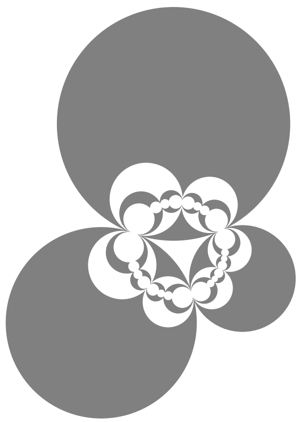
Геометрия: классика и современность

Тройственная симметрия фрактального калейдоскопа

Р. Р. Пименов



Оглавление

Предисловие

Эта статья — развитие и применение идей эстетической геометрии, которыми я стараюсь «заразить» и научное сообщество, и любителей. Она развивает идеи книги [3].

Эстетическая геометрия возникает как способ построения красивых образов на основании симметрий относительно окружности и плодотворна в разных разделах математики, таких как фрактальная геометрия, неевклидовы геометрии, проективная геометрия, основания геометрии. Её основы изложены в электронном учебнике [7] и сопутствующих материалах на сайте. Занимаясь эстетической геометрией, я пришёл к тезису: любое математическое понятие может быть содержательно проиллюстрировано в её рамках, а иллюстрация получается красивой (потому и возникло название «эстетическая» геометрия). Тезис хорошо себя показал на занятиях для школьников по основам теории групп.

В статье изучаются свойства трёх касающихся друг друга окружностей. Конечно, хочется провести ещё четвёртую окружность, касающуюся трёх исходных. Свойства такого «четырёхокружника» разбираются в учебнике [7]. Сейчас мы ограничиваемся изучением трёх касающихся друг друга окружностей, но делаем это разносторонне и порою глубоко. Что здесь можно изучать? Сверх того, что рассказывается о них в учебнике? Что вообще в рамках эстетической геометрии можно построить на основе трёх окружностей?

Можно отражать их симметрично друг относительно друга, получая первые фигуры фрактальной геометрии. Многократные отражения позволяют кодировать слова трёхбуквенного алфавита, моментально зримо представляя их многие важные свойства. Этот метод называется «Инверсионной диаграммой», и он полезен при изучении случайных процессов, а вероятно и шире — в семиотике. Количественный анализ «фрактального калейдоскопа», который при этом получается, требует специальной меры или координат, естественных для геометрии окружности. Для этого в начале статьи излагается понятие «изгиба» или координат, созданных парой касающихся окружностей. Нам даны три касающиеся окружности, поэтому можно ввести шесть таких систем координат. Правила перехода от одной системы координат к другой просты, элегантны и основаны на *тройственной симметрии* (симметрии третьего порядка). Эта симметрия возникает оттого, что все три исходные окружности абсолютно равноправны, её изучение — простой случай теории групп.

Введя числовую меру для эстетической геометрии, мы можем измерять и решать уравнения. Неожиданно, но первое нетривиальное уравнение, которое удаётся составить в эстетической геометрии, имеет решением золотое сечение, известное из прямолинейной геометрии Евклида. А сама мера изгиба оказывается эквивалентной двойному отношению в проективной геометрии. Проективная геометрия — раздел геометрии прямых, возникший из изучения перспективы во многом под влиянием эстетических устремлений художников и архитекторов. Оказывается, свойства двойного отношения определяются и проявляются в геометрии окружности ярче и естественней, чем в геометрии прямых. Затем уравнения эстетической геометрии записываются через дробно-линейные функции (частное двух линейных функций) и делаются новые приложения к поиску закономерностей в последовательности знаков. Также удаётся численно описать окружности, возникающие во «фрактальном калейдоскопе». Для этого оказываются полезны цепные дроби. В конце статьи намечаются пути развития эстетической геометрии.

Из-за разнообразия появляющихся математических идей и методов возникает мысль, что материал эстетической геометрии годится для прояснения содержательного, неформального единства математики. С конца XIX века единство математики устанавливалось и изучалось на формальном языке: с помощью теории множеств и математической логики. Это привело к бурному развитию математики, но и затруднило взаимопонимание учёных: изобилие и изощрённость формальных, языковых средств сделало современную науку доступной только для узкого круга специалистов. Эстетическая геометрия может изменить это положение, отвечая на полузапретный в мире профессиональных математиков вопрос: «Что же мы на самом деле изучаем в математике?». Я считаю, она — лучший путь к изучению богатейшего мира математики.

Методологически она связана с двумя идеями Лейбница. Его мысль об исчислении, связанном не с внешней системой координат, а с самими изучаемыми фигурами, находит разрешение в исчислении симметрий, постоянно используемом в эстетической геометрии. Подобное делал и Бахман в своих работах по основаниям геометрии на основе симметрий относительно прямых (см. [1]). Вторая связанная с эстетической геометрией мысль Лейбница — его мысль о монадах. Естественной моделью монады является окружность или сфера, относительно которой отражаются остальные монады-окружности (сферы). Это ясно видно при изучении «фрактального калейдоскопа», который в статье описывается как слова трёх- (или более) буквенного алфавита. Удобное наглядное представление любой последовательности знаков с помощью фрактального калейдоскопа или инверсионной диаграммы указывает на связь геометрии окружности и семиотики.

Сегодня малоизвестны даже основы геометрии окружности. Поэтому — несколько слов о том, как удобней читать эту статью. Все основные понятия подробно изложены в книге [3] и в учебнике [7]. Но я старался, чтобы главные идеи были ясны и без постоянного просмотра учебника. Единственное, что надо знать заранее, — *инверсия, или симметрия относительно окружности*. Если окружность получается из окружности A симметрией (инверсией) относительно окружности B, то мы обозначаем её $B(A)$. Все остальные упоминания теорем эстетической геометрии можно пропустить при первом чтении и вернуться к этим местам позднее, используя уже материал учебника. И, пожалуй, не помешает раньше времени заглянуть в самый конец. Надеюсь, что моя работа не только расскажет об интересных свойствах трёх окружностей и фрактального калейдоскопа, но и послужит введением в высшую математику для любителей. А профессионалы найдут в ней новые связи между известными им теориями.

В известной книге «Ожерелье Индры» [2] инверсии описываются алгебраически, с привлечением вспомогательной техники комплексных чисел и других мощных негеометрических средств. В моей же статье ведётся чисто геометрическая работа. И показано, как на её основе получаются и фундаментальные понятия теории групп, и вычислительные результаты.

Полная версия статьи с цветными иллюстрациями выложена в открытый доступ на сайте [5]. Данная версия статьи адаптирована к условиям публикации в настоящем сборнике. Автор будет рад обсуждению и развитию высказанных в ней идей и результатов. Наука, всякое познание, может твориться в одиночестве, но развивается и цветёт в сообществе.

§ 0. Три касающиеся окружности в искусстве

Поскольку мы занимаемся «эстетической геометрией», начнём изучение трёх сокасающихся окружностей с живописи и орнаментов. Я нашёл не очень много примеров. Лучше всего три касающиеся окружности видны в работах Рериха (см. [9]).

Три и даже четыре касающиеся окружности мы видим на древних кельтских орнаментах, также приведённых в [5].

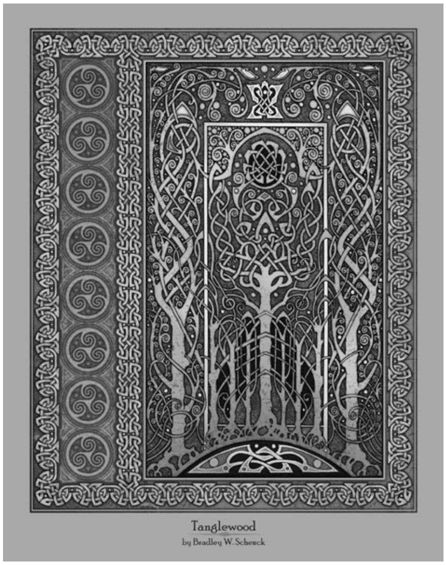


Рис. 1

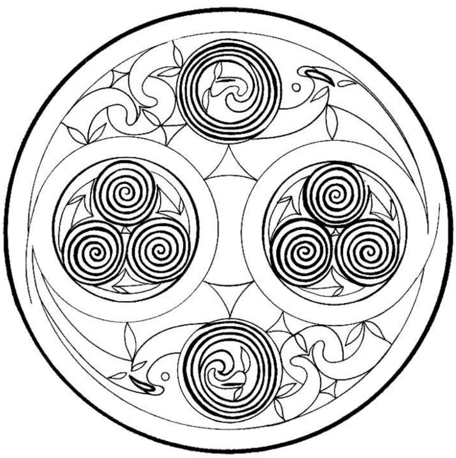


Рис. 2

Они изображены на левом поле замечательной работы, показанной на рис. 1. Здесь четыре касающиеся окружности, причём они возникают из спиралей. В центре работы сложный узловидный орнамент. Как показано на сайте [6], подобные орнаменты удобно строятся в рамках геометрии окружности, это демонстрирует и программа dodecaLook, имеющаяся на этом же сайте. А на рис. 2 мы видим более простое и подробное изображение четырёх касающихся окружностей-спиралей в кельтском искусстве.

Во всех подобных изображениях три касающиеся окружности явно имеют не только эстетическое значение, но выражают какую-то мистическую или философско-религиозную мысль. Я же буду исследовать их исключительно как математик-геометр.

На рис. 1 и 2 ни одна из трёх маленьких окружностей не лежит внутри другой. В дальнейшем во всей статье предполагается именно такое расположение исходных окружностей A, B, C. Возможен и другой вариант: окружности B и C лежат внутри A. Все наши рассуждения работают и в этом случае, но иногда становятся более громоздкими, поэтому такой вариант расположения окружностей не рассматривается. Читатель может обдумать его самостоятельно (это имеет особый интерес при построении фрактального калейдоскопа).

§ 1. Определение изгиба

Рассмотрим пучок окружностей, определённый окружностями A и B, т. е. совокупность окружностей, касающихся A и B одновременно в точке s. Представим возникновение этого пучка: маленькая, трудноотличимая от точки s окружность X увеличивается в размере — всё время касаясь A в s: вначале она лежит внутри A, в какой-то момент сравнивается с A, касается A с *другой стороны,* продолжает изгибаться и уже сравнивается с B, затем она и B касается с *другой стороны*, затем X уменьшается и снова почти превращается в точку s. На своём пути *из точки в точку* она в какой-то момент становится прямой линией.

В каждый момент окружность X изгибается. Хочется ввести *численную меру* этого изгиба. Чтобы сделать обычную линейку, нам прежде всего надо указать на ней меру: отметки 0 и 1. Чтобы ввести меру изгиба окружности, действуем аналогично: будем считать, что у окружности A — нулевой изгиб, а у окружности B — единичный. Раньше мы считали, что окружность X стартует с какой-то маленькой, неотличимой почти от точки касания окружности. Это наглядно, но не очень удобно. Вводя деления на обычной линейке, нулевую отметку часто помещают в центр линейки. Аналогично мы будем измерять *изгиб* окружности. Окружность A, как мы договорились, имеет нулевой изгиб. Положительный изгиб — это изгиб по направлению к окружности B, самой окружности B мы приписали единичный изгиб. Окружность C, лежащая строго посередине между A и B (т. е. такая окружность, относительно которой A и B симметричны), естественно, получит половинный изгиб. Окружность D, серединная между A и C, имеет изгиб, равный 0,25. И так далее. Окружности с отрицательным изгибом лежат по другую сторону от A. Окружность $A(B)$ имеет изгиб, равный -1, изгиб $A(C)$ равен $-0{,}5$, изгиб $A(D)$ равен $-0{,}25$ и так далее: если какая-то окружность пучка K имеет изгиб x, то окружность $A(K)$ имеет изгиб $-x$. Заметим: окружность A *разделяет* окружности K и $A(K)$.

Пока все рассмотренные нами окружности имели изгиб от -1 (окружность $A(B)$) до $+1$ (окружность B). Где лежат окружности большего изгиба? Они лежат по другую сторону от B. Например, окружность $B(A)$ имеет изгиб 2. Если мы продолжим сгибать окружность в том же направлении, её изгиб будет возрастать. Если же мы будем сгибать окружность A в *другом* направлении, то получим сначала окружность $A(B)$ с изгибом -1. При продолжении работы изгиб будет всё большим по абсолютной величине, но отрицательным по знаку, а окружность будет всё меньше отличаться от точки s, как это было вначале. Но мы приходим к этой точке *с другой стороны.*

Мы можем представлять кузнеца за работой: перед ним прут, зажатый в точке s. В начальный момент пусть прут совпадает с окружностью A, кузнецу вовсе не надо изгибать прут, чтобы получить A, поэтому изгиб A и назван нулевым; усилие кузнеца, нужное, чтобы изогнуть прут до окружности B, мы считаем единичным; работая дальше и прикладывая всё больше сил, кузнец сгибает прут в точку; сгибая прут в другую сторону, он в конце концов получает то же самое.

На рис. 3 мы видим несколько окружностей из пучка касающихся окружностей (A, B). Окружность F имеет отрицательный изгиб, окружности C и D имеют положительный изгиб, меньший чем 1, окружность E имеет изгиб, больший чем 1.

После инверсии относительно произвольной окружности S с центром в q все окружности пучка перейдут в семейство параллельных прямых (рис. 4). Причём благодаря разобранному в учебнике [3] принципу «симметричное

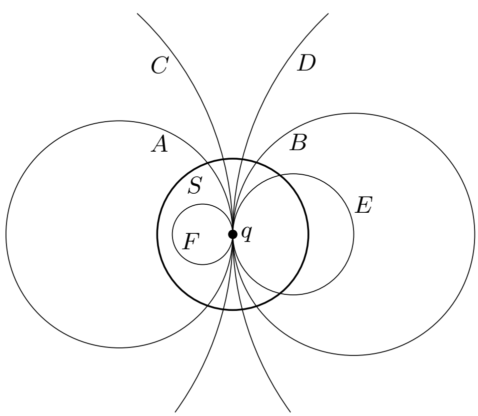


Рис. 3

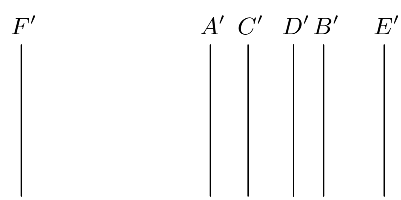


Рис. 4

симметрично симметричному» есть простое соответствие расстояний между получающимися прямыми и (A–B) изгибом окружностей, а именно: $S(F) = F'$, $S(A) = A'$, $S(C) = C'$, $S(D) = D'$, $S(B) = B'$, $S(E) = E'$.

Например, (A–B) изгиб окружности C есть отношение расстояния между прямыми A' и C' к расстоянию между A' и B'. Аналогично можно определить изгиб других окружностей пучка. При этом прямые, лежащие слева от A', мы считаем лежащими в отрицательной области (как и на линейке точки слева от нуля мы считаем отрицательными). Разумеется, так происходит потому, что прямая B' лежит справа от A'. Заметим ещё, что изгиб окружностей не меняется при симметрии относительно произвольной окружности W (не обязательно с центром в q). Это опять-таки легко следует из принципа «симметричное симметрично симметричному» (или из рассмотрения автоморфизма, определённого симметрией относительно W).

Назовём введённый изгиб «(A–B) координатой» окружности пучка. И введём обозначение: пусть $(A–B)(X)$, где A, B — касающиеся окружности, а X — окружность из созданного ими пучка, означает изгиб окружности X относительно пары окружностей A и B. Мы можем обобщить наше обозначение на случай, когда рассматривается не окружность, а произвольная точка плоскости. Ведь через точку x можно провести единственную окружность X, касающуюся A и B в точке s. Именно (A–B) координату этой окружности X мы назовём и (A–B) координатой *точки* x. Особый случай, когда x совпадает с s, мы можем исключить из рассмотрения, а можем считать в этом случае $(A–B)(x)$ равным бесконечности. Отметим, что каждая окружность пучка *однозначно* характеризуется своим изгибом, а точки плоскости — нет. Возьмём произвольную точку x и проведём через неё окружность X из пучка, созданного A и B. Все точки, лежащие на этой окружности, имеют тот же изгиб, что и точка x.

