

数理科学基礎 (線形代数学) / 線形代数学
レポート解答解説

穂坂 秀昭

2015 年 5 月 19 日

目次

S1 ターム	数理科学基礎 (線形代数学)	1
第 1 回	複素数と代数学の基本定理	3
1	はじめに	3
2	複素 (数) 平面の幾何	4
3	多項式の性質	7
第 2 回	種々の函数	13
1	全体的な講評	13
2	逆函数について	13
3	逆三角函数と円周率の近似公式	14
4	双曲線函数と逆双曲線函数	18
第 3 回	座標空間と数ベクトル	21
1	情報の調べ方	21
2	空間内における直線と平面の取り扱い	25
3	ベクトルの 1 次独立性と行列式	28
4	残りの問題	30
S2 ターム	線形代数学 1	31

S1 ターム

数理科学基礎 (線形代数学)

数理科学基礎 (線形代数学) 第 1 回

複素数と代数学の基本定理

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 4 月 15 日 1 限

1 はじめに

ごあいさつ みなさん、はじめまして。この授業の TA (ティーチング・アシスタント) をすることになりました、数理科学研究科博士課程の穂坂といいます。これから 1 セメスターの間、よろしくお願いします。

この授業では毎回レポート問題が課され、それを TA の穂坂が添削して返却します。また答案返却に合わせて、この文書のような解説プリントを配付していく予定です。何か疑問要望等があれば、提出するレポートの片隅にメッセージを書くなり、メールを hosaka@ms.u-tokyo.ac.jp に送るなり、授業後に聞くなりしてください。

課題提出時のお願い 答案が消えたら困りますので、次の 2 点は必ず守ってください。

- 氏名、学生証番号の両方を書いてください。
- 答案が複数枚に渡るときは、左上をホチキス止めしてください。

問題を解くにあたって 毎週出題される問題はレポート課題ですから、とにかくにも期限までに提出しないといけません。とは言っても、どうせ解くなら 1 問 1 問からなるべく多くの教訓を引きずり出したいし、何よりなるべく楽しい問題を解きたいものです。レポートに取り組むときは、次のようなことを意識してください。

- 簡単な計算問題は、授業で扱った定理などを確かめるための具体例です。常に「どの定理を、どう使っているのか」を考えながら解きましょう。
- どんな問題であっても、ただ解くだけでなく、見通しの良い解法を探すべきです。計算問題ならなるべく手間を減らし、証明問題なら本質的な部分を捉えるよう努力しないといけません。

このプリントの作り方について せっかくなので、このプリントをどう作っているかについて説明しておきます。

ふつう「コンピュータで文書を作る」というと、大抵の人が Microsoft Word とか一太郎といったワープロソフトを連想すると思います。ところが残念なことに、市販のワープロソフトでは数式を入力するのに大変な苦勞を強いられてしまいます。そこで数学が専門の人はどうするかというと、そういったワープロソフトの代わりに“ \LaTeX ”というソフトウェアを使います。これは D. E. Knuth^{*1} という非常に有名な数学者・計算機科学者が作った“ \TeX ”というソフトウェアを、色々な人が改良してできあがったものです。

\LaTeX はワープロソフトとはちょっと違い、文字のスタイルを変えたり見出しをつけたりするのに「コマンド」というものを使います。ですから \LaTeX を使うにはコマンドの使い方を覚えなければいけません。加えてキーボードが打ち込んだものが、見た目通りに出てくるわけでもありません。一旦コマンドも含めて打ち込んだ文書を \LaTeX というプログラムその他色々処理させることによって、やっと整形された文書がでできます。ですから使い始めるにはちょっとハードルが高いのですが、使い慣れるとワープロソフトよりも手際よく文書が書けるし、また数式を中心として文書の仕上がり美しいというメリットもあります。

もしかしたら皆さんの中には \LaTeX を既に知っている人がいるかもしれませんが、また将来 \LaTeX を使う必要に迫られる人がいるかもしれません。そこで https://github.com/HideakiHosaka/2015_linear_algebra に、この文書の PDF ファイルと \LaTeX ソースコードを置いておきます。もし \LaTeX の方に興味がある人は、ページ右下の“Download ZIP”のボタンから一式をダウンロードしてください。また東京大学が持つ情報処理システムのオンライ

^{*1} <http://www-cs-faculty.stanford.edu/~uno/>

ン自習教材「はいぱーワークブック」の第 27 章^{*2}に、 \LaTeX の説明があります。 \LaTeX を使う人は、一度読んでおく
と良いと思います。

ちなみにソースコードの公開には “GitHub” というサービス^{*3}を利用しています。上に貼った URL を開くと、古い
バージョンのプリントや、そうしたプリントがどう更新されていったかを見ることができます。授業自体の役には立た
ないと思いますが、興味があれば見てみてください。

2 複素 (数) 平面の幾何

2.1 複素数と複素平面

複素数の定義 まず、最初に複素数の定義をおさらいしましょう。 $i^2 = -1$ という規則で i という「数」^{*4}を定めま
す。このとき、2 つの実数 $x, y \in \mathbb{R}$ を用いて $z = x + yi$ と表される数を複素数と言うのでした。また $z = x + yi$ の x
を実部、 y を虚部と言うことも知っているはずですが、足し算と掛け算はそれぞれ、分配法則などが上手く成り立つよう

$$(x + yi) + (x' + y'i) := (x + x') + (y + y')i, \quad (x + yi)(x' + y'i) := (xx' - yy') + (xy' + x'y)i$$

と決めていました^{*5}。また、これらのルールがあれば、 $x' + y'i \neq 0$ のとき

$$\frac{x + yi}{x' + y'i} = \frac{(x + yi)(x' - y'i)}{(x' + y'i)(x' - y'i)} = \frac{(xx' + yy') + (x'y - xy')i}{(x')^2 + (y')^2} = \frac{xx' + yy'}{(x')^2 + (y')^2} + \frac{x'y - xy'}{(x')^2 + (y')^2}i$$

というように割り算もできます。こうして複素数では四則演算が全部できると分かりました。

複素平面 さて、全ての複素数は $x + yi$ の形に、 (x, y) という二つの実数 x, y のペアを用いて表せます。また複素
数を二つの実数 x, y で $x + yi$ の形に表す方法がただ一通りなことも明らかでしょう。これらの事実^{*6}から、複素数
 $z = x + yi$ と座標平面上の点 (x, y) とを 1 : 1 に対応させられることが分かります。このように平面 \mathbb{R}^2 上の点一つ
一つを複素数と見なしたとき、平面 \mathbb{R}^2 のことを複素 (数) 平面^{*7}と呼びます。また x 軸、 y 軸をそれぞれ実軸 (real axis)、
虚軸 (imaginary axis) と言います。

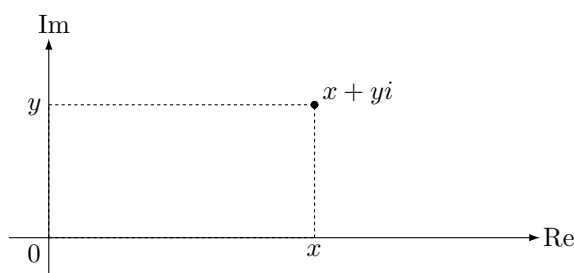


図 1.1 複素平面における数と点の対応

^{*2} <http://hwb.ecc.u-tokyo.ac.jp/current/applications/latex/>

^{*3} もしあなたが既に “GitHub” を知っているなら、きっと “pull request” の機能も知っているはずですが、プリントに対して何か意見があれば、積極的に pull request を送ってください。:D

^{*4} 誰もが一度は「 -1 の平方根を数と呼んでいいのか」という疑問を抱いたことがあると思います。その疑問に対する答えをまだ書いていなかったたので、ここでは括弧つきで「数」と書きました。でも一々こう書くのが面倒なので、以下では括弧をつけないことにします。

^{*5} ここで出てくる $:=$ は「左辺のものを右辺で定義する」という意味です。式変形をするときの $=$ とは意味が違います。定義の際には $:=$ が使われることがあります。また左右を入れ替えて $=$ とすると「右辺のものを左辺で定義する」という意味になります。

^{*6} 集合と写像の言葉できちんと書くと、 $(x, y) \in \mathbb{R}^2$ に $x + yi \in \mathbb{C}$ を対応させる写像 $\mathbb{R}^2 \rightarrow \mathbb{C}$ が全単射、ということです。

^{*7} 複素平面と複素数平面は、どちらの言葉も同じ意味です。複素平面の方を使う人が多いですが、複素数平面と言っても誤解を招くことはないですし、昔は複素数平面という言葉も割と良く使われていたそうです。関西学院大学の示野信一先生のブログに詳しい事情が書いてありますので、気になる人は読んでみてください: <http://mathsci.blog41.fc2.com/blog-entry-60.html>

少し大げさに言うと、我々は「複素数という代数的なもの」と「平面という幾何的なもの」を対応付けました。このことには、非常に重要な意味があります。なぜなら代数の観点と幾何の観点を行ったりきたりすることで、色々なことが分かるようになるからです。たとえば長さなどの幾何的な情報を代数的な操作で捉えたり、逆に掛け算などの代数的な操作を図形的に捉えたりというように。これから、それをやってみましょう。

2.2 幾何を代数で捉える

複素数の大きさと共役 幾何的な情報の最も典型的なものとして「2点間の距離」が挙げられます。複素平面の場合、原点 0 と $z \in \mathbb{C}$ との距離を z の絶対値または大きさと言い、 $|z|$ で表します。三平方の定理から、すぐに $|x + yi| = \sqrt{x^2 + y^2}$ が従います。またベクトルのときと同様、複素数 $z, z' \in \mathbb{C}$ の間の距離は $|z - z'|$ となります。

幾何的な観点からは、「線/点对称移動」といった操作を考えられるというメリットもあります。たとえば「実軸に対する線対称移動」で z が写る点を \bar{z} と書くと、 $\overline{x + yi} = x - yi$ です。この \bar{z} を z の共役と言います*⁸。共役は $\overline{z + w} = \bar{z} + \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$ という性質を満たすことが、計算で確かめられます。また共役を用いて、複素数の大きさを表すことができます。

複素数と複素数平面: 問1の解答 $z = x + iy$ のとき、 $z\bar{z} = (x + yi)(x - yi) = x^2 + y^2 = |z|^2$ となる。 $z = 0$ であることは z と原点 0 との距離が 0 であることと同値なので、直ちに $z = 0 \Leftrightarrow |z| = 0$ を得る。 ■

複素数と複素数平面: 問4の解答 (1) $z_1 = x_1 + y_1i$, $z_2 = x_2 + y_2i$ とおく。絶対値は 0 以上の実数だから、 $|z_1 + z_2| \leq |z_1| + |z_2|$ の代わりに $|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$ を示せばよい。実際計算すると

$$\begin{aligned} (|z_1| + |z_2|)^2 - |z_1 + z_2|^2 &= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 - |z_1 + z_2|^2 \\ &= x_1^2 + y_1^2 + 2\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2} + x_2^2 + y_2^2 - ((x_1 + x_2)^2 + (y_1 + y_2)^2) \\ &= 2\left(\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2} - (x_1x_2 + y_1y_2)\right) \end{aligned}$$

となる。そして括弧の中は、平方根が非負であることと次の計算とを組み合わせれば、 0 以上と分かる。

$$\left(\sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}\right)^2 - (x_1x_2 + y_1y_2)^2 = x_1^2y_2^2 + y_1^2x_2^2 - 2x_1x_2y_1y_2 = (x_1y_2 - x_2y_1)^2 \geq 0$$

これで示すべきことが言えた。(2) は (1) を使えば $|z_1 - z_2| = |z_1 + (-z_2)| \leq |z_1| + |-z_2| = |z_1| + |z_2|$ となる。 ■

2.3 代数を幾何で捉える

続いて、四則演算という代数的な操作を複素平面で見えてみましょう。複素数の足し算や引き算がベクトルの足し算や引き算と全く同じであることは、すぐに分かると思います。非自明なのは平面上の点やベクトルには掛け算が定義されていないのに対し、複素数には掛け算があるという点です*⁹。そこで「2つの複素数を掛け算した結果は、複素平面上ではどのように見えるか」が問題となってきます。まずは問題を1つ解いてみましょう。

複素数と複素数平面: 問2の解答 $z = 2 + i$ とおくと

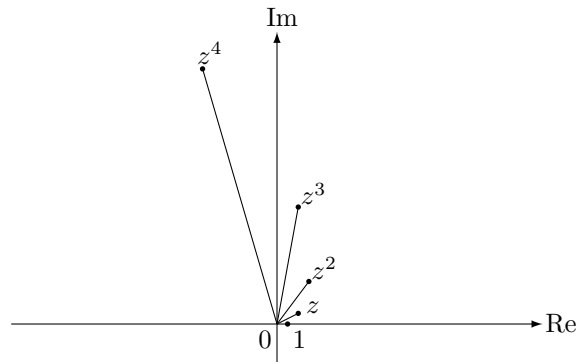
$$z^2 = 3 + 4i, \quad z^3 = 2 + 11i, \quad z^4 = -7 + 24i, \quad z^5 = -38 + 41i, \quad z^6 = -117 + 44i$$

である。これらをプロットした結果*¹⁰は次の図の通り。

*⁸ この部分は嘘ではないですが、若干語弊があります。今は「実軸に関する線対称移動」という幾何学的な操作として共役を定義しましたが、本来「共役」とは、実数 \mathbb{R} から複素数 \mathbb{C} を作る「拡大」という代数的な操作に伴って定義されるものです。

*⁹ 「ベクトルにも内積や外積があるじゃん」という声が聞こえてきそうですが、内積や外積は、いわゆる普通の「積」とは違う性質を持ちます。2つのベクトルの内積は数になってしまい、ベクトルにはなりません。また2つのベクトルの外積はベクトルになりますが、ベクトルの外積は順序を入れ替わると結果が変わります。この辺が数の掛け算と全然違うところです。

*¹⁰ z^5 と z^6 は図から激しくはみ出るので描いていません。

図 1.2 $z = 2 + i$ のべき乗のプロット

極形式表示 いま複素数 $z \neq 0$ に対し $z^0 = 1, z^1 = z, z^2, z^3, z^4$ を平面上にプロットし、これらの点を原点と結んだ結果を眺めると、隣り合う角が全て同じ大きさであるように見えます。それを実際に確かめてみましょう。

角度を計算したいので、極座標を使うのが筋がよさそうです。そこで $z = x + yi$ の表す点 (x, y) を、極座標 (r, θ) を用いて $x = r \cos \theta, y = r \sin \theta$ と表します。これに対応する複素数の方で表すと

$$z = x + yi = r \cos \theta + (r \sin \theta)i = r(\cos \theta + i \sin \theta)$$

となります。この書き方を、複素数の極形式表示と呼びます。極形式表示のもとで $r = |z|$ です。また θ は、複素平面の半直線 $0z$ と実軸の非負の部分がなす角を、反時計回りに測った角度となっています。この θ を複素数 z の偏角といい、 $\theta = \arg z$ と書きます。 $\arg z$ は一通りでなく 2π の整数倍だけずれますが、今は気にしないでおきます。

