Задача Эрдёша – Гинзбурга – Зива и ее окрестности

А. М. Райгородский

1. Введение

В этой статье мы поговорим об одной из самых красивых задач современной комбинаторики. Ее предложили в далеком 1961 году П. Эрдёш, А. Гинзбург и А. Зив, которым удалось доказать следующее замечательное утверждение.

Теорема 1 (Эрдёш, Гинзбург, Зив, 1961). *Из любого множества $A = \{a_1, \ldots, a_{2n-1}\}$, состоящего из целых чисел, можно выбрать n чисел, сумма которых делится на n.*

Разумеется, числа в множестве A из формулировки теоремы не обязаны быть различными. Более того, нам лишь нужно знать, какой остаток от деления на n дают эти числа. Понятно сразу, что в некотором смысле теорема 1 неулучшаема: величину $2n-1$ в ней нельзя заменить на $2n-2$. Действительно, если взять числа

$$a_1 = \ldots = a_{n-1} = 0, \qquad a_n = \ldots = a_{2n-2} = 1,$$

то среди них уже не будет n чисел, сумма которых делится на n.

С одной стороны, теорема 1 послужила отправной точкой для развития целого направления в комбинаторной математике. Науку, которая выросла из нее, принято относить к «аддитивной комбинаторике», т. е. к разделу комбинаторики, в котором изучаются задачи, связанные с отысканием различных множеств чисел, чьи суммы обладают теми или иными интересными свойствами.

С другой стороны, теорема 1 весьма «олимпиадна», и потому ее сразу полюбили популяризаторы математики. По-видимому, особенный всплеск ее популярности пришелся на начало 70-х годов XX века, когда, например, в «Кванте» была опубликована статья с ее доказательством (см. [4]). Позже в «олимпиадной» среде задачу Эрдёша – Гинзбурга – Зива слегка подзабыли, о чем свидетельствует еще одна недавняя статья в том же «Кванте» (см. [5]).

В этой статье нам хочется рассказать о той удивительной науке, которая выросла из теоремы 1 и которую незаслуженно мало знают у нас в стране. Но прежде всего мы изложим одно очень красивое доказательство теоремы, основанное на применении малой теоремы Ферма. Оно не самое простое из многочисленных известных доказательств (см. [4–7]), но для наших целей оно будет исключительно полезно: именно его главная идея ляжет в основу всех дальнейших изысканий.

2. Доказательство теоремы 1

Ниже мы рассмотрим лишь случай, когда $n = p$, где p — простое число. Переход к общему случаю достаточно прост, и о нем можно прочесть в статье [5]. Поскольку мы собираемся пользоваться малой теоремой Ферма, в первом параграфе этого раздела мы напомним читателю все необходимые сведения из классической арифметики, а также приведем пару утверждений о делимости биномиальных коэффициентов. Во втором параграфе мы изложим доказательство теоремы 1.

2.1. Немного терминологии и несколько вспомогательных фактов

Выражение «a делится на p» мы часто будем записывать в виде «$a \equiv 0 \pmod p$» (читается «a сравнимо с нулем по модулю p»). Вообще, запись «$a \equiv b \pmod p$» подразумевает, что $b - a$ делится на p, т. е. у чисел a и b одинаковые остатки от деления на p.

Пусть дано множество чисел $A = \{a_1, \ldots, a_k\}$. Рассмотрим множество индексов $I = \{i_1, \ldots, i_l\} \subseteq \{1, \ldots, k\}$. Тогда сумма вида $\sum\limits_{i \in I} a_i$ — это просто $a_{i_1} + \ldots + a_{i_l}$.

Количество элементов в данном множестве X обозначим через $|X|$. Под записью

$$\sum_{\substack{I \subset \{1,\ldots,k\}: \\ |I|=l}} \sum_{i \in I} a_i$$

будем понимать сумму по всем l-элементным подмножествам множества $\{1, \ldots, k\}$ уже известных нам сумм $\sum\limits_{i \in I} a_i$. Например,

$$\sum_{\substack{I \subset \{1,2,3,4\}: \\ |I|=2}} \sum_{i \in I} a_i = (a_1+a_2)+(a_1+a_3)+(a_1+a_4)+(a_2+a_3)+(a_2+a_4)+(a_3+a_4).$$

В новых терминах теорема Эрдёша – Гинзбурга – Зива звучит так: *для любого множества* $A = \{a_1, \ldots, a_{2p-1}\}$, *состоящего из целых чисел*,

существует такое множество $I \subset \{1, \dots, 2p-1\}$, *что* $|I| = p$ *и* $\sum\limits_{i \in I} a_i \equiv 0 \pmod p$.