Пусть мы знаем (A–B) координату окружности (или точки) X. Как найти (B–A) координату X? Пучок тот же самый, изменились лишь условия подсчёта изгиба: теперь окружность B имеет нулевой изгиб (кузнец начинает гнуть окружность B, а не A), а окружность A — единичный изгиб. Обозначим исходный изгиб z, а искомый $h(z)$. Функция h должна обладать *двойственной симметрией* (симметрией второго порядка), иначе говоря, быть инволютивной: $h(h(z)) \equiv z$. Подсчитаем эту функцию для нескольких простых случаев: $h(0) = 1$, $h(1) = 0$, $h(0{,}5) = 0{,}5$. В самом деле, нулевой изгиб был у окружности A, но после переобозначения ей присвоен единичный изгиб; единичный изгиб был у окружности B, после переобозначения у неё нулевой изгиб. Половинный изгиб имела серединная окружность C, она и останется серединной, если A и B поменять местами. Возникает гипотеза, что $h(z) = 1 - z$. Это можно доказать, рассмотрев инверсию с центром в s и применив тезис «симметричное симметрично симметричному». Запишем этот вывод так: $(A–B)(X) = 1 - (B–A)(X)$ или $(A–B)(X) + (B–A)(X) = 1$.

Мы столкнулись с двойственной симметрией пучка окружностей. Теперь мы перейдём к тройственной симметрии.

§ 2. Три сокасающиеся окружности и тройственные координаты

Пусть три окружности A, B, C касаются друг друга. Три точки их касания обозначим s, p, q. При этом A и B касаются в точке s, B и C — в точке p, A и C — в точке q. Каждая пара окружностей определяет описанную выше систему координат для измерения изгиба. Если окружность X лежит в пучке (A, B), то мы можем измерить её $(A–B)$-изгиб, но тогда она не лежит ни в пучке (B, C), ни в пучке (A, C) и мы не можем определить её $(B–C)$ или $(A–C)$ координаты. Но у *точки* плоскости x можно найти три координаты: $(A–B)(x)$, $(B–C)(x)$ и $(C–A)(x)$, проводя через x окружности, лежащие в нужных пучках. Таким образом, каждой точке плоскости можно сопоставить три числа: её $(A–B)$, $(B–C)$ и $(C–A)$ координаты. Можно ли, зная одно из этих чисел, определить два других? Нет.

На рис. 5 изображены точка x и три проходящих через неё окружности. Одна лежит в пучке (A, B), другая в пучке (B, C), третья в пучке (A, C). Обратите внимание, что третья окружность обязательно проходит через точку пересечения двух других.

Если мы сдвинем точку x по большой правой окружности, то её $(A–C)$ координата не изменится, а $(A–B)$ и $(B–C)$ координаты изменятся. Поэтому

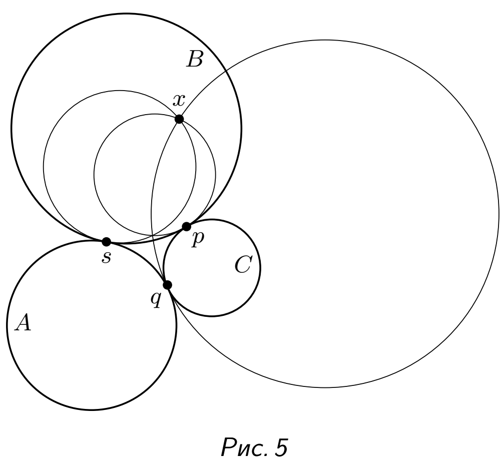


Рис. 5

не может быть способа определить по одной координате точки x две другие её координаты.

Но мы будем изучать (A–B) координаты не всех точек плоскости, а только некоторых. Проведём окружность I через точки s, p, q. Она перпендикулярна трём исходным окружностям A, B, C. Именно (A–B) координаты точек, лежащих на этой окружности, мы и будем изучать. Через произвольную точку x, лежащую на I, мы проводим окружность X, лежащую в пучке (A, B), и изгиб X в этом пучке и будет (A–B) координатой точки x. Легко видеть, что *различные* точки x и y, лежащие на I, обязательно имеют разные (A–B) координаты — поэтому точка на I однозначно задаётся своими координатами. Мы уже говорили о кузнеце, стягивающем и разгибающем прут. В каждый момент этот «прут» (окружность пучка (A, B)) пересекает I в двух точках: одна точка — s, точка касания A и B, она общая для всех «прутьев», а вторая точка пробегает всю окружность I, двигаясь через A и B.

Чтобы разобраться с этой и некоторыми подобными конструкциями, удобно воспользоваться стереографической проекцией, точнее её простейшим, одномерным случаем. Стереографической проекцией сферы называется проекция сферы на плоскость (на которой лежит сфера) из высшей точки сферы. В нашем случае не сфера проектируется на плоскость, а окружность на прямую, как изображено на рис. 6: окружность I проектируется на касающуюся её горизонтальную прямую из точки s. Точка s — высшая точка окружности, диаметрально противоположная точке касания прямой и окружности.

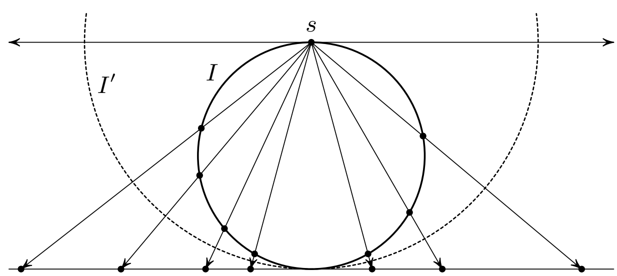


Рис. 6

Точке окружности ставится в соответствие точка на прямой. Но для самого центра проектирования s такой точки нет. Мы можем считать, что точке s соответствует бесконечно удалённая точка на прямой. Заметим, что такое же соответствие точек задаёт и инверсия относительно окружности I' с центром в s, касающейся прямой проектирования.

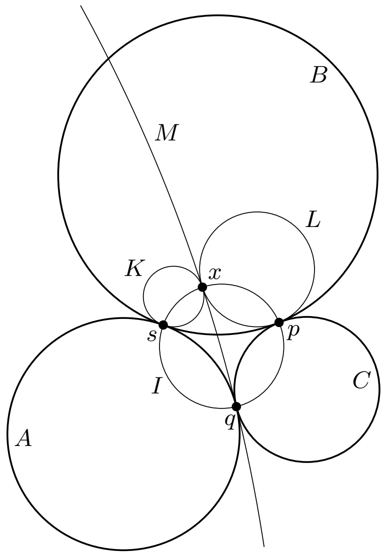


Рис. 7

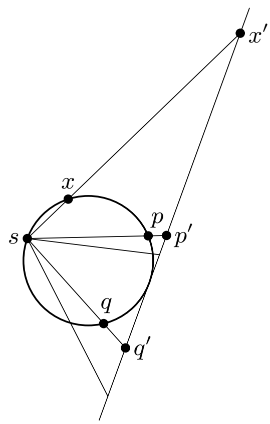


Рис. 8

Мы построили окружность I по трём сокасающимся окружностям A, B, C и выяснили, как определить $(A–B)$ координату произвольной точки x на I. Но точка x на I имеет не только $(A–B)$ координату, но и $(A–C)$ и $(C–B)$ координаты. Как связаны эти координаты между собой?

Через точку x на I проведены (см. рис. 7) три окружности K, L, M, лежащие в соответствующих пучках (между собой эти окружности обязательно касаются, что следует, например из теоремы о трёх окружностях). Интересующие нас изгибы — это $(A–B)(K)$, $(B–C)(L)$ и $(A–C)(M)$.

Как уже говорилось, значение каждого изгиба однозначно задаёт точку x на I. Величину изгиба легко представить с помощью стереографической проекции. Сделаем такую проекцию из точки s (см. рис. 8).

Аналогично связи между изибом окружностей и расстоянием между параллельными прямыми есть связь между $(A–B)$ координатой точки x и расстояниями между точками на прямой проектирования. Именно, $(A–B)(x) = |q', x'| : |q', p'|$. Это можно уяснить из принципа «симметричное симметрично симметричному». В данном случае речь идёт о связи (изоморфизме) между симметриями окружностей в пучке и симметрией точек на прямой проектирования друг относительно друга. Спроектировав из точек p и q на соответствующие прямые, мы получим два других числа для изгибов, измеренных в других системах координат.

Введём обозначения:

$$(A–B)(x) = k, \quad (A–C)(x) = f(k), \quad (B–A)(x) = h(k).$$

Для измерения изгиба мало указать пучок (две касающиеся окружности), нужно ещё указать порядок, направление изгибания. Поэтому каждая пара окружностей определяет две возможные системы координат («прямую» и «обратную»), всего получается 6 возможных систем координат. И, зная хоть одно из этих шести чисел, можно вычислить оставшиеся пять возможных координат. Как это сделать, какая связь между координатами — мы и будем сейчас выяснять.

Об одной связи мы уже говорили:

$$(A–B)(x) = 1 - (B–A)(x) \quad \text{или} \quad (A–B)(x) + (B - A)(x) = 1.$$

Эта формула была указана ранее для случая, когда x — окружность, она очевидно верна и если x — точка. Поэтому $h(k) = 1 - k$. Но пока мы будем писать всюду $h(k)$, а не $1 - k$, потому что так ясней видны закономерности, связывающие координаты. Мы заметили, что $h(h(k)) = k$. Аналогично $f(f(k)) = k$: функция перехода от $(A–B)$ координаты к $(A–C)$ координате должна быть такой же, как функция перехода от $(A–C)$ координаты к $(A–B)$ координате. Через функции f и h можно выразить координаты точки x в любой системе координат, которую мы составим из трёх исходных окружностей.

Пусть мы хотим выразить $(B–C)(x)$ через $(A–B)(x) = k$. Применим сначала преобразование h, меняющее координаты местами. Мы положили $h(k) = (B–A)(x)$. Выразить $(B–C)(x)$ через $(B–A)(x)$ просто: преобразование f как раз и заменяет *вторую* координату, потому $(B–C)(x)=f(h(k))(x)$, что и требовалось. А как заменить *первую* координату? Поступим так: перевернём координаты $(A–B)$, заменим *вторую* координату и снова перевернём. «Переворачивание» координат обеспечивает функция h, *вторую* координату заменяет f, поэтому $(C–B)(x)=h(f(h(k)))$. Аналогично можно поступать и в других случаях. Сведём результаты в таблицу.

Система координат	$(A–B)(x)$	$(B–A)(x)$	$(A–C)(x)$	$(C–A)(x)$	$(B–C)(x)$	$(C–B)(x)$
Значение	k	$h(k)$	$f(k)$	$h(f(k))$	$h(f(h(f(k))))$	$f(h(f(k)))$

Мы видим, что взаимосвязь между разными координатами выражается с помощью композиций всего двух функций: f и h. Функции f и h инволютивны, то есть дважды применённые дают тождественное преобразование: $f(f(k))=h(h(k))=k$. А что происходит при их композиции? Если мы применим функцию f, то вторая окружность будет заменена оставшейся окружностью. А после применения h первая координата станет второй, а вторая — первой. В общем итоге на первом месте будет стоять оставшаяся окружность, а на втором — стоявшая на первом: $(C–A)(x) = h(f(A–B)(x))$. Композиция, применённая трижды, ничего не меняет. Положим $w(k) = h(f(k))$ (или, используя специальный знак для композиции, $h * f = w$). При этом $w(w(w(k))) = k$ (или $w * w * w = e$, где e — тождественная функция). Так мы пришли к тройственной симметрии. Выпишем соотношения подробнее:

$$
\begin{aligned}
&w((A–B)(x)) = (C–A)(x),\\
&w((C–A)(x)) = (B–C)(x),\\
&w((B–C)(x))(A–B)(x).
\end{aligned}
$$

Напомним, что x — точка, а k — число, координата точки x, и все введённые функции w, f, h действуют на числах. Каждый раз переход происходит по следующему правилу: на место первой окружности ставится та, которой нет среди двух используемых, а та, которая была первой, ставится на второе место. Если мы представим A, B, C точками на окружности, то w можно мыслить просто сдвигом.

Но чему равна f — функция, обеспечивающая переход от $(A–B)$ координаты к $(A–C)$ координате? Перебирая разные простые функции, а в качестве первых гипотез всегда стоит пробовать простые варианты, мы натыкаемся на функцию $y = 1/z$. Эта функция инволютивна. Рассматривая точки на I, заметим, что точка с $(A–B)$ координатой 0 (точка q) имеет $(A–C)$ координату, равную бесконечности. Точка с $(A–B)$ координатой 1 (точка p) имеет $(A–C)$ координату 1. Наконец, точка с $(A–B)$ координатой бесконечность (точка s) имеет $(A–C)$ координату 0. Но именно таковы значения функции $y = 1/z$. Далее, как мы показали, композиция $w(z) = (h * f)(z) = h(f(z))$ обладает тройственной симметрией: $w(w(w(z))) = z$ для всех z. При этом $h(z) = 1 - z$. Проверим, удовлетворяет ли функция $y = 1/z$ этому тождеству. Подставим $h(z) = 1 - z$ и получим: $w(z) = 1 - 1/z$. В этом случае в самом деле имеется тождество $w(w(w(z)))=z$. Итак, функция $y=1/z$ — очень хороший кандидат на роль f. Строгое доказательство того, что $f(z) = 1/z$, можно провести либо с помощью проектирования, о котором говорилось недавно (но это требует некоторой вычислительной работы), либо с помощью связи между заменой координат и симметрией относительно биссектрисы, о которой будет сказано в § 9. Заметим, что в поисках функции f мы наткнулись на важное и интересное тройственное тождество для композиции функций $1 - z$ и $1/z$, функции $w(z) = 1 - 1/z$, которое почему-то не часто упоминают при обучении математике. Разумеется, аналогичное тождество есть и для функции, обратной к w: $y(z) = 1/(1 - z)$. Изучение этих и связанных тождеств — прекрасное введение в теорию групп. Теперь мы составим таблицу перехода от одной системы координат к другой, используя найденный вид функции f.

Система координат	$(A–B)(x)$	$(B–A)(x)$	$(A–C)(x)$
Значение	k	$h(k) = 1 - k$	$f(k) = \frac{1}{k}$
Система координат	$(C–A)(x)$	$(B–C)(x)$	$(C–B)(x)$
Значение	$h(f(k)) = 1 - \frac{1}{k}$	$h(f(h(f(k)))) = \frac{1}{1-k}$	$f(h(f(k))) = \frac{k}{k-1}$

Полезно проверить эти несложные вычисления самостоятельно. Очевидно, $h(f(h(f(h(f(k))))))=k$ (это просто перезапись того, что $w(w(w(k)))=k$), из чего следует ряд тождеств, например: $h(f(h(k))) = f(h(f(k)))$. А вот факт, что функция $v(k) = h(f(h(k))) = 1 - 1/(1-k)$ тоже инволютивна, т. е. $v(v(k)) = k$, можно вывести уже из того, что f и h инволютивны.

Наши функции f и h и все функции, которые можно получить их композициями (включая тождественную функцию), в совокупности образуют группу. Отметим — идея рассмотрения одного и того же объекта в разных координатах играет ключевую роль во многих разделах математики и физики. Без неё немыслимо векторное исчисление, а теория относительности даже получила своё название из-за неё: она исходит из того, что один и тот же равномерно двигающийся предмет рассматривается разными, также равномерно двигающимися, наблюдателями. Правила перехода от одного наблюдателя к другому, как и здесь, образуют группу, а свойства этой группы и лежат в основе специальной теории относительности. Рассматриваемая смена координат, образованных двумя окружностями, — простейший, но и значительный случай применения этого метода. Кстати, между геометрией окружности и геометризацией теории относительности (метод Минковского) есть прямая связь, но эта тема выходит за рамки нашей статьи.

§ 3. Точки и калибровка окружности

Три касающиеся друг друга окружности A, B, C можно задать, указав точки их касания: s, p, q. Действительно, проведём через эти точки окружность I. Мы это уже сделали: именно на I лежат точки x, $(A–B)$ координаты которых мы изучали. Возьмём из трёх точек s, p, q какие-нибудь две и проведём через них окружность, перпендикулярную I. Это можно сделать одним и только одним способом. Мы получим одну из трёх данных окружностей. Так же строятся и две другие окружности.

Выберем упорядоченную пару окружностей, например (A, B). Мы можем обозначить эту пару окружностей *упорядоченной* тройкой точек: начнём с точки q, в которой A касается C, на второе место поставим точку s, в которой A касается B, а на третье — точку p, в которой B касается C. Таким образом, паре окружностей (A, B) соответствует последовательность точек (q, s, p). Мы можем по упорядоченной тройке точек однозначно восстановить упорядоченную пару окружностей (проходящих через эти точки). Именно: проведём окружность, перпендикулярную I, через две первые точки (q и s), она будет первой окружностью искомой пары. Вторую окружность мы проведём также ортогонально I через вторую и третью точки (s, p) данной тройки.