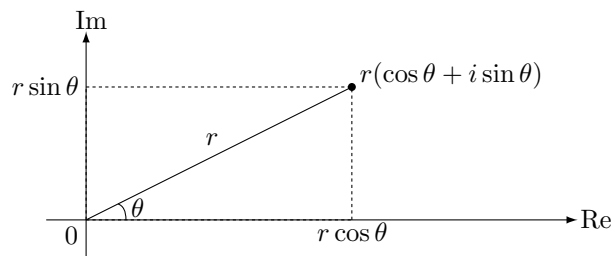


図 1.3 複素数の極形式表示

さて、極形式で表された 2 つの複素数を掛け算すると

$$\begin{aligned} r(\cos \theta + i \sin \theta) \times r'(\cos \theta' + i \sin \theta') &= rr' \{ (\cos \theta \cos \theta' - \sin \theta \sin \theta') + i(\sin \theta \cos \theta' + \cos \theta \sin \theta') \} \\ &= rr' (\cos(\theta + \theta') + i \sin(\theta + \theta')) \end{aligned}$$

となります。最後の式変形は、もちろん三角函数^{*11}の加法定理を使っています。この式は非常に重要なことを示唆しています。それは複素数 z, z' の積 zz' について

- 大きさは、 $|zz'| = |z||z'|$ で与えられる
- 偏角は $\arg zz' = \arg z + \arg z'$ で与えられる

ということです。言い換えれば、複素数 z に対して別の複素数 z' を掛け算する操作は

- z の大きさを $|z'|$ 倍し
- z の偏角に $\arg z'$ を足し算する

^{*11} 函数と関数は同じ意味です。少々古臭い言い回しですが、好みの問題でこちらを使います。

ということに他ならないからです。このように極形式を使うことによって、複素数の掛け算が「拡大縮小」と「回転」の組み合わせという図形的意味を持つことが読み取れるのです。

なお、一々 $\cos \theta + i \sin \theta$ と書いているのは長ったらしくて大変なので、以下では実数 θ に対して $e^{i\theta} = \cos \theta + i \sin \theta$ や $\exp i\theta = \cos \theta + i \sin \theta$ と表すことにします。たとえば $e^{i\pi} = \cos \pi + i \sin \pi = -1$ という感じです。指数関数 e^x と同じ記法を用いることには実は意味がある^{*12}のですが、今は「単なる記号」だと思っていてください。この記号を使うと、極形式表示での掛け算は

$$re^{i\theta} \times r'e^{i\theta'} = rr'e^{i(\theta+\theta')}$$

と書けます。スッキリしてていいですね。

極形式表示は計算面でも、非常に威力を発揮することがあります。

複素数と複素数平面: 問3の解答 $z = e^{2k\pi i/n}$ ($k = 0, 1, \dots, n-1$) とおくと、

$$z^n = (e^{\frac{2k\pi i}{n}})^n = \exp\left(\frac{2k\pi i}{n} + \frac{2k\pi i}{n} + \dots + \frac{2k\pi i}{n}\right) = \exp\left(\frac{2k\pi i}{n} \times n\right) = e^{2k\pi i} = (e^{2\pi i})^k = 1$$

となる。ここで k が $k = 0, 1, \dots, n-1$ を動けば、 n 個の異なる複素数が得られる^{*13}。また多項式 $z^n - 1$ は n 次式だから、 $n+1$ 個以上の根を持つことはない。ゆえに $z = e^{2k\pi i/n}$ ($k = 0, 1, \dots, n-1$) が全ての根を与える。 ■

2.4 残りの問題

ここまでで紹介していなかった問題の解答を記します。

複素数と複素数平面: 問6の解答 $x, y \in M$ とすると $x = a^2 + b^2$, $y = c^2 + d^2$ となる整数 $a, b, c, d \in \mathbb{Z}$ が取れる。このとき $x = |a + bi|^2$, $y = |c + di|^2$ なので $xy = |a + bi|^2 |c + di|^2 = |(a + bi)(c + di)|^2 = |(ac - bd) + (ad + bc)i|^2 = (ac - bd)^2 + (ad + bc)^2$ となる。 $a, b, c, d \in \mathbb{Z}$ より $ac - bd, ad + bc \in \mathbb{Z}$ である。よって $xy \in M$ である。 ■

複素数と複素数平面: 問7の解答 次の式の t に好きな有理数を代入すればよい。

$$\left(\frac{2t}{1+t^2}\right)^2 + \left(\frac{1-t^2}{1+t^2}\right)^2 = 1$$

$t = \tan \theta$ とおくと、この式は $\sin^2 2\theta + \cos^2 2\theta = 1$ に化ける。したがって $0 \leq t \leq 1$ の範囲で $2t/(1+t^2)$ は短調増加する。これと 0 以上 1 以下の有理数が無限個存在することから、有理点も無限個存在することが分かる。 ■

3 多項式の性質

3.1 多項式の割り算

実数係数や複素数係数^{*14}の多項式では、割り算と余りの計算ができます。 $f(x)$ を m 次多項式、 $g(x)$ を n 次多項式、 $m \geq n$ とすると、 $f(x) = g(x)q(x) + r(x)$ を満たす多項式 $g(x)$ と n 次未満の多項式 $r(x)$ がただ一つだけ存在します。実際 $f(x) = a_mx^m + a_{m-1}x^{m-1} + \dots + a_0$, $g(x) = b_nx^n + b_{n-1}x^{n-1} + \dots + b_0$ とおくと、 $f(x) - b_nx^{m-n}g(x)/a_m$ の次数は $m-1$ 次以下になります。こうやって「 $g(x)$ に上手い数と x のべき乗をかけ、 $f(x)$ の項を次数が高い順に消していく」という操作をすれば、商と余りの計算ができます。実際の計算には筆算を使ったり、あるいは $g(x)$ が 1 次式のときは「組立除法」という技が使えたりしますが、その辺は割愛します。

^{*12} そのうち微分積分学の授業で、関数の Taylor 展開というものを習うはずですが、その後に巾 (べき) 級数で指数関数 e^x を定義し直すと、元々の「 e のなんとか乗」という意味を越えて、 e^x の x に複素数を代入できるようになります。そうして初めて $e^{i\theta} = \cos \theta + i \sin \theta$ という式に正しく意味を与えることができます。

^{*13} 複素平面上にプロットすれば、異なることが直ちに分かります。

^{*14} 有理数係数でも大丈夫です。より一般に、係数が「体」と呼ばれるものであれば、複素数係数と同様に割り算ができます。

特に $g(x)$ が 1 次式 $x - \alpha$ のとき、 $f(x) = (x - \alpha)q(x) + r(x)$ に出てくる $r(x)$ は 0 次以下^{*15}の式、つまり定数です。なので $r(x)$ の代わりに r と書くと、 $f(x) = (x - \alpha)q(x) + r$ の両辺に $x = \alpha$ を代入して $r = f(\alpha)$ が得られます。よって、 f が $x - \alpha$ で割り切れることと $f(\alpha) = 0$ が同値になります。この事実を因数定理と呼ぶのでした。

因数定理などを用いて解ける問題を、まとめて片付けてしまいましょう。

多項式: 問 1 の解答 $f(x)$ が $x - \alpha$ で割り切れるので、 $f(x) = (x - \alpha)g(x)$ と書ける。また $f(x)$ は $x - \beta$ でも割り切れるので、 $f(\beta) = 0$ である。よって $0 = f(\beta) = (\beta - \alpha)g(\beta)$ となるが、 $\alpha \neq \beta$ より $g(\beta) = 0$ でないといけない。よって $g(x)$ は $(x - \beta)$ で割り切れる。これより $f(x)$ は $(x - \alpha)(x - \beta)$ で割り切れる。 ■

多項式: 問 2 の解答 $f(x) = x^3 + 10x^2 + ax - 2$ に $x = 2, -3$ を代入した値が等しい。よって $f(2) = 46 + 2a$ と $f(-3) = -3a + 61$ が等しいのだから、 $a = 3$ が得られる。求める余りは 52 である。 ■

多項式: 問 3 の解答 $f(x) = mx^3 + nx^2 - 5$ とおく。 $f(-\frac{1}{2}) = 0$ より $-\frac{m}{8} + \frac{n}{4} - 5 = 0$ である。また $f(\frac{2}{3}) = 7$ より $\frac{8}{27}m + \frac{4}{9}n - 5 = 7$ である。これより $m = 6, n = 23$ である。 ■

多項式: 問 4 の解答 $F(x)$ を $2x^2 + x - 1$ で割った余りを $px + q$ と書くと、何か多項式 $P(x)$ を用いて

$$F(x) = (2x^2 + x - 1)P(x) + px + q$$

と書ける。 $F(x)$ を $x + 1$ で割った余りが 6 なので $F(-1) = 6$ である、よって上式に $x = -1$ を代入して $6 = -p + q$ を得る。同様に $F(x)$ を $2x - 1$ で割った余りが 3 なので、 $F(\frac{1}{2}) = 3$ である。これより $3 = \frac{1}{2}p + q$ を得る。こうして p, q の連立 1 次方程式が得られたので、解くと $p = -2, q = 4$ が得られる。よって余りは $-2x + 4$ である。 ■

多項式: 問 6 の解答 3 次多項式 $f(x)$ は、適当な多項式 $g(x)$ と $h(x)$ によって $f(x) = g(x)(x^2 - 1) + 5x - 8 = h(x)(x^2 - x - 6) + 17x + 4$ と書ける。これに $x = 1, -1, -2, 3$ を代入すると、それぞれ $f(1) = -3, f(-1) = -13, f(-2) = -30, f(3) = 55$ が得られる。そこで $f(x) = ax^3 + bx^2 + cx + d$ とおくと

$$\begin{cases} a + b + c + d &= -3 \\ -a + b - c + d &= -13 \\ -8a + 4b - 2c + d &= -30 \\ 27a + 9b + 3c + d &= 55 \end{cases}$$

という連立一次方程式が得られる。これを a, b, c, d について解けば $f(x) = 2x^3 + 3x - 8$ と分かる。 ■

3.2 有名な多項式

今回の問題の中にはいくつか有名な多項式が出てくるので、問題を解きつつ紹介します。

複素数と複素数平面: 問 5 の解答 $f(z) = z^4 + z^3 + z^2 + z + 1$ である。

(1) z を $f(z) = 0$ の根とする。 $z^5 - 1 = (z - 1)f(z) = 0$ なので、 $z \neq 0$ である。よって $f(z)/z^2 = 0$ である。また $f(z)/z^2 = z^2 + z + 1 + z^{-1} + z^{-2} = (z + z^{-1})^2 + (z + z^{-1}) - 1 = t^2 + t - 1$ である。よって $t^2 + t - 1 = 0$ となるので、これを解いて $t = \frac{-1 \pm \sqrt{5}}{2}$ を得る。

(2) $t = z + z^{-1}$ より $z^2 - tz + 1 = 0$ である。 $t = \frac{-1 \pm \sqrt{5}}{2}$ のとき、この方程式を解くと

$$z = \frac{t \pm \sqrt{t^2 - 4}}{2} = \frac{-1 + \sqrt{5} \pm \sqrt{(1 - \sqrt{5})^2 - 16}}{4} = \frac{-1 + \sqrt{5} \pm \sqrt{-10 - 2\sqrt{5}}}{4} = \frac{-1 + \sqrt{5} \pm \sqrt{10 + 2\sqrt{5}}i}{4}$$

^{*15} 0 でない定数は 0 次式ですが、0 だけは次数を $-\infty$ と定めます。これは多項式の次数を \deg で表すとき、 $\deg fg = \deg f + \deg g$ が常に成り立つようにするためです。

となる。同様に $t = \frac{-1-\sqrt{5}}{2}$ のとき

$$z = \frac{t \pm \sqrt{t^2 - 4}}{2} = \frac{-1 - \sqrt{5} \pm \sqrt{(-1 - \sqrt{5})^2 - 16}}{4} = \frac{-1 - \sqrt{5} \pm \sqrt{-10 + 2\sqrt{5}}}{4} = \frac{-1 - \sqrt{5} \pm \sqrt{10 - 2\sqrt{5}}i}{4}$$

が得られる。これらが全ての解である。

(3) $z = e^{2k\pi i/5}$ ($k = 0, 1, 2, 3, 4$) が $z^5 = 1$ の全ての解である。複素平面にプロットすれば、(2) で求めた解のうち実部と虚部がともに正なものが $e^{2\pi i/5}$ だと分かる。これと $\sin(\frac{\pi}{2} - \theta) = \cos \theta$, $\cos(\frac{\pi}{2} - \theta) = \sin \theta$ より

$$\frac{-1 + \sqrt{5} + \sqrt{10 + 2\sqrt{5}}i}{4} = e^{2\pi i/5} = \cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5} = \sin \frac{\pi}{10} + i \cos \frac{\pi}{10}$$

となる。この式の実部と虚部を見ればよい。 ■

多項式: 問 5 の解答 (1) $x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + a^2x^{n-3} + \cdots + a^{n-1})$

(2) もし $x^n - a^n$ が $x + a$ で割り切れることは、 x に $-a$ を代入した結果が 0 になることと同値である。すなわち $0 = (-a)^n - a^n = ((-1)^n - 1)a^n$ より、 $(-1)^n = 1$ が必要十分条件である。これは n が偶数であることに他ならない。

(3) $x^n + a^n$ に $x = -a$ を代入すると $(-a)^n + a^n = ((-1)^n + 1)a^n$ となる。これが 0 になることは n が奇数であることと同値である。よって n が奇数なら $x^n + a^n$ は $x + a$ で割り切れる。 ■

円周等分多項式 今の問題で出てきた因数分解 $x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + \cdots + a^{n-1})$ は非常に良く見かけます。特に $a = 1$ と置いてできる多項式 $x^n - 1$ の複素根は、極形式で考えれば直ちに $e^{2k\pi i/n}$ ($k = 0, 1, \dots, n-1$) と分かります。これらの解をプロットすると、原点を中心とする半径 1 の円周が n 等分されます。そういう理由で n が素数のとき^{*16}、 $z^n - 1$ を $z - 1$ で割ってできる多項式 $z^{n-1} + z^{n-2} + \cdots + 1$ のことを円周等分多項式と呼びます。

多項式: 問 7 の解答 答えのみ記す。

$$(1) 4(ab + cd)^2 - (a^2 + b^2 - c^2 - d^2)^2 = (a + b + c - d)(a + b - c + d)(a - b + c + d)(-a + b + c + d)$$

$$(2) x^3 - (a + b + c)x^2 + (ab + bc + ca)x - abc = (x - a)(x - b)(x - c)$$
 ■

基本対称式 今の問題の (2) で出てきた $a + b + c$, $ab + bc + ca$, abc はいずれも a, b, c について対称な多項式です。このような多項式を a, b, c の対称式と言う^{*17}のでした。特に、ここに出てきた 3 つの対称式は基本対称式という名前がついています。これは「全ての対称式は基本対称式たちの定数倍の和と積で書ける」という重要な定理があるからです。遠くない将来にまた登場すると思いますので、その時に改めて詳しく解説します。

3.3 複素数について

“Was sind und was sollen die Zahlen?”