Хорошо известна следующая теорема.

Теорема 2 (Малая теорема Ферма). *Если* $a \not\equiv 0 \pmod p$, *то* $a^{p-1} \equiv 1 \pmod p$.

Доказательство теоремы 2 можно найти в любой книжке по арифметике или теории чисел. Например, оно есть в книге [3]. Однако мы приведем ниже одно симпатичное рассуждение, которое также доказывает теорему и вместе с тем идейно близко ко всему, о чем пойдет речь дальше.

Доказательство теоремы 2. Достаточно показать, что $a^p \equiv a \pmod p$, если $a > 0$. Имеем

$$a^p = (1 + 1 + \ldots + 1)^p = 1^p + 1^p + \ldots + 1^p + \sum \frac{p!}{n_1! \cdot \ldots \cdot n_a!} 1^{n_1} \cdot \ldots \cdot 1^{n_a},$$

где последнее суммирование идет по всем упорядоченным наборам чисел $n_1 < p, \dots, n_a < p$, для которых $n_1 + \ldots + n_a = p$. Поскольку число p простое, величина $\dfrac{p!}{n_1! \cdot \ldots \cdot n_a!}$ делится на p (ведь в знаменателе нет ни одного числа, делящегося на p). Таким образом,

$$a^p \equiv 1^p + 1^p + \ldots + 1^p \pmod p, \qquad a^p \equiv a \pmod p,$$

и теорема доказана.

Нам потребуются следующие два факта относительно биномиальных коэффициентов.

Утверждение 1. *Для любого простого числа* p *выполнено* $C_{2p-1}^p \equiv 1 \pmod p$.

Утверждение 2. *Для любого простого числа* p *и любого* $q \in \{1, \dots, p-1\}$ *выполнено* $C_{2p-1-q}^{p-q} \equiv 0 \pmod p$.

Докажем, для примера, утверждение 1; второе утверждение доказывается аналогично.

Доказательство утверждения 1. Посмотрим на известное тождество

$$C_{2p}^0 + C_{2p}^1 + \ldots + C_{2p}^p + \ldots + C_{2p}^{2p} = 4^p.$$

В нем первое и последнее слагаемые равны единице. Остальные слагаемые, кроме C_{2p}^p, имеют вид $C_{2p}^i = \dfrac{(2p)!}{i!(2p-i)!}$ с $i \neq p$. Ясно, что в каждом из них числитель делится на p^2, а знаменатель не делится. Значит, $C_{2p}^i \equiv 0 \pmod p$, т. е.

$$C_{2p}^0 + C_{2p}^1 + \ldots + C_{2p}^p + \ldots + C_{2p}^{2p} \equiv 2 + C_{2p}^p \pmod p.$$

В то же время по малой теореме Ферма $4^p \equiv 4 \pmod p$. Получается, что $C_{2p}^p \equiv 2 \pmod p$. Но

$$C_{2p-1}^p = \frac{1}{2} C_{2p}^p \equiv 1 \pmod p,$$

и утверждение доказано.

Отметим, что в связи с теоремой 2 и утверждениями 1 и 2 очень полезны статьи [1, 2].

2.2. Само доказательство

Будем работать с переформулировкой теоремы 1, которую мы привели в параграфе 2.1. Предположим, что, вопреки ее утверждению, для любого множества $I \subset \{1, \ldots, 2p-1\}$, имеющего p элементов, $\sum\limits_{i \in I} a_i \not\equiv 0 \pmod p$. Тогда по малой теореме Ферма

$$\Big(\sum_{i \in I} a_i \Big)^{p-1} \equiv 1 \pmod p.$$

Следовательно, с учетом утверждения 1

$$S = \sum_{\substack{I \subset \{1,\ldots,2p-1\}: \\ |I|=p}} \Big(\sum_{i \in I} a_i \Big)^{p-1} \equiv C_{2p-1}^p \equiv 1 \pmod p.$$

