугольные, образуют подгруппу $B(n,R)$ в $\mathrm{GL}(n,R)$, называемую **группой верхних треугольных матриц**. Аналогично определяется и **группа нижних треугольных матриц** $B^-(n,R)$

Предостережение. В большинстве книг группы $B(n,R)$ и $B^-(n,R)$ определяются *неправильно*, как группы **всех** обратимых верхних треугольных или нижних треугольных матриц, соответственно. Однако в случае некоммутативных колец нетрудно привести примеры верхних треугольных матриц, обратные к которым не являются верхними треугольными и даже таких верхних треугольных матриц, обратные к которым нижние треугольные. Поэтому условие на обратную матрицу в этом определении является совершенно необходимым!

• **Унитреугольные матрицы.** Верхняя треугольная матрица $x = (x_{ij}) \in M(n,R)$ называется **верхней унитреугольной**, если все ее элементы на главной диагонали равны 1, т.е. $x_{ij} = \delta_{ij}$ для всех $i \geq j$. Аналогично, нижняя

треугольная матрица x называется **нижней унитреугольной**, если все ее элементы на главной диагонали равны 1, т.е. если $x_{ij} = \delta_{ij}$ для всех $i \leq j$. Как множество $U(n, R)$ верхних унитреугольных матриц, так и множество $U^-(n, R)$ нижних унитреугольных матриц образуют подгруппы в $\mathrm{GL}(n, R)$.

• **Мономиальные матрицы.** Матрица $x = (x_{ij}) \in M(n, R)$ называется **мономиальной**, если в каждой строке и каждом столбце этой матрицы ровно 1 обратимый элемент. Множество $N(n, R)$ всех мономиальных матриц образует подгруппу в $\mathrm{GL}(n, R)$. Изобразим, для примера, группу $N(2, R)$:

$$\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \cup \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix}.$$

• **Матрицы перестановки.** Матрица $x = (x_{ij}) \in M(n, R)$ называется мономиальной, если в каждой строке и каждом столбце этой матрицы ровно 1 элемент равен 1, в то время как все остальные элементы равны 0. Матрицы перестановки образуют подгруппу W_n в $\mathrm{GL}(n, R)$, изоморфную S_n.

Задача. Убедитесь, что обратная к матрице перестановки совпадает с транспонированной.

§ 29. Блочные матрицы

Сейчас мы изучим важнейший прием вычислений с матрицами. А именно, все наиболее часто используемые вычисления используют разбиение матриц на блоки.

1. Разбиение матриц на блоки. Напомним, что набор натуральных чисел $(n_1, \ldots, n_t)$ называется **разбиением** натурального n, если $n = n_1 + \ldots + n_t$. Каждое такое разбиение задает отношение эквивалентности на множестве $\underline{n}$. А именно, скажем, что $i \sim j$, если найдется такое $h = 1, \ldots, t$, что

$$n_1 + \ldots + n_{h-1} + 1 \leq i, j \leq n_1 + \ldots + n_h.$$

Обозначим через $X_1, \ldots, X_t$ классы этой эквивалентности. Ясно, что все эти классы состоят из последовательных натуральных чисел, причем $|X_h| = n_h$ для всех $h = 1, \ldots, t$.

Рассмотрим теперь модуль прямоугольных матриц $M(m, n, R)$ и зафиксируем разбиение $\mu = (m_1, \ldots, m_s)$ числа m, и разбиение $\nu = (n_1, \ldots, n_t)$ количества числа n. Пусть $X_1 \sqcup \ldots \sqcup X_s = \underline{m}$ и $Y_1 \sqcup \ldots \sqcup Y_t = \underline{n}$ — соответствующие разбиения множеств строчных и столбцовых индексов. Тогда подматрица $x^{hk} = (x_{ij})$, $i \in X_h$, $j \in Y_k$, матрицы x называется ее блоком в h-й горизонтальной и k-й вертикальной полосе. По определению $x^{hk} = M(m_h, n_k, R)$. Теперь матрицу x можно записать в виде

$$\begin{pmatrix} x^{11} & \ldots & x^{1t} \\ \ldots & \ldots & \ldots \\ x^{s1} & \ldots & x^{st} \end{pmatrix}$$

Такая запись называется разбиением x на блоки в соответствии с разбиением строк μ и разбиением столбцов ν. Подматрица (x_{ij}), $i \in X_h$, $j \in Y$, называется

h-й **горизонтальной полосой** матрицы x, а подматрица (x_{ij}), $i \in X$, $j \in Y_k$, — ее k-й **вертикальной полосой**.

2. Операции с блочными матрицами. Предположим вначале, что $x, y \in M(m, n, R)$ — матрицы, *одинаково* разбитые на блоки, скажем в соответствии с разбиением строк $\mu = (m_1, \ldots, m_s)$ и разбиением столбцов $\nu = (n_1, \ldots, n_t)$. Тогда в матрице $x + y$, рассматриваемой в соответствии с тем же разббиением на блоки, имеем $(x + y)^{hk} = x^{hk} + y^{hk}$. Также, конечно и $(\alpha x)^{hk} = \alpha x^{hk}$ и $(x\alpha)^{hk} = x^{hk}\alpha$.

С другой стороны, пусть $x \in M(l, m, R)$, $y \in M(m, n, R)$, причем x разбита на блоки в соответствии с разбиением строк $\lambda = (l_1, \ldots, l_r)$ и разбиением столбцов $\mu = (m_1, \ldots, m_s)$, а y разбита на блоки в соответствии с *тем же самым* разбиением строк $\mu = (m_1, \ldots, m_s)$ и разбиением столбцов $\nu = (n_1, \ldots, n_t)$. При этом количество вертикальных полос матрицы x равно количеству горизонтальных полос матрицы y, причем ширина каждой из вертикальных полос матрицы x равна ширине соответствующей горизонтальной полосы матрицы y. Будем называть такие разбиения матриц x и y на блоки *согласованными*. В этом случае определено произведение $x^{jh}y^{hk} \in M(l_j, n_k, R)$.

Задача. Проверьте, что

$$(xy)^{jk} = x^{j1}y^{1k} + \ldots + x^{js}y^{sk}.$$

Иными словами, входящие в произведение блочных матриц x и y блоки вычисляется при помощи обычных формул для произведения матриц так, как будто эти матрицы состоят из элементов кольца R, а не из блоков, каждый из которых сам является матрицей. Это правило особенно удобно, если среди блоков матриц x и/или y много равных 0, e или чему-нибудь в таком духе!

3. Изоморфизм $M(l, m, M(n, K)) \cong M(ln, mn, R)$. Особенно важен случай, когда все горизонтальные и все вертикальные полосы имеют одинаковую ширину n. В этом случае каждый из блоков матрицы x можно истолковать как матрицу $M(n, R)$. Таким образом,

§ 30. Прямая сумма матриц

1. Прямая сумма матриц. Следующая операция позволяет построить по матрицам $x \in M(k, l, R)$ и $y \in M(m, n, R)$ матрицу $x \oplus y \in M(k + m, l + n, R)$, называемую их прямой суммой:

$$x \oplus y = \operatorname{diag}(x, y) = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}.$$

Легко видеть, что эта операция ассоциативна, $(x \oplus y) \oplus z = x \oplus (y \oplus z)$, так что мы можем писать просто $x \oplus y \oplus z$. Обычно мы применяем эту операцию к квадратным матрицам.

Если $x_1, \ldots, x_t$ квадратные матрицы степеней $n_1, \ldots, n_t$ соответственно, то матрица $x = x_1 \oplus \ldots \oplus x_t$ степени $n = n_1 + \ldots + n_t$ называется **блочно-диагональной** матрицей, или, если нужно явно указать размеры диагональных блоков, — блочно диагональной матрицей типа $\nu = (n_1, \ldots, n_t)$.

Легко видеть, что прямая сумма связана со сложением и умножением матриц тождествами взаимной дистрибутивности

$$(x + y) \oplus (z + w) = (x \oplus z) + (y \oplus w), \qquad (xy) \oplus (zw) = (x \oplus z)(y \oplus w).$$

Эти тождества показывают, что множество

$$M(n_1, R) \oplus \ldots \oplus M(n_t, R) \leq M(n, R)$$

всех блочно диагональных матриц типа ν является подкольцом в $M(n, R)$, изоморфным прямой сумме колец $M(n_1, R), \ldots, M(n_t, R)$. Тем самым, оба смысла, в которых можно понимать знак прямой суммы в этой формуле, согласованы.

Блочно треугольные матрицы

блочно мономиальные матрицы

строго блочно треугольные матрицы

§ 31. Кронекерово произведение матриц

В этом параграфе кольцо R предполагается коммутативным. Хотя формально определение кронекеровского произведения можно дать и для некоммутативных колец, это произведение обладает хорошими свойствами только если R коммутативно.