これは、ドイツの数学者 Richard Dedekind が記した本^{*18}のタイトルです。『数とは何か、また何であるべきか?』この問題を、考えてみましょう。

複素数は数なのか? 我々が日常生活の中で出会う「数」にはどんなものがあるでしょうか? たとえば物の個数を数えるときは自然数を使いますし、お金の計算をするときは収入と支出を表すのに正の数と負の数を使います。また料理をすればレシピの中に分数が出てきますし、円周の長さを測ろうとしたら π のような無理数も現れます。これらに登場する数は、いずれも実数の範囲に収まっていますね。

^{*16} n が素数でないときも円周等分多項式は定義されますが、諸々の事情で $z^{n-1} + z^{n-2} + \cdots + 1$ そのものにはなりません。

^{*17} 変数の数が増えても同様で、多項式 $f(x_1, \dots, x_n)$ がどの 2 つの x_i, x_j を入れ替えても変わらないとき、 f を n 変数の対称式と言います。

^{*18} ちなみに、1893 年に出版されたドイツ語原著の第 2 刷が、東大駒場キャンパスの数理科学図書室にあります。

一方、複素数は実数の範囲を超えるものです。そのため $i^2 = -1$ となる数 i を我々の世界で目にすることはありません。たとえばおつかいに行った子供が「ママー！ おつりで $300 + 250i$ 円もらったよ！」なんて言うわけないですね。複素数を「気持ち悪い」と感じる主たる理由は、おそらくここにあるのではないのでしょうか。英語では空想上の数 “imaginary number” と呼ばれますし、日本語ではさらにネガティブな含みを持つ「虚」数という呼び名もあります。

ですが「我々の身の回りに見当たらないから」というだけの理由で、複素数を数と呼ぶべきではないのでしょうか。ここで一度、我々の身近にある数について「何が数たらしめているのか」を考えてみましょう。数を考える上で何よりも大事なことは「計算」です。たとえば自然数だったら足し算と掛け算ができます。整数なら、引き算がいつでもできます。有理数や実数なら、0 以外の数による割り算もできます。また計算とは別に「大小の比較」ができることも、数の大きな特徴でしょう。

複素数は残念なことに「大小の比較」をすることはできません。ですが既に見てきたとおり、複素数では四則演算の全てを行うことができます。これをもって「数」と呼んでも良いのではないのでしょうか。また「 -1 の平方根」というと気味が悪いかもしれませんが、「形式的に i という数を付け加え、 $i^2 = -1$ というルールで計算を行う」という風に思えば、 i の存在も受け入れられる気がします。

実際、現代数学では今のような方法で複素数を捉えています。一般に実数や有理数のように「四則演算ができる数の範囲」を、数学では体 (たい) と呼びます。そしていま、実数係数の 1 変数多項式全体を $\mathbb{R}[x]$ と書きます。このとき多項式 $f(x)$ に対して「 x^2 に -1 を代入する」という操作、より正確には「多項式 $f(x)$ を $x^2 + 1$ で割った余りを考える」という操作^{*19}をすれば、多項式の変数だった x が -1 の平方根であるかのように振る舞います。これを「多項式 $x^2 + 1$ の根を添加して実数体 \mathbb{R} を拡大する」と言います。このように体が与えられたとき、その中に入らない元を付け加えて大きい体を作る拡大という操作を定義することによって、きちんと複素数体 \mathbb{C} を定義できるのです。

複素数: 問 8 の解答 大体の人が「実数の範疇を飛び出る」とか「大小がなくなる」と書いてくれました。もちろん、どちらも数の性質を踏まえた、真っ当な感じ方だと思います。 ■

3.4 代数学の基本定理

今回最後の話題は、複素数係数の多項式の根です。多項式 $P(x)$ に対し、方程式 $P(x) = 0$ の解を根と言うのでした。まずは「係数が全て実数」という特別な場合に、複素数の根が「共役とペアで現れる」という事実を確認しましょう。

代数学の基本定理: 問 1 の解答 $f(z) = a_0 + a_1z + \cdots + a_nz^n$ とおく。このとき $a_0, a_1, \dots, a_n \in \mathbb{R}$ である。よって $\alpha \in \mathbb{C}$ が $f(z)$ の解であるとき、 $f(\alpha) = 0$ の共役を取ると

$$0 = \overline{f(\alpha)} = \overline{a_0 + a_1\alpha + \cdots + a_n\alpha^n} = \overline{a_0} + \overline{a_1}\bar{\alpha} + \cdots + \overline{a_n}\bar{\alpha}^n = a_0 + a_1\bar{\alpha} + \cdots + a_n\bar{\alpha}^n = f(\bar{\alpha})$$

となる。よって $\bar{\alpha}$ も解である。 ■

代数学の基本定理 さて多項式が根を持つかどうかは、「考える数の範囲」によって変わってきます。たとえば $P(x) = x^2 + 1$ は実数の範囲で根を持ちませんが、複素数の範囲に広げると $x = \pm i$ という根を持ちます。このように「多項式がいつも根を持つか」は係数の範囲に依存するのですが、実はなんと、複素数で考える限り定数でない多項式は必ず 1 個は根を持つのです^{*20}。これが代数学の基本定理と呼ばれる内容です。

^{*19} 代数学の用語では「多項式環 $\mathbb{R}[x]$ をイデアル $(x^2 + 1)$ で割る」という操作に相当します。

^{*20} ただし「解を持つ」とこと「解が計算で求まる」ことは全く別の問題です。たとえば f を連続関数とします。このとき $a < b$ とすると、 f は閉区間 $[a, b]$ 上のどこかの点 c で最大値を取ります。この事実はグラフを描けばすぐ納得できますが、でも「最大値を取る c がどこにあるか」について、全く教えてくれません。代数学の基本定理も、これと同じ状況になっています。

ちなみに 2 次方程式は解の公式を使えば常に解けますが、Galois 理論というものを使うと、5 次以上の多項式に対する「解の公式」が存在しないことまで証明できます。

そして「根が1個は存在する」ことが分かった、そこから「定数でない n 次多項式はいつも n 個の根を持つ」ことまで言えてしまいます。このことを問題で確かめてから、最後に代数学の基本定理の完全な証明を与えましょう*21。

代数学の基本定理: 問2の解答 f の次数に関する帰納法で示す。まず $\deg f = 1$ のときは1次式1個の積である。次に、 $n-1$ 次以下の全ての複素係数多項式が1次式の積に分解すると仮定する。このとき $f(z)$ を n 次多項式とすると、代数学の基本定理より $f(z)$ の根が存在する。それを α とすると $f(z) = (z - \alpha)g(z)$ と書け、 $g(z)$ は $n-1$ 次多項式となる。帰納法の過程から $g(z)$ は1次式の積に分解するので、 $f(z)$ 全体も1次式の積となる。

代数学の基本定理: 問3の解答 $g(0) \neq 0$ より、 $g(z)$ の定数項は0でない。定数項の次に低い次数の項が $a_k z^k$ だとすると、 $g(z) = a_0 + a_k z^k + a_{k+1} z^{k+1} + \cdots + a_n z^n$ とおく。うまく z の偏角を調節すれば a_0 と $a_k z^k$ の偏角が π だけずれ、原点から見て逆向きになるようにできる。その上で $|z|$ を十分小さくすれば、 $|a_0 + a_k z^k|$ は $|a_0|$ より小さくなるはずである。さらに $|z|$ を小さくすれば、 z の $(k+1)$ 乗以上の項の絶対値は、 z^k の項の絶対値に比べていくらかでも小さくなる。そうすれば $|g(z_0)| < |g(0)|$ となるはずである。この考え方にに基づき、 z の適切な偏角と絶対値を見出そう。

いま $a_0 = r_0 e^{i\theta_0}$, $a_k = r_k e^{i\theta_k}$ ($r_0, r_k > 0$) と書ける。これらを用いて $z_0 := r e^{i(\theta_0 - \theta_k + \pi)/k}$ と定めると

$$\begin{aligned} g(z_0) &= a_0 + a_k z_0^k \left(1 + \frac{a_{k+1}}{a_k} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k}\right) = r_0 e^{i\theta_0} + r_k e^{i\theta_k} \left(r e^{i(\theta_0 - \theta_k + \pi)/k}\right)^k \left(1 + \frac{a_{k+1}}{a_k} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k}\right) \\ &= r_0 e^{i\theta_0} + r_k r^k e^{i(\theta_0 + \pi)} \left(1 + \frac{a_{k+1}}{a_k} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k}\right) = (r_0 - r_k r^k) e^{i\theta_0} - r_k r^k e^{i\theta_0} \left(\frac{a_{k+1}}{a_k} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k}\right) \end{aligned}$$

となる。よって $|e^{i\theta_0}| = 1$ と三角不等式より、 $|f(z_0)|$ は

$$\begin{aligned} |g(z_0)| &\leq |r_0 - r_k r^k| + r_k r^k \left| \frac{a_{k+1}}{a_k} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k} \right| = |r_0 - r_k r^k| + r_k r^k |z_0| \left| \frac{a_{k+1}}{a_k} + \frac{a_{k+2}}{a_{k+1}} z_0 + \cdots + \frac{a_n}{a_k} z_0^{n-k-1} \right| \\ &\leq |r_0 - r_k r^k| + r_k r^{k+1} \left\{ \left| \frac{a_{k+1}}{a_k} \right| + \left| \frac{a_{k+2}}{a_k} z_0 \right| + \cdots + \left| \frac{a_n}{a_k} z_0^{n-k-1} \right| \right\} \end{aligned}$$

となる。この式で

$$M := \max \left\{ \left| \frac{a_{k+1}}{a_k} \right|, \left| \frac{a_{k+2}}{a_k} z_0 \right|, \dots, \left| \frac{a_n}{a_k} z_0^{n-k-1} \right| \right\}$$

とおくと、 $|g(z_0)| \leq |r_0 - r_k r^k| + r_k r^{k+1} (n - k - 1) M$ が得られる。そこで

$$r := \min \left\{ \left| \frac{r_0}{r_k} \right|^{\frac{1}{k}}, \frac{1}{2(n - k - 1) M} \right\}$$

と取れば、 $r_0 - r_k r^k \geq 0$ かつ $r \leq \frac{1}{2(n - k - 1) M}$ なので

$$|g(z_0)| \leq r_0 - r_k r^k + r_k r^{k+1} (n - k - 1) M \leq r_0 - r_k r^k (1 - r(n - k - 1) M) \leq r_0 - \frac{r_k r^k}{2} < r_0$$

である。これより $|g(z_0)| < r_0 = |g(0)|$ となることが分かった。 ■

代数学の基本定理の証明 せっかく問3の解答を与えたので、代数学の基本定理の証明を与えておきます。以下、 $f(z)$ を複素係数の多項式とします。証明のアイデアは次の通りです。

1. $f(z) = 0$ となる z が存在することを「 z が複素平面全体を動く時の $|f(z)|$ の最小値が0」と言い換える。
2. 原点を中心とする十分大きい半径 R の円板を考え、その外側には $|f(z)|$ の最小値が決して現れないことを言う。
3. 半径 R の円板内のどこかで、 $|f(z)|$ が最小値を取ることを言う。
4. $|f(z)|$ の最小値が0でないと、矛盾が生じることを示す。

*21 この証明には微分積分学の知識が必要なので、今は読み解くのが難しいかもしれませんが。連続関数の性質を習った後で読むと、手頃な勉強になるでしょう。

step 1. $|z| > R$ なる全ての複素数 z に対して $|f(z)| > |f(0)|$ が成り立つような正の実数 $R > 0$ が取れることを示す。

$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ とおく。このとき

$$|f(z)| = |a_n||z|^n \left| 1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n} \right|$$

である。ここで三角不等式 $|z_1| - |z_2| \leq |z_1 - z_2|$ を使うと、 $|z_1 + z_2| \geq |z_1| - |z_2| = |z_1| - |z_2|$ なので

$$\left| 1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n} \right| \geq \left| 1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_1}{a_n z^{n-1}} \right| - \left| \frac{a_0}{a_n z^n} \right| \geq \cdots \geq 1 - \left| \frac{a_{n-1}}{a_n z} \right| - \cdots - \left| \frac{a_0}{a_n z^n} \right|$$

となる。ここで

$$R_0 := \max \left\{ \frac{n|a_{n-1}|}{0.1|a_n|}, \left(\frac{n|a_{n-2}|}{0.1|a_n|} \right)^{\frac{1}{2}}, \dots, \left(\frac{n|a_0|}{0.1|a_n|} \right)^{\frac{1}{n}} \right\}$$

とおく^{*22} と、 $|z| > |R_0|$ のとき、全ての $1 \leq k \leq n-1$ に対し

$$\left| \frac{a_{n-k}}{a_n z^k} \right| = \frac{|a_{n-k}|}{|a_n||z|^k} < \frac{|a_{n-k}|}{|a_n|R_0^k} \leq \frac{|a_{n-k}|}{|a_n|} \left(\frac{0.1|a_n|}{n|a_{n-k}|} \right)^{\frac{1}{k} \cdot k} = \frac{0.1}{n}$$

となる。よって

$$\left| 1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n} \right| \geq 1 - \left| \frac{a_{n-1}}{a_n z} \right| - \cdots - \left| \frac{a_0}{a_n z^n} \right| > 1 - \frac{0.1}{n} - \cdots - \frac{0.1}{n} = 0.9$$

となる。さらに正の実数 R を

$$R := \max \left\{ \left(\frac{|f(0)|}{0.9|a_n|} \right)^{\frac{1}{n}}, R_0 \right\}$$

とおくと、 $|z| > R$ のとき

$$|f(z)| = |a_n||z|^n \left| 1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n} \right| > |a_n| \left(\frac{|f(0)|}{0.9|a_n|} \right)^{\frac{1}{n} \cdot n} \cdot 0.9 = |f(0)|$$

となる。

step 2. 複素平面上の実数値関数 $|f(z)|$ が最小値を持つことを示す。

半径 R の閉円板 $\Delta_R := \{z \in \mathbb{C} \mid |z| \leq R\}$ を考える。 $|f(z)|$ は $\mathbb{C} = \mathbb{R}^2$ 上の連続関数である。そして Δ_R は平面 \mathbb{R}^2 内の有界閉集合だから、 $|f(z)|$ は Δ_R 上で最小値を取る。その点を z_0 とする。いま $|z| > R$ とすると、 R の取り方から $|f(z)| > |f(0)|$ が従う。一方 $0 \in \Delta_R$ より $|f(z_0)| \leq |f(0)|$ である。これらより $|z| > R$ のときも $|f(z)| > |f(z_0)|$ が従う。かくして $|f(z_0)|$ は、複素平面全体における $|f(z)|$ の最小値である。

step 3. 方程式 $f(z) = 0$ が解を持つことを示す。

もし $|f(z_0)| \neq 0$ だとすると、 $f(z_0) \neq 0$ である。そこで新しい多項式 $g(z)$ を $g(z) := f(z + z_0)/f(z_0)$ で定めることができる。このとき $g(0) = 1$ なので、問3の結果より、適当な複素数 $z_1 \in \mathbb{C}$ を取ると $|g(z_1)| < 1$ とできる。ところが $|g(z_1)| = |f(z_1 + z_0)|/|f(z_0)|$ と合わせると $|f(z_1 + z_0)| < |f(z_0)|$ が導かれてしまう。これは $|f(z_0)|$ が複素平面全体で $|f(z)|$ の最小値になることに矛盾する。ゆえに $|f(z_0)| = 0$ である。つまり、 z_0 は $f(z)$ の根である。 ■