Теперь мы можем говорить не только про (A, B) координаты произвольной точки x, но и про (q, s, p) координаты точки x. Мы можем ввести обозначение $(q, s, p)(x)$, где q, s, p, x — произвольные точки, лежащие на одной окружности. Тогда $(q, s, p)(x)$ есть изгиб окружностей, касающихся друг друга в точке s (второй точки в тройке), точнее говоря, $(q, s, p)(x) = (A, B)(X)$, где A — окружность, проходящая через (q, s), B — проходящая через (s, p), X — проходящая через (s, x). Все названные окружности, естественно, должны быть перпендикулярны I. Важно, что переход от окружностей к точкам оказался возможен только потому, что любая пара точек на окружности I однозначно определяет окружность, проходящую через эту пару и перпендикулярную I.

Определив координаты с помощью трёх упорядоченных точек, вернёмся к вопросу о «замене координат». Ранее для этого у нас были функции f и h, где h меняет местами координаты, f заменяет вторую окружность на оставшуюся. Эти две функции показывали, как меняются значения координат фиксированной точки x. Ранее координаты определялись упорядоченной парой окружностей, теперь они выражаются упорядоченной тройкой точек. Укажем связь в таблице: к уже используемой таблице добавим строку, выражающую координаты в точечном виде.

A и B касаются в s, B и C в p, A и C в q						
Система координат, *окружности*	$(A–B)(x)$	$(B–A)(x)$	$(A–C)(x)$	$(C–A)(x)$	$(B–C)(x)$	$(C–B)(x)$
Система координат, *точки*	$(q, s, p)(x)$	$(p, s, q)(x)$	$(s, q, p)(x)$	$(p, q, s)(x)$	$(s, p, q)(x)$	$(q, p, s)(x)$
Значение координаты	k	$h(k)$	$f(k)$	$h(f(k))$	$h(f(h(f(k))))$	$f(h(f(k)))$

В точечном виде удобно объяснить смысл функций h и f: функция h определяет, как меняется координата фиксированной точки x, если мы меняем местами первую и последнюю точки (среди трёх выбранных q, s, p); функция f определяет, как меняется координата точки x, если мы меняем местами первую и вторую точку. Какой геометрический и алгебраический смысл в этих заменах и точках q, s, p? Посмотрим на первый столбец таблицы.

Первая точка, q, — точка, где окружность A пересекает I и касается C. Окружность A в пучке (A, B) имеет нулевую координату, также и точка q имеет координату, равную нулю. Вторая точка, s, — точка, где A и B касаются. Как мы говорили, эта точка имеет бесконечную (A–B) координату, она и есть *бесконечно удалённая* точка. Наконец, третья точка, p, — точка пересечения B и I. Окружность B в пучке (A, B) имеет единичную координату, потому p также имеет единичную координату. Это можно выразить так: $(q, s, p)(q) = 0$, $(q, s, p)(s) = \infty$, $(q, s, p)(p) = 1$. Процесс установления координат на окружности I (калибровки окружности) оказывается прост: мы можем любую точку на окружности *назвать* нулём (началом). Затем любую другую точку *назвать* бесконечно удалённой, и, наконец, выбрать третью точку и *назвать* её единицей (мерой). От калибровки прямой (или разметки линейки) это отличается тем, что на прямой надо указать *две* точки (нуль и единицу), а на окружности — *три* (нуль, единицу и бесконечность). Можно считать, что прямая — это окружность, на которой одна точка (бесконечность) уже выделена.

Если мы фиксируем точку x и как-то меняем на окружности положение точек, названных нулём, единицей и бесконечностью (точек p, q, s), то будет меняться и значение координаты точки x (так как меняется сама система координат). Если на линейке мы поменяем местами точки 0 и 1, то как изменится координата k фиксированной точки x? Она станет равной $1 - k$. У нас такой сменой заведует функция h. Калибровка прямой достигается указанием всего *двух* точек, поэтому есть всего одна перестановка: они меняются местами. Калибровка окружности достигается указанием *трёх* точек, а три точки можно переставлять по-разному, потому и появляется функция f и композиции функций h и f.

Обратим теперь внимание, как переставляются (p, q, s) координаты под действием w. Это закрепит наше знакомство с тройственной симметрией и выражением её в символическом виде:

$$w((p, q, s)(x)) = (s, p, q)(x),$$
$$w((s, p, q)(x)) = (q, s, p)(x),$$
$$w((q, s, p)(x)) = (p, q, s)(x).$$

Функция w определяет изменение координаты точки при такой «перезаписи» точек: третья точка становится первой, первая — второй, вторая — третьей.

§ 4. Перестановки трёх точек и связанные с ними функции

Итак, у нас есть три точки, которые мы *называем* Нуль, Единица и Бесконечность. А введённые функции f и h (и их композиции) управляют тем, как меняются координаты фиксированной точки x в результате того, что мы *переименовывает* три выбранные точки. То, что мы раньше называли *переименованием,* теперь будем называть *перестановкой* элементов A, B, C. А как записывать перестановку? Пусть, например, первая буква тройки указывает, куда переходит A, вторая — куда переходит B, третья — куда переходит C. Есть и другой удобный способ записи. Пусть мы хотим выразить какую-то перестановку w. Возьмём букву A, следующей запишем букву $w(A)$, в которую A переходит, затем букву $w(w(A))$, в которую переходит $w(A)$, и так далее. Скажем, последовательность (A, B, C) теперь определяет перестановку: A переходит в B, B переходит в C, C переходит в A. Последовательность (A, B) выражает перестановку A и B.

Заметим, что в теории представлений обычно изучают представление групп линейными преобразованиями евклидова пространства. Здесь мы получили представление группы перестановок трёх элементов преобразованиями окружности.

§ 5. Алфавит, слова и композиции симметрий

Подготовив базу, то есть систему координат или меру изгиба, подходящую для количественного анализа фрактального калейдоскопа, начнём строить калейдоскоп и анализировать его качественно. Мы рассматриваем три сокасательные окружности A, B, C. Что будет происходить при композиции симметрий относительно них? Возьмём произвольную точку x, лежащую вне исходных окружностей A, B, C. Будем отражать её относительно A, B, C в каком-то порядке, получим последовательность точек, например: x, $A(x)$, $B(A(x))$, $C(B(A(x)))$, $B(C(B(A(x))))$, $\ldots$ С каждым отражением точка x приближается к окружности I. Подобную картину мы также видим при отражениях относительно прямых (точнее плоскостей), а не окружностей (или сфер).

Пусть имеется последовательность симметрий относительно A_1, A_2, $A_3, \ldots, A_k$, где все окружности A_i — одна из трёх исходных окружностей A, B или C. Заметим, что в этой последовательности мы считаем все *соседние* окружности различными. Ведь если рядом стоят две одинаковые окружности, то композиция симметрий относительно них — тождественное движение и их обе можно просто исключить из последовательности. Докажем, что результат композиции симметрий относительно этих окружностей обязательно лежит внутри *последней* окружности, окружности A_k. В самом деле, исходная точка x (напомним, она лежит вне окружностей A, B, C) после симметрии относительно одной из окружностей обязательно лежит внутри неё. Но тогда x будет *вне* двух оставшихся. Поэтому после симметрии относительно какой-то из этих двух точка окажется *внутри* неё (и вне двух оставшихся). И так далее — точка оказывается всё время внутри той окружности, относительно которой проведена последняя симметрия. Что и требовалось.

Пусть теперь имеется вторая последовательность $B_1, B_2, B_3, \ldots, B_s$, где все окружности B_i — также A, B или C, и мы знаем, что

$$A_k\big(A_{k-1}(\ldots(A_2(A_1(x)))\ldots)\big) = B_s\big(B_{s-1}(\ldots(B_2(B_1(x)))\ldots)\big).$$

Что можно сказать про A_i и B_i? При каких условиях во втором случае точка x попадёт туда же, куда и в первом? Как было доказано,

$$A_k\big(A_{k-1}(\ldots(A_2(A_1(x)))\ldots)\big)$$

обязательно лежит внутри A_k, а

$$B_s\big(B_{s-1}(\ldots(B_2(B_1(x)))\ldots)\big)$$

— внутри B_s. Поэтому равенство возможно, только если $A_k = B_s$. Так как при симметрии разные точки переходят в разные, из $A_k = B_s$ следует

$$A_{k-1}(\ldots(A_2(A_1(x)))\ldots) = B_{s-1}(\ldots(B_2(B_1(x)))\ldots).$$

Но из этого равенства следует, что $A_{k-1} = B_{s-1}$. И так далее. Мы получаем, что равенство возможно, только если справа и слева стоят одинаковые окружности, т. е. $k = s$ и $A_i = B_i$.

Рассуждение нетрудно обобщить. Пусть у нас есть не три сокасающиеся окружности, а несколько таких непересекающихся окружностей, что любые две из них лежат по одну сторону от любой другой. Можно представлять их небольшими окружностями-монетами, такими, что ни одна не лежит внутри другой. Как и раньше, возьмём произвольную точку x (лежащую вне всех окружностей), отразим её относительно какой-то одной из исходных окружностей, результат — относительно другой, и так далее. Точка x при этом путешествует по внутренности исходных окружностей, каждый раз находясь внутри той окружности, относительно которой была последняя симметрия. Как и в рассмотренном ранее случае трёх сокасающихся окружностей A, B, C, если в ходе какой-то последовательности симметрий точка x попала туда же, что и в ходе другой последовательности симметрий, то обе эти последовательности симметрий совпадают. Доказательство аналогично.

Как удобно записывать эти композиции? Назовём исходные окружности (например, A, B, C, D, E) *алфавитом.* Тогда каждое *слово*, составленное из этих букв, можно *прочитать* как последовательность симметрий. Например, $ABCDAE$ читается как композиция симметрий: сначала относительно E, потом относительно A, потом относительно D и так далее по записи. Обратите внимание — мы читаем слово *справа налево.* Если в каком-то слове рядом стоят две одинаковые буквы — их можно «сократить», то есть удалить обе эти буквы из слова. Очень важно, что два разных слова задают *различные* преобразования: это следует из того, что, как было показано, даже на одну точку x, лежащую вне всех окружностей, включённых в алфавит, разные слова действуют по-разному. Заметим, что если бы среди окружностей алфавита были пересекающиеся, то разные композиции относительно пересекающихся окружностей *могли бы* дать один и тот же результат. Например, если P и Q перпендикулярны, то $P * Q = Q * P$ и слово PQ означает то же преобразование, что и QP.

Как выглядит слово, обратное данному (то есть задающее обратное преобразование)? Нужно просто записать данное слово «задом наперёд», перевернуть его. При этом выделится совокупность слов, которые одинаково читаются слева направо и справа налево: *симметричные* слова, например *A*, *ABA*, *ABCBABCBA* и так далее. Какие преобразования плоскости задают такие слова? Очевидно, это *инволютивные* преобразования, т. е. если симметричное слово задаёт преобразование f, то $f * f = e$ (e — тождественное преобразование). Все симметрии инволютивны. Появляется догадка, что симметричные слова означают какие-то симметрии на плоскости. Какие именно? Разберёмся, как выглядит симметричное слово. Предположим, что в нём чётное число букв, $2k$. Тогда буквы с номерами k и $k+1$ должны быть одинаковы. Но две соседние одинаковые буквы можно удалить из слова, после их удаления рядом снова окажутся одинаковые буквы. И так далее — все буквы в симметричном слове чётной длины будут сокращаться. Пустое слово можно считать тождественным преобразованием e. Перейдём к словам нечётной длины $2k + 1$. У них имеется буква, стоящая в середине слова на $(k+1)$-м месте. Буквы, стоящие на равном расстоянии от неё, должны совпадать. Это даёт лёгкий способ построения произвольного симметричного слова длины $2k + 1$. Возьмём произвольное слово q из k букв. Припишем к нему справа (к концу слова) произвольную букву W (не совпадающую, разумеется, с последней буквой слова q), затем перевернём слово q и припишем полученный «перевёртыш» справа. Мы получим симметричное слово.

Какое преобразование выражает слово такого вида? Как рассказывается в курсе эстетической геометрии (или любом учебнике по группам преобразований), это слово задаёт симметрию относительно $q(W)$, где q — преобразование, определённое словом q (будем пользоваться одним обозначением для слова и для определённого им пребразования), а W — окружность, обозначенная буквой W. Итак, каждое симметричное слово определяет симметрию относительно окружности. Ранее отмечалось, что окружность и симметрию относительно окружности мы обозначаем одинаково. Теперь же появляется и *слово*, которое выражает то симметрию относительно окружности, а то саму окружность, в зависимости от контекста. В дальнейшем изложении все строчные латинские буквы означают симметричные слова. А исходные окружности, названные буквами алфавита, по-прежнему будут обозначаться прописными латинскими буквами.

На рис. 9 мы видим часть фрактального калейдоскопа, созданного тремя окружностями A, B, C (сплошной, штриховой, пунктирной). Исходные окружности представляют слова длины 1: A, B, C. Следующие по уменьшению окружности — слова длины 3: ABA, ACA, BAB и т. п. — всего шесть таких слов. Ещё меньшие окружности — слова длины 5, например: $ABCBA$ или $CBCBC$. Самые маленькие окружности — слова длины 7, как $ABACABA$ или $CBACABC$.

Как, глядя на окружность, прочитать слово, которое она означает? Посмотрим, скажем, на маленькую сплошную окружность внутри штриховой окружности B. Всякая окружность, лежащая внутри B, получена симметрией относительно B, потому первая (и последняя) буква в этом слове — B. Находим образ этой маленькой окружности после симметрии относительно B. Это сплошная окружность внутри окружности A. Всякая окружность, лежащая в A, получена симметрией относительно A. Следовательно, вторая (и предпоследняя) буква интересующего нас слова — это A. После симметрии относительно A мы получим сплошную окружность уже побольше, лежащую в B, так что третья с начала (и с конца) буква B. После симметрии полученной окружности относительно B у нас будет окружность A. Поэтому в центре слова написана буква A. Само же искомое слово имеет вид $BABABAB$. Аналогично можно прочитать любую изображённую окружность. Каждый раз процесс заканчивается, когда после симметрии мы попадаем в одну из исходных окружностей алфавита.

Слова длины 1 — это исходные окружности A, B, C.

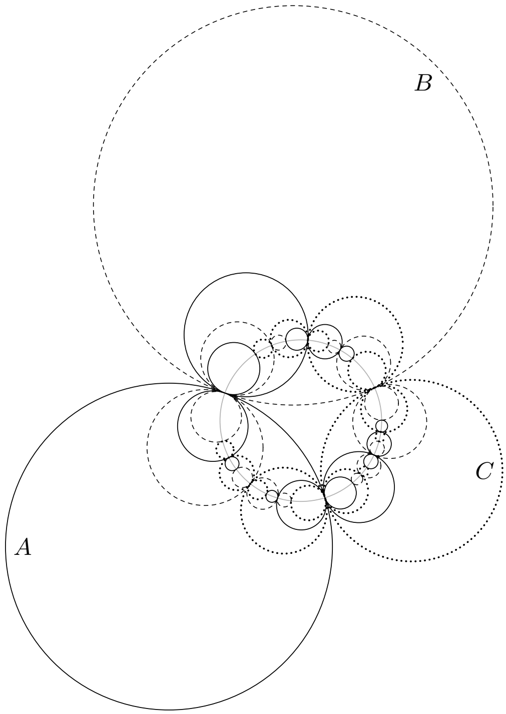


Рис. 9

Чем длиннее слово, тем меньше представляющая его окружность, и тем ближе она подходит к окружности I, перпендикулярной исходным окружностям A, B, C. Все окружности, представленные симметричными словами, никогда не пересекаются. Поэтому слово f (точнее, «окружность, представленная словом f», но для краткости мы будем это опускать) обязательно разделяет слово h и слово fhf (так как эти слова представляют окружности h и $f(h)$, т. е. пару окружностей, симметричную относительно f). Мы выбрали окружности алфавита так, что ни одна окружность не лежит внутри другой. Поэтому если X — какая-то буква алфавита, то любое симметричное слово, начинающееся с буквы X, лежит внутри X (это слово представляет результат симметрии относительно X). Мы обсуждали подобный вопрос ещё до того, как ввели понятие алфавита: надо было убедиться, что разные слова представляют разные преобразования. Пусть симметричное слово f начинается с букв XY, т. е. имеет вид $XYhYX$. Представленную им окружность можно выразить композицией симметрий: $f = X(Y(h))$. Так как YhY лежит внутри Y, окружность f лежит внутри XYX. Аналогично слово $XYZhZYX$ лежит внутри не только внутри X и XYX, но и внутри $XYZYX$. Таким образом, чтобы понять, внутри каких окружностей лежит слово f, нам не всегда нужно дочитывать f до конца (точнее, до середины, потому что f симметрично). Как в десятичной записи числа слева стоят бóльшие разряды и мы можем приближённо оценить число, не обращая внимания на стоящие справа цифры (число тысяч нам важней числа единиц или десятков), так можно примерно оценить расположение окружности f, не дочитывая слово f до конца. В общем случае, если слово f имеет вид $hXgXh$ (где g — некое симметричное слово меньшей длины), а X — буква алфавита, то f лежит внутри слова hXh, где h — некое не обязательно симметричное слово, с которого начинается слово f.

