

# 解析学ノート

久島広幸 (h2@hisasima.jp)

2015年12月3日  
revised 2019年3月31日

# 目次

プロローグ	4
<b>第 1 章 近代解析流儀での導関数の導出</b>	<b>7</b>
1.1 極限の算術	7
1.1.1 極限の定義と記号	7
1.1.2 極限の算術	8
1.2 微分係数の定義と「微分可能」という言葉	9
1.3 関数が連続であるということ	10
1.4 関数の連続性と「微分可能」との関係	10
1.5 導関数の導出を支える数学的事実	11
1.6 導関数の導出	11
1.6.1 導関数の導出	11
1.6.2 典型的な関数の導関数	11
1.6.2.1 巾乗の導関数	11
1.6.2.2 三角関数の導関数	12
1.6.2.3 指数関数の導関数	13
1.7 導関数関連の定理	13
1.7.1 基本的な定理	13
1.7.2 合成関数の導関数の定理	15
1.7.3 逆関数の導関数の定理	17
1.8 高階導関数	20
1.9 多変数関数と偏微分記号	20
<b>第 2 章 古典解析（無限小解析）流儀での微分法と導関数の導出</b>	<b>22</b>
2.1 微分量とその計算ルール（公理系）	22
2.1.1 変化量と微分量	22
2.1.2 微分量に関する公理	22
2.1.3 独立変化量と、従属変化量（関数）	23
2.1.4 基本的な定理	23
2.2 微分量の公理系を支える数学的事実	25
2.3 微分量の具体的な計算例	25
2.4 導関数	26
2.5 高階導関数と独立変化量（独立変数）	28
2.5.1 高階導関数	28
2.5.2 独立変化量（独立変数）の導入	28
2.5.3 微分記号 $d$ の記法の拡張	29

2.6	多変数関数と全微分	30
<b>第3章</b>	<b>Taylor 展開</b>	<b>31</b>
3.1	近代解析的導出	31
3.1.1	Maclaurin 展開	31
3.1.2	Taylor 展開	32
3.2	古典解析的導出	32
<b>第4章</b>	<b>積分</b>	<b>35</b>
4.1	不定積分	35
4.1.1	不定性	35
4.1.2	計算規則を与える定理	36
4.1.3	導関数での計算規則	37
4.2	定積分	38
4.2.1	定義か定理か	38
4.2.2	Riemann 和の極限 ( $dx$ の寄せ集めのイメージ)	39
4.2.3	定積分の性質	40
4.3	不定積分と定積分をつなぐもの	41
4.4	積分の例題	41
<b>第5章</b>	<b>変分法</b>	<b>42</b>
5.1	変分の世界の構築	42
5.1.1	汎関数と変分量	42
5.1.2	微分と変分	43
5.2	変分量とその計算ルール (公理系)	43
5.2.1	変分量の記述	43
5.2.2	変分の公理系	44
5.2.3	基本的な定理	45
5.2.4	変分量の公理系を支える数学的事実	46
5.2.5	基本的な関数の変分の例	47
5.3	微分と変分の関係	48
5.3.1	形式と操作	48
5.3.2	微分と変分をつなぐもの	50
5.4	積分と変分の関係	51
5.5	閑話休題 — 関数概念と汎関数 —	52
<b>第6章</b>	<b>微分と変分の実践</b>	<b>55</b>
6.1	長方形の等周問題	55
6.1.1	2次関数	55
6.1.2	微分の応用	55
6.1.3	ベクトルの利用 (未定乗数法一歩手前)	55
6.1.4	未定乗数法微分版	56
6.2	確率変数のエントロピー	56
6.2.1	離散確率版	57

6.2.2	連続確率版	58
6.3	2点間の最短経路	59
6.4	Fermatの原理(幾何光学)	62
6.4.1	反射の法則	62
6.4.2	屈折の法則	64
6.4.3	補足	65
付録A	合成関数の導関数の導出はなぜこんなにややこしいのか	66
	エピローグ	69