1. Определение кронекерова произведения. Пусть l, m, n, p — четыре натуральных числа, тогда определена операция

$$\otimes : M(l, m, R) \times M(n, p, R) \longrightarrow M(lm, np, R),$$

сопоставляющая паре матриц $x \in M(l, m, R)$, $y \in M(n, p, R)$ их **тензорное произведение** $x \otimes y$, которую проще всего представлять себе как блочную матрицу вида

$$x \otimes y = \begin{pmatrix} x_{11}y & \ldots & x_{1m}y \\ & \ldots & \\ x_{l1}y & \ldots & x_{lm}y \end{pmatrix} \in M(lm, np, K).$$

Мы уже упоминали эту конструкцию в главе 1, где изображена матрица $x \otimes y$ для случая, когда $x, y \in M(2, R)$.

Особенно часто эта конструкция используется для квадратных матриц. При этом если $x \in M(m, R)$, $y \in M(n, R)$ то $x \otimes y \in M(m, M(n, R))$, в то время как $y \otimes x \in M(n, M(m, R))$. Как мы знаем, и то и другое кольцо можно отождествить с $M(mn, R)$. Тем не менее, вообще говоря, $x \otimes y \neq y \otimes x$. В действительности, матрица $y \otimes x$ получается из $x \otimes y$ перестановкой строк и столбцов.

Все тождества следующего пункта *можно* доказывать в терминах явного вида элементов тензорного произведения $x \otimes y$. Собственно, так примерно это и делается в большинстве обычных учебников теории матриц для прикладников, даже самых удачных[82]. Хотя я думаю, что мы не зря изучали операции с блочными матрицами, все же поймем, чему равен элемент $x \otimes y$ в позиции (r, s), где $1 \leq r, s \leq mn$.

Задача. Вычислите $(x \otimes y)_{rs}$.

[82] П.Ланкастер, Теория матриц. — Наука, М., 1978, с.1–280, с.235–236. Впрочем, в книге М.Маркус, Х.Минк, Обзор по теории матриц и матричных неравенств. — Наука, М., 1972, с.1–232. стр.20 приводится правильное доказательство в терминах умножения блоков.

Решение. По определению $x_{ij}y_{hk}$ стоит в позиции $((i-1)n+h, (j-1)n+k)$ для любых $1 \leq i, j \leq m$, $1 \leq h, k \leq n$. Прочтите эту формулу в обратную сторону.

2. Основные тождества. Кронекерово произведение связано с другими матричными операциями весьма интересными тождествами. При этом мы считаем, что произведение матриц связывает сильнее, чем кронекерово произведение, так что $x \otimes yz$ интерпретируется как $x \otimes (yz)$, а не как $(x \otimes y)z$. В то же время, кронекерово произведение связывает сильнее, чем сумма, так что $x + y \otimes z$ интерпретируется как $x + (y \otimes z)$, а не $(x + y) \otimes z$.

Теорема. *Кронекерово произведение матриц обладает следующими свойствами:*

1) *Ассоциативность. Пусть* $x \in M(l, R)$, $y \in M(m, R)$, $z \in M(n, R)$. *Тогда*

$$(x \otimes y) \otimes z = x \otimes (y \otimes z)$$

(*после отождествления обеих частей с матрицами в* $M(lmn, R)$).

2) *Дистрибутивность относительно сложения*

$$(x + y) \otimes z = x \otimes z + y \otimes z, \qquad x \otimes (y + z) = x \otimes y + x \otimes z.$$

3) *Однородность*

$$\alpha x \otimes y = \alpha(x \otimes y) = x \otimes \alpha y.$$

4) *Взаимная дистрибутивность кронекерова произведения и умножения матриц*

$$xy \otimes uv = (x \otimes u)(y \otimes v).$$

Доказательство. Все свойства, кроме 4, непосредственно вытекают из определения. Для доказательства 4 заметим, что блок матрицы $(x \otimes u)(y \otimes v)$, стоящий на пересечении i-й горизонтальной и j-й вертикальной полосы, равен $\sum(x_{il}u)(y_{lj}v)$, где сумма берется по $1 \leq l \leq m$. Однако по предположению кольцо R коммутативно, так что эта сумма равна $\sum(x_{il}y_{lj})(uv) = (xy)_{ij}(uv)$, как и утверждалось.

Задача. Проверьте, что если x и y обе обратимы, то $(x \otimes y)^{-1} = x^{-1} \otimes y^{-1}$.

Задача. Чему равно $(x \otimes y)^t$? Варианты ответа: 1) $x^t \otimes y^t$, 2) $y^t \otimes x^t$, 3) none of the above.

Задача. Что из следующих включений верно?

1) $D(m, R) \otimes D(n, R) \leq D(mn, R)$;

2) $N(m, R) \otimes N(n, R) \leq N(mn, R)$;

3) $U(m, R) \otimes U(n, R) \leq U(mn, R)$;

4) $B(m, R) \otimes B(n, R) \leq B(mn, R)$.

§ 32. Умножение матриц в терминах кронекеровского произведения

Определим отображение $M(mn, R) \longrightarrow R^{mn}$, $x \mapsto x_{*1} \oplus \ldots \oplus x_{*n}$, сопоставляющее матрице x вектор, являющийся прямой суммой ее столбцов.

§ 33. След матрицы

Сейчас мы определим аддитивный гомоморфизм $\mathrm{tr} : M(n, R) \longrightarrow R$.

Задача. Пусть $x \in M(m, n, \mathbb{R})$. Докажите, что $x^t x = 0$, то $x = 0$. Верно ли аналогичное утверждение для матриц $x \in M(m, n, \mathbb{C})$? Как нужно модифицировать формулировку в комплексном случае?

$\mathrm{tr}(x + y) = \mathrm{tr}(x) + \mathrm{tr}(y)$,
$\mathrm{tr}(\alpha x) = \alpha\,\mathrm{tr}(x)$,
$\mathrm{tr}(x\alpha) = \mathrm{tr}(x)\alpha$,
$\mathrm{tr}(xy) = \mathrm{tr}(yx)$,
$\mathrm{tr}(x \otimes y) = \mathrm{tr}(x)\,\mathrm{tr}(y)$.
$\mathrm{tr}(\bigwedge^m(x)$= сумма главных миноров m-го порядка матрицы x

Тема 9: ИДЕАЛЫ

При модном слове *идеал*
Невольно Ленский задремал.
Пушкин, 'Евгений Онегин', Гл. VI.

Поник я буйной головой, погибли идеалы.
Некрасов

§ 1. Односторонние идеалы

Сейчас мы введем одно из важнейших понятий, связанных с кольцами, понятие идеала. Идеалы играют такую же роль в теории колец, как нормальные подгруппы в теории групп.

1. Левые и правые идеалы. Пусть R – произвольное ассоциативное кольцо с 1.

Определение. *Непустое подмножество $I \subseteq R$ называется* **левым идеалом** *в R, если*

AI) *I является аддитивное подгруппой в R, т.е., $\forall x, y \in I$, $x - y \in I$;*

LI) *I* **устойчиво** *относительно умножения на элементы R* **слева***: $\forall x \in I$, $\forall y \in R$, $yx \in I$.*

Непустое подмножество $I \subseteq R$ называется **правым идеалом** *в R, если оно удовлетворяет аксиоме I1 и следующей*

RI) *I* **устойчиво** *относительно умножения на элементы R* **справа***: $\forall x \in I$, $\forall y \in R$, $xy \in I$.*

Левые и правые идеалы в R называются **односторонними**. Пользуясь умножением по Минковскому аксиомы LI и RI можно записать в виде $RI \subseteq I$ и $IR \subseteq I$, соответственно. Любой элемент вида yx, где $y \in R$, называется **левым кратным** элемента x, а элемент вида xy – **правым кратным** x. Таким образом, левый/правый идеал – это такая аддитивная подгруппа, которая вместе с каждым своим элементом содержит все его левые/правые кратные.

2. Главные левые/правые идеалы. Пусть $x \in R$. Обозначим через $Rx = \{yx \mid y \in R\}$ – множество всех *левых* кратных элемента x. Легко видеть, что Rx – левый идеал кольца R. В самом деле, Rx непусто (так как содержит $0 = 0x$ и $x = 1x$), разность двух левых кратных x тоже является его левым кратным, $yx - zy = (y - z)x$ и, наконец, левое кратное левого кратного x само является левым кратным x так как $z(yx) = (zy)x$ (ассоциативность!) Множество Rx называется **главным левым идеалом** кольца R, порожденным x. Точно так же проверяется что множество $xR = \{xy \mid y \in R\}$ всех *правых* кратных элемента x является правым идеалом в R, называемым **главным правым идеалом** в R, порожденным x.

Особую роль в структурной теории колец играют левые/правые идеалы, порожденные идемпотентами

Задача.– СДВИНУТЬ!!! Покажите, что если R – коммутативное кольцо, то для любого $x \in R$ и любого идемпотента $e \in R$ имеем $Rx \cap Re = Rxe$.

Решение. Ясно, что правая часть содержится в левой (кольцо R коммутативно!) Обратно, если $y \in Rx \cap Re$, то y имеет вид $y = ue$ для некоторого $u \in R$

так что $ye = ue^2 = ue = y$. С другой стороны, $y = vx$ для некоторого $v \in R$, так что $y = ye = vxe$, что и утверждалось.

