РЕАЛИЗУЕМОСТЬ ГИПЕРГРАФОВ: ПРИМЕРЫ И АЛГОРИТМЫ [1]

А. Зимин, А. Рухович, А. Скопенков и М. Скопенков

Представляют: А. Зимин, Д. Зунг, А. Рухович, А. Скопенков и Г. Челноков

1 Задачи до промежуточного финиша

О чем этот цикл задач

Хорошо известна проблема реализуемости графов в плоскости: можно ли граф расположить на плоскости так, чтобы его ребра не пересекались и не самопересекались? Этот цикл задач посвящен реализуемости двумерных аналогов графов (называемых гиперграфами) в трехмерном и четырехмерном пространствах. Важнейшие результаты — простые доказательства нереализуемости в четырехмерном пространстве полного гиперграфа с 7 вершинами и (решение проблемы Менгера) декартова произведения $K_5 \times K_5$. [2] См. задачи 1.19.a,mn; другие наиболее интересные задачи — 1.11, 1.12 и 1.13. При этом красивые нетривиальные результаты о реализуемости гиперграфов сформулированы на языке систем точек. Поэтому понятие гиперграфа не понадобится. Работать с четырехмерным пространством придется только в конце, когда это уже будет не страшно (например, потому что на трехмерных примерах будет отработано умение сводить геометрические задачи к задачам меньшей размерности; подробнее см. начало пункта 'реализуемость в четырехмерном пространстве').

Общие соглашения

Если условие задачи является утверждением, то задача состоит в том, чтобы это утверждение доказать.

Школьник (или команда школьников, работающих вместе над задачей) получает „звездочку " за каждое записанное решение, оцененное в + или +.. *Большая понятная преподавателю картинка оценивается как записанное построение примера* в задаче 1.6, а также в тех пунктах задач 1.13 и 1.19, где ответ 'да'. Жюри будет также награждать дополнительными „звездочками" за красивые решения, решения сложных задач и за (некоторые) решения, записанные в TEX-е. „Звездочек" у жюри бесконечно много. Можно сдавать задачи устно, теряя „звездочку" за каждую попытку.

Мы приглашаем всех школьников, решающих этот цикл задач, *консультироваться* по поводу возникающих вопросов и идей решения.

Задачи 1.2.abc, 1.4.ab и 1.7 будут разобраны на представлении, сдавать их можно только до.

Школьники, успешно решающие задачи, смогут получить *дополнительные задачи.*

Реализуемость в плоскости

Набор(=подмножество) точек на плоскости называется *набором общего положения*, если никакие 3 из них не лежат на одной прямой.

Под *n точками на плоскости (в пространстве)* подразумевается *n-элементное подмножество плоскости (пространства).* Т.е., считается, что эти n точек различны.

Следующим утверждением можно пользоваться в дальнейшем без доказательства.

1.1. *Теорема о четности.* Если 6 вершин двух треугольников на плоскости находятся в общем положении, то контуры этих треугольников пересекаются в четном числе точек.

[1] После ЛКТГ обновляемая версия будет поддерживаться в качестве части книги www.mccme.ru/circles/oim/algor.pdf. Благодарим А. Сосинского за перевод частей текста на английский, П. Кожевникова за полезные обсуждения И. Богданова за изготовление некоторых рисунков.

[2] Все эти объекты определены далее. Обычно эти результаты доказываются с использованием сложной техники [Pr06]. Впрочем, *топологическая* нереализуемость — в отличие от *линейной*, о которой речь выше, и *кусочно-линейной*, для которой доказательства аналогичны — действительно доказывается сложнее [Sk03], [Sk08, §5].

Подмножество плоскости (или пространства) называется *выпуклым*, если оно содержит вместе с любыми двумя точками соединяющий их отрезок. *Выпуклой оболочкой* множества X называется наименьшее (по включению) выпуклое множество, содержащее X.

1.2. (a) Существуют такие 4 точки на плоскости, что для любого их разбиения на две пары отрезок, соединяющий точки в первой паре, не пересекает отрезок, соединяющий точки во второй паре.

(b) Любые 4 точки на плоскости можно разбить на две группы так, что выпуклая оболочка точек первой группы пересекает выпуклую оболочку точек второй группы.

(c) Из любых 5 точек на плоскости можно выбрать две такие непересекающиеся пары точек, что отрезок, соединяющий точки в первой паре, пересекает отрезок, соединяющий точки во второй паре.

(d) Даны две тройки точек на плоскости. Тогда существуют два пересекающихся отрезка, не имеющие общих вершин, каждый из которых соединяет точки из разных троек.

Семейство отрезков (на плоскости или в пространстве) называется *вложенным*, если

• отрезки, не имеющие общих вершин, не пересекаются, и

• отрезки, имеющие общую вершину, пересекаются только в этой вершине.

Возможно, перед решением пунктов (c) и (d) Вам захочется решить их ослабленные версии:

(c') Для любых 5 точек на плоскости семейство всех отрезков, их соединяющих, не является вложенным.

(d') Для любых двух троек точек на плоскости семейство всех отрезков, соединяющих точки из разных троек, не является вложенным. [3]

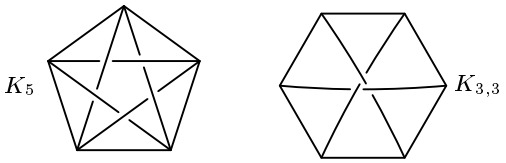


Рис. 1: Непланарные графы

1.3. Найдите количество всех неупорядоченных пар $\{\{i,j\},\{k,l\}\}$ непересекающихся двухэлементных подмножеств пятиэлементного множества.

1.4. Пусть на плоскости дан набор $f := \{1,2,3,4,5\}$ пяти точек общего положения. Для любых четырех различных точек i,j,k,l из них отрезки ij и kl либо не пересекаются, либо пересекаются в одной точке. Определим $v(f)$ как четность количества точек пересечения отрезков ij и kl для всех неупорядоченных пар $\{\{i,j\},\{k,l\}\}$ непересекающихся двухэлементных подмножеств $\{i,j\},\{k,l\} \subset f$:

$$v(f) := \sum \{|ij \cap kl| \ : \ \{\{i,j\},\{k,l\}\} \subset \binom{f}{2},\ \{i,j\} \cap \{k,l\} = \emptyset\} \mod 2.$$

(a) Для набора f_0 пяти точек на плоскости, изображенного на рис. 1 слева, $v(f_0) = 1$.

(b) $v(f)$ не зависит от f.

1.5. (a,b) Сформулируйте и докажите аналог задач 1.4.a,b для шести точек общего положения на плоскости, разбитых на две тройки.

(c,d) Сформулируйте и докажите аналог задач 1.2.c,d для точек на сфере.

[3] Конечно, эти утверждения — версии непланарности графов K_5 и $K_{3,3}$. Но доказываются они проще: достаточно задачи 1.1 вместо нетривиальных версий теоремы Жордана. Если в Вашем решении такие версии используются, не забудьте их доказать.

1.6. Можно ли нарисовать без самопересечений графы K_5 и $K_{3,3}$ (рис. 1)
(a) на сфере? (b) на боковой поверхности цилиндра (рис. 2)?
(c) на торе (рис. 2)? (d) на листе Мебиуса (рис. 2)?

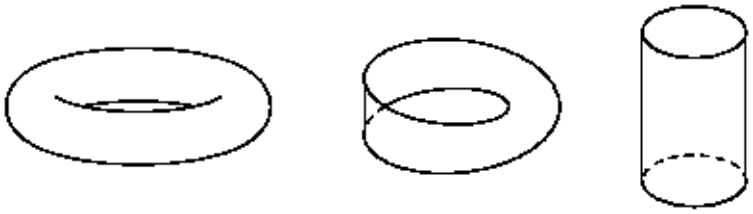


Рис. 2: Тор, лист Мебиуса и цилиндр

Тором называется поверхность бублика, рис. 2 слева. Или, эквивалентно, фигура, полученная из (двумерного) квадрата склейкой его пар противоположных сторон 'с одинаковыми направлениями', т.е. без поворота. *Листом Мебиуса* называется фигура, полученная из длинной прямоугольной полоски склейкой ее двух противоположных сторон 'с противоположным направлением', т.е. с поворотом на 180°, рис. 2. Эти (и другие) фигуры предполагаются *прозрачными*, т.е. точка (или подмножество), 'лежащая на одной стороне поверхности', 'лежит и на другой стороне'. Это аналогично тому, что при изучении геометрии мы говорим, например, о треугольнике на плоскости, а не о треугольнике на верхней (или нижней) стороне плоскости.

Реализуемость в пространстве

1.7. Существуют 100 точек в пространстве таких, что семейство всех отрезков, их соединяющих, является вложенным.

Набор точек в пространстве находится *в общем положении*, если никакие 4 точки из этого набора не лежат в одной плоскости.

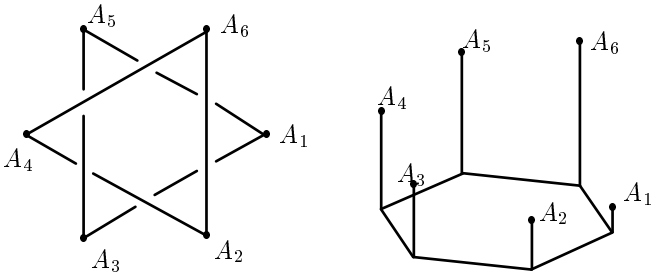


Рис. 3: Точки общего положения

1.8. (a) (Рис. 3.) Рассмотрим в горизонтальной плоскости правильный шестиугольник. Набор точек A_1, A_2, A_3, A_4, A_5, A_6, расположенных в точности над вершинами на высотах 1, 2, 3, 4, 5, 6 соответственно, находится в общем положении.

(b) То же для набора точек (t, t^2, t^3) в декартовой системе координат, где $t \in [0, 1]$.

1.9. (a) Существуют 4 точки в пространстве, которые нельзя разбить на две группы так, что выпуклая оболочка точек первой группы пересекает выпуклую оболочку точек второй группы.

(b) Любые 5 точек в пространстве можно разбить на две группы так, что выпуклая оболочка точек первой группы пересекает выпуклую оболочку точек второй группы.

1.10. В пространстве отмечено несколько точек общего положения и точка O. Известно, что для любых трех отмеченных точек A, B, C найдется отмеченная точка D такая, что O лежит строго внутри тетраэдра $ABCD$. Докажите, что отмечено ровно 4 точки.

1.11. (a) Из любых 6 точек в пространстве можно выбрать 5 точек O, A, B, A', B' так, что двумерные треугольники OAB и $OA'B'$ имеют некоторую общую точку, кроме O.

(b) Для 5 точек аналогичное утверждение неверно.

Семейство двумерных треугольников в пространстве называется *вложенным*, если

• треугольники, не имеющие общих вершин, не пересекаются, и

• треугольники, имеющие ровно одну общую вершину, пересекаются только в этой вершине, и

• треугольники, имеющие общую сторону, пересекаются только по этой стороне.

Эти условия формализуют 'отсутствие самопересечений в конструкции'. При доказательстве вложенности семейства выполнение этих условий нужно аккуратно проверить только в первый раз и в тех случаях, когда жюри попросит это сделать (если выполнение очевидно из конструкции, то жюри не будет просить его проверять).

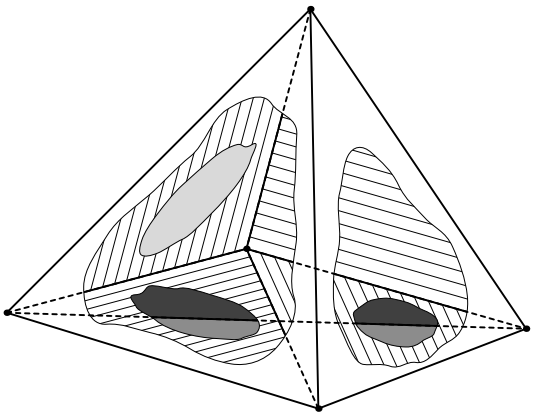


Рис. 4: 5 точек в пространстве: вершины и центр тетраэдра

Например, на рисунке 4 изображены такие 5 точек в пространстве, что семейство всех образованных ими треугольников является вложенным. Задача 1.11.a показывает, что 6 точек с таким свойством не существует.

1.12. (a) Существует 6 точек $A_0, A_1, \ldots, A_5$ в пространстве, для которых семейство треугольников $A_0A_jA_k$, $1 \le j < k \le 5$, $k \ne 2$, является вложенным.

(b) Если в пространстве имеются 5 точек и семейство S всех образованных ими треугольников является вложенным, то для любой точки пространства один из 5 отрезков, соединяющих эту точку с данными, пересекает один из треугольников, образованных данными точками.

Доказать существование можно, явно указав нужные точки. При обосновании обязательного наличия пересечений можно пользоваться без доказательства всеми верными четко сформулированными фактами типа 'поверхность выпуклого многогранника разбивает пространство на куски', справедливость которых подтвердило жюри.

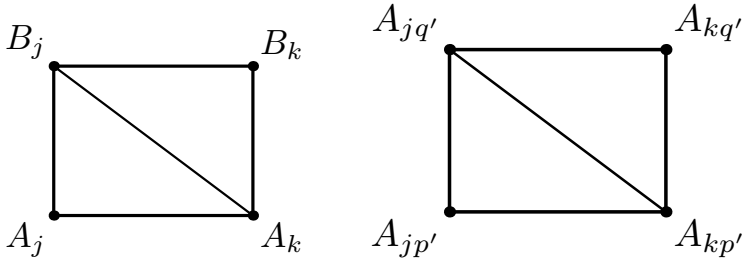


Рис. 5: К задачам о цилиндре и о декартовом произведении

1.13. (n) *Конус.* Для каких n существуют $n+1$ точек $O, A_1, \dots, A_n$ в пространстве таких, что семейство всех треугольников

$$OA_jA_k, \quad 1 \le j < k \le n,$$

является вложенным? Отдельно принимаются пункты (4), (5).

(lmn) *Джойн.* Для каких l, m, n существуют $l+m+n$ точек $A_1, \dots, A_l$, $B_1, \dots, B_m$, $C_1 \dots, C_n$ в пространстве таких, что семейство всех треугольников

$$A_iB_jC_k, \quad 1 \le i \le l, \quad 1 \le j \le m, \quad 1 \le k \le n,$$

является вложенным? Отдельно принимаются пункты (222), (223), (233).

(2n) *Цилиндр.* Для каких n существуют $2n$ точек $A_1, \dots, A_n$, $B_1, \dots, B_n$ в пространстве таких, что семейство всех треугольников

$$A_jB_jA_k \quad \text{и} \quad A_kB_kB_j, \quad 1 \le j < k \le n,$$

является вложенным? Отдельно принимаются пункты (24), (25).

(mn) *Декартово произведение.* Для каких m, n существует mn точек $A_{j,p}$, $j \in \{1, 2, \dots, m\}$, $p \in \{1, 2, \dots, n\}$, в пространстве, что семейство всех треугольников

$$A_{j,p}A_{j,q}A_{k,p} \quad \text{и} \quad A_{k,p}A_{k,p}A_{j,q}, \quad 1 \le j < k \le m, \quad 1 \le p < q \le n,$$

является вложенным? Отдельно принимаются пункты (33), (34), (35), (44).

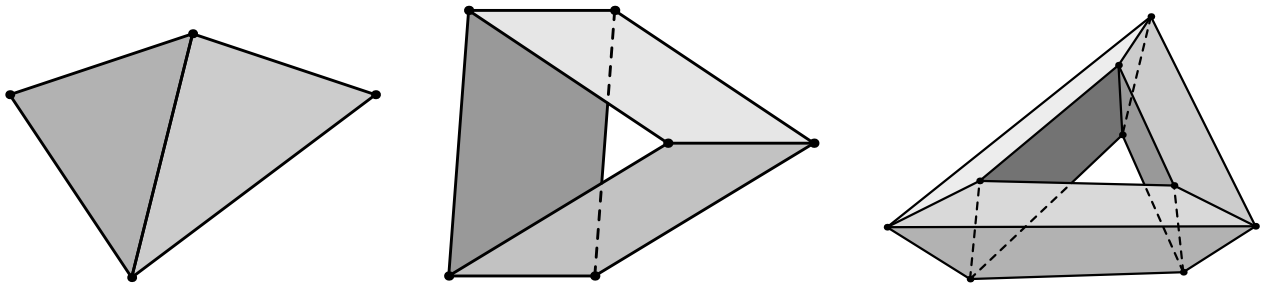


Рис. 6: Так выглядят (m, n)-реализации

Назовем (m, n)-*реализацией в* $\mathbb{R}^3$ вложенное семейство двумерных треугольников из задачи 1.13.mn. (Общепринятый термин — *линейное вложение комплекса* $K_m \times K_n$.) Ясно, что $(1, n)$-реализация пуста, $(2, 2)$-реализация является прямоугольником, перегнутым по диагонали, $(2, 3)$-реализация выглядит как цилиндр и $(3, 3)$-реализация выглядит как тор, см. рис. 6.

Реализуемость в четырехмерном пространстве

Как работать с четырехмерным пространством? Можно определить

• прямую $\mathbb{R}$ как множество всех вещественных чисел,

• плоскость $\mathbb{R}^2$ как множество всех упорядоченных пар (x, y) вещественных чисел,

• трехмерное пространство $\mathbb{R}^3$ как множество всех упорядоченных троек (x, y, z) вещественных чисел,

• четырехмерное пространство $\mathbb{R}^4$ как множество всех упорядоченных четверок (x, y, z, t) вещественных чисел.

Далее можно 'аналитически' в $\mathbb{R}^2$ определить прямые, в $\mathbb{R}^3$ — прямые и плоскости, а в $\mathbb{R}^4$ — прямые, плоскости и (трехмерные) гиперплоскости. Можно выводить из указанного аналитического определения (или называем аксиомами) только простейшие свойства геометрических объектов. Более сложные можно выводить из простейших 'синтетически' (т.е. как в школьной геометрии,

не используя аналитического определения). При этом плоскую задачу часто удобно сводить к линейной (т.е. к задаче на прямой), а пространственную — к плоской. Точно так же важнейший метод решения следующих четырехмерных задач — сведение к пространственным. При решении задач об $\mathbb{R}^4$ можно пользоваться без доказательства всеми верными четко сформулированными фактами о множествах решений систем линейных уравнений или о заменах систем координат, справедливость которых подтвердило жюри.

Определение вложенного семейства двумерных треугольников в $\mathbb{R}^4$ аналогично трехмерному случаю. Нужно лишь заменить $\mathbb{R}^3$ на $\mathbb{R}^4$.

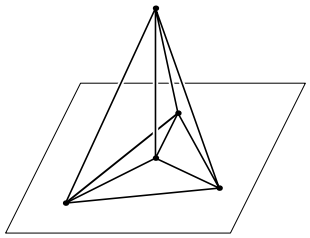


Рис. 7: 101 точка в $\mathbb{R}^4$. Четырехмерное пространство изображено в виде трехмерного, а (трехмерная) гиперплоскость в $\mathbb{R}^4$ — в виде двумерной плоскости в $\mathbb{R}^3$.

Приведем пример рассуждения с четырехмерным пространством. Докажем, что существуют 101 точка $O, A_1, \dots, A_{100} \in \mathbb{R}^4$ такая, что семейство всех треугольников OA_jA_k, $1 \le j < k \le 100$, является вложенным. Доказательство аналогично решению задачи 1.13.4 (рис. 7). Возьмем 100 точек $O, A_1, \dots, A_{100}$ в (трехмерной) гиперплоскости в $\mathbb{R}^4$ таких, что семейство всех соединяющих их отрезков является вложенным (см. задачу 1.7). Возьмем в $\mathbb{R}^4$ точку O, не лежащую в этой гиперплоскости. Тогда точки $O, A_1, \dots, A_{100}$ — искомые.

1.14. (a) Для любых двух точек, не лежащих на плоскости $x = y = 0$ в $\mathbb{R}^4$, существует соединяющая их ломаная, не пересекающая этой плоскости.

(b) Для любой гиперплоскости в $\mathbb{R}^4$ найдутся две не лежащие на ней точки такие, что любая ломаная, соединяющая эти две точки, пересекает гиперплоскость.

1.15. (a) Существуют 5 точек в $\mathbb{R}^4$, которые нельзя разбить на две группы так, что выпуклая оболочка точек первой группы пересекает выпуклую оболочку точек второй группы.

(b) Любые 6 точек в $\mathbb{R}^4$ можно разбить на две группы так, что выпуклая оболочка точек первой группы пересекает выпуклую оболочку точек второй группы.

В задачах 1.16 и 1.17 достаточно привести верные ответы.

1.16. Что получается в пересечении *трехмерной сферы*

$$S^3 := \{(x, y, z, t) \in \mathbb{R}^4 \mid x^2 + y^2 + z^2 + t^2 = 1\}$$

со следующими множествами:

(a) прямой $x = y = z = 0$, проходящей через центр сферы;

(b) плоскостью $x = y = 0$, проходящей через центр сферы;

(c) гиперплоскостью $x = 0$, проходящей через центр сферы;

(d) пересечением положительной шестнадцатки и объединения двумерных координатных плоскостей, т.е.

$$\{(x, y, z, t) \in \mathbb{R}^4 \mid x \ge 0,\ y \ge 0,\ z \ge 0,\ t \ge 0 \text{ и два из четырех чисел } x, y, z, t \text{ нулевые}\}.$$

Набор точек в $\mathbb{R}^4$ находится *в общем положении*, если никакие 5 точек из этого набора не лежат в одной гиперплоскости. Например, точки (t, t^2, t^3, t^4), $t \in [0, 1]$, находятся в общем положении.

1.17. Даны точки 1,2,3,4,5,6,7,8 общего положения в $\mathbb{R}^4$. По какому множеству пересекаются

(a) прямая 12 и гиперплоскость 5678? (b) прямая 12 и плоскость 567?

(c) плоскость 123 и гиперплоскость 5678? (d) гиперплоскости 1234 и 5678?

(e) плоскости 123 и 567?

(Заметим, что наше определение общего положения отличается от того, что общепринято для таких задач.)

1.18. (a) Существуют такие 6 точек в $\mathbb{R}^4$, что семейство всех образованных ими треугольников является вложенным.

(b) Существуют такие 7 точек в $\mathbb{R}^4$, что семейство всех образованных ими треугольников, кроме одного, является вложенным.

Первый пункт следующей задачи показывает, что для 7 точек утверждение, аналогичное задаче 1.18.a, неверно.

1.19. Основные задачи. (a) Из любых 7 точек в $\mathbb{R}^4$ можно выбрать две такие непересекающиеся тройки точек, что образованные этими тройками двумерные треугольники пересекаются.

(b) Даны три тройки точек в $\mathbb{R}^4$. Тогда существуют два пересекающихся треугольника, не имеющие общих вершин, вершины каждого из которых образованы точками разных троек.

(mn) Для каких m, n существует (m, n)-реализация в $\mathbb{R}^4$?

(Определение аналогично (m, n)-реализации в $\mathbb{R}^3$, только точки берутся в $\mathbb{R}^4$.)

Отдельно принимаются пункты (35), (3n), (44), (45), (4n), (55).

Доказать обязательность наличия пересечений в этой задаче может быть трудно без задач-подсказок, которые будут даны после промежуточного финиша.

1.20. Существуют 100 точек в пятимерном пространстве таких, что семейство всех образованных ими треугольников является вложенным.

2 Указания и решения, выдаваемые на представлении

1.2. (a) Треугольник и точка внутри него.

(b) Рассмотрим четверку точек A, B, C, D на плоскости.

Если какие-то 3 из них лежат на одной прямой, то некоторая из них, скажем B, лежит на отрезке между двумя другими, например, между A и C. Обозначим через $[XY]$ отрезок с вершинами X, Y. Тогда $[AC] \cap [BD] \neq \emptyset$.

Значит, никакие 3 точки не лежат на одной прямой. Если одна из этих точек лежит внутри треугольника, образованного остальными, то задача решена. Иначе каждая из этих точек лежит снаружи треугольника, образованного остальными. Поскольку точка D снаружи треугольника ABC, то она либо внутри одного из углов, вертикальных углам треугольника ABC, либо внутри одного из углов треугольника ABC.

Случай 1. Точка D внутри одного из углов, вертикальных углам треугольника ABC. Без ограничения общности, D внутри угла, вертикального углу $\angle ACB$. Тогда точка C внутри ABD, противоречие.

Случай 2. Точка D внутри одного из углов треугольник ABC, скажем $\angle BAC$. Поскольку точка D вне треугольника ABC и внутри угла BAC, то точки D и A лежат по разные стороны от прямой BC. Следовательно, отрезки $[AD]$ и $[BC]$ пересекаются.

(c) *Первое решение* вытекает из задачи 1.4.

(c) *Другое решение.* Предположим, напротив, что существуют такие 5 точек $OABCD$ на плоскости, что нужную пару выбрать нельзя. Тогда $A \notin OB$ и $B \notin OA$. Значит, A не лежит на луче OB. Поэтому можно считать, что точки A, B, C, D идут в том порядке, в котором они видны из O. Тогда треугольники OAC и OBD пересекаются в единственной точке O. Их пересечение

'трансверсально'. Значит, по теореме о четности (т.е. аналогично задаче 1.1) $AC \cap BD \neq \emptyset$. Противоречие.

(d) Аналогично первому решению пункта (c), см. задачу 1.5.

1.4. (b) Достаточно доказать, что для любых точек общего положения $1, 2, 3, 4, s, s'$ и множеств

$$A := \{1, 2, 3, 4\}, \quad f := A \cup \{s\} \quad \text{и} \quad f' = A \cup \{s'\} \quad \text{выполнено} \quad v(f) = v(f').$$

Докажем это. Для каждого $i \in A$ обозначим через A_i треугольник с вершинами из $A - \{i\}$. Тогда утверждение задачи следует из

$$v(f') - v(f) = \sum_{i \in A} (|si \cap A_i| - |s'i \cap A_i|) = \sum_{i \in A} |ss' \cap A_i| = 0 \mod 2.$$

Второе равенство следует из того, что число $|ss'i \cap A_i|$ четно по теореме о четности. Последнее равенство следует из того, что для для каждой неупорядоченной пары $\{i, j\} \subset A$ существует ровно два треугольника с вершинами из A, которые содержат отрезок ij. Значит, для каждой неупорядоченной пары $\{i, j\} \subset A$ число $|ss' \cap ij|$ входит в сумму для двух треугольников A_i, A_j.

1.5. *Формулировка.* Пусть на плоскости дано шесть точек общего положения, разбитых на две тройки $f_1 = \{1, 2, 3\}$ и $f_2 = \{4, 5, 6\}$. Для любых двух точек $i, j \in f_1$ и двух точек $i', j' \in f_2$ отрезки ii' и jj' либо не пересекаются, либо пересекаются в одной точке. Определим $v(f_1, f_2)$ как четность количества точек пересечения отрезков ii' и jj', ij' и ji' для всех двухэлементных подмножеств $\{i, j\} \subset f_1$, $\{i', j'\} \subset f_2$

$$v(f_1, f_2) := \sum \{|ii' \cap jj'| + |ij' \cap ji'| \ : \ \{i, j\} \in \binom{f_1}{2}, \ \{i', j'\} \in \binom{f_2}{2}\} \mod 2.$$

(a) Для набора f_1, f_2, изображенного на рис. 1 справа, $v(f_1, f_2) = 1$.

(b) $v(f_1, f_2)$ не зависит от f_1, f_2.

(c,d) *Сферической прямой* называется пересечение сферы и плоскости, проходящей через центр сферы.

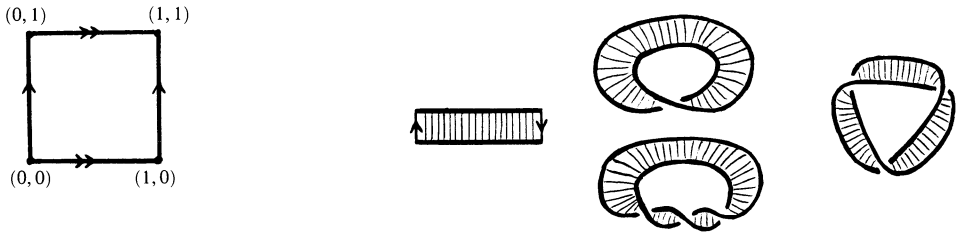


Рис. 8: Склейки прямоугольной полоски, дающие лист Мебиуса и тор

1.6. Можно рисовать не только на торе и листе Мебиуса с рис. 2, но и на их плоской развертке, рис. 8.

1.7. Расположим 3 точки в пространстве, не лежащие на одной прямой. Тогда они в общем положении. Пусть есть $n \geq 3$ точек общего положения. Тогда существует лишь конечное число плоскостей, проходящих через тройки этих точек. Значит, существует точка, не принадлежащая ни одной из этих плоскостей. Добавив эту точку к нашему набору из n точек, получим набор из $n + 1$ точки, никакие 4 из которых не лежат в одной плоскости. Таким образом, мы доказали, что для любого n существуют n точек общего положения в пространстве.

Рассмотрим 100 точек общего положения в пространстве. Назовем A множество всех отрезков с концами в этих точках. Если два отрезка из A с попарно различными концами пересекаются, то 4

их конца лежат в одной плоскости, противоречие. Если два отрезка с общим концом пересекаются более чем в одной точке, то их концы лежат на одной прямой, противоречие.

1.8. Используйте координаты.

1.13. (n),(lmn), (44) Используйте задачу 1.11.

3 Решения, выдаваемые на промежуточном финише

Если текст по задаче начинается со слов 'ответ' или 'указание', то детали (в частности, доказательства сформулированных утверждений или завершение решений) остаются для самостоятельной работы. Соответствующая задача принимается и после промежуточного финиша.

1.3. *Ответ-указание:* $5 \cdot \binom{4}{2}/2 = 15$.

1.6. (a), (b). Нельзя.

(a) Если бы граф K_5 был нарисован на сфере без самопересечений, то выкинув из сферы точку, не лежащую на K_5, мы получили бы плоскость, содержащую K_5. Противоречие.

(b) Граф K_5 не планарен, а цилиндр можно спроектировать без самопересечений на плоскость.

(c), (d). Можно. Красивые реализации графа K_5 на торе и графа $K_{3,3}$ на листе Мебиуса изображены на рис. 9.

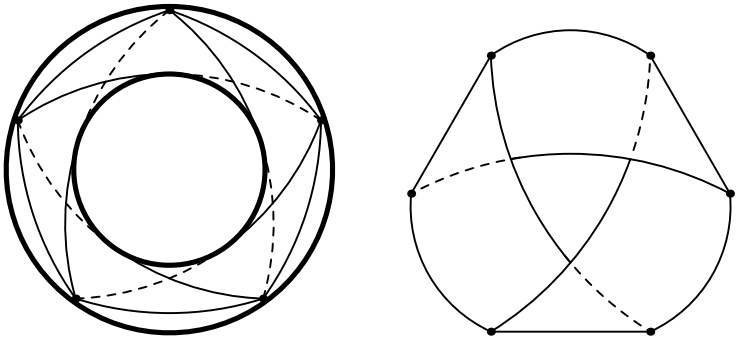


Рис. 9: Реализация графов Куратовского

Имеются также другие решения. Например, можно нарисовать граф Куратовского на плоскости с *одним* самопересечением и...

1.9. (a) Любые четыре точки, не лежащие в одной плоскости, удовлетворяют условию.

(b) См. теорему Радона в конце §3.

1.10. *Указание.* Обозначим через $A_1, A_2, \ldots, A_n$ отмеченные точки. Обозначим через l_i луч с началом в точке O, содержащий точку симметричную A_i относительно точки O. Докажите следующее утверждение.

Утверждение. Рассмотрим треугольник $A_iA_jA_k$. Точка O лежит внутри тетраэдра $XA_iA_jA_k$ тогда и только тогда, когда точка X лежит внутри трехгранного угла с вершиной O и сторонами l_i, l_j, l_k.

Рассмотрим тетраэдр с вершинами в отмеченных точках, внутри которого лежит O. Без ограничения общности будем считать, что это тетраэдр $A_1A_2A_3A_4$. Объединение всех трехгранных углов со сторонами $l_1, ..., l_4$ — трехмерное пространство. Внутри каждого из этих углов должна быть отмеченная точка по Утверждению. Луч l_5 лежит внутри ровно одного из тех трехгранных углов и 'разбивает' его на три трехгранных угла, внутри каждого из которых должна быть отмеченная точка. Луч l_6 лежит внутри ровно одного из тех шести трехгранных углов и 'разбивает' его на три трехгранных угла, внутри каждого из которых должна быть отмеченная точка. Итак, для $n > 4$ такой процесс бесконечен.

1.11. (a) Рассмотрим маленькую сферу вокруг любой точки O из данных. Пересечение этой сферы с объединением треугольников OAB для всех пар A, B данных точек — граф K_5. Противоречие.

Другое решение следует из теоремы Конвея-Гордона-Закса (задача 4.5). Впрочем, приведенное ниже доказательство теоремы Конвея-Гордона-Закса фактически повторяет вышеописанную редукцию к непланарности графа K_5.

(b) См. рис. 4.

1.13. *Ответы:* (n) $n \leq 4$;

(lmn) не более одного из чисел l, m, n больше 2;

(2n) для любых n;

(mn) либо одно из чисел m, n меньше 3, либо одно равно 3, а другое меньше 5.

(35) *Решение.* Предположим, существует (3,5)-реализация в $\mathbb{R}^3$.

Треугольник $i \times pqr$ — это треугольник $A_{i,p}, A_{i,q}, A_{i,r}$.

Кольцо $ij \times pqr$ — это (2,3)-реализация, 'отвечающая индексам i, j и p, q, r'.

Тор $ijk \times pqr$ — это (3,3)-реализация 'отвечающая индексам i, j, k и p, q, r'.

По теореме Жордана тор 123×123 разбивает пространство на две части. Объединение $14 \times 123 \cup 24 \times 123$ колец — кольцо. Это кольцо пересекает тор 123×123 по треугольникам 1×123 и 2×123. Кольцо 34×123 пересекает кольцо $14 \times 123 \cup 24 \times 123$ по треугольнику 4×123. Кольцо 34×123 пересекает тор 123×123 по треугольнику 3×123. Обозначим (4,3)-реализацию 'отвечающую индексам $1, 2, 3, 4$ и $1, 2, 3$', через $K_4 \times K_3$. По версии теоремы Жордана $K_4 \times K_3$ разбивает пространство на 4 части. В данной (5,3)-реализации вершина $A_{5,1}$ соединена

- с вершиной $A_{i,1}$ отрезком $A_{5,1}A_{i,1}$ для каждого $1 \leq i \leq 4$;
- с вершиной $A_{i,j}$ ломаной $A_{5,1}A_{5,j}A_{i,j}$ для каждого $1 \leq i \leq 4$, $2 \leq j \leq 3$.

Поскольку индекс 5 не 'участвует' в $K_4 \times K_3$, каждый из этих отрезков и каждая из этих ломанных пересекает $K_4 \times K_3$ только по концевым точкам. Рассмотрим связную компоненту дополнения $\mathbb{R}^3 - K_4 \times K_3$, содержащую точку $A_{5,1}$. На границе этой компоненты лежит точка $A_{i,j}$ для каждого $1 \leq i \leq 4$ и $1 \leq j \leq 3$, поскольку эта точка соединена с $A_{5,1}$ отрезком или ломаной, внутренность которой не пересекает $K_4 \times K_3$. Значит, граница этой компоненты содержит 12 точек $A_{i,j}$, для $1 \leq i \leq 4$ и $1 \leq j \leq 3$, из 15 данных. С другой стороны, эта граница — тор, т. е. (3,3)-реализация. Значит, эта граница содержит только 9 точек из 15 данных. Противоречие.

1.15. (a) Пять точек, не лежащие в одном трехмерном пространстве.

(b) См. теорему Радона в конце §3.

1.18. (a) Возьмем 5 вершин четырехмерного симплекса и точку внутри него.

(b) Пусть $ABCD$ — правильный тетраэдр в $\mathbb{R}^4$, а E — его центр. Выберем точку X внутри тетраэдра $ABCE$ так, чтобы точки A, B, C, D, E, X находились в общем положении в $\mathbb{R}^3$. Построим перпендикуляр l к гиперплоскости $ABCD$ в точке X. Наконец, выберем точки X_1, X_2 на прямой l по разные стороны от X. Докажем, что набор $V = \{A, B, C, D, E, X_1, X_2\}$ из семи точек — искомый, т.е., что семейство $\binom{V}{3} \backslash \triangle X_1X_2D$ треугольников вложенное.

Пусть α, β, γ — различные точки из $\{A, B, C, D, E\}$. Тогда имеем три класса треугольников:

- $\triangle X_1X_2\alpha$ для $\alpha \neq D$;
- $\triangle X_i\alpha\beta$ для $i \in \{1, 2\}$;
- $\triangle\alpha\beta\gamma$.

Легко проверить, что семейство треугольников из каждого класса является вложенным.

Треугольник из класса 1 пересекает треугольник из класса 2 по общей вершине X_i или общей стороне $X_i\alpha$.

Треугольник из класса 2 пересекает треугольник из класса 3 по общей вершине α или общей стороне $\alpha\beta$.

Наконец, рассмотрим пересечение треугольников из класса 1 и класса 3. Треугольник $\triangle X_1X_2\alpha$ пересекает гиперплоскость $ABCD$ по отрезку $X\alpha$. Поскольку X лежит внутри $ABCE$, то отрезки XA, XB, XC и XE пересекают треугольники из класса 3 только по вершинам.

Итак, семейство $\binom{V}{3}\backslash\triangle X_1X_2D$ треугольников является вложенным.

1.19. (a) Аналогично задачам 1.2.c и 1.4. Следует из задачи 4.3.

(b) Аналогично задачам 1.2.d и 1.5. Следует из задачи 4.4.

(mn) *Ответ:* $\min\{m,n\}\le 4$.

(4n) *Указание.* Докажем, что существует $(4,n)$-реализация в $\mathbb{R}^4$. Возьмем точки $A_{j,1}$, $1\le j\le n$, общего положения в $\mathbb{R}^4$. Возьмем упорядоченное множество K из четырех точек на плоскости в $\mathbb{R}^4$, четвертая из которых лежит внутри двумерного треугольника, образованного остальными. Например, $K:=((0,0,0,0),(2,0,0,0),(1,2,0,0),(1,1,0,0))$. Возьмем образы этого множества при переносах на векторы $A_{j,1}$, $1\le j\le n$. Т.е. обозначим $(A_{j,1},A_{j,2},A_{j,3},A_{j,4}):=K+A_{j1}$. Тогда

• для каждого $1\le j\le n$ точка $A_{j,4}$ лежит внутри двумерного треугольника $A_{j,1}A_{j,2}A_{j,3}$;

• для любых i,j множество $K+A_{i,1}$ совмещается с множеством $K+A_{j,1}$ параллельным переносом на вектор $A_{j,1}-A_{i,1}$;

• для любых i,j,k все 12 точек из $(K+A_{i,1})\cup(K+A_{j,1})\cup(K+A_{k,1})$ не лежат в одном трехмерном пространстве, поскольку точки $A_{j,1}$, $1\le j\le n$ находятся в общем положении.

Выведите из этого, что точки A_{ij} образуют $(4,n)$-реализацию в $\mathbb{R}^4$.

(55) *Указание.* См. задачи 4.14 и 4.15.

1.20. Аналогично задаче 1.7. Набор точек в $\mathbb{R}^5$ находится в *общем положении*, если никакие шесть из этих точек не лежат в одной четырехмерной гиперплоскости.

Для множества V его выпуклую оболочку будем обозначать $conv(V)$.

Теорема Радона. *Любые $n+2$ точки в $\mathbb{R}^n$ можно разбить на два множества $\{X_1,\dots X_k\}$ и $\{X_{k+1},\dots,X_{n+2}\}$ так, что $conv\{X_1,\dots,X_k\}\bigcap conv\{X_{k+1},\dots,X_{n+2}\}\ne\emptyset$.*

Доказательство. Отождествим точки с их радиус-векторами. Сначала докажем, что найдутся такие $c_1,\dots,c_{n+2}\in\mathbb{R}$, не все из которых равны нулю, что

$$c_1X_1+c_2X_2+\dots+c_{n+2}X_{n+2}=0\quad\text{и}\quad c_1+\dots+c_{n+2}=0.$$

В самом деле, рассмотрим набор векторов $X_1-X_{n+2},X_2-X_{n+2},\dots,X_{n+1}-X_{n+2}$. Поскольку это $n+1$ вектор в $\mathbb{R}^n$, то для этих векторов найдется нетривиальная линейная зависимость с коэффициентами $c_1,\dots c_{n+1}$. Тогда набор $c_1,\dots c_{n+1},-c_1-\dots-c_{n+1}$ — требуемый.

Теперь перенумеруем точки так, чтобы сначала шли положительные значения c_i. Перенесем слагаемые с отрицательными коэффициентами в правую часть. Получим: $c_1X_1+\dots+c_kX_k=-c_kX_k-\dots-c_{n+2}X_{n+2}$. Домножим это равенство на такой коэффициент, чтобы сумма коэффициентов справа и слева стала равна 1. После этого равенство будет утверждать, что $conv\{X_1,\dots,X_k\}\cap conv\{X_{k+1},\dots,X_{n+2}\}\ne\emptyset$. □

4 Новые задачи, выдаваемые на промежуточном финише

Трудные задачи о нереализуемости можно решать двумя способами. Первый способ — обобщение доказательства непланарности графа K_5 (т.е. решения задачи 1.2.c) при помощи препятствия Ван Кампена (задача 1.4). Он реализован в первом пункте этого параграфа и в дополнении. Второй способ — обобщение другого доказательства непланарности графа K_5 (т.е. решения задачи 1.2.c). Он основан на редукции к меньшей размерности и реализован во втором и третьем пунктах этого параграфа.

Задачи 1.19.a,b вытекают из нижеследующих задач 4.3 и 4.4. Несуществование $(5,5)$-реализации в $\mathbb{R}^4$ вытекает из сферической версии теоремы Закса (задача 4.9.b) и следующих задач 4.14, 4.15

(или из задач дополнения). Чтобы подойти к этому расуждению, сначала можете по-другому решить задачу 1.13.53, при помощи задач 4.10 и 4.11, а также задачу 1.19.a при помощи сферической версии теоремы Конвея-Гордона-Закса (задача 4.9.a) и задач 4.12, 4.13.

Мы называем *треугольником* (в $\mathbb{R}^3$ или в $\mathbb{R}^4$) его контур (т.е. замкнутую ломаную), а *двумерным треугольником* выпуклую оболочку его вершин. Аналогично, мы называем *тетраэдром* (в $\mathbb{R}^3$ или в $\mathbb{R}^4$) объединение его двумерных граней, а *трехмерным тетраэдром* выпуклую оболочку его вершин.

Будем называть (m,n)-реализацией (в $\mathbb{R}^3$ или в $\mathbb{R}^4$) также объединение вложенного семейства треугольников из определения (m,n)-реализации.

Обобщения препятствия Ван Кампена

4.1. (a) Для любых семи точек 1,2,3,4,5,6,7 общего положения в трехмерном пространстве треугольник 123 и двумерный тетраэдр 4567 пересекаются в конечном четном числе точек. [4]

(b) Для любых восьми точек 1,2,3,4,5,6,7,8 общего положения в четырехмерном пространстве двумерные тетраэдры 1234 и 5678 пересекаются в конечном четном числе точек.

4.2. Найдите количество всех неупорядоченных пар $\{\{i,j,k\},\{l,m,n\}\}$ непересекающихся трехэлементных подмножеств семиэлементного множества.

4.3. Пусть в четырехмерном пространстве дан набор $f := \{1,2,3,4,5,6,7\}$ семи точек общего положения. Для любых шести различных точек i,j,k,l,m,n из них двумерные треугольники ijk и lmn либо не пересекаются, либо пересекаются в одной точке. Определим $v(f)$ как четность количества точек пересечения двумерных треугольников ijk и lmn для всех неупорядоченных пар $\{\{i,j,k\},\{l,m,n\}\}$ непересекающихся трехэлементных подмножеств $\{i,j,k\},\{l,m,n\} \subset f$:

$$v(f) := \sum \{|ijk \cap lmn| \ : \ \{\{i,j,k\},\{l,m,n\}\} \subset \binom{f}{3},\ \{i,j,k\} \cap \{l,m,n\} = \emptyset\} \mod 2.$$

(a) Для набора f_0 семи точек в $\mathbb{R}^4$, придуманного Вами при решении задачи 1.18.b, $v(f_0) = 1$.

(b) $v(f)$ не зависит от f.

4.4. (a,b) Сформулируйте и докажите аналог задач 4.3.a,b для девяти точек общего положения в четырехмерном пространстве, разбитых на три тройки.

Элементы рамсеевской теории зацеплений

В этом пункте мы наметим доказательство 'линейных' случаев теорем Конвея–Гордона–Закса и Закса (задачи 4.5 и 4.8). Они понадобятся для доказательства невозможности в четырехмерных основных примерах и интересны сами по себе. Такие утверждения и методы их доказательства составляют *рамсеевскую теорию зацеплений*. См. подробнее [PS05].

Треугольники Δ и Δ' в пространстве, шесть вершин которых находятся в общем положении, называются *зацепленными*, если Δ пересекает двумерный треугольник Δ' в единственной точке. Например, треугольники $A_1A_3A_5$ и $A_2A_4A_6$ из задачи 1.8.a зацеплены.

4.5. *Теорема Конвея–Гордона–Закса для линейных вложений.* Для любых 6 точек общего положения в пространстве найдутся два зацепленных треугольника с вершинами в этих точках.

Следующая задача 4.6 не используются для доказательства теорем Конвея-Гордона-Закса и Закса. Но она поясняет понятие зацепленности.

4.6. (a) Если один из треугольников Δ, Δ', шесть вершин которых находятся в общем положении, не пересекает плоскость другого треугольника, то Δ и Δ' не зацеплены.

[4] При решении этой задачи нельзя пользоваться без доказательства теоремой Жордана о том, что многогранник разбивает пространство на куски (ибо сама эта задача используется при доказательстве теоремы Жордана).

(b) Пусть на прямой отмечено 2 красных и 2 синих точки, причем точки попарно различны. Будем говорить, что эти пары *зацеплены*, если они перемежаются, т.е. расположены на прямой в порядке (красная, синяя, красная, синяя) или (синяя, красная, синяя, красная), смотря с какой стороны идти.

Треугольники Δ и Δ' зацеплены $\Leftrightarrow$ прямая l пересечения их плоскостей пересекает каждый из них по паре точек и эти пары зацеплены на прямой l.

(c) Зацепленность треугольников не изменяется, если их вершины движутся в пространстве, оставаясь в общем положении.

(d) Треугольники Δ и Δ' зацеплены $\Leftrightarrow$ треугольники Δ' и Δ зацеплены.

(e) Для каких положений точки A_1 из задачи 1.8.a на вертикальной прямой треугольники $A_1A_3A_5$ и $A_2A_4A_6$ зацеплены?

Плоскость находится *в общем положении* относительно набора точек в $\mathbb{R}^3$, если ортогональные проекции точек набора на плоскость находятся в общем положении.

4.7. (a) Дана проекция пары треугольников на плоскость общего положения (относительно шести вершин треугольников), причем в местах пересечения двух линий показано, какая из них проходит выше (как на рис. 3 слева). Треугольники зацеплены тогда и только тогда, когда количество точек пересечения их проекций, в которых первый проходит над вторым, нечетно.

(b) Пусть в пространстве даны 6 точек общего положения. Назовем *разбиением* неупорядоченную пару треугольников с вершинами в этих точках, не имеющих общих вершин. Тогда количество зацепленных разбиений нечетно.

Замкнутые четырехзвенные ломаные $ABCD$ и $A'B'C'D'$ в пространстве, восемь вершин которых находятся в общем положении, называются *зацепленными*, если число точек пересечения ломаной $ABCD$ с объединением внутренностей треугольников $A'B'C'$ и $A'D'C'$ нечетно.

4.8. *Теорема Закса для линейных вложений.* Пусть в пространстве даны 4 красные и 4 синие точки, причем никакие два отрезка с разноцветными концами не имеют общих внутренних точек. Тогда найдутся две зацепленные замкнутые четырехзвенные ломаные с вершинами в этих точках, каждое звено которых соединяет точки разных цветов.

4.9. (a,b) Сформулируйте и докажите аналоги задач 4.5 и 4.8 с заменой пространства на трехмерную сферу S^3. (Зацепленность определяется аналогично случаю $\mathbb{R}^3$. Треугольники Δ и Δ' в трехмерной сфере, среди шести вершин которых никакие четыре не лежат на двумерной подсфере трехмерной сферы, называются *зацепленными*, если треугольник Δ пересекает любой двумерный сферический треугольник, натянутый на Δ', в единственной точке. Заметим, что имеется два таких натянутых треугольника.)

Применения рамсеевской теории зацеплений

4.10. Пусть в $\mathbb{R}^3$ даны замкнутая ломаная длины 3 и $(3,3)$-реализация N, пересекающиеся ровно в одной точке x, являющейся их общей вершиной. Тогда любая достаточно малая сфера S^2 с центром в x пересекает ломаную по паре точек, находящихся в одной компоненте связности дополнения $S^2 - N$. (До промежуточного финиша такими фактами разрешалось пользоваться без доказательства. Но мы предлагаем доказать этот факт здесь, чтобы подойти к доказательству аналогичного факта для $\mathbb{R}^4$.)

4.11. Предположим, что в $\mathbb{R}^3$ имеется $(5,3)$-реализация.

(a) Пересечение любой достаточно малой сферы с центром в $A_{1,1}$ и данной $(5,3)$-реализации — граф $K_{4,2}$, 'линейно' вложенный в сферу.

(b) Пусть даны любые несамопересекающийся цикл в этом графе и пара вершин этого графа, не лежащих на цикле. Тогда в данной $(5,3)$-реализации существуют замкнутая ломаная длины 3 и $(3,3)$-реализация, пересекающиеся друг с другом ровно по точке, а со сферой — по данным паре вершин и циклу, соответственно.

4.12. Пусть в $\mathbb{R}^4$ даны два тетраэдра, пересекающиеся ровно в одной точке x, являющейся их общей вершиной. Тогда любая достаточно малая сфера S^3 с центром в x пересекает тетраэдры по паре сферических треугольников, не зацепленных в сфере S^3.

4.13. Предположим, что в $\mathbb{R}^4$ имеется 7 точек 0,1,2,3,4,5,6, среди которых нельзя выбрать две такие непересекающиеся тройки точек, что образованные этими тройками двумерные треугольники пересекаются.

(a) Пересечение достаточно малой трехмерной сферы с центром в 0 и объединения всех треугольников $0ij$, $1 \le i < j \le 7$ — полный граф K_6, вложенный в S^3.

(b) Для любого разбиения $\{1,2,3,4,5,6\} = \{i,j,k\} \cup \{p,q,r\}$ тетраэдры $0ijk$ и $0pqr$ пересекаются ровно в одной точке.

4.14. Сформулируйте и докажите аналог задачи 4.12 для двух $(3,3)$-реализаций в $\mathbb{R}^4$.

4.15. Предположим, что в $\mathbb{R}^4$ имеется $(5,5)$-реализация.

(a) Пересечение любой достаточно малой трехмерной сферы с центром в $A_{1,1}$ и $(5,5)$-реализации — граф $K_{4,4}$, 'линейно' вложенный в трехмерную сферу.

(b) Для любых двух непересекающихся несамопересекающихся циклов в этом графе существуют две $(3,3)$-реализации в $(5,5)$-реализации, пересекающиеся друг с другом ровно по точке, а со сферой — по данным циклам.

5 Решения, выдаваемые на окончательном финише

1.1. Ср. [BE82, §5]. Если контуры треугольников не пересекаются, то задача доказана. Если пересечение контуров двух треугольников непусто, и вершины этих треугольников находятся в общем положении, то пересечением контура первого треугольника с выпуклой оболочкой вершин второго треугольника является объединением конечного числа незамкнутых ломаных, каждая из которых является подмножеством контура первого треугольника. Значит, концевые точки каждой из этих ломаных являются точками пересечения контуров наших треугольников. Поскольку у каждой незамкнутой ломаной есть ровно две концевые точки, то контуры наших треугольников пересекаются в четном числе точек.

1.13. (mn) *Для $m = n = 4$.* Докажем, что $(4,4)$-реализации в $\mathbb{R}^3$ не существует. Аналогично решению задачи 1.11.a. Пересечение (4,4)-реализации с достаточно малой сферой с центром $A_{1,1}$ есть граф $K_{3,3}$, линейно вложенный в эту сферу. Противоречие.

Другое решение $m = 5$ и $n = 3$. Докажем, что $(5,3)$-реализации в $\mathbb{R}^3$ не существует. Аналогично доказательству непланарности графа K_5 (т.е. другому решению задачи 1.2.c). Пусть в $\mathbb{R}^3$ имеется $(5,3)$-реализация. Рассмотрим достаточно малую сферу S^2 с центром в точке $A_{1,1}$. Пересечение сферы S^2 и $(5,3)$-реализации — граф $K_{4,2}$.

При любом вложении графа $K_{4,2}$ в сферу некоторые две вершины X, Y из 4-вершинной доли находятся по разные стороны от цикла Σ, образованного четырьмя оставшимися вершинами. Без ограничения общности считаем, что отрезки, соответствующие вершинам X, Y соединяют вершину $A_{1,1}$ с вершинами $A_{2,1}$ и $A_{3,1}$. Обозначим через γ_{XY} ломаную $A_{1,1}A_{2,1}A_{3,1}$ и через γ тор 145×123.

Так как на сфере S^2 точки $\{X, Y\} = \gamma_{XY} \cap S^2$ находятся по разные стороны от цикла $\Sigma = \gamma \cap S^2$, то пересечение в точке $A_{1,1}$ трансверсально. Противоречие с теоремой о четности (задача 4.10).

4.1. Аналогично задаче 1.1.

4.2. *Ответ-указание:* $7 \cdot \binom{6}{3}/2 = 70$.

4.3. (a) См. решение задачи 1.18.a. Двумерный треугольник X_1X_2D пересекает трехмерный тетраэдр $ABCD$ по отрезку $[DX]$. Поскольку точка X лежит внутри тетраэдра $ABCE$, а точка

D — вне, то отрезок $[DX]$ пересекает поверхность тетраэдра $ABCE$ ровно в одной точке. Значит, двумерный треугольник X_1X_2D пересекается только с одним из двумерных треугольников, натянутых на остальные вершины, т.е. $v(f) = 1$.

(b) Достаточно доказать, что для 8 точек $1, 2, 3, 4, 5, 6, v, v'$ общего положения в $\mathbb{R}^4$ и множеств

$$A := \{1, 2, 3, 4, 5, 6\} \quad f := A \cup \{v\}, \quad f' := A \cup \{v'\} \quad \text{выполнено} \quad v(f) = v(f').$$

Обозначим через T_{ij} тетраэдр с вершинами из $A - \{i, j\}$.

$$v(f) - v(f') = \sum_{\{i,j\} \in \binom{A}{2}} (|vij \cap T_{ij}| - |v'ij \cap T_{ij}|) = \sum_{\{i,j\} \in \binom{A}{2}} |vv'i \cap T_{ij}| = 0 \mod 2.$$

Второе равенство следует из задачи 4.1.b. Последнее равенство выполняется, так как для любых $a, b, c \in A - \{i\}$ найдутся ровно два тетраэдра T_{ik}, T_{ij}, $j, k \in A - \{i, a, b, c\}$, содержащих двумерный треугольник abc. Таким образом, каждое число $|vv'i \cap abc|$ входит в эту сумму дважды.

4.4. *Формулировка.* Пусть в четырехмерном пространстве даны 9 точек общего положения, разбитых на три тройки $f_1 := \{1, 2, 3\}$, $f_2 := \{4, 5, 6\}$, $f_3 := \{7, 8, 9\}$, такие что для любых двух различных точек $i, i' \in \{1, 2, 3\}$, двух различных точек $j, j' \in \{4, 5, 6\}$ и двух различных точек $k, k' \in \{7, 8, 9\}$ двумерные треугольники ijk и $i'j'k'$ либо не пересекаются, либо пересекаются в одной точке. Определим $v(f_1, f_2, f_3)$ как четность количества точек пересечения двумерных треугольников ijk и $i'j'k'$ для всех упорядоченных пар $(i, i') \in f_1^2$, $(j, j') \in f_2^2$, $(k, k') \in f_3^2$:

$$v(f_1, f_2, f_3) := \left(\frac{1}{2} \sum \{|ijk \cap i'j'k'| \ : \ (i, i') \in f_1^2, \ (j, j') \in f_2^2 \ (k, k') \in f_3^2\} \right) \mod 2.$$

(a) *Пример.* Вложим граф $K_{3,3}$ с долями $\{1, 2, 5\}$, $\{3, 4, 6\}$ в $\mathbb{R}^3$ так, чтобы цикл 1234 был квадратом и никакие четыре точки, кроме $1, 2, 3, 4$, не лежали в одной плоскости. Расположим точки 7 и 8 в $\mathbb{R}^4$ по разные стороны от трехмерной гиперплоскости, в которую вложен граф $K_{3,3}$. Расположим внутри пирамиды 71234 точку 9. Тогда для $f_1 := \{1, 2, 5\}$, $f_2 := \{3, 4, 6\}$ и $f_3 := \{7, 8, 9\}$ выполнено $v(f_1, f_2, f_3) = 1$.

(b) Аналогично задачам 4.3.b и 1.5.

4.5. Следует из задачи 4.7.b.

4.6. (a) Достаточно доказать, что если треугольники Δ и Δ' зацеплены, то Δ' пересекает плоскость треугольника Δ и Δ пересекает плоскость треугольника Δ'. Первое очевидно. Поэтому внутренность треугольника Δ пересекает плоскость треугольника Δ'. Значит, и Δ пересекает плоскость треугольника Δ'.

(b) Пусть треугольник Δ зацеплен с треугольником Δ'. Так как треугольник Δ пересекает плоскость треугольника Δ', то по (a) $\Delta \cap l \neq \emptyset$. Из соображений общего положения следует, что $\Delta \cap l$ не может состоять из одной точки. Докажем теперь зацепленность пар. Пусть множество $\Delta \cap l = \{A, B\}$ и $\Delta' \cap l = \{A', B'\}$. Все точки пересечения треугольника Δ и внутренности треугольника Δ' лежат на отрезке $A'B'$. Значит, ровно одна из точек A и B принадлежит отрезку $A'B'$. Поэтому пары A, B и A', B' зацеплены на l.

Обратно, пусть пары A, B и A', B' зацеплены на l. Тогда согласно сказанному выше Δ пересекает внутренность треугольника Δ' в единственной точке, то есть Δ зацеплен с Δ'.

(c) Воспользуемся пунктом (b). Равносильное условие зацепленности сохраняется при таком перемещении вершин, при котором плоскости треугольников Δ и Δ' не параллельны. Если же в некоторый момент плоскости треугольников Δ и Δ' параллельны, то из (a) следует, что в этот момент, а также непосредственно до и после него, треугольники Δ и Δ' не зацеплены.

(d) Каждое из двух условий равносильно одному и тому же условию по пункту (b).

(e) Обозначим через t высоту точки A_1 над горизонтальной плоскостью. Треугольники зацеплены для $t \in (-\infty;\, 2) \cup (3,5;\, 4,5) \cup (6;\, +\infty)$ и не зацеплены для $t \in (2;\, 3,5) \cup (4,5;\, 6)$. Это следует из зацепленности треугольников $A_1A_3A_5$ и $A_2A_4A_6$ из задачи 1.8.a и следующей леммы (которая, в свою очередь, вытекает из (b)).

Лемма о движении. *Пусть вершина A треугольника Δ движется равномерно по отрезку в пространстве, а остальные две вершины и треугольник Δ' неподвижны. Обозначим через Δ_t положение треугольника в момент времени t, где $0 \le t \le 2$. Предположим, что 6 вершин треугольников Δ_t и Δ' находятся в общем положении при всех t, кроме $t = 1$.*

- *Если $\Delta_1 \cap \Delta' = \emptyset$, то пары (Δ_0, Δ') и (Δ_2, Δ') зацеплены или нет одновременно.*
- *Если Δ_1 и Δ' пересекаются в единственной точке, не совпадающей ни с одной из их вершин, то ровно одна из пар (Δ_0, Δ') и (Δ_2, Δ') является зацепленной.*

4.7. (b) Рассмотрим проекцию на произвольную плоскость общего положения. Обозначим одну из точек через A. Докажите, что четность количества зацепленных разбиений равна четности количества пересечений проекций неупорядоченных пар несоседних (т.е. не имеющих общих вершин) отрезков, не содержащих вершины A. Для этого преобразуйте сумму по разбиениям количеств пересечений из 4.7.b в сумму по упорядоченным парам непересекающихся ребер, проекции которых пересекаются и проекция первого из которых проходит выше проекции второго. Так как $K_6 - A \cong K_5$, то по сферическому аналогу задачи 1.4 (задача 1.5.c) последняя четность равна 1.

4.8. Аналогично доказательству теоремы Конвея–Гордона–Закса (задача 4.5), намеченному в задаче 4.7. Рассмотрите проекцию на маленький эллипсоид вокруг ребра e графа $K_{4,4}$. Вместо $K_6 - A \cong K_5$ используйте $K_{4,4} - e \cong K_{3,3}$.

4.7 и **4.8.** См. подробности в [Zi].

4.9. *Сферическим отрезком* называется пересечение трехмерной сферы и двумерного угла с вершиной в центре сферы.

Сферический аналог теоремы Конвея-Гордона-Закса есть следующее утверждение.

Для любых 6 точек общего положения на трехмерной сфере найдутся две зацепленные замкнутые сферические ломаные длины 3 с вершинами в этих точках.

Замкнутые четырехзвенные ломаные $ABCD$ и $A'B'C'D'$ на трехмерной сфере называются *зацепленными*, если число точек пересечения сферической ломаной $ABCD$ с объединением двумерных сферических треугольников $A'B'C'$ и $A'D'C'$ нечетно.

Сферический аналог теоремы Закса есть следующее утверждение.

Пусть на трехмерной сфере даны 4 красные и 4 синие точки, причем никакие два сферических отрезка с разноцветными концами не имеют общих внутренних точек. Тогда найдутся две зацепленные замкнутые четырехзвенные ломаные на сфере S^3 с вершинами в этих точках, каждое звено которых соединяет точки разных цветов.

Доказательства этих теорем аналогичны доказательствам их аналогов для трехмерного пространства.

4.10, 4.12. См. [Zu].

4.11. (a) Для каждой грани (5, 3)-реализации N, не содержащего вершины $A_{1,1}$, существует шар с центром в вершине $A_{1,1}$, не пересекающий этой грани (потому что существует точка этой грани, в которой функция 'расстояние до точки $A_{1,1}$' принимает минимум на этой грани, и этот минимум не равен нулю). Поскольку граней в N конечное число, то существует шар с центром в вершине $A_{1,1}$, не пересекающий ни одну грань из N, не содержащую $A_{1,1}$. Рассмотрим граничную сферу этого шара. Она пересекается только с теми отрезками и треугольниками в N, которые содержат вершину $A_{1,1}$. Это отрезки $A_{1,1}A_{1,a}$, $A_{1,1}A_{b,1}$ и треугольники $A_{1,1}A_{a,1}A_{b,1}$ для всех $2 \le a \le 5$ и всех $2 \le b \le 3$. Их пересечение со сферой состоит из 4 точек, соответствующих ребрам

$A_{1,1}A_{1,a}$, двух точки, соответствующих ребрам $A_{1,1}A_{b,1}$, и дуг больших кругов, соединяющие точки первой группы с точками второй группы. Это и есть граф $K_{4,2}$, линейно вложенный в сферу.

(b) См. второй абзац решения задачи 1.13.

4.13. (a) Пересечение треугольника $0ij$ с S^3 — дуга большой окружности S^3, соединяющая точки пересечения прямых $0i$ и $0j$ с S^3 (ср. с задачей 1.16.b). Значит, пересечения треугольников $0ij$, $1 \le i < j \le 6$, с S^3 образуют граф K_6. Он вложен в S^3, т.к. точки $0, 1, \dots, 6$ находятся в общем положении.

(b) По условию треугольники с различными вершинами не пересекаются. А поскольку точки в общем положении, то треугольники с общей вершиной или стороной пересекаются только по этой вершине или стороне соответственно.

4.14. *Формулировка.* Пусть в $\mathbb{R}^4$ даны две (3,3)-реализации, пересекающиеся ровно в одной точке, являющейся их общей '(1,1)-вершиной'. Тогда любая достаточно малая сфера S^3 с центром в этой точке пересекает эти (3,3)-реализации по паре сферических многоугольников, не зацепленных по модулю два в сфере S^3 .

4.15. (a) Аналогично задаче 4.11.a.

(b) Оба цикла имеют длину 4. Не уменьшая общности, вершины первого цикла соответствуют точкам $A_{1,2}$, $A_{2,1}$, $A_{1,3}$ и $A_{3,1}$, а вершины второго — точкам $A_{1,4}$, $A_{4,1}$, $A_{1,5}$ и $A_{5,1}$. Тогда первый искомый тор есть 123×123, а второй — 145×145.

Список литературы

[BE82] *В. Г. Болтянский и В. А. Ефремович.* Наглядная топология М.: Наука, 1982.

[Pr06] *В. В. Прасолов.* Элементы теории гомологий. М.: МЦНМО, 2006. См. `http://www.mccme.ru/prasolov/`.

[PS05] *В. В. Прасолов и М. Скопенков.* Рамсеевская теория зацеплений // Мат. Просвещение. 2005. 9. С. 108-115.

[Sk03] *M. Skopenkov.* Embedding products of graphs into Euclidean spaces // Fund. Math. 2003. 179. P. 191-198.

[Sk08] *A. Skopenkov.* Embedding and knotting of manifolds in Euclidean spaces // London Math. Soc. Lect. Notes, 347 (2008) 248–342; arxiv:math/0604045.

[Zi] *A. Zimin.* A short proof of the Conway-Gordon-Sachs and Sachs Theorems, unpublished.

[Zu] *J. Zung.* A non-general-position Parity Theorem, unpublished.

6 Дополнительные задачи

Звездочками отмечены задачи, решение которых авторам неизвестно.

6.1. Для любых точки O и $(4,3)$-реализации в $\mathbb{R}^3$ один из отрезков $OA_{1,1}$, $OA_{2,1}$, $OA_{3,1}$, $OA_{4,1}$ пересекает $(4,3)$-реализацию.

6.2. Для любых шести точек $0,1,2,3,4,5 \in \mathbb{R}^3$ если семейство треугольников

(a) $0jk$, $1 \le j < k \le 5$, $k \ne 2$, является вложенным, то треугольники 012 и 345 зацеплены.

(b) $0jk$, $1 \le j < k \le 5$, $(j,k) \notin \{(1,2),(1,3)\}$, является вложенным, то зацеплены либо треугольники 012 и 345, либо треугольники 013 и 245.

(c) $0jk$, $1 \le j < k \le 5$, $(j,k) \notin \{(1,2),(1,3),(1,4)\}$, является вложенным, то зацеплены либо треугольники 012 и 345, либо треугольники 013 и 245, либо треугольники 014 и 235.

6.3. Существует ли в $\mathbb{R}^4$ вложенное семейство всех треугольников $(5,5)$-реализации, кроме одного?

6.4. * Для каких семейств трехэлементных подмножеств 6-элементного множества в $\mathbb{R}^3$ существует вложенное семейство треугольников, отвечающих этому подмножеству?

6.5. Любую ли склейку сторон (плоского двумерного) многоугольника можно осуществить (a) в $\mathbb{R}^3$? (b)* в $\mathbb{R}^4$?

Приведем идею альтернативного доказательства нереализуемости

6.6. * Пусть f — набор из 15 точек $A_{i,j}, 1 \le i \le 3, 1 \le j \le 5$ общего положения в $\mathbb{R}^4$. Для любого $a \in \{3,4,5\}$ рассмотрим множество M_a всех упорядоченных пар, первым элементом которых является грань (т.е. двумерный треугольник) тора $123 \times 12a$, а вторым — несмежный ей отрезок треугольника $1 \times (\{1,3,4,5\} - \{a\})$. Обозначим

$$v(f) := \sum_{a \in \{3,4,5\}} \sum_{(X,Y) \in M_a} |X \cap Y| \mod 2.$$

(a) $v(f)$ не зависит от f.

(b) $v(f) = 1$ для некоторого набора f из 15 точек общего положения в $\mathbb{R}^3$.

6.7. * Пусть f — набор из 25 точек $A_{i,j}, 1 \le i,j \le 5$ общего положения в $\mathbb{R}^4$. Для любых $a \in \{3,4,5\}$ и $b \in \{3,4,5\}$ рассмотрим множество $M_{a,b}$ всех упорядоченных пар, первым элементом которых является грань (т.е. двумерный треугольник) тора $12a \times 12b$, а вторым — несмежная ей грань тора $(\{1,3,4,5\} - \{a\}) \times (\{1,3,4,5\} - \{b\})$. Обозначим

$$v(f) := \sum_{a,b \in \{3,4,5\}} \sum_{(X,Y) \in M_{a,b}} |X \cap Y| \mod 2.$$

(a) $v(f)$ не зависит от f.

b) $v(f) = 1$ для некоторого набора f из 25 точек общего положения в $\mathbb{R}^4$.

Указания

6.1. Аналогично задаче 1.13.35.

6.2. Следует из теоремы Конвея-Гордона-Закса (задача 4.5).

6.3. Не существует. Можно даже удалить некоторые 48 треугольников, все равно не будет существовать. См. решение задачи 1.19.55. Рассмотрите объединение треугольников, которые в нем используются.

REALIZABILITY OF HYPERGRAPHS: EXAMPLES AND ALGORITHMS [1]

A. Zimin, A. Rukhovich, A. Skopenkov and M. Skopenkov

Presented by: A. Zimin, J. Zung, A. Rukhovich, A. Skopenkov and G. Chelnokov

1 Problems up to the semi-final

What are these problems about?

The following problem is well known: can a given graph be embedded in the plane, i.e., can the graph be drawn on the plane so that its edges have no pairwise intersections and no self intersections except at end points? The present cycle of problems is about the embedding of the two-dimensional analogs of graphs (called hypergraphs) in three-dimensional space and even in four-dimensional space. Here some beautiful and nontrivial results on the embedding of hypergraphs will be stated in in the language of certain systems of points, so that we will not need the notion of hypergraph. We will work in four-dimensional space only at the end, when this will not seem scary, because by then we will have learned how to reduce geometric problems to problems of lesser dimension (for more details, see the beginning of the section "Embedding into four-dimensional space").

The most important results are simple proofs of the non-realizability in four-dimensional space of the complete hypergraph on seven vertices and (solution of the Menger Problem) of the product $K_5 \times K_5$. [2] See Problems 1.19.a,mn; another most interesting problems are 1.11, 1.12 and 1.13.

Participants who succeed in solving the problems can obtain *additional problems* about piecewise linear realizability and algorithmic questions of realizability.

General conventions

Whenever the formulation of the problem is an assertion, the problem is to prove the assertion.

A student (or a team of students working together on a problem) obtains a 'bonus point' for each written solution of any problem, for which he/she/it received a + or a +.. *A large size picture that jury member can understand counts as a written solution* for those items of Problems 1.13 and 1.19 where the answer is 'yes'. The Jury will also award more bonus points for beautiful solutions, for solutions of difficult problems and for (some) solutions typeset in TeX. The Jury has an unlimited number of bonus points. The student (or the team) can try to present solutions orally (without handing in a written text), but it loses one bonus point at each such attempt.

We ask all participants working on our problems to feel free to *discuss with us* any questions that may arise or ideas of solutions.

The solutions of Problems 1.2.abc, 1.4.ab and 1.7 will be given at the presentation, you may submit these problems only before the presentation.

Realizability in the plane

A collection (=subset) of points in the plane is *in general position*, if no 3 points of the collection belong to one line.

By *n points on the plane (in space)* we mean an *n-element subset of the plane (space).* So these n points are considered to be different.

The following assertion can be used further without proof.

[1] After the Summer Conference a Russian up-to-date version will be available as a part of the book www.mccme.ru/circles/oim/algor.pdf. We are grateful to A. Sossinsky for English translation of parts of the text, to P. Kozhevnikov for useful discussions and to I. Bogdanov for preparing some figures.

[2] Usually the proof of these classical examples involves complicated techniques [Pr06]. Actually the proofs of the *topological* unrealizability is indeed harder compared to the *linear* unrealizability mentioned above, and the *piecewise-linear* unrealizability, for which the proofs are analogous. [Sk03], [Sk08, §5].

1.1. *Parity Theorem.* If the 6 vertices of two triangles in the plane are in general position, then the contours of the triangles intersect in an even number of points.

A subset of the plane or space is called *convex*, if for any two points from this subset the segment joining these two points is in this subset. The *convex hull* of set X is the minimal convex set that contains X.

1.2. (a) There exist 4 points in the plane such that for any of their decompositions into two pairs , the segment joining the points of the first pair does not intersect the segment joining the points of the second pair.

(b) Any 4 points in the plane can be decomposed into two groups such that the convex hull of the points of the first group intersects the convex hull of the points of the second group.

(c) From any 5 points in the plane, one can choose two disjoint pairs of points such that the segment joining the points of the first pair intersects the segment joining the points of the second pair.

(d) Two triples of points are given in the plane. Then there exist two intersecting segments without common vertices such that each segment joins the points from distinct triples.

A set of segments (in the plane or in space) is called *embedded*, if the following conditions hold:

• segments without common vertices are disjoint; and

• segments with a common vertex intersect only at this vertex.

Perhaps you would like to solve the following weaker versions before solving points (c) and (d) themselves:

(c') For any 5 points in the plane, the set of all the segments joining them is not embedded.

(d') For any two triples of points in the plane, the set of all the segments joining the points from distinct triples is not embedded. [3]

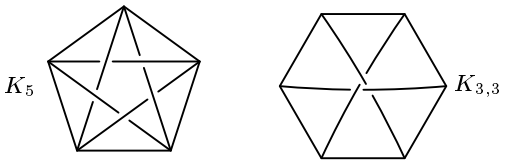


Figure 1: Nonplanar graphs

1.3. Count the number of unordered pairs $\{\{i,j\},\{k,l\}\}$ of disjoint two-element subsets $\{i,j\},\{k,l\} \subset \{1,2,3,4,5\}$.

1.4. Let a collection $f := \{1,2,3,4,5\}$ of five points in general position in the plane be given. For any four distinct points i,j,k,l of the collection, the segments ij and kl either are disjoint or have a unique common point. Define $v(f)$ to be the parity of the number of intersection points of the segments ij and kl for all unordered pairs $\{\{i,j\},\{k,l\}\}$ of disjoint two-element subsets $\{i,j\},\{k,l\} \subset f$:

$$v(f) := \sum \{|ij \cap kl| \ : \ \{i,j\},\{k,l\} \subset \binom{f}{2},\ \{i,j\} \cap \{k,l\} = \emptyset\} \mod 2.$$

(a) For the collection f_0 of five points in the plane shown in Figure 1 to the left we have $v(f_0) = 1$.

(b) $v(f)$ does not depend on f.

1.5. (a,b) State and prove the analogues of Problems 1.4.a,b for six points (in general position in the plane) decomposed into two triples.

(c,d) State and prove the analogues of Problems 1.2.c,d for points in the sphere.

[3] Of course these assertions are versions of the nonplanarity of K_5 and $K_{3,3}$. But they are easier to prove: it is sufficient to use Problem 1.1 instead of nontrivial versions of the Jordan theorem. If your solution uses such versions, then please do not forget to prove them.

1.6. Is it possible to draw without self-intersections graphs K_5 and $K_{3,3}$ (fig. 1)
(a) on the sphere? (b) on the lateral surface of the cylinder (Fig. 2)?
(c) on the torus (Fig. 2)? (d) on the Möbius strip (Fig. 2)?

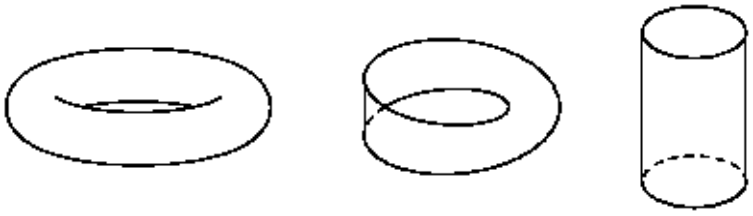


Figure 2: Torus, Möbius strip and a cylinder

The *torus* is the surface of a doughnut (Fig. 2, left). Or, equivalently, the figure obtained by gluing the opposite sides of the square in 'the same directions', i.e. without a twist. The *Möbius strip* is the figure obtained by gluing the short opposite sides of a long rectangular strip in 'opposite directions', i.e., after a 180° twist (Fig. 2, middle).

Realizability in space
The main problems of this subsection are 1.11, 1.12 and 1.13.

1.7. There exist 100 points in space (i.e. in $\mathbb{R}^3$) such that the set of all the segments joining the points is embedded.

A set of points in space is *in general position*, if no 4 points of the set belong to one plane.

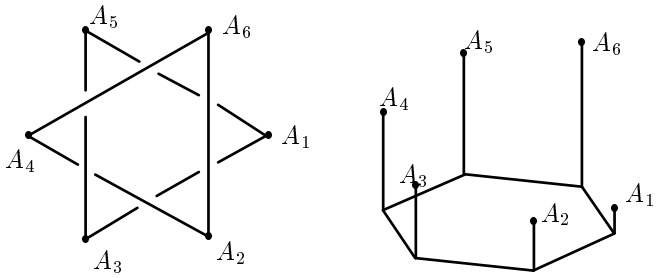


Figure 3: A set of points in general position

1.8. The following sets of points are in general position:
(a) (See Figure 3.) Consider a regular hexagon in a horizontal plane. The set of points A_1, A_2, A_3, A_4, A_5, A_6 exactly above the vertices of the hexagon at the heights 1, 2, 3, 4, 5, 6, respectively.
(b) The points with *Cartesian coordinates* $(t; t^2; t^3)$, where $t \in (0, 1)$.

1.9. (a) There exist 4 points in space which cannot be decomposed into two groups such that the convex hull of the points of the first group intersects the convex hull of the points of the second group.
(b) Any 5 points in space can be decomposed into two groups such that the convex hull of the points of the first group intersects the convex hull of the points of the second group.

1.10. Several points in general position and a point O are marked in space. It is known that for any three marked points A, B, C there is a marked point D such that the point O belongs to the interior of the tetrahedron $ABCD$. Prove that exactly 4 points are marked.

1.11. (a) From any 6 points in space one can choose 5 points O, A, B, A', B' such that the two-dimensional triangles OAB and $OA'B'$ have a common point other than O.
(b) For 5 points an analogous assertion is not true.

A set of two-dimensional triangles in space is *embedded*, if the following conditions hold:

- triangles without common vertices are disjoint;
- triangles with exactly one common vertex intersect only at this vertex; and
- triangles with a common side intersect only along this side.

These conditions formalize the 'non-existence of self-intersections in the construction'. Accurate checking of these conditions in the proofs is required only the first time or when the jury asks to do it (this will not be the case when these conditions are clear from the construction).

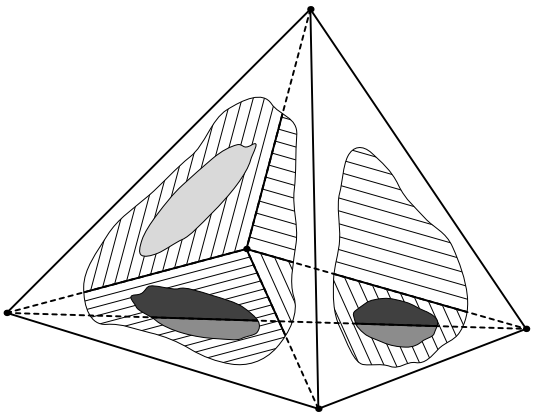


Figure 4: Five points in space: the vertices and the center of a tetrahedron

For instance, in Figure 4 one can see 5 points in space such that the set of all triangles with the vertices at these points is embedded. Problem 1.11.a shows that no 6 points with this property exist.

1.12. (a) There exist 6 points $A_0, A_1, \ldots, A_5 \in \mathbb{R}^3$ such that the set of all triangles $A_0A_jA_k, 1 \le j \le k \le 5, k \ne 2$ is embedded.

(b) Suppose that we have 5 points in $\mathbb{R}^3$. If the set of all triangles with vertices at these points is embedded, then for each point of $\mathbb{R}^3$ one of the 5 segments connecting this point and one of given ones intersects at least one triangle with vertices at given points.

You may prove the existence by an explicit construction of required points. While proving that the intersection is unavoidable, you can use without proof facts like 'the surface of a convex polyhedron splits $\mathbb{R}^3$ into 2 parts' if these facts are confirmed by Jury.

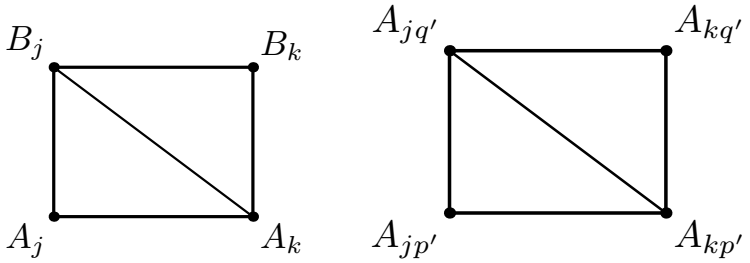


Figure 5: To the problem on a cylinder and on a Cartesian product

1.13. (n) *Cone.* For which n does there exist $n+1$ points $O, A_1, \ldots, A_n \in \mathbb{R}^3$ such that the set of all the triangles

$$OA_jA_k, \quad 1 \le j < k \le n,$$

is embedded? (Items (4), (5) of the problem are accepted separately.)

(lmn) *Join.* For which l, m, n does there exist $l+m+n$ points $A_1, \ldots, A_l, B_1, \ldots, B_m, C_1 \ldots, C_n \in \mathbb{R}^3$ such that the set of all the triangles

$$A_iB_jC_k, \quad 1 \le i \le l, \quad 1 \le j \le m, \quad 1 \le k \le n,$$

is embedded? (Items (222), (223), (233) are accepted separately.)

(2n) *Cylinder.* For which n does there exist $2n$ points $A_1, \dots, A_n, B_1, \dots, B_n \in \mathbb{R}^3$ such that the set of all the triangles

$$A_jB_jA_k \quad \text{and} \quad A_kB_kB_j, \quad 1 \le j < k \le n,$$

is embedded? (Items (24), (25) are accepted separately.)

(mn) *Cartesian product.* For which m, n does there exist mn points $A_{j,p} \in \mathbb{R}^3$, $j \in \{1, 2, \dots, m\}$, $p \in \{1, 2, \dots, n\}$, such that the set of all the triangles

$$A_{j,p}A_{j,q}A_{k,p} \quad \text{and} \quad A_{k,q}A_{k,p}A_{j,q}, \quad 1 \le j < k \le m, \quad 1 \le p < q \le n,$$

is embedded? (Items (33), (34), (35), (44) are accepted separately.)

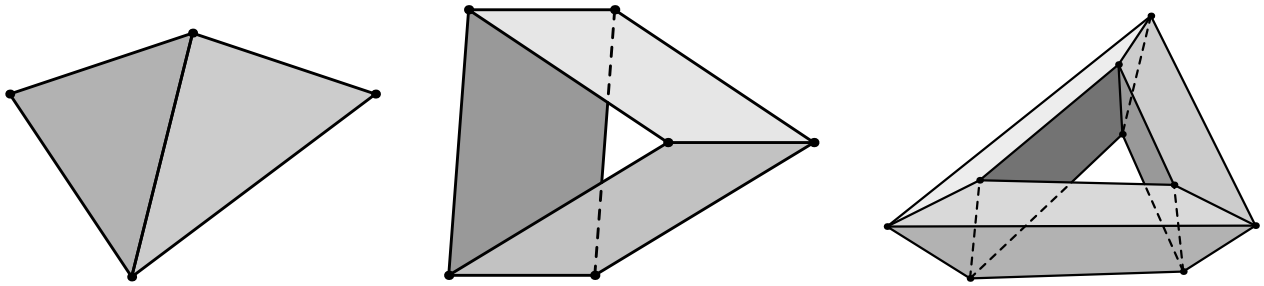


Figure 6: So (m, n)-realizations look like

We will call the embedded set of two-dimensional triangles from Problem 1.13.mn an (m, n)-*realization in* $\mathbb{R}^3$. (A more common term is a *linear embedding of the complex* $K_m \times K_n$.) An (m, n)-realization is an *annulus* if $m = 3$ and $n = 2$, or a *torus* if $m = n = 3$.

Realizability in four-dimensional space

How to work with four-dimensional space? One can define

- the *line* as the set of all real numbers;
- the *plane* as the set of all ordered pairs (x, y) of real numbers x and y;
- *three-dimensional space* as the set of all ordered triples (x, y, z) of real numbers;
- *four-dimensional space* as the set of all ordered quadruples (x, y, z, t) of real numbers.

Then one can 'analytically' define lines in a plane, lines and planes in three-dimensional space, lines, planes and (three-dimensional) hyperplanes in three-dimensional space. However, only the simplest properties of planar and spatial geometric objects are deduced from the analytic definition (or just accepted as axioms). More complicated properties can be deduced 'synthetically' from the simplest ones (i.e., as in school geometry, without using the analytic definition). Often it is convenient to reduce a planar problem to a linear one (i.e., to a problem in a line), and a spatial problem to a planar one. Similarly, the most important approach to the following four-dimensional problems is a reduction to spatial ones. While solving problems about $\mathbb{R}^4$, you can use without proof all rigorously formulated facts about solutions of systems of linear equations, if these facts are confirmed by Jury.

The definition of an embedded set of two-dimensional triangles in four-dimensional space is analogous to the three-dimensional case. One should only replace 'space' by 'four-dimensional space'.

An example of an argument with four-dimensional space.

Let us prove that there exist 101 points $O, A_1, \dots, A_{100}$ in four-dimensional space such that the set of all the triangles OA_jA_k, $1 \le j < k \le 100$, is embedded. This proof is analogous to the solution of Problem 1.13.4 (see Figure 7). Take 100 points $O, A_1, \dots, A_{100}$ in a three-dimensional hyperplane in four-dimensional space such that the set of all the segments joining them is embedded (see Problem 1.7). Take a point O in four-dimensional space not belonging to the three-dimensional hyperplane. Then the points $O, A_1, \dots, A_{100}$ are the required ones.

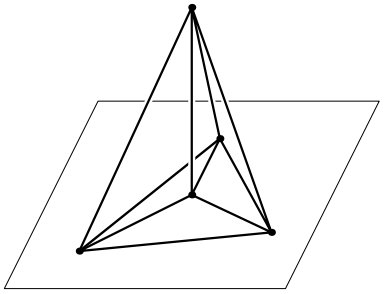


Figure 7: 101 points in four-dimensional space; four-dimensional space is shown as three-dimensional space and a (three-dimensional) hyperplane in four-dimensional space is shown as a two-dimensional plane in three-dimensional space

1.14. (a) For each two points, which are not in the plane $x = y = 0$ in $\mathbb{R}^4$, there exists a broken line which connects these points and does not intersect this plane.

(b) For each hyperplane in $\mathbb{R}^4$, there exist two points not in this hyperplane such that each broken line connecting them intersects this hyperplane.

1.15. (a) There exist 5 points in $\mathbb{R}^4$ which cannot be decomposed into two groups such that the convex hull of the first group intersects the convex hull of the second group.

(b) Any 6 points in $\mathbb{R}^4$ can be decomposed into two groups such that the convex hull of the first group intersects the convex hull of the second group.

In Problems 1.16 and 1.17 it is sufficient to write the correct answers.

1.16. What is the intersection of the *3-dimensional sphere*

$$S^3 := \{(x, y, z, t) \in \mathbb{R}^4 \mid x^2 + y^2 + z^2 + t^2 = 1\}$$

with the following sets:

(a) the line $x = y = z = 0$, containing the center of the sphere;

(b) the plane $x = y = 0$, containing the center of the sphere;

(c) the (3-dimensional) hyperplane $x = 0$, containing the center of the sphere;

(d) the intersection of the positive sixteenth of $\mathbb{R}^4$ and the union of the 2-dimensional coordinate planes, i.e.

$$\{(x, y, z, t) \in \mathbb{R}^4 \mid x \geq 0,\ y \geq 0,\ z \geq 0,\ t \geq 0 \text{ and two of four numbers } x, y, z, t \text{ are zeros}\}.$$

A set of points in $\mathbb{R}^4$ is in *general position* if no 5 points from the set are in one hyperplane. For example, points $(t, t^2, t^3, t^4), t \in (0, 1)$ are in general position.

1.17. Eight points 1,2,3,4,5,6,7,8 in general position in $\mathbb{R}^4$ are given. What is the intersection of:

(a) the line 12 and the hyperplane 5678? (b) the line 12 and the plane 567?

(c) the plane 123 and the hyperplane 5678? (d) the hyperplanes 1234 and 5678?

(e) the planes 123 and 567?

(Note that our definition of a general position is different from what is usually accepted in such problems.)

1.18. (a) There exist 6 points in $\mathbb{R}^4$ such that the set of all triangles with vertices at these points is embedded.

(b) There exist 7 points in $\mathbb{R}^4$ such that the set of all triangles, except one, with vertices at these points is embedded.

The first item of the following problem shows that an analogous assertion for 7 points is not true.

1.19. Main examples. (a) From any 7 points in $\mathbb{R}^4$ one can choose two disjoint triples such that the triangles formed by the triples intersect each other.

(b) Three triples of points in $\mathbb{R}^4$ are given. Then there exist two intersecting triangles without common vertices such that the vertices of each triangle belong to distinct triples.

(mn) For which m, n does there exist an (m, n)-realization in $\mathbb{R}^4$?

(The definition is analogous to (m, n)-realization in $\mathbb{R}^3$, only one should take points from $\mathbb{R}^4$.)

(Items (35), (3n), (44), (45), (4n), (55) are accepted separately.)

It may be rather difficult to prove that the intersection is unavoidable without the hints (in the form of new problems) which will be given after the semi-final.

1.20. There are 100 points in five-dimensional space such that the set of all the triangles with vertices at these points is embedded.

2 Solutions and hints suggested at the presentation

1.2. (a) A triangle and a point inside it.

(b) Consider points A, B, C, D in the plane.

If some three of these points are on one straight line, then one point, say B, is in the segment with vertices at the two other points, say A, C. Denote by $[XY]$ the segment with vertices at points X, Y. Then $[AC] \cap [BD] \neq \emptyset$.

Suppose that no three of these points are on one straight line. If one of the four points is inside the triangle with vertices at the other three points, then the problem is solved. Suppose that each point of these four is outside the triangle with vertices at the other three points. Then point D is outside triangle ABC. So D is either inside one of the angles symmetric to the angles of ABC with respect to the corresponding vertex of ABC or inside one of the angles of triangle ABC.

Case 1. D is inside one of the angles symmetric to the angles of ABC with respect to the corresponding vertex of ABC. Suppose that D is inside the angle symmetric to $\angle ACB$ with respect to point C. Then point C is inside triangle ABD, a contradiction.

Case 2. D is inside one of the angles of triangle ABC, say $\angle BAC$, by the previous arguments and because no three of these four points are on one line. Point D is outside triangle ABC and inside $\angle BAC$, so points A and D are in different half-planes bounded by line BC. Then $[AD]$ and $[BC]$ intersect.

(c) *First solution.* It follows from 1.4.

(c) *Second solution.* Assume the converse, i.e., there exist 5 points $OABCD$ in the plane such that it is impossible to choose such a pair. Then $A \notin OB$ and $B \notin OA$. So A is not in the half-line OB. So we can think that points A, B, C, D are in the order A, B, C, D if we look at them from point O. Then triangles OAC and OBD intersect in one point (O). This intersection is 'transversal'. So, by the Parity Theorem, (i.e. as in Problem 1.1) $AC \cap BD \neq \emptyset$. Contradiction.

(d) Analogous to the first solution of (c), see Problem 1.5.

1.4. (b) It suffices to prove that for any points $1, 2, 3, 4, s, s' \in \mathbb{R}^2$ in general position and sets

$$A := \{1, 2, 3, 4\}, \quad f := A \cup \{s\} \quad \text{and} \quad f' = A \cup \{s'\} \quad \text{we have} \quad v(f) = v(f').$$

Let us prove this fact. For each $i \in A$ denote by A_i the triangle with vertices from $A - \{i\}$. Then the problem follows from

$$v(f') - v(f) = \sum_{i \in A} (|si \cap A_i| - |s'i \cap A_i|) = \sum_{i \in A} |ss' \cap A_i| = 0 \mod 2.$$

The second equality holds because $|ss'i \cap A_i|$ is even for each $i \in A$ by the Parity Theorem. The last equality holds because for each non-ordered pair $\{i, j\} \subset A$ there exist exactly two triangles with

vertices from A containing the segment ij. So for each non-ordered pair $\{i,j\} \subset A$ the number $|ss' \cap ij|$ appears in the sum twice for two triangles A_i, A_j.

1.5. *Statement.* Assume that there are six points in general position in the plane. Split them into two triples $f_1 = \{1,2,3\}$ and $f_2 = \{4,5,6\}$. For each two points $i,j \in f_1$ and two points $i',j' \in f_2$, the segments ii' and jj' either do not intersect or have only one intersection point. Define $v(f_1,f_2)$ as the parity of the number of intersection points of segments ii' and jj', ij' and ji' for each 2-element subsets $\{i,j\} \subset f_1$, $\{i',j'\} \subset f_2$

$$v(f_1,f_2) := \sum \{|ii' \cap jj'| + |ij' \cap ji'| \ : \ \{i,j\} \in \binom{f_1}{2},\ \{i',j'\} \in \binom{f_2}{2}\} \mod 2.$$

(a) For the sets f_1, f_2 in Figure 1, right, $v(f_1,f_2) = 1$.

(b) $v(f_1,f_2)$ does not depend on f_1, f_2.

(c,d) *A spherical line* is the intersection of the sphere with a plane containing the center of the sphere.

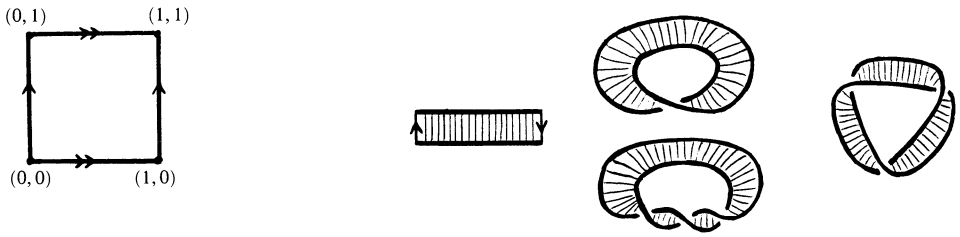


Figure 8: Gluings of a rectangular strip that yield the torus and the Möbius strip

1.6. You can draw a graph not only on the torus or on the Möbius band from fig. 2, but also on a square 'respecting' gluings, Fig. 8.

1.7. Choose three points in 3-space that do not belong to one straight line. Suppose that we have chosen $n \geq 3$ points in general position. Then there is a finite number of planes containing triples of these n points. Then we can choose a point that is neither of these planes. Add this point to our set of n points. The obtained set of $n+1$ points has no four points in one plane, because the new point is not in one plane with any three of these n points. So for each n there exist n points in 3-space that are in general position.

Consider 100 points in 3-space that are in general position. Denote by A the set of all segments joining pairs of these points. If some two segments from A with different endpoints intersect, then four endpoints of these two segments are in one plane. If some two segments from A with common endpoint intersect not only at their common endpoint, then the three endpoints of these two segments are on one line.

1.8. Use coordinates.

1.13. (n),(lmn), (44) Use Problem 1.11.

3 Solutions presented at the semi-final

If the text on a problem starts with a word *hint* of *answer*, then the details (for example, proving propositions stated or completing solutions) remain for your individual work. You can hand in solutions of such problems even after the semi-final.

1.3. Answer-solution: $5 \cdot \binom{4}{2}/2 = 15$.

1.6. (a), (b). Impossible.

(a) Let there be a graph K_5 drawn on the sphere without self-intersections. Remove one point, that does not belong to K_5, from the sphere. We obtain a plane and a graph K_5 on it. A contradiction.

(b) The graph K_5 is non-planar, and a cylinder can be projected onto a plane without self-intersections.

(c), (d). Yes, it is possible. The beautiful realizations of a graph K_5 on the torus and a graph $K_{3,3}$ on the Möbius strip are shown in Fig. 9.

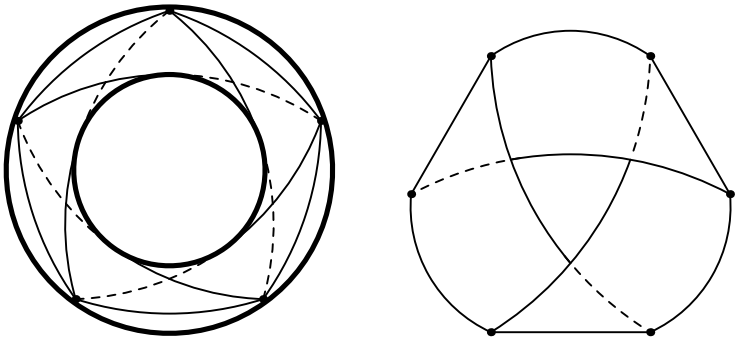


Figure 9: Realization of Kuratowski graphs

Also there are other solutions. For example, draw one of Kuratowski graphs on the plane with exactly *one* intersection and ...

1.9. (a) Any four points that are not in one plane satisfy the statement.

(b) See Radon Theorem at the end of Section 3.

1.10. *Hint.* Denote by $A_1, A_2, ...A_n$ the marked points. Denote by l_i the half-line with endpoint O that contains point symmetric to A_i with respect to point O. Prove the following fact.

Proposition. Consider triangle $A_iA_jA_k$. Point O is inside the tetrahedron $XA_iA_jA_k$ if and only if point X is inside the trihedral angle with vertex O whose sides are half-lines l_i, l_j, l_k.

Consider a tetrahedron whose vertices are marked points containing point O. Without loss of generality we may assume that this is the tetrahedron $A_1A_2A_3A_4$. The union of all trihedral angles with sides $l_1, ..., l_4$ is the 3-space. Each of these angles must contain a marked point by the Proposition. The half-line l_5 is inside exactly one of those trihedral angles. This half-line 'splits' that trihedral angle into 3 angles each of whom must contain a marked point. The half-line l_6 is inside exactly one of those 6 angles. This half-line 'splits' that angle into 3 angles each of whom must contain a marked point. So for $n > 4$ this process is infinite.

1.11. (a) Consider a small sphere with center at any point O of the given ones. The intersection of this sphere with the union of triangles OAB for all pairs A, B of given points is a graph K_5. A contradiction.

Another solution follows from the Conway-Gordon-Sachs Theorem (Problem 4.5). Note that the following proof of the Conway-Gordon-Sachs Theorem in fact repeats the reduction to the non-planarity of the graph K_5 as above.

(b) See Fig. 4.

1.13. *Answers*: (n) $n \le 4$; (lmn) at most one of numbers l, m, n is greater than 2;
(2n) for each n; (mn) either $m < 3$, or $n < 3$, or $m = n = 3$, or $\{m, n\} = \{3, 4\}$.

(35) *Solution.* Suppose to the contrary that there exists a $(3, 5)$-realization in $\mathbb{R}^3$.

A *triangle* $i \times pqr$ is the triangle $A_{i,p}A_{i,q}A_{i,r}$.

An *annulus* $ij \times pqr$ is the $(2, 3)$-realization 'corresponding to subscripts i, j and p, q, r'.

A *torus* $ijk \times pqr$ is the $(3, 3)$-realization 'corresponding to subscripts i, j, k and p, q, r'.

By the Jordan Theorem the torus 123×123 splits the 3-space into two parts. The union $14 \times 123 \cup 24 \times 123$ of the annuli 14×123 and 24×123 is an annulus. This annulus intersects the torus 123×123

along the triangles 1×123 and 2×123. The annulus 34×123 intersects the annulus $14\times 123\cup 24\times 123$ along the triangle 4×123. The annulus 34×123 intersects the torus 123×123 along the triangle 3×123. Denote by $K_4\times K_3$ the $(4,3)$-realization 'corresponding to the subscripts $1,2,3,4$ and $1,2,3$'. By a version of the Jordan Theorem $K_4\times K_3$ splits the 3-space into 4 parts. In the $(5,3)$-realization the vertex $A_{5,1}$ is joined

• to the vertex $A_{i,1}$ by the segment $A_{5,1}A_{i,1}$, for each $1\le i\le 4$;

• the vertex $A_{i,j}$ by the broken line $A_{5,1}A_{5,j}A_{i,j}$, for each $1\le i\le 4$, $2\le j\le 3$.

Since subscript 5 does not 'participate' in $K_4\times K_3$, each of these segments and broken lines intersects $K_4\times K_3$ only at the endpoints. Consider the connected component of $\mathbb{R}^3-K_4\times K_3$ that contains the point $A_{5,1}$. The boundary of this component contains the point $A_{i,j}$ for each $1\le i\le 4$ and $1\le j\le 3$, because this point is joined to $A_{5,1}$ by a segment or a broken line whose interior is disjoint with $K_4\times K_3$. Thus this boundary contains 12 points $A_{i,j}$ for $1\le i\le 4$ and $1\le j\le 3$. On the other hand, this boundary is a torus, i.e. a (3,3)-realization. Hence this boundary has only 9 points of the given 15 points. A contradiction.

1.15.(a) Five points that are not in one 3-space.

(b) See Radon Theorem at the end of Section 3.

1.18.(a) Take 5 vertices of 4-dimensional simplex and a point inside it.

(b) Let $ABCD$ be a regular tetrahedron in $\mathbb{R}^4$ and let E be centre of this tetrahedron. Choose a point X on the interior of $ABCE$ so that points A,B,C,D,E,X are in general position in $\mathbb{R}^3$. Erect a line l which is perpendicular to the hyperplane $ABCD$ and intersects $ABCD$ at X. Finally, choose points X_1, X_2 on l which are on opposite sides of X. Let us prove that the set $V=\{A,B,C,D,E,X_1,X_2\}$ of seven points is as required, i.e. the set $\binom{V}{3}\backslash\triangle X_1X_2D$ of triangles is embedded.

Let α,β,γ be distinct points from $\{A,B,C,D,E\}$. Now there are three classes of triangles:

• $\triangle X_1X_2\alpha$ for $\alpha\ne D$;

• $\triangle X_i\alpha\beta$ for $i\in\{1,2\}$;

• $\triangle\alpha\beta\gamma$.

It is easy to check that the set of triangles from each class are embedded.

A triangle from class 1 intersects a triangle from class 2 either in a common vertex X_i or in a common edge $X_i\alpha$.