^{*22} 0.1 という数に本質的な意味はありません。この 0.1 は、1 未満の任意の正の数 ε に置き換えて構いません。

数理科学基礎 (線形代数学) 第 2 回 種々の函数

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 4 月 22 日 1 限

1 全体的な講評

今回のテーマは「高校生のうちに一応習いはするけど、でもあんまり深入りはしない函数たち」です。逆三角函数や双曲線函数について、名前自体は授業前から知っていた人が多かったのではないのでしょうか。いずれの問題も、そうした函数の性質を知るためのものです。

問題は比較的易しいものが多かったので、ほとんどの人が良く解けていました。中には一つの問題に複数の方法で取り組んだり、あるいは関連する問題をまとめて解いたりした人がいました。大変良いことだと思います。普通に問題を解くのに加え、問題を解く過程から何かを見出せるよう頑張ってください。

また今回、答案の中に質問やリクエストを書いてくれた人もいました。そういうことが書いてあると解説プリントを作るときの指針になりますので、何か思うことがあれば、ぜひ答案内に書き込んでおいてください。

それでは、今回のテーマを一つずつ見ていきましょう。

2 逆函数について

2.1 逆函数の定義

函数とは、数に対して何かしらの数に対応させる規則のことです。たとえば $f(x) = x^2$ という函数は、実数 $x \in \mathbb{R}$ に $x^2 \in \mathbb{R}$ という数に対応させています。他にも三角函数、対数函数や指数函数といった有名な函数があり、そしてこれらの函数を合成するとさらに色々な函数を作ることができますが、どの函数も数に数に対応させることには間違いありません。別の言い方をすれば、僕たちが普段「函数」と呼んでいるものは、定義域と値域が実数全体 \mathbb{R} (の部分集合) であるような写像です^{*1}。

さて函数 $f(x)$ が与えられると、数 x が与えられた時に f の値 $f(x)$ が定まります。そしてしばしば「数 y が与えられたとき、 $y = f(x)$ となる数 x を求めたい」という状況が生じます。対応 $x \mapsto f(x)$ ではなく、その逆の対応を求めたいわけです。このように函数 f が与えられたとき、「数 y に対し、 $y = f(x)$ となる x を対応させる」という方法で定義される函数を f の逆函数と言い、 f^{-1} と書きます。 $x \mapsto f(x)$ の対応が逆になるので、逆函数のグラフは元の函数のグラフを直線 $y = x$ について線対称に折り返したものになります。

2.2 逆函数の主値

逆函数を定義するための条件 残念なことに、与えられた函数に対していつでも逆函数が無条件で定義できるわけではありません。上で述べたように、函数は写像ですから「1 つの数に、1 つの数に対応させる」ことが大原則です。だから「1 つの数に 2 つの数に対応させる」という操作になってしまっただけでは、函数とは呼べないのです。

いま、函数 f が異なる実数 x_1, x_2 に対して $f(x_1) = f(x_2)$ を満たしたとしましょう。そして $y = f(x_1) = f(x_2)$ とおきます。すると x_1, x_2 のどちらに対しても y が対応するので、逆函数を作ろうとしたら「 y に x_1, x_2 の両方が対応する」というマズい状況が実際に起きます。こうならないためには「異なる実数 x_1, x_2 に対して必ず $f(x_1) \neq f(x_2)$

^{*1} ただし、この用語の使い方は慣用的なもので、厳密な取り決めがあるわけではありません。たとえば定義域や値域がベクトルの集合であっても、函数と呼ぶことはままあります。

が成り立つ」という条件が必要です。こういう性質を 単射性 というのです。

また単射なら良いかというと、そうでもありません。たとえば指数函数 $\exp x := e^x$ を考えましょう。指数函数を $\exp: \mathbb{R} \rightarrow \mathbb{R}$ という写像と思えば、これは単射です。グラフを描けば狭義短調増加ですから、異なる値に対して同じ値が対応しようがありません。そして正の実数 $y > 0$ に対しては、 $e^x = y$ となる x が $x = \log y$ で与えられます。ですが $y \leq 0$ のとき、 $y = \exp x$ となる $x \in \mathbb{R}$ は存在しません。さっきは「1つの数に2個(以上)の数に対応してしまう」ことが問題でしたが、今度は「ある数に対して、対応させられる数がなくなってしまった」という問題が起きています。これももちろん写像の定義を満たさないので、指数函数 \exp を \mathbb{R} から \mathbb{R} への写像と思うと、逆函数は作れません。ですが \exp の値域を正の整数全体の集合 $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$ だと思えば、どんな正の数 $y > 0$ に対しても $x := \log y$ とおくことで $y = e^x$ なる実数 x を見つけられます。このような、函数 f の「どんな値域の元 y に対しても、 $y = f(x)$ となる x が存在する」という性質を、全射性と呼ぶのです。

まとめると、 f の逆函数が存在するための必要条件は、 f が全単射であることです。また f の逆函数 g が存在すれば、 f は全単射になると示せます。かくして f の逆函数が存在することと、 f が全単射であることが同値だと分かります。

逆函数の主値 ただし f が全単射でないときも、部分的には f の逆函数を作れることが多いです。さっきの指数函数の例では、値域を \mathbb{R} から $\mathbb{R}_{>0}$ に取り換えれば逆函数が作れました。また f の定義域を制限してしまうのも一つの手です。たとえば $f(x) = x^2$ という函数を考えましょう。正の実数 $x > 0$ に対し $f(-x) = (-x)^2 = x^2 = f(x)$ という式が成り立ってしまうので、 f は単射ではありません。ですが f の定義域を非負実数全体の集合 $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ に制限してしまえば、 f は単射になります。そしてどんな非負の実数 $y \in \mathbb{R}_{\geq 0}$ に対しても、 $x := \sqrt{y}$ とおけば $y = x^2 = f(x)$ が成り立ちます。こうして f を制限して $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ という写像にすれば、 f の逆函数 $y \mapsto \sqrt{y}$ が得られます。 f の定義域を制限して逆写像 f^{-1} を作るとき、その値を f^{-1} の主値と言います。

ちなみに複素数の偏角 \arg の「主値」という言葉遣いも、函数の主値と同じ意味です。0 以外の複素数全体の集合を $\mathbb{C}^\times := \{z \in \mathbb{C} \mid z \neq 0\}$ と書きます。すると $z \in \mathbb{C}^\times$ の偏角 $\arg z$ は 2π 単位でずらせるので、主値を決めないと普通の函数にはなりません。そこで主値を半開区間 $[0, 2\pi)$ とすれば、偏角は $\arg: \mathbb{C}^\times \rightarrow [0, 2\pi)$ という函数になります*2。

2.3 逆函数の微分

函数 f が微分可能で逆函数 f^{-1} が定義される場合、 f^{-1} も大体は微分可能になります。いま $y_0 = f(x_0)$ だったとします。このとき $f'(x_0)$ は、 $y = f(x)$ のグラフの点 (x_0, y_0) における接線 ℓ の傾きです。一方、 $y = f(x)$ のグラフを直線 $y = x$ について折り返せば逆函数 $x = f^{-1}(y)$ のグラフが得られます。したがって $y = f(x)$ のグラフと一緒に接線 ℓ を折り返せば、 $x = f^{-1}(y)$ の点 (y_0, x_0) における接線が得られます。その接線の傾きは ℓ の接線の傾きの逆数になります。したがって $(f^{-1})'(y_0) = 1/f'(x_0)$ となります。

逆函数の微分公式をちゃんと示すには少し神経を使いますから、ここでは証明はしません。でも、式自体は使いやすいと思います。また忘れても「グラフの接線を $y = x$ について折り返す」というイメージさえ覚えておけば、必要な場面では思い出せるはずです。

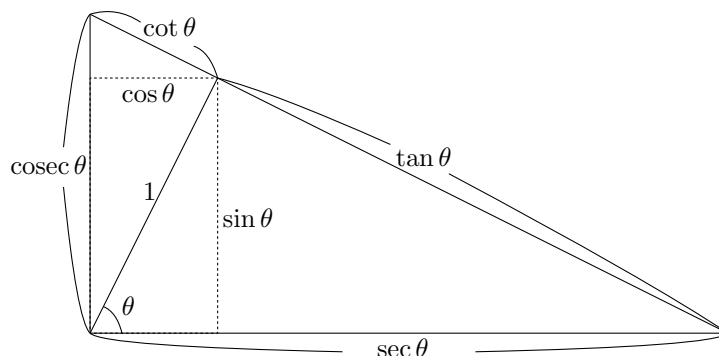
3 逆三角函数と円周率の近似公式

3.1 逆三角函数の定義

念の為に三角函数の定義を確認しておきます。単位円周上の点 P に対し、 x 軸の非負の部分と半直線 OP のなす角が θ であるとき、 P の座標を $(\cos \theta, \sin \theta)$ と表すのです。そして残りは $\tan \theta := \sin \theta / \cos \theta$, $\cot \theta := 1 / \tan \theta$,

*2 ですが無理やり函数にした代償で、実軸正の部分での連続性が破れることに注意しましょう。偏角 0 付近で連続性を保ちたい場合は、主値を $(-\pi, \pi]$ などに取り直す必要があります。また一般に、どのような主値を選んでも \arg を \mathbb{C}^\times 全体で連続な函数にはできません。

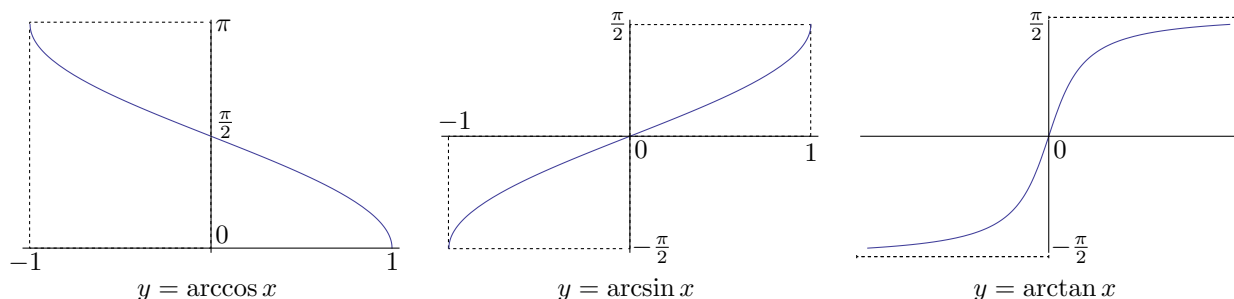
$\sec \theta := 1/\cos \theta$, $\operatorname{cosec} \theta := 1/\sin \theta$ と定めます。これを図に描くと次のようになります。



三角函数の逆函数が逆三角函数です。cos, sin, tan の逆函数をそれぞれ arccos, arcsin, arctan と書きます。ただし三角函数はすべて周期函数なので、ある意味「単射から最も遠い函数」です。たとえば $\sin x = 1/2$ となる実数 x は2個どころか無限個存在してしまいます。ですから先ほど説明したように、逆三角函数を考えるには、三角函数の定義域を制限して主値を選ぶ必要があります。通常、主値は次のように選びます。

函数	定義域	主値
arccos	$[-1, 1]$	$[0, \pi]$
arcsin	$[-1, 1]$	$[-\pi/2, \pi/2]$
arctan	\mathbb{R}	$(-\pi/2, \pi/2)$

この主値に関するグラフは次の通りです。直線 $y = x$ について折り返すと、元の三角函数のグラフになりますね。



この中で一番よく使うのは arctan だと思います。というのも直線 tan は角度から傾きを計算するのに使う函数ですから、arctan は「与えられた傾きを実現する角度は何度か？」を表しています。いかにも役立ちそうですね。

3.2 arctan 1 の公式たち

三角函数に色々な公式があるように、逆三角函数にも色々な公式があります。その中でも arctan の計算公式は比較的有名です。というのも後で見ると、arctan の計算ができると、 $\pi/4 = \arctan 1$ を使って円周率を求めることができるからです。arctan 1 を色々な方法で表してみましょう。

問 2, 3 の解答 tan の加法定理

$$\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$$

に $u = \tan x$, $v = \tan y$ を代入すると

$$\arctan u + \arctan v = x + y = \arctan \tan(x + y) = \arctan \frac{\tan x + \tan y}{1 - \tan x \tan y} = \arctan \left(\frac{u + v}{1 - uv} \right)$$

を得る。この式で $u = 1/2$, $v = 1/3$ とおくと

$$\arctan \frac{1}{2} + \arctan \frac{1}{3} = \arctan \left(\frac{\frac{1}{2} + \frac{1}{3}}{1 - \frac{1}{6}} \right) = \arctan 1 = \frac{\pi}{4}$$

が得られる。これは Euler の公式と呼ばれる。

また $u = v = 1/5$ とおくと、同じように計算することで

$$2 \arctan \frac{1}{5} = \arctan \frac{\frac{2}{5}}{1 - \frac{1}{25}} = \arctan \frac{5}{12}$$

となる。さらに $u = v = 5/12$ に対しては

$$2 \arctan \frac{5}{12} = \arctan \frac{\frac{10}{12}}{1 - \frac{25}{144}} = \arctan \frac{120}{119}$$

である。そして $u = 120/119$ に対し $(u + v)/(1 - uv) = 1$ となる v を求めると、 $v = (1 - u)/(1 + u) = -1/239$ となる。これらをまとめると、Machin の公式

$$\frac{\pi}{4} = \arctan \frac{\frac{120}{119} + \frac{-1}{239}}{1 - \frac{120}{119} \cdot \frac{-1}{239}} = \arctan \frac{120}{119} + \arctan \frac{-1}{239} = 2 \arctan \frac{5}{12} - \arctan \frac{1}{239} = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}$$

を得る*3。

問 4 の解答 計算自体は簡単である。

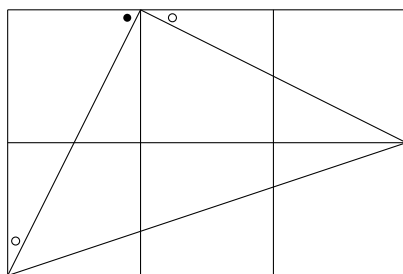
$$(1) (3 + i)(7 + i) = 20 + 10i, \quad (2) (2 + i)(3 + i) = 5 + 5i, \quad (3) (5 + i)^4 / (239 + i) = 2 + 2i$$