С другой стороны, раскроем скобки в выражении $\Big(\sum\limits_{i \in I} a_i \Big)^{p-1}$. Понятно, что каждое слагаемое в результирующей сумме будет иметь вид

$$a_{i_1}^{k_{i_1}} \cdot \ldots \cdot a_{i_q}^{k_{i_q}},$$

где $1 \leqslant q \leqslant p-1$, $k_{i_1} \geqslant 1$, $\ldots$, $k_{i_q} \geqslant 1$, а $k_{i_1} + \ldots + k_{i_q} = p-1$. Значит, можно написать

$$S = \sum_{\substack{I \subset \{1,\ldots,2p-1\}: \\ |I|=p}} \sum a_{i_1}^{k_{i_1}} \cdot \ldots \cdot a_{i_q}^{k_{i_q}}.$$

Переставим в последней записи порядки суммирования. Иными словами, сперва зафиксируем произвольное выражение

$$a_{i_1}^{k_{i_1}} \cdot \ldots \cdot a_{i_q}^{k_{i_q}},$$

а затем посчитаем, сколько есть множеств $I \subset \{1, \ldots, 2p-1\}$ из p элементов, для каждого из которых это выражение могло возникнуть при раскрытии скобок в записи $\Big(\sum\limits_{i \in I} a_i \Big)^{p-1}$. Нетрудно видеть, что такие множества I должны содержать набор $\{i_1, \ldots, i_q\}$ в качестве подмножества,

т. е. их ровно C_{2p-1-q}^{p-q} штук. Ввиду утверждения 2 это количество делится на p. Иначе говоря, каждое слагаемое в сумме S делится на p, а стало быть, $S \equiv 0 \pmod{p}$.

Однако выше, исходя из предположения противного, мы получили $S \equiv \equiv 1 \pmod{p}$. Имеем противоречие, и теорема 1 доказана.

3. Первое обобщение задачи Эрдёша – Гинзбурга – Зива: проблема Кемница

В этом разделе мы расскажем об исключительно красивом «двумерном» аналоге теоремы Эрдёша – Гинзбурга – Зива.

3.1. Немного об истории

В 1983 году А. Кемниц придумал одно очень естественное обобщение задачи Эрдёша – Гинзбурга – Зива. Он предложил вместо целых чисел рассматривать пары целых чисел, которые представляют собой точки на обычной плоскости, имеющие целые координаты.

Действительно, рассмотрим произвольное множество точек $A = \{(a_1, b_1), \ldots, (a_f, b_f)\}$. Сложим любые n точек из множества A покоординатно. Скажем, что полученная сумма (также являющаяся точкой на плоскости) делится на n, если обе ее координаты делятся на n. Иными словами, полученная при сложении точка должна иметь вид (na, nb), где $a, b \in \mathbb{Z}$. Вопрос: при каком f мы можем гарантировать наличие в A подмножества мощности n, сумма элементов которого делится на n?

Нетрудно видеть, что в качестве f нельзя взять $4n-4$. В самом деле, пусть A содержит $n-1$ точек $(0,0)$, столько же точек $(0,1)$, столько же точек $(1,0)$ и столько же $(1,1)$. Очевидно, в A нет n-элементных подмножеств с суммой элементов, делящейся на n.

Кемниц высказал гипотезу: $f = 4n-3$. Эта гипотеза оказалась весьма «крепким орешком». Для нескольких маленьких значений n ее «вручную» проверили сам Кемниц и Х. Харборт. Затем Н. Алон и М. Дубинер в 1993 году показали, что $f \leqslant 6n-5$. В 2000 году Л. Роньяи установил неравенство $f \leqslant 4n-2$, и лишь в 2003 году гипотезу «дожал» Х. Райхер.

И Роньяи, и Райхер действовали примерно одинаково, и в основе их рассуждений было далеко идущее обобщение идей, которые мы использовали в параграфе 2.2. Рассуждение Райхера довольно трудное (см. [6]), а рассуждение Роньяи — это, пожалуй, одно из самых изящных рассуждений в комбинаторике. Его-то мы здесь и изложим (опять же, для простых n). Но сперва нам понадобится ряд вспомогательных понятий и фактов.

3.2. Еще немного терминологии

Во-первых, мы будем писать $(a_1, b_1) + \ldots + (a_k, b_k) \equiv 0 \pmod p$, подразумевая, что сумма точек делится на p.