A triangle from class 2 intersects a triangle from class 3 either in a common vertex α or a common edge $\alpha\beta$. Finally, let's consider the intersection of class 1 triangles and class 3 triangles.

Consider the intersection of triangles from class 1 and class 3. A triangle $\triangle X_1X_2\alpha$ intersects hyperplane $ABCD$ in the line segment $X\alpha$. Since X lies in $ABCE$, we have that XA, XB, XC, and XE intersect class 3 triangles in at most a common vertex.

Therefore the set $\binom{V}{3}\backslash\triangle X_1X_2D$ of triangles is embedded.

1.19. (a) Analogous to Problems 1.2.c and 1.4. Follows from Problem 4.3.

(b) Analogous to Problems 1.2.d and 1.5. Follows from Problem 4.4.

(mn) Answer: $\min\{m,n\}\le 4$.

(4n) *Hint.* Let us prove that there exist a $(4,n)$-realization in $\mathbb{R}^4$. Take points $A_{j,1}$, $1\le j\le n$ in general position in $\mathbb{R}^4$. Take an ordered set K of four points in the plane in $\mathbb{R}^4$ such that the fourth of them is inside two-dimensional triangle, formed by other three points. For example, $K:=((0,0,0,0),(2,0,0,0),(1,2,0,0),(1,1,0,0))$. Take the images of this set under translations by vectors $A_{j,1}$, $1\le j\le n$. I.e. denote $(A_{j,1},A_{j,2},A_{j,3},A_{j,4}):=K+A_{j1}$. Then:

• for each $1\le j\le n$ point $A_{j,4}$ is inside two-dimensional triangle $A_{j,1}A_{j,2}A_{j,3}$;

• for each i,j sets $K+A_{i,1}$ and $K+A_{j,1}$ are congrous by translation by vector $A_{j,1}-A_{i,1}$;

• for each i,j,k all 12 points from the set $(K+A_{i,1})\cup(K+A_{j,1})\cup(K+A_{k,1})$ are not in one 3-space because points $A_{j,1}$, $1\le j\le n$ are in general position.

Deduce from these facts that these points form the required $(4,n)$-realization in $\mathbb{R}^4$.

(55) *Hint.* See Problems 4.14 and 4.15.

1.20. Analogous to Problem 1.7. A set of points in $\mathbb{R}^5$ is *in general position* if no six of these point are in one four-dimensional hyperplane.

Denote convex hull of set V by $conv(V)$.

Radon theorem. *Given $n+2$ points in $\mathbb{R}^n$, one can split into two sets $\{X_1, \dots X_k\}$ and $\{X_{k+1}, \dots, X_{n+2}\}$, such that $conv\{X_1, \dots, X_k\} \bigcap conv\{X_{k+1}, \dots, X_{n+2}\} \neq \emptyset$.*

Proof. We identify points and vectors in $\mathbb{R}^n$. Lets prove that there exist $c_1, \dots c_{n+2} \in \mathbb{R}$, some of which are nonzero, such that

$$c_1X_1 + c_2X_2 + \dots + c_{n+2}X_{n+2} = 0 \quad \text{and} \quad c_1 + \dots c_{n+2} = 0.$$

Indeed, consider vectors $X_1 - X_{n+2}, X_2 - X_{n+2}, \dots, X_{n+1} - X_{n+2}$. Since this is a set of $n+1$ vectors in $\mathbb{R}^n$, there exists a non-trival linear dependence $c_1(X_1 - X_{n+2}) + \dots + c_{n+1}(X_{n+1} - X_{n+2}) = 0$. Thus the set $c_1, \dots, c_{k+1}, -c_1 - \dots - c_{n+1}$ is as required.

Rearrange our points so that all positive c_i will be in the beginning. Bring the summands with negative c_i to the right side: $c_1X_1 + \dots + c_kX_k = -c_kX_k - \dots - c_{n+2}X_{n+2}$. Multiply this equation by a positive constant such that the sums of coefficients on the left side and on the right side equal 1. The obtained equation implies that $conv\{X_1, \dots, X_k\}$ and $conv\{X_{k+1}, \dots, X_{n+2}\}$ have a common point.

4 Problems suggested after the semi-final

Hard Problems about non-realizability could be solved by two different ways. The first way is to generalize a proof of non-planarity of graph K_5 (i.e. the first solution of Problem 1.2.c) using the Van Kampen obstruction (Problem 1.4). This way is realized in the first subsections of this section. The second way is to generalize another proof of non-planarity of graph K_5 (i.e. the second solution of Problem 1.2.c). It is based on the reduction to the lower dimension. This way is realized in the second and the third subsections of this section.

Problems 1.19.a,b are implied by the following Problems 4.3, 4.4. The nonexistence of the $(5, 5)$-realization in $\mathbb{R}^4$ follows from a spherical version of the Sachs Theorem (Problem 4.9.b) and the following Problems 4.14, 4.15. To get closer to this idea, first you could solve the Problem 1.13.53 in other way using Problems 4.10, 4.11 and also 1.19.a using a spherical version of the Conway-Gordon-Sachs Theorem (Problem 4.9.a) and Problems 4.12, 4.13.

We call *a triangle* (in $\mathbb{R}^3$ or in $\mathbb{R}^4$) its contour (i.e. a clozed broken line), and *a two-dimensional triangle* a convex hull of its vertices. Analogously we call *a tetrahedron* (in $\mathbb{R}^3$ or in $\mathbb{R}^4$) the union of its two-dimensional faces, and *a three-dimensional tetrahedron* a convex hull of its vertices.

Let us call a (m, n)-realization (in $\mathbb{R}^3$ or in $\mathbb{R}^4$) also the union of triangles of a (m, n)-realization.

Generalizations of the Van Kampen obstruction

4.1. (a) For each points 1,2,3,4,5,6,7 in general position in $\mathbb{R}^3$ triangle 123 and two-dimensional tetrahedron 4567 intersect by finite set of points. [4]

(b) For each points 1,2,3,4,5,6,7,8 in general position in $\mathbb{R}^4$ two-dimensional tetrahedrons 1234 and 5678 intersect by finite set of points.

4.2. Find the number of all non-ordered pairs $\{\{i, j, k\}, \{l, m, n\}\}$ of disjoint three-element subsets of a seven-element set.

[4] In your solutions of this problem you must not use without proof the Jordan Theorem, that a polyhedron splits $\mathbb{R}^3$ into two parts (because the Jordan Theorem is proved using this Problem).

4.3. Let a set $f := \{1, 2, 3, 4, 5, 6, 7\}$ of seven general position points in $\mathbb{R}^4$ be given. For any six different points i, j, k, l, m, n two-dimensional triangles ijk and lmn do not intersect or intersect at a unique point. Denote $v(f)$ as the parity of the number of intersection points of two-dimensional triangles ijk and lmn for all non-ordered pairs $\{\{i, j, k\}, \{l, m, n\}\}$ of disjoint three-element subsets $\{i, j, k\}, \{l, m, n\} \subset f$:

$$v(f) := \sum \{|ijk \cap lmn| \ : \ \{\{i, j, k\}, \{l, m, n\}\} \subset \binom{f}{3}, \ \{i, j, k\} \cap \{l, m, n\} = \emptyset\} \mod 2.$$

(a) For set f_0 of seven points from the solution of Problem 1.18.b, $v(f_0) = 1$.
(b) $v(f)$ does not depend on f.

4.4. (a,b) State and proof the analogs of Problems 4.3.a,b for three triples of points in four-dimensional space such that all nine points are in general position.

Elements of Ramsey linking theory

In this section, we will sketch the proof of the linear cases of the Conway–Gordon–Sachs and Sachs theorems (Problems 4.5 and 4.8). They will be needed in the impossibility proof in the main four-dimensional examples and, at the same time, are interesting in themselves. Such statements, as well as their methods of proof constitute *Ramsey linking theory*. For more details, see [PS05].

Triangles Δ and Δ' in space whose six vertices are in general position are said to be *linked* if Δ intersects the interior of triangle Δ' in exactly one point. For example, triangles $A_1A_3A_5$ and $A_2A_4A_6$ from Problem 1.8.a are linked.

4.5. *Conway–Gordon–Sachs Theorem for linear embeddings.* For any 6 points in general position in space, there are two linked triangles with vertices at these points.

The next problem 4.6 is not necessary for the proof of the Conway–Gordon–Sachs Theorem, but it clarifies the notion of linking.

4.6. (a) If one of the triangles Δ, Δ' whose six vertices are in general position, does not intersect the plane of the other triangle, then Δ and Δ' are not linked.

(b) Suppose that two red points and two blue points are marked on a straight line, the 4 points being pairwise distinct. We say that the four points are *linked* if they alternate: red-bue-red-blue or vice-versa.

Triangles Δ and Δ' are linked $\Leftrightarrow$ the common line l of the planes of the triangles intersects each of them in two points and these pairs of points are linked.

(c) If the vertices of two triangles in space are continuously moved so that they remain in general position, then the triangles remain linked or unlinked.

(d) Triangles Δ and Δ' are linked if and only if Δ' and Δ are linked.

(e) For what positions of the point A_1 on the vertical line are the triangles $A_1A_3A_5$ and $A_2A_4A_6$ from Problem 1.8.a are linked?

A plane is in *general position* w.r.t. a set of points in $\mathbb{R}^3$ if orthogonal projections of these points onto the plane are in general position.

4.7. (a) Assume that we have the projection of two triangles on a general position plane, and on the projection it is shown which of the sides passes above the other at the intersection points of the projections (as in Fig. 3, left). Then the triangles are linked if and only if the number of intersection points of the projection at which the first triangle passes above the second triangle, is odd.

(b) Suppose 6 points in general position are given. We say that a non-ordered pair of triangles with vertices at these points with no common vertices is a *splitting* of the 6 points. Then the number of linked splittings is odd.

Two closed quadrangular broken lines $ABCD$ and $A'B'C'D'$ in space whose 8 vertices are in general position in space are called *linked* if the number of transversal intersection points of the broken line

$ABCD$ with the union of the interiors of the triangles $A'B'C'$ and $A'D'C'$ is odd.

4.8. *The Sachs Theorem for linear embeddings.* Suppose we are given 4 red points and 4 blue points in space such that any two line segments with endpoints of different colors have no common interior points. Then there are two linked closed quadrangular broken lines with vertices at these points each edge of which has endpoints of different colors.

4.9. (a,b) State and prove analogs of problems 4.5 and 4.8 replacing space by S^3. (Linking is defined similarly to the case of $\mathbb{R}^3$. Triangles Δ and Δ' such that no four of their six vertices are in one two-dimensional sphere with center in the center of S^3, are called *linked* if Δ intersects a 2-dimensional spherical triangle spent by Δ' in exactly one point. Note that there are exactly two such 2-dimensional spherical triangles.)

Applications of Ramsey link theory

4.10. Suppose that a closed broken line of length 3 and a $(3,3)$-realization N in $\mathbb{R}^3$ have a unique common point x, which is their common vertex. Then any sufficiently small sphere S^2 with the center x intersects the broken line at a pair of points belonging to one connected component of the complement $S^2 - N$. (Before the half-final the facts like this could be used without proof. But here we suggest you to prove it to prepare for proving analogous fact for $\mathbb{R}^4$.)

4.11. Assume that there exists a $(5,3)$-realization in $\mathbb{R}^3$.

(a) The intersection of any sufficiently small sphere with the center $A_{1,1}$ and the $(5,3)$-realization is the graph $K_{4,2}$ 'linearly' embedded into the sphere.

(b) Suppose that in this graph we have a cycle without self-intersections and a pair of vertices not belonging to the cycle. Then in the given $(5,3)$-realization there exist a closed broken line of length 3 and a $(3,3)$-realization intersecting each other in a unique point and intersecting the sphere at the given cycle and the given pair of vertices, respectively.

4.12. Assume that two (two-dimensional) tetrahedra in $\mathbb{R}^4$ have a unique intersection point x, which is their common vertex. Then each sufficiently small three-dimensional sphere S^3 with the center x intersects the tetrahedra by a pair of spherical triangles that are not linked in S^3.

4.13. Assume that there are 7 points $0,1,2,3,4,5,6$ in $\mathbb{R}^4$, among which one cannot choose two disjoint triples such that the two-dimensional triangles formed by these triples intersect each other.

(a) The intersection of a sufficiently small three-dimensional sphere with the center 0 with the union of all the triangles $0ij$, $1 \le i < j \le 7$, is a complete graph K_6 embedded into S^3.

(b) For any decomposition $\{1,2,3,4,5,6\} = \{i,j,k\} \cup \{p,q,r\}$ the tetrahedra $0ijk$ and $0pqr$ intersect at a unique point.

4.14. State and prove an analogue of Problems 4.12 for two $(3,3)$-realizations in $\mathbb{R}^4$.

4.15. Assume that there is a $(5,5)$-realization in $\mathbb{R}^4$.

(a) The intersection of each sufficiently small three-dimensional sphere with the center $A_{1,1}$ and the $(5,5)$-realization is the graph $K_{4,4}$ linearly embedded into the three-dimensional sphere.

(b) For each two disjoint non-self-intersecting cycles in this graph there exist two $(3,3)$-realizations in the $(5,5)$-realization, which intersect each other at a unique point and intersect the sphere at the given cycles.

5 Solutions presented after the final

1.1. Cf.[BE82, §5]. Let A, B, C, D, E, F be 6 general position points on a plane. Note that the intersection of the triangle ABC and two-dimensional triangle DEF is the union of finite number of broken lines, each of whom is a subset of ABC. The endpoints of these broken lines form the set

$ABC \cap DEF$. A closed broken line has zero endpoints, an unclosed one has 2, thus the number $|ABC \cap DEF|$ is even.

1.13. (mn) *For $m = n = 4$.* Prove that (4,4)-realization in $\mathbb{R}^3$ does not exist. Analogous to the solution of Problem 1.11.a. The intersection of a small sphere with center at $A_{1,1}$ with the (4,4)-realization is the graph $K_{3,3}$ linearly embedded to this sphere. A contradiction.

For $m = 3, n = 5$. Prove that (3,5)-realization in $\mathbb{R}^3$ does not exist. *Another solution.* Analogous to the proof of non-planarity of graph K_5 (i.e. the second solution of Problem 1.2.c). Suppose that there exists a $(5, 3)$-realization in $\mathbb{R}^3$. Consider a small sphere S^2 with center at $A_{1,1}$. The intersection of S^2 and $(5, 3)$-realization is a linear embedding of the graph $K_{4,2}$

In every embedding the of graph $K_{4,2}$ on the sphere there exist two vertices X, Y of a part with 4 vertices that are in different pcomponents bounded by cycle Σ formed by other four vertices. Without loss of generality assume that the segments, corresponding to vertices X, Y connect the vertex $A_{1,1}$ with vertices $A_{2,1}$ and $A_{3,1}$. Denote by γ_{XY} a broken line $A_{1,1}A_{2,1}A_{3,1}$ and by γ a torus 145×123. Then $\gamma_{XY} \cup S^2 = \{X, Y\}$ and $\gamma \cup S^2 = \Sigma$.

Because on the sphere S^2 vertices X, Y are in different regions bounded by cycle Σ, the intersection of γ and γ_{XY} at the point $A_{1,1}$ is transversal. A contradiction with the Parity Theorem (Problem 4.10).

1.17. (a), (e) A point or the empty set.

(b) The empty set.

(c) A line or the empty set.

(d) A plane or the empty set.

4.1. Similar to 1.1.

4.2. Answer-hint: $7 \cdot \binom{6}{3}/2 = 70$.

4.3. (a) See the solution of Problem 1.18.b. The intersection of two-dimensional triangle X_1X_2D and three-dimensional tetrahedron $ABCD$ is a segment $[DX]$. Since X is inside three-dimensional tetrahedron $ABCD$ and D is outside it, we have that $[DX]$ intersects $\square ABCE$ at a unique point. So the two-dimensional triangle X_1X_2D intersects only one of the two-dimensional triangles with vertices in other points, i.e. $v(f) = 1$.

(b) It suffices to prove that for general position points $1, 2, 3, 4, 5, 6, v, v'$ in $\mathbb{R}^4$ and sets

$$A := \{1, 2, 3, 4, 5, 6\}, f := A \cup \{v\}, f' := A \cup \{v'\} \quad \text{we have} \quad v(f) = v(f').$$

Denote by T_{ij} the 2-dimensional tetrahedron with vertices from $A - \{i, j\}$.

$$v(f) - v(f') = \sum_{\{i,j\} \in \binom{A}{2}} (|vij \cap T_{ij}| - |v'ij \cap T_{ij}|) = \sum_{\{i,j\} \in \binom{A}{2}} vv'i \cap T_{ij} = 0 \mod 2$$

The second equality follows from the problem 4.1.(b). The last equality holds because for any $a, b, c \in A - \{i\}$, there exist exactly two tetrahedrons T_{ij}, T_{ik}, $j, k \in (A - \{i, a, b, c\})$, that contain this triangle. So each summand $|vv'i \cap abc|$ enters this sum twice. Then $v(f) = v(f')$.

4.4. *Statement.* Consider 9 points $1, 2, \dots 9$ in $\mathbb{R}^4$ satisfying the following condition: for each $i, i' \in f_1 = \{1, 2, 3\}$ $i \neq i'$, each $j, j' \in f_2 = \{4, 5, 6\}$ $j \neq j'$ and each $k, k' \in f_3 = \{7, 8, 9\}$ $k \neq k'$ we have that the two-dimensional triangles ijk and $i'j'k'$ intersect in at most one point. *Remark:* this condition is weaker than the condition that points $1, \dots 9$ are in general position. Define

$$v(f_1, f_2, f_3) := \left(\frac{1}{2} \sum \{|ijk \cap i'j'k'| : (i, i') \in f_1^2, (j, j') \in f_2^2, (k, k') \in f_3^2\} \right) \mod 2.$$

(a) *Example.* Consider the graph $K_{3,3}$ and call the vertices of one of its parts $1, 2, 5$, of the other part $3, 4, 6$. Embed this graph in $\mathbb{R}^3$ in such a way that $1, 2, 3, 4$ are vertices of a square and no other 4

points are in one plane. Let 7 and 8 be in different half-spaces cut out by a hyperplane containing the embedding of the $K_{3,3}$, in $\mathbb{R}^4$. Finally, take a point, say 9, inside the pyramid 12347.

Check that $v(\{1,2,3\},\{4,5,6\},\{7,8,9\}) = 1$.

(b) Analogously to the Problems 4.3.b and 1.5.

4.5. Follows from Problem 4.7.b.

4.6. (a) It is sufficient to prove that if triangles Δ and Δ' are linked, then triangle Δ' intersects plane containing Δ and triangle Δ intersects plane containing Δ'.

The first fact is obvious. Then the interior of Δ intersects plane containing Δ'. So Δ intersects plane containing Δ'.

(b) Let triangle Δ be linked with triangle Δ'. By (a) we have that Δ intersects plane containing Δ', so $\Delta \cap l \neq \emptyset$. By general position, the intersection $\Delta \cap l$ is exactly two points. Now let us prove that the pairs of points are proved. Denote by A, B the intersection points $\Delta \cap l$ and by A', B' the intersection points $\Delta' \cap l$. The intersection of Δ and two-dimensional triangle Δ' is a subset of segment $A'B'$. So exactly one of points A and B is in the segment $A'B'$. So pairs A, B and A', B' are linked in l.

Let us prove the converse. Let the pairs A, B and A', B' be linked in l. Then according to the previous assertion, Δ intersects the two-dimensional triangle Δ' by exactly one point, i.e. Δ is linked with Δ'.

(c) Let us use (b). While moving points so that they remain in general position and the planes containing triangles are not parallel, the four points in the line l move continuously, so they are either always linked or always unlinked. If there is a moment when the planes containing the triangles are parallel, then by (a) at this moment and during some time before and after this moment the triangles are unlinked.

(d) Each of these two conditions is equal to the condition from (b).

(e) Denote by t height of point A_1 above the horizontal plane. The triangles are linked if $t \in (-\infty;\, 2) \cup (3,5;\, 4,5) \cup (6;\, +\infty)$ and are unlinked if $t \in (2;\, 3,5) \cup (4,5;\, 6)$. This follows from Problem 1.8.a and the following lemma (which is implied by (b)).

The motion Lemma. *Let the vertex A of the triangle Δ be moving with constant velocity along a segment in $\mathbb{R}^3$, and let other two vertices and triangle Δ' be fixed. Denote by Δ_t the position of triangle at the moment t, for $0 \leq t \leq 2$. Suppose that 6 vertices of triangles Δ_t and Δ' are in general position for each t, except $t = 1$.*

- *If $\Delta_1 \cap \Delta' = \emptyset$ then pairs (Δ_0, Δ') and (Δ_2, Δ') are either both linked or both unlinked.*
- *If Δ_1 and Δ' intersect in exactly one point, that is not a vertex of these triangles, then exactly one pair of (Δ_0, Δ') and (Δ_2, Δ') is linked.*

4.7 and **4.8.** See [Zi].

4.9. Denote by a *spherical segment* an intersection of a 3-dimentional sphere and 2-dimentional angle with vertex at the center of the sphere in $\mathbb{R}^4$.

A spherical analogue of the Conway-Gordon-Sachs Theorem is the following statement.

For each 6 points in general position in the three-dimensional sphere there exist two linked closed trinangle broken line with vertices at these points.

Two closed 4-segment broken lines $ABCD$ and $A'B'C'D'$ in the three-dimensional sphere are called *linked* if the number of intersections of a broken line $ABCD$ with the union of two-dimensional spherical triangles $A'B'C'$ and $A'D'C'$ is odd.

A spherical analog of the Sachs Theorem is the following statement.

Suppose we are given 4 red points and 4 blue points in S^3 such that any two spherical segments with endpoints of different colors have no common interior points. Then there are two linked closed quadrangular broken lines with vertices at these points each edge of which has endpoints of different colors.

The proofs of these Theorems are absolutely similar to their analogs.

4.10, 4.12. See [Zu].

4.11. (a) For each triangle of (5, 3)-realization N, not containing vertex $A_{1,1}$, there exists a ball with center at $A_{1,1}$ that does not intersect this triangle (because there exist a point of this triangle for which the function 'distance to $A_{1,1}$' reaches its minimum and this minimum is not equal to zero). Since the number of triangles of N is finite, there exists a ball with center at $A_{1,1}$ that does not intersect any triangle of N, that does not contain $A_{1,1}$. Consider the bounding sphere of this ball. It intersects only those segments and triangles of N, which contain vertex $A_{1,1}$. They are segments $A_{1,1}A_{1,a}$, $A_{1,1}A_{b,1}$ and triangles $A_{1,1}A_{a,1}A_{b,1}$ for each $2 \le a \le 5$ and each $2 \le b \le 3$. The intersection of these segments and triangles with a sphere is

- four points corresponding to segments $A_{1,1}A_{1,a}$,
- two points corresponding to segments $A_{1,1}A_{b,1}$,
- eight spherical segments connecting points of the first set to points if the second set.

This is a graph $K_{4,2}$, linearly embedded to a sphere.

(b) See the second paragraph of the solution of Problem 1.13.

4.13. (a) By i' denote $Oi \cap S^3$. By the Problem 1.16.b the intersection of a two-dimensional triangle Oij and sphere S^3 is a spherical segment $i'j'$. Thus the intersection of S^3 with the union of triangles $Oij, 1 \le i < j \le 6$ is a graph K_6. It is an embedding since $0, 1, \dots, 6$ are in general position.

(b) By the statement, the triangles with different vertices do not intersect. Since points are in general position, we have that the triangles with one or two common vertices intersect in the corresponding vertex or side.

4.14. *Statement.* Given two (3,3)-realizations $\mathbb{R}^4$ such that their intersection is exactly one point, their common '(1,1)-vertex'. Then each sufficiently small sphere S^3 with center at this vertex intersects these (3,3)-realizations in a pair of spherical polygons that are unlinked mod 2 in S^3.

4.15. (a) Analogous to Problem 4.11.a.

(b) The length of both cycles is 4. Without loss of generality vertices of the first cycle correspong to points $A_{1,2}$, $A_{2,1}$, $A_{1,3}, A_{3,1}$, and the vertices of the second — to points $A_{1,4}$, $A_{4,1}$, $A_{1,5}$ and $A_{5,1}$. Then the first of the required two $(3,3)$-realizations is a torus 123×123. And the second of the required two $(3,3)$-realizations is a torus 145×145.

References

[BE82] *V. G. Boltyansky and V. A. Efremovich.* Visual topology, Moscow: Nauka, 1982.

[Pr06] *V. V. Prasolov.* Elements of gomology theory. Moscow: MCCME, 2006. See. `http://www.mccme.ru/prasolov/`.

[PS05] *V. V. Prasolov and M. Skopenkov.* Ramsey link theory // Math. Prosveschenie. 2005. 9. Pages. 108-115.

[Sk03] *M. Skopenkov.* Embedding products of graphs into Euclidean spaces // Fund. Math. 2003. 179. P. 191-198.

[Sk08] *A. Skopenkov.* Embedding and knotting of manifolds in Euclidean spaces // London Math. Soc. Lect. Notes, 347 (2008) 248–342; http://arxiv.org/abs/0604045.

[Zi] *A. Zimin.* A short proof of the Conway-Gordon-Sachs and Sachs Theorems, unpublished.

[Zu] *J. Zung.* A non-general-position Parity Theorem, unpublished.

6 Additional material: intuitively visual embedding problems

This section is independent of the other ones.

The requirements concerning mathematical rigor in this section are less strict as in the other ones.

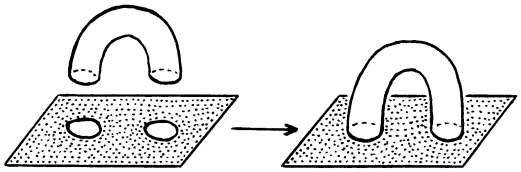


Figure 10: Gluing of a handle

A *sphere with g handles* is the surface obtained by making $2g$ circular holes in the sphere and gluing g copies of the lateral surface of the cylinder along their boundary circles to the boundaries of the holes (Fig. 10 and 2,right for $g = 3$).

Consider n rectangles XYB_kA_k, $k = 1, 2, \ldots, n$ in the three-dimensional space such that any two of them intersect exactly at the segment XY. An *n-page book* is the union of such rectangles as it is shown in Fig. 11 for $n = 3$, left.

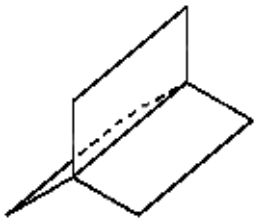


Figure 11: A 3-page book

6.1. Any graph can be drawn without self intersections
(a) in 3-space;
(b) on a sphere with a certain number of handles depending on the graph;
(c) on a book with a certain number of pages depending on the graph;
(d) on the book with three pages. (Fig. 11).

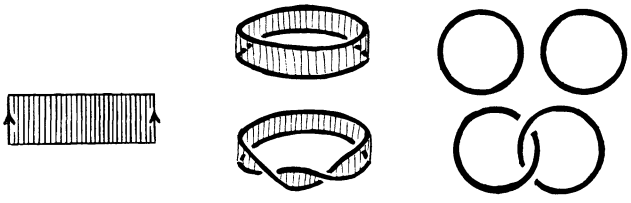


Figure 12: An annulus

An *annulus* is the surface obtained by gluing the short opposite sides of a long rectangular strip in "the same direction", i.e., without twisting (Fig. 12).

6.2. (a) Can one cut the Möbius strip so as to obtain a cylinder?
(b) Can one cut a cylinder and a Möbius strip out of the Möbius strip?
(c) Can one cut the Möbius strip so as to obtain a cylinder and a Möbius strip?

The *Klein bottle* is the figure obtained by gluing one pair of opposite sides of the square "in the same direction" and the other pair "in opposite directions", (Fig. 13).

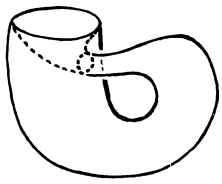


Figure 13: Klein bottle: gluing of a rectangular strip (!) and image in $\mathbb{R}^3$

6.3. (a) Cut the Klein bottle (Fig. 13)into two Möbius strips;

(b) Cut the Klein bottle so as to obtain one Möbius strip.

6.4. From the 3-page book (Fig. 11) cut out

(a) a Möbius strip;

(b) a torus with a hole.

6.5. For any 6 points $0, 1, 2, 3, 4, 5 \in \mathbb{R}^3$, if the set of triangles

(a) $0jk$, $1 \le j < k \le 5$, $k \ne 2$, is embedded, then the triangles 012 and 345 are linked;

(b) $0jk$, $1 \le j < k \le 5$, $(j,k) \notin \{(1,2),(1,3)\}$, is embedded, then either the triangles 012 and 345, or the triangles 013 and 245, are linked;

(c) $0jk$, $1 \le j < k \le 5$, $(j,k) \notin \{(1,2),(1,3),(1,4)\}$, is embedded, then either the triangles 012 and 345, or the triangles 013 and 245, or the triangles 014 and 235 are linked.

Solution of Problem 6.1. (a) Draw this graph (possibly, with self-intersection) on the plane such that the edges are not self-intersecting. If there are points (except vertices) that belong to more than two edges, then move some edges so that only points of intersection of two edges remain. For each two intersecting edges lets raise one on them up, so that the intersection dissapear.

Or see solution of Problem 1.7.

(b) Use the idea from the solution of (a).

(d) We may assume that all the intersection points are 'good' and are in one line. Let us glue the third page along this line. Then for each intersection point let us raise one of the edges up into the third page like a 'bridge'. So we can delete all intersection points.

A non-general-position intersection parity theorem

Jonathan Zung (jonathanzung@gmail.com)

August 11, 2013

1 Introduction

Two piecewise-linear closed surfaces in $\mathbb{R}^4$ generically intersect in an even number of points. When the two surfaces are not in general position but share a common vertex x, we will show that this theorem may be corrected by adding the linking number between the restrictions of the two surfaces to a small sphere around x. The main technique is to decompose our surfaces into tetrahedra, whose pairwise intersections may be studied in $\mathbb{R}^2$. Analogous results hold in the case of a closed surface and a closed curve in $\mathbb{R}^3$.

For a triangle $\mathcal{T} = ABC$, we define $\blacktriangle\mathcal{T}$ to be the 2-dimensional triangle with these vertices. For a closed broken line $\mathcal{T} = A_0A_1 \dots A_n$ in general position in $\mathbb{R}^3$, define $\blacktriangle\mathcal{T} = \blacktriangle A_0A_1A_2 + \blacktriangle A_0A_2A_3 + \dots + \blacktriangle A_0A_{n-1}A_n$, where $+$ denotes the symmetric difference of sets. We make the same definition for spherical polygons.

For two closed broken lines $\mathcal{T}_1$ and $\mathcal{T}_2$ in general position in $\mathbb{R}^3$ or S^3, we define their linking number by $link(\mathcal{T}_1, \mathcal{T}_2) = |\blacktriangle\mathcal{T}_1 \cap \mathcal{T}_2| \mod 2$. A priori, the linking number may depend on the order of vertices we chose in order to construct $\blacktriangle\mathcal{T}_1$.

(Main result) Parity theorem for 2 closed surfaces in $\mathbb{R}^4$. *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be the two piecewise-linear, closed surfaces in $\mathbb{R}^4$ sharing a vertex x. Assume that their vertices are in general position. Further, let $\widetilde{\mathcal{T}_1}$ and $\widetilde{\mathcal{T}_2}$ be the intersections of $\mathcal{T}_1$ and $\mathcal{T}_2$ with a small sphere around x. Then $|\mathcal{T}_1 \cap \mathcal{T}_2| + link(\widetilde{\mathcal{T}_1}, \widetilde{\mathcal{T}_2})$ is even.*

2 Proofs

Fix a point x in $\mathbb{R}^4$. Let S be a small 3-sphere around x. For any object T, let $\widetilde{T}$ denote the intersection of T with S.

Lemma 1 (easy case)**.** *Any two tetrahedra T_1 and T_2 in general position in $\mathbb{R}^4$ intersect in an even number of points.*

Proof. The hyperplanes containing T_1 and T_2 respectively intersect in a 2-dimensional plane π. Let T_1 and T_2 intersect π in the convex polygons P_1 and P_2 respectively. By our hypothesis of general position, the vertices of P_1 and P_2 are in general position in the plane. Therefore, $P_1 \cap interior(P_2)$ in a union of broken paths or a cycle. The intersections between P_1 and P_2 are precisely the endpoints of broken paths in this set, of which there are an even number as desired. □

Lemma 2 (harder case)**.** *Consider two tetrahedra T_1 and T_2 in $\mathbb{R}^4$ sharing a common vertex x. Assume that the seven vertices are in general position. Let a small 3-sphere S around x intersect T_1 and T_2 in two spherical triangles $\widetilde{T}_1$ and $\widetilde{T}_2$. Then $link(\widetilde{T}_1, \widetilde{T}_2) + |T_1 \cap T_2 \setminus x|$ is even.*

Proof. As in lemma 1, let P_1 and P_2 be the intersections of T_1 and T_2 with their common 2-dimensional plane π. Again, $P_1 \cap interior(P_2)$ is a union of broken paths or a cycle. The intersections between P_1 and P_2 are precisely the endpoints of broken paths in this set, with the possible exception of x. Let S intersect π in the circle s. In other words, s is the common circle of $\widetilde{T}_1$ and $\widetilde{T}_2$. Observe that $\widetilde{T}_1$ and $\widetilde{T}_2$ are linked if and only if they intersect s in an alternating pattern. This is true if and only if x is an endpoint of a broken path in $P_1 \cap interior(P_2)$, which yields our theorem. □

Proof of the main result.

Proof. We may assume that $\mathcal{T}_1$ and $\mathcal{T}_2$ have only triangular faces. Let $T^1 = \{T_1^1, T_2^1, \ldots, T_{n_1}^1\}$ be a triangulation of a (possibly singular) Seifert surface with boundary $\mathcal{T}_1$. By this we mean a set of tetrahedra in $\mathbb{R}^4$ such that each face of $\mathcal{T}_1$ appears once among the faces of tetrahedra in T^1, and every other triangle appears 0 or 2 times in among these faces. One way to do this is to choose an arbitrary point v in general position and let T^1 consist of tetrahedra with base a face of $\mathcal{T}_1$ and last vertex v. Similarly, give $\mathcal{T}_2$ a triangulation $T^2 = \{T_1^2, T_2^2, \ldots, T_{n_2}^2\}$.

Now we claim that

$$|\mathcal{T}_1 \cap \mathcal{T}_2 \setminus x| \equiv \sum_{i,j} |T_i^1 \cap T_j^2 \setminus x| \qquad \text{mod } 2$$

This is because intersections counted on the right side which lie on a face not in $\mathcal{T}_1$ or $\mathcal{T}_2$ are counted an even number of times due to the triangulation condition. Conversely, intersections between two faces of $\mathcal{T}_1$ and $\mathcal{T}_2$ are counted on the right side precisely once.

Now observe that, triangulations T^1 and T^2 descend to triangulations $\widetilde{T^1}$ and $\widetilde{T^2}$ of $\widetilde{\mathcal{T}_1}$ and $\widetilde{\mathcal{T}_2}$ respectively. Again by the triangulation condition, we have

$$link(\widetilde{\mathcal{T}_1}, \widetilde{\mathcal{T}_2}) \equiv |\blacktriangle \widetilde{\mathcal{T}_1} \cap \widetilde{\mathcal{T}_2}| \equiv \sum_{i,j} |\blacktriangle \widetilde{T_i^1} \cap \widetilde{T_j^2}| \equiv \sum_{i,j} link(\widetilde{T_i^1}, \widetilde{T_j^2}) \quad \text{mod } 2$$

Combining our two equations, we get

$$\begin{aligned} |\mathcal{T}_1 \cap \mathcal{T}_2 \setminus x| + link(\widetilde{\mathcal{T}_1}, \widetilde{\mathcal{T}_2}) &\equiv \sum_{i,j} \left(|T_i^1 \cap T_j^2 \setminus x| + link(\widetilde{T_i^1}, \widetilde{T_j^2}) \right) && \text{mod } 2 \\ &\equiv 0 && \text{mod } 2 \end{aligned}$$

The last congruence holds by the two lemmas. Therefore, the left side vanishes, as desired. □

These results generalize in a straightforward manner to the case where our surfaces share several common vertices but no common edges. Here, we should simply sum the linking numbers at each common vertex.

We can also generalize to the case of the intersection between a k-dimensional and an l-dimensional closed surfaces intersecting in $\mathbb{R}^{k+l}$. Take as an example the case of a 2-dimensional surface $\mathcal{T}_1$ and a closed curve $\mathcal{T}_2$ intersecting in $\mathbb{R}^3$. The definition of the linking number remains $|\blacktriangle \widetilde{\mathcal{T}_1} \cap \widetilde{\mathcal{T}_2}| \mod 2$, where $\widetilde{\mathcal{T}_2}$ is now simply 2 points. T_1 and T_2 still share a common plane π, and the proofs of the parity theorems may proceed as above.

A short proof of the Conway-Gordon-Sachs and Sachs Theorems

Arseny Zimin

Abstract

In this paper we present a short and apparently new proof of the Conway-Gordon-Sachs Theorem about the complete graph at 6 vertices embedded to $\mathbb{R}^3$ and the the Sachs Theorem about the the complete biparted graph at 8 vertices. We reduce this theorems to certain property of the complete graph at 5 vertices and the complete biparted graph at 6 vertices maped to a sphere or a plane.

Two triangles in the 3-dimensional space whose six vertices are in general position are *linked* if the outline of the first triangle intersects the interior of the second triangle exactly at one point.

Points in 3-dimensional space are in *general position* if no four of them are in one plane.

Rectilinear Conway-Gordon-Sachs Theorem. *Assume that six points in the* 3*-dimensional space are in general position. Then there exist two linked triangles with vertices at these points.*

Define a 2*-dimensional complex* as a set of triangles, segments and points in $\mathbb{R}^3$ that satisfies the following conditions:

- sides of any triangle from the complex are in the complex
- endpoints of any segment from the complex are in the complex

Two closed broken lines a and b without self-intersections in the 3-dimensional space are *linked* if there exist a 2-dimensional complex, denote it by A, embedded to $\mathbb{R}^3$, with a boundary a such that the number of intersection points of A with b is odd and vertices of the broken line b and the complex $A - a$ are in general position.

Denote by K_n the complete graph at n vertices. Denote by $K_{n,n}$ the complete biparted graph at $2n$ vertices.

Conway-Gordon-Sachs Theorem. *Assume that the graph K_6 is piecewise-linear embedded in the* 3*-dimensional space. Then in this graph exist two linked* 3*-length cycles.*

Remark. The statement of the theorem is meaningful because any 3-lenth cycle in this graph is a closed broken line.

Sachs Theorem. *Assume that the graph $K_{4,4}$ is piecewise-linear embedded in the* 3*-dimensional space. Then there exist two linked* 4*-length cycles in this grpah.*

Proof of the rectilinear Conway-Gordon-Sachs Theorem.

Let a, b be segments in the 3-dimensional space, S^2 be a sphere whose center is denoted by O. Let $f : \mathbb{R}^3 - \{O\} \to S^2$ be the central projection with the center O. A segment a is *higher* than a segment b, if

- $|f(a) \cap f(b)| = 1$, and
- O is closer to $f^{-1}(f(b)) \cap a$ than to $f^{-1}(f(a)) \cap b$.

Remark. The set $f^{-1}(f(b))$ is a 2-dimensional angle with the vertex O and sides joining O with the endpoints of b.

Lemma 1. *Assume that vertices of two triangles are in general position. Denote by $A_1A_2A_3$ the first triangle. Denote by S^2 a sphere with the center A_1 and radius so small that all the vertices of the triangles except A_1 are outside S^2. If the number of the sides of the second triangle that are lower than A_2A_3 is odd then these two triangles are linked.*

Remark. The condition that the vertices of the triangles except A_1 are outside the sphere could be avoided at the price of some complications both in the statement and the proof.

Proof of Lemma 1.

Denote by A_4, A_5, A_6 the vertices of the second triangle. Let $f : \mathbb{R}^3 - \{A_1\} \to S^2$ be the central projection with the center A_1. By the assertion of the lemma there exists a side, say A_4A_5, of triangle $A_4A_5A_6$ such that A_2A_3 is higher than A_4A_5. Then the point $f^{-1}(f(A_2A_3)) \cap A_4A_5$ is inside the 2-dimensional triangle $A_1A_2A_3$. Since $f(A_2A_3)$ is a ark of a circle on S^2 and $f(A_4A_5A_6)$ is a spherical triangle on S^2, $f(A_2A_3)$ intersects the projection of the outline of the triangle $A_4A_5A_6$ at most at 2 points. So there is a unique side A_4A_5 of the triangle $A_4A_5A_6$ that is lower than A_2A_3. Since the vertices of these two triangles are in general position the outlines of triangles $A_1A_2A_3$ and $A_4A_5A_6$ do not intersect. This implies that the outline of the triangle $A_4A_5A_6$ intersects the interior of the triangle $A_1A_2A_3$ at a unique point $f^{-1}(f(A_2A_3)) \cap A_4A_5$. So these two triangles are linked. QED

Continuation of the proof. Suppose that points $A_1, A_2, A_3, A_4, A_5, A_6$ are in general position in the 3-dimensional space. Consider the complete graph K_5 whose vertices are points A_2, A_3, A_4, A_5, A_6 and edges are segments joining pairs of these points. Consider a sphere S^2 with center A_1. Let this sphere be enough small to make points A_2, A_3, A_4, A_5, A_6 be outside the sphere. Consider the central projection $f : \mathbb{R}^3 - A_1 \to S^2$ with the center A_1. For ordered pair (e, e') of $e, e' \in K_5$ denote

$$e \circ e' := \begin{cases} 1, & \text{if } e \text{ is higher than } e' \\ 0, & \text{otherwise} \end{cases}.$$

For any edge $e \in K_5$ define its *linking number*

$$S_e := \sum_{e' \in (K_5 - e)} e \circ e'$$

Then

$$\sum_{e \in K_5} S_e \equiv \sum_{(e,e'), e, e' \in K_5} e \circ e' \equiv$$

$$\equiv \sum \{|f(e) \cap f(e')| : \{e, e'\} \text{ is a non-ordered pair of nonadjacent edges of } K_5\} \equiv 1 \mod 2.$$

Hence the linking number of some edge, say A_2A_3, is odd. Then Lemma implies that triangles $A_1A_2A_3$ and $A_4A_5A_6$ are linked.

The first equality follows from definition of S_e. The second equality holds because

- for any two edges $e, e' \in K_5$ $|f(e) \cap f(e')| \leq 1$ because vertices of K_5 are in general position
- if edges $e, e' \in K_5$ are nonadjacent and $f(e) \cap f(e') \neq \emptyset$ then $e \circ e' + e' \circ e = 1$
- if edges $e, e' \in K_5$ are adjacent or $f(e) \cap f(e') = \emptyset$ then $e \circ e' + e' \circ e = 0$.

The third equality follows from Lemma 2.

Lemma 2. *For any general position linear map $f : K_5 \to S^2$ the number of self-intersections of $f(K_5)$ is odd.*

This Lemma is known, see e.g. [Sk, §1].QED

Proof of the Conway-Gordon-Sachs Theorem.

Consider a general position plane. Define what means that a segment a is *higher* than a segment b analogous to the definition in the linear case but replacing a 'sphere' with the 'general position plane' and the 'central projection' to the 'orthogonal projection to this plane'.

Lemma 3. *Consider two closed broken lines, denote them by A, B. Consider a general position plane. Assume that the number of ordered pairs (a, b) of sides $a, \in A$, $b \in B$ such that a is higher than b is odd. Then these two broken lines are linked.*

This Lemma is known, see [?].

Consider a general position plane π. Consider orthogonal projection $f : \mathbb{R}^3 \to \pi$.

For any ordered pair of broken lines (A, B) in the 3-dimensional space denote

$$A \circ B := \begin{cases} 1, & \text{if the number ordered pairs } (a,b) \text{ of sides } a \in A,\ b \in B \text{ such that } a \text{ is higher than } b \text{ is odd} \\ 0, & \text{otherwise} \end{cases}.$$

Denote by a one of the vertices of graph K_6. Denote by C_{ij} the cycle of edges of the graph $K_6 - \{a\}$ that does not contain the edge ij. Then the problem follows from

$$\sum_{bc \in K_6 - \{a\}} abc \circ C_{bc} \equiv \sum_{bc \in K_6 - \{a\}} (ab \circ C_{bc} + ac \circ C_{bc}) + \sum_{bc \in K_6 - \{a\}} bc \circ C_{bc} \equiv \sum_{bc \in K_6 - \{a\}} bc \circ C_{bc} \equiv$$

$$\equiv \sum \{|f(e) \cap f(e')| : \{e, e'\} \text{ is a non-ordered pair of nonadjacent edges of } K_6 - \{a\}\} \equiv 1 \mod 2,$$

Hence for some two cycles abc, C_{bc} of graph K_6 the number $abc \circ C_{bc}$ is equal to 1 and Lemma 2 implies that these cycles are linked. QED

Proof of the second equality.

Note that $ab \circ C_{bc} = \sum\limits_{e \in C_{bc}} ab \circ e$

For each $i \in K_6 - \{a\}$ and for each edge $e \in (K_6 - \{a\})$ there exist exactly two 3-length cycles in $K_6 - \{a\}$ containing this edge. So for each edge $ij \in K_6 - \{a, b\}$ the number $ab \circ ij$ appears twice in the sum $\sum\limits_{bc \in K_6 - \{a\}} (ab \circ C_{bc} + ac \circ C_{bc})$. Analogous for each edge $ij \in K_6 - \{a, c\}$ the number $ac \circ ij$ appears twice in this sum. Then this sum is even. QED

The proof of the third equality is the same to the *proof of the second equality* in the linear case.

The last equality follows from Lemma 4.

Lemma 4. *For any general position piecewise-linear map $f : K_5 \to \pi$ the number of self-intersections of $f(K_5)$ is odd.*

This lemma is the generalization of Lemma 2, see [Sk, §1].

Proof of the Sachs Theorem. Consider a general position plane π. Consider the orthogonal projection $f : \mathbb{R}^3 \to \pi$. Denote by a, b two vertices of graph $K_{4,4}$ from different parts. Denote by C_{ij} the cycle of edges of the graph $K_{4,4} - \{a, b\}$ nonadjacent to edge $ij \in K_{4,4} - \{a, b\}$. Denote by $xyzt$ a 4-length cycle of edges $xy, yz, zt, tx \in K_{4,4}$

$$\sum_{ij \in K_{4,4} - \{a,b\}} abij \circ C_{ij} \equiv \sum_{ij \in K_{4,4} - \{a,b\}} ab \circ C_{ij} + \sum_{ij \in K_{4,4} - \{a,b\}} (aj \circ C_{ij} + bi \circ C_{ij}) \equiv$$

$$\equiv \sum \{|f(e) \cap f(e')| : \{e, e'\} \text{ is a non-ordered pair of nonadjacent edges of } K_{4,4} - \{a, b\}\} \equiv 1 \mod 2$$

The second equality holds because for each $i \in K_{4,4} - \{a, b\}$ and for each edge of $K_{4,4} - \{a, b, i\}$ there exist four 4-length cycles containing this edge. So for each edge $kl \in K_{4,4} - \{a, b, j\}$ the number $aj \circ kl$ appears four times in the sum $\sum\limits_{ij \in K_{4,4} - \{a,b\}} (aj \circ C_{ij} + bi \circ C_{ij})$. And analogous for each edge $kl \in K_{4,4} - \{a, b, i\}$ the number $bi \circ kl$ appears four times in this sum. Hence this sum is even.

The proof of the second equality is the same to the *proof of the second equality* in the proof of the rectilinear Conway-Gordon-Sachs Theorem.

The last equality follows from Lemma 5.

Lemma 5. *For any general position piecewise-linear map $f : K_{3,3} \to \pi$ the number of self-intersections of $f(K_{3,3})$ is odd.*

This lemma is known, see [**?**].

The author is grateful to Arqady Skopenkov for productive discussions. The author is also grateful to Mikhail Skopenkov for reading this paper.

References

[Sk] A. Skopenkov, Algorithms for recognition of the realizability of hypergraphs, in Russian, http://www.mccme.ru/circles/oim/algor.pdf

[BE] V. Boltyansky, V. Efremovich, Visual topology, in Russian, http://ilib.mccme.ru/djvu/geometry/bo nagl-topo.htm

REALIZABILITY OF HYPERGRAPHS AND RAMSEY LINK THEORY

A. SKOPENKOV

Abstract. We present short simple proofs of Conway-Gordon-Sachs' theorem on graphs in 3-dimensional space, as well as van Kampen-Flores' and Ummel's theorems on nonrealizability of certain hypergraphs (or simplicial complexes) in 4-dimensional space. The proofs use a reduction to lower dimensions which allows to exhibit relation between these results.

We present a simplified exposition accessible to non-specialists in the area and to students who know basic geometry of 3-dimensional space and who are ready to learn straightforward 4-dimensional generalizations. We use elementary language (e.g. collections of points) which allows to present the main ideas without technicalities (e.g. without using the formal definition of a hypergraph).

Contents

1. Introduction 2
1.1. Impossible constructions and intrinsic linking 2
1.2. Realizability of hypergraphs 3
1.3. Linking and intersection in higher dimensions 6
1.4. Cartesian product and the Menger conjecture 6
1.5. Linear, piecewise-linear (PL) and topological versions 6
1.6. Comparison with other expositions 7
1.7. Further generalizations 7
2. Proofs and further results 7
2.1. Intersection in the plane 7
2.2. Linking and intersection in three-dimensional space 9
2.3. Linking and intersection in four-dimensional space 10
3. Realizability of products and the Menger conjecture 12
3.1. Realizability of products 12
3.2. Realizability of products in three-dimensional space 13
3.3. Parity Lemmas 15
3.4. Realizability of products in four-dimensional space 16
References 18

Homepage: www.mccme.ru/~skopenko.

Supported in part by RFBR Grant No. 15-01-06302, the D. Zimin Dynasty foundation, and the Simons-IUM fellowships.

This text is based on the author's lectures at Moscow Institute of Physics and Technology, Independent University of Moscow, Institute of Science and Technology (Austria), summer school 'Modern Mathematics', Summer Conference of Tournament of Towns, Kirov region summer school, Moscow 'olympic' school, math circle 'Olympiads and Mathematics'. I am grateful to M. Skopenkov for joint writing an earlier version of this text, to A. Zimin for notes on Example 3.4.ab, and to G. Chelnokov, I. Izmestiev, R. Karasev, A. Matushkin, A. Rukhovich, A. Shapovalov, M. Skopenkov, A. Sossinsky, S. Tabachnikov, O. Viro, A. Zimin, J. Zung for useful discussions.

'It's too difficult.'
'Write simply.'
'That's hardest of all.'
I. Murdoch, The Message to the Planet.

1. Introduction

1.1. **Impossible constructions and intrinsic linking.** 'Impossible constructions' like the impossible cube, the Penrose triangle, the blivet etc (see Figure 1 and [Io]) are well-known, mainly due to pictures by M.C. Escher, see also [Br26, CKS+, GSS+]. The pictures do not allow the global spatial interpretation because of collision between local spatial interpretations to each other. In geometry, topology and graph theory there are also famous basic examples of 'impossible constructions' (of which local parts are 'possible').

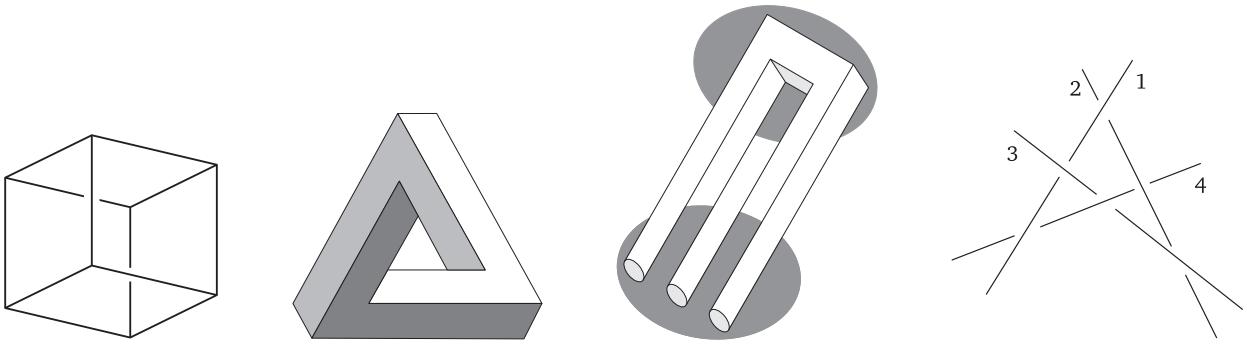


Figure 1. The impossible cube, the Penrose triangle, the blivet, an impossible projection

In this paper we exhibit a striking relation of 'impossible constructions' in four-dimensional space to 'intrinsic linking' results in three-dimensional space. Such a relation was found by M. Skopenkov in [Sk03] and used there to obtain a short proof of the Menger 1929 conjecture and its generalizations, see Remark 1.5 and §1.4. Let us give a beautiful example of 'intrinsic linking'.

We abbreviate 'three-dimensional space $\mathbb{R}^3$' to '3-space'. Analogous meaning has '4-space'.

By a *triangle* we mean 'the interior' of a triangle (more accurately, the convex hull[1] of three points).

Take two triangles in 3-space no 4 of whose 6 vertices lie in the same plane. The triangles are called **linked**, if the outline of the first triangle intersects the second triangle exactly at one point. It is not obvious from the definition that the property of being linked is symmetric. For a proof see e.g. [Sk, Symmetry Lemma 4.2].

E.g. the triangles $A_1A_3A_5$ and $A_2A_4A_6$ in Figure 2 are linked. (The distance from the point A_j to the projection plane equals j, see Figure 2, left. So the projection in Figure 2, right, is realizable, as opposed to Figure 1, right.)

Theorem 1.1 (Linear Conway–Gordon–Sachs Theorem; [Sa81, CG83])**.** *If no 4 of 6 points in 3-space lie in the same plane, then there are two linked triangles with vertices at these 6 points.*

[1]A subset of the plane or of $\mathbb{R}^d$ is called *convex*, if for any two points from this subset the segment joining these two points is in this subset. The *convex hull* of a subset X of the plane or $\mathbb{R}^d$ is the minimal convex set that contains X.

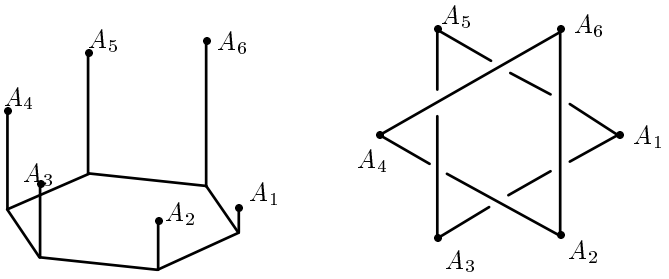


FIGURE 2. Linked triangles

Moreover, the number of linked unordered pairs of triangles with vertices at these 6 points is odd.

See idea of a short proof in Remark 1.5. Formally, Theorem 1.1 is reduced to Proposition 1.2 below in §2.2. See more results on linking in 3-space in §2.2 and in [Sk, §4.1 'Linking of triangles in three-dimensional space'].

1.2. **Realizability of hypergraphs.** Another example of an 'impossible construction' is that one cannot construct 3 houses and 3 wells in the plane and join each house to each well by a path so that paths intersect only at their starting points or endpoints.[2]

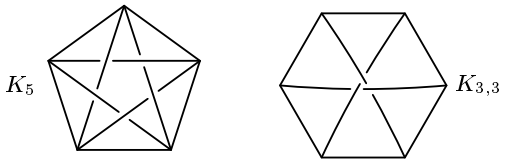


FIGURE 3. Nonplanar graphs K_5 and $K_{3,3}$

Proposition 1.2 (see proof in §2.1)**.** *From any* 5 *points in the plane one can choose two disjoint pairs such that the segment joining the first pair intersects the segment joining the second pair.*[3]

Moreover, if no 3 *of* 5 *points in the plane lie in the same line, then the number of intersection points of interiors of segments joining the* 5 *points is odd.* [4]

In this paper we present a natural interesting generalization: beautiful and nontrivial examples of *two-dimensional* analogues of graphs non-realizable in three- and four-dimensional space.

Remark 1.3 (why this expository paper might be interesting)**.** We present a simplified exposition accessible to non-specialists in the area, see also the second paragraph of §1.6. We state the examples in terms of certain systems of points, see Theorem 1.4 below. So we

[2]In graph-theoretic terms this means that the complete bipartite graph $K_{3,3}$ is not planar, see Figure 3, right.

[3]This is a 'linear' version of the nonplanarity of the complete graph K_5 on 5 vertices, see Figure 3, left.

[4]The first sentence of Proposition 1.2 indeed follows by the 'moreover' part. This is true because for non-general-position points the first sentence is obvious: if points A, B, C among given 5 points lie in the same line, B between A and C, and D is any other given point, then segments AC and BD intersect. This is also true because we can make a small shift so that no 3 of 5 shifted points lie in the same line, and no intersection points of segments with disjoint vertices are added. Analogous remarks can be made for Theorems 1.1, 1.4 below; such remarks are omitted.

do not use the notions of a hypergraph and its realizability neither for the statements nor for the proofs. (We do mention hypergraphs because the problem of their realizability helps to understand *the motivation* of the results.) For understanding most of the paper it suffices to know basic geometry of 3-dimensional space and to be ready to learn straightforward 4-dimensional generalizations. We believe that describing simple applications of topological methods in elementary language makes these methods more accessible (although this is called 'detopologization' in [MTW12, §1]).

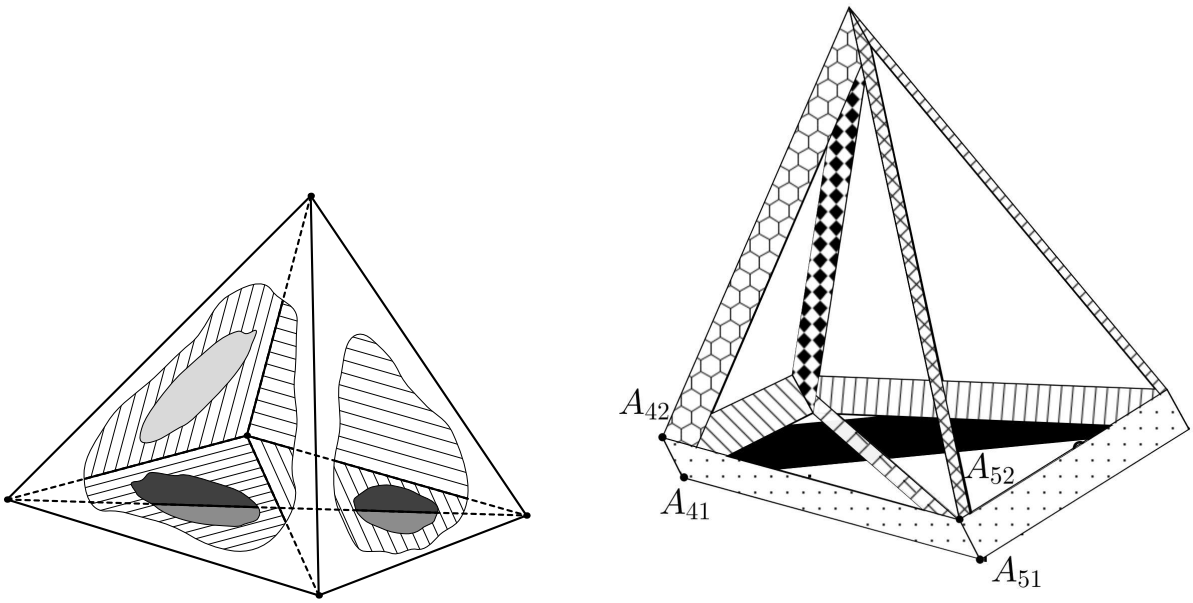


FIGURE 4. Left: Realization in $\mathbb{R}^3$ of the complete 3-homogeneous hypergraph on 5 vertices.
Right: Realization in $\mathbb{R}^3$ of the *product* of the complete graphs on 5 and on 2 vertices.

Such analogues are *3-homogeneous*, or *2-dimensional hypergraphs* defined as collections of 3-element subsets of a finite set.[5] For brevity, we omit '3-homogeneous, or '2-dimensional'. For instance, *a complete hypergraph* on k vertices is the collection of all 3-element subsets of a k-element set. *Realizability* of a hypergraph in *d-dimensional Euclidean space* $\mathbb{R}^d$ is defined similarly to the realizability of a graph in the plane (one 'draws' a triangle for every three-element subset; see Figures 4 and 5).[6] Hypergraphs (and simplicial complexes) play an important role in mathematics. One cannot imagine topology and combinatorics without them. They are also used in computer science and bioinformatics, see, e.g.[PS11].

A 'small shift' (or 'general position') argument shows that every graph is realizable in $\mathbb{R}^3$. A straightforward generalization shows that every hypergraph is realizable in $\mathbb{R}^5$.

It is easy to see that the complete hypergraph on 6 vertices is non-realizable in $\mathbb{R}^3$ (Proposition 2.4.a). Already in the early history of topology (1920s) mathematicians tried to construct hypergraphs non-realizable in $\mathbb{R}^4$. Egbert van Kampen and A. Flores in 1932-34

[5]In topology such objects are called *pure, or dimensionally homogeneous, 2-dimensional simplicial complexes*, but I hope the term hypergraph is more convenient to generic mathematician or computer scientist.

[6]Here is a rigorous definition. A hypergraph $(V, F \subset \binom{V}{3})$ is *linear realizable* in $\mathbb{R}^d$ if there is a set of non-degenerate triangles in $\mathbb{R}^d$ whose vertices correspond to V, whose triangles correspond to F, and every two triangles either are disjoint, or intersect only at a common vertex, or intersect only by a common side.

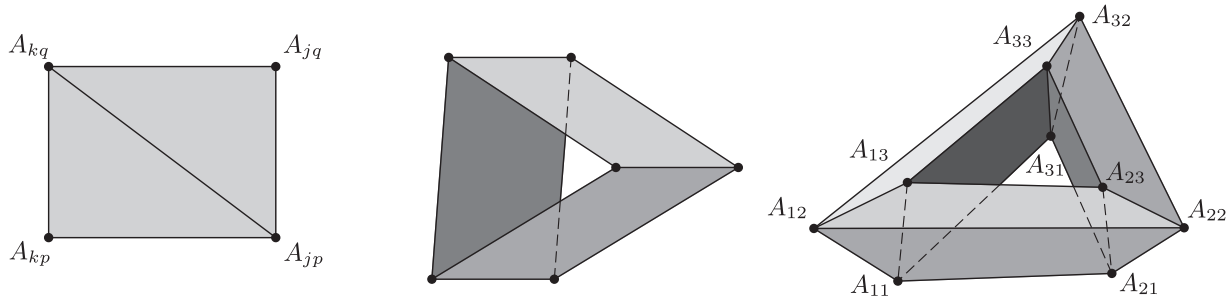


FIGURE 5. Left: Realization in $\mathbb{R}^2$ of the *square* of the complete graph on 2 vertices.
Middle: Realization in $\mathbb{R}^3$ of the *product* of the complete graphs on 2 and on 3 vertices.
Right: Realization in $\mathbb{R}^3$ of the *square* of the complete graph on 3 vertices.

proved that the complete hypergraph on 7 vertices is not realizable in $\mathbb{R}^4$ (Theorem 1.4). It is both an early application of *combinatorial topology* (nowadays called algebraic topology) and one of the first results of *topological combinatorics* (also an area of ongoing active research).

Before stating Theorem 1.4 observe that 'typical' intersection of two segments in the plane is either empty set or a point. Analogously, 'typical' intersection of two triangles in 4-space is either empty set or a point. More intuition on 4-space can be developed by reading e.g. [Sk, §4.7 'How to work with four-dimensional space?'], see also Remark 1.5 below.

Theorem 1.4 (Linear Van Kampen-Flores Theorem; [vK32, Fl34])**.** *From any 7 points in 4-space one can choose two disjoint triples such that the two triangles with vertices at the triples intersect.*

Moreover, if no 5 of 7 points in 4-space lie in the same 3-dimensional hyperplane, then the number of intersection points of triangles with vertices at these points is odd.

See idea of a short proof in Remark 1.5. Formally, Theorem 1.4 is reduced to Theorem 1.1 in §2.3.

An analogue of Theorem 1.4

• is true for 5 points in the plane or 6 points in 3-space (Propositions 1.2 and 2.4.b);

• is false for 4 points in the plane, 5 points in 3-space or 6 points in 4-space (in $\mathbb{R}^n$ take the $n+1$ vertices and an interior point of an n-simplex, see Figure 4, left).

Remark 1.5 (lowering of dimension)**.** A striking idea is that the nonrealizability of hypergraphs in $\mathbb{R}^4$ can be reduced to 3-dimensional results due to John Conway, Cameron Gordon and Horst Sachs (Theorems 1.1 and 2.5). Before reducing the 4-dimensional results to the 3-dimensional results (§2.3), we reduce the 3-dimensional results to certain 2-dimensional results (§2.2), and the 2-dimensional results to certain 1-dimensional result (§2.1). Thus Proposition 1.2 is reduced to Proposition 2.1 below, Theorem 1.1 to Proposition 1.2, and Theorem 1.4 to Theorem 1.1. This pattern is generalized by Theorem 1.6 below. Because of such 'lowering of dimension' the reader not familiar with 4-dimensional space need not be scared. See also Historical Remark 2.8.

The non-realizability results may be called 'Ramsey intersection theory', just as the Conway–Gordon–Sachs theorem is departure point of *Ramsey linking theory*. See surveys [RA05, PS05].

1.3. **Linking and intersection in higher dimensions.** The above relation between intrinsic linking in dimension 3 and non-realizability (i.e. intrinsic intersection) in dimension 4 generalizes to a relation between intrinsic linking and non-realizability in consecutive dimensions. That is, the above results on intrinsic linking and non-realizability turn out to be particular cases of a result in arbitrary dimensions (Theorem 1.6). For simplicity we mention dimensions higher than 4 only in that theorem and state only the 'quantitative', 'moreover' parts, omitting the 'existence' parts.

Take two k-dimensional simplices in $(2k-1)$-space of whose $2k+2$ vertices no $2k$ lie in the same $(2k-2)$-dimensional hyperplane. The two simplices are called **linked**, if the boundary of the first simplex intersects the convex hull of the second simplex exactly at one point.

Theorem 1.6. *Take any $n+3$ points in $\mathbb{R}^n$ of which no $n+1$ points lie in the same $(n-1)$-dimensional hyperplane.*

For n even mark the intersection points of the interiors of convex hulls of $n/2$-simplices with vertices at these points. Then the number of marked points is odd.

If n is odd, then the number of linked unordered pairs of $(n+1)/2$-simplices with vertices at these points is odd.

This is Proposition 1.2 for $n=2$, is the Linear Conway–Gordon–Sachs Theorem 1.1 for $n=3$, is the linear version of a result by Lovas-Schrijver-Taniyama for odd $n>3$ [LS98, Corollary 1.1], [Ta00], and is the linear version of the van Kampen-Flores Theorem for n even [vK32, Fl34].

Theorem 1.6 is proved by induction on n. The base is $n=1$ and is trivial. The inductive step is proved in §2 for $n=2,3,4$; the proof for the general case is analogous.

There is also an 'intersection property' of odd-dimensional space (Proposition 2.4.b is an analogue of Theorems 1.2, 1.4, 1.6). It is weaker than the corresponding 'linking property' (Theorems 1.1, 1.6). For 'unlinking properties' see Remark 2.9.

1.4. **Cartesian product and the Menger conjecture.** The *(Cartesian) product* $F\times F'$ of two figures F,F' in $\mathbb{R}^3$ is the set of points $(x,y,z,x',y',z')\in\mathbb{R}^6$ such that $(x,y,z)\in F$ and $(x',y',z')\in F'$. A combinatorial version of this notion is *product* of two graphs (not necessarily planar). This product can be considered (although not canonically) as a hypergraph; see Figure 5, left. In Figure 5, middle and right, splitting of quadrilaterals into triangles is not shown.

Karl Menger conjectured in 1929 that the square of a nonplanar graph is not realizable in $\mathbb{R}^4$ [Me29]. This was proved only in 1978 by Brian Ummel [Um78] (Theorem 3.3). A simpler proof was obtained in 2003 by Mikhail Skopenkov [Sk03]. There is a short formula for the minimal number d such that given product of several graphs is realizable in $\mathbb{R}^d$ [Sk03].[7] The argument of [Sk03] is based on discovery and use of the relation between linking and non-realizability phenomena in dimensions 3 and 4 (illustrated in §3.2 and §3.4).

1.5. **Linear, piecewise-linear (PL) and topological versions.** We present elementary statements and simple proofs of the *linear* versions of the above classical results. *PL and topological* realizations (=embeddings) of hypergraphs are defined and discussed e.g. in [Sk18, §3.2], [Sk, §5]. Our proofs are easily generalized to the PL versions [Sk03, Zi13]. The 'quantitative' PL versions of Proposition 1.2 and Theorems 1.1, 1.4, [Sk18, Theorem 3.1.2]

[7]This formula (generalizing the Menger conjecture) was announced in a 1992 preprint of Marek Galecki. However, after an extensive search Robert J. Daverman kindly informed the authors of a corresponding result for manifolds [ARS01] that there is no longer any copy of Galecki's dissertation (presumably containing a proof) available at the University of Tennessee.

(analogous to their 'moreover' parts) imply the PL versions for *almost-embeddings* (see the PL case of [Sk18, Theorem 1.4.1 and 3.1.6]). The latter imply the *topological* versions (see explanation in [Sk18, the paragraph after Theorem 1.4.1]).

Proof of the Menger conjecture (see §1.4) in [Um78] works for the topological version but is complicated (one computes an obstruction via spectral sequences). Proof in [Sk03] is much simpler but for the topological version uses *the Bryant approximation theorem* which is not easy. A simpler proof could possibly be obtained by proving 'quantitative' PL version of the Menger conjecture (i.e. improvements of Proposition 3.1 and Theorems 3.2, 3.3 analogous to the 'moreover' parts of Proposition 1.2 and Theorems 1.1, 1.4, see Problem 3.9).

1.6. **Comparison with other expositions.** The (linear, PL and topological) van Kampen-Flores theorem has an alternative simple proof using *the van Kampen number*, see e.g. [Sk18, §1.4], [Sk, §1.4, §5]. That proof and the proof sketched in this paper, are presumably the simplest known proofs ('proofs from the Book'). Proofs of the Menger conjecture (see §1.4) using an analogue of the van Kampen number or the Borsuk-Ulam theorem are not known.

Usually the van Kampen-Flores theorem is proved using the Borsuk-Ulam theorem [Pr07, §10.3], [Ma03, §5]. As opposed to this paper (and to the alternative simple proof using the van Kampen number), this requires some knowledge of algebraic topology. And this knowledge does not make things simpler: no known proof of the Borsuk-Ulam theorem (see [Ma03] and the references therein) is easier than direct proof of the van Kampen-Flores theorem (presented here or in [Sk18, §1.4], [Sk, §1.4, §5]). E.g. the Borsuk-Ulam theorem is usually proved using *the degree* analogously to the direct proof of the van Kampen-Flores theorem using the van Kampen number.

Short algebraic proofs of the linear versions of the van Kampen-Flores and the Conway–Gordon–Sachs in the spirit of the 'standard' proof of the Radon theorem are given in [BM15]. However, those proofs do not generalize to PL (or topological) versions.

1.7. **Further generalizations.** The results discussed in this survey are in the basis of ongoing research.

An important area is study of realizability of (higher-dimensional) hypergraphs, including applications of algebraic topology to algorithmic problems. For recent surveys see [Sk08, §4, §5], [MTW11, §1], [Sk18, §3.2]. For a recent application of the relation between intrinsic linking and non-realizability in computer science see [Pa15, Sk18o].

Realizations (=embeddings) are maps without self-intersections. For topological combinatorics and discrete geometry it is interesting to study of maps whose self-intersections are 'not too complicated'. This is similar to study of smooth maps where one needs to study maps whose singularities are 'not too complicated', i.e. to develop singularity theory. An important particular case is studying maps *without triple intersections* and, more generally, maps *without r-tuple intersections*, see e.g. survey [Sk18, §3.3]. For relation of this subject to the topological Tverberg conjecture see survey [Sk16] and references therein.

For analogous problem on embedding dynamical systems see [LT14] and references therein.

2. Proofs and further results

By *k points in $\mathbb{R}^d$* (in this paper mostly $d \leq 4$) we mean a k-element subset of $\mathbb{R}^d$; so these k points are assumed to be pairwise distinct.

2.1. **Intersection in the plane.** Proposition 1.2 is easily proved by analyzing the convex hull of the points. In order to illustrate the 'lowering of dimension' argument in the simplest

situation, let us present another proof of Proposition 1.2 based on reduction to the following obvious 1-dimensional result.

Take 4 points on a line, 2 red and 2 blue. The red and the blue pairs of points are called **linked** if they alternate: red-blue-red-blue or blue-red-blue-red. The following result is obvious:

Proposition 2.1. *Every 4 points in a line can be colored in 2 red and 2 blue so that the red pair is linked with the blue pair.*

Moreover, the number of linked unordered pairs of pairs with vertices at these 4 points is odd.

Proof of the first sentence in Proposition 1.2. We may assume that O is the unique point among given ones whose first coordinate a is maximal. Consider a line $x = b$, where b is slightly smaller than a. Denote by A, B, C, D the remaining points.

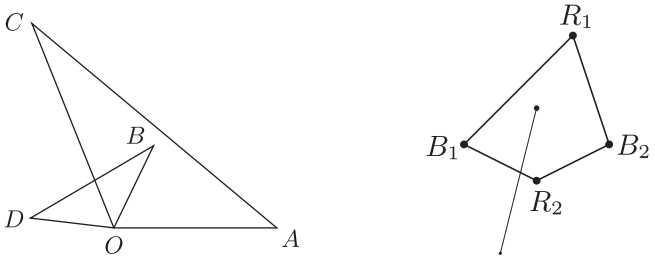


FIGURE 6. Left: to the proof of Proposition 1.2. Right: to Proposition 2.3.b.

If for some two points $X, Y \in \{A, B, C, D\}$ the point X belongs to the segment OY, then we are done. Otherwise we can assume that the points A, B, C, D are seen from O in this order, see Figure 6. Then by the following Lemma 2.2 the outlines of the triangles OAC and OBD have an intersection point different from O. Hence some two sides of the triangles have disjoint vertices and intersect. □

Lemma 2.2 (See figure 6, left)**.** *Assume that two triangles Δ, Δ' in the plane have a common vertex O, and no 3 of their vertices lie in the same line. Then the outlines $\partial\Delta, \partial\Delta'$ of the triangles intersect at an even number of points if and only if the intersection $\partial\Delta \cap \Delta'$ contains only one segment with vertex O.*

This lemma is trivial. It is explicitly stated in order to illustrate higher-dimensional generalizations (Lemmas 2.6 and 3.8).

The 'moreover' part of Proposition 1.2 follows by a simple additional counting analogous to the proof of the Linear Conway-Gordon-Sachs Theorem 1.1 in §2.2 and using the 'moreover' part of Proposition 2.1.

The following propositions are proved analogously to Proposition 1.2. They are used for some 3-dimensional results (Proposition 3.1 and Theorems 3.2, 2.5) in §2.2 and §3.2.

Proposition 2.3. *(a) (See figure 3, right, and Theorem 2.7.) Two triples of points are given in the plane. Then there exist two intersecting segments without common vertices and such that each segment joins the points from distinct triples.*

(b) (See figure 6, right) Suppose that there are 4 red and 2 blue points B_1, B_2 in the plane. Suppose further that any two segments joining points of different colors either are disjoint or intersect at their common vertex. Then there are 2 red points R_1, R_2 such that the quadrilateral $R_1B_1R_2B_2$ does not have self-intersections and the remaining 2 red points

lie on different sides of the quadrilateral. (I.e. a general position polygonal line joining the remaining 2 red points intersects the outline of the quadrilateral at an odd number of points.)

See more results in [Sk18, §1.1].

2.2. **Linking and intersection in three-dimensional space.** First we illustrate the 'lowering of the dimension' idea (see Remark 1.5) of proof of the Linear Conway–Gordon–Sachs Theorem 1.1 by proving its weaker versions.

Proposition 2.4. *(a) From any 6 points in 3-space one can choose 5 points O, A, B, C, D such that the triangles OAB and OCD have a common point other than O.*

(b) From any 6 points in 3-space one can choose disjoint pair and triple such that the segment joining points of the pair intersects the triangle spanned by the triple.

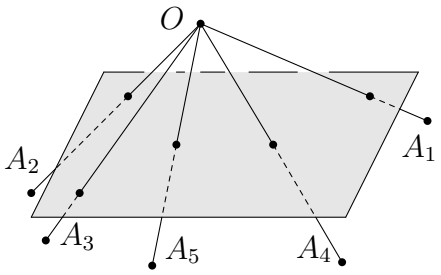


FIGURE 7. To the proofs of Proposition 2.4.a and Theorem 1.1. A plane in $\mathbb{R}^3$ intersects the segments $OA_1, \dots, OA_5$ by points $A'_1, \dots, A'_5$.

Proof of (a). Without loss of generality we may assume that there is a unique 'highest' point O among the given ones. Consider a 'horizontal' plane slightly below the point O. Consider the intersection of this plane with the union of triangles OAB for all pairs A, B of given points. Now the assertion follows by Proposition 1.2. □

Part (b) follows from (a). Part (b) is an improvement of (a) and is a spatial analogue of Proposition 1.2 (without the 'moreover' part).

Figure 4, left, shows that the analogue of (a) for 5 points is false.

Proof of Theorem 1.1. We may assume that O is the unique point among given ones whose first coordinate a is maximal. Consider a plane $x = b$, where b is slightly smaller than a. Denote by $A'_1, \dots, A'_5$ the intersection points of this plane and segments joining O to other given points. See Figure 7.

In 3-space *a segment p is below a segment q* (looking from point O), if there exists a half-line OX with the endpoint O that intersects the segment p at a point $P := p \cap OX$, the segment q at a point $Q := q \cap OX$, $P \neq Q$, so that Q is in the segment OP. So in the plane $x = b$ we can draw a figure analogous to Figure 2, right. Since no 4 of the given points $O, A_1, \dots, A_5$ lie in the same plane, the number of those sides of the triangle $A_3A_4A_5$ that are higher than A_1A_2 equals to the number of intersection points of the outline of the triangle $A_3A_4A_5$ with the triangle OA_1A_2. Also, a segment cannot intersect a triangle by more than 2 points. All this implies that *the triangles OA_1A_2 and $A_3A_4A_5$ are linked if and only if A_1A_2 is below an odd number of sides of the triangle $A_3A_4A_5$.*

For the existence of linked triangles it suffices to prove that *if no 3 of 5 points in the plane lie in the same line and the intersection points (different from vertices) of segments joining these points are marked so as to show that one segment 'passes below the other', then*

there is a segment that is below exactly one side of its 'complementary' triangle. This can be proved by considering all possible cases. Instead of giving details, let us present a counting argument that gives the 'moreover' part.

The following numbers have the same parity:

• the number of linked unordered pairs of triangles formed by given points;

• the number of segments A_iA_j that are below an odd number of sides of their 'complementary' triangles $A_kA_lA_m$, $\{i,j,k,l,m\}=\{1,2,3,4,5\}$;

• the number of ordered pairs (A_iA_j, A_kA_l) of segments of which the first is below the second;

• the number of intersection points of segments whose vertices are $A'_1,\dots,A'_5$.

By Proposition 1.2 the latter number is odd. □

The following version of Theorem 1.1 is analogously reduced to Proposition 2.3.b [Zi13] and is used for some 4-dimensional result (Theorem 3.3) in §3.4.

Take two space quadrilaterals (i.e. closed quadrangular polygonal lines) $ABCD$ and $A'B'C'D'$ in 3-space no 4 whose 8 vertices lie in the same plane. The quadrilaterals are called *linked modulo 2* if the number of intersection points of the polygonal line $ABCD$ with the union of the triangles $A'B'C'$ and $A'D'C'$ is odd. (As opposed to triangles, there are space quadrilaterals *linked* but not linked modulo 2 [Wl].) Proposition 3.7 illustrates this notion of linking.

Theorem 2.5 (Linear Sachs Theorem; [Sa81])**.** *Suppose that there are 8 general position points in 3-space, 4 red and 4 blue. Then there are two linked space quadrilaterals with vertices at these points consisting of segments joining points of different colors.*

2.3. **Linking and intersection in four-dimensional space.** This and the following two subsections are independent of each other (except that §3.4 uses the statement of Lemma 3.8), so they can be read in any order.

Proof of the first sentence in the Linear Van Kampen-Flores Theorem 1.4. We may assume that no 5 of the given 7 points $O, A_1,\dots,A_6$ lie in the same 3-dimensional hyperplane (see the sentence after Proposition 1.2). We may also assume that O is the unique point among them whose first coordinate a is maximal. Consider a 3-dimensional hyperplane $x=b$, where b is slightly smaller than a.

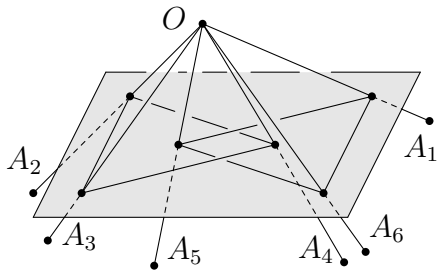


FIGURE 8. To the proof of Theorem 1.4. A hyperplane in $\mathbb{R}^4$ (shown as a plane in $\mathbb{R}^3$) intersects the segments $OA_1,\dots,OA_6$ at 6 points $A'_1,\dots,A'_6$ which are vertices of two linked triangles.

Take the 6 intersection points $A'_1,\dots,A'_6$ of the hyperplane with the segments $OA_1,\dots,OA_6$; see Figure 8. Clearly, no 4 of the obtained 6 points lie in the same plane. Hence by the Linear Conway–Gordon–Sachs Theorem 1.1 there are two linked triangles with vertices at

these points. Without loss of generality, the vertices of the first triangle belong to the segments joining O to A_2, A_3, A_4, and the vertices of the second triangle belong to the segments joining O to A_1, A_5, A_6. The above triangles are the intersections with the hyperplane of the tetrahedra $OA_2A_3A_4$ and $OA_1A_5A_6$.

Since the triangles are linked, the outline of $A'_2A'_3A'_4$ intersects the triangle $A'_1A'_5A'_6$ at exactly one point. Hence either triangles $A_2A_3A_4$ and $A_1A_5A_6$ intersect (then we are done) or the surface of $OA_2A_3A_4$ intersects the convex hull of $OA_1A_5A_6$ at exactly one segment. In the second case by the following Lemma 2.6 the surfaces of the tetrahedra have an intersection point distinct from O. Since no 5 of the given 7 points lie in the same 3-dimensional hyperplane, any two triangles spanned by the 7 points and having one common vertex intersect only at the vertex. Hence some two faces of the tetrahedra $OA_2A_3A_4$ and $OA_1A_5A_6$ have disjoint vertices and intersect. □

Lemma 2.6. *Assume that two tetrahedra Δ, Δ' in 4-space have a common vertex O, and no 5 of their 7 vertices lie in the same 3-dimensional hyperplane. Then the surfaces $\partial\Delta, \partial\Delta'$ of the tetrahedra intersect at an even number of points if and only if the intersection $\partial\Delta \cap \Delta'$ contains only one segment with vertex O.*

This lemma (and Lemma 3.8 below) is not as obvious as its low-dimensional analogues (Lemma 2.2 and analogous result for a triangle and a tetrahedron in 3-space) because *the surface of a tetrahedron in 4-space does not split 4-space*. Lemma 2.6 is reduced to Lemma 2.2 by proving that the intersection plane of 3-dimensional hyperplanes spanned by the tetrahedra intersects each tetrahedron by a triangle.

The condition on $\partial\Delta\cap\Delta'$ of Lemma 2.6 is equivalent to the following: a small 3-dimensional sphere containing O in its interior intersects Δ and Δ' by two triangles which are *linked* in the sphere. Cf. Lemma 3.8.

The 'moreover' part of Theorem 1.4 follows by a simple additional counting (analogous to the proof of Theorem 1.1 in §2.2) using the 'moreover' part of Theorem 1.1.

The following result can perhaps be deduced analogously to Theorem 1.4 from some 3-dimensional linking result and some 4-dimensional parity lemma.

Theorem 2.7 (cf. Proposition 2.3.a; [Fl34])**.** *Three triples of points in 4-space are given. Then there exist two intersecting triangles without common vertices such that the vertices of each triangle belong to distinct triples.*

Remark 2.8 (historical)**.** Of course general 'lowering of dimension' or 'the link construction' ideas are simple and well-known. Proofs of the *Radon theorem* on convex hulls[8] based on this idea are given in [Pe72, Ko]. For a recent application in computer science see [DE94, proof of 2.3.i]. Also well-known is relation between linking and intersection in consecutive dimensions (e.g. the linking number of two disjoint closed polygonal lines in 3-dimensional sphere ∂D^4 equals to the intersection number of two *general position* 2-dimensional disks in 4-dimensional ball D^4 spanning the two polygonal lines). An elaboration of this idea to a relation between intrinsic linking and non-realizability in consecutive dimensions is non-trivial (cf. the difference between Proposition 2.4.a and Theorem 1.1). Proofs that discover and use that relation seem to have not been published

• before [RST, RST'], Alexander Shapovalov's 2003 solution of an olympic problem, [RSS+, Zi13] for reduction of intrinsic linking to non-realizability in lower dimension (the Conway–Gordon–Sachs theorem),

[8]See e.g. [Sk16, §1] for the statement of the Radon theorem. See [Sk16, §4] for relations between the Radon, the van Kampen-Flores and the Conway–Gordon–Sachs theorems.

• before [Sk03, Example 2, Lemmas 2 and 1'], [RSS+] for reduction of non-realizability to intrinsic linking in lower dimension (the van Kampen-Flores theorem and the Menger conjecture, see below).

Remark 2.9 (unlinking properties)**.** (2) *There are 5 general position points in the plane such that every segment joining 2 of these points intersects the outline of the triangle formed by the remaining points at an even number of points.*

This means that every pair of points is unlinked with the triangle formed by the remaining points. We do not spell out analogous interpretations of properties (3), (4-2) and (4-3) below.

(2') *For every 5 general position points in the plane the number of those segments joining 2 of these points that intersect the outline of the triangle formed by the remaining points exactly at one point, is even.*

Proofs of (2,2') are easy and are left to the reader.

In 3-space instead of unlinking properties (2,2') there are a linking property (Theorem 1.1) and the following unlinking properties.

(3) *There are 6 general position points in 3-space such that every segment joining 2 of these points intersects the surface of the tetrahedron formed by the remaining points at an even number of points.*

(3') *For every 6 general position points in 3-space the number of those segments joining 2 of these points that intersect the surface of the tetrahedron formed by the remaining points exactly at one point, is even.*

For (3) we can take points on a helix, see Figure 2. For (3') we can use the symmetry of linking [Sk, Symmetry Lemma 4.2] to prove that this number is twice the number from the 'moreover' part of Theorem 1.1.

The odd-dimensional analogue of the 'moreover' part of Proposition 1.2 fails by (3'). So under transition from dimension 2 to dimension 3 the property of the existence of intersection is preserved, while the parity of the number of intersections change.

It would be interesting to prove the following conjectures and their higher-dimensional analogues. (I am grateful to M. Tancer for sending me proof of the PL version of (4-3).)

(4-3) *There are 7 general position points in 4-space such that every triangle formed by 3 of these points intersects the surface of the tetrahedron formed by the remaining points at an even number of points.*

(4'-3) *For every 7 general position points in 4-space the number of those triangles spanned by 3 of these points that intersect exactly at one point the surface of the tetrahedron formed by the remaining points, is even.*

(4-2) *There are 7 general position points in 4-space such that every segment joining 2 of these points intersects the surface of the 4-simplex formed by the remaining points at an even number of points.*

(4'-2) *For every 7 general position points in 4-space the number of those segment joining 2 of these points that intersect exactly at one point the surface of the 4-simplex formed by the remaining points, is even.*

3. Realizability of products and the Menger conjecture

3.1. **Realizability of products.** For motivations see §1.4. Suppose that we have mn points A_{jp}, where $j \in [m] := \{1, 2, \dots, m\}$ and $p \in [n]$, in 3- or 4-space. For two-element subsets $\{j, k\} \subset [m]$, $j < k$, and $\{p, q\} \subset [n]$, $p < q$, denote by $jk \times pq$ the collection, or the union, of two triangles $A_{jp}A_{kq}A_{jq}$ and $A_{jp}A_{kq}A_{kp}$ having a common side (see Figure 5, left). This union could be, but need not be, a plane quadrilateral. An (m, n)**-product** is a collection

of $2mn$ triangles from

$$jk \times pq, \quad \text{where} \quad 1 \le j < k \le m, \quad 1 \le p < q \le n.$$

The union of triangles of (m, n)-product is a polyhedral and possibly self-intersecting
- square, if $m = n = 2$ (Fig. 5, left);
- lateral surface of a cylinder, if $m = 3$ and $n = 2$ (Fig. 5, middle);
- torus, if $m = n = 3$ (Fig. 5, right).

A typical example is the Cartesian product of m points in the plane and n points in the line (or in the plane).

Proposition 3.1. *Any $(4, 4)$-product in 3-space has two triangles which have disjoint vertices but intersect.*

Proposition 3.1 is reduced to Proposition 2.3.a in §3.2.

In terms of hypergraphs or complexes Proposition 3.1 implies that $K_4 \times K_4$ is not linearly realizable in 3-space. We do not spell out analogous corollaries of the following two theorems.[9]

Theorem 3.2 (Product; [Sk03])**.** *Any $(5, 3)$-product in 3-space has two triangles which have disjoint vertices but intersect.*

The Product Theorem 3.2 is reduced to Proposition 2.3.b in §3.2.

Theorem 3.3 (Square; [Um78, Sk03])**.** *Any $(5, 5)$-product in 4-space has two triangles which have disjoint vertices but intersect.*

The Square Theorem 3.3 is reduced to the Linear Sachs Theorem 2.5 in §3.4.

Example 3.4. *The analogues of Theorems 3.2 and 3.3 are false for*

(a) $(2, n)$-products in 3-space for every n (for $n \le 4$ this is obvious; for $n = 5$ see Figure 4, right: the vertices of the parallelograms are the required 10 points; for $n \ge 6$ the construction is analogous, see §3.2; cf. [RSS', Theorem 1.5]*);*

(b) $(3, n)$-products in 3-space for every $n \le 4$ (for $n \le 3$ this is obvious, see Figure 5, right; for $n = 4$ the construction is analogous, see §3.2);

(c) $(4, n)$-products in 4-space for every n (see §3.4).

3.2. **Realizability of products in three-dimensional space.**

Proof of Example 3.4.a. Let $(0, 0, 0), V, A_{11}, \dots A_{1n}$ be points in $\mathbb{R}^3$ of which no 4 lie in the same plane. For every $p \in [n]$ denote $A_{2p} := V + A_{1p}$. If V is close enough to $(0, 0, 0)$, then the points A_{jp}, $j \in \{1, 2\}$, $p \in [n]$, are as required: there are no two triangles with vertices at these points which have disjoint vertices but intersect.

Indeed, $12 \times pq$ is a parallelogram for every $p \ne q$. Since no 4 of the points $(0, 0, 0)$, V, $A_{11}, \dots A_{1n}$ lie in the same plane, for any distinct p, q, r, s the segments $A_{1p}A_{1q}$ and $A_{1r}A_{1s}$ are disjoint. Since V is close enough to $(0, 0, 0)$, the same holds for 1 replaced by 2. Then any two (convex hulls of) parallelograms $12 \times pq$ and $12 \times rs$ that have no common side are disjoint. Therefore the points A_{jp} are as required. □

Proof of Example 3.4.b. Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the rotation through $\frac{2\pi}{3}$ w.r.t. x-axis. Let

$$(A_{11}, A_{12}, A_{13}, A_{14}) = ((1, 0, 1), (-1, 0, 1), (0, 0, 2), (0, 0, 3)).$$

[9]Proof of Proposition 3.1 shows that even $K_{3,1} \times K_{3,1}$ is not linearly realizable in 3-space. Analogous improvements of the following two theorems are false.

Let $A_{2p} = f(A_{1p})$ and $A_{3p} = f(f(A_{1p}))$ for every $p \in [4]$. Cf. Figure 5, right. Then the points A_{jp}, $j \in [3]$, $p \in [4]$, are as required: there are no two triangles with vertices at these points which have disjoint vertices but intersect.

Indeed, $jk \times pq$ is a parallelogram for every $j \neq k, p \neq q$. Since every two segments joining points A_{1p} either are disjoint or intersect at a common vertex, any two of such parallelograms that have no common side are disjoint. Therefore the points A_{jp} are as required. □

Proof of Proposition 3.1. (The proof is analogous to Proposition 2.4.) Take a small tetrahedron containing A_{11} in its interior. For every $j = 2, 3, 4$ color in red the intersection point of the surface S of the tetrahedron with the segment $A_{11}A_{j1}$, see Figure 9, left. For every $k = 2, 3, 4$ color in blue the intersection point of S with the segment $A_{11}A_{1k}$. (The intersection of S with the union of the triangles of the $(4, 4)$-product is the image of a *piecewise linear map* of the graph $K_{3,3}$ to S.) Then by an analogue of Proposition 2.3.a (cf. [Sk18, Remark 1.5.1.d]) there are $2 \leq j < k \leq 4$ and $2 \leq p < q \leq 4$ such that the triangles $A_{11}A_{1p}A_{j1}$ and $A_{11}A_{1q}A_{k1}$ have a common point other than A_{11}. Hence without loss of generality the segment $A_{1p}A_{j1}$ intersects the triangle $A_{11}A_{1q}A_{k1}$. So the triangles $A_{jp}A_{1p}A_{j1}$ and $A_{11}A_{1q}A_{k1}$ have disjoint vertices but intersect. [10] □

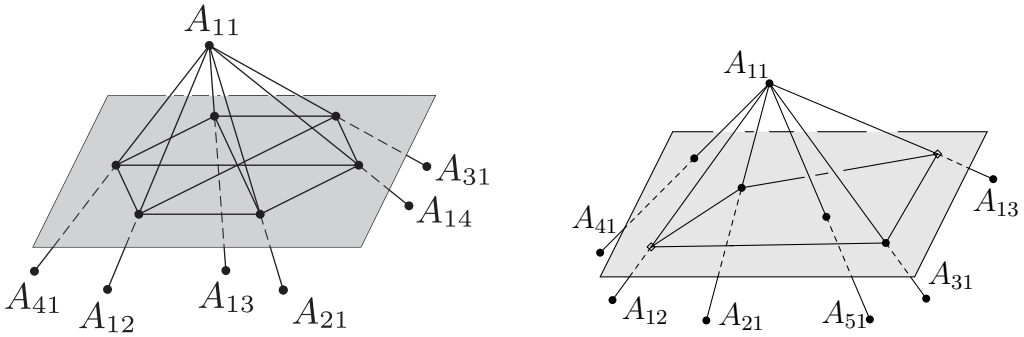


FIGURE 9. To the proofs of Proposition 3.1 (left) and the Product Theorem 3.2 (right)

Given 9 points A_{jk}, $j, k \in \{u, v, w\}$, in 3- or 4-space denote by T_{uvw} the *body* of the corresponding $(3, 3)$-product, i.e. the union of products $jk \times pq$ (defined at the beginning of §3.1) taken for every 2-element subsets $\{j, k\}, \{p, q\} \subset \{u, v, w\}$. See Figure 5, right. (As opposed to the figure, T_{uvw} can have self-intersections.) We abbreviate 'the body of a $(3, 3)$-product' to 'a $(3, 3)$-product'.

Proof of the Product Theorem 3.2. Take a small tetrahedron containing A_{11} in its interior. For every $j = 2, 3, 4, 5$ color in red the intersection point of the surface S of the tetrahedron with the segment $A_{11}A_{j1}$, see Figure 9, right. For every $k = 2, 3$ color in blue the intersection point of S with the segment $A_{11}A_{1p}$. (The intersection of S with the union of the triangles of the $(5, 3)$-product is the image of a piecewise linear map of the graph $K_{4,2}$ to S.)

Denote the blue points by B_1, B_2. The intersection of a triangle $A_{11}A_{j1}A_{1p}$ with S is called an *arc*. Analogously to the last sentences from the proof of Proposition 3.1 either some two triangles of the $(5, 3)$-product have disjoint vertices and intersect, or any two arcs joining points of different colors can only intersect at their common vertex. In the second case

[10] It is here that we use a specific triangulation of $K_4 \times K_4$. Thus the point A_{11} is not interchangeable with other A_{jp}. So we have to consider a tetrahedron instead of a (hyper)plane as in Theorems 1.1, 1.4 and 3.3. Analogous remark applies for the proof of the Product Theorem 3.2 below.

by an analogue of Proposition 2.3.b there are 2 red points R_1, R_2 such that the polygonal line $R_1B_1R_2B_2$ formed by arcs does not have self-intersections and the remaining two red points R_3, R_4 lie in S on different sides of the polygonal line. Without loss of generality, R_1, B_1, R_2, B_2 belong to the segments joining A_{11} to A_{21}, A_{12}, A_{31}, A_{13}, respectively, and R_3, R_4 belong to the segments joining A_{11} to A_{41}, A_{51}, respectively. Then the points R_3 and R_4 are intersection points of S and the outline of the triangle $A_{11}A_{41}A_{51}$. The intersection of $S \cap A_{11}A_{41}A_{51}$ is a polygonal line joining R_3 and R_4. The polygonal line $R_1B_1R_2B_2$ is the intersection of S and the $(3,3)$-subproduct T_{123}. Since R_3, R_4 lie in S on different sides of the polygonal line, $(S \cap A_{11}A_{41}A_{51}) \cap (S \cap T_{123}) \neq \emptyset$. Thus $A_{11}A_{41}A_{51} \cap T_{123} \neq \emptyset$. Hence one of the two triangles $A_{11}A_{41}A_{51}$, $A_{45}A_{41}A_{51}$ and some triangle from T_{123} have disjoint vertices but intersect. □

3.3. **Parity Lemmas.** For the proof of the Square Theorem 3.3 we need Lemma 3.8 whose simpler analogues were already used above (see Lemmas 2.2, 2.6 and an argument on a triangle and a $(3,3)$-product in 3-space from the proof of the Product Theorem 3.2).

Proof of Lemma 3.8 allows to exhibit a basic idea of homology theory (i.e. Poincaré Lemma on the homology of Euclidean space) in an elementary language accessible to non-specialists. See a similar alternative proof in [Zu] and more on parity lemmas in [Sk18, §1.3], [Sk, §4].

In order to illustrate the idea in the simplest situation, we start with a planar version of a 3-dimensional 'general position' parity lemma (Lemma 3.6) required for Lemma 3.8.

Some points in the plane **are in general position**, if no three of them lie in the same line and no three segments joining them have a common interior point.

Lemma 3.5 (Parity; [Sk18, Parity Lemma 1.3.7])**.** *Any two closed polygonal lines in the plane whose vertices are in general position intersect at an even number of points.*

We need a generalization of the following evident fact: *if no 4 of the vertices of a polygonal line and of a tetrahedron in 3-space lie in the same plane, then the polygonal line and the surface of the tetrahedron intersect at an even number of points.*

Some points in 3-space **are in general position**, if no 4 of them lie in the same plane, and for every pair, triple and triple of the points the common points of their convex hulls is the same as the convex hull of the set of their common points. (In particular, if the pair, triple and triple are pairwise disjoint, then their convex hulls do not have a common point.) E.g. in general position are

• the set of 6 points in Figure 2. (Consider a regular hexagon in a horizontal plane. Point A_j lies exactly above the vertices of the hexagon at the height $j = 1, 2, \dots 6$.)

• the set of points with *Cartesian coordinates* $(t; t^2; t^3)$, where $t \in (0, 1)$ ('moment curve').

A **2-cycle** is a collection of (different) triangles such that every segment is the side of an even number (possibly, zero) of triangles from the collection. *The vertices* of a 2-cycle are the vertices of its triangles. *The body* of a 2-cycle is the union of its triangles.

An example of a 2-cycle is the surface of a tetrahedron (possibly, degenerate). Also, the $(3,3)$-product T_{uvw} defined in §3.2 is the body of a 2-cycle.

Lemma 3.6 (Parity)**.** *If the vertices of a polygonal line and a 2-cycle in 3-space are in general position, then the polygonal line intersects the body of the 2-cycle at an even number of points.*

Sketch of the proof. The lemma follows by its particular case when the closed polygonal line is a triangle (analogously to [Sk18, §1.3, proof of the Parity Lemma 1.3.7]). This particular case is reduced to (the case when one polygonal line is a triangle of) the Parity Lemma 3.5

by proving that the intersection of the 2-cycle and the plane containing the triangle is the union of closed polygonal lines. □

Proposition 3.7. *Let $ABCD$ and $A'B'C'D'$ be two closed quadrangular polygonal lines in 3-space no 4 of whose 8 vertices lie in the same plane.*

(a) The polygonal lines are linked if and only if an odd number among the following pairs of triangles are linked pairs:

$$(ABC, A'B'C'), \quad (ABC, A'D'C'), \quad (ADC, A'B'C'), \quad (ADC, A'D'C').$$

(b) Assume that $\Delta_1, \dots, \Delta_k$ are triangles in 3-space such that $\Delta_1, \dots, \Delta_k, ABC, ADC$ is a 2-cycle and the union of their vertices is in general position. (Such a collection of triangles is called a coboundary of $ABCD$.) Assume that $\Delta'_1, \dots, \Delta'_{k'}$ is an analogous collection of triangles for $A'B'C'D'$. The polygonal lines are linked if and only if an odd number among the kk' pairs $(\Delta_j, \Delta'_{j'})$ of triangles are linked pairs.

Proof. Part (a) is a particular case of (b) for $k = k' = 2$, $\Delta_1 = ABC$, $\Delta_2 = ADC$, $\Delta'_1 = A'B'C'$, $\Delta'_2 = A'D'C'$.

Denote by $\partial\Delta$ the outline of a triangle or a quadrilateral Δ. For a finite set S denote by $|S|$ the number of elements in S. By $\underset{2}{\equiv}$ denote congruence modulo 2. Part (b) follows because

$$|ABCD \cap (A'B'C' \cup A'D'C')| \underset{2}{\equiv} \sum_{j'=1}^{k'} |ABCD \cap \Delta'_{j'}| \underset{2}{\equiv} \sum_{j=1,\ j'=1}^{k,k'} |(\partial\Delta_j) \cap \Delta'_{j'}|.$$

Here the first congruence follows by the Parity Lemma 3.6. □

Lemma 3.8. *Assume that two $(3,3)$-products T_{123} and T_{145} in 4-space intersect at a unique point A_{11}, which is their common vertex, no 5 of their vertices lie in the same 3-dimensional hyperplane, and the triangles of $(3,3)$-products having disjoint vertices are disjoint. Consider the intersection of the union of triangle of T_{123} containing A_{11} and the union of (the convex hulls of) tetrahedra $A_{11}A_{14}A_{41}A_{15}$ and $A_{11}A_{14}A_{41}A_{51}$. Then this intersection contains an even number of segments with vertex A_{11}.*

Proof. The conclusion of the lemma is equivalent to the following: a small 3-dimensional sphere containing O in its interior intersects T_{123} and T_{145} by two quadrangular polygonal lines which are *linked* in the sphere.