3. Конечно порожденный левый/правый идеал. Обобщим предыдущий пример. Пусть $x_1, \dots, x_n \in R$ – любое конечное семейство элементов кольца R. Рассмотрим множество всех **левых линейных комбинаций** элементов $x_1, \dots, x_n$ с коэффициентами из R:

$$Rx_1 + \dots + Rx_n = \{y_1x_1 + \dots + y_nx_n \mid y_1, \dots, y_n \in R\}.$$

Точно такая же выкладка, как в предыдущем примере, показывает, что разность двух левых линейных комбинаций снова является линейной комбинацией и кратное линейной комбинации элементов $x_1, \dots, x_n$ само является левой линейной комбинацией тех же элементов. Поэтому $Rx_1 + \dots + Rx_n$ – левый идеал кольца R. Он называется **левым идеалом, порожденным** $x_1, \dots, x_n$. Совершенно аналогично проверяется и что множество всех **правых линейных комбинаций** элементов $x_1, \dots, x_n$ с коэффициентами из R:

$$x_1R + \dots + x_nR = \{x_1y_1 + \dots + x_ny_n \mid y_1, \dots, y_n \in R\}.$$

является правым идеалом в R, который называется **правым идеалом, порожденным** $x_1, \dots, x_n$. Левый/правый идеал называется **конечно порожденным**, если он совпадает с левым/правым идеалом, порожденным своим конечным подмножеством $\{x_1, \dots, x_n\}$.

4. Порождение левых/правых идеалов. Приведем первые примеры левых/правых идеалов.

§ 2. Двусторонние идеалы

1. Идеалы. Пусть по прежнему R – произвольное ассоциативное кольцо с 1.

Определение. *Аддитивная подгруппа I кольца R называется* **идеалом**, *если она является одновременно как левым, так и правым идеалом кольца R.*

Таким образом, идеал, это непустое подмножество в R, удовлетворяющее аксиомам AI, LI и RI, т.е. такая аддитивная подгшруппа, что $RI, IR \subseteq I$ или, что то же самое, $RIR \subseteq I$. Чтобы обозначить, что I – идеал в R обычно пишут $I \trianglelefteq R$. Иногда, если нужна особая точность, чтобы подчеркнуть, что $RIR \subseteq I$ говорят, что I – **двусторонний идеал**. В классических книгах идеалы обычно обозначаются фрактурой: $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, ... Сегодня фрактура практически вышла из употребления, за исключением трех стандартных обозначений: $\mathfrak{p}$ для простых идеалов, $\mathfrak{q}$ для примарных идеалов и $\mathfrak{m}$ для максимальных идеалов.

Если кольцо R *коммутативно*, то понятия левого идеала, правого идеала и двустороннего идеала совпадают. В частности, в этом случае вместо главных левых и главных правых идеалов можно говорить просто о **главных идеалах** кольца R. Если $x \in R$, то множество $(x) = Rx = xR$, всех **кратных** x называется главным идеалом, **порожденным** x. Аналогично, если $x_1, \dots, x_n \in R$, то вместо левых и правых линейных комбинаций этих элементов можно говорить просто о **линейных комбинациях** с коэффициентами из R. Все такие комбинации образуют идеал

$$(x_1, \dots, x_n) = Rx_1 + \dots + Rx_n = x_1R + \dots + x_nR \trianglelefteq R.$$

Замечание. Можно было бы назвать произведения yxz, где $y, z \in R$, *двусторонними кратными* x в кольце R, но это понятие не очень полезно. Дело в том, что в отличие от левых и правых кратных, a priori совершенно непонятно, почему сумма двусторонних кратных x сама будет кратным x. В самом деле, как, например, выразить $yx + xz$ как кратное x? В действительности, рассматривая ранги матриц, легко убедиться, что сделать это **невозможно**: в кольце $M(n, K)$ матриц степени $n \geq 2$ над полем K любое двустороннее кратное стандартной матричной единицы e_{11} имеет ≤ 1, в то время как $e_{21}e_{11} + e_{11}e_{12} = e_{21} + e_{12}$ имеет ранг 2 и, значит, не является кратным e_{11}. Таким образом, *неверно*, что двусторонний идеал, порожденный элементом x, обязан состоять из кратных x. В действительности, он состоит из всевозможных конечных сумм таких кратных. Поэтому для некоммутативных колец обычно не говорят о двусторонних кратных элемента x и главных идеалах.

• Пусть R коммутативное кольцо. Тогда все нильпотентные элементы образуют идеал кольца R обозначаемый $\mathrm{Nil}(R)$ и называемый **нильрадикалом**. Например, в кольце двойных чисел $R = K[d]$ имеем $\mathrm{Nil}(R) = \{dx \mid x \in K\} = dR$.

§ 2. Двусторонние идеалы

1. Очевидные идеалы, простые кольца. В любом кольце есть два **очевидных** идеала: **тривиальный** или **нулевой** идеал $0 = \{0\} \trianglelefteq R$, и **несобственный** или **единичный** идеал $R \trianglelefteq R$. Все идеалы I кольца R, отличные от R, называются **собственными**. Ясно, что если кольцо $R = T$ является телом, то в нем нет уже никаких других *односторонних* идеалов. Оказывается, верно и обратное.

Предложение. *Кольцо R в том и только том случае является телом, когда в нем ровно два* **левых** *идеала.*

Доказательство. Если $R = T$ – тело, то любой ненулевой элемент $x \neq 0$ обратим и, значит, если $x \in I$, где I левый идеал в R, то для любого $y \in R$ имеем $y = (yx^{-1})x \in I$.

Обратно, так как в кольце R **два** левых идеала, то $0 \neq 1$. Возьмем теперь $x \in R$, $x \neq 0$, и рассмотрим главный левый идеал идеал Rx. Так как $Rx \neq 0$, то $Rx = R$, так что x обратим слева, пусть, например, $yx = 1$. Так как y в свою очередь обратим слева, то y, а вместе с тем и x, обратим.

Таким образом, коммутативное кольцо является полем в том и только том случае, когда в нем **ровно** два идеала.

Определение. *Кольцо R называется* **простым**, *если в ровно два двусторонних идеала.*

2. Простота матричных колец. Мы только что доказали, что каждое *коммутативное* простое кольцо является полем. Можно ли и в некоммутативном случае утверждать, что каждое простое кольцо является телом? Оказывается, это **безнадежно** неверно.

Теорема. *Кольцо матриц $M(n, T)$ любой степени $n \geq 1$ над телом T является простым кольцом.*

Лрказательство. Пусть $I \trianglelefteq M(n, T)$, $I \neq 0$, и $x = (x_{ij}) \in I$ – ненулевой элемент из I. Это значит, что найдутся такие индексы h, k, что $x_{hk} \neq 0$.

Тогда для любых i, j имеем $e_{ij} = x_{hk}^{-1} e_{ih} x e_{kj}$. Тем самым, $y = \sum y_{ij} e_{ij} \in I$ для любой матрицы $y = (y_{ij}) \in M(n, T)$.

Одна из ключевых классических теорем теории колец, **теорема Веддербарна-Артина**, с которой мы познакомимся в 3-м семестре, утверждает, что кольцами $M(n, T)$ исчерпываются все простые **артиновы** кольца (см. Главу VI по поводу определения). На условие артиновости здесь нужно смотреть как на своего рода 'конечномерность': любое кольцо, конечномерное над полем, автоматически является артиновым. Однако бесконечномерные простые кольца устроены *весьма* замысловато и никакого столь же простого описания для них нет.

§ 4. Кольца главных идеалов: erste Fassung

4. Кольца главных идеалов. Опишем все идеалы в кольце $\mathbb{Z}$ целых рациональных чисел. Ясно, что каждому $m \in \mathbb{Z}$ отвечает главный идеал $m\mathbb{Z}$, причем главные идеалы $m\mathbb{Z}$ и $n\mathbb{Z}$ равны только если $m = \pm n$. Оказывается, никаких других идеалов в R нет.

Предложение. *Каждый идеал кольца $\mathbb{Z}$ имеет вид $m\mathbb{Z}$ для некоторого $m \in \mathbb{Z}$.*

Доказательство. В самом деле, пусть $I \trianglelefteq \mathbb{Z}$, $I \neq 0$, – ненулевой идеал в $\mathbb{Z}$. Заметим, что I содержит натуральное число n. В самом деле, пусть $l \in I$, $l \neq 0$. Если $l > 0$, то положим $n = l$, в противном случае положим $n = -l$. Множество $I \cap \mathbb{N}$ непусто и в силу полной упорядоченности $\mathbb{N}$ содержит наименьший элемент, который мы обозначим через m. Мы утверждаем, что $I = m\mathbb{Z}$. В самом деле, пусть $n \in I$. Поделим n с остатком на m: $n = qm + r$, где $0 \leq r < m$. Если $r \neq 0$, то $r = n - mq \in I \cap \mathbb{Z}$ элемент множества $I \cap \mathbb{N}$, что противоречит минимальности m. Поэтому n делится на m и, значит $n \in m\mathbb{Z}$.