ただ、この問題で重要なのは計算結果から \arctan の公式が得られる点である。複素数 $z = x + yi$ の偏角が $\arctan(y/x)$ で与えられること、また偏角が $\arg zz' = \arg z + \arg z'$ という式を満たすことを思い出そう。両辺の偏角を取れば

$$\arctan \frac{1}{3} + \arctan \frac{1}{7} = \arctan \frac{1}{2}, \quad \arctan \frac{1}{2} + \arctan \frac{1}{3} = \arctan 1 = \frac{\pi}{4}, \quad 4 \arctan \frac{1}{5} - \arctan \frac{1}{239} = \arctan 1 = \frac{\pi}{4}$$

が得られる。

問 5 の解答 次の図を見れば、求める角度は直角二等辺三角形の内角だから 45 度と分かる。



この図もまた、 $\arctan(1/2) + \arctan(1/3) = \pi/4$ を表している。

*3 途中で 1 回、主値 $(-\pi/2, \pi/2)$ のもとで \arctan が奇関数であることを使います。

3.3 円周率の計算

\arctan には、実は次のような近似公式があります：

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \quad (\text{if } -1 \leq x \leq 1)$$

この公式は、形式的には

$$\arctan x = \int_0^x (\arctan t)' dt = \int_0^x \frac{1}{1+t^2} dt = \int_0^x (1-t^2+t^4-t^6+\cdots) dt = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

のようにして求められます*⁴。右辺は気合で頑張ればいくらでも精度よく計算できますから、この公式をガリガリ計算することで $\arctan x$ の値を求めることができます。たとえば $x = 1$ とすれば

$$\frac{\pi}{4} = \arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

という風にして円周率が求められます。これは Leibniz の公式と呼ばれているようです。

ただ、今の $\pi/4$ 公式は非常に効率が悪いです。というのも第 50 番目の項が $-1/99 = 0.010101 \cdots$ ですから、50 番目の項を計算すると小数第 2 位の値が変わります。円周率を求める上では「ちょこっと計算しただけで、上の方の桁が正確に分かる」ような公式が望ましいわけで、こんな「50 項も計算してもまだ小数第 2 位が分からない」なんて公式は役に立ちません。そこで登場するのが Euler の公式や Machin の公式です。たとえば Euler の公式なら

$$\begin{aligned} \frac{\pi}{4} &= \arctan \frac{1}{2} + \arctan \frac{1}{3} = \left(\frac{1}{2} - \frac{1}{3 \cdot 2^3} + \frac{1}{5 \cdot 2^5} - \cdots \right) + \left(\frac{1}{3} - \frac{1}{3 \cdot 3^3} + \frac{1}{5 \cdot 3^5} - \cdots \right) \\ &= \left(\frac{1}{2} - \frac{1}{24} + \frac{1}{160} - \cdots \right) + \left(\frac{1}{3} - \frac{1}{81} + \frac{1}{1215} - \cdots \right) \end{aligned}$$

のように、分母にある x^n の形の項が効いてきて、足し引きされる項の値が急激に小さくなっていきます。今の場合、最初から 3 つずつの項を拾うだけで $\pi = 3.14 \cdots$ が求まります。さっきより断然楽ですね。Machin の公式に至っては $1/239$ という項がありますから、少ない手間でもっと精度よい値を求められます。William Jones という数学者が 1706 年に著した “Synopsis Palmariorum Mathesos” という本*⁵に、Machin が求めたとされる円周率が結果だけ 100 桁以上載っていますが、今の公式を使ったのでしょうか？

もちろん \arctan に放り込む数がい小さくなれば小さくなるほど、計算はどんどん楽になります。したがって「 π が上手く出てくる \arctan の公式を見つけて、君だけのオリジナルの円周率近似公式をつくろう！」... という話になりそうですが、実はこの公式は頑張ってもそんなに精度が出ません。いまの公式では分母に x^n が出てくるのがキモだったので「1 つ先の項を計算したときに、何桁分だけ精度が良くなるか」は毎回同じです。もし x が 100 だったら、次の項を考えてもせいぜい 2 桁程度しか精度が良くならないわけです。ところが世の中には「計算を 1 手進めるだけで、それまでの桁数の倍の桁数だけ精度が良くなる」という、圧倒的に強い公式があります。Gauss–Legendre アルゴリズムといえますので、興味のある人は調べてみてください*⁶。

ちなみに、プリントをアップロードしている GitHub のページ*⁷に、Excel で円周率の近似計算の実験をした結果 `pi_approx.xlsx` を載せてみました。いくつかの公式に対し n 番目の項の値、 n 番目の近似値と真の値からの誤差を表にしています。興味がある人は、このデータの誤差を対数プロットしてみると良いでしょう。ただし Excel は高精度計算には向かないので、途中で精度が打ち止めになっています。本気で円周率を求めるには、プログラムを書き、必要に応じて高精度計算のためのライブラリを援用しないといけません。

*⁴ この「証明」は全然厳密ではありません。何も考えずに無限和と積分の順序を入れ替えてはいけません。Taylor 展開を習った後なら、こんなことをしなくても公式が導けるようになります。

*⁵ ちなみに本郷キャンパスにある総合図書館の書庫内に、この本があるそうです。OPAC で検索すると出てきます。

*⁶ たとえば四ツ谷 晶二・村井 実『楕円関数と仲良くなろう』（日本評論社）の第 2 章に詳しい説明があります。

*⁷ https://github.com/HideakiHosaka/2015_linear_algebra

4 双曲線関数と逆双曲線関数

最後のテーマは、双曲線関数と呼ばれる函数たちです。これらは指数函数の四則演算で書けるという意味では、そこまで目新しいものではありません。ですが三角函数と良く似た性質を持ち、かつ置換積分などの際に役立つので、性質を知っておくと便利です。

4.1 双曲線関数

定義 次に挙げる函数を、双曲線関数といいます。

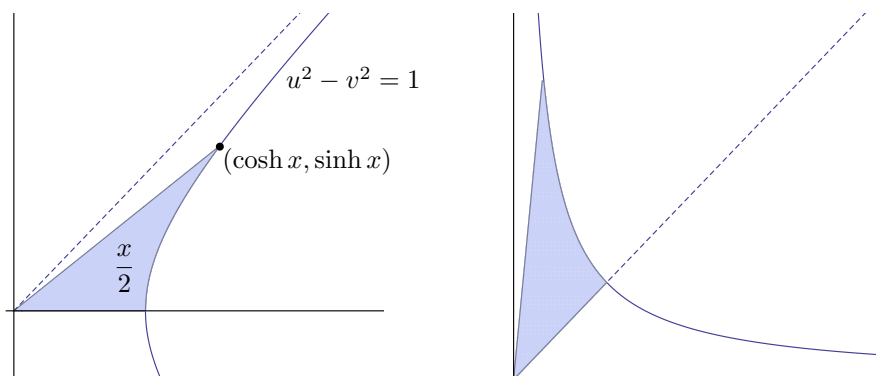
$$\cosh x := \frac{e^x + e^{-x}}{2}, \quad \sinh x := \frac{e^x - e^{-x}}{2}, \quad \tanh x := \frac{\sinh x}{\cosh x}$$

名前の付け方のルールは三角函数と同じです*⁸。

これらが双曲線関数と呼ばれる所以の一つは、双曲線のパラメータ表示を与えるからです。三角函数は $\cos^2 t + \sin^2 t = 1$ という恒等式を満たしましたが、双曲線関数では $\cosh^2 t - \sinh^2 t = 1$ という式が成り立ちます。したがって t が動けば、点 $(\cosh t, \sinh t)$ は双曲線 $x^2 - y^2 = 1$ の上を動きます。ただし $\cosh t > 0$ なので、このパラメータ表示は双曲線の右側だけしか与えていません。双曲線の左側のパラメータ表示は $(-\cosh t, \sinh t)$ で与えられます。

他にも、双曲線を使って \cosh, \sinh に図形的方法で定義を与えることもできます。

問 6 の解答 双曲線 $u^2 - v^2 = 1$ 上に、下図で塗りつぶされた部分の面積が $x/2$ となるような点を取る。この点の座標を $(\cosh x, \sinh x)$ と書く。 $\cosh x$ と $\sinh x$ の明示的な式を求めよう。面積の計算には積分が必要である。積分計算を楽にするため、この双曲線を正の向き*⁹に 45° 回転させる。そうすると双曲線 $u^2 - v^2 = 1$ は、直線 $v = u$ 上で原点からの距離が 1 である点、すなわち点 $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ を通る反比例のグラフになる。よって回転後の双曲線の式は $\eta = \frac{1}{2\xi}$ である*^{10*11}。



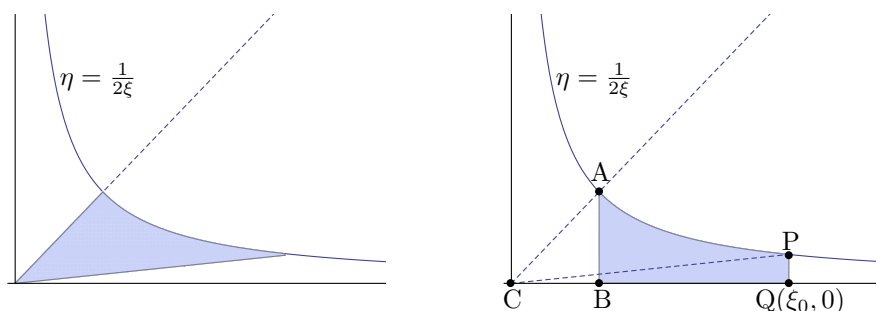
縦より横に積分した方が見やすいので、面積を求めるべき部分を直線 $\eta = \xi$ について折り返す。すると、次の二つの図で示された面積は等しいことが分かる。 $AB \cdot BC = PQ \cdot CQ = \frac{1}{2}$ が成り立ち、 $\triangle ABC$ と $\triangle CQP$ の面積が等しくなるからである。

*⁸ 三角函数の $\cot, \sec, \operatorname{cosec}$ に対応して双曲線関数でも $\coth, \operatorname{sech}, \operatorname{cosech}$ というのが定義されますが、使う場面はまず無いでしょう。

*⁹ 数学でいう「正の向き」は、反時計回りです。

*¹⁰ ξ, η はそれぞれギリシャ文字のグザイとエータです。

*¹¹ 軸を u, v から ξ, η に変えたことに、深い意味はありません。



さて点 Q の ξ 座標を ξ_0 とおく。B の ξ 座標は $\frac{1}{\sqrt{2}}$ なので、積分で面積を計算すると

$$\frac{x}{2} = \int_{\frac{1}{\sqrt{2}}}^{\xi_0} \frac{d\xi}{2\xi} = \frac{1}{2} \left(\log \xi_0 - \log \frac{1}{\sqrt{2}} \right) = \frac{1}{2} \log(\sqrt{2}\xi_0)$$

と求まる。一方で点 P は、点 $(\cosh x, \sinh x)$ を x 軸について折り返してから正の向きに 45° 度回転することで得られるのであった。ゆえに、この平面を複素平面と同一視することで P の座標が

$$\overline{\cosh x + i \sinh x} \times e^{i\pi/4} = (\cosh x - i \sinh x) \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i \right) = \frac{\cosh x + \sinh x}{\sqrt{2}} + \frac{\cosh x - \sinh x}{\sqrt{2}}i$$

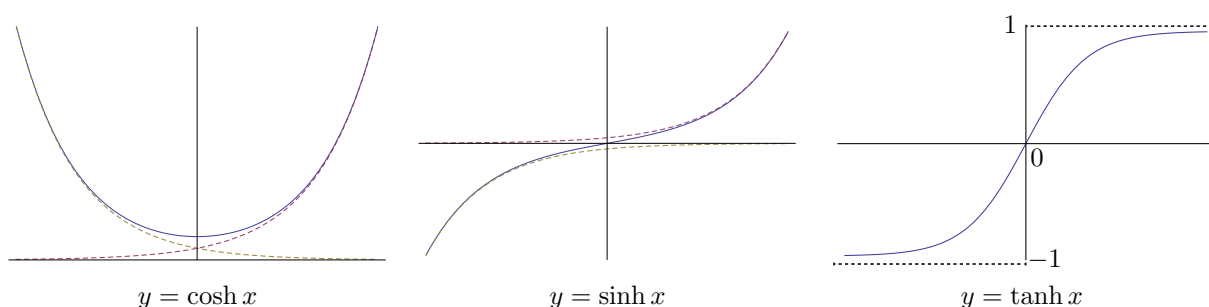
と分かる^{*12}。この実部が ξ_0 なので、面積の計算と合わせて $x = \log(\sqrt{2}\xi_0) = \log(\cosh x + \sinh x)$ が従う。つまり $\cosh x + \sinh x = e^x$ である。

最後に、点 $(\cosh x, \sinh x)$ は双曲線 $u^2 - v^2 = 1$ の上にあったことを思い出す。これより

$$1 = \cosh^2 x - \sinh^2 x = (\cosh x - \sinh x)(\cosh x + \sinh x) = (\cosh x - \sinh x)e^x$$

という式が得られる。すなわち $\cosh x - \sinh x = e^{-x}$ である。かくして $\cosh x + \sinh x = e^x$ と合わせて $\cosh x$ と $\sinh x$ の連立 1 次方程式が得られたので、これを解いて $\cosh x$ と $\sinh x$ の求める式を得る。 ■

グラフ 双曲線函数のグラフは次の通りです。 $\cosh x$ と $\sinh x$ のグラフに点線で描き込まれているのは、函数 $e^x/2$ と $\pm e^{-x}/2$ です。



本体だけ見ると $\cosh x$, $\sinh x$ はそれぞれ 2 次、3 次函数のグラフに似てなくもないです。しかし中に指数函数があるので、 $x \rightarrow \pm\infty$ での絶対値の増大度が全然違います。気を付けましょう。また $\cosh x$ のグラフには「懸垂線」という名前が付いています。これは、糸の両端を持ってぶら下げた時にできる曲線が \cosh で書けることに由来します。

問 6 (d) の解答と導函数 計算すると一瞬で $(\cosh x)' = \sinh x$, $(\sinh x)' = \cosh x$ が分かります。これも符号の付き方が微妙に違うだけで、三角函数とよく似ています。また今の計算から、 $y = \cosh x$ と $y = \sinh x$ は両方とも微分方程式 $y'' = y$ の解だとが分かります。これも、単振動の方程式 $y'' = -y$ の解が三角函数で得られることと似ています。

^{*12} 回転行列のことを知っているなら、それを使ってももちろん同じ計算ができます。

問 6 (b) の解答と加法定理 双曲線函数に関しても、三角函数と同様の加法定理が成り立ちます。ただし、符号の付き方が三角函数のときと少しだけ変わります。

$$\begin{aligned}\cosh x \cosh y + \sinh x \sinh y &= \frac{e^x + e^{-x}}{2} \frac{e^y + e^{-y}}{2} + \frac{e^x - e^{-x}}{2} \frac{e^y - e^{-y}}{2} \\ &= \frac{e^{x+y} + e^{x-y} + e^{-x+y} + e^{-(x+y)}}{4} + \frac{e^{x+y} - e^{x-y} - e^{-x+y} + e^{-(x+y)}}{4} \\ &= \cosh(x+y) \\ \sinh x \cosh y + \cosh x \sinh y &= \frac{e^x - e^{-x}}{2} \frac{e^y + e^{-y}}{2} + \frac{e^x + e^{-x}}{2} \frac{e^y - e^{-y}}{2} \\ &= \frac{e^{x+y} + e^{x-y} - e^{-x+y} - e^{-(x+y)}}{4} + \frac{e^{x+y} + e^{x-y} - e^{-x+y} - e^{-(x+y)}}{4} \\ &= \sinh(x+y)\end{aligned}$$

