журнал© Квант СЕНТЯБРЬ ОКТЯБРЬ 2008 № 5

НАУЧНО-ПОПУЛЯРНЫЙ ФИЗИКО-МАТЕМАТИЧЕСКИЙ ЖУРНАЛ

ИЗДАЕТСЯ С ЯНВАРЯ 1970 ГОДА

В номере:

В праздновании 100-летнего юбилея академика И.К.Кикоина финансовое участие принимает ОАО «ТЕХСНАБЭКСПОРТ»

Учредители — Российская академия наук, Фонд поддержки фундаментальной науки и образования (Фонд Осипьяна), ИФТТ РАН

Бюро Квантум

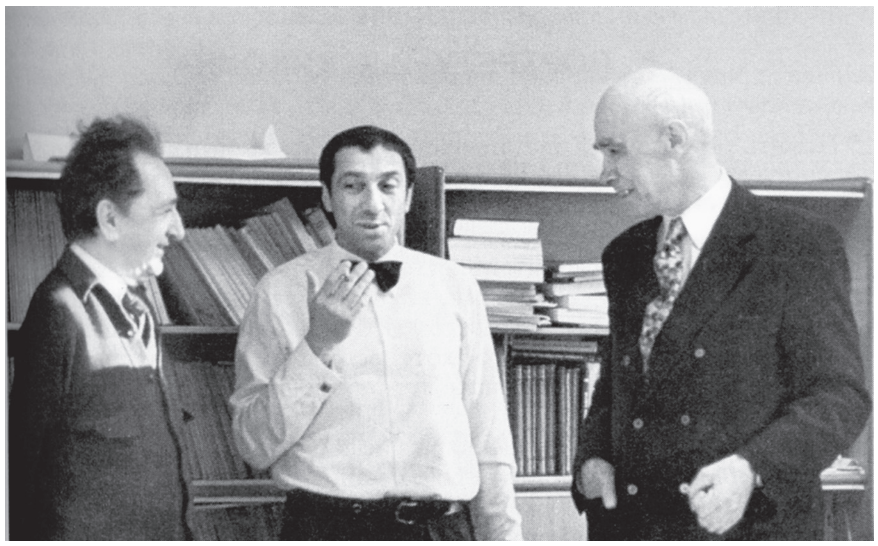
К 100-ЛЕТИЮ И.К.КИКОИНА

Физик и инженер

Я.СМОРОДИНСКИЙ

Первая моя встреча с Исааком Константиновичем была необычной. Я, аспирант Ландау, жил в одной из квартир Института физических проблем. В этой же квартире порой останавливались и гости – физики, которые приезжали в Москву. Это был 1940 год. Вечером, довольно поздно, раздался звонок в дверь. Я пошел открывать. Весь проем двери заслоняла огромная меховая медвежья шуба – никаких признаков ее владельца не было видно. Откуда-то из центра шубы появилась рука, и голос (как мне показалось, с высоты) прозвучал: «Здравствуйте, я – Кикоин».

Так я познакомился с человеком, с которым мне было суждено провести много лет.



И.К.Кикоин (справа) и Я.А.Смородинский с С.Ю.Юрским (в центре) (1973 г.)

Настоящая работа началась в 1944 году. Для новой темы, нового задания правительства Кикоин подбирал людей, которые, как он надеялся, составят сильный и дружный коллектив, способный быстро и эффективно решить целую серию задач, никогда не решаемых физиками. Теоретические направления возглавили замечательные люди: член-корреспондент АН СССР И.Н.Вознесенский, который посвятил нас в идеи регулирования, и академик С.Л.Соболев, принесший с собой высочайшую математическую культуру. Работали много и до позднего вечера. Вечером собирались у И.К.Кикоина, и наступало самое интересное и плодотворное время: начиналась дискуссия об основных успехах в сделанном и задачах на будущее.

Исаак Константинович обладал многими талантами: он был внимательным слушателем, помнил и знал все, что было связано с задачей, и мгновенно оценивал правильность и эффективность того, о чем ему рассказывали. Он был не только великим физиком, что общеизвестно, но и превосходным инженером. Дело заключалось не только в том, что он понимал конструкции и мог вносить в них важные усовершенствования, но и в том, что он ясно представлял различие между лабораторией и заводом. Он всегда четко знал ту границу, до которой надо доводить лабораторные исследования, чтобы обеспечить практический успех. Он понимал также, какую теорию надо развивать, чтобы ее можно было использовать на промышленном объекте.

Но не только о промышленных вопросах шла речь. Исаак Константинович не ослаблял своего интереса к физике, и разговоры о новостях науки всегда завершали наши неофициальные семинары. Старые воспоминания об открытии фотомагнитного эффекта часто приводили к обсуждению новых эффектов в кристаллах, которые можно исследовать в лабораториях. Многие из высказанных идей Кикоин реализовал.

Замечательным качеством Исаака Константиновича было то, что он никогда не оставлял работы в лаборатории, и занятие физикой составляло для него жизненную потребность. Поэтому он до конца своих дней оставался физиком, и это качество делало из него организатора коллектива, который был способен к интенсивной целенаправленной работе.

Интересным периодом в его жизни было время поиска фильтра для разделения изотопов. Были задействованы самые разные направления: на одном этапе медную фольгу кололи иголками, на другом – подбирали порошки. Фильтры приготовили быстро, их выбор был сделан Исааком Константиновичем безошибочно, хотя вначале казалось, что задача почти неразрешимая. Так же быстро был проведен и выбор типа машин из предлагавшихся самых разных конструкций (некоторые из них уже работали в других лабораториях).

Как можно было угадать правильное направление конструирования и отбросить неперспективные типы? Для этого надо было иметь редкую комбинацию инженерного и физического мышления. Именно этим отличалось мышление И.К.Кикоина.

С большим удовольствием представляем читателям статью академика Льва Борисовича Окуня. Он всегда уделял большое внимание популяризации науки (достаточно вспомнить его замечательную книгу «$\alpha, \beta, \gamma \ldots Z$: элементарное введение в физику элементарных частиц», вышедшую в том числе и в серии «Библиотечка «Квант»).

Несколько публикаций Л.Б.Окуня связаны с проблемами преподавания теории относительности. В январе 2008 года он сделал на эту тему доклад на сессии Российской академии наук, посвященной 100-летию со дня рождения Л.Д.Ландау. Мы предлагаем вам текст этого доклада, подготовленный для презентации в виде слайдов.

Каждый слайд – это маленькое законченное утверждение, мысль, идея, как бы мгновенная вспышка-фотография одного из ракурсов проблемы. Хотя некоторые слайды при первом чтении могут показаться сложными, прочитав текст целиком, вы получите яркое и цельное представление о современном взгляде на теорию относительности Эйнштейна.

Теория относительности и теорема Пифагора

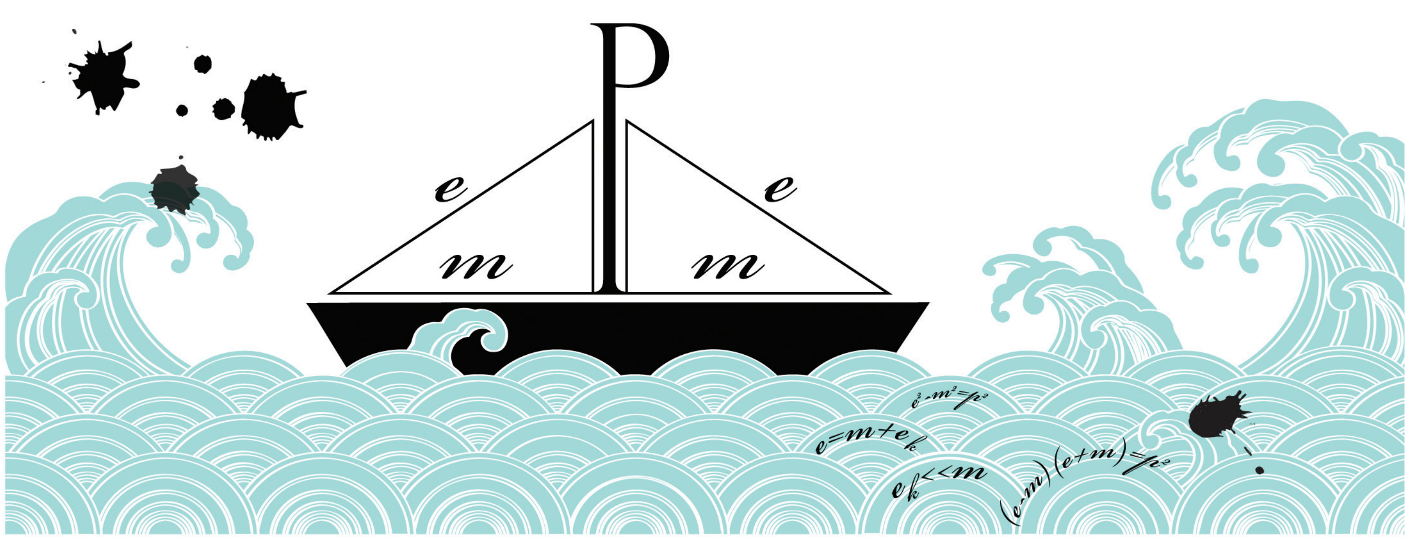
Л.ОКУНЬ

Физика XX века в корне изменила представление о том, что такое пустота и материя, и по-новому связала между собой такие свойства материи, как энергия, импульс и масса. Тем важнее не забывать высказывание, приписываемое Фоку: «Физика – наука, по существу, простая. Главная проблема в ней – понимать, что какая буква означает». Формирование фундаментальных понятий физики еще не завершено и вряд ли завершится в обозримом будущем. В частности и поэтому так важно выбирать адекватные слова и буквы при обсуждении физических явлений и теорий.

Постараюсь рассказать об основных формулах теории относительности максимально просто, пользуясь в основном теоремой Пифагора.

Относительность

Высшая точка зрения. Понятие массы в физике имеет многовековую историю. История эта очень интересна, но здесь мы не будем ее касаться. Вместо этого постараемся взглянуть на массу «с высшей точки зрения». Я использую здесь знаменитое название книги Феликса Клейна «Элементарная математика с высшей точки зрения» (оно традиционно неправильно переводится на русский язык как «Элементарная математика с точки зрения высшей»). Такая высшая,



современная точка зрения, в основе которой лежат принципы симметрии вообще и теория относительности в частности, позволяет избежать терминологической путаницы и связанных с ней парадоксов.

Принцип относительности. Со времен Галилея и Ньютона термин «принцип относительности» выражает невозможность оставаясь внутри замкнутого пространства (например, внутри корабля) любыми экспериментами обнаружить поступательное (равномерное и прямолинейное) движение этого пространства. Эйнштейн в 1905 году обобщил этот принцип на случай существования предельной скорости распространения сигналов. (Конечная скорость распространения света была впервые определена Рёмером еще в 1676 году.) Планк назвал построенную таким образом теорию теорией относительности Эйнштейна.

Механика и оптика. Ньютон пытался построить как теорию движения массивных объектов (тел) – механику, так и теорию распространения света – оптику. Единую теорию вещества и света удалось построить только в XX веке. При этом оказалось, что свет – это тоже вещество, но его частицы безмассовы. Такой взгляд на частицы света – фотоны – до сих пор вызывает сопротивление многих физиков.

Размерности

Единицы, в которых $c = 1$. Максимальную скорость в природе обычно называют скоростью света и обозначают c. При рассмотрении формул теории относительности удобно пользоваться такой системой единиц, в которой c выбрана в качестве единицы скорости. Поскольку $c/c = 1$, в такой системе единиц во всех формулах следует положить $c = 1$, что очень упрощает их. Если измерять время в секундах, то в этой системе единиц расстояние следует измерять в световых секундах: одна световая секунда равна $3 \cdot 10^{10}$ см.

Пуанкаре и c. Один из создателей теории относительности Пуанкаре, говоря в 1904 году о том, что скорость света c входит во все уравнения электродинамики, сравнил ситуацию с геоцентрической теорией эпициклов Птолемея, в которой в соотношения между движениями небесных тел входил земной год. Он высказал надежду, что будущий Коперник избавит электродинамику от c. Но уже в следующем году Эйнштейн показал, что в теории относительности c должна играть ключевую роль.

Сравнение СИ и $c = 1$. В международной системе единиц – СИ – единица скорости 1 м/с навязывается соображениями удобства, стандартизацией производства и торговли, но не законами природы. В отличие от этого, c в качестве единицы скорости навязывается самой природой, когда мы хотим рассматривать фундаментальные процессы в ней.