Забегая вперед, скажем, что коммутативное кольцо с 1 называется **кольцом главных идеалов**, если оно яаляется областью целостности и все идеалы в нем главные. Доказанный только что результат означает, что $\mathbb{Z}$ является кольцом главных идеалов.

Используя алгоритм деления многочленов, мы покажем, что для любого поля K кольцо многочленов $K[x]$ тоже является кольцом главных идеалов, т.е. любой идеал в нем имеет вид $(f) = fK[x]$ для некоторого многочлена $f \in K[x]$. Во втором семестре мы тщательно изучим арифметику колец главных идеалов и приведем много других примеров. В частности, у нас будет случай вернуться с другой точки зрения к следующей задаче.

Задача. Докажите, что все подкольца в $\mathbb{Q}$ являются кольцами главных идеалов.

Указание. Пусть R такое кольцо и I – идеал в нем. Тогда $I = R(I \cap \mathbb{Z})$.

Задача. Пусть R – коммутативное кольцо, $x, y \in R$, $xy = 0$. Предположим, что идеал $Rx + Ry$ содержит неделитель 0. Показать, что тогда $Rx \cap Ry = 0$ и уже элемент $x + y$ не является делителем 0.

Решение. Если $Rx \cap Ry \neq 0$ и $z \in Rx \cap Ry$, $z \neq 0$, то $xz = 0$ (так как $z \in Ry$) и $xy = 0$ (так как $z \in Rx$). Поэтому для любых $a, b \in R$ выполняется равенство $(ax + by)z = 0$, так что $Rx \cap Ry$ целиком состоит из делителей 0. С

другой стороны, предположим, что $x+y$ является делителем 0. Тогда найдется $z \in R^\bullet$ такое, что $(x+y)z = 0$ или, что то же самое, $xz = -yz$. Ясно, что при этом $xz = -yz \neq 0$, так как иначе все элементы $ax + by$ были бы делителями 0. Но тогда $(ax + by)zx = (b - a)yzx = 0$ для любых $a, b \in R$, противоречие.

5. Пример неглавного идеала. Рассмотрим кольцо многочленов $R = K[x, y]$ от двух переменных. Мы будем изучать идеалы этого кольца (и вообще колец многочленов от произвольного числа переменных) в Главе ?. Пока заметим, что не все идеалы этого кольца главные. В самом деле, пусть I – множество многочленов без свободного члена, т.е.

$$I = \{f = \sum a_{ij}x^i y^j \mid a_{ij} \in K, i, j \in \mathbb{N}_0, a_{00} = 0\}.$$

Ясно, что I *собственный* идеал в $K[x, y]$, причем $I = xK[x, y]+yK[x, y]$ порождается *двумя* элементами. Однако идеал I не может быть главным. В самом деле, если $I = fK[x, y]$ для некоторого $f \in K[x, y]$, то как x так и y являются кратными f. Но тогда f должно быть ненулевой константой, что невозможно так как тогда $I = K[x, y]$.

Идеал (x, p) в кольце $R = \mathbb{Z}[x]$ – ОБРАБОТАТЬ!!

§ 5. Покрытие шахматной доски

Опишем чрезвычайно эффектное применение идеалов в кольце $K[x, y]$ к задаче покрытия шахматной доски (см. Т.Саати 'Целочисленные методы оптимизации и связанные с ними экстремальные проблемы', М. 1973). Для этого занумеруем вертикали степениями x, от 1 до x^7, а горизонтали – степенями y, от 1 до y^7. Тогда каждой клетке соответствует одночлен вида $x^i y^j$, $0 \leq i, j \leq 7$. например, клетка, обычно обозначаемая e4 теперь обозначена x^4y^3. Таким образом, каждому подмножеству Z шахматной доски отвечает некоторая линейная комбинация $f = F_Z$ этих одночленов с коэффициентами $0, 1$. Положим теперь на шахматную доску костяшку домино так, чтобы она покрывала две соседние клетки и пусть $x^i y^j$ – нижняя левая из них. Тогда вторая клетка равна $x^{i+1}y^j$, если костяшка лежит горизонтально, и $x^i y^{j+1}$, если костяшка лежит вертикально. Это значит, что костяшка покрывает множество $x^i y^j(1 + x)$ в первом случае и множество $x^i y^j(1 + y)$ во втором. Тем самым, если Z – какое-то подмножество шахматной доски, которое можно покрыть костяшками домино, то f_Z лежит в идеале кольца $\mathbb{Z}[x, y]$, порожденном $1 + x$ и $1 + y$.

Покажем, например, как отсюда получается, что множество Z, представляющее собой шахматную доску с вырезанными левым нижним и правым верхним углами, нельзя покрыть костяшками домино. В самом деле,

$$f_Z = (1 + x + \ldots + x^7)(1 + y + \ldots + y^7) - 1 - x^7y^7.$$

Однако очевидно, что этот многочлен не лежит в идеале $(1 + x)\mathbb{Z}[x, y] + (1 + y)\mathbb{Z}[x, y]$, порожденном $1 + x$ и $1 + y$. В самом деле, допустим $f_Z = (1 + x)f + (1+y)g$ для некоторых многочленов $f, g \in \mathbb{Z}[x, y]$. В частности, отсюда следует, что для любого гомоморфизма $\phi : \mathbb{Z}[x, y] \longrightarrow R$ образ $\phi(f_Z)$ должен лежать в идеале, порожденном $\phi(1+x)$ и $\phi(1+y)$. Рассмотрим, например, гомоморфизм $\phi : \mathbb{Z}[x, y] \longrightarrow \mathbb{Z}$, $f \mapsto f(-1, -1)$, который каждому многочлену $f \in \mathbb{Z}[x, y]$ сопоставляет его значение в точке $(x, y) = (-1, -1)$. Тогда $\phi(f_Z) = -2$, в то

время как $\phi(1+x) = \phi(1+y) = 0$. Это значит, что f_Z не может лежать в идеале, порожденном $1 + x$ и $1 + y$ и, тем самым, Z невозможно покрыть костяшками домино.

Конечно, это игрушечный пример, и в данном случае невозможность такого покрытия была уже нам известна, но заметим, что теперь мы владеем **общим методом** решения **всех** подобных вопросов, пригодным для любого подмножества и костяшек любой формы: не только домино, но и тримино, тетрамино, пентамино, и т.д. Этот метод сразу же обобщается на покрытие n-мерной шахматной доски (для этого достаточно перейти к кольцу $\mathbb{Z}[x_1, \dots, x_n]$ многочленов от n переменных), шахматной доски на торе (рассмотрите кольцо $\mathbb{Z}[x, y]/(x^m - 1, y^n - 1)$) и т.д. Кроме того, мы нигде не использовали, что коэффициенты линейной комбинации являются целыми числами. В действительности, мы доказали, что шахматную доску с вырезанными углами нельзя равномерно покрыть костяшками домино, имеющими различную вещественную толщину, даже если допустить, что толщина некоторых костяшек может быть отрицательной. С учетом теории базисов Гребнера, которую мы изучим в Главе ?, этот метод дает абсолютно рабочий алгоритм для решения всех реально возникающих задач упаковки, который позволяет в каждом конкретном случае либо построить требуемую упаковку, либо доказать ее отсутствие.

§ 6. Сравнения по модулю идеала.

Обсуждая конструкцию фактор-операции в Главе I мы уже встречались со сравнениями в кольце целых чисел. В этом параграфе мы излагаем общую теорию.

1. Сравнения по модулю идеала. Пусть R – произвольное ассоциативное кольцо с 1, не обязательно коммутативное, I – двусторонний идеал в R.

Определение. *Говорят, что два элемента $x, y \in R$ сравнимы по модулю идеала I и пишут $x \equiv y \pmod I$, или просто $x \equiv y(I)$, если $x - y \in I$.*

Отношение сравнимости по модулю I иногда обозначается также $\equiv_I$. Основные примеры, которые будут нам встречаться, связаны с коммутативными кольцами R, в первую очередь, как обычно, кольцом $\mathbb{Z}$ и кольцом $K[x]$. В обоих этих случаях каждый идеал I имеет вид $I = (z)$ для некоторого $z \in R$ и, поэтому, вместо сравнимости по модулю I говорят обычно просто о сравнимости по модулю z и пишут $x \equiv y \pmod z$ – мы вскоре вернемся R этой теме. Вот еще один пример.

Рассмотрим кольцо $R = \mathbb{R}^X$ вещественнозначных функций на множестве X и пусть $I = I_Y$ – идеал функций, обращающихся в 0 на подмножестве $Y \subseteq X$ (см. пункт 7 предыдущего параграфа). Тогда $f \equiv g \pmod I$, в том и только том случае, когда для любого $y \in Y$ имеем $f(y) = g(y)$.

2.

Предложение. *Фиксируем идеал I в R. Отношение сравнимости по модулю I является отношением эквивалентности на R.*

Доказательство. Все свойства эквивалентности легко вытекают из определения идеала.