Во-вторых, нам понадобятся многочлены от нескольких переменных. Конечно, все отлично знают, что многочлен — это запись вида $f(x) = a_n x^n + \ldots + a_1 x + a_0$. Но это многочлен степени n от одной переменной x. Что же такое, например, многочлен степени n от двух переменных x, y? Если $n = 0$, то по-прежнему речь идет о произвольной константе. Если $n = 1$, то общий вид многочлена таков: $f(x, y) = a_1 x + a_2 y + a_3$. При $n = 2$ имеем $f(x, y) = a_1 x^2 + a_2 y^2 + a_3 xy + a_4 x + a_5 y + a_6$. И так далее. Иными словами, $f(x, y) = \sum a_{i,j} x^i y^j$, где $a_{i,j} \in \mathbb{R}$, а суммирование идет по некоторым парам натуральных чисел i и j. При этом степень многочлена — это $\max(i + j)$.

Аналогично определяется многочлен произвольного числа переменных:

$$f(x_1, \ldots, x_n) = \sum a_{i_1,\ldots,i_n} x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}.$$

Его степень равна $\max(i_1 + \ldots + i_n)$. Степень многочлена f обозначается $\deg f$. Числа $a_{i_1,\ldots,i_n} \in \mathbb{R}$ называются коэффициентами многочлена.

Далее, у малой теоремы Ферма есть следующее уточнение: *для любого простого $p > 2$ и любого $k \in \{1, \ldots, p-2\}$ найдется такое число a, что $a^k \not\equiv 0 \pmod p$ и $a^k \not\equiv 1 \pmod p$*. Смысл уточнения в том, что в малой теореме Ферма нельзя заменить степень $p - 1$ на $k \leqslant p - 2$. Оставим его читателю в качестве упражнения.

Наконец, совсем легкое упражнение состоит в том, чтобы доказать такое утверждение: *если a взаимно просто с p и x пробегает от 1 до p, то ax тоже пробегает от 1 до p (но, возможно, в ином порядке)*.

3.3. Теорема Варнинга – Шевалле

Имеет место следующий замечательный факт.

Теорема 3 (Варнинг – Шевалле). *Пусть f — многочлен от n переменных с целыми коэффициентами степени строго меньше n. Пусть, далее, p — простое число. Обозначим через N количество различных наборов $(x_1, \ldots, x_n)$, таких, что $x_i \in \{1, \ldots, p\}$ для каждого i и $f(x_1, \ldots, x_n) \equiv 0 \pmod p$. Тогда $N \equiv 0 \pmod p$.*

Теорема 3 означает, что число решений сравнения $f(x_1, \ldots, x_n) \equiv 0 \pmod p$ делится на p, лишь бы было выполнено неравенство $\deg f < n$. Довольно забавный и, на вид, нетривиальный факт.

Доказательство теоремы 3. С учетом малой теоремы Ферма

$$N \equiv \sum_{x_1=1}^{p} \dots \sum_{x_n=1}^{p} \left(1 - f^{p-1}(x_1, \dots, x_n)\right) \pmod{p}.$$

Значит, достаточно показать, что

$$\sum_{x_1=1}^{p} \dots \sum_{x_n=1}^{p} f^{p-1}(x_1, \dots, x_n) \equiv 0 \pmod{p}.$$

При возведении многочлена в степень и раскрытии скобок возникнет сумма выражений вида

$$a_{i_1,\dots,i_n} x_1^{i_1} \cdot \dots \cdot x_n^{i_n},$$

где $a_{i_1,\dots,i_n} \in \mathbb{Z}$, $i_1 + \dots + i_n < n(p-1)$ (ср. §2.2). Следовательно, достаточно убедиться в том, что для любых натуральных $i_1, \dots, i_n$, удовлетворяющих неравенству $i_1 + \dots + i_n < n(p-1)$, выполнено соотношение

$$S = \sum_{x_1=1}^{p} \dots \sum_{x_n=1}^{p} x_1^{i_1} \cdot \dots \cdot x_n^{i_n} \equiv 0 \pmod{p}.$$

Имеем

$$\sum_{x_1=1}^{p} \dots \sum_{x_n=1}^{p} x_1^{i_1} \cdot \dots \cdot x_n^{i_n} = \Big(\sum_{x_1=1}^{p} x_1^{i_1} \Big) \cdot \dots \cdot \Big(\sum_{x_n=1}^{p} x_n^{i_n} \Big).$$

Возможны два случая. В первом случае какое-то i_ν равно нулю. Тогда, очевидно, $S \equiv 0 \pmod{p}$, и теорема доказана.