Размерные множители. Рассмотрим некую физическую величину a. Обозначим $[a]$ размерность этой величины. Умножение a на любую степень мировой постоянной c несомненно меняет ее размерность, но не меняет ее физической сути. В дальнейшем я поясню, почему это так.

Скорость, импульс, энергия, масса. Обычно размерности импульса, массы и скорости частицы связаны соотношением $[p] = [m][v]$, а размерности энергии, массы и скорости – соотношением $[E] = [m][v^2]$. Введем безразмерную скорость $\vec{v}/c$ и, начиная с этого момента, будем именно ее обозначать $\vec{v}$. Аналогично, будем называть импульсом $\vec{p}$ то, что обычно обозначают $\vec{p}/c$. А энергией будем называть величину $e = E/c^2$. Очевидно, что после этого размерности p, e и m станут одинаковыми, и потому эти величины можно будет измерять в одних и тех же единицах, например – в граммах или в электронвольтах, как это принято в физике элементарных частиц.

О букве e для энергии. Выбор буквы e для обозначения энергии частицы может вызвать недовольство читателя, поскольку обычно этим символом обозначают электрон и элементарный электрический заряд. Однако к путанице такой выбор обычно не приводит, зато позволяет компактно записать формулы для одной частицы, напоминая, что эти формулы написаны в системе единиц, в которой $c = 1$. С другой стороны, как будет видно из дальнейшего, букву E удобно использовать для обозначения энергии двух или большего числа частиц.

О различии между энергией и частотой. Я только что настаивал на том, что $e = E/c^2$, так же, как и E, является энергией несмотря на то, что имеет размерность массы. Но тогда естественно спросить: а почему $\omega = E/\hbar$ является не энергией, а частотой? Ведь квант действия $\hbar$, подобно скорости света c, является мировой константой. Ответ на этот вопрос легко найти, если рассмотреть, как измеряют e и ω. Величины E и e измеряют одинаковым образом – скажем, с помощью калориметра. А частоту измеряют принципиально другим способом – например, с помощью часов. Поэтому равенство $\omega = E/\hbar$ говорит нам о связи между результатами двух различных типов измерений, в то время как равенство $e = E/c^2$ такой информации не содержит. Соображения, аналогичные высказанным относительно частоты, справедливы и для длины волны.

Одна частица

Относительные величины. Кинетическая энергия любого тела – величина относительная: она зависит от того, в какой системе отсчета ее измеряют. То же относится и к импульсу тела. В отличие от них масса тела – величина абсолютная: она характеризует тело само по себе, безотносительно к наблюдателю. Абсолютной величиной является и энергия покоя тела, поскольку в ней система отсчета раз и навсегда фиксирована.

Инвариантная масса. В теории относительности масса тела (частицы) определяется соотношением

$$m^2 = e^2 - p^2 . \qquad (1)$$

Заметьте, что энергия и импульс данного тела не ограничены сверху, а его масса фиксирована. Формула (1) является простейшим соотношением между энергией, импульсом и массой, которое можно написать «из головы». (Соотношение между e, $\vec{p}$ и m не может быть линейным, поскольку $\vec{p}$ – вектор, а e и m – скаляры в трехмерном пространстве.)

4-мерный импульс. Как впервые указал Минковский, теория относительности приобретает наиболее простой вид, если рассматривать ее в четырехмерном пространстве-времени. В теории относительности энергия и импульс тела образуют 4-мерный вектор энергии-импульса

$$p_i \; (i = 0, a), \text{ где } p_0 = e, \; p_a = \vec{p}, \; a = 1, 2, 3.$$

Масса является лоренцовым скаляром, характеризующим длину 4-мерного вектора, или просто 4-вектора, p_i:

$$m^2 = p_i^2 = e^2 - \vec{p}^2.$$

4-мерное пространство псевдоевклидово – отсюда знак «минус» в формуле для квадрата длины.

Другой способ понять знак «минус» – это ввести мнимый импульс $i\vec{p}$. Тогда

$$m^2 = e^2 + (i\vec{p})^2,$$

и мы имеем дело с теоремой Пифагора для такого псевдоевклидова прямоугольного треугольника, в котором гипотенуза m короче большего катета e.

Связь импульса и скорости. Импульс тела связан с его скоростью формулой

$$\vec{p} = e\vec{v}. \tag{2}$$

Эта формула самым простым образом удовлетворяет тому, что 3-мерный вектор (3-вектор) импульса должен быть пропорционален 3-мерному вектору скорости, а размерный коэффициент пропорциональности не должен обращаться в ноль для безмассового фотона.

В теории относительности сохранение так определенного импульса является следствием однородности 3-пространства, а сохранение энергии – следствием однородности времени (теорема Нётер).

Теорема Пифагора. Формула (1) представлена на рисунке 1 обычным евклидовым прямоугольным треугольником, в котором m и p – катеты, а e – гипотенуза.

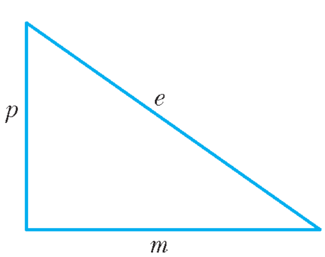


Рис. 1

Переход от $m \neq 0$ к $m = 0$. Формула (1) очевидным образом справедлива и при $m = 0$. А формула (2) справедлива и при $v = 1$. Отсюда следует, что существует плавный переход к безмассовым частицам от массивных, когда энергия последних намного превосходит их массу.

Физика от $p = 0$ до $p = e$. Рассмотрим формулы (1) и (2) сначала при импульсе, равном нулю, затем в пределе очень малых импульсов, когда $p \ll m$, потом в пределе очень больших импульсов, когда $p \sim e \gg m$, и, наконец, в случае безмассовых фотонов. Будем называть случай малых импульсов и скоростей ньютоновым, а случай очень больших импульсов и скоростей, близких к скорости света, – ультрарелятивистским. Но начнем с покоящегося тела.

Энергия покоя

Тело в покое. Если тело покоится, то его скорость и импульс равны нулю, а его энергия e по определению равна энергии покоя e_0. (Индекс 0 напоминает, что здесь мы имеем дело не вообще с энергией данного тела, а именно с его энергией тогда, когда оно покоится!) В этом случае из уравнения (1) следует, что

$$e_0 = m. \tag{3}$$

Горизонтальный «двуугольник». Таким образом, для покоящегося тела треугольник из рисунка 1 «схлопывается» в горизонтальный «двуугольник» (рис.2).

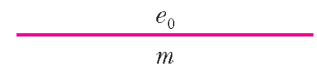


Рис. 2

Великое открытие Эйнштейна. В единицах, в которых $c \neq 1$, уравнение (3) имеет вид

$$E_0 = mc^2. \tag{4}$$

Осознание, что покоящееся тело обладает колоссальной энергией, было великим открытием Эйнштейна.

«Знаменитая формула». Очень часто (особенно в научно-популярной литературе) уравнение (4) записывают в виде «знаменитого уравнения Эйнштейна», опуская индекс 0:

$$E = mc^2. \tag{5}$$

Такое, на первый взгляд невинное, упрощение приводит к недопустимой путанице в понимании основ физики. В частности, к абсолютно ложному представлению о том, что, согласно теории относительности, масса тела эквивалентна его полной энергии и потому зависит от его скорости.

Вопрос не вкуса, а понимания. Часто можно услышать, что введение понятия массы, зависящей от скорости, это «вопрос вкуса». Разумеется, E/c^2 можно обозначить буквой m и даже назвать массой, хотя это не более разумно, чем обозначить E/c буквой p и назвать импульсом. Но это «переодевание» приводит к введению излишних понятий – релятивистской массы и массы покоя m_0 – и затрудняет понимание теории относительности. Вспоминается известная русская поговорка: «Назови хоть горшком, только в печь не ставь». Но люди, называющие E/c^2 массой, ставят этот горшок в печь преподавания физики.

Продольная и поперечная массы. В начале XX века наряду с релятивистской массой широко обсуждались поперечная и продольная массы: m_t и m_l. Последняя росла как $\left(e^3/m^3\right)m$ и «объясняла» с помощью ньютоновой формулы $F = ma$, почему массивное тело нельзя разогнать до скорости света. Потом о ней забыли, и такие популяризаторы теории относительности, как Хокинг, стали внушать своим читателям, что даже гораздо более медленный рост массы со скоростью ($(e/m)m$) якобы способен объяснить, почему скорость массивного тела не может достичь c.

Ложная интуиция. А.Н.Скринский рассказал мне о том, что представление о релятивистской массе мешало его собеседнику – известному физику – понять, как релятивистский электрон при столкновении может передать всю свою энергию покоящемуся электрону. Ну как может тяжелая бита отдать всю свою энергию пинг-понговому шарику? А ведь в физике, как и в обыденной жизни, мы очень часто опираемся на интуицию.

Механика Ньютона

Импульс в механике Ньютона. Ньютонова механика с высокой точностью описывает движение макроскопических тел в земных условиях, поскольку их скорости гораздо меньше скорости света. Так, скорость пули имеет порядок 1 км/с, что соответствует $v = 1/300000$, а $v^2 = 10^{-11}$. В этих условиях уравнение (2) сводится к такому:

$$\vec{p} = m\vec{v}\,. \tag{6}$$

Уравнение (1) в ньютоновом пределе в схематическом виде изображено на рисунке 3. Отрезок, изображающий p, на этом рисунке непропорционально велик. При соблюдении масштаба он должен был бы измеряться микронами.

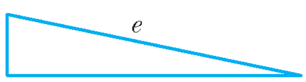


Рис. 3

Кинетическая энергия e_k. При малых скоростях разумно переписать формулу (1) в виде

$$e^2 - m^2 = p^2 \tag{7}$$

и представить ее так:

$$(e - m)(e + m) = p^2\,. \tag{8}$$

Это позволяет получить нерелятивистское выражение для кинетической энергии, не прибегая к обычному разложению в ряд квадратного корня. Учтем, что полная энергия e равна сумме энергии покоя e_0 и кинетической энергии e_k, и, следовательно,

$$e = m + e_k\,.$$

Энергия в механике Ньютона. В ньютоновом пределе $e_k \ll m$ (для летящей пули, например, $e_k/m = 10^{-11}$). Поэтому e с высокой точностью можно заменить на m в выражении (2) для импульса и в сомножителе $(e + m)$ в уравнении (8). Из последнего сразу же следует выражение для кинетической энергии e_k в механике Ньютона:

$$e_k = \frac{p^2}{2m} = \frac{mv^2}{2}\,. \tag{9}$$

Потенциальная энергия. Наряду с кинетической энергией, зависящей от скорости, в нерелятивистской механике тела, находящегося во внешнем силовом поле, важную роль играет потенциальная энергия, зависящая только от положения (координаты) тела. В стационарных процессах сумма кинетической и потенциальной энергий тела сохраняется. Потенциальная энергия определена с точностью до произвольной аддитивной постоянной, поскольку сила, действующая на тело, равна градиенту потенциальной энергии. Потенциальная энергия является сугубо нерелятивистским понятием.

Ньютон и современная физика. Гений Ньютона ознаменовал рождение современной науки. Ее достижения со времен Ньютона фантастичны. Современные взгляды на строение материи радикально отличаются от взглядов Ньютона. Тем не менее, даже в XXI веке многие учебники по физике продолжают использовать уравнения Ньютона при энергиях, на много порядков превосходящих пределы применимости механики Ньютона $e_k \ll e_0$.

Если некоторые профессора предпочитают настаивать на продолжении этой традиции с ее массой, зависящей от скорости, они должны по крайней мере познакомить своих студентов и с фундаментальными понятиями инвариантной массы и энергии покоя, и с истинным уравнением Эйнштейна $E_0 = mc^2$.

Ультрарелятивизм

Физика высоких энергий. Рассмотрим теперь другой предельный случай, когда $e/m \gg 1$. Именно такое соотношение между энергией и массой осуществляется в физике высоких энергий. Например, для электронов в коллайдере LEP (Large Electron-Positron Collider, коллайдер – ускоритель) в Европейском центре ядерных исследований (ЦЕРН) $e/m = 10^5$, поскольку $m = 0{,}5$ МэВ, а $e = 50$ ГэВ. И для протонов в большом адронном коллайдере LHC (Large Hadron Collider), который расположен в том же тоннеле, где раньше работал LEP, $e/m \sim 10^4$ (здесь $m \sim 938$ МэВ, а $e \sim 7$ ТэВ.)