Рассмотрим простейший случай: алфавит состоит всего из двух букв A, B. Тогда любое слово состоит из последовательности чередующихся букв A и B. Все слова можно разделить на два множества: в одном множестве слова, начинающиеся с A, во втором — начинающиеся с B. Все слова нечётной длины оказываются симметричны, поэтому каждое слово нечётной длины в алфавите из двух букв представляет собой окружность. Как устроены эти окружности? Они показаны на рис. 3, иллюстрирующем понятие *пучка*, поскольку все они лежат в одном пучке, в котором лежат A и B. Рассмотрим все слова нечётной длины, начинающиеся с A: A, ABA, $ABABA$. Эти слова представляют череду вложенных друг в друга окружностей. Существует только одна точка, которая лежит *внутри всех* этих окружностей. Эта точка называется «центром пучка окружностей» Теперь рассмотрим слова нечётной длины, начинающиеся с B: B, BAB, $BABABAB$, … Этот ряд слов-окружностей совершенно аналогичен уже рассмотренному. Существует только одна точка, лежащая внутри всех рассматриваемых окружностей. Эта точка будет вторым центром пучка окружностей. Заметим одно важное для нас исключение. Если окружности A и B *касаются* друг друга, а именно с такого случая мы и начали наше рассмотрение, то центры пучка совпадают друг с другом и с точкой касания этих окружностей. Эта точка лежит не внутри рассматриваемой последовательности окружностей, а на всех этих окружностях. Введённое раньше понятие *изгиба* окружностей даёт численную меру для всех симметричных слов в этом случае. Каждому слову, симметричному слову-окружности, приписывается (A–B) координата; слова нечётной длины, начинающиеся с буквы A, имеют координаты $0, -1, -2, -3, \dots$, а слова нечётной длины, начинающиеся с буквы B, имеют (A–B) координаты $1, 2, 3, \dots$ Этот случай можно увидеть на рис. 9, если рассматривать только сплошные и штриховые окружности.

Между двумя рассмотренными последовательностями слов-окружностей есть простое соответствие. Заменим в каждом слове единовременно букву A на букву B, а букву B на A. Тогда A соответствует B, ABA

соответствует BAB, и так далее. Какой *геометрический смысл* этого соответствия? Произведённая замена букв наглядно изображается симметрией относительно биссектрисы (серединной окружности) между A и B. При этой симметрии меняются местами окружности A и B и все окружности двух разобранных последовательностей. Также меняются местами центры пучка. А что происходит с (A–B) координатами (для случая, когда A и B касаются друг друга)?

Можно сказать, хотя это и не очень строго: чем длиннее симметричное слово, тем меньше представленная им окружность. Более того, если мы рассмотрим симметричные слова длины n, то с увеличением n совокупный объём этих слов-окружностей становится всё меньше, стремясь к нулю, хотя самих окружностей становится всё больше и больше. Если окружностей алфавита всего три: A, B, C, то обязательно существует окружность I, перпендикулярная им всем. Окружности, представленные симметричными словами из этого алфавита, также обязательно будут перпендикулярны I. Это уже помогает нам представить расположение окружностей.

Что происходит, если A, B, C касаются друг друга? Тогда I проходит через точки их касания, а окружности A, B, C накрывают I полностью, всякая точка I лежит внутри какой-то из этих трёх окружностей. Ко всякой точке x, лежащей на I, можно приблизиться охватывающими её симметричными словами-окружностями. Это делается так: x лежит либо внутри A, либо внутри B, либо внутри C. Пусть это окружность A. Слова-окружности ABA и ACA покрывают целиком ту часть I, которая лежит в A, следовательно, x лежит внутри одной из этих окружностей. И так далее. Мы получим последовательность слов-окружностей, стягивающихся к произвольной точке x. Это напоминает запись точки в десятичной (или троичной) системе счисления. В десятичной системе счисления разные последовательности знаков выражают разные числа, за некоторыми исключениями. Десятичная дробь $0{,}99999999\ldots$, где девяток бесконечно много, означает то же, что и число 1. Подобное случается и когда мы рассматриваем бесконечное симметричное окружность-слово. Это связано с точками касания исходных A, B, C между собой. Будем рассматривать слова из двух букв A и B. Слово AB означает преобразование, стягивающее все точки плоскости к s — точке касания A и B. Преобразование $ABAB$ — стягивает к s в два раза быстрее, $ABABAB$ стягивает в три раза быстрее, и так далее. Преобразования, начинающиеся с буквы B: BA, $BABA$, $BABABA$ также стягивают всё к точке s (но с другой стороны). Поэтому преобразование-слово чётной бесконечной длины стягивает всю плоскость в точку s.

Разумеется, эту мысль можно и нужно уточнять. Мы рассмотрим симметричные слова-окружности из двух букв. Любые две такие окружности касаются друг друга в точке s, причём окружности-слова, начинающиеся с A, касаются с одной стороны, а начинающиеся с B — с другой. С увеличением длины такого слова оно означает окружность всё меньшего размера, которая становится малоотличима от точки s. Таким образом, к точке s можно приблизиться *двумя* последовательностями слов-окружностей. Аналогичными свойствами обладают и две другие точки касания p и q. И если мы будем отражать любую из точек касания относительно трёх окружностей алфавита, то к получаемым точкам также можно приближаться с разных сторон. Семейство точек, полученных из исходных точек касания симметриями относительно окружностей алфавита, постепенно распространяется по всей окружности I. Они не заполняют целиком никакой дуги окружности I, но на произвольной, сколь угодно малой дуге окружности I обязательно найдутся точки из этого семейства, подобно тому как рациональные числа есть внутри сколь угодно малого отрезка числовой прямой.

Теперь рассмотрим случай, когда три исходные окружности алфавита A, B, C не имеют общих точек. Тогда перпендикулярная им окружность I *не полностью* накрывается окружностями $A, B, C, \ldots$ Что происходит с симметриями трёх исходных окружностей друг относительно друга? Они выражаются симметричными словами-окружностями. Все симметричные слова длины 3 (их шесть, внутри каждой исходной окружности лежат 2 другие) накрывают симметричные слова длины 5 (снова внутри каждой окружности длины 3 лежит 2 окружности длины 5, всего таких окружностей 12) и так далее: с увеличением длины слова на 2 число симметричных слов данной длины удваивается. При этом каждая новая окружность перпендикулярна I. Мы можем отмечать на I «дыры». Окружности-слова длины 1 не накрывают все I. Окружности-слова длины 3 накрывают ещё меньшую часть I: внутри дуги I, накрытой окружностью A, появляются дыры. Окружности-слова длины 5 накрывают ещё меньшую часть, образуя новые дыры. И так далее. Постепенно «дыр» становится всё больше и их совокупная длина оказывается всё ближе к длине всей окружности I. Соответственно, длина дуг окружности I, накрытых симметричными словами-окружностями длины k, становится всё меньше и меньше, стремясь к нулю. В отличие от случая, когда A, B, C касались друг друга, вовсе не к каждой точке, лежащей на I, удастся сколь угодно близко подобраться с помощью симметричных слов-окружностей длины k. Наоборот, вероятность того, что к точке x можно так подобраться, равна нулю.

Что получается в пределе, если отражать окружности бесконечно долго? Они стянутся в некоторые точки, лежащие на I. Легко понять, что этих точек бесконечно много, но они лежат «порознь» друг от друга. Похожими свойствами обладает, например, множество чисел на отрезке от 0 до 1, в десятичной записи которых нет цифры 9. А множество «дыр» напоминает, наоборот, множество таких чисел, в десятичной записи которых есть цифра 9. Хотя на первый взгляд чисел «без девятки» куда больше, чем чисел «с девяткой», на самом деле вероятность того, что среди десятичной записи произвольной точки нет девятки (ни в одном разряде!) — равна нулю.

Совокупность точек, к которым стягиваются все окружности после бесконечных отражений друг относительно друга, обозначим W. Все эти точки, напомним, лежат на I.

Теперь посмотрим на все слова в алфавите A, B, C. Эти слова означают некоторые преобразования плоскости. Каждое из этих преобразований имеет предельную точку (куда стягивается вся плоскость при многократном применении этого преобразования). Эта предельная точка обязательно лежит в W. Чем длиннее слово, тем быстрее означаемое им преобразование стягивает всё к предельной точке, если же слово бесконечной длины, то оно отображает все точки плоскости сразу в одну, лежащую в W. Мы можем так и мыслить W — как совокупность точек плоскости, куда обязательно попадёт любая точка плоскости под действием какого-нибудь слова бесконечной длины (отражаясь достаточно долго от трёх окружностей алфавита). Это множество W принято называть «фрактальной пылью». Во фрактальную пыль стирает всю плоскость действие бесконечных слов-преобразований.

Пока мы анализировали свойства окружностей, возникающих из симметрий относительно нескольких данных окружностей алфавита, *качественно.* Хочется ввести какую-то количественную меру, позволяющую записывать точные равенства и уравнения. Этим, используя определённый ранее *изгиб* окружностей, мы и займёмся в следующем параграфе.

§ 6. Симметрии относительно трёх окружностей в (A–B) координатах. Дробно-линейные преобразования и цепные дроби. Уравнение золотого сечения

Как мы видели, понятие изгиба или (A–B) координат даёт возможность ввести координаты для любой точки плоскости. Если у нас есть три касающиеся окружности A, B, C, то любая точка окружности I, перпендикулярной им всем, *однозначно* задаётся своей (A–B) координатой, а также своей (B–C) или (C–A) координатой. Правила перехода от одной системы координат к другой мы свели в таблицы. Причём были указаны два варианта введения координат: с помощью окружностей A, B, C и с помощью точек касания этих окружностей: p, q, s. Теперь мы изучим, как меняются координаты точек на I под воздействием симметрий относительно A,

B и C. Это позволит нам описать, как действуют слова-преобразования, составленные из алфавита A, B, C, на точки окружности I.

Начнём с простейшего. Пусть x — и точка, и её (A–B) координата на окружности I. Чему равна (A–B) координата точки $A(x)$? Легко видеть, она равна $-x$. Чему равна (A–B) координата точки $B(x)$? Также нетрудно видеть, она равна $2 - x$. Теперь мы легко увидим, как меняются координаты точки x при композиции симметрий относительно A и B: $A(B(x)) = -(2 - x) = x - 2$; $B(A(x)) = x + 2$. Таким образом, точка с бесконечной координатой остаётся на месте, а остальные точки сдвигаются на 2 в ту или другую сторону. Что означают слова, составленные из двух букв A и B? Слово чётной длины $2k$, начинающееся с A, означает функцию $x - 2k$, начинающееся с B — функцию $x + 2k$. Слово нечётной длины при двухбуквенном алфавите всегда означает симметрию относительно некоторой окружности и потому должно выражаться инволютивной функцией. Установив, как работают симметрии относительно A и B и какие функции задают слова, составленные из этих двух букв, перейдём к симметрии относительно окружности C. Чему равно $C(x)$? Ответ не очевиден. Воспользуемся следующим приёмом: мы легко можем выразить эту зависимость в (A–C) координатах. Это точно такая же зависимость, как для $B(x)$ в (A–B) координатах. Поэтому найдём (A–C) координаты точки x, это нами обозначено было как $f(x)$, затем найдём $B(f(x))$, а полученное выразим в (A–B) координатах, результат будет $C(x) = f^{-1}(B(f(x)))$. Мы знаем, что $f(x) = 1/x$, $B(f(x)) = 2 - f(x) = 2 - 1/x$, подставим: $f^{-1}(B(f(x))) = 1/(2 - 1/x) = x/(2x - 1)$. Мы сталкиваемся с подобным приёмом, не только работая с координатами или системами отсчёта, но и в более разнообразных ситуациях, например при переводе. Общаясь с англичанином, мы, чтобы узнать ответ на свой вопрос, переводим вопрос с русского на английский, а ответ переводим с английского на русский. Функция f в этом случае означает «перевод с русского на английский» а обратная ей f^{-1} — «перевод с английского на русский». При этом методе у нас неизбежно появляется понятие сопряжённых элементов или сопряжённых функций: если f — правило перехода из одной системы координат в другую, а Q — некое построение или вычисление, какое нам удобней делать в этой новой системе координат, то $H(x) = f^{-1}(Q(f(x))$ называется сопряжённым с $Q(x)$.

Мы выразили, как действуют симметрии относительно трёх данных касающихся окружностей A, B, C на точки, лежащие на I: $A(x) = -x$, $B(x) = 2 - x$, $C(x) = x/(2x - 1)$. Составляя композиции, мы будем использовать дробно-линейные функции. Правда, не все дробно-линейные функции выражаются такими словами: их коэффициенты не произвольные вещественные числа, а целые числа особого вида. Сейчас мы не будем углубляться в изучение того, какими должны быть эти коэффициенты, а заметим связь между трёхбуквенными словами и цепными дробями.

Мы знаем, что $C(x)=x/(2x-1)$. Поделим числитель и знаменатель на x, получим $C(x) = 1/(2 - 1/x)$. Применяя к этому выражению композиции, составленные из симметрий относительно A и B, получим выражение вида $2k \pm \dfrac{1}{2-1/x}$. Применив к нему симметрию относительно C, получим

$$\cfrac{1}{2-\cfrac{1}{2k\pm\cfrac{1}{2-\cfrac{1}{x}}}}.$$

Мы видим, что получается цепная дробь. Таким образом, каждому слову из трёх букв A, B, C можно сопоставить цепную дробь, коэффициенты которой однозначно определены. Количество знаков деления в этой дроби в два раза больше, чем вхождений буквы C в слово.

После аналитического, численного выражения симметрии относительно окружности с помощью (A–B) координат появляется возможность составлять уравнения, решениями которых будут числа. До этого мы могли составлять уравнения, решениями которых были бы точка или окружность. Самое простое уравнение: $A(x) = x$. Этому условию удовлетворяют две точки пересечения A с I: q и s. Координаты этих точек — 0 и ∞. Аналогично разбираются два других уравнения. В случае $B(x) = x$ или $2 - x = x$ решения этого уравнения 1 и ∞, им соответствуют точки p и s (точки пересечения B с I). Уравнение $C(x) = x$ означает $x/(2x - 1) = x$, корни этого уравнения 0 и 1, точки q и p, по которым C пересекается с I. Мы видим, что решения уравнений согласуются с геометрией, но, разумеется, не узнаем ничего нового. Так же обстоит дело при рассмотрении уравнения $A(B(x))=x$. У этого уравнения единственное решение: ∞ или точка s, где касаются A и B. Аналогично обстоит дело и с уравнениями $B(C(x)) = x$ или $A(C(x))=x$. Аналитическое, численное решение лишь подтвердит очевидный геометрический факт: неподвижной остаётся только точка касания окружностей.

Первое содержательное уравнение, решение которого не очевидно геометрически, выглядит так: $C(B(A(x)))=x$. В (A–B) координатах оно записывается как $(2-(-x))/(2(2-(-x))-1)=x$. Раскрывая скобки и приводя подобные члены, получим: $(2 + x)/(2x + 3) = x$. Это уравнение приводится к квадратному $x^2 + x - 1 = 0$. Это то же самое уравнение, которое возникает при определении *золотого сечения*. Неожиданно, что первое интересное уравнение, которое можно написать, используя понятие изгиба, имеет своим решением золотое сечение. Ведь золотое сечение возникает в геометрии прямых и казалось не связанным с геометрией окружности.