Denote by $\Delta_1, \dots, \Delta_9$ ($\Delta'_1, \dots, \Delta'_9$) those triangles of T (of T') that do not contain O. Let $OX = \operatorname{conv}\{\{O\} \cup X\}$ be the cone over X with the center O. Then $(T \cap T') - \{O\} = \emptyset$ consists of an even number of points. Hence there is an even number of pairs $(j, j') \in [9]^2$ such that the surfaces of tetrahedra $O\Delta_j$ and $O\Delta'_{j'}$ intersect at an odd number of points. By (a spherical analogue of) Lemma 2.6 the latter number has the same parity as the number of pairs $(j, j') \in [9]^2$ such that the triangles $\pi \cap O\Delta_j$ and $\pi \cap O\Delta'_{j'}$ are linked. So the lemma follows by (a spherical analogue of) Proposition 3.7.b. □

3.4. **Realizability of products in four-dimensional space.**

Sketch of the proof of a weaker version of Example 3.4.c: $(3,5)$-product in 4-space. Take a 3-dimensional hyperplane in $\mathbb{R}^4$ (shown in Figure 10, left, as a plane in 3-space). In this hyperplane take 10 vertices A_{jp}, where $j \in [5]$, $p \in \{1,2\}$, shown in Figure 4, right. Take a vector v not parallel to the hyperplane. Set $A_{j3} := A_{j1} + v$. (In Figure 10, left, we see the lateral surface of the prismoid $A_{41}A_{42}A_{43}A_{51}A_{52}A_{53}$.) Then the points A_{jp}, $j \in [5]$, $p \in [3]$,

are as required: there are no two triangles with vertices at these points which have disjoint vertices but intersect. □

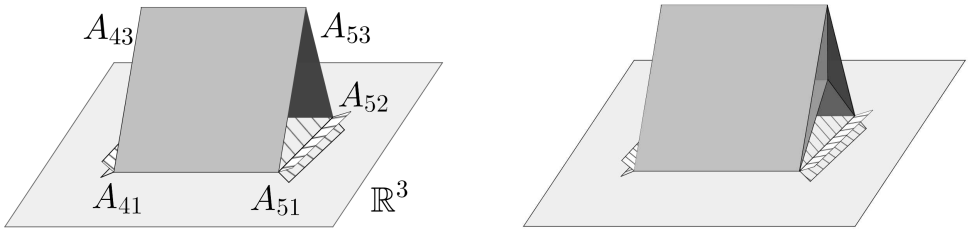


FIGURE 10. Left: to realization in $\mathbb{R}^4$ of the *product* of the complete graphs on 5 and on 3 vertices.
Right: to realization in $\mathbb{R}^4$ of the *product* of the complete graphs on 5 and on 4 vertices.

Sketch of the proof of Example 3.4.c. See Figure 10, right. Take points $A_{jp} \in \mathbb{R}^3 \subset \mathbb{R}^4$, $j \in \{1,2\}$, $p \in [n]$ from the proof of Example 3.4.a. Then $\overrightarrow{A_{1p}A_{1q}} = \overrightarrow{A_{2p}A_{2q}}$ for every $p \neq q$. Take vectors $v_3, v_4 \in \mathbb{R}^4$ not parallel to the hyperplane $\mathbb{R}^3 \subset \mathbb{R}^4$. Denote $A_{jp} := A_{1p} + v_j$, $j \in \{3,4\}$. We can take v_3, v_4 so that A_{14} is an interior point of the triangle $A_{11}A_{12}A_{13}$. Then the points A_{jp}, $j \in [4]$, $p \in [n]$, are as required: there are no two triangles with vertices at these points which have disjoint vertices but intersect. □

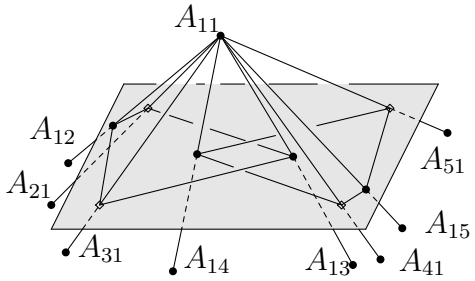


FIGURE 11. To the proof of the Square Theorem 3.3.

Proof of the Square Theorem 3.3. We may assume that no 5 of the given 25 points A_{jp} lie in the same 3-dimensional hyperplane (see the sentence after Proposition 1.2). We may also assume that A_{11} is the unique point among them whose first coordinate a is maximal. Consider a 3-dimensional hyperplane $x = b$, where b is slightly smaller than a.

For every $j = 2,3,4,5$ color in red the intersection point of the hyperplane with the segment $A_{11}A_{1j}$; see Figure 11. For every $p = 2,3,4,5$ color in blue the intersection point of the hyperplane with the segment $A_{11}A_{p1}$. Clearly, no 5 of the 8 colored points in the hyperplane lie in the same plane. Hence by the Linear Sachs Theorem 2.5 there are two linked closed quadrangular polygonal lines whose vertices are the colored points and whose edges have endpoints of different colors. Without loss of generality, the vertices of the first polygonal line belong to the segments joining A_{11} to $A_{12}, A_{21}, A_{13}, A_{31}$, and the vertices of the second polygonal line belong to the segments joining A_{11} to $A_{14}, A_{41}, A_{15}, A_{51}$. Then the polygonal lines are the intersections with the hyperplane of the $(3,3)$-products T_{123} and T_{145}. By Lemma 3.8 T_{123} and T_{145} have an intersection point distinct from A_{11}. Hence analogously

to the last two sentences in the proof of Theorem 1.4, some two triangles of T_{123} and T_{145} have disjoint vertices but intersect. □

Problem 3.9. *Find a subset*

$$M \subset \left\{ \{(X,Y),(X',Y')\} \ : \ X,Y,X',Y' \in \binom{[5]}{2},\ X\cap X'=\emptyset \text{ or } Y\cap Y'=\emptyset \right\}$$

such that for any 25 general position points A_{jp}, $j,p\in[5]$, in 4-space there is an odd number of pairs $\{(X,Y),(X',Y')\}\in M$ for which the intersection $(X\times Y)\cap(X'\times Y')$ consists of an odd number of points.

This problem is a particular case of the following generalized Menger problem: Complexes K,L have non-trivial van Kampen obstructions to embeddability in $\mathbb{R}^m$ and in $\mathbb{R}^n$, respectively (see definition e.g. in [Fo04], [Sk18, §1.5] [Sk, §5]*). Does the cartesian product $K\times L$ of K and L have non-trivial van Kampen obstruction to embeddability in $\mathbb{R}^{m+n}$?*

References

[AMS+] *S. Avvakumov, I. Mabillard, A. Skopenkov and U. Wagner.* Eliminating Higher-Multiplicity Intersections, III. Codimension 2, arxiv:1511.03501.

[ARS01] *P. Akhmetiev, D. Repovš and A. Skopenkov*, Embedding products of low-dimensional manifolds in $\mathbb{R}^m$, Topol. Appl. 2001. 113. P. 7-12.

[BE82] * *V.G. Boltyansky and V.A. Efremovich.* Intuitive Combinatorial Topology. Springer.

[BM15] *I. Bogdanov and A. Matushkin.* Algebraic proofs of linear versions of the Conway–Gordon–Sachs theorem and the van Kampen–Flores theorem, arXiv:1508.03185.

[Br26] *P. Bruegel.* The Magpie on the Gallows, 1526, `https://en.wikipedia.org/wiki/The_Magpie_on_the_Gallows`.

[CG83] *J. H. Conway and C. M. A. Gordon*, Knots and links in spatial graphs, J. Graph Theory 7 (1983), 445–453.

[CKS+] * New ways of weaving baskets, presented by G. Chelnokov, Yu. Kudryashov, A.Skopenkov and A. Sossinsky, `http://www.turgor.ru/lktg/2004/lines.en/index.htm`.

[DE94] *T.K. Dey and H. Edelsbrunner.* Counting triangle crossings and halving planes, Discrete Comput. Geom, 12 (1994), 281–289.

[Fl34] *A. Flores*, Über n-dimensionale Komplexe die im E^{2n+1} absolut selbstverschlungen sind, Ergeb. Math. Koll. 6 (1934) 4–7.

[Fo04] * *R. Fokkink.* A forgotten mathematician, Eur. Math. Soc. Newsletter 52 (2004) 9–14.

[GSS+] * Projections of skew lines, presented by A. Gaifullin, A. Shapovalov, A. Skopenkov and M. Skopenkov, `http://www.turgor.ru/lktg/2001/index.php`.

[Io] `https://en.wikipedia.org/wiki/Category:Impossible_objects`

[Ko] * *E. Kolpakov.* A proof of Radon Theorem via lowering of dimension, Mat. Prosveschenie, submitted.

[LS98] *L. Lovasz and A. Schrijver,* A Borsuk theorem for antipodal links and a spectral characterization of linklessly embeddable graphs, Proc. of AMS 126:5 (1998), 1275-1285.

[LT14] *E. Lindenstrauss and M. Tsukamoto,* Mean dimension and an embedding problem: an example, Israel J. Math. 199 (2014).

[Ma03] * *J. Matoušek.* Using the Borsuk-Ulam theorem: Lectures on topological methods in combinatorics and geometry. Springer Verlag, 2008.

[Me29] *K. Menger.* Über plättbare Dreiergraphen und Potenzen nicht plättbarer Graphen, Ergebnisse Math. Kolloq., 2 (1929) 30–31.

[MTW11] *J. Matoušek, M. Tancer, U. Wagner.* Hardness of embedding simplicial complexes in $\mathbb{R}^d$, J. Eur. Math. Soc. 13:2 (2011), 259–295. arXiv:0807.0336.

[MTW12] *J. Matoušek, M. Tancer, U. Wagner.* A geometric proof of the colored Tverberg theorem, Discr. and Comp. Geometry, 47:2 (2012), 245–265. arXiv:1008.5275.

[Pa15] *S. Parsa,* On links of vertices in simplicial d-complexes embeddable in the euclidean $2d$-space, Discrete Comput. Geom. 59:3 (2018), 663–679. arXiv:1512.05164v4.

[Pe72] * *B. B. Peterson.* The Geometry of Radon's Theorem, Amer. Math. Monthly 79 (1972), 949-963.

[Pr07] * *V. V. Prasolov.* Elements of homology theory. 2007, GSM 74, AMS, Providence, RI.

[PS05] * *V. V. Prasolov and M.B. Skopenkov.* Ramsey link theory, Mat, Prosvescheniye, 9 (2005), 108–115.

[PS11] *Y. Ponty and C. Saule.* A combinatorial framework for designing (pseudoknotted) RNA algorithms, Proc. of the 11th Intern. Workshop on Algorithms in Bioinformatics, WABI'11, 250–269.

[RA05] * *J. L. Ramírez Alfonsín.* Knots and links in spatial graphs: a survey. Discrete Math., 302 (2005), 225–242.

[RSS'] *D. Repovš, A. B. Skopenkov and E. V Ščepin.* On embeddability of $X \times I$ into Euclidean space, Houston J. Math. 1995. 21. P. 199-204.

[RSS+] * *A. Rukhovich, A. Skopenkov, M. Skopenkov, A. Zimin*, Realizability of hypergraphs, `http://www.turgor.ru/lktg/2013/1/index.htm`.

[RST] *N. Robertson, P. Seymour, and R. Thomas*, A survey of linkless embeddings, Graph Structure Theory (Seattle, WA, 1991), Contemp. Math. 147, (1993) 125–136.

[RST'] *N. Robertson, P. Seymour, and R. Thomas*, Linkless embeddings of graphs in 3-space, Bull. of the AMS, 21 (1993) 84–89.

[Sa81] *H. Sachs.* On spatial representation of finite graphs, in: Finite and infinite sets, Colloq. Math. Soc. Janos Bolyai, North Holland, Amsterdam (37) 1981.

[Sk03] *M. Skopenkov.* Embedding products of graphs into Euclidean spaces, Fund. Math. 2003. 179. P. 191-198.

[Sk08] * *A. Skopenkov.* Embedding and knotting of manifolds in Euclidean spaces, London Math. Soc. Lect. Notes, 347 (2008) 248–342; arXiv:math/0604045.

[Sk16] * *A. Skopenkov,* A user's guide to the topological Tverberg Conjecture, Russian Math. Surveys, 73:2 (2018), 323–353. Earlier version: arXiv:1605.05141v4. §4 available as *A. Skopenkov*, On van Kampen-Flores, Conway-Gordon-Sachs and Radon theorems, arXiv:1704.00300.

[Sk18] * *A. Skopenkov.* Invariants of graph drawings in the plane, arXiv:1805.10237.

[Sk] * *A. Skopenkov.* Algebraic Topology From Algorithmic Viewpoint, draft of a book, mostly in Russian, `http://www.mccme.ru/circles/oim/algor.pdf`.

[Sk18o] * *A. Skopenkov.* A short exposition of S. Parsa's theorem on intrinsic linking and non-realizability.

[Ta00] *K. Taniyama,* Higher dimensional links in a simplicial complex embedded in a sphere, Pacific Jour. of Math. 194:2 (2000), 465-467.

[Um78] *B. Ummel.* The product of nonplanar complexes does not imbed in 4-space, Trans. Amer. Math. Soc., 242 (1978) 319–328.

[vK32] *E. R. van Kampen*, Komplexe in euklidischen Räumen, Abh. Math. Sem. Hamburg, 9 (1932) 72–78; Berichtigung dazu, 152–153.

[Wl] `https://en.wikipedia.org/wiki/Whitehead_link`

[Zi13] *A. Zimin.* Alternative proofs of the Conway-Gordon-Sachs Theorems, arXiv:1311.2882.

[Zu] *J. Zung.* A non-general-position Parity Lemma, `http://www.turgor.ru/lktg/2013/1/parity.pdf`.

Books, surveys and expository papers in this list are marked by the stars.

Апериодические замощения

Илья Иванов-Погодаев, Алексей Канель-Белов, Иван Митрофанов, Тома Ферник

Этот проект посвящен интересной области математики, связанной с замощениями. Обычно задан некоторый конечный набор плиток (фигур на плоскости), и мы пытаемся замостить плоскость, прикладывая плитки друг к другу так, чтобы между ними не оставалось пустых мест. Типовой задачей является выяснение вопроса, возможно ли замощение с помощью данного набора и какие свойства могут быть или не быть у этого замощения. Для некоторых наборов, например, возможно лишь непериодическое замощение плоскости.

Вам предлагаются задачи разбитые на несколько циклов.

Среди задач встречаются как простые, так и сложные, в том числе есть открытые вопросы, решения которых никто пока не нашел. Будет здорово, если на конференции будет получено продвижение по какому-нибудь такому вопросу.

A Размерность один

Рассмотрим одноразмерные мозаики. *Плитками* тут являются буквы в конечном алфавите, а *краевыми условиями* являются запреты для некоторых букв стоять друг за другом. Обобщая это понятие, можно сказать, что задается конечное число запрещенных слов – конечных последовательностей букв. *Разрешенными* являются слова, не содержащие запрещенных подслов. Аналогом *замощения* является существование бесконечного в обе стороны слова. Нас будет интересовать вопрос, как конечным числом запретов задавать различные структуры слов. Например, запреты aa и bb задают бесконечное периодическое слово с периодом (ab). Ясно, что других разрешенных бесконеных слов с данными запретами не может быть.

A.1 Пусть задано множество S бесконечных слов в алфавите $\{a, b\}$: это слова содержащие одну, две или три буквы b подряд, а остальные буквы a (при этом серий букв b может быть сколько угодно). Верно ли что существует конечный набор запрещенных слов, задающих множество S?

A.2 Посчитайте минимальное необходимое число запретов, чтобы задать следующие периодические последовательности, указаны их периоды: (ab), (aab), $(aabaabab)$, $(aabaababaabaababaabab)$.

A.3 Рассмотрим множеств бесконечных слов в алфавите $\{a, b\}$. Обозначим через a^n слово, где буква a написана n раз подряд. Существует ли конечный набор запрещеных слов, такой что разрешенными словами являются все слова не содержащие кусков ba^nb для различных n и только они?

A.4 Пусть теперь разрешено раскрашивать буквы в конечное число цветов, то есть, например, a_1b может быть запрещенным словом, а a_2b уже нет. Можно ли теперь задать конечное число запрещенных слов так, чтобы разрешенными были все слова не содержащие кусков ba^nb (где в качестве букв a и b могут встречаться любые их оттенки) и только они?

A.5 Пусть множество S состоит из бесконечных слов, содержащих подряд серии лишь из четного количества букв a. Можно ли задать S конечным числом запретов? Изменится ли ответ, если разрешается раскрашивать буквы в конечное число цветов?

A.6 Те же вопросы, если серии состоят из нечетного числа букв a.

A.7 Пусть $u_0 = a$, $u_1 = ab$, $u_{n+2} = u_n u_{n+1}$. Посчитайте необходимое минимальное число запретов, чтобы задать бесконечную периодическую последовательность с периодом u_n.

B Мозаики на плоскости

2D слово это бесконечная квадратная решетка на плоскости, где в каждой клетке может быть написана одна из букв a и b. *Паттерн* это такая же решетка заполненная буквами, но конечная.

B.1 Найдите конечный набор запрещенных паттернов, такой что единственным разрешенным бесконечным замощением является шахматное, где роль белых и черных клеток играют буквы a и b.

B.2 Рассмотрим множество S бесконечных $2D$-слов, таких что любая связная по стороне компонента состоящая из букв b, (вокруг которой расположены только буквы a) состоит из четного числа букв. Можно ли выбрать конечное множество запрещенных паттернов, чтобы разрешенными бесконечными $2D$-словами были только слова из S? Какой будет ответ, если разрешается раскрашивать буквы в конечное число цветов?

B.3 Какой будет ответ в случае нечетных компонент?

B.4 Докажите, что если с помощью плиток можно замостить область включающую круг произвольного размера, то можно замостить и плоскость.

Будем называть *плитками* конечные многоугольники, с помощью которых будем замощать плоскость. Каждый тип плитки можно раскрашивать в конечное число цветов. Таким образом, бесконечные 2D-слова это замощения квадратными плитками двух типов a и b. Локальные правила примыкания плиток в этом смысле соответствуют заданию запрещенных паттернов.

B.5 Рассмотрим шестиугольные плитки, стороны которых либо прямые, либо содержат выпуклость или вогнутость одинаковой формы, несколько примеров указаны на рисунке 1. Выясните, для различных вариантов плиток, можно ли получить замощение бесконечной плоскости, используя только плитки заданного типа?

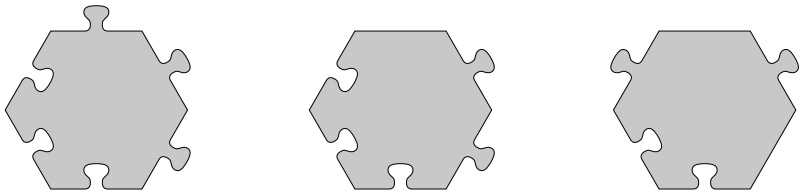


Рис. 1: Правильные шестиугольники с насечками на сторонах.

C Периодичность и квазипериодичность

Вернемся в одномерную ситуацию. Сосредоточимся на вопросе периодичности, и рассмотрим новое свойство, *квазипериодичность*.

C.1 Докажите, что если последовательность периодична, то ее можно задать с помощью конечного числа запретов.

C.2 Пусть можно раскрашивать буквы в конечное число цветов. Пусть задан конечный набор запрещенных слов, и известно, что есть разрешенные бесконечные слова. Докажите, что есть периодическое бесконечное разрешенное слово.

Таким образом, периодичности нельзя избежать в одноразмерной ситуации. Но как далеко мы можем зайти?

C.3 Пусть мы можем задать не более n запретов. Пусть p – наименьший период разрешенного слова. Какого наибольшего значения p мы можем добиться?

C.4 Пусть теперь мы можем задавать сколько угодно запретов, но количество букв в каждом из них не более n. Аналогичный вопрос, какого наибольшего значения p мы можем добиться?

Слово называется *квазипериодичным*, если для любого паттерна P, встречающегося в нем, существует число r, такое что P содержится в любом круге радиуса r.

C.5 Докажите, что каждое периодическое слово является квазипериодичным.

C.6 Найдите непериодическое слово, являющееся квазипериодичным, а также неквазипериодическое слово.

Теперь рассмотрим двумерную ситуацию. Понятия периодичности и квазипериодичности может быть легко обобщены (сделайте это). Как мы увидим позже, периодичности теперь можно избежать. А вот квазипериодичности нельзя:

C.7 Докажите, что если с помощью набора можно замостить плоскость, то это можно сделать с помощью квазипериодического замощения.

D Замощения Робинсона

Первый набор, с помощью которого можно составлять только непериодические замощения был открыт в 1964 году Робертом Бергером и содержал 20426 плиток. Более простой набор показан на рисунке 2. Он был открыт в 1971 Рафаэлем Робинсоном. Выпуклости и выемки (которые могут иметь или не иметь симметрию) вынуждают вариант замощения, при этом линии на плитках образуют интересный рисунок.

D.1 Докажите, что плитками, изображенными на рисунке 2 можно замостить плоскость.

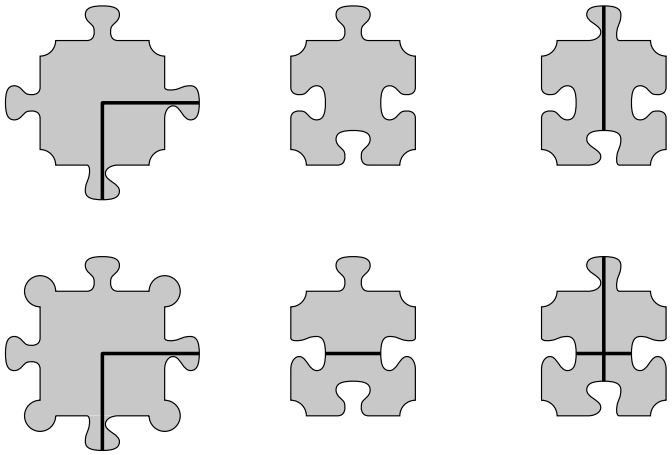


Рис. 2: Шесть квадратов с насечками. Обратите внимание на разную форму насечек.

D.2 Докажите, что любое такое замощение непериодично.

D.3 Назовем *родителем* квадрат, образованный отрезками на плитках, если он пересекает меньший квадрат. *Родословная* квадрата может быть закодирована бесконечным словом в четырех буквенном алфавите, где каждая буква угол, пересекаемый родительским квадратом. В каких случаях два квадрата имеют общего предка?

D.4 Докажите, что существует несчетное количество замощений, даже с учетом изометрии.

E Иерархические замощения

E.1 Покажите, как замостить плоскость с помощью плиток, изображенных на рисунке 3. Цвета точек должны совпадать в месте соприкосновения плиток. Этот набор придуман в 1974 году.

E.2 Аналогичный вопрос, покажите, как замостить плоскость с помощью плиток, изображенных на рисунке 4. Этот набор придуман в начале 90-ых.

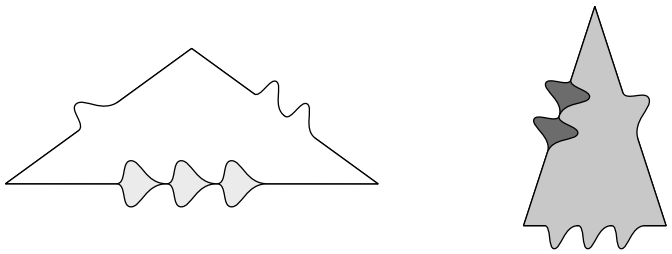


Рис. 3: Треугольники с различными насечками. Верхние углы соответственно 108° и 36°.

E.3 Можно ли задать замощения в двух предыдущих задачах конечным набором локальных паттернов, если нельзя использовать цветовые метки, насечки на сторонах и тому подобное?

Пусть мы составили из плиток несколько макроплиток. Если каждая макроплитка представляет собой фигуру подобную соответствующей плитке (с общим для всех коэффициентом подобия) то мы можем продолжить разбиение и составить макроплитки второго уровня по тому же принципу. Продолжая в том же

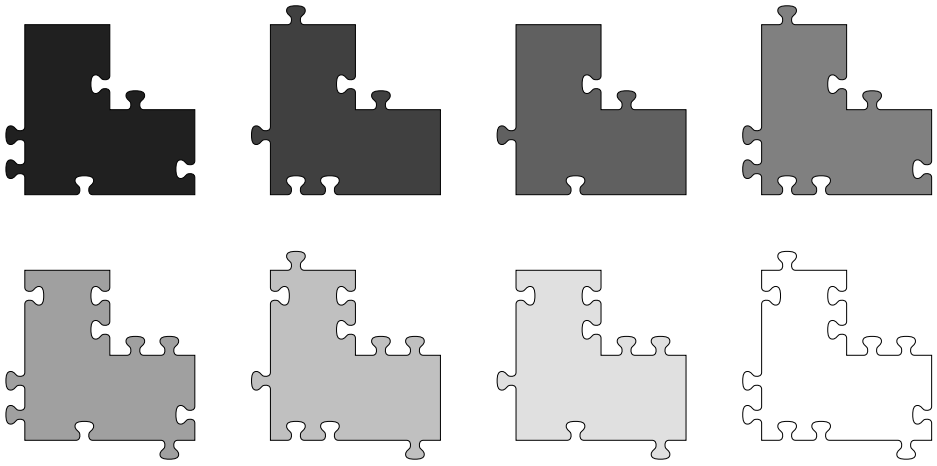


Рис. 4: Notched L-shaped tiles.

духе, мы можем получить замощение плоскости. Способ, с помощью которого макроплитки составляются из плиток, будем называть *подстановкой*. Такой способ построения замощений называется *иерархическим*.

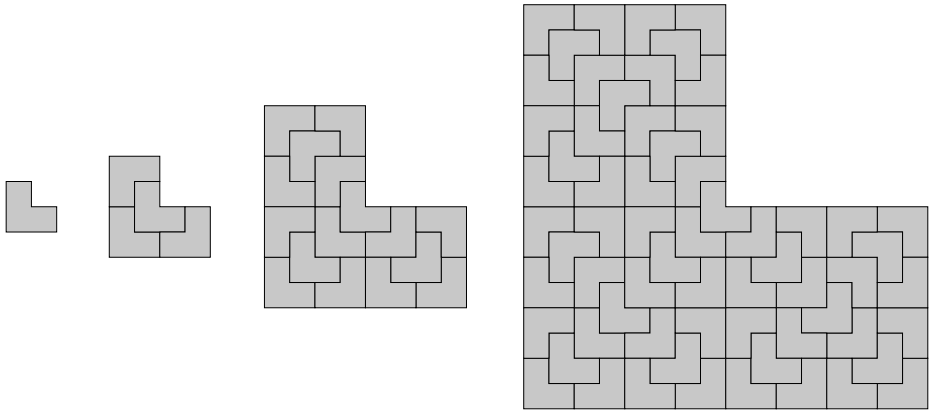


Рис. 5: Пример подстановки.

E.4 Докажите, что иерархические замощения непериодичны.

E.5 Каждая из изображенных на рисунке 6 подстановок задает множество замощений. Для каждого из этих случаев выясните, можно ли определить соответствующее множество с помощью конечного набора запрещенных паттернов – связных сочетаний неперекрывающихся плиток?

Важный результат в математике замощений, полученный в 1998 году, заключается в том, **что для заданной подстановки плиток, их можно раскрасить в конечное число цветов так, что иерархическое замощение может быть определено с помощью конечного числа запрещенных паттернов.** В этом случае говорят, что плитки могут быть декорированы.

E.6 Можете ли вы декорировать плитки для подстановок из предыдущей задачи так, чтобы соответствующие замощения задавались конечным числом запрещенных паттернов?

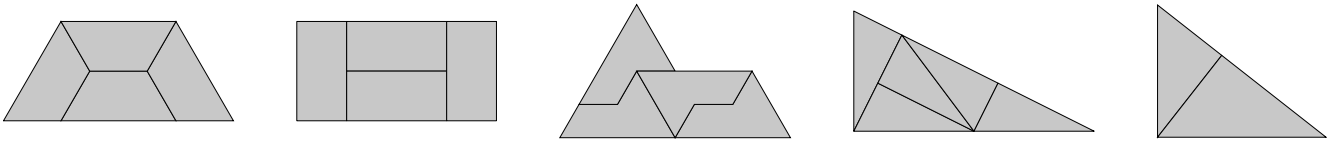


Рис. 6: Подстановки (крайняя справа: плитки гомотетичны но не изометричны).

Апериодические замощения

Илья Иванов-Погодаев, Алексей Канель-Белов, Иван Митрофанов, Тома Ферник

F Самосборка

F.1 *Паттерном* называется связная область без дыр, покрытая несколькими плитками, выложенными по правилам. Паттерн называется *мертвым*, если он не появляется ни в каком замощении плоскости. Найдите мертвый паттерн для набора плиток на рисунках. 2–4.

F.2 Докажите, что в любом наборе плиток, который может замостить плоскость только непериодически, существует мертвый паттерн.

Пусть определена некоторая глобальная константа *температура* и задан *вес* для каждого плиточного ребра. *Схема самосборки* для набора декорированных многоугольных плиток это процесс проводимый по следующим правилам. На первом шаге процесса выкладывается любая плитка. На $N+1$ шаге процесса к уже выложенному куску по правилам прикладывается плитка, причем сумма весов на ее ребрах, которые совпали с ребрами уже установленных ранее плиток, должна быть не менее чем температура. Процесс останавливается, когда к уже построенному куску нельзя добавить ни одной плитки.

F.3 Пусть под весом понимается количество пятен, а температура равна 2. Какие паттерны могут быть построены с помощью схемы самосборки, используемой с плитками на рисунке 7 ?

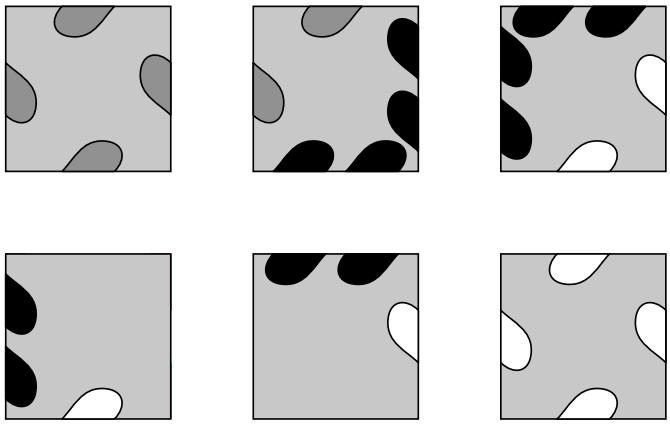


Рис. 7: Схема самосборки для температуры 2, с весами на ребрах равными числу пятен.

Общий результат утверждает, что если задана подстановка на наборе плиток, то можно выбрать температуру, веса, и существует схема самосборки для этого набора, такая что все замощения могут быть построены это иерархические замощения, заданные этой подстановкой.

F.4 Можете ли вы найти схемы самосборки для подстановок из циклов D и E?

G Замощения ромбами

G.1 Докажите, что можно замостить плоскость с помощью квадрата и ромба изображенных на рисунке 8.

G.2 В продолжение предыдущей задачи, придумайте замощение, использующее как можно меньшее количество квадратов.

G.3 Верно ли, что все замощения из первой задачи этого цикла являются квазипериодичными.

G.4 Решите задачи G1-G3 для набора плиток, состоящего из ромбов на Рис. 9

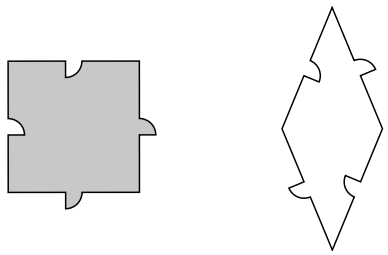


Рис. 8: Серый квадрат и белый ромб (острый угол 45°).

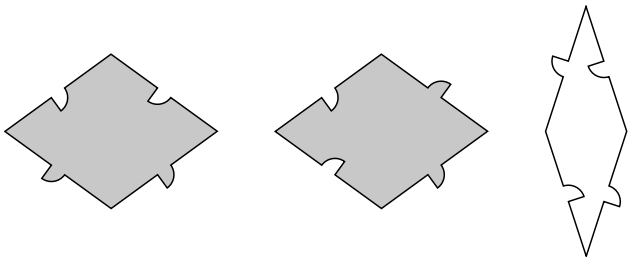


Рис. 9: Серые ромбы (острый угол 72°) и белый ромб (острый угол 36°).

Апериодические замощения

Илья Иванов-Погодаев, Алексей Канель-Белов, Иван Митрофанов, Тома Ферник

Дополнительные подсказки для построения декораций иерархических замощений.

Рассмотрим подробнее подстановку с разбиением трапеции на четыре трапеции меньшего размера (Рисунок 1). Можно заметить, что четыре плитки образуют макроплитку, четыре макроплитки образуют макроплитку следующего уровня и так далее. Мы хотим ввести некоторые локальные правила, чтобы из плиток можно было составлять только непериодические мозаики. Для этого нужно проконтролировать, чтобы плитки приставлялись друг к другу только в соотвествии с правилами подстановки. Мы будем использовать немного другой способ задания локальных правил.

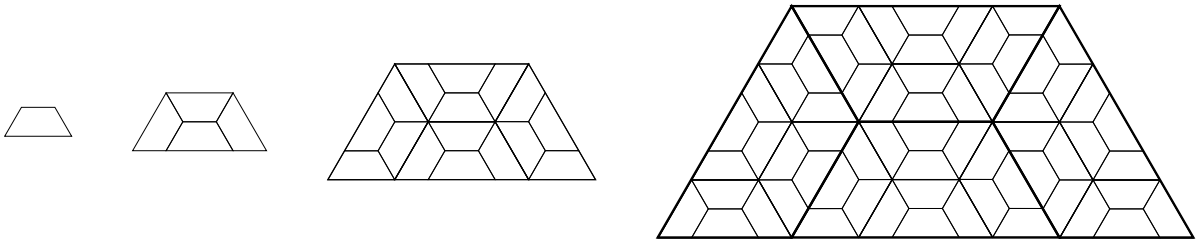


Рис. 1:

Введем некоторый язык, который кодирует различные виды вершин и ребер у наших плиток. *Правильной* мозаикой будем называть замощение плоскости полученное с помощью указанной подстановки с трапециями. *Правильными* будем называть те паттерны, которые присутствуют в правильной мозаике.

С одной стороны, мы определяем разметку нашей подстановочной системы, а с другой доказываем (индукцией по рангу) что разметка делает возможными только правильные мозаики.

E.7 Покажите, что любая вершина на правильном паттерне относится к одному из трех типов, изображенных на рисунке 2. (Трапеции, окружающие вершины могут быть макроплитками любого размера). Также покажите, что любое ребро макроплитки на таком паттерне относится к одному из пяти типов, показанных на рисунке 3. Покажите, что каждая сторона макроплитки является большим основанием в двух каких-то (макро)плитках.

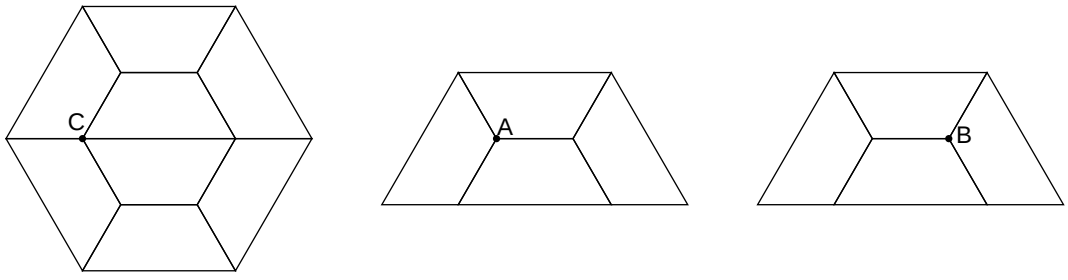


Рис. 2:

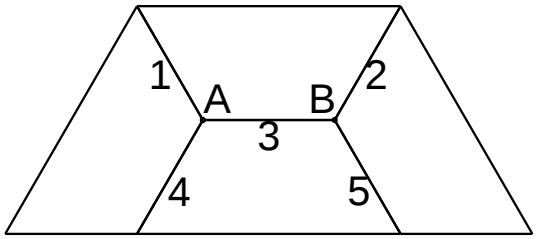


Рис. 3:

Как кодировать ребра плиток? Ребро каждой плитки будет иметь тот же тип, как у максимального ребра, частью которого оно является. Таким образом, у плитки четыре вершины и четыре ребра, типов которых конечное количество. Идея запретов получает новое звучание – можно рассматривать пути на мозаике, кодированные последовательностями типов ребер и вершин, и конструировать запреты таких путей

E.8 Приведите пример последовательностей вида XYZ, где X, Z – типы ребер, Y тип вершины, которые не могут кодировать никакой путь, лежащий на правильной мозаике.

Наша цель состоит в том, чтобы оперируя запретами путей, обеспечить, чтобы единственным возможным замощением была бы правильная мозаика. Но для этого метод запретов путей требуется доработать.

Введем дополнительнй тип вершин D, он будут обозначать вершины на большом основании, учавствующие в разбиении плитки на следующем уровне разбиения (рис. 4). При дальнейших разбиениях на этом большом основании будут образовываться уже вершины типа C.

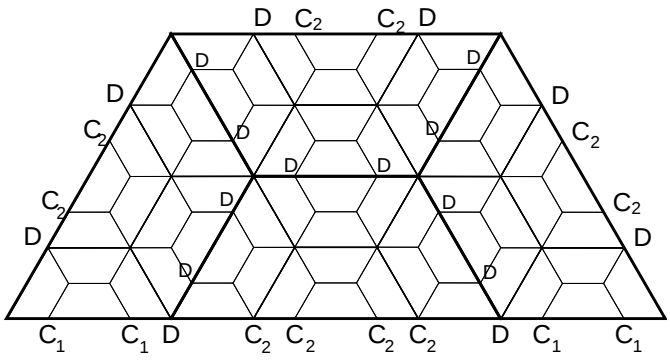


Рис. 4:

Также введем дополнительные цвета вершин C. Для каждой такой вершины можно найти плитку, на большом основании которой она лежит. Будем считать, что вершины, расположенные между краем основания и вершиной типа D имеет цвет один, а вершины, расположенные между двумя вершинами типа D (может быть далеко от них) имеют цвет два.

Также нам потребуются дополнительные цвета ребер. Любое ребро либо является ребром самой мелкой плитки, либо лежит на каком-то большом основании некоторой макроплитки. Цветом ребра будем считать упорядоченную пару "боссов"(X, Y), где X и Y – типы вершин (с цветами) в концах этого большого основания (или в концах самого ребра, если это ребро самой мелкой плитки). Поскольку мы будем говорить о путях, ребро имеет направление прохождения и упоминая ребро, мы можем сказать какой "босс"остался позади, а какой – босс впереди.

E.9 На картинке с четырьмя выполнеными подстановками укажите типы вершин и ребер с их цветами.

E.10 Какие типы ребер выходят из вершин каждого типа?

Таким образом, мы можем пронумеровать входящие ребра для каждого типа узла. То есть, можно следить откуда путь вошел в вершину и куда вышел. Будем запрещать достаточно короткие пути, которые не встречаются на правильных паттернах, рассматривая каждый путь как конечную последовательность чередующихся вершин и ребер, с запоминанием по какому ребру мы входим в каждую вершину и по какой выходим.

E.11 Какие пути надо объявить запрещенными, чтобы путь по ребрам, составляющим большое основание некоторой плитки, обязательно содержал ровно две вершины C и остальные D, не считая крайних вершин?

E.12 Попробуйте описать набор запрещенных путей (последовательностей) гарантирующих, что из плиток можно будет составить замощение и это будет обязательно правильная мозаика.

Aperiodic Tilings

Thomas Fernique, Ilya Ivanov-Pogodaev, Alexei Kanel-Belov and Ivan Mitrofanov

This project is devoted to an interesting area of mathematics called tilings. We consider some finite set of tiles (figures on the plane) and try to cover the plane by placing our tiles next to each other such that there are no empty spaces between the tiles. The main question is to understand how the local constraints (due to the way tiles can fit next to each other) enforce global properties of tilings. In particular, an interesting global property is non-periodicity, and a tile set which can form only non-periodic tilings is said to be *aperiodic*.

There are a lot of problems separated into several cycles. Some problems are simple, some other more complicated, and some are even open questions: it would be great if someone make an advancement for these problems.

A One dimension

Let us consider the one dimensional case. *Tiles* can be seen as *letter* over a finite alphabet, and *tilings* as *two-side infinite words*. We call *pattern* of a word any of its finite subwords. Letter are allowed to stay next to each other or not, and a common way to specify such constraints is to give a finite set of forbidden patterns. For example, if aa and bb are forbidden, then the only infinite word that can be formed is periodic with period (ab).

A.1 Let us consider the set S of infinite words in the alphabet $\{a, b\}$ which contain runs of b's of length at most three. Does there exist a finite set of forbidden words determining the S?

A.2 Count the minimal number of forbidden words to determine the following infinite words (given by its periods): (ab), (aab), $(aabaabab)$, $(aabaababaabaababaabab)$.

A.3 Consider the set of infinite words over the alphabet $\{a, b\}$ where the patterns (subword) ba^nb are forbidden, for any $n > 1$. Can you find a finite set of forbidden patterns which defines the same set of words?

A.4 What about the previous question if, in addition, you are now allowed to color letters (using finitely many different colors), that is, for example, to make a difference between a blue a and a green a?

A.5 Let S be the set of infinite words whose finite runs of a's are all of even length only. Is it possible to determine the S with some finite number of forbidden words? And if we can color letters using finitely many different colors?

A.6 Same questions if the length of finite runs of a's is asked to be odd.

A.7 Let $u_0 = a$, $u_1 = ab$, $u_{n+2} = u_n u_{n+1}$. Count the minimal number of necessary forbidden words to enforce infinite periodical words with period u_n.

B Two dimensions

A *2D infinite word* over a given alphabet is obtained by writing a letter of the alphabet in each cell of the 2D grid, and a *pattern* is a restriction of such a word to a finite region.

B.1 Find a small set of forbidden patterns which defines 2D infinite words where the a's and b's alternate as the black and white squares on a checkerboard.

B.2 Consider the set of 2D infinite words over the alphabet $\{a, b\}$ such that any finite connected (by side) component of a's has even size. Can you find a finite set of forbidden patterns which defines this set? And if colors are allowed?

B.3 Same question if the size of the connected component of a's is asked to be odd.

B.4 Show that if a finite set of forbidden words allow to form patterns which cover arbitrarily large disk, then one can form a 2D word without any such forbidden words.

Given finitely many polygons called *tiles*, a *tiling* is a covering of the plane by isometric copies of these tiles that can intersect only on vertices or along whole edges. Tiles can in addition be colored using finitely many different colors. 2D infinite words are thus tilings: tiles are square with a color corresponding to the letter a or b. Convince yourself that colors and forbidden patterns are equivalent to notching of tiles or decorations that must match where tiles intersect.

B.5 Consider the hexagonal tiles whose side can have a bump or a dent (see, for example, Fig. 1): for which tiles can we form a tiling using only one of these tiles?

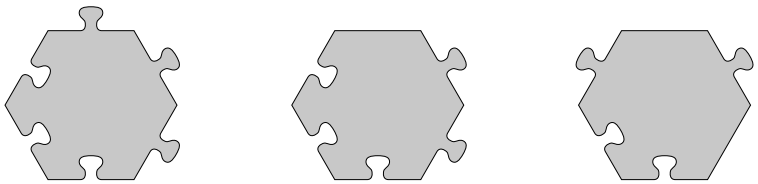


Figure 1: Some notched regular hexagons.

C Periodicity and Quasiperiodicity

Let us return to one dimensional situation. We shall here focus on the question of periodicity, and introduce a much larger notion, *quasiperiodicity*.

C.1 Consider a periodic infinite word. Prove that it can be determined by a finite set of forbidden words.

C.2 Show that any finite set of forbidden words which allows at least one infinite word also allows a periodic word (even if finitely many colors are allowed).

Hence, periodicity is unavoidable in dimension one. But how far can we go?

C.3 Assume that one can use only n forbidden words. Let p be the smallest period of periodical words. How large p could be?

C.4 Now we can use any number of forbidden words, but each contains at most n letters. How large p could be?

An infinite word is said to be *quasiperiodic* if, for any pattern w that appears somewhere in this word, there is a number r such that w appears at distance at most r from any letter of the word.

C.5 Show that any periodic word is quasiperiodic.

C.6 Find a non-periodic but quasiperiodic word, and a non-quasiperiodic word.

Consider now the two-dimensional case. Periodicity and quasiperiodicity can be easily generalized (do it). As we will later see, periodicity is however no more unavoidable! But quasiperiodicity does:

C.7 Show that any finite set of forbidden patterns which allows at least one tiling also allows a quasiperiodic tiling.

D Robinson tilings

The first aperiodic tile set has been discovered in 1964 by Robert Berger and contains 20426 tiles. The much simpler tile set on Fig. 2 have been discovered in 1971 by Rafael Robinson. Bumps or dents (which can have a symmetry or not) enforce the way tiles can be assembled, while the segments drawn on tiles form an interesting picture.

D.1 Show that the tiles on Fig. 2 can form a tiling of the plane.

D.2 Show that no tiling by the tiles on Fig. 2 is periodic.

D.3 When segments drawn on tiles form a square which intersects a smaller square, then the former is called the *parent* of the latter. The *lineage* of a square is then be encoded by an infinite word over a four-letter alphabet, with each letter coding the corner which is intersected by the parent square. At which condition any two squares of a tiling have a common ancestor?

D.4 Show that there are uncountably many different tilings, even up to an isometry.

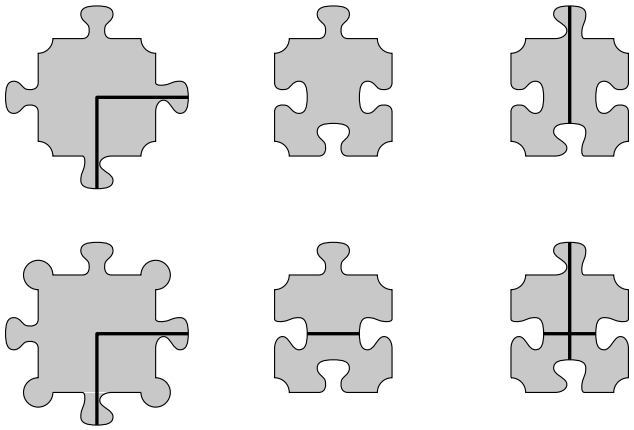


Figure 2: Six notched squares (note the two different type of bump/dent).

E Hierarchical tilings

We shall introduce *hierarchical tilings*, which provide the first general way to find aperiodic tile set, that is, finite tile sets forming only non-periodic tilings.

E.1 Show that the tiles on Fig. 3, discovered in 1974, do tile the plane

E.2 Same question for the tiles on Fig. 4, discovered in the early 90's.

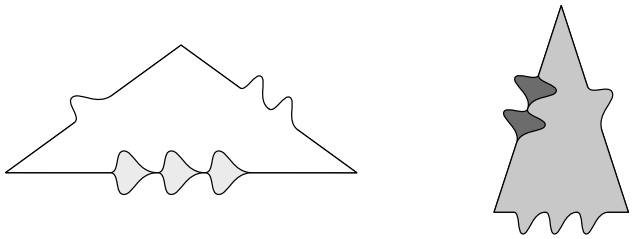


Figure 3: Notched isosceles triangles. Top angles are respectively 108° and 36°.

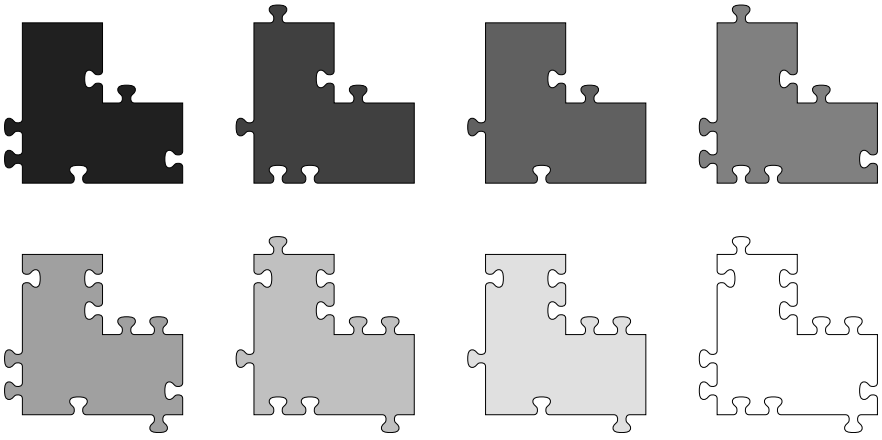


Figure 4: Notched L-shaped tiles.

E.3 If we cannot use colorings or notches on edges, can we nevertheless characterize the tilings of two previous questions characterized by finitely many of their patterns?

A *substitution* is a map which first inflates (by a common factor) the tiles of a given tile set and then subdivide them in non-overlaping tiles of this tile set (Fig. 5 This is thus a map on the tilings by this tile set. With a given substitution are then associated so-called *hierarchical tilings*: they are the tilings which admit a uniquely defined infinite sequence of preimages by this substitution.

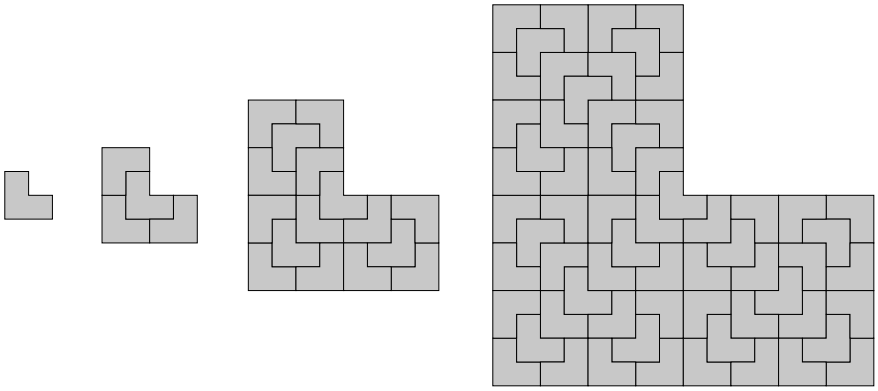


Figure 5: Example of substitution.

E.4 Show that hierarchical tilings are non periodic.

E.5 Are the hierarchical tilings associated with the substitutions on Fig. 6 characterized by finitely many patterns?

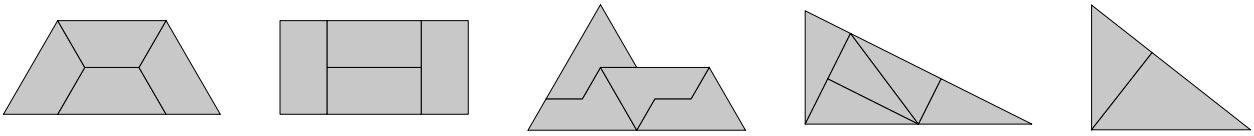


Figure 6: Substitutions (rightmost: tiles are homothetic but not isometric).

A general result discovered in 1998 states that, **given a substitution over a set of tiles, these tiles can be colored so that the hierarchical tilings associated with are characterized by finitely many patterns** (equivalently, tiles can be notched or decorated).

E.6 Can you find such decorations for the hierarchical tilings associated with the substitutions the previous question?

F Self-assembly

F.1 A pattern is said to be *dead* if it appears in no tiling of the whole plane. Find a dead pattern by the tile sets depicted on Fig. 2–4.

F.2 Show that any finite tile set which can form only non-periodic tilings can also form dead patterns.

A *self-assembly scheme* for a set of decorated polygonal tiles consists in giving a *weight* to each tile edge and a global parameter called *temperature*. Then, to form a pattern or a tiling of the whole plane, one first put a tile, then add other tiles one at a time so that, when a tile is added, the sum of the weigths of its edges whoses decorations match is greater or equal to the temperature. Process stops when no tile could be further added.

F.3 Which patterns do form the self-assembly scheme depicted on Fig. 7?

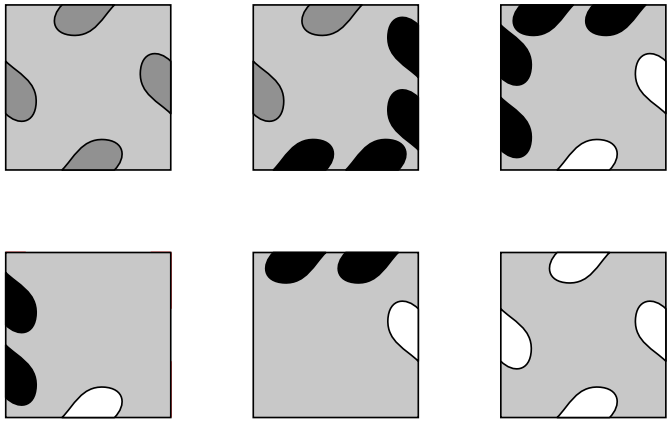


Figure 7: Self-assembly scheme at temperature 2, with the weight of an edge being the number of marks.

A general result states that, given a substitution over a tile set, there is a self-assembly scheme for this tile set such that tilings that can be assembled are exactly the hierarchical tilings associated with the substitution.

F.4 Can you find a self-assembly scheme for substitutions of Sections D or E?

G Alternating rhombi

G.1 Show that one can tile the plane with the notched square and rhombus of Fig. 8.

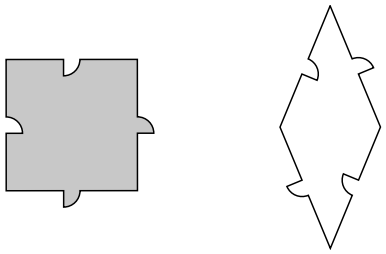


Figure 8: A grey square and a white rhombus (acute angle 45°).

G.2 Which tiling can you form that use as few as possible grey tiles?

G.3 Is it true that any tiling of the plane by the tiles of Fig. 8 is quasiperiodic?

G.4 Same questions with the two rhombi depicted on Fig. 9.

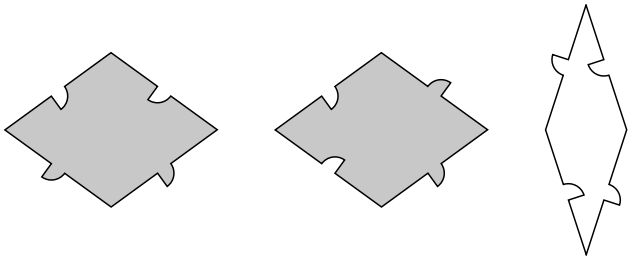


Figure 9: Two grey rhombi (acute angle 72°) and one white (acute angle 36°).

Aperiodic Tilings

Thomas Fernique, Ilya Ivanov-Pogodaev, Alexei Kanel-Belov and Ivan Mitrofanov

Additional hints for the construction of hierarchial tiling decorations.

Let us consider the trapezoid substitution in more detail (Fig 1). We can see that four tiles form a macrotile, four macrotiles form a higher level macrotile and so on. We want to assign some local rules to enforce that tiles can only form hierarchical tilings associated with the trapezoid substitution We will use another method for local rules setting.

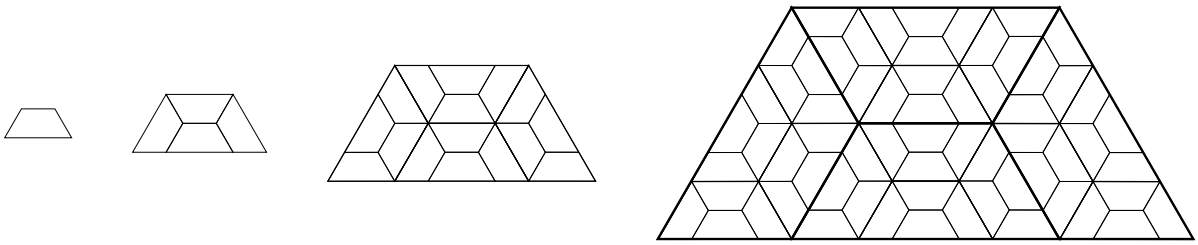


Рис. 1:

Let us set up a language that encodes the different tile types into the vertices and edges. We say that a tiling of the plane is *correct* if it hierarchical for the trapezoid substitution. We say that a pattern is *correct* if it appears in correct tiling. We assign a code for our substitution system and from the other side we prove (by rank induction) that our code enforce the correct patterns.

E.7 Show that every vertex on a correct pattern has one of the three types on Fig. 2. (Trapezoids around the vertices can be macrotiles of any level.) Show that every macrotile edge on a correct pattern has one of the five types on Fig. 3. Show that every macrotile edge is the base-edge of two (macro)tiles.

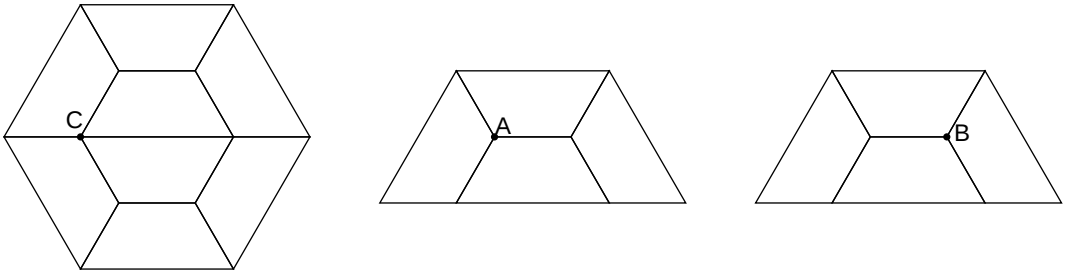


Рис. 2:

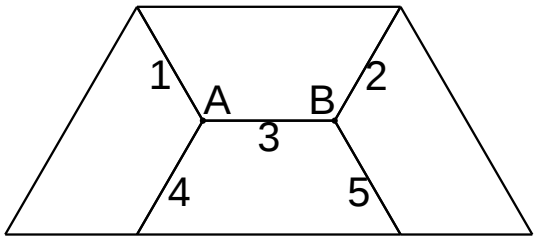


Рис. 3:

How can one code tile edges? Every edge has the same type as the maximal macrotile edge that can contain it. So, every tile has four vertices and four edges. The number of vertices (and edges) types is finite. The idea of forbidden words could be used here again – we can consider paths on a tiling encoded by sequences of vertices and edges types and construct lists of forbidden paths.

E.8 Find an example of sequence having the form XYZ, (there X, Z –are types of edges, Y is a type of vertex) which could not code any path on a correct pattern.

Our goal is to use forbidden paths to ensure that the only possible tilings are correct ones. But we need to work a bit.

Let us add a new type of vertex, D. This type codes vertices on a large base-edge that are part of the next lower hierarchy level of the tiling (Fig. 4). The vertices on this base-edge which are part of further lower levels of the hierarchy are of type C.

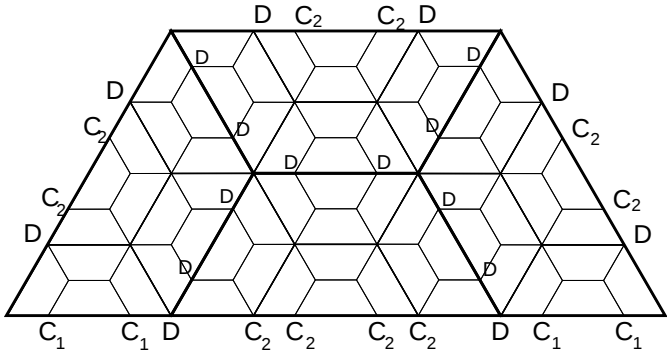


Рис. 4:

Also we assign new colors to C vertices. For every such vertex, we can find the tile on which base it is located We can say that the vertices located between an edge of the base and a D vertex are colored in a first color, while vertices between D vertices (they could be located far from them) are colored in a second color.

We also need additional colors for edges. Every edge is a side of a smallest tile or is situated on a big base of some macrotile. Let X,Y be types of vertices (with colors) in the ends of that big base (or the ends of the smallest tile side, if our edge is the side of smallest tile). We assign an ordered pair of "bosses" (X, Y) as the color of our edge. While we speak about paths, we have a direction on each edge, so we known which are the previous and the next bosses.

E.9 Look at the picture with four levels of substitutions and mark types of all vertices and edges with their colors.

E.10 Which types of edges start at vertices of each type?

So, we can number incoming edges for any type of vertex. Thus, we can know where are the incoming path from and where it goes. We will forbid paths which do not appear on a correct tiling by looking each path as a finite sequence of traveled vertices and edges and by memorizing through which edge we enter and exit each vertex.

E.11 Find the list of forbidden paths to obtain the following property: every path along the edges that form the big base of some macrotile contains exactly two vertices of type D (except the ends of that path).

E.12 Try to construct a finite list of forbidden paths to obtain the following properties: we can construct a tiling using our tiles, and every such tiling is correct.

A7. Слово с периодом (a) можно задать запретом слова b, при этом меньше чем одним запретом обойтись, очевидно, не удастся.

Для всех конечных слов v, состоящих из букв a и b, определим $\varphi(v)$ по следующему правилу: $\varphi(a) = ab$, $\varphi(b) = a$, $\varphi(a_1a_2\dots a_n) = \varphi(a_1)\varphi(a_2)\dots\varphi(a_n)$. Индукцией по k несложно доказать, что $u_{k+1} = \varphi(u'_k)$, где u'_k – циклический сдвиг слова u_k на одну букву.

Лемма. Пусть период $(v) \neq (b)$ задаётся k запретами и не задаётся $k-1$ запретом. Тогда период $(\varphi(v))$ задаётся $k+1$ запретами и не задаётся k.

□ 1. Пусть период (v) задаётся минимальной системой запрещённых слов $v_1, v_2, \dots, v_k$.

Построим систему запретов для $\varphi(v)$. Во-первых, включим туда bb. Во-вторых, каждое слово вида v_i представляется в виде w_ia или w_ib. В первом случае включим в систему запретов $\varphi(w_i)ab$, а во втором случае – $\varphi(w_i)aa$.

Слова, не содержащие bb – это в точности те слова, которые представляются в виде $\varphi(x)$ для некоторого бесконечного x. Заметим, что x можно определить однозначно.

Слова, не содержащие bb и $\varphi(w_i)ab$ – это в точности все слова, представляющиеся в виде $\varphi(x)$ при условии, что x не содержит w_ia.

Слова, не содержащие bb и $\varphi(w_i)aa$ – это в точности все слова, которые представляются в виде $\varphi(x)$ при условии, что x не содержит w_ib.

Значит, слова, разрешаемые новой системой запретов – в точности те, которые представляются в виде $\varphi(x)$, где x не содержит слова из набора v_i, то есть имеют вид $(\varphi(v))$.

2. Покажем, что потребуется хотя бы k запретов. Пусть есть некоторая минимальная система запретов, задающая $(\varphi(v))$. В этой системе должно быть слово вида b^n. Если $n > 2$, то система не минимальная, так как это слово можно заменить на bb (так как bb не встречается во всём слове, и здесь мы используем, что запрета b нет).

Тогда слова, запрещаемые словом bb – в точности те, которые представляются в виде $\varphi(x)$.

Никакой из оставшихся запретов не может оканчиваться на ba, иначе у него можно отбросить последнюю букву.

Если какой-то из запретов *начинается* на b, то перед ним можно написать a – получится не менее сильная система запретов (на этом шаге мы отказываемся от минимальности). Тогда каждый из них представляется одним из двух видов: $\varphi(w)aa$ или $\varphi(w)ab$.

Повторяем рассуждения: слова, не содержащие bb и $\varphi(w)ab$ – это в точности все слова, представляющиеся в виде $\varphi(x)$ при условии, что x не содержит wa.

Слова, не содержащие bb и $\varphi(w)aa$ – это в точности все слова, которые представляются в виде $\varphi(x)$ при условии, что x не содержит wb.

Для каждого из запрещающих слов (кроме bb) строится некоторое слово v_i, и известно, что слова, избегающие запрещающих – это в точности те слова, которые представляются в виде $\varphi(x)$, где x не содержит слов из набора v_i. С другой стороны, это по условию все те слова, которые представляются

в виде $\varphi(x)$, где x периодично с периодом v. Значит, набор $\{v_i\}$ однозначно задаёт слова с периодом v, то есть их не менее $k+1$, а всего не менее $k+1$.

□

Из леммы и наблюдения, что слова $\varphi(u_k)$ и u_{k+1} являются циклическими сдвигами друг друга, следует ответ $k+1$.

C4. Докажем, что длина периода не может быть более 2^{k-1}. Рассмотрим в бесконечном слове $\dots a_{-1}a_1a_2\dots$ $2^{k-1}+1$ подслов длины $k-1$, первые буквы которых имеют номера 0, 1, 2, ..., 2^{k-1}. По принципу Дирихле среди них есть два одинаковых. Пусть это слова, начинающиеся с позиций i и j. Тогда слово $(a_ia_{i+1}\dots a_{j-1})$ имеет период не больший, чем 2^{n-1}. Так как все его слова длины k являются подсловами исходного, то оно не запрещено.

Докажем, что существует периодичное слова с длиной периода ровно 2^{k-1} такое, что в его периоде встречаются все возможные подслова длины $k-1$ по одному разу. Тогда в качестве запрещённых возьмём все не встречающиеся в периоде слова длины k, и эти слова будут определять период слова, так как мы куску длины $k-1$ всегда можно будет определить следующую букву.

Итак, пример. Рассмотрим ориентированный граф на 2^{k-2} вершине. Его вершины – это слова длины $k-2$, они соединены стрелкой, если пересекаются по слову длины $k-3$. Пример: $k=7$, из вершины *abbab* ведут рёбра в *bbaba* и *bbabb*.

Этот граф сильносвязен и исходящие степени всех вершин одинаковы. Следовательно, существует циклический обход его рёбер (Эйлеров обход). Рассмотрим бесконечный путь, повторяющий этот обход. Этому пути соответствует искомое слово с длиной периода 2^{k-1}.

A7. The word (a) can be defined by forbidden word b. It is obvious we can not use empty set of forbidden words.

For a finite word v, if it consists of letters a and b, define $\varphi(v)$ according the rule

$$\varphi(a)=ab;\varphi(b)=a;\varphi(a_1a_2\dots a_n)=\varphi(a_1)\varphi(a_2)\dots\varphi(a_n)$$

. We can prove (induction on k) that $u_{k+1}=\varphi(u'_k)$ if u'_k is the one-letter cyclic shift of u_k.

Lemma. Suppose $(v)\neq(b)$, this word can be defined by k forbidden words and can not be defined by $k-1$. Then the period $(\varphi(v))$ can be defined by $k+1$ f.w. and can not be defined by k.

□ 1. Suppose $v_1,v_2,\dots,v_k$ is the minimal system of f.w. that defines (v).

Construct system od forbidden words for $(\varphi(v))$.

The first word will be *bb*. For each v_i we take a forbiden word as follows: if v_i has form $v_i=w_ia$, we add $\varphi(w_i)ab$. In other case (i.e. $v_i=w_ib$) we take $\varphi(w_i)aa$ as a forbidden word.

The set of words y that do not include *bb* is exactly the set of words that can be expressed as $y=\varphi(x)$ for some infinite x. Notice, that x is defined unique by y.

Words without subwords *bb* and $\varphi(w_i)ab$ are all words of form $\varphi(x)$ with condition "x has not the factor w_ia".

Words without subwords bb and $\varphi(w_i)aa$ are all words of form $\varphi(x)$ with condition "x has not the factor w_ib".

So, the set of words that are not forbidden is exactly the set of words of form $\varphi(x)$ for x that has not words v_i as subwords. And such words have period $(\varphi(v))$.

2. Now we'll prove we need k forbiden words. Let set $\{v_i\}$ be a set of forbiden words that defines $(\varphi(v))$. This set forbids (b) so it includes a word b^n for some n. Case $n = 1$ is not interesting. If there is forbiden word b^n for $n > 2$ then we can change this ford by bb, and the system will forbid the same set of infinite words.

The set of words y that do not include bb is exactly the set of words that can be expressed as $y = \varphi(x)$ for some infinite x.

Suppose the set contains a forbiden word u and u ends by ba. Then we delete the last letter of v_i, this this operation will not change the set of allowed infinite words (because we also have bb).

If some v_i has its first letter b, we change v_i by av_i.

After these operations we get the same number of words, and they forbid the same set of infinite words. Each of v_i has form $\varphi(w_i)aa$ either $\varphi(w_i)ab$ for some w_i.

As before: words without subwords bb and $\varphi(v_i)aa$ are all words of form $\varphi(x)$ with condition "x has not the factor w_ib".

Words without subwords bb and $\varphi(v_i)ab$ are all words of form $\varphi(x)$ with condition "x has not the factor w_ia".

Each of $\{v_i\}$ (besides bb) defines a word w_i. Infinite words without $\{v_i\}$ are exactly the words of form $\varphi(x)$, where x has not any w_i as a factor. From the other hand this is the set of periodic words with period $(\varphi(v))$. It means that the set $\{w_i\}$ defines the period v and there are at least k of them, and there are al least $k + 1$ words in $\{v_i\}$. $\square$

This lemma and the fact before the lemma imply that the answer is $k + 1$.

C4. At first we prove the period can not be larger then 2^{k-1}. Conside in a (not forbiden) infinite word $\ldots a_{-1}a_1a_2\ldots$ a set of $2^{k-1} + 1$ subwords of length $k - 1$, (these words start at positions 0, 1, 2, ..., 2^{k-1}). Dirichlet principle: there are two equal among them. Suppose these two words start at positions i and j. Then the word $(a_ia_{i+1}\ldots a_{j-1})$ is periodic with period equal or less then чем 2^{n-1}. Since all its factors of length k are factors of $\ldots a_{-1}a_1a_2\ldots$, it is not forbiden.

Now se prove there exists a periodic word with period of length 2^{k-1} that contains all possible subwords of length $k-1$ (one occurence each). Then forbiden words of lenght k are all words that are not in period. These words will define the period, because every pattern with length at least $k - 1$ is prolongable right uniqe way.

How we construct such a word? Consider oriented graph on 2^{k-2} vertices and 2^{k-1} edges. Vertices are words of length $k - 2$, two words are connected with an arrow iff they overlap by a word of length $k - 3$. Example: $k = 7$, vertex $abbab$ has outdegree 2 and edges from it go to $bbaba$ and $bbabb$.

This graph is strictly connected and all outdegrees and indegrees are equal to two. It follows the existence of a cyclic path that goes each edge once (Eurler circuit). We turn this path to a periodic one repeating is infinitly manu times. Last letters of vertices on this path form a periodic word we want.

Модел. и анализ информ. систем. Т.**14**, №**2** (2007) 12–16

УДК 512.552.4+519.115.1

О числе запретов, задающих периодическую последовательность

Челноков Г.Р. [1]
Ярославский государственный университет,
150 000, Ярославль, Советская, 14,

получена 29 апреля 2007

Аннотация

Рассматриваются последовательности W периода u над алфавитом из l букв. Требуется однозначно определить последовательность W, указывая слова, не являющиеся ее подсловами. Для $n \in \mathbb{N}$ обозначим за U_n множество слов u длины n, не являющихся степенями (т.е. не представимых в виде $u = v^k$, $k > 1$). Пусть $T(u^\infty)$ — минимальное число запретов, задающих последовательность u^∞. Обозначим

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

Доказаны следующие теоремы:

Теорема 1. $m_n \le n(l-1)$.

Отметим, что оценка точна при бесконечно многих n и реализуется, например, для периода, содержащего все слова некоторой фиксированной длины t (т.е. $n = l^t$).

Теорема 2. $r_n \ge \log_2 n + 1$.

Теорема 3. *Существует возрастающая последовательность n_i, такая, что*

$$r_{n_i} \le \log_\phi n_i, \quad \textit{где} \quad \phi = \frac{1+\sqrt{5}}{2}.$$

1. Введение

Исследование комбинаторных свойств периодических последовательностей играет важную роль в проблемах бернсайдовского типа. Соответствующие вопросы исследовались рядом авторов, см. [1-3].

При изучении мономиальных алгебр (алгебр, заданных соотношениями вида: моном равен 0) важную роль играют алгебры A_u, заданные соотношениями $v = 0$, где v — не подслово u^∞, см. [3]. Это — первичная конечно определенная PI-алгебра, и все первичные конечно определенные мономиальные алгебры имеют такой вид. Можно показать, что все определяющие соотношения алгебры A_u имеют вид $v = 0$, где $|v| \le |u| - 1$.

Представляет интерес более точное исследование структуры соотношений, задающих A_u. Этому и посвящена настоящая работа.

2. Основные результаты

Далее все рассматриваемые слова есть над фиксированным алфавитом A из l букв.

Ниже некоторые множества слов будут называться *системами запретов*. Определим понятия, связанные с системой запретов.

Пусть дана некоторая система запретов $V = \{v_i\}$. Будем говорить, что w *удовлетворяет системе запретов* V, если слово w не содержит ни одно из v_i в качестве подслова; *система запретов определяет бесконечное в обе стороны слово* w, если w есть единственное бесконечное слово, удовлетворяющее системе запретов v_i.

Легко показать, что только периодические слова могут быть определены конечной системой запретов, см. [1].

Под *минимальной* системой запретов, определяющей слово w, ниже будет пониматься минимальная по количеству слов система.

Рассмотрим слово u, не содержащее ни один из запретов в качестве подслова. Слово uX, где $X \in A$, называется *продолжением вправо* слова u, если uX также не содержит ни один из запретов в качестве подслова. Слово u называется *однозначно продолжаемым вправо*, если u имеет ровно одно продолжение

[1] Исследование выполнено при поддержке РФФИ, грант №06-01-00648.

вправо, *неоднозначно продолжаемым*, если имеет больше одного продолжения, и *непродолжаемым*, если не имеет ни одного продолжения. Аналогичным образом определяются продолжения слова на любое количество букв, а также бесконечные продолжения. Продолжения влево определяются тоже аналогично.

Слово u называется *началом* слова w, если слово w представимо в виде us, *собственным началом*, если слово s непусто. Аналогично определяется *конец* и *собственный конец*.

Для $n \in \mathbb{N}$ обозначим за U_n множество слов u длины n, не являющихся степенями (т.е. не представимых в виде $u = v^k$, $k > 1$). Очевидно, что эти и только эти слова являются периодами последовательностей, наименьший период которых есть n. Обозначим через $T(u^\infty)$ минимальное число запретов, задающих последовательность u^∞. Будем обозначать

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

Рассмотрим бесконечное в обе стороны слово w. Множество всех слов v_i, таких, что v_i не есть подслово w, но любое собственное подслово v_i является подсловом w, будем называть *канонической системой запретов* и обозначать $C(w)$. Легко показать, что $C(u^\infty)$ конечна, см. [1].

Лемма 1. *Одна из минимальных систем запретов, задающих слово u^∞, есть $C(u^\infty)$.*

Доказательство. Пусть дана система запретов V, задающая слово u^∞. Каждый запрет является невстречающимся подсловом. Если есть запреты, содержащие невстречающиеся подслова в качестве собственных подслов, то заменим каждый из таких запретов на его минимальное по включению невстречающееся подслово, и выкинем из системы повторяющиеся запреты, если они появились. Полученная система V_1 не слабее исходной. Теперь докажем, что каждое минимальное невстречающееся подслово входит в систему V_1 в качестве запрета. Пусть слово v является минимальным невстречающимся в u^∞, но не входит в V_1. Возможны 2 случая.

1. Пусть существует бесконечное влево слово s и бесконечное вправо слово t, такие, что слова sv и vt не содержат запретов из V_1. Тогда рассмотрим бесконечное в обе стороны слово svt. Любое его подслово есть или подслово одного из слов sv и vt, и тогда не может быть запретом V_1 по предположению, или содержит подсловом слово v, и тогда не может быть запретом из V_1, ибо тогда этот запрет или не минимален, или совпадает с v. Итак, в этом случае существует удовлетворяющее системе запретов бесконечное в обе стороны слово svt, следовательно, система запретов V_1 не задает слово u^∞.

2. Пусть, без ограничения общности, не существует бесконечного вправо слова t, такого, что vt не содержит запретов. Пусть первая буква слова v есть X, обозначим $v = Xs$. Заметим, что s есть подслово в u^∞ из минимальности v, следовательно, s допускает бесконечное продолжение вправо st (такое, как в u^∞). Слово vt содержит запрет $v_i \in V_1$, следовательно, этот запрет содержит первую букву слова v, тогда v_i или содержит слово v, что противоречит минимальности v_i, или содержится в слове v, что противоречит минимальности слова v. □

Теорема 1. $m_n \le n(l-1)$.

Доказательство. Рассмотрим слово u^∞, минимальный период которого n. Если у произвольного запрета $v_i \in C(u^\infty)$ отрезать последнюю букву, он станет подсловом u^∞. Таким образом, каждому запрету соответствует пара из буквы алфавита (являющейся последней буквой этого запрета) и места в периоде, на которое попадает правый конец этого запрета без последней буквы. Заметим, что если бы двум запретам соответствовала одна пара, то один из них был бы подсловом другого, значит, запретов не больше, чем пар. Всего позиций в периоде n, для каждой запрещено может быть не больше $l-1$ продолжений, значит, всего запретов не больше $n(l-1)$.

□

В качестве примера u, для которого оценка становится точной, построим такое u, что при некотором фиксированном натуральном k каждое слово длины k встречается ровно 1 раз на периоде u^∞. Тогда период этого слова имеет длину l^k. Минимальных невстречающихся слов есть ровно $l^k(l-1)$, так как для каждого слова длины k есть его единственное продолжение, а все остальные продолжения не встречаются. Но все эти слова длины $k+1$ есть минимальные невстречающиеся, ибо каждое слово длины k встречается.

Построим такое u. Рассмотрим ориентированный граф, вершины которого суть слова длины $k-1$ над A, то есть количество вершин l^{k-1}. Из вершины X в вершину Y ведет стрелка, если последние $k-2$ буквы слова X есть первые $k-2$ буквы слова Y. Таким образом, стрелки этого графа биективно соответствуют словам длины k. Будем говорить, что на каждой стрелке написана буква — последняя буква слова, соответствующего вершине, в которую входит эта стрелка. Входящая степень каждой вершины в этом графе равна выходящей, ибо обе равны l, кроме того, граф, очевидно, связен. Тогда в нем есть эйлеров цикл (см., например, [4]). Последовательность букв, соответствующих стрелкам цикла, и есть искомое слово u.

Теорема 2. $r_n \geq \log_2 n + 1$.

Доказательство.
Докажем сначала несколько лемм.

Лемма 2. *Для натуральных чисел* $k_1, \ldots, k_i$ *из* $k_1 k_2 \cdots k_i \geq n$ *следует*

$$(k_1 - 1) + \cdots + (k_i - 1) \geq \log_2 n.$$

Доказательство. Предположим противное и рассмотрим контрпример с минимальным значением суммы $\max(0; k_1 - 2) + \max(0; k_2 - 2) + \cdots + \max(0; k_i - 2)$. Если все k_j для $j \in [1; i]$ равны 1 или 2 — утверждение очевидно. Если некоторое $k_j \geq 3$, то заменим его на $k_j - 1$ и добавим $k_{i+1} = 2$. Произведение $k_1, \ldots, k_{i+1}$ увеличилось, сумма $(k_1 - 1) + \cdots + (k_{i+1} - 1)$ не изменилась, значение суммы $\max(0; k_1 - 2) + \max(0; k_2 - 2) + \cdots + \max(0; k_{i+1} - 2)$ уменьшилось, что противоречит минимальности контрпримера. □

Рассмотрим бесконечное в обе стороны слово w. Слово v будем называть *развилкой*, если оно относительно системы запретов $C(w)$ неоднозначно продолжается как влево, так и вправо. Количество продолжений слова v вправо назовем *правой кратностью* или просто *кратностью* развилки v.

Два слова, ни одно из которых не является подсловом другого, назовем *несравнимыми*.

Лемма 3. *Для каждого слова* v, *являющегося подсловом* u^∞, *существует наименьшая развилка* w, *содержащая* v *в качестве подслова, причем единственного, если* w *не есть* u^∞.

Доказательство. Пусть есть две несравнимые развилки w_1 и w_2, содержащие v, обозначим $w_1 = s_1 v t_1$ и $w_2 = s_2 v t_2$. Пусть s — наибольший общий конец слов s_1 и s_2, а t — наибольшее общее начало слов t_1 и t_2. Тогда $w = svt$ — развилка, меньшая w_1 и w_2. В самом деле, t или есть собственное начало t_1 и t_2, и тогда svt продолжается вправо минимум двумя способами: так, как оно продолжается в w_1, и так, как оно продолжается в w_2; или t совпадает, не ограничивая общности, с t_1, и тогда svt продолжается вправо неоднозначно, потому что w_1 неоднозначно продолжается вправо, а svt есть конец w_1. Аналогично svt неоднозначно продолжается влево. □

Лемма 4. $k_1 k_2 \cdots k_i \geq n$.

Пусть все развилки в u^∞ занумерованы $v_1, \ldots, v_i$, их кратности $k_1, \ldots, k_i$, кроме того, для каждой развилки v_j пронумерованы все ее продолжения вправо x_{jm} $m \in [1; k_j]$ (так, к примеру, нумерация продолжений пустой развилки есть просто нумерация всех букв, встречающихся в u). Рассмотрим некоторый циклический сдвиг слова u, то есть подслово w слова u^∞, такое, что $|w| = |u|$. Будем считать пустую развилку v_1 началом слова w. Посмотрим, продолжением с каким номером для развилки v_1 является первая буква слова w, обозначим номер за x_1. По лемме 3 для первой буквы однозначно определена минимальная развилка v_j, содержащая ее, то есть имеем $v_j = s_j t_j$, $w = t_j r_j$. Пусть первая буква слова r_j в списке продолжений развилки v_j имеет номер x_j, тогда по лемме 3 однозначно определена минимальная развилка, содержащая слово $v_j x_j$. Будем повторять этот процесс, пока минимальной развилкой не станет u^∞. Таким образом, для некоторых номеров $j \in J \subset [1; i]$ мы определили соответствующие им продолжения x_j, доопределим произвольным образом x_j, $j \notin J$. Из алгоритма построения очевидно, что выбором x_j однозначно определен сдвиг w слова u, а поскольку различных сдвигов n, лемма доказана. □

Рассмотрим дерево (связный граф без циклов), некоторую его вершину назовем корнем, ориентируем все ребра в направлении от корня (поскольку граф без циклов, направление определено корректно). Такой граф будем называть *ориентированным деревом*.

Лемма 5. *Количество запретов не меньше* $(k_1 - 1) + \cdots + (k_i - 1) + 1$.

Доказательство. Построим ориентированное дерево. Множество его вершин есть объединение множеств развилок и некоторых (возможно, не всех) запретов канонической системы, а исходящая степень каждой развилки в этом дереве есть в точности кратность этой развилки; вершины, соответствующие запретам — тупиковые.

Для построения дерева рассмотрим развилку v_j, $j \in [1; i]$ и ее продолжение x_m, $m \in [1; k_j]$, а также два различных продолжения влево y_1 и y_2 развилки v_j. Если слово $v_j x_m$ однозначно продолжается влево, то одно из слов $y_1 v_j x_m$ и $y_2 v_j x_m$ не встречается. Пусть без ограничения общности это $y_1 v_j x_m$. Тогда, поскольку и $y_1 v_j$, и $v_j x_m$ встречаются, то $y_1 v_j x_m$ — запрет. Каждый построенный таким образом запрет соответствует только одной паре (v_j, x_m). Если слово $v_j x_m$ неоднозначно продолжается влево, то минимальная развилка v_h, содержащая слово $v_j x_m$, имеет его своим началом.

Следовательно, v_h однозначно определяется парой (v_j, x_m). В самом деле, пусть при $v_{j_1} \neq v_j$ мы построили ту же развилку v_h, тогда из двух слов v_j и v_{j_1} одно является началом другого. Пусть v_j, тогда и $v_j x_m$ является началом v_{j_1}, что противоречит минимальности v_h.

Проведем из каждой развилки стрелки, соответствующие ее продолжениям вправо, в построенные ранее запреты или развилки.

В полученном дереве сумма исходящих степеней есть $k_1 + k_2 + \cdots + k_i$, v_1 имеет входящую степень 0, остальные $i-1$ развилка имеют входящую степень 1, значит, есть

$$k_1 + k_2 + \cdots + k_i - (i-1) = (k_1 - 1) + \cdots + (k_i - 1) + 1$$

висячих вершин, соответствующих запретам. □

Приступим к доказательству теоремы 2. По лемме 4 кратности развилок $k_1, k_2, \ldots, k_i$ слова u^∞ удовлетворяют неравенству $k_1 k_2 \cdots k_i \geq n$. Откуда по лемме 2 имеем

$$(k_1 - 1) + \cdots + (k_i - 1) \geq \log_2 n.$$

А по лемме 5 число запретов не меньше

$$(k_1 - 1) + \cdots + (k_i - 1) + 1 \geq \log_2 n + 1.$$

□

Теорема 3. *Существует возрастающая последовательность n_i такая, что*

$$r_n \leq \log_\phi n_i, \quad \text{где} \quad \phi = \frac{1+\sqrt{5}}{2}.$$

Доказательство. Построим бесконечную серию слов u, для которых $T(u^\infty) \leq \log_\phi n_i$, где $\phi = \frac{1+\sqrt{5}}{2}$. Рассмотрим алфавит из двух букв 0 и 1. Будем строить две вспомогательные последовательности слов s_i и t_i. Обозначим $s_1 = 0$, $t_1 = 1$; $t_{i+1} = s_i t_i$, $s_{i+1} = s_i s_i t_i$ $\forall i \in \mathbb{N}$. Рассмотрим систему запретов $v_{2i-1} = t_1 t_2 \cdots t_{i-1} t_i t_i$ и $v_{2i-1} = t_1 t_2 \cdots t_{i-1} s_i s_i s_i$ при $i \in [1; n-1]$ для некоторого $n \in \mathbb{N}$, и $v_{2n-1} = t_1 t_2 \cdots t_{n-1} t_n t_n$, $v_{2n} = t_1 t_2 \cdots t_{n-1} s_n s_n$.

Докажем, что система запретов $v_1, \ldots, v_{2n}$ задает слово $(t_{n+1})^\infty$.

Лемма 6. *Слово $t_1 t_2 \cdots t_{i-1}$ является концом любого слова, представимого в виде произведения слов s_i и t_i.*

Доказательство. Обоснование леммы проведем индукцией по i. База индукции при $i = 2$ очевидна. Выполним шаг индукции. Пусть слово w представимо в виде произведения слов s_i и t_i. Тогда оно также представимо в виде произведения слов s_{i-1} и t_{i-1}, причем последним словом в произведении будет t_{i-1}. Обозначим $w = w_1 t_{i-1}$, тогда то, что слово $t_1 t_2 \cdots t_{i-1}$ является концом слова w, равносильно тому, что слово $t_1 t_2 \cdots t_{i-2}$ является концом слова w_1 (представимого в виде произведения слов s_{i-1} и t_{i-1}, как отмечалось выше), но это в точности индукционное предположение. □

Лемма 7. *Всякое бесконечное слово, не содержащее запретов $v_1, \ldots, v_n$, есть $(t_{n+1})^\infty$.*

Доказательство. Индукцией по i докажем, что любое бесконечное слово, удовлетворяющее запретам $v_1, \ldots, v_{2i}$, $i < n$, разбивается на слова s_{i+1} и t_{i+1}. База очевидна. Пусть утверждение доказано для $i-1$, докажем его для i. Бесконечное слово w разбивается на слова s_i и t_i по предположению индукции. Докажем, что разбиение не может содержать два слова t_i подряд. Предположив противное, обозначим $w = w_1 t_i t_i w_2$, где w_1 само некоторым образом разбито на слова s_i и t_i, тогда по лемме 6 слово $t_1 t_2 \cdots t_{i-1}$ есть конец w_1, а значит, слово $v_{2i-1} = t_1 t_2 \cdots t_{i-1} t_i t_i$ есть подслово w, что противоречит условию.

Аналогично доказывается, что разбиение слова w на слова s_i и t_i не может содержать три слова s_i подряд. Следовательно, в разбиении слова w слова s_i и t_i можно объединить в слова s_{i+1} и t_{i+1}.

Аналогично доказывается, что бесконечное слово, не содержащее запретов $v_1, \ldots, v_n$, разбивается на слова (t_{n+1}), то есть что оно есть $(t_{n+1})^\infty$. □

Очевидно, что слово $(t_{n+1})^\infty$ удовлетворяет системе запретов $v_1, \ldots, v_{2n}$. Заметим, что слова t_i и s_i являются конечными шагами в итерационном построении равномернорекуррентного слова Штурма (Sturmian prime word, см. [5],[6]), связанного с Bergman's gap, см. [3],[6].

Рассмотрим последовательность f_i, заданную условиями $f_1 = f_2 = 1$, $f_i = f_{i-1} + f_{i-2}$ при $i > 2$. Легко видеть, что $\forall i \quad f_i$ и f_{i+1} взаимнопросты. Поскольку слово t_{n+1} содержит f_{2n-1} букв 1 и f_{2n} букв 0 (это утверждение легко доказывается индукцией по n вместе с утверждением, что s_{n+1} содержит f_{2n} букв 1

и f_{2n+1} букв 0), t_{n+1} не является степенью большей 1 никакого слова, а значит, t_{n+1} есть минимальный период слова $(t_{n+1})^\infty$. Длина периода, таким образом, равна $f_{2n+1} = \frac{1}{\sqrt{5}}(\phi^{2n+1} + \psi^{2n+1})$, где $\phi = \frac{1+\sqrt{5}}{2}$, $\psi = \frac{1-\sqrt{5}}{2}$.

При этом, поскольку последовательность $(t_{n+1})^\infty$ задана $2n$ запретами, $T((t_{n+1})^\infty) \le 2n$.

□

Автор выражает благодарность А. Я. Белову за постановку задачи, И. И. Богданову, за ценные замечания, приведшие к существенному упрощению многих доказательств, а также В. Л. Дольникову за помощь в работе над статьей.

Список литературы

[1] В. А. Уфнаровский, "Комбинаторные и асимптотические методы в алгебре", *Итоги науки и техники. Сер. Совр. пробл. математики. Фундаментальные направления*. Т. 57, ВИНИТИ, М., 1990, 5–177.

[2] А. Г. Курош, "Проблемы теории колец, связанные с проблеммой Бернсайда о периодических группах", *Изв. АН СССР сер. мат.*, **5** (1941), 233–240.

[3] А. Я. Белов, В. В. Борисенко, В. Н. Латышев, "Мономиальные алгебры", *Итоги науки и техники. Сер. Современная математика и ее приложения. Тематические обзоры*. Т. 26, ВИНИТИ, М., 2002, 35–214.

[4] Р. Уилсон, *Введение в теорию графов*, Мир, М., 1977, 208 с.

[5] J.-P. Allouche, J. Shallit, *Automatic sequences. Theory, applications, generalizations*, Cambridge Univercity Press, Cambridge, 2003, 571 pp.

[6] J. P. Bell, "Examples in finite Gel'fand-Kirilov dimension", *J. Algebra*, **263**:1 (2003), 159–175.

On the Number of Restrictions Determining a Periodical Sequence

Chelnokov G.R.

We consider sequences W of the period u over an alphabet consisting of l letters. It is required to determine unambiguously the sequence W picking out words which are not subwords of the sequence. For $n \in \mathbb{N}$ we denote by U_n the set of words u of length n, which are not powers (i.e. are not represented in form $u = v^k$ $k > 1$).

Let $T(u^\infty)$ be the minimal number of restrictions determining the sequence u^∞.

Denote

$$m_n = \max_{u \in U_n} T(u^\infty), \quad r_n = \min_{u \in U_n} T(u^\infty).$$

We prove that

1. $m_n \le n(l-1)$.

The estimate is precise for infinite values of n. For instance, it takes place for a period which contains all the words of some given length t (i.e. $n = l^t$).

2. $r_n \ge \log_2 n + 1$.

3. There exists an increasing sequence n_i so that

$$r_{n_i} \le \log_\phi n_i, \quad \text{where} \quad \phi = \frac{1+\sqrt{5}}{2}.$$

Наибольшая длина периода слова, задаваемого n запретами

И.И. Богданов, Г.Р. Челноков

3 мая 2013 г.

1 Введение

Исследование комбинаторных свойств периодических последовательностей (слов) играет важную роль в проблемах бернсайдовского типа, см., например, [1, 2, 3].

Алгебра $A = F\langle x_1, \ldots, x_n\rangle/I$ называется *мономиальной*, если идеал I свободной алгебры $F\langle x_1, \ldots, x_n\rangle$ порождён мономами. При изучении мономиальных алгебр важную роль играют алгебры вида A_u, где u — непериодичное слово: алгебра A_u задана соотношениями $v = 0$, где v пробегает множество всех слов, не являющихся подсловами в u^∞, см. [3]. Алгебрами A_u исчерпывается класс первичных конечно определенных мономиальных PI-алгебр. В то же время, не все слова v необходимы для задания такой алгебры. Достаточно, например, ограничиться всеми словами длины, не превосходящей длины u.

Представляет интерес более точное исследование структуры соотношений, задающих A_u. В данной работе исследуется вопрос о возможной длине слова u, при котором алгебра A_u может быть задана n мономиальными соотношениями. Мы показываем (см. теорему 2.5), что в случае алфавита из двух букв наибольшая длина слова равен числу Фибоначчи $F(n)$.

Работа является продолжением статьи [4], в которой получены экспоненциальные оценки на длину слова u. Мы используем некоторые понятия и результаты из этой статьи.

Как авторам стало известно, в настоящее время П. Лавров предложил другое доказательство этого факта [5]. Было бы интересно сравнить методы доказательств.

2 Предварительные сведения

Пусть $X = \{x_1, \ldots, x_k\}$ — конечный алфавит (в большей части статьи мы полагаем $k = 2$). Под *конечным (бесконечным вправо/влево/в обе стороны) словом* мы понимаем любую конечную (бесконечную вправо/влево/в обе стороны) последовательность букв алфавита; пустая последовательность Λ также является словом. *Длиной* $|u|$ конечного слова u называется количество букв в нём. Все конечные слова образуют моноид относительно конкатенации.

Определение 2.1. *Слово u называется* подсловом *слова w, если $w = v_1uv_2$ для некоторых слов v_1, v_2. Слово u является* началом (концом) *слова w, если $v_1 = \Lambda$ ($v_2 = \Lambda$). Подслово (начало, конец) u слова v является* собственным, *если $u \neq v$.*

Введём на множестве конечных слов частичный порядок: скажем, что $u \preceq v$, если u является подсловом слова v.

Непустое слово u называется периодическим, *если $u = v^n$ для некоторого слова v и некоторого $n \geq 2$. В противном случае оно называется* непериодическим.

Определение 2.2 ([4])**.** Системой запретов *назовём конечное множество слов $V = \{v_1, \ldots, v_n\}$ в алфавите X. Будем говорить, что (конечное или бесконечное) слово w* удовлетворяет системе запретов V, *если $v \not\preceq w$ для любого $v \in V$.*

[1]The work is supported by the Russian government project 11.G34.31.0053.

Пусть W — бесконечное в обе стороны слово. Будем говорит, что система запретов V определяет *слово W, если W — единственное бесконечное слово, удовлетворяющее этой системе запретов.*

Пусть u — конечное слово. Определим *бесконечное слово с периодом u* как $u^\infty = \dots uuuu \dots$. Если существует такое слово u, что $W = u^\infty$, то бесконечное слово W назовём *периодичным*. Нетрудно видеть, что если система запретов определяет слово W, то оно периодично.

Для каждого периодичного бесконечного в обе стороны слова W существует в определённом смысле оптимальная система запретов. Слово v назовём *каноническим запретом* для W, если v не является подсловом W, а любое его собственное подслово — является. Множество всех канонических запретов для W назовём *канонической системой запретов* для W; она обозначается $C(W)$.

Лемма 2.3 (см. [4, Лемма 1])**.** *Каноническая система запретов $C(W)$ определяет слово W. При этом любая система запретов, задающая W, содержит не меньше элементов, чем $C(W)$.* □

Замечание. Можно показать также, что $C(W)$ — единственная система запретов, задающая W, с минимальной возможной суммой длин входящих в неё слов.

Отметим ещё одно полезное свойство системы $C(W)$.

Предложение 2.4. *Конечное слово v удовлетворяет $C(W)$ тогда и только тогда, когда v — подслово в W.*

Доказательство. Если v — подслово в W, то, очевидно, оно удовлетворяет $C(W)$. Обратно, предположим, что v не является подсловом в W. Тогда существует минимальное по длине подслово $v' \preceq v$, не являющееся подсловом W; оно по определению лежит в $C(W)$. Значит, v не удовлетворяет $C(W)$. □

Определим числа Фибоначчи по следующему правилу: $F(0) = F(1) = 1$, $F(k+1) = F(k) + F(k-1)$. Мы продолжим эту последовательность на отрицательные индексы; так, $F(-1) = 0$, $F(-2) = 1$, $F(-3) = -1$.

Цель данной статьи — нахождение точных верхних оценок на длину периода слова, если известна мощность системы запретов, его задающая. Основным результатом статьи является следующая теорема.

Теорема 2.5. *Пусть $|X| = 2$. Пусть система запретов V определяет слово $W = u^\infty$, где слово u непериодично. Тогда $|u| \le F(|V|)$.*

Замечание. В силу леммы 2.3 можно ограничиться случаем $V = C(W)$.

Определение 2.6. *Пусть слово u удовлетворяет системе запретов V, и $x \in X$. Назовём слово $u' = ux$ ($u' = xu$)* продолжением *слова u* вправо (влево) *относительно V, если u' также удовлетворяет V.*

Слово u назовём неоднозначно продолжимым вправо (влево)*, если у него существуют хотя бы два разных продолжения вправо (влево).*

Наконец, назовём u развилкой *(относительно V), если u неоднозначно продолжимо как вправо, так и влево.* Кратностью *развилки u назовём количество её продолжений вправо.*

Назовём слово u развилкой *относительно бесконечного в обе стороны слова W, если u является развилкой относительно $C(W)$. Само слово W также назовём* развилкой *относительно W.*

Пример. Пусть $|X| = 2$, $X = \{a, b\}$. Тогда конечное слово u является развилкой относительно W тогда и только тогда, когда все четыре слова ua, ub, au, bu являются подсловами в W. При этом все развилки имеют кратность 2. Стоит отметить, что слова aua, aub, bua, bub

уже не обязательно являются подсловами в W; с другой стороны, нетрудно видеть, что хотя бы два из них должны удовлетворять $C(W)$.

В работе [4] задача оценки количества запретов, задающих слово W, была сведена к задаче оценки количества развилок в слове W. Мы также будем использовать этот результат.

Лемма 2.7 (см. [4, Лемма 3])**.** *Для каждого подслова u слова W существует наименьшая (относительно порядка $\preceq$) развилка $v = r(u)$, содержащая u.* □

Замечание. Если $v \preceq v'$ — подслова в W, то, очевидно, $r(v) \preceq r(v')$.

Лемма 2.8 (см. [4, Лемма 5])**.** *Пусть $v_1, \ldots, v_n$ — все конечные развилки в периодичном слове $W = u^\infty$, а $k_1, \ldots, k_n$ — их кратности. Тогда*

$$|C(W)| \geq 1 + (k_1 - 1) + (k_2 - 1) + \cdots + (k_n - 1).$$

Следствие 2.9. $|C(W)| \geq n + 1$. □

Замечание. Можно показать, что при $|X| = 2$ в лемме 2.8 и в следствии 2.9 всегда достигается равенство.

Это следствие позволяет свести теорему 2.5 к следующей.

Теорема 2.10. *Пусть $|X| = 2$. Рассмотрим периодическое слово $W = u^\infty$, где слово u непериодично. Пусть для этого слова существует n конечных развилок. Тогда $|u| \leq F(n + 1)$.*

Именно этот вариант мы и доказываем в конце раздела 4.

В заключение приведём пример, показывающий, что оценка в теореме 2.5 (и, следовательно, в теореме 2.10) неулучшаемы ни при каком $n \geq 2$.

Пример. Пусть $X = \{a, b\}$. Построим последовательности слов (s_i), (t_i) по следующему правилу. Положим $s_0 = a$, $t_0 = b$; далее, при всех $i \geq 0$ положим $s_{i+1} = s_i s_i t_i$, $t_{i+1} = s_i t_i$. Нетрудно видеть, что $|s_i| = F(2i + 1)$, $|t_i| = F(2i)$. В работе [4, Теорема 3] показано, что при $i \geq 1$ слово $W_{2i} = (t_i)^\infty$ задаётся $2i$ запретами; значит, $|C(W_{2i})| \leq 2i$ по лемме 2.3. Тогда ясно, что слово W_{2i} показывает неулучшаемость оценок в теоремах 2.5 и 2.10 при чётном n.

Аналогично можно показать, что при $i \geq 1$ слово $W_{2i+1} = (s_i)^\infty$ задаётся $2i + 1$ запретом; это показывает неулучшаемость оценок при нечётном n.

3 Комбинаторика

На протяжении этого и последующего разделов мы рассматриваем фиксированное непустое конечное непериодическое слово u, и слово $W = u^\infty$. Развилки и запреты относительно слова W мы называем просто развилками и запретами.

Определение 3.1. *Назовем* значимостью $z(v)$ *подслова v количество раз, которое оно встречается на периоде; формально говоря, если $u = u_1 \ldots u_d$, где $u_1, \ldots, u_d \in X$, и $|v| = t$, то*

$$z(v) = |\{1 \leq i \leq d : u_i \ldots u_{i+t-1} = v\}|,$$

где мы полагаем $u_{t+d} = u_i$ при $1 \leq i \leq d$.

Напомним, что для подслова v слова W через $r(v)$ обозначается наименьшая развилка, содержащая v.

Предложение 3.2. *Если $v \preceq v'$, то $z(v) \leq z(v')$. Кроме того, $z(v) = z(r(v))$.* □

Предложение 3.3. *Пусть v — произвольная конечная развилка. Тогда*

$$z(v) = \sum_{x \in X} z(vx) = \sum_{x \in X} z(r(vx)).$$ □

Пусть $v_0, v_1, \ldots, v_n$ — все развилки, упорядоченные по значимости, то есть $z_0 \le z_1 \le \cdots \le z_n$, где $z_i = z(v_i)$. При этом мы считаем, что $v_0 = W$ (и $z_0 = 1$), а $v_n = \Lambda$ (и $z_n = |u|$). Таким образом, наша цель — получить верхнюю оценку на z_n.