## プロローグ

はじめは、「微分と変分の順序が交換できる」ことを示すだけの目的で筆をとった（というか、キーボードを叩いた）。その時書いたものは

関数  $f(x)$  が各点  $x$  においてすこしずれたものを  $g(x)$  としよう。各点でのずれは、 $h(x) = g(x) - f(x)$  と表せる。 $x$  での微分演算をすれば、

$$h'(x) = g'(x) - f'(x)$$

基本はこれだけで、あとはこれから用いる「ずれ」をあらわす  $\delta$  記号に惑わされずにいけばよい。

点  $x$  における任意の関数  $p(x)$  の「ずれ」を  $\delta p(x)$  と表すことにしよう。すると上の例では、

$$g(x) = f(x) + \delta f(x)$$

となって、結果、

$$h(x) = \delta f(x),$$

$$h'(x) = (\delta f(x))'$$

である。この論法を  $g'(x)$  にもあてはめる。 $g'(x)$  は  $f'(x)$  がずれたものであるとして<sup>a</sup>

$$g'(x) = f'(x) + \delta f'(x),$$

これより、

$$h'(x) = \delta f'(x)$$

となる。 $h(x)$  はその前でももとめておいた。両者は同じものであるべきだから、

$$(\delta f(x))' = \delta f'(x).$$

Leipniz 流にかけば、

$$\frac{d}{dx} \delta f(x) = \delta \frac{d}{dx} f(x).$$

<sup>a</sup>むむ。この論旨はちょっと都合よすぎるか？危うさを感じる。

というものである。脚注に書いてあるとおり、当時から少しばかりの危うさを感じていた。でもしばらく放っておいた。それに「変分だといって面倒なことをやらなくても、この導出で美しいじゃないか」という自惚れもあった。けれど、脚注に書いたくらいだから、どこかでひっかかっていたのだろう。ある時読み直してみて、

脚注したその危うさの理由がわかった。

危うさはどこからくるかというところ、天真爛漫に  $g'(x) = f'(x) + \delta f'(x)$  とした所からくる。  $f'(x)$  の「ずれ」を  $\delta f'(x)$  と記すことには問題はない。ただ  $f'(x) + \delta f'(x)$  が  $g'(x)$  であるという保証はない。いいかえれば、これを保証することこそが証明すべきものなのである。

なのに、先に  $g'(x) = f'(x) + \delta f'(x)$  であるとし、かつ、導関数の計算により  $g'(x) = f'(x) + (\delta f(x))'$  を導いておいて、「両者が等しい」という論理から微分と変分の順序交換の結論を導出している。つまり証明すべき順序交換を暗黙に実行していることになっている。論点先取である。これがいけなかった。

やはり、きちんと「変分」の概念を明確にしてそこから始めないとだめなのだと考えなおし、新たに書き始めたのが本書である<sup>1</sup>。

第1章「近代解析流儀での導関数の導出」は、近代的流儀の初等的（高校数学的）な説明を試みている。極限の算術を所与のものとして微分係数を導き、微分係数を関数化して導関数を導くという道筋である。そしてこれらの道具立てのもとで微分に関する定理を導出している。合成関数の導関数は少しややこしいので、付録Aに詳細をまとめた。

この道中で、もうひとつの数学的事実として、代表的な超越関数にまつわる性質をも所与のものとした。  $\theta$  が微小な時  $\sin \theta \approx \theta$  であるという事柄は、 $\sin \theta$  を Taylor 展開して説明されることがままある。けれども、この  $\sin \theta$  の Taylor 展開自身が  $\sin \theta$  の導関数をもとにしており、その導関数導出には  $\sin \theta \approx \theta$  という事実が必要である、という循環した構造になっている。なのでうまくない。したがってここでは、数学的事実として無条件で認めることにした<sup>2</sup>。

第2章「古典解析（無限小解析）流儀での微分法と導関数の導出」は微小な微分量についての公理を定め、無限小解析的な理路を追った。2階導関数の考察から「独立変化量」が導入され、独立変化量の微分が定量、すなわち、 $d(dx) = 0$  が導かれる。微分量にまつわる演算操作は圧巻である。なおここでも、超関数にまつわる微分量については、数学的事実として所与のものとした。

本章は、高瀬著『dx と dy の解析学—オイラーに学ぶ』[1]に寄る所大である。とても勉強になる本であった。

第3章「Taylor 展開」は、文字通り Taylor 展開についてのまとめである。古典解析流儀での導出方法については、ちょっと疑義があるけれども、手並みとしては鮮やかである。

第4章「積分」は、不定積分積分と定積分についてのまとめである。特に、定積分における「平均値の定理」の威力には驚く。

第5章「変分法」が、筆者がいちばん整理しなかった（理解を深めたかった）変分についての章である。変分といえども微小量であるということから、第2章の理路を参考にし、同一の公理系を用いてまとめてみた。《変数と関数》という対象から《関数と汎関数》という対象への移行である。そして関数が微分と変分の橋渡しになる。

<sup>1</sup>結果として、 $f'(x)$  の変分は  $\delta f'(x)$  であることになるのだが、それは変分概念を明確にした結果そうなるのであって、この証明もどきのように利用すべきではないのである。