Отметим простой способ наглядно находить приближённое решение уравнения $C(B(A(x))) = x$. Возьмём произвольную точку плоскости. Отразим её последовательно относительно окружностей A, B, C в указанном порядке. С полученной точкой сделаем то же самое, и т. д. В итоге мы построим последовательность точек, сходящуюся к неподвижной точке преобразования $C * B * A$, — одному из корней изучаемого уравнения. Взяв член этой последовательности с большим номером, мы получим неплохое приближённое решение уравнения. На рис. 10 это и показано.

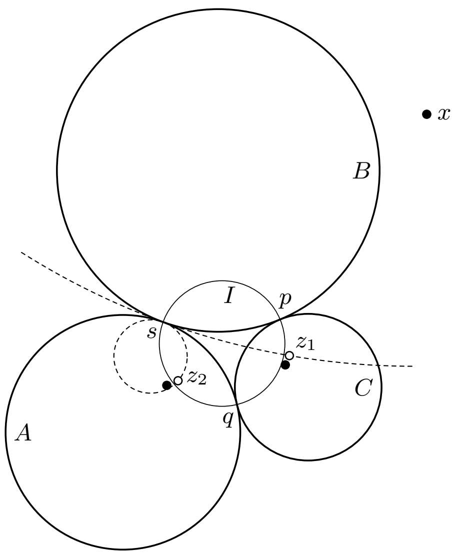


Рис. 10

Выбрана произвольная точка x, с неё начинается последовательность чёрных точек $C(B(A(x)))$, $C(B(A(C(B(A(x))))))$, ... Эта последовательность быстро прижимается к окружности I, и уже после второго шага её члены мало отличаются друг от друга. Член этой последовательности с большим номером изображён белым кружком z_1, с хорошей точностью z_1 является искомым решением уравнения $C(B(A(x))) = x$, а изгиб точки z_1 в $(A–B)$ координатах равен золотому сечению. Штрихами указана окружность из пучка (A, B), проходящая через z_1. Уравнение $C(B(A(x))) = x$ имеет и второе решение; чтобы приблизиться к нему, нужно двигаться от x в другую сторону, то есть рассмотреть последовательность точек: x, $A(B(C(x)))$, $A(B(C(A(B(C(x))))))$, ... Её предел есть также решение уравнения $C(B(A(x))) = x$, белый кружок z_2 — член этой последовательности с большим номером — является приближённым решением уравнения. Штрихами показана окружность из пучка (A, B), проходящая через z_2. Её изгиб отрицателен, он приближённо равен второму корню полученного ранее квадратного уравнения (к первому корню близка $(A–B)$ координата точки z_1, приближённо равная золотому сечению).

Вспомним, что мы начали с трёх касающихся окружностей, и сформулируем результат так: если достаточно долго отражать произвольную точку относительно трёх касающихся окружностей (в указанном порядке), то $(A–B)$ координата точки будет мало отличаться от золотого сечения. Это — предельное свойство золотого сечения в геометрии окружности. А указанный способ приближённого решения уравнения подходит для многих уравнений геометрии окружности. И часто его графическая реализация приводит к эстетически значимым результатам. Таким образом строятся интереснейшие барочные спирали, о которых рассказывается на сайте [7]. В анализе подобный метод решения уравнений вида $f(x) = x$ называется итерационным методом Ньютона, обычно он применяется к решению уравнений в вещественных числах.

§ 7. Пары точек и гармоническое отношение

Ранее мы уже отметили, что окружность, перпендикулярная I, однозначно задаётся своими точками пересечения с I и что каждые две точки на I определяют перпендикулярную I окружность. Поэтому, как было показано в § 3, можно говорить не об окружностях, перпендикулярных I, а о парах точек, лежащих на I. Пара точек (w, v), лежащих на I, задаёт такое отображение t точек на I, что $t(x)$ есть точка, симметричная x относительно окружности Q, проходящей через w и v (см. рис. 11). Мы можем назвать это отображение сужением симметрии относительно Q на окружность I. Хотя всё действие происходит на I, геометрически

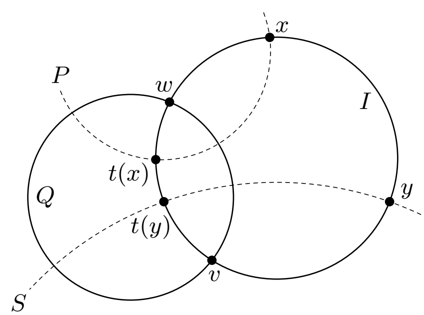


Рис. 11

(как сделать это аналитически — мы покажем позднее) построить по произвольной точке x точку $t(x)$ проще, выйдя за пределы I. Есть два способа. Первый: провести через (w, v) окружность, перпендикулярную I, а затем провести через x окружность, перпендикулярную I и новопостроенной. Вторая точка пересечения этой окружности с I (первая — исходная точка x) и будет искомой $t(x)$. Второй способ возвращает нас к миру прямых. Проведём через w и v касательные (см. рис. 12). Точка их пересечения A будет центром окружности, перпендикулярной I и проходящей через w и v. Проведём через A и x прямую. Вторая точка пересечения этой прямой и I (первая — исходная точка x) и будет искомой $t(x)$. Хорошее упражнение — доказать, что оба способа дают один и тот же результат.

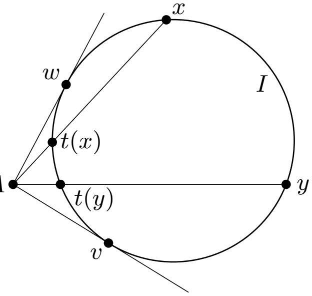


Рис. 12

Окружность Q перпендикулярна I, вспомогательные окружности P и S перпендикулярны им обеим. Разумеется, $t(t(x)) = x$, $t(t(y)) = y$. Поскольку оба способа дают один и тот же результат, можно сформулировать теорему планиметрии: выбрав произвольные точки x и y, мы получим точки $t(x)$ и $t(y)$, как показано на рис. 11, найдём A — точку пересечения соответствующих прямых, как показано на рис. 12. Теорема утверждает, что прямые, проходящие через A и w или через A и v, касаются I. Можно по-разному переформулировать эту теорему, например: прямые, проходящие через пары «образ-прообраз» отображения t, обязательно пересекаются в одной точке A.

Первый способ связан с *биплетной* симметрией плоскости: так в геометрии окружности называется симметрия плоскости относительно *пары точек.* Второй способ основан на A-отображениях. О них подробно говорится в [8]. Среди прочего там рассказывается о моделировании A-отображениями проективной геометрии. Описанное только что понятие симметрии одной пары точек относительно другой пары точек в проективной геометрии называется гармоническим расположением (или отношением) пар точек. Обратим внимание, что симметрию относительно пары точек можно мыслить как выворачивание наизнанку дуги, ограниченной этими двумя точками. Точке, лежащей по одну сторону от пары точек, симметрична точка, лежащая по другую сторону от этой пары.

Вернёмся на окружность I, где пара точек x и $t(x)$ симметрична относительно пары w, v. Свяжем эту симметрию с $(A–B)$ координатами. Как было показано в § 3, можно говорить не про $(A–B)$ координаты, а про (q, s, p) координаты, где q, s, p — точки касания исходных окружностей между собой. При этом точка q соответствует нулю, точка s соответствует бесконечности, точка p соответствует единице. Какова координата точки, симметричной s относительно p и q? Ответ прост: через 0 и 1 проходит ортогональная I окружность C, поэтому мы ищем такую точку s, что $C(s) = C(\infty) = 1/2$. Это можно получить, используя формулу $C(x) = x/(2x - 1)$. В данном случае $x = \infty$, поэтому нужно вычислить предел. Можно этот же результат получить геометрически, используя тот факт, что окружность, проходящая через s и $C(s)$ (перпендикулярная I, как и все рассматриваемые нами окружности), будет биссектрисой между A и B, а (A–B) координата биссектрисы равна $-1/2$. А как найти точку (её координату), симметричную p относительно q и s? Можно воспользоваться заменой координат (перейти к (A–C) координатам), но проще заметить, что эта точка есть $A(1) = -1$. Аналогично $C(0) = 2 - 0 = 2$.

Полученные три точки также можно отражать относительно исходных окружностей. Полученные точки — снова отразить. И так далее. Ранее мы изучали этот процесс качественно в § 5, а в § 6 связали композиции относительно A, B, C с цепными дробями. Заметим, что на каждом шаге (кроме первого) точку можно отразить относительно *двух* окружностей, хотя всего окружностей *три*. Ведь мы попали в данную точку после симметрии относительно какой-то окружности, и симметрия относительно неё же вернёт нас обратно. Полученное множество точек всюду плотно распределено на окружности I.

Термин «гармоническое отношение» возник в проективной геометрии для описания двух пар точек, лежащих на одной прямой. Проективная геометрия возникла в связи с изучением перспективы в живописи: точку «схода» всех «параллельных» на картине изучают методами проективной геометрии. Точнее — изучение этой точки, понимание законов перспективы художниками подталкивало математиков к развитию проективной геометрии, расцветшей в XIX веке. Многое для неё сделал ещё Паскаль. Хотя проективная геометрия изучает прямые и гармоническое отношение определяется для точек, лежащих на прямой, а мы сейчас изучаем геометрию окружности, перейти от одного к другому несложно. От геометрии окружности перейти к геометрии прямых в данном случае просто, так как все симметрии происходят на окружности I. Достаточно поэтому рассмотреть инверсию с центром в точке, лежащей на I. Тогда I перейдёт в прямую, а точки, симметричные относительно пары точек, — снова в симметричные точки. Напоминаю подробно разобранный в учебнике тезис: симметричное симметрично симметричному. В проективной геометрии можно довольно неожиданным образом определить если не окружность, то её обобщения — кривые второго порядка. Но куда удобней глядеть из геометрии окружности на проективную геометрию, чем наоборот.

Теперь введём определение *гармонического отношения* двух пар точек на окружности: если пара точек (x, y) симметрична относительно пары точек (w, v), то мы называем пары точек (x, y) и (w, v) гармоническими парами или «гармонически разделяющими друг друга». В этом случае точки x и y обязательно лежат по разные стороны от каждой из точек w

и v, а точки w и v — по разные стороны от каждой из точек x и y. Укажем на важнейшее свойство гармонических пар: свойство *взаимности*. Если x и y симметричны (гармонически разделены) относительно w и v, то обязательно точки w и v симметричны (гармонически разделены) относительно x и y. Лучше всего это видно из первого способа построения симметричных точек, связанного, как было сказано, с биплетным исчислением. Проведём через w и v окружность J, ортогональную I (на которой лежат все наши точки), и окружность U, проходящую через x и перпендикулярную J и I. Вторая точка пересечения U и I (обозначенная как y) будет симметрична x относительно w и v. Рассмотрим теперь симметрию относительно U (на окружности I это будет совпадать с симметрией относительно пары точек x и y). Тогда $U(w) = v$, так как $U(I) = I$ и $U(J) = J$, поэтому точки пересечения I с J (w и v) меняются местами. Что и требовалось. Можно понять свойство взаимности немного иначе. Заметим, что окружности J и U абсолютно равноправны: они обе перпендикулярны друг другу и I, и свойство взаимности лишь переформулирует это геометрическое свойство.

Теперь посмотрим на гармоническое отношение с помощью (q, s, p) координат. Как мы видели, если точка x симметрична точке s относительно q и p, то её (q, s, p) координата равна $1/2$. Можно обозначить это так: $g(q, s, p, x) = 1/2$. После перестановок получим: $g(s, q, p, x) = 2$, $g(p, q, s, x) = -1$. Другие перестановки трёх точек новых значений не дадут. Заметьте, что два новых полученных значения -1 и 2 совпадают с полученными ранее значениями $A(1)$ и $B(0)$. В этом есть простой геометрический смысл. Обратим внимание, что в последней записи $g(p, q, s, x) = -1$ точки следуют именно в том порядке, в каком входят в пары.

Введённое обозначение $g(q, s, p, x)$ само подсказывает нам, что его можно истолковать более общим способом: речь идёт о функции, сопоставляющей любым четырём точкам на окружности вещественное число, равное координате последнего аргумента (точки x), исчисленной относительно первых трёх точек (взятых именно в указанном порядке). Это открывает нам возможность искать новые зависимости: теперь мы можем переставлять и четвёртую переменную x. Поэтому возникает вопрос: как, например, связаны $g(q, s, p, x)$ и $g(q, s, x, p)$? На этот и подобные вопросы мы ответим в следующем параграфе.

§ 8. Подробнее о калибровке окружности. Двойное отношение. Соотношения между изгибами

В § 3 уже говорилось о связи изгиба и смены координат с калибровкой линейки. Чтобы превратить прямую в линейку, достаточно выделить на ней две точки A, B и написать у одной «0», а у другой «1». Введём функцию трёх переменных $k(A, B, C)$, означающую координату точки C, если 0 отмечен в точке A, а 1 — в точке B. Легко понять, что $1-k(A, B, C)=k(B, A, C)$. Функция $k(A, B, C)$ выражается формулой

$$k(A, B, C) = \frac{|A, C|}{|A, B|},$$

где $|\ |$ означает расстояние между точками. Эта функция называется ещё «простым отношением» (трёх точек A, B, C). При этом расстояния надо считать со знаком, ориентированными: $|A, B|$ (мера длины) всегда положительно, а $|A, C|$ положительно, если C лежит на прямой по ту же сторону от A, что и B. Если же A разделяет B и C, то $|A, C|$ — отрицательно. Теперь поступим аналогично тому, как мы поступили, изучая изгиб. Будем считать $k(A, B, C)$ не только «координатой» точки C (при единице измерения $|A, B|$), но функцией трёх точек. Тогда естественно изучать не только связь между $k(A, B, C)$ и $k(B, A, C)$, но и связь между $k(A, B, C)$ и $k(A, C, B)$. Она тривиально получается из формул

$$k(A, B, C) = \frac{|A, C|}{|A, B|}, \quad k(A, C, B) = \frac{|A, B|}{|A, C|}.$$

Те же самые функции, которые управляли сменой координат изгиба окружностей, появляются и здесь. Это не очень удивительно, ведь в обоих случаях речь идёт о перестановках трёх объектов, каждой перестановке отвечает функция, управляющая значениями координат. Отсюда сразу следует инволютивность функций, управляющих переменой мест двух букв, и тройственная симметрия их композиций.

Вернёмся к рассмотрению окружности. В предыдущем разделе введена функция четырёх переменных: $g(q, s, p, x)$. Свяжем её с только что сделанными построениями на прямой. Это можно сделать, например, с помощью проектирования из точки s на рис. 8. В этом случае точка s переходит в бесконечно удалённую точку. Её мы трогать пока не будем. Посмотрим на g как на функцию *трёх* переменных q, p, x. На прямой проектирования она совпадает с рассмотренной только что функцией k. На прямой проектирования точка q' играет роль нуля (точки A), p' — роль единицы (точки B), а точка x' — роль точки C (см. рис. 8). Тогда

$$g(q, s, p, x) = k(q', p', x') = \frac{|q', x'|}{|q', p'|}.$$

Теперь мы возвращаемся к функции g, так как функция k выполнила своё первое предназначение. Ранее мы изучили и даже показали в таблицах, как меняется функция g, если менять местами три первые её переменные q, s, p. Рассмотрев, как меняется функция k при перестановках переменных, мы понимаем, что происходит при перемене местами двух последних переменных: p и x. Именно,

$$g(q,s,x,p)=k(q',x',p')=\frac{|q',p'|}{|q',x'|}=\frac{1}{g(q,s,p,x)}.$$

При перестановке третьего и четвёртого аргумента g снова возникает функция $1/z$, так как $f(g(q,s,p,x))=g(q,s,x,p)=g(s,q,p,x)$ (последнее равенство было установлено уже давно). Поэтому с помощью композиций функций f и h можно выразить изменение значений функции g. Обратим внимание на появившееся равенство $g(q,s,x,p)=g(s,q,p,x)$. Перемена мест сразу в двух парах аргументов не меняет значение функции g. Подобное происходит при любом выборе пар аргументов. Равенства можно доказать, рассматривая композиции, управляющие заменами аргументов. А можно доказать и геометрически.