Вертикальный треугольник. Треугольник для протона в LHC весьма схематически изображен на рисунке 4. Его основание на четыре порядка меньше гипотенузы.

Нейтрино. Еще более ультрарелятивистскими частицами являются нейтрино: их массы составляют доли электронвольта, а энергии – мегаэлектронвольты у нейтрино, летящих из Солнца и ядерных реакторов, и гигаэлектронвольты у нейтрино от распадов частиц, рождающихся в космических лучах и на ускорителях. При таких высоких энергиях нейтрино основание треугольника, схематически изображенного на рисунке 4, на много порядков меньше и его вертикального катета, и его гипотенузы.

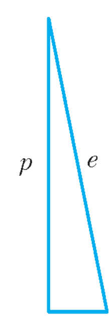


Рис. 4

Осцилляции нейтрино и $m^2/(2e)$. Из уравнения $(e - p)(e + p) = m^2$ сразу же следует, что

$$e - p \approx \frac{m^2}{2e}\,.$$

Различие масс трех нейтрино ν_1, ν_2, ν_3, обладающих определенными массами в вакууме, приводит к осцилляциям между нейтрино, не обладающими определенными массами, но представляющими три разных вида – электронное, мюонное, тауонное: ν_e, ν_μ, ν_τ. Данные по осцилляциям нейтрино дают

$$\Delta m_{21}^2 = (0{,}8 \pm 0{,}04) \cdot 10^{-4}\ \text{эВ}^2\,,$$

$$\Delta m_{32}^2 = (25 \pm 6) \cdot 10^{-4}\ \text{эВ}^2\,.$$

Фотон. Как известно, масса фотона равна нулю.

Поэтому для фотона $e = p$ и треугольник из рисунка 4 схлопывается в вертикальный «двуугольник», изображенный на рисунке 5.

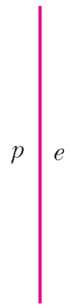


Рис. 5

Фотон и энергия покоя? Уместно заключить рассмотрение одночастичной механики вопросом: совместимо ли с понятием энергии покоя существование безмассового фотона? Легко видеть, что совместимо. Ведь чем меньше масса частицы, тем ближе к c скорость, с которой должен двигаться наблюдатель, чтобы перейти в систему ее покоя. Для очень легких одиноких частиц эта система отсчета становится не относящейся к сути обсуждаемого вопроса вместе с самим понятием e_0. Поэтому вполне согласуется с предельным переходом значение $e_0 = 0$ при $m = 0$.

Две свободные частицы

Столкновение двух частиц. В случае столкновения двух частиц сравнение системы покоя одной из них с системой покоя их общего центра инерции демонстрирует преимущества последней. Это обстоятельство используется в коллайдерах.

Зачем нужны коллайдеры? Если импульсы сталкивающихся частиц равны и противоположны, как, например, в коллайдерах LHC или LEP, то практически вся энергия сталкивающихся частиц идет на рождение новых частиц.

Масса системы частиц. Как известно, для изолированной системы частиц полная энергия E и полный импульс $\vec{P}$ сохраняются. Так как энергия и импульс аддитивны, то для двух свободных частиц

$$E = e_1 + e_2 , \qquad (10)$$

$$\vec{P} = \vec{p}_1 + \vec{p}_2 . \qquad (11)$$

Определим величину M следующей формулой:

$$M^2 = E^2 - \vec{P}^2 . \qquad (12)$$

Массы аддитивны при $v = 0$. Уравнение (12) инвариантно относительно лоренцевых преобразований, как и уравнение (1). Поэтому величину M естественно назвать массой системы двух частиц. В статическом пределе, когда p_1 и p_2 равны нулю, из уравнения (12) следует, что

$$M = e_{01} + e_{02} = m_1 + m_2 . \qquad (13)$$

Таким образом, в ньютоновом пределе M действительно равно сумме масс двух частиц, т.е. массы аддитивны.

Массы не аддитивны при $v \neq 0$. Однако при больших скоростях связь между M и массами m_1 и m_2 отсутствует. Например, в коллайдерах ЦЕРН величина M на четыре порядка превышает массу электронов или протонов. Величина M кардинальным образом зависит от относительного направления импульсов двух частиц, поскольку сумма двух векторов зависит от угла между ними. Так, для двух фотонов, летящих в одном направлении,

$$P = p_1 + p_2 . \qquad (14)$$

Коллинеарные фотоны. Для фотонов $p_1 = e_1$ и $p_2 = e_2$. И потому для пары фотонов, летящих в одном направлении,

$$P = p_1 + p_2 = e_1 + e_2 = E . \qquad (15)$$

Тогда из уравнения (12) следует, что в этом случае масса пары фотонов $M = 0$. А это значит, что масса «игольного» пучка света равна нулю.

А если фотоны разлетаются? Однако если фотоны летят в противоположные стороны с одинаковыми энергиями, то $\vec{p}_1 = -\vec{p}_2$, а $\vec{P} = 0$. В этом случае энергия покоя системы двух фотонов просто равна сумме энергий этих фотонов, а масса этой системы равна

$$M = E_0 = 2e . \qquad (16)$$

Шок. Разумеется, утверждение о том, что пара двух безмассовых частиц имеет огромную массу, может вызвать шок у неподготовленного читателя. Какой смысл имеет понятие энергии покоя для двух фотонов, каждому из которых «покой лишь только снится»? Что покоится в этом случае?

Ответ очевиден. Действительно, очевиден: покоится центр инерции двух фотонов. Но если для одной покоящейся частицы энергия покоя это энергия, скрытая в ее массе, то для двух фотонов это просто сумма их энергий (кинетических!) в системе отсчета, в которой их импульсы равны по величине и противоположны по направлению. Скрытой энергии в этом случае нет!

Что значит слово «сохраняется»? Когда мы говорим, что энергия сохраняется, мы имеем в виду, что сумма энергий частиц, вступающих в реакцию, равна сумме энергий частиц, возникших в результате реакции. Аналогичный смысл имеет утверждение о сохранении импульса. Сохраняются величины $E = \sum e_i$ и $\vec{P} = \sum \vec{p}_i$. А следовательно, сохраняется и масса M системы (совокупности) частиц, определяемая формулой $M^2 = E^2 - \vec{P}^2$. Но, в отличие от энергии и импульса, масса не аддитивна: $M \neq \sum m_i$. Некоторые авторы говорят о неаддитивности массы как о ее несохранении. Но для отдельных частиц, участвующих в реакции, вообще говоря, не сохраняются не только их массы, но и энергии, и импульсы, да и они сами. Поэтому говорить о несохранении массы неправильно.

Мысленный опыт Эйнштейна. Конечно, понятие массы системы двух разлетающихся фотонов весьма непривычно. Но именно с его помощью Эйнштейн в 1905 году открыл энергию покоя массивного тела. Он заметил, что излучив «два количества света» в противоположные стороны, покоящееся тело остается в покое, но его масса в этом мысленном эксперименте уменьшается.

Аннигиляция позитрония. Как известно, nihil по-латыни значит *ничто*. При аннигиляции позитрония (связанной водородоподобной системы, состоящей из электрона и позитрона)

$$e^+e^- \to \gamma\gamma$$

величина M сохраняется, поскольку сохраняются E и

$\vec{P}$. В начальном состоянии M равно сумме масс электрона и позитрона, а в конечном – сумме энергий двух фотонов в системе покоя позитрония. Таким образом, энергия покоя электрона и позитрона полностью переходит в кинетическую энергию фотонов, но массы начального и конечного состояний в этом процессе одинаковы, как того требует сохранение полной энергии и полного импульса.

Распады мезонов. Аналогичным образом при распаде K-мезона на два π-мезона или на три π-мезона энергия покоя каона переходит в сумму полных энергий пионов, каждая из которых имеет вид $e = e_k + m$. Но масса системы двух или трех пионов, возникших при распаде каона, равна массе каона.

Что называть материей? Во всех распадах энергия покоя переходит в энергию движения, а полная энергия изолированной системы сохраняется. Сохраняется и масса системы, но не массы отдельных ее частиц. Массивные частицы переходят в менее массивные, а то и вовсе в безмассовые. Разумно называть частицами материи не только такие массивные частицы, как протоны и электроны, но и очень легкие нейтрино, и безмассовые фотоны, и даже гравитоны. В современной квантовой теории поля все они рассматриваются на равных основаниях.

Чеширский кот? В реакциях распада и аннигиляции материя не исчезает подобно чеширскому коту, от которого остается только его улыбка – энергия. Носителями энергии здесь являются частицы материи. Во всех изученных до сих пор процессах энергии без материи («чистой энергии») не бывает.

Это, правда, не относится к так называемой темной энергии, которая была открыта в последние годы XX века. Темная энергия проявляется в ускоряющемся расширении Вселенной. (Об этом ускоряющемся расширении свидетельствуют скорости далеких сверхновых звезд.) Похоже, что носителем темной энергии является вакуум и что она составляет три четверти всей энергии Вселенной. Оставшаяся четверть заключена в массе обычной материи (5%) и темной материи (20%). В лабораторных опытах с обычным веществом темная энергия не проявляется.

Распад в среде. Если распад (или реакция)происходит в среде, то в результате распада система перестает быть изолированной из-за взаимодействия продуктов распада с частицами среды. При этом среда будет нагреваться, а ее энтропия – расти.

Несвободные частицы

Тела и частицы. Все физические тела состоят из элементарных частиц. Такие элементарные частицы, как протон и нейтрон, сами состоят из еще более элементарных частиц – кварков и глюонов. А электрон и нейтрино на современном уровне познания выглядят как истинно элементарные.

Масса газа. Во всех рассмотренных выше случаях, когда масса системы частиц была больше суммы их масс, эти частицы свободно разлетались. Обратимся теперь к ситуации, когда они разлететься не могут. Такая ситуация осуществляется, например, в часто обсуждаемом мысленном опыте с газом молекул или фотонов в замкнутом покоящемся сосуде. Полный импульс такого газа равен нулю, поскольку газ изотропен: $\vec{P} = \sum \vec{p}_i = 0$. Поэтому полная масса M такого газа равна его полной энергии E и равна сумме энергий отдельных частиц:

$$M = E = \sum e_i .$$

Масса нагретого газа. При нагревании газа полный импульс не меняется, а полная энергия увеличивается, так как растет кинетическая энергия отдельных частиц. В результате масса газа в целом растет, в то время как масса каждой отдельной частицы остается неизменной. (В литературе можно встретить неверное утверждение, что при увеличении кинетической энергии частиц (фотонов) растет их масса.)

Масса горячего утюга. Таким же образом, как масса газа, должна расти при нагревании и масса утюга, хотя массы колеблющихся атомов остаются неизменными. Однако совокупность формул (10) – (12) для системы свободных частиц здесь неприменима, поскольку частицы, в этом случае атомы, не свободны, а связаны в кристаллической решетке металла. Разумеется, увеличение массы утюга слишком мало, чтобы его можно было реально измерить.

Масса атомного ядра. Аналогичным образом связаны нуклоны в атомных ядрах. Но здесь эффект настолько велик, что от него зависит судьба жизни на Земле. Начнем с ядра тяжелого водорода. Масса дейтрона меньше суммы масс протона и нейтрона, составляющих его. Энергия связи нуклонов в дейтроне составляет 2,2 МэВ. Чтобы развалить дейтрон на нуклоны, надо затратить энергию, равную энергии связи или превышающую ее. Атомные ядра остальных элементов таблицы Менделеева тоже существуют благодаря энергии связи нуклонов в этих ядрах.

О формулах (10) – (12). Поясним, почему для таких несвободных частиц, как электроны в атомах или нуклоны в атомных ядрах, неприменимы формулы (10) – (12). Прежде всего потому, что эти частицы не имеют определенных импульсов, в силу соотношения неопределенностей. Чем меньше радиус сферы, в которой заключены частицы, тем больше неопределенность их импульсов.

Соотношение неопределенностей. И для атомов, и для ядер очень важны закономерности квантовой механики, и в частности – соотношение неопределенностей. Как известно, произведение неопределенностей импульса Δp и координаты Δx должно быть больше, чем квант действия $\hbar$. Поэтому частицы в атомах не имеют определенных индивидуальных импульсов, а имеют только определенный суммарный импульс.