<sup>2</sup>とはいえ、実際の所、何かから導出したいという思いはある。例えば  $\lim_{h \rightarrow 0} \sin \theta / \theta = 1$  であるが、単位円をもとにして三角形と扇型の面積の関係を利用して（ついでに  $\epsilon$ - $\delta$  も利用して）求めることができるのではと想像している。

第 6 章「微分と変分の実践」で、実際的な計算を練習した。Fermat の定理は懐かしい。

正直、筆者は、物理の世界（応用数学の世界？）での解析の計算のおおらかさについて、実用にあたっては確かに問題はないのだけれども、どうもじっくりと来ない部分があるなあ、と感じていた。微分量の縦横無尽な操作についての一抔の不安の解消、そして、変分法についての理論武装もどきができないか、と考えていた。高瀬 [1] の p.210 に『今日では、理論的枠組みは近代解析に求め、実際の計算では古典解析を踏襲している。このちぐはぐさが露呈して議論の循環が起こるのである』とある。実際の計算のおおらかさはここからくるのかと合点がいった。第 1 章と第 2 章、そして第 5 章で、筆者なりの「近代解析」と「古典解析」の整合性を試みたのだが、うまくいっているだろうか？それとも「甘い」だろうか？

なお、言うまでもないことであるが、本書の記述においては、筆者の理解の浅さ・誤解・執筆時の息切れなどによるおかしい所が、必ずあるはずである。気づいたときには折にふれ修正していきたいと考えている。

# 第1章 近代解析流儀での導関数の導出

微分と導関数の概念の初等的（高校数学的）導出について、まとめてみる。その際、四則演算、巾乗などの通常の算術は前提される。つぎに関数概念も認める。代数的関数の他、超越関数も存在している。実数の連続性ももちろん前提される。こころへんは数学的常識の範疇にいて、所与のものとしてあつかっていく。本章での微分と導関数の対象は、関数である。出発点は極限の算術であり、それをもちいた「微分係数」である。極限の算術と、いくつかの「数学的事実」から、微分と導関数を演繹する、というのがここで説明する初等的導出の論理構造である。

なお、取り扱う関数については、独立変数は実数であるものとする（つまり実変数関数）。

## 1.1 極限の算術

$x$  が  $a$  に近づくの時の関数  $f(x)$  の振る舞いを観察するという数学的行為がまず基本にある。この時の  $f(x)$  のあり様には様々な場合があるだろうが、 $f(x)$  がある一定値に近づいてくる、という場合があることが観察される（我々が馴染んでいるものは圧倒的にこのケースである）。極限とはこの特殊な場合を定義（名前付け）したものである。

ここでは、初等的に、「近づく」という言葉でもって内容を説明していく。さらに極限の算術については、それを天下りに認めてしまう。きちんと（非初等的？）にやるには、あの  $\varepsilon$ - $\delta$  による概念の噛み砕きが必要となる。

### 1.1.1 極限の定義と記号

ある実数点  $a$  に対して、『独立変数  $x$  を  $a$  に近づけたとき（ただし  $x$  は  $a$  にはならない）、 $f(x)$  がある一定の値に近づく』、という状況が存在する場合を考える<sup>1</sup>。『 $f(x)$  が一定の値に近づく』という事柄が、これから述べる極限の定義の本質である。

$x$  を  $a$  に近づけることを  $x \rightarrow a$  と表そう。実数の数直線を思い浮かべると、 $x \rightarrow a$  には2通りの向きがあることがわかる。 $a$  より大きい方（大きい方を右側に置くのが普通だろうか）から向かう場合と、 $a$  より小さい方（左側）から向かう場合の2つである。大きい方から近づけた時に  $f(x)$  が近づくその一定の値を

$$\lim_{x \rightarrow a^+} f(x), \quad \lim_{x \searrow a} f(x), \quad \lim_{x \downarrow a} f(x)$$

などと、方向感がでるような記号を用いて書く。同様に、小さい方から近づけた時のものを

$$\lim_{x \rightarrow a^-} f(x), \quad \lim_{x \nearrow a} f(x), \quad \lim_{x \uparrow a} f(x)$$

などと書く<sup>2</sup>。そして、この記号によって表される「一定の値」を、それぞれ「右極限」「左極限」と定義する。

<sup>1</sup>この場合、 $a$  は  $f(x)$  の定義域に含まれていなくても良い。たとえば、 $f(x) = x^2$  ( $0 < x < 2$ ) のような関数であっても、 $\lim_{x \rightarrow 2} f(x) = \lim_{x \rightarrow 2} x^2 = 4$  である。ただし、この関数の定義では  $f(2)$  は存在しない（意地悪関数）。

<sup>2</sup>これらの3種の使い分けは何分気分的なもので、それぞれで意味合いが変わることはもちろんない。



右極限, 左極限がおのおの存在し, かつ, 同一の値である場合には, それを「極限」と定義し,

$$\lim_{x \rightarrow a} f(x)$$

と書く.