1. Рефлексивность: $x \equiv x(I) \iff x - x = 0 \in I$;
2. Симметричность: $x \equiv y(I) \iff x - y \in I \iff y - x \in I \iff y_x(I)$;

3. Транзитивность: $x \equiv y(I)\ y\ y \equiv z(I) \Longleftrightarrow x-y, y-z \in I \Longrightarrow x-z = (x-y)+(y-z) \in I \Longleftrightarrow x_z(I)$.

Здесь использовалось лишь то, что I является аддитивной подгруппой, однако то, что I действительно идеал, будет в полной мере использоваться при доказательстве того, что отношение сравнимости по модулю I является конгруэнцией на кольце R. Напомним, прежде всего, определение.

Определение. *Отношение эквивалентности на кольце R называется* **конгруэнцией**, *если оно совместимо с алгебраическими операциями, т.е., иными словами, для любых двух пар эквивалентных элементов $x \sim y$, $z \sim w$ соответствующие суммы, разности и произведения также эквивалентны: $x \pm z \sim y \pm w$ (знаки в правой и левой части пробегаются одновременно) и $xz \sim yw$.*

Предложение. *Фиксируем идеал I в R. Отношение сравнимости по модулю I является конгруэнцией на R.*

Доказательство. Совместимость со сложением снова вытекает просто из того, что I – аддитивная подгруппа. В самом деле, пусть $x \equiv y \pmod I$, $z \equiv w \pmod I$. Тогда $(x+z)-(y+w) = (x-y)+(z-w) \in I$ так как оба слагаемых принадлежат I, но это и значит, что $x+z \equiv y+w \pmod I$ и $x-z \equiv y-w \pmod I$.

Для доказательства же совместимости с умножением потребуется условие, что I – двусторонний идеалю В самом деле, в тех же обозначениях, $xz-yw = xz-xw+xw-yw = x(z-w)+(x-y)w$. Первое слагаемое в правой части принадлежит I в силу того, что $z-w \in I$ и I является левым идеалом, а второе слагаемое – в силу того, что $x-y \in I$ и I является правым идеалом. Тем самым, окончательно, $xz-yw \in I$ и, значит, $xz \equiv yw \pmod I$.

§ 7. Фактор-кольцо по модулю идеала.

Мы продолжаем считать, что R – любое ассоциативное кольцо с 1 и I - двусторонний идеал в R. Как мы только что показали, отношение сравнимости по модулю I является отношением эквивалентности на R и, следовательно, с ним можно связать фактор-множество R/I, состоящее из классов этой эквивалентности. По определению $x \equiv y \pmod I$ в том и только том случае, когда $x-y \in I$, так что класс эквивалентности элемента $x \in R$ состоит из всех y, имеющих вид $y = x+z$ для некоторого $z \in I$, и мы коротко обозначим его через $\overline{x} = x+I = \{x+z \mid z \in I\}$. Таким образом, R/I имеет вид: $R/I = \{x+I \mid x \in R\}$.

В действительности, мы знаем, что сравнимость по модулю I не просто эквивалентность, а конгруэнция. Это значит, что существует единственный способ, задать на R/I сложение и умножение так, чтобы каноническая проекция $\pi : R \longrightarrow R/I$, $x \mapsto x+I$ являлась гомоморфизмом колец, т.е. сохраняла как аддитивную, так и мультипликативную структуру.

Определим на R/I операции, полагая $\overline{x}+\overline{y} = \overline{x+y}$ и $\overline{x} \cdot \overline{y} = \overline{xy}$ для любых двух классов $x, y \in R/I$. Мы должны еше, разумеется, показать, что так определенные операции определены корректно, т.е. не зависят от выбора представителей в данных классах. Но именно это и содержится в утверждении, что $\equiv_I$ – конгруэнция.

В самом деле, пусть $\overline{z} = \overline{x}$ и $\overline{w} = \overline{y}$. Тогда $z \equiv x \pmod I$ и $w \equiv y \pmod I$ и, следовательно, $z+w \equiv x+y \pmod I$ и $zw \equiv xy \pmod I$. Тем самым, $\overline{z+w} =$

$\overline{x+y}$ и $\overline{zw} = \overline{xy}$, так что сумма и произведение двух классов действительно определяются равенствами $\overline{x}+\overline{y} = \overline{x+y}$ и $\overline{x}\cdot\overline{y} = \overline{xy}$ корректно и не зависят от выбора представителей.

Теорема. *Фактор-множество R/I с так определенными операциями представляет собой кольцо и каноническая проекция $\pi : R \longrightarrow R/I$ задает сюръективный гомоморфизм колец с ядром I.*

Доказательство. Проверим, прежде всего, выполнение для R/I аксиом кольца. По сложению R/I образует абелеву группу: действительно, выполнение ассоциативности и коммутативности в R/I гарантируется соответствующими свойствами сложения в R, в качестве нейтрального элемента по сложению в R/I выступает класс $0 = I$, а в качестве противоположного к x – класс $-x$. Также и ассоциативность умножения в R/I и двусторонняя дистрибутивность умножения относительно сложения вытекают из соответствующих свойств умножения в R, а в качестве нейтрального элемента по умножению в R/I выступает класс 1. Тем самым, R/I действительно представляет собой ассоциативное кольцо с 1.

По самому построению R/I каноническая проекция $\pi : R \longrightarrow R/I$ является гомоморфизмом колец. В самом деле, $\pi(x+y) = \overline{x+y} = \overline{x}+\overline{y} = \pi(x)+\pi(y)$ и $\pi(xy) = \overline{xy} = \overline{x}\cdot\overline{y} = \pi(x)\pi(y)$. Как вытекает из определения фактор-множества, этот гомоморфизм сюръективен. Наконец, его ядро состоит из тех $x \in R$, для которых $x + I = \pi(x) = 0 = I$, т.е. из всех $x \in I$. Теорема полностью доказана.

Примеры
$\mathbb{Z}/m\mathbb{Z}$
$K[x]/(f)$, f – неприводимый многочлен, поле разложения многочлена f
$\mathbb{C} = \mathbb{R}[x]/(x^2+1)$
$\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[x]/(x^2-2)$.
$K[x]/(x^n)$ – кольцо усеченных многочленов
$K[C_n] \cong K[x]/(x^n-1)$ – групповая алгебра циклической группы
$K[C_m \times C_n] \cong K[x,y]/(x^m-1, y^n-1)$
$K[\zeta_n] \cong \mathbb{Z}[x]/(\Phi_n)$ (теорема Кронекера–Дедекинда)
аффинное кольцо алгебраического множества

Предложение. *Пусть $R_1 \oplus \ldots \oplus R_s$ и $I_i \trianglelefteq R_i$ для всех i. Тогда $I_1 \oplus \ldots \oplus I_s \trianglelefteq R$ и*

$$R/I \cong R_1/I_1 \oplus \ldots \oplus R_s/I_s.$$

Задача. Пусть K – поле. Покажите, что в кольце K^n (с покомпонентными операциями) ровно 2^n идеалов и что каждое фактор-кольцо кольца K^n изоморфно K^m для некоторого $0 \le m \le n$.

Задача. Идеал кольца R, порожденный аддитивными коммутаторами $[x,y] = xy-yx$, $x, y \in R$, – это наименьший идеал кольца Rб фактор-кольцо по которому коммутативно.

Следующие примеры фактор-колец естественно возникают при изучении плоских алгебраических кривых (см., например, Хартсхорн [На], задача 1.1 на стр. 24).

Задача (аффинное кольцо параболы). Покажите, что $K[x,y]/(x^2-y) \cong K[z]$.

Задача (аффинное кольцо гиперболы). Покажите, что $K[x,y]/(xy-1) \cong K[z,z^{-1}]$.

Задача (аффинное кольцо коники). Пусть $f \in K[x,y]$ – неприводимый многочлен степени 2. Показать, что $R = K[x,y]/(f)$ изоморфно либо $K[z]$, либо $K[z,z^{-1}]$. Как узнать, которому из этих колец изоморфно R?

Ответ можно найти в любом учебнике "аналитической геометрии". Нужно представить f в виде $f = f_2 + f_1 + f_0$ и привести f_2 к главным осям.

§ 8. Теоремы об изоморфизме

Пусть $\pi : R \longrightarrow R/I$ – каноническая проекция на фактор-кольцо по модулю идеала $I \trianglelefteq R$.

Теорема Нетер об изоморфизме. *Если $A \leq R$ – подкольцо, а $I \trianglelefteq R$ – идеал, то $A \cap I$ – идеал в A и*

$$A/(A \cap I) \cong (A+I)/I \leq R/I.$$

Доказательство. То, что это изоморфизм аддитивных групп, вытекает уже из теоремы Нетер об изоморфизме для групп. Этот изоморфизм определяется сопоставлением $A \longrightarrow (A+I)/I$, $x \mapsto x+I$. Поэтому остается лишь проверить, что это действительно гомоморфизм колец: $xy \mapsto xy+I = (x+I)(y+I)$. Ядро гомоморфизма колец автоматически является идеалом.