Энергия поля. Другой причиной, по которой в атомах неприменимы формулы (10) – (12), является то, что пространство между отдельными частицами в них по существу не является пустым: оно заполнено материальной средой – физическими полями. Пространство внутри атома заполнено электромагнитным полем, а внутри ядра – гораздо более плотным и сильным полем, которое часто называют мезонным.

Энергия связи. Наличие энергии поля приводит к тому, что для двух тесно взаимодействующих частиц, скажем, в атомном ядре тяжелого водорода – дейтрона – в формуле (10) $E = e_1 + e_2$ необходимо учитывать энергию поля. В результате получается $M < m_1 + m_2$. Величину $m_1 + m_2 - M$ называют энергией связи.

Слияние и деление ядер. Энергия связи в начале таблицы Менделеева максимальна у ядра гелия, а в середине – у ядра железа. Именно поэтому при образовании гелия из водорода в реакциях слияния на Солнце и в водородных бомбах выделяется большое количество кинетической энергии. А в ядерных реакторах и атомных бомбах эта энергия образуется в реакциях деления при развале тяжелых ядер урана и плутония на более легкие ядра в середине таблицы Менделеева.

Химические реакции. Существенно меньшая энергия, измеряемая единицами электронвольт, выделяется в химических реакциях за счет различия энергий связи электронов в разнообразных химических соединениях. Однако источником кинетической энергии и в химических, и в ядерных реакциях является разность масс начальных и конечных частиц, участвующих в этих реакциях.

Закон Кулона. Энергия связи электронов в атомах много меньше массы электрона. Поэтому понятие энергии связи в атомах можно пояснить с помощью нерелятивистского понятия потенциальной энергии. Энергия связи равна сумме положительной кинетической энергии связанной частицы и ее отрицательной потенциальной энергии. Потенциальная энергия электрона, скажем, в атоме водорода определяется законом Кулона:

$$U = -\frac{\alpha}{r}, \tag{17}$$

где $\alpha = e^2/(\hbar c) = 1/137$, а e – заряд электрона.

Еще о потенциальной энергии. Понятие потенциальной энергии определено только в ньютоновом пределе. Если одна из двух взаимодействующих частиц или обе они существенно релятивистские, то понятие потенциальной энергии неприменимо.

Электромагнитное поле. В теории относительности кулоново поле является 0-компонентой 4-потенциала электромагнитного поля A_i $(i = 0,1,2,3)$. Источником поля является электромагнитный ток, пропорциональный 4-вектору ep_i, где p_i – 4-импульс частицы с зарядом e. Взаимодействие двух движущихся зарядов осуществляется путем распространения поля от одного заряда к другому. Оно описывается так называемой функцией Грина, или пропагатором электромагнитного поля.

Гравитация

Гравитационные орбиты. На различных эмблемах часто изображают орбиты электронов в атомах, подобные орбитам планет. Из только что сказанного должно быть ясно, что таких орбит в атомах нет – это запрещено законами квантовой механики. Для макроскопических же тел и уж, конечно, для таких тяжелых, как планеты, квантовые эффекты ничтожны, поэтому их орбиты прекрасно описываются классической механикой.

Константа Ньютона. Потенциальная энергия Земли в гравитационном поле Солнца определяется законом Ньютона:

$$U = -\frac{GMm}{r}, \tag{18}$$

где M – масса Солнца, m – масса Земли, r – расстояние между их центрами, а G – константа Ньютона (гравитационная постоянная):

$$G = 6{,}71 \cdot 10^{-39}\,\hbar c \ \left(\text{ГэВ}/c^2\right)^{-2}. \tag{19}$$

(Здесь использованы единицы, в которых $c \neq 1$.)

Тензор энергии-импульса частицы. В ньютоновой физике источником гравитации является масса. В теории относительности источником гравитации является тензор энергии-импульса $p_i p_k$, который, будучи деленным на энергию e, служит как бы «гравитационным зарядом». (Напомню, что p_i – 4-вектор энергии-импульса, $i = 0, 1, 2, 3$.)

Распространение поля от источника до «стока» описывается функцией Грина, или пропагатором гравитационного поля. Пропагатор гравитационного поля пропорционален $g^{il}g^{km} + g^{im}g^{kl} - g^{ik}g^{lm}$, где g^{ik} – метрический тензор.

Масса Планка. В физике частиц используется понятие массы Планка:

$$m_P = \sqrt{\frac{\hbar c}{G}}. \tag{20}$$

В единицах, в которых c, $\hbar = 1$, масса Планка равна $m_P = 1/\sqrt{G} = 1{,}22 \cdot 10^{19}$ ГэВ.

Гравитационное взаимодействие двух ультрарелятивистских частиц растет как квадрат их энергии E в системе их центра инерции. Оно становится предельно сильным при $E \sim m_P$ и при расстоянии между частицами $r \sim 1/m_P$. Но вернемся от этих фантастически больших энергий и малых расстояний к яблокам и фотонам в гравитационных полях Земли и Солнца.

Яблоко и фотон. Рассмотрим частицу, находящуюся в статическом гравитационном поле, для которого в силу статичности l, $m = 0$. В этом случае пропагатор гравитационного поля пропорционален $2g^{i0}g^{k0} - g^{ik}g^{00}$, а тензор энергии-импульса частицы, умноженный на пропагатор, сводится к простому выражению $p_i p_k \Rightarrow \left(2e^2 - m^2\right)$. Следовательно, для нерелятивистского яблока с массой m «гравитационный заряд» равен m, а для фотона с энергией e он равен $2e$. Обратите внимание на множитель 2. Кинетическая энергия «притягивается» сильнее, чем скрытая энергия, заключенная в массе.

Фотон в поле Солнца. Взаимодействие фотона с гравитационным полем должно вызывать отклонение луча света далекой звезды вблизи солнечного диска. В 1915 году Эйнштейн вычислил этот угол отклонения и показал, что он должен составлять $4GM/\left(c^2R\right) \approx 1{,}75''$. (Здесь M и R – масса и радиус Солнца.) Это предсказание было подтверждено во время солнечного затмения в 1919 году, что вызвало огромный интерес к теории относительности.

Атом в поле Земли. При подъеме тела над Землей его потенциальная энергия растет пропорционально его массе. Поэтому разность энергий двух уровней атомного ядра должна быть тем больше, чем выше этаж дома, на котором ядро находится.

Энергия фотона сохраняется. С другой стороны, частота ω фотона, летящего в статическом гравитационном поле, а следовательно, и его полная энергия $e = \hbar\omega$ не должны меняться.

В результате фотон, испущенный на нижнем этаже дома при переходе между двумя уровнями ядра, не сможет вызвать обратного перехода в том же ядре на верхнем этаже. Это теоретическое предсказание было подтверждено в 1960-х годах после открытия эффекта Мёссбауэра, позволяющего измерять ничтожно малые смещения ядерных уровней.

Но длина волны меняется. Фотон, летящий в статическом гравитационном поле, подобно летящему камню, сохраняет свою полную энергию e и частоту ω. Но его импульс, а следовательно, и его длина волны меняются с изменением расстоянии до гравитирующего тела.

Показатель преломления. При удалении фотона от источника гравитационного поля его скорость возрастает, стремясь к c, а при приближении к источнику – падает. Гравитационное поле, как прозрачная среда, имеет коэффициент преломления. Это наглядно объясняет отклонение света в поле Солнца и в гравитационных линзах галактик. Уменьшение скорости фотона вблизи Солнца было обнаружено в опытах по измерению запаздывания радарного эха от планет.

Часы или секундомеры. Пусть на первом этаже рядом находятся синхронно идущие часы A и B. Если поднять на второй этаж сначала часы A, а, скажем, назавтра – B, то часы A будут впереди B. Но в качестве синхронизованных секундомеров они будут одинаково хороши.

Когда каждую точку в пространстве снабжают своими собственными часами, то по существу говорят не о часах, а о секундомерах. Если теперь с помощью локальных секундомеров измерять частоту света, то она была бы тем меньше, чем выше этаж.

Эпистемика и лингвистика

Эпистемика. Эпистемика или, как принято было говорить раньше, эпистемология – это теория знания и познания. Очевидно, что обсуждаемые в этой статье вопросы относятся не только к физике, но и к эпистемике.

Физика и семантика. Но какое отношение к физике имеет наука о языке – лингвистика и ее раздел семантика – наука о словах и символах?

Тут уместно вспомнить высказывание Фока, упомянутое в начале доклада.

«Склееные понятия». В «Началах» Ньютона «склеены» понятия массы и материи (вещества): «масса пропорциональна плотности и объему». В статьях Эйнштейна масса «склеена» с инерцией и гравитацией (инертная и гравитационная массы). А энергия «склеена» с материей.

Архетип. Согласно словарям, архетип – это исторически исходная форма (протоформа), исходное понятие, слово или исходный тип (прототип). Понятием архетипа очень интересовался Паули, опубликовавший в 1952 году статью о влиянии архетипических представлений на создание естественно-научных теорий Кеплером. Возможно, понятие массы является именно таким архетипическим понятием, которое «склеило» в себе понятия вещества, инерции и тяжести.

Атом и архетип. «Atom and Archetype» – под этим заглавием была переведена с немецкого на английский язык книга, содержащая переписку Вольфганга Паули и ведущего немецкого психоаналитика Карла Юнга, длившаяся с 1932 по 1958 год. Они обсуждали, в частности, материальную природу времени и возможность общаться с людьми, жившими века и тысячелетия тому назад.

Широко известно, что Паули серьезно относился к эффекту, носящему его имя: когда он входил в экспериментальную лабораторию, приборы в ней ломались.

Поэты о терминах. Д.Самойлов о словах: «Их протирают, как стекло. И в этом наше ремесло».

В.Маяковский: «Улица корчится безъязыкая. Ей нечем кричать и разговаривать».

Нехватка точных терминов и неумение пользоваться ими восполняется многими авторами обильным использованием слов-паразитов типа «масса покоя», придающих текстам гладкость и «энергетику», подобно тому как в обычной современной речи это делают слова типа «блин».

Как преподавать физику. Надо «протирать» и «расклеивать» термины. «Пуповина», соединявшая современную физическую теорию с предшествующей «материнской» теорией, должна бережно обрезаться при преподавании. (В случае теории относительности «матерью» был «кентавр», состоявший из теории поля Максвелла и механики Ньютона, а «пуповина» – релятивистская масса.)

Вспомним название знаменитой книги Ф.Клейна «Элементарная математика с высшей точки зрения». На ландшафт современной физики надо смотреть не из оврага истории, а с вершины принципов симметрии.

Мне представляется, что недопустимо выдавать зависимость массы от скорости за экспериментальный факт, скрывая от студента, что она является интерпретационным «фактоидом». (Согласно словарям, фактоид похож на факт, но считается достоверным только потому, что встречается в печатных текстах.)

Заключение

О заглавии. Мой хороший друг и большой эксперт по теории относительности, прочитав слайды этого доклада, посоветовал мне укоротить его название, выбросив вторую половину. Если бы мне пришлось сокращать название, я бы скорей выбросил первую половину: ведь рассмотрения прямоугольных треугольников без извлечения квадратного корня я в литературе по теории относительности не встречал.

Комплексные числа

С.ДОРИЧЕНКО

Комплексные числа появились впервые в XVI веке при решении кубических уравнений. Чтобы найти действительные решения кубического уравнения, иногда приходилось в промежуточных вычислениях оперировать «несуществующими» квадратными корнями из отрицательных чисел. Но в итоге получался верный ответ. Удивительно и непонятно.

Постепенно обнаруживалось, что комплексные числа могут быть очень полезны в самых разных областях математики: алгебре, геометрии, теории чисел, анализе.

Тем не менее, комплексные числа очень долго вызывали недоверие математиков, считались таинственными и сомнительными для использования величинами. Тут есть некоторая аналогия с отрицательными числами – когда-то давно люди их не знали. Сейчас же мы легко пользуемся отрицательными числами, да еще применяя «странные» правила вроде «минус на минус дает плюс». Почему? Наивное объяснение состоит в том, что отрицательные числа с такими правилами помогают правильно решать уравнения. Если верно составлено уравнение (например, в задаче на движение), можно в этом уравнении раскрывать (по правилам) скобки, можно переносить числа (или переменные) из одной части в другую, меняя при этом знак на противоположный, – такие преобразования не приведут к ошибке. Это удобно: действуя таким образом, мы можем упрощать уравнение, не задумываясь в каждый момент, какой смысл имеют выражения справа и слева от знака равенства. Например, в Европе отрицательные числа так и входили в обращение в эпоху Возрождения – постепенно, по мере того как люди стали оперировать с буквенными выражениями. Точно так же и к комплексным числам математики привыкали очень медленно.