Из предложения 3.3 следует, что $z_1 = z_0 + z_0 = 2$. Из предложений 3.2 и 3.3 следует следующее предложение.

Предложение 3.4. *Пусть $x \in X$, $0 \le i \le n$. Тогда $z(v_i x) < z(v_i)$. В частности, если $r(v_i x) = v_j$, то $j < i$. Наконец, из $v_i \prec v_j$ следует $z_i > z_j$.* □

Далее мы работаем со словами в алфавите $X = \{a, b\}$. В этом случае кратность каждой развилки равна 2. Из предложения 3.3 теперь вытекает следующее предложение.

Предложение 3.5. $z_i \le 2z_{i-1}$, *и* $\max\{r(v_i a), r(v_i b)\} \ge z_i/2$. □

Определение 3.6. *Назовем развилку v_i ($i \ge 2$)* исключительной, *если $z_i > z_{i-1} + z_{i-2}$. В противном случае назовем v_i* регулярной. *Развилки v_0 и v_1 также будем считать регулярными. Индекс i назовем* исключительным (регулярным), *если развилка v_i исключительна (регулярна). Обозначим множество исключительных развилок через $\mathcal{I}$.*

Неформально говоря, в регулярных случаях последовательность (z_i) растет не быстрее чисел Фибоначчи.

Предложение 3.7. *Если развилка v_i исключительна, то $z_i = 2z_{i-1}$, $z_{i-1} > z_{i-2}$, и $r(v_i a) = r(v_i b) = v_{i-1}$.*

Доказательство. Пусть $r(v_i a) = r(v_i b) = v_{i-1}$; тогда $z_i = 2z_{i-1}$, и исключительность развилки v_i равносильна тому, что $2z_{i-1} > z_{i-1} + z_{i-2}$, то есть $z_{i-1} > z_{i-2}$. В противном случае можно считать, что $r(v_i a) = v_j$ при $j \le i - 2$. Тогда по предложению 3.3 $z_i = z(v_i) = z(r(v_i a)) + z(r(v_i b)) \le z_{i-2} + z_{i-1}$, то есть v_i регулярна. □

Замечание. Исключительные развилки могут существовать. Например, в слове u^∞, где $u = (ababbabbabbb)^n a$, развилка $v = babbabb$ исключительна при $n \ge 2$. Действительно, нетрудно проверить, что $z(v) = 2n$, $r(va) = r(vb) = ababbabbabbba = w$, $z(w) = n$; значимость же любой другой развилки либо не меньше $3n - 1$, либо не больше $n - 1$.

Остаток этого раздела посвящён изучению исключительных развилок.

Определение 3.8. *Пусть $v_i \in \mathcal{I}$. Пусть v_j — максимальное собственное начало развилки v_{i-1}, являющееся развилкой. Назовём развилку v_j и её индекс j* штрафными *для исключительной развилки v_i и её индекса i; мы будем обозначать $v_j = \Psi(v_i)$.*

Замечание. В принципе, определением не запрещена ситуация $i = j$; но в дальнейшем мы увидим, что она невозможна, см. предложение 3.11.

Из предложения 3.3 вытекает

Предложение 3.9. *Пусть $v_i \in \mathcal{I}$ и $v_j = \Psi(v_i)$. Тогда $z_j \le z_{j-1} + z_{i-1}$.* □

Предложение 3.10. *Пусть $v_i \in \mathcal{I}$ и $v_j = \Psi(v_i)$, причём $v_{i-1} = r(v_j a)$. Тогда $v_{i-1} \succeq v_j b$.*

Доказательство. Построим последовательность развилок (s_k) следующим образом. Положим $s_0 = v_j$, $s_1 = r(v_jb)$; заметим, что $z(s_1) = z(r(v_jb)) = z(v_j) - z(r(v_ja)) = z(v_j) - z(v_{i-1}) \geq z(v_i)/2$, так как $z(v_j) \geq z(v_i) = 2z(v_{i-1})$. При $k \geq 1$ через s_{k+1} обозначим такую из развилок $r(s_ka)$ и $r(s_kb)$, для которой $z(s_{k+1}) \geq z(s_k)/2$; она существует согласно предложению 3.5 (по замечанию выше, неравенство $z(s_{k+1}) \geq z(s_k)/2$ выполнено и при $k = 0$). Заметим, что $v_jb \preceq s_k$ при каждом $k \geq 1$.

Пусть k — максимальное число, для которого $z(s_k) \geq z(v_{i-1})$; пусть $s_k = v_m$. Предположим, что $m \neq i-1$. По предложению 3.7 имеем $z_{i-2} < z_{i-1}$; значит, случай $m < i-1$ невозможен. Поэтому $m \geq i$, то есть $z(s_k) \geq z_i = 2z_{i-1}$. Но тогда $z(s_{k+1}) \geq z(s_k)/2 \geq z_{i-1}$, что противоречит выбору k. Итак, $m = i-1$, поэтому $v_{i-1} = s_k \succeq v_jb$. □

Предложение 3.11. *Пусть $v_i \in \mathcal{I}$ и $v_j = \Psi(v_i)$, причём $v_{i-1} = r(v_ja)$. Тогда существует такое k $(i < k < j)$, что $z_k \leq z_{k-1} + z_{i-2}$ и $z_k < z_j$. В частности, $j \geq i+2$, и развилка v_k регулярна. Кроме того, в канонической системе запретов существует запрет вида yv_ka, где y — буква.*

Доказательство. Согласно определению слова v_j и предложению 3.10, слово v_{i-1} можно представить как $v_{i-1} = v_jat_1 = t_2v_jbt_3$ для некоторых слов t_1, t_2, t_3. Слово t_2, очевидно, непусто; пусть x — его последняя буква, $t_2 = t_2'x$. Поскольку $v_{i-1} = r(v_ja)$, любое вхождение v_ja в слово W продолжается до $v_jat_1 = v_{i-1}$; в частности, оно продолжается до t_2v_jb. Это значит, что слово t_2v_ja (начинающееся с v_ja) не встречается в W.

Тогда слово $t_2v_ja = t_2'xv_ja$ должно содержать некоторый запрет yv_kz из канонической системы (здесь y, z — буквы, v_k — некоторая развилка). Этот запрет не может быть подсловом слова t_2v_j, ибо оно встречается в W. Также он не может являться подсловом слова xv_ja. Действительно, поскольку v_{i-1} является развилкой, слова av_{i-1} и bv_{i-1} встречаются в W; значит, и их подслова av_ja и bv_ja также в нем встречаются и потому не могут содержать запретов.

Итак, наш запрет yv_kz не содержится в подсловах t_2v_j и xv_ja. Это значит, что он является концом слова t_2v_ja, строго содержащим xv_ja; таким образом, $z = a$, а $v_k = s'v_j$ для некоторого непустого слова s'. Рассмотрим теперь развилку $v_\ell = r(v_ka)$. Слово v_ka заканчивается на v_ja; значит, развилка v_ℓ должна содержать развилку $r(v_ja) = v_{i-1}$. Более того, согласно определению, слово v_ja является началом развилки v_{i-1} и находится не в начале развилки v_ℓ; значит, v_{i-1} — собственное подслово в v_ℓ, то есть $v_\ell \succ v_{i-1}$. Поскольку и v_ℓ, и v_{i-1} — развилки, получаем, что $z(v_\ell) < z(v_{i-1})$ и $\ell \leq i-2$.

Итого, мы нашли развилку $v_k = s'v_j$ такую, что $z(r(v_ka)) \leq z_{i-2}$; значит, $z_k = z(v_k) = z(r(v_ka)) + z(r(v_kb)) \leq z_{i-2} + z_{k-1}$. Заметим, что $v_j \prec v_k \prec v_{i-1}$, поэтому $i-1 < k < j$ и $z_k = z(v_k) < z(v_j) = z_j$. Кроме того, $k \neq i$, ибо $z(r(v_ia)) = z_{i-1} > z_{i-2} \geq z_\ell = z(r(v_ka))$. Значит, $i < k < j$ (и, значит, $j \geq i+2$), и требуемое k найдено. Наконец, поскольку $z_k \leq z_{k-1} + z_{i-2} < z_{k-1} + z_{i-1} \leq z_{k-1} + z_{k-2}$, развилка v_k регулярна. □

Определение 3.12. *Пусть $v_i \in \mathcal{I}$, $v_j = \Psi(v_i)$. Пусть v_k — развилка, построенная в предложении 3.11. Назовем эту развилку v_k и ее индекс k* пеневыми *для исключительной развилки v_i и ее индекса i; обозначим $v_k = \Pi(v_i)$.*

Отметим некоторые свойства штрафных и пеневых развилок.

Предложение 3.13. *Пусть $v_i \in \mathcal{I}$, $v_j = \Psi(v_i)$ и $v_{i-1} = r(v_ja)$. Тогда $z(r(v_ja)) < z(r(v_jb))$. В частности, развилка v_j регулярна.*

Доказательство. Первое утверждение следует из того, что $r(v_ja) = v_{i-1}$, а $z(r(v_ja)) + z(r(v_jb)) = z(v_j) = z_j > z_k \geq z_i = 2z(r(v_ja))$. Тогда v_j не исключительна по предложению 3.7. □

Предложение 3.14. *Пусть $v_i, v_{i'} \in \mathcal{I}$, $v_j = \Psi(v_i)$, $v_{j'} = \Psi(v_{i'})$. Тогда, если $i \neq i'$, то и $j \neq j'$.*

Доказательство. Предположим противное; пусть $i > i'$ и $v_{i-1} = r(v_j a)$. Тогда из предложений 3.7 и 3.13 следует, что $z_{i'-1} < z_{i-1} = z(r(v_j a)) < z(r(v_j b))$, и потому $v_{i'-1}$ не может являться $r(v_j b)$. Таким образом, $v_{i'-1} = r(v_j a) = v_{i-1}$, и $i = i'$. Противоречие. □

Предложение 3.15. *Пусть $v_i, v_{i'} \in \mathcal{I}$, $v_k = \Pi(v_i)$. Тогда $v_k \neq \Psi(v_{i'})$.*

Доказательство. Пусть $v_j = \Psi(v_i)$, причём $v_{i-1} = r(v_j a)$. По предложению 3.11), в канонической системе запретов существует запрет вида $yv_k a$, где y — буква, при этом $z(r(v_k a)) \leq z_{i-2}$, а $z(v_k) \geq z_i > 2z_{i-2}$. Значит, $z(r(v_k b)) = z(v_k) - z(r(v_k a)) > z(r(v_k a))$. Поэтому, если развилка $v_k = \Psi(v_{i'})$, то по предложению 3.13 $v_{i'-1} = r(v_k a)$, и $v_k a$ является началом слова $v_{i'-1}$ по определению штрафной развилки. Но поскольку $v_{i'-1}$ является развилкой, то подслово $yv_{i'-1}$ (и тем более $yv_k a$) встречается в W и потому не может являться запретом — противоречие. □

Суммируем результаты предложений 3.9, 3.11, 3.13, 3.14 и 3.15 в следующей теореме.

Теорема 3.16. *Для каждого исключительного индекса i существуют штрафной и пеневой индексы j и k такие, что $i < k < j$, $z_j \leq z_{j-1} + z_{i-1}$ и $z_k \leq z_{k-1} + z_{i-2}$. При этом штрафные индексы для разных исключительных также различны, а пеневой не может являться штрафным. Кроме того, штрафные и пеневые индексы регулярны (т.е. не исключительны).* □

Определение 3.17. *Назовем индекс r* рядовым*, если он не является ни исключительным, ни штрафным, ни пеневым.*

4 Оценки

В этом разделе мы оцениваем рост последовательности (z_i). Для этого мы сначала введём класс абстрактных (не обязательно связанных со словами) последовательностей, мажорирующих последовательности вида (z_i), а затем будем оценивать эти последовательности.

4.1 Допустимые последовательности

Определение 4.1. *Пусть $n \geq 2$ — натуральное число. Пусть в множестве $\{2, 3, \ldots, n\}$ выделены три попарно непересекающихся подмножества I, J и K, $|I| = |J| \geq |K|$ (элементы этих подмножеств будем называть соответственно* исключительными, штрафными и пеневыми*; индекс, не лежащий ни в одном из подмножеств, назовем* рядовым*). Наконец, пусть вдобавок зафиксированы биекция $\psi : I \to J$ и сюръекция $\pi : I \to K$, причем $i < \pi(i) < \psi(i)$ для любого $i \in I$. Назовём набор $\mathcal{S} = (n, I, J, K, \psi, \pi)$* системой*. Для $k \in K$ определим $d(k) = \min \pi^{-1}(k)$; элементы множества $d(K) \subseteq I$ назовём* плохими *для системы $\mathcal{S}$.*

Самой простой системой является «пустая» система $\mathcal{O}_n = (n, \varnothing, \varnothing, \varnothing, \varnothing, \varnothing)$.

Пусть $\Pi = (x_i)_{i=0}^n$ — последовательность неотрицательных чисел. Будем говорить, что Π соответствует *системе $\mathcal{S}$, если выполнено условие*

(1) для любого $2 \leq r \leq n$, $x_r = x_{r-1} + x_{\theta(r)}$, где

$$\theta(r) = \begin{cases} r-2, & \text{если } r \text{ — рядовой;} \\ r-1, & \text{если } r \text{ — исключительный;} \\ \psi^{-1}(r) - 1, & \text{если } r \text{ — штрафной;} \\ d(r) - 2, & \text{если } r \text{ — пеневой.} \end{cases}$$

Ясно, что такая последовательность задаётся начальными членами x_0 и x_1; будем обозначать её $\Pi_{\mathcal{S}}(x_0, x_1)$.

Назовём последовательность $\Pi_{\mathcal{S}}(a, b)$ допустимой *для $\mathcal{S}$, если $0 \leq a \leq b \leq 2a$; наконец, будем говорить, что последовательность $\Pi_{\mathcal{S}} = \Pi_{\mathcal{S}}(1, 2)$* порождена *системой $\mathcal{S}$.*

Пусть теперь $W = u^\infty$ — бесконечное периодичное слово, и (z_i) — последовательность значимостей, определённая в предыдущем разделе. Результаты этого раздела позволяют выписать порождённую последовательность, мажорирующую последовательность (z_i). Именно, пусть индексы $i_1 < \dots < i_m$ являются исключительными для слова W, индексы $j_1, \dots, j_m$ — штрафными (причем j_t — штрафной для i_t), а индексы $k_1, \dots, k_s$ — пеневыми (напомним, что пеневой индекс может соответствовать нескольким исключительным). Положим $I = \{i_1, \dots, i_m\}$, $J = \{j_1, \dots, j_m\}$, $K = \{k_1, \dots, k_s\}$; по теореме 3.16 эти множества попарно не пересекаются. Далее, для всех $1 \le r \le m$ положим $\psi(i_r) = j_r$ и определим $\pi(i_r)$ как пеневой индекс, соответствующий исключительному i_r. Тогда (n, I, J, K, ψ, π) — система согласно теореме 3.16. Из этой же теоремы вытекает следующее предложение.

Предложение 4.2. *Пусть последовательность (y_i) порождена системой $\mathcal{S} = (n, I, J, K, \psi, \pi)$. Тогда для любого индекса $r = 0, \dots, n$ выполнено неравенство $z_r \le y_r$.*

Доказательство. Индукция по r. При $r = 0, 1$ утверждение очевидно, так как $z_r = y_r$. Пусть $z_s \le y_s$ при всех $s < r$. Тогда по теореме 3.16 имеем $z_r \le z_{r-1} + z_{\theta(r)} \le y_{r-1} + y_{\theta(r)} = y_r$, что и требовалось. □

Отметим сразу некоторые свойства любой допустимой последовательности (y_i), аналогичные свойствам последовательности (z_i) из предыдущего раздела.

Предложение 4.3. *Пусть последовательность (y_i) соответствует системе $\mathcal{S}$. Тогда $0 \le y_i \le y_{i+1}$ при всех $1 \le i \le n$. Если, вдобавок, (y_i) допустима для $\mathcal{S}$, то $0 \le y_i \le y_{i+1} \le 2y_i$ при всех $0 \le i < n$.*

Доказательство. Неравенство $y_i \ge 0$ (и поэтому $y_{i+1} \ge y_i$) следует из определения. Осталось доказать неравенство $y_i \le y_{i+1} \le 2y_i$ для допустимой последовательности (y_i). Применим индукцию по i. При $i = 0$ все утверждения верны. Далее, при $i \ge 1$ имеем $y_{i+1} = y_i + y_r$ при некотором $r = \theta(i+1) \le i$. По предположению индукции имеем $0 \le y_r \le y_i$, откуда $y_i \le y_i + y_r \le 2y_i$, что и требовалось доказать. □

Напомним, что числа Фибоначчи заданы условиями $F(0) = F(1) = 1$ и $F(n+1) = F(n) + F(n-1)$ при всех целых n.

Предложение 4.4. *Пусть $k \ge 2$, $t \ge -1$, и индексы $k, k+1, \dots, k+t$ — рядовые. Тогда $y_{k+t} = F(t+1)y_{k-1} + F(t)y_{k-2}$.*

Доказательство. Индукция по t. При $t = -1, 0, 1$ имеем

$$y_{k-1} = F(0)y_{k-1} + F(-1)y_{k-2}, \quad y_k = F(1)y_{k-1} + F(0)y_{k-2}, \quad y_{k+1} = y_{k-1} + y_k = F(2)y_{k-1} + F(1)y_{k-2}.$$

Если же $t \ge 2$, то

$$y_{k+t} = y_{k+t-1} + y_{k+t-2} = (F(t) + F(t-1))y_{k-1} + (F(t-1) + F(t-2))y_{k-2} = F(t+1)y_{k-1} + F(t)y_{k-2},$$

что и требовалось. □

Пусть последовательность (x_i) порождена системой $\mathcal{O}_n$; тогда, ясно, $x_i = F(i+1)$ при всех $0 \le i \le n$. Наша цель — показать, что для любой порождённой последовательности $y_0, \dots, y_n$ выполняется неравенство $y_n \le x_n = F(n+1)$. С этой целью мы будем перестраивать систему (n, I, J, K, ψ, π), сводя её к пустой, так, чтобы значение y_n не уменьшалось.

4.2 Элементарные улучшения порождённой последовательности

Здесь и далее, если не оговорено противное, $\mathcal{S} = (n, I, J, K, \psi, \pi)$ — произвольная система, а $y_0, \dots, y_n$ — последовательность, ею порождённая. Обозначим через $L = I \cup J \cup K$ множество всех нерядовых индексов.

Каждый раз мы будем перестраивать систему $\mathcal{S}$, получая систему $\mathcal{S}' = (I', J', K', \psi', \pi')$ и порождённую ею последовательность $(y'_i)_{i=0}^n$, для которой $y'_n \geq y_n$ (функции d и θ, а также множество L для системы $\mathcal{S}'$ также будем помечать штрихами). Такую последовательность (y'_i) (систему $\mathcal{S}'$) мы будем называть *улучшением* последовательности (y_i) (системы $\mathcal{S}$). Достаточные условия для улучшения обеспечивает следующая лемма.

Лемма 4.5 (об улучшении)**.** *Пусть $\ell \geq 2$, и выполнены следующие условия:*
(1) из $\theta(i) \geq \ell - 1$ следует $\theta'(i) = \theta(i)$;
(2) $y'_{\ell-1} \geq y_{\ell-1}$, $y'_\ell \geq y_\ell$;
(3) $y'_{\theta'(i)} \geq y_{\theta(i)}$ при всех $i > \ell$ таких, что $\theta(i) < \ell - 1$.
Тогда $y'_i \geq y_i$ при всех $i \geq \ell - 1$; в частности, (y'_i) является улучшением (y_i).

Доказательство. Индукция по i. База для $i = \ell - 1, \ell$ есть условие *(2)*; пусть $i > \ell$. Если $\theta(i) \geq \ell - 1$, то $y_i = y_{i-1} + y_{\theta(i)} \leq y'_{i-1} + y'_{\theta'(i)} = y'_i$ по предположению индукции и условию *(1)*. Если же $\theta(i) < \ell - 1$, то $y_i = y_{i-1} + y_{\theta(i)} \leq y'_{i-1} + y'_{\theta'(i)} = y'(i)$ по предположению индукции и условию *(3)*. □

В этом подразделе мы приведём несколько элементарных улучшений. Первые два из них можно схематично изобразить так:

$$\text{ИР}\dots\text{РН} \to \text{Р}\dots\text{РИН}; \qquad \text{ИКР} \to \text{КРИ}, \quad \text{ИКН} \to \text{КИН},$$

где через И, Р, К, Н обозначены соответственно исключительный индекс, рядовой индекс, штрафной или пеневой индекс, нерядовой неплохой индекс (Напомним, что индекс плох, если он лежит в множестве $d^{-1}(K)$).

Предложение 4.6 (сдвиг исключительного индекса вправо)**.** *Пусть r — исключительный индекс, а $\ell = \min\{t : r < t \in L\}$ — следующий за r нерядовой индекс. Пусть индекс ℓ неплохой. Обозначим через I' множество, полученное из I заменой r на $r' = \ell - 1$. Соответственно изменим функции ψ, π, полагая $\psi'(r') = \psi(r)$, $\pi'(r') = \pi(r)$. Тогда система $\mathcal{S}' = (n, I', J, K, \psi', \pi')$ — улучшение системы $\mathcal{S}$.*

Доказательство. При $\ell = r + 1$ доказывать нечего; пусть $\ell \geq r + 2$. Очевидно, что после замены получается система. Напомним, что через $(y'_i)_{i=0}^n$ мы обозначаем последовательность, порождённую $\mathcal{S}'$. Заметим, что $y'_i = y_i$ при $i < r$. Положим $t = \ell - r - 1 \geq 1$.

Обозначим $a = y_{r-2} = y'_{r-2}$, $b = y_{r-1} = y'_{r-1}$; заметим, что $b \leq 2a$ по предложению 4.3. Тогда $y_r = 2b$, и по предложению 4.4 получаем

$$y_{\ell-1} = y_{r+t} = F(t)y_r + F(t-1)y_{r-1} = (2F(t) + F(t-1))b = F(t+2)b.$$

Аналогично имеем

$$\begin{aligned} y'_{\ell-2} = y'_{r+t-1} &= F(t)y_{r-1} + F(t-1)y_{r-2} = F(t)b + F(t-1)a \geq \\ &\geq F(t)b + \frac{F(t-1)}{2}b = \frac{F(t)}{2}b, \\ y'_{\ell-1} = 2y'_{\ell-2} &\geq F(t+2)b = y_{\ell-1}. \end{aligned}$$

Далее, $y_\ell = y_{\ell-1} + y_{\theta(\ell)}$, $y'_\ell = y'_{\ell-1} + y'_{\theta'(\ell)}$. При этом, если $\ell = \pi(r)$, то $y'_{\theta'(\ell)} = y'_{\ell-3} \geq y'_{r-2} = y_{r-2} = y_{\theta(r)}$. Иначе $\theta'(\ell) = \theta(\ell)$, и либо $\theta(\ell) = \ell - 1$, либо $\theta(\ell) \leq r - 1$, ибо индексы $r+1, \dots, \ell-1$ рядовые. В любом случае получаем $y'_{\theta'(\ell)} \geq y_{\theta(\ell)}$, а потому и $y'_\ell \geq y_\ell$.

Мы готовы проверить, что условия леммы 4.5 об улучшении выполнены, откуда будет следовать требуемое. Условие *(2)* уже проверено; условие *(1)* очевидно. Условие же *(3)* очевидно для всех $i \notin \{\psi(r), d^{-1}(r), d^{-1}(\ell)\}$, ибо тогда из $\theta(i) < \ell - 1 \le i - 2$ следует $\theta(i) = \theta'(i) \le r - 1$ и $y_{\theta(i)} = y'_{\theta'(i)}$. Рассмотрим оставшиеся случаи. При $i = \psi(r) = \psi'(\ell - 1)$ имеем $y'_{\theta'(i)} = y'_{\ell-2} \ge y'_{r-1} = y_{r-1} = y_{\theta(i)}$. При $i = d^{-1}(r)$ имеем $y'_{\theta'(i)} = y'_{\ell-3} \ge y'_{r-2} = y_{r-2} = y_{\theta(i)}$. Наконец, случай $i = d^{-1}(\ell)$ невозможен, ибо ℓ — неплохой, т.е. $\ell \notin d(K)$. □

Замечание. Подобную операция *замены* одного индекса другим мы будем описывать многократно. В дальнейшем мы не будем описывать соответствующую замену функций ψ, π, считая её подразумевающейся.

Предложение 4.7 (перемена мест)**.** *Пусть $2 \le r \le n - 2$, $r \in I$, $r + 1 \in J \cup K$, причём $r + 1 \notin \{\psi(r), \pi(r)\}$, а индекс $r + 2$ — неплохой. Если $r + 2$ — регулярный, то заменим в I индекс r на $r + 2$, а в J или в K — индекс $r + 1$ на r. Иначе заменим в I индекс r на $r + 1$, а в J или в K — индекс $r + 1$ на r.*

Тогда полученная система $\mathcal{S}'$ улучшает $\mathcal{S}$.

Доказательство. Заметим сразу, что $y_i = y'_i$ при $i < r$. Кроме того, условие $r+1 \notin \{\psi(r), \pi(r)\}$ гарантирует, что после замены получается система. Обозначим $b = y_{r-1}$, $p = y_{\theta(r+1)}$. Возможны три случая.

1. Пусть индекс $r + 2$ — регулярный. Тогда

$$\begin{aligned} &y_r = 2b, && y_{r+1} = 2b + p, && y_{r+2} = 4b + p, \\ &y'_r = b + p, && y'_{r+1} = 2b + p = y_{r+1}, && y'_{r+2} = 4b + 2p \ge y_{r+2}. \end{aligned}$$

Проверим условия леммы об улучшении при $\ell = r + 2$. Условие *(1)* очевидно, а *(2)* уже проверено. Условие *(3)* требует проверки лишь при $i \in \{\psi(r), d^{-1}(r)\}$ (иначе $\theta(i) = \theta(i') < r$ и $y_{\theta(i)} = y'_{\theta'(i)}$). Если $i = \psi(r) = \psi'(r + 2)$, то $y'_{\theta'(i)} = y'_{r+1} \ge y'_{r-1} = y_{\theta(i)}$. Если же $r = d(i)$, то $y'_{\theta'(i)} = y'_r \ge y'_{r-2} = y_{\theta(r)}$.

2. Пусть теперь индекс $r + 2$ — исключительный и неплохой. Тогда

$$\begin{aligned} &y_r = 2b, && y_{r+1} = 2b + p, && y_{r+2} = 4b + 2p, \\ &y'_r = b + p, && y'_{r+1} = 2b + 2p \ge y_{r+1}, && y'_{r+2} = 4b + 4p \ge y_{r+2}. \end{aligned}$$

Опять проверим условия леммы об улучшении при $\ell = r+2$. Условия *(1)* и *(2)* верны. Условие *(3)* требует проверки лишь при $i \in \{\psi(r), d^{-1}(r)\}$ (напомним, что $r + 2$ — неплохой); эта проверка производится аналогично предыдущему случаю.

3. Наконец, пусть индекс $r + 2$ — штрафной или пеневой. Обозначим $q = y_{\theta(r+2)}$. Тогда

$$\begin{aligned} &y_r = 2b, && y_{r+1} = 2b + p, && y_{r+2} = 2b + p + q, \\ &y'_r = b + p, && y'_{r+1} = 2b + 2p \ge y_{r+1}, && y'_{r+2} = 2b + 2p + q \ge y_{r+2}. \end{aligned}$$

Проверка условий леммы об улучшении проводится аналогично первому случаю. □

Следствие 4.8 (о разделении)**.** *Пусть $2 \le p \le q < n$, причём $q + 1$ — неплохой нерядовой индекс. Обозначим $T = [p, q]$. Предположим, что $T \cap K = \varnothing$, и все индексы из множества $I \cap T$, кроме, возможно, наименьшего из них — неплохие.*

Тогда существует система $\mathcal{S}' = (n, I', J', K', \psi', \pi')$, улучшающая $\mathcal{S}$; при этом $\mathcal{S}$ и $\mathcal{S}'$ отличаются лишь на отрезке T (формально говоря, $I \setminus T = I' \setminus T$, $J \setminus T = J' \setminus T$, $K \setminus T = K' \setminus T$, и $\psi(i) = j \iff \psi'(i) = j$, $\pi(i) = k \iff \pi'(i) = k$ для любых $i, j, k \notin T$), и множество $I' \cap T$ находится правее, чем $J' \cap T$. Более того, $|I' \cap T| = |I \cap T|$, $|J' \cap T| = |J \cap T|$, и $K' \cap T = \varnothing$.

Доказательство. Если $I_1 = I \cap T$ уже находится правее $J_1 = J \cap T$ (в частности, если одно из этих множеств пусто), то можно положить $\mathcal{S}' = \mathcal{S}$. Иначе выберем

$$i_0 = \min I_1, \qquad j = \min\{j \in J_1 : j > i_0\}, \qquad i = \max\{i \in I_1 : i < j\}.$$

Тогда i можно заменить на $j-1$ по предложению 4.6, а затем поменять их местами по одному из вариантов предложения 4.7. Последнее возможно, так как $j+1$ не может быть плохим индексом по условию, а также $\pi(i) > j$ (ибо $I \cap K = \varnothing$). Поскольку сумма исключительных индексов строго возрастает, серией таких замен мы рано или поздно добьёмся требуемого.

Осталось заметить, что при каждой замене мощности множеств $I \cap T$, $J \cap T$ и $K \cap T$ не менялись. □

Ещё одно преобразование связано только с изменением функции ψ, то есть с «переназначением» штрафных индексов.

Предложение 4.9 (о переназначении двух штрафов)**.** *Пусть $i_1 < i_2$ — некоторые исключительные индексы, а $j_1 < j_2$ — соответствующие им штрафные (т.е. $\psi(i_s) = j_s$ при $s = 1, 2$), причём $j_1 > \pi(i_2)$. Изменим функцию ψ на элементах i_1, i_2, полагая $\psi'(i_s) = j_{3-s}$ при $s = 1, 2$. Тогда получилась система $\mathcal{S}'$, являющаяся улучшением системы $\mathcal{S}$.*

Доказательство. Условия $i_1 < i_2 < \pi(i_2) < j_1 < j_2$ гарантируют, что $\mathcal{S}'$ — система. Заметим, что $y_t = y'_t$ при $t < j_1$. Обозначим $a_s = y_{i_s - 1} = y_{\theta(j_s)} = y'_{\theta'(j_{3-s})}$ при $s = 1, 2$, $\delta = a_2 - a_1 \geq 0$. Тогда $y'_{j_1} = y_{j_1 - 1} + a_2 = y_{j_1} + \delta$. Непосредственная индукция показывает, что $y'_t \geq y_t + \delta$ при $j_1 \leq t < j_2$. Тогда $y'_{j_2} = y'_{j_2 - 1} + a_1 \geq y_{j_2 - 1} + a_1 + \delta = y_{j_2}$. Тогда нетрудно видеть, что все условия леммы 4.5 об улучшении при $\ell = j_2$ выполнены. □

Следствие 4.10 (о переназначении штрафов)**.** *Пусть $k \geq 2$, $i_1 < \dots < i_k$ — некоторые исключительные индексы, а $j_1 > \dots > j_k$ — штрафные индексы, причём $\psi(\{i_1, \dots, i_k\}) = \{j_1, \dots, j_k\}$. Изменим функцию ψ на элементах $i_1, \dots, i_k$, полагая $\psi'(i_s) = j_s$ при $s = 1, \dots, k$. Тогда, если $\mathcal{S}' = (n, I, J, K, \psi', \pi)$ — система, то $\mathcal{S}'$ — улучшение системы $\mathcal{S}$.*

Доказательство. Индукция по k. При $k = 2$ это — предыдущее предложение. Пусть $k > 2$. Если $\psi(i_k) = j_k$, то можно непосредственно применить предположение индукции. Пусть теперь $\psi(i_k) = j_s$ при $s < k$; тогда $j_k = \psi(i_t)$ при некотором $t < k$. Применим предложение 4.9 к индексам i_t, i_k, j_k, j_s. Поскольку $\mathcal{S}'$ — система, то $i_\ell < i_k < \pi(i_k) < j_k < j_s$, поэтому при замене функции ψ переназначением $\psi''(i_t) = j_s > j_k$, $\psi''(i_k) = j_k$ также получается система $\mathcal{S}''$, являющаяся улучшением $\mathcal{S}$. Для неё опять можно применить предположение индукции, ибо $\psi''(i_k) = j_k$. □

Замечание. Для того, чтобы в условиях следствия 4.10 $\mathcal{S}'$ оказалась системой, достаточно, например, чтобы выполнялось условие $|\pi(\{i_1, \dots, i_k\})| = 1$.

4.3 Случай единственной пени

Разберём сначала случай, когда $|K| = 1$. В этом случае оказывается верна следующая лемма.

Лемма 4.11. *Пусть последовательности (y_i) и (x_i) порождены системами $\mathcal{S} = (n, I, J, K, \psi, \pi)$ и $\mathcal{O}_n$ оответственно, причём $2 \in I$, $|K| = 1$ и $x_n \geq y_n$. Тогда для любых допустимых последовательностей $\Pi_{\mathcal{O}_n}(a, b) = (x'_i)$ и $\Pi_{\mathcal{S}}(a, b) = (y'_i)$ имеем $x'_n \geq y'_n$.*

Доказательство. Обозначим $(a_i) = \Pi_{\mathcal{O}_n}(1, 0)$, $(b_i) = \Pi_{\mathcal{S}}(1, 0)$; пусть $K = \{k\}$. Заметим, что $b_i = 0$ при $i < k$, поскольку при всех таких индексах $\theta(i) > 0$. Кроме того, поскольку $k \geq 3$, мы имеем $a_k \geq 1 = b_k$ и $a_{k+1} \geq 2 = b_k + b_0 = b_k + b_{\theta(k)} = b_{k+1}$.

Далее, никакой индекс $i \geq k$ — не исключительный. Покажем индукцией по $i \geq k$, что $b_i \leq a_i$. Действительно, при $i = k$ и $i = k + 1$ утверждение уже доказано; если же $i \geq k + 2$,

то $b_i = b_{i-1} + b_{\theta(i-1)}$; если $\theta(i-1) \neq i-2$, то $b_{\theta(i-1)} = 0$ и $b_i = b_{i-1} \leq a_{i-1} \leq a_i$; иначе $b_i \leq b_{i-1} + b_{i-2} = a_{i-1} + a_{i-2} = a_i$, что и требовалось.

Итак, мы получаем, что $b_n \leq a_n$. Наконец, заметим, что $\Pi_{\mathcal{O}_n}(a,b) = (b/2)\Pi_{\mathcal{O}_n}(1,2) + (a - b/2)\Pi_{\mathcal{O}_n}(1,0)$ и $\Pi_{\mathcal{S}}(a,b) = (b/2)\Pi_{\mathcal{S}}(1,2) + (a-b/2)\Pi_{\mathcal{S}}(1,0)$, причём $a - b/2 \geq 0$, поскольку эти последовательности допустимы. Значит,

$$x'_n = (b/2)x_n + (a-b/2)a_n \geq (b/2)y_n + (a-b/2)b_n = y'_n,$$

что и требовалось доказать. □

Определение 4.12. *Для числового множества X определим его* сдвиг влево *как $X^- = \{x-1 : x \in X\}$. Для числовой функции ϕ определим её* сдвиг влево *формулой $\phi^-(x) = \phi(x+1) - 1$.*

Пусть $\mathcal{S} = (n, I, J, K, \psi, \pi)$ — система, в которой 2 — регулярный индекс. Определим её сдвиг влево *как систему $\mathcal{S}^- = (n-1, I^-, J^-, K^-, \psi^-, \pi^-)$.*

Лемма 4.13. *Пусть $\mathcal{S} = (n, I, J, K, \psi, \pi)$ — система с $|K| = 1$. Пусть $(x_i) = \Pi_{\mathcal{S}}(a,b)$ и $(y_i) = \Pi_{\mathcal{O}_n}(a,b)$ — допустимые последовательности. Тогда $x_n \geq y_n$.*

Доказательство. Предположим противное; выберем из всех допустимых последовательностей $\Pi_{\mathcal{S}}(a,b)$ (при всевозможных $\mathcal{S}$, a и b), противоречащих лемме, ту, для которой n минимально, а из таких — ту, для которой минимально $|I|$. Согласно лемме 4.11, можно считать, что $a = 1$, $b = 2$. Если $2 \notin I$, то 2 — регулярный индекс. Значит, для последовательностей $(x_{i+1})_{i=0}^{n-1} = \Pi_{\mathcal{S}^-}(b, a+b)$ и $(y_{i+1})_{i=0}^{n-1} = \Pi_{\mathcal{O}_{n-1}}(b, a+b)$ утверждение леммы верно, то есть $x_n \geq y_n$, что не так. Итак, $2 \in I$.

Пусть $K = \{k\}$. Предположим, что $I \neq \{2, 3, \ldots, k-1\}$. Положим $i_0 = \min\{i \geq 2 : i \notin I\}$, $j_0 = \min\{\ell \in L : \ell > i_0\}$, $t = j_0 - i_0$. По лемме 4.6, можно последовательно сдвинуть все индексы $i_0 - 1, i_0 - 2, \ldots, 2$ в индексы $i_0 + t - 1, \ldots, 2 + t$ соответственно, улучшив систему $\mathcal{S}$. При этом 2 не является исключительным для новой системы, что невозможно. Значит, $I = \{2, \ldots, k-1\}$. Далее, согласно следствию 4.10 о переназначении штрафов, можно считать, что $\psi(2) > \psi(3) > \cdots > \psi(k-1)$.

Предположим, что $|I| \geq 2$. Положим $I' = (I \setminus \{2\})^-$, $J' = (J \setminus \{\psi(2)\})^-$, $K' = \{k-1\}$, $\psi' = \psi^-\big|_{I'}$, $\pi'(i) = k-1$ для всех $i = 2, \ldots, k-2$. Тогда $\mathcal{S}' = (n, I', J', K', \psi', \pi')$ — система; пусть $(y'_i) = \Pi_{\mathcal{S}'}$. Покажем, что $y'_n \geq y_n$; это будет противоречить исходному выбору, ибо $|I'| < |I|$.

Положим $t = \max J$, $t' = \max J'$; тогда $d = t - t' \geq 2$. Положим $\mathcal{S}'' = (t', I', J', K', \psi', \pi')$. Пусть $(a_i) = \Pi_{\mathcal{S}''} = (y'_i)_{i=0}^{t'}$. Положим $p = a_{t'-1}$, $q = a_{t'}$. Заметим, что $p \geq 5$, $q = a_{t'-1} + a_1 = p + 2 \geq 7$. Тогда по предложению 4.4, $y'_{t-1} = F(d-1)q + F(d-2)p$, $y'_t = F(d)q + F(d-1)p$. С другой стороны, имеем $y_1 = 2$, $y_2 = 4$; значит, отрезок последовательности $(y_{i+1})_{i=0}^{t'}$ строится так же, как и $\Pi_{\mathcal{S}''}(2,4) = (2a_i)$, за единственным исключением: $y_k = y_{k-1} + 1$, в то время как $2a_{k-1} = 2a_{k-2} + 2$. Тогда нетрудно видеть, что $y_{t'} \leq 2p$, $y_{t'+1} \leq 2q$. Теперь, снова по предложению 4.4, получаем $y_{t-1} \leq 2F(d-2)q + 2F(d-3)p$. Тогда

$$y_t = y_{t-1} + y_1 \leq 2F(d-2)q + 2F(d-3)p + 2 \leq F(d)q + F(d-1)p = y'_t,$$

поскольку $F(d) \geq 2F(d-2)$, $F(d-1) \geq 2F(d-3)$, причём хотя бы одно из этих неравенств строгое. Далее, пусть $t < n$; покажем тогда, что $y'_{t+1} \geq y_{t+1}$. Поскольку $q = p + 2$, имеем

$$\begin{aligned} y'_{t+1} - y_{t+1} &= (y'_t + y'_{t-1}) - (y_t + y_{t-1}) \geq (F(d+1)q + F(d)p) - (4F(d-2)q + 4F(d-3)p + 2) = \\ &= (F(d+2) - 4F(d-1))p + 2(F(d+1) - 4F(d-2) - 1) = \\ &= (2F(d-2) - F(d-1))(p-2) + 2(2F(d-1) - F(d) - 1). \end{aligned}$$

Поскольку $2F(d-2) \geq F(d-1)$ и $2F(d-1) \geq F(d)$, причём хотя бы одно из этих неравенств строгое, получаем $y'_{t+1} - y_{t+1} \geq \min\{p-4, 0\} = 0$. Итак, $y_{t+1} \leq y'_{t+1}$, $y_t \leq y'_t$, откуда и следует, что $y_n \leq y'_n$.

Наконец, пусть $|I| = 1$. Тогда аналогично $\mathcal{S}$ заменяется на пустую систему с увеличением последнего члена. □

4.4 Общая оценка

Теперь мы готовы к доказательству общей оценки. Сначала докажем лемму, позволяющую сделать ключевой индукционный шаг с применением леммы 4.13.

Для каждого $k \in K$ определим его *отрезок влияния* $A(k) = [d(k), \max \psi(\pi^{-1}(k))]$. Иными словами, отрезок влияния пеневого индекса k — это минимальный отрезок, содержащий все исключительные и штрафные индексы, соответствующие k. Назовём индекс $k \in K$ *выделенным*, если на отрезке $A(k)$ нет индексов, соответствующих другому пеневому индексу (иначе говоря, $\pi(I \cap A(k)) = \pi(\psi^{-1}(J \cap A(k))) = K \cap A(k) = \{k\}$).

Лемма 4.14 (о выделении)**.** *Пусть $|K| \geq 2$. Тогда существует система $\mathcal{S}'$, улучшающая $\mathcal{S}$, такая, что в ней $|K'| \leq |K|$, причём либо в $\mathcal{S}'$ существует выделенный пеневой индекс k_0, либо $|K'| < |K|$. При этом в первом случае имеем $d(K) \cap [k_0, n] = \varnothing$.*

Доказательство. Процесс улучшения будет проходить в несколько шагов. На каждом шаге мы будем улучшать систему, добиваясь выполнения некоторого свойства (своего для каждого шага). При этом это свойство будет сохраняться и при всех последующих шагах.

Пусть $i_0 = \max d(K)$, $k_0 = \pi(i_0)$.

Шаг 1. Предположим, что существует пеневой индекс k, лежащий на интервале (i_0, k_0). Переопределим функцию π, полагая $\pi'(i) = k$ для всех $i \in [i_0, k] \cap \pi^{-1}(k_0)$. Из определения i_0 следует, что значение $d(k)$ не изменилось. Если после этого оказалось, что $\pi'^{-1}(k_0) = \varnothing$, то выкинем k_0 из K. Нетрудно видеть, что получилась система, причём являющаяся улучшением исходной (все члены порождённой последовательности вплоть до k_0-го не изменились, последующие не уменьшились); при этом, если $|K'| = |K|$, то значение $d(k_0)$ увеличилось. После нескольких таких шагов мы либо уменьшим $|K|$ (тем самым добившись требуемого), либо добьёмся того, что $(i_0, k_0) \cap K = \varnothing$.

Итого, можно считать, что $(i_0, k_0) \cap K = \varnothing$.

Шаг 2. Переопределим функцию π, полагая $\pi'(i) = k_0$ для всех $i \in [i_0, k_0] \cap I$. Получилась снова система, причём, поскольку множество $(i_0, k_0] \cap I$ не содержит плохих индексов, порождённая последовательность не изменилась.

Итого, можно считать, что $\pi([i_0, k_0] \cap I) = k_0$.

Шаг 3. Рассмотрим отрезок $T = [i_0, k_0 - 1]$; на нём нет пеневых индексов, а плохим является только индекс i_0. По следствию 4.8 о разделении можно так перестроить систему $\mathcal{S}$ на отрезке T, что на отрезке T все штрафные индексы будут левее всех исключительных. При этом значение i_0 может только увеличиться, а свойство шага 2 сохраняется.

Иными словами, можно считать, что $[i_0, k_0] \cap J = \varnothing$.

Шаг 4. Обозначим $I_1 = [i_0, k_0] \cap I = \{i_0, \ldots, i_t\}$ $(i_0 < \cdots < i_t)$, $J_1 = \psi(I_1)$. По следствию 4.10 о переназначении штрафов, можно считать, что $\psi(i_s) > \psi(i_r)$ при $0 \leq s < r \leq t$. Обозначим $j_s = \psi(i_s)$ при $0 \leq s \leq t$.

Шаг 5. Предположим, что на отрезке $[i_0, j_0]$ содержится ещё какой-то пеневой индекс $k \neq k_0$ (тогда $k > k_0$; мы выбираем k наименьшим возможным). Пусть $s_0 = \max\{s : j_s > k\}$. Тогда можно переопределить функцию π на элементах $i_0, \ldots, i_{s_0}$, полагая $\pi'(i_s) = k$ при $0 \leq s \leq s_0$ (при этом, если $s_0 = t$, то надо ещё выкинуть k_0 из K', уменьшив тем самым $|K|$; в противном случае будем иметь $d'(k_0) = i_{s_0+1}$). При этом значение $d(k)$ не изменится по выбору i_0. Нетрудно видеть, что получилась система, улучшающая исходную: члены порождённой последовательности вплоть до k_0-го не изменились, а дальнейшие не уменьшились.

При этом в изменённой системе (для новых значений i_0, k_0) выполнено условие $[i_0, j_0] \cap K = \{k_0\}$.

Шаг 6. Пусть теперь $J_1 = J \cap [k_0, j_0]$. Покажем, что можно переназначить штрафы так, чтобы индексы $j_0, \ldots, j_t$ являлись минимальными индексами в J_1 (иначе говоря, чтобы $\{j_0, \ldots, j_t\} = J \cap [k_0, j_0]$). Пусть это не так, то есть для некоторого $0 \leq s \leq t$ существует $j \in J \setminus \{j_0, \ldots, j_t\}$ такой, что $j < j_s$; пусть $i = \psi^{-1}(j) < i_0$. Можно считать, что s — максимальный индекс с этим

свойством, а j — минимальный для этого s. Тогда по предложению 4.9 о переназначении двух штрафов можно переназначить $\psi'(i) = j_s$, $\psi'(i_s) = j$; ясно, что получится система, причём в ней для индекса s уже не будет существовать таких j. Повторяя процедуру, в конце концов добьёмся требуемого. Заметим, что в процессе переназначений порядок элементов $\psi(i_0), \dots, \psi(i_t)$ остаётся неизменным.

Итак, мы добились того, что $\{j_0, \dots, j_t\} = [i_0, j_0] \cap J$ (напомним, что $[i_0, k_0] \cap J = \varnothing$ уже после Шага 3).

Шаг 7. Предположим, наконец, что $[i_0, j_0] \cap L \neq \{i_0, \dots, i_t, j_0, \dots, j_t, k_0\}$; по результатам предыдущих шагов, «лишними» элементами могут быть только исключительные индексы, лежащие на отрезке $[k_0, j_0]$. Тогда пеневые индексы, им соответствующие, больше j_0. Положим $f_0 = \min\{f \in L : f > j_0\}$. Из выбора i_0 следует, что f_0 — неплохой. Тогда по следствию 4.8 о разделении, существует улучшение $\mathcal{S}'$ нашей системы, отличающееся от неё лишь на отрезке $[k_0 + 1, f_0 - 1]$, в котором уже $\left[k_0, \max\limits_{0 \le s \le t} \psi(i_s)\right] \cap I = \varnothing$.

Суммируя предыдущие результаты, видим, что в полученной системе выполняется соотношение

$$[i_0, j_0] \cap L = \{i_0, \dots, i_t, j_0, \dots, j_t, k_0\}.$$

Таким образом, индекс k_0 в ней является выделенным. Кроме того, по определению i_0, мы имеем $d(K) \cap [k_0, n] = \varnothing$. □

Теорема 4.15. $y_n \le F(n+1)$.

Доказательство. Индукция по $|K|$. При $|K| = 0$ доказывать нечего, при $|K| = 1$ утверждение следует из леммы 4.13.

Пусть $|K| \ge 2$. Применяя лемму 4.14 о выделении, мы либо уменьшим $|K|$ (после чего применимо предположение индукции), либо получим систему с тем же значением $|K|$, в которой некоторый пеневой индекс k_0 — выделенный. Полученную систему опять будем обозначать через $\mathcal{S}$.

Итого, пусть индекс $k_0 \in K$ выделен. Пусть $\mathcal{S}' = (n, I', J', K', \psi', \pi')$, где

$$K' = K \setminus \{k_0\}, \qquad I' = I \setminus \pi^{-1}(k_0), \qquad J' = J \setminus \psi(\pi^{-1}(k_0)), \qquad \psi' = \psi|_{I'}, \qquad \pi' = \pi|_{I'};$$

нетрудно видеть, что $\mathcal{S}'$ — система. Пусть (y'_i) — последовательность, порождённая $\mathcal{S}'$. Поскольку $|K'| = |K| - 1$, по предположению индукции $y'_n \le F(n+1)$. Для завершения доказательства достаточно доказать, что $\mathcal{S}'$ улучшает $\mathcal{S}$.

Положим

$$I_0 = \pi^{-1}(k_0), \quad i_0 = \min I_0, \qquad J_0 = \psi(I_0), \quad j_0 = \max J_0.$$

Заметим, что $y'_i = y_i$ при $i < i_0$. Пусть j — максимальный индекс такой, что все индексы из полуинтервала $(j_0, j]$ — рядовые (таким образом, если $j_0 = \max L$, то $j = n$, иначе $j = \min\{j \in L : j > j_0\} - 1$).

Неформально говоря, поскольку индекс k_0 — выделенный, последовательность (y_i) ведёт себя на отрезке $[i_0, j]$ в точности как последовательность, допустимая для «сдвинутой» системы $(j - i_0 + 2, I_0, J_0, K_0, \psi|_{I_0}, \pi|_{I_0})$. Формализуем это утверждение.

Положим $i_* = i_0 - 2$. Определим систему $\mathcal{S}_1 = (j - i_*, I_1, J_1, K_1, \psi_1, \pi_1)$ следующим образом:

$$I_1 = I_0 - i_*, \qquad J_1 = J_0 - i_*, \qquad K_1 = \{k_0 - i_*\},$$
$$\psi_1(i - i_*) = \psi(i) - i_*, \qquad \pi_1(i - i_*) = \pi(i) - i_*.$$

Пусть $(x_i)_{i=0}^{j-i_*} = \Pi_{\mathcal{S}_1}(y_{i_*}, y_{i_*+1})$. Тогда непосредственная индукция показывает, что $x_i = y_{i+i_*}$ при всех $0 \le i \le j - i_*$, ибо оба члена получаются из предыдущих по одинаковым правилам.

Аналогично, если $(x'_i)_{i=0}^{j-i_*}$ — допустимая последовательность для пустой системы с теми же начальными условиями, то $x'_i = y'_{i+i_*}$ при всех $0 \le i \le j - i_*$. По следствию 4.13, имеем теперь $y_j = x_{j-i_*} \le x'_{j-i_*} = y'_j$. Если $j = n$, то это и есть требуемое неравенство.

Пусть, наконец, $j < n$. Покажем, что последовательности (y_i) и (y'_i) удовлетворяют условиям леммы 4.5 об улучшении при $\ell = j + 1$. Заметим, что $y_{j+1} = y_j + y_{\theta(j+1)}$, $y'_{j+1} = y'_j + y'_{\theta'(j+1)}$, причём индекс $\theta = \theta(j + 1) = \theta'(j + 1)$ либо равен j (если $j + 1 \in I$), либо не превосходит $i_0 - 1$. Значит, $y_{\theta(j+1)} \le y'_{\theta(j+1)}$, откуда следуют условия *(2)*. Напомним, что из утверждения леммы 4.14 вытекает, что индекс $j + 1$ неплохой.

Далее, при любом $i \ge j+1$ имеем $\theta = \theta(i) = \theta(i)'$, что доказывает *(1)*. Кроме того, при этих же значениях i либо $\theta(i) \ge j$, либо $\theta(i) \le i_0 - 1$ (это следует из того, что $j + 1$ — неплохой, а k_0 — выделенный). Поэтому условие *(3)* также выполнено, ибо $y_s = y'_s$ при $s \le i_0 - 1$.

Итак, по лемме об улучшении $y_n \le y'_n$, что и требовалось доказать. □

Теперь мы можем доказать основную теорему.

Доказательство теоремы 2.10. Пусть $v_1, \dots, v_n$ — все развилки в слове W, упорядоченные по неубыванию значимости, $z_i = r(v_i)$. По предложению 4.2, $z_n \le y_n$ для некоторой порождённой последовательности $(y_i)_{i=0}^n$. По теореме 4.15, $y_n \le F(n+1)$. Значит, и $|u| = z_n \le F(n+1)$, что и требовалось доказать. □

5 Алфавит из произвольного количества букв

Список литературы

[1] В.А. Уфнаровский. Комбинаторные и асимптотические методы в алгебре // ВИНИТИ, 1990 Сер. Совр. пробл. Математики. Фундаментальные направления. Т.57. М. С.5 — 177

[2] А.Г. Курош. Проблеммы теории колец, связанные с проблеммой Бернсайда о периодических группах //Изв. АН СССР сер. мат. 1941. Т.5. С. 233 — 240

[3] А.Я. Белов, В.В. Борисенко, В.Н. Латышев. Мономиальные алгебры //ВИНИТИ, 2002. Итоги науки и техники. Сер. Современная математика и ее приложения. Тематические обзоры. Т.26. М. С.35 — 214

[4] Г.Р. Челноков. О числе запретов, задающих периодическую последовательность. // Модел. и анализ информ. систем. Т. 13, № 3 (2007), 66–70.

[5] P. Lavrov. Number of restrictions required for periodic words in finite alphabet. `arXiv:1209.0220.`

О числе запретов, задающих периодическую последовательность.

1 Предисловие

При изучении алгебраических объектов используется их представление в виде образующих и определяющих соотношений. Способов представлять элемент через образующие может быть много, так что изучается *каноническая форма* представления. Например, пусть A – ассоциативная алгебра, $a_1, \ldots, a_s$ – ее образующие. Порядок $a_1 \prec \cdots \prec a_s$ индуцирует порядок на множестве мономов от a_i (сперва по длине, потом лексикографически). Множество слов, не являющихся линейной комбинацией меньших, образует *нормальный базис* алгебры A. *Функция роста* $V_A(n)$ есть размерность пространства, порожденного мономами степени не выше n и совпадает с числом элементов нормального базиса степени не выше n. Функции роста и нормальные формы определяются для разных алгебраических систем, им посвящена обширная литература. Обзор – см. [1]

Рассматривая идеал соотношений I, рассматривают старшие члены его элементов или *редуцируемые слова*. Пусть $\{f_i\}$ – образующие I. Тогда надслово редуцируемого слова редуцируемо. *Обструкцией* называется минимальное редуцируемое слово, т.е. без редуцируемых подслов. Если все обструкции входят в старшие члены базиса, то он называется *базисом Гребнера-Ширшова* идеала I. Впервые это понятие было введено А.И.Ширшовым. Аналогичные понятия для полилинейных слов ввел В.Н.Латышев (минимальность понимается также и в том что слово не является изотонным образом редуцированного). А.И.Ширшов ввел понятие *композиции* и предложил критерий того, что $\{f_i\}$ является базисом гребнера. Это легло в основе знаменитой Diamond-леммы Бергмана.

Базис гребнера даже конечно определенной алгебры может быть бесконечен. А.Я.Белов ввел понятие *короста* в алгебе или функции, $B(n)$ выражающей количество обструкций длины не выше n.

Данная работа посвящена исследованию *кодлины* периода или количеству запретов, которыми можно задать периодическую последовательность.

Основной результат данной работы состоит в следующем:

Теорема 1. *В случае двухсимвольного алфавита $A = \{a, b\}$ если слово из I_n задается с запретами, то $\varphi_c \geq n$, где φ_c — это с-е число Фибоначчи ($\varphi_1 = 1$, $\varphi_2 = 2$, $\varphi_3 = 3$, $\varphi_4 = 5$ и т.д.).*

Отметим, что логарифмическая оценка, а также основные примеры и оценка сверху на количество запретов были получены ранее в работе Г. Р. Челнокова [2]. Мы вычисляем более точную ассимптотику (множитель при логарифме).

В основе доказательства лежит работа с *графами и схемами Рози*

Последовательности схем Рози исследовались в ряде работ. С их помощью удается решить ряд алгоритмических проблем. см. например [3], [4] и [5]

Отметим,что результаты работ [3] и [4] другим методом независимо получены Ф.Дюрандом: [6], [7]

В терминах размеченных схем Рози удается получить критерий того, что слово отвечает перекладыванию отрезков [8]. Подробнее см. обзор [9]

Перейдем к доказательству основного результата.

2 Введение

В данной работе изучаются свойства бесконечных в обе стороны периодических последовательностей (слов, строк — будем считать эти термины эквивалентными) над фиксированным конечным алфавитом A.

Обозначения:

F — множество всех конечных слов в алфавите A.
F_n — слова длины n.
U_n — слова длины не больше n.
F_w — все конечные подслова слова w.
$S_w = F \backslash F_w$ (все конечные слова, не встречающиеся в w).
I — бесконечные в обе стороны слова в алфавите A заданные с точностью до сдвига (формально: реализованные, например, как отображения из $\mathbb{Z}$ в A факторизованные по отношению эквивалентности сдвига).
$I_n \subset I$ — слова с наименьшим периодом n (которые можно себе представлять как n букв стоящие по кругу).
$I_\infty \subset I$ — непериодические слова. (Очевидно, $I = \left(\bigcup\limits_{n=1}^{\infty} I_n \right) \cup I_\infty$)

Определение 2.1. Будем называть любое множество $S \subseteq F$ системой запретов, а его элементы — запретами.

Определение 2.2. Будем говорить, что бесконечное слово $w \in I$ удовлетворяет системе запретов S если $\forall s \in S$ s не является подсловом w, или, что то же самое, $S \subseteq S_w$

Определение 2.3. Будем говорить, что система запретов S задает бесконечное слово $w \in I$ если оно и только оно удовлетворяет этой системе запретов.

Предложение 2.1. *Если система запретов S задает какое-то слово $w \in I$ и $|S| < \infty$ то w — периодическое.*

Это будет доказано ниже.

Предложение 2.2. $\forall n$ $\forall w \in I_n$ $\exists S,\ |S| < \infty$ *и* S *задает* w.

Например, $S = S_w \cap U_{n+1}$. И даже $S = S_w \cap F_{n+1}$. Это также будет доказано ниже.

В данной работе исследуется вопрос о наименьшем возможном количестве элементов в конченой системе запретов, задающей какое-нибудь слово из I_n в зависимости от n. Доказана

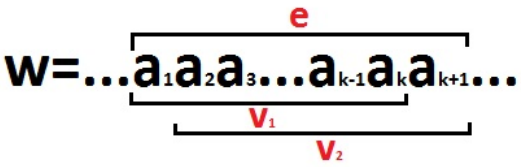
Теорема 1. *В случае двухсимвольного алфавита $A = \{a, b\}$ если слово из I_n задается с запретами, то $\varphi_c \geq n$, где φ_c — это с-е число Фибоначчи ($\varphi_1 = 1$, $\varphi_2 = 2$, $\varphi_3 = 3$, $\varphi_4 = 5$ и т.д.).*

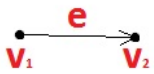
Также рассматривается вопрос минимальности этой оценки для каждого n и случай k-буквенного алфавита.

3 Доказательство оценки. Теоретическая часть.

3.1 Определение и простейшие свойства графов Рози

Определение 3.1.1. k-м графом Рози слова $w \in I$ называется граф G_k^w множеством вершин которого является множество $V = F_w \cap F_k$, а множеством ребер — множество $E = F_w \cap F_{k+1}$, где ребро $e \in E$ ведет из вершины v_1 равной первым k буквам e в вершину v_2 равную последним k буквам e — это естественная ориентация на G_k:





В случае, если понятно о графе Рози какого слова идет речь, верхний индекс будет опускаться. Чтобы не путаться, введем **Обозначения:**

$l(v)$ или $l(e)$ — слово, соответствующее вершине v или ребру e соответственно,
$v(l) \in G_{|l|}$ и $e(l) \in G_{|l-1|}$ — ребро и вершина для данного слова.

Для некоторой конечной строки $l \in F$ **Обозначения:**

$l[i]$ — i-я буква l
$l[i:j]$ — строка состоящая из букв l с i-й по j-ю в том же порядке.

Заметим, что переход от одной вершины к другой в направлении ориентации соединяющего их ребра соответствует переходу к рассмотрению подслова, сдвинутого на одну букву. Это позволяет понять что по путям в графах Рози можно выписывать слова, а по словам — строить пути в некотороых графах Рози. В частности:

Предложение 3.1.1. *Любому замкнутому пути $e_1 \cdots e_m$ в некотором графе Рози G_k^w соответствует слово $w_1 \in I_m$: $w_1 = \cdots l(e_m)[1] + l(e_1)[1] + l(e_2)[1] + \cdots + l(e_m)[1] + l(e_1)[1] \cdots$. Поскольку ребра последовательные, то все $l(e_i)$ встречаются как подслова в w_1 (начинающиеся в соответствующем $l(e_i)[1]$): $l(e_i) = l(e_i)[1] + l(e_{i+1})[1] \cdots l(e_{(i+k) \mod m})[1]$.*

Также с точностью до сдвига выписывается бесконечное слово (возможно, непериодическое) по любому бесконечному пути в G_k^w. С точностью до сдвига — т.к. кроме первой можно выписывать любую наперед заданную по порядку букву каждого ребра.

Предложение 3.1.2. *В любом графе Рози G_k^w есть бесконечный в обе стороны путь проходящий по всем ребрам и всем вершинам, соответствующий слову w.*

Доказательство. Выберем в слове w $k+1$ букву идущую подряд (подслово l_1). $e(l_1)$ — первое ребро пути. Сместим выбор на 1 букву вправо (подслово l_2). $e(l_2)$ — это второе. И так далее. Аналогично, влево. Очевидно, конец $e(l_i)$ совпадает с началом $e(l_{i+1})$. Очевидно этот путь проходит по всем ребрам и всем вершинам: он перебирает все подслова слова w длин k и $k+1$. □

В случае периодического слова $l_{i+n} = l_i$, и $e(l_{i+n}) = e(l_i)$, и можно рассматривать просто замкнутый путь длины n, т.к. остальное — его повторение.

Определение 3.1.2. Будем называть систему запретов S приведенной системой запретов, если она задает некоторое слово $w \in I$, не содержит дубликатов и любое из собственных подслов любого слова $s \in S$ является подсловом w.

Предложение 3.1.3. *По любой конечной системе запретов S, задающей некоторое бесконечное слово w можно построить приведенную приведенную систему запретов задающую то же слово и содержащую не большее число элементов.*

Доказательство. Если какое-то собственное подслово s_1 какого-то слова $s \in S \subseteq S_w$ так же не содержится в w — можно заменить s на s_1 и удалить дубликаты. Очевидно, w подходит новой системе (никакое слово в нем не содержится) и если какое-то слово подходит новой системе, то оно подходило и старой (т.к. запреты мы только укоротили и не добавляли) — а значит новая система тоже задает w. В силу конечности количества запретов и их длин — процесс когда-нибудь завершится и мы получим приведенную систему запретов. □

Изучим, как связанны G_k^w и G_{k+1}^w для некоторого периодического $w \in I_n$. Понятно, что вершины G_{k+1} взаимно соответствуют ребрам G_k — т.к. это подслова w длины $k+2$. А ребра G_{k+1} соответствуют путям длины 2 в G_k (парам последовательных ребер): ребру e в G_{k+1} соответствует пара ребер $e(\ l(e)[1:k+1]\ )$ и $e(\ l(e)[2:k+2]\ )$ — с общей вершиной $v(\ l(e)[2:k+1]\ )$.

Предложение 3.1.4. *А тем путям длины 2 в G_k, которым не соответствует ребро графа G_{k+1}, соответствует запрет в любой приведенной системе запретов S, задающей данное слово, точнее — существует биекция между запретами из S длины $k+1$ и такими путями в G_k, причем каждой такой паре последовательных ребер e_1 и e_2 соответствует запрет $l(e_1)+l(e_2)[k+1]$, а запрету s — пара ребер $e(\ s[1:k+1]\ )$ и $e(\ s[2:k+2]\ )$.*

Доказательство. Предположим, что найдутся два последовательно идущих ребра G_k e_1 и e_2 таких, что слово $x = l(e_1)+l(e_2)[k+1]$ не является подсловом w и не принадлежит S. Значит, никакой запрет из S не содержит x как подслово. Пусть слева от $l(e_1)$ в w написано бесконечное влево слово u, а справа от $l(e_2)$ — бесконечное вправо слово v. Тогда слово $w_1 = uxv$ не равное w т.к. в w не встречается x — будет удовлетворять системе S (если какой-то запрет из S содержится в w_1, то либо он содержит x, либо он содержится в $u + l(e_1)$ или $l(e_2) + v$ — чего быть не может). Противоречие. Для любого же запрета $s \in (S \cap F_{k+2})$, s не является подсловом w а $s[1:k+1]$ и $s[2:k+2]$ (поскольку система приведенная) — являются. Значит запрету s соответствует пара последовательных ребер в G_k, которым не соответствует ребра G_{k+1}. □

Обозначения:

$\tilde{v}(e) = v(l(e))$
$\tilde{e}(v) = e(l(v))$
$\tilde{e}(e_1, e_2) = e(l(e_1) + l(e_2)[k+1])$
$\tilde{p}(e) = (e(\ l(e)[1:k+1]\ ), e(\ l(e)[2:k+2]\ )$ — в последних двух функциях аргументом или значением в соответствующих случаях могут выступать элементы приведенной системы запретов.

Это значит, что по G_k очень легко нарисовать G_{k+1}:

Метод 0. В центре каждого ребра поставить по вершине, соеденить ребрами все пары последовательных ребер G_k, стереть все старое (от G_k) и зачеркнуть (точнее, стереть) все запрещенные ребра (ребра, принадлежащие $S \cap F_{k+2}$).

Определение 3.1.3. Будем называть произведение первой части данных действий „полушагом“, а второй (стирания ребер, соответствующих словам из S) — „применением запретов“. Всю совокупность этих действий (а точнее — переход от от рассмотрения G_k к G_{k+1}) будем называть „шагом“.

Обозначения:

$h(G_k)$ — граф, получающийся из G_k после полушага. (вообще, аналогичные операции можно проделать с любым ориентированным графом G. Результат — тоже ориентированный граф — будем обозначать $h(G)$). За $\tilde{p}$ будем обозначать естественное вложение вершин, ребер и развилок всех типов из G_{k+1} в $h(G_k)$.

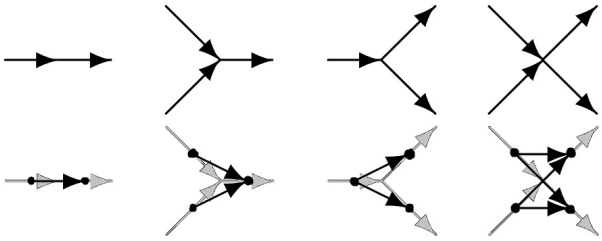
3.2 Развилки в графах Рози. Их количество.

Определения развилок

Заметим, что поскольку букв в нашем алфавите всего две, входящая и исходящая степень каждой вершины не больше двух (есть всего два способа продолжить на одну букву слово соответствующее данной вершине как влево, так и вправо). А поскольку есть замкнутый путь, проходящий через все ребра и все вершины (предложение 3.1.2) — эти степени не меншье одного. То есть есть четыре типа вершин (указаны соотв. входящая и исходящая степени): (1,1), (2,1), (1,2) и (2,2).

Определение 3.2.1. Назовем такие вершины соответственно дорога, входящая развилка, исходящая развилка и перекресток.

При полушаге они преобразуются так:



В дальнейшем для нас будет важно понятие развилок. Стоит обратить внимание что можно ввести три разных (однако почти эквивалентных по свойствам) определения развилок, которые каждое по своему удобны для доказательства некоторых утверждений. Первое уже введено.

Определение 3.2.2. Будем называть входящей d-развилкой пару различных ребер, имеющих общий конец, а исходящей — общее начало.

Определение 3.2.3. Будем называть входящей t-развилкой пару из входящей d-развилки и „направляющего“ ребра (возможно — принадлежащего этой d-развилке), начало которого совпадает с ее концом , а исходящей — наоборот.

При использовании каждого определения будем употреблять соответственно слова „развилка“, „d-развилка“ и „t-развилка“ за исключением случаев, которые будут оговорены отдельно.

Обозначения:

c_k^w — количество перекрестков в G_k^w

f_k^w — количество входящих развилок (на самом деле, входящих или исходящих — не важно, т.к. они равны — это будет доказано в следующем абзаце)

d_k^w — количество входящих d-развилок

t_k^w — количество входящих t-развилок.

Если понятно о каком слове идет речь, верхний индекс будет опускаться.

Предложение 3.2.1. *$\forall w$ f_k^w равно количеству исходящих развилок в G_k^w.*

Доказательство. Тривиально следует из того что сумма входящих степеней в любом G_k^w равна сумме исходящих. □

Предложение 3.2.2. *$d_k^w = f_k^w + c_k^w$, $t_k^w = f_k^w + 2 * c_k^w$. Для исходящих d- и t-развилок аналогично.*

Доказательство. Сопоставим каждой d-развилке ее конец. Его исходящая степень может быть 1 или 2. Для t-развилок аналогично. Для исходящих d- и t-развилок аналогично. □

Значит и количества входящих d- и t-развилок тоже равны соответствующим количествам исходящих.

Соответствия между развилками, перекрестками, d- и t-развилками в одном и различных графах Рози

Как уже упоминалось, определения развилок, d-развилок и t-развилок почти эквивалентны. В частности — между ними есть естественные соответствия:

Развилке — соответствуют одна d-развилка и одна t-развилка (входящей — входящие, исходящей — исходящие): $\tilde{d}(f)$ и $\tilde{t}(f)$ (поскольку просто развилками мы называем вершины графов Рози, аргументами могут быть соответствующие вершины).

Перекрестку — две d-развилки (входящая и исходящая) и четыре t-развилки (две входящие и две исходящие). $\tilde{d}(c)$ и $\tilde{t}(c)$ (Аналогичное замечание. Кроме того, образом многозначного соответствия мы естественно будем считать множество из всех соответствующих элементов).

d-развилке — ее вершина (развилка соотв. типа или перекресток) и соответствуеющие ей t-развилки: $\tilde{v}(d)$ и $\tilde{t}(d)$.

t-развилке — тоже: $\tilde{v}(t)$ и $\tilde{d}(t)$.

Заметим, что в случае развилки это соответствие взаимно однозначное, а в случае перекрестка — нет.

Заметим также, что все свойства графов Рози для $h(G_k)$ выполняются (что и понятно — он мог бы быть графом Рози какого-нибудь слова, если бы S не содержала запретов длины $k+2$). Значит можно говорить о развилках в нем.

Предложение 3.2.3. *Существует биекция между t-развилками в G_k и d-развилками в $h(G_k)$ (и значит их количества равны), причем t-развилке соответствует d-развилка, вершина которой соответствует направляющему ребру этой t-развилки.*

Доказательство. Сопоставим естественным образом вершины и ребра $h(G_k)$ (k+1)- и (k+2)-буквенным словам (вершине поставленной в центре ребра e — слово $l(e)$, а ребру из e_1 в e_2 — слово $l(e_1)+l(e_2)[k+1]$. Теперь зададим биекцию: каждой t-развилке t_1 в G_k состоящей из d-развилки d (из ребер d_1 и d_2) и ребра e_1 будет соответствовать d-развилка из ребер соотв. путям d_1e_1 и d_2e_1, а d-развилке в $h(G_k)$ будет соответствовать t-развилка из пары ее начальных вершин и конечной вершины — т.к. начальные вершины не совпадают: нигде кроме G_0 не может два разных ребра начинаться и заканчиваться в одинаковых вершинах — иначе у конечной вершины последняя буква должна быть равна и „a“, и „b“, а такого быть не может (заметим, что с тем же успехом можно было рассмотреть первую букву начальной вершины). □ **Обозначения:**

Обозначим это соответствие за $\tilde{h}(t)$ (Обратное можно обозначать и как $\tilde{h}^{-1}(d)$, и как $\tilde{h}(d)$ — т.к. оно однозначно определяется типом аргумента)

Заметим, что если вершина d-развилки в $h(G_k)$ — развилка, то ей соответствует единственная t-развилка. Аналогично если вершина t-развилки в G_k — развилка, то ей соответствует единственная d-развилка. Значит, на множестве всех d-развилок во всех G_k данного слова можно ввести отношение эквивалентности порожденное следующими равенствами: две d-развилки эквивалентны, если они находятся в последовательных графах Рози G_k и G_{k+1} и однозначно друг другу соответствуют в указанном естественном смысле (если вершина той что в G_k — развилка). Аналогично для t-разивлок. Формально: $d_1\ d_2$, если $\tilde{p}(d_2)=\tilde{h}(\tilde{t}(d_1))$, (или наоборот).

Определение 3.2.4. Соответствующие классы эквивалентности будем называть md-развилками и mt-развилками.

В дальнейшем нам будет удобно говорить о какой-то конкретной md- или mt-развилке просто указав один из ее элементов. Заметим, что d- или t-развилки из одного класса эквивалентности содержатся в некотором количестве подряд идущих графов Рози (по 1 в каждом). Для удобства терминологии будем представлять себе k как ось дискретного времени (для G_k) и говорить, что md- или mt-развилка „существует в момент времени t“ или „принадлежит G_t“, если какой-то ее элемент содержится в G_t.

Очевидно, что существование каждой md- и mt-развилки начинается перекрестком и заканчивается перекрестком или запретом (на самом деле, заканчиваться оно может тоже только перекрестком, т.к. как будет видно далее, в случае приведенной системы запретов, запреты могут происходить только сразу через шаг после перекрестков, а значит этой развилки бы просто не существовало вообще).

Количество развилок

Предложение 3.2.4. *Количество d-развилок при полушаге увеличивается на c_k.*

Доказательство. Очевидно следует из предложений 3.2.2 и 3.2.3. □

Предложение 3.2.5. *При применении запретов количество входящих d-развилок уменьшается на количество примененных запретов (стертых ребер).*

Доказательство. Пусть ребро x с вершинами v_1 и v_2 из G_{k+1} должно быть стерто (содержится в приведенной с.з. S). Исходящая степень v_1 как и входящая степень v_2 в G_{k+1} не меньше одного.

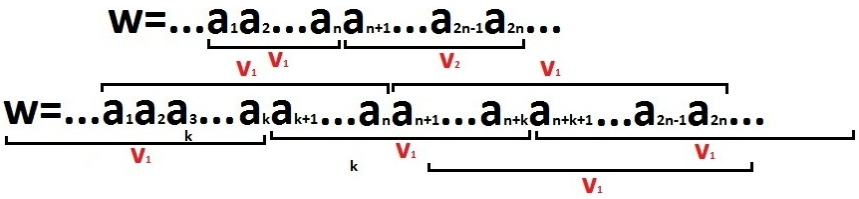
Значит в $h(G_k)$ эти степени не меньше 2. Но они не могут быть больше 2. Значит они ровно 2. Очевидно, одно ребро не может содержаться в двух разных входящих d-развилках. Значит количество входящих d-развилок при стирании этого ребра уменьшится ровно на 1. Далее индукция. □

Лемма 1. *$\forall w \in I\ \forall k \geq 1,\ d_k^w = d_{k-1}^w + c_{k-1}^w - |S \cap F_{n+1}|$, где S — приведенная система запретов, задающая w.*

Доказательство. Очевидно следует из предложений 3.2.4, 3.1.4 и 3.2.5. □

Лемма 2. *Если $w \in I_n$, то G_n^w — циклический граф содержащий n ребер (то есть $d_n^w = 0$ а $|G_n^w| = n$).*

Доказательство. По условию $v_1 = v_2$ (см. рисунок), значит в G_n есть замкнутый путь по n ребрам (т.к. при n сдвигах рассматриваемого подслова на одну букву мы получим то же подслово). Он не самопересекается, иначе у w есть меньший период ("$a_{2k+1}...a_{k+n}$"="$a_{k+1}...a_n$"="$a_1...a_{n-k}$"; "$a_{k+n+1}...a_{2k+n}$"="$a_1...a_k$"="$a_{n-k+1}...a_n$") — равный k (где k это количество ребер до самопересечения, а v_1 — вершина, в которой путь самопересекается). □



Долг.

Теперь настало время вернуть долг. Докажем следующие утверждения:

Предложение 2.1: если конечная система запретов S задает какое-то слово $w \in I$, то оно периодическое.

Доказательство. Пусть $k+1$ — наибольшая длинна запрета в S. Рассмотрим G_k^w. Если это цикл — то слово w периодическое, если нет, то помимо наименьшего замкнутого пути l соответствующего w в этом графе есть еще хотя бы 1: поскольку l самопересекается, можно рассмотреть замкнутый подпуть l от точки самопересечения до нее же (меньший чем l — иначе это не точка самопересечения) — и ходить по нему до бесконечности. Слово w_1 соответствующее этой циркуляции не равно w (т.к. в нем встречаются не все слова длины $k+1$, которые встречаются в w). С другой стороны — оно удовлетворяет S, т.к. содержит только те подслова длины $k+1$, которые встречаются в w, и если какое-то слово из S содержится в w_1, то оно содержится в каком-то его подслове длины $k+1$ — а значит и в w. Противоречие.

Заметим, что этим мы доказали еще один общий факт: любое бесконечное периодическое слово, соответствующее некоторому замкнутому пути в G_k^w удовлетворяет $S \cap F_{k+1}$ (если S — приведенная система запретов задающая слово w). □

Предложение 2.2: для любого $w \in I_n$ существует конечная система запретов, задающая его.

Доказательство. Расмотрим n-й граф Рози G_n^w этого слова. Мы доказали, что это цикл длины n. Очевидно, слово w удовлетворяет системе запретов $S = S_w \cap F_{n+1} \subset S_w$. Если какое-то другое слово w_1 тоже удовлетворяет ей, то $G_n^{w_1} \subseteq G_n^w$, т.к. он не может содержать никаких других ребер. Чтобы в нем был хоть один бесконечный в обе стороны путь (соответствующий, собственно, слову w_1) необходимо $G_n^{w_1} = G_n^w$. Тогда этоть путь единственен, а слово w_1 соответствующее ему — совпадает с w. Противоречие. □

Предложение 3.2.6. *Для данного $w \in I_n$ существует единственная приведенная система запретов, задающая его, и состоит она в точности из тех слов, которые обнаруживаются при рассмотрении графов Рози.*

Доказательство. Тривиально следует из предложений 2.2, 3.1.3 и 3.1.4 □

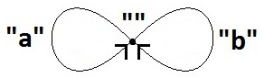
Заметим, что в этом разделе, как и в разделе 2.1 не используется двухсимвольность алфавита A

Основной вывод

Из лемм 1 и 2, а также предложения 3.2.6 очевидно

Следствие 3. $|S| = \sum\limits_{k=1}^{\infty} |S \cap F_k| = \sum\limits_{n=1}^{\infty} (d_{k-1}^w - d_k^w + c_{k-1}^w) = \sum\limits_{k=0}^{\infty} c_k + d_0$

Начинать рассмотрение целесообразно с G_0 (вершина — пустое слово, и два ребра — „a“ и „b“). При этом можно не рассматривать тривиальные случаи когда запрещено a или b, т.к. тогда оценка выполняется и точна (период слова состоящего только из буквы „a“ или „b“ равен 1). Тогда начальное количество развилок равно 1. G_0:



3.3 Размер графов Рози

Изучим как изменяется $|G_k|$ (количество ребер в G_k).

Лемма 4. $|G_{k+1}^w| = |G_k^w| + (d_k^w + c_k^w) - |S_w \cap F_{n+1}|$

Доказательство. Поскольку ребра $h(G_k)$ взаимно соответствуют путям длины два в G_k, их количества равны. Сопоставим каждому пути его первое ребро. Значит, количество таких путей равно $|G_k|$ плюс количество ребер, конечная вершина которых имеет исходящюю степень 2. Количество таких ребер равно $2c_k + f_k = t_k = d_k + c_k$.
$|G_{k+1}^w| - |h(G_k)^w| = |S_w \cap F_{n+1}|$: на каждый запрет мы стираем по ребру. □

Значит, согласно леммам 2 и 4, а также следствию 3 верно:

Следствие 5. *$\forall w \in I_n$ верно равенство:*

$$n = |G_n^w| = |G_0| + \sum_{k=0}^{n-1} (|G_{k+1}^w| - |G_k^w|) =$$

$$= |G_0| + \sum_{k=0}^{n-1} ((d_k^w + c_k^w) - |S_w \cap F_{n+1}|) =$$

$$= |G_0| + \sum_{k=0}^{n-1} d_k^w + \sum_{k=0}^{n-1} c_k^w - |S| = |G_0| + \sum_{k=1}^{n-1} d_k^w$$

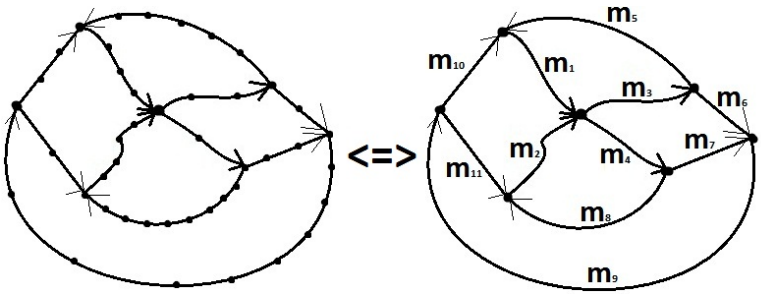
По сути это значит что на каждую d-развилку в каждом графе Рози приходится увеличение конечного количества ребер (и, соотвтетственно, периода) на 1. Остается только найти разумный способ сопоставить некоторые группы d-развилок запретам (или, поскольку их количества равны, перекресткам) так, чтобы получить необходимую оценку — что мы в конечном счете и сделаем (на самом деле, эти группы — те самые md-развилки, и каждая из них будет сопоставляться перекрестку, которым заканчивается, но как видно из утверждения теоремы рост конечного количества ребер от количества перекрестков может быть экспоненциальным, и значит важно в каком порядке

3.4 Другие способы нарисовать G_{k+1} по G_k

Сейчас мы опишем и докажем корректность другого, необходимого для доказательства способа построения G_{k+1} по G_k. Для лучшего понимания, упрощенного обоснования и ясной мотивировки мы приведем 8 разных способов построения, изменяющихся постепенно от имеющегося способа к необходимому.

Для первого способа нам понадобится понятие схемы.

Определение 3.4.1. Схемой ориентированного графа G будем называть взвешенный ориентированный граф $\widetilde{G}$ множеством вершин которого является множество вершин G не являющихся дорогой, а множеством ребер — пути в G соединяющие такие вершины и не содержащие ни одной из них кроме как своим началом и концом (с весами равными длинам этих путей). Естественно, ориентация ребер в путях должна быть согласована и наследуется соответствующему ребру (из этого ясно, что такие пути не самопересекаются).



Другими словами, схема графа Рози — это взешенный ориентированный граф, полученный из графа Рози заменой цепочек идущих подряд без развилок ребер на одно ребро с весом равным длине этой цепочки (и соответствующей ориентацией). Она наглядно отражает структуру графов.

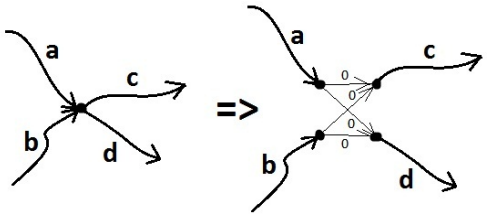
Формально: рассмотрим отношение эквивалентности порожденное следующими равенствами: два последовательных ребра эквивалентны, если их общая вершина — дорога. Очевидно, в каждом классе эквивалентности (если только граф не является циклом) будет ровно по одному ребру не начинающемуся и не заканчивающемуся в „дороге"(b и e). Построим новый взвешенный ориентированный граф с множеством вершин равным множетсву развилок и перекрестков, а множеством ребер — для каждого класса эквивалентности по ребру: из начальной вершины ребра b в конечную вершину ребра e с весом равным количеству ребер в классе эквивалентности. Также легко производится обратная операция (при этом вершины, соединенные ребром с весом 0 отождествляются).

Заметим, что на схемы корректно переносятся все определения развилок — вес можно просто не учитывать.

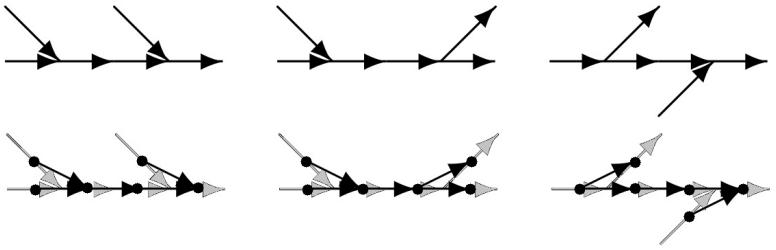
Мотивация введения схем: это необходимо для введения фиктивных ребер с весом ноль в некоторых местах с целью удобства подсчета периода равного сумме приростов количеств ребер, соответствующих md-развилкам, соответствующих запретам.

Используя понятие схемы предложим первый новый способ нарисовать G^w_{k+1} по G^w_k:

Метод 1. Рассмотрим $\widetilde{G^w_k}$. Заменим в нем перекрестки на фиктивную конструкцию эквивалентную той, что получается из перекрестка при полушаге, только с весами 0 (см. рисунок) — назовем эту операцию 0-заменой. Тогда перекрестков не останется — только развилки. $\widetilde{G^w_k}$ изменим по правилу: если ребро закначивается однотипными развилками, его вес не изменится, если разными — то $+1$, если его начало — входящая развилка и -1, если исходящая (по сути — это тот же полушаг). Теперь по схеме восстановим обратно граф (заменой ребер с весами на цепочки — и получится точно такой же граф, как после полушага), а после этого применим запреты.



Доказательство. Для доказательства корректности этого метода достаточно доказать что после перехода обратно от схемы к графу мы получим такой же граф, как после полушага. Заметим, что количество t-развилок в G_k очевидно равно количеству развилок в схеме после 0-замены, и они естественным образом отождествляются: если вершина t-развилки была развилкой, то ей она и соответствует, а если перекрестком, то ей соответствует та развилка, в которой содержится ее „направляющее“ ребро (оно принадлежит ровно одному из путей a,b,c,d — см. рисунок). Каждой t-развилке соответствует d-развилка в $h(G_k)$. Остается заметить, что длины путей между d-развилками в $h(G_k)$ — точно такие же как веса ребер между соответствующими им развилками в нашей схеме после преобразования — возьмем пару развилок в $\widetilde{G_k}$ соединенных некоторым ребром. Если его вес 0, то после преобразования он будет равен 1 (т.к. это могли быть только развилки получившиеся заменой перекрестка) — как и после полушага, а если нет, то см. рисунок. □



Обозначения:

$h^*(G)$ — результат применения к графу G такой операции (пути между однотипными развилками не меняем, а между разнотипными — сокращаем на 1 если начало — входящая, а конец — исходящая, и увеличиваем на 1, если наоборот).

Заметим, что метод 1 принципиально отличается от метода 0 тем, что описывает преобразование графа не локально, а глобально: становится видно, что структура графа меняется только в местах перекрестков, а в остальных местах только изменяются расстояния. Причем противоположные развилки как бы движутся навстречу друг другу в направлении стрелок которые изображают. Заметим также, что перекрестки образуются по сути при встречах противоположных развилок.

Метод 2. Второй способ отличается от первого только тем, что запреты применяются до перехода от схемы обратно к графу Рози (удаляется соответствующее ребро).

Доказательство. Для доказательства корректности этого метода вспомним, что как уже было объяснено в доказательстве предложения 3.2.5 запреты происходят с ребрами для которых исходящая степень начальной вершины и входящая степень конечной равны 2. Но это значит что до перехода обратно от схемы к графу начальной и конечной вершинам соответствовали исходящая и входящая развилка, а нашему ребру — ребро с весом 1 (и именно его мы удалим в нашей операции). Более того, это значит что до преобразования вес этого ребра был 0, и оно появилось в схеме после 0-замены. □

Метод 3. А третий — тем, что теперь запреты вообще применяются до преобразования схемы. Результат будет эквивалентен, поскольку запреты ведь происходят только сразу после размножения, то есть в местах перекрестков — а значит их можно сделать в самих наших фиктивных конструкциях.

Доказательство. Поскольку уже показано что запреты будут применяться только к ребрам с весом 0 добавленным при 0-замене перекрестков, необходимо доказать только что если сначала применить запреты, а потом сделать преобразование схемы — получится то же самое, как если сначала сделать преобразование, а потом применить запреты. Для этого достаточно доказать что графы будут иметь одинаковую структуру и соответствующие веса будут равны. Один случай, для наглядности, разобран на картинке. Поскольку структура схемы не меняется при нашем переходе (только веса) — утверждение про идентичность структур очевидно (будут зачеркнуты одни и те же ребра). Заметим, что согласно нашей операции сумма весов на любом пути от одной развилки до другой при переходе меняется ровно так, как если бы на этом пути было одно ребро (очевидно из рассмотрения типов развилок на этом пути или картинки в доказательстве к первому методу — путь между развилками там мог быть любой, не обязательно не содержащий других развилок) — значит расстояния между оставшимися после наших действий развилками не зависят от последовательности этих действий. □

Вообще, весь этот переход к схемам не обязателен, и для решения важно только иметь возможность разрешать перекрестки заменой данной врешины фиктивной структурой размера 0, но описывать преобразование графа „там где развилки разнотипные одно ребро с весом один исчезло, а в другом месте ребро появилось“ все-таки менее удобно чем говорить про изменение веса.

Метод 4. Теперь все-таки попробуем обойтись без перехода к схемам. Просто изменим каждый путь между двумя развилками в графе соответствующим образом (полушаг), а потом применим запреты. При этом перекрестки можно обрабатывать по-разному: можно считать что у всех ребер бывших в G_k веса 1, а с перекрестками делать 0-замену, а затем делать описанное только что преобразованее (при этом ребро веса 0 — как бы путь длины 0 — заменяется на 1 ребро с весом 1 — путь длины 1). Или можно считать что в месте перекрестка находятся 4 t-развилки на расстоянии 0 (между противоположными из них), и когда рассматривается какой-то путь содержащий своим концом перекресток — он считается путем до той развилки, направляющее ребро которой в нем содержится).

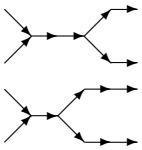
Доказательство. Этот метод очевидно эквивалентен методу 1. □

Метод 5. Теперь сначала полностью разрешим перекрестки, а потом сделаем полушаг.

Доказательство. Подробнее: Рассмотрим граф Рози G_k. Каждому его ребру присвоим вес равный 1. Произведем 0-замену. Удалим ребра, соответствующие ребрам удаляемым в $h(G_k)$ (мы знаем, что удаляемые из $h(G_k)$ ребра соответствуют ребрам в схеме $h(\tilde{G}_k)$ с весом 1 и противоположными развилками на концах, а они соответствуют ребрам с весом 0 в текущем графе). Изменим длины путей в графе по описанным ранее правилам и отождествим вершины соединенные ребрами с весом 0. Веса забудем. Получим G_{k+1} — это следует из метода 3. □

Согласно замечанию о том, что длина любого пути между двумя развилками меняется только в зависимости от их типов, можно представить себе, что, согласно тому что расстояния между однотипными развилками не меняются, одни из них стоят на месте, а другие — едут им на встречу со скоростью один.

Метод 6. С одной стороны — этот способ принципиально отличается от всех предыдущих, и содержит существенно важную для решения идею, а с другой — отличается от предыдущих только взглядом на вещи. Суть его вот в чем: сначала, как уже много раз делали, заменим все перекрестки на фиктивные конструкции размера 0 (очевидно, перекрестков у нас после этого не останется). А затем сделаем некоторую новую операцию, результатом которой будет то же самое что при полушаге, но процедура другая: выберем тип развилок (входящие или исходящие). Предположим без о.о. что входящие. Будем считать, что входящие t-развилки стоят на месте, а исходящие — сдвигаются на 1 в ту сторону в которую они „указывают“, поглощая одно ребро из пути перед собой, и выплевывая два — по одному в каждый из хвостов за собой.

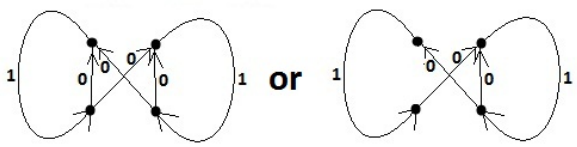


Доказательство. Заметим, что все пути между развилками изменятся строго по нашему правилу. Значит этот метод очевидно эквивалентен методам 1, 2 и 4. □

Метод 7. То же самое, но теперь будем сначала сразу разрешать все перекрестки, а потом сдвигать исходящие развилки к входящим.

Доказательство. Эквивалентен методам 3 и 5. □

Метод 8. Строго говоря, это уже не способ получить следующий граф Рози по предыдущему, и графы Рози данного слова будут появляться только как промежуточная фаза каждой такой операции, но поскольку для исследуемого нами вопроса важно только количество произведенных за все операции запретов и количество ребер в конечном цикле — это не важно. Здесь мы будем сначала делать полушаг в новом смысле, а потом разрешать все образовавшиеся перекрестки. Начинать в этом случае будем с графа $\widehat{G_0} = \mathbb{G}_2$ полученного из G_0 разрешением перекрестка. Варианты $\mathbb{G}_2$ (такой выбор индекса будет пояснен чуть позднее):



Доказательство. Однако нас интересует только общее количество запретов (или, эквивалентно, общее количество перекрестков) встретившихся в графах в течение всего процесса и (так как очевидно, как и графы Рози наши графы станут с некоторого момента одинаковыми циклами) — конечное количество ребер. Они, очевидно, будут равны соответствующим количествам для последовательности графов Рози некоторого данного слова (поскольку последовательность, фактически, выглядит так: $\widehat{G_0} \to G_1 \to \widehat{G_1} \to G_2 \to \cdots$ □

3.5 Что мы получили

Имеется некоторое бесконечное в обе стороны периодическое слово $w \in I_n$ и задающая его приведенная система запретов S. Имеется последовательность графов Рози этого слова $G_0^w, G_1^w, \cdots, G_n^w$ (заканчивающаяся циклическим графом длина которого равна периоду слова n) и совокупность вспомогательных графов $\{h(G_i^w)\}$, с естественными биециями между множествами $E_{G_i^w} \leftrightarrow V_{h(G_i^w)}$ и $L^2_{G_i^w} \leftrightarrow E_{h(G_i^w)}$ ($L^2_{G_i^w}$ — пути длины 2 в G_i^w, а также множествами $V_{h(G_i^w)} \leftrightarrow V_{G_{i+1}^w}$ и $E_{h(G_i^w)} \leftrightarrow E_{G_{i+1}^w} \sqcup (S \cap F_{i+1})$

Также имеется последоваетельность графов $\widehat{G_0^w}, \widehat{G_1^w}, \cdots, \widehat{G_n^w}$, таких что $h^*(\widehat{G_i^w}) = G_{i+1}^w$, которые строятся индуктивно по следующим правилам: $\widehat{G_i^w} \to h^*(\widehat{G_i^w}) \overset{0-exchange}{\longrightarrow} \widehat{G_i^w}' \overset{restrictions}{\longrightarrow} \widehat{G_{i+1}^w}$, где запреты происходят в тех же ребрах, где и в $h(G_i^w)$, структурно равном $\widehat{G_i^w}'$.

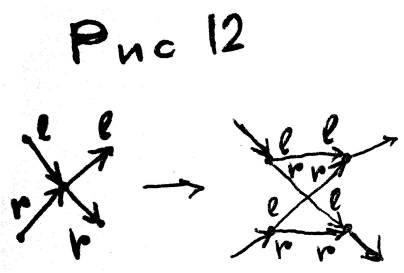
3.6 Что дальше

В процессе применения 8-го метода абсолютно наглядно на каждую из d_k^w d-развилок какого-то типа за 1 шаг (от $\widehat{G_{k-1}^w}$ к $\widehat{G_k^w}$) приходится прирост общего количества ребер на 1, на каждый перекресток (из c_{k+1}^w) — прирост количества d-развилок на 1, на каждый запрет — уменьшение количества d-развилок на 1, и все эти приросты-уменьшения происходят параллельно. Ключевая идея доказательства оценки — в том чтобы делать это последовательно, на каждом шагу с какой-то одной развилкой, с которой это сделать возможно: если в данном графе перед какой-то исходящей (в направлении ее движения) развилкой до ближайшей входящей нет других исходящих — единомоментно проведем ее до этой входящей развилки через этот путь и разрешим образовавшийся перекресток (применив запреты так же, как мы собирались это сделать при планомарном изменении графа — параллеельном движении развилок). Заметим, что поскольку количество запретов равно количеству перекрестков — мы можем считать только последние. Так, при каждой такой операци количество перекрестков увеличивается на 1, а $|\mathbb{G}_k|$ увеличивается на длину пути, по которому мы провели развилку (вот мы и сопоставили каждому запрету прирост количества ребер — ту самую md-развилку). Соответственно целесообразно индексировать наши графы количеством учтенных перекрестков (поэтому и индекс начального графа — 2: для превращения $\mathbb{G}_2$ в цикл необходимо два запрета). Как уже можно было заметить, термин „шаг“ используется автором для любой минимальной циклически повторяющейся операции над графом. Теперь мы будем использовать его для обозначения последней из описанных — передвижения развилки и разрешения появившегося перекрестка.

Теперь фиксируем слово ($w \in I_n$) и его приведенную систему запретов S (впрочем, и так единственную).

Рассмотрим $\mathbb{G}_2 = \widehat{G_0}$. Как мы уже говорили, есть два, с точностью до перемены a и b местами варианта. Рассмотрим в последовательности $\widehat{G_i}$ l-развилки (развилки естественным образом отождествляются с наследниками в методе 8 — если не произошло 0-замены). Вне завсисмости от

того как выглядит $\widehat{G_0}$ для каждой развилки в нем каждому (из двух) ее хвостов (входящему в нее ребру для входящей или исходящему для исходящей) сопсотавим метку „left“ или „right“ (одну одному, а другую — другому). Опишем как наследуются эти метки: при h^* — естественным образом (последнему ребру в пути, в котором было ребро „left“ — метку „left“, аналогично для „right“). При 0-замене — тоже довольно естественно:



(После применения запретов какие-то из нарисованных развилок исчезнут, а какие-то останутся. Вместе с метками). В общем-то получается, что на время существования l-развилки в графе для нее вполне естественно определены левый и правый хвост. Понятно, что с помощью этих меток однозначно определяется какое ребро запрещается после 0-замены (даже — меньшей информацией — достаточно пометить развилки только одного типа — допустим, только входящие). Также заметим, что для каждой l-развилки однозначно определяется ее история (после взаимодействия каких пар развилок противоположного типа она появилась и левым или правым хвостом являлись соответствующие развилки) и то с какой развилкой она в итоге размножилась (из множества всех l-развилок противоположного типа), да и вообще — определена полная история взаимодействий (какие пары развилок размножились, какие ребра были запрещены и кто были их дети).

Возьмем $\mathbb{G}_2$. Расставим в нем как-нибудь метки (и установим соответствие с l-развилками в $\widehat{G_0}$). Будем производить в произвольном порядке следующие (уже ранее описывавшиеся) операции: если в текущем графе на пути между двумя данными развилками (входящей и исходящей — в указанном порядке) нет других развилок (причем путь проходит через их направляющие ребра), то сдвинем конкретно эту исходящую (для определенности) развилку по описанным правилам (одно ребро поглощает, два выдает) до входящей, произведем 0-замену образовавшегося перекрестка, расставим метки „left“ и „right“, удалим соответствующие ребра (такие же, как те, которые были удалены в изначальной последовательности графов $\widehat{G_i}$), оставшимся развилкам поставим в соответствие l-развилки этой последовательности — те, которые получились при размножении соответствовавших текущим развилкам l-развилок.

Предложение 3.1. *В таком случае эта операция всегда определена корректно (те l-развилки которые соответствуют двум развилкам с которыми мы производим данную операцию в исходной последовательности размножались именно друг с другом), в конце эта последовательность станет циклом (эта часть очевидно следует из предыдущей) и длина цикла будет равна длине финального цикла исходной последовательности графов — вне зависимости, в частности, от выбора на каждом шаге конкретной пары развилок между которыми ничего нет.*

Доказательство. Вообще, это очевидно — если представить, что в исходной последовательности развилки двигаются паралельно (не обгоняя друг друга), встречаются, размножаются, происходят какие-то запреты — то, если все происходит идентично — очевидно не важно в какой последовательности это происходит если они друг друга не обгоняют (т.к. при встрече развилок пути переклеиваются идентично).

Рассмотрим схемы графов из процесса применения метода 8. Как мы уже говорили, при переходе от графа к графу меняются только веса ребер и небольшие фрагменты графа в местах встречи развилок (когда ребро становится равным 0). Заметим, что согласно с этим ребра могут появляться, исчезать и объединяться. Всё это — только при встрече развилок. Рассмотрим множество всех ребер когда-либо появившихся в этих графах $\mathbb{E}$ (можно считать что при встрече развилок появляются 4 ребра, а потом мы сразу некоторые из них убиваем, а можно — что появляются только те из них которые остаются, и некоторые сразу объединяются с теми между которыми убита развилка. Для удобства и ясности примем первое) . Мы помним, что у нас уже есть множество размеченных развилок, и соответственно легко установить в последовательности $\mathbb{G}_k$ какие ребра каким соответствуют

при появлении (опять же рассматривая схемы). При объединении двух ребер будем писать на нем слово в алфавите $\mathbb{E}$ получающееся приписыванием текущего слова на первом ребре к текущему слову на втором (если объединяются три подряд — тоже понятно как). При появлении на ребре будем писать однобуквенное слово из того ребра что ему соответствует. Заметим, что в каждом графе на концах ребра находятся какие-то развилки. Если развилка „смотрит“ в сторону ребра, концом которого она является — она, очевидно, будет его концом до исчезновения этого ребра (возможно после его объединения с какими-нибудь другими ребрами на другом конце) из-за встречи этой развилки с противоположной. Если эта развилка смотрит „от“ — она встретится с кем-нибудь там с той стороны, и у этого ребра либо поменяется конец на развилку смотрящую „к“, либо припишется с этой стороны какое-нибудь ребро. Утверждение: в каждом графе $\mathbb{G}_k$ на ребрах будут написаны только слова, являющиеся подсловами слов написанных на ребрах процесса метода 8 в момент их смерти (то есть после максимального объединения слов), а также у каждого ребра на конце будет та развилка, которая должна быть — либо смотрящая в его сторону и одновременно та, с которой он должен умереть (можно считать что развилки тоже буквы — тогда просто это то самое условие на подслова), либо смотрящая в противоположную сторону и та, которая при размножении заменится соотв. на нужную смотрящую в его сторону, либо приведет к приклеиванию однобуквенного слова — того, которое идет перед данным словом (в том самом длинном слове в котором оно встречается как подслово — очевидно такое одно, т.к. буквы во всех словах разные и не повторяются).