Теорема. *Пусть $I \trianglelefteq R$. Тогда соответствие $A \mapsto \pi^{-1}(A)$ устанавливает изоморфизм решетки левых/правых/двусторонних идеалов кольца R/I и решетки левых/правых/двусторонних идеалов кольца R, содержащих I.*

Доказательство. Нужно проверить лишь, что прообраз левого/правого/двустороннего идеала кольца R/I тоже будет левым/правым/двусторонним идеалом и что при этом

$$\pi^{-1}(A \cap B) = \pi^{-1}(A) \cap \pi^{-1}(B), \qquad \pi^{-1}(A+B) = \pi^{-1}(A) + \pi^{-1}(B).$$

Это предоставляется читателю.

Точно так же устанавливается биекция между множеством всех подколец в R/I и множеством всех подколец в R, содержащих I.

Вторая теорема об изоморфизме. *Если $I, A \trianglelefteq R$, $U \leq A$, то $R/A \cong (R/I)/(A/I)$.*

§ 9. Дубль кольца вдоль идеала

В приложениях теории колец очень часто возникают *пары*, состоящие из кольца R и идеала I в нем. Оказывается, в большинстве случаев изучение такой пары (R,I) полностью сводится к изучению некоторого нового кольца $R \times_I R$, называемого дублем кольца R вдоль идеала I. Сейчас мы рассмотрим эту важнейшую конструкцию.

1. Дубль кольца. Определим **дубль** $R\times_I R$ кольца R по отношению к идеалу $I \leq R$ посредством **декартова квадрата**

$$\begin{array}{ccc} R \times_I R & \xrightarrow{\pi_1} & R \\ {\scriptstyle \pi_2}\downarrow & & \downarrow{\scriptstyle \pi} \\ R & \xrightarrow[\pi]{} & R/I \end{array}$$

Иными словами, $R \times_I R$ состоит их всех пар $(a, b) \in R \times R$ таких, что $a \equiv b \pmod I$ с покомпонентными операциями и $\pi_1(a, b) = a$, $\pi_2(a, b) = b$. Ясно, что $\operatorname{Ker}\pi_1 = (0, I)$ и $\operatorname{Ker}\pi_2 = (I, 0)$. Диагональное вложение $\delta : R \longrightarrow R \times_I R$, задаваемое посредством $\delta(a) = (a, a)$ расщепляет как π_1, так и π_2.

Для пары (R, I) можно определить полупрямое произведение $R \leftthreetimes I$ кольца R и I как множество пар (a, c), $a \in R$, $c \in I$, с покомпонентным сложением и умножением, определенным формулой $(a, c)(b, d) = (ab, ad + cb + cd)$.

Лемма. *Кольцо $R \times_I R$ изоморфно полупрямому произведению $R \leftthreetimes I$ кольца $\delta(R) \cong R$ и идеала $\operatorname{Ker}\pi_1 \cong I$.*

Доказательство. Определим отображение из $R \times_I R$ в $R \leftthreetimes I$, полагая $(a, b) \mapsto (a, b - a)$. Из определения умножения в $R \leftthreetimes I$ следует, что это гомоморфизм. Обратный гомоморфизм определяется посредством $(a, c) \mapsto (a, a + c)$.

В дальнейшем мы отождествляем I с $\operatorname{Ker}\pi_1$.

Тема 10: ПРОСТЫЕ И МАКСИМАЛЬНЫЕ ИДЕАЛЫ

§ 1. Характеристика области целостности

> "Can you do addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one and one?" "I don't know," said Alice, "I lost count."
>
> Charles Dodgson "Through the looking glass"

В настоящем параграфе мы введем важнейший инвариант коммутативного кольца.

1. Характеристика кольца. Пусть R – коммутативное кольцо с 1. Рассмотрим гомоморфизм колец $\psi : \mathbb{Z} \longrightarrow R$, определенный посредством

$$n \mapsto n \cdot 1 = \underbrace{1 + \ldots + 1}_{n \text{ слагаемых}} .$$

Пусть $I = \mathrm{Ker}(\psi)$ – ядро этого гомоморфизма. Так как $\mathbb{Z}$ – кольцо главных идеалов, то $I = m\mathbb{Z}$ для некоторого $m \in \mathbb{N}_0$. Для $m \neq 0$ в кольце R выполняется $1 + \ldots + 1 = 0$, где число единиц равно m.

Определение. *Если* $\mathrm{Ker}(\psi) = m\mathbb{Z}$*, то говорят, что характеристика* R *равна* m *и пишут* $\mathrm{char}(R) = m$.

Иными словами, $\mathrm{char}(R) = 0$, если $\psi : \mathbb{Z} \longrightarrow R$ является мономорфизмом и $\mathrm{char}(R) = m$, если m – наименьшее натуральное число такое, что

$$\underbrace{1 + \ldots + 1}_{m \text{ слагаемых}} = 0.$$

Чтобы различать случаи $\mathrm{char}(R) = 0$ и $\mathrm{char}(R) > 0$, в первом из них говорят, что R кольцо **нулевой** характеристики, а во втором, что R кольцо **положительной** характеристики. Особенно важен случай, когда $\mathrm{char}(R) = p > 0$ – простое число, такое R называется кольцом **простой** характеристики.

Вот примеры колец различных характеристик.

• Характеристика $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ равна 0.

• Характеристика кольца классов вычетов $\mathbb{Z}/m\mathbb{Z}$ равна m, в частности, характеристика простого поля $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ равна p.

• Характеристика булева кольца множеств $R = 2^X$, $X \neq \emptyset$, (относительно симметрической разности и пересечения) равна 2.

2. Характеристика области целостности. Оказывается, тот факт, что характеристика всех известных нам полей либо равна 0, либо является простым числом, не случаен.

Теорема. *1) Характеристика неразложимого кольца либо равна 0, либо является примарным числом.*

2) Характеристика области целостности либо равна 0, либо является простым числом.

Доказательство. 1) В самом деле, пусть $\mathrm{char}(R) = mn > 0$, где $\gcd(m, n) = 1$. Мы утверждаем, что тогда $R = mR \oplus nR$. В самом деле, если $x \in mRcapnR$,

то $mx = nx = 0$ так что $x = \gcd(m, n)x = 0$. Поэтому $mR \cap nR = 0$. С другой стороны, так как $1 \in mR + nR$, то $mR + nR = R$, так что кольцо R разложимо.

Предположим, теперь, что R область целостности и сумма m единиц в R равна 0 для некоторого $m \in \mathbb{N}$. Допустим, что $m = kl$ составное. Это значит, что в R выполняется равенство

$$(\underbrace{1 + \ldots + 1}_{k \text{ слагаемых}})(\underbrace{1 + \ldots + 1}_{l \text{ слагаемых}}) = \underbrace{1 + \ldots + 1}_{m \text{ слагаемых}} = 0.$$

Отсюда следует, что по крайней мере один из сомножителей равен 0, так что характеристика R меньше, чем m.

В некоторых результатах нам придется накладывать различные ограничения на характеристику, например, требовать, чтобы она равнялась 0, была отлична от 2. etc. Кольца характеристики 0 ближе к классическим числовым образованиям, рассматриваемым в школьной математике, чем кольца положительной характеристики.

§ 2. Эндоморфизм Фробениуса

1. Эндоморфизм Фробениуса. Кольца положительной характеристики обладают многими непривычными свойствами. Укажем одно из наиболее замечательных свойств колец простой характеристики. Начнем со следующей элементарной леммы.

Лемма. *Если p – простое число, то для любого m, $1 \leq m \leq p-1$, выполнено сравнение $\binom{p}{m} \equiv 0 \pmod{p}$.*

Доказательство. В самом деле,

$$\binom{p}{m} = \frac{p(p-1)\ldots(p-m+1)}{1 \cdot 2 \cdot \ldots \cdot m},$$

причем p в числителе не может сократиться, так как оно взаимно просто со всеми i, $1 \leq i \leq p-1$.

Теорема. *Пусть R коммутативное кольцо простой характеристики p. Тогда отображение $F_p : R \longrightarrow R$, $x \mapsto x^p$, является эндоморфизмом кольца R.*

Доказательство. Так как R коммутативно, то $(xy)^p = x^p y^p$ для любых $x, y \in R$. С другой стороны, по формуле бинома Ньютона (которая также справедлива для произвольного коммутативного кольца), имеем

$$(x+y)^p = x^p + \binom{p}{1}x^{p-1}y + \ldots + \binom{p}{p-1}xy^{p-1} + y^p,$$

причем по лемме все коэффициенты $\binom{p}{1}, \ldots, \binom{p}{p-1}$ делятся на p и, значит, обращаются в 0 в кольце R. Тем самым, действительно, $(x+y)^p = x^p + y^p$, так что F_p является гомоморфизмом колец.