Благодаря комплексным числам были решены сложные задачи, в формулировках которых участвовали только действительные числа, появились новые интересные задачи и теории. Поэтому сейчас комплексные числа прочно вошли в математику и физику, где имеют множество применений. Название «чи́сла» оправдано потому, что у них много общего с действительными числами: их можно складывать, вычитать, умножать, делить одно на другое (как обычно, нельзя делить только на ноль). Но есть и отличия: например, из любого комплексного числа можно извлечь квадратный корень.

Мы расскажем о том, что такое комплексные числа и как они помогают решать самые разные задачи. Не пытайтесь сразу все усвоить – прочитайте определения, решите несколько задач, и постепенно вы все поймете.

Определения и основные свойства

Комплексное число – это выражение вида $a + bi$, где a и b – действительные числа, а i – так называемая мнимая единица, специальный символ, квадрат которого равен минус единице: $i^2 = -1$.

У комплексного числа $z = a + bi$ число a называют его *действительной* (или *вещественной*) *частью* (пишется $a = \operatorname{Re} z$), а число b – его *мнимой частью* (пишется $b = \operatorname{Im} z$). Два комплексных числа считаются равными, если равны их действительные части и равны их мнимые части. Если мнимая часть равна нулю, т.е. число имеет вид $a + 0i$, пишут просто a – мы получаем обычное действительное число.

Комплексные числа складывают, вычитают и умножают, раскрывая скобки и приводя подобные (заменяя при этом i^2 на -1). Найдем, например, сумму комплексных чисел $5 - 4i$ и $2 + 3i$:

$$(5 - 4i) + (2 + 3i) = (5 + 2) + (-4 + 3)i = 7 - i.$$

Найдем произведение этих чисел:

$$(5 - 4i) \cdot (2 + 3i) = 5 \cdot 2 + 5 \cdot 3i - 4 \cdot 2i - 4 \cdot 3 \cdot i^2 = 22 + 7i.$$

Упражнение 1. Найдите действительную и мнимую части суммы, разности и произведения комплексных чисел $a + bi$ и $c + di$.

Число $a - bi$ называется *комплексно-сопряженным* числу $z = a + bi$ и обозначается $\bar{z}$. Произведение $z\bar{z}$ равно $a^2 + b^2$, т.е. всегда является действительным числом. Это замечание поможет нам научиться делить комплексные числа друг на друга. Вычислим, например, $\frac{5 - 4i}{2 + 3i}$. Сделаем небольшой алгебраический трюк: домножим числитель и знаменатель на число $2 - 3i$ (сопряженное к знаменателю). В числителе получим $-2 - 23i$, а в знаменателе получим $4 - 9i^2$, т.е. 13 – действительное число. Значит, наша дробь равна $-\frac{2}{13} - \frac{23}{13}i$.

Упражнение 2. Найдите действительную и мнимую части отношения комплексных чисел $a + bi$ и $c + di$ (где $c + di \neq 0$).

У комплексных чисел есть наглядная геометрическая интерпретация.

Возьмем плоскость с декартовой системой координат. Пусть v_1 – вектор с координатами $(1;\ 0)$, а v_2 – вектор с координатами $(0;\ 1)$. Как вы знаете, любой вектор на плоскости выражается единственным способом через векторы v_1 и v_2 (в виде суммы этих векторов, взятых с числовыми множителями). А именно, вектор с координатами $(a;\ b)$ равен сумме векторов с координатами $(a;\ 0)$ и $(0;\ b)$, т.е. равен сумме

$av_1 + bv_2$. Обозначим вектор v_2 символом i, а вектор v_1 – единицей (которую писать не будем), тогда запись $av_1 + bv_2$ как раз превратится в $a + bi$. Фактически мы просто сопоставляем каждому комплексному числу $z = a + bi$ вектор с координатами $(a; b)$ (рис.1). Длина этого вектора (т.е. число $\sqrt{a^2+b^2}$) называется *модулем* комплексного числа z и обозначается $|z|$. Пусть $z \neq 0$. Угол, отсчитанный против часовой стрелки от вектора с координатами (1; 0) до вектора с координатами $(a; b)$, называется *аргументом* комплексного числа z и обозначается $\operatorname{Arg} z$. (Углы измеряем в радианах.) Аргумент комплексного числа определен с точностью до прибавления числа вида $2\pi n$, где n – целое. Например, для числа i аргументами будут числа $\pi/2$, $\pi/2 + 2\pi$, $\pi/2 - 2\pi$ и т.д. Мы говорим, что аргументы равны, если они отличаются на число вида $2\pi n$, где n – целое.

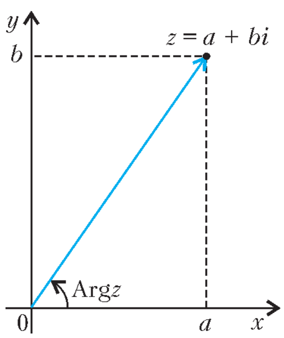


Рис. 1

Упражнения

3. Найдите модуль и аргумент чисел -4, i, $1+i$, $1-i\sqrt{3}$, $\sin\alpha + i\cos\alpha$.

4. Выразите модуль и аргумент числа $\bar{z}$ через модуль и аргумент числа z.

Если $z = a + bi$, длина вектора z равна r, а угол наклона к оси абсцисс равен φ, то координаты вектора есть $a = r\cos\varphi$ и $b = r\sin\varphi$ (рис.2). Поэтому для любого ненулевого комплексного числа z имеет место равенство $z = r(\cos\varphi + i\sin\varphi)$, где $r = |z|$, $\varphi = \operatorname{Arg} z$. Это так называемая *тригонометрическая форма записи комплексного числа*. Например, $1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$.

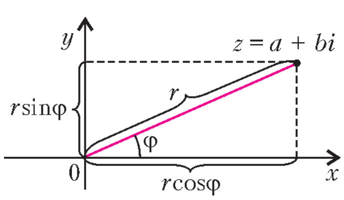


Рис. 2

Итак, можно представлять себе комплексное число как вектор, координаты которого – действительная и мнимая части этого числа. А можно сопоставить комплексному числу конец этого вектора и считать комплексное число точкой на плоскости. Поэтому мы будем называть комплексным числом (и обозначать одной и той же буквой) и выражение $a + bi$, и вектор с координатами $(a; b)$, и точку плоскости с координатами $(a; b)$.

При этом замечательно, что сложение комплексных чисел и сложение векторов (по правилу параллелограмма) – это одно и то же (что сразу следует из определения).

Упражнение 5. Докажите, что $|z + w| \le |z| + |w|$ при любых комплексных числах z и w.

А каков геометрический смысл умножения? Чтобы ответить на этот вопрос, отметим одно очень важное свойство комплексных чисел, которое не раз пригодится в дальнейшем, – они тесно связаны с преобразованиями плоскости. Например, чтобы получить число, комплексно-сопряженное числу z, надо просто отразить вектор, соответствующий z, относительно оси абсцисс (рис.3). Таким образом, перевод числа в комплексно-сопряженное отвечает симметрии плоскости относительно оси абсцисс.

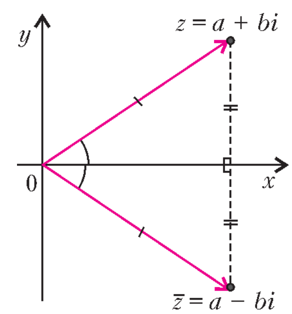


Рис. 3

Что будет, если к каждой точке плоскости мы прибавим комплексное число z? Все точки сдвинутся на этот вектор, т.е. мы получим сдвиг плоскости в направлении нашего вектора на его длину.

Умножение на ненулевое действительное число – это просто гомотетия. А какое получится преобразование плоскости, если мы каждую ее точку умножим на ненулевое комплексное число z?

Разберем сначала случай, когда модуль числа z равен 1, т.е. z можно записать в виде $\cos\varphi + i\sin\varphi$. Посмотрим, куда перейдут (при умножении на z) точки на оси абсцисс. Любое число a на оси абсцисс действительное, и оно перейдет в число $a\cos\varphi + i(a\sin\varphi)$, т.е. просто повернется на угол φ относительно начала координат (рис.4). А куда перейдут точки на оси ординат? Любая такая точка имеет вид bi, где число b действительное, и значит, перейдет в число $-b\sin\varphi + i(b\cos\varphi)$, т.е. тоже повернется на угол φ (рис.4). Докажем, что и любой вектор плоскости при умножении на z повернется на угол φ. Действительно, пусть нам дано число $w = a + bi$. Тогда $z \cdot w = z \cdot a + z \cdot (bi)$. Мы уже знаем, что вектор $z \cdot a$ – это вектор a, повернутый на угол φ, и вектор $z \cdot (bi)$ – это вектор bi, повернутый на угол φ, но в таком случае и сумма $z \cdot a + z \cdot (bi)$ – это сумма векторов a и bi, повернутая на угол φ (рис.5). Что и требовалось доказать.

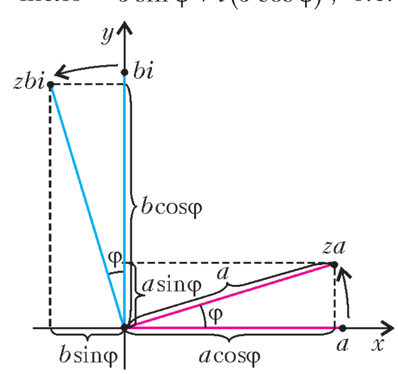


Рис. 4

Если же умножить вектор w на ненулевой вектор z не обязательно единичной длины, то w, очевидно, тоже повернется на угол $\operatorname{Arg} z$, и еще его длина умножится на $|z|$ (кстати, в геометрии такое преобразование называется поворотной гомотетией). Значит, мы доказали такую замечательную теорему.

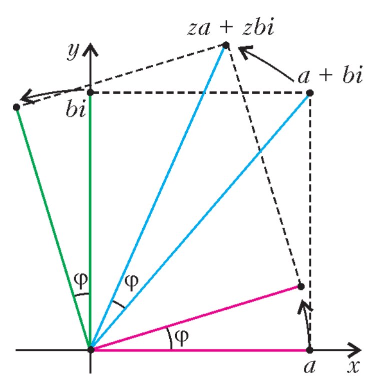


Рис. 5

Теорема. *При перемножении ненулевых комплексных чисел их модули перемножаются, а аргументы складываются:* $|zw| = |z| \cdot |w|$, $\operatorname{Arg} zw = \operatorname{Arg} z + \operatorname{Arg} w$.

Приведем одно следствие из этой теоремы. Возьмем два числа $\cos\alpha + i\sin\alpha$ и $\cos\beta + i\sin\beta$ и перемножим их. По только что доказанной теореме мы должны получить число $\cos(\alpha+\beta) + i\sin(\alpha+\beta)$. С другой стороны, если мы просто перемножим эти два числа, раскрывая скобки и приводя подобные, мы получим число

$$\cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\cos\alpha\sin\beta + \sin\alpha\cos\beta).$$

Но ведь должно получиться то же самое! Значит, мы доказали тригонометрические формулы:

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta,$$
$$\sin(\alpha+\beta) = \cos\alpha\sin\beta + \sin\alpha\cos\beta.$$

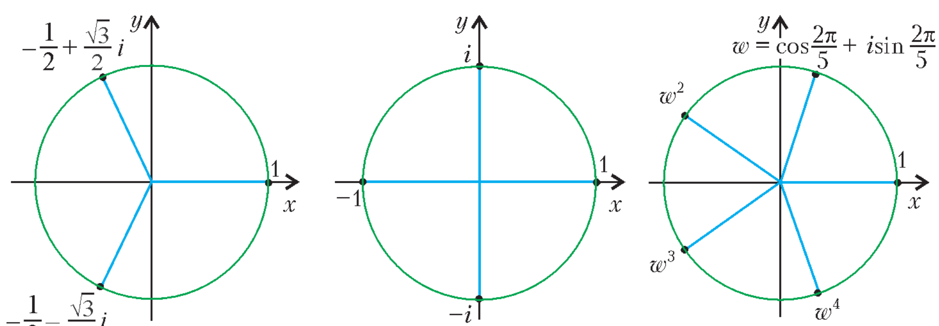
Упражнения

6. Пусть $z = r(\cos\varphi + i\sin\varphi)$. Докажите, что для любого натурального n выполнено равенство (формула Муавра) $z^n = r^n(\cos n\varphi + i\sin n\varphi)$.