Заметим, что по сути наше утверждение (о том что операции можно делать в любом порядке) — есть утверждение об ассоциативности произведения операций. Соответственно его доказательство также будет сродни классическим доказательствам ассоциативности.

Всё это доказывается по индукции. Для $\mathbb{G}_2$ это верно, т.к. там всё так же как в G_1. Пусть для $\mathbb{G}_k$ данное утверждение верно. При переходе к $\mathbb{G}_{k+1}$ мы производим всего одну операцию — встречаются две развилки находящиеся на концах ребра, с которым они должны умереть. Очевидно значит, они размножались именно друг с другом в основном процессе (т.к. вместе с одним ребром умирают только две конкретные развилки). Нас интересуют четыре ребра концами которых являлись эти развилки с других сторон. Условия на граф обеспечивают то что при 0-замене к уже имевшимся ребрам приклеились именно те развилки или ребра, которые нужно. Остается только проверить, что для четырех новопоявившихся ребер верно следующее: либо они исчезнут, либо у них на концах те развилки которые нужно. Это просто: поскольку появление ребер и развилок естественным образом привязано друг к другу (мы знаем что в основной последовательности графов эти ребра и развилки появились вместе — мы их так сопоставляем) — то на концах ребер развилки смотрящие „от“, после которых прибавится то что нужно — либо развилка с которой мы встретим сметрь, либо однобуквенное слово, которое и должно быть (потому что так было в основном процессе). Корректность встречи как подслов индуцируется тем, что слова в основном процессе приписывались с „таким“ упорядоченьем букв — хоть может быть и в разной очередности во времени.

Конечное количество ребер всегда будет одно и то же, т.к. оно просто равно сумме начального количества ребер плюс для каждого ребра на котором написано финальное слово (то, с которым это ребро умрет, а не к нему что-нибудь приклеится) — количество раз сколько его левая граница сдвигалась вправо, и количество раз сколько его правая граница сдвигалась влево (в нашем предположении одни из них всегда стоят).

Будем считать, что все наши действия разбиты на малюсенькие части — сдвиги одной развилки на 1. При этом происходит -1 к длине одного ребра и +1 к длинам двух. Ребра делятся на те, которые когда-нибудь умрут, и те которые не умрут (по сути — будут частью слова написанного на конечном цикле графа). Заметим, что количество раз сколько концевые развилки для данного слова будут придвинуты друг к другу равно количеству раз сколько до этого развилки в которых заканчиваются подслова этого слова будут отодвинуты (это почти очевидно). По сути — конечная длина цикла — количество раз, которые будут отодвинуты развилки на концах слов, которые не умрут.

Если развилки одного типа стоят, а другого — двигаются, то нам не важно (для количества) в каком порядке они двигаются (если они не обгоняют друг друга) — т.к. очевидно вне зависимости от порядка каждая развилка пройдет до своей пары одно и то же расстояние — это можно доказать по индукции. Заметим, что любая развилка за время своего существования не меняет ближайшие к ней буквы написанные на ребрах заканчивающихся в ней. Ну, а поскольку всё корректно — значит эти ребра являются подсловами одних и тех же конечных слов (т.к. данные буквы встречаются каждая только в одном конечном слове). Значит каждая развилка с момента появления до момента

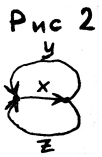
смерти дает прирост по $+1$ к словам содержащим как подслова оба ее хвоста. А время ее жизни равно сумме времен жизни всех тех развилок, которые давали прирост к ее слову (может быть — удвоенным — если оба хвоста содержались как подслово). По индукции по множеству развилок утверждение очевидно. □

4 Доказательство оценки. Техническая часть.

4.1 Утверждение оценки

Сформулируем утверждение оценки:

$|\mathbb{G}_k| \leq \frac{\varphi_{k+d_k+I_{(d_k>1)}-1}}{2^{d_k+I_{(d_k>1)}-1}}$, где $|\mathbb{G}_k|$ — сумма весов всех ребер $\mathbb{G}_k$ (некорректно говорить об их количестве, при наличии ребер весом 0), d_k — количество входящих развилок развилок в $\mathbb{G}_k$, а $I_{(d_k>1)}$ равно 1, если $d_k > 1$, и 0 иначе. Если $d_k = 1$ то дополнительно $2x + y + z \leq \varphi_{k+1}$.



Заметим, что для графа $\mathbb{G}_1$ это утверждение верно:

Мы будем начинать рассмотрение с этого графа и рассматривать его как базу индукции. Запреты при первом шаге будем делать так, чтобы получилось $\mathbb{G}_2$.

Шаг индукции сформулируем следующим образом: если для некоторого $\mathbb{G}_k$ оценка верна, то можно сделать еще несколько шагов так, что для нового графа оценка также будет верна. Также будем утверждать, что для любого $\mathbb{G}_k$, такого что $d_k = 1$ — оценка верна.

4.2 План доказательства

Доказательство естественным образом делится на три части. Заметим, что если $d_k = 1$ — у нас есть только один вариант какую пару развилок схлопнуть — единственную имеющуюся в графе (под словом „схлопнуть“ мы будем понимать произведение описанной ранее операции над данной парой развилок соединенных путем без других развилок). Первая, самая простая часть доказательства касается случая, когда после этого единственного возможного схлопывания останется снова одна развилка ($d_{k+1} = 1$) — в этой ситуации возможны только два случая, и в обоих оценка остается верной. Если $d_{k+1} > 1$, то оценка резко усиливается — мы требуем запаса на одно гипотетическое удвоение. Однако, в силу того что начальная структура графа строго фиксирована, все возможные случаи перебираются, и либо оценка оказывается верна, либо мы возвращаемся к одноразвилочному состоянию и тогда она тоже верна. Третья часть доказательства соответственно касается случая, когда $d_k > 1$. В этом случае мы не знаем как выглядит граф, однако должны отследить только отклонение оценки в зависимости от количества произведенных операций и изменения количества развилок — это позволяет обобщить большой класс ситуаций, когда мы можем сразу сделать несколько операций с сохранением верности оценки, а остальные опять разобрать перебором.

4.3 Необходимые численные данные и мотивация

Мы будем изучать величину $\frac{|\mathbb{G}_j|}{|\mathbb{G}_i|}$ и оценивать ее некоторой константой c. Из верности утверждения оценки для $\mathbb{G}_i$ и неравенства $c \leq \frac{\varphi_{k+j-i+d_j-d_i}}{\varphi_k} * 2^{d_i-d_j}$ для c (где k — соответствующий индекс для $\mathbb{G}_i$) будет следовать верность утверждения для $\mathbb{G}_j$. Если поделить и домножить правую часть этого неравенства на $\alpha^{j-i+d_j-d_i}$ где α — золотое сечение ($\frac{1+\sqrt{5}}{2}$), то получится произведение двух выражений, значения которых легко анализировать: $\frac{\varphi_{k+j-i+d_j-d_i}}{\varphi_k * \alpha^{j-i+d_j-d_i}} * \left(\alpha^{j-i}\left(\frac{\alpha}{2}\right)^{d_j-d_i}\right)$. Очевидно, что правая часть неравенства всегда будет больше чем второе выражение умноженное на минимальное значение первого. И почти всегда — если умножить на 2-е по порядку значение первого (именно это значение нам будет удобнее всего, а оставшийся частный случай можно рассмотреть отдельно).

Соответственно, приведем значения выражения $\frac{\varphi_u}{\varphi_v * \alpha^{u-v}}$ (Приложение B, по вертикали откладывается u). Согласно нашей индексации $u, v \geq 1$. Нас интересует только нижний угол, где $u \geq v$, т.к. $j - i + d_j - d_i \geq 0$. Легко понять (из таблицы, а также из аналитической формулы для чисел Фибоначчи: $\varphi_c = \frac{\alpha^{c+1}-(-\alpha)^{-c-1}}{\sqrt{5}}$), что 0.9270509831 (вычислено с бóльшей точностью) является наименьшим значением в нижнем углу таблицы, а 0.9442719100 — вторым по порядку.

Заметим также, что числа фибоначчи представляют из себя не чисто геометрическую прогрессию с показателем α, а таковую с некоторыми небольшими (но существенными для асимптотики в начале) отклонениями, и таблица B характеризует отклонение увеличение значения φ от такового для геометрической прогрессии.

Значения выражения $\alpha^c * \left(\frac{\alpha}{2}\right)^k$ при $c > 0$ и $-c \leq k \leq c + 1$, отражающего максимальное допустимое увеличение размера графа при c шагах и изменении при них количества развилок на k приведены в таблице A (в каждом элементе таблицы указано значение данного выражения, значение $c+1$ — обычно, но не всегда в нашем доказательстве мы будем стремиться чтобы именно такое количество раз ограничивало увеличение размера графа за указанное количество шагов — и значение k.)

Приведем также таблицу значений этого выражения, домноженного на 0.9442719100 (небольшое количество случаев, когда $\frac{\varphi_u}{\varphi_v * \alpha^{u-v}}$ все-таки равно 0.9270509831 будет прокомментировано позднее): приложение C. Первый столбец, где $k = c + 1 > c$ соответствует как раз начальному увеличению числа развилок (когда была 1, а стало больше) и понадобится в соответствующем доказательстве.

Также заметим, что по данным таблицам (а также из очевидных соображений о производных) легко сделать выводы о росте соответствующих значений за их пределами.

4.4 Одноразвилковая ситуация

Известно, что $d_k = d_{k+1} = 1, 2x_k + y_k + z_k \leq \varphi_{k+1}, x_k + y_k + z_k \leq \varphi_k$. Без ограничения общности возможны два случая: $x_{k+1} = x_k + y_k$, $y_{k+1} = x_k + z_k$, $z_{k+1} = 0$ или $x_{k+1} = x_k + z_k$, $y_{k+1} = x_k + y_k$, $z_{k+1} = 0$. В любом из них $|\mathbb{G}_{k+1}| = 2x_k + y_k + z_k \leq \varphi_{k+1}$, $2x_{k+1} + y_{k+1} + z_{k+1} \leq 3x_k + y_k + z_k + max(y_k, z_k) \leq (2x_k + y_k + z_k) + (x_k + y_k + z_k) \leq \varphi_{k+1} + \varphi_k = \varphi_{k+2}$

4.5 Многоразвилковая ситуация

Заметим, что после любой операции количество развилок в графе может либо уменьшиться на 1, либо остаться тем же, либо увеличиться на 1. При этом размер графа увеличится на длину схлопнутого пути.

Уменьшение количества развилок

Если количество развилок после некоторой операции уменьшилось на один, то оценка останется верна: очевидно, длина несамопересекающегося пути в графе не может превышать общего количества ребер в нем. Как мы уже отмечали, за одну операцию размер графа увеличивается на длину схлопнутого пути, то есть $|\mathbb{G}_{k+1}| = |\mathbb{G}_k| + s \leq 2 * |\mathbb{G}_k| = 2 * \frac{\varphi_{k+d_k+I_{(d_k>1)}-1}}{2^{d_k+I_{(d_k>1)}-1}} = \frac{\varphi_{(k+1)+d_{k+1}+1}}{2^{d_{k+1}}}$ (здесь очевидно $I_{(d_k>1)} = 1$).

Заметим, что если в графе останется одна развилка, то у нас будет запас на одно удвоение для асимптотики — это не случайно. Это используется для случая, когда d_k было равно 2, и, соответственно, d_{k+1} стало равно 1. Мы знаем, что $x_{k+1} + y_{k+1} + z_{k+1} \leq \frac{\varphi_{k+2}}{2} \leq \varphi_{k+1}$. Из этого следует, что $2x_{k+1} + y_{k+1} + z_{k+1} \leq 2(x_{k+1} + y_{k+1} + z_{k+1}) \leq \varphi_{k+2}$.

Однако это еще не полное доказательство того, что для одноразвилкового графа оценка всегда верна. Дело в том, что у нас могло не быть промежуточной остановки на графе с больше чем одной развилкой с верной оценкой: если была одна развилка, потом стало больше (и усиленная оценка не выполнена), а потом обратно одна. Это случай будет рассмотрен в последней части доказательства.

В силу простоты рассуждения в этом случае условимся, что если для некоторого графа оценка верна и в нем есть пара развилок, после схлопывания которой количество развилок уменьшается на 1 — мы будем производить операцию с ней.

Пусть теперь в графе нет таких пар развилок.

Когда пар развилок, которые можно схлопнуть много

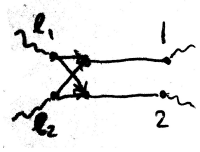
Рассмотрим некоторый текущий граф для которого оценка верна. Схлопнем в нем все пары развилок которые есть (не схлопывая вновь образовавшиеся). В этих операциях размер графа увеличится не более чем в два раза (т.к. пути которые мы схлопнули и тем самым удвоили одновременно присутствовали в нашем начальном графе и не пересекались). Посмотрим на таблицу C. То, что оценка перестала быть верна означает, что мы сделали 1 или 2 операции (причем если 2 — то количество развилок увеличилось ровно на 2). Во всех остальных случаях оценка осталась верна. Это значит, что остаются только случаи когда в графе в паре находятся только 1 или 2 пары развилок. Их и будем рассматривать.

Заметим, что вообще не быть схлапываемых пар развилок в графе не может: пойдем по ребрам от какой-нибудь входящей развилки в направлении ориентации. Пока на этом пути встречаются только входящие развилки путь однозначен. Идти всегда есть куда — у нас нет вершин исходящей степени 0. Раз есть входящие, значит есть и исходящие. В силу конечности графа мы когда-нибудь в них придем — иначе мы получим цикл, в который ребра только входят. Значит эти входящие развилки никогда бы не встретились с исходящими, и никогда не могли бы быть уничтожены — противоречие, т.к. в конце мы получаем циклический граф.

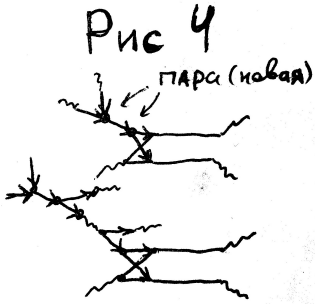
Увеличение количества развилок при схлопывании

Сначала рассмотрим случай когда пар развилок 2. Схлопнем одну из развилок.

Рис. 3

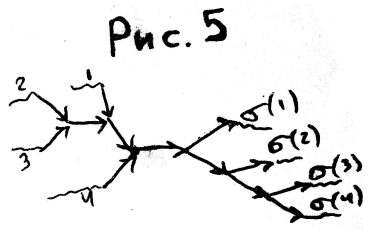


По одному из путей l_1 и l_2 точно есть какие-нибудь другие развилки (другой путь может непосредственно вести в вершину 1 или 2, но оба — не могут, т.к. в графе другие развилки все-таки есть). Пойдем по этому пути пока не встретим входящую развилку. Если перед этим мы встречали исходящие развилки, то, очевидно, между этой входящей и последней исходящей ничего нет, и это та самая вторая пара, что есть в графе и что тоже схлопывается без запретов. Если нет — то сейчас ничего нет между этой развилкой и развилкой в начале нашего пути. Во втором случае просто схлопнем эту пару и пару которая еще есть в графе из тех двух что были в начале. Получится, что граф не более чем удвоился. За три операции. Значит, исходя из нашей таблицы С — оценка осталась верна. В первом же случае — сдалем так: в этой (второй) паре не исходящюю быстро проведем ко входящей, а наоборот. То, что так делать можно — почти очевидно. Во-первых, потому что в самом начале мы могли решить остановить не входящие, а исходящие. Во-вторых, потому что структурно граф меняется точно так же, как если двигать исходящую ко входящей, а значит остается только понять корректность в смысле количественных характеристик — количества ребер в конечном графе. Это мы докажем чуть позже. Но мы знаем, что до схлопнутой последней исходящей развилки на нашем пути по которому мы шли (в направлении противоположном ориентации ребер) были еще исходящие развилки (может быть, только одна — в самом начале) Значит, поскольку мы не запрещали ребер при обоих схлопываниях, сейчас есть еще одна пара развилок. Схлопнем их. Три операции, не более чем удвоение (засчет того что мы удваивали только ребра присутствовавшие в графе до начала операций) — оценка верна.



Докажем просто, что если в некоторый момент какую-то входящую развилку подвинуть на 1 вправо (по направлению ориентации ее направляющего ребра), а потом все делать как обычно (понятно, что так можно — т.к. структуру графа это не поменяло) — то конечное количество ребер будет тем же самым. Действительно, когда мы подвинем эту развилку, размер графа увеличится на 1. Но зато той развилке, которая схлопнется с этой будет идти на 1 меньше, и засчет этого мы потеряем 1. Когда эти развилки схлопнутся, останется сколько-то входящих и столько же исходящих развилок. Заметим, что теперь всем исходящим идти на 1 больше, чем в обычном процессе, а тем исходящим, которые схлопнутся с образовавшимися входящими — на 1 меньше — баланс опять 0. И так далее для каждого схлапывания, а всё остальное очевидно будет так же как обычно — значит в конце ребер столько же.

Остается случай одной пары развилок. В этом случае граф выглядит так:

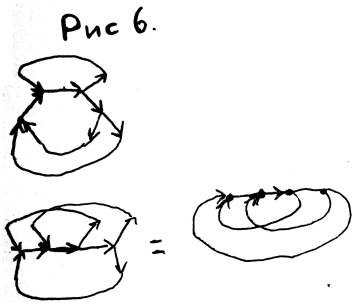


Пояснение — слева за входящей развилкой некоторое дерево входящих развилок. Справа за исходящей — дерево исходящих. Их листья соединены между собой в некотором порядке (перестановка σ).

Соответственно мы можем только схлопнуть эту пару развилок. В этом пункте мы рассмотрим ситуацию, когда при этом не происходит запретов и количество развилок увеличивается. Тогда, по-

скольку в графе сейчас более одной развилки, хотя бы по одному из путей (аналогично предыдущим рисункам l_1 и l_2) будет входящая развилка. Схлопнем соответвующую пару. Две операции. Если во второй раз запреты были — оценка верна (см. таблицу C). Если нет — мы, если в графе было более двух развилок, можем опять схлопнуть какую-нибудь вновь образовавшуюся по одному из путей пару развилок. Три операции. Не более чем удвоение. Оценка точно верна. Значит остается случай когда в графе ровно две развилки.

Без ограничения общности в этих случаях граф может иметь один из двух следующих видов (доказывается тривиальным перебором возможных перестановок с точностью до изоморфизма графов):



В обоих случаях начнем с того (с чего же еще?) что схлопнем единственную имеющуюся в графе пару развилок между которыми ничего нет.

Далее в первом случае, поскольку запретов по нашему предположению не произошло — мы можем схлопнуть развилки 1, 2 и 3.

Легко понять, что при этом граф увеличится не более чем в 4 раза. Заглянем в таблицу C, и осознаем, что нас интересует только случай когда за эти 4 операции произошло не больше одного запрета (в противном случае оценка стала верна).

Теперь схлопнем те из развилок 1 и 2 (на последнем рисунке) которые можно. Это либо одна, либо две. Опять понятно что в итоге все равно будет не более чем учетверение (т.к. по каждому ребру исходного графа мы проехали в общей сложности не более чем 4 раза). Если это только одна, значит был запрет. 5 операций, $\leq +4$ развилки, не более чем учетверение. См. таблицу C — оценка верна. Иначе — 6 операций, не более чем учетверение — оценка верна.

Теперь второй случай.

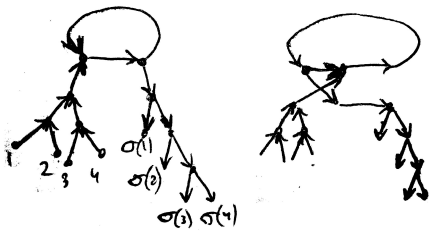
Предположим, что развилки пронумерованы цифрами 1,2,3,4 в порядке слева направо как они располагаются на рисунке. Схлопнем развилки 2 и 3. Запретов не было. Теперь схлопнем развилку 1 с её новой парой. Если запреты были, то 2 операции, не более чем +1 развилка, не более чем удвоение — оценка верна. Если не было, то схлопнем еще и развилку 4 с её текущей парой, а затем те две развилки, одна из которых образовалась при этой операции, а вторая — при схлопывании развилки 1. Если запреты были, то 4 операции, $\leq +3$ развилок, не более чем утроение — оценка верна. Если запретов не было, то схлопнем теперь все имеющиеся в графе пары, которых три. Получим в итоге не более чем ушестерение, 7 операций, не более чем +7 развилок. Оценка верна.

Неизменность количества развилок при схлапывании

Помним, что в этом пункте нам осталось только рассмотреть случай когда в графе только одна схлапываемая пара развилок. Если после первой операции схлапывания оставшаяся развилка (не важно, входящая или исходящая) — попала на поддерево, где были еще развилки — мы при первой операции двигаем развилку в соответствующую сторону и потому схлапываем новообразовавшуюся пару. Две операции, не более чем удвоение, запреты были — оценка верна.

Значит, остается случай когда с каждой стороны одно из ребер из корневой вершины ведет сразу в другую вершину, и на это ребро попадают обе развилки после первого схлапывания.

Рис 9



Видно, что новый граф имеет структуру абсолютно аналогичную предыдущему. Ну снова схлопнем единственную имеющуюся развилку, и вообще будем продолжать это делать пока не случится иначе. А когда случится — сделаем как описывалось ранее — схлопнем с кем-нибудь из поддерева. Если это произошло c раз, то граф увеличился не более чем в $c+1$ раз (надо понимать, что мы как бы заранее планировали с каким поддеревом будем схлапывать последнюю развилку, и в соответствующую сторону сдвигали нашу при схлопываниях). $c+1$ операция, $c \geq 2$, не более чем +2 развилки. Оценка верна (очевидно из таблицы и анализа производной соотв. выражения) всегда кроме случая $c = 2$, +2 развилки. В этом случае схлопнем пару, аналогичную той которую уже схлопнули два раза (она есть, т.к. последих запретов не было — отсюда было +2) — сместились по таблице на 1 вправо вниз — оценка верна. Тем более оценка верна, если в конце наших идентичных операций не +1 развилка, а -1 (c операций, -1 развилка, увеличение не более чем в $c+1$ раз). Если вдруг после этого „-1“ осталась только одна развилка — то аналогично доказательствам приведенным ранее — поскольку верна усиленная оценка на размер графа — верно и утверждение нашей основной оценки для одноразвилковых графов.

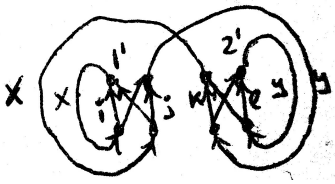
4.6 Начальное увеличение количества развилок

Пусть в $\mathbb{G}_k$ ровно одна развилка, и для него утверждение оценки верно. Если после схлопывания этой развилки не произойдет запретов, то количество развилок увеличится и станет равным двум, а нам будет необходимо сделать еще несколько операций так, чтобы асимптотика размера графа отстала в нужное количество раз или мы вернулись к одноразвилковой ситуации и утверждение оценки было верно. Без ограничения общности $\mathbb{G}_{k+1}$ будет выглядеть так:

Рис. 10

Нам известно, что $x + y \leq \varphi_{k+1}$, а $x + y + max(x, y) \leq \varphi_{k+2}$ (из утверждения оценки для $\mathbb{G}_k$). Схлопнем развилки 1 и 2. Получится такой граф (возможно, некоторые ребра из i,j,k и l запрещены — но только они, иначе есть цикл со входящими развилками, но без входящих, или наоборот):

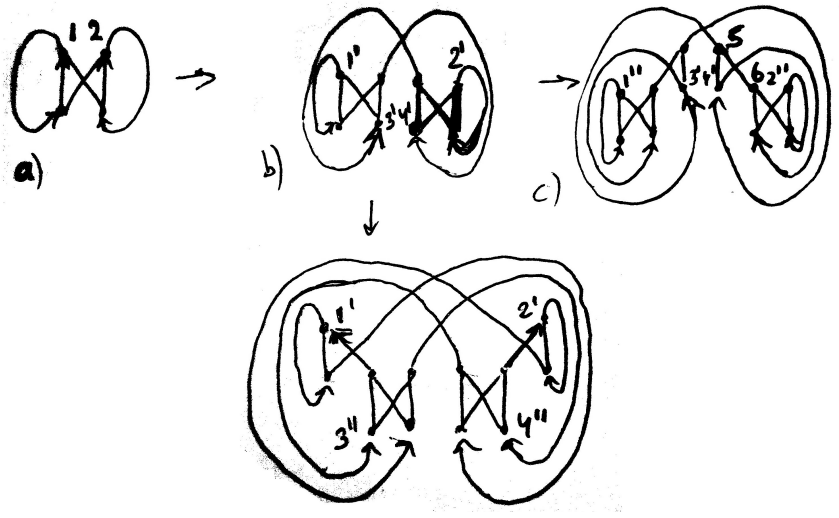
Рис 11



Если запрещены ребра i и j, то мы могли схлопнуть только развилку 1 и получить одноразвилковый граф с верным утверждением оценки. Аналогично если запрещены ребра k и l. Если осталась развилка $1'$ — схлопнем её. Если $2'$ осталась — её. Аналогично после этого появятся развилки 1" или 2". Будем продолжать схлопывать эти развилки пока они остаются. Когда-нибудь это закончится. Пусть с развилками $1^{(t)}$ это произведено n_1 раз, а с развилками $2^{(t)}$ n_2 раз. Граф увеличился не более чем в $max(n_1, n_2) + 1$ раз. Сделано $n_1 + n_2$ операций. Количество развилок — стало не более чем $n_1 + n_2$ (вспомним, что оно определяет k в таблице). Получается, что в таблице нас интересует второй столбец, когда $c = k$. Заметим, что $n_1, n_2 \geq 1$. Без ограничения общности можно предположить, что $n_1 \leq n_2$. Получается, что граф увеличился в $n_2 + 1$ раз. Тогда оценка не верна только для случаев $n_1 = 1, n_2 \in [1, 6]; n_1 = 2, n_2 \in [2, 3]$. Заметим, что если за эти $n_1 + n_2$ операций был хотя бы один запрет (кроме последних двух), то оценка не верна только когда $n_1 = 1, n_2 \in [2, 5]$, а если хотя бы два лишних запрета — то в любом случае верна.

Теперь схлопнем те из развилок 3' и 4', которые можно, а затем — те из развилок 3" и 4" которые можно (на рисунке приведены результаты операций схлопывания развилок 1' и 2' — c) и схлопывания 3' и 4' — d)). Теперь начнем оценивать числовые характеристики наших действий.

Рис 13



Если на первом этапе был ровно 1 запрет, то $n_1 = 1, n_2 \in [2, 5]$. Поскольку ребра i и j как мы условились одновременно быть запрещены не могли, то развилка 3' есть. Поскольку $n_2 \geq 2$, а запрет был ровно 1, то хотя бы одна из развилок 5, 6 есть. Значит можно с ней схлопнуть нашу развилку 3'. Числовые характеристики как если бы n_1 было равно 2 и был хотя бы один запрет, значит оценка верна.

Если на первом этапе запретов не было — значит все развилки есть. В частности — 3' и 4'. Схлопнем их. Если запреты были, то числовые характеристики как если $n_1, n_2 \geq 2$, значит оценка верна. Если запретов не было опять, то есть 3" и 4", схлопнем их. Теперь числовые характеристики как если $n_1, n_2 \geq 3$, и оценка опять-таки верна.

Заметим, что если в итоге стала одна развилка, то для этого графа верна усиленная оценка, а если 0 — то рассуждение окончено.

4.7 Замечание

По сути данное доказательство — есть формальное обоснование того факта, что если развилок много, то их встречи происходят очень часто относительно диаметра графа. Но поскольку в случае большого количества развилок структура графа становится непонятной и сложной, нельзя утверждать что не случится настолько удачной ситуации, что увеличение размера графа не окупит все предыдущие потери (а искуственно ситуацию когда на каждое размножение приходится по удвоению построить можно. Больше чем удвоение — нет, т.к. длина несамопересекающегося пути в графе очевидно не может превышать общего количества ребер в нем). Поэтому приходится делать такую усиленную оценку в зависимости от текущего количества перекрестков.

Рассмотрим оставшийся случай, когда $\frac{\varphi_u}{\varphi_v * \alpha^{u-v}}$ все-таки равно минимальному значению 0.9270509831. Это возможно только когда мы на основании оценки для $|\mathbb{G}_2|$ пытаемся оценить $|\mathbb{G}_3|$. Если в $\mathbb{G}_3$ одна развилка — утверждение верно (в $\mathbb{G}_3$ в любом случае 3 ребра), если нет — мы, по нашему доказательству, делали дополнительные шаги (т.к. усиленная оценка была не верна — т.к. верна только самая мягкая оценка — для одноразвилкового случая).

4.8 Заключение

Как мы доказали, как бы мы ни делали наши операции — в конце получится циклический граф с количеством ребер равным периоду слова. Поскольку за одну операцию количество развилок в графе уменьшается не более чем на 1, а циклический граф — граф без развилок, то на предпоследнем шаге в графе $\mathbb{G}_k$ была ровно одна развилка. Значит для этого графа оценка была верна. Тогда $|\mathbb{G}_{k+1}| = 2x_k + y_k + z_k \leq \varphi_{k+1}$, в чем и заключается теорема 1.

5 Доказательство минимальности оценки

Несложно строится пример последоваетльности слов с бесконечно возрастающим периодом задаваемых соответствующим минимально возможным количеством запретов: просто будем сохранять в графах Рози одну развилку и делать так, что $y_{k+1} = x_k, x_{k+1} = x_k + y_k$ — тогда размер графа будет увеличиваться ровно по последовательности фибоначчи. На некотором шаге уничтожим обе развилки и получим соответствующее слово. Небольшие отклонения в этом процессе позволяют получить слова с другими периодами задаваемые соответствующим минимальным количеством запретов.

6 Случай многобуквенного алфавита

Случай многобуквенного алфавита довольно удачно сводится к случаю двухбуквенного. Заметим, что в графах Рози слов содержащих k букв могут встречаться вершины входящей и исходящей степени от 1 до k. Аналогично двухбуквенному случаю все сводится к рассмотрению графов, в которых одна из входящей и исходящей степеней любой вершины равна 1.

Вершина входящей степени l может быть представлена как бинарное дерево (не важно какой формы) имеющее l концов (например, $l-1$ последовательных входящих развилок), каждое ребро которого имеет вес 0. Легко понять как происходит эволюция графов Рози в случае k-буквенного алфавита: развилки аналогично едут навстречу друг другу, а при встрече, допустим, n-валентной входящей развилки и m-валентной исходящей, получается m n-валентных входящих развилок и n m-валентных исходящих. После этого как-то происходят запреты.

Заметим, что если сначала столкнуть развилки, а потом заменить на бинарные - получится $m*(n-1)$ входящих двоичных развилок, и $n*(m-1)$ исходящих, а если сначала заменить на деревья, а потом схлопнуть все эти развилки друг с другом, то получится столько же (каждая исходящая при прохождении через каждую входящюю дает $+1$ входящую развилку, тоесть всего $+(m-1)*(n-1)$). Только вот запреты уже надо делать не в конце, а по ходу схлопывания двоичных - и тут их получится меньше, при той же итоговой конфигурации.

Получается, если в случае многобуквенного в начальном графе заменить многовалентные развилки на бинарные деревья, и провести эволюцию с аналогичными запрещениями до конца, то ребер будет столько же, а запретов меньше. Значит, максимальный конечный размер графа увеличивается не быстрее чем экспонента с основанием золотое сечение и степень - количество запретов, и интересна только начальная константа (очевидно, из изложенного ранее рассуждения, что в начальном графе $k-1$ развилка, а его размер — 1, таким образом начальная константа $\left(\frac{2}{\alpha}\right)^{k-1}$ подходит).

Пример же аналогично строится, если с самого начала убить все развилки (не совсем с самого - иначе это будет просто запрещение букв), кроме 1 входящей и 1 исходящей, и дальше сделать рост как последовательность Фибоначчи.

Следовательно, возникает интересный вопрос: поскольку требование наличия k букв элементарно (по крайней мере на уровне идеи) сводится к случаю наличия 2 букв благодаря уничтожению большинства валентностей у развилок, интересно исследовать асимптотику роста максимального размера графа в зависимости от количества запретов при более строгих ограничениях, допустим, таких: суммарная валентность входящих развилок за вычетом их количества всегда не меньше k (по сути - это количество эквивалентных двоичных развилок).

Она (асимптотика), очевидно, будет существенно меньше — то есть для сложных слов требуется большое количество запретов. Впрочем, может быть, что она тоже будет экспоненциальной.

7 Замечание.

Как мне стало известно от А.Я.Белова аналогичный результат был независимо получен И.И.Богдановым и Г.Р.Челноковым. Будет интересно сравнить доказательства когда появится их текст.

Список литературы

[1] *Belov A., Borisenko V., Latyshev V. Monomial algebras*. NY, Plenum. 1998, p. 5 – 190. J.Math.Sci., NY, 87, 1997, no3, pp.3463–3575

[2] *Г. Р. Челноков О числе запретов, задающих периодическую последовательность* Модел. и анализ информ. систем, 14:2 (2007), 12-16 http://mi.mathnet.ru/mais128

[3] *Ivan Mitrofanov On uniform recurrence of HD0l systems* Comments: 20 pages, in Russian Subjects: Combinatorics (math.CO); Logic (math.LO)

[4] *Ivan Mitrofanov A proof for the decidability of HD0L ultimate periodicity* arXiv:1110.4780

pdf, *ps*, *other*

Comments: 33 pages, in Russian Subjects: Combinatorics (math.CO); Logic (math.LO)

[5] *Alexei Kanel-Belov, Ivan Mitrofanov Periodicity of Rauzy scheme and substitutional systems* arXiv:1107.0185

pdf, *ps*, *other*

[6] *Fabien Durand (LAMFA) Decidability of uniform recurrence of morphic sequences* arXiv:1204.5393

pdf, *ps*, *other*

Subjects: Combinatorics (math.CO); Discrete Mathematics (cs.DM)

[7] Fabien Durand (LAMFA) Decidability of the HD0L ultimate periodicity problem arXiv:1111.3268

pdf, *ps*, *other*

Subjects: Combinatorics (math.CO); Discrete Mathematics (cs.DM)

[8] *A.Ya. Kanel-Belov, A.L. Chernyat'ev. Describing the set of words generated by interval exchange transformation.* Comm. in Algebra, Vol. 38, No 7, July 2010, pages 2588–2605.

[9] *A.Ya.Belov, G.V.Kondakov, I.Mitrofanov. Inverse problems of symbolic dynamics.* Banach Center Publ. 94 (2011), 43–60.

A Таблица 1

$$\begin{pmatrix}
\begin{pmatrix} 1.05902 \\ 2 \\ 2 \end{pmatrix} & \begin{pmatrix} 1.30902 \\ 2 \\ 1 \end{pmatrix} & \begin{pmatrix} 1.61803 \\ 2 \\ 0 \end{pmatrix} & \begin{pmatrix} 2. \\ 2 \\ -1 \end{pmatrix} & \begin{pmatrix} 2.47214 \\ 2 \\ -2 \end{pmatrix} & \begin{pmatrix} 3.05573 \\ 2 \\ -3 \end{pmatrix} \\
\begin{pmatrix} 1.38627 \\ 3 \\ 3 \end{pmatrix} & \begin{pmatrix} 1.71353 \\ 3 \\ 2 \end{pmatrix} & \begin{pmatrix} 2.11803 \\ 3 \\ 1 \end{pmatrix} & \begin{pmatrix} 2.61803 \\ 3 \\ 0 \end{pmatrix} & \begin{pmatrix} 3.23607 \\ 3 \\ -1 \end{pmatrix} & \begin{pmatrix} 4. \\ 3 \\ -2 \end{pmatrix} \\
\begin{pmatrix} 1.81465 \\ 4 \\ 4 \end{pmatrix} & \begin{pmatrix} 2.24303 \\ 4 \\ 3 \end{pmatrix} & \begin{pmatrix} 2.77254 \\ 4 \\ 2 \end{pmatrix} & \begin{pmatrix} 3.42705 \\ 4 \\ 1 \end{pmatrix} & \begin{pmatrix} 4.23607 \\ 4 \\ 0 \end{pmatrix} & \begin{pmatrix} 5.23607 \\ 4 \\ -1 \end{pmatrix} \\
\begin{pmatrix} 2.37541 \\ 5 \\ 5 \end{pmatrix} & \begin{pmatrix} 2.93617 \\ 5 \\ 4 \end{pmatrix} & \begin{pmatrix} 3.62931 \\ 5 \\ 3 \end{pmatrix} & \begin{pmatrix} 4.48607 \\ 5 \\ 2 \end{pmatrix} & \begin{pmatrix} 5.54508 \\ 5 \\ 1 \end{pmatrix} & \begin{pmatrix} 6.8541 \\ 5 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 3.10945 \\ 6 \\ 6 \end{pmatrix} & \begin{pmatrix} 3.8435 \\ 6 \\ 5 \end{pmatrix} & \begin{pmatrix} 4.75082 \\ 6 \\ 4 \end{pmatrix} & \begin{pmatrix} 5.87234 \\ 6 \\ 3 \end{pmatrix} & \begin{pmatrix} 7.25861 \\ 6 \\ 2 \end{pmatrix} & \begin{pmatrix} 8.97214 \\ 6 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 4.07033 \\ 7 \\ 7 \end{pmatrix} & \begin{pmatrix} 5.0312 \\ 7 \\ 6 \end{pmatrix} & \begin{pmatrix} 6.21891 \\ 7 \\ 5 \end{pmatrix} & \begin{pmatrix} 7.68699 \\ 7 \\ 4 \end{pmatrix} & \begin{pmatrix} 9.50164 \\ 7 \\ 3 \end{pmatrix} & \begin{pmatrix} 11.7447 \\ 7 \\ 2 \end{pmatrix} \\
\begin{pmatrix} 5.32813 \\ 8 \\ 8 \end{pmatrix} & \begin{pmatrix} 6.58593 \\ 8 \\ 7 \end{pmatrix} & \begin{pmatrix} 8.14065 \\ 8 \\ 6 \end{pmatrix} & \begin{pmatrix} 10.0624 \\ 8 \\ 5 \end{pmatrix} & \begin{pmatrix} 12.4378 \\ 8 \\ 4 \end{pmatrix} & \begin{pmatrix} 15.374 \\ 8 \\ 3 \end{pmatrix} \\
\begin{pmatrix} 6.97461 \\ 9 \\ 9 \end{pmatrix} & \begin{pmatrix} 8.62109 \\ 9 \\ 8 \end{pmatrix} & \begin{pmatrix} 10.6563 \\ 9 \\ 7 \end{pmatrix} & \begin{pmatrix} 13.1719 \\ 9 \\ 6 \end{pmatrix} & \begin{pmatrix} 16.2813 \\ 9 \\ 5 \end{pmatrix} & \begin{pmatrix} 20.1248 \\ 9 \\ 4 \end{pmatrix} \\
\begin{pmatrix} 9.12988 \\ 10 \\ 10 \end{pmatrix} & \begin{pmatrix} 11.2852 \\ 10 \\ 9 \end{pmatrix} & \begin{pmatrix} 13.9492 \\ 10 \\ 8 \end{pmatrix} & \begin{pmatrix} 17.2422 \\ 10 \\ 7 \end{pmatrix} & \begin{pmatrix} 21.3125 \\ 10 \\ 6 \end{pmatrix} & \begin{pmatrix} 26.3437 \\ 10 \\ 5 \end{pmatrix} \\
\begin{pmatrix} 11.9512 \\ 11 \\ 11 \end{pmatrix} & \begin{pmatrix} 14.7725 \\ 11 \\ 10 \end{pmatrix} & \begin{pmatrix} 18.2598 \\ 11 \\ 9 \end{pmatrix} & \begin{pmatrix} 22.5703 \\ 11 \\ 8 \end{pmatrix} & \begin{pmatrix} 27.8984 \\ 11 \\ 7 \end{pmatrix} & \begin{pmatrix} 34.4844 \\ 11 \\ 6 \end{pmatrix}
\end{pmatrix}$$

B Таблица 2

$$\begin{pmatrix}
1.000000 & 0.8090170 & 0.8726780 & 0.8472136 & 0.8567627 & 0.8530900 & 0.8544891 & 0.8539542 \\
1.236068 & 1.000000 & 1.078689 & 1.047214 & 1.059017 & 1.054477 & 1.056207 & 1.055545 \\
1.145898 & 0.9270510 & 1.000000 & 0.9708204 & 0.9817627 & 0.9775541 & 0.9791574 & 0.9785444 \\
1.180340 & 0.9549150 & 1.030057 & 1.000000 & 1.011271 & 1.006936 & 1.008588 & 1.007956 \\
1.167184 & 0.9442719 & 1.018576 & 0.9888544 & 1.000000 & 0.9957132 & 0.9973463 & 0.9967219 \\
1.172209 & 0.9483372 & 1.022961 & 0.9931116 & 1.004305 & 1.000000 & 1.001640 & 1.001013 \\
1.170290 & 0.9467844 & 1.021286 & 0.9914855 & 1.002661 & 0.9983626 & 1.000000 & 0.9993739 \\
1.171023 & 0.9473775 & 1.021926 & 0.9921066 & 1.003289 & 0.9989880 & 1.000626 & 1.000000
\end{pmatrix}$$

C Таблица 3

$$\begin{pmatrix}
\begin{pmatrix} 1. \\ 2 \\ 2 \end{pmatrix} & \begin{pmatrix} 1.23607 \\ 2 \\ 1 \end{pmatrix} & \begin{pmatrix} 1.52786 \\ 2 \\ 0 \end{pmatrix} & \begin{pmatrix} 1.88854 \\ 2 \\ -1 \end{pmatrix} & \begin{pmatrix} 2.33437 \\ 2 \\ -2 \end{pmatrix} & \begin{pmatrix} 2.88544 \\ 2 \\ -3 \end{pmatrix} \\
\begin{pmatrix} 1.30902 \\ 3 \\ 3 \end{pmatrix} & \begin{pmatrix} 1.61803 \\ 3 \\ 2 \end{pmatrix} & \begin{pmatrix} 2. \\ 3 \\ 1 \end{pmatrix} & \begin{pmatrix} 2.47214 \\ 3 \\ 0 \end{pmatrix} & \begin{pmatrix} 3.05573 \\ 3 \\ -1 \end{pmatrix} & \begin{pmatrix} 3.77709 \\ 3 \\ -2 \end{pmatrix} \\
\begin{pmatrix} 1.71353 \\ 4 \\ 4 \end{pmatrix} & \begin{pmatrix} 2.11803 \\ 4 \\ 3 \end{pmatrix} & \begin{pmatrix} 2.61803 \\ 4 \\ 2 \end{pmatrix} & \begin{pmatrix} 3.23607 \\ 4 \\ 1 \end{pmatrix} & \begin{pmatrix} 4. \\ 4 \\ 0 \end{pmatrix} & \begin{pmatrix} 4.94427 \\ 4 \\ -1 \end{pmatrix} \\
\begin{pmatrix} 2.24303 \\ 5 \\ 5 \end{pmatrix} & \begin{pmatrix} 2.77254 \\ 5 \\ 4 \end{pmatrix} & \begin{pmatrix} 3.42705 \\ 5 \\ 3 \end{pmatrix} & \begin{pmatrix} 4.23607 \\ 5 \\ 2 \end{pmatrix} & \begin{pmatrix} 5.23607 \\ 5 \\ 1 \end{pmatrix} & \begin{pmatrix} 6.47214 \\ 5 \\ 0 \end{pmatrix} \\
\begin{pmatrix} 2.93617 \\ 6 \\ 6 \end{pmatrix} & \begin{pmatrix} 3.62931 \\ 6 \\ 5 \end{pmatrix} & \begin{pmatrix} 4.48607 \\ 6 \\ 4 \end{pmatrix} & \begin{pmatrix} 5.54508 \\ 6 \\ 3 \end{pmatrix} & \begin{pmatrix} 6.8541 \\ 6 \\ 2 \end{pmatrix} & \begin{pmatrix} 8.47214 \\ 6 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 3.8435 \\ 7 \\ 7 \end{pmatrix} & \begin{pmatrix} 4.75082 \\ 7 \\ 6 \end{pmatrix} & \begin{pmatrix} 5.87234 \\ 7 \\ 5 \end{pmatrix} & \begin{pmatrix} 7.25861 \\ 7 \\ 4 \end{pmatrix} & \begin{pmatrix} 8.97214 \\ 7 \\ 3 \end{pmatrix} & \begin{pmatrix} 11.0902 \\ 7 \\ 2 \end{pmatrix} \\
\begin{pmatrix} 5.0312 \\ 8 \\ 8 \end{pmatrix} & \begin{pmatrix} 6.21891 \\ 8 \\ 7 \end{pmatrix} & \begin{pmatrix} 7.68699 \\ 8 \\ 6 \end{pmatrix} & \begin{pmatrix} 9.50164 \\ 8 \\ 5 \end{pmatrix} & \begin{pmatrix} 11.7447 \\ 8 \\ 4 \end{pmatrix} & \begin{pmatrix} 14.5172 \\ 8 \\ 3 \end{pmatrix} \\
\begin{pmatrix} 6.58593 \\ 9 \\ 9 \end{pmatrix} & \begin{pmatrix} 8.14065 \\ 9 \\ 8 \end{pmatrix} & \begin{pmatrix} 10.0624 \\ 9 \\ 7 \end{pmatrix} & \begin{pmatrix} 12.4378 \\ 9 \\ 6 \end{pmatrix} & \begin{pmatrix} 15.374 \\ 9 \\ 5 \end{pmatrix} & \begin{pmatrix} 19.0033 \\ 9 \\ 4 \end{pmatrix} \\
\begin{pmatrix} 8.62109 \\ 10 \\ 10 \end{pmatrix} & \begin{pmatrix} 10.6563 \\ 10 \\ 9 \end{pmatrix} & \begin{pmatrix} 13.1719 \\ 10 \\ 8 \end{pmatrix} & \begin{pmatrix} 16.2813 \\ 10 \\ 7 \end{pmatrix} & \begin{pmatrix} 20.1248 \\ 10 \\ 6 \end{pmatrix} & \begin{pmatrix} 24.8756 \\ 10 \\ 5 \end{pmatrix} \\
\begin{pmatrix} 11.2852 \\ 11 \\ 11 \end{pmatrix} & \begin{pmatrix} 13.9492 \\ 11 \\ 10 \end{pmatrix} & \begin{pmatrix} 17.2422 \\ 11 \\ 9 \end{pmatrix} & \begin{pmatrix} 21.3125 \\ 11 \\ 8 \end{pmatrix} & \begin{pmatrix} 26.3437 \\ 11 \\ 7 \end{pmatrix} & \begin{pmatrix} 32.5626 \\ 11 \\ 6 \end{pmatrix}
\end{pmatrix}$$

C.7 Let X be the (non-empty) set of $2D$ words that can be formed. We write $c \prec u$ if the finite pattern c appears in $u \in X$, and $u \prec v$ if any finite patterns of $u \in X$ also appears in $v \in X$. A finite pattern c of $u \in X$ is said to be *critical* if there is a sequence (S_n) of patterns of u which cover arbitrarily large disks and such that, for any n, $c \not\prec S_n$.

1. A $2D$ word is quasiperiodic if and only if it has no critical pattern.

2. Assume that $u_0 \in X$ is not quasiperiodic. Let c_0 be a critical pattern of u_0 and (S_n) be the associated sequence of patterns. We use **B.4** to build from (S_n) a $2D$ word $u_1 \in X$. One has $c_0 \not\prec u_1 \prec u_0$.

3. While u_n has a critical pattern c_n, we find as above $u_{n+1} \in X$ such that $c_n \not\prec u_{n+1} \prec u_n$. We moreover take for c_n the smallest one among the critical patterns of u_n. Note that c_n is also critical for u_k for $k \leq n$.

4. If we eventually find $u_n \in X$ without critical pattern, then we are done.

5. Otherwise, we again use **B.4** to build from (u_n) a $2D$ word $u_\infty \in X$. One has, for any n, $u_\infty \prec u_n$. Let us show that u_∞ is quasiperiodic.

6. If u_∞ has a critical pattern c, then c is a critical pattern of any u_n.

7. One has $c_n \notin \{c_0, \ldots, c_{n-1}\}$ because c_n is also critical for u_k, $k \leq n$. The size of c_n is thus not uniformly bounded. For c_n larger than c, this contradicts the minimality of c_n among critical patterns of u_n.

А.Заславский, А.Акопян,
О.Заславский, П.Кожевников

ТОЧКИ БРОКАРА

1 Точки Брокара в треугольнике

1. Дан треугольник ABC. Докажите, что существует единственная точка P, такая что $\angle PAB = \angle PBC = \angle PCA = \phi_1$, и единственная точка Q, такая что $\angle QBA = \angle QCB = \angle QAC = \phi_2$.

Определение 1. Точки P и Q называются *точками Брокара* треугольника ABC.

2.
a) Докажите, что $\phi_1 = \phi_2 = \phi$.
b) Выразите ϕ через углы треугольника ABC.

Определение 2. Угол ϕ называется *углом Брокара* треугольника ABC.

3. Докажите, что проекции точек Брокара на стороны треугольника лежат на одной окружности (На самом деле, это верно для любых двух изогонально сопряженных точек).

4. Пусть O — центр описанной окружности ABC.
a) Докажите, что $OP = OQ$.
b) Докажите, что $\angle POQ = 2\phi$.

Определение 3. Прямые, симметричные медианам треугольника относительно соответствующих биссектрис, называются *симедианами*. Можно доказать, что три симедианы пересекаются в одной точке L, которая называется *точкой Лемуана* треугольника.

5. Докажите, что P и Q лежат на окружности с диаметром OL.

6. (К.Кноп) Рассмотрим два треугольника: один образован центрами описанных окружностей треугольников PAB, PBC, PCA; другой — центрами описанных окружностей треугольников QAB, QBC, QCA. Докажите, что эти треугольники
a) подобны треугольнику ABC;
b) равны.
c) Найдите центр и угол поворота, переводящего один из этих треугольников в другой.

7. Пусть C' — такая точка на стороне AB, что прямая AB является внешней биссектрисой угла $PC'Q$. Докажите, что CC' — симедиана треугольника. (Т.е. можно построить эллипс с фокусами в точках Брокара, касающийся сторон треугольника в основаниях его симедиан).

8. Пусть T_1, T_2 — такие точки на прямой OL, что $\angle LPT_1 = \angle LPT_2 = 60^\circ$. Докажите, что проекции каждой из этих точек на стороны треугольника ABC образуют правильный треугольник (эти точки называются *точками Аполлония*).

2 Точки Брокара в четырехугольнике

9. Дана выпуклая ломаная $ABCD$. Докажите, что существует единственная точка P, такая что $\angle PAB = \angle PBC = \angle PCD = \phi$.

Определение 4. Точку P и угол ϕ будем называть *точкой и углом Брокара* ломаной $ABCD$ и обозначать $P(ABCD)$ и $\phi(ABCD)$.

10. Выразите $\phi(ABCD)$ через длины звеньев ломаной и углы между ними.

11. Докажите, что $\phi(ABCD) = \phi(DCBA)$ тогда и только тогда, когда точки A, B, C, D лежат на одной окружности.

В дальнейшем все рассматриваемые многоугольники предполагаются вписанными.

12. Пусть $P_1 = P(ABCD)$, $P_2 = P(BCDA)$, $P_3 = P(CDAB)$, $P_4 = P(DABC)$. Докажите, что четырехугольник $P_1P_2P_3P_4$ — вписанный.

13. Пусть $Q_1 = P(DCBA)$, $Q_2 = P(ADCB)$, $Q_3 = P(BADC)$, $Q_4 = P(CBDA)$. Докажите, что $P_1P_2/Q_1Q_2 = BC/CD$, $P_2P_3/Q_2Q_3 = CD/DA$ и т.д.

14. (Д.Белев) Пусть M_1, M_2 — такие точки на прямых AD, AB соответственно, что $BM \parallel CD$, $CM_2 \parallel DA$.

a) Докажите, что описанные окружности треугольников BAM_1 и BCM_2 пересекаются в точке P_1.

b) Опишите аналогичное построение для точек $P_i, i = 2, \ldots, 4$, $Q_i, i = 1, \ldots, 4$.

15. (Д.Белев) Докажите, что прямые CP_1, DP_2, AP_3 и BP_4 пересекаются в одной точке, и прямые BQ_1, CQ_2, DQ_3 и AQ_4 также пересекаются в одной точке.

16. (Д.Белев) Обозначим точки, полученные в предыдущей задаче через P_0, Q_0.

a) Докажите, что $S_{P_1P_2P_0} = S_{Q_1Q_2Q_0}$

b) Докажите, что площади четырехугольников $P_1P_2P_3P_4$ и $Q_1Q_2Q_3Q_4$ равны.

17. Докажите, что $\phi(ABCD) = \phi(BCDA)$ тогда и только тогда, когда $AB{\cdot}CD = AD{\cdot}BC$.

Определение 5. Вписанный четырехугольник, произведения противоположных сторон которого равны называется *гармоническим*. Из последней задачи следует, что в гармоническом четырехугольнике существуют такие точки P и Q, что $\angle PAB = \angle PBC = \angle PCD = \angle PDA = \angle QDC = \angle QCB = \angle QBA = \angle QAD = \phi$. Точки P, Q будем называть *точками Брокара*, а угол ϕ *углом Брокара* четырехугольника $ABCD$.

18. Докажите, что каждое из следующих условий равносильно тому, что четырехугольник $ABCD$ гармонический.

a) Касательные к описанной окружности в точках A и C пересекаются на прямой BD.

b) Диагональ BD является симедианой треугольника ABC.

c) Расстояния от точки пересечения диагоналей L до сторон четырехугольника пропорциональны этим сторонам.

d) Существует инверсия, переводящая точки A, B, C, D в вершины квадрата.

e) Существует центральная проекция, при которой описанная окружность $ABCD$ проектируется в окружность, а сам четырехугольник в квадрат.

19. Выразите угол Брокара через углы гармонического четырехугольника.

20. Докажите, что $OP = OQ$ и $\angle POQ = 2\phi$.

21. Докажите, что P и Q лежат на окружности с диаметром OL.

3 Точки Брокара в многоугольниках

22. Пусть дана окружность, точка P внутри нее и угол ϕ. Для произвольной точки X_0 окружности построим точку X_1, для которой ориентированный угол PX_0X_1 равен ϕ. Аналогично по X_1 построим точку X_2 и т.д. Докажите, что, если $X_n = X_0$, то это выполняется и для любой другой начальной точки.

23. Выведите условие замыкания в предыдущей задаче.

Напоминаем, что все рассматриваемые многоугольники вписанные.

Определение 6. Многоугольник $A_1 \dots A_n$ будем называть *брокаровским* если существует такая точка P, что $\angle PA_1A_2 = \angle PA_2A_3 = \dots \angle PA_nA_1 = \phi$.

24. Докажите, что в брокаровском многоугольнике существует также такая точка Q, что $\angle QA_1A_n = \angle QA_nA_{n-1} = \dots \angle QA_2A_1 = \phi$.

Определение 7. Точки P, Q и угол ϕ будем называть *точками и углом Брокара* многоугольника $A_1 \dots A_n$.

25. Докажите, что брокаровость равносильна каждому из следующих условий.

a) Существует точка L, расстояния от которой до сторон многоугольника пропорциональны этим сторонам.

b) Симедианы треугольников $A_1A_2A_3, A_2A_3A_4, \dots, A_nA_1A_2$, проведенные из вершин $A_2, A_3, \dots, A_1$, пересекаются в одной точке.

c) Точки пересечения прямых $A_1A_3, A_2A_4, \dots, A_nA_2$ с касательными к описанной окружности многоугольника в точках $A_2, A_3, \dots, A_1$ лежат на одной прямой.

d) Существует инверсия, переводящая точки $A_1, \dots, A_n$ в вершины правильного многоугольника.

e) Существует центральная проекция, переводящая описанную окружность многоугольника в окружность, а сам многоугольник в правильный.

26. Докажите, что точки Брокара лежат на окружности с диаметром OL и $\angle POL = \angle QOL = \phi$.

27.

a) Докажите, что существуют две точки T_1, T_2, инверсия с центром в которых переводит точки $A_1, \dots, A_n$ в вершины правильного многоугольника.

b) Докажите, что T_1, T_2 лежат на прямой OL и $\angle T_1PL = \angle T_2PL = \frac{\pi}{n}$.

28. Выразите угол Брокара через отношение OL/R.

А.Заславский, А.Акопян,
О.Заславский, П.Кожевников

ТОЧКИ БРОКАРА

Решения

1 Точки Брокара в треугольнике

1. Так как $\angle PAB = \angle PBC$, $\angle BPA = \pi - \angle B$, т.е. P лежит на окружности, проходящей через A и B и касающейся BC. Так как $\angle PBC = \angle PCA$, P лежит на окружности, проходящей через B и C и касающейся CA. Следовательно, P — точка пересечения этих окружностей, отличная от B. Точка Q строится аналогично.

2. Ответ $\operatorname{ctg}\phi = \operatorname{ctg}A + \operatorname{ctg}B + \operatorname{ctg}C$ следует из формулы, доказываемой ниже.

3. Пусть A', B', C' — точки, симметричные P относительно BC, CA, AB. Так как $CA' = CP = CB'$ и $\angle PCA = \angle QCB$, CQ — серединный перпендикуляр к отрезку $A'B'$, т.е. Q — центр описанной окружности треугольника $A'B'C'$. Значит, середина PQ — центр описанной окружности треугольника, образованного проекциями P на стороны ABC. Аналогично, середина PQ — центр описанной окружности треугольника, образованного проекциями Q. Очевидно, что радиусы этих окружностей равны.

4. Пусть прямые AP, BP, CP вторично пересекают описанную около ABC окружность в точках A', B', C'. Тогда дуги BA', CB' и AC' равны, т.е. треугольник $B'C'A'$ получается из треугольника ABC поворотом вокруг O на угол 2ϕ. Точка P для треугольника $A'B'C'$ является второй точкой Брокара, откуда следуют оба пункта задачи.

5. Пусть C' — точка пересечения прямых AP и BQ. Так как угол между этими прямыми равен 2ϕ, то по предыдущей задаче получаем, что C' лежит на окружности OPQ. Кроме того, очевидно, что $OC' \perp AB$. Поэтому утверждение задачи равносильно тому, что $C'L \parallel AB$. Определим точки A', B' аналогично точке C'. Так как треугольники ABC', BCA' и CAB' подобны, расстояния от A', B', C' до соответствующих сторон треугольника ABC пропорциональны этим сторонам. Так же относятся и расстояния до сторон треугольника от диаметрально противоположной O точки окружности OPQ. Но этим условиям удовлетворяет только точка Лемуана.

6. a) , b) **Указание.** Рассмотрите поворотные гомотетии с центрами P (Q), переводящие Q (P) в O.

c) **Ответ.** Середина отрезка OL, $\pi - 2\phi$.

7. Так как $\angle PAC' = \angle QBC' = \phi$ и $\angle PC'A = \angle QC'B$, треугольники APC' и BQC' подобны, т.е. $AC'/BC' = AP/BQ$. Но из треугольников ACP, BCQ $AP/\sin\phi = AC\sin A$, $BQ/\sin\phi = BC/\sin B$. Следовательно, $AC'/BC' = AC^2/BC^2$ и CC' — симедиана.

8. Это частный случай задачи 27.

2 Точки Брокара в четырехугольнике

9. Доказательство такое же, как в задаче 1.

10. Так как $\angle APB = \pi - \angle B$, $\angle BPC = \pi - \angle C$, то применяя теорему синусов к треугольникам APB и BPC, получаем

$$\frac{PB}{\sin\phi} = \frac{AB}{\sin B}, \quad \frac{PB}{\sin(C-\phi)} = \frac{BC}{\sin C}.$$

Разделив первое уравнение на второе, после преобразований получаем

$$\operatorname{ctg}\phi = \frac{AB}{BC\sin B} + \operatorname{ctg}C.$$

11. Запишем условие $\phi(ABCD) = \phi(DCBA)$ в виде

$$\frac{AB}{BC\sin B} - \operatorname{ctg}B = \frac{CD}{BC\sin C} - \operatorname{ctg}C.$$

Приведя обе части этого равенства к общему знаменателю, возведя их в квадрат и прибавив по единице, получим

$$\frac{AB^2 + BC^2 - 2AB\cdot BC\cos B}{\sin^2 B} = \frac{CD^2 + BC^2 - 2CD\cdot BC\cos C}{\sin^2 C},$$

т.е. $\frac{AC}{\sin B} = \frac{BD}{\sin C}$, ч.т.д.

12. Из построения точек P_i следует, что четырехугольники BCP_1P_2, CDP_2P_3, DAP_3P_4, ABP_4P_1 — вписанные. Отсюда легко вывести, что $\angle P_1P_4P_3 + \angle P_3P_2P_1 = \angle A + \angle C = \pi$.

13. Из вписанного четырехугольника BCP_1P_2 получаем

$$\frac{P_1P_2}{BC} = \frac{\sin(\phi(ABCD) - \phi(BCDA)}{\sin(C + \phi(ABCD) - \phi(BCDA)} = \frac{Q_1Q_2}{CD}.$$

14. См. http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

15. См. http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

16. См. http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

17. Так как $\phi(ABCD) = \phi(DCBA)$, требуемое равенство можно записать в виде

$$\frac{AB}{BC\sin B} + \operatorname{ctg}C = \frac{AD}{DC\sin D} + \operatorname{ctg}C.$$

Поскольку $\sin B = \sin D$, это равносильно искомому.

18. Так как $ABCD$ вписанный, $AB\cdot CD + AD\cdot BC = AC\cdot BD$, т.е. $AB\cdot CD = AC\cdot BD/2$. Пусть M — середина AC. Тогда $CM\cdot BD = BC\cdot AD$, т.е. $BC/CM = BD/AD$. Поскольку $\angle BCM = \angle BDA$, треугольники BCM и BDA подобны. Следовательно, $\angle MBC = \angle ABD$ и BD — симедиана треугольника ABC, что доказывает пп. а)-в).

Для доказательства п. г) достаточно заметить, что любые 4 точки можно инверсией перевести в вершины паралелограмма. При этом вершины вписанного четырехугольника перейдут в вершины прямоугольника, а отношение произведений противоположных сторон не изменится. Следовательно, вершины гармонического четырехугольника перейдут в вершины квадрата.

Для доказательства п. д) рассмотрим центральную проекцию, сохраняющую окружность $ABCD$ и переводящую точку пересечения диагоналей четырехугольника в центр. Тогда четырехугольник перейдет в прямоугольник. Поскольку касательные к окружности в противоположных вершинах прямоугольника должны быть параллельны его диагонали, прямоугольник является квадратом.

19. Так как $\operatorname{ctg}\phi = \frac{AB}{BC\sin B} + \operatorname{ctg}C = \frac{BC}{AB\sin B} + \operatorname{ctg}A$, $\operatorname{ctg}^2\phi - \operatorname{ctg}^2 A = \frac{1}{\sin^2 B}$ или

$$\frac{1}{\sin^2\phi} = \frac{1}{\sin^2 A} + \frac{1}{\sin^2 B}.$$

20. Доказательство такое же, как в задаче 4.

21. Доказательство такое же, как в задаче 5.

3 Точки Брокара в многоугольниках

22. Указание. Докажите, что все прямые X_iX_{i+1} касаются одного эллипса.

23. Ответ.

$$\frac{OL^2}{R^2} + \operatorname{tg}^2\phi\operatorname{tg}^2\frac{\pi}{n} = 1.$$

24. Указание. Рассмотрите повороты многоугольника вокруг O на $\pm\phi$.

25. Доказательство такое же, как при $n = 4$.

26. Доказательство такое же, как при $n = 3$.

27. Указание. T_1, T_2 — предельные точки пучка, порожденного описанной окружностью многоугольника и окружностью OPQ.

A.Zaslavsky, A.Akopyan,
P.Kozhevnikov, A.Zaslavsky

Brocard points

1 Brocard points in triangles

1. Let a triangle ABC be given. Prove that there exists a unique point P, such that $\angle PAB = \angle PBC = \angle PCA = \phi_1$, and a unique point Q, such that $\angle QBA = \angle QCB = \angle QAC = \phi_2$.

Definition 1. Points P and Q are called the *Brocard points* of triangle ABC.

2.
a) Prove that $\phi_1 = \phi_2 = \phi$.
b) Find ϕ as a function of the angles of ABC.

Definition 2. Angle ϕ is called the *Brocard angle* of triangle ABC.

3. Prove that the projections of Brocard points to the sidelines of ABC are concyclic. (This is true for any pair of isogonally conjugated points).

4. Let O be the circumcircle of ABC.
a) Prove that $OP = OQ$.
b) Prove that $\angle POQ = 2\phi$.

Definition 3. The reflections of the medians of a triangle in its correspondent bisectors are called the *symmedians*. Three symmedians concur in point L, which is called the *Lemoine point* of the triangle.

5. Prove that P and Q lie on the circle with diameter OL.

6. (K.Knop) Consider two triangles: one of them is formed by the circumcenters of triangles PAB, PBC, PCA; the second one is formed by the circumcenters of triangles QAB, QBC, QCA. Prove that these triangles are
a) similar to ABC;
b) equal.
c) Find the center and the angle of the rotation transforming one of these triangles to the second one.

7. Let C' be a point of segment AB, such that AB is the external bisector of angle $PC'Q$. Prove that CC' is the symmedian of ABC. (I.e. there exists an ellipse with foci P and Q touching the sides of the triangle in the bases of its symmedians).

8. Let T_1, T_2 be points of line OL, such that $\angle LPT_1 = \angle LPT_2 = 60^\circ$. Prove that the projections of each of these points to the sidelines of ABC form a regular triangle (these points are called the *Apollonius points*).

2 Brocard points in quadrilaterals

9. Let $ABCD$ be a convex broken line. Prove that there exists a unique point P, such that $\angle PAB = \angle PBC = \angle PCD = \phi$.

Definition 4. We will call P and ϕ the *Brocard point* and the *Brocard angle* of broken line $ABCD$. We will denote them as $P(ABCD)$ and $\phi(ABCD)$.

10. Find $\phi(ABCD)$ as a function of the lengths of segments AB, BC, CA and the angles between them.

11. Prove that $\phi(ABCD) = \phi(DCBA)$ iff A, B, C, D are concyclic.

Now we will consider only cyclic polygons.

12. Let $P_1 = P(ABCD)$, $P_2 = P(BCDA)$, $P_3 = P(CDAB)$, $P_4 = P(DABC)$. Prove that $P_1P_2P_3P_4$ is a cyclic quadrilateral.

13. Let $Q_1 = P(DCBA)$, $Q_2 = P(ADCB)$, $Q_3 = P(BADC)$, $Q_4 = P(CBDA)$. Prove that $P_1P_2/Q_1Q_2 = BC/CD$, $P_2P_3/Q_2Q_3 = CD/DA$ etc.

14. (D.Belev) Let M_1, M_2 be points on lines AD, AB respectively such that $BM_1 \parallel CD$, $CM_2 \parallel DA$.

a) Prove that the circumcircles of triangles BAM_1 and BCM_2 meet in P_1.

b) Define the similar construction for $P_i, i = 2, \ldots, 4$, $Q_i, i = 1, \ldots, 4$.

15. (D.Belev) Prove that lines CP_1, DP_2, AP_3, BP_4 concur, and lines BQ_1, CQ_2, DQ_3, AQ_4 concur.

16. (D.Belev) Denote the points obtained in the previous problem as P_0, Q_0.

a) Prove that $S_{P_1P_2P_0} = S_{Q_1Q_2Q_0}$

b) Prove that the areas of $P_1P_2P_3P_4$ and $Q_1Q_2Q_3Q_4$ are equal.

17. Prove that $\phi(ABCD) = \phi(BCDA)$ iff $AB \cdot CD = AD \cdot BC$.

Definition 5. A cyclic quadrilateral with equal products of opposite sides is called *harmonic*. From the last problem we obtain that in the harmonic quadrilateral there exist points P and Q, such that $\angle PAB = \angle PBC = \angle PCD = \angle PDA = \angle QDC = \angle QCB = \angle QBA = \angle QAD = \phi$. We will call P, Q and ϕ the *Brocard points* and the *Brocard angle* of quadrilateral $ABCD$.

18. Prove that each of the following conditions is true iff $ABCD$ is harmonic.

a) The tangents to the circumcircle in A and C meet on BD.

b) BD is a symmedian of ABC.

c) The distances from the common point L of the diagonals to the sides are proportional to these sides.

d) There exists an inversion transforming A, B, C, D to the vertices of a square.

e) There exists a central projection transforming $ABCD$ and its circumcircle to a square and a circle.

19. Find the Brocard angle of a harmonic quadrilateral as a function of its angles.

20. Prove that $OP = OQ$ and $\angle POQ = 2\phi$.

21. Prove that P and Q lie on the circle with diameter OL.

3 Brocard points in polygons

22. Let a circle, a point P inside it and an angle ϕ be given. For an arbitrary point X_0 on the circle construct a point X_1, such that the oriented angle PX_0X_1 is equal to ϕ. Similarly for X_1 construct X_2 etc. Prove that if $X_n = X_0$, then this is true for any other initial point.

23. Find the closure condition in the previous problem.

Remind that all considered polygons are cyclic.

Definition 6. We will call a polygon $A_1 \dots A_n$ a *Brocard* polygon if there exists a point P, such that $\angle PA_1A_2 = \angle PA_2A_3 = \dots \angle PA_nA_1 = \phi$.

24. Prove that in a Brocard polygon there exists a point Q such that $\angle QA_1A_n = \angle QA_nA_{n-1} = \dots \angle QA_2A_1 = \phi$.

Definition 7. We will call P, Q and ϕ the *Brocard points* and the *Brocard angle* of $A_1 \dots A_n$.

25. Prove that each of the following conditions is true iff $A_1 \dots A_n$ is the Brocard polygon.

a) There exists a point L, such that the distances from it to the sides of the polygon are proportional to these sides.

b) The symmedians of triangles $A_1A_2A_3, A_2A_3A_4, \dots, A_nA_1A_2$ from $A_2, A_3, \dots, A_1$ concur.

c) The common points of lines $A_1A_3, A_2A_4, \dots, A_nA_2$ with the tangents to the circumcircle in $A_2, A_3, \dots, A_1$ respectively are collinear.

d) There exists an inversion transforming $A_1, \dots, A_n$ to the vertices of a regular triangle.

e) There exists a central projection transforming the polygon and its circumcircle to a regular polygon and a circle.

26. Prove that the Brocard points lie on the circle with diameter OL and $\angle POL = \angle QOL = \phi$.

27.

a) Prove that there exist two points T_1, T_2 such that the inversion with the center in any of them transforms $A_1, \dots, A_n$ to the vertices of a regular triangle.

b) Prove that T_1, T_2 lie on OL and $\angle T_1PL = \angle T_2PL = \frac{\pi}{n}$.

28. Find the Brocard angle as a function of OL/R.

A.Zaslavsky, A.Akopyan,
P.Kozhevnikov, A.Zaslavsky

Brocard points
Solutions

1 Brocard points in triangles

1. Since $\angle PAB = \angle PBC$ we have $\angle BPA = \pi - \angle B$, i.e. P lies on the circle passing through A and B and touching BC. Since $\angle PBC = \angle PCA$, P lies on the circle passing through B and C and touching CA. Therefore P is the common point of these circles distinct from B. Point Q is constructed similarly.

2. Answer $\mathrm{ctg}\phi = \mathrm{ctg}A + \mathrm{ctg}B + \mathrm{ctg}C$ follows rom the formula which will be proved later.

3. Let A', B', C' be the reflections of P in BC, CA, AB. Since $CA' = CP = CB'$ and $\angle PCA = \angle QCB$, we obtain that CQ is the perpendicular bisector to segment $A'B'$, i.e. Q is the circumcenter of $A'B'C'$. Thus the midpoint of PQ is the center of the circle passing through the projections of P to the sidelines of ABC. Similarly the midpoint of PQ is the center of the circle passing through the projections of Q. It is clear that the radii of these circles are equal.

4. Let AP, BP, CP meet for the second time the circumcircle of ABC in points A', B', C'. Then the arcs BA', CB' and AC' are equal, i.e. triangle $B'C'A'$ is the rotation of ABC around O to angle 2ϕ. Then P is the second Brocard point of $A'B'C'$ and this yields both assertions of the problem.

5. Let C' be a common point of lines AP and BQ. Since the angle between these lines is equal to 2ϕ we obtain by previous problem that C' lie on the circle OPQ. Also it is evident that $OC' \perp AB$. Therefore it is sufficient to prove that $C'L \parallel AB$. Let points A', B' be defined similarly as C'. Since triangles ABC', BCA', and CAB' are similar the distances from A', B', C' to the correspondent sides of ABC are proportional to the lengths of these sides. The ratio from the point of circle OPQ opposite to O to these sides are the same. Thus this point coincide with L.

6. a) , b) **Hint.** Consider the spiral similarities with center P (Q), transforming Q (P) to O.

c) **Answer.** The midpoint of OL, $\pi - 2\phi$.

7. Since $\angle PAC' = \angle QBC' = \phi$ and $\angle PC'A = \angle QC'B$, triangles APC' and BQC' are similar, i.e. $AC'/BC' = AP/BQ$. But from triangles ACP, BCQ we have $AP/\sin\phi = AC\sin A$, $BQ/\sin\phi = BC/\sin B$. Therefore, $AC'/BC' = AC^2/BC^2$ and CC' is a symmedian.

8. This is a partial case of problem 27.

2 Brocard point in quadrilaterals

9. The proof is the same as in problem 1.

10. Since $\angle APB = \pi - \angle B$, $\angle BPC = \pi - \angle C$, we obtain using the sinus theorem to triangles APB and BPC

$$\frac{PB}{\sin\phi} = \frac{AB}{\sin B}, \quad \frac{PB}{\sin(C-\phi)} = \frac{BC}{\sin C}.$$

Dividing the first equation to the second one we have

$$\mathrm{ctg}\phi = \frac{AB}{BC\sin B} + \mathrm{ctg}C.$$

11. The condition $\phi(ABCD) = \phi(DCBA)$ is equivalent to

$$\frac{AB}{BC\sin B} - \mathrm{ctg}B = \frac{CD}{BC\sin C} - \mathrm{ctg}C.$$

Adding the unit to the squares of both parts we obtain

$$\frac{AB^2 + BC^2 - 2AB\cdot BC\cos B}{\sin^2 B} = \frac{CD^2 + BC^2 - 2CD\cdot BC\cos C}{\sin^2 C},$$

i.e. $\frac{AC}{\sin B} = \frac{BD}{\sin C}$, ч.т.д.

12. By the construction of P_i we obtain that quadrilaterals BCP_1P_2, CDP_2P_3, DAP_3P_4, ABP_4P_1 are cyclic. From this $\angle P_1P_4P_3 + \angle P_3P_2P_1 = \angle A + \angle C = \pi$.

13. Since BCP_1P_2 is a cyclic quadrilateral we obtain

$$\frac{P_1P_2}{BC} = \frac{\sin(\phi(ABCD) - \phi(BCDA)}{\sin(C + \phi(ABCD) - \phi(BCDA)} = \frac{Q_1Q_2}{CD}.$$

14. See http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

15. See http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

16. See http://jcgeometry.org/Articles/Volume2/Belev_Brocard_points.pdf

17. Since $\phi(ABCD) = \phi(DCBA)$, the sought equality is equivalent to

$$\frac{AB}{BC\sin B} + \mathrm{ctg}C = \frac{AD}{DC\sin D} + \mathrm{ctg}C.$$

Since $\sin B = \sin D$ we obtain the assertion of the problem.

18. Since $ABCD$ is cyclic, $AB\cdot CD + AD\cdot BC = AC\cdot BD$, i.e. $AB\cdot CD = AC\cdot BD/2$. Let M be the midpoint of AC. Then $CM\cdot BD = BC\cdot AD$, i.e. $BC/CM = BD/AD$. Since $\angle BCM = \angle BDA$, triangles BCM and BDA are similar. Therefore, $\angle MBC = \angle ABD$ and BD is the symmedian of ABC, which yields a)-c).

For prove d) note that four arbitrary points can be transformed by an inversion to the vertices of a parallelogram. If the given points are concyclic this parallelogram will be a rectangle with the same ratio of the products of the opposite sides. Thus the vertices of a harmonic quadrilateral will be transformed to the vertices of a square.

For prove e) consider a central projection conserving the circumcircle of $ABCD$ and transforming the common point of its diagonals to the center. Then the image of the quadrilateral will be a rectangle. Since the tangents to the circumcircle in the opposite vertices of this rectangle are parallel to its diagonal the rectangle is a square.

19. Since $\operatorname{ctg}\phi = \frac{AB}{BC\sin B} + \operatorname{ctg}C = \frac{BC}{AB\sin B} + \operatorname{ctg}A$, $\operatorname{ctg}^2\phi - \operatorname{ctg}^2 A = \frac{1}{\sin^2 B}$ or

$$\frac{1}{\sin^2\phi} = \frac{1}{\sin^2 A} + \frac{1}{\sin^2 B}.$$

20. The proof is the same as in problem 4.
21. the proof is the same as in problem 5.

3 Brocard points in polygons

22. Hint. Prove that all lines X_iX_{i+1} are the tangents to the same ellipse.
23. Answer.

$$\frac{OL^2}{R^2} + \operatorname{tg}^2\phi\operatorname{tg}^2\frac{\pi}{n} = 1.$$

24. Hint. Consider the rotations around O to $\pm\phi$.
25. The proof is the same as for $n = 4$.
26. The proof is the same as for $n = 3$.
27. Hint. T_1, T_2 are the limit points of the pencil containing the circumcircle of the polygon and the circumcircle of OPQ.

Диофантовы уравнения–1

Теорема (Гаусс)**.** Натуральное число представимо в виде суммы трёх квадратов, если и только если оно не представимо в виде $4^n(8m-1)$.

Вводные задачи

Задача 1. Докажите, что уравнения а) $2x^2+2xy-y^2=1$, b) $x^2-xy+y^2=2$ не имеют решений в целых числах.

Задача 2. Докажите, что уравнения а) $x^2-2y^2=1$, b) $x^2-3y^2=1$, c) $x^2-6y^2=1$ имеют бесконечно много решений в целых числах.

Задача 3. Докажите, что уравнение $x^2+1000xy+1000y^2=2001$ имеет бесконечно много решений в целых числах.

Задача 4.* Фиксируем нечётное простое число p. Докажите, что уравнение $x^2-py^2=-1$ имеет решение в целых числах, если и только если p имеет остаток 1 при делении на 4.

Задача 5. Докажите, что для всякого m количества решений в целых числах уравнений

$$x^2-xy+y^2=m \quad \text{и} \quad 3x^2+9xy+7y^2=m$$

одинаковы.

Задача 6. Докажите, что для всякого целого числа n уравнение $x^2+y^2=n$ имеет решение в целых числах, если и только если оно имеет решения в рациональных числах.

Задача 7. Приведите пример квадратичного уравнения с целыми коэффициентами, имеющего решения в рациональных числах, но не имеющего решений в целых числах.

Задача 8. Докажите, что для любых целых положительных чисел a и b существует бесконечно много натуральных чисел m, для которых уравнение $ax^2+by^2=m$ не имеет решений в целых числах.

Задача 9. Докажите, что для всякого целого числа m уравнение $x^2+2y^2-3z^2=m$ имеет решение в целых числах.

Квадратичные формы

Квадратичная форма — это однородный многочлен степени 2. По определению, квадратичная форма f *представляет* число m, если уравнение $f=m$ имеет ненулевые решения в целых числах (тонкость: не всякая форма представляет 0). Две квадратичные формы называются *эквивалентными*, если они представляют одно и то же множество чисел.

Задача 10. Опишите все целые числа, которые представляются формами а) x^2+y^2; b) x^2-y^2; c)* x^2+xy+y^2.

Задача 11. Докажите, что квадратичные формы

$$f(x,y), \quad f(x-y,y), \quad f(x,y-x), \quad f(-x,y) \quad \text{и} \quad f(x,-y)$$

попарно эквивалентны.

Задача 12. a) Докажите, что формы x^2+y^2 и x^2+xy+y^2 не эквивалентны.

b) Докажите, что форма $4x^2-6xy+5y^2$ не эквивалентна форме вида ax^2+by^2 ни для каких целых чисел a и b.

Определение 1. Квадратичная форма называется

a) *положительно определённой*, если она представляет только положительные числа,

b) *неотрицательно определённой*, если она представляет только неотрицательные числа,

c) *неопределённой*, если она представляет и положительные, и отрицательные числа.

Задача 13. Приведите пример неотрицательно определённой формы, которая не является положительно определённой.

Расширенная арифметика: p-адические числа

Теорема (Лежандр)**.** Всякое положительное целое число представимо в виде суммы четырёх квадратов целых чисел.

Задача 14. Пусть m и n – целые числа, свободные от квадратов. Если уравнение

$$z^2-mx^2-ny^2=0 \tag{1}$$

имеет ненулевое решение в рациональных числах, то выполнены следующие условия

a) хотя бы одно из чисел m,n положительно,

b) m является квадратичным вычетом при делении на n,

c) n является квадратичным вычетом при делении на m.

Задача 15. Сведите метатеорему для двух переменных к решению уравнений вида (1).

Определение 2. Выражение вида

$$a_{-k}p^{-k}+a_{-k+1}p^{-k+1}+\ldots+a_np^n+\ldots \tag{2}$$

(k — произвольное целое число, $a_i\in\mathbb{Z}$) называется *p-адическим числом*. Если $k\leqslant 0$, то мы называем (2) *целым* p-адическим числом.

Задача 16. Уравнение с целыми коэффициентами $f=0$ имеет решение в $\mathbb{Z}_p$, если и только если оно имеет решение в остатках при делении на p^n для всякого $n\in\mathbb{Z}_{\geqslant 0}$.

Задача 17. Когда p-адическое число в форме (2) равно 0?

Задача 18. Докажите, что произведение двух ненулевых p-адических чисел не равно 0.

Задача 19. Докажите, что $\mathbb{Q}\subset\mathbb{Q}_p$ для всякого простого числа p (докажите, что для всякой пары ненулевых целых чисел m,n существует p-адическое число x такое, что $nx=m$).

Задача 20. Докажите, что -1 является полным квадратом в p-адических числах тогда и только тогда, когда p имеет остаток 1 при делении на 4.

Задача 21. Придумайте описание для p-адических чисел, являющихся полными квадратами.

Задача 22. Докажите, что любое ненулевое 3-адическое число m есть или x^2, или $2x^2$, или $3x^2$, или $6x^2$ для какого-то 3-адического числа x.

Задача 23. Пусть p — нечётное простое число, а $x_1,\ldots,x_5$ — ненулевые p-адические числа. Докажите, что x_i/x_j есть полный квадрат в p-адических числах для каких-то i,j ($1\leqslant i<j\leqslant 5$).

Задача 24. Докажите, что для всякого нечётного простого числа p существуют ненулевые p-адические числа $x_1, \ldots, x_{p-1}$ такие, что $x_1^2 + \ldots + x_{p-1}^2 + 1 = 0$.

Задача 25. Докажите, что уравнение $x^2 + x + 1 = 0$ имеет ровно два решения в целых 7-адических числах.

Задача 26. Докажите, что уравнение $x^2 + y^2 = -1$ имеет решения в p-адических числах для всякого нечётного простого числа p.

Теорема (Принцип Минковского-Хассе)**.** Квадратное уравнение $f = 0$ от нескольких переменных имеет решение в рациональных числах, если и только если оно одновременно имеет решения

- в вещественных числах,
- в p-адических числах ($:=\mathbb{Q}_p$) для всякого простого числа p.

Задача 27. Докажите принцип Минковского-Хассе для уравнений от одной и двух переменных.

Определение 3. Положим $(a, b)_p = 1$, если $z^2 - ax^2 - by^2 = 0$ имеет p-адические решения, и положим $(a, b)_p = -1$ иначе. Значение $(a, b)_p$ называется *символом Гильберта* пары (a, b) относительно простого числа p.

Задача 28. Докажите, что для символа Гильберта выполнены следующие соотношения

1) $(a, b)_p = (b, a)_p$, 2) $(a, c^2)_p = 1$,
3) $(a, -a)_p = 1$, $\quad (a, 1 - a)_p = 1$, 4) $(a, b)_p = (a, -ab)_p = (a, (1 - a)b)_p$.

Задача 29. Пусть $(a, b)_p = 1$. Тогда $(a', b)_p = (aa', b)_p$ для любого a'.

Определение 4. Чтобы компактно записать явную формулу для символа Гильберта, нам потребуется *символ Лежандра* $\left(\frac{x}{p}\right)$, определённый для любых целого x и простого p. Он равен 1, -1 или 0 в зависимости от того, является x ненулевым квадратичным вычетом, невычетом или нулём по модулю p. Для нечётного простого p символ Лежандра вычисляется по формуле

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}} \pmod{p}.$$

Задача 30. Пусть p – нечётное простое число, $a = p^\alpha u$, $b = p^\beta v$, где α, β, u, v – это целые числа такие, что u и v взаимно просты с p. Докажите, что

$$(a, b)_p = (-1)^{\alpha\beta\varepsilon(p)} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha,$$

где $\varepsilon(p) := \frac{p-1}{2}$.

Задача 31. Найдите явную формулу для $(a, b)_2$ при всех целых числах a, b.

Задача 32. Докажите, что $(a, b)_p(a, b')_p = (a, bb')_p$ для любых целых чисел a, b, b'.

Задача 33. Докажите, что уравнение $ax^2 + by^2 = c$ (a, b, c – это параметры; x, y – это переменные) имеет решение в p-адических числах, если и только если $(c, -ab)_p = (a, b)_p$.

Задача 34*. Фиксируем однородный многочлен $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2 (n \geqslant 2)$, где $a_1, \ldots, a_n \neq 0$. Положим

$$d = a_1a_2 \ldots .a_n \quad \text{и} \quad \varepsilon = \prod_{i<j}(a_i, a_j)_p. \tag{3}$$

Докажите, что уравнение $f = 0$ имеет ненулевое решение в p-адических числах тогда и только тогда, когда выполнено одно из следующих условий

1) $n = 2$, а число $-d$ является полным квадратом в $\mathbb{Q}_p$;

2) $n = 3$ и $(-1, d)_p = \varepsilon$;

3) $n = 4$ и $d \neq \alpha^2$, или же $d = \alpha^2$ и $\varepsilon = (-1, -1)_p$;

4) $n \geqslant 5$. (т.е., если f зависит от 5 и более переменных, то уравнение $f = 0$ имеет ненулевое решение в $\mathbb{Q}_p$ для любого p.)

Выведите из задачи 34 следующее утверждение.

Задача 35. Фиксируем однородный многочлен $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$ $(n \geqslant 2)$, где $a_1, \ldots, a_n \neq 0$, а также целое число $a \neq 0$. Определим d и ε формулой (3). Докажите, что уравнение $f = a$ имеет решение в p-адических числах тогда и только тогда, когда выполнено одно из следующих условий:

1) $n = 1$, а число a/d является полным квадратом в $\mathbb{Q}_p$;

2) $n = 2$ и $(a, -d)_p = \varepsilon$;

3) $n = 3$ и: ad не является точным квадратом в $\mathbb{Q}_p$ или ad является точным квадратом и $\varepsilon = (-1, -d)_p$;

4) $n \geqslant 4$. (Т.е., если f зависит от 4 и более переменных, то уравнение $f = a$ имеет ненулевое решение в $\mathbb{Q}_p$ для любого p).

Задача 36. Докажите принцип Минковского-Хассе.

Задача 37. Используя задачу 35 и принцип Минковского-Хассе, докажите, что целое число n представимо в виде суммы трёх квадратов рациональных чисел, если и только если оно не представимо в виде $4^a(8b - 1)$, т.е. если $-n$ не является полным квадратом в $\mathbb{Q}_2$.

Задача 38. Фиксируем целое число n. Докажите, что если существуют рациональные числа x, y, z такие, что $x^2 + y^2 + z^2 = n$, то существуют и целые числа x', y', z' такие, что

$$(x')^2 + (y')^2 + (z')^2 = n.$$

Выведите из этого утверждения теорему Гаусса.

Задача 39. Выведите из теоремы Гаусса теорему Лежандра.

Важные свойства символа Гильберта (ДУ-2)

Цель этого раздела — доказать, что для фиксированной пары ненулевых целых чисел (a, b) символ Гильберта $(a, b)_p$ равен 1 для почти всех (=всех, кроме конечного числа) простых чисел p. Как водится, это утверждение является частным случаем более общего утверждения.

Задача 40. a) Пусть f – это однородный многочлен степени n от k переменных, где $k > n$. Тогда число решений f (включая нулевое) в остатках при делении на p делится на p (Подсказка: примените малую теорему Ферма и рассмотрите случай $p = 2$).

b) Пусть f – это многочлен степени не более n от k переменных, где $k > n$. Тогда число решений уравнения $f = 0$ в остатках при делении на p делится на p.

Задача 41. Выведите из предыдущей задачи, что уравнение $ax^2+by^2+cz^2 = 0$ от переменных x, y, z имеет ненулевое решение в остатках при делении на p.

Задача 42. Выведите из предыдущей задачи, что для пары ненулевых целых чисел (a, b) и нечётного простого числа p символ Гильберта $(a, b)_p$ равен единице, если $a, b \not\vdots p$. Объясните, почему символ Гильберта $(a, b)_p$ равен 1 для почти всех p.

Задача 43. Выведите из задачи 41, что уравнение $ax^2+by^2+cz^2+dv^2+ew^2 = 0$ от переменных x, y, z, v, w (a, b, c, d, e – это целые параметры) имеет ненулевое решение в $\mathbb{Q}_p$ для всех простых чисел p.

Задача 44. Докажите, что для всякой пары ненулевых целых чисел (a, b) имеет место равенство

$$\prod_p (a, b)_p = (a, b)_{-1},$$

где произведение берётся по всем простым числам p, а

$$(a, b)_{-1} = \begin{cases} 1, & \text{если уравнение } z^2 - ax^2 - by^2 = 0 \text{ имеет решение в } \mathbb{R}, \\ -1 & \text{иначе.} \end{cases}$$

Имеет место следующий «далёкий аналог китайской теоремы об остатках»: оказывается, что по значениям символа Гильберта может быть построен элемент с данными значениями.

Задача 45. Зафиксируем конечный набор ненулевых целых чисел a_i и для каждого простого p зададим значения $\varepsilon_{i,p} = \pm 1$. Тогда система уравнений

$$(a_i, x)_p = \varepsilon_{i,p} \quad \forall i,\ \forall p,$$

имеет решение, если и только если

a) почти все (=все кроме конечного числа) $\varepsilon_{i,p} = 1$,

b) для каждого простого числа p существует ненулевое p-адическое число x_p такое, что

$$(a_i, x_p) = \varepsilon_{i,p}.$$

Уравнения от двух переменных и карты (ДУ-3)

Рассматривается уравнение

$$E_m: \quad ax^2 + bxy + cy^2 = m \tag{4}$$

от целых переменных x, y, где a, b, c, m — какие-то целые числа (параметры).

Задача 46 (Суперзадача)**.** Докажите, что если уравнение E_m имеет решения при каком-то положительном числе m, при каком-то отрицательном числе m и не имеет решений при $m = 0$, то для всякого m или E_m не имеет решений, или же E_m имеет бесконечно много решений.

Задача 47 (Суперзадача)**.** Верно ли, что если уравнение E_m имеет решения в целых числах при

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

то E_m имеет решения при всяком целом числе m?

Задача 48 (Суперзадача)**.** Докажите, что если уравнения E_1, E_2, E_3, E_5 имеют решения в целых числах, то уравнение E_m имеет решения при каком-то $m < 0$.

Рисуем картинки

Задача 49. Покажите, что если $\{w_1, w_2\}$ — это базис $\mathbb{Z}^2$, то пары

$$\{w_2, w_1\},\ \{w_1 - w_2, w_2\},\ \{w_1 + w_2, w_2\},\ \{-w_1, w_2\} \tag{5}$$

также являются базисами $\mathbb{Z}^2$.

Задача 50. Покажите, что преобразованиями (5) можно из любого базиса получить любой другой.

Задача 51. Покажите, что квадратичная форма может записываться одинаково в нескольких разных базисах.

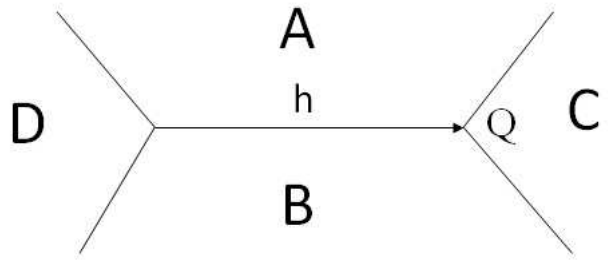
Задача 52. Укажите квадратичную форму, для которой любым двум разным базисам $\mathbb{Z}^2$ соответствуют различные квадратичные формы.

Упражнение 1. Выпишите все расширения данного базиса $\{w_1, w_2\}$. Выпишите все специализации данного супербазиса $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$.

Упражнение 2. Нарисуйте (ориентированные) карты для квадратичных форм

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

В следующих двух задачах числа A, B, C, D, h относятся к картинке



Задача 53. Покажите, что для чисел A, B, C, D, h выполнены соотношения:

$$C = A + B + h, \quad D = A + B - h.$$

Задача 54. Пусть A, B, C положительны, а ребро h направлено от C к D. Покажите, что тогда число D также положительно, а стрелки на двух остальных рёбрах, инцидентных вершине Q, направлены прочь от Q.

Задача 55. Докажите, что граф, задаваемый точками-супербазисами и рёбрами-базисами, является деревом, т.е. не содержит циклов.

Задача 56. Пусть Q — единственный колодец положительно определённой квадратичной формы f, а p, q, r — это числа, записанные в областях, примыкающих к Q. Покажите, что в любой другой области карты f написано число, большее, чем любое из чисел p, q, r.

Задача 57. Докажите, что всякая положительно определённая квадратичная форма обладает колодцем.

Задача 58. a) Докажите, что положительно определённая квадратичная форма имеет не более двух колодцев.

b) Укажите квадратичную форму, обладающую двумя колодцами.

Задача 59. Объясните, как решить уравнение $ax^2 + bxy + cy^2 = m$ (a, b, c, m — параметры, x, y, z — переменные) в предположении, что форма $ax^2+bxy+cz^2$ положительно определена.

Задача 60 (Классификация положительно определённых квадратичных форм)**.**
a) Покажите, что каждая положительно определённая квадратичная форма эквивалентна квадратичной форме вида

$$(p + q)x^2 + 2qxy + (q + r)y^2 \tag{6}$$

для какого-то набора положительных чисел p, q, r.

b) Покажите, что две квадратичные формы, соответствующие наборам

$$(p_1, q_1, r_1) \text{ и } (p_2, q_2, r_2),$$

эквивалентны тогда и только тогда, когда эти наборы совпадают как множества.

c) Определите, какие наборы (p, q, r) задают целую квадратичную форму.

d) Определите, какие наборы (p, q, r) задают положительно определённую квадратичную форму.

Часть 3: Малая мафусаилова форма

Целью этой части является доказательство следующей теоремы.

Теорема (Конвей)**.** Малая мафусаилова форма $x^2 + 2y^2 + yz + 4z^2$ представляет все числа от 1 до 30. Всякая другая положительно определённая форма f, представляющая все числа от 1 до 30, эквивалентна с точностью до линейной замены малой мафусаиловой форме.

Для того, чтобы доказать теорему Конвея, мы предлагаем участникам развить теорию положительно определённых квадратичных форм от трёх переменных, отталкиваясь от теории положительно определённых форм от двух переменных.

Начнём мы с проработки новой точки зрения на положительно определённые квадратичные формы от двух переменных. Пусть $f(x, y) = ax^2+bxy+cy^2$ – это некоторая положительно определённая квадратичная форма. Мы будем задавать такую форму с помощью таблиц 2x2 и 3x3

$$F := \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \hat{F} := \begin{pmatrix} a & \frac{b}{2} & -(a+\frac{b}{2}) \\ \frac{b}{2} & c & -(c+\frac{b}{2}) \\ -(a+\frac{b}{2}) & -(c+\frac{b}{2}) & (a+b+c) \end{pmatrix}.$$

Задача 61. Докажите, что

$$f(x, y) = -\frac{b}{2}(x - y)^2 + (a + \frac{b}{2})x^2 + (c + \frac{b}{2})y^2. \tag{7}$$

Задача 62. Докажите, что таблицы

$$\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \begin{pmatrix} c & \frac{b}{2} \\ \frac{b}{2} & a \end{pmatrix}, \quad \begin{pmatrix} a & -\frac{b}{2} \\ -\frac{b}{2} & c \end{pmatrix} \tag{8}$$

задают эквивалентные квадратичные формы.

Задача 63. Докажите, что квадратичные формы, соответствующие таблицам

$$\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \begin{pmatrix} a & -(a+\frac{b}{2}) \\ -(a+\frac{b}{2}) & a+b+c \end{pmatrix}, \quad \begin{pmatrix} c & -(c+\frac{b}{2}) \\ -(c+\frac{b}{2}) & a+b+c \end{pmatrix}, \tag{9}$$

эквивалентны (заметим, что все таблицы (9) получаются из $\hat{F}$ выбором 2 строк и 2 соответствующих им столбцов).

Далее мы отождествляем квадратичную форму f с её таблицами F и $\hat{F}$.

Задача 64. Используя (8) и (9), докажите что всякая положительно определённая квадратичная форма эквивалентна форме

$$\begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix},$$

для которой $0 \leqslant -b' \leqslant a' \leqslant c'$. Для формы такого вида правая часть (7) представляет собой сумму трёх неотрицательных чисел.

Задача 64 является аналогом задачи 60 второй части проекта. Мы хотим доказать аналог задачи 64 для квадратичных форм от трёх переменных. Мы будем действовать по той же схеме, но нам потребуется больше обозначений. Фиксируем квадратичную форму

$$f(x,y,z) = a_{xx}x^2 + a_{yy}y^2 + a_{zz}z^2 + a_{xy}xy + a_{yz}yz + a_{xz}xz.$$

Мы будем задавать такую форму с помощью следующих таблиц 3x3 и 4x4

$$F := \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} \end{pmatrix},$$

$$\hat{F} := \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}.$$

Задача 65. Докажите, что

$$f(x,y,z) = -\frac{a_{xy}}{2}(x-y)^2 - \frac{a_{xz}}{2}(x-z)^2 - \frac{a_{yz}}{2}(y-z)^2 + \\ + (a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2})x^2 + (a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2})y^2 + (a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2})z^2. \quad (10)$$

Задача 66. Докажите, что квадратичные формы, соответствующие таблицам

$$\begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} \end{pmatrix}, \quad \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xy}}{2} & a_{yy} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}, \quad (11)$$

$$\begin{pmatrix} a_{xx} & \frac{a_{xz}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}, \quad (12)$$

$$\begin{pmatrix} a_{yy} & \frac{a_{yz}}{2} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ \frac{a_{yz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}, \quad (13)$$

эквивалентны (заметим, что все таблицы (13) получаются из $\hat{F}$ выбором 3 строк и 3 соответствующих им столбцов).

Далее мы отождествляем квадратичную форму f с её таблицами F и $\hat{F}$.

Задача 67. Используя (13), докажите что всякая положительно определённая квадратичная форма эквивалентна форме

$$\begin{pmatrix} a'_{xx} & \frac{a'_{xy}}{2} & \frac{a'_{xz}}{2} \\ \frac{a'_{xy}}{2} & a'_{yy} & \frac{a'_{yz}}{2} \\ \frac{a'_{xz}}{2} & \frac{a'_{yz}}{2} & a'_{zz} \end{pmatrix},$$

для которой

$$0 < a'_{xx} \leqslant a'_{yy} \leqslant a'_{zz},$$

$$|a'_{xy}|, |a'_{xz}| \leqslant |a'_{xx}|, |a'_{yz}| \leqslant |a'_{zz}|.$$

Задача 68. Используя задачу (67), докажите, что всякая положительно определённая квадратичная форма $f(x, y, z)$ эквивалентна квадратичной форме f' с таблицей 4x4

$$\hat{F} := \begin{pmatrix} a'_{xx} & \frac{a'_{xy}}{2} & \frac{a'_{xz}}{2} & -(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) \\ \frac{a'_{xy}}{2} & a'_{yy} & \frac{a'_{yz}}{2} & -(a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) \\ \frac{a'_{xz}}{2} & \frac{a'_{yz}}{2} & a'_{zz} & -(a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) \\ -(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) & -(a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) & -(a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) & a'_{xx} + a'_{yy} + a'_{zz} + \\ & & & + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2} + \frac{a'_{xz}}{2} \end{pmatrix}, \tag{14}$$

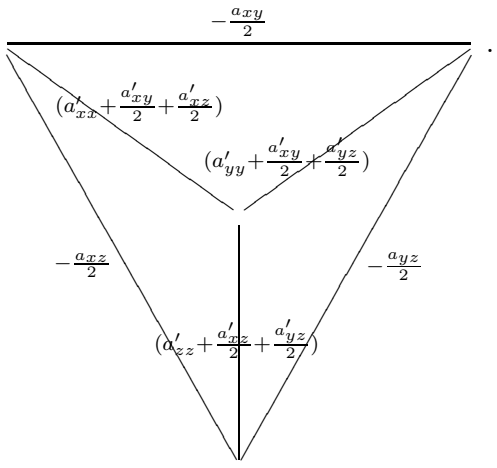
для которой

$$a'_{xy},\ a'_{yz},\ a'_{xz} \leqslant 0, \tag{15}$$

$$(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) \geqslant 0,\ (a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) \geqslant 0,\ (a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) \geqslant 0. \tag{16}$$

Для формы такого вида правая часть (10) представляет собой сумму положительных квадратов.

Квадратичные формы, соответствующие таблице (14), удовлетворяющей условиям (16), мы будем изображать с помощью графа на четырёх вершинах, как показано на рисунке

$\mathcal{F} =$ 

Если на каком-то ребре написано число 0, то это ребро стирается.

Задача 69. Докажите, что если на рисунке $\mathcal{F}$, построенном по квадратичной форме f, присутствуют все рёбра, то форма f не представляет 1.

Задача 70. Докажите, что если на рисунке $\mathcal{F}$ формы f есть вершина, не соединённая рёбрами с 2 другими вершинами, то f эквивалентна форме вида

$$ax^2 + g(y, z) \tag{17}$$

для какого-то положительного целого числа a и положительно определённой квадратичной формы g.