「一定の値」に近づく時それを  $\lim_{x \rightarrow a} f(x)$  と書いた. これを, 『 $\lim_{x \rightarrow a} f(x)$  が存在する』という風に言うことも多い. もちろんこれはある一定値を示している (右極限, 左極限の場合も同じ).

「一定の値」に近づかないのならば, 右極限, 左極限, 極限それらは皆存在しない.

## 1.1.2 極限の算術

以下極限の場合のみを記すけれど, これらは右極限や左極限の時も同様に成り立つものである.

$\lim_{x \rightarrow a} f(x) = F$ ,  $\lim_{x \rightarrow a} g(x) = G$  ( $F, G$  は実数定数) であるとき (極限值が存在しているという前提が重要)

$$\lim_{x \rightarrow a} \alpha f(x) = \alpha \cdot \lim_{x \rightarrow a} f(x) = \alpha F \quad (\alpha \text{ は定数}) \quad (1.1.1)$$

$$\lim_{x \rightarrow a} \{f(x) \pm g(x)\} = \lim_{x \rightarrow a} f(x) \pm \lim_{x \rightarrow a} g(x) = F \pm G \quad (1.1.2)$$

$$\lim_{x \rightarrow a} \{f(x) \cdot g(x)\} = \lim_{x \rightarrow a} f(x) \cdot \lim_{x \rightarrow a} g(x) = FG \quad (1.1.3)$$

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow a} f(x)}{\lim_{x \rightarrow a} g(x)} = \frac{F}{G} \quad (\text{ただし, } G \neq 0 \text{ の時}). \quad (1.1.4)$$

(1.1.1), (1.1.2) で示されている線型性は極限の定義から素直に受け入れられよう. (1.1.3), (1.1.4) については次のように説明 (証明) できる. まず  $FG - f(x) \cdot g(x)$  は

$$FG - f(x) \cdot g(x) = FG - f(x)G + f(x)G - f(x) \cdot g(x) = (F - f(x))G + f(x)(G - g(x))$$

とあらわせるので

$$\lim_{x \rightarrow a} \{FG - f(x) \cdot g(x)\} = \lim_{x \rightarrow a} \{(F - f(x))G + f(x)(G - g(x))\}$$

であり, (1.1.1), (1.1.2) から

$$\begin{aligned} (\text{右辺}) &= \lim_{x \rightarrow a} (F - f(x))G + \lim_{x \rightarrow a} f(x)(G - g(x)) = 0 \\ & \quad (x \rightarrow a \text{ のとき, } (F - f(x)) \rightarrow 0, (G - g(x)) \rightarrow 0 \text{ であるから}). \end{aligned}$$

したがって

$$\lim_{x \rightarrow a} \{FG - f(x) \cdot g(x)\} = 0 \iff FG = \lim_{x \rightarrow a} f(x) \cdot g(x).$$

同様な論法を用いると,

$$\frac{1}{G} - \frac{1}{g(x)} = \frac{g(x) - G}{Gg(x)}$$

であるから

$$\lim_{x \rightarrow a} \left\{ \frac{1}{G} - \frac{1}{g(x)} \right\} = \lim_{x \rightarrow a} \frac{g(x) - G}{Gg(x)} = 0 \iff \lim_{x \rightarrow a} \frac{1}{g(x)} = \frac{1}{G}.$$

これと上の結果を合わせると

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} f(x) \lim_{x \rightarrow a} \frac{1}{g(x)} = \frac{F}{G}.$$

$h$  にまったく関係のない関数を  $p(!h)$  と記すことにする。以下の算術は有用である。

$$\begin{aligned} \lim_{h \rightarrow 0} 1 &= 1 \\ \lim_{h \rightarrow 0} p(!h) &= p(!h) \cdot \lim_{h \rightarrow 0} 1 = p(!h) \\ \lim_{h \rightarrow 0} p(!h)h &= p(!h) \cdot \lim_{h \rightarrow 0} h = 0 \\ \lim_{h \rightarrow 0} p(!h)h^n &= p(!h) \cdot \lim_{h \rightarrow 0} h^n = 0 \quad (n \text{ は正の整数}) \end{aligned}$$