7. Вычислите $\left(1+i\sqrt{3}\right)^{150}$ и $\dfrac{\sqrt{3}+i}{(1-i)^{30}}$.

8. Выразите $\cos nx$ и $\sin nx$ через $\cos x$ и $\sin x$ для $n = 2$, $n = 3$, $n = 4$, …

Заметим, что отношение $\dfrac{z}{w}$ комплексных чисел z и w – это, по определению, такое число, которое при умножении на w дает z. Тогда, по доказанной теореме, модуль этого числа равен $\dfrac{|z|}{|w|}$, а аргумент равен $\operatorname{Arg} z - \operatorname{Arg} w$ и равен углу между векторами z и w (с точностью до прибавления чисел вида $2\pi n$, где n – целое).



Корни уравнения $z^3 = 1$ Корни уравнения $z^4 = 1$ Корни уравнения $z^5 = 1$

Рис. 7

Корни из комплексных чисел

Теперь давайте научимся извлекать корни из комплексных чисел. *Корнем степени n* из комплексного числа z называется любое такое комплексное число w, что $w^n = z$. Корень любой степени из нуля равен нулю, поэтому рассмотрим ненулевое комплексное число z с модулем r и аргументом φ. Найдем все квадратные корни из z, т.е. все такие комплексные числа w, что $w^2 = z$. Ясно, что модуль числа w должен равняться корню (арифметическому) из модуля z (т.е. $\sqrt{r}$). Аргумент же числа w, умноженный на 2, должен дать φ (или отличаться от φ на число вида $2\pi n$, где n – целое). Нетрудно убедиться, что таких w ровно два: одно с аргументом $\varphi/2$ и другое с аргументом $\varphi/2 + \pi$. Эти числа противоположны (рис. 6). Например, уравнение $w^2 = -1$ имеет ровно два решения: i и $-i$.

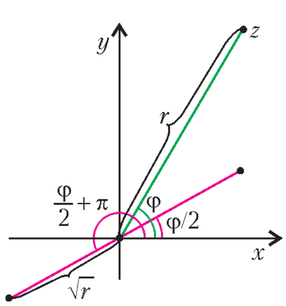


Рис. 6

Объяснить, что квадратных корней из ненулевого числа z не может быть больше двух, можно и алгебраически: если a – один из корней, то для любого другого корня w будет выполнено $z = w^2 = a^2$, откуда $(w-a)(w+a) = 0$, и, значит, либо $w = a$, либо $w = -a$.

Чтобы извлечь квадратный корень из комплексного числа $z = a + bi$, на практике иногда удобнее действовать алгебраически: ищем w такое, что $w^2 = z$, в виде $w = c + di$. Подставляем в уравнение, получаем $(c+di)^2 = a + bi$, раскрываем скобки и решаем систему.

Теперь мы можем решить любое квадратное уравнение (даже с комплексными коэффициентами) – оно решается по обычной школьной формуле, только надо уметь извлекать корень из отрицательного (или любого комплексного) числа.

Упражнение 9. Решите уравнения:
а) $z^2 = i$; б) $z^2 = 5 - 12i$; в) $z^2 + (2i - 7)z + 13 - i = 0$.

Аналогично можно извлекать корни бóльших степеней. На рисунке 7 изображены корни уравнений $z^3 = 1$, $z^4 = 1$, $z^5 = 1$ – они образуют правильные многоугольники, вписанные в окружность единичного радиуса с центром в начале координат.

Нетрудно проверить, что у уравнения $z^n = 1$ ровно n корней. Как и в приведенных выше примерах, они расположены в вершинах правильного n-угольника. Один из корней равен $\cos\dfrac{2\pi}{n} + i\sin\dfrac{2\pi}{n}$, обозначим его через w. Тогда остальные корни нашего уравнения – это числа w^2, w^3, …, w^n (причем $w^n = 1$).

Упражнение 10. Докажите это.

Теперь докажем, что $w + w^2 + \ldots + w^n = 0$. Обозначим сумму $w + w^2 + \ldots + w^n$ через S. Домножим эту сумму на w. Каждое слагаемое, кроме последнего, при домножении на w перейдет в следующее слагаемое, а последнее слагаемое перейдет в первое (поскольку $w^{n+1} = w$), т.е. мы получим ту же самую сумму! Значит, $Sw = S$, откуда $S(1-w) = 0$. Но $w \neq 1$, поэтому сумма S равна нулю. Заодно мы доказали такую теорему: сумма векторов, ведущих из центра правильного многоугольника в его вершины, равна нулю.

Упражнение 11. а) Докажите, что произведение всех корней степени n из единицы равно 1.

б) Пусть s – любое натуральное число. Найдите сумму s-х степеней всех корней уравнения $z^n = 1$.

Комплексные числа в геометрии

Прежде чем читать дальше, выполните такое упражнение.

Упражнение 12. Докажите, что:

а) $|z|^2 = z\bar{z}$ для любого комплексного числа z;

б) $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$ и $\overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2$ для любых комплексных чисел z_1, z_2, и $\overline{\left(\dfrac{z_1}{z_2}\right)} = \dfrac{\bar{z}_1}{\bar{z}_2}$, если $z_2 \neq 0$.

Расстояние между точками плоскости z_1 и z_2 – это длина вектора $z_1 - z_2$, а значит, квадрат расстояния между точками z_1 и z_2 равен $|z_1 - z_2|^2 = (z_1 - z_2)(\bar{z}_1 - \bar{z}_2)$. Это простое замечание помогает решить много планиметрических задач. Докажем, например, теорему косинусов.

Теорема. *В треугольнике ABC с углом α при вершине A выполнено равенство $BC^2 = AB^2 + AC^2 - 2AB \cdot AC \cos\alpha$.*

Расположим треугольник ABC на координатной плоскости так, показано на рисунке 8 (точка A совпадает с началом координат, а точка C лежит на оси абсцисс справа от нуля). Пусть точки B и C отвечают комплексным числам z и w (при этом число w действительное и положительное, откуда $w = \bar{w} = |w|$). Длина стороны CB есть модуль числа $w - z$, откуда

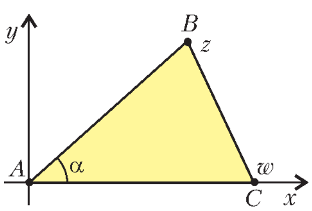


Рис. 8

$$|CB|^2 = (w - z)(\bar{w} - \bar{z}) = w\bar{w} + z\bar{z} - w\bar{z} - \bar{w}z = |w|^2 + |z|^2 - |w|(\bar{z} + z).$$

Осталось заметить, что векторы z и $\bar{z}$ симметричны относительно оси абсцисс, и, значит, их сумма равна удвоенной проекции вектора z на ось абсцисс, т.е. равна $2|z|\cos\alpha$, и теорема косинусов доказана.

Решим теперь такую задачу.

Задача 1. *Пусть $A_1A_2A_3$ и $B_1B_2B_3$ – равносторонние треугольники, причем их вершины занумерованы в порядке обхода против часовой стрелки. Докажите, что середины отрезков A_1B_1, A_2B_2 и A_3B_3 – вершины равностороннего треугольника (или совпадают).*

Пусть z_1, z_2, z_3 – соответственно вершины первого треугольника, w_1, w_2, w_3 – соответственно вершины второго (рис.9). Пусть $s = \cos\dfrac{\pi}{3} + i\sin\dfrac{\pi}{3}$ – число с модулем 1 и аргументом $\dfrac{\pi}{3}$. Условие того, что треугольник $A_1B_1C_1$ равносторонний, можно записать так: $z_3 - z_1 = s(z_2 - z_1)$ (сторона A_1A_3 получается поворотом стороны A_1A_2 на $60°$ против часовой стрелки). Аналогично записываем, что треугольник $B_1B_2B_3$ – равносторонний: $w_3 - w_1 = s(w_2 - w_1)$. Серединами отрезков A_1B_1, A_2B_2 и A_3B_3 будут точки $\dfrac{z_1 + w_1}{2}$, $\dfrac{z_2 + w_2}{2}$ и $\dfrac{z_3 + w_3}{2}$. Но

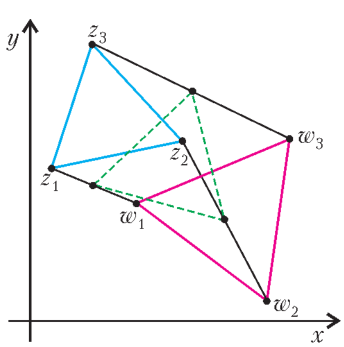


Рис. 9

$$\frac{z_3 + w_3}{2} - \frac{z_1 + w_1}{2} = \frac{z_3 - z_1}{2} + \frac{w_3 - w_1}{2} = s\frac{z_2 - z_1}{2} + s\frac{w_2 - w_1}{2} = s\left(\frac{z_2 + w_2}{2} - \frac{z_1 + w_1}{2}\right),$$

т.е. середины тоже образуют равносторонний треугольник (или совпадают), что и требовалось доказать.

В дальнейшем нам понадобится несложное, но очень полезное наблюдение, которое мы сформулируем в виде упражнения.

Упражнение 13. Докажите, что три различные точки z_1, z_2, z_3 лежат на одной прямой тогда и только тогда, когда число $\dfrac{z_1 - z_3}{z_2 - z_3}$ действительное.

Решим еще одну вспомогательную задачу.

Задача 2. *Дана окружность с центром в точке 0 и радиусом 1. Через точки z_1 и z_2 этой окружности провели касательные к ней. Найдите точку пересечения этих касательных.*

Пусть w – искомая точка. По теореме Пифагора для треугольника с вершинами в точках 0, z_1 и w получаем $z_1\bar{z}_1 + (w - z_1)(\bar{w} - \bar{z}_1) = w\bar{w}$, откуда $2z_1\bar{z}_1 = w\bar{z}_1 + \bar{w}z_1$. По условию $z_1\bar{z}_1 = 1$, и предыдущее равенство переписывается в виде $2 = \dfrac{w}{z_1} + \bar{w}z_1$. Аналогично, $2 = \dfrac{w}{z_2} + \bar{w}z_2$. Из двух последних равенств находим $w = \dfrac{2z_1z_2}{z_1 + z_2}$. Если $z_1 = -z_2$, касательные параллельны, и задача не имеет решения.

Теперь мы готовы доказать следующую замечательную теорему, носящую имя Исаака Ньютона.

Теорема Ньютона. *В описанном около окружности четырехугольнике середины диагоналей и центр этой окружности лежат на одной прямой.*

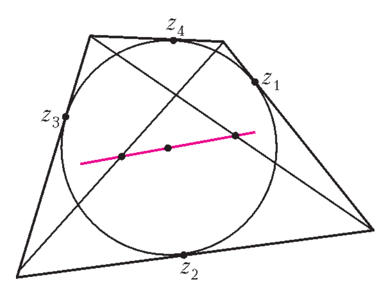


Рис. 10

Выберем систему координат и масштаб так, чтобы центр окружности был в начале координат, а радиус равнялся 1. Пусть z_1, z_2, z_3, z_4 – последовательные точки касания сторон четырехугольника с окружностью (рис.10). Вершинами четырехугольника будут тогда точки $\dfrac{2z_1z_2}{z_1 + z_2}$, $\dfrac{2z_2z_3}{z_2 + z_3}$, $\dfrac{2z_3z_4}{z_3 + z_4}$, $\dfrac{2z_4z_1}{z_4 + z_1}$, а серединами диагоналей – точки $w_1 = \dfrac{z_1z_2}{z_1 + z_2} + \dfrac{z_3z_4}{z_3 + z_4}$ и $w_2 = \dfrac{z_2z_3}{z_2 + z_3} + \dfrac{z_4z_1}{z_4 + z_1}$. Нетрудно проверить, что $\frac{w_1}{w_2} = \frac{(z_4+z_1)(z_2+z_3)}{(z_1+z_2)(z_3+z_4)}$. С другой стороны, $z_1 + z_2 = \frac{1}{\bar z_1} + \frac{1}{\bar z_2} = \frac{\bar z_1 + \bar z_2}{\bar z_1 \bar z_2}$. Заменив в предыдущем равенстве каждую скобку на аналогичное выражение, получим $\frac{w_1}{w_2} = \frac{(\bar z_4+\bar z_1)(\bar z_2+\bar z_3)}{(\bar z_1+\bar z_2)(\bar z_3+\bar z_4)}$, т.е. $\frac{w_1}{w_2} = \overline{\left(\frac{w_1}{w_2}\right)}$. Это означает, что число $\frac{w_1}{w_2}$ действительное, откуда, в силу упражнения 13, точки w_1, w_2, 0 лежат на одной прямой, что и требовалось доказать.