Задача 71. Докажите, что если форма f вида (17) представляет все числа от 1 до 30, то она эквивалентна малой мафусаиловой форме.

Назовём форму $f(x, y)$, в каком-то рисунке которой каждая вершина соединена с 2 или более рёбрами, *неразложимой*.

Задача 72. Опишите рисунки всех неразложимых квадратичных форм $f(x, y, z)$, представляющих числа

a) 1;
b) 1, 2;
c) 1, 2, 3, 5.

Задача 73. Завершите доказательство теоремы Конвея.

Диофантовы уравнения второй степени

Проект посвящён изучению диофантовых уравнений второй степени. Мы надеемся, что участники проекта разовьют теорию, которая позволит им решать достаточно большой класс задач. Наиболее яркие из этих задач приведены ниже.

До промежуточного финиша мы будем работать с уравнениями второй степени в рациональных числах. Мы выпишем явный алгоритм, с помощью которого можно эффективно определить, имеет ли данное уравнение решение или нет. Как приложение развитой техники, мы докажем следующую теорему, принадлежащую Карлу Фридриху Гауссу.

Теорема (Гаусс)**.** Натуральное число представимо в виде суммы трёх квадратов, если и только если оно не представимо в виде $4^n(8m-1)$.

После промежуточного финиша мы сосредоточимся на уравнениях в целых числах от двух переменных. Мы научимся эффективно решать такие уравнения, используя карты квадратичных форм. Также мы докажем следующее утверждение.

Теорема (Дж. Конвей)**.** Существует единственный [1] однородный многочлен $f(x,y,z)$ степени 2, для которого уравнение $f(x,y,z)=m$ имеет решение для всякого $m=1,\ldots,30$ и не имеет решений при $m<0$.

Вводные задачи

В этой части мы собрали задачи на целочисленные квадратичные формы, которые могут быть решены с помощью единого и общего алгоритма решения квадратичных уравнений (который, как мы надеемся, будет построен участниками). Впрочем, все эти задачи могут быть решены и непосредственно. Заметим, что общего алгоритма решения диофантовых уравнений произвольной степени не существует и принципиально существовать не может (это 10-ая проблема Гильберта, решенная отрицательно Ю. Матиясевичем в 1970-м году).

Задачи 1–9.

Если какие-то из этих задач у Вас не получилось решить, скажем, за час — не огорчайтесь. Вы всегда можете вернуться к ним позднее, имея больше технических средств.

Квадратичные формы

Определение 1. Мы называем *квадратичной формой* однородный многочлен степени 2 от какого-то числа переменных. Примерами квадратичных форм являются многочлены

$$2x^2+2xy-y^2 \text{ и } x^2-xz+y^2-2z^2.$$

Для всякого натурального числа d мы обозначим через $\mathbb{Z}^d$ множество наборов из d целых чисел. Например, множество пар целых чисел обозначается нами $\mathbb{Z}^2$. Всякая квадратичная форма от двух переменных x,y задаёт функцию на $\mathbb{Z}^2$, т.е. сопоставляет каждой паре чисел (x,y) число $f(x,y)$. В дальнейшем, мы часто будем заменять элемент $(x,y)\in\mathbb{Z}^2$ одной буквой (скажем, v) и писать $f(v)$ вместо $f(x,y)$.

Определение 2. Мы будем говорить, что квадратичная форма *представляет целое число* n, если $\exists v\in\mathbb{Z}^d \mid f(v)=n$, или, что то же самое, если уравнение

$$f(x,y)=n$$

имеет решение в целых числах.

[1] Это утверждение формально неверно. Оно будет уточнено позже.

Задачи 10–11.

Определение 3. Мы назовём две квадратичные формы *эквивалентными.* если они представляют одно и то же множество целых чисел.

Задача 12.

С какими-то квадратичными формами работать проще, а с какими-то — сложнее. Хотелось бы найти для каждой квадратичной формы как можно более удобного представителя её класса эквивалентности (скажем, квадратичную форму вида $ax^2 + by^2$). Для этого полезно иметь какие-то разумные критерии того, когда две квадратичные формы эквивалентны. Более того, полезно иметь какие-то легко вычислимые инварианты квадратичных форм. Мы предложим некоторый набор таких инвариантов.

Определение 4. Квадратичная форма f называется *положительно определённой*, если $f(v) > 0$ для всякого $v \neq 0$. Квадратичная форма называется *неотрицательно определённой*, если $f(v) \geqslant 0$ для всякого $v \in \mathbb{Z}^2$. Наконец, квадратичная форма называется *неопределённой*, если $f(u) > 0$ и $f(v) < 0$ для некоторых $u, v \in \mathbb{Z}^2$.

Задача 13.

Расширенная арифметика: p-адические числа

В этой части проекта мы займёмся доказательствами следующей теоремы.

Теорема (Метатеорема)**.** Квадратное уравнение от произвольного числа переменных имеет решения в рациональных числах тогда и только тогда, когда к этому нет препятствий, связанных с остатками при делении на простые числа.

С помощью Метатеоремы мы докажем теорему (Гаусса) и следующую теорему Лежандра.

Теорема (Лежандр)**.** Всякое положительное целое число представимо в виде суммы четырёх квадратов целых чисел.

В проекте мы предлагаем подразбиение теоремы (как и теоремы Гаусса/Лежандра) на несколько подзадач, каждая из которых может быть решена самостоятельно. Для начала нам нужно придать какой-нибудь формальный смысл формально неопределённой (и неформально неверной) метатеореме.Начнём мы с примера, а именно со следующего утверждения, являющегося частным случаем метатеоремы.

Определение 5. Число m называется *квадратичным вычетом* в остатках при делении на n, если существует целое число t такое, что $m \equiv t^2 (\mathrm{mod}\ n)$.

Задачи 14–15.

Если НОД$(m, n) = 1$, то из условий a)-c) задачи 14 следует существование ненулевого рационального решения уравнения

$$ax^2 + by^2 = c.$$

Если НОД$(m, n) \neq 1$, то на пару (m, n) должны быть наложены дополнительные условия, связанные с простыми делителями НОД(m, n). Эти условия достаточно просты, но громоздки. Изящный способ записи этих условий связан с понятием p-адического числа. Для начала мы определим p-адические числа и выявим их простейшие ($\sim$ важнейшие) свойства.

Для каждого простого числа p определим $\mathbb{Z}_p$ как множество формальных выражений вида

$$a_0 + a_1 p + \ldots + a_n p^n + \ldots \quad (a_i \in \mathbb{Z}) \tag{1}$$

(число слагаемых может быть бесконечно). Два таких выражения считаются равными, если они совпадают с точностью до членов порядка p^n для любого n. К примеру,

$$1 = (p+1) - (p+1)p + (p+1)p^2 - (p+1)p^3 + \dots.$$

Выражения вида (1) можно складывать, вычитать и умножать. Таким образом, для каждого уравнения $f = 0$ с коэффициентами в целых числах можно рассматривать его решения в $\mathbb{Z}_p$ (=: целых p-адических числах). Следующая задача указывает на связь решений в целых числах и решений в целых p-адических числах.

Задача 16.

Целые p-адические числа представляют собой расширение понятия целого числа. Такое же p-расширение имеет и понятие рационального числа. Для каждого простого p определим $\mathbb{Q}_p$ как множество формальных выражений вида

$$a_{-k}p^{-k} + a_{-k+1}p^{-k+1} + \ldots + a_n p^n + \ldots \tag{2}$$

(k — произвольное целое число, $a_i \in \mathbb{Z}$). Очевидно, что всякое целое p-адическое число имеет вид (2) с $a_{-k} = ... = a_{-1} = 0$.

Для того, чтобы помочь участникам привыкнуть к понятию p-адического числа, ниже мы приводим несколько задач.

Задачи 17–26.

Теперь мы готовы к формальной версии Метатеоремы.

Теорема (Принцип Минковского-Хассе)**.** Квадратное уравнение $f = 0$ от нескольких переменных имеет решение в рациональных числах, если и только если оно имеет решения

а) в вещественных числах,

б) p-адических числах ($:=\mathbb{Q}_p$) для всякого простого числа p.

Задача 27.

Принцип Минковского-Хассе сводит решение уравнений в рациональных числах к решению тех же уравнений в числах p-адических. При этом подразумевается, что решать уравнения в числах p-адических много проще. Для начала мы сформулируем в виде набора задач алгоритм решения квадратного уравнения от двух переменных в p-адических числах. Начнём мы с однородного уравнения

$$z^2 - ax^2 - by^2 = 0. \tag{3}$$

Определение 6. Положим $(a, b)_p = 1$, если уравнение (3) имеет ненулевое решение в целых p-адических числах. В противном случае положим $(a, b)_p = -1$. Значение $(a, b)_p$ называется *символом Гильберта* пары (a, b) по отношению к простому числу p.

Таким образом, для решения уравнения (3) хочется научиться находить значения $(a, b)_p$.

Задачи 28–29.

Чтобы компактно записать явную формулу для символа Гильберта, нам потребуется *символ Лежандра* $\left(\dfrac{x}{p}\right)$, определённый для любых целого x и простого p. Он равен 1, -1 или 0 в зависимости от того, является x ненулевым квадратичным вычетом, невычетом или нулём по модулю p. Для нечётного простого p символ Лежандра вычисляется по формуле

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}} \pmod{p}.$$

Задачи 30–39.

Уравнения от двух переменных и карты

В этой части проекта мы разовьём технику, позволяющую эффективно решать в целых числах уравнение

$$E_m : ax^2 + bxy + cy^2 = m \tag{4}$$

от целых переменных x, y, где a, b, c, m — какие-то целые числа (параметры). Для этого мы сопоставим каждой квадратичной форме от двух переменных карту и выразим свойства уравнения (4) через свойства этой карты. Мы надеемся, что с помощью этого подхода участники проекта смогут также решить следующие (супер)задачи.

Задача 46 (Суперзадача)**.** Докажите, что если уравнение E_m имеет решения при каком-то положительном числе m, при каком-то отрицательном числе m и не имеет решений при $m = 0$, то для всякого m или E_m не имеет решений, или же E_m имеет бесконечно много решений.

Задача 47 (Суперзадача)**.** Верно ли, что если уравнение E_m имеет решения в целых числах при

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

то E_m имеет решения при всяком целом числе m?

Задача 48 (Суперзадача)**.** Докажите, что если уравнения E_1, E_2, E_3, E_5 имеют решения в целых числах, то уравнение E_m имеет решения при каком-то $m < 0$.

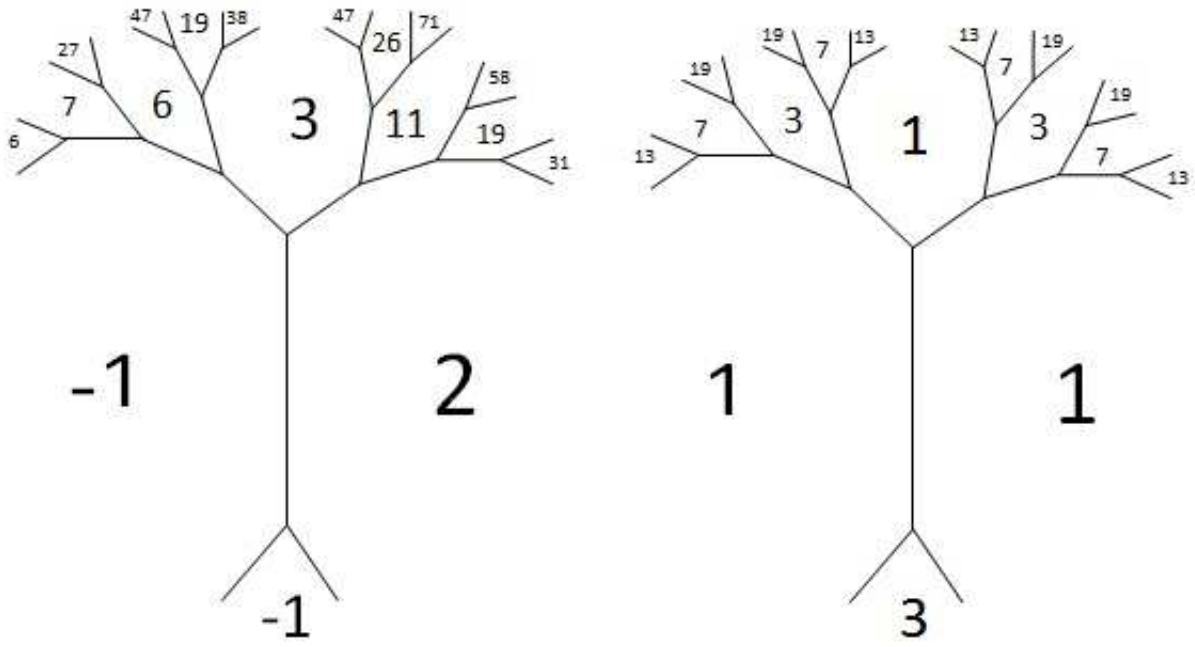
Начнём же мы с примера решения уравнения с помощью карты.

Пример работы с картами

Цель этого пункта — помахав руками, показать, что многочленам от двух переменных можно сопоставлять карты. К примеру, многочлены

$$2x^2 + 2xy - y^2 = 1 \text{ и } x^2 - xy + y^2 = 2$$

обладают картами



и из этих карт видно, что соответствующие уравнения не имеют решений в целых числах.

Рисуем картинку [2]

Для того, чтобы найти что-то общее среди множества чего-то очень разного, бывает полезно рассмотрение «всего и сразу» одновременно и снабжение этого «всего» какой-либо дополнительной структурой. В соответствии с этой стратегией для каждой квадратичной формы f мы рассмотрим все квадратичные формы, ей эквивалентные с точностью до так называемой линейной замены базиса, и снабдим их дополнительной структурой (мы разместим «точки» квадратичных форм на плоскости и соединим их отрезками). Для осуществления этого плана нам потребуются понятия базиса и супербазиса $\mathbb{Z}^2$.

Определение 7. *Базисом* $\mathbb{Z}^2$ называется такой набор $w_1, w_2 \in \mathbb{Z}^2$, что для любого $v \in \mathbb{Z}^2$ существуют числа $m, n \in \mathbb{Z}$, для которых

$$v = mw_1 + nw_2.$$

У нас было понятие эквивалентных форм. К сожалению, при работе с картами более правильным является следующее понятие.

Определение 8. Две формы f_1, f_2 называются *эквивалентными с точностью до линейной замены базиса*, если $\exists$ a, b, c, d, для которых $ad - bc = 1$ и

$$f_1(x, y) = f_2(ax + by, cx + dy).$$

Задачи 49–52.

Определение 9. *Супербазисом* $\mathbb{Z}^2$ называется набор $\{\pm w_1,\ \pm w_2,\ \pm(w_1+w_2)\}$, где $\{w_1, w_2\}$ — какой-то базис $\mathbb{Z}^2$. Назовём базис $\{w_1, w_2\}$ *специализацией* супербазиса $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$. Назовём супербазис $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$ *расширением* базиса $\{w_1, w_2\}$.

Упражнение 1. Выпишите все расширения данного базиса $\{w_1, w_2\}$. Выпишите все специализации данного супербазиса $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$.

Теперь мы готовы нарисовать карту квадратичной формы f. Начнём мы с той её части, что от f никак не зависит:

(1) каждому супербазису $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$ мы сопоставляем точку на плоскости (вершину будущего графа),

(2) каждому базису $\{w_1, w_2\}$ мы сопоставляем отрезок на плоскости (ребро графа), соединяющий точки-супербазисы

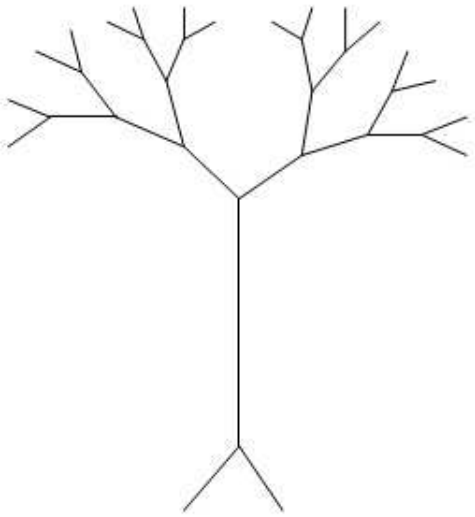
$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\} \text{ и } \{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\},$$

(базисам $\{w_1, w_2\}$, $\{-w_1, w_2\}$, $\{w_1, -w_2\}$ и $\{-w_1, -w_2\}$ соответствует одно и то же ребро),

(3) каждому набору $w \in \mathbb{Z}^2$ мы сопоставляем область плоскости, границей которого являются отрезки-базисы, набор w содержащие (наборам w и $-w$ соответствует одна и та же область).

Оказывается, что всю эту картинку можно нарисовать на плоскости.

[2] Те, кто хочет сразу увидеть, как нарисовать карту квадратичной формы, могут заглянуть в дополнение к этой части проекта. А потом подумать о том, как же доказываются свойства этой самой карты.



(5)

Заметим, что (5) пока никак не зависит от квадратичной формы f. Мы напишем в каждой области (5) (а потом и рядом с каждым отрезком (5)) числа, и по этому набору чисел класс эквивалентности с точностью до линейной замены формы f будет однозначно восстанавливаться. Мы используем следующие правила.

(1) В области, соответствующей набору $w \in \mathbb{Z}^2$, мы напишем число $f(w)$.

(2) Рядом с отрезком I, соответствующим базису $\{w_1, w_2\}$, мы напишем положительное число, по модулю равное

$$f(w_1 + w_2) - f(w_1) - f(w_2).$$

Если $f(w_1 + w_2) > f(w_1 - w_2)$, началом отрезка I мы считаем точку-супербазис

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\},$$

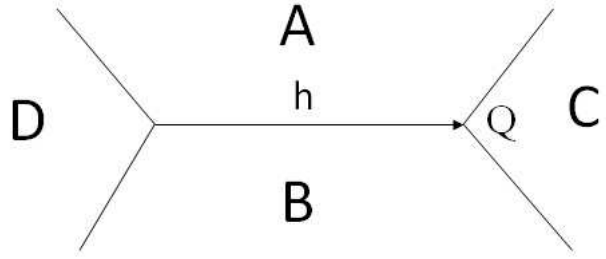
а концом считаем точку-супербазис $\{\pm w_1,\ \pm w_2,\ \pm(w_1+w_2)\}$. Если $f(w_1+w_2) < f(w_1-w_2)$ — то наоборот. Если $f(w_1 + w_2) = f(w_1 - w_2)$, то направление на I мы не задаём (и никакого числа рядом с I обычно не пишем).

Результат этой процедуры мы называем *ориентированной картой квадратичной формы* f. Если с числами около рёбер и их направлениями мы не заморачиваемся, то соответствующую картинку мы называем *картой квадратичной формы*. Для форм $2x^2 + 2xy - y^2$ и $x^2 - xy + y^2$ карты изображены на рисунках первой страницы проекта.

Упражнение 2. Нарисуйте (ориентированные) карты для квадратичных форм

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

В следующих двух задачах числа A, B, C, D, h относятся к картинке



Задачи 53–55.

Следующее определение является ключевым в решении большинства вопросов о положительно определённых квадратичных формах.

Определение 10. *Колодцем* называется вершина ориентированной карты квадратичной формы, все рёбра, инцидентные которой, направлены не к ней.

Задачи 56–60.

Нам бы хотелось обратить внимание на следующие несколько фактов:

1) идейно решения суперзадач 1, 2 и общего уравнения вида (4) близки к задачам 59, 60.

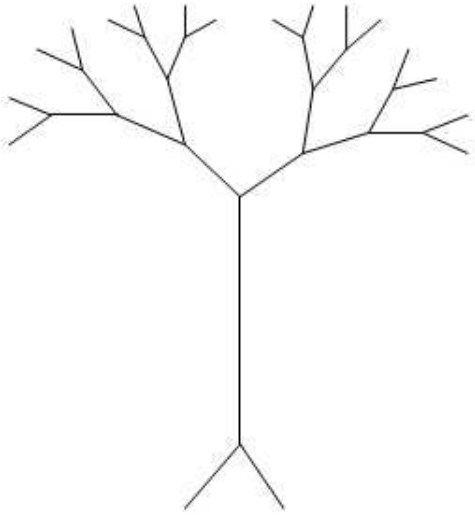
2) в тяжёлых трудовых математических буднях никто (кроме Вас самих) не будет писать для Вас ряд (достаточно простых) упражнений, который вёл бы к решению той или иной сколько-нибудь интересной задачи. Вам очень повезёт, если Вам (чаще всего, условно случайно) расскажут сколько-нибудь заметную часть нужных при решении задачи идей и приёмов.

3) цель данной конференции – познакомить Вас в той или иной форме с трудовыми математическими буднями.

Если вы ещё не догадались, к чему мы клоним, то на этом упражнения, помогающие Вам справиться с суперзадачами 1, 2, 3 и научиться решать общее уравнение вида (4), закончились. Пытаясь ещё немного облегчить Вашу участь, мы приготовили пару картинок, которые (может быть) могут Вам в (чём-нибудь) помочь или (что-нибудь) подсказать.

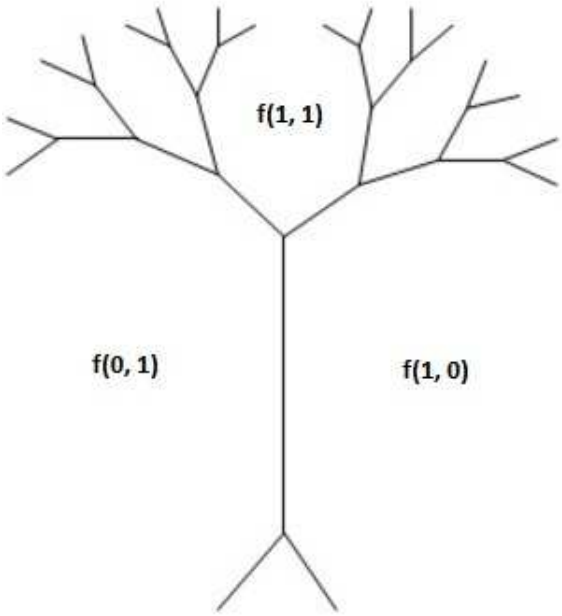
Рабоче-крестьянское описание карты квадратичной формы

Алгоритм $f \to \Gamma_f$: Рассмотрим бесконечное плоское троичное дерево (т.е. нарисованный на плоскости связный граф без циклов, степень каждой вершины которого равна 3). Часть такого графа идёт ниже.



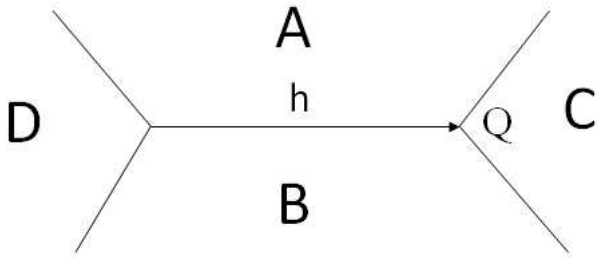
(6)

Рассмотрим какую-то вершину этого графа и впишем в три инцидентные ей грани числа $f(1,0)$, $f(0,1)$ и $f(1,1)$, как показано на рисунке.



(7)

Значения во всех остальных гранях определяются по правилу



. (8)

Правило 1: Если три значения на гранях около одного ребра (как показано на рисунке (8)) известны, то четвёртое определяется формулой $2(A+B)=C+D$.

Легко видеть, что это правило определяет карту Γ_f. Теперь по карте Γ_f мы построим ориентированную карту $\vec{\Gamma}_f$ по следующим правилам (см. рисунок (8)):

Правило 2: Над ребром h мы напишем число

$$|(A+B)-C| = |(A+B)-D|\,.$$

Правило 3: Если $C<D$, то мы заменяем ребро h на стрелку из C в D; если $C<D$, то мы заменяем ребро h на стрелку из D в C; если $C=D$, то мы оставляем ребро h неориентированным.

Свойства карты:

1) точкам графа Γ_f соответствуют квадратичные формы, эквивалентные форме f с точностью до линейной замены;

2) грани Γ_f находятся во взаимно однозначном соответствии с ненулевыми рациональными числами $\frac{m}{n}$;

3) внутри каждой грани написано число, равное значению $f(m,n)$ на соответствующей $\frac{m}{n}$ паре целых чисел (m,n);

4) квадратичные формы f и g эквивалентны с точностью до линейной замены, если и только карты Γ_f и Γ_g одинаковы.

Вводные задачи

Задача 1. Докажите, что уравнения a) $2x^2 + 2xy - y^2 = 1$, b) $x^2 - xy + y^2 = 2$ не имеют решений в целых числах.

Доказательство. a) Анализ остатков при делении на три показывает, что уравнение

$$2x^2 + 2xy - y^2 = 3x^2 - (y - x)^2 = 1$$

не имеет решений в целых числах.

b) Легко видеть, что если $|x| \geq 3$ или $|y| \geqslant 3$ уравнение $x^2 - xy + y^2 = \frac{1}{2}(x^2 + y^2 + (x - y)^2) = 2$ не имеет решений в целых числах. Перебор оставшихся 25 вариантов показывает, что решений в целых числах у этого уравнения нет.

b) Заметим, что $x^2 - xy + y^2 = (x - y/2)^2 + 3/4y^2 = 2$, то есть $y^2 \leqslant 8/3$, $|y| \leqslant 1$, аналогично для x. Однако хотя бы одно из x и y должно быть чётное, то есть нулевое, пусть $x = 0$, тогда $y^2 = 2$, противоречие. □

Задача 2. Докажите, что уравнения a) $x^2 - 2y^2 = 1$, b) $x^2 - 3y^2 = 1$, c) $x^2 - 6y^2 = 1$ имеют бесконечно много решений в целых числах.

Доказательство. a) Для всякого целого n наборы

$$x = \frac{(3 + 2\sqrt{2})^n + (3 - 2\sqrt{2})^n}{2}, \quad y = \frac{(3 + 2\sqrt{2})^n - (3 - 2\sqrt{2})^n}{2\sqrt{2}}$$

являются решениями уравнения.

b) Для всякого целого n наборы

$$x = \frac{(2 + \sqrt{3})^n + (2 - \sqrt{3})^n}{2}, \quad y = \frac{(2 + \sqrt{3})^n - (2 - \sqrt{3})^n}{2\sqrt{3}}$$

являются решениями уравнения.

c) Для всякого целого n наборы

$$x = \frac{(5 + 2\sqrt{6})^n + (5 - 2\sqrt{6})^n}{2}, \quad y = \frac{(5 + 2\sqrt{6})^n - (5 - 2\sqrt{6})^n}{2\sqrt{6}}$$

являются решениями уравнения. □

Задача 3. Докажите, что уравнение $x^2 + 1000xy + 1000y^2 = 2001$ имеет бесконечно много решений в целых числах.

Доказательство. Дискриминант $1000^2 - 4 \cdot 1000$ уравнения $x^2 + 1000xy + 1000y^2 = 2001$ больше 0 и не является полным квадратом. Следовательно, форма $x^2 + 1000xy + 1000y^2 = 2001$ неопределена, не представляет 0 и представляет 2001 при $x = y = 1$. По задаче 46 отсюда следует, что у уравнения

$$x^2 + 1000xy + 1000y^2 = 2001$$

бесконечно много решений. □

Задача 4. Фиксируем нечётное простое число p. Докажите, что уравнение $x^2 - py^2 = -1$ имеет решение в целых числах, если и только если p имеет остаток 1 при делении на 4.

Доказательство. Пусть уравнение $x^2 - py^2 = -1$ имеет целое решение. Докажем, что p имеет остаток 1 при делении на 4. Действительно, в этом случае -1 является квадратичным вычетом по модулю p, т.е. $p \equiv 1 \pmod 4$.

Наоборот, пусть $p \equiv 1 \pmod 4$. По суперзадаче 46 существует нетривиальное решение уравнения $x^2 - py^2 = 1$.

Пусть S_+ — множество решений (x_0, y_0) уравнения $x^2 - py^2 = 1$ таких, что $x_0, y_0 > 0$. Пусть (x_0, y_0) — это решение из S_+ с наименьшим значением y_0. Тогда

$$(x_0 - 1)(x_0 + 1) = py_0^2. \tag{1}$$

Из (1) следует, что или $2(x_0+1)$, или $2(x_0-1)$ является полным квадратом. Рассмотрим оба варианта.

Пусть $2(x_0+1) = d^2$ для какого-то целого положительного числа d. Тогда d является чётным числом и $d \mid y_0$. Положим $d = 2d_0$. Положим

$$x_1 = (x_0+1)/d = d_0,\ y_1 = y_0/d.$$

Тогда

$$x_1^2 - py_1^2 = \frac{1}{d^2}((x_0+1)^2 - py_0^2) = \frac{2(x_0+1)}{d^2} = 1.$$

Таким образом, (x_1, y_1) также принадлежит S_+. Очевидно, что $y_1 < y_0$, что противоречит предположению о минимальности пары (x_0, y_0).

Пусть $2(x_0-1) = d^2$ для какого-то целого положительного числа d. Тогда d является чётным числом и $d \mid y_0$. Положим $d = 2d_0$. Положим

$$x_1 = (x_0-1)/d = d_0,\ y_1 = y_0/d.$$

Тогда

$$x_1^2 - py_1^2 = \frac{1}{d^2}((x_0-1)^2 - py_0^2) = \frac{2(1-x_0)}{d^2} = -1.$$

Таким образом, (x_1, y_1) является решением уравнения $x^2 - py^2 = -1$. Его и требовалось найти. □

Задача 5. Докажите, что для всякого m количества решений в целых числах уравнений

$$x^2 - xy + y^2 = m \quad \text{и} \quad 3x^2 + 9xy + 7y^2 = m$$

одинаковы.

Доказательство. Мы покажем, что всякому целочисленному решению уравнения $x^2 - xy + y^2 = m$ соответствует целочисленное решение уравнения $3x^2 + 9xy + 7y^2 = m$. И наоборот.

Пусть $v = x + y$, $u = -x - 2y$, очевидно, эти выражения целые. Подставим в $x^2 - xy + y^2$ вместо x выражение $u + 2v$, вместо y выражение $-u - v$. Получим $3u^2 + 9uv + 7v^2$. Таким образом, всякому решению уравнения $3x^2 + 9xy + 7y^2 = m$ соответствует решение уравнения $x^2 - xy + y^2 = m$.

Обратно, подставим в $3x^2 + 9xy + 7y^2$ вместо x выражение $u + 2v$, вместо y выражение $-u - v$. Получим $u^2 - uv + v^2$. Таким образом, всякому решению уравнения $x^2 - xy + y^2 = m$ соответствует решение уравнения $3x^2 + 9xy + 7y^2 = m$. □

Задача 6. Докажите, что для всякого целого числа n уравнение $x^2 + y^2 = n$ имеет решение в целых числах, если и только если оно имеет решения в рациональных числах.

Доказательство. Пусть рациональные числа x, y таковы, что $x^2 + y^2 = n$. Знаменатель чисел x и y в приведённой форме один и тот же. Мы обозначим его d и допустим, что мы выбрали x и y так, чтобы d было наименьшим возможным. Предположим, что $d > 1$ (т.е. что x, y нецелы и у уравнения $x^2+y^2 = n$ нет целых решений). Пусть r_x, r_y – это ближайшие к x, y целые числа соответственно, $s_x := x - r_x$, $s_y := y - r_y$. Тогда

$$|s_x|,\ |s_y| \leqslant \frac{1}{2}, \quad s_x^2 + s_y^2 = n - (r_x^2 + r_y^2) - 2(s_x r_x + s_y r_y). \tag{2}$$

Положим

$$x' = r_x - \frac{s_x(n - r_x^2 - r_y^2)}{s_x^2 + s_y^2}, \qquad y' = r_y - \frac{s_y(n - r_x^2 - r_y^2)}{s_x^2 + s_y^2}.$$

Из (2) следует, что $s_x^2 + s_y^2 = d'/d$, причём $0 < d' < d$. Отсюда следует, что в приведённой форме знаменатель d' дробей x', y' делит d, в частности, меньше d. Имеем $x'^2 + y'^2 = n$. Следовательно, знаменатель дробей x', y' меньше знаменателя дробей x, y. Противоречие. Следовательно, $d = 1$, т.е. уравнение $x^2 + y^2 = n$ имеет решения в целых числах. □

Задача 7. Приведите пример квадратичного уравнения с целыми коэффициентами, имеющего решения в рациональных числах, но не имеющего решений в целых числах.

Доказательство. Рассмотрим уравнение $4x^2 = 1$. Значение $\frac{1}{2}$ даёт рациональное решение этого уравнения, а вот целых решений у этого уравнения нет. □

Задача 8. Докажите, что для любых целых положительных чисел a и b существует бесконечно много натуральных чисел m, для которых уравнение $ax^2 + by^2 = m$ не имеет решений в целых числах.

Доказательство. Пусть N — какое-то целое число. Если для какого-то $n \leqslant N$ уравнение $ax^2 + by^2 = n$ имеет решение, то

$$|x| \leqslant \sqrt{\frac{N}{a}}, \quad |y| \leqslant \sqrt{\frac{N}{b}}.$$

Таким образом, если для всякого $n \leqslant N$ уравнение $ax^2 + by^2 = n$ имеет решение, то существует N пар чисел (x, y) таких, что $0 \leqslant x \leqslant \sqrt{\frac{N}{a}}$, $0 \leqslant y \leqslant \sqrt{\frac{N}{b}}$. Тогда

$$N \leqslant \frac{N}{\sqrt{ab}}.$$

Очевидно, что это неравенство не выполнено для достаточно больших N, если $ab > 1$. С учётом этого осталось рассмотреть случай $a = b = 1$. Ни одно число, имеющее остаток 3 при делении на 4, не представляется в виде $x^2 + y^2$, что завершает доказательство. □

Задача 9. Докажите, что для всякого целого числа m уравнение $x^2 + 2y^2 - 3z^2 = m$ имеет решение в целых числах.

Доказательство. Достаточно показать, что $x^2 + 2y^2 - 3z^2$ представляет 0, все нечётные числа и все числа, имеющие остаток 2 при делении на 4.

Подставим $x = y = z = 1$ в $x^2 + 2y^2 - 3z^2$. Получим 0. Следовательно, $x^2 + 2y^2 - 3z^2$ представляет 0.

Подставим $x = u + 1$, $y = u$, $z = u$ в $x^2 + 2y^2 - 3z^2$. Получим $2u + 1$. Следовательно, $x^2 + 2y^2 - 3z^2$ представляет все нечётные числа.

Подставим $x = u$, $y = u + 1$, $z = u$ в $x^2 + 2y^2 - 3z^2$. Получим $4u + 2$. Следовательно, $x^2 + 2y^2 - 3z^2$ представляет все числа, имеющие остаток 2 при делении на 4.

Если m делится на 4, то поделим все переменные на 2 и сведём задачу к одному из разобранных случаев. □

Квадратичные формы

Задача 10. Опишите все целые числа, которые представляются формами a) $x^2 + y^2$; b) $x^2 - y^2$; c) $x^2 + xy + y^2$.

Доказательство. a) $n = x^2 + y^2$ если и только если в разложении n на простые множители все простые делители, входящие в n в нечётной степени, имеют остаток 1 при делении на 4.

b) $(u+1)^2 - u^2 = 2u + 1$. Следовательно, все нечётные числа представляются формой $x^2 - y^2$. Также $(u + 1)^2 - (u - 1)^2 = 4u$.

Как итог, все целые числа, имеющие остатки 0, 1, 3 при делении на 4 представляются формой $x^2 - y^2$. Анализ остатков при делнии на 4 показывает, что числа с остатками 2 при делении на 4 не представимы в виде $x^2 - y^2$.

c) Фиксируем число n. Рассуждение, аналогичное рассуждению, приведённому в задаче 6, показывает, что $x^2 + xy + y^2 = n$ имеет решение в рациональных числах если и только если оно имеет решение в числах целых. В рациональных числах форма $x^2 + xy + y^2$ эквивалентна форме $x^2 + 3y^2$ ($x^2 + xy + y^2 = (x + \frac{y}{2})^2 + 3(\frac{y}{2})^2$). Покажем, что форма $x^2 + 3y^2$ представляет число n в рациональных числах (см. часть про символ Гильберта), если и только если в разложении числа n на простые множители все простые делители, входящие в нечётной степени, имеют остаток 1 или 0 при делении на 3.

Действительно, $x^2 + 3y^2 = n$ имеет решение в $\mathbb{Q}$ тогда и только тогда, когда $x_1^2 + 3y_1^2 - nz^2 = 0$ имеет решение в $\mathbb{Z}$ с ненулевым z. Это уравнение (по теореме Минковского-Хассе) имеет решение тогда и только тогда, когда символ Гильберта $(n, -3)_p$ равен 1 для всех простых p. Вычислим его непосредственно.

Пусть $p > 3$. Запишем $n = p^\alpha \cdot u$, $3 = p^0 \cdot (-3)$. При помощи явной формулы для символа Гильберта и квадратичного закона взаимности Гаусса (Серр, Глава 1, § 3 Теорема 6) получаем, что

$$(n, -3)_p = \left(\frac{-3}{p}\right)^\alpha = \left(\frac{-1}{p}\right)^\alpha \left(\frac{3}{p}\right)^\alpha = \left(\frac{-1}{p}\right)^\alpha \cdot \left((-1)^{\frac{p-1}{2}} \left(\frac{p}{3}\right)\right)^\alpha,$$

то есть данное выражение равно 1 всегда, если α чётно, а при нечётном α оно равно $\left(\frac{p}{3}\right)$, то есть равно 1 тогда и только тогда, когда p имеет остаток 1 при делении на 3. Получили, что если уравнение

$$x_1^2 + 3y_1^2 - nz^2 = 0$$

имеет решения по модулю p, при этом p имеет вид $3k + 2$, то в разложение числа n сомножитель p входит в чётной степени.

Случай $p = 2$ оставим читателю в качестве упражнения.

Случай $p = 3$. Предположим, что $n = 3^\alpha \cdot u$, при этом $\beta = 1$, $v = -1$. Имеем:

$$(n, -3)_3 = (-1)^\alpha \left(\frac{u}{3}\right) \left(\frac{-1}{3}\right)^\alpha = \left(\frac{u}{3}\right).$$

Это выражение равно 1 тогда и только тогда, когда u имеет остаток 1 при делении на 3. Но если мы уже знаем, что все простые делители вида $3k + 2$ входят в чётной степени, то это условие не даёт ничего нового. □

Определение 1. Назовём две формы *эквивалентными*, если они представляют один и тот же набор целых чисел.

Задача 11. Докажите, что квадратичные формы

$$f(x, y), \quad f(x - y, y), \quad f(x, y - x), \quad f(-x, y) \quad \text{и} \quad f(x, -y) \tag{3}$$

попарно эквивалентны.

Доказательство. Если число m представляется формой $f(x, y)$ при $x = x_0$, $y = y_0$, то m представляется формой $f(x-y, y)$ при $x = x_0+y_0$, $y = y_0$, формой $f(x, y-x)$ при $x = x_0$, $y = y_0+x_0$, формой $f(-x, y)$ при $x = -x_0$, $y = y_0$, формой $f(x, -y)$ при $x = x_0$, $y = -y_0$. Таким образом, все числа представляемые формой $f(x, y)$ представляются и другими формами списка (3). Аналогично доказывается и то, что все числа, представляемые любой другой формой списка (3), представляются и всеми остальными формами списка (3). Таким образом, все формы списка (3) эквивалентны. □

Задача 12. a) Докажите, что формы $x^2 + y^2$ и $x^2 + xy + y^2$ не эквивалентны.

b) Докажите, что форма $4x^2 - 6xy + 5y^2$ не эквивалентна форме $ax^2 + by^2$ ни для каких целых чисел a и b.

Доказательство. a) Форма $x^2 + y^2$ представляет 2, а $x^2 + xy + y^2$ — нет. Следовательно, они не эквивалентны.

b) Значения вокруг единственного колодца формы $4x^2 - 6xy + 5y^2$ равны 3, 4, 5. Следовательно, числа 3, 4, 5 являются тремя наименьшими значениями формы $4x^2 - 6xy + 5y^2$.

Пусть $a, b \geqslant 0$. Тогда тремя наименьшими значениями формы $ax^2 + by^2$ могут быть наборы чисел

$$\{a,\ b,\ a + b\}, \quad \{a,\ 2a,\ b\}, \quad \{a,\ b,\ 2b\}, \quad \{a,\ 2a,\ 4a\}, \quad \{b,\ 2b,\ 4b\}. \tag{4}$$

Очевидно, что набор 3, 4, 5 не является ни одним из (4) ни для каких a и b.

Следовательно, формы $4x^2 - 6xy + 5y^2$ и $ax^2 + by^2$ не эквивалентны ни при каких неотрицательных числах a, b. □

Задача 13. Приведите пример неотрицательно определённой формы, которая не является положительно определённой.

Доказательство. Пример: $f(x, y) = x^2$. □

Расширенная арифметика

Задача 14. Пусть m и n – целые числа, свободные от квадратов. Если уравнение

$$z^2 - mx^2 - ny^2 = 0 \tag{5}$$

имеет ненулевое решение в рациональных числах, то выполнены следующие условия

a) хотя бы одно из чисел m, n положительно,

b) m является квадратичным вычетом при делении на n,

c) n является квадратичным вычетом при делении на m.

Доказательство. Фиксируем ненулевое рациональное решение (x_0, y_0, z_0) уравнения (5). Мы можем считать (и считаем), что числа x_0, y_0, z_0 взаимно просты в совокупности.

a) Если $m, n \leqslant 0$, то $x_0^2 - my_0^2 - nz_0^2 \geqslant 0$, и равенство достигается только при $x_0 = y_0 = z_0 = 0$. Противоречие.

b) Достаточно доказать, что для всякого простого делителя p числа m число n является полным квадратом в остатках при делении на p.

Пусть p — это какой-то простой делитель m. Если $n \vdots p$, то утверждение задачи, очевидно, выполнено. Пусть $n \not\vdots p$. Рассмотрим два случая: $y_0 \vdots p$ и $y_0 \not\vdots p$.

Допустим, что $y_0 \vdots p$. Тогда x_0, z_0 также делятся на p, что противоречит тому, что

$$\text{НОД}(x_0, y_0, z_0) = 1.$$

Следовательно, $y_0 \not\vdots p$. Тогда в остатках при делении на p выполнено равенство

$$n \equiv (z/y)^2 \pmod{p},$$

что завершает доказательство пункта b).

Пункт c) доказывается аналогично пункту b). □

Задача 15. Сведите метатеорему для двух переменных к решению уравнений вида (5).

Доказательство. Всякое квадратное уравнение имеет вид

$$f(X_1, X_2) = f_2(X_1, X_2) + f_1(X_1, X_2) + f_0 = 0,$$

где f_2 — это однородный многочлен степени 2, f_1 — степени 1, f_0 — степени 0 (т.е. константа).

Начнём с некоторого общего утверждения: уравнения

$$f(X_1, X_2) = 0 \text{ и } f(X_1 + cX_2 + t, X_2) = 0$$

либо одновременно имеют решения, либо одновременно не имеют решений в рациональных числах для всяких рациональных чисел c, t. Мы оставляем это утверждение в качестве простого упражнения.

Очевидно, что замены вида

$$f(X_1, X_2) \to f(X_1 + cX_2, X_2) \tag{6}$$

действуют независимо на компоненты f_1, f_2 и сохраняют f_0.

Представим f_2 в виде

$$c_1X_1^2 + c_{12}X_1X_2 + c_2X_2^2,$$

где c_1, c_2, c_{12} — это параметры.

Если $f_2 \neq 0$, то, сделав несколько замен вида (6), мы можем считать, что $c_1 \neq 0$.

Рассмотрим функцию

$$f(X_1 - \frac{c_{12}}{2c_1}X_2, X_2). \tag{7}$$

Легко видеть, что (7) имеет вид

$$c_1X_1^2 + c_2'X_2^2$$

для некоторого рационального числа c_2'. Итак, мы можем считать

$$f_2(X_1, X_2) = c_1 X_1^2 + c_2 X_2^2$$

для каких-то рациональных чисел c_1, c_2. Если $c_2 = 0$, а $c_1 \neq 0$, то уравнение $f = 0$ принимает вид

$$c_1 X_1^2 = -rX_2 - f_0$$

и решается элементарно. Таким образом, далее мы считаем, что $c_1 \neq 0$. По аналогичным соображениям мы считаем, что $c_2 \neq 0$.

Линейная функция $f_1(X_1, X_2)$ представима в виде $r_1 X_1 + r_2 X_2$. Рассмотрим замену

$$f(X_1, X_2) \to f\left(X_1 - \frac{r_1}{2c_1}, X_2 - \frac{r_2}{2c_2}\right).$$

У функции $f(X_1 - \frac{r_1}{2c_1}, X_2 - \frac{r_2}{2c_2})$ часть f_1 равна 0. В этом случае уравнение $f = 0$ принимает вид

$$c_1 X_1^2 + c_2 X_2^2 + f_0 = 0.$$

Это уравнение эквивалентно однородному уравнению

$$z^2 + \frac{c_2}{c_1} y^2 + \frac{f_0}{c_1} z^2 = 0.$$

Что и требовалось доказать. □

Задача 16. Уравнение с целыми коэффициентами $f = 0$ имеет решение в $\mathbb{Z}_p$, если и только если оно имеет решение в остатках при делении на p^n для всякого $n \in \mathbb{Z}_{\geqslant 0}$.

Доказательство. Пусть $x_1, \ldots, x_n$ — это решение в $\mathbb{Z}_p$ уравнения $f = 0$. Тогда классы $x_1, \ldots, x_n$ в остатках при делении на p^n являются решениями в остатках при делении на p^n. В частности, $f \equiv 0$ имеет решение в остатках при делении на p^n для всякого целого положительного числа n.

Наоборот, положим, что уравнение $f \equiv 0$ имеет решение в остатках при делении на p^m для любого целого положительного числа m. Обозначим через S_m для каждого m множество решений в остатках при делении на p^m уравнения $f \equiv 0$. По предположению это множество остатков не пусто для всякого $m \geqslant 0$.

Так как любой остаток при делении на p^{m+1} можно рассмотреть по модулю p^m, то имеется отображение $S_{m+1} \to S_m$. Обозначим через S_m^∞ пересечение образов S_{m+k} для всех $k \geqslant 0$. Так как $S_{m+k} \neq 0$, то $S_m^\infty \neq 0$. Для каждого $s_m \in S_m^\infty$ существует $s_{m+1} \in S_{m+1}^\infty$ такой, что s_m есть образ s_{m+1}. Таким образом, можно построить бесконечную цепочку

$$s_1, \ldots, s_m, \ldots, \tag{8}$$

где s_m есть набор из n остатков при делении на p^m и s_m есть проекция s_{m+1} в остатки при делении на p^m. Последовательность (8) задаёт единственный набор из n целых p-адических чисел $x_1, \ldots, x_n$, обладающих заданными наборами остатков $s_1, \ldots, s_m, \ldots$ при делении на $p, \ldots, p^m, \ldots$.

Числа $x_1, \ldots, x_n$ являются решениями уравнения $f = 0$. □

Задача 17. Когда p-адическое число в форме (2) равно 0?

Доказательство. Ответ, следующий из определения: «Когда $a_{-k} + \ldots + a_{-k+i} p^i \equiv 0 (\mathrm{mod}\ p^{i+1})\ \ \forall i$». □

Задача 18. Докажите, что произведение двух ненулевых p-адических чисел не равно 0.

Доказательство. Рассмотрим два ненулевых p-адических числа a, b. Без ограничения общности мы считаем, что $a, b \in \mathbb{Z}_p$ и $a, b \not\equiv 0 (\mathrm{mod}\ p)$. Тогда $ab \not\equiv 0 (\mathrm{mod}\ p)$ и, следовательно, $ab \neq 0$. □

Задача 19. Докажите, что $\mathbb{Q} \subset \mathbb{Q}_p$ для всякого простого числа p (докажите, что для всякой пары ненулевых целых чисел m, n существует p-адическое число x такое, что $nx = m$).

Доказательство. Без ограничения общности считаем, что m, n взаимно просты с p. В этом предположении задача 16 следует из задачи 20. □

Задача 20. Докажите, что -1 является полным квадратом в p-адических числах тогда и только тогда, когда p имеет остаток 1 при делении на 4.

Доказательство. Следует из задачи 21. □

Задача 21. Придумайте описание для p-адических чисел, являющихся полными квадратами.

Доказательство. Мы рассмотрим два случая: $p = 2$ и $p \neq 2$.

Пусть $p = 2$. Всякое 2-адическое число x представимо в виде $2^n(2m + 1)$, где n — это целое число, а m — это целое 2-адическое число. Имеем $x^2 = 2^{2n}(1 + 8\frac{m(m+1)}{2})$. Положим $m' = \frac{m(m+1)}{2}$. Тогда

$$x^2 = 2^{2n}(1 + 8m'), \tag{9}$$

где m' — это целое 2-адическое число.

Докажем, что всякое 2-адическое число вида (9) является полным квадратом в 2-адических числах. Для этого достаточно показать, что всякое целое 2-адическое число m' представимо в виде $\frac{m(m+1)}{2}$.

Благодаря задаче 16, достаточно доказать, что сравнение $x(x + 1) \equiv 2m'$ имеет решения в $\mathbb{Z}$ для всякого $i \in \mathbb{Z}_{\geqslant 0}$. Мы докажем это утверждение по индукции.

База: $i = 1$, очевидно, выполнена.

Переход: $i \to i+1$. Пусть $m_i \in \mathbb{Z}$ — это решение уравнения $x(x+1) \equiv 2m' (\mathrm{mod}\ 2^i)$. Выполнено одно из двух условий

1) $m_i(m_i + 1) \equiv 2m' \pmod{2^{i+1}}$,

2) $m_i(m_i + 1) \equiv 2m' + 2^i \pmod{2^{i+1}}$.

В случае 1) m_i является решением сравнения $x(x+1) \equiv 2m' \pmod{2^{i+1}}$. В случае 2) m_i+2^i является решением сравнения $x(x + 1) \equiv 2m' \pmod{2^{i+1}}$.

Пусть $p \neq 2$, т.е. p — нечётное простое число. Всякое p-адическое число x представимо в виде $p^n m$, где n — это целое число, а m — это целое p-адическое число, не делящееся на p. Имеем $x^2 = p^{2n}m^2$. Положим $m' = m^2$. Тогда

$$x^2 = p^{2n}m', \tag{10}$$

где m' — это целое p-адическое число, остаток при делении на p которого является квадратичным вычетом и не равен 0.

Докажем, что всякое p-адическое число вида (9) является полным квадратом в p-адических числах. Для этого достаточно показать, что всякое целое p-адическое число m' такое, что

3) m' не делится на p,

4) остаток при делении m' на p является квадратичным вычетом и не равен 0,

представимо в виде m^2.

Благодаря задаче 16 достаточно доказать, что уравнение $x^2 \equiv m' \pmod{p^i}$ имеет решения в $\mathbb{Z}$ для всякого $i \in \mathbb{Z}_{\geqslant 0}$. Мы докажем это утверждение по индукции.

База: $i = 1$ выполнена, так как остаток при делении на p числа m' является квадратичным вычетом.

Переход: $i \to i + 1$. Пусть $m_i \in \mathbb{Z}$ — это решение уравнения $x^2 \equiv m' \pmod{p^i}$. Тогда

$$m_{i+1}^2 \equiv m' + rp^i \pmod{p^{i+1}},$$

где $r \in \mathbb{Z}$ — это какое-то число. Так как m' не делится на p и $p \neq 2$, то существует $r' \in \mathbb{Z}$ такое, что $2m_i r' \equiv r \pmod{p^{i+1}}$. Положим $m_{i+1} := m_i - r'p^i$. Тогда $m_{i+1}^2 \equiv m' \pmod{p^{i+1}}$. Что завершает переход. □

Задача 22. Докажите, что любое ненулевое 3-адическое число m есть или x^2, или $2x^2$, или $3x^2$, или $6x^2$ для какого-то 3-адического числа x.

Доказательство. Заметим, что для любого p произвольное p-адическое число представляется в виде $p^i \cdot a \cdot y$, где a — целое число от 1 до $p - 1$, p^i — степень p, а y — целое p-адическое число, сравнимое с единицей по модулю p (по задаче 21 оно автоматически является полным квадратом). Тогда, подставляя $p = 3$, получаем, что в зависимости от чётности числа i число p^i — или квадрат, или утроенный квадрат, $a \in \{1,\ 2\}$, y — точный квадрат. Поэтому их произведение является одним из предложенных вариантов. □

Задача 23. Пусть p — нечётное простое число, а $x_1, \ldots, x_5$ — ненулевые p-адические числа. Докажите, что x_i/x_j есть полный квадрат в p-адических числах для каких-то i, j $(1 \leqslant i < j \leqslant 5)$.

Доказательство. Аналогично решению предыдущей задачи любое число представляется в виде $p^i \cdot a \cdot y$. Разобьём числа на 2 группы. В первой группе i четно, во второй — нечетно. После чего разобьём каждую на 2 подгруппы: в первой подгруппе a является квадратичным вычетом, а во второй — нет. Заметим, что по принципу Дирихле хотя бы 2 из взятых нами чисел попадет в одну группу. Их отношение имеет вид:

$$p^j \cdot a_{new} \cdot \frac{y_1}{y_2},$$

где j делится на 2, a_{new} является квадратичным вычетом (как отношение двух вычетов либо двух невычетов), а $\frac{y_1}{y_2}$ является p-адическим числом, начинающимся с единицы. Теперь очевидно, что это число является квадратом, как произведение трех квадратов. □

Задача 24. Докажите, что для всякого нечётного простого числа p существуют ненулевые p-адические числа $x_1, \ldots, x_{p-1}$ такие, что $x_1^2 + \ldots + x_{p-1}^2 + 1 = 0$.

Доказательство. Заметим, что по задаче 21 число $1 - p$ является полным квадратом в p-адических чисел. Следовательно, $-1 = 1 + 1 \ldots + 1$ ($p - 2$ единицы)$+1 - p$ есть сумма $(p - 1)$-го квадрата в p-адических числах. Что и требовалось доказать. □

Задача 25. Докажите, что уравнение $x^2 + x + 1 = 0$ имеет ровно два решения в целых 7-адических числах.

Доказательство. Решения уравнения $x^2 + x + 1 = 0$ выражается формулой $x_{1,2} = \frac{-1 \pm \sqrt{-3}}{2}$. Так как $\sqrt{-3} \in \mathbb{Z}_7$ (см. задачу 21), то $x^2 + x + 1 = 0$ имеет два различных решения в $\mathbb{Z}_7$. □

Задача 26. Докажите, что уравнение $x^2 + y^2 = -1$ имеет решения в p-адических числах для всякого нечётного простого числа p.

Доказательство. Достаточно доказать, что это уравнение имеет решение по модулю p. Проверим это. Выражение x^2 может принимать $(p + 1)/2$ различных значений: ноль и все квадратичные вычеты. Перечислим все значения выражения $-x^2 - 1$. Если хотя бы одно из них может быть представлено в виде y^2, то задач решена. Если же ни одно из них не имеет вид y^2, то для возможных значений y^2 (в остатках по модулю p) имеется максимум $(p - 1)/2$ возможных значений. Противоречие.

Второе решение. Достаточно доказать, что уравнение $x^2 + y^2 = -1$ имеет ненулевое решение в $\mathbb{Z}/p$. Всякий квадратичный вычет представим в виде $x^2 + y^2$. Если в виде $x^2 + y^2$ представимы только квадратичные вычеты, то в виде $x_1^2 + \ldots + x_{p-1}^2$ представимы только квадратичные вычеты. Но в виде $x_1^2 + \ldots + x_{p-1}^2$ представимы все остатки при делении на p. Следовательно, $x^2 + y^2$ представляет квадратичные невычеты. Следовательно, $x^2 + y^2$ представляет все элементы $\mathbb{Z}/p$. В том числе, -1. □

Задача 27. Докажите принцип Минковского-Хассе для уравнений от одной и двух переменных.

Доказательство. А) Уравнения от одной переменной. Уравнение имеет вид $ax^2 = b$. Достаточно доказать что, если уравнение не имеет решений в $\mathbb{Q}$, то оно не имеет решений или в $\mathbb{R}$ или в $\mathbb{Q}_p$ для какого-то p. Если $ax^2 = b$ не имеет решений в $\mathbb{Q}$, то b/a не является полным квадратом в $\mathbb{Q}$, т.е. или $b/a < 0$, или какой-то простой делитель p входит в b/a нечётное число раз. В первом случае $ax^2 = b$ не имеет решений в $\mathbb{R}$, а во втором — в $\mathbb{Q}_p$.

В) Уравнения от двух переменных. По задаче 15 всякое уравнение в рациональных числах от двух переменных эквивалентно уравнению $ax^2 + by^2 = 1$. При этом можно считать, что

1) числа a, b целые и свободны от квадратов,

2) $|a| \leqslant |b|$.

Достаточно доказать, что если уравнение $ax^2 + by^2 = 1$ имеет решение в $\mathbb{Q}_p$ для всякого p и в $\mathbb{R}$, то оно имеет решение в $\mathbb{Q}$. Положим $m(a, b) := |a| + |b|$. Будем доказывать наше утверждение индукцией по $m(a, b)$.

База: $m(a, b) = 2$ проверяется непосредственно.

Переход: $m \to m + 1$. Пусть a, b — это какие-то числа, удовлетворяющие условиям 1), 2) и такие, что

• $m(a, b) = m + 1$

• уравнение $ax^2 + by^2 = 1$ имеет решения в p-адических числах для всякого p и имеет решение в $\mathbb{R}$.

Рассмотрим два случая: $|a| = |b|$ и $|a| < |b|$. Если $|a| = |b|$, то уравнение $ax^2 + by^2 = 1$ эквивалентно уравнению

$$-(b/a)y^2 + az^2 = 1. \tag{11}$$

Более того, уравнение (11) имеет решения в $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}_p$, если и только если уравнение $ax^2 + by^2 = 1$ имеет решение в том же множестве чисел. Так как $m(-\frac{b}{a}, a) < m(a, b) = m + 1$, то $-(b/a)y^2 + az^2 = 1$ имеет решения в $\mathbb{Q}$ по предположению индукции. Следовательно, и исходное уравнение имеет решение в $\mathbb{Q}$.

Пусть теперь $|a| < |b|$. Из условия • следует, что a является полным квадратом в остатках при делении на b, т.е.

$$a + bb' = t^2,$$

где b', t — это какие-то целые числа и $b' \geqslant 0$. Без ограничения общности считаем, что

$$|t| \leqslant \frac{|b|}{2}.$$

Уравнение $ax^2 + by^2 = 1$ имеет решения в $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{Q}_p$, если и только если в тех же множествах чисел имеет решение уравнение

$$ax^2 + b'y^2 = 1, \qquad b' = \frac{t^2 - a}{b}.$$

Имеем $|b'| \leqslant \frac{|b|}{4}$ и, следовательно $m(a, b') < m(a, b) = m + 1$. Так $m(a, b') \leqslant m$, то $ax^2 + b'y^2 = 1$ имеет решения в $\mathbb{Q}$ по предположению индукции. Следовательно, и исходное уравнение имеет решение в $\mathbb{Q}$. Доказательство завершено. □

Задача 28. Докажите, что для символа Гильберта выполнены следующие соотношения

1) $(a, b)_p = (b, a)_p$,
2) $(a, c^2)_p = 1$,
3) $(a, -a)_p = 1$, $\quad (a, 1 - a)_p = 1$,
4) $(a, b)_p = (a, -ab)_p = (a, (1 - a)b)_p$.

Доказательство. 1) очевидно из определения.

2) Уравнение $z^2 - ax^2 - c^2y^2 = 0$ имеет ненулевое решение $z = c, x = 0, y = 1$.

3) Уравнение $z^2 - ax^2 + ay^2 = 0$ имеет ненулевое решение $z = 0, x = y = 1$.

Уравнение $z^2 - ax^2 - (1 - a)y^2 = 0$ имеет ненулевое решение $x = y = z = 1$.

4) Следует из пункта 3) и задачи 29. □

Задача 29. Пусть $(a, b)_p = 1$. Тогда $(a', b)_p = (aa', b)_p$ для любого a'.

Доказательство. Пусть b является полным квадратом. Тогда $(a, b)_p = (aa', b)_p = 1$.

Пусть теперь b не является полным квадратом. Мы воспользуемся следующей леммой.

Лемма 1. Если

• b не является полным квадратом,

• $(a, b)_p = 1$ и $(a', b)_p = 1$,

то $(aa', b) = 1$.

Завершим, используя лемму 1, решение задачи 29. Если $(a', b)_p = 1$, то $(aa', b)_p = 1$по лемме 1. Если $(aa', b)_p = 1$, то $(a', b)_p = (a^2a', b)_p = 1$ по лемме 1. Таким образом, если одно из чисел $(aa', b)_p, (a', b)_p$ равно 1, то и второе равно 1. Следовательно, эти числа равны.

Доказательство леммы 1. Пусть x_0, y_0, z_0 — это ненулевое решение уравнения $z_0^2 - ax_0^2 - by_0^2 = 0$. Так как b не является полным квадратом в $\mathbb{Q}_p$, то $x_0 \neq 0$. Тогда мы можем считать, что $x_0 = 1$ и $a = z_0^2 - by_0^2$. По аналогичным соображениям существуют z_1, y_1 такие, что $a' = z_1^2 - by_1^2$. Тогда

$$aa' = (z_0z_1 - by_0y_1)^2 - b(z_0y_1 + z_1y_0)^2.$$

Следовательно, $(aa', b) = 1$. □

□

Чтобы компактно записать явную формулу для символа Гильберта, нам потребуется *символ Лежандра* $\left(\frac{x}{p}\right)$, определённый для любых целого x и простого p. Он равен 1, -1 или 0 в зависимости от того, является x ненулевым квадратичным вычетом, невычетом или нулём по модулю p.

Задача 30. Пусть p – нечётное простое число, $a = p^\alpha u$, $b = p^\beta v$, где α, β, u, v – это целые числа такие, что u и v взаимно просты с p. Докажите, что

$$(a,b)_p = (-1)^{\alpha\beta\varepsilon(p)} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha,$$

где $\varepsilon(p) := \frac{p-1}{2}$.

Доказательство. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 3, § 1, Теорема 1. □

Задача 31. Найдите явную формулу для $(a,b)_2$ при всех целых числах a, b.

Доказательство. Пусть $a = 2^\alpha u$, $b = 2^\beta v$, где α, β, u, v — это целые числа такие, что u и v нечетны. Символ Гильберта $(a,b)_2$ задаётся формулой

$$(-1)^{\varepsilon(u)\varepsilon(v)+\alpha\omega(v)+\beta\omega(u)},$$

где $\varepsilon(u) = \frac{u-1}{2}$, а $\omega(u) = \frac{u^2-1}{8}$. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 3, § 1, Теорема 1. □

Задача 32. Докажите, что $(a,b)_p(a,b')_p = (a,bb')_p$ для любых целых чисел a, b, b'.

Доказательство. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 3, § 1, Теорема 2. □

Задача 33. Докажите, что уравнение $ax^2 + by^2 = c$ (a, b, c – это параметры; x, y – это переменные) имеет решение в p-адических числах, если и только если $(c, -ab)_p = (a,b)_p$.

Доказательство. Если уравнение $ax^2 + by^2 = c$ имеет решение, то уравнение

$$z^2 - \frac{a}{c}x^2 - \frac{b}{c}y^2 = 0$$

также имеет решение. Что по определению значит, что $(a/c, b/c)_p = 1$. Преобразуя, имеем

$$1 = (a/c, b/c)_p = (a,b)_p(a,c)_p(b,c)_p(c,c)_p = (a,b)_p(ab,c)_p(c,-1)_p. \tag{12}$$

Откуда $(a,b)_p = (c,-ab)_p$.

Теперь положим, что $(a,b)_p = (c,-ab)_p$, и докажем, что уравнение $ax^2 + by^2 = c$ имеет ненулевое решение. Тогда из выкладок (12) следует, что $(a/c, b/c)_p = 1$. Тогда уравнение $z^2 - \frac{a}{c}x^2 - \frac{a}{b}y^2 = 0$ имеет решения. Обозначим одно такое решение (x_0, y_0, z_0). Если $z_0 \neq 0$, то $(\frac{x_0}{z_0}, \frac{y_0}{z_0})$ — это решение уравнения $ax^2 + by^2 = c$.

Таким образом, если $z_0 \neq 0$, то задача решена. Далее мы считаем, что $z_0 = 0$. Для всяких r_x, r_y рассмотрим уравнение

$$a(tx_0 + r_x)^2 + b(ty_0 + r_y)^2 = c.$$

Оно эквивалентно уравнению

$$(ar_x^2 + br_y^2) + 2t(ax_0r_y + by_0r_x) = c. \tag{13}$$

Для общей пары рациональных чисел (r_x, r_y) имеем $(ax_0r_y + by_0r_x) \neq 0$ и $t_0 = \frac{c-(ar_x^2+br_y^2)}{2(ax_0r_y+by_0r_x)}$ является решением уравнения (13). Соответственно, уравнение $ax^2 + by^2 = c$ имеет бесконечно много рациональных решений. □

Используя свойства символа Гильберта, решите следующую задачу.

Задача 34. Фиксируем однородный многочлен $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2 (n \geqslant 2)$, где $a_1, \ldots, a_n \neq 0$. Положим

$$d = a_1a_2 \ldots .a_n \quad \text{и} \quad \varepsilon = \prod_{i<j}(a_i, a_j)_p. \tag{14}$$

Докажите, что уравнение $f = 0$ имеет ненулевое решение в p-адических числах тогда и только тогда, когда выполнено одно из следующих условий

1) $n = 2$, а число $-d$ является полным квадратом в $\mathbb{Q}_p$;

2) $n = 3$ и $(-1, d)_p = \varepsilon$;

3) $n = 4$ и $d \neq \alpha^2$, или же $d = \alpha^2$ и $\varepsilon = (-1, -1)_p$;

4) $n \geqslant 5$. (т.е., если f зависит от 5 и более переменных, то уравнение $f = 0$ имеет ненулевое решение в $\mathbb{Q}_p$ для любого p.)

Доказательство. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 4, § 2, Теорема 6. □

Выведите из задачи 34 следующее утверждение.

Задача 35. Фиксируем однородный многочлен $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$ $(n \geqslant 2)$, где $a_1, \ldots, a_n \neq 0$, а также целое число $a \neq 0$. Определим d и ε формулой (14). Докажите, что уравнение $f = a$ имеет решение в p-адических числах тогда и только тогда, когда выполнено одно из следующих условий:

1) $n = 1$, а число a/d является полным квадратом в $\mathbb{Q}_p$;

2) $n = 2$ и $(a, -d)_p = \varepsilon$;

3) $n = 3$ и: ad не является точным квадратом в $\mathbb{Q}_p$ или ad является точным квадратом и $\varepsilon = (-1, -d)_p$;

4) $n \geqslant 4$. (Т.е., если f зависит от 4 и более переменных, то уравнение $f = a$ имеет ненулевое решение в $\mathbb{Q}_p$ для любого p).

Доказательство. 1) $a_1x_1^2 = a \Leftrightarrow x_1^2 = \frac{a}{a_1}$. Очевидно, у него есть решения в p-адических числах тогда и только тогда, когда $\frac{a}{a_1}$ является точным квадратом в $\mathbb{Q}_p$.

2) $a_1x_1^2 + a_2x_2^2 = a$. Условие $(a, -d)_p = \varepsilon$ эквивалентно $(a, -a_1a_2)_p = (a_1, a_2)$, что совпадает с задачей 33 для $a = a_1$, $b = a_2$, $c = a$.

3) Требуется решить уравнение $a_1x_1^2 + a_2x_2^2 + a_3x_3^2 - a = 0$. Очевидно, оно эквивалентно уравнению $a_1x_1^2 + a_2x_2^2 + a_3x_3^2 - ax_4^2 = 0$, поскольку получается заменой

$$(x_1, x_2, x_3, x_4) \leftrightsquigarrow (\frac{x_1}{x_4}\frac{x_2}{x_4}\frac{x_3}{x_4}, 1),$$

а это уравнение имеет решения в p-адических числах.

Теперь докажем, что если есть нетривиальное решение с $x_4 = 0$, то существует и нетривиальное решение с $x_4 \neq 0$. Без ограничения общности $x_1 \neq 0$. Пусть (C, D) — решение уравнения $C^2 - D^2 = -\frac{a}{a_1}$ (например, $C = \frac{1-a/a_1}{2}$, $D = \frac{-1-a/a_1}{2}$)).

Очевидно, можно умножить наше решение на $\frac{C}{x1}$, получится $(C,\, x_2,\, x_3,\, 0)$. Легко проверить, что

$$f(C, x_2, x_3, 0) = f(D, x_2, x_3, 1),$$

и мы свели задачу к задаче 34c.

Первый случай: $ad \neq -m^2$ в $\mathbb{Q}_p$. Это эквивалентно $d \neq m^2$.

Второй случай: $ad = -m^2$.

Требуется: $(a_1, a_2)_p(a_1, a_3)_p(a_2, a_3)_p = (-1, -a_1a_2a_3)_p \Leftrightarrow (a_1, a_2)_p(a_1, a_3)_p(a_2, a_3)_p(-a, d) = (-1, -1)_p$.

Очевидно, для решения задачи вида $a = b \Leftrightarrow c = d$, где (a, b, c, d) из множества $\{1, -1\}$, достаточно проверить $ac = bd$, то есть

$$(-a, d)_p = (-1, -d)_p(-1, -1)_p \quad \Leftrightarrow \quad (-a, d)_p = (-1, -1)_p(-1, -1)_p(-1, d)_p \quad \Leftrightarrow$$
$$\Leftrightarrow \quad (-a, d)_p = (-1, d)_p \quad \Leftrightarrow \quad (-1, d)_p(-a, d)_p = 1 \quad \Leftrightarrow \quad (a, d)_p = 1 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad (a, a)_p(a, \frac{d}{a})_p = 1 \quad \Leftrightarrow \quad 1 \cdot 1 = 1,$$

поскольку $\frac{d}{a}$ является точным квадратом в $\mathbb{Q}_p$.

4) Как и в 3), требуется решить уравнение

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_4^2 - ax_5^2 = 0.$$

Но у него всегда есть решения по задаче 34d.

Также доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 4, § 2, Следствие Теоремы 6 □

Задача 36. Докажите принцип Минковского-Хассе.

Доказательство. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра "A course in arithmetic", Глава 4, § 3, Теорема 8. □

Задача 37. Используя задачу 35 и принцип Минковского-Хассе, докажите, что целое число n представимо в виде суммы трёх квадратов рациональных чисел, если и только если оно не представимо в виде $4^a(8b-1)$, т.е. если $-n$ не является полным квадратом в $\mathbb{Q}_2$.

Доказательство. По теореме Минковского-Хассе достаточно проверить, есть ли у уравнения $x^2+y^2+z^2 = n$ p-адические решения или нет. Сохраняем обозначения решения задачи 35. $a_1 = a_2 = a_3 = 1$, $d = 1$, $\varepsilon = (1,1)_p^3 = (1,1)_p$, $a = n$. Сначала докажем, что для $p > 2$ уравнение решается в p-адических числах.

Если n не имеет вид $-m^2$, то всё доказано. Если $n = -m^2$, то $\varepsilon = (1,1)_p = [Problem30] = 1 = [Problem30] = (-1,-1)_p$, то есть у него есть решения.

Осталось рассмотреть случай $p = 2$. Если $n \neq -m^2$, то задача решена. Теперь пусть $n = -m^2$.

Если у уравнения есть решение, то

$$\varepsilon = (1,1)_2 = (-1,-1)_2,$$

что противоречит задаче 31. □

Задача 38. Фиксируем целое число n. Докажите, что если существуют рациональные числа x, y, z такие, что $x^2 + y^2 + z^2 = n$, то существуют и целые числа x', y', z' такие, что

$$(x')^2 + (y')^2 + (z')^2 = n.$$

Выведите из этого утверждения теорему Гаусса.

Доказательство. Пусть рациональные числа x, y, z таковы, что $x^2+y^2+z^2 = n$. Пусть d — это общий знаменатель чисел x, y, z. Выберем x, y, z так, чтобы d было наименьшим возможным. Допустим, что $d > 1$ (т.е. что одно из чисел x, y или z нецело и целых решений у уравнения $x^2 + y^2 + z^2 = n$ нет). Пусть r_x, r_y, r_z — это ближайшие к x, y, z числа соответственно, $s_x := x - r_x, s_y := y - r_y, s_z := z - r_z$. Тогда

$$|s_x|, |s_y|, |s_z| \leqslant \frac{1}{2}, \qquad s_x^2 + s_y^2 + s_z^2 = n - (r_x^2 + r_y^2 + r_z^2) - 2(s_xr_x + s_yr_y + s_zr_z). \tag{15}$$

Положим

$$x' = r_x - \frac{s_x(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}, \qquad y' = r_y - \frac{s_y(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}, \qquad z' = r_z - \frac{s_z(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}.$$

Из (15) следует, что $s_x^2 + s_y^2 + s_z^2 = d'/d$, причём $0 < d' < d$. Откуда следует, что общий знаменатель x', y', z' делит d', то есть меньше d. Заметим, что $x'^2 + y'^2 + z'^2 = n$. Противоречие. Следовательно, $d = 1$ и уравнение $x^2 + y^2 + z^2 = n$ имеет решение в целых числах.

По задаче 37 любое целое положительное число N, не представимое в виде $4^n(8m-1)$ является суммой трёх квадратов рациональных чисел. Благодаря текущей задаче, такое число N есть и сумма трёх квадратов целых чисел. □

Задача 39. Выведите из теоремы Гаусса теорему Лежандра.

Доказательство. Из теоремы Гаусса следует, что всякое положительное целое число, имеющее остатки 1, 2, 3, 5, 6 при делении на 8 представимо в виде суммы трёх квадратов (а, значит, и виде суммы четырёх квадратов). Таким образом, достаточно доказать, что любое положительное целое число, имеющее остатки 0, 4, 7 при делении на 8 представимо в виде суммы 4 квадратов.

Если число n представимо в виде суммы четырёх квадратов, то $4n$ тоже представимо. Отсюда следует, что достаточно доказать, что любое число, имеющее остаток 7 при делении на 8 представимо в виде суммы четырёх квадратов.

Фиксируем положительное целое число n, имеющее остаток 7 при делении на 8. Так как $n-1$ имеет остаток 6 при делении на 8, то $n-1$ представимо по теореме Гаусса в виде суммы трёх квадратов. Следовательно, n представимо в виде суммы четырёх квадратов. □

Важные свойства символа Гильберта

Цель этого раздела — доказать, что для фиксированной пары ненулевых целых чисел (a, b) символ Гильберта $(a, b)_p$ равен 1 для почти всех (=всех, кроме конечного числа) простых чисел p. Как водится, это утверждение является частным случаем более общего утверждения.

Задача 40. a) Пусть f – это однородный многочлен степени n от k переменных, где $k > n$. Тогда число решений f (включая нулевое) в остатках при делении на p делится на p (Подсказка: примените малую теорему Ферма и рассмотрите случай $p = 2$).

b) Пусть f – это многочлен степени не более n от k переменных, где $k > n$. Тогда число решений уравнения $f = 0$ в остатках при делении на p делится на p.

Доказательство. Из пункта b) пункт a) очевидно следует. Мы докажем пункт b). Рассмотрим многочлен $f(x_1, \ldots, x_k)$ степени n. Рассмотрим сумму

$$\sum_{x_1,\ldots,x_n} f(x_1, \ldots, x_n)^{p-1}, \tag{16}$$

где $x_1, \ldots, x_n$ пробегают все остатки при делении на p. Заметим, что все элементы суммы (16) равны 0 или 1 в остатках при делении на p. Соответственно, значение суммы (16) определяет остаток при делении на p числа решений уравнения $f(x_1, \ldots, x_k)$ в остатках при делении на p. Заметим что степень многочлена $f(x_1, \ldots, x_k)^{p-1}$ равна $(p-1)n$. Следовательно, в каждый моном $f(x_1, \ldots, x_k)^{p-1}$ хотя бы одна из $k > n$ переменных входит в степени меньшей $p - 1$. Заметим, что в для такого монома суммирование (16) даёт ноль, т.к.

$$\sum_{x_i} x_i^l \equiv 0 \pmod{p},$$

если $l < p - 1$. Следовательно, (16)$\equiv 0 \pmod{p}$ и число решений уравнения $f(x_1, \ldots, x_n) = 0$ сравнимо с 0 по модулю p. □

Задача 41. Выведите из предыдущей задачи, что уравнение $ax^2 + by^2 + cz^2 = 0$ от переменных x, y, z имеет ненулевое решение в остатках при делении на p.

Доказательство. Многочлен $ax^2 + by^2 + cz^2$ имеет степень 2 и зависит от трёх переменных. Следовательно, число решений уравнения $ax^2 + by^2 + cz^2 \equiv 0$ сравнимо с 0 по модулю p. В частности, это значит, что у уравнения $ax^2 + by^2 + cz^2 \equiv 0$ есть ненулевое решение. □

Задача 42. Выведите из предыдущей задачи, что для пары ненулевых целых чисел (a, b) символ Гильберта $(a, b)_p$ равен единице, если $a, b \not\vdots p$. Объясните, почему символ Гильберта $(a, b)_p$ равен 1 для почти всех p.

Доказательство. Фиксируем $p \neq 2$ и такое, что a, b не делится на p. Докажем, что $(a, b)_p = 1$.

Пусть (x_0, y_0, z_0) — какое-то решение уравнения $z^2 - ax^2 - by^2 \equiv 0 \pmod{p}$ такое, что

$$(x_0, y_0, z_0) \not\equiv (0, 0, 0) \pmod{p}$$

(такое решение существует по задаче 41). Без ограничения общности считаем, что $z_0 \not\equiv 0 (\mathrm{mod}\ p)$. Тогда в силу задачи 21 число $ax_0^2 + by_0^2$ является полным квадратом в p-адических числах и, следовательно, уравнение $z^2 - ax^2 - by^2 = 0$ имеет ненулевое решение в p-адических числах, т.е. $(a, b)_p = 1$. □

Задача 43. Выведите из задачи 41, что уравнение $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$ от переменных x, y, z, v, w (a, b, c, d, e – это целые параметры) имеет ненулевое решение в $\mathbb{Q}_p$ для всех простых чисел p.

Доказательство. Без ограничения общности считаем, что числа a, b, c, d, e целые и свободны от квадратов. Докажем, что мы можем считать, что никакие три из чисел a, b, c, d, e не имеют общего делителя. Действительно, пусть p — это общий простой делитель трёх или более чисел. Тогда, умножая уравнение $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$ на p и сокращая полные квадраты, приходим к уравнению, в

котором p делит не более 2 из чисел a, b, c, d, e. Очевидно, что эта процедура «сокращения простого делителя» применима независимо ко всем простым делителям чисел a, b, c, d, e. Таким образом, для каждого простого числа p по крайней мере 3 из чисел a, b, c, d, e не делятся на p.

Фиксируем простое число p. Без ограничения общности считаем, что a, b, c не делятся на p.

Если p — нечётное простое число, то, используя задачу 42, получаем, что уравнение

$$ax^2 + by^2 + cz^2 = 0$$

(а, следовательно, и уравнение $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$) имеет ненулевое решение в $\mathbb{Q}_p$.

Если $p = 2$, то задачу можно решать перебором остатков чисел a, b, c, d, e по модулю 8. □

Задача 44. Докажите, что для всякой пары ненулевых целых чисел (a, b) имеет место равенство

$$\prod_p (a, b)_p = (a, b)_{-1}, \tag{17}$$

где произведение берётся по всем простым числам p, а

$$(a, b)_{-1} = \begin{cases} 1, & \text{если уравнение } z^2 - ax^2 - by^2 = 0 \text{ имеет решение в } \mathbb{R}, \\ -1 & \text{иначе.} \end{cases}$$

Доказательство. В силу мультипликативности символа Гильберта (задача 32), достаточно проверить (17) для случая, когда a, b — это простые числа или -1.

Начнём со случая, в котором $a = b = -1$. Тогда $(a, b)_p = (-1, -1)_p$ может быть не равно 1, только если $p = 2$. В этом случае непосредственно проверяется, что $(-1, -1)_2 = (-1, -1)_{-1} = -1$.

Следующий случай: $a = -1$, b — это простое число. Тогда $(a, b)_p = (-1, b)_p$ может быть не равно 1, только если $p = b$ или $p = 2$. В этом случае непосредственно проверяется, что $(-1, p)_p = (-1, p)_2$ для $p \neq 2$ и $(-1, 2)_2 = 1$. Откуда следует, что левая часть (17) равна правой части (17) и равна 1. □

Имеет место следующий «далёкий аналог китайской теоремы об остатках»: оказывается, что по значениям символа Гильберта может быть построен элемент с данными значениями.

Задача 45. Зафиксируем конечный набор ненулевых целых чисел a_i и для каждого простого p зададим значения $\varepsilon_{i,p} = \pm 1$. Тогда система уравнений

$$(a_i, x)_p = \varepsilon_{i,p} \quad \forall i, \ \forall p,$$

имеет решение, если и только если

а) почти все (=все кроме конечного числа) $\varepsilon_{i,p} = 1$,

b) для каждого простого числа p существует ненулевое p-адическое число x_p такое, что

$$(a_i, x_p) = \varepsilon_{i,p}.$$

Доказательство. Доказательство этого факта можно прочитать в книжке Ж.-П. Серра “A course in arithmetic”, Глава 3, § 2, Теорема 4. □

Уравнения от двух переменных и карты (ДУ-3)

Рассматривается уравнение

$$E_m: \quad ax^2 + bxy + cy^2 = m \tag{18}$$

от целых переменных x, y, где a, b, c, m — какие-то целые числа (параметры).

Задача 46 (Суперзадача)**.** Докажите, что если уравнение E_m имеет решения при каком-то положительном числе m, при каком-то отрицательном числе m и не имеет решений при $m = 0$, то для всякого m или E_m не имеет решений, или же E_m имеет бесконечно много решений.

Доказательство. Из условия задачи следует, что $f(x, y) = ax^2 + bxy + cy^2$ является знакопеременной квадратичной формой и не представляет 0. Следовательно, карта f разделяется на положительную и отрицательную часть периодической рекой. Следовательно, карта f периодична и, следовательно, любое значение, выписанное на карте, повторяется на ней бесконечно много раз. Откуда и следует утверждение суперзадачи. □

Задача 47 (Суперзадача)**.** Верно ли, что если уравнение E_m имеет решения в целых числах при

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

то E_m имеет решения при всяком целом числе m?

Доказательство. Имеется контрпример: $f(x, y) = x^2 + xy - 18y^2$. □

Задача 48 (Суперзадача)**.** Докажите, что если уравнения E_1, E_2, E_3, E_5 имеют решения в целых числах, то уравнение E_m имеет решения при каком-то $m < 0$.

Доказательство. Допустим обратное, то есть допустим, что f положительно определена или неотрицательно определена. Мы рассмотрим эти два случая отдельно.

Пусть f неотрицательно определена. Тогда $f(x, y) = r(px + qy)^2$ для каких-то целых чисел r, p, q. Так как f представляет 1, то $r = 1$. Но тогда f не представляет 5.

Пусть теперь f положительно определена. Без ограничения общности считаем, что

$$f = px^2 + qy^2 + r(x - y)^2$$

для каких-то неотрицательных чисел p, q, r (см. задачу 60). Числа p, q, r или одновременно целые или одновременно полуцелые. Без ограничения общности считаем, что $p \geq q \geqslant r$.

Наименьшее ненулевое значение f равно $q + r$. Так как f представляет 1, $q + r = 1$. Следовательно, или $q = 1$, $r = 0$, или $q = r = \frac{1}{2}$. Мы рассмотрим оба этих случая.

Положим, что $q = 1$, $r = 0$. Так как f представляет 2, то $p = 1$ или $p = 2$. В первом случае f не представляет 3, а втором случае f не представляет 5.

Положим теперь, что $q = r = \frac{1}{2}$. Тогда для всех положительных p имеем

$$f(x, y) \geqslant x^2 - xy + y^2.$$

Так как f представляет 2, то $f(x, y) = 2$ для каких-то целых чисел x, y. В частности, $x^2 - xy + y^2 \leq 2$. Это неравенство выполнено для следующих пар (x, y):

$$(0, 1), \qquad (1, 0), \qquad (1, 1).$$

Так как $f(0, 1) = 1$, остаётся два варианта: $f(1, 0) = 2$, $f(1, 1) = 2$. Мы рассмотрим оба этих варианта.

Положим $f(1, 0) = 2$. Тогда $p = \frac{3}{2}$, $f = y^2 - xy + 2y^2$. В этом случае f не представляет 3.

Положим $f(1, 1) = 2$. Тогда $p = \frac{3}{2}$, $f = y^2 - xy + 2y^2$. В этом случае f не представляет 3. □

Рисуем картинки

Задача 49. Покажите, что если $\{w_1, w_2\}$ — это базис $\mathbb{Z}^2$, то пары

$$\{w_2, w_1\},\ \{w_1 - w_2, w_2\},\ \{w_1 + w_2, w_2\},\ \{-w_1, w_2\} \tag{19}$$

также являются базисами $\mathbb{Z}^2$.

Доказательство. По аналогии с решением задачи 11. □

Задача 50. Покажите, что преобразованиями (19) можно из любого базиса получить любой другой.

Доказательство. Пусть $\{u, v\} := \{(a, b), (c, d)\}$ — какой-то базис $\mathbb{Z}^2$. Мы докажем, что преобразованиями (19) можно получить из любого базиса $\{u, v\}$ базис $\{(1, 0), (0, 1)\}$. Рассмотрим квадратичную форму $f(x, y) := x^2 - xy + y^2$. В соответствии с задачей 60 в каком-то базисе $\{u', v'\}$, получаемом из $\{u, v\}$ преобразованиями (19) форма f будет эквивалентна форме вида

$$\begin{gathered} px^2 + qy^2 + r(x + y)^2, \\ 2p = f(v') + f(u' + v') - f(u') \geqslant 0, \\ 2q = f(u') + f(u' + v') - f(v') \geqslant 0, \\ 2r = f(u' + v') - f(u') - f(v') \geqslant 0. \end{gathered} \tag{20}$$

Наименьшие значения формы (20) достигаются на парах (x, y) равных

$$(0, 1), \qquad (1, 0), \qquad (1, 1), \tag{21}$$

и значение на любой другой паре (x, y) больше хотя бы одного из этих значений. Так как $x^2 - xy + y^2$ положительно определена, значение (20) на всех парах, кроме пар списка (21), больше 1. Следовательно, значение (20) на парах (21) равно 1. Что влечёт одно утверждение из трёх идущих ниже

$$\{u', v'\} = \{(0, 1), (1, 0)\}, \qquad \{u', v'\} = \{(0, 1), (1, 1)\}, \qquad \{u', v'\} = \{(1, 0), (1, 1)\}. \tag{22}$$

Следовательно, базис $\{u, v\}$ эквивалентен одному из базисов (22). Что и требовалось доказать. □

Задача 51. Покажите, что квадратичная форма может записываться одинаково в нескольких разных базисах.

Доказательство. Форма $x^2 - 2y^2$ имеет один и тот же вид в базисах $\{(3, 2), (4, 3)\}$ и $\{(1, 0), (0, 1)\}$. □

Задача 52. Укажите квадратичную форму, для которой любым двум разным базисам $\mathbb{Z}^2$ соответствуют различные квадратичные формы.

Доказательство. Положим $f(x, y) := 2x^2 - xy + 3y^2$. Покажем, что разным базисам соответствуют разные формы. Допустим обратное, т.е. что существуют два разных базиса, в которых форма имеет один и тот же вид. Будем последовательно и синхронно восстанавливать по этим базисам всю карту квадратичной формы. Рассмотрим первое пересечение областей восстановления. Это пересечение есть или ребро или вершина. Рассмотрим эти случаи отдельно.

Пусть это пересечение — вершина. Тогда карта квадратичной формы, симметрична относительно одного из рёбер, выходящих из этой точки. Как следствие, единственный колодец формы $2x^2 - xy + 3y^2$ также симметричен относительно одного из рёбер, выходящих из него. Что неверно, так как значения вокруг колодца равны 2, 3, 4.

Пусть это пересечение — ребро. Тогда карта квадратичной формы, сохраняется при перемене местами вершин ребра. Как следствие, вершинами ребра являются колодцы, в частности, форма $2x^2 - xy + 3y^2$ имеет два колодца. Противоречие.

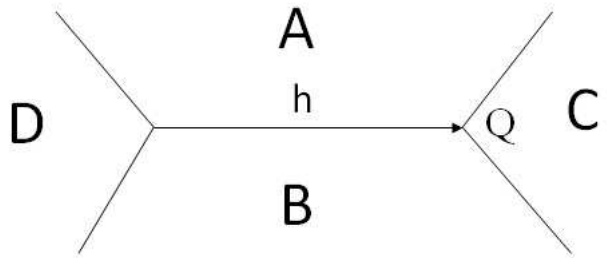
Следовательно, форма $2x^2 - xy + 3y^2$ имеет разный вид в разных базисах. □

Упражнение 1. Выпишите все расширения данного базиса $\{w_1, w_2\}$. Выпишите все специализации данного супербазиса $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$.

Упражнение 2. Нарисуйте (ориентированные) карты для квадратичных форм

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

В следующих двух задачах числа A, B, C, D, h относятся к картинке



Задача 53. Покажите, что для чисел A, B, C, D, h выполнены соотношения:

$$C = A + B + h, \quad D = A + B - h.$$

Доказательство. Утверждение задачи эквивалентно следующему тождеству для квадратичных форм.

$$f(x + y) + f(x - y) = 2(f(x) + f(y)).$$

□

Задача 54. Пусть A, B, C положительны, а ребро h направлено от C к D. Покажите, что тогда число D также положительно, а стрелки на двух остальных рёбрах, инцидентных вершине Q, направлены прочь от Q.

Доказательство. Так как $D, h > 0$, $C = D + 2h > 0$. Числа в областях, инцидентных рёбрам, инцидентным вершине D равны

$$4A + 2h + B, 4B + 2h + A \ (> A, B). \tag{23}$$

Следовательно, оставшиеся две стрелки направлены от вершины D. □

Задача 55. Докажите, что граф, задаваемый точками-супербазисами и рёбрами-базисами, является деревом, т.е. не содержит циклов.

Доказательство. Рассмотрим квадратичную форму $f(x, y) = x^2 + xy + y^2$. Её карта обладает единственным колодцем Q и по задаче 54 все стрелки этой карты направлены прочь от Q. Если бы на карте существовал цикл, то все стрелки не могли бы быть направлены прочь от Q. Следовательно, карта формы f не содержит циклов. Но сам граф карты не зависит от формы, поэтому для любой f он будет деревом. □

Задача 56. Пусть Q — единственный колодец положительно определённой квадратичной формы f, а p, q, r — это числа, записанные в областях, примыкающих к Q. Покажите, что в любой другой области карты f написано число, большее, чем любое из чисел p, q, r.

Доказательство. Фиксируем область A такую, что

a) A не граничит с колодцем,

b) значение f на A наименьшее среди всех областей, удовлетворяющих a).

Мы докажем, что $f(A) > p, q, r$. Это завершит решение задачи. Выделим наименьший по длине путь W от A до колодца Q. Так как W кратчайший, то он упирается в A. По определению колодца все стрелки с ним граничащие направлены от него. Из этого и задачи 54 следует, что все стрелки W направлены от Q. Следовательно, последняя стрелка в пути W указывает на A. В силу условий a), b), это возможно только если W состоит из одного ребра. Для областей, соединённых ребром с колодцем, утверждение задачи проверяется непосредственно (см. формулу (23)). □

Задача 57. Докажите, что всякая положительно определённая квадратичная форма обладает колодцем.

Доказательство. Выберем вершину карты Q, для которой сумма значений в соседних областях наименьшая. Она и будет колодцем (см. также задачу 53). □

Задача 58. a) Докажите, что положительно определённая квадратичная форма имеет не более двух колодцев.

b) Укажите квадратичную форму, обладающую двумя колодцами.

Доказательство. b) Форма $x^2 + y^2$ положительно определена и обладает двумя колодцами.

a) Пусть Q — это какой-то колодец, а $p \geqslant q \geqslant r$ — это значения в областях вокруг него. Есть два случая $q + r > p$ или $q + r = p$. Мы рассмотрим эти случаи отдельно.

Положим, что $q + r > p$. Тогда все три стрелки, выходящие из Q, направлены прочь от Q. Фиксируем вершину Q' и предположим, что она является колодцем. Пусть W — это кратчайший путь, соединяющий Q и Q'. По задаче 54, все стрелки W направлены от Q к Q'. Следовательно Q' — не колодец.

Положим, что $q+r = p$. Тогда второй конец ребра E, отделяющего значение q от значения r, также является колодцем, и мы обозначим его Q'. Наборы чисел вокруг Q и Q' одинаковы. Пусть существует ещё один колодец Q'' и пусть W — это кратчайший из путей, соединяющих Q'' либо с Q, либо с Q'. Тогда W не проходит через второй колодец. Без ограничения общности мы считаем, что W соединяет Q с Q''. Тогда все рёбра на W направлены от Q, и, следовательно, Q'' не является колодцем. □

Задача 59. Объясните, как решить уравнение $ax^2 + bxy + cy^2 = m$ (a, b, c, m — параметры, x, y, z — переменные) в предположении, что форма $ax^2 + bxy + cz^2$ положительно определена.

Доказательство. После замены переменных, мы можем считать, что $f = px^2 + q(x - y)^2 + ry^2$ для положительных чисел p, q, r (см. задачу 60). Если $f(x, y) = n$ имеет решение в целых числах, то

$$px^2 \leqslant n, ry^2 \leqslant n. \tag{24}$$

Число пар целых чисел (x, y), для которых x, y удовлетворяют неравенствам (24), конечно. Проверив их все, мы определим, имеет ли уравнение $f(x, y) = n$ решение или нет. □

Задача 60 (Классификация положительно определённых квадратичных форм)**.**

a) Покажите, что каждая положительно определённая квадратичная форма эквивалентна квадратичной форме вида

$$(p + q)x^2 + 2qxy + (q + r)y^2 \tag{25}$$

для какого-то набора положительных чисел p, q, r.

b) Покажите, что две квадратичные формы, соответствующие наборам

$$(p_1, q_1, r_1) \text{ и } (p_2, q_2, r_2),$$

эквивалентны тогда и только тогда, когда эти наборы совпадают как множества.

c) Определите какие наборы (p, q, r) задают целую квадратичную форму.

d) Определите какие наборы (p, q, r) задают положительно определённую квадратичную форму.

Доказательство. Пусть Q — это какой-то колодец f, а m, n, k — это числа вокруг колодца. Положим

$$p = \frac{m + n - k}{2}, q = \frac{m + k - n}{2}, r = \frac{k + n - m}{2}.$$

Тогда f эквивалентна форме

$$px^2 + qy^2 + r(x - y)^2 = (25).$$

Пункт b) в наших обозначениях неверен. Контрпример: $x^2 + 3y^2$ и $x^2 + xy + y^2$. В условии требуется исправить «эквивалентны» на «линейно эквивалентны».

Ответ пункта c): когда числа p, q, r целы, или когда числа $p - \frac{1}{2}$, $q - \frac{1}{2}$, $r - \frac{1}{2}$ целы.

Ответ пункта d): форма f положительно определена, если $p, q, r \geqslant 0$ и хотя бы два из чисел p, q, r отличны от 0. □

Diophantine equations–1

Theorem (Gauss)**.** A positive integer d can be written as a sum of three squares if and only if d cannot be represented in the form $4^n(8m-1)$.

Introductory problems

Problem 1. Prove that the equations a) $2x^2+2xy-y^2=1$, b) $x^2-xy+y^2=2$ have no integer solutions.

Problem 2. Prove that each of the equations a) $x^2-2y^2=1$, b) $x^2-3y^2=1$, and c) $x^2-6y^2=1$ has infinitely many integer solutions.

Problem 3. Prove that the equation $x^2+1000xy+1000y^2=2001$ has infinitely many integer solutions.

Problem 4***.** Fix an odd prime p. Prove that equation $x^2-py^2=-1$ has an integer solution if and only if $p\equiv 1 \pmod 4$.

Problem 5. Prove that for every integer m, the numbers of integer solutions of equations

$$x^2-xy+y^2=m \quad \text{and} \quad 3x^2+9xy+7y^2=m$$

are equal.

Problem 6. Prove that for every integer n the equation $x^2+y^2=n$ has an integer solution if and only if it has a rational solution.

Problem 7. Provide an example of a quadratic equation with integer coefficients which has a rational solution but has no integer solutions.

Problem 8. Prove that for every positive integers a and b there exist infinitely many positive integers m such that the equation $ax^2+by^2=m$ has no integer solutions.

Problem 9. Prove that for every integer m the equation $x^2+2y^2-3z^2=m$ has an integer solution.

Quadratic forms

By definition, a *quadratic form* is a homogeneous polynomial of second degree. We say that f *represents* an integer m if the equation $f=m$ has a nonzero integer solution (thus not every form represents 0.) Two quadratic forms are called *equivalent* if they represent the same set of numbers.

Problem 10. Describe all integers which are represented by forms a) x^2+y^2; b) x^2-y^2; c)* x^2+xy+y^2.

Problem 11. Prove that the quadratic forms

$$f(x,y), \quad f(x-y,y), \quad f(x,y-x), \quad f(-x,y), \quad \text{and} \quad f(x,-y)$$

are equivalent.

Problem 12. a) Prove that the forms x^2+y^2 and x^2+xy+y^2 are not equivalent.

b) Prove that the form $4x^2-6xy+5y^2$ is not equivalent to any form ax^2+by^2 with integer a and b.

Definition 1. A quadratic form is called

a) *positive definite*, if it represents only positive integers,

b) *non-negative definite*, if it represents only non-negative integers,

c) *indefinite*, if it represents the set of integers containing both positive and negative ones.

Problem 13. Provide an example of a non-negative definite form which is not positive definite.

Extended arithmetics: p-adic numbers

Theorem (Legendre)**.** Any integer is a sum of four squares.

Problem 14. Let m and n be square-free integers. Assume that the equation

$$z^2 - mx^2 - ny^2 = 0 \tag{1}$$

has a nontrivial rational solution. Prove that

a) either m or n is positive,

b) m is a quadratic residue modulo n,

c) n is a quadratic residue modulo m.

Problem 15. Reduce Metatheorem for the equations in two variables to the case of equations of the form (1).

Definition 2. An expression of the form

$$a_{-k}p^{-k} + a_{-k+1}p^{-k+1} + \ldots + a_n p^n + \ldots \qquad (k \in \mathbb{Z}, \quad a_i \in \mathbb{Z}) \tag{2}$$

is called a *p-adic number*. Such an expression is a *p-adic integer* if $k \leq 0$.

Problem 16. Let f be a polynomial with integer coefficients. Prove that the equation $f = 0$ has a solution in $\mathbb{Z}_p$ if and only if it has a solution modulo p^n for every positive integer n.

Problem 17. When is a p-adic number in the form (2) equal to 0?

Problem 18. Prove that the product of two nonzero p-adic numbers is also nonzero.

Problem 19. Prove that $\mathbb{Q} \subset \mathbb{Q}_p$ for any prime p (i.e., prove that for every pair of nonzero integers m and n there exists a p-adic number x such that $nx = m$).

Problem 20. Prove that -1 is a square in $\mathbb{Q}_p$ if and only if $p \equiv 1 \pmod 4$.

Problem 21. Find a description of all perfect squares in $\mathbb{Q}_p$.

Problem 22. Prove that for any nonzero 3-adic number m there exists a 3-adic number x such that m is equal to one of the numbers x^2, $2x^2$, $3x^2$, or $6x^2$.

Problem 23. Let p be an odd prime, and let $x_1, \ldots, x_5$ be nonzero p-adic numbers. Prove that there exist indices i and j with $1 \leq i < j \leq 5$ such that x_i/x_j is a perfect square in $\mathbb{Q}_p$.

Problem 24. Prove that for every odd prime p there exist p-adic numbers $x_1, \ldots, x_{p-1}$ such that $x_1^2 + \ldots + x_{p-1}^2 + 1 = 0$.

Problem 25. Prove that the equation $x^2 + x + 1 = 0$ has exactly two solutions in $\mathbb{Z}_7$.

Problem 26. Prove that the equation $x^2 + y^2 = -1$ has a p-adic solution for every odd prime p.

Theorem (the Hasse–Minkowski principle)**.** A quadratic equation $f = 0$ in several variables has rational solutions if and only if the equation $f = 0$ has simultaneously solutions

- in real numbers,
- in $\mathbb{Q}_p$ for every prime p.

Problem 27. Prove the Hasse–Minkowski principle for equations in one and two variables.

Definition 3. Set $(a,b)_p = 1$ if the equation $z^2 - ax^2 - by^2 = 0$ has a nonzero solution in p-adic integers; otherwise set $(a,b)_p = -1$. The value $(a,b)_p$ is the *Hilbert symbol* of the pair (a,b) with respect to the prime p.

Problem 28. Prove the following properties of the Hilbert symbol:

1) $(a,b)_p = (b,a)_p$,
2) $(a,c^2)_p = 1$,
3) $(a,-a)_p = 1$, $\quad (a,1-a)_p = 1$,
4) $(a,b)_p = (a,-ab)_p = (a,(1-a)b)_p$.

Problem 29. Let $(a,b)_p = 1$. Show that $(a',b)_p = (aa',b)_p$ for any a'.

Definition 4. To write down an expression for the Hilbert symbol in a compact form, we will use the *Legendre symbol* $\left(\frac{x}{p}\right)$ defined for any integer x and prime p. It equals to 1, -1, or 0 depending on whether x is a nonzero quadratic residue, a quadratic non-residue, or zero. For an odd prime p, one may calculate it using the formula

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}} \pmod{p}$$

Problem 30. Let p be an odd prime; let $a = p^\alpha u$, $b = p^\beta v$, where α, β, u, v are integers such that u and v are not divisible by p. Prove that

$$(a,b)_p = (-1)^{\alpha\beta\cdot\varepsilon(p)} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha,$$

where $\varepsilon(p) = \frac{p-1}{2}$.

Problem 31. Find an explicit formula for $(a,b)_2$ for every nonzero integers a and b.

Problem 32. Prove that $(a,b)_p(a,b')_p = (a,bb')_p$ for every nonzero integers a, b, b'.

Problem 33. Prove that the equation $ax^2 + by^2 = c$ in variables x and y (with parameters a, b, and c) has a solution in p-adic numbers if and only if $(c,-ab)_p = (a,b)_p$.