Во втором случае все i_ν не меньше единицы. Но их сумма строго меньше $n(p-1)$. Значит, какое-то из них строго меньше $p-1$, т. е. оно лежит в пределах от 1 до $p-2$ (и, в частности, $p > 2$). Обозначим его k (ср. уточнение малой теоремы Ферма из параграфа 3.2). Исходя из упомянутого только что уточнения, возьмем то самое a, с которым $a^k \not\equiv 0 \pmod{p}$ и $a^k \not\equiv 1 \pmod{p}$. Тогда

$$a^k \cdot \sum_{x_\nu=1}^{p} x_\nu^k = \sum_{x_\nu=1}^{p} (a x_\nu)^k = \sum_{x_\nu=1}^{p} x_\nu^k.$$

Последнее равенство выполнено ввиду «легкого упражнения», завершающего параграф 3.2. В итоге $\sum\limits_{x_\nu=1}^{p} x_\nu^k \equiv 0 \pmod{p}$, а стало быть, $S \equiv 0 \pmod{p}$. Теорема доказана.

У доказанной теоремы есть очевидное следствие.

Следствие из теоремы 3. *Пусть f — многочлен от n переменных с целыми коэффициентами степени строго меньше n. Пусть, далее, p —*

простое число. Если набор $(p,\dots,p)$ *удовлетворяет сравнению* $f(p,\dots,p)\equiv 0 \pmod p$, *то найдется и набор* $(x_1,\dots,x_n)$, *в котором не все* x_i *делятся на* p *и с которым также* $f(x_1,\dots,x_n)\equiv 0 \pmod p$.

Наконец, имеет место следующее обобщение, доказательство которого мы предоставим читателю.

Обобщение следствия из теоремы 3. *Пусть* $f_1,\dots,f_m$ — *многочлены от* n *переменных с целыми коэффициентами, сумма степеней которых строго меньше* n. *Пусть, далее,* p — *простое число. Если набор* $(p,\dots,p)$ *одновременно удовлетворяет всем сравнениям*

$$f_1(p,\dots,p)\equiv 0 \pmod p,\qquad \dots,\quad f_m(p,\dots,p)\equiv 0 \pmod p,$$

то найдется и набор $(x_1,\dots,x_n)$, *в котором не все* x_i *делятся на* p *и с которым также*

$$f_1(x_1,\dots,x_n)\equiv 0 \pmod p,\qquad \dots,\qquad f_m(x_1,\dots,x_n)\equiv 0 \pmod p.$$

3.4. Основная лемма

Сейчас мы с помощью теоремы Варнинга – Шевалле установим следующий важный факт.

Лемма 1. *Пусть* p — *простое число. Пусть, далее,* $(a_1,b_1),\dots,(a_{3p},b_{3p})$ — *произвольные точки с целыми координатами, сумма которых делится на* p. *Тогда найдется такое множество* $I\subset\{1,\dots,3p\}$, *что* $|I|=p$ *и* $\sum\limits_{i\in I}(a_i,b_i)\equiv 0 \pmod p$.

Доказательство леммы 1. Зафиксируем произвольные точки (a_1,b_1), $\dots$, (a_{3p},b_{3p}), удовлетворяющие соотношению $\sum\limits_{i=1}^{3p}(a_i,b_i)\equiv 0 \pmod p$. Рассмотрим три многочлена

$$f_1(x_1,\dots,x_{3p-1})=\sum_{i=1}^{3p-1}a_i x_i^{p-1},$$

$$f_2(x_1,\dots,x_{3p-1})=\sum_{i=1}^{3p-1}b_i x_i^{p-1},$$

$$f_3(x_1,\dots,x_{3p-1})=\sum_{i=1}^{3p-1}x_i^{p-1}.$$

Эти многочлены зависят от $3p-1$ переменных, и сумма их степеней, равная $3p-3$, строго меньше величины $3p-1$. Более того, набор $(p,\dots,p)$, очевидно, таков, что на нем все три многочлена принимают значения, делящиеся

на p. Значит, в силу обобщения следствия из теоремы 3, существует набор $(x_1, \dots, x_{3p-1})$, в котором не все числа делятся на p и на котором, тем не менее, f_1, f_2, f_3 одновременно обнуляются по модулю p. Пусть J — множество, состоящее из всех индексов $i \in \{1, \dots, 3p-1\} \subset \{1, \dots, 3p\}$, с которыми $x_i \not\equiv 0 \pmod p$.