Разберем еще одну замечательную теорему.

Теорема Птолемея. *Произведение длин диагоналей четырехугольника, вписанного в окружность, равно сумме произведений длин его противоположных сторон.*

Сначала заметим, что для любых чисел z_1, z_2, z_3, z_4 выполнено равенство

$$(z_1-z_2)(z_4-z_3)+(z_2-z_3)(z_4-z_1) = (z_2-z_4)(z_3-z_1). \tag{1}$$

Его легко проверить, просто раскрыв скобки.

Пусть теперь z_1, z_2, z_3, z_4 – вершины произвольного четырехугольника (занумерованные по порядку). Из равенства (1) следует, что

$$|(z_1-z_2)(z_4-z_3)+(z_2-z_3)(z_4-z_1)| = |(z_2-z_4)(z_3-z_1)|.$$

Воспользовавшись упражнением 5, получаем, что

$$|(z_1-z_2)(z_4-z_3)| + |(z_2-z_3)(z_4-z_1)| \geq$$
$$\geq |(z_2-z_4)(z_3-z_1)|. \tag{2}$$

Тем самым мы уже доказали, что в любом четырехугольнике произведение длин диагоналей не превосходит сумму произведений длин противоположных сторон.

А когда неравенство (2) обращается в равенство? Сумма длин двух векторов на плоскости равна длине их суммы ровно в том случае, когда эти векторы сонаправлены, т.е. неравенство (2) обращается в равенство ровно в том случае, когда у чисел $(z_1-z_2)(z_4-z_3)$ и $(z_2-z_3)(z_4-z_1)$ аргументы одинаковы. Но это равносильно тому, что у отношения $\frac{(z_1-z_2)(z_4-z_3)}{(z_2-z_3)(z_4-z_1)}$ этих чисел аргумент равен 0, т.е. это отношение является положительным действительным числом. Обозначим его λ. Тогда $\frac{z_4-z_3}{z_2-z_3} \cdot \frac{z_2-z_1}{z_4-z_1} = -\lambda < 0$ (мы поменяли знак в одной из скобок) – полученное равносильно тому, что аргументы чисел $w_1 = \frac{z_4-z_3}{z_2-z_3}$ и $w_2 = \frac{z_2-z_1}{z_4-z_1}$ дают в сумме π.

Пусть теперь z_1, z_2, z_3, z_4 лежат на окружности (рис.11). На рисунке отмечены аргументы чисел w_1 и w_2 – видно, что они дают в сумме π в силу того, что четырехугольник вписан в окружность. Теорема Птолемея доказана.

Можно доказать и обратное: если неравенство (2) обращается в равенство, то четырехугольник вписан в окружность.

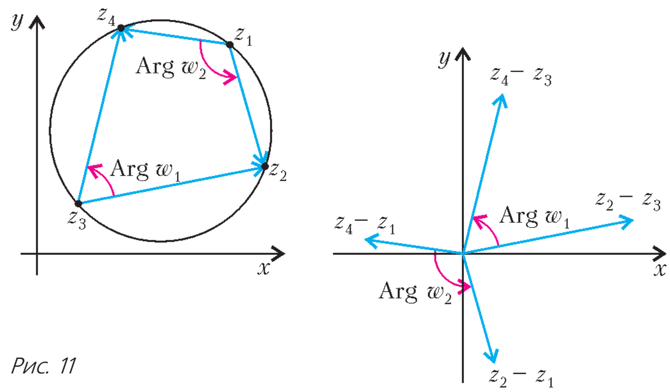


Рис. 11

Упражнение 14. *Двойным отношением* четверки различных точек z_1, z_2, z_3 и z_4 называется число $\frac{z_1-z_3}{z_2-z_3} : \frac{z_1-z_4}{z_2-z_4}$. Докажите, что двойное отношение действительно тогда и только тогда, когда точки z_1, z_2, z_3, z_4 лежат на одной окружности или прямой.

Задача о неподвижной точке и преобразования плоскости

Рассмотрим теперь такую задачу.

Задача 3. *На прямоугольную карту положили карту той же местности, но меньшего масштаба, причем меньшая карта целиком лежит внутри большей. Докажите, что можно проткнуть иголкой одновременно обе карты так, что точка прокола будет изображать на обеих картах одну и ту же точку местности.*

Попробуем записать алгебраически тот факт, что вторая карта – меньшего масштаба и как-то положена на первую. Можно считать, что наши карты лежат на плоскости. Введем на ней систему координат так, что началом будет одна из вершин первой карты, а оси будут направлены по ее сторонам. Тогда вторая карта получается из первой следующим образом: сначала мы изменяем масштаб в r раз, где r – некоторое число, меньшее 1 (т.е. умножаем все точки на r, см. рис.12),

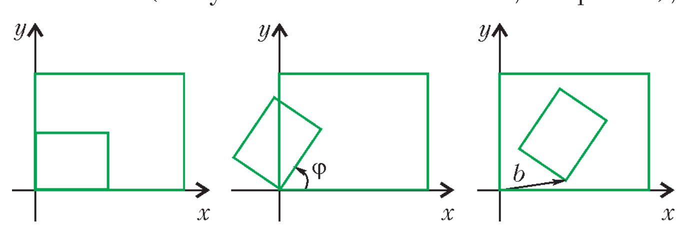


Рис. 12

затем поворачиваем уменьшенную карту на нужный угол φ (т.е. умножаем все точки на комплексное число $\cos\varphi + i\sin\varphi$), после чего сдвигаем повернутую карту на нужный вектор (прибавляем комплексное число b). В итоге получаем: если точка z плоскости обозначает на большой карте некоторую точку местности, то на маленькой карте эту же точку местности обозначает точка $az + b$ плоскости (где $a = r(\cos\varphi + i\sin\varphi)$).

Чтобы найти такую точку плоскости, которая изображает на обеих картах одну и ту же точку местности, надо просто решить уравнение $z = az + b$. Откуда получаем $z = \frac{b}{1-a}$ – нужная точка прокола найдена.

(Заметим, что $a \neq 1$ – иначе масштаб карт одинаков, что противоречит условию.)

Неужели уже все доказано? Не совсем. Ведь правило $z \to az + b$ преобразует не только точки большой карты, но и все точки плоскости. И значит, мы пока доказали, что при этом преобразовании где-то на плоскости найдется *неподвижная* точка (т.е. точка, которая остается на месте, переходит сама в себя). Но надо еще доказать, что эта точка действительно принадлежит обеим картам. Да ведь мы пока и не использовали одно важное условие – что вторая карта целиком поместилась в первой.

Заметим, что среди точек первой карты, не принадлежащих второй карте, не может быть неподвижной (поскольку они перейдут во вторую карту, а сначала были вне нее). А почему никакая точка снаружи первой карты не может остаться на месте?

Чтобы это понять, заметим сначала, что при нашем преобразовании расстояния между точками уменьшаются: ведь мы сначала уменьшаем масштаб (и значит, уменьшаем расстояния), а потом (при повороте и сдвиге) расстояния не меняются. Пусть неподвижная точка z лежит снаружи первой карты. Тогда она находится на некотором расстоянии от нее (расстояние до карты – это расстояние до ближайшей точки на границе этой карты). Найдем на границе первой карты точку w, ближайшую к z (рис.13). После преобразования расстояние между точками должно уменьшиться – но w остается в пределах первой карты, а значит, точка z должна сдвинуться, чтобы расстояние от нее до карты уменьшилось, т.е. точка z не может быть неподвижной.

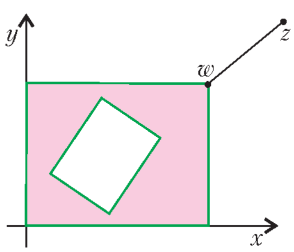


Рис. 13

Задача решена.

Здесь уместно вернуться к движениям плоскости. Как записать поворот на угол φ против часовой стрелки относительно произвольной точки a? Можно поступить так же, как в задаче про карты. Пусть $s = \cos\varphi + i\sin\varphi$. Сделаем сначала сдвиг в начало координат (z переходит в $z - a$), потом поворот относительно начала координат (умножение на s), а затем обратный сдвиг. В итоге получаем: точка z при повороте на угол φ относительно точки a переходит в точку $s(z - a) + a$. Будем считать, что $s \neq 1$.

А что будет, если мы после этого поворота сделаем еще сдвиг на вектор b? Тогда z в итоге перейдет в точку $s(z - a) + a + b$. Но оказывается, это составное движение (сначала поворот относительно a, потом сдвиг) является просто поворотом на тот же угол, но относительно некоторой другой точки (рис.14). Чтобы доказать это, обозначим искомую точку через c. При повороте на угол φ относительно c точка z переходит в $s(z - c) + c$. Значит, нам надо решить уравнение $s(z - a) + a + b = s(z - c) + c$. Решаем и находим $c = \dfrac{-sa + a + b}{-s + 1} = a + \dfrac{b}{1 - s}$.

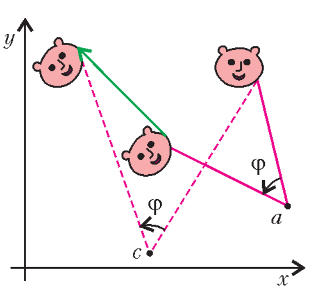


Рис. 14

Упражнение 15. Докажите, что составное движение, состоящее из поворота на угол φ относительно точки a и затем еще поворота на угол ψ относительно точки b, является либо поворотом на угол $\varphi + \psi$ относительно некоторой точки c, либо сдвигом на некоторый вектор, если $\varphi + \psi$ имеет вид $2\pi n$, где n – целое.

Комплексные числа и суммы

Покажем теперь, как с помощью комплексных чисел найти сумму $S = \cos\varphi + \cos 2\varphi + \ldots + \cos n\varphi$. Пусть $z = \cos\varphi + i\sin\varphi$. Тогда S будет равна действительной части числа $z + z^2 + \ldots + z^n$. А это число – сумма геометрической прогрессии, правда состоящей из комплексных чисел. Но формула для суммы такой прогрессии точно такая же, как и в случае действительных чисел: $z + z^2 + \ldots + z^n = \dfrac{1 - z^{n+1}}{1 - z}$ (проверьте, домножив обе части равенства на $1 - z$). Поэтому S равна действительной части числа

$$\frac{1 - \cos(n+1)\varphi - i\sin(n+1)\varphi}{1 - \cos\varphi - i\sin\varphi}.$$

Домножим числитель и знаменатель этой дроби на сопряженное к знаменателю число $1 - \cos\varphi + i\sin\varphi$ и выпишем действительную часть, получим

$$S = \frac{(1 - \cos(n+1)\varphi)(1 - \cos\varphi) + \sin(n+1)\varphi\sin\varphi}{(1 - \cos\varphi)^2 + \sin^2\varphi}.$$

Это уже ответ. Желающие могут упростить его (используя формулы тригонометрии) в качестве упражнения.

Упражнение 16. Найдите сумму

$$\sin\varphi + \frac{1}{2}\sin 2\varphi + \ldots + \frac{1}{2^n}\sin n\varphi.$$

Если вы знаете, что такое бином Ньютона и биномиальные коэффициенты, советуем решить следующее упражнение. Дадим подсказку к пункту а): запишите число $(1 - i)^n$ двумя способами – просто раскрыв скобки (по формуле бинома Ньютона) и используя формулу Муавра.

Упражнение 17. Вычислите суммы:

а) $C_n^1 - C_n^3 + C_n^5 - C_n^7 + \ldots$; б) $C_n^0 + C_n^4 + C_n^8 + C_n^{12} + \ldots$

Основная теорема алгебры

Как мы уже знаем, любой квадратный трехчлен с действительными (и даже с комплексными) коэффициентами имеет комплексный корень. А что можно сказать про многочлены более высоких степеней? Одной из жемчужин математики является следующая знаменитая теорема.

Теорема (основная теорема алгебры). *Любой многочлен положительной степени с действительными (и даже с комплексными) коэффициентами имеет комплексный корень.*

Мы не будем доказывать здесь эту теорему – она довольно сложная. Желающие могут прочитать разные доказательства этой теоремы, опубликованные в «Кванте» (см., например, статьи А.Тоома, В.Тихомирова, Л.Понтрягина в списке литературы).