Именно геометрический смысл этих тождеств мы сейчас разберём. Вернёмся к рис. 7. При исходном порядке аргументов $g(q,s,p,x)$ равно изгибу окружности K в пучке (A,B). Если мы поменяем местами первый с третьим аргументом, второй с четвёртым, то получим $g(p,x,q,s)$. Это будет изгиб окружности K в пучке (L,M). Формула утверждает, что эти изгибы совпадают. Более наглядно сравнение изгибов окружностей, одни из которых касаются друг друга в точке s, а другие в точке p. Изгиб окружности B в пучке (K,A) равен $g(x,s,q,p)$, снова переставим местами аргументы: первый с третьим, второй с четвёртым — и получим равенство $g(x,s,q,p)=g(q,p,x,s)$. Правая часть выражает изгиб той же самой окружности B в пучке (C,L). Аналогично изгиб окружности M в пучке (A,C) равен изгибу окружности M в пучке (L,K). Так же можно проследить связи между другими изгибами, во всех случаях взаимосвязь определяется через перестановки аргументов функции g и легко считается с помощью композиции функций f и h (т. е. $1/z$ и $1-z$).

Детальный анализ изменения значений функции g после перестановок аргументов проводится методами теории групп. Среди всех перестановок аргументов выделяются такие перестановки, которые не изменяют значения функции g. Как было сказано, это такие перестановки, которые меняют местами две пары аргументов. Такие перестановки сами образуют группу. Это нормальная подгруппа группы всех перестановок аргументов.

Вернёмся к функции $k(A,B,C)$ (по сути это просто функция g с фиксированным вторым аргументом). До сих пор мы изучали, как меняются функции k и g при перестановках аргументов. Теперь же мы рассмотрим новые переменные. Мы уже указывали на связь функции k с переходом от одной системы мер к другой. Выполняется тождество

$$k(A, B, C) \cdot k(A, C, D) = k(A, B, D).$$

Оно равносильно тождеству

$$k(A, B, C) \cdot k(A, C, D) \cdot k(A, D, B) = 1.$$

Функция k есть «усечённая» функция g (с фиксированным вторым аргументом). Поэтому имеет место следующее тождество:

$$g(q, s, p, x) \cdot g(q, s, x, y) \cdot g(q, s, y, p) = 1.$$

Здесь первые два аргумента в каждом сомножителе неизменны, а остальные три аргумента циклически сменяют друг друга, точно так, как это было для функции k.

Геометрический смысл этого тождества иллюстрируется на рис. 13. На окружности I выделены пять точек q, s, p, y, x и проведены все возможные перпендикулярные I окружности через пары этих точек. Всего имеется десять таких окружностей. Если рассматривать внутренность I, то дуги этих окружностей образуют интересную фигуру. Окружности, имеющие общую точку, обязательно в ней касаются друг с другом. В каждой из пяти точек появляются пучки окружностей, в которых можно измерять изгибы.

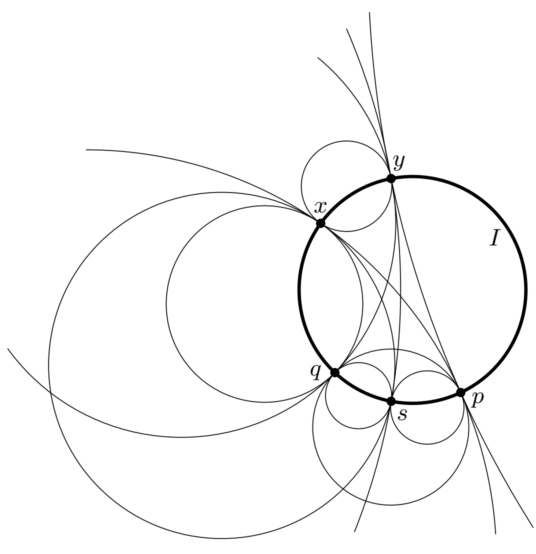


Рис. 13

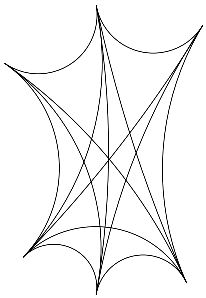


Рис. 14

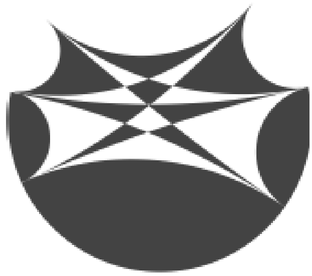


Рис. 15

Интересующее нас тождество связывает изгибы окружностей, проходящих через s. Пользуясь им и правилами замены аргументов функции g, можно установить много других тождеств, связывающих изгибы построенных окружностей.

Приведём аналогичное построение, сделанное на шести точках (см. рис. 14). От всего рисунка оставлены только дуги касающихся окружностей, лежащие внутри перпендикулярной им всем I. Тут есть много интересных соотношений, но я сейчас предлагаю просто порадоваться красивой фигуре. А на рис. 15 изображена эта же фигура, но немного обработанная и повёрнутая.

Вернёмся к функции g. Мы нашли много её свойств, хочется выразить её формулой. Численное выражение функции $g(q, s, p, x)$ естественно искать в координатном виде. Пусть теперь буквы q, s, p, x означают не только точки, но и их координаты. Скажем, если $q = 0$, $s = \infty$, $p = 1$, то мы можем записать $g(0, \infty, 1, x) = x$. В общем случае ответ оказывается удивительно прост:

$$g(q, s, p, x) = \frac{(x-q)/(x-s)}{(p-q)/(p-s)}.$$

Или, используя обозначения, употреблённые в выражении для k:

$$g(a, b, c, d) = \frac{(d-a)/(d-b)}{(c-a)/(c-b)}.$$

С помощью приведённой формулы удобно проследить, как меняются значения g при перестановке переменных. Обратим внимание, что функция g является *инвариантом*: g выражает свойство четырёх точек, не зависящее от системы координат. *Инвариантность* функции g проявляется и в том, что если мы отразим четыре входящие в неё точки относительно какой-нибудь окружности — значение g на получившихся точках будет тем же самым, что и на их прообразах.

Как доказать, что функция g имеет приведённый вид? Есть много вычислительных способов, связанных с переходом от окружности I к прямой. На самом деле функция g, полагаю, первична. Выше мы указали тождество для пяти точек и функции g. Это тождество можно положить в основу определения умножения. Также, опираясь на g, можно определить и сложение, и вычитание. Здесь мы не будем этим заниматься, а в последнем параграфе укажем, как из геометрии окружности построить комплексные числа и более общие структуры. В завершение параграфа заметим: функцию g называют также *двойным отношением* (определённым на точках a, b, c, d), и эта функция широко используется в проективной геометрии. Геометрия окружности и понятие изгиба позволяют ввести его очень естественно.

§ 9. Уравнения и предельные точки. Геометрический смысл замены букв в слове

А какие вообще уравнения можно составить, оперируя симметриями относительно нескольких окружностей? Не так-то много: если Q и V — слова из нескольких букв-окружностей алфавита, то все уравнения имеют вид $Q(x) = V(x)$. Это уравнение равносильно уравнению $F(x) = x$, где $F = V^{-1}Q$. Как найти решение такого уравнения?

Прежде чем искать решение, надо уяснить, *что* мы считаем *решением* такого уравнения. Разумно считать x не произвольным объектом, а точкой или окружностью. Есть и третий вариант, качественно отличный от предыдущих: F — некое преобразование, и мы можем считать x тоже неким преобразованием. Тогда запись $F(x)$ приобретает новый смысл, её нужно прочитать как $F * x * F^{-1} = x$, что означает коммутативность преобразований F и x. Тем самым решение уравнения в этом случае сводится к поиску коммутирующих преобразований, что мы не будем разбирать, так как это уведёт от геометрии к чистой теории групп. Вернёмся к двум геометрическим случаям: x — точка или x — окружность. Прежде всего заметим, что если в алфавите всего три буквы-окружности, то окружность $x = I$, перпендикулярная им всем, является решением *любого* уравнения $F(x) = x$, так как I остаётся неподвижной при симметриях относительно всех трёх букв алфавита.

Рассмотрим случай, когда x — точка, т. е. мы ищем неподвижные точки преобразования F. Как показано в § 5, буквы-окружности алфавита задают некое множество предельных точек W, куда стягиваются все симметричные слова достаточно большой длины. Туда стягиваются и все точки плоскости. Решения любого уравнения $F(x) = x$ лежат в этом множестве. Нетрудно видеть, что к x будут стягиваться близкие (на самом деле не только близкие) точки.

Теперь нетрудно найти приближённый метод решения уравнения $F(x)=x$. Мы уже использовали его для частного случая $C(B(A(x))) = x$. Возьмём произвольную точку z и начнём на неё действовать словом F, получится последовательность точек z, $F(z)$, $F(F(z))$, $F(F(F(z)))$; эта последовательность имеет предел, и этот предел обязательно является решением уравнения $F(x) = x$. Рассмотрим уравнение $F^{-1}(x) = x$. Ясно, что решение нового уравнения будет решением старого, и наоборот. Обозначим для удобства $G = F^{-1}$ и рассмотрим последовательность z, $G(z)$, $G(G(z))$, $\ldots$ Предел этой последовательности даёт ещё одно решение рассматриваемого уравнения. Итак, мы видим, что уравнение $F(x) = x$ всегда имеет решение и может иметь два решения. Возможен случай, когда оба эти решения совпадают: например, в рассмотренном ранее случае $F = AB$, $F^{-1} = BA$, где A и B — касающиеся друг друга окружности.

Мы видим некоторое сходство уравнений типа $F(x) = x$ с квадратными уравнениями: они тоже могут иметь два корня или один корень (когда оба корня совпадают). Правда, в отличие от квадратных, наши уравнения обязательно имеют корень. И мы не выяснили — могут ли они иметь *больше* двух корней? Нам станет ясно, как обстоит дело, после рассмотрения основного случая: когда алфавит состоит из трёх взаимокасающихся окружностей A, B, C. В этом случае легко показать, какую функцию выражает то или иное слово трёхбуквенного алфавита. Эти функции суть композиции симметрий относительно A, B, C, а эти симметрии выражены формулами в § 6. Их композиции в (A–B) координатах, как показано там же, обязательно имеют вид дробно-линейного преобразования:

$$F(x) = \frac{Mx + K}{Nx + T},$$

где M, K, N, T — какие-то вещественные (в нашем случае — целые, но для дальнейшего это неважно) коэффициенты. Наше уравнение сводится к виду $x = (Mx + K)/(Nx + T)$; домножив обе его части на знаменатель дроби, получим квадратное уравнение. Итак, в данном случае всякое уравнение сводится к квадратному, которое мы легко решаем.

В этом рассуждении есть неточность. Ведь мы свели слово-преобразование F к дробно-линейной функции только для того случая, когда точка x лежит на I — окружности, перпендикулярной A, B, C. Но *все* решения этого уравнения обязательно лежат на I, ведь, как было показано ранее, решения такого уравнения обязательно лежат в предельном множестве W, куда стягиваются при многократном действии симметрий относительно окружностей — букв алфавита все точки плоскости. А в нашем случае W и есть окружность I.

Мы показали в § 6, что действие всякого слова в данном трёхбуквенном алфавите на точки окружности I выражается численно с помощью дробно-линейной функции. Оказывается, почти то же самое верно в гораздо более общем случае. Если точка не лежит на окружности I, то всякое слово также будет действовать на неё как некоторое дробно-линейное преобразование. Но — с комплексными коэффициентами. Также и сама точка плоскости выражается комплексным числом. На самом деле всякое слово в *любом* алфавите, оказывается, можно представить дробно-линейным преобразованием (считая точку плоскости комплексным числом). Мы здесь не будем в это углубляться, связь комплексных чисел, вообще структуры *поля* и кватернионов с исчислением симметрий и биплетным исчислением — тема отдельного разговора.

Сведение слов-преобразований к дробно-линейным преобразованиям полезно, но при этом теряются некоторые важные идеи. Одной из них мы сейчас коснёмся. Мы рассматривали уравнение $C(B(A(x))) = x$ и нашли его решение в $(A–B)$ координатах. С приведённым уравнением связано ещё несколько подобных ему: $A(B(C(x)))=x$ или $B(A(C(x)))=x$ — каждый раз мы как-то переставляем буквы A, B, C. Достаточно решить одно из них — то, с которого мы начали, и с помощью замены координат найти решение требуемого. Пусть теперь у нас есть произвольное уравнение $F(x) = x$, где F — некое слово трёхбуквенного алфавита. Три буквы алфавита можно менять шестью перестановками, и каждая перестановка букв определяет (индуцирует) изменение слова F. Таким образом, из F мы получим некое слово F'. Назовём полученное слово F' сопряжённым с F (с помощью данной перестановки). Тогда и F' также сопряжено с F (с помощью обратной перестановки). Ещё одно свойство сопряжения слов — транзитивность. Если первое слово сопряжено со вторым, а второе с третьим, то первое также сопряжено с третьим при помощи композиции подстановок, сопрягающих первое со вторым, а второе — с третьим. Уравнение $F'(x) = x$ будет называться сопряжённым с уравнением $F(x) = x$.

Верно и более общее утверждение. Если у нас есть произвольные слова $F, G, H, \ldots$ и мы можем про эти слова, про составленные из них уравнения что-то сказать, тогда то же самое можно сказать и про сопряжённые с ними слова F', G', H' (разумеется, сопрягать все слова должна одна и та же перестановка букв алфавита).

Теперь мы воспользуемся равноправностью исходных окружностей и геометрически проиллюстрируем сопряжение слов. Перестановка, меняющая буквы A и B, геометрически реализуется симметрией, меняющей местами окружности A и B (и оставляющей окружность C на своём месте). Это — симметрия относительно серединной окружности или биссектрисы между A и B. Как показано в учебнике [7], окружность C обязательно перпендикулярна этой биссектрисе. Её легко описать численно с помощью (A–B) координат: она пересекает I в точке касания A и B, то есть в точке s, имеющей координату ∞, и в точке с координатой $1/2$, которая симметрична ∞ относительно пары точек 0, 1. Перестановка, меняющая местами окружности B и C, геометрически реализуется биссектрисой между B и C, при симметрии относительно этой биссектрисы окружность A остаётся на месте. Перестановка, меняющая местами A и C, реализуется биссектрисой между ними. Как и в первом случае, координаты точек пересечения этих биссектрис с I легко вычисляются. Пользуясь симметриями относительно этих трёх биссектрис, мы можем как угодно переставлять буквы-окружности A, B, C на чертеже или в слове-преобразовании.

Обозначим биссектрису между A и B буквой D: $D(A) = B$. Тогда результат перестановки букв A и B в произвольном слове w можно выразить как DwD. Аналогично можно поступить и с перестановками, меняющими местами другие буквы. Тем самым абстрактная операция замены букв в слове моделируется геометрическими преобразованиями — симметрией между окружностями алфавита. Мы могли бы просто ввести в алфавит букву-окружность D и ничего не говорить о перестановках букв A и B — это обеспечивалось действием симметрии относительно D, чисто геометрически. Правда, ранее требовалось, чтобы окружности алфавита не пересекали друг друга и были определённым образом расположены на плоскости. Это было нужно, чтобы слова, составленные из разных букв, означали разные преобразования. Включение в алфавит окружности D нарушает это требование: D пересекает C, разделяет A и B — то и другое было запрещено для окружностей алфавита. После нарушения запрета стали появляться и равенства: теперь есть разные слова, означающие одно и то же преобразование, например $DAD = B$ просто в силу определения биссектрисы. Аналогично обстоит дело с двумя другими биссектрисами-перестановками (между B и C и между A и C).

Замена букв в слове — чисто формальная операция. Возможность геометрически её истолковать и проиллюстрировать появляется потому, что любая из окружностей алфавита перпендикулярна биссектрисе между двумя другими. Если бы исходные окружности этим свойством не обладали (а, например, окружности A, $B(A)$, $B(C)$ этим свойством не обладают и ничто не мешает выбрать именно эти окружности в качестве алфавита) — то перестановки букв в словах не имели бы геометрического истолкования. Заметим, что *четыре* окружности на *плоскости* могут быть расположены таким образом, чтобы все они были «равноправны и одинаково расположены». Таковы четыре сокасающиеся окружности. А вот пять окружностей на плоскости не могут быть «равноправны и одинаково расположены». Поэтому если мы хотим геометрически рассматривать слова пятибуквенного алфавита и перестановки букв в них — нам надо перейти в трёхмерное пространство и рассматривать пятибуквенный алфавит, составленный из сфер. Четыре сокасательные сферы также обладают нужным свойством: они «равноправны и одинаково расположены». Сфера, относительно которой симметричны две сферы-буквы такого алфавита, обязательно перпендикулярна двум оставшимся сферам-буквам алфавита. Поэтому все перестановки четырёх или пяти сокасающихся сфер-букв можно рассматривать как композиции симметрий относительно сфер-биссектрис.