Мы знаем, что $f_1(x_1, \dots, x_n) \equiv 0 \pmod p$. Стало быть, с учетом малой теоремы Ферма, имеем

$$\sum_{i=1}^{3p-1} a_i x_i^{p-1} \equiv \sum_{i \in J} a_i x_i^{p-1} \equiv \sum_{i \in J} a_i \equiv 0 \pmod p.$$

Аналогично получаем

$$\sum_{i \in J} b_i \equiv 0 \pmod p,$$

т. е.

$$\sum_{i \in J} (a_i, b_i) \equiv 0 \pmod p.$$

Наконец, из соотношения $f_3(x_1, \dots, x_n) \equiv 0 \pmod p$ вытекает сравнение $|J| \equiv 0 \pmod p$, т. е. либо $|J| = p$ и тогда лемма доказана (полагаем $I = J$), либо $|J| = 2p$. В последнем случае берем $I = \{1, \dots, 3p\} \setminus J$. Мощность множества I равна p, и

$$\sum_{i \in I} (a_i, b_i) = \sum_{i=1}^{3p} (a_i, b_i) - \sum_{i \in J} (a_i, b_i) \equiv 0 \pmod p.$$

Лемма доказана.

3.5. Доказательство неравенства $f \leqslant 4p - 2$

Для краткости введем обозначение $m = 4p - 2$. Зафиксируем произвольные m точек $(a_1, b_1), \dots, (a_m, b_m)$. Мы хотим доказать существование такого множества $I \subset \{1, \dots, m\}$, что $|I| = p$ и $\sum\limits_{i \in I} (a_i, b_i) \equiv 0 \pmod p$. Предположим, однако, что такого множества нет, т. е. для любого $I \subset \{1, \dots, m\}$, состоящего из p элементов, $\sum\limits_{i \in I} (a_i, b_i) \not\equiv 0 \pmod p$. С учетом леммы 1 можно предположить даже больше: для любого $I \subset \{1, \dots, m\}$, состоящего из p или $3p$ элементов, $\sum\limits_{i \in I} (a_i, b_i) \not\equiv 0 \pmod p$. Постараемся прийти к противоречию.

Обозначим через σ_p многочлен

$$\sigma_p(x_1, \dots, x_n) = \sum_{\substack{I \subset \{1, \dots, n\}: \\ |I| = p}} \prod_{i \in I} x_i.$$

Например,

$$\sigma_2(x_1, x_2, x_3, x_4) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4.$$

Положим

$$g(x_1, \ldots, x_m) = \\ = \Big(\Big(\sum_{i=1}^{m} a_i x_i\Big)^{p-1} - 1\Big)\Big(\Big(\sum_{i=1}^{m} b_i x_i\Big)^{p-1} - 1\Big)\Big(\Big(\sum_{i=1}^{m} x_i\Big)^{p-1} - 1\Big)\cdot \\ \cdot \big(\sigma_p(x_1, \ldots, x_m) - 2\big).$$

Будем подставлять в многочлен g произвольные наборы $(x_1, \ldots, x_m)$, состоящие из нулей и единиц. Рассмотрим несколько отдельных случаев.

Пусть $x_1 + \ldots + x_m \in \{p, 3p\}$. В этом случае обозначим через I множество всех индексов i, с которыми $x_i = 1$, так что $|I| \in \{p, 3p\}$. Сделанное нами предположение противного означает, что для такого множества I, как и для любого другого аналогичного множества, выполнено $\sum\limits_{i\in I}(a_i, b_i) \not\equiv 0 \pmod p$. Следовательно, либо $\sum\limits_{i\in I} a_i \not\equiv 0 \pmod p$ и тогда за счет малой теоремы Ферма обнуляется (по модулю p) первая скобка в записи многочлена g, либо $\sum\limits_{i\in I} b_i \not\equiv 0 \pmod p$ и тогда обнуляется вторая скобка в записи многочлена g. Иными словами, $g(x_1, \ldots, x_m) \equiv 0 \pmod p$.

Пусть $x_1 + \ldots + x_m = 2p$. В этом случае $\sigma_p(x_1, \ldots, x_m) = C_{2p}^p$. Доказывая утверждение 1 (см. §2.1), мы по ходу дела поняли, что $C_{2p}^p \equiv 2 \pmod p$. Значит, обнуляется четвертая скобка в записи g, т. е. снова $g(x_1, \ldots, x_m) \equiv \equiv 0 \pmod p$.

Пусть, $x_1 + \ldots + x_m \not\equiv 0 \pmod p$. В этом случае обнуляется третья скобка, и опять $g(x_1, \ldots, x_m) \equiv 0 \pmod p$.