Problem 34*. Let us fix a homogeneous polynomial $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$ with $n \geq 2$, where $a_1, \ldots, a_n \neq 0$. Set

$$d = a_1a_2\ldots a_n \quad \text{and} \quad \varepsilon = \prod_{i<j}(a_i,a_j)_p. \tag{3}$$

Prove that the equation $f = 0$ has a nonzero p-adic solution if and only if one of the following conditions is satisfied:

1) $n = 2$ and $-d$ is a square in $\mathbb{Q}_p$;
2) $n = 3$ and $(-1,d)_p = \varepsilon$;
3) $n = 4$ and $d \neq \alpha^2$, or $d = \alpha^2$ and $\varepsilon = (-1,-1)_p$;
4) $n \geq 5$. (i.e., if f depends on 5 or more variables, then $f = 0$ has a nonzero solution in $\mathbb{Q}_p$ for any p.)

Deduce the following problem from problem 34.

Problem 35. Fix a homogeneous polynomial $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$, where $a_1, \ldots, a_n \neq 0$, and an integer $a \neq 0$. Define d, ε by formula (3). Then the equation $f = a$ has a p-adic solution if and only if one of the following conditions is satisfied:

1) $n = 1$ and a/d is a square in $\mathbb{Q}_p$;

2) $n = 2$ and $(a, -d)_p = \varepsilon$;

3) $n = 3$ and ad is not a perfect square in $\mathbb{Q}_p$, or ad is a perfect square and $\varepsilon = (-1, -d)_p$;

4) $n \geq 4$. (i.e., if f depends on 4 or more variables, then the equation $f = a$ has a nonzero solution in $\mathbb{Q}_p$ for any p.)

Problem 36. Prove the Hasse–Minkowski principle.

Problem 37. Using problem 35 and the Hasse–Minkowski principle, show that an integer d is a sum of 3 squares in rational numbers if and only if the number d cannot be represented in the form $4^a(8b - 1)$, i.e. if $-d$ is not a perfect square in $\mathbb{Q}_2$.

Problem 38. Fix an integer n. Prove that if there exist rational numbers x, y, and z such that $x^2 + y^2 + z^2 = n$, then there also exist integers x', y', and z' such that $(x')^2 + (y')^2 + (z')^2 = n$. Deduce the Gauss theorem from this statement.

Problem 39. Deduce the Legendre theorem from the Gauss theorem.

Some properties of the Hilbert symbol (DE-2)

The goal of this section is to show that, for a pair of nonzero integers (a, b), the Hilbert symbol $(a, b)_p$ equals 1 for almost all (=all except finite number) primes p. We deduce this statement from a more general statement presented below.

Problem 40. a) Let f be a homogeneous polynomial of degree n, depending on k variables, where $k > n$. Then the number of solutions of the equivalence $f \equiv 0$ (including 0-solution!) modulo p is divisible by p (Hint: apply the little Fermat theorem and consider case $p = 2$).

b) Let f be a polynomial of degree n depending on k variables, where $k > n$. Then the number of solutions of the equivalence $f \equiv 0$ modulo p is divisible by p.

Problem 41. Deduce from the previous problem that for any integers a, b, c the equivalence $ax^2 + by^2 + cz^2 \equiv 0$ in variables x, y, z has a nonzero solution modulo p .

Problem 42. Deduce from the previous problem that, for a pair of nonzero integers (a, b) and an odd prime p, $(a, b)_p = 1$ if $a, b \not\vdots p$. Explain why $(a, b)_p = 1$ for all primes p except a finite number.

Problem 43. Deduce from Problem 41 that the equation $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$ in variables x, y, z, v, w (a, b, c, d, e are parameters) has a nonzero solution in $\mathbb{Q}_p$ for any odd prime p.

Problem 44. Prove that, for any pair of nonzero integers (a, b), we have

$$\prod_p (a, b)_p = (a, b)_{-1},$$

where the product is taken over all primes p and

$$(a, b)_{-1} = \begin{cases} 1, & \text{if the equation } z^2 - ax^2 - by^2 = 0 \text{ has a real solution,} \\ -1 & \text{otherwise.} \end{cases}.$$

As a last problem of this list, we mention an "analogue" of the Chinese Remainder Theorem: it turns out that one can construct a rational number with the prescribed values of the Hilbert symbol.

Problem 45. Fix a finite set of nonzero integers a_i and for every prime p define the values $\varepsilon_{i,p} = \pm 1$. Show that the system of equations

$$(a_i, x)_p = \varepsilon_{i,p}, \quad \forall i, \ \forall p,$$

has a solution if and only if

a) almost all (=all except finite number) $\varepsilon_{i,p} = 1$,

b) for any prime p, there exists a nonzero p-adic number x_p such that

$$(a_i, x_p) = \varepsilon_{i,p}.$$

Two variables: maps of quadratic forms (DE-4)

In this section we study the equation

$$E_m: \quad ax^2 + bxy + cy^2 = m \tag{4}$$

depending on integer variables x, y, where a, b, c, m are integer parameters.

Problem 46 (Superproblem)**.** Prove that if the equation E_m has a solution for some positive m, has a solution for some negative m, has no non-trivial solutions for $m = 0$, then for every m either E_m has no solutions, or E_m has infinitely many solutions.

Problem 47 (Superproblem)**.** Is it true that if the equation E_m has solutions for

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

then in this case E_m has solutions for any integer m?

Problem 48 (Superproblem)**.** Prove that if the equations E_1, E_2, E_3, E_5 have integer solutions, then the equation E_m has an integer solution for some $m < 0$.

Drawing a map

Problem 49. Prove that, if $\{w_1, w_2\}$ is a basis of $\mathbb{Z}^2$, then pairs

$$\{w_2, w_1\}, \{w_1 - w_2, w_2\}, \{w_1 + w_2, w_2\}, \{-w_1, w_2\} \tag{5}$$

are also bases of $\mathbb{Z}^2$.

Problem 50. Show that, using transformations (5), it is possible to transform any basis to any other one.

Problem 51. Show that a quadratic form can have the same representations in several different bases.

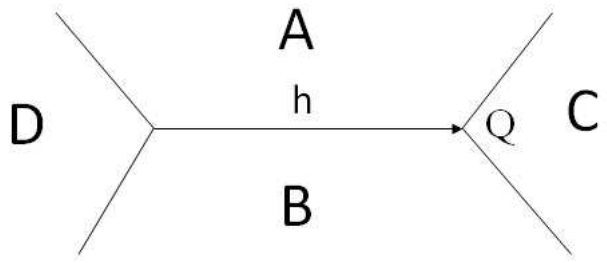
Problem 52. Find a quadratic form which has different representations in any two different bases of $\mathbb{Z}^2$.

Excercise 1. Write down all the extensions of a basis $\{w_1, w_2\}$. Write down all the specializations of a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$.

Excercise 2. Draw (oriented) maps of the following quadratic forms:

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

In two problems below, the values A, B, C, D, and h are related to the following picture.



Problem 53. Show that A, B, C, D, and h satisfy

$$C = A + B + h, \qquad D = A + B - h.$$

Problem 54. Assume that A, B, C are positive and the edge h goes from C to D. Show that in this case D is also positive and that the arrows on two other edges which are incident to Q go out of Q.

Problem 55. Show that the graph determined by the points-superbases and edges-bases is a tree, i.e., it has no cycles.

Problem 56. Let Q be a unique well of a positive definite quadratic form f, and p, q, r be integers written in the regions adjacent to Q. Show that the number in any other region of a map related to f is strictly greater than $\max(p, q, r)$.

Problem 57. Prove that every positive definite form has a well.

Problem 58. a) Prove that a positive definite form has not more than two wells.
b) Find a positive definite form with two wells.

Problem 59. Provide an algorithm which solves the equation $ax^2 + bxy + cy^2 = m$ (a, b, c, m are parameters, x, y, z are variables), under the assumption that $ax^2 + bxy + cz^2$ is positive definite.

Problem 60 (Classification of positive definite quadratic forms)**.**

a) Show that any positive definite quadratic form is equivalent to the form

$$(p + q)x^2 + 2qxy + (q + r)y^2 \tag{6}$$

for some non-negative numbers p, q, r.
b) Show that the quadratic forms corresponding to

$$(p_1, q_1, r_1) \text{ and } (p_2, q_2, r_2)$$

are equivalent if and only if these triples coincide as multisets.
c) Find out which triples (p, q, r) determine an integer quadratic form.
d) Find out which triples (p, q, r) determine a positive definite quadratic form.

Part 3: Little Methuselah form

The goal of this section is to prove the following theorem.

Theorem. (Conway) Little Methuselah form $x^2+2y^2+yz+4z^2$ represents all the integers from 1 to 30. Any other positive definite form $f(x,y,z)$ which represent all the integers from 1 to 30 is linearly equivalent to the little Methuselah form.

To prove the Conway theorem, we try to develop a theory of positive definite quadratic forms in three variables. First we revisit the theory of quadratic forms in two variables.

Let $f(x,y)=ax^2+bxy+cy^2$ be a quadratic form. We assign the following 2x2 and 3x3 tables

$$F:=\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \hat{F}:=\begin{pmatrix} a & \frac{b}{2} & -(a+\frac{b}{2}) \\ \frac{b}{2} & c & -(c+\frac{b}{2}) \\ -(a+\frac{b}{2}) & -(c+\frac{b}{2}) & (a+b+c) \end{pmatrix}$$

to such a form. It is easy to see that f may be uniquely recovered from the tables F and $\hat{F}$.

Problem 61. Prove that

$$f(x,y)=-\frac{b}{2}(x-y)^2+(a+\frac{b}{2})x^2+(c+\frac{b}{2})y^2. \tag{7}$$

.

Problem 62. Prove that the tables

$$\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \begin{pmatrix} c & \frac{b}{2} \\ \frac{b}{2} & a \end{pmatrix}, \quad \begin{pmatrix} a & -\frac{b}{2} \\ -\frac{b}{2} & c \end{pmatrix} \tag{8}$$

determine equivalent quadratic forms.

Problem 63. Prove that the quadratic forms corresponding to the tables

$$\begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix}, \quad \begin{pmatrix} a & -(a+\frac{b}{2}) \\ -(a+\frac{b}{2}) & a+b+c \end{pmatrix}, \quad \begin{pmatrix} c & -(c+\frac{b}{2}) \\ -(c+\frac{b}{2}) & a+b+c \end{pmatrix} \tag{9}$$

are equivalent (note that tables (9) can be obtained from $\hat{F}$ by a choice of 2 rows and corresponding 2 columns).

Below, we identify a quadratic form f with its tables F and $\hat{F}$.

Problem 64. Using (8) and (9), show that any positive definite quadratic form is equivalent to a form

$$\begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix},$$

for which $0 \le -b' \le a' \le c'$. Note that under this restriction, the right hand side of (7) is a sum of 3 non-negative numbers.

Problem 64 is an analogue of Problem 60. We wish to prove the analogue of Problem 64 for quadratic forms in three variables. We will use the same scheme but we need more notation. Fix a quadratic form

$$f(x,y,z) = a_{xx}x^2 + a_{yy}y^2 + a_{zz}z^2 + a_{xy}xy + a_{yz}yz + a_{xz}xz.$$

We identify the form f with the following 3x3 and 4x4 tables:

$$F := \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} \end{pmatrix},$$

$$\hat{F} := \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}.$$

Problem 65. Prove that

$$f(x,y,z) = -\frac{a_{xy}}{2}(x-y)^2 - \frac{a_{xz}}{2}(x-z)^2 - \frac{a_{yz}}{2}(y-z)^2 + \\ + (a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2})x^2 + (a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2})y^2 + (a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2})z^2. \quad (10)$$

Problem 66. Prove that the quadratic forms

$$\begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & \frac{a_{xz}}{2} \\ \frac{a_{xy}}{2} & a_{yy} & \frac{a_{yz}}{2} \\ \frac{a_{xz}}{2} & \frac{a_{yz}}{2} & a_{zz} \end{pmatrix}, \quad \begin{pmatrix} a_{xx} & \frac{a_{xy}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xy}}{2} & a_{yy} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}, \quad (11)$$

$$\begin{pmatrix} a_{xx} & \frac{a_{xz}}{2} & -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) \\ \frac{a_{xz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{xx} + \frac{a_{xy}}{2} + \frac{a_{xz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix}, \quad (12)$$

$$\begin{pmatrix} a_{yy} & \frac{a_{yz}}{2} & -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) \\ \frac{a_{yz}}{2} & a_{zz} & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) \\ -(a_{yy} + \frac{a_{xy}}{2} + \frac{a_{yz}}{2}) & -(a_{zz} + \frac{a_{xz}}{2} + \frac{a_{yz}}{2}) & a_{xx} + a_{yy} + a_{zz} + \\ & & + \frac{a_{xy}}{2} + \frac{a_{yz}}{2} + \frac{a_{xz}}{2} \end{pmatrix} \quad (13)$$

are equivalent (note that tables (13) can be obtained from $\hat{F}$ by a choice of 3 rows and 3 corresponding columns).

Problem 67. Using (13), show that a positive definite form f is equivalent to a form

$$\begin{pmatrix} a'_{xx} & \frac{a'_{xy}}{2} & \frac{a'_{xz}}{2} \\ \frac{a'_{xy}}{2} & a'_{yy} & \frac{a'_{yz}}{2} \\ \frac{a'_{xz}}{2} & \frac{a'_{yz}}{2} & a'_{zz} \end{pmatrix},$$

for which

$$0 < a'_{xx} \le a'_{yy} \le a'_{zz},$$

$$|a'_{xy}|, |a'_{xz}| \le |a'_{xx}|, |a'_{yz}| \le |a'_{zz}|.$$

Problem 68. Using Problem 67, show that every positive definite quadratic form $f(x, y, z)$ is equivalent to

$$\hat{F} := \begin{pmatrix} a'_{xx} & \frac{a'_{xy}}{2} & \frac{a'_{xz}}{2} & -(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) \\ \frac{a'_{xy}}{2} & a'_{yy} & \frac{a'_{yz}}{2} & -(a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) \\ \frac{a'_{xz}}{2} & \frac{a'_{yz}}{2} & a'_{zz} & -(a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) \\ -(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) & -(a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) & -(a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) & a'_{xx} + a'_{yy} + a'_{zz} + \\ & & & + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2} + \frac{a'_{xz}}{2} \end{pmatrix} \tag{14}$$

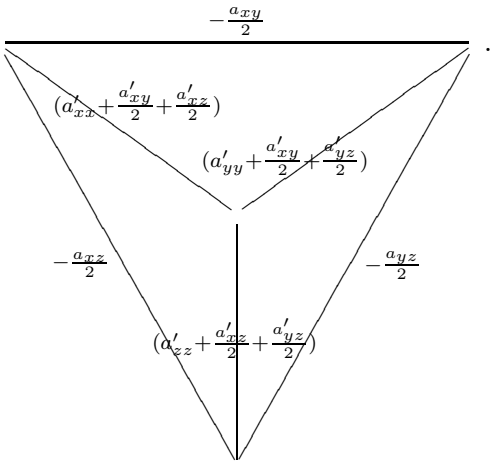
such that

$$a'_{xy},\ a'_{yz},\ a'_{xz} \le 0, \tag{15}$$

$$(a'_{xx} + \frac{a'_{xy}}{2} + \frac{a'_{xz}}{2}) \ge 0,\ (a'_{yy} + \frac{a'_{xy}}{2} + \frac{a'_{yz}}{2}) \ge 0,\ (a'_{zz} + \frac{a'_{xz}}{2} + \frac{a'_{yz}}{2}) \ge 0. \tag{16}$$

Note under these conditions the right hand side of (10) is a sum of squares.

To every quadratic form (14) satisfying conditions (16), we assign a graph with 4 vertices as it is shown on Figure below.

$\mathcal{F} =$ 

If the number on some edge is 0, then we delete this edge.

Problem 69. Prove that if the graph $\mathcal{F}$ has all possible edges, then f does not represent 1.

Problem 70. Prove that if some vertex of $\mathcal{F}$ is incident to strictly less than 2 edges, then f is equivalent to a form

$$ax^2 + g(y, z) \tag{17}$$

for some positive integer a and a positive definite quadratic form g in 2 variables.

Problem 71. Prove that if a quadratic form (17) represents all the integers from 1 to 30, then (17) is equivalent to the little Methuselah form.

We say that f is *indecomposable* if any vertex of $\mathcal{F}$ is incident to at least 2 edges.

Problem 72. Describe the graphs of all indecomposable quadratic forms $f(x, y, z)$ which represent
a) 1;
b) 1, 2;
c) 1, 2, 3, 5.

Problem 73. Finish the proof of the Conway theorem.

Diophantine equations of second degree

In this project we study some properties of Diophantine equations of second degree. Those who advance in the project will develop a theory allowing one to solve a large (and interesting) class of problems. Some exciting examples are presented below.

We start with second degree equations in rational numbers. We will elaborate an algorithm which effectively determines whether an equation has a solution. As an application of this theory, we prove the following theorem by Carl Friedrich Gauss.

Theorem (Gauss)**.** A positive integer number d can be written as a sum of three squares if and only if d cannot be represented in the form $4^n(8m-1)$.

After the semifinal, we will focus on integral solutions of degree 2 equations in two variables. To investigate the solutions of these equations, we will introduce the *maps* of quadratic forms. We will also prove the following statement.

Theorem (J. Conway)**.** There exists a unique[1] homogeneous polynomial $f(x,y,z)$ of degree 2 such that all the equations $f(x,y,z)=m$ with $m=1,\dots,30$ admit integral solutions, but any equation of the form $f(x,y,z)=m$ with $m<0$ has no integral solutions.

Introductory problems

In this subsection we collect several easy problems on (integral) quadratic forms. These problems may be solved using a general algorithm of solution of such equations; we believe that some participants will construct such an algorithm. Nevertheless, all these introductory problems may be solved in a direct way.

Notice that there is no such algorithm for Diophantine equations of an arbitrary degree; the fact that it cannot exist was proved by Yu. Matiyasevich in 1970; by proving this fact he has solved the 10th Hilbert problem.

See Problems 1–9.

If you are stuck on some of these problems, you may proceed to the next sections and return to these problem later, after obtaining some technical background.

The quadratic forms

Definition 1. A *quadratic form* is a homogeneous polynomial of degree 2. Here are two examples: $2x^2+2xy-y^2$ and $x^2-xz+y^2-2z^2$.

For every positive integer d, we denote by $\mathbb{Z}^d$ the set of d-tuples of integers. E.g., the set of pairs of integers is denoted by $\mathbb{Z}^2$. Any quadratic form in two variables x and y determines a function on $\mathbb{Z}^2$ mapping a pair (x,y) to the number $f(x,y)$. Hereafter we will frequently denote a pair $(x,y)\in\mathbb{Z}^2$ by one letter (say, v) and write $f(v)$ for $f(x,y)$.

Definition 2. We say that a quadratic form f *represents an integer* m if there exists a pair $v\in\mathbb{Z}^2$ with $v\neq(0,0)$ and $f(v)=n$. In other words, f represents m if the equation $f(x,y)=n$ has a nonzero integer solution (thus not any quadratic form represents 0).

See Problems 10–11.

Definition 3. We say that two quadratic forms are *equivalent* if each number represented by one of these forms can also be represented by the other one.

[1]Formally, this statement is wrong; this polynomial is unique up to some equivalence which will be described later.

See Problem 12.

It appears that some quadratic forms are easier to deal with than some other ones. One of our aims is the following: Given a quadratic form f, we wish to find some convenient form equivalent to it (e.g., a form like $ax^2 + by^2$). For that, we need to work out some necessary and sufficient conditions on the two quadratic forms to be equivalent. In particular, we will find some explicitly computable invariants of quadratic forms.

Definition 4. We say that a quadratic form f is
a) *positive definite* if $f(v) > 0$ for all $v \neq 0$,
b) *non-negative definite* if $f(v) \geq 0$ for all $v \in \mathbb{Z}^2$,
c) *indefinite* if $f(u) > 0$ for some $u \in \mathbb{Z}^2$ and $f(v) < 0$ for some $v \in \mathbb{Z}^2$.

See Problem 13.

Extended arithmetics: p-adic numbers

The main goal of this section is to impart some sense to the following Metatheorem.

Theorem (Metatheorem)**.** A quadratic equation has a solution in rational numbers if and only if there are no obstacles modulo any prime p.

Using this Metatheorem, one can prove, for instance, the Gauss theorem and the following theorem by Legendre.

Theorem (Legendre)**.** Every positive integer is a sum of four squares of integers.

In our project we split the proof of Metatheorem (as well as of theorems by Gauss and Legendre) into several problems. Any such problem can be solved independently. To start with, we need to impart a formal sense to our Metatheorem (in the previous formulation, it is ambiguous; moreover, it remains wrong after any easy attempt to formalize it). Let us present some example.

Definition 5. We say that m is a *quadratic residue* modulo n if there exists an integer t such that $m \equiv t^2 \pmod{n}$.

See Problems 14–15.

In the case $\gcd(m, n) = 1$, the conditions a)–c) of Problem 14 imply that the equation

$$ax^2 + by^2 = c$$

has a rational solution. On the other hand, in the case $\gcd(m, n) \neq 1$ one needs to introduce additional conditions on m and n which are related to prime divisors of $\gcd(m, n)$. If one writes them down directly, these conditions would look a bit long, although simple.

An elegant (and short) way to present such conditions is based on the notion of *p-adic numbers*. We follow this approach.

For any prime p, a *p-adic integer* is defined as any formal sum of the form

$$a_0 + a_1 p + \ldots + a_n p^n + \ldots \qquad (a_i \in \mathbb{Z}) \tag{1}$$

where the number of summands may be infinite. Two p-adic integers are assumed to be equal if they coincide modulo p^n for any n. For example,

$$1 = (p+1) - (p+1)p + (p+1)p^2 - (p+1)p^3 + \ldots.$$

The set of p-adic integers is denoted by $\mathbb{Z}_p$.

One may add, subtract, and multiply p-adic integers in an obvious way. Therefore, given an equation $f = 0$ with integer coefficients, one may consider its solutions in $\mathbb{Z}_p$. The following problem provides a connection between the sets of solutions of $f = 0$ in integers and in p-adic integers.

See Problem 16.

The notion of a p-adic integer is an extension of a notion of an integer. A similar extension exists for the rational numbers. Namely, for any prime p we define a *p-adic number* (or a *p-adic rational*) as a formal expression of the form

$$a_{-k}p^{-k} + a_{-k+1}p^{-k+1} + \ldots + a_n p^n + \ldots \qquad (k \in \mathbb{Z}, \quad a_i \in \mathbb{Z}); \tag{2}$$

the equality of two p-adic numbers is defined as above. The set of all p-adic numbers is denoted by $\mathbb{Q}_p$. Obviously, any p-adic integer can be represented in the form (2) with $a_{-k} = ... = a_{-1} = 0$ (or with $k \leq 0$).

In order to get acquainted with the notion of p-adic numbers, it is useful to solve the following problems.

See Problems 17–26.

Now we are ready to present a formal version of Metatheorem.

Theorem (the Hasse–Minkowski principle)**.** A quadratic equation $f = 0$ has a rational solution if and only if it simultaneously has solutions

- in real numbers,
- in p-adic numbers for every prime p.

See Problem 27.

The Hasse–Minkowski principle reduces solving an equation in rational numbers to solving the same equation in p-adic numbers. The advantage is that equations in p-adic numbers are much easier to solve. To show this, we first describe an algorithm which allows one to check whether an equation in two variables has a rational solution. Let us first deal with an equation of the form

$$z^2 - ax^2 - by^2 = 0. \tag{3}$$

Definition 6. Consider a prime p and a pair of integers (a, b). Let us define the *Hilbert symbol* $(a, b)_p$ of a pair (a, b) with respect to p as follows: If the equation (3) has a nonzero solution in p-adic integers, then we set $(a, b)_p = 1$; otherwise we set $(a, b)_p = -1$.

Thus, for finding the solutions of (3) it is helpful to learn how to find $(a, b)_p$.

See Problems 28–29.

To write down an expression for the Hilbert symbol in a compact form, we will use the *Legendre symbol* $\left(\frac{x}{p}\right)$ defined for any integer x and prime p. It equals to 1, -1, or 0 depending on whether x is a nonzero quadratic residue, a quadratic non-residue, or zero. For an odd prime p, one may calculate it using the formula

$$\left(\frac{x}{p}\right) = x^{\frac{p-1}{2}} \pmod{p}$$

See Problems 30–39.

Two variables: maps of quadratic forms

In this section of the project we develop a technique which allows us to solve the equation

$$E_m : ax^2 + bxy + cy^2 = m \tag{4}$$

effectively, here x and y are integer variables and a, b, c, m are integer parameters. To do this, we assign a *map* to any quadratic form in two variables and show how to read properties of the form out of this map. We believe that using this approach the participants will be able to solve the following (super)problems. By a solution in this section we always mean a nonzero integer solution if not mentioned otherwise.

Problem 46 (Superproblem)**.** Assume that the equation E_m has a solution for some positive m, for some negative m and has no solutions for $m = 0$. Prove that in this case either E_m has no solutions, or E_m has infinitely many solutions for any m.

Problem 47 (Superproblem)**.** Is it true that if the equation E_m has solutions for

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

then in this case E_m has solutions for any integer m?

Problem 48 (Superproblem)**.** Assume that the equations E_1, E_2, E_3, E_5 have solutions. Show that in this case the equation E_m has solutions for some $m < 0$.

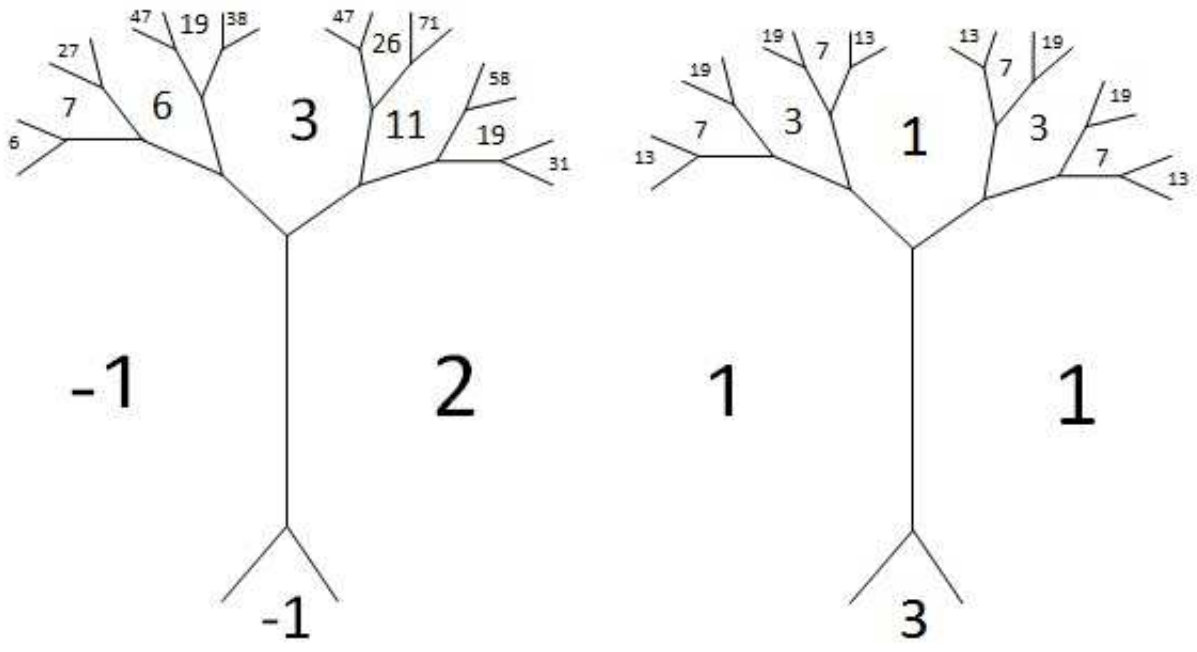
Now we treat two examples to show how the map of a quadratic form may help to solve equations.

Examples of maps

The goal of this subsection is to show that it might be interesting to consider maps of quadratic forms. Given two polynomials

$$2x^2 + 2xy - y^2 = 1 \text{ and } x^2 - xy + y^2 = 2, \tag{5}$$

we assign the following pictures to them, they are called *maps*:



From these maps we see that equations (5) have no integer solutions.

Drawing a map [2]

To find something common in a variety of something very different it was a good idea (from time to time) to consider all this different (some)things simultaneously and providing this "all" by some additional structure. Following this approach, we consider all forms which are linearly equivalent (see definition below) to a form f and provide this set with an oriented graph structure (we put points of quadratic forms on the plane and connect them by edges in some way). To do this we need a notion of basis/superbasis of $\mathbb{Z}^2$.

Definition 7. A *basis* of $\mathbb{Z}^2$ is a pair $w_1, w_2 \in \mathbb{Z}^2$ such that for any $v \in \mathbb{Z}^2$ there exist $m, n \in \mathbb{Z}$, for which

$$v = mw_1 + nw_2.$$

Before semifinal, we had the notion of equivalent forms. Unfortunately, if we work with maps of quadratic forms, it is more natural to use the following notion.

Definition 8. Two forms f_1, f_2 are called *linearly equivalent*, if $\exists$ a, b, c, d, such that $ad - bc = 1$ and

$$f_1(x, y) = f_2(ax + by, cx + dy).$$

See/solve Problems 49–52.

Definition 9. A *superbasis* of $\mathbb{Z}^2$ is a collection $\{\pm w_1, \pm w_2, \pm(w_1 + w_2)\}$, where $\{w_1, w_2\}$ is a basis of $\mathbb{Z}^2$. We say that a basis $\{w_1, w_2\}$ is a *specialization* of a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$. We say that a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$ is an *extension* of a basis $\{w_1, w_2\}$.

Example 1. Write down all the extensions of a given basis $\{w_1, w_2\}$. Write down all the specializations of a given superbasis $\{\pm w_1, \pm w_2, \pm(w_1 + w_2)\}$.

Now we are able to describe the map f. We start from a part of this map which does not depend on f at all:

(1) to any superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$, we assign a point on the plane (the vertex of the graph),

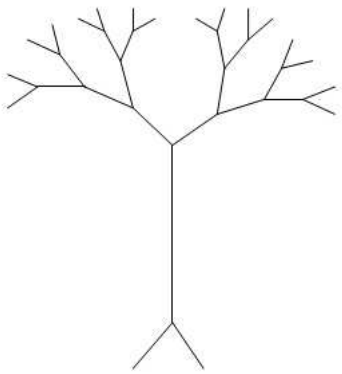
(2) to any basis $\{w_1, w_2\}$, we assign a segment on the plane (the edge of the graph), which connects

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\} \text{ and } \{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\}$$

(we assign the same edge to $\{w_1, w_2\}$, $\{-w_1, w_2\}$, $\{w_1, -w_2\}$, and $\{-w_1, -w_2\}$);

(3) to any $w \in \mathbb{Z}^2$, we assign the region on the plane such that its border consists of edges corresponding to bases containing w (we assign the same region to w and $-w$).

It turns out that it is possible to draw the following picture without self-intersections on the plane.



(6)

[2] If you wish to see a much shorter way to draw a map of a form you could go to the appendix of this section. Try to prove why a map defined in this way satisfies the desired properties.

Note that (6) does not depend on f. Now we will mark the graph with integers depending on f. Integers will be assigned to every region and to every edge of (6) in such a way that it will be possible to restore the class of f up to linear equivalence in the unique way. We use the following rules.

(1) If a region corresponds to $w \in \mathbb{Z}^2$, then $f(w)$ will be assigned to it.

(2) If an edge I corresponds to a basis $\{w_1, w_2\}$, then we assign to it the positive integer

$$|f(w_1 + w_2) - f(w_1) - f(w_2)|.$$

Also, we make I directed: if $f(w_1 + w_2) > f(w_1 - w_2)$, then edge I starts at vertex-superbasis

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 - w_2)\}$$

and ends in

$$\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\};$$

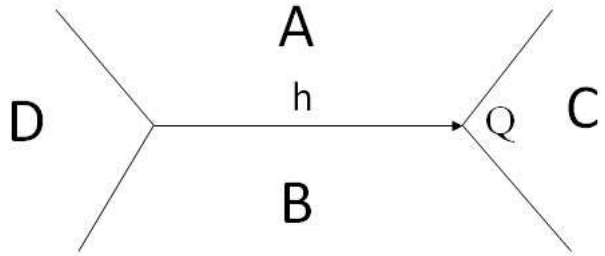
if $f(w_1+w_2) < f(w_1-w_2)$, then otherwise. If $f(w_1+w_2) = f(w_1-w_2)$, then we do not determine the direction of I (and usually omit 0 at I).

The resulting picture will be called the *oriented map* of a quadratic form f. If we omit numbers attached to the edges in this picture, then it will be called the *map* of a quadratic form. For example, the maps of the forms $2x^2 + 2xy - y^2$ and $x^2 - xy + y^2$ are presented on page 4.

Example 2. Draw the (oriented) maps for the quadratic forms

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

In two following problems, the integers A, B, C, D, h are related to the picture



See/solve Problems 53–55.

For positive definite quadratic forms, the following definition plays a key role.

Definition 10. A *well* is a vertex Q of the oriented map of a quadratic form such that all the edges which are incident to Q go out of Q.

See/solve Problems 56–60.

We want to give you an advice:

1) The ideas of proofs of Superproblems 1, 2, and even of every equation of type (4), is very close to Problems 59, 60.

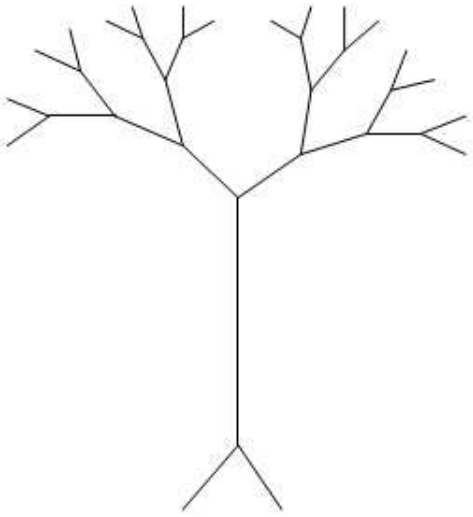
2) In real mathematical life, no one (except yourself) would give you a sequence of (relatively simple) exercises which lead to a proof of any mathematical Problem. You will be very lucky if you learn (most probably, occasionally) a significant piece of the desired methods and ideas somewhere.

3) The goal of this conference is to let you know something about real mathematical life.

If you did not guess, we end up with problems helping you to solve Superproblems 1, 2, 3 and determine when equations (4) have solutions. To simplify your life, we have prepared several pictures which can help you to solve or to guess something.

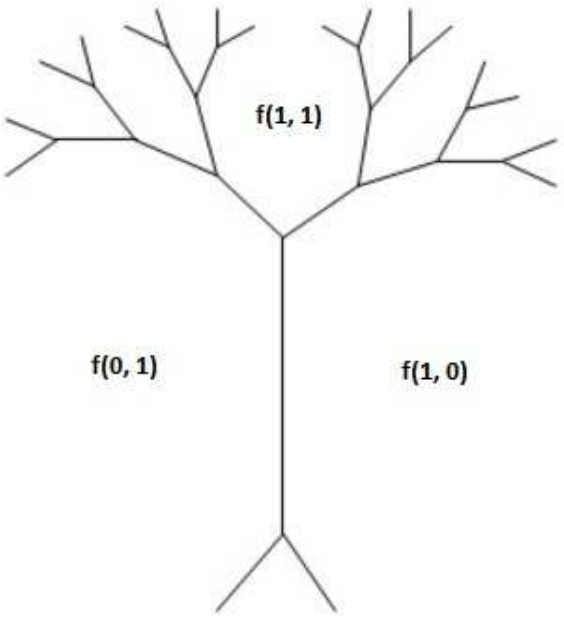
A quick way to describe the map of a quadratic form

Algorithm $f \to \Gamma_f$ **:** We consider an infinite tree (a connected graph without cycles) on the plane such that every its vertex is incident to exactly 3 edges. A part of such a tree is presented below



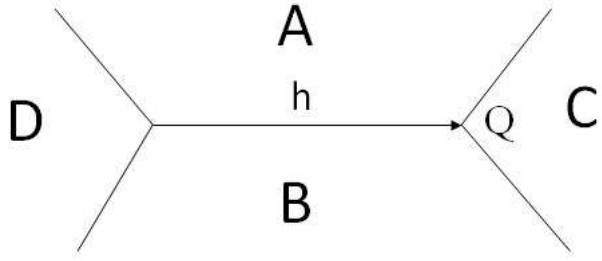
. (7)

We take any vertex of this graph and write integers $f(1,0)$, $f(0,1)$, and $f(1,1)$ on three regions which meet at this vertex.



. (8)

The values at all the other regions are determined by the following rule



. (9)

Rule 1: For a given edge, if 3 values which are adjacent to it (see Figure (9)) are already known, then the fourth one is determined by the formula $2(A+B) = C+D$.

It is easy to see that Rule 1 determines the map Γ_f. Now we need to construct the oriented map $\vec{\Gamma}_f$. We use the following rules (see Figure (9)):

Rule 2: We write $|2(A+B)-C| = |2(A+B)-D|$ at the edge h.

Rule 3: If $C < D$, then the edge h is replaced by an arrow from C to D; if $C < D$ then h is replaced by an arrow from D to C; if $C = D$ then the edge h is unoriented.

Some properties of Γ_f:

1) the points of Γ_f correspond to quadratic forms which are linearly equivalent to f;

2) the regions Γ_f are in one-to-one correspondence with the nonzero rational numbers $\frac{m}{n}$;

3) if the region corresponds to a reduced fraction $\frac{m}{n}$, then integer $f(m,n)$ is written in it;

4) two quadratic forms f and g are linearly equivalent if and only if their maps Γ_f and Γ_g coincide.

Introductory problems

Problem 1. Prove that the equations a) $2x^2+2xy-y^2=1$, b) $x^2-xy+y^2=2$ have no integer solutions.

Proof. a) Analyzing residues modulo three, we see that the equation

$$2x^2+2xy-y^2=3x^2-(y-x)^2=1$$

has no integer solutions.

b) First solution. Notice that $x^2-xy+y^2=(x-y/2)^2+3/4y^2=2$, hence $y^2\leqslant 8/3$, $|y|\leqslant 1$, analogously for x. From the other hand, at least one of x and y must be even, so it must be zero; if we put $x=0$ then $y^2=2$, we get a contradiction.

Second solution. Clearly, if $|x|\geqslant 3$ or $|y|\geqslant 3$, the equation $x^2-xy+y^2=\frac{1}{2}(x^2+y^2+(x-y)^2)=2$ has no integer solutions. A case-by-case consideration of the remaining 25 possibilities shows that this equation has no integer solutions. □

Problem 2. Prove that each of the equations a) $x^2-2y^2=1$, b) $x^2-3y^2=1$, and c) $x^2-6y^2=1$ has infinitely many integer solutions.

Proof. a) For every integer n the pair

$$x=\frac{(3+2\sqrt{2})^n+(3-2\sqrt{2})^n}{2},\quad y=\frac{(3+2\sqrt{2})^n-(3-2\sqrt{2})^n}{2\sqrt{2}}$$

is a solution of the equation.

b) For every integer n the pair

$$x=\frac{(2+\sqrt{3})^n+(2-\sqrt{3})^n}{2},\quad y=\frac{(2+\sqrt{3})^n-(2-\sqrt{3})^n}{2\sqrt{3}}$$

is a solution of the equation.

c) For every integer n the pair

$$x=\frac{(5+2\sqrt{6})^n+(5-2\sqrt{6})^n}{2},\quad y=\frac{(5+2\sqrt{6})^n-(5-2\sqrt{6})^n}{2\sqrt{6}}$$

is a solution of the equation. □

Problem 3. Prove that the equation $x^2+1000xy+1000y^2=2001$ has infinitely many integer solutions.

Proof. The discriminant $1000^2-4\cdot 1000$ of the equation $x^2+1000xy+1000y^2=2001$ is greater than 0 and is not a perfect square. Hence, the quadratic form $x^2+1000xy+1000y^2=2001$ is indefinite, does not represent 0 and represents 2001 for $x=y=1$. From Problem 46 it follows that the equation

$$x^2+1000xy+1000y^2=2001$$

has infinitely many solutions. □

Problem 4. Fix an odd prime p. Prove that equation $x^2-py^2=-1$ has an integer solution if and only if $p\equiv 1 \pmod 4$.

Proof. Suppose that the equation $x^2-py^2=-1$ has an integer solution. Let us prove that p is equivalent to 1 modulo 4. Indeed, in this case -1 is a quadratic residue modulo p, i.e., $p\equiv 1 \pmod 4$.

Now let us prove the converse, let $p\equiv 1 \pmod 4$. By Problem 46, the equation $x^2-py^2=1$ has a nontrivial solution.

Define by S_+ the set of solutions (x_0,y_0) of the equation $x^2-py^2=1$ such that $x_0,y_0>0$. Let (x_0,y_0) be a solution from S_+ with minimal y_0. Then

$$(x_0-1)(x_0+1)=py_0^2. \tag{1}$$

It follows from (1) that either $2(x_0+1)$ or $2(x_0-1)$ is a perfect square. Consider two cases.

First, assume that $2(x_0+1) = d^2$ for a positive integer d. Then d is even ands $d \mid y_0$. Let $d = 2d_0$, and let

$$x_1 = (x_0+1)/d = d_0,\ y_1 = y_0/d.$$

Then

$$x_1^2 - py_1^2 = \frac{1}{d^2}((x_0+1)^2 - py_0^2) = \frac{2(x_0+1)}{d^2} = 1.$$

Hence, (x_1, y_1) also belongs to S_+. Clearly, $y_1 < y_0$, which contradicts the minimality of the pair (x_0, y_0).

Now consider the second case, let $2(x_0-1) = d^2$ for some positive integer d. Then d is even and $d \mid y_0$. Let $d = 2d_0$, and let

$$x_1 = (x_0-1)/d = d_0,\ y_1 = y_0/d.$$

Then

$$x_1^2 - py_1^2 = \frac{1}{d^2}((x_0-1)^2 - py_0^2) = \frac{2(1-x_0)}{d^2} = -1.$$

We see that (x_1, y_1) is a required solution of $x^2 - py^2 = -1$. □

Problem 5. Prove that for every integer m, the numbers of integer solutions of equations

$$x^2 - xy + y^2 = m \quad \text{and} \quad 3x^2 + 9xy + 7y^2 = m$$

are equal.

Proof. Let us show that there is a bijection between integer solutions of the equation $x^2 - xy + y^2 = m$ and integer solutions of $3x^2 + 9xy + 7y^2 = m$.

Let $v = x + y$, $u = -x - 2y$, clearly, they are integers. Substitute x by $u + 2v$, and y by $-u - v$ in $x^2 - xy + y^2$. We get $3u^2 + 9uv + 7v^2$. So for every solution of $3x^2 + 9xy + 7y^2 = m$ there is a corresponding solution of $x^2 - xy + y^2 = m$.

Conversely, u and v in $u^2 - uv + v^2$ can be replaced by an appropriate integer linear combination of x and y. Hence, for every solution of the equation $x^2 - xy + y^2 = m$ we can produce the solution of the equation $3x^2 + 9xy + 7y^2 = m$. □

Problem 6. Prove that for every integer n the equation $x^2 + y^2 = n$ has an integer solution if and only if it has a rational solution.

Proof. Let x, y be rational numbers such that $x^2 + y^2 = n$. Reduce x and y to a common denominator d and choose the pair (x, y) with the minimal possible d. We assume that $d > 1$ (this means that x, y are not integers and the equation $x^2 + y^2 = n$ has no integer solutions). Let r_x, r_y be the integers closest to x and y, respectively. Denote $s_x := x - r_x$, $s_y := y - r_y$. Then

$$|s_x|,\ |s_y| \leqslant \frac{1}{2}, \quad s_x^2 + s_y^2 = n - (r_x^2 + r_y^2) - 2(s_x r_x + s_y r_y). \tag{2}$$

Let

$$x' = r_x - \frac{s_x(n - r_x^2 - r_y^2)}{s_x^2 + s_y^2}, \qquad y' = r_y - \frac{s_y(n - r_x^2 - r_y^2)}{s_x^2 + s_y^2}.$$

It follows from (2) that $s_x^2 + s_y^2 = d'/d$, and $0 < d' < d$. Hence, if we write x', y' with common denominator d', then it divides d, in particular, it is less than d. Meanwhile $x'^2 + y'^2 = n$. It contradicts the minimality of (x, y). Hence, $d = 1$, i.e., the equation $x^2 + y^2 = n$ has integer solutions. □

Problem 7. Provide an example of a quadratic equation with integer coefficients which has a rational solution but has no integer solutions.

Proof. The equation $4x^2 = 1$ is the required example. It has a rational solution $x = \frac{1}{2}$. Clearly, it has no integer solutions. □

Problem 8. Prove that for every positive integers a and b there exist infinitely many positive integers m such that the equation $ax^2 + by^2 = m$ has no integer solutions.

Proof. Let N be an integer. If the equation $ax^2 + by^2 = n$ has integer solutions for some $n \leqslant N$, then

$$|x| \leqslant \sqrt{\frac{N}{a}}, \quad |y| \leqslant \sqrt{\frac{N}{b}}.$$

If the equation $ax^2 + by^2 = n$ has an integer solution for each $n \leqslant N$, then there exist N pairs (x, y) such that $0 \leqslant x \leqslant \sqrt{\frac{N}{a}}$, $0 \leqslant y \leqslant \sqrt{\frac{N}{b}}$. We obtain

$$N \leqslant \frac{N}{\sqrt{ab}}.$$

Clearly, if $ab > 1$, then this inequality is not held for N great enough. So it remains to consider the case $a = b = 1$. But if n is equivalent to 3 modulo 4, then it cannot be represented in form $x^2 + y^2$, which ends the proof. □

Problem 9. Prove that for every integer m the equation $x^2 + 2y^2 - 3z^2 = m$ has an integer solution.

Proof. It is enough to show that $x^2 + 2y^2 - 3z^2$ represents 0, every odd number and every number equivalent to 2 modulo 4.

If $x = y = z = 1$, then $x^2 + 2y^2 - 3z^2$ equals 0. Hence, $x^2 + 2y^2 - 3z^2$ represents 0.

If $x = u + 1$, $y = u$, $z = u$, then $x^2 + 2y^2 - 3z^2$ equals $2u + 1$. Hence, $x^2 + 2y^2 - 3z^2$ represents all the odd numbers.

If $x = u$, $y = u + 1$, $z = u$, then $x^2 + 2y^2 - 3z^2$ equals $4u + 2$. Hence, $x^2 + 2y^2 - 3z^2$ represents all the numbers equivalent to 2 modulo 4.

If m is divisible by 4, then we factor it out and reduce the problem to one of the already considered cases. □

Quadratic forms

Problem 10. Describe all integers which are represented by forms a) $x^2 + y^2$; b) $x^2 - y^2$; c)* $x^2 + xy + y^2$.

Proof. a) $n = x^2 + y^2$ if and only if in the factorization of n into primes, every prime divisor entering in n in odd power, is equivalent to 1 modulo 4.

b) $(u + 1)^2 - u^2 = 2u + 1$. Hence, $x^2 - y^2$ represents all the odd numbers. Also $(u + 1)^2 - (u - 1)^2 = 4u$.

We see that $x^2 - y^2$ represents all the integers equivalent to 0, 1, and 3 modulo 4. Analyzing this form modulo 4, we see that residue 2 cannot be represented .

c) Let us fix n. Analogously the proof of Problem 6, we can show that the equation $x^2 + xy + y^2 = n$ has integer solutions if and only if it has rational solutions. In rational numbers, $x^2 + xy + y^2$ is linearly equivalent to $x^2 + 3y^2$ ($x^2 + xy + y^2 = (x + \frac{y}{2})^2 + 3(\frac{y}{2})^2$). We show here (read the section about the Hilbert symbol!), that $x^2 + 3y^2$ represents n in rational numbers if and only if every prime entering in n in odd power is equivalent to 0 or 1 modulo 3.

Indeed, $x^2 + 3y^2 = n$ has solutions in $\mathbb{Q}$ if and only if $x_1^2 + 3y_1^2 - nz^2 = 0$ has solutions in $\mathbb{Z}$ with nonzero z. By the Minkowski-Hasse theorem, this equation has solutions if and only if the Hilbert symbol $(n, -3)_p$ equals 1 for every prime p. Let us find it.

Consider $p > 3$. Let $n = p^\alpha \cdot u$, $3 = p^0 \cdot (-3)$. Using the formula for the Hilbert symbol and the quadratic reciprocity law (Serre, Chapter 1, § 3, Theorem 6) we get that

$$(n, -3)_p = \left(\frac{-3}{p}\right)^\alpha = \left(\frac{-1}{p}\right)^\alpha \left(\frac{3}{p}\right)^\alpha = \left(\frac{-1}{p}\right)^\alpha \cdot \left((-1)^{\frac{p-1}{2}} \left(\frac{p}{3}\right)\right)^\alpha,$$

so it always equals 1 for even α, and for odd α it equals $\left(\frac{p}{3}\right)$, i.e., equals 1 if and only if p has residue 1 modulo 3. So, if the equation

$$x_1^2 + 3y_1^2 - nz^2 = 0$$

has solutions modulo p, for p of the form $3k + 2$, then p enters in pair degree into the decomposition of n into primes.

The case $p = 2$ is let to the reader.

Consider $p = 3$. Let $n = 3^\alpha \cdot u$, here $\beta = 1$, $v = -1$. We have:

$$(n, -3)_3 = (-1)^\alpha \left(\frac{u}{3}\right) \left(\frac{-1}{3}\right)^\alpha = \left(\frac{u}{3}\right).$$

This expression equals 1 if and only if u has residue 1 modulo 3. But we have already checked that all the prime divisors of the form $3k + 2$ enter in even degree, so this condition gives nothing new. □

Definition 1. Two quadratic forms are called *equivalent* if they represent the same set of numbers.

Problem 11. Prove that the quadratic forms

$$f(x, y), \qquad f(x - y, y), \qquad f(x, y - x), \qquad f(-x, y), \qquad f(x, -y) \tag{3}$$

are equivalent.

Proof. If m is represented by the form $f(x, y)$ for $x = x_0$, $y = y_0$, then m can be represented by the form $f(x - y, y)$ for $x = x_0 + y_0$, $y = y_0$, by the form $f(x, y - x)$ for $x = x_0$, $y = y_0 + x_0$, by the form $f(-x, y)$ for $x = -x_0$, $y = y_0$, by the form $f(x, -y)$ for $x = x_0$, $y = -y_0$. Hence, every integer which can be represented by the form $f(x, y)$ can also be represented by any other form from the list (3). One can analogously prove that every integer represented by one of these forms can be represented also by any other form from the list (3). We obtain that all the forms (3) are equivalent. □

Problem 12. a) Prove that the forms $x^2 + y^2$ and $x^2 + xy + y^2$ are not equivalent.

b) Prove that the form $4x^2 - 6xy + 5y^2$ is not equivalent to any form $ax^2 + by^2$ with integer a and b.

Proof. a) The form $x^2 + y^2$ represents 2, while $x^2 + xy + y^2$ not. Hence, they are not equivalent.

b) The form $4x^2 - 6xy + 5y^2$ has a unique well, and the values around it equal 3, 4, and 5. Hence, 3, 4, and 5 are three minimal values of the form $4x^2 - 6xy + 5y^2$.

Let $a, b \geqslant 0$. Then three minimal values of the form $ax^2 + by^2$ can be the following sets of numbers:

$$\{a,\ b,\ a + b\}, \quad \{a,\ 2a,\ b\}, \quad \{a,\ b,\ 2b\}, \quad \{a,\ 2a,\ 4a\}, \quad \{b,\ 2b,\ 4b\}. \tag{4}$$

Clearly, we cannot find a and b to represent the set $\{3, 4, 5\}$ in any of forms (4).

We conclude that there do not exist nonnegative integers a and b such that the form $4x^2 - 6xy + 5y^2$ is equivalent to $ax^2 + by^2$. □

Problem 13. Provide an example of a non-negative definite form which is not positive definite.

Proof. Example: $f(x, y) = x^2$. □

Extended arithmetics: p-adic numbers

Problem 14. Let m and n be square-free integers. Assume that the equation

$$z^2 - mx^2 - ny^2 = 0 \tag{5}$$

has a nontrivial rational solution. Prove that

a) either m or n is positive,

b) m is a quadratic residue modulo n,

c) n is a quadratic residue modulo m.

Proof. We fix a nonzero rational solution (x_0, y_0, z_0) of equation (5). Let us assume that x_0, y_0, z_0 have no common divisor greater than 1.

a) If $m, n \leqslant 0$, then $x_0^2 - my_0^2 - nz_0^2 \geqslant 0$, and the equality can be obtained only for $x_0 = y_0 = z_0 = 0$. We get a contradiction.

b) It is enough to show that for every prime divisor p of m, the integer n is a quadratic residue modulo p.

Fix a prime divisor p of m. If $n \vdots p$, then there is nothing to prove. Now let $n \not\vdots p$. Consider two cases: $y_0 \vdots p$ and $y_0 \not\vdots p$.

First let $y_0 \vdots p$. Then x_0 and z_0 are also divisible by p, which contradicts our assumption that

$$\gcd(x_0, y_0, z_0) = 1.$$

Hence, $y_0 \not\vdots p$. Then the following equivalence is true modulo p:

$$n \equiv (z/y)^2 \pmod{p},$$

which ends the proof of b).

The proof of c) is analogous. □

Problem 15. Reduce Metatheorem for the equations in two variables to the case of equations of the form (5).

Proof. Every quadratic equation has form

$$f(X_1, X_2) = f_2(X_1, X_2) + f_1(X_1, X_2) + f_0 = 0,$$

where f_2 is a homogeneous polynomial of degree 2, f_1 of degree 1, f_0 is a constant.

First let us prove a general statement: either both

$$f(X_1, X_2) = 0 \text{ and } f(X_1 + cX_2 + t, X_2) = 0$$

have rational solutions, or have no rational solutions for all the pairs of rational numbers (c, t). We leave the proof of this fact as an exercise.

Clearly, the following changes of variables

$$f(X_1, X_2) \to f(X_1 + cX_2, X_2) \tag{6}$$

change f_1 and f_2 independently and preserve f_0.

Let us present f_2 in the form

$$c_1X_1^2 + c_{12}X_1X_2 + c_2X_2^2,$$

where c_1, c_2, and c_{12} are parameters.

If $f_2 \neq 0$, then we can perform several changes of the form (6) and assume that $c_1 \neq 0$.

Consider the function

$$f(X_1 - \frac{c_{12}}{2c_1}X_2, X_2). \tag{7}$$

It is easy to see that (7) has form

$$c_1X_1^2 + c_2'X_2^2$$

for some rational number c_2'. Hence, we may assume that

$$f_2(X_1, X_2) = c_1X_1^2 + c_2X_2^2$$

for some rational numbers c_1, c_2. If $c_2 = 0$, but $c_1 \neq 0$, then the equation $f = 0$ can be written as

$$c_1X_1^2 = -rX_2 - f_0$$

which can be easily solved. So from now on we assume that $c_1 \neq 0$. Analogously, we may assume that $c_2 \neq 0$.

The linear part $f_1(X_1, X_2)$ has the form $r_1X_1 + r_2X_2$. Consider the following change of variables:

$$f(X_1, X_2) \to f\left(X_1 - \frac{r_1}{2c_1}, X_2 - \frac{r_2}{2c_2}\right).$$

If now we expand the function $f(X_1 - \frac{r_1}{2c_1}, X_2 - \frac{r_2}{2c_2})$, we obtain that its linear part f_1 equals 0. Now the equation $f = 0$ takes the form

$$c_1X_1^2 + c_2X_2^2 + f_0 = 0.$$

This equation is equivalent to a homogeneous equation

$$z^2 + \frac{c_2}{c_1}y^2 + \frac{f_0}{c_1}z^2 = 0.$$

This ends the proof. □

Problem 16. Let f be a polynomial with integer coefficients. Prove that the equation $f = 0$ has a solution in $\mathbb{Z}_p$ if and only if it has a solution modulo p^n for every positive integer n.

Proof. Let $x_1, \ldots, x_n, \ldots$ be the solution of the equation $f = 0$ in $\mathbb{Z}_p$. Then for every n the residue of x_n modulo p^n is a solution of the equivalence modulo p^n. In particular, $f \equiv 0$ has a solution modulo p^n for every positive integer n.

Now prove the other implication. Suppose that the equation $f \equiv 0$ has a solution modulo p^m for every positive integer m. For every m, we denote by S_m the set of solutions of the equation $f \equiv 0$ modulo p^m. By our assumption, S_m is non-empty for every $m \geqslant 0$.

Since every residue modulo p^{m+1} can be treated as a residue modulo p^m, we have a projection $S_{m+1} \to S_m$. Let us denote by S_m^∞ the intersection of images of S_{m+k} for all $k \geqslant 0$. Since $S_{m+k} \neq 0$, the set $S_m^\infty \neq 0$. For every $s_m \in S_m^\infty$ there exists a $s_{m+1} \in S_{m+1}^\infty$ such that s_m is the image of s_{m+1} with respect to the projection defined above. Proceeding in such a way, we can construct an infinite chain

$$s_1, \ldots, s_m, \ldots, \tag{8}$$

where s_m is a set containing n residues modulo p^m and s_m is the projection of s_{m+1} to the residues modulo p^m. The sequence (8) defines the unique set of n p-adic integers $x_1, \ldots, x_n$, having the prescribed sets of residues $s_1, \ldots, s_m, \ldots$ modulo $p, \ldots, p^m, \ldots$.

The numbers $x_1, \ldots, x_n$ are the solutions of the equation $f = 0$. □

Problem 17. When is a p-adic number in the form (2) equal to 0?

Proof. The answer follows from the definition: If $a_{-k} + \ldots + a_{-k+i}p^i \equiv 0 \,(\mathrm{mod}\ p^{i+1})\ \ \forall i$. □

Problem 18. Prove that the product of two nonzero p-adic numbers is also nonzero.

Proof. Consider two non-zero p-adic numbers a, b. We assume without loss of generality that $a, b \in \mathbb{Z}_p$ and $a, b \not\equiv 0 (\mathrm{mod}\ p)$. But it means that $ab \not\equiv 0 (\mathrm{mod}\ p)$, hence, $ab \neq 0$. □

Problem 19. Prove that $\mathbb{Q} \subset \mathbb{Q}_p$ for any prime p (i.e., prove that for every pair of nonzero integers m and n there exists a p-adic number x such that $nx = m$).

Proof. Without loss of generality we may assume that m, n are coprime with p. But now the statement of Problem 16 follows from Problem 20. □

Problem 20. Prove that -1 is a square in $\mathbb{Q}_p$ if and only if $p \equiv 1 \pmod 4$.

Proof. It follows from Problem 21. □

Problem 21. Find a description of all perfect squares in $\mathbb{Q}_p$.

Proof. We consider two cases $p = 2$ and $p \neq 2$ separately.

First consider $p = 2$. Every 2-adic number x can be represented as $2^n(2m + 1)$, where n is an integer, and m is an integer 2-adic number. We have $x^2 = 2^{2n}(1 + 8\frac{m(m+1)}{2})$. Let $m' = \frac{m(m+1)}{2}$. Then

$$x^2 = 2^{2n}(1 + 8m'), \tag{9}$$

where m' is a 2-adic integer.

Let us prove that every 2-adic integer of the form (9) is a perfect square in 2-adic numbers. It is sufficient to show that every 2-adic integer m' can be represented in the form $\frac{m(m+1)}{2}$.

By Problem 16, is is enough to show that the equivalence $x(x + 1) \equiv 2m'$ has solutions in $\mathbb{Z}$ for every $i \in \mathbb{Z}_{\geqslant 0}$. We will prove this by induction.

Base $i = 1$ is true.

The step of the induction: $i \to i + 1$. Let $m_i \in \mathbb{Z}$ be a solution of the equation $x(x + 1) \equiv 2m' (\mathrm{mod}\ 2^i)$. There are two possibilities:

1) $m_i(m_i + 1) \equiv 2m' \pmod{2^{i+1}}$,

2) $m_i(m_i + 1) \equiv 2m' + 2^i \pmod{2^{i+1}}$.

In Case 1), m_i is also the solution of the equation $x(x + 1) \equiv 2m' \pmod{2^{i+1}}$. In Case 2), $m_i + 2^i$ is the solution of the equation $x(x + 1) \equiv 2m' \pmod{2^{i+1}}$.

Now let $p \neq 2$, i.e., p is an odd prime. Every p-adic number x can be represented in the form $p^n m$, where n is an integer and m is a p-adic integer which is not divisible by p. We have $x^2 = p^{2n} m^2$. Let $m' = m^2$. Then

$$x^2 = p^{2n} m', \tag{10}$$

where m' is a p-adic integer such that its residue modulo p is a nonzero quadratic residue.

Let us prove that every p-adic number of the form (9) is a perfect square in p-adic numbers. It is enough to show that every p-adic integer m' such that

3) m' is not divisible by p,

4) the residue of m' modulo p is a nonzero quadratic residue,

can be represented in the form m^2.

Using Problem 16, it is enough to prove that the equation $x^2 \equiv m' \pmod{p^i}$ has solutions in $\mathbb{Z}$ for every $i \in \mathbb{Z}_{\geqslant 0}$. We prove this statement by induction.

Base $i = 1$ is fulfilled since the residue of m' modulo p is a quadratic residue.

Step: $i \to i + 1$. Let $m_i \in \mathbb{Z}$ be the solution of the equation $x^2 \equiv m' \pmod{p^i}$. Then

$$m_{i+1}^2 \equiv m' + rp^i \pmod{p^{i+1}},$$

for some $r \in \mathbb{Z}$. Since m' is not divisible by p and $p \neq 2$, there exists a $r' \in \mathbb{Z}$ such that $2m_i r' \equiv r \pmod{p^{i+1}}$. Let $m_{i+1} := m_i - r'p^i$. Then $m_{i+1}^2 \equiv m' \pmod{p^{i+1}}$, which ends the proof of the induction step. □

Problem 22. Prove that for any nonzero 3-adic number m there exists a 3-adic number x such that m is equal to one of the numbers x^2, $2x^2$, $3x^2$, or $6x^2$.

Proof. Note that for every p, any p-adic number can be represented in the form $p^i \cdot a \cdot y$, where a is an integer from 1 to $p-1$, p^i is a power of p, and y is a p-adic integer, equivalent to 1 modulo p (by Problem 21, it is a perfect square). In our case $p = 3$, and we obtain that, depending on parity of i, p^i is either a perfect square or 3 times a perfect square, $a \in \{1,\ 2\}$, and y is a perfect square. Clearly, their product has the required form. □

Problem 23. Let p be an odd prime, and let $x_1, \ldots, x_5$ be nonzero p-adic numbers. Prove that there exist indices i and j with $1 \leqslant i < j \leqslant 5$ such that x_i/x_j is a perfect square in $\mathbb{Q}_p$.

Proof. As in the proof of the previous problem, every p-adic number has the form $p^i \cdot a \cdot y$. Divide the given numbers into 2 groups: in the first group i is even and in the second odd. Now divide each group into two smaller groups depending on whether a is a quadratic residue or not. By pigeonhole principle, since we have 5 numbers and 4 groups, there are two numbers in the same group. Their ratio has the form

$$p^j \cdot a_{new} \cdot \frac{y_1}{y_2},$$

where j is even, a_{new} is a quadratic residue (since it is the ratio of either two quadratic residues or two quadratic non-residues), and $\frac{y_1}{y_2}$ is a p-adic number starting with 1. So this ratio is a perfect square as a product of three perfect squares. □

Problem 24. Prove that for every odd prime p there exist p-adic numbers $x_1, \ldots, x_{p-1}$ such that $x_1^2 + \ldots + x_{p-1}^2 + 1 = 0$.

Proof. By Problem 21, $1-p$ is a perfect square in p-adic numbers. Hence, $-1 = 1+1\ldots+1$ ($p-2$ ones)$+1-p$ is a sum of $(p-1)$ perfect squares in p-adic numbers. □

Problem 25. Prove that the equation $x^2 + x + 1 = 0$ has exactly two solutions in $\mathbb{Z}_7$.

Proof. The solutions of the $x^2+x+1 = 0$ can be found with the help of the standard formula $x_{1,2} = \frac{-1 \pm \sqrt{-3}}{2}$. Since $\sqrt{-3} \in \mathbb{Z}_7$ (see Problem 21), the equation $x^2 + x + 1 = 0$ has two distinct solutions in $\mathbb{Z}_7$. □

Problem 26. Prove that the equation $x^2 + y^2 = -1$ has a p-adic solution for every odd prime p.

Proof. First solution. It is enough to show that this equation can be solved modulo p. Indeed, x^2 can take $(p+1)/2$ different values: zero and all the quadratic residues. Now write down the list of all the values of $-x^2-1$. If any of them has the form y^2, we are done. But if no one of them has the form y^2, then there are at most $(p-1)/2$ possible values for y^2, a contradiction.

Second solution. It is enough to show that this equation can be solved modulo p. Every quadratic residue can be represented as x^2+y^2. If only quadratic residues can be represented in the form x^2+y^2, then for every i we show by induction that only quadratic residues can be represented in the form $x_1^2+\ldots+x_i^2$. But by Problem 24 every residue modulo p can be represented in the form $x_1^2+\ldots+x_{p-1}^2$. Hence, x^2+y^2 represents also quadratic non-residues. It means that x^2+y^2 represents all the elements of $\mathbb{Z}/p$, in particular, -1. □

Problem 27. Prove the Hasse–Minkowski principle for equations in one and two variables.

Proof. A) Equations in one variable. Any such equation has the form $ax^2=b$. It is enough to show that if it has no solutions in $\mathbb{Q}$, then either it has no solutions in $\mathbb{R}$, or it has no solutions in $\mathbb{Q}_p$ for some p. If $ax^2=b$ has no solutions in $\mathbb{Q}$, then b/a is not a perfect square in $\mathbb{Q}$, i.e., either $b/a<0$, or there exists a prime number p which enters in b/a in odd power. In the first case $ax^2=b$ has no solutions in $\mathbb{R}$, in the second in $\mathbb{Q}_p$.

B) Equations in two variables. By Problem 15, every equation (of degree two!) in two variables in rational numbers is equivalent to the equation $ax^2+by^2=1$. We may assume that:

1) the coefficients a, b are square-free integers,

2) $|a|\leqslant|b|$.

It is enough to show that if the equation $ax^2+by^2=1$ has a solution in $\mathbb{Q}_p$ for every p and in $\mathbb{R}$, then it has solutions in $\mathbb{Q}$. Let $m(a,b):=|a|+|b|$. We prove this statement by induction on $m(a,b)$.

Base $m(a,b)=2$ can be verified directly.

Step: $m\to m+1$. Let a, b be some integers verifying the conditions 1), 2) and such that

• $m(a,b)=m+1$

• the equation $ax^2+by^2=1$ has solutions in p-adic numbers for every p and has a solution in $\mathbb{R}$.

Consider two cases: $|a|=|b|$ and $|a|<|b|$. If $|a|=|b|$, then the equation $ax^2+by^2=1$ is equivalent to the equation

$$-(b/a)y^2+az^2=1. \tag{11}$$

Moreover, the equation (11) has solutions in $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{Q}_p$, if and only if the equation $ax^2+by^2=1$ has solutions in the same set. Since $m(-\frac{b}{a},a)<m(a,b)=m+1$, the equation $-(b/a)y^2+az^2=1$ has solutions in $\mathbb{Q}$ by the induction assumption. Hence, the initial equation has solutions in $\mathbb{Q}$.

Now let $|a|<|b|$. It follows from the condition • that a is a perfect square modulo b, i.e.,

$$a+bb'=t^2,$$

where b', t are some integers and $b'\geqslant 0$. We assume without loss of generality that

$$|t|\leqslant\frac{|b|}{2}.$$

The equation $ax^2+by^2=1$ has solutions in $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{Q}_p$ if and only if the equation

$$ax^2+b'y^2=1, \qquad b'=\frac{t^2-a}{b}$$

has solutions in the same set. We have $|b'|\leqslant\frac{|b|}{4}$ and $m(a,b')<m(a,b)=m+1$. Since $m(a,b')\leqslant m$, the equation $ax^2+b'y^2=1$ has solutions in $\mathbb{Q}$ by the induction hypothesis. Hence, the initial equation also has solutions in $\mathbb{Q}$. □

Problem 28. Prove the following properties of the Hilbert symbol:

1) $(a,b)_p=(b,a)_p$,

2) $(a,c^2)_p=1$,

3) $(a,-a)_p=1$, $\quad(a,1-a)_p=1$,

4) $(a,b)_p=(a,-ab)_p=(a,(1-a)b)_p$.

Proof. 1) Obvious from the definition.
2) The equation $z^2 - ax^2 - c^2y^2 = 0$ has a nonzero solution $z = c, x = 0, y = 1$.
3) The equation $z^2 - ax^2 + ay^2 = 0$ has a nonzero solution $z = 0, x = y = 1$.
The equation $z^2 - ax^2 - (1-a)y^2 = 0$ has a nonzero solution $x = y = z = 1$.
4) Using Problem 29, one can deduce it from 3). □

Problem 29. Let $(a, b)_p = 1$. Show that $(a', b)_p = (aa', b)_p$ for any a'.

Proof. Let b be a perfect square. Then $(a, b)_p = (aa', b)_p = 1$.
Now suppose that b is not a perfect square. We need the following lemma.

Lemma 1. If
- b is not a perfect square,
- $(a, b)_p = 1$ and $(a', b)_p = 1$,

then $(aa', b) = 1$.

Finish the proof of Problem 29 using Lemma 1. If $(a', b)_p = 1$, then by Lemma 1 we have $(aa', b)_p = 1$. If $(aa', b)_p = 1$, then by Lemma 1 $(a', b)_p = (a^2a', b)_p = 1$. Hence, if one of the numbers $(aa', b)_p$ and $(a', b)_p$ equals 1, then the second one either equals 1. Hence, they are equal.

Proof of Lemma 1. Let (x_0, y_0, z_0) be a nonzero solution of the equation $z_0^2 - ax_0^2 - by_0^2 = 0$. Since b is not a perfect square in $\mathbb{Q}_p$, $x_0 \neq 0$. So we may assume that $x_0 = 1$ and $a = z_0^2 - by_0^2$. By the similar reasoning, there exist z_1, y_1 such that $a' = z_1^2 - by_1^2$. Then

$$aa' = (z_0z_1 - by_0y_1)^2 - b(z_0y_1 + z_1y_0)^2.$$

Hence, $(aa', b) = 1$. □

□

To write down an expression for the Hilbert symbol in a compact form, we will use the *Legendre symbol* $\left(\frac{x}{p}\right)$ defined for any integer x and prime p. It equals to 1, -1, or 0 depending on whether x is a nonzero quadratic residue, a quadratic non-residue, or zero.

Problem 30. Let p be an odd prime; let $a = p^\alpha u$, $b = p^\beta v$, where α, β, u, v are integers such that u and v are not divisible by p. Prove that

$$(a, b)_p = (-1)^{\alpha\beta\cdot\varepsilon(p)} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha,$$

where $\varepsilon(p) = \frac{p-1}{2}$.

Proof. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 3, § 1, Theorem 1. □

Problem 31. Find an explicit formula for $(a, b)_2$ for every nonzero integers a and b.

Proof. Let $a = 2^\alpha u$, $b = 2^\beta v$, where α, β, u, v are integers such that u and v are odd. The Hilbert symbol $(a, b)_2$ is given by the formula

$$(-1)^{\varepsilon(u)\varepsilon(v)+\alpha\omega(v)+\beta\omega(u)},$$

where $\varepsilon(u) = \frac{u-1}{2}$, and $\omega(u) = \frac{u^2-1}{8}$. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 3, § 1, Theorem 1. □

Problem 32. Prove that $(a, b)_p(a, b')_p = (a, bb')_p$ for every nonzero integers a, b, b'.

Proof. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 3, § 1, Theorem 2. □

Problem 33. Prove that the equation $ax^2 + by^2 = c$ in variables x and y (with parameters a, b, and c) has a solution in p-adic numbers if and only if $(c, -ab)_p = (a, b)_p$.

Proof. If the equation $ax^2 + by^2 = c$ has a solution, then the equation

$$z^2 - \frac{a}{c}x^2 - \frac{b}{c}y^2 = 0$$

also has a solution. By definition, this means that $(a/c, b/c)_p = 1$. We can re-write it as follows.

$$1 = (a/c, b/c)_p = (a, b)_p(a, c)_p(b, c)_p(c, c)_p = (a, b)_p(ab, c)_p(c, -1)_p. \tag{12}$$

So $(a, b)_p = (c, -ab)_p$.

Now let $(a, b)_p = (c, -ab)_p$ and show that the equation $ax^2 + by^2 = c$ has a non-zero solution. It follows from (12) that $(a/c, b/c)_p = 1$. Hence, the equation $z^2 - \frac{a}{c}x^2 - \frac{a}{b}y^2 = 0$ has a solution. Let us denote by (x_0, y_0, z_0) one of its solutions. If $z_0 \neq 0$, then $(\frac{x_0}{z_0}, \frac{y_0}{z_0})$ is a solution of the equation $ax^2 + by^2 = c$.

We have solved the problem if $z_0 \neq 0$. Now we assume that $z_0 = 0$. For any r_x, r_y consider the equation

$$a(tx_0 + r_x)^2 + b(ty_0 + r_y)^2 = c.$$

It is equivalent to the equation

$$(ar_x^2 + br_y^2) + 2t(ax_0r_y + by_0r_x) = c. \tag{13}$$

For a generic pair of rational numbers (r_x, r_y), we see that $(ax_0r_y + by_0r_x) \neq 0$ and $t_0 = \frac{c-(ar_x^2+br_y^2)}{2(ax_0r_y+by_0r_x)}$ is a solution of the equation (13). Hence, the equation $ax^2 + by^2 = c$ has infinitely many rational solutions. □

Using the properties of the Hilbert symbol, solve the following problem.

Problem 34*. Let us fix a homogeneous polynomial $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$ with $n \geqslant 2$, where $a_1, \ldots, a_n \neq 0$. Set

$$d = a_1a_2 \ldots a_n \quad \text{and} \quad \varepsilon = \prod_{i<j}(a_i, a_j)_p. \tag{14}$$

Prove that the equation $f = 0$ has a nonzero p-adic solution if and only if one of the following conditions is satisfied:

1) $n = 2$ and $-d$ is a square in $\mathbb{Q}_p$;
2) $n = 3$ and $(-1, d)_p = \varepsilon$;
3) $n = 4$ and $d \neq \alpha^2$, or $d = \alpha^2$ and $\varepsilon = (-1, -1)_p$;
4) $n \geqslant 5$. (i.e., if f depends on 5 or more variables, then $f = 0$ has a nonzero solution in $\mathbb{Q}_p$ for any p.)

Proof. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 4, § 2, Theorem 6. □

Deduce the following problem from problem 34.

Problem 35. Fix a homogeneous polynomial $f = a_1x_1^2 + a_2x_2^2 + \ldots + a_nx_n^2$, where $a_1, \ldots, a_n \neq 0$, and an integer $a \neq 0$. Define d, ε by formula (14). Then the equation $f = a$ has a p-adic solution if and only if one of the following conditions is satisfied:

1) $n = 1$ and a/d is a square in $\mathbb{Q}_p$;
2) $n = 2$ and $(a, -d)_p = \varepsilon$;
3) $n = 3$ and ad is not a perfect square in $\mathbb{Q}_p$, or ad is a perfect square and $\varepsilon = (-1, -d)_p$;
4) $n \geqslant 4$. (i.e., if f depends on 4 or more variables, then the equation $f = a$ has a nonzero solution in $\mathbb{Q}_p$ for any p.)

Proof. 1) $a_1x_1^2 = a \Leftrightarrow x_1^2 = \frac{a}{a_1}$. Obviously, it has solutions in p-adic numbers if and only if $\frac{a}{a_1}$ is a perfect square in $\mathbb{Q}_p$.

2) $a_1x_1^2 + a_2x_2^2 = a$. The condition $(a, -d)_p = \varepsilon$ is equivalent to $(a, -a_1a_2)_p = (a_1, a_2)$ which is just Problem 33 for $a = a_1$, $b = a_2$, $c = a$.

3) We are solving the equation $a_1x_1^2 + a_2x_2^2 + a_3x_3^2 - a = 0$. Obviously, it is equivalent to the equation $a_1x_1^2 + a_2x_2^2 + a_3x_3^2 - ax_4^2 = 0$ because of transformation

$$(x_1, x_2, x_3, x_4) \leftrightsquigarrow (\frac{x_1}{x_4}\frac{x_2}{x_4}\frac{x_3}{x_4}, 1),$$

which is solvable in p-adic numbers.

Now we prove that if there is a nontrivial solution with $x_4 = 0$, then there exists a nontrivial solution with $x_4 \neq 0$. Without loss of generality $x_1 \neq 0$. Let (C, D) be a solution of the equation $C^2 - D^2 = -\frac{a}{a_1}$ (for example, $C = \frac{1-a/a_1}{2}$, $D = \frac{-1-a/a_1}{2}$)).

Obviously, we can multiply our solution to $\frac{C}{x1}$, and we get $(C,\ x_2,\ x_3,\ 0)$. It is easy to check that

$$f(C, x_2, x_3, 0) = f(D, x_2, x_3, 1),$$

so we reduced this problem to Problem 34c.

First case: $ad \neq -m^2$ in $\mathbb{Q}_p$. It is equivalent to $d \neq m^2$.

Second case: $ad = -m^2$.

We need: $(a_1, a_2)_p(a_1, a_3)_p(a_2, a_3)_p = (-1, -a_1a_2a_3)_p \Leftrightarrow (a_1, a_2)_p(a_1, a_3)_p(a_2, a_3)_p(-a, d) = (-1, -1)_p$.

Obviously, if we have a problem like $a = b \Leftrightarrow c = d$ where (a, b, c, d) from $\{1, -1\}$, then we need to prove $ac = bd$, so we need

$$(-a, d)_p = (-1, -d)_p(-1, -1)_p \quad \Leftrightarrow \quad (-a, d)_p = (-1, -1)_p(-1, -1)_p(-1, d)_p \quad \Leftrightarrow$$
$$\Leftrightarrow \quad (-a, d)_p = (-1, d)_p \quad \Leftrightarrow \quad (-1, d)_p(-a, d)_p = 1 \quad \Leftrightarrow \quad (a, d)_p = 1 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad (a, a)_p(a, \frac{d}{a})_p = 1 \quad \Leftrightarrow \quad 1 \cdot 1 = 1,$$

because $\frac{d}{a}$ is a perfect square in $\mathbb{Q}_p$.

4) As in 3), we only need to solve the equation

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + a_4x_4^2 - ax_5^2 = 0.$$

But by Problem 34d this equation is always solvable.

Also, the proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 4, § 2, Corollary of Theorem 6. □

Problem 36. Prove the Hasse–Minkowski principle.

Proof. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 4, § 3, Theorem 8. □

Problem 37. Using problem 35 and the Hasse–Minkowski principle, show that an integer d is a sum of 3 squares in rational numbers if and only if the number d cannot be represented in the form $4^a(8b - 1)$, i.e. if $-d$ is not a perfect square in $\mathbb{Q}_2$.

Proof. By the Hasse-Minkowski principle, we only need to consider whether the equation $x^2 + y^2 + z^2 = n$ has a p-adic solution or not. We work in terms of Problem 35. $a_1 = a_2 = a_3 = 1$, $d = 1$, $\varepsilon = (1, 1)_p^3 = (1, 1)_p$, $a = n$. First, we prove that if $p > 2$, then it is solvable in p-adic numbers.

If n is not a $-m^2$ then we are done. If $n = -m^2$ then $\varepsilon = (1, 1)_p = [Problem30] = 1 = [Problem30] = (-1, -1)_p$, so a solution exists.

It remains to consider the case $p = 2$. If $n \neq -m^2$ then we are done. Now let $n = -m^2$. If a solution exists, then

$$\varepsilon = (1, 1)_2 = (-1, -1)_2,$$

which leads to a contradiction by Problem 31. □

Problem 38. Fix an integer n. Prove that if there exist rational numbers x, y, and z such that $x^2+y^2+z^2 = n$, then there also exist integers x', y', and z' such that $(x')^2 + (y')^2 + (z')^2 = n$. Deduce the Gauss theorem from this statement.

Proof. Let rational numbers x, y, z be such that $x^2 + y^2 + z^2 = n$. Denote by d the common denominator of x, y, z. Choose a triple (x, y, z) with the minimal value of d. Let us assume that $d > 1$ (i.e., that one of x, y, and z is not integer and that the equation $x^2 + y^2 + z^2 = n$ has no integer solutions). Let r_x, r_y, r_z be the integers closest to x, y, z, and let $s_x := x - r_x$, $s_y := y - r_y$, $s_z := z - r_z$. Then

$$|s_x|, |s_y|, |s_z| \leqslant \frac{1}{2}, \qquad s_x^2 + s_y^2 + s_z^2 = n - (r_x^2 + r_y^2 + r_z^2) - 2(s_xr_x + s_yr_y + s_zr_z). \tag{15}$$

Let

$$x' = r_x - \frac{s_x(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}, \qquad y' = r_y - \frac{s_y(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}, \qquad z' = r_z - \frac{s_z(n - r_x^2 - r_y^2 - r_z^2)}{s_x^2 + s_y^2 + s_z^2}.$$

It follows from (15) that $s_x^2 + s_y^2 + s_z^2 = d'/d$, and moreover $0 < d' < d$. It means that the least common denominator of x', y', z' divides d', i.e., is less than d. Notice that $x'^2 + y'^2 + z'^2 = n$. We get a contradiction. Hence, $d = 1$, and the equation $x^2 + y^2 + z^2 = n$ has integer solutions.

By Problem 37, every positive integer N which does not have the form $4^n(8m - 1)$ is a sum of three squares of rational numbers. But we proved above that in this case N is also a sum of three squares of integers. □

Problem 39. Deduce the Legendre theorem from the Gauss theorem.

Proof. It follows from the Gauss theorem that every positive integer, equivalent to 1, 2, 3, 5, or 6 modulo 8, can be represented as a sum of three (and, consequently, of four) perfect squares. It remains to show that every integer which is equivalent to 0, 4, or 7 modulo 8, can be represented as a sum of 4 perfect squares.

If n can be represented as a sum of four squares, then $4n$ either. So it is enough to consider the case $n \equiv 7 \mod 8$.

Fix n, $n \equiv 7 \mod 8$. Since $n - 1$ is equivalent to 6 modulo 8, by Gauss theorem $n - 1$ can be represented as a sum of three squares. Hence, n can be represented as a sum of four squares. □

Some properties of the Hilbert symbol (DE-2)

The goal of this section is to show that, for a pair of nonzero integers (a, b), the Hilbert symbol $(a, b)_p$ equals 1 for almost all (=all except finite number) primes p. We deduce this statement from a more general statement presented below.

Problem 40. a) Let f be a homogeneous polynomial of degree n, depending on k variables, where $k > n$. Then the number of solutions of the equivalence $f \equiv 0$ (including 0-solution!) modulo p is divisible by p (Hint: apply the little Fermat theorem and consider case $p = 2$).

b) Let f be a polynomial of degree n depending on k variables, where $k > n$. Then the number of solutions of the equivalence $f \equiv 0$ modulo p is divisible by p.

Proof. Obviously, case b) is a generalization of case a). We prove here b). Consider a polynomial $f(x_1, \ldots, x_k)$ of degree n. Consider the following sum:

$$\sum_{x_1, \ldots, x_n} f(x_1, \ldots, x_n)^{p-1}, \tag{16}$$

where $x_1, \ldots, x_n$ run all the residues modulo p. Note that every element of the sum (16) equals 0 or 1 modulo p. The key idea is that the residue modulo p of the number of solutions of the equation $f(x_1, \ldots, x_k) \equiv 0$ equals (16) modulo p. But the degree of the polynomial $f(x_1, \ldots, x_k)^{p-1}$ is $(p - 1)n$. Since we have $k > n$ variables, for every monomial of $f(x_1, \ldots, x_k)^{p-1}$ there exists a variable entering in this monomial in degree less than $p - 1$. But for such a monomial, the summation in (16) gives 0, because

$$\sum_{x_i} x_i^l \equiv 0 \pmod{p}$$

for all $l < p - 1$. Hence, the sum in (16) is equivalent to 0 (mod p), and the number of solutions of the equation $f(x_1, \ldots, x_n) = 0$ has residue 0 modulo p. □

Problem 41. Deduce from the previous problem that for any integers a, b, c the equivalence $ax^2 + by^2 + cz^2 \equiv 0$ in variables x, y, z has a nonzero solution modulo p .

Proof. The polynomial $ax^2 + by^2 + cz^2$ has degree 2 and depends on three variables. Hence, the number of solutions of the equation $ax^2 + by^2 + cz^2 \equiv 0$ has residue 0 modulo p. In particular, this means that the equation $ax^2 + by^2 + cz^2 \equiv 0$ has a nonzero solution. □

Problem 42. Deduce from the previous problem that, for a pair of nonzero integers (a, b) and an odd prime p, $(a, b)_p = 1$ if $a, b \not\vdots p$. Explain why $(a, b)_p = 1$ for all primes p except a finite number.

Proof. Let us fix $p \neq 2$ and relatively prime with a and b. We prove that $(a, b)_p = 1$.

Let (x_0, y_0, z_0) be a solution of the equation $z^2 - ax^2 - by^2 \equiv 0 \pmod p$ such that

$$(x_0, y_0, z_0) \not\equiv (0, 0, 0) \pmod p$$

(by Problem 41, such a solution exist). We assume without loss of generality that $z_0 \not\equiv 0 (\mathrm{mod}\ p)$. Then by Problem 21 the value $ax_0^2 + by_0^2$ is a perfect square in p-adic numbers, hence, the equation $z^2 - ax^2 - by^2 = 0$ has a nonzero solution in p-adic numbers, i.e., $(a, b)_p = 1$. □

Problem 43. Deduce from Problem 41 that the equation $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$ in variables x, y, z, v, w (a, b, c, d, e are parameters) has a nonzero solution in $\mathbb{Q}_p$ for any odd prime p.

Proof. We may assume without loss of generality that a, b, c, d, e are integer and square-free (i.e., every prime divisor enters in the 1st power). First we show that we may assume that no three of a, b, c, d, e have a nontrivial common divisor. Indeed, let p be a prime divisor of three or more parameters.Then we multiply the equation $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$ by p and factor out perfect squares, thus obtaining an equation in which p divides at most 2 of the parameters a, b, c, d, e. Clearly, we may apply this procedure for all the prime divisors of a, b, c, d, e. Finally we may assume that for each prime p, at least three numbers among a, b, c, d, e are not divisible by p.

Fix a prime p. Without loss of generality a, b, c are not divisible by p.

If p is odd, then we use Problem 42: the equation

$$ax^2 + by^2 + cz^2 = 0$$

(and, consequently, the equation $ax^2 + by^2 + cz^2 + dv^2 + ew^2 = 0$) has a non-zero solution in $\mathbb{Q}_p$.

If $p = 2$, then one has to consider case-by-case all the residues of a, b, c, d, e modulo 8. □

Problem 44. Prove that, for any pair of nonzero integers (a, b), we have

$$\prod_p (a, b)_p = (a, b)_{-1},$$

where the product is taken over all primes p and

$$(a, b)_{-1} = \begin{cases} 1, & \text{if the equation } z^2 - ax^2 - by^2 = 0 \text{ has a real solution,} \\ -1 & \text{otherwise.} \end{cases} . \tag{17}$$

Proof. Since the Hilbert symbol is multiplicative (Problem 32), is is enough to check (17) in the case when a, b are prime numbers or -1.

Consider the case $a = b = -1$. Then $(a, b)_p = (-1, -1)_p$ is not 1 only if $p = 2$. But in this case one can directly verify that $(-1, -1)_2 = (-1, -1)_{-1} = -1$.

The next case: $a = -1$, b is a prime number. Then $(a, b)_p = (-1, b)_p$ is not 1 only if $p = b$ or $p = 2$. The direct check shows that $(-1, p)_p = (-1, p)_2$ for $p \neq 2$ and $(-1, 2)_2 = 1$. It means that the left-hand side of (17) equals the right-hand side of (17) and equals 1. □

As a last problem of this list, we mention an "analogue" of the Chinese Remainder Theorem: it turns out that one can construct a rational number with the prescribed values of the Hilbert symbol.

Problem 45. Fix a finite set of nonzero integers a_i and for every prime p define the values $\varepsilon_{i,p} = \pm 1$. Show that the system of equations

$$(a_i, x)_p = \varepsilon_{i,p}, \quad \forall i,\ \forall p,$$

has a solution if and only if

a) almost all (=all except finite number) $\varepsilon_{i,p} = 1$,

b) for any prime p, there exists a nonzero p-adic number x_p such that

$$(a_i, x_p) = \varepsilon_{i,p}.$$

Proof. The proof of this fact can be found in the book "A course in arithmetic" by J.-P. Serre, Chapter 3, § 2, Theorem 4. □

Two variables: maps of quadratic forms (DE-4)

In this section we study the equation

$$E_m: \quad ax^2 + bxy + cy^2 = m \tag{18}$$

depending on integer variables x, y, where a, b, c, m are integer parameters.

Problem 46 (Superproblem)**.** Prove that if the equation E_m has a solution for some positive m, has a solution for some negative m, has no non-trivial solutions for $m = 0$, then for every m either E_m has no solutions, or E_m has infinitely many solutions.

Proof. It follows from the problem statement that $f(x, y) = ax^2 + bxy + cy^2$ is an indefinite quadratic form which does not represent 0. Hence, the map of f splits into a positive domain and a negative domain by a periodic river. Consequently, the map of f is periodic; it means that every value written on the map appears infinitely many times. □

Problem 47 (Superproblem)**.** Is it true that if the equation E_m has solutions for

$$m = \pm 1,\ \pm 2,\ \pm 3,$$

then in this case E_m has solutions for any integer m?

Proof. We give a counterexample $f(x, y) = x^2 + xy - 18y^2$. □

Problem 48 (Superproblem)**.** Prove that if the equations E_1, E_2, E_3, E_5 have integer solutions, then the equation E_m has an integer solution for some $m < 0$.

Proof. Suppose the contrary. Then there are two cases: either f is positive definite or f is non-negative definite. We consider these cases independently.

Let f be non-negative definite. Then $f(x, y) = r(px + qy)^2$ for some integers r, p, q. Since f represents 1, we have $r = 1$. But then f does not represent 5.

Now let f be positive definite. Without loss of generality we assume that

$$f = px^2 + qy^2 + r(x - y)^2$$

for some non-negative integers p, q, r (see Problem 60). The values p, q, r are either all integers or all semi-integers. Without loss of generality we assume that $p \geqslant q \geqslant r$.

The minimal nonzero value of f is $q + r$. Since f represents 1, $q + r = 1$. Hence, either $q = 1$, $r = 0$, or $q = r = \frac{1}{2}$. We consider both these cases.

In the first case $q = 1$, $r = 0$. Since f represents 2, either $p = 1$, or $p = 2$. In the first case f does not represent 3, in the second case f does not represent 5.

In the second case $q = r = \frac{1}{2}$. Then for all positive p we have

$$f(x, y) \geqslant x^2 - xy + y^2.$$

Since f represents 2, $f(x, y) = 2$ for some integers x, y. In particular, $x^2 - xy + y^2 \leqslant 2$. This inequality is held for the following pairs (x, y):

$$(0, 1), \qquad (1, 0), \qquad (1, 1).$$

Since $f(0, 1) = 1$, there are only two cases: $f(1, 0) = 2$, $f(1, 1) = 2$. Let us consider both these cases.

Let $f(1, 0) = 2$. Then $p = \frac{3}{2}$, $f = y^2 - xy + 2y^2$. In this case f does not represent 3.

Let $f(1, 1) = 2$. Then $p = \frac{3}{2}$, $f = y^2 - xy + 2y^2$. In this case f does not represent 3 neither. □

Drawing a map

Problem 49. Prove that, if $\{w_1, w_2\}$ is a basis of $\mathbb{Z}^2$, then pairs

$$\{w_2, w_1\}, \{w_1 - w_2, w_2\}, \{w_1 + w_2, w_2\}, \{-w_1, w_2\} \tag{19}$$

are also bases of $\mathbb{Z}^2$.

Proof. Analogous to the proof of Problem 11. □

Problem 50. Show that, using transformations (19), it is possible to transform any basis to any other one.

Proof. Let $\{u, v\} := \{(a, b), (c, d)\}$ be a basis of $\mathbb{Z}^2$. We show that by transformations (19) we may send every basis $\{u, v\}$ to the basis $\{(1, 0), (0, 1)\}$. Consider the quadratic form $f(x, y) := x^2 - xy + y^2$. By Problem 60, in some basis $\{u', v'\}$ which can be obtained from $\{u, v\}$ via a series of transformations (19), the quadratic form f will be equivalent to a quadratic form of the form

$$\begin{gathered} px^2 + qy^2 + r(x + y)^2, \\ 2p = f(v') + f(u' + v') - f(u') \geqslant 0, \\ 2q = f(u') + f(u' + v') - f(v') \geqslant 0, \\ 2r = f(u' + v') - f(u') - f(v') \geqslant 0. \end{gathered} \tag{20}$$

The least values of the quadratic form (20) are attained for (x, y) equal to

$$(0, 1), \quad (1, 0), \quad (1, 1), \tag{21}$$

and the value on any other pair (x, y) is greater than at least one of these values. Since $x^2 - xy + y^2$ is positive definite, the value (20) is greater than 1 on all the pairs besides the pairs of the list (21). Hence, the value (20) on the pairs (21) is equal to 1. This implies one of the three statements below:

$$\{u', v'\} = \{(0, 1), (1, 0)\}, \qquad \{u', v'\} = \{(0, 1), (1, 1)\}, \qquad \{u', v'\} = \{(1, 0), (1, 1)\}. \tag{22}$$

Hence, the basis $\{u, v\}$ is equivalent to one of the bases (22). □

Problem 51. Show that a quadratic form can have the same representations in several different bases.

Proof. The quadratic form $x^2 - 2y^2$ is the same in bases $\{(3, 2), (4, 3)\}$ and $\{(1, 0), (0, 1)\}$. □

Problem 52. Find a quadratic form which has different representations in any two different bases of $\mathbb{Z}^2$.

Proof. Let $f(x, y) := 2x^2 - xy + 3y^2$, and let us show that for different bases, this form is written differently. Suppose the converse: assume that there exist two different bases such that this quadratic form has the same form in these bases. Let us reconstruct the map of this form step by step, starting from these bases, synchronously. Consider the first moment when the reconstructed domains intersect. Their intersection is either an edge or a vertex. Consider these cases independently.

If this intersection is a vertex, then the map of f is symmetric with respect to one of the edges incident to this vertex. Hence, the only well of the form $2x^2 - xy + 3y^2$ is also symmetric with respect to one of the edges incident to it. But this is not true since the values around the well are 2, 3, and 4.

If this intersection is an edge, then the map of f does not change, if we interchange the vertices of the edge. Hence, the vertices of this edge are wells, in particular, the form $2x^2 - xy + 3y^2$ has two wells. We get a contradiction.

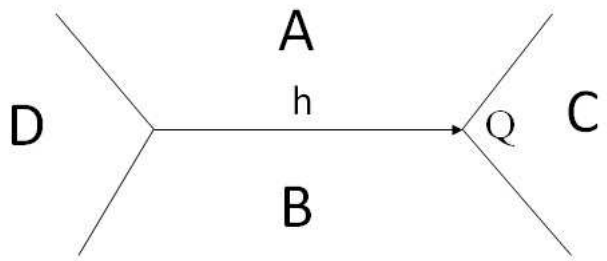
We proved that the quadratic form $2x^2 - xy + 3y^2$ is different in different bases. □

Exercise 1. Write down all the extensions of a basis $\{w_1, w_2\}$. Write down all the specializations of a superbasis $\{\pm w_1,\ \pm w_2,\ \pm(w_1 + w_2)\}$.

Exercise 2. Draw (oriented) maps of the following quadratic forms:

$$f_1 = 3x^2 + 9xy + 7y^2, \quad f_2 = x^2 - 2y^2, \quad f_3 = x^2 - 3y^2.$$

In two problems below, the values A, B, C, D, and h are related to the following picture.



Problem 53. Show that A, B, C, D, and h satisfy

$$C = A + B + h, \qquad D = A + B - h.$$

Proof. The problem statement is equivalent to the following identity for quadratic forms:

$$f(x + y) + f(x - y) = 2(f(x) + f(y)).$$

□

Problem 54. Assume that A, B, C are positive and the edge h goes from C to D. Show that in this case D is also positive and that the arrows on two other edges which are incident to Q go out of Q.

Proof. Since $D, h > 0$, $C = D + 2h > 0$. The values in the regions incident to the edges incident to D equal

$$4A + 2h + B, 4B + 2h + A \ (> A, B). \tag{23}$$

Hence, two remaining arrows go out of D. □

Problem 55. Show that the graph determined by the points-superbases and edges-bases is a tree, i.e., it has no cycles.

Proof. Consider the quadratic form $f(x, y) = x^2 + xy + y^2$. Its map has a unique well Q, and by Problem 54 all the arrows of this map go out of Q. If there were a cycle on this map, then it would be impossible that all the arrows go out of Q. Hence, the map of the quadratic form f does not contain cycles. But the underlying graph does not depend on the quadratic form, so for every f its map is a tree. □

Problem 56. Let Q be a unique well of a positive definite quadratic form f, and p, q, r be integers written in the regions adjacent to Q. Show that the number in any other region of a map related to f is strictly greater than $\max(p, q, r)$.

Proof. We fix a region A such that
a) A is not adjacent to a well,
b) the value of f on A is minimal among all the regions satisfying a).
We are going to prove that $f(A) > p, q, r$. It will end the proof. Let us find the path of the smallest length W from A to the well Q. Since W is the shortest, its last arrow bottoms at A. By the definition of the well all the arrows incident to it go out of it. This together with Problem 54 implies that all the arrows of W are oriented from Q. Hence, the last edge of W is oriented to A. So it follows from the conditions a) and b) that W contains only one edge. The statement of the problem for the regions connected with a well by an edge can be verified directly (see formula (23)). □

Problem 57. Prove that every positive definite form has a well.

Proof. Choose a vertex Q of the map, for which the sum of values of the neighbor regions is minimal. This vertex will be a well (cf. also Problem 53). □

Problem 58. a) Prove that a positive definite form has not more than two wells.
b) Find a positive definite form with two wells.

Proof. b) The form $x^2 + y^2$ is positive definite and has two wells.

a) Let Q be a well, and $p \geqslant q \geqslant r$ be the values written in the regions around it. There are two cases: $q + r > p$ or $q + r = p$. Consider these cases independently.

Let $q + r > p$. Then all the three arrows incident to Q go out of Q. Fix a vertex Q' and assume that it is a well. Let W be the shortest path joining Q and Q'. It follows from the proof of Problem 54 that all the edges of W are directed from Q to Q'. Hence, Q' is not a well.

Now let $q + r = p$. Then the second vertex of the edge E, separating the value q from the value r, is also a well. We denote it by Q'. The collections of values around Q and around Q' coincide. Assume that there exists one more well Q'', and let W be the shortest path, joining Q'' either with Q or with Q'. Then W does not pass through the second one. Without loss of generality we assume that W joins Q with Q''. Then all the edges of W are directed from Q, hence, Q'' is not a well. □

Problem 59. Provide an algorithm which solves the equation $ax^2+bxy+cy^2 = m$ (a, b, c, m are parameters, x, y, z are variables), under the assumption that $ax^2 + bxy + cz^2$ is positive definite.

Proof. After an appropriate change of variables we may assume that $f = px^2 + q(x - y)^2 + ry^2$ for some positive values p, q, r (see Problem 60). If $f(x, y) = n$ has integer solutions, then

$$px^2 \leqslant n, ry^2 \leqslant n. \tag{24}$$

The number of pairs of integers (x, y) for which x, y satisfy (24) is finite. If we check them all, we will detect whether the equation $f(x, y) = n$ has solutions or not. □

Problem 60 (Classification of positive definite quadratic forms)**.**

a) Show that any positive definite quadratic form is equivalent to the form

$$(p + q)x^2 + 2qxy + (q + r)y^2 \tag{25}$$

for some non-negative numbers p, q, r.

b) Show that the quadratic forms corresponding to

$$(p_1, q_1, r_1) \text{ and } (p_2, q_2, r_2)$$

are equivalent if and only if these triples coincide as multisets.

c) Find out which triples (p, q, r) determine an integer quadratic form.

d) Find out which triples (p, q, r) determine a positive definite quadratic form.

Proof. Let Q be a well of f, and m, n, k be the values around this well. Let

$$p = \frac{m + n - k}{2}, q = \frac{m + k - n}{2}, r = \frac{k + n - m}{2}.$$

Then f is equivalent to the form

$$px^2 + qy^2 + r(x - y)^2 = (25).$$

In the case b) there is a counterexample in our notation $x^2 + 3y^2$ и $x^2 + xy + y^2$. In the statement of the problem, "equivalent" must be replaced by "linearly equivalent".

The answer of c): when either p, q, r are integers, or $p - \frac{1}{2}$, $q - \frac{1}{2}$, $r - \frac{1}{2}$ are integers.

The answer of d): the form f is positive definite, if $p, q, r \geqslant 0$ and at least two of numbers p, q, r are nonzero. □

Половинчатые графы

Задачу представляют И. Богданов, К. Кохась

В этой серии задач мы будем иметь дело с *геометрическими графами*, т. е. с графами, изображенными на плоскости. В геометрическом графе никакие три вершины не лежат на одной прямой, а ребра изображаются отрезками (которые могут пересекаться не в вершине). Вот два примера геометрических графов, представляющих для нас основной интерес.

Определение. 1) *Половинчатым* графом $G(n)$ назовем граф, получающийся следующей конструкцией. Пусть на плоскости нарисовано n (четное число) точек, никакие три из которых не лежат на одной прямой. Это вершины графа. Ребро соединяет две вершины, если прямая, содержащая это ребро, делит множество вершин на две равные части (т. е. в ее полуплоскостях поровну оставшихся вершин).

2) Рассмотрим более общую ситуацию — определим *k-отделяющий граф* $G_k(n)$. Пусть k и n — натуральные числа, $n > 2k + 2$ (при этом n уже не обязательно четное). Нарисуем на плоскости множество $\mathcal{S}$ из n точек, никакие три из которых не лежат на одной прямой. Ориентированное ребро соединяет две точки A и B, если прямая AB делит множество точек на две части, в одной из которых k точек. Эта часть должна находиться в полуплоскости, которая расположена *справа* от прямой, если двигаться от A к B. Множество вершин графа $G_k(n)$ — это те точки множества $\mathcal{S}$, из которых выходит или в которые входит хотя бы одно ребро. Если нужно подчеркнуть зависимость построенного графа $G_k(n)$ от начальной конфигурации точек, будем его обозначать также $G_k(\mathcal{S})$.