加法定理が成り立ちますから、当然ながら n 倍角の公式や和積・積和公式の類も三角函数のやつを少し修正するだけで得られます。これらの公式は、たまに双曲線函数の掛け算を積分する際に役立ちます。大学入試のときみたいに頑張って公式を暗記する必要まではありませんが、「三角函数と同じ公式がある」ことは頭の片隅に置いておきましょう。

双曲線函数と三角函数の関係 さて「三角函数と双曲線函数は似ている」ということを延々見てきたわけですが、なぜこんなにも似ているのでしょうか。それは定義域を複素数まで広げてみると分かります。前回 $e^{i\theta} = \cos \theta + i \sin \theta$ という式を紹介しました。そして双曲線函数は指数函数を使って定義されていますから、 $e^{i\theta}$ の式を使えば、双曲線函数に複素数を放り込めます。「何で指数函数に複素数を入れていいの」は全く説明していませんが、深いことは気にしないでおきましょう。実際にやってみると

$$\cosh i\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} = \cos \theta, \quad \sinh i\theta = \frac{e^{i\theta} - e^{-i\theta}}{2} = i \sin \theta$$

というように、なんと三角函数が出てきます。つまり複素数 $z \in \mathbb{C}$ に対し $(e^z \pm e^{-z})/2$ という函数を考えたとき^{*13}、この函数は実軸上では双曲線函数に、虚軸上では三角函数 (の i 倍) に化けるのです。このように \cosh と \cos , \sinh と \sin の共通の親玉となる函数がいるので、似たような挙動を示していたのです。

4.2 逆双曲線函数

双曲線函数の逆函数を逆双曲線函数といい、三角函数の時と同様、頭に “arc” を付けて $\operatorname{arccosh}$ などと表します。これらの函数は、 \log と平方根で表せます。したがって微分も、今まで知っている知識だけでできます。

問 6 (c) の解答 $y = \cosh x = (e^x + e^{-x})/2$ とおくと、 $(e^x)^2 + 1 = 2ye^x$ である。よって e^x の 2 次方程式 $(e^x)^2 - 2ye^x + 1 = 0$ を解いて $e^x = y \pm \sqrt{y^2 - 1}$ を得る。よって $\operatorname{arccosh} y = x = \log(y \pm \sqrt{y^2 - 1})$ である。± の符号はそれぞれグラフの上の枝と下の枝に対応する。

同様に $y = \sinh x = (e^x - e^{-x})/2$ を e^x について解くと $e^x = y \pm \sqrt{y^2 + 1}$ を得る。ただし $e^x > 0$ なので − の符号は不適である。よって $x = \operatorname{arsinh} x = \log(y + \sqrt{y^2 + 1})$ が得られる。 ■

問 6 (d) の解答 地道に計算すると、次のようになる。

$$(\operatorname{arccosh} x)' = \frac{1 + \frac{2x}{2\sqrt{x^2-1}}}{x + \sqrt{x^2-1}} = \frac{1}{\sqrt{x^2-1}}, \quad (\operatorname{arsinh} x)' = \frac{1 + \frac{2x}{2\sqrt{x^2+1}}}{x + \sqrt{x^2+1}} = \frac{1}{\sqrt{x^2+1}}$$

逆函数の微分法を使っても良い。 $y = \cosh x$ のとき $(\operatorname{arccosh} y)' = 1/(\cosh x)' = 1/\sinh x = 1/\sqrt{y^2 - 1}$ となる。 ■

^{*13} 一般の複素数 $z = a + i\theta$ に対しては、 $e^z := e^a e^{i\theta}$ と定めます。

数理科学基礎 (線形代数学) 第 3 回 座標空間と数ベクトル

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 5 月 7 日 1 限

1 情報の調べ方

1.1 文献調査の一般論

数学的なことではありませんが、今回の「vector の語源を調べる」という問題に対して不適切な答案が多かったの
で、文献調査に関する最低限の基礎を述べておきます。きっと他の授業でも同じようなことを教わると信じています
が、念のため一読しておいてください。

まず調査に使った文献を明記しないのは問答無用でアウトです。今回は「調査すること」自体が課題ですから、どの
ような答案であってもどこから引用されていることは分かります。ですが他のレポートで、出典を書かずに他人の
文章を取りこんだら、それは剽窃と呼ばれることになります。他人の文章を引っ張るときは「引用」という決まった作
法に従い、自分の文章でないことを明らかにした上で、「誰がどこに書いたものなのか」を確実に分かるようにしてく
ださい^{*1}。「たかが授業のレポートごときで」などと甘く見てはいけません。事が極限まで大きくなると、STAP の人
がやらかしたように「学位論文の 1 つの章が丸ごとコピペ」という事態になるのですから^{*2}。

次に、出典を明記するにしても信頼できる文献をきちんと選んで引用しましょう。特に Wikipedia は便利なもので、
キーワードをインターネットでググると頻繁にヒットします。ですが Wikipedia の記事は専門家が書いているとは限
りません。大抵の場合 Wikipedia の記事には出典が付いていますから、引用する場合は元の文献にきちんと当たり、孫
引き (引用の引用) にならないようにしましょう。引用でなくとも [要出典] タグが付いた場所など、出典がはっきりし
ない部分を鵜呑みにしてはいけません。

Wikipedia 以外の web サイトでも同じです。「公的機関の web サイトだから信頼できる」等の特別な事情がない限
り、インターネット上の文書は「誰だか知らない人が書いたもの」に過ぎません。しばらく昔に Yahoo!知恵袋を使っ
て京都大学の受験生がカンニングしたことが話題になりましたが、その際 Yahoo!知恵袋で「英文和訳問題の答」を教
えた人は、翻訳ソフトを使っていたそうです^{*3}。この程度の低品質な情報も平然と転がっているのですから、インター
ネット上で調べた情報を使うときは、その情報が信頼できるという裏付けを取ってください。

実際 “vector” の語源を一つとっても、インターネット上には不正確な情報があふれていることが垣間見えます。
せっかくの機会ですから、少し気合を入れて情報を調べてみましょう。

1.2 “vector” の語源

ネット上の記事 最終的には信頼できる文献を探さねばいけないとはいえ、調査の最初の段階ではインターネットを
使うのが楽です。手始めに「vector 語源」などでググってみましょう。見つかったサイトの記述から、関連個所をいく
つか引用してみます^{*4}。

- <http://www.nn.iij4u.or.jp/~hsat/techterm/vector.html>

辞書を紐解くと、数学用語の vector は 1843 年に数学者 William Rowen Hamilton (ハミルトン)(1805–1865)

^{*1} <http://www.bengo4.com/topics/1332/> に、著作権の視点からの分かりやすい説明があります。

^{*2} http://stapcells.blogspot.jp/2014/02/blog-post_2064.html

^{*3} <http://www.47news.jp/CN/201102/CN2011022801000578.html>

^{*4} 以下での web サイトの引用は、いずれも 2015/5/14 に閲覧したときの情報に基づいています。

がラテン語 *vectum* から作ったとある。この *vectum* は「運ぶ」という意味であり、ギリシャ語の ‘ὄσος’ が語源。更に遡ると *vehēre* に行き着くらしい。サンスクリットのヴァーハナは北インドで子供を守る女神の乗り物を意味するということである。

- <http://hooktail.sub.jp/vectoranalysis/restudyVector1/>
vector という言葉は、1843 年に Hamilton がラテン語の *vectum* から作った造語ですが、*vectum* とは運ぶもの、という意味です。更に遡ると、サンスクリット語で女神の乗り物を意味するヴァーハナと同根のようです。
- http://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1485274904
語源はラテン語です。ラテン語の意味は ‘carrier’ 「運び屋」です。
- <http://ja.wikipedia.org/wiki/ベクトル>
ベクトルは ドイツ語: *Vektor* ^{*5} に由来し、ベクターは 英語: *vector* に由来する。...(中略)... *vector* は「運ぶ」を意味するラテン語: *vehere* に由来し、18 世紀の天文学者によってはじめて使われた

これらの記事から、どうもラテン語の “*vehere*” とか “*vectum*” といった単語と関係があるようです。また Hamilton という数学者が関わっている気配も見えてとれます。ところが *vector* という言葉が生まれた時期について、1843 年および 18 世紀という、2 種類の食い違う記述が見られます。どちらが正しいのでしょうか？ もう少し Wikipedia を調べてみましょう。<http://ja.wikipedia.org/wiki/空間ベクトル> を見てみます。

ハミルトンは 1846 年に四元数の複素数における実部と虚部に相当するものとしてスカラーとベクトルという用語を導入した (今日の用法とは異なることに注意されたい):

この記事は「ベクトルという用語を、Hamilton が 1846 年に導入した」と書いてあります。18 世紀でも 1843 年でもありません。ふと某名探偵の「真実はいつもひとつ!」という言葉が脳裏をよぎります。何が真実なのでしょう？

書籍や論文の調査 こういう時は、きちんとした文献の出番です。まず英単語としての語源を調べるなら、やはり使うべきは辞書でしょう。研究社から出ている『新英和大辞典 第 6 版』^{*6}を引いてみると、p. 2727 に大体こんなことが書いてあります^{*7}。この辞書は「18 世紀に *vector* が誕生」という説を支持するようです。

vector は元々ラテン語で ‘carrier’ という意味の単語。そして *vector* は *vehere* の過去分詞形 *vectus* から派生している^{*8}。初出は 1704 年^{*9}。

一方 Wikipedia の「空間ベクトル」の記事には、Hamilton が 1846 年に導入したという事実がきちんと出典付で書いてあります。それによれば “London, Edinburgh & Dublin Philosophical Magazine 3rd series 29 27” が出典だそうです。皆さんはまだ見慣れないかもしれませんが、これは学術雑誌と呼ばれるものです。ですからこの文献を直接当たれば、1846 年に “*vector*” という言葉を Hamilton が使ったか確認できますね。

ここで皆さんの中に「1846 年の文献なんてどこで探すんだよ」と思った人がいるかもしれません。最近は便利なもので、なんと昔の学術雑誌がスキャンされてインターネット上に公開されていたりするのです。実際、今回の学術雑誌のタイトルでググれば見つかります^{*10}。その上、英語版 Wikipedia で Hamilton の記事 ^{*11} を見てみると、References

^{*5} (引用者注) これはスペルミスです。ドイツ語での「ベクトル」は “*Vector*” ではなく “*Vektor*” になります。

^{*6} 辞書のチョイスは個人的な好みによるものです。ちゃんとした辞書なら、何でも OK です。受講生の皆さんの中には、大修館書店の『ジーニアス英和大辞典』を引用した人もいました。

^{*7} まるっきりこのままではありません。辞書の方では、凡例記号で色々省略がされています。

^{*8} (引用者注) 水谷智洋編『改訂版 羅和辞典』(研究社) の記述と一致することを確認しました。また *vectus* は「動詞 *vehō* の不定法現在形」だそうです。ラテン語を読めないで、この辺の事情を深く調べるのは諦めました。

^{*9} (引用者注) ただし凡例によると、初出は「利用可能な文献による限りの初出例の年代」ということで、絶対的なものではない」とのこと。

^{*10} <http://www.biodiversitylibrary.org/bibliography/58679#/summary> にある “ser.3:v.31 (1846:July-Dec)” です。

^{*11} http://en.wikipedia.org/wiki/William_Rowan_Hamilton

の節に “David R. Wilkins’s collection of Hamilton’s Mathematical Papers”^{*12} というリンクがあります。なんと Hamilton の論文を、全てまとめて (しかも $\text{T}_\text{E}\text{X}$ で打ち直して！) web サイトを作ってくれた人がいるのです。そこで <http://www.maths.soton.ac.uk/EMIS/classics/Hamilton/OnQuat.pdf> の 16 ページを見ると

“... the algebraically imaginary part, being geometrically constructed by a straight line, or radius vector, which has, in general, for each determined quaternion, a determined length and determined direction in space, may be called the vector part ...”

という記述が見られます。Wikipedia の「空間ベクトル」の記事で引用されてたやつですね。これは 1846 年 7 月の記事らしいので、確かに 1846 年に Hamilton が vector という用語を使ったことは間違いありません^{*13}。

2 種類の問題 ここまでで、2 つ確実に分かったことがあります。

- 新英和大辞典によると、vector はラテン語の「運び屋」という意味の単語で、初出は 1704 年である
- Hamilton は 1846 年に vector part という用語を導入した

これらの事実は疑いようがありません。そこで浮かび上がってくるのは「vector という英単語が誕生したのは 1704 年で、それを数学の文脈で最初に用いたのが Hamilton だった」という可能性です。

このような「一般的に使われている単語が、数学用語として用いられるようになる」という事案は、現代でも日常的に起こっています。たとえば “cellular algebra” という数学の言葉がありますが、これは 1996 年、J. J. Graham と G. I. Lehrer という 2 人の数学者によって定義されたことがはっきりしています^{*14}。ですが cellular という単語は、1996 年よりずっとずっと昔から存在していました^{*15}。そこで vector についても、同じことが起きているのではないかと推測されるわけです。

インターネットで見つけた 2 つの web ページには、これらの記述と食い違う内容が書かれています。一番最初に挙げた web サイトを見直すと「辞書を紐解いたら」Hamilton が vector を作ったと書かれているそうです。ところが参考文献として挙げられている、三省堂の『カレッジクラウン英和辞典』^{*16}および岩波書店の『岩波理化学辞典 第 5 版』を実際に紐解いても、ベクトルの語源とか意味だけが書いてあって、「Hamilton が作った」とはどこにも書いてありません。したがって、一番最初に挙げた web サイトの著者は vector について「言葉が作られたこと」と「言葉を数学に持ち込んだこと」を混同していると推測されます^{*17}。また、1843 年に出版された Hamilton の論文全部を検索しても vector という単語は出ず、1844 年にやっと “radius vector” という単語が飛び出します。1843 年説も間違いでしたね。このようにインターネットには、著者の勘違い等に基づく不正確な情報が転がっていることがしばしばあるのです。

ついでに言えばページが作成された日時と内容を見るに、「物理のかぎしっぽ」の記事は 1 つ目の文献を剽窃した上、精査していないようにも見えます。「クロ」とまで言い切る証拠はないですが、皆さんはこのような疑わしい文章を書かないように心がけましょう。

vector の語源と初出は？ これまでの調査で、vector の語源がラテン語の動詞 “vectus” であること、元々の意味が「運び屋」であることはほぼはっきりしました。ですが初出がまだはっきりしません。せっかくだから、もう少し頑張ってみましょう。こういう時に岩波書店の『数学辞典』が役立ちそうですが、残念ながら今回は欲しい情報が載っていません

^{*12} <http://www.maths.soton.ac.uk/EMIS/classics/Hamilton/>

^{*13} さらに東京大学 OPAC で検索すると、本郷にある理学部の物理図書室に London, Edinburgh & Dublin Philosophical Magazine の現物が眠っていることが分かります。1846 年の文献でありながら、直接現物を確認することもできてしまうのです。

^{*14} 正確な文献情報は次の通り: J. J. Graham, G. I. Lehrer, “Cellular algebras”, *Inventiones Mathematicae* **123**, 1–34.