Остается лишь случай, когда $x_1 = \ldots = x_m = 0$. Здесь уже $g(x_1, \ldots, x_m) = 2$.

Теперь представим себе, что мы раскрыли все четыре скобки в записи g. Разумеется, возникнет, как обычно, сумма выражений вида $x_{i_1}^{k_{i_1}} \cdot \ldots \cdot x_{i_r}^{k_{i_r}}$, где $k_{i_1} \geqslant 1, \ldots, k_{i_r} \geqslant 1$. Формально заменим каждое такое выражение выражением $x_{i_1} \cdot \ldots \cdot x_{i_r}$. Получится новый многочлен g', значения которого на наборах из нулей и единиц в точности совпадают со значениями исходного многочлена g.

Итак, у нас есть многочлен g', который обнуляется на всех наборах нулевых и единичных аргументов, кроме набора из одних нулей, на котором его значение равно 2. При этом каждая переменная входит в этот многочлен в степени не выше 1. Читателю предлагается доказать, что тогда

$$g'(x_1, \ldots, x_m) = 2 \cdot (1 - x_1) \cdot (1 - x_2) \cdot \ldots \cdot (1 - x_m).$$

Ясно, что $\deg g' = m$. В то же время

$$\deg g' \leqslant \deg g = (p-1)+(p-1)+(p-1)+p = 4p-3 < 4p-2 = m.$$

Это и есть искомое противоречие, завершающее доказательство неравенства $f \leqslant 4p-2$.

4. Второе обобщение задачи Эрдёша – Гинзбурга – Зива

По прочтении всего предшествующего текста сразу возникает вопрос: ну, хорошо, мы изучили наборы целых чисел и пар целых чисел; а почему бы не рассмотреть задачу про наборы последовательностей целых чисел произвольной длины d? Разумеется, этим много занимались, и ниже мы об этом поговорим.

4.1. Постановка задачи

Итак, даны числа n и d. Обозначим через $f(n,d)$ наименьшее f, при котором для любого множества A, состоящего из f последовательностей целых чисел $\left(a_1^1,\dots,a_d^1\right)$, $\dots$, $\left(a_1^f,\dots,a_d^f\right)$, найдется такое подмножество $I \subset \{1,\dots,f\}$, что $|I| = n$ и сумма $\sum\limits_{i\in I}\left(a_1^i,\dots,a_d^i\right)$ делится на n (по каждой «координате»). Ясно, что $f(n,1) = 2n-1$ и $f(n,2) = 4n-3$. Что же будет, если $d \geqslant 3$?

Прежде всего очень легко обобщить примеры, свидетельствовавшие о том, что $f(n,1) \geqslant 2n-1$ и $f(n,2) \geqslant 4n-3$. Для этого надо взять все последовательности из d нулей и единиц и каждую такую последовательность проитерировать $n-1$ раз. Получится множество A из $f = 2^d(n-1)$ последовательностей, в котором нет n-элементных подмножеств с суммой элементов, делящейся на n. Иными словами, $f(n,d) \geqslant 2^d(n-1)+1$.

При $d=1$ и $d=2$ оценка $f(n,d) \geqslant 2^d(n-1)+1$ оказывалась точной, т. е. удавалось показать, что $f(n,d) = 2^d(n-1)+1$. Естественное предположение состоит в том, что последнее равенство выполнено для всех d. Однако и тут нас ожидает сюрприз: уже при $d=3$ предположение неверно!

Действительно, рассмотрим девять последовательностей

$$(2,1,2),\quad (0,0,0),\quad (0,0,1),\quad (0,1,0),\quad (0,1,1),$$
$$(1,0,0),\quad (1,0,1),\quad (1,1,2),\quad (1,2,2).$$

Возьмем каждую из этих последовательностей дважды, в результате чего образуется множество A из восемнадцати последовательностей. Нетрудно убедиться в том, что в A нет трех последовательностей, сумма которых делится на 3. Иными словами, $f(3,3) \geqslant 19$. В то же время неравенство $f(n,d) \geqslant 2^d(n-1)+1$ дает лишь оценку $f(3,3) \geqslant 17$.

Описанная элементарная конструкция принадлежит Х. Харборту. В 2004 году Х. Элсхольц сумел обобщить ее, доказав в итоге, что $f(n,3) \geqslant$ $\geqslant 9n-8$ при всех нечетных n (см. [8]).

Видно, что задача становится исключительно трудной. Не ясно даже, каковы предположительные значения величины $f(n,d)$. В следующем параграфе мы перечислим некоторые из известных результатов.