Заметим, что $F_p^n : x \mapsto x^{p^n}$, причем по индукции из теоремы сразу вытекает, что

$$(xy)^{p^n} = x^{p^n} y^{p^n}, \qquad (x+y)^{p^n} = x^{p^n} + y^{p^n}.$$

Эндоморфизм F_p, а часто и его степени F_p^n, называется **эндоморфизмом Фробениуса**. Этот эндоморфизм является одним из важнейших инструментов при изучении конечных полей. Мы могли бы не исключать здесь кольца характеристики 0, введя понятие **характеристичекой экспоненты**. Характеристическая экспонента кольца равна p, если его характеристика $p > 0$ и равна 1, если его характеристика равна 0.

2. Характеризация унипотентных элементов. Следующий результат часто (и обычно без всяких специальных ссылок) используется в случае кольца матриц $R = M(n, K)$, где K – поле характеристики p.

Задача. Если R – кольцо примарной характеристики p^n, то множество унипотентных элементов совпадает с p-частью группы R^*:

$$U(R) = \{u \in R^* \mid \exists m \in \mathbb{N}, o(u) = p^m.\}$$

В частности, если R приведенное, то порядок любого элемента $x \in R^*$ взаимно прост с p. Если R – коммутативное, то $U(R) = R_p^*$ есть p-примарная компонента группы R^*.

3. Простые поля. Оказывается, среди всех полей данной характеристики существует единственное с точностью до изоморфизма самое маленькое.

Определение. *Поле K, не содержащее собственных подполей, называется* **простым**.

Теорема. *Простое поле изоморфно либо $\mathbb{Q}$, если его характеристика равна 0, либо $\mathbb{F}_p$, если его характеристика равна $p > 0$.*

Доказательство. Рассмотрим подкольцо в K, порожденное 1. Оно совпадает либо с $\mathbb{Z}$, либо с $\mathbb{Z}/p\mathbb{Z}$ для некоторого $p \in \mathbb{P}$. Во втором случае оно само является подполем, в первом случае оно порождает подполе, изоморфное полю $\mathbb{Q}$.

p-многочлены. Пусть характеристика поля K равна p. В этом случае **p-многочленами** называются многочлены вида $a_0x + a_1x^p + a_2x^{p^2} + \ldots$. По поводу следующей задач см.

Р.Лидл, Х.Нидеррайтер "Введение в теорию конечных полей и их приложений", § 3.4)

Задача. Показать, что множество p-многочленов образует коммутативное ассоциативное кольцо относительно обычного сложения многочленов и композиции $\circ$.

Построим отображение $\widehat{\ }\colon K[x] \longrightarrow K[x]$, $\sum a_i x^i \mapsto \sum a_i x^{p^i}$.

Задача. Показать, что отображение $f \mapsto \widehat{f}$ задает биекцию кольца многочленов на кольцо p-многочленов. При этом $\widehat{f+g} = \widehat{f} + \widehat{g}$, $\widehat{fg} = \widehat{f} \circ \widehat{g}$, так что в действительности это **изоморфизм колец**.

§ 3. Простые идеалы

1. Простые идеалы. Сейчас мы введем важнейший класс идеалов коммутативных колец.

Определение. *Идеал I кольца R называется* **собственным**, *если $I \neq R$.*

Определение. *Собственный идеал $\mathfrak{p}$ кольца R называется* **простым**, *если из того, что $xy \in \mathfrak{p}$ вытекает, что по крайней мере один из элементов x, y лежит в $\mathfrak{p}$.*

Множество всех простых идеалов кольца R обозначается $\mathrm{Spec}(R)$ и называется **спектром кольца** R. Это важнейший **геометрический** объект, связанный с кольцом R. Он снабжается естественной структурой топологического пространства (топология Зариского) и пучком локальных колец, которые превращают его в локально окольцованное пространство.

По определению нулевой идеал (0) в том и только том случае прост, когда R является областью целостности. Обобщая это наблюдение, получаем следующий результат, в котором $\mathfrak{a}, \mathfrak{b}$ обозначают произвольные идеалы кольца R.

Предложение. *Следующие условия эквивалентны:*

1) Идеал $\mathfrak{p}$ прост;

2) Фактор-кольцо $R/\mathfrak{p}$ является областью целостности;

3) Из того, что $\mathfrak{ab} \subseteq \mathfrak{p}$ вытекает, что по крайней мере один из идеалов $\mathfrak{a}, \mathfrak{b}$ содержится в $\mathfrak{p}$.

Доказательство. 1) $\Longrightarrow$ 2) Пусть $(x+\mathfrak{p})(y+\mathfrak{p}) = 0$ в $R/\mathfrak{p}$. Это означает, что $xy \in \mathfrak{p}$. По определению простого идеала $x \in \mathfrak{p}$ или $y \in \mathfrak{p}$. так что по крайней мере один из классов $x+\mathfrak{p}$ или $y+\mathfrak{p}$ равен 0.

2) $\Longrightarrow$ 3) Пусть $\mathfrak{a}, \mathfrak{b} \trianglelefteq R$. Рассмотрим их образы $\overline{\mathfrak{a}} = (\mathfrak{a}+\mathfrak{p})/\mathfrak{p}$ и $\overline{\mathfrak{b}} = (\mathfrak{b}+\mathfrak{p})/\mathfrak{p}$ при канонической проекции в $R/\mathfrak{p}$. Так как $\mathfrak{ab} \subseteq \mathfrak{p}$, то $\overline{\mathfrak{a}}\overline{\mathfrak{b}} = 0$. Если теперь $\mathfrak{a} \not\subseteq \mathfrak{p}$ и $\mathfrak{b} \not\subseteq \mathfrak{p}$, то это противоречит целостности $R/\mathfrak{p}$.

3) $\Longrightarrow$ 1) Условие 3 формально сильнее условия 1, чтобы убедиться в этом, достаточно рассмотреть главные идеалы $=(x)$, $\mathfrak{b} = (y)$, порожденные элементами $x, y \in R$.

Условия в пунктах 1 и 3 сразу распространяются на любое конечное число сомножителей. Скажем, если $x_1 \dots x_n \in \mathfrak{p}$, то по крайней мере один из сомножителей x_i лежит в $\mathfrak{p}$.

Задача. Докажите, что если в коммутативном кольце все собственные идеалы простые, то это кольцо поле.

Решение. Так как идеал 0 прост, то R является областью целостности. С другой стороны, если $a \neq 0$, то, так как $a^2 \in (a^2)$, то $a \in (a^2)$. Тем самым, $(a) = (a^2)$. Тем самым, $a = a^2x$ для некоторого $x \in R$. Это значит, что $a(1-ax) = 0$ и, окончательно, $ax = 1$.

Примеры простых идеалов. Приведем несколько очевидных примеров простых идеалов.

• 0 является простым идеалом кольца R в том и только том случае, когда R область целостности.

• Простыми идеалами в кольце $\mathbb{Z}$ являются главные идеалы (p), $p \in \mathbb{P}$, и идеал (0).

• Простыми идеалами в кольце $K[x]$ являются главные идеалы (f), где f неприводим над K, и идеал (0).

• Пусть $R = K[x, y]$. Тогда главный идеал (y), порожденный переменной y, прост так как фактор-кольцо $K[x, y]/(y) = K[x]$ целостное. С другой стороны,

(y) не максимален. Он содержится, например, в максимальном идеале $xR+yR$, состоящем из всех многочленов без свободного члена (а также в идеалах x^2+yR, x^3R+yR, и т.д.)

• Единственным простым идеалом кольца $\mathbb{Z}/p^n\mathbb{Z}$, $p \in \mathbb{P}$, $n \in \mathbb{N}$, является $p\mathbb{Z}/p^n\mathbb{Z}$.

2. Неприводимые идеалы. Двусторонний идеал $I \trianglelefteq R$ называется **неприводимым**, если не существует пары идеалов $\mathfrak{a}, \mathfrak{b} \trianglelefteq R$, $\mathfrak{a}, \mathfrak{b} \neq I$, таких, что $\mathfrak{a} \cap \mathfrak{b} = I$. Так как $\mathfrak{ab} \subseteq \mathfrak{a} \cap \mathfrak{b}$, то каждый простой идеал кольца R неприводим.

Задача. Каждый двусторонний идеал I кольца R есть пересечение всех содержащих его неприводимых идеалов.

Решение. По теореме об изоморфизме вопрос сводится к кольцу R/I так что без потери общности можно считать $I = 0$. Пусть $x \neq 0$. Множество X всех идеалов, не содержащих x, индуктивно упорядочено. Из леммы Куратовского-Цорна вытекает, что в этом множестве существует максимальный элемент $\text{Ш} \trianglelefteq R$, $x \notin \text{Ш}$. Ясно, что Ш неприводим. Но это и значит, что пересечение всех неприводимых идеалов кольца R равно 0.

Лемма. *Пусть* $\mathfrak{p}_1, \ldots, \mathfrak{p}_n \in \operatorname{Spec}(R)$ *и* $I \leq \mathfrak{p}_1 \cup \ldots \cup \mathfrak{p}_n$. *Тогда* $I \leq \mathfrak{p}_i$ *для какого-то* $i = 1, \ldots, n$.