^{*15} 「Wikipedia によれば」1665 年に Robert Hooke が細胞を “cell” と名付け、1674 年に Leewenhoek が詳しく観察したそうです。気になる人はきちんと調べてみてください。

^{*16} カレッジクラウン英和辞典第 2 版は 1986 年に出版された古い辞書ですが、頑張って現物を調達しました。

^{*17} Hamilton が 1833 年に “Introductory Lecture on Astronomy” という本を出版していることを考えると、Hamilton が天文学の ‘vector’ という言葉を知らなかった可能性はほとんど無いでしょう。

んでした。そこで Wikipedia の「ベクトル」のページにある参考文献、Daniel Fleisch 著、河辺哲次訳『物理のためのベクトルとテンソル』（岩波書店、2013 年）の p. 1 を開いてみます。

ベクトル (vector) という言葉は「運ぶ」を意味するラテン語 *vehere* に由来します。この言葉は、太陽の周りを「運ばれる」惑星の運動を研究していた 18 世紀の天文学者たちによって、初めて使われました。

新英和大辞典と合わせると、初出 1704 年で、かつ最初に天文学で使われたとみてよさそうです。ですがこの本には「誰が使ったか」まで書いてないので、ここで Oxford English Dictionary (OED) の第 2 版を持ち出してみましょう。19 巻の p. 470 で “vector” を引くと、こんな風書いてあります。

†1. *Astr.* (See quot. 1704) Also *vector radius*, = *radius vector* RADIUS 3 e. *Obs.* 1704 J. HARRIS *Lex. Techn.* I. s.v., A Line supposed to be drawn from any Planet...

これでようやく、初出文献が分かりました。J. Harris という人の “Lex. techn.” という本だそうです。略されているため正確なタイトルが分からないので、東京大学 OPAC で検索してみましょう。「著者 Harris, 出版年 1704」で検索をかけると何も出ないのですが、他大学の図書まで含めた検索をかけると、放送大学図書館の蔵書がヒットします。そのレコードで著者とタイトルが John Harris, “Lexicon technicum, or, An universal English dictionary of arts and sciences” (1704) だと確定します。これを見れば、1704 年に本当に vector が登場したか分かります。

あとはどうやってこの本を調べるか、という問題が残ります。東京大学内の図書館に所蔵がありません。一つには、東京大学の図書館を経由して放送大学付属図書館に文献の複写を依頼する手があります。また国立国会図書館の郵送複写サービス等も利用可能です。ところが今回「国立国会図書館サーチ」*18に書籍タイトルを突っ込むと、なんとこの本のスキャン画像が出てきます：<http://jairo.nii.ac.jp/0108/00002692> この頭文字 “V” のページを見れば、“Vector” の項目に

A Line supposed to be drawn from any Planet moving round a Center, or the Focus of an Ellipsis, to that Center or Focus, is by some Writers of the New Astronomy, called the Vector; because 'tis that Line by which the Planet seems to be carried round its Center, and with which it describes proportional Areas in proportional Times.

と説明が書いてあります。確かに 1704 年に “vector” が登場したことがはっきりします*19。

ちなみに OED には「数学の “vector” は、Hamilton が 1846 年に使った」とも書かれていて、そこでは上で挙げた Hamilton の論文が引用されています。頑張れば、独自の調査で OED の初出年代にたどり着くこともできるんですね。

1.3 教訓

このように、普段「ネットでちょこっと調べればいいや」と思いがちな情報検索も、頑張ろうとすると結構時間と手間がかかるものなのです。ただ、ここまで読んだ皆さんは「何か無駄が多くなかったか？」と思うかもしれません。それは正しいです。実際 web サイトの記述を検証するために余分な手間がかかっているかもしれませんが、それを差し引いても

- 一番最初に OED と『新英和大辞典』をセットで調べてみる
- 出てきた参考文献の原本を、東京大学/国会図書館の OPAC や Hamilton の業績コレクションを使って検索する

という順で作業してしまえば、もっと手早く正確な情報にたどり着けたはずです。また今回、「昔の書物がオンラインで見られること」が非常に重要な役割を果たしました。ですから、今回の例を教訓に

*18 <http://www.ndl.go.jp/index.html> の左上です。NDL-OPAC とは別なので注意してください。

*19 ちなみに vector の初出については、Harris の本よりさらに遡ろうとしてみました。しかし Harris の本にある “the New Astronomy” がどこにあるのか、天文学専攻の友人に聞いても良く分からなかったのです。諦めました。

- 情報検索「使える」文献や手段 (今回は OED など) を知っておくこと
- オンラインで見られる文献を頑張って探すこと

の大事さを認識しておいてください。

最後に、これまでの“vector”の語源の話は全て検証可能です。ですからもし TA が嘘をついているとすれば、挙げた文献に当たればそれを見抜くことができます。今回そこまでは皆さんにお任せしますが、先生や TA が言ったことも必要に応じて検証せねばならないし、もし検証不可能なことがあれば皆さんは考え、質問をしなければいけません。そのことを忘れないでください。

2 空間内における直線と平面の取り扱い

皆さんは既に、平面 \mathbb{R}^2 内の直線の扱いについては十分習熟しているはずです。たとえば 2 点を通る直線の方程式を求めるとか、直線の法線ベクトルを求めるといった計算は、すぐにできるでしょう。同じようなことを 1 次元上げてもできるようになろう、というのが今回のお題の一つです。

まず最初に、平面を決定するための条件を確認しておきます。平面は

- 3 つの相異なる点
- 1 つの点と、平面を張る 2 本の 1 次独立なベクトル
- 1 つの点と、1 本の法ベクトル

が与えられれば定まります。これは頭の中で想像すればすぐ分かります。また、平面の表し方には

- 方程式による表示
- パラメータ表示

という 2 通りのやり方がありました。ですから

- 与えられた平面の決定条件から、平面の方程式やパラメータ表示を求めること
- 平面の決定条件のうちどれか 1 つから、残りの 2 つの条件を導くこと
- 平面の方程式とパラメータ表示をいったりきたりすること

がスムーズにできなければいけません。

それから、今回の問題では「2 次元版と 3 次元版」が並列されていることに気づいた人が多いと思います。以下で見るように、問題の次元が 1 個変わっても、解き方はほとんど変わりません。そして皆さんが後で学ぶ「線型代数」とは、そのような次元によらず同じように使えるテクニックを身に着けることだと言えます。そのような「何で同じように問題が解けるのか」という原理に関しても、注意してみてください。

2.1 直線と平面のパラメータ表示

まず最初に、直線のパラメータ表示を復習しましょう。空間内の直線は、通る点 $x_0 = (x_0, y_0, z_0)$ と方向ベクトル ${}^t(a, b, c)$ が^{*20}与えられれば決まります。そのとき直線上の点 x に対し $x - x_0$ は ${}^t(a, b, c)$ と平行なので、適当な実数 $t \in \mathbb{R}$ を用いて $x - x_0 = t {}^t(a, b, c)$ と書けます。逆に全ての实数 t に対して $x = x_0 + t {}^t(a, b, c)$ は直線上の点です。こうして $x = x_0 + t {}^t(a, b, c)$ ($t \in \mathbb{R}$) が直線のパラメータ表示を与えることが分かりました。平面の場合も同様です。

^{*20} 記号 t は、縦ベクトルを横ベクトルに直す記号です。縦ベクトルを使うと紙面上で面積を消費するので、印刷物ではこんな書き方をすることがしばしばあります。また「別に横ベクトルでもいいじゃん」と思うかもしれませんが、後で行列の話をするとき、縦ベクトルと横ベクトルを区別する必要が出てきます。なので以下、空間内のベクトルは転置記号をつけて縦ベクトルにします。

通る点 x_0 と平面を張るベクトル u, v が与えられると、 $x = x_0 + au + bv$ ($a, b \in \mathbb{R}$) がパラメータ表示です。

さて今度は、2 点 p_1, p_2 が与えられたとします。この 2 点を通る直線方向ベクトルは $p_2 - p_1$ なので、直線のパラメータ表示は $x = p_1 + t(p_2 - p_1) = (1-t)p_1 + tp_2$ となります。したがってベクトル p_1 と p_2 を係数が 1 になるよう足し合わせることで、直線の方程式が得られます。書き方を変えると、 $sp_1 + tp_2$ ($s+t=1$) という表示もできます。

平面についても同じように、 p_1, p_2, p_3 を通る平面のパラメータ表示は $x = p_1 + s(p_2 - p_1) + t(p_3 - p_1) = (1-s-t)p_1 + sp_2 + tp_3$ ($s, t \in \mathbb{R}$) あるいは $x = sp_1 + tp_2 + up_3$ ($s+t+u=1$) となります。

これらの式は結構利用価値が高いので、覚えておいてください。

2.2 平面の方程式と法ベクトル

次に、平面の法ベクトルについて考えてみましょう。平面は、通る点と法ベクトルがあれば一通りに決定されます。そこで $x_0 = {}^t(x_0, y_0, z_0)$ を通り $n = {}^t(a, b, c)$ を法ベクトルとする平面を考えます。すると平面上の任意の点 x に対し、 $x - x_0$ は n と垂直になります。したがって

$$0 = n \cdot (x - x_0) = a(x - x_0) + b(y - y_0) + c(z - z_0)$$

です。この式で $d := -ax_0 - by_0 - cz_0$ とおけば、平面上の点 x は方程式 $ax + by + cz + d = 0$ を満たします。

逆に方程式 $ax + by + cz + d = 0$ が与えられたとき、この方程式を満たす点全体の集合は平面になります。まず $(a, b, c) \neq (0, 0, 0)$ なので、 $ax + by + cz + d = 0$ の解が一つは存在します。たとえば $a \neq 0$ なら、 $x_0 := {}^t(-d/a, 0, 0)$ とすれば良いです。そして $x = {}^t(x, y, z)$ が方程式の解だとすれば、 $ax + by + cz + d = 0$ なので

$$0 = ax + by + cz + d = {}^t(a, b, c) \cdot x - {}^t(a, b, c) \cdot x_0 = {}^t(a, b, c) \cdot (x - x_0)$$

となります。こうして全ての解ベクトルは ${}^t(a, b, c)$ と垂直なので、方程式 $ax + by + cz + d = 0$ の定める集合は平面だと分かりました。

この事実を踏まえれば、問 6 と問 7 が解けます。

問 6, 7 の解答 平面 $\Pi: ax + by + cz + d = 0$ ^{*21} の法ベクトルは ${}^t(a, b, c)$ である。また点 (p, q, r) が与えられたとき、この点を通り方向ベクトルが ${}^t(a, b, c)$ であるような直線 ℓ のパラメータ表示は ${}^t(p, q, r) + \alpha {}^t(a, b, c)$ ($\alpha \in \mathbb{R}$) である。いま ${}^t(a, b, c)$ は平面 Π の法ベクトルだから、平面 Π と直線 ℓ は 1 点で交わる。すなわち平面 $ax + by + cz + d = 0$ にベクトル ${}^t(p, q, r) + \alpha {}^t(a, b, c)$ が乗るような $\alpha \in \mathbb{R}$ が唯一存在する。この α を用いて、 (p, q, r) から平面までの距離は $\|\alpha {}^t(a, b, c)\| = |\alpha| \sqrt{a^2 + b^2 + c^2}$ と書ける。他方 ${}^t(p, q, r) + \alpha {}^t(a, b, c)$ が方程式 $ax + by + cz + d = 0$ を満たすので

$$a(a\alpha + p) + b(b\alpha + q) + c(c\alpha + r) + d = 0 \quad \text{すなわち} \quad \alpha(a^2 + b^2 + c^2) = -(ap + bq + cr + d)$$

よって求める距離は

$$|\alpha| \sqrt{a^2 + b^2 + c^2} = \frac{|ap + bq + cr + d|}{\sqrt{a^2 + b^2 + c^2}}$$

で与えられる。この式で $c = r = 0$ と置けば、平面内の場合の式が得られる。 ■

与えられた点を通る平面 これまでの議論で、与えられた 3 点を通る平面を求める方法が分かりました。

- まずその方程式を $ax + by + cz + d = 0$ とおき
- この方程式に 3 点の座標を代入して、 a, b, c の連立 1 次方程式を得る
- 得られた方程式を解く

^{*21} Π は、ギリシャ文字 π (パイ) の大文字です。

とすれば OK です。ただ、ここで出てくる方程式は 3 変数の連立 1 次方程式なので、一般の場合に解くのは面倒です。問 2～問 5 のように、通る点の座標に 0 がたくさん出てくる場合は解くのが楽ですが、そうでない場合は

- パラメータ表示を使う
- 後で説明するベクトルの外積を使い、法線ベクトルを求めてから方程式を得る

など別の手を使う方が賢いでしょう。

問 2, 3 の解答 平面の方程式を $\alpha x + \beta y + \gamma z = \delta$ とおき、この式に $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$ を代入して、方程式

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$$

が直ちに得られる。この式で z/c の項を落とせば、直線の式が得られる。 ■

問 4, 5 の解答 原点を通るので、方程式の定数項は 0 である。あとは座標 $(1, 0, a)$ と $(1, 0, b)$ を代入すれば、求める式は $z = ax + by$ と分かる。平面上の直線の場合は、同様にして $y = ax$ が求まる。 ■

2.3 平面の方程式と決定条件

問 8, 10 の解答 2 点 (p_1, q_1) と (p_2, q_2) を通る直線がただ一つに決まる条件は $(p_1, q_1) \neq (p_2, q_2)$ である。同様に 3 点 (p_1, q_1, r_1) , (p_2, q_2, r_2) , (p_3, q_3, r_3) を通る平面がただ一つに決まる条件は、3 点が同一直線上にないことである。 ■

代数的な考察 この問題の結果自体は、直感的にはほとんど明らかです。そもそも Euclid の『原論』にも、公理として「相異なる 2 点を通る直線がただ 1 つ存在する」と書かれています。が、ここでちょっと視点を変え「直線がただ一つに決まる」という事実を、代数的に捉えてみましょう。

平面 \mathbb{R}^2 内の直線 $ax + by + c = 0$ が点 (p_1, q_1) と (p_2, q_2) を両方通るとします。このとき $ap_1 + bq_1 + c = 0$ および $ap_2 + bq_2 + c = 0$ という式が成り立ちます。この 2 つの式の差を取れば

$$a(p_1 - p_2) + b(q_1 - q_2) = 0$$

という式が得られます。逆にこの式を a と b の満たすべき方程式だと思えば、この 0 でない解の一つ一つが (p_1, q_1) と (p_2, q_2) の両方を通る直線に対応することになります。ただし 0 でない任意の実数 $\alpha \in \mathbb{R}$ に対し、方程式 $ax + by + c = 0$ と $\alpha ax + \alpha by + \alpha c = 0$ は同じ直線を表します。ですから $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の解 (a, b) に 0 でない定数 α をかけたとき、 (a, b) と $(\alpha a, \alpha b)$ は同じ直線に対応します。