§ 10. Поиск закономерностей в знаковых системах. Инверсионная диаграмма

Теперь мы применим концепцию алфавита из букв-окружностей и слов-преобразований к моделированию и поиску закономерностей. Азбука Морзе построена всего на двух сигналах: точке и тире, коротком и длинном. Но в ней, как и на письме, встречаются два одинаковых символа, стоящих подряд. Если в исходный набор (алфавит) входит три или более символа (буквы), то можно соблюсти ограничение, накладываемое геометрией окружности: два одинаковых символа не могут стоять рядом.

Пусть у нас есть какое-то устройство, испускающее в каждый момент времени один из трёх данных сигналов (позднее мы увидим, что те же методы годятся и для случая, когда исходных сигналов не три, а сколь угодно много). Под «сигналом» мы можем понимать всё что угодно: например, скрипач играет на скрипке и мы записываем нотами его музыку. Заметим, что если скрипач играет непрерывно, то два одинаковых звука (ноты) не могут стоять в записи одновременно: звук ноты «до» сольётся со звуком следующей ноты «до». Или сигналы подаёт оператор, нажимая на одну из трёх (или большее число) кнопок, а мы фиксируем, какую кнопку в данный момент он нажимает. В этом случае ещё ясней, что запрет одному и тому же сигналу стоять в записи два раза подряд имеет смысл. На клавиатуре мы, конечно, можем дважды подряд нажать на букву «A» и получить «AA». Но ведь перед тем, как нажать второй раз на букву «A», мы должны её отпустить, а это самостоятельное действие, для выражения которого надо бы создать специальный знак (сигнал).

Пусть у нас имеется три возможных сигнала, обозначим их, как обычно: A, B, C. Тогда всё происходящее можно представлять двояко. Первый, простейший вариант: в каждый момент мы регистрируем один какой-то сигнал из трёх. Тогда вся запись происходящего имеет примерно такой вид:

$A, B, C, A, C, B, \ldots$ — последовательность букв A, B, C, в которой никакая буква не записана два раза подряд. Второй вариант нам по сути не понадобится. Но метод, который будет предложен для первого варианта, годится и в этом случае (в слегка изменённом виде). Во втором варианте в каждый момент времени мы получаем не одну букву, а целое слово трёхбуквенного алфавита.

Чтобы моделировать происходящее методами эстетической геометрии, возникающие слова удобней писать «справа налево». То есть сигнал-символ, полученный первым, писать справа, слева к нему приписывать следующий символ, и так далее. В итоге сигнал, пришедший первым, читается последним, а всё слово переворачивается. В рамках нашей модели мы представим каждый сигнал-букву окружностью алфавита, а каждое слово после этого окажется зримо представленным *точкой* на плоскости. Наша цель — установить закономерности, описывающие возникающее слово.

Простейший случай закономерности — периодичность. Как из хаоса случайных слов трёхбуквенного алфавита научиться быстро выделять слова, обладающие периодичностью? Этот вопрос легко формализовать, а геометрия окружности даёт и простой ответ. Слово q называется периодическим, если оно представимо в виде $vvv\ldots = q$, где v — какое-то слово меньшей длины. Это означает, что слово q получается просто многократным повторением одного и того же слова v.

Пусть наши три буквы-окружности алфавита A, B, C — три сокасающиеся окружности. На самом деле то, что они касаются друг друга, не очень существенно, но это упрощает дело, к тому же данный случай подробно разобран. Возьмём для начала произвольную точку x, лежащую вне всех окружностей алфавита, и построим, исходя из слова, получающегося при записи поступающих сигналов, последовательность точек следующим естественным образом: «нулевая» точка последовательности — это сама точка x, первая точка — $A(x)$, вторая $B(A(x))$, третья $C(B(A(x)))$, четвёртая $A(C(B(A(x))))$, и так далее. После фиксации каждого сигнала A, B, C мы отражаем точку относительно соответствующей окружности. Таким образом, всякое слово (последовательность сигналов) означает не только композицию симметрий, но и результат действия этой композиции на точку x. В данном случае мы изобразили слово $\ldots BCACBA$, которое уже приводилось в пример. Назовём это построение *Инверсионной диаграммой* слова $\ldots BCACBA$.

На рис. 16 мы видим инверсионную диаграмму слова $CABCBA$. Видно, что уже под действием слова длины три CBA точка x очень близка к I, перпендикулярной A, B, C. Следующие точки инверсионной диаграммы будут ещё ближе к I. Внутри каждой окружности алфавита лежит по две

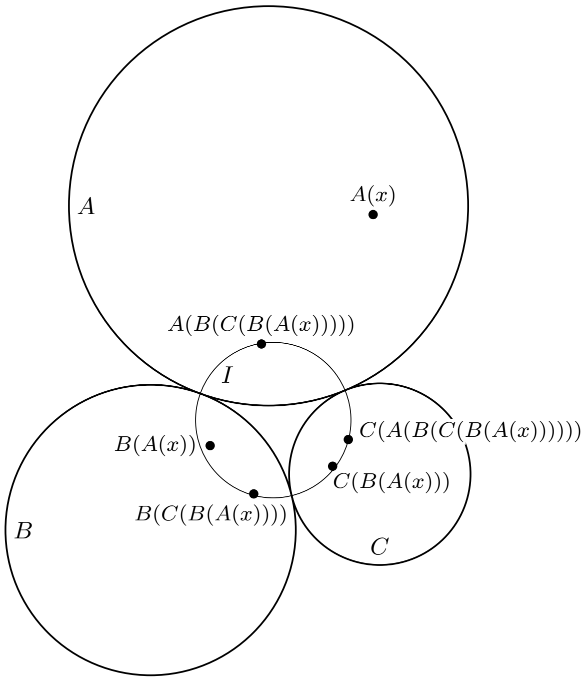


Рис. 16

точки инверсионной диаграммы, и каждая буква входит в исследуемое слово по два раза. Никаких областей сгущения на диаграмме не наблюдается (кроме неизбежного приближения точек к I).

Поясним определение *инверсионной диаграммы*. Пусть у нас есть слово q. Его инверсионной диаграммой называется последовательность точек, полученных из исходной точки x композицией симметрий относительно букв q. Композиция осуществляется в том порядке, в каком эти буквы-окружности расположены в слове q. Обратим внимание на два обстоятельства: исходная точка x выбирается почти произвольно (важно только, чтобы она была вне окружностей алфавита), следует указать направление, в котором читается слово q (справа налево или слева направо), обычно это ясно из контекста. Каждая точка диаграммы представляет «подслово» слова q, т. е. часть слова, прочитанную от его начала. Обозначим подслово длины k через q_k.

Расположение точек диаграммы слова q сообщает информацию о самом слове q и о его подсловах. Чтобы считывать эту информацию, надо воспользоваться соображениями, изложенными в § 5, где анализируются качественные характеристики слов-преобразований. Диаграмма сразу сообщает нам, сколько раз в слово q входит та или иная буква алфавита:

столько точек на диаграмме лежит внутри соответствующей окружности алфавита. В самом деле, буква входит в слово столько раз, сколько существует подслов, начинающихся с неё. А всякая точка диаграммы, представляющая подслово, начинающееся с данной буквы, — лежит внутри данной буквы-окружности. Что и требовалось.

Мы можем усложнить задачу: сколько раз в данное слово q входит какое-то слово меньшей длины q'? Начнём со случая, когда слово q' симметрично, то есть представляет собой некоторую окружность. Эта окружность легко может быть получена композицией симметрий. Количество точек диаграммы, лежащих внутри неё, и даёт ответ на вопрос. Пусть слово q' не симметрично. Внутри каких слов-окружностей обязательно лежит точка вида $q's(x)$, где s — произвольное слово? Пусть r — последняя буква слова q', $q' = q''r$ (где r, разумеется, есть A, B или C). Всякая точка y (лежащая вне r) под воздействием слова-преобразования $q''r$ перейдёт в точку, лежащую внутри $q''(r)$. Эта окружность выражается словом $q''rq''$. Поэтому и точка $s(x)$ лежит внутри этой окружности, и в ней под действием слова $q = q's$ окажется x. Итак, для того чтобы узнать, сколько раз слово q' входит в слово q, достаточно изобразить окружность $q''rq'' = q''(r)$ и подсчитать, сколько точек инверсионной диаграммы слова q лежит внутри неё. Изобразить её просто, эта окружность есть результат симметрий окружностей алфавита A, B, C друг относительно друга.

Обычно нам интересно применять инверсионную диаграмму, если слово q формируется длинной последовательностью входящих сигналов. Представим себя оператором, следящим за дисплеем, где динамически отображается инверсионная диаграмма: после приёма нового сигнала сразу появляется новая точка на диаграмме. В этом случае места, где точки расположены гуще, сразу становятся видны, и мы легко определим, внутри каких слов-окружностей эти места лежат. А это определяет, какие буквы алфавита или какие комбинации этих букв чаще входят в длинное слово q, чем другие буквы или комбинации.

Используем теперь инверсионную диаграмму для поиска периодичности в q. Для этого снова надо присмотреться к густоте точек на диаграмме. Пусть слово q периодично, т. е. имеет вид $q = www \ldots w$ (n раз). Тогда $q = w^n$ и $q(x) = w^n(x)$. В инверсионную диаграмму слова q входят точки x, $w(x)$, $ww(x)$, $www(x)$ и так далее. Эти точки сгущаются к неподвижной точке слова-преобразования w. Мы можем найти эту точку, решив уравнение $w(x) = x$. Как решать такие уравнения, разобрано в предыдущем параграфе. Фактически мы не знаем слово w, мы знаем длинное слово q и ищем w. Мы даже не знаем, существует ли w (т. е. периодично ли слово q), мы хотим иметь способ быстро узнавать, есть ли периодичность, и если есть — каков период (т. е. найти слово w). Мы установили, что если периодичность есть, то на инверсионной диаграмме обязательно есть точка сгущения (корень уравнения $w(x)=x$). Это — не единственная точка сгущения, будет ещё $k-1$ точка сгущения, где k — длина слова. Обозначим искомое решение z, тогда точки $w_1(z)$, $w_2(z)$, $w_3(z)$, $\ldots$, $w_{k-1}(z)$, где все w_i — подслова в w, также будут точками сгущения на инверсионной диаграмме слова q, а других точек сгущения у периодического слова быть не может. Обнаружив их на инверсионной диаграмме, можно определить и само слово w. Мы можем представить себе не слово q, а последовательность сигналов, поступающих оператору. Все сигналы отображаются на экране с помощью инверсионной диаграммы. Если сигналов очень много и меняются они быстро, то видеть изображение всех сигналов сразу неудобно, мы можем увидеть лишь «грязь». Что мы увидим, если изображается только последний сигнал? Если последовательность сигналов периодическая, мы будем видеть непрерывно мигающие точки экрана в k местах — описанных ранее точках сгущения.

Мы рассматривали инверсионную диаграмму в случае, когда алфавит (базовые сигналы) состоит из трёх букв. Если букв-сигналов больше, можно поступить двумя способами. Можно промоделировать трёхбуквенным алфавитом алфавит с любым числом букв (как с помощью двух символов морзянки можно выразить любую букву любого алфавита). А можно увеличить число букв-окружностей в алфавите, чтобы каждый сигнал соответствовал какой-то букве. Если в алфавите четыре и более окружностей, то инверсионная диаграмма может выглядеть очень интересно. Ранее все точки диаграммы обязательно прижимались к окружности I (перпендикулярной всем трём окружностям алфавита), теперь уже нет никакой окружности, перпендикулярной всем окружностям алфавита, и инверсионная диаграмма может составить интересный рисунок фрактального типа. Тема эта столь широка, что мы не будем в неё входить. Заметим лишь, что важные свойства инверсионной диаграммы *помехоустойчивы*. Это означает, что если при приёме сигналов оператор напутал и внёс сигнал-букву в слово неправильно, то это локальное искажение не повлияет на общие выводы.

§ 11. Подведение итогов и пути развития эстетической геометрии

Здесь мы обсудим не только перспективы эстетической геометрии, но и её связи с другими разделами математики (напомню тезис из предисловия: «всякое математическое понятие может содержательно иллюстрироваться в эстетической геометрии»). И даже возможные связи эстетической геометрии с другими науками. Начнём с таблицы, собирающей воедино некоторые результаты вычислений изгибов. В таблице — базовые зависимости. Окружности задаются двумя точками (точками пересечения с I). Всего есть шесть координат, но здесь мы приводим лишь три, остальные три можно получить перестановкой окружностей, с помощью функции $h(x) = 1 - x$.

№	Объект	Базис: окружности A, B, C		
		Координата A–B	Координата C–A	Координата B–C
1	Точка x на окружности I	x	$1-\frac{1}{x}$	$\frac{1}{1-x}$
2	Окружность A	∞, 0	1, ∞	0, 1
3	Окружность B	∞, 1	1, 0	0, ∞
4	Окружность C	1, 0	0, ∞	∞, 1
5	Точка касания C и A	0	∞	1
6	Точка касания B и C	1	0	∞
7	Точка касания A и B	∞	1	0
8	Произвольная окружность H, ортогональная I	k_1, k_2	$1-\frac{1}{k_1}$, $1-\frac{1}{k_2}$	$\frac{1}{1-k_1}$, $\frac{1}{1-k_2}$
9	$H(x)$	$\frac{2k_1k_2-x(k_1+k_2)}{k_1+k_2-2x}$		
10	$A(x)$	$-x$	$1+\frac{1}{x}$	$\frac{1}{x+1}$
11	$B(x)$	$2-x$	$1-\frac{1}{2-x}$	$\frac{1}{x-1}$
12	$C(x)$	$\frac{x}{2x-1}$	$1+\frac{1}{x}$	$2+\frac{1}{x-1}$
13	Окружность W пучка (A,B)	∞, k	1, $1-\frac{1}{k}$	0, $\frac{1}{1-k}$
14	$W(x)$	$2k-x$	$1-\frac{1}{2k-x}$	$\frac{1}{1-2k+x}$

В первых строках показано, как при смене окружностей меняются координаты выбранной точки (произвольной точки x и точек касания окружностей A, B, C). В восьмой строке парой точек k_1 и k_2 задаётся некоторая окружность. Далее вычисляется, как выглядит симметрия относительно пары точек (т. е. относительно проходящей через них окружности) в разных случаях. Обратим внимание на строку 9. Заполнить две недостающие её графы я предлагаю читателю самостоятельно. Эта формула ранее не приводилась. Её несложно получить, например инвертируя I в окружность, а можно использовать свойства функции g. Эта формула требует осмысления. Во-первых, обратим внимание, что $H(x)$ — частное линейных функций, т. е. дробно-линейное преобразование. Во-вторых, $H(x)$ симметрично относительно k_1 и k_2. Это естественно, ведь окружность H задаётся парой точек k_1 и k_2 и не важно, в каком порядке они указаны.

В-третьих: введём обозначение

$$y = \frac{2k_1k_2 - x(k_1 + k_2)}{k_1 + k_2 - 2x}.$$

Здесь y — координата точки, симметричной x относительно H. Избавимся от знаменателя, перенесём всё в правую часть, получим

$$0 = 2xy - (x + y)(k_1 + k_2) + 2k_1k_2.$$

Полученное выражение симметрично не только относительно k_1 и k_2, но и относительно x и y. Это неудивительно, ведь $H(x)$ — симметрия, поэтому в уравнение, связывающее $H(x) = y$ и x, переменные x и y *должны* входить симметрично. Ранее было указано, что всякая композиция симметрий представима дробно-линейным преобразованием. Пользуясь этим рассуждением (но начав с другого конца, с общей симметричной зависимости x и y), легко показать, какие дробно-линейные преобразования задают симметрию.

В каких случаях $x = y$ (т. е. $H(x) = x$)? Подставив $x = y$ в уравнение, получаем: $0 = x^2 - x(k_1 + k_2) + k_1k_2$. Решая квадратное уравнение, получаем два корня $x = k_1$, $x = k_2$. Результат легко было предсказать: точки k_1 и k_2 в самом деле неподвижны при симметрии относительно H. Собственно, именно таким образом и была найдена формула строки 9. Теперь обратим внимание на следующее: мы начали рассуждение с точек k_1 и k_2. Поэтому полученное уравнение заведомо имеет решение в вещественных числах. Но мы могли бы начать с дробно-линейного преобразования или с какого-то симметричного уравнения, связывающего x и y. Тогда, находя неподвижные точки этого преобразования, мы бы получили квадратное уравнение $0 = x^2 - vx + w$. Если это уравнение имеет решения в действительных числах, то они дают координаты неподвижных точек. А если решения — комплексные числа? Есть ли какой-то геометрический смысл у этих комплексных решений? Оказывается — есть. Решения можно изобразить парой точек, симметричных относительно I, а само преобразование оказывается биплетной симметрией относительно этой пары точек (точки на I при такой симметрии переходят друг в друга).