Упражнение 18. Докажите теорему Безу: если $R(x)$ – многочлен с комплексными коэффициентами и s – любой его корень, то $R(x)$ делится на $x - s$, т.е. его можно представить в виде $(x-s)Q(x)$, где $Q(x)$ – некоторый многочлен.

Из этого упражнения и основной теоремы алгебры легко следует такая теорема.

Теорема. *Любой многочлен степени n с комплексными коэффициентами можно представить в виде* $a(x-b_1)(x-b_2)\ldots(x-b_n)$, *где* $a, b_1, \ldots, b_n$ *– комплексные числа (такое представление называют разложением на линейные множители).*

Применим эту теорему к многочлену с действительными коэффициентами. Пусть $P(x)$ – такой многочлен. Разложим его на линейные множители. Пользуясь свойствами комплексного сопряжения, можно вывести следующий результат.

Упражнение 19. Пусть $P(x)$ – многочлен с действительными коэффициентами, b – комплексный корень этого многочлена, т.е. $P(b)=0$. Докажите, что тогда и $P(\bar{b})=0$.

Значит, если среди множителей многочлена $P(x)$ есть $(x - b)$, причем число b не является действительным, то среди его множителей есть и $(x-\bar{b})$. Перемножим их – получим многочлен $x^2-(b+\bar{b})+b\bar{b}$ с действительными коэффициентами (ведь $b+\bar{b}$ – это удвоенная действительная часть числа b, а $b\bar{b}$ – это квадрат модуля числа b). Но тогда можно представить $P(x)$ в виде $(x^2-(b+\bar{b})+b\bar{b})P_1(x)$, где $P_1(x)$ – многочлен меньшей степени с действительными коэффициентами. Применим аналогичные рассуждения к многочлену $P_1(x)$, и так далее. В итоге мы докажем следующую теорему.

Теорема. *Любой многочлен положительной степени с действительными коэффициентами раскладывается на множители с действительными коэффициентами, среди которых только линейные и квадратные.*

В формулировке никаких комплексных чисел нет. А доказательство удалось получить благодаря использованию комплексных чисел.

Рекуррентные соотношения

Мы умеем теперь решать любое квадратное уравнение. Но где могут пригодиться комплексные корни квадратных уравнений? Разберем один пример.

В математике часто встречаются последовательности, которые заданы не явной формулой, а рекуррентно – т.е. известны несколько первых членов последовательности, и есть правило, как получать следующие члены из предыдущих. (Яркий пример – последовательность чисел Фибоначчи, где первые два числа равны 1, а каждое последующее число равно сумме двух предыдущих.)

Пусть, например, $a_1=2$, $a_2=5$ и

$$a_{n+2}=3a_{n+1}+4a_n \text{ при } n\ge 1 \qquad (3)$$

(коротко вся последовательность обозначается символом (a_n)). Мы хотим найти явную формулу для членов этой последовательности.

Найдем сначала все геометрические прогрессии вида q^n, удовлетворяющие уравнению (3). Подставляя в уравнение (3), получаем $q^{n+2}=3q^{n+1}+4q^n$. Сокращая на q^n, получим квадратное уравнение $q^2=3q+4$, решая его, находим $q = -1$ или $q = 4$. Значит, последовательности $(-1)^n$ и 4^n удовлетворяют уравнению (3). Заметим, что для любых чисел c и d последовательность $b_n=c\cdot(-1)^n+d\cdot 4^n$ тоже будет удовлетворять уравнению (3).

Теперь уже нетрудно найти формулу для членов последовательности (a_n). Если мы найдем такие числа c и d, что будет выполнено $a_1=b_1$ и $a_2=b_2$, то последовательности (a_n) и (b_n) окажутся одинаковыми. Ведь первые два члена последовательности, удовлетворяющей уравнению (3), полностью ее определяют. Так как $a_1=2$ и $a_2=5$, осталось найти такие c и d, что $2 = -c + 4d$ и $5 = c + 16d$. Решая эту систему, находим $d=\frac{7}{20}$, $c=-\frac{3}{5}$, откуда

$$a_n=-\frac{3}{5}\cdot(-1)^n+\frac{7}{20}\cdot 4^n.$$

Попробуем таким же методом найти формулу для последовательности (a_n), где $a_1=2$, $a_2=4$, и

$$a_{n+2}=2a_{n+1}-2a_n \text{ при } n\ge 1. \qquad (4)$$

Какие прогрессии вида q^n удовлетворяют уравнению (4)? Подставляя, получаем $q^{n+2}=2q^{n+1}-2q^n$, т.е. $q^2-2q+2=0$. Это квадратное уравнение не имеет действительных корней, но у него есть два комплексных корня: $1 + i$ и $1 - i$. Ищем a_n в виде $c\cdot(1+i)^n+d\cdot(1-i)^n$, получаем $a_1=2=c\cdot(1+i)+d\cdot(1-i)$, $a_2=4=c\cdot(1+i)^2+d\cdot(1-i)^2=2ci-2di$, откуда находим $c = -i$, $d = i$. Поэтому $a_n=-i(1+i)^n+i(1-i)^n$ – последовательность действительных чисел 2, 4, 4, 0, −8, −16, −16, 0, 32, 64, … – задается простой формулой, в которой участвуют комплексные числа. Впрочем, теперь от них уже можно избавиться: поскольку $1+i=\sqrt{2}\left(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}\right)$, $1-i=\sqrt{2}\left(\cos\left(-\frac{\pi}{4}\right)+i\sin\left(-\frac{\pi}{4}\right)\right)$, то, по формуле Муавра, $a_n=2\left(\sqrt{2}\right)^n\sin\frac{\pi n}{4}$ (проверьте).

Упражнение 20. Найдите явные формулы для последовательности чисел Фибоначчи: $a_0=a_1=1$, $a_{n+2}=a_{n+1}+a_n$.

* * *

Мы затронули только небольшую часть математики, где встречаются комплексные числа, даже не сформулировали некоторые интересные вопросы – можно ли, например, возводить числа в комплексную степень или брать логарифм комплексного числа. Мы только упомянем здесь одну красивую формулу, доказанную Эйлером:

$$e^{i\varphi} = \cos\varphi + i\sin\varphi$$

(она справедлива для любого действительного числа φ). Чтобы объяснить ее смысл, потребовалась бы еще целая статья. Но можно рассматривать $e^{i\varphi}$ просто как сокращенное обозначение для $\cos\varphi + i\sin\varphi$. При таком понимании нами доказано, например, равенство $e^{i\varphi}\cdot e^{i\psi} = e^{i(\varphi+\psi)}$, что соответствует правилу, по которому перемножаются степени.

Те, кто захотят продолжить свое знакомство с комплексными числами, найдут ниже подборку задач для самостоятельного решения, а также список некоторых замечательных книг и статей об этих удивительных числах.

Задачи для самостоятельного решения

1. Нарисуйте множество комплексных чисел z, для которых: а) $z^n + 1 = 0$; б) $|z - i| \le 2$; в) $\operatorname{Re}\left(\frac{1}{z}\right) = 1$; г) $|z-1| = |z-i|$; д) $|z-1| = 2|z-i|$; е) $|z-1| + |z+1| = 2$.

2. Докажите, что сумма квадратов длин сторон четырехугольника отличается от суммы квадратов диагоналей на учетверенный квадрат длины отрезка, соединяющего середины диагоналей.

3. На сторонах A_2A_3, A_1A_3, A_1A_2 треугольника $A_1A_2A_3$ во внешнюю сторону построены квадраты с центрами B_1, B_2, B_3 соответственно. Докажите, что отрезки B_1B_2 и A_3B_3 равны по длине и перпендикулярны.

4. Докажите, что в любом треугольнике центр тяжести треугольника, его ортоцентр и центр описанной окружности лежат на одной прямой (прямой Эйлера) .

5*. Каждую сторону n-угольника в процессе обхода против часовой стрелки продолжили на ее длину. Оказалось, что концы построенных отрезков служат вершинами правильного n-угольника. Докажите, что исходный n-угольник – тоже правильный.

6. Пусть P – многочлен степени k с комплексными коэффициентами. Докажите, что среднее арифметическое значений P в вершинах правильного n-угольника равно значению P в центре этого многоугольника, если $n > k$.

7. а) Пусть $z = (3+4i)/5$. Найдется ли такое натуральное число n, что $z^n = 1$?

б) Докажите, что число $\frac{1}{\pi}\operatorname{arctg}\frac{4}{3}$ иррационально.

8. Сумма трех комплексных чисел равна нулю, сумма их квадратов равна нулю и сумма их кубов тоже равна нулю. Верно ли, что и сумма четвертых степеней этих чисел равна нулю?

9. Изобразите на плоскости множество точек, соответствующих комплексным числам $z = \frac{1+ti}{1-ti}$, где t пробегает все действительные числа.

10. Мы уже знаем, что некоторые преобразования плоскости задаются функциями комплексного переменного: например, функция $f(z) = \bar{z}$ задает симметрию относительно оси абсцисс, а функция $f(z) = (\cos\varphi + i\sin\varphi)z$ задает поворот на угол φ против часовой стрелки. Запишите как функцию комплексного переменного:

а) ортогональную проекцию на ось абсцисс;
б) симметрию относительно оси ординат;
в) центральную симметрию с центром A;
г) гомотетию с коэффициентом k и центром A;
д) симметрию относительно прямой $y = 3$ со сдвигом на 1 влево;
е) поворот, переводящий ось абсцисс в прямую $y = 2x + 1$;
ж) симметрию относительно прямой $y = 2x + 1$;
з) инверсию относительно единичной окружности (этот пункт для тех, кто знаком с понятием инверсии).

11. Дан многочлен $P(x)$ с действительными коэффициентами, который при всех действительных x принимает только неотрицательные значения. Выведите из результатов раздела «Основная теорема алгебры», что этот многочлен можно представить в виде суммы нескольких квадратов многочленов с действительными коэффициентами.

12*. Пусть $c_1, c_2, \ldots, c_n, z$ – такие комплексные числа, что

$$\frac{1}{z-c_1} + \frac{1}{z-c_2} + \ldots + \frac{1}{z-c_n} = 0,$$

причем точки $c_1, c_2, \ldots, c_n$ являются вершинами выпуклого многоугольника. Докажите, что точка z лежит внутри этого многоугольника.

13. Для действительных чисел есть понятие «меньше». Докажите, что это понятие нельзя распространить на комплексные числа так, чтобы выполнялись условия:

1) для действительных чисел понятие не изменилось;
2) из двух разных комплексных чисел ровно одно меньше другого, причем если $x < y$ и $y < z$, то $x < z$;
3) при умножении неравенства на число, меньшее нуля, неравенство меняет знак.

14. а) Найдите все общие комплексные корни многочленов $x^n - 1$ и $x^m - 1$ и наибольший общий делитель этих многочленов.

б) Решите эту же задачу для многочленов $x^n + 1$ и $x^m + 1$.

15. Рассмотрим последовательности (a_n), удовлетворяющие уравнению

$$a_{n+2} = 4a_{n+1} - 4a_n \text{ при } n \ge 1. \qquad (5)$$

а) Докажите, что среди геометрических прогрессий вида q^n уравнению (5) удовлетворяет только одна (какая?).

б) Докажите, что любую последовательность (a_n), удовлетворяющую уравнению (5), можно задать формулой вида $c\cdot 2^n + dn\cdot 2^n$, где c и d – некоторые числа, и найдите их для случая последовательности с начальными членами $a_1 = 1$, $a_2 = 2$.

Литература

1. *М.Балк, Г.Балк, А.Полухин*. Реальные применения мнимых чисел. (Киев: Радянська школа, 1988)
2. *А.Тоом*. Дама с собачкой. («Квант» №2 за 1990 г.)
3. *В.Тихомиров*. Теоремы существования и основная теорема алгебры. («Квант» №4 за 2005 г.)
4. *Л.Понтрягин*. Основная теорема алгебры. («Квант» №4 за 1982 г.)
5. *В.Сендеров, А.Спивак*. Суммы квадратов и целые гауссовы числа. («Квант» №3 за 1999 г.)
6. *В.Алексеев*. Теорема Абеля в задачах и решениях. (М.: МЦНМО, 2001)
7. *Я.Понарин*. Алгебра комплексных чисел в геометрических задачах. (М.: МЦНМО, 2004)
8. *А.Киселев*. Алгебра. Часть II. (М.: Физматлит, 2005).
9. *Ю.Соловьев*. Комплексные числа. («Квант» №7 за 1991 г.)