Доказательство. Выбросив, если нужно, несколько $\mathfrak{p}_i$, можно считать, что между $\mathfrak{p}_i$ нет включений, т.е. $\mathfrak{p}_i \not\leq \mathfrak{p}_j$ для $i \neq j$. Предположим, вопреки ожиданиям, что $I \not\leq \mathfrak{p}_i$ для всех i. Тогда $I \cap \bigcap_{i \neq j} \mathfrak{p}_i \not\leq \mathfrak{p}_j$. Для каждого $j = 1, \ldots, n$ выберем $x_j \in \left(I \cap \bigcap_{i \neq j} \mathfrak{p}_i\right) \setminus \mathfrak{p}_j$ и рассмотрим $x = x_1 + \ldots + x_n$. Тогда $x \in I$, но $x \equiv x_i \not\equiv 0 \pmod{\mathfrak{p}_i}$, так что $x \notin \cup \mathfrak{p}_i$, вопреки предположению.

3. Функториальность. Пусть $\phi : R \longrightarrow S$ – гомоморфизм коммутативных колец. Сейчас мы построим отображение $\phi^* : \operatorname{Spec}(S) \longrightarrow \operatorname{Spec}(R)$.

Теорема. *Для любого идеала* $\mathfrak{p} \in \operatorname{Spec}(S)$ *идеал* $\phi^{-1}(\mathfrak{p}) \in \operatorname{Spec}(R)$ *прост. Если* ϕ *сюръективен, то отображение* $\mathfrak{p} \mapsto \phi^{-1}(\mathfrak{p})$ *устанавливает взаимно однозначное соответствие между всеми простыми идеалами кольца* S *и теми простыми идеалами кольца* R, *которые содержат* $\operatorname{Ker}(\phi)$.

Доказательство. По теореме о гомоморфизме для любого идеала $I \trianglelefteq S$ гомоморфизм ϕ определяет вложение $R/\phi^{-1}(I)$ в S/I. Ненулевое подкольцо области целостности само является областью целостности. Это вложение является биекцией, если ϕ сюръективно.

Таким образом, мы можем положить $\phi^*(\mathfrak{p}) = \phi^{-1}(\mathfrak{p})$. Это сопоставление **функториально**, точнее, **контравариантно** в том смысле, что и $(\phi\psi)^* = \psi^*\phi^*$. Отметим важнейший частный случай предшествующего результата.

Следствие. *Если* S/R *– расширение колец и* $\mathfrak{p} \in \operatorname{Spec}(S)$, *то* $\mathfrak{p} \cap R \in \operatorname{Spec}(R)$.

Предостережение. Аналогичное утверждение для максимальных идеалов не имеет места!!! Например, $\mathbb{Z} \leq \mathbb{Q}$, причем идеал $0 \trianglelefteq \mathbb{Q}$ максимален, но его прообраз $0 \trianglelefteq \mathbb{Z}$ прост, но не является больше максимальным. Поэтому инвариантный геометрический смысл имеет лишь пространство простых идеалов, а вовсе не пространство максимальных идеалов. В действительности, максимальные идеалы отвечают **замкнутым точкам** спектра, но, как правило, в этом пространстве имеется и много незамкнутых точек.

§ 4. Максимальные идеалы

1. Максимальные идеалы. Ясно, что ненулевой главный идеал $\mathfrak{p} = (f)$ в том и только том случае прост, когда f – простой элемент кольца R. Таким образом, если $\mathfrak{p}$ – простой идеал в кольце главных идеалов R, то либо $\mathfrak{p} = 0$, либо $\mathfrak{p} = (f)$ где f – неприводимый элемент кольца R. В действительности, как мы знаем из Главы ?, идеал (f), порожденный неприводимым элементом, максимален. Напомним определение.

Определение. *Идеал $\mathfrak{m}$ кольца R называется* **максимальным***, если он собственный, и не содержится ни в каком большем собственном идеале, т.е. иными словами, $\mathfrak{m} < R$ и не существует идеала I такого, что $\mathfrak{m} < I < R$.*

Если $\mathfrak{m}$ – максимальный идеал R и $x \in R \setminus \mathrm{maxi}$, то $I + (f) = R$. Таким образом все ненулевые элементы $x + \mathfrak{m}$ кольца $R/\mathfrak{m}$ обратимы, или, что то же самое, $R/\mathfrak{m}$ является полем.

Лемма. *Фактор-кольцо $R/\mathfrak{m}$ в том и только том случае является полем, когда идеал $\mathfrak{m}$ максимален.*

Так как поле является областью целостности, то каждый максимальный идеал прост.

- Идеал 0 максимален в поле K.
- Идеал $p\mathbb{Z}$, $p \in \mathbb{P}$, максимален в $\mathbb{Z}$.
- Единственными максимальными идеалами в K^n являются $\mathfrak{m}_i = \{x = (x_1, \dots, x_n) \in K^n \mid x_i = 0\}$.

Задача. Докажите, что в булевом кольце всякий простой идеал максимален.

Задача. Пусть X – компактное хаусдорфово пространство. Докажите, что в кольце $C(X)$ всякий максимальный идеал имеет вид

$$\mathfrak{m}_x = \{f \in C(X) \mid f(x) = 0\}.$$

2. Идеал точки в аффинном пространстве. Приведем *важнейший* пример максимального идеала. Над любым полем K для любой точки $c = (c_1, \dots, c_n) \in K^n$ идеал

$$\mathfrak{m}_c = R(x_1 - c_1) + \dots + R(x_n - c_n)$$

в кольце многочленов $R = K[x_1, \dots, x_n]$, порожденный многочленами $x_i - c_i$, максимален. Для этого достаточно заметить, что $R/\mathfrak{m}_c \equiv K$. В действительности, $\mathfrak{m}_c$ это в точности ядро гомоморфизма эвалюации $\mathrm{ev}_c : R \longrightarrow K$, $f \mapsto f(c)$. Иными словами, многочлен $f \in R$ в том и только том случае принадлежит $\mathfrak{m}_c$, когда $f(c) = 0$.

Одна из форм **теоремы Гильберта о нулях** утверждает, что если поле K алгебраически замкнуто, то верно и обратное: каждый максимальный идеал в кольце $R = K[x_1, \dots, x_n]$ имеет вид $\mathfrak{m}_c$ для некоторой точки $c \in R^n$.

Задача. Пусть K – произвольное поле, $\mathfrak{m} \trianglelefteq R = K[x_1, \dots, x_n]$ – максимальный идеал. (ГДЕ это используется в доказательстве – ДОЛЖНО ??) Тогда $\mathfrak{m}$ порождается n элементами $\mathfrak{m} = (f_1, \dots, f_n)$, причем f_i можно выбрать так, чтобы $f_i \in K[x_1, \dots, x_i]$.

Решение. Индукция по n. Пусть $\mathfrak{n} = \mathfrak{m} \cap A$, где $A = K[x_1, \ldots, x_{n-1}]$. ТОгда нужно показать, что $\mathfrak{m}/\mathfrak{n}R$ – главный идеал в $R/\mathfrak{n}R \cong (A/\mathfrak{n})[x_n]$ – ЗАКОНЧИТЬ!!

3. Существование максимальных идеалов. Следующее утверждение в действительности эквивалентно аксиоме выбора.

Теорема Крулля. *Каждый собственный левый идеал I кольца R содержится в некотором максимальном левом идеале.*

Доказательство. Пусть Ω – множество собственных левых идеалов кольца R, содержащих I упорядочененое по включению. Так как $I \in X$, то X непусто. Покажем, что X индуктивно упорядочено, т.е. всякое линейно упорядоченного подмножества в Ω имеет мажоранту. В самом деле, если $X \subseteq \Omega$ – такое подмножество, мы покажем, что $C = \cup A$, $A \in X$, собственный левый идеал. Ясно, что тогда C мажоранта этой цепи. В самом деле, если $x \in C$, то найдется $A \in X$ такое, что $x \in A$. Тогда для любого $y \in R$, имеем $yx \in A \leq C$. Если $x, y \in C$, то найдутся $A, B \in X$ такие, что $x \in A$, $y \in B$. Один из идеалов A, B содержится во втором. Пусть, например, $A \leq B$. Тогда $x, y \in B$ и, значит, $x - y \in B \leq C$. Это значит, что C действительно является левым идеалом. Так как $1 \notin A$ для всех $A \in Y$, то $1 \notin$ C, поэтому C собственный идеал и, значит, C $\in \Omega$. По лемме Куратовского-Цорна существует максимальный элемент в Ω, а это как раз и есть максимальный левый идеал, содержащий I.

Предостережение. Обратите внимание, что мы существенно использовали в доказательстве наличие 1 в R. Без этого предположения теорема Крулля **безнадежно** неверна! Скажем, если R – кольцо с нулевым умножением, то левые идеалы – это в точности подгруппы аддитивной группы. Ясно, что легко построить абелевы группы вообще без максимальных подгрупп, таковы, скажем $\mathbb{Q}$ или $\mathbb{Q}/\mathbb{Z}$.

Следствие. *Если R – коммутативное кольцо, то каждый собственный идеал содержится в некотором максимальном идеале.*