4.2. Некоторые известные результаты

Прежде всего заметим, что всегда $f(n,d) \leqslant n^d(n-1)+1$. В самом деле, пусть A — множество последовательностей, состоящее из $f = n^d+1$ элементов. Тогда по принципу Дирихле в этом множестве есть две последовательности, сравнимые по модулю n, т. е. такие последовательности $(a_1,\ldots,a_d)$, $(b_1,\ldots,b_d)$, что $a_i \equiv b_i \pmod n$ для любого i. Значит, если в A не меньше $n^d(n-1)+1$ элементов, то в A есть, как минимум, n попарно сравнимых последовательностей. Сумма любых n из них и делится на n.

Разумеется, оценка $f(n,d) \leqslant n^d(n-1)+1$ крайне далека от оценки $f(n,d) \geqslant 2^d(n-1)+1$. Однако, ввиду результатов Харборта и Элсхольца, описанных в предыдущем параграфе, трудно сказать даже, к какой из этих оценок ближе истинное значение величины $f(n,d)$.

Гораздо больших продвижений к настоящему времени удалось достичь в вопросе уточнения верхней оценки. Так, Алон и Дубинер доказали, что $f(n,d) \leqslant c(d)n$, где $c(d)$ — некоторая функция, зависящая от d, но не зависящая от n. Понятно, что при больших n (возможно, очень больших) оценка Алона – Дубинера значительно сильнее неравенства $f(n,d) \leqslant$ $\leqslant n^d(n-1)+1$. Тем не менее, $c(d)$ гораздо больше, нежели 2^d, а потому до решения задачи еще крайне далеко (доказано, например, что $c(d) \geqslant 2^d 1.125^{[d/3]}$).

Впрочем, при n, равном степени двойки, задача решена полностью. В 1973 году Харборт показал (см. [9]), что для любого a выполнено $f\,(2^a,d) = 2^d\,(2^a-1)+1$.

В остальном, известны лишь частные результаты. Перечислим их:

- $f(3,3) = 19$ (Харборт, 1973);
- $f(3,4) = 41$ (Дж. Пеллегрино, Т. Браун, Й. Булер, Дж. Бреннер, Кемниц, 1983);
- $f(3,5) = 91$ (И. Эдель, С. Ферре, И. Ландьев, Л. Сторм, 2002);
- $225 \leqslant f(3,6) \leqslant 229$ (Эдель, Ферре, Ландьев, Сторм, 2002);
- $f(3,18) \geqslant 300 \cdot 2^{12}$ (Р. Грехем, П. Франкл, В. Рёдль, 1987);
- $f(3,d) \geqslant 2.217389^d$ при достаточно больших d (Эдель, 2004);

- $f(3,d) \leqslant 2\frac{3^d}{d}$ (Р. Мешулам, 1995);
- $f(n,d) \geqslant 2.08^d$ для нечетных n и достаточно больших d (Элсхольц, 2004).

Список литературы

[1] Э.Б. Винберг. *Удивительные арифметические свойства биномиальных коэффициентов* // Математическое просвещение. Третья серия, вып. 12. 2008. С. 33–42.

[2] Э.Б. Винберг. *Малая теорема Ферма и ее обобщения* // Математическое просвещение. Третья серия, вып. 12. 2008. С. 43–54.

[3] И.М. Виноградов. *Основы теории чисел*. Москва – Ижевск: НИЦ «Регулярная и хаотическая динамика». 2003.

[4] *Задачник Кванта. Математика.* // Квант, №9. 1970. С. 49; Квант, №7. 1971. С. 30.

[5] А. Толпыго. *Об одной забытой задаче* // Квант, №2. 2010. С. 45–49.

[6] А.М. Райгородский. *Линейно-алгебраический метод в комбинаторике*. М.: МЦНМО. 2007.

[7] N. Alon, M. Dubiner. *Zero-sum sets of prescribed size* // Combinatorics: Paul Erdős is Eighty. Bolyai Society, Mathematical Studies, Keszthely, Hungary, 1993. P. 33–50.

[8] C. Elsholtz. *Lower bounds for multidimensional zero sums* // Combinatorica. Vol. 24. 2004. P. 351–358.

[9] H. Harborth. *Ein Extremalproblem für Gitterpunkte* // J. Reine Angew. Math. Bd. 262/263. 1973. S. 356–360.

А. М. Райгородский