Мы пришли к двум важным темам: биплетная симметрия и связь эстетической геометрии с комплексными числами. Тема биплетной симметрии подробней рассмотрена на сайте [7]. В статье же даётся определение симметрии относительно пары точек или биплета только для частного случая, когда все точки лежат на I. Это же определение можно дать на основе гармонического разделения точек. В исчислении биплетов (то есть в исчислении композиции симметрий относительно них) важнейшую роль играет понятие перпендикулярных или коммутирующих биплетов. По сути это переименование гармонического разделения: два биплета перпендикулярны (коммутируют), если точки одного из них симметричны относительно другого. Тут действует *принцип взаимности*. Зачем же переименовывать симметричные или гармонические биплеты в «перпендикулярные»? Дело в том, что свойства симметрий биплетов похожи на свойства симметрий относительно прямых в трёхмерном евклидовом пространстве. И полностью совпадают со свойствами симметрий относительно прямых в трёхмерном пространстве Лобачевского.

В трёхмерном пространстве для двух прямых в общем случае существует одна прямая, перпендикулярная им обеим. Точно так же для двух биплетов существует третий, перпендикулярный (коммутирующий) с ними обоими. (За одним исключением — когда у пар точек, образующих два биплета, есть общая точка: тогда нет симметрии, коммутирующей одновременно с биплетами (v, w) и (v, z)). Заметим, что композиция симметрий относительно двух перпендикулярных биплетов сама является биплетной симметрией (так же как композиция симметрий относительно двух перпендикулярных прямых есть симметрия относительно перпендикулярной им обеим прямой), а шесть точек, образующих три упомянутых биплета, являются «великолепной шестёркой» точек, разбираемой подробно на сайте [7]: это точки касания четырёх взаимокасающихся окружностей или точки пересечения трёх взаимоперпендикулярных. Последним свойствам биплетной симметрии уже нет аналогов в мире прямых.

Наличие у двух биплетов перпендикулярного им позволяет сразу установить важнейшее свойство функции g. Вернёмся к рис. 13, где мы сравниваем изгибы на четырёх точках p, s, q, x. Утверждалось, что значение функции g не меняется, если поменять местами две пары переменных. Сейчас мы более геометрично объясним, почему это происходит. Пусть мы меняем местами p и s, q и x. Существует биплет, меняющий местами указанные точки, — это биплет, перпендикулярный биплетам (p, s) и (q, x). Но при биплетной симметрии все изгибы сохраняются, все свойства симметричных точек совпадают. Следовательно, значение функции g не изменится.

Рассмотрим (на плоскости или на окружности I) совокупность биплетов, перпендикулярных данному. Оказывается, композиция симметрий относительно любых трёх из них — снова биплетная симметрия. Рассмотрим совокупности композиций относительно любых пар биплетов, коммутирующих с данным. Все полученные композиции коммутируют друг с другом. Они образуют группу, которая изоморфна группе умножения комплексных чисел (если мы рассматриваем только биплеты, лежащие на I, то будет изоморфизм с группой умножения действительных чисел). Чтобы определить сложение, надо рассмотреть совокупность биплетов, у которых есть одна общая точка. Кстати, привычную центральную симметрию можно рассматривать как биплетную: это симметрия относительно пары точек, одна из которых перед нами (центр симметрии), а вторая бесконечно удалена. В биплетном исчислении много интересных возможностей. Заметим, что некоторые свойства плоской геометрии теряются при переходе в трёхмерное пространство: тогда совокупность биплетов, коммутирующих с данным, не обладает указанными выше свойствами. Но на основании этой совокупности также можно построить интересную структуру, видимо совпадающую с кватернионами.

Прежде чем отойти от «таблицы координат», укажем на её важность для плоской геометрии Лобачевского, а именно модели Кэли — Клейна. В ней каждой прямой геометрии Лобачевского ставится в соответствие хорда окружности I. У нас эта хорда определяется парой точек на I (определяющих биплетную симметрию или симметрию относительно окружности, перпендикулярной I). Таким образом, свойства симметрий относительно прямых геометрии Лобачевского сводятся к свойствам симметрий относительно пары точек, которые описываются строкой 9 таблицы.

Теперь подойдём к комплексным числам и геометрии окружности с другой стороны. Известно, что с помощью композиции симметрий любые три точки плоскости можно отобразить в любые три другие точки. Кстати, простейшее доказательство основано на рассмотрении окружностей, касающихся друг друга в этих точках. Оно дано на сайте [4] (статья 9, теорема об отображении трёх точек). Из этого следует, что с точки зрения геометрии окружности любые три точки равноправны. Если мы можем в рамках геометрии окружности сделать какое-то построение, используя три *данные* точки, то точно такое же построение даст точно такой же результат на *любых* трёх точках. А вот четвёрки точек отличаются друг от друга в геометрии окружности. Мы видели, что для четырёх точек, лежащих на одной окружности, можно ввести «координаты» или изгиб. Каждой четвёрке точек на окружности функция g сопоставляет *вещественное* число, и это число является инвариантом данной четвёрки точек. После симметрий данная четвёрка точек перейдёт в какую-то другую четвёрку, но значение функции g не изменится. Соответственно, если $g(A, B, C, D) \neq g(A', B', C', D')$, то и невозможно композицией каких-либо симметрий отобразить одновременно точки A в A', B в B', C в C', D в D'. Если же точки A, B, C, D *не лежат* на одной окружности, то им можно сопоставить *комплексное* число, причём это сопоставление происходит при помощи функции g', по всем своим свойствам аналогичной функции g, только она определяется на комплексных числах и её значение — комплексное число. При этом значение функции g' для случая, когда четыре её аргумента лежат на одной окружности, совпадает со значением функции g.

Связать комплексные числа и эстетическую геометрию можно также с помощью двухцентровых спиралей. Такие спирали встречаются, например, в барокко. Их примеры и свойства приводятся на сайте [7]. Совокупность спиралей, имеющих одни и те же центры (неподвижные точки, откуда всё вытягивается и куда всё втягивается), можно рассматривать как множество комплексных чисел. Умножению комплексных чисел соответствует композиция преобразований, определяющих спирали, эти преобразования коммутируют между собой. А сложение комплексных чисел определяется через биплетную симметрию.

Связь с комплексными числами и кватернионами — интересный, но внешний для эстетической геометрии вопрос. Важней изучать эстетическую геометрию, исходя из её внутренней логики, отвечать на вопросы, возникающие внутри неё самой. Попутно это может прояснить связи между уже известными областями алгебры и геометрии. Рассмотрим следующий шаг в изучении изгиба и систем координат. Мы начали с трёх касающихся окружностей A, B, C. Вернёмся к ним, рассмотрим точку x на плоскости, у неё есть три координаты $(A–B)(x)$, $(B–C)(x)$, $(C–A)(x)$. Как было сказано, никакая из них не определяет однозначно две другие. Именно чтобы избавиться от неоднозначности, рассматривались не все точки плоскости, а только лежащие на окружности I. В этом случае по одной координате точки x можно однозначно восстановить две другие, что и показано в таблице. Но в общем случае надо знать *две* координаты, например $(A–B)(x)$ и $(B–C)(x)$, чтобы найти третью, $(C–A)(x)$ в данном случае. При этом, даже найдя третью координату (она уже находится однозначно), мы не находим однозначно саму точку x. У неё остаются *две* возможности (поскольку в данном случае все три возникающие окружности проходят через две точки, по обобщённой теореме о трёх окружностях). Поэтому, чтобы точно указать точку x, надо воспользоваться ещё чем-то. Проведём четвёртую окружность D, касающуюся трёх исходных, и рассмотрим, например, $(A–D)(x)$. Зная и эту координату точки x, мы находим саму точку уже однозначно.

До этого мы шли от точки к координатам. Поставим противоположную задачу. Пусть мы знаем координаты и хотим по ним восстановить точку x. Это можно сделать не всегда. Пусть нам известна $(A–B)(x)$, это определяет окружность из пучка (A, B), на которой лежит x. Пусть нам известна ещё и $(B–C)(x)$, это определяет окружность из пучка (B, C), в которой лежит x. Но эти две окружности могут не иметь общих точек, и потому точки с такими координатами просто не существует! А что происходит, если эти окружности касаются? Точка x обязательно лежит на I, т. е. этот случай мы как раз рассмотрели. Если же окружности не имеют общих точек, то, хотя *точки* с такими координатами на плоскости нет, вопрос о том, какой же будет координата $(C–A)$, имеет смысл, поскольку существует единственная окружность из пучка (C, A), лежащая в том же *пучке*, что и две новопроведенные окружности. Мы видели многообразие возможных координат. Оно особенно интересно, если мы систематически включаем в рассмотрение окружность D (касающуюся трёх исходных) и рассматриваем пучки (A, D), (B, D), (C, D) и координаты-изгибы в них. Появляется задача: как описать все возникающие связи между координатами точки в этих пучках и что происходит, если координаты есть, а точки — нет (соответствующие окружности не имеют общих точек)? Рассматривать эту задачу стоит в свете теории групп, в терминах перестановок окружностей A, B, C, D между собой. Разумеется, задача обобщается на трёхмерный и многомерный случай. Там надо рассматривать изгибы сфер, касающихся друг друга, и в трёхмерном пространстве существует пять сокасательных сфер (в n-мерном пространстве таких сфер $n + 2$). Каждая пара сокасающихся сфер задаёт систему координат, в которой можно мерить изгибы сфер, лежащих в этом пучке. Какие есть связи между этими координатами в общем случае?

Этот вопрос имеет отношение к теории представлений. Обычное введение в задачи теории представлений требует больше математической техники и менее геометрично.

Всё проведённое исследование базировалось на тройственной симметрии, т. е. на том, что три исходные окружности (и любые их пары) обладают абсолютно одинаковыми свойствами. Аналогичное свойство есть и у четырёх сокасающихся окружностей, и у сокасающихся сфер. Но именно аналогия, а не полное сходство имеет место. Скажем, точки касания трёх окружностей можно выбрать абсолютно произвольно. А уже точки касания четырёх окружностей (великолепная шестёрка) — вовсе не произвольны. Поэтому тройственная симметрия (а не пятерная или шестерная) занимает особое место.

Тройственная симметрия, правда иначе, проявляется и в геометрии Евклида. Общеизвестнейшие примеры — теоремы о пересечении высот, биссектрис, медиан. У всех этих теорем есть аналоги в геометрии окружности, из теорем о треугольнике они превращаются в теоремы о трёхокружнике, это разобрано на сайте [7]. Сейчас мы свяжем тройственную симметрию с биплетами, т. е. с симметрией относительно пары точек. Прежде всего заметим, что произвольная пара биплетов A и B задаёт пучок биплетов. Это совокупность таких биплетов X, что $A * B * X$ — снова инволютивное преобразование, симметрия относительно какого-то четвёртого биплета. Эту совокупность легко описать геометрически. Найдём биплет Y, перпендикулярный (коммутирующий с) A и B. Тогда *все* биплеты, перпендикулярные Y, лежат в пучке, созданном A и B. Мы уже делали подобное построение, говоря о связи геометрии окружности с комплексными числами. Теперь же мы с его помощью превратим три произвольных биплета в аналог треугольника, по крайней мере научимся строить высоты в «трехбиплетнике». Пусть даны биплеты A, B, C. Назовём высотой, опущенной на C, биплет c, лежащий в пучке (A, B) и перпендикулярный C. Нетрудно доказать, что такой биплет с существует и единственен. Аналогично определяются высоты на B и A. Теорема о перпендикулярах «трехбиплетника» утверждает, что три биплета-высоты a, b, c сами лежат в одном пучке («лежать в одном пучке» в данном случае аналогично понятию «пересекаться в одной точке» для прямых). В этой теореме особенно ярко выражается тройственная симметрия. Она имеет интересные обобщения на многомерное пространство: ведь три биплета могут не лежать в одной плоскости. Биплет состоит из двух точек, в общем случае в трёх биплетах 6 точек и они лежат в пятимерном пространстве. Что можно сказать о высотах этого пятимерного трехбиплетника? Вопрос заслуживает изучения. Как и обобщения тройственной симметрии высот и биссектрис.

Есть совершенно другой подход к геометрии окружности. Его можно назвать «абстрактно алгебраическим». В математике есть различные теории (например, теория полугрупп), где изучаются записи из букв конечного алфавита и возможные равенства между составленными таким образом словами. Геометрию окружности можно строить, начиная именно с последовательности слов трёх- или четырёхбуквенного алфавита. Далее нужно рассматривать симметрии слов, возникающие после замены во всех словах одной буквы на другую, как это показано в § 9. И включать в рассмотрение эти симметрии как новые буквы (или слова?). Получающуюся структуру можно геометризировать. Ключевыми здесь, вероятно, окажутся простые соображения: между композициями симметрий относительно окружностей возможны равенства только в особых геометрических ситуациях. Если есть всего две окружности, то эти равенства могут выполняться, только если окружности пересекаются. Если они касаются — есть равенство в пределе (при композиции, проведённой бесконечное число раз). Но пока можно лишь наметить эту тему.

Возникает вопрос об аксиоматизации геометрии окружности. Его можно решить довольно просто в рамках классического подхода. Можно воспринимать геометрию окружности как раздел геометрии Лобачевского. На сайте [7] показано, как исходя из геометрии окружности определить неевклидовы геометрии. Чаще поступают наоборот, хотя этот путь и более сложен и менее нагляден. Плоскую геометрию окружности можно определить через трёхмерную геометрию Лобачевского. Но мне этот подход представляется неудачным: аксиомы должны как-то пояснять суть предмета, а при таком методе аксиоматизации ничего подобного не происходит. Мне представляется вопрос об аксиоматизации преждевременным, интереснее исследовать геометрию окружности как таковую и в связи с другими разделами математики. А может быть, подходящих со всех точек зрения аксиом и не выбрать: в конце концов у самой же окружности нет начала.

Я наметил несколько интересных вопросов, связанных с эстетической геометрией и её местом в математике. Укажу на её возможные связи с другими науками. Многие, видевшие образы, сделанные методами эстетической геометрии, «бесконечные картины» видеоарта программы dodecaLook высказывались о её значении для биологии. В самом деле, антропоморфные и зооморфные образы возникают в геометрии окружности, когда их и не ждёшь. Я полагаю, что законы эстетической геометрии пригодятся для понимания морфогенеза. С другой стороны, приведённая в §§ 5, 9, 10 связь геометрии окружности с простейшими знаковыми системами наводит на мысль о её полезности для семиотики. Я не специалист в обеих науках и не берусь подробно развивать эти соображения, но хочу привлечь к ним внимание.

И, разумеется, есть возможности для применения эстетической геометрии к классической эстетике и к психологии восприятия нами красоты.

Есть огромный простор для поиска закономерностей в произведениях искусства: живопись, скульптура, архитектура в свете симметрии относительно окружности. Особенно напрашиваются связи образов эстетической геометрии с архитектурой барокко и религиозным искусством Востока, например мандалами. Я полагаю, что систематическое изучение изобразительных искусств подобными методами может далеко расширить наше знание о красоте, а это — интересная и достойная задача для науки.

Список литературы

[1] *Бахман Ф.* Построение геометрии на основе понятия симметрии. М.: Наука, 1969.

[2] *Мамфорд Д., Райт Д., Сирис К.* Ожерелье Индры. Видение Феликса Клейна. М.: МЦНМО, 2011.

[3] *Пименов Р. Р.* Эстетическая геометрия, или Теория симметрий. СПб.: Школьная лига, 2014.

[4] `http://bogemnyipeterburg.net/revolt/matem/Bereg_site/index.html`

[5] `http://bogemnyipeterburg.net/revolt/matem/index.htm`

[6] `http://bogemnyipeterburg.net/revolt/matem/pict/index.html`

[7] `http://bogemnyipeterburg.net/revolt/matem/teachpictures/index.html`

[8] `http://bogemnyipeterburg.net/revolt/matem/Bereg_site/bereg_4.html`

[9] `http://www.centre.smr.ru/win/pics/pic0084/p0084_800.htm`

Р. Р. Пименов, ФТШ им. Иоффе «Академический Университет», Санкт-Петербург

revoltp@mail.ru