1 Вершины

1.1. Докажите, что в половинчатом графе нет изолированных вершин.

1.2. Если n — число точек конфигурации — фиксировано, то какие значения может иметь степень по задаче вершины в графе $G(n)$?

1.3. Может ли половинчатый граф $G(50)$ иметь ровно 50 вершин: 25 вершин степени 1 и 25 вершин степени 3?

1.4. Докажите, что любой k-отделяющий граф $G_k(n)$ содержит не менее $2k + 3$ вершин.

1.5. a) Докажите, что при $n \geqslant 6$ половинчатый граф имеет не более трех вершин степени $n - 3$.
b) Сколько именно вершин степени $n - 3$ может иметь половинчатый граф?

1.6. Существует ли половинчатый граф, у которого 8 вершин и 9 ребер?

1.7. Докажите, что в половинчатом графе со 100 вершинами не более 60 вершин имеют степень 41.

2 Свойства графов

2.1. Докажите, что если в половинчатом графе $G(2n)$ ровно n ребер, то отрезки, изображающие ребра, попарно пересекаются.

2.1$\frac{1}{2}$. На плоскости нарисован геометрический ориентированный граф. Докажите, что ни при каких n и n' он не может оказаться одновременно графом $G_{10}(n)$ и графом $G_{15}(n')$.

2.2. Докажите, что в половинчатом графе не может быть гамильтонова пути (т.е. пути, проходящего по всем вершинам ровно по одному разу).

2.3. Пусть $n = 103$. При каких k можно утверждать, что k-отделяющий граф $G_k(n)$ заведомо является связным?

2.4. Докажите, что каждая компонента связности k-отделяющего графа $G_k(n)$ имеет эйлеров путь (путь, проходящий по всем ребрам по одному разу).

2.5. Докажите, что если при некотором n построены половинчатые графы с k_1 и k_2 ребрами, то для любого m, $k_1 \leqslant m \leqslant k_2$, существует половинчатый граф с n вершинами и m ребрами.

2.6. a) Пусть на плоскости нарисован k-отделяющий граф $G_k(\mathcal{S})$ и пусть $\mathcal{S}'$ — множество его вершин. Докажите, что этот граф есть граф вида $G_{k'}(\mathcal{S}')$ при подходящем выборе k'.
b) Докажите, что каждая компонента связности графа $G_k(n)$ является графом вида $G_{k'}(n')$.

2.7. a) Докажите, что любой граф является подграфом подходящего половинчатого графа.
b) Докажите, что любой граф является индуцированным подграфом подходящего половинчатого графа.

3 Выпуклые цепочки и мельницы

Разобьем половинчатый граф на «выпуклые цепочки». Для этого развернем его так, чтобы ни одна из прямых, соединяющих вершины, не была вертикальной, и проведем вертикальную прямую (не проходящую через вершины), по сторонам от которой расположено поровну вершин графа. Теперь проведем вертикальную прямую ℓ через самую левую вершину V_1. Будем вращать прямую ℓ по часовой стрелке вокруг V_1, пока она не совпадет с каким-нибудь ребром половинчатого графа, скажем, с V_1V_2. Продолжим вращение прямой теперь уже вокруг вершины V_2, пока она не наткнется на следующее ребро V_2V_3 и т. д. Если в какой-то момент прямая снова станет вертикальной, остановим процесс, и скажем, что построенная ломаная $V_1V_2V_3\ldots$ является *выпуклой цепочкой*. После этого начнем строить новую цепочку, взяв в качестве стартовой точки самую левую из вершин, у которой есть ребро, не принадлежащее ни одной из уже построенных цепочек. В результате весь граф окажется разбит на выпуклые цепочки. Это разбиение зависит от того, какое направление мы изначально выбрали в качестве вертикального.

3.1. Докажите, что в результате описанного процесса половинчатый граф будет представлять собой объединение $n/2$ выпуклых цепочек, каждая из которых начинается в левой полуплоскости, заканчивается в правой, и при этом никакие две цепочки не имеют общих ребер.

3.2. Докажите, что для любых двух вершин половинчатого графа сумма их степеней не превосходит n.

В следующих задачах мы считаем, что на плоскости дано конечное множество точек $\mathcal{S}$, никакие три из которых не лежат на одной прямой.

Определение. *Мельницей* будем называть следующий процесс. Выберем прямую ℓ, проходящую через ровно одну точку $T \in \mathcal{S}$. Будем вращать эту прямую по часовой стрелке до того момента, когда на ней впервые появится еще одна точка из множества $\mathcal{S}$, назовем ее U. Далее будем продолжать вращать прямую по часовой стрелке, но теперь уж вокруг точки U, пока на ней не появится еще одна точка множества $\mathcal{S}$ и т. д.

3.3. Докажите, что можно выбрать точку T и начальную прямую ℓ так, что каждая точка множества $\mathcal{S}$ побывает центром вращения мельницы бесконечно много раз.

3.4. Докажите, что для любого множества $\mathcal{S}$ можно так запустить мельницу, что каждая точка плоскости в какой-то момент побывает на прямой ℓ.

3.5. Докажите, что в любом множестве $\mathcal{S}$ есть такая точка, что любая мельница, запущенная из нее, посещает все вершины.

3.6. Пусть на плоскости отмечены точки общего положения и прямая a, не проходящая через них. Покрасим точки с одной стороны от нее (пусть их K штук) в красный цвет, а с другой стороны (пусть их M штук) — в синий. Докажите, что для любых $k < K$, $m < M$ существует прямая b такая, что по одну сторону от нее лежат ровно k красных и ровно m синих точек.

4 Экстремальные задачи

4.1. a) Может ли при каком-нибудь n половинчатый граф с n вершинами иметь больше $2013n$ ребер?
b) Может ли граф $G_{10}(n)$ при каком-нибудь n иметь $2013n$ ребер?

4.2. Какое наибольшее число ребер может иметь путь в половинчатом графе $G(n)$?

4.3. Какое наибольшее число ребер может иметь цикл в половинчатом графе $G(n)$?

4.4. Докажите, что полный подграф с k вершинами может быть подграфом половинчатого графа, в котором a) порядка k^3 вершин; b) порядка k^2 вершин.

4.5. Докажите, что оценка из пункта b) предыдущей задачи асимптотически точная, а именно, если половинчатый граф с n вершинами содержит клику с k вершинами, то $n \geqslant [k^2/2]$.

4.6. Обозначим через $e_{k,n}$ наибольшее число ребер, которое может иметь k-отделяющий граф $G_k(n)$. Докажите, что $e_{2n,2k+1} \geqslant 2e_{n,k} + n$.

4.7. Докажите, что в графе $G_k(n)$ при $k < (n-2)/2$ не более $4\sqrt{(k+1)(n-k-1)n}$ ребер.

5 Круговые последовательности

В этой серии задач мы предлагаем один подход к задаче 4.7; он связан с переформулировкой на другом языке. Поэтому задачи этой серии почти не связаны с остальными задачами.

Определение. 1) *Круговой n-последовательностью* будем называть (конечную) последовательность перестановок множества $\{1, 2, \ldots, n\}$, состоящую из $C_n^2 + 1$ перестановок, в которой

а) каждая следующая перестановка отличается от предыдущей только положением каких-то двух рядом стоящих чисел, будем называть это изменение порядка чисел *флипом*; флип, переставляющий элементы на местах с номерами k и $k+1$, будем называть *k-флипом*;

б) в последней перестановке все элементы расставлены в противоположном порядке (в сравнении с первой перестановкой);

в) любые два числа из исходного множества присутствовали вместе ровно в одном флипе.

2) *Двойная круговая n-последовательность* — это последовательность перестановок множества $\{1, 2, \ldots, n\}$, состоящая из $2C_n^2+1$ перестановок, в которой первые C_n^2+1 перестановок представляют собой круговую последовательность, а последние C_n^2+1 перестановок — ее «зеркальное отражение». Точнее говоря, если в i-й перестановке элементы множества были расставлены в некотором порядке и выполнялся k-флип, меняющий местами числа a и b, то в (C_n^2+i)-й перестановке элементы множества расставлены в противоположном порядке и выполнется $(n-k)$-флип, меняющий местами эти же числа a и b.

Некоторые круговые последовательности могут быть получены геометрически: рассмотрим в плоскости множество $\mathcal{S}$ из n точек общего положения; мы предполагаем, что прямые, соединяющие эти точки, попарно непараллельны. Будем проектировать наши точки на прямую, поворачивающуюся вокруг фиксированного центра. Порядок проекций точек при повороте прямой на 180° — это круговая последовательность, а при повороте прямой на 360° — двойная круговая последовательность.

3) Множество $\mathcal{P} \subset \{1, 2, \ldots, n\}$ будем называть *полуплоскостью* (относительно круговой n-последовательности $\mathcal{T}$), если элементы множества $\mathcal{P}$, расположенные в некотором порядке, являются начальным фрагментом хотя бы одной из перестановок, составляющих $\mathcal{T}$. Если множество $\mathcal{P}$ состоит из k элементов, будем называть его также k-множеством. Количество k-множеств в двойной круговой последовательности $\mathcal{T}$ будем обозначать $s_k(\mathcal{T})$.

5.1. Докажите, что в любой двойной круговой последовательности $\mathcal{T}$ число k-флипов в точности равно числу k-множеств $s_k(\mathcal{T})$.

5.2. Докажите, что для любой двойной круговой n-последовательности $\mathcal{T}$ выполнено равенство $\sum\limits_{k=1}^{n-1} s_k(\mathcal{T}) = n(n-1)$.

5.3. Пусть $S_k(n)$ — наибольшая из сумм вида $\sum\limits_{i=1}^{k} s_i(\mathcal{T})$, где максимум берется по всем двойным круговым n-последовательностям $\mathcal{T}$. Докажите, что $S_k(n) = kn$ при $1 \leqslant k < n/2$.

5.4. Пусть двойная круговая n-последовательность $\mathcal{T}$ построена по множеству $\mathcal{S}$. Докажите, что количество ребер в графе $G_k(\mathcal{S})$ равно $s_{k+1}(\mathcal{T})$ при $1 \leqslant k \leqslant (n-2)/2$.

5.5. Верно ли, что всякая круговая n-последовательность реализуется геометрически?

5.6. Пусть $\mathcal{T}$ — круговая n-последовательность, P — первая перестановка из $\mathcal{T}$, $P = XYZ$, причем длина подстроки Y равна y (а подстроки X и Z могут быть и пустыми), наконец, пусть $1 \leqslant k \leqslant n-1$. Докажите, что количество k-флипов в $\mathcal{T}$, затрагивающих числа из Y, не превосходит $C_y^2 + 2k$.

5.7. Докажите, что существует такое число C, что для всех двойных круговых n-последовательностей $\mathcal{T}$ при $1 \leqslant k \leqslant n/2$ выполнено неравенство $s_k(\mathcal{T}) \leqslant Cn\sqrt{k}$.

6 Пересечения и максимальные количества ребер

Определение. Назовем *пересечением* пару ребер геометрического графа, имеющих общую точку, не являющуюся их общим концом. В этой серии мы изучаем возможное количество пересечений в графах.

6.1. Пусть V — множество вершин половинчатого графа, $|V| = n$, n — четное. Для каждой вершины $v \in V$ обозначим через d_v степень этой вершины. Пусть X — количество пересечений в этом графе. Докажите, что

$$X + \sum_{v \in V} C^2_{(d_v+1)/2} = C^2_{n/2}.$$

6.2. Пусть $10^6 n < e < 10^{-6} n^2$. Докажите, что существует геометрический граф, в котором n вершин, e ребер, а количество пересечений не превосходит $10^6 e^3/n^2$.

6.3. Докажите, что существует такое число $c > 0$, что при $e > 100n$ в геометрическом графе с n вершинами и e ребрами имеется не менее $c \cdot e^3/n^2$ пересечений.

6.4. Докажите, что число ребер в половинчатом графе с n вершинами не превосходит $Cn^{4/3}$, для некоторой константы C, не зависящей от n и от графа.

Половинчатые графы

Задачу представляют И. Богданов, К. Кохась

В этой серии задач мы будем иметь дело с *геометрическими графами*, т. е. с графами, изображенными на плоскости. В геометрическом графе никакие три вершины не лежат на одной прямой, а ребра изображаются отрезками (которые могут пересекаться не в вершине). Вот два примера геометрических графов, представляющих для нас основной интерес.

Определение. 1) *Половинчатым* графом $G(n)$ назовем граф, получающийся следующей конструкцией. Пусть на плоскости нарисовано n (четное число) точек, никакие три из которых не лежат на одной прямой. Это вершины графа. Ребро соединяет две вершины, если прямая, содержащая это ребро, делит множество вершин на две равные части (т. е. в ее полуплоскостях поровну оставшихся вершин).

2) Рассмотрим более общую ситуацию — определим *k-отделяющий граф* $G_k(n)$. Пусть k и n — натуральные числа, $n > 2k + 2$ (при этом n уже не обязательно четное). Нарисуем на плоскости множество $\mathcal{S}$ из n точек, никакие три из которых не лежат на одной прямой. Ориентированное ребро соединяет две точки A и B, если прямая AB делит множество точек на две части, в одной из которых k точек. Эта часть должна находиться в полуплоскости, которая расположена *справа* от прямой, если двигаться от A к B. Множество вершин графа $G_k(n)$ — это те точки множества $\mathcal{S}$, из которых выходит или в которые входит хотя бы одно ребро. Если нужно подчеркнуть зависимость построенного графа $G_k(n)$ от начальной конфигурации точек, будем его обозначать также $G_k(\mathcal{S})$.

1 Вершины

1.1. Докажите, что в половинчатом графе нет изолированных вершин.

1.2. Если n — число точек конфигурации — фиксировано, то какие значения может иметь степень по задаче вершины в графе $G(n)$?

1.3. Может ли половинчатый граф $G(50)$ иметь ровно 50 вершин: 25 вершин степени 1 и 25 вершин степени 3?

1.4. Докажите, что любой k-отделяющий граф $G_k(n)$ содержит не менее $2k + 3$ вершин.

1.5. a) Докажите, что при $n \geqslant 6$ половинчатый граф имеет не более трех вершин степени $n - 3$.
b) Сколько именно вершин степени $n - 3$ может иметь половинчатый граф?

1.6. Существует ли половинчатый граф, у которого 8 вершин и 9 ребер?

1.7. Докажите, что в половинчатом графе со 100 вершинами не более 60 вершин имеют степень 41.

2 Свойства графов

2.1. Докажите, что если в половинчатом графе $G(2n)$ ровно n ребер, то отрезки, изображающие ребра, попарно пересекаются.

2.1$\frac{1}{2}$. На плоскости нарисован геометрический ориентированный граф. Докажите, что ни при каких n и n' он не может оказаться одновременно графом $G_{10}(n)$ и графом $G_{15}(n')$.

2.2. Докажите, что в половинчатом графе не может быть гамильтонова пути (т.е. пути, проходящего по всем вершинам ровно по одному разу).

2.3. Пусть $n = 103$. При каких k можно утверждать, что k-отделяющий граф $G_k(n)$ заведомо является связным?

2.4. Докажите, что каждая компонента связности k-отделяющего графа $G_k(n)$ имеет эйлеров путь (путь, проходящий по всем ребрам по одному разу).

2.5. Докажите, что если при некотором n построены половинчатые графы с k_1 и k_2 ребрами, то для любого m, $k_1 \leqslant m \leqslant k_2$, существует половинчатый граф с n вершинами и m ребрами.

2.6. a) Пусть на плоскости нарисован k-отделяющий граф $G_k(\mathcal{S})$ и пусть $\mathcal{S}'$ — множество его вершин. Докажите, что этот граф есть граф вида $G_{k'}(\mathcal{S}')$ при подходящем выборе k'.
b) Докажите, что каждая компонента связности графа $G_k(n)$ является графом вида $G_{k'}(n')$.

2.7. a) Докажите, что любой граф является подграфом подходящего половинчатого графа.
b) Докажите, что любой граф является индуцированным подграфом подходящего половинчатого графа.

3 Выпуклые цепочки и мельницы

Разобьем половинчатый граф на «выпуклые цепочки». Для этого развернем его так, чтобы ни одна из прямых, соединяющих вершины, не была вертикальной, и проведем вертикальную прямую (не проходящую через вершины), по сторонам от которой расположено поровну вершин графа. Теперь проведем вертикальную прямую ℓ через самую левую вершину V_1. Будем вращать прямую ℓ по часовой стрелке вокруг V_1, пока она не совпадет с каким-нибудь ребром половинчатого графа, скажем, с V_1V_2. Продолжим вращение прямой теперь уже вокруг вершины V_2, пока она не наткнется на следующее ребро V_2V_3 и т. д. Если в какой-то момент прямая снова станет вертикальной, остановим процесс, и скажем, что построенная ломаная $V_1V_2V_3\ldots$ является *выпуклой цепочкой*. После этого начнем строить новую цепочку, взяв в качестве стартовой точки самую левую из вершин, у которой есть ребро, не принадлежащее ни одной из уже построенных цепочек. В результате весь граф окажется разбит на выпуклые цепочки. Это разбиение зависит от того, какое направление мы изначально выбрали в качестве вертикального.

3.1. Докажите, что в результате описанного процесса половинчатый граф будет представлять собой объединение $n/2$ выпуклых цепочек, каждая из которых начинается в левой полуплоскости, заканчивается в правой, и при этом никакие две цепочки не имеют общих ребер.

3.2. Докажите, что для любых двух вершин половинчатого графа сумма их степеней не превосходит n.

В следующих задачах мы считаем, что на плоскости дано конечное множество точек $\mathcal{S}$, никакие три из которых не лежат на одной прямой.

Определение. *Мельницей* будем называть следующий процесс. Выберем прямую ℓ, проходящую через ровно одну точку $T \in \mathcal{S}$. Будем вращать эту прямую по часовой стрелке до того момента, когда на ней впервые появится еще одна точка из множества $\mathcal{S}$, назовем ее U. Далее будем продолжать вращать прямую по часовой стрелке, но теперь уж вокруг точки U, пока на ней не появится еще одна точка множества $\mathcal{S}$ и т. д.

3.3. Докажите, что можно выбрать точку T и начальную прямую ℓ так, что каждая точка множества $\mathcal{S}$ побывает центром вращения мельницы бесконечно много раз.

3.4. Докажите, что для любого множества $\mathcal{S}$ можно так запустить мельницу, что каждая точка плоскости в какой-то момент побывает на прямой ℓ.

3.5. Докажите, что в любом множестве $\mathcal{S}$ есть такая точка, что любая мельница, запущенная из нее, посещает все вершины.

3.6. Пусть на плоскости отмечены точки общего положения и прямая a, не проходящая через них. Покрасим точки с одной стороны от нее (пусть их K штук) в красный цвет, а с другой стороны (пусть их M штук) — в синий. Докажите, что для любых $k < K$, $m < M$ существует прямая b такая, что по одну сторону от нее лежат ровно k красных и ровно m синих точек.

4 Экстремальные задачи

4.1. a) Может ли при каком-нибудь n половинчатый граф с n вершинами иметь больше $2013n$ ребер?
b) Может ли граф $G_{10}(n)$ при каком-нибудь n иметь $2013n$ ребер?

4.2. Какое наибольшее число ребер может иметь путь в половинчатом графе $G(n)$?

4.3. Какое наибольшее число ребер может иметь цикл в половинчатом графе $G(n)$?

4.4. Докажите, что полный подграф с k вершинами может быть подграфом половинчатого графа, в котором a) порядка k^3 вершин; b) порядка k^2 вершин.

4.5. Докажите, что оценка из пункта b) предыдущей задачи асимптотически точная, а именно, если половинчатый граф с n вершинами содержит клику с k вершинами, то $n \geqslant [k^2/2]$.

4.6. Обозначим через $e_{k,n}$ наибольшее число ребер, которое может иметь k-отделяющий граф $G_k(n)$. Докажите, что $e_{2n,2k+1} \geqslant 2e_{n,k} + n$.

4.7. Докажите, что в графе $G_k(n)$ при $k < (n-2)/2$ не более $4\sqrt{(k+1)(n-k-1)n}$ ребер.

5 Круговые последовательности

В этой серии задач мы предлагаем один подход к задаче 4.7; он связан с переформулировкой на другом языке. Поэтому задачи этой серии почти не связаны с остальными задачами.

Определение. 1) *Круговой n-последовательностью* будем называть (конечную) последовательность перестановок множества $\{1, 2, \ldots, n\}$, состоящую из $C_n^2 + 1$ перестановок, в которой

а) каждая следующая перестановка отличается от предыдущей только положением каких-то двух рядом стоящих чисел, будем называть это изменение порядка чисел *флипом*; флип, переставляющий элементы на местах с номерами k и $k+1$, будем называть *k-флипом*;

б) в последней перестановке все элементы расставлены в противоположном порядке (в сравнении с первой перестановкой);

в) любые два числа из исходного множества присутствовали вместе ровно в одном флипе.

2) *Двойная круговая n-последовательность* — это последовательность перестановок множества $\{1, 2, \ldots, n\}$, состоящая из $2C_n^2+1$ перестановок, в которой первые C_n^2+1 перестановок представляют собой круговую последовательность, а последние C_n^2+1 перестановок — ее «зеркальное отражение». Точнее говоря, если в i-й перестановке элементы множества были расставлены в некотором порядке и выполнялся k-флип, меняющий местами числа a и b, то в (C_n^2+i)-й перестановке элементы множества расставлены в противоположном порядке и выполнется $(n-k)$-флип, меняющий местами эти же числа a и b.

Некоторые круговые последовательности могут быть получены геометрически: рассмотрим в плоскости множество $\mathcal{S}$ из n точек общего положения; мы предполагаем, что прямые, соединяющие эти точки, попарно непараллельны. Будем проектировать наши точки на прямую, поворачивающуюся вокруг фиксированного центра. Порядок проекций точек при повороте прямой на 180° — это круговая последовательность, а при повороте прямой на 360° — двойная круговая последовательность.

3) Множество $\mathcal{P} \subset \{1, 2, \ldots, n\}$ будем называть *полуплоскостью* (относительно круговой n-последовательности $\mathcal{T}$), если элементы множества $\mathcal{P}$, расположенные в некотором порядке, являются начальным фрагментом хотя бы одной из перестановок, составляющих $\mathcal{T}$. Если множество $\mathcal{P}$ состоит из k элементов, будем называть его также k-множеством. Количество k-множеств в двойной круговой последовательности $\mathcal{T}$ будем обозначать $s_k(\mathcal{T})$.

5.1. Докажите, что в любой двойной круговой последовательности $\mathcal{T}$ число k-флипов в точности равно числу k-множеств $s_k(\mathcal{T})$.

5.2. Докажите, что для любой двойной круговой n-последовательности $\mathcal{T}$ выполнено равенство $\sum\limits_{k=1}^{n-1} s_k(\mathcal{T}) = n(n-1)$.

5.3. Пусть $S_k(n)$ — наибольшая из сумм вида $\sum\limits_{i=1}^{k} s_i(\mathcal{T})$, где максимум берется по всем двойным круговым n-последовательностям $\mathcal{T}$. Докажите, что $S_k(n) = kn$ при $1 \leqslant k < n/2$.

5.4. Пусть двойная круговая n-последовательность $\mathcal{T}$ построена по множеству $\mathcal{S}$. Докажите, что количество ребер в графе $G_k(\mathcal{S})$ равно $s_{k+1}(\mathcal{T})$ при $1 \leqslant k \leqslant (n-2)/2$.

5.5. Верно ли, что всякая круговая n-последовательность реализуется геометрически?

5.6. Пусть $\mathcal{T}$ — круговая n-последовательность, P — первая перестановка из $\mathcal{T}$, $P = XYZ$, причем длина подстроки Y равна y (а подстроки X и Z могут быть и пустыми), наконец, пусть $1 \leqslant k \leqslant n-1$. Докажите, что количество k-флипов в $\mathcal{T}$, затрагивающих числа из Y, не превосходит $C_y^2 + 2k$.

5.7. Докажите, что существует такое число C, что для всех двойных круговых n-последовательностей $\mathcal{T}$ при $1 \leqslant k \leqslant n/2$ выполнено неравенство $s_k(\mathcal{T}) \leqslant Cn\sqrt{k}$.

6 Пересечения и максимальные количества ребер

Определение. Назовем *пересечением* пару ребер геометрического графа, имеющих общую точку, не являющуюся их общим концом. В этой серии мы изучаем возможное количество пересечений в графах.

6.1. Пусть V — множество вершин половинчатого графа, $|V| = n$, n — четное. Для каждой вершины $v \in V$ обозначим через d_v степень этой вершины. Пусть X — количество пересечений в этом графе. Докажите, что

$$X + \sum_{v \in V} C^2_{(d_v+1)/2} = C^2_{n/2}.$$

6.2. Пусть $10^6 n < e < 10^{-6} n^2$. Докажите, что существует геометрический граф, в котором n вершин, e ребер, а количество пересечений не превосходит $10^6 e^3/n^2$.

6.3. Докажите, что существует такое число $c > 0$, что при $e > 100n$ в геометрическом графе с n вершинами и e ребрами имеется не менее $c \cdot e^3/n^2$ пересечений.

6.4. Докажите, что число ребер в половинчатом графе с n вершинами не превосходит $Cn^{4/3}$, для некоторой константы C, не зависящей от n и от графа.

Решения

1.1. Требуется доказать, что через каждую вершину графа $G(n)$ можно провести прямую, делящую множество вершин пополам. Возьмем произвольную вершину V, проведем через нее прямую ℓ, и будем вращать прямую ℓ вокруг точки V. При вращении другие вершины графа по одной переходят из левой полуплоскости в правую или из правой в левую. После поворота на 180° левая и правая полуплоскости поменяются местами. Поэтому в какой-то момент в этих полуплоскостях будет поровну вершин (и, поскольку общее число вершин четно, наша прямая будет содержать еще одну точку нашего множества, кроме V).

1.2. Ответ: степень вершины всегда нечетна; она может быть любым нечетным числом от 1 до $n-1$.

Как мы проверили в предыдущей задаче, через каждую вершину графа $G(n)$ можно провести прямую, делящую множество вершин пополам. Поэтому степень вершины не может быть равна нулю. При этом в любом графе с n вершинами степень вершины не превосходит $n-1$.

Докажем, что степень любой вершины графа $G(n)$ нечетна. Возьмем произвольную вершину V, проведем через нее прямую ℓ, делящую множество вершин пополам. Вершина V делит эту прямую на два луча, на одном из которых есть еще одна вершина, скажем, P_1. Луч, содержащий P_1, покрасим в красный цвет, а второй луч — в синий.

Лемма. Пусть VP_1, VP_2, …, VP_k — прямые, последовательно получающиеся при вращении прямой ℓ вокруг точки V, делящие множество вершин графа пополам. Тогда любые две соседние точки P_i и P_{i+1} находятся на лучах разного цвета.

Доказательство. Проверим, что если точка P_i находилась на красном луче, то следующая точка P_{i+1} окажется на синем луче. Это так, поскольку, повернув немного прямую ℓ из положения VP_i, мы увеличиваем на 1 число число вершин графа в той полуплоскости, куда попала точка P_i. Если при вращении еще какая-либо вершина переходит в другую полуплоскость через красный луч, разность количеств вершин в полуплоскостях еще больше увеличивается, а если через синий — уменьшается. В момент, когда точка P_{i+1} появляется на прямой ℓ, эта разность опять становится равной нулю. Очевидно, это событие может произойти только на синем луче. □

После поворота на 180° прямая ℓ опять встретит точку P_1, но теперь уже на синем луче. Поэтому число k нечетно.

Осталось описать конструкцию, которая для каждого $k < n/2$ позволяет построить граф $G(n)$, содержащий вершину степени $2k+1$. Для этого нам понадобятся две геометрические конструкции.

1) «*Толстая линия*». Изображая какой-нибудь геометрический граф (на ограниченной части плоскости), мы можем изменить масштаб картинки по вертикали, в результате чего она станет очень узкой, половинчатые прямые почти сольются, и вся картинка станет похожа на слегка утолщенную, прорисованную несколько раз линию (рис. 1).

2) «*Крест*». Возьмем две конфигурации, задающих половинчатые графы G_1 и G_2, изобразим каждую из них в виде толстой линии, наметим на каждой линии «среднюю точку» (по сторонам от которой поровну точек данной конфигурации) и пересечем эти две толстые прямые в средних точках (рис. 2). В результате мы получим конфигурацию точек, половинчатый граф которой состоит из двух компонент связности G_1 и G_2.

Возвращаясь к решению задачи, заметим, что крест, одна из компонент которого «звезда» (правильный $(2k+1)$-угольник плюс одна вершина в центре), а другая — выпуклый $(n-2k-2)$-угольник, дает нам граф, в котором одна вершина степени $2k+1$, а остальные — степени 1.

1.3. Ответ: да.

Мы взяли это утверждение в [6, лемма 5.4]. Приведем два возможных примера.

Один пример таков. Возьмем вершины правильного 25-угольника; также возьмем вершины его образа при гомотетии с центром в центре 25-угольника и коэффициентом, близким к 1. Все вершины внешнего 25-угольника имеют степень 1, все вершины внутреннего — степень 3.

Рис. 1. Толстая линия

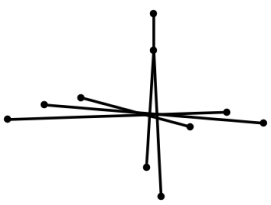


Рис. 2. Крест

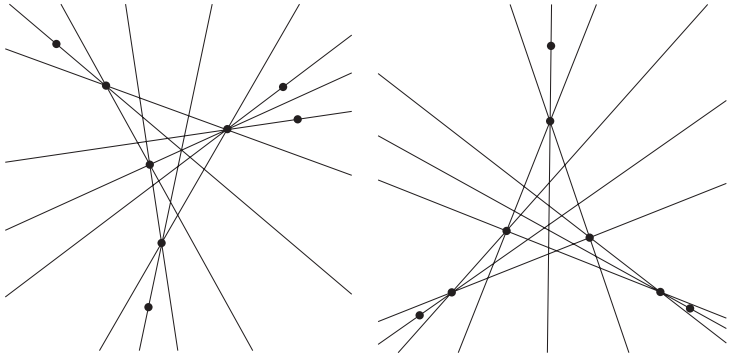


Рис. 3. 8 вершин, 9 ребер

Другой пример таков. Достаточно взять «крест» двух конфигураций: первая конфигурация — из решения задачи 4.2 при $n = 28$ (все вершины V-образного пути, построенного в той задаче, кроме крайних, имеют степень 3 в половинчатом графе), а вторая — вершины выпуклого 22-угольника (все вершины половинчатого графа этой конфигурации имеют степень 1).

1.4. Мы взяли это утверждение в [5, теорема 3.3].

Лемма. Пусть $\mathcal{S}$ — конфигурация из n точек. Точка $A \in \mathcal{S}$ тогда и только тогда является вершиной графа $G_k(\mathcal{S})$, когда существует прямая ℓ, проходящая через A, для которой одна из полуплоскостей содержит не более k точек. (Напомним, что $n > 2k + 2$.)

Доказательство. В одну сторону (если A — вершина, то существует прямая) это утверждение очевидно. В другую сторону оно доказывается вращением прямой ℓ вокруг A. Ориентируем прямую ℓ. В начальный момент в одной из полуплоскостей, для определенности, слева, находится не более k точек, а справа — не менее $k + 1$ точки. После поворота на 180° полуплоскости поменялись местами. Значит, был момент, когда прямая проходила через еще одну точку B и в полуплоскости слева было ровно k точек. Значит, $\overrightarrow{BA}$ — ребро нашего графа, и, в частности, A — его вершина. □

Теперь докажем утверждение задачи. Возьмем произвольную вершину A выпуклой оболочки исходного множества точек. У этой вершины есть входящее ребро $\overrightarrow{CA}$ и исходящее ребро $\overrightarrow{AB}$ (это легко доказать, рассмотрев опорную прямую в точке A и повернув ее на 180°). Прямые AB и AC отсекают от конфигурации $2k + 3$ точки, включая A, B и C. По лемме, каждая из этих точек является вершиной графа $G_k(n)$.

1.5. a) Мы взяли это утверждение в [6, лемма 5.2].

Ясно, что каждая вершина выпуклой оболочки исходного набора точек имеет в половинчатом графе степень 1. Поскольку выпуклая оболочка представляет собой выпуклый многоугольник, он содержит не менее трех вершин, и значит, не менее трех вершин половинчатого графа имеют степень 1 («лист»). Поскольку $n > 4$, вершины степени $n - 3$ не являются листами; более того, каждая из них соединена со всеми листами, кроме как максимум двух. Следовательно, количество вершин степени $n - 3$ не может быть больше 3.

b) Ответ: 0 или 1.

То, что количество вершин степени $n - 3$ не превосходит 1, сразу следует из утверждений задач 3.2 и 6.1. Построение графа с одной вершиной степени $n - 3$ (и с нулем таких вершин) обсуждалось в задаче 1.2.

1.6. Ответ: да, см. рис. 3. Мы взяли эти картинки в [4].

1.7. Это сразу следует из утверждения задачи 6.1, но мы приведем простое рассуждение с выпуклыми цепочками [6, лемма 6.3].

Возьмем разбиение графа на выпуклые цепочки. По утверждению задачи 3.1 каждая вершина является концом ровно одной цепочки. Упорядочим вершины левой полуплоскости слева направо. Тогда самая левая вершина имеет степень 1. Вторая вершина имеет степень не более 3, так как через нее может проходить цепочка, выпущенная из первой вершины, и кроме того в ней начинается вторая цепочка. Третья вершина имеет степень не более 5, так как через нее проходит не более двух цепочек и еще одна в ней начинается, и т. д. Таким образом, в левой полуплоскости лишь 30 вершин — вершины с 21-й по 50-ю — могут иметь степень 41. Аналогично в правой полуплоскости.

2.1. Это утверждение из [7].

Заметим, что никакие два ребра не имеют общего конца, иначе нашлась бы изолированная вершина. Предположим, что ребра P_1P_2 и P_3P_4 не пересекаются. Тогда можно считать, что отрезок P_3P_4 не пересекается с прямой P_1P_2; далее, можно считать, что прямая $\ell = P_1P_2$ горизонтальна, точки P_3 и P_4 лежат выше нее, и вершина P_4 находится дальше от ℓ, чем P_3. Будем вращать прямую P_3P_4 вокруг точки P_3 против часовой

стрелки, пока она не станет параллельна ℓ. Сразу после начала вращения справа от прямой будет больше половины точек, а в конце — меньше половины. Значит, через точку P_3 проходит еще одно ребро, отличное от P_3P_4. Противоречие.

2.1$\frac{1}{2}$. По лемме из решения задачи 1.4, каждая точка исходной конфигурации, отсеченная ребром, является вершиной графа. Поэтому число k однозначно определяется по графу $G_k(n)$.

2.2. Как мы отмечали в решении задачи 1.5, в половинчатом графе имеется не менее трех вершин степени 1. Если бы существовал гамильтонов путь, то у каждой вершины графа, кроме, быть может, концов пути, степень была бы не меньше 2.

2.3. Ответ: только при $k = 50$.

Этот сюжет взят в [5, теоремы 2.3, 2.4].

При $k = 50$ граф связен, как показано в решении задачи 3.3 (нечетный случай).

Пусть теперь $k < 50$; построим несвязный граф $G_k(n)$. Пусть Ω — окружность с центром O, а AA', BB' и CC' — три ее диаметра. Выберем достаточно маленький треугольник XYZ со сторонами, параллельными этим диаметрам, как показано на рис. 4; точки X, Y и Z — первые вершины нашей конфигурации.

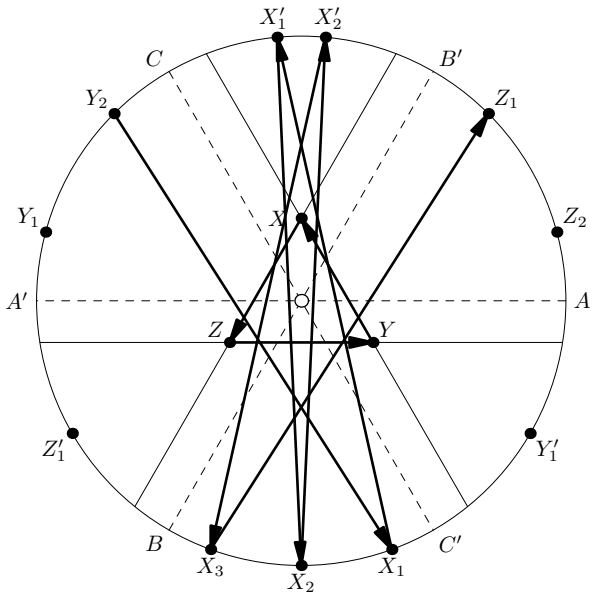


Рис. 4. Несвязный граф $G_k(n)$: здесь $k = 5 = 1 + (2 + 1 + 1)$

Теперь разложим k в сумму $k = 1 + p + q + r$ с неотрицательными целыми p, q и r. Выберем на дуге $C'B$ точки X_1, …, X_{p+1} (упорядоченные по часовой стрелке), а затем выберем на дуге CB точки X_1', …, X_p' так, чтобы точки O и X лежали внутри каждого из треугольников $X_iX_i'X_{i+1}$ (это возможно, если точка X достаточно близка к O). Аналогично мы выберем точки $Y_1, \ldots, Y_{q+1}$ на дуге $A'C$, и т. д.; все выбранные точки также являются вершинами нашей конфигурации (на рис. 4 представлен случай $p = 2$, $q = r = 1$). Наконец, поставим оставшиеся $n - (2k + 4) \geqslant 0$ точек достаточно близко к центру O.

Нетрудно проверить (например, запустив подходящую мельницу), что все ребра соответствующего графа $G_k(n)$ разбиваются на два цикла: $X \to Y \to Z \to X$ и

$$X_1 \to X_1' \to X_2 \to X_2' \to \cdots \to X_{p+1} \to Z_1 \to Z_1' \to \cdots \to Z_{r+1} \to Y_1 \to Y_1' \to \cdots \to Y_{q+1} \to X_1.$$

Замечание. В исходной работе [5] пример несвязного графа был неверным. Здесь мы приведем этот пример (исправленный так, чтобы он подходил) и обсудим, для каких k он не подходит.

Пусть $k < (n-3)/2$, и пусть d — некоторое натуральное число. Отметим вершины некоторого правильного $(2k + 2 + d)$-угольника (предполагая, что $2k + 2 + d \leqslant n$) и несколько точек достаточно близко к его центру так, чтобы всего получилось n точек. Тогда вершинами соответствующего графа $G_k(n)$ будут все вершины $(2k + 2 + d)$-угольника; занумеруем их против часовой стрелки. Все ребра будут выглядеть как

$$i \to (i + k + 1) \pmod{2k + 2 + d};$$

так что эти ребра образуют несколько «звезд»; их количество равно $\text{НОД}(2k+2+d, k+1) = \text{НОД}(k+1, d)$, так что минимальное возможное d, при котором граф несвязен — это наименьший простой делитель числа $k + 1$. Будем считать, что d и есть этот делитель.

Итак, эта конструкция не работает при $2k+2+d>n$. В случае $n=103$, такими исключениями являются $k=36$, 40, 42, 46 и 49.

2.4. Существование эйлерова цикла в каждой компоненте связности вытекает из того, что у каждой вершины v графа $G_k(n)$ количество выходящих и входящих ребер одинаково. Докажем это, следуя [5, теорема 2.5].

Выберем какую-либо ориентированную прямую ℓ, проходящую через v, и будем вращать ее против часовой стрелки. Достаточно проверить, что между двумя положениями прямой ℓ, в которых она содержала исходящие ребра, обязательно было положение, в котором она содержала входящее ребро (и наоборот, между двумя «входящими» положениями было хотя бы одно «исходящее»).

Пусть ℓ_1 и ℓ_2 — два «исходящих» положения. При вращении от положения ℓ_1 к положению ℓ_2 количество точек справа от прямой ℓ могло меняться: если прямая проходила через вершину внутри ориентированного угла $\angle(\ell_1,\ell_2)$, число точек возрастало на 1, а при прохождении через вертикальный угол $\angle(-\ell_1,-\ell_2)$ — уменьшалось на 1. Вблизи положения ℓ_1 число точек справа от прямой ℓ было равно $k+1$, а вблизи положения ℓ_2 — это число равно k. Значит, в какой-то момент, количество точек должно было уменьшиться с $k+1$ до k. Вершина u, через которую в этот момент прошла прямая, задает входящее ребро $\overrightarrow{uv}$.

2.5. Это утверждение из [6, теорема 3.5].

Нарисуем два данных графа так, чтобы никакие три вершины их объединения не лежали на одной прямой. Можно считать, что вершины графа пронумерованы. Будем по одной передвигать каждую вершину первого графа на место, занимаемое соответственной вершиной второго графа. При таком передвижении ребра половинчатого графа могут исчезать или появляться только в момент, когда движущаяся вершина пересекает прямую, проходящую через две другие вершины. Как нетрудно видеть, количество ребер при таком пересечении изменяется не более чем на 1 (пример такого изменения показан на рис. 17).

Итак, графы, получающиеся в этом процессе, имеют любое количество ребер от k_1 до k_2.

2.6. Опишем «мельничную» конструкцию построения графа $G_k(\mathcal{S})$ из [5, § 2].

Возьмем произвольную ориентированную прямую ℓ, справа от которой лежит $k+1$ точка множества $\mathcal{S}$. Подвинем прямую влево, чтобы она наткнулась на ближайшую слева точку $P_1\in\mathcal{S}$. Далее будем действовать как в определении мельницы: будем вращать прямую ℓ против часовой стрелки до того момента, когда на ней впервые появится еще одна точка $P_2\in\mathcal{S}$. Продолжим вращать прямую ℓ против часовой стрелки, но теперь уже вокруг точки P_2, пока на ней не появится еще одна точка множества $P_3\in\mathcal{S}$ и т. д. В результате мы построим последовательность точек P_1, P_2, …

Ясно, что в каждый момент, когда прямая ℓ проходит лишь через одну точку P_i, справа от нее лежит $k+1$ точка множества $\mathcal{S}$; когда же прямая ℓ совпадает с прямой P_iP_{i+1}, справа от нее может лежать $k+1$ или k точек. Как нетрудно проверить, *если вектор $\overrightarrow{P_{i+1}P_i}$ сонаправлен направлению прямой ℓ, то справа от ℓ лежит k точек и, следовательно, этот вектор есть ребро нашего графа*, а в противном случае справа от ℓ лежит $k+1$ точка. При выполнении полного оборота прямая ℓ пробегает все возможные направления; более того, из описанных ее свойств следует, что, когда она сонаправлена с ребром графа $G_k(n)$, она должна его содержать. Значит, среди векторов $\overrightarrow{P_{i+1}P_i}$ встречаются все ребра графа $G_k(\mathcal{S})$, и это в точности ребра, описанные в предыдущем предложении. (Нетрудно заметить, что все остальные векторы вида $\overrightarrow{P_{i+1}P_i}$ — это ровно все ребра графа $G_{k+1}(\mathcal{S})$.)

Теперь вернемся к задаче. Мы решаем одновременно пункты a) и b); разница в решениях заключена в следующем обозначении. В пункте a) мы обозначим через G и $\mathcal{S}'$ граф $G_k(\mathcal{S})$ и множество его вершин, соответственно. В пункте b) мы обозначим через G и $\mathcal{S}'$ нашу компоненту и множество ее вершин, соответственно.

Выполним мельничную конструкцию, дополнительно отслеживая, сколько точек из $\mathcal{S}'$ лежит справа от прямой ℓ. Проверим, что это число не меняется при вращении прямой (кроме случаев, когда она проходит через две точки). Пусть в начальный момент, когда прямая ℓ проходила только через точку P_1, справа от ℓ лежало $k'+1$ точка из $\mathcal{S}'$. Рассмотрим момент, когда центр вращения меняется с P_1 на P_2. Возможны несколько случаев.

1) Пусть вектор $\overrightarrow{P_2P_1}$ не является ребром графа G и противонаправлен с ℓ. Тогда точка P_2 лежала слева от ℓ, а точка P_1 уходит влево от ℓ после смены центра. Таким образом, количество точек из $\mathcal{S}'$ справа от ℓ не меняется.

2) Пусть вектор $\overrightarrow{P_2P_1}$ не является ребром графа G, но теперь он сонаправлен с ℓ. Это невозможно в пункте a) согласно описанию мельничной конструкции; в пункте b) это значит, что $\overrightarrow{P_2P_1}$ — ребро в $G_k(\mathcal{S})$, но не в G, так что обе вершины P_1 и P_2 не лежат в $\mathcal{S}'$. Значит, исследуемое число не изменяется.

3) Пусть, наконец, $\overrightarrow{P_2P_1}$ — ребро в G; тогда $P_1,P_2\in\mathcal{S}'$, точка P_2 была справа от ℓ до смены центра, а после нее точка P_1 уходит вправо от ℓ. Значит, количество точек из $\mathcal{S}'$ справа от ℓ опять же не изменяется.

Итак, количество точек из $\mathcal{S}'$ справа от ℓ всегда равно $k'+1$, кроме случаев, когда ℓ проходит через две точки из $\mathcal{S}'$, соединенные ребром. В этих же случаях это число становится равным k'. Значит, в нашем процессе ребра графа $G_{k'}(\mathcal{S}')$ появляются одновременно с ребрами графа G. Наконец, при этом мы получаем *все* ребра

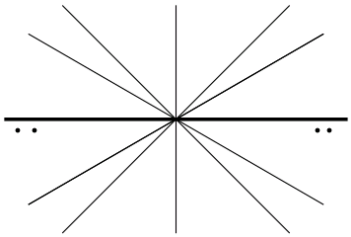


Рис. 5. Добавляем точки

графа $G_{k'}(\mathcal{S}')$: прямая принимает все возможные направления, и в любой момент, когда она параллельна какому-то ребру в $G_{k'}(\mathcal{S}')$, она должна содержать это ребро.

2.7. Достаточно проверить утверждение b). Это утверждение из [6, теорема 5.8].

Нарисуем требуемый граф на плоскости (вершины — в общем положении, ребра — отрезки). Проведем красные прямые, проходящие через ребра, а синие прямые — через все пары точек, не соединенных ребром; можно считать, что среди этих прямых нет параллельных. Выберем настолько мелкий масштаб, что в результате картинка будет выглядеть как набор прямых, пересекающихся «почти в одной точке». Добавляя, если надо, одну точку, можно считать, что общее число точек четно.

Теперь разность количества вершин в полуплоскостях каждой прямой тоже четна. Будем последовательно добавлять точки так, чтобы каждая из красных прямых стала половинной, а каждая синяя стала неполовинной. Для этого заметим, что для любой прямой ℓ можно добавить четное число точек с одной стороны от нее (и очень близко к ней) так, чтобы для других прямых разность количеств точек в полуплоскостях не изменилась. Пример такого добавления показан на рис. 5 — добавляются 4 точки. Ясно, что после серии таких добавлений мы получим требуемую конфигурацию.

3.1. Это утверждение из [6].

По лемме из решения задачи 1.2, для любых двух ребер AC и BC внутри угла R, вертикального к углу ACB, лежит еще хотя бы одно ребро нашего графа. Это значит, что вершины выпуклой цепочки $V_1, V_2, \dots$ идут слева направо. Далее, отсюда же следует, что разные цепочки не могут иметь общих ребер: если, например, две цепочки содержат ребра AC и BC (здесь C лежит правее вершин A и B), то одна из этих цепочек продолжается ребром, лежащим в угле R, а другая — нет.

Кстати, отсюда же следует, что построение цепочки обратимо (нужно просто вращать прямую в обратную сторону). Аналогично проверяется, что в одной вершине не могут начинаться две цепочки. Действительно, если бы это произошло, мы могли бы одну из этих цепочек подходящим образом продолжить влево.

Далее, ориентируем вращающуюся прямую в определении выпуклой цепочки (изначально она направлена вверх). Непосредственно перед тем, как она приходит в положение V_1V_2, количество точек слева от нее меньше, чем справа, а до этого она не проходила через ребра половинчатого графа; значит, и изначально количество слева меньше, то есть точка V_1 лежит в левой половине всей конфигурации. Аналогично, рассматривая последнее ребро, убеждаемся в том, что выпуклая цепочка обязана заканчиваться в правой половине.

В результате каждое ребро принадлежит какой-либо цепочке. Наконец, как мы знаем из решения задачи 1.2, степень каждой вершины нечетна, поэтому каждая вершина слева служит началом некоторой цепочки, причем, как мы уже проверили, ровно одной (а каждая вершина справа — концом ровно одной из них). Таким образом, количество цепочек равно $n/2$.

3.2. Это утверждение из [6, теорема 6.10].

Пусть P и Q — произвольные вершины половинчатого графа. Повернем граф так, чтобы отрезок PQ стал почти вертикальным, и проекции на горизонталь всех остальных вершин графа лежали бы вне отрезка между проекциями вершин P и Q.

Если P и Q не принадлежат одной цепочке, то каждая цепочка проходит максимум через одну из этих точек, поэтому она дает вклад, не превосходящий 2, в сумму их степеней. Кроме того, две цепочки, которые начинаются или заканчиваются в вершинах P и Q, дают вклад 1. Итого получаем, что сумма степеней не более $n-2$.

Если же P и Q принадлежат оной цепочке, то она содержит ребро PQ (и поэтому такая цепочка единственна). Эта цепочка дает вклад 4; значит, к подсчитанной оценке следует прибавить 2.

3.3. Это — задача 2 с IMO-2011, автор Дж. Смит.

Заметим сначала, что каждая мельница продолжается однозначно как в будущее, так и в прошлое; значит, она чисто периодична. Поэтому достаточно построить мельницу, посещающую все точки хотя бы по разу.

Случай 1. Пусть $|\mathcal{S}| = 2n + 1$. Возьмем какую-нибудь прямую ℓ, которая делит множество $\mathcal{S}$ на две равные половины. Такая прямая однозначно задается своим направлением (кроме конечного числа направлений прямых, соединяющих две точки из $\mathcal{S}$) и проходит через какую-нибудь точку $T \in \mathcal{S}$.

Запустим мельницу, начиная с этой прямой, одну из полуплоскостей для удобства будем красить в черный цвет, другую — в белый. Если при вращении прямой ℓ, скажем, вокруг точки A прямая ℓ наткнулась на точку B, до того находившуюся в белой полуплоскости, то при продолжении вращения прямой (теперь уже вокруг точки B) точка A «уйдет» в белую же полуплоскость. Таким образом, при работе мельницы прямая ℓ все время делит множество $\mathcal{S}$ на две равные половины. Следовательно, когда прямая повернется на 180°, она вернется в исходное положение, но черная и белая полуплоскости поменяются местами. Значит, в какой-то момент каждая из остальных точек перешла из полуплоскости одного цвета в другую. Это могло быть, только если она побывала центром вращения мельницы.

Случай 2. Пусть теперь $|\mathcal{S}| = 2n$. Запустим мельницу, начиная с какой-нибудь прямой, проходящей через некоторую точку $T \in \mathcal{S}$, по сторонам от которой лежат n и $n - 1$ точек. После поворота на 180° прямая будет проходить через некоторую точку R, причем все точки, кроме T и R, при вращении сменили цвет полуплоскости. Таким образом, опять каждая точка побывала центром вращения мельницы.

3.4. В качестве такой мельницы можно взять практически ту же конструкцию, что и в предыдущей задаче. Для случая $|\mathcal{S}| = 2n + 1$ вообще не требуется изменения рассуждений, ибо фраза «каждая точка перешла из полуплоскости одного цвета в другую» может быть отнесена к любой точке плоскости вообще. В случае $|\mathcal{S}| = 2n$ достаточно запустить мельницу, начав с какой-нибудь прямой TR, по сторонам от которой по $n - 1$ точке; тогда вышеупомянутая фраза оказывается верна и в этом случае.

3.5. Эта задача была опубликована в Задачнике «Кванта» [1].

Зафиксируем точку $A \in \mathcal{S}$. Для каждой прямой, проходящей через A и не проходящей через другие точки множества $\mathcal{S}$, подсчитаем, сколько точек множества $\mathcal{S}$ оказалось в «меньшей» из двух полуплоскостей этой прямой. Наименьшее из таких чисел назовем *глубиной* точки A.

Пусть теперь A — это точка с самой большой глубиной m. Проверим, что любая мельница, запущенная из точки A, проходит по всем точкам множества $\mathcal{S}$. Как и в решении задачи 3.3, считаем что полуплоскости мельницы раскрашены в два цвета, причем меньшая полуплоскость — белая (по определению глубины, в белой полуплоскости не меньше, чем m точек). Рассмотрим произвольную точку $B \in \mathcal{S}$. Глубина точки B не больше m, значит, существует (ориентированная) прямая ℓ, проходящая через B и не проходящая через другие точки множества $\mathcal{S}$, слева от которой лежит не более m точек. Слегка повращав прямую ℓ вокруг точки B, мы можем считать, что прямая ℓ не параллельна никакой прямой, соединяющей две точки из $\mathcal{S}$. Далее будем считать, что выбранная прямая ℓ направлена вертикально вверх.

Рассмотрим момент, когда мельница была параллельна прямой ℓ и белая полуплоскость была тоже слева. Если мельница в этот момент не проходила через точку B, то B лежала в белой полуплоскости (так слева от ℓ не больше m точек, а в белой полуплоскости – не меньше m). А в тот момент, когда мельница была параллельна прямой ℓ и белая полуплоскость была справа, точка B должна находиться в черной полуплоскости, так как справа от ℓ лежит бóльшая половина всех точек множества $\mathcal{S}$, а в белой полуплоскости — меньшая. Таким образом, точка B побывала и белой полуплоскости и в черной, следовательно, через нее проходила мельница.

3.6. Авторы задачи — Л. Радзивилловский, Д. Кармон.

Пусть прямая a вертикальна, а красные точки лежат слева от нее. Запустим мельницу, в белой полуплоскости которой обычно находится $k + \ell$ точек. Когда прямая мельницы вертикальна, а белая полуплоскость находится слева, то в этой белой полуплоскости находится $\min(k+\ell, K) \geqslant k$ красных точек. Когда же прямая вертикальна, но белая полуплоскость — слева, в ней находится $\min(k+\ell, L) \geqslant \ell$ синих точек. Поскольку количество красных точек в белой полуплоскости изменяется каждый раз не более, чем на 1, в некоторый момент в белой полуплоскости будет k красных и ℓ синих точек. Сдвинув прямую так, чтобы она не содержала точек набора, получаем требуемое.

4.1. а) Ответ: да.

Опишем «Y-образную конструкцию», которая позволяет строить последовательность графов, в которых количество ребер растет быстрее количества вершин. Пусть половинчатый граф G имеет v вершин и e ребер. Рассмотрим три копии графа G и разместим их на толстых лучах, расположенных под углом 120° друг к другу (рис. 6). Все старые ребра останутся половинчатыми, но добавятся $3v/2$ ребер между вершинами разных лучей. Итак, получившаяся конфигурация содержит $3v$ вершин и $3e + 3v/2$ ребер.

Теперь будем последовательно применять эту конструкцию, начиная с конфигурации из двух точек: $v_0 = 2$, $e_0 = 1$. Мы будем получать половинчатые графы в которых

$$v_1 = 2 \cdot 3^1,\ e_1 = 6; \quad v_2 = 2 \cdot 3^2,\ e_2 = 27, \quad \ldots$$

По индукции легко проверяется, что

$$v_k = 2 \cdot 3^k, \quad e_k = k \cdot 3^k.$$

Таким образом, отношение e_k/v_k может быть сколь угодно велико.

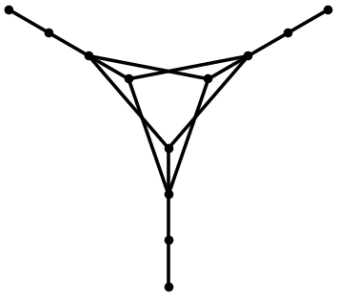


Рис. 6. Y-образная конструкция

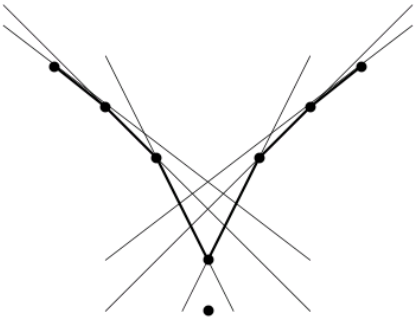


Рис. 7. V-образный путь длины $n-1$

b) Ответ: не может.

Проверим, следуя [5, теорема 4.2], что степени вершин в графе $G_k(n)$ (суммарно исходящие плюс входящие) не превосходят $2k+2$. Для нашей задачи это означает, что в графе $G_{10}(n)$ не более $11n$ ребер. Заметим, что это также вытекает из задачи 5.3.

Пусть сначала n четно. Возьмем произвольную вершину P графа $G_k(n)$. Так как через P проходит хотя бы одно половинчатое ребро, среди отрезков, соединяющих P с другими вершинами, не более $n-2$ отрезков могут быть ребрами хоть в каком-нибудь из графов $G_{k'}(n)$, $1 \leqslant k' \leqslant \frac{n-4}{2}$. По лемме из решения задачи 1.4 точка P является вершиной каждого из графов $G_{k+1}(n)$, $G_{k+2}(n)$, …, $G_{(n-4)/2}(n)$ и имеет в каждом из них степень не меньше 2. Поскольку эти графы не имеют общих ребер, мы получаем оценку на степень вершины P:

$$\deg P \leqslant n-2-2\Big(\frac{n-4}{2}-k\Big) = 2k+2.$$

Для нечетного n в этих рассуждениях следует заменить $n-2$ на $n-1$, а $\frac{n-4}{2}$ на $\frac{n-3}{2}$.

4.2. Ответ: максимальный возможный путь содержит $n-1$ ребро.

Мы взяли это утверждение в [6, лемма 5.5].

По утверждению задачи 2.2 никакой путь в половинчатом графе не может содержать n вершин. С другой стороны, нетрудно предъявить конструкцию пути с $n-1$ вершиной. Рассмотрим $(n-2)/2$ точки, лежащие на графике вогнутой функции. Реализуем эту конфигурацию в виде толстой линии, идущей вдоль прямой $y=x$ в первом квадранте. Аналогичную конфигурацию разместим вдоль прямой $y=-x$ во втором квадранте. Добавим две точки: $(0;-1)$ и $(0;-2)$. Половинчатый граф такой конфигурации содержит V-образный путь, проходящий через все вершины, кроме $(0;-2)$ (рис. 7).

4.3. Ответ: при четном $n \notin \{2,8\}$ максимальная возможная длина цикла равна $n-3$.

Как мы знаем, при $n>2$ по крайней мере три вершины половинчатого графа (крайние точки выпуклой оболочки) имеют степень 1. Поэтому никакой цикл не может иметь больше $n-3$ ребер.

Приведем пример для $n=6k$, $k \geqslant 2$ [6, теорема 5.7]. Возьмем V-образную конструкцию на $2k$ точках, описанную в предыдущей задаче; построим её так, чтобы вся левая ветвь пути A–B при проекции на вертикаль попадала в проекцию первого отрезка правой ветви (рис. 8). Реализуем 3 одинаковых экземпляра такой V-образной конструкции в виде «толстых линий», пересекающихся под углом 120°. Ребра этих трех путей остаются половинными ребрами построенной конфигурации точек. Кроме того, соединим точку B каждого пути с точкой A следующего по часовой стрелке пути — эти отрезки тоже будут половинными ребрами (рис. 9). Получился цикл длины $6k-3$.

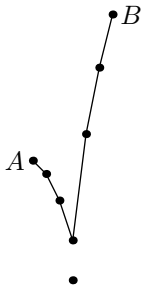


Рис. 8. Подготовим V-образную конструкцию

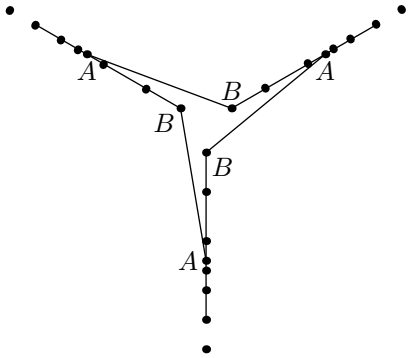


Рис. 9. Склеим 3 пути

Покажем, как можно модифицировать эту конструкцию, чтобы получился пример для $n = 6k + 2$, $k \geqslant 2$. Добавим две новые точки — C_1 и C_2, поместив их внутри закрашенных секторов вблизи отрезков B_2A_3 и B_3A_1, как показано на рис. 10. Как нетрудно видеть, для полученной конфигурации половинные ребра, находившиеся внутри толстых линий, так и остались половинными ребрами.

Заменим ребро B_2A_3 на пару ребер B_2C_2, C_2A_3, а B_3A_1 — на пару ребер B_3C_1, C_1A_1. Если точка C_2 достаточно близка к B_2A_3, то снизу от прямой B_2C_2 находятся те же точки, что и снизу от B_2A_3, а также точка A_3, а сверху от B_2C_2 находятся те же точки, что сверху от B_2A_3, а также новая точка C_1. Таким образом, B_2C_2 — половинчатое ребро новой конфигурации. Аналогично, C_2A_3, B_3C_1 и C_1A_1 — тоже половинчатые ребра.

Нам осталось позаботиться о том, чтобы ребро A_2B_1 также стало половинчатым, потому что прямо сейчас снизу от прямой A_2B_1 на две точки больше, чем сверху. Заменим фрагмент B_2–A_2 нашего пути, на чуть-чуть исправленный вариант исходной V-конструкции (рис. 11), а именно, потребуем, чтобы вся левая половина пути A_2B_2, кроме точки A_2, при проекции на вертикаль попадала внутрь проекции первого звена D_1D_2 правой половины, а проекция точки A_2 при проекции на вертикаль попадала бы внутрь проекции второго звена D_2D_3. Теперь, по сравнению с ситуацией до исправления, одна точка снизу от исправленного ребра B_1A_2 (точка D_2) перешла в верхнюю полуплоскость, и ребро B_1A_2 тоже стало половинным.

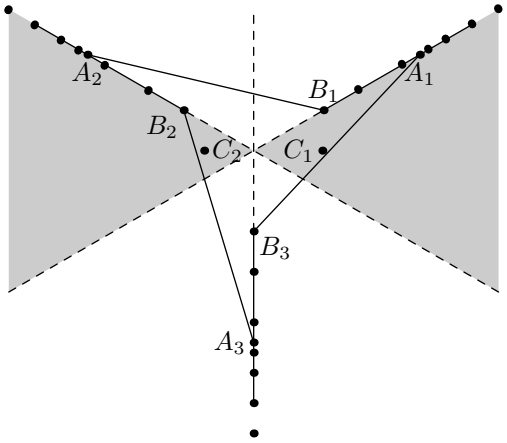


Рис. 10. Добавим две точки

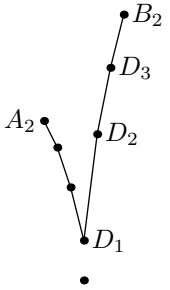


Рис. 11. Приподнимем точку A

Похожим образом можно построить цикл для $n = 6k - 2$. Ограничимся примером для $n = 10$ (рис. 12). В общем случае вместо фрагментов B_1–C_1–A_1, B_2–C_2–A_2, B_3–C_3–A_3 следует взять V-образные конструкции из $2k$ точек. Следует правильно подобрать высоту отдельных половинок этих конструкций, чтобы ребро A_1B_2 стало половинчатым.

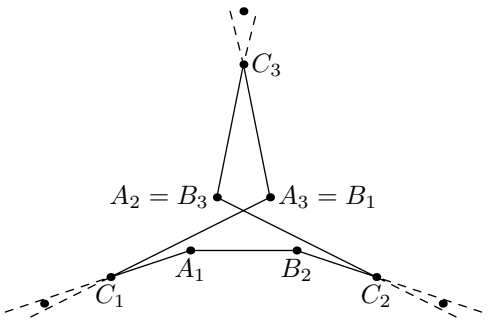


Рис. 12. Случай $n = 6k + 4$

Разбор случая $n = 8$ оставляется читателю; в этом случае максимальный возможный цикл имеет длину 4.

4.4. a) В решении задачи 2.7b) описано, как для любого заданного графа K построить половинчатый граф, содержащий K как индуцированный подграф. Конструкция состоит в том, что сначала рисуют граф K, а потом для каждого его ребра (или антиребра) добавляют новые точки. На каждом шаге число добавляемых точек не превосходит числа вершин графа K. Взяв в этой конструкции в качестве K полный граф на k вершинах, мы построим половинчатый граф, в котором не больше k^3 новых вершин.

b) Это утверждение из [6, теорема 5.9].

Можно считать, что k четное. Для начала расположим k точек будущей клики в вершинах правильного k-угольника. Тогда стороны и диагонали многоугольника разбиваются на k групп параллельных между собой линий по $k/2$ или $k/2-1$ линий в одной группе. Сделаем проективное преобразование, при котором бесконечно удаленная прямая переходит, скажем, в вертикальную линию ℓ, расположенную «далеко» справа от образа многоугольника. Тогда вершины многоугольника окажутся в общем положении, а семейство прямых, которые раньше было параллельным, теперь будет «почти параллельным» и будет пересекаться в какой-то точке прямой ℓ. Опишем, как добавить к этой конфигурации некоторое количество точек, чтобы все стороны и диагонали k-угольника стали половинчатыми линиями.

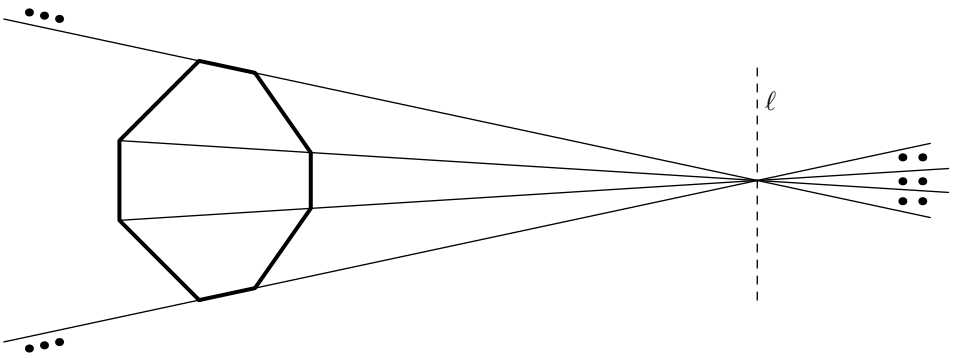


Рис. 13. Строим клику

Рассмотрим одно семейство почти параллельных прямых. Между каждыми двумя соседними прямыми чуть справа от прямой ℓ поставим по две дополнительные точки — всего их будет $k-2$ или $k-4$ в зависимости от семейства (на рис. 13 $k=8$ и справа добавлено $k-2$ точки). В результате каждая прямая из нашего семейства стала половинчатой, но нужно еще позаботиться о том, чтобы выполнение этой конструкции не мешало делать половинчатыми прямые из других семейств.

Для этого добавим еще $k-2$ или $k-4$ точки далеко слева от многоугольника. Эти точки добавим двумя равными половинами — половину точек чуть сверху от самой верхней прямой рассматриваемого семейства, другую половину — чуть снизу от самой нижней (на рис. 13 слева добавлено по $(k-2)/2=3$ точек сверху и снизу). В результате этого прямые нашего семейства так и остались половинчатыми, а по отношению к остальным прямым разность количества точек в полуплоскостях после всех этих добавлений не изменилась.

Выполним эту конструкцию для каждого семейства почти параллельных прямых. В результате мы построим k-клику, добавив в сумме менее $2k^2$ точек.

4.5. Это утверждение из [6, следствие 6.13].

Представим себе, что прямая ℓ пересекает m ребер половинчатого графа. Сделаем ее вертикальной и построим разбиение нашего графа на выпуклые цепочки. Тогда ребра, пересеченные прямой ℓ, принадлежат разным выпуклым цепочкам. Значит, в графе не менее m цепочек, и следовательно, не менее $2m$ вершин. Для решения задачи осталось в качестве прямой ℓ выбрать прямую, не проходящую через вершины графа, в полуплоскостях которой находятся по $k/2$ точек клики, если k четное, и по $(k\pm1)/2$ точек, если k нечетное. Такая прямая пересекает $k^2/4$ ребер (соответственно, $(k^2-1)/4$ ребер), откуда получаем, что число вершин графа не меньше $[k^2/2]$.

4.6. Это утверждение из [5, теорема 4.5].

Сначала заметим, что в графе $G_k(\mathcal{S})$ с максимальным числом ребер должно быть n вершин. Действительно, если точка A не является вершиной графа $G_k(\mathcal{S})$ (по лемме из решения задачи 1.4 это значит, что точка A лежит «глубоко внутри» конфигурации $\mathcal{S}$), начнем двигать ее «наружу», как это делалось в решении задачи 2.5. Это приведет к увеличению числа ребер.

Чтобы перейти от конфигурации $\mathcal{S}$ к конфигурации с удвоенным числом точек, сопоставим каждой вершине P какое-нибудь выходящее из нее ребро E_P. Теперь заменим каждую вершину P конфигурации $\mathcal{S}$ на две близкие точки P_1 и P_2 так, чтобы вектора $\overrightarrow{P_1P}$ и $\overrightarrow{PP_2}$ были сонаправлены с E_P и имели малую длину. Пусть $\mathcal{S}'$ — полученное множество точек.

Рассмотрим граф $G_{2k+1}(\mathcal{S}')$. Пусть $E_P=\overrightarrow{PQ}$, тогда лишь одна из точек Q_1, Q_2 лежит справа от E_P, пусть это будет Q_1. Очевидно, $\overrightarrow{P_1P_2}$ — ребро графа $G_{2k+1}(\mathcal{S}')$, так как мы удвоили число точек в каждой полуплоскости по сторонам от прямой PQ, и еще справа от нее добавилась точка Q_1. Далее, $\overrightarrow{P_2Q_1}$ и $\overrightarrow{P_2Q_2}$ — тоже ребра $G_{2k+1}(\mathcal{S}')$. Это так, например, для $\overrightarrow{P_2Q_1}$, поскольку прямая P_2Q_1 почти параллельна прямой P_1P_2 и наборы точек из $\mathcal{S}$ слева от этих прямых одинаковы. Наконец, если $\overrightarrow{PR}\ne\overrightarrow{PQ}$ — еще какое-нибудь ребро графа $G_k(\mathcal{S})$, то среди векторов $\overrightarrow{P_iR_j}$ нетрудно отыскать еще пару ребер графа $G_{2k+1}(\mathcal{S}')$ (на рис. 14 это ребра $\overrightarrow{P_1R_2}$ и $\overrightarrow{P_2R_1}$).

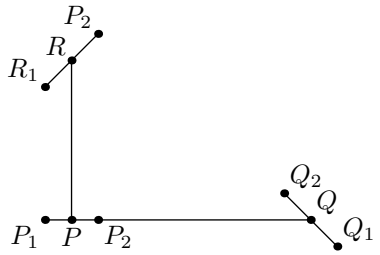


Рис. 14. Удваиваем вершины

Итак, каждому старому ребру в новом графе соответствует два ребра, а для ребер вида E_P (число которых равно числу вершин P, т. е. n) — даже три ребра! Поэтому $e_{2n,2k+1} \geqslant 2e_{n,k} + n$.

4.7. Это — теорема 4.3 в [5]. В процессе написания решения оценка несколько улучшилась; более того, и это решение ещё можно чуть улучшить.

Пусть $\mathcal{S}$ — конфигурация, задающая наш граф $G_k(n)$. По задаче 2.6, можно считать, что множество вершин графа — также $\mathcal{S}$. Можно считать, что ни одна из прямых, соединяющих точки $\mathcal{S}$, не вертикальна. Проведем $n-1$ вертикальных прямых ℓ_1, …, ℓ_{n-1}, делящих плоскость на n частей, содержащих по точке из $\mathcal{S}$. Следующая лемма — это обобщение наблюдения из решения задачи 4.5.

Лемма. Прямая ℓ_i пересекает ровно $m_i = \min\{i, k+1, n-i\}$ ребер графа $G_k(\mathcal{S})$, идущих справа налево, и столько же ребер, идущих слева направо.

Доказательство. Запустим процесс, аналогичный построению выпуклых цепочек, в нашем графе; вращающаяся прямая будет ориентированной (в начале цепочки — вверх, в конце — вниз), и мы будем переходить к следующей вершине по ребру, лежащему на нашей прямой и *противонаправленному* с ней. Тогда, аналогично решению задачи 3.1, можно показать, что начальные точки цепочек — это ровно левые $k+1$ точек, конечные — ровно правые $k+1$ точек, и каждое ребро, идущее справа налево, лежит ровно в одной цепочке. (При этом используется полный аналог леммы из решения задачи 1.2.) Итак, каждая «средняя» прямая ℓ_i (при $k+1 \leqslant i \leqslant n-k-1$) пересекает ровно $k+1$ цепочку; каждая «левая» прямая ℓ_i при $i \leqslant k$ — ровно i цепочек, начинающихся левее; наконец, каждая «правая» прямая ℓ_i при $i \geqslant n-k$ — ровно $n-i$ цепочек, кончающихся правее. Отсюда и следует утверждение леммы. □

Согласно лемме, общее количество пересечений прямых ℓ_i с ребрами графа (идущими в обе стороны) равно

$$N = 2(m_1 + \cdots + m_{n-1}) = 2 \cdot 2 \cdot (1 + 2 + \cdots + k) + (n - 2k - 1) \cdot 2(k+1) = 2(k+1)(n-k-1).$$

Зафиксируем теперь некоторое натуральное ℓ; пусть E_1 — множество ребер, каждое из которых пересекает хотя бы ℓ прямых, а E_2 — множество всех остальных ребер. Тогда количество ребер в E_1 не превосходит N/ℓ, а количество ребер в E_2 не превосходит количества пар вершин, между которыми меньше ℓ прямых, то есть $n(\ell-1) - \ell(\ell-1)/2 < n(\ell-1)$ (заметим, что каждая пара соединена максимум одним ребром). Итак, общее число ребер не превосходит $n(\ell-1) + N/\ell$. Наконец, полагая $\ell = \lceil \sqrt{N/n} \rceil$, мы получаем, что общее число ребер не превосходит

$$n\sqrt{N/n} + \frac{N}{\sqrt{N/n}} = 2\sqrt{Nn} = 2\sqrt{2n(k+1)(n-k-1)}.$$

5.1. Ясно, что k-множество меняется ровно тогда, когда происходит k-флип; значит, количество k-множеств не превосходит количества k-флипов (поскольку в начале и в конце двойной круговой последовательности стоит одна и та же перестановка).

Осталось показать, что после всех k-флипов появляются разные множества. Это следует из пункта в) определения круговой последовательности. Действительно, пусть некоторое k-множество A появилось дважды, и пусть после его первого появления первый k-флип менял местами числа $x \in A$ и $y \notin A$. Чтобы множество A появилось повторно, число x должно стать левее, чем y, то есть они должны поменяться еще раз, что произойдет только через C_n^2 флипов. Таким образом, между появлениями множества A произошло хотя бы $C_n^2 + 1$ флипов. Но тогда, если продолжить нашу двойную последовательность периодично, мы увидим, что между вторым и третьим появлениями A проходит меньше, чем C_n^2 флипов; это невозможно, как показано выше.

5.2. Это вопрос на понимание определений. По определению, всякая двойная круговая последовательность получается с помощью $2C_n^2$ флипов. А по утверждению предыдущей задачи число k-флипов в двойной круговой последовательности $\mathcal{T}$ равно числу k-множеств $s_k(\mathcal{T})$. Таким образом, доказываемое равенство двумя способами подсчитывает количество флипов.

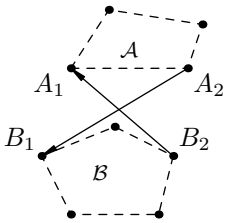


Рис. 15. Соответствие ребер графа $G_k(\mathcal{S})$ и $(k+1)$-множеств

5.3. Это утверждение из [3, теорема 3.14].

Для выпуклого n-угольника при каждом k имеется n k-множеств. Поэтому $S_k(n) \geqslant kn$.

Докажем неравенство $S_k(n) \leqslant kn$. По утверждению задачи 5.1, вместо k-множеств можно подсчитывать k-флипы. Возьмем произвольную двойную круговую последовательность. Каждое фиксированное число m, $1 \leqslant m \leqslant n$, поучаствовало в этой последовательности ровно в $2n-2$ флипах, поскольку оно дважды менялось местами с каждым из остальных чисел. Кроме того, в силу симметрии второй половины последовательности по отношению к первой половине, m участвовало i-флипах столько же раз, сколько оно поучаствовало $(n-i)$-флипах.

Обозначим через p номер самого левого места, на котором побывало число m, тогда номер самого правого его положения — это $(n-p+1)$. Поскольку нас интересуют лишь i-флипы при $i \leqslant k$, разберемся, что происходит, если $p \leqslant k < n/2$. В этом случае число m по дороге от p-й позиции к $(n-p)$-й и обратно по крайней мере дважды побывало в столбцах с номерами $k+1$, $k+2$, …, $n-k$. Значит, оно поучаствовало не менее чем в двух i-флипах для всех i, $k+1 \leqslant i \leqslant n-k-1$, т. е. в сумме в $2n-4k+2$ флипах (из $2n-2$ возможных). Таким образом, число m поучаствовало не более чем в $4k$ i-флипах для $i \leqslant k$ и $i \geqslant n-k$. Тогда в силу уже упомянутой симметрии получаем, что для $i \leqslant k$ число m поучаствовало не более чем в $2k$ i-флипах. Таким образом, общее количество флипов не более $2kn$. При этом каждый флип мы подсчитали дважды, поскольку во флипе участвуют 2 числа.

5.4. Эта задача — непосредственное следствие из задачи 5.1, поскольку две вершины, участвующие в $(k+1)$-флипе, как раз и определяют ребро графа $G_k(n)$. Тем не менее, мы приведём ее непосредственное решение.

Рассмотрим произвольное $(k+1)$-множество $\mathcal{A}$, пусть $\mathcal{B} = \mathcal{S} \setminus \mathcal{A}$. Тогда множества $\mathcal{A}$ и $\mathcal{B}$ лежат в разных полуплоскостях, их выпуклые оболочки не пересекаются, и значит, имеют две внутренние совместные опорные прямые, скажем, A_1B_2 и A_2B_1. Ориентируем отрезки так, чтобы множество $\mathcal{A}$ лежало справа от этих прямых. Тогда лишь один из отрезков окажется ориентирован в сторону вершины из $\mathcal{A}$ (на рис. 15 это отрезок $\overrightarrow{B_2A_1}$), и этот отрезок является ребром графа $G_k(\mathcal{S})$.

Мы сопоставили каждому $(k+1)$-множеству ребро в графе $G_k(\mathcal{S})$. Нетрудно понять, что это соответствие взаимно однозначно.

5.5. Ответ: нет. Заметим, что (двойная) круговая последовательность, построенная по множеству точек $\mathcal{S}$, позволяет восстановить многие свойства множества $\mathcal{S}$. Приведем два таких свойства.

(1) Точка a является вершиной выпуклой оболочки множества $\mathcal{S}$ ровно тогда, когда в одной из перестановок она стоит с краю.

(2) Для трех прямых, соединяющих пары точек i и j, k и ℓ, p и q, можно определить их «порядок» по часовой стрелке (то есть, можно установить, направление какой из второй и третьей прямой получится раньше, если начать вращать первую по часовой стрелке). Действительно, это — просто порядок, в котором в нашей круговой последовательности происходят флипы соответствующих пар точек.

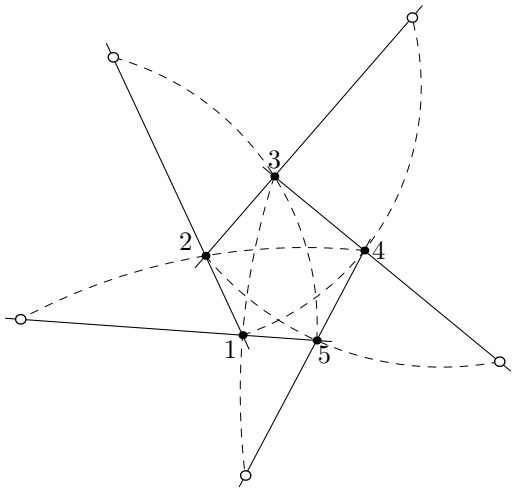
Итак, для построения примера мы рассмотрим (гипотетическую) конфигурацию точек $\mathcal{S}$ с несовместимыми свойствами и построим круговую последовательность по ней. Эти конфигурация и последовательность изображены на рис. 16. Из свойств (1) и (2) мы понимаем, что точки 1–5 образуют выпуклый пятиугольник (именно в этом порядке), и отмеченные точки пересечения прямых находятся именно там, где обозначено на чертеже.

Осталось показать, что это невозможно. Рассмотрим пять треугольников, образованных тремя последовательными вершинами пятиугольника; пусть, например, 125 — треугольник наименьшей площади. Тогда луч 53 не может пересекать прямую 12 — противоречие.

5.6. Это утверждение из [3, лемма 3.16].

Количество k-флипов с участием сразу двух чисел из Y, очевидно, не превосходит C_y^2 (на самом деле, количество всех таких флипов равно C_y^2). Оценим число k-флипов, затрагивающих лишь одно число из Y.

Покрасим все числа из X в белый цвет, все числа из Z — в черный. Далее будем считать, что мы не различаем между собой различные белые и различные черные числа. Реорганизуем последовательность $\mathcal{T}$.



(2 1 3 5 4)
(2 3 1 5 4)
(2 3 1 4 5)
(2 3 4 1 5)
(3 2 4 1 5)
(3 4 2 1 5)
(3 4 2 5 1)
(3 4 5 2 1)
(4 3 5 2 1)
(4 5 3 2 1)
(4 5 3 1 2)

Рис. 16. Так не бывает

Сначала выбросим из нее все флипы, которые меняют местами два белых или два черных числа. После этого заново пронумеруем в первой перестановке все белые и черные числа (и перенумеруем числа во всех остальных перестановках так, чтобы полученные перестановки отличались на флипы с теми же номерами, что и раньше). Тогда при выполнении последующих флипов белые числа будут стоять в порядке возрастания номеров и черные числа будут стоять в порядке возрастания номеров; более того, белые числа двигаются только вправо (меняясь с черными и с числами из Y), а черные — только влево.

Полученная последовательность перестановок $\mathcal{T}^*$ является начальным куском некоторой круговой последовательности, в которой осталось выполнить флипы между белыми числами и между черными числами. Поскольку белые числа двигаются только вправо, каждое черное и каждое белое число участвует не более, чем в одном k-флипе. Более точно, лишь k самых левых белых чисел и k самых левых черных чисел могут участвовать в k-флипе. Таким образом, число k-флипов, затрагивающих лишь одно число из Y, не превосходит $2k$.

5.7. Это утверждение из [3, теорема 3.17].

Пусть круговая последовательность $\mathcal{T}_1$ — это первая половина последовательности $\mathcal{T}$. Обозначим $r = \lceil \sqrt{k} \rceil$. Разобьем $\mathcal{T}_1$ на $m = \lceil n/r \rceil$ подпоследовательностей, каждая из которых имеет длину r (кроме последней, которая короче); элементы каждой подпослеовательности — последовательные перестановки. По утверждению предыдущей задачи элементы одной подстроки в сумме участвуют не более чем в $C_r^2 + 2k < 3k$ k-флипах. Следовательно, общее число k-флипов не превосходит

$$3km \leqslant 3k\Big(\frac{n}{\sqrt{k}} + 1\Big) = 3n\sqrt{k} + 3k < 4n\sqrt{k}.$$

Тогда во всей последовательности $\mathcal{T}$ количество k-флипов не превосходит $8n\sqrt{k}$.

6.1. Это утверждение доказано в [2].

Проверяем, что левая часть не меняется, если двигать точки. Будем двигать вершину r половинчатого графа G. Как нетрудно видеть, граф не меняется, если при движении вершина r не оказывается на прямой, проходящей через две другие вершины — p и q. А если это событие происходит, изменения касаются только ребер, проходящих между вершинами p, q и r.

Рассмотрим сначала случай, когда pq — ребро в графе G, а вершина r при движении пересекает отрезок pq (рис. 17). Тогда до момента пересечения отрезки qr и pr не были, а после пересечения стали ребрами нашего графа. В то же время до момента пересечения отрезок pq был ребром графа, а после пересечения — перестал им быть. Степени вершин графа в результате не изменились, кроме вершины r, у которой степень увеличилась на 2, т.е. $d'_r = d_r + 2$, где d'_r — степень вершины r после момента пересечения.

Как изменилось количество пересечений ребер? Для ребер, не инцидентных вершине r, число пересечений не изменилось: пересечение с ребром pq, если оно было, заменилось на пересечение с ребром pr или qr. А вот число пересечений, в которых участвуют ребра, выходящие из r, изменилось. Проведем через r прямую ℓ, параллельную pq. При движении вершины r в полуплоскости относительно ℓ, содержащей p и q, всегда находится на одну точку больше, чем во второй полуплоскости. Пусть до момента пересечения из вершины r в эту вторую полуплоскость выходило x ребер (которые давали x неучтенных точек пересечения с pq). Как и в решении задачи 1.2, это значит, что во вторую полуплоскость из вершины r выходило $x - 1$ ребер. Таким образом, $x = \frac{d_r+1}{2}$. В результате движения вершины r эти x точек пересечения исчезли, и никаких новых точек пересечения не появилось. Итак после момента пересечения количество пересечений равно $X' = X - x$.

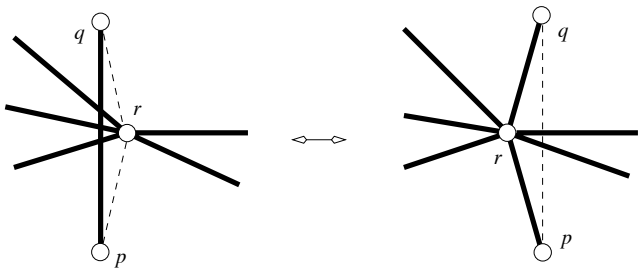


Рис. 17. Изменение графа при движении вершины

Тогда

$$X + C^2_{(d_v+1)/2} = X + C^2_x = X - x + C^2_{x+1} = X' + C^2_{(d'_v+1)/2}$$

Следовательно, в рассматриваемом случае сумма в левой части доказываемого равенства не меняется при движении точки.

Осталось заметить, что только разобранный случай и «противоположный» ему влияют на изменение графа G и конфигурации пересечений. (В ситуации, когда r пересекает, скажем, продолжение отрезка pq за точку q, можно локально считать, что q движется, а r остается неподвижной.) Для завершения доказательства передвинем исходные вершины в выпуклое положение и заметим, что в этом случае доказываемое равенство тривиально.

6.2. Эта конструкция взята из [8]. Мы проведем «качественное» рассуждение, оставляя количественные оценки читателю.

Возьмем все узлы квадратной решетки в квадрате $\sqrt{n}\times\sqrt{n}$; его можно выбрать так, чтобы точек оказалось $\lfloor\sqrt{n}\rfloor^2 \leqslant n$. Впоследствии мы добавим несколько изолированных вершин.

Далее, соединим все пары узлов, находящиеся на расстоянии не больше $d = \sqrt{2e/n}$. Из каждого узла, находящегося далеко от вершины квадрата, выходит не менее d^2 ребер, так что всего ребер будет не меньше e; выкинем несколько ребер так, чтобы осталось ровно e. Далее, каждое ребро имеет в окрестности радиуса d порядка d^2 узлов, так что его пересекает как максимум порядка d^4 ребер, и общее количество пересечений по порядку не больше $ed^4 \sim e^3/n^2$. Осталось добавить несколько вершин, чтобы их число стало равно n.

6.3. Это — классическая теорема дискретной геометрии; её доказали T. Leighton и, независимо, четыре автора: M. Ajtai, V. Chvátal, M. Newborn, E. Szemerédi. Суть этого доказательства — вероятностная; мы изложим его на комбинаторном языке. Ясно, что достаточно получить оценку только для больших значений n. Мы начнем с гораздо более слабой оценки.

Лемма. Пусть в геометрическом графе n вершин и e ребер. Тогда в нем не менее $e - 3n$ пересечений.

Доказательство. Пусть количество пересечений равно d. Удалив из каждого пересечения по ребру, получим плоский граф на n вершинах с $e' \geqslant e - d$ ребрами. Пусть f — количество граней в нем; тогда, поскольку все они — хотя бы треугольники, имеем $3f \leqslant 2e'$, а из формулы Эйлера (для, возможно, несвязного графа) получаем $0 < n - e' + f \leqslant n - e'/3$; итак, $n > e'/3 > (e - d)/3$, откуда $d > e - 3n$. □

Перейдем к решению задачи. Пусть H — произвольный индуцированный подграф в G на k вершинах; число k мы выберем позже. Пусть $e(H)$ — число ребер в H. По лемме, в этом графе хотя бы $e(H) - 3k$ пересечений. Теперь мы собираемся оценить число пересечений в G, просуммировав эту оценку по всем графам H.

Количество таких графов равно C^k_n. Каждое ребро из G участвует в C^{k-2}_{n-2} таких графах, так что сумма всех величин $e(H)$ равна $e \cdot C^{k-2}_{n-2}$. Наконец, каждая пара пересекающихся ребер войдет в C^{k-4}_{n-4} графов, ибо все 4 их вершины должны туда войти. Итого, число пересечений в G не меньше, чем

$$\frac{e \cdot C^{k-2}_{n-2} - 3k \cdot C^k_n}{C^{k-4}_{n-4}} = \frac{e(n-2)(n-3)}{(k-2)(k-3)} - 3 \cdot \frac{n(n-1)(n-2)(n-3)}{(k-1)(k-2)(k-3)} \approx \frac{ekn^2 - 3n^4}{k^3}.$$

Полагая $k \approx 4n^2/e$, получаем, что наше количество не меньше, чем

$$\sim \frac{4n^4 - 3n^4}{64n^6/e^3} = \frac{e^3}{64n^2},$$

что и требовалось.

6.4. Из задачи 6.1 следует, что количество пересечений X в половинчатом графе не превосходит $C^2_{n/2} < n^2/8$. По предыдущей задаче, оно не меньше $c \cdot e^3/n^2$; значит, $e \leqslant \sqrt[3]{cn^4/8}$, что и требовалось.

Литература

[1] Задачник «Кванта», задача 2245 // Квант. 2011. № 5–6. Квант. 2012. № 2.

[2] *Andrzejak A., Aronov B., Har-Peled S., Seidel R., WelzlE.* Results on k-Sets and j-Facets via Continuous Motion // Proc. 14th Ann. ACM Symp. on Comput. Geom. 1998. P. 192–199.

[3] *Edelsbrunner H.* Algorithms in combinatorial geometry. Springer-Verlag, 1987.

[4] *Eppstein D.* Sets of points with many halving lines.
`http://www.ics.uci.edu/~eppstein/pubs/Epp-TR-92-86.pdf`

[5] *Erdős P., Lovász L., Simmons A., Straus E.G.* Dissection graphs of planar point sets. In A Survey of Combinatorial Theory (J. N. Srivastava, editor). North-Holland, 1973. P. 139–154.

[6] *Hovanova T., Day Yang.* Halving Lines and Their Underlying Graphs // arXiv.org:1210.4959v1

[7] *Pach J., Solymosi J.* Halving lines and perfect cross-matchings
`http://www.renyi.hu/~pach/publications/halving.pdf`

[8] *Pach J.* Crossing numbers
`http://www.renyi.hu/~pach/publications/halving.pdf`

Halving graphs

I. Bogdanov, K. Kokhas

In this set of problems we deal with *geometrical graphs*, i.e. graphs drawn in the plane. We assume that no three vertices of a geometrical graph are collinear, and all the edges are depicted by segments; these segments may have common points different from vertices. The following two geometrical graphs are of the main interest.

Definition. 1) A *halving graph* $G(n)$ is a graph constructed as follows. Consider a set $\mathcal{S}$ of n points in the plane (where n is even), no three points being collinear; they form the set of vertices of our graph. A pair of vertices is connected by an edge exactly if the line passing through these points splits the whole set of vertices into two equal parts (that means that each open halfplane defined by this line contains $(n-2)/2$ points).

2) Consider also a more general case; let us define a *k-separating graph* $G_k(n)$. Let k and n be nonnegative integers with $n > 2k+2$ (now n is not necessarily even). Consider a set $\mathcal{S}$ of n points in the plane, no three of them being collinear. An **oriented** edge is drawn from A to B exactly if one of the two (open) halfplanes defined by the line AB contains exactly k points; this halfplane must be to the *right* of the line, when we go along the line from A to B. The set of vertices of the graph $G_k(n)$ consists of those points that have at least one edge (either ingoing or outgoing one). Sometimes we will also denote this graph by $G_k(\mathcal{S})$, when we wish to stress its dependence of the initial set $\mathcal{S}$.

1 Vertices

1.1. Prove that a halving graph contains no isolated vertices.

1.2. For a fixed value of n, determine all the values a degree of a vertex in a graph $G(n)$ can attain.

1.3. Does there exist a graph $G(50)$ containing 25 vertices of degree 1 and 25 vertices of degree 3?

1.4. Prove that each k-separated graph $G_k(n)$ contains at least $2k+3$ vertices.

1.5. a) Prove that a halving graph contains at most 3 vertices of degree $n-3$.
b) How many vertices of degree $n-3$ a halving graph may contain?

1.6. Does there exist a halving graph containing exactly 8 vertices and exactly 9 edges?

1.7. Prove that any halving graph $G(100)$ contains at most 60 vertices of degree 41.

2 Properties of graphs

2.1. Assume that a halving graph $G(2n)$ contains exactly n edges. Prove that each two of the segments representing these edges have a common point.

2.1$\frac{1}{2}$. A geometrical oriented graph is drawn in the plane. Prove that it cannot have both types $G_{10}(n)$ and $G_{15}(n')$.

2.2. Prove that no halving graph $G(2n)$ contains a Hamiltonian path (i.e. a path that passes through every vertex exactly once).

2.3. Let $n = 103$. Find all k for which a k-separated graph $G_k(n)$ is necessarily connected.

2.4. Prove that each connected component in a k-separated graph $G_k(n)$ contains an Eulerian path (i.e. a path that passes along every edge exactly once).

2.5. Assume that there exist two halving graphs on n vertices that contain k_1 edges and k_2 edges, respectively. Prove that for each m with $k_1 \leqslant m \leqslant k_2$ there exists a halving graph on n vertices having m edges.

2.6. a) Consider a k-separated graph $G_k(n)$ depicted on the plane; let $\mathcal{S}'$ be the set of its vertices. Prove that this graph coincides with the graph $G_{k'}(\mathcal{S}')$ for some k'.
b) Prove that each connected component of a graph $G_k(n)$ is also of the form $G_{k'}(n')$ for some k' and n'.

2.7. a) Prove that each (abstract) graph is a subgraph of some halving graph.
b) Prove that each (abstract) graph is an induced subgraph of some halving graph.

3 Convex chains and windmills

We will group all the edges of a halving graph into several *convex chains*. For this, let us first rotate a graph so that none of its edges is vertical, and draw a vertical line (not passing through the vertices) that has equal number of vertices on both sides of it. Now, we draw a vertical line ℓ through the leftmost vertex V_1. Rotate this line clockwise around V_1 until it passes through some edge, say V_1V_2. Next, we continue rotation of the line ℓ around the vertex V_2 clockwise until it passes through a next edge, say V_2V_2, and so on. If the line becomes vertical, we terminate the process. Now we say that a polyline $V_1V_2V_3\ldots$ is a *convex chain*. After that we start a new process taking as a starting point the leftmost vertex that contains no edge lying in the chains yet constructed.

As a result, after there are no more unused edges, all the edges will be partitioned into several convex chains. Notice that this partition depends on the direction initially chosen as vertical.

3.1. Prove that the process above partitions the edges of a halving graph into exactly $n/2$ convex chains, each chain starting in the left halfplane and ending in the right halfplane; moreover, no two chains have a common edges.

3.2. Prove that the sum of degrees of any two vertices of a halving graph does not exceed n.

In the next series of problems we deal with a finite set $\mathcal{S}$ of points in the plane, no three of which are collinear.

Definition. A *windmill* is the following process. Choose a line ℓ that passes through a single point $T \in \mathcal{S}$. This line rotates clockwise about the pivot T until the first time that the line meets some other point belonging to $\mathcal{S}$. This point, U, takes over as a new pivot, and the line ℓ now rotates clockwise about U, until it next meets a point of $\mathcal{S}$, and so on.

3.3. Prove that one can choose a point T and a starting line ℓ so that the resulting windmill uses each point of $\mathcal{S}$ as a pivot infinitely many times.

3.4. Prove that for each set $\mathcal{S}$ there exists a windmill such that the windmill line sweeps all the points of the plane.

3.5. Prove that for each set $\mathcal{S}$ there exists a point in it such that each windmill starting from this point passes through all the points of $\mathcal{S}$.

3.6. Consider a finite set of points in the plane, no three being collinear, and a line a that passes through no points of the set. Let us colour all the points on one side of a in red, and all the other marked points in blue; assume that there are K red and M blue points. Prove that for each $k < K$ and $m < M$ there exists a line b such that one of halfplanes defined by b contains exactly k red and m blue points.

4 Extremal problems

4.1. a) Does there exist a positive integer n and a halving graph with n vertices and $2013n$ edges?
b) Does there exist a positive integer n and a graph $G_{10}(n)$ with n vertices and $2013n$ edges?

4.2. Find the maximal number of edges in a path in a halving graph $G(n)$.

4.3. Find the maximal number of edges in a cycle in a halving graph $G(n)$.

4.4. Prove that a clique of size k can be a subgraph of a halving graph that has
a) $O(k^3)$ vertices; b) $O(k^2)$ vertices.

4.5. Prove that the statement b) of the previous problem is asymptotically exact; more exactly, prove that if a halving graph with n vertices contains a clique of size k, then $n \geqslant \lfloor k^2/2 \rfloor$.

4.6. Denote by $e_{k,n}$ the maximum of number of edges in a k-separated graph $G_k(n)$. Prove that

$$e_{2n,2k+1} \geqslant 2e_{n,k} + n.$$

4.7. Prove that a graph $G_k(n)$ for $k < (n-2)/2$ contains at most $4\sqrt{(k+1)(n-k-1)n}$ edges.

5 Circular sequences

In this section we suggest some approach to problem 4.7; in this approach, we rephrase our problem in different terms. Thus, problems in this section look quite different from the other ones.

Definition. 1) A *circular n-sequence* is a sequence of $\binom{n}{2}+1$ permutations of the set $\{1, 2, \ldots, n\}$ satisfying the following properties:

a) for every two consecutive permutations, one of them may be obtained from the other by swapping two neighboring numbers (such a swapping hereafter is called a *flip*, more specifically a *k-flip* if it swaps the numbers on kth and $(k+1)$st positions);

b) the last permutation of a sequence is obtained from the first one by reversing the order of all elements;
c) every two numbers from the set $\{1, 2, \dots, n\}$ participate in exactly one flip of the sequence.

2) A *double circular n-sequence* is a sequence of $2\binom{n}{2}+1$ permutations of the same set, such that the first $\binom{n}{2}+1$ its elements form a circular n-sequence, and the last elements form the "reflection" of this sequence. I.e., the $(\binom{n}{2}+i)$th permutation is obtained from the ith one by reversing the order of all elements; thus, if the ith and $(i+1)$st permutations differ by a k-flip, then their reflections differ by a $(n-k)$-flip, and these flips swap the same numbers.

Some circular sequences may be constructed in the following geometrical way. In the plane, consider a set $\mathcal{S}$ of n points in a general position (we assume also that the lines connecting these points are pairwise non-parallel). Let us take some line and project these points onto this line; we will obtain some permutation of our points on the line. Now, we start rotating the line around some fixed center; the order of the points will sometimes change by a flip, so we will receive some sequence of the permutations. When a line rotates at 180° (360°), we will get a (double) circular sequence.

3) Let $\mathcal{T}$ be some (double) circular sequence. A set $\mathcal{P} \subset \{1, 2, \dots, n\}$ is a *halfplane* (with respect to $\mathcal{T}$) if there exists a permutation σ in $\mathcal{T}$ such that the elements of $\mathcal{P}$ (in some order) form the leftmost piece of σ. A halfplane consisting of k elements is also called a *k-set*. Denote the number of k-sets in a double circular sequence $\mathcal{T}$ by $s_k(\mathcal{T})$.

5.1. Consider a double circular sequence $\mathcal{T}$. Prove that the number of its k-flips coincides with the number of its k-sets.

5.2. For every double circular sequence $\mathcal{T}$ prove that $\sum\limits_{k=1}^{n-1} s_k(\mathcal{T}) = n(n-1)$.

5.3. Let $S_k(n)$ be the maximal possible sum of the form $\sum\limits_{i=1}^{k} s_i(\mathcal{T})$, where $\mathcal{T}$ is a double circular n-sequence. Prove that $S_k(n) = kn$ for all $1 \leqslant k < n/2$.

5.4. Assume that a double circular n-sequence $\mathcal{T}$ is constructed from a set of points S. Prove that the number of edges in a graph $G_k(S)$ is equal to $s_{k+1}(\mathcal{T})$ (for every k, $1 \leqslant k \leqslant (n-2)/2$).

5.5. Determine whether every circular sequence arises from some set of points in the plane.

5.6. Consider any circular n-sequence $\mathcal{T}$. Let P be its first permutation, and let $P = XYZ$ with $|Y| = y > 0$ (the substrings X and Z may happen to be empty). Prove that for every k with $1 \leqslant k \leqslant n-1$, the number of k-flips in $\mathcal{T}$ involving at least one element of Y does not exceed $\binom{y}{2} + 2k$.

5.7. Prove that there exists a constant C (depending on none of n and k) such that for every double circular n-sequence and for every $1 \leqslant k \leqslant n/2$ the inequality $s_k(\mathcal{T}) \leqslant Cn\sqrt{k}$ holds.

6 Intersections and the maximal number of edges

D e f i n i t i o n. We say that a pair of distinct edges in a geometrical graph is an *intersection* if these edges have a common point (different from a common vertex). In this section, we investigate the number of intersections.

6.1. Let V be the set of vertices of some halving graph with $|V| = n$ (surely, n is even). Denote by d_v the degree of a vertex $v \in V$. Let X be the number of intersections in our graph. Prove that

$$X + \sum_{v \in V} \binom{(d_v+1)/2}{2} = \binom{n/2}{2}.$$

6.2. Assume that the numbers n and e satisfy the conditions $10^6 n < e < 10^{-6} n^2$. Prove that there exists a (geometrical) graph on n vertices having e edges and at most $10^6 e^3/n^2$ intersections.

6.3. Prove that there exists a positive constant c such that every geometrical graph on n vertices with $e > 100n$ edges has at least $c \cdot e^3/n^2$ intersections.

6.4. Prove that the number of edges in a halving graph on n vertices does not exceed $Cn^{4/3}$, where C is some absolute constant (not depending on n).