いま $(p_1, q_1) \neq (p_2, q_2)$ だとしたら、 $(p_1 - p_2)a + (q_1 - q_2)b = 0$ は、原点を通る (a, b) 平面の直線を表します。したがって、 $(0, 0)$ 以外のどの二つの解も定数倍でうつります。ゆえに直線はただ一つに定まります。ところが $(p_1, q_1) = (p_2, q_2)$ なら、全ての (a, b) が解となってしまいます。これらの解のうちには定数倍でうつり合わないものが無数にありますから、 (p_1, q_1) と (p_2, q_2) の両方を通る直線は無数に存在します。

このように「相異なる 2 点 (p_1, q_1) , (p_2, q_2) を通る直線がただ一つ存在する」という条件が「方程式 $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の $(0, 0)$ でない解が、定数倍を除いてただ一つに決まる」という代数的な言葉に言い換えられたわけです。若干言い方がまどろっこしいですが、少しして線型代数の勉強をすると「方程式 $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の解空間が 1 次元」という、もっとすっきりした表現が使えるようになります。

空間の場合も話は同じです。平面の方程式 $ax + by + cz + d = 0$ に通る 3 点の座標を代入すると、 a, b, c, d に関する 3 つの連立 1 次方程式が得られます。変数が 4 つで式が 3 つですから、この連立方程式の解はただ 1 つに決まらず、解を動かす自由度が残ります。ただ 0 でない定数を (a, b, c, d) にかけても平面自体は変わらないので、「1 次元分の自由度」が「平面が 1 個に決まること」に相当します。

2.4 直線の方程式

ここまで使いませんでしたが、直線の方程式についても少しだけ触れておきます。パラメータ表示された直線 $\ell: {}^t(x_0, y_0, z_0) + t^t(a, b, c)$ が与えられたとします。このとき直線上の点 $x = {}^t(x, y, z)$ は、ある実数 $t \in \mathbb{R}$ に対し

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}$$

を満たします。この式を t について解くことで、直線上の点が満たす式

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

が得られるのです。ちなみにこの式は a, b, c の中に 0 が混ざっているとマズいですが、そういう時も使えるような式を得るなら、分母を払っておけば大丈夫です。 a, b, c の中に、少なくとも 1 つは 0 でないものがあります。それが a だった場合は、今の方程式を

$$b(x - x_0) = a(y - y_0), c(x - x_0) = a(z - z_0)$$

と書き換えれば問題ありません。逆にこの方程式が与えられていれば、 $t = (x - x_0)/a$ とおくことで、 $a \neq 0$ と合わせて $y - y_0 = bt$, $z - z_0 = ct$ が得られます。こうして直線の方程式からパラメータ表示を復元することもできました。

ここで、直線の方程式は

$$\begin{cases} \Pi_1: bx - ay - (bx_0 - ay_0) = 0 \\ \Pi_2: cx - az - (cx_0 - az_0) = 0 \end{cases}$$

という 2 本の平面の方程式を連立して得られています。これらの平面は

- いずれも点 (x_0, y_0, z_0) を通り
- それぞれの法ベクトル $(b, -a, 0)$ と $(c, 0, -a)$ が共に ℓ の方向ベクトル (a, b, c) と垂直

なことに気を付けましょう。したがって平面 Π_1, Π_2 は共に直線 ℓ を含んでいることが分かります。すなわち直線 ℓ の方程式は、 ℓ を 2 枚の平面 Π_1, Π_2 の共通部分として表しているわけです。

このような手は、他の場合にも応用できます。空間内の図形をつかむには、その図形をいくつかの分かりやすい図形の共通部分として表しておいて、それぞれのパーツの方程式を作れば良いのです。そうすれば連立方程式の形で、求める図形を表す方程式が得られます。

3 ベクトルの 1 次独立性と行列式

3.1 ベクトルの 1 次独立性

1 次独立性 空間内の n 本のベクトル v_1, \dots, v_n が 1 次独立であることの定義は、 $\alpha_1 v_1 + \dots + \alpha_n v_n$ となる $(\alpha_1, \dots, \alpha_n)$ が $(0, \dots, 0)$ 以外に存在しないことでした。1 次独立でなく、たとえばもし $\alpha_1 = 1$ だったなら、 $v_1 = -\alpha_2 v_2 - \dots - \alpha_n v_n$ というように、 v_1 が他のベクトルのスカラー倍と足し算で表されます。なので 1 次独立性は、大体「ベクトルたちがバラバラの方向を向いている」ということに相当します。また 1 次独立でないことを 1 次従属と呼びます。これを踏まえた上で、問 9, 11 の解答を見てみましょう。

共線条件と共面条件 (問 9, 11 の解答) 2 点 $(p_1, q_1), (p_2, q_2)$ を通る直線のパラメータ表示は、 $s^t(p_1, q_1) + t^t(p_2, q_2)$ ($s + t = 1$) で与えられます。したがって 3 点が同一直線上にあるためには、 $(p_3, q_3) = s^t(p_1, q_1) + t^t(p_2, q_2)$, $s + t = 1$ という連立方程式が解を持つか、 $(p_1, q_1) = (p_2, q_2)$ であることが必要十分です。

ただ、この方程式は「ベクトルのパラメータ表示」と「1次方程式」がバラバラに出てきてちょっと扱いづらいです。そこで平面ベクトルの話ですがわざと次元を上げて、 ${}^t(1, p_1, q_1)$, ${}^t(1, p_2, q_2)$ というベクトルを考えてみます。そうすると $s {}^t(1, p_1, q_1) + t {}^t(1, p_2, q_2) = {}^t(s+t, sp_1+tp_2, sq_1+tq_2)$ となるので、求める条件が「 $s {}^t(1, p_1, q_1) + t {}^t(1, p_2, q_2) = {}^t(1, p_3, q_3)$ となる s, t が存在すること」でまとまります*22。この式は移項すれば

$$s {}^t(1, p_1, q_1) + t {}^t(1, p_2, q_2) - {}^t(1, p_3, q_3) = \mathbf{0}$$

となります。すなわち「 ${}^t(1, p_1, q_1)$, ${}^t(1, p_2, q_2)$, ${}^t(1, p_3, q_3)$ が1次従属であること」が求める条件です。

空間の場合も、議論は全く同じです。求める共面条件は「ベクトル ${}^t(1, p_1, q_1, r_1)$, ${}^t(1, p_2, q_2, r_2)$, ${}^t(1, p_3, q_3, r_3)$, ${}^t(1, p_4, q_4, r_4)$ が1次従属であること」となります。

3.2 ベクトルの外積

問12の解答 ${}^t(p_1, q_1)$ に垂直なベクトルは、明らかに $\alpha {}^t(-q_1, p_1)$ ($\alpha \neq 0$) である。実際内積を取れば、 ${}^t(p_1, q_1) \cdot \alpha {}^t(-q_1, p_1) = \alpha(-p_1q_1 + q_1p_1) = 0$ となる。

ベクトルの外積 (問13の解答) 平面 \mathbb{R}^2 の場合はこのように垂直なベクトルはすぐ求まるわけですが、空間の場合はどうなるでしょうか。すなわち空間の場合、2本のベクトル u, v が与えられたとき、両方のベクトルに垂直なベクトルを作るにはどうすれば良いでしょうか。実は「ベクトルの外積」というものを使うと、それを求めることができます。

$u = {}^t(p_1, q_1, r_1)$, $v = {}^t(p_2, q_2, r_2)$ に対し、 u と v の外積は $u \times v := (q_1r_2 - q_2r_1, r_1p_2 - r_2p_1, p_1q_2 - p_2q_1)$ という式で与えられるベクトルです。そうすると

$$u \cdot (u \times v) = p_1(q_1r_2 - q_2r_1) + q_1(r_1p_2 - r_2p_1) + r_1(p_1q_2 - p_2q_1) = 0$$

$$v \cdot (u \times v) = p_2(q_1r_2 - q_2r_1) + q_2(r_1p_2 - r_2p_1) + r_2(p_1q_2 - p_2q_1) = 0$$

となるので、確かに $u \times v$ は u および v と垂直になることが分かります。今は天下りな定義を与えましたが、これから行列式を使い、もう少しマシな説明をします。

3.3 行列式

ベクトルの1次独立性 / 従属性や外積の話は「行列式」というものを知っていると良く理解できます。本格的に勉強するのは後ですから、ここではラフに話をしておきます。

色々な定義がありますが、 \mathbb{R}^2 内の2本の縦ベクトル u, v の張る平行四辺形の符号付き面積*23を $\det(u, v)$ と書きます。また \mathbb{R}^3 内の3本の縦ベクトル u, v, w の張る平行六面体の符号付き体積*24を $\det(u, v, w)$ と書きます。このとき、 \det をベクトルの成分で書き下すと

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc, \quad \det \begin{pmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ r_1 & r_2 & r_3 \end{pmatrix} = p_1 \det \begin{pmatrix} q_2 & q_3 \\ r_2 & r_3 \end{pmatrix} - q_1 \det \begin{pmatrix} p_2 & r_3 \\ r_2 & p_3 \end{pmatrix} + r_1 \det \begin{pmatrix} p_2 & p_3 \\ q_2 & q_3 \end{pmatrix}$$

となることが知られています。

さて平面内に2本のベクトルがあったとき、これらが1次独立だったら平行四辺形が張れます。逆に1次従属だったら、平行四辺形を張ろうとすると面積が0になります。したがって u, v が1次独立性であることと $\det(u, v) \neq 0$ が同値になります。3次元以上の場合でも同様に \det を使えば、すごくややこしくなりますが「ベクトルが1次独立であるか」を成分計算する式が得られるのです。これを使えば共線・共面条件をたった1本の式で書き下せます。

*22 この「一度次元を上げてから、見やすいように書き直す」という手法は、他の場面でもしばしば用いられます。

*23 符号は、 u から v に向かう方向が正の向きであるときに + とし、負の向き - とします。

*24 符号は、 u, v, w がこの順で右手系をなすときに +, 左手系をなすときに - とします。

さらに 3 次元の \det を注意深く観察すると、

$$\det \begin{pmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ r_1 & r_2 & r_3 \end{pmatrix} = {}^t(p_1, q_1, r_1) \cdot {}^t \left(\det \begin{pmatrix} q_2 & q_3 \\ r_2 & r_3 \end{pmatrix}, -\det \begin{pmatrix} p_2 & p_3 \\ r_2 & r_3 \end{pmatrix}, \det \begin{pmatrix} p_2 & p_3 \\ q_2 & q_3 \end{pmatrix} \right)$$

のように、 \det が 2 本のベクトルの内積で書けることが分かります。ここで 2 次元のときの \det を見ると、列の入れ替えで符号が (-1) 倍されると分かります。よって今の式で \det の前の $-$ が消せて

$$\det \begin{pmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ r_1 & r_2 & r_3 \end{pmatrix} = {}^t(p_1, q_1, r_1) \cdot {}^t \left(\det \begin{pmatrix} q_2 & q_3 \\ r_2 & r_3 \end{pmatrix}, \det \begin{pmatrix} r_2 & r_3 \\ p_2 & p_3 \end{pmatrix}, \det \begin{pmatrix} p_2 & p_3 \\ q_2 & q_3 \end{pmatrix} \right)$$

となります。この右辺に出てくる、成分に \det が入ったベクトルが、さっき外積と呼んだものです。したがって、ここまでの議論で $\det(\mathbf{u} \ \mathbf{v} \ \mathbf{w}) = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ が分かりました。この式を使えば、なぜ $\mathbf{v} \times \mathbf{w}$ が \mathbf{v} と直交するのか一目瞭然ですね。 $\mathbf{v} \cdot (\mathbf{v} \times \mathbf{w})$ はベクトル \mathbf{v} , \mathbf{v} , \mathbf{w} の張る平行六面体の体積ですが、この平行六面体はぺっちゃんこなので、どう見ても体積 0 です。こうして内積の値が 0 となり、 \mathbf{v} と $\mathbf{v} \times \mathbf{w}$ が直交すると分かります。 \mathbf{w} についても同じです。

このようにベクトルの 1 次独立性を \det と関連付けることによって、ベクトルが一段と便利な道具になります。また \det を用いてベクトルの情報を得る過程には、 \det が持つ諸々の計算公式が重要な役割を果たします。そういう事も追々勉強しますので、楽しみにしててください。

4 残りの問題

問 14 の解答 3 つの点が同一直線上にないことが必要十分条件です。このとき 3 点は三角形をなすから、外接円が 3 点全てを通過します。続いて円の方程式を $(x-a)^2 + (y-b)^2 = r^2$ とおく。これに $(x, y) = (p_1, q_1), (p_2, q_2), (p_3, q_3)$ を代入すると、 a, b, r に関する 3 変数の連立一次方程式が得られる...のですが、これを真面目に解いてもあまり徳しません。 $\triangle ABC$ の 3 点の位置ベクトルを $\mathbf{a}, \mathbf{b}, \mathbf{c}$ とすると、外心が $\frac{\sin 2\angle A \mathbf{a} + \sin 2\angle B \mathbf{b} + \sin 2\angle C \mathbf{c}}{\sin 2\angle A + \sin 2\angle B + \sin 2\angle C}$ と表せます。こっちの式の方が使い勝手が良いと思います*25。

問 15 の解答 存在しないことを背理法で示す。全ての頂点が格子点である正三角形があったとする。このとき

- 成分が整数であるようなベクトルに沿って格子点を平行移動したものは、再び格子点になる
- 図形を平行移動すると、元の図形と合同になる

という事実があるので、正三角形の頂点の 1 つが原点であると仮定して差支えない。そこで正三角形の頂点を $(0, 0)$, (a, b) , (c, d) ($a, b, c, d \in \mathbb{Z}$) とする。

いま正三角形の一辺の長さは $\sqrt{a^2 + b^2}$ である。したがって正三角形の面積は

$$\frac{1}{2} \times \sqrt{a^2 + b^2} \times \frac{\sqrt{3}}{2} \sqrt{a^2 + b^2} = \frac{\sqrt{3}(a^2 + b^2)}{4}$$

である。一方、座標を使って計算すると面積が $|ad - bc|/2$ と求まる。したがって

$$\frac{\sqrt{3}(a^2 + b^2)}{4} = \frac{|ad - bc|}{2} \quad \text{より} \quad \sqrt{3} = \frac{2|ad - bc|}{a^2 + b^2} \in \mathbb{Q}$$

となる。これは $\sqrt{3}$ が無理数である事実に反する。 ■

*25 頂点の取り方を工夫していくらか計算してみましたが、座標ベースではあんまり綺麗な式や面白い式が得られませんでした。すいません。

S2 ターム

線形代数学