さらに、深淵な関係として

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{0}{h} &= 0 \\ \lim_{h \rightarrow 0} \frac{0}{h^n} &= 0 \quad (n \text{ は正の整数}) \end{aligned}$$

がある。極限の定義から、 $h$  はドンピシャと  $0$  にはならない。幾分か「距離」はある。なので、分子が  $0$  ならば無条件に極限も  $0$  になる<sup>3</sup>。

## 1.2 微分係数の定義と「微分可能」という言葉

極限の算術をもちいて、「微分係数」を定義する。もちろんこれも、関数の観察という数学的行為から導かれるものであるのだろうが、先の極限の場合のように観察結果から直ちにでてくるような筋合いのものではないだろう。ここに「微分係数」として述べるような量を思い浮かべるといふ行為には、ある種の思考の飛躍が感じられる<sup>4</sup>。とにかく、この「微分係数」は定義である。何かから導出されるという筋合いのものではない。その数学的実体の存在を感じ取り、理路を構成するために定義というものに昇華したというのが実状なのではないだろうか。もちろん関数によっては微分係数が存在しない（極限が定まらない）場合もありうる。

関数  $f(x)$  に対して、次のようにあらわされる一定値となる極限を、実数点  $a$  での「右微分係数」と定義（名前付け）し<sup>5</sup>、それを  $f'_+(a)$  とあらわす：

$$f'_+(a) := \lim_{x \downarrow a} \frac{f(x) - f(a)}{x - a}.$$

同様に、左微分係数は次の様に定義される：

$$f'_-(a) := \lim_{x \uparrow a} \frac{f(x) - f(a)}{x - a}.$$

<sup>3</sup>ではどのくらいの距離があるのか、というと、ほとんど  $0$ 。次章で述べる無限小解析での最初のハードルであり、かつ、非常に神秘的なものなのである。

<sup>4</sup>思考を飛躍させることができる、というのが才能とよばれるものなんだろうという気がする。

<sup>5</sup>「微分」というものをこれから考えて行くはずなのに、それに先立って「微分係数」という名前をつけている。論点先取気味ではあるが、結局関連が出てくるので、名前は名前と割り切って気にしないことにしよう。

右微分係数と左微分係数が等しい時、つまり、 $f'_+(a) = f'_-(a)$ の時、単に微分係数と呼んで、

$$f'(a) := \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

と表す。また、 $h := x - a$  とすることによりもとまる次の形式も重宝される：

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}.$$

この  $f'(a)$  が存在する時（つまり、微分係数が一定値として定まる時）、『関数  $f(x)$  は実数点  $a$  で微分可能である』、という。

### 1.3 関数が連続であるということ

$f(x)$  が、ある実数点  $a$  で連続であるとは、

$$\begin{aligned} & f(a) \text{ が存在する,} \\ & \lim_{x \rightarrow a} f(x) \text{ が存在する,} \\ & \lim_{x \rightarrow a} f(x) = f(a) \text{ である,} \end{aligned}$$

という事柄全てが成立することである。どれか一つ欠けても  $x = a$  では連続ではない。連続関数とは、定義域内の全ての  $a$  においてこれが成立する関数のことである。

### 1.4 関数の連続性と「微分可能」との関係

関数  $f(x)$  が実数点  $a$  で微分可能であれば、 $f(x)$  は実数点  $a$  で連続である。しかしながら、その逆は一般には成立しない。

ひとつ実例を見ておこう。 $f(x) = |x|$  を思い浮かべると、 $x = 0$  で連続ではある。ところが左右微分係数をみると、絶対値の性質から、

$$\begin{aligned} f'_+(0) &= \lim_{h \downarrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \downarrow 0} \frac{|h|}{h} = \lim_{h \downarrow 0} \frac{h}{h} = 1, \\ f'_-(0) &= \lim_{h \uparrow 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \uparrow 0} \frac{|h|}{h} = \lim_{h \uparrow 0} \frac{-h}{h} = -1, \end{aligned}$$

となり、一致しない。微分係数は存在しないので微分可能ではない。  
スローガンのように書けば、

$$(\text{微分可能}) \implies (\text{連続}).$$

## 1.5 導関数の導出を支える数学的事実

以下に述べる極限に関する等式は、微分と関わりなく成立する数学的事実である。必要な時は無条件に利用する<sup>6</sup>。

$$\sin x \text{ は連続関数であり } \lim_{h \rightarrow 0} \frac{\sin h}{h} = 1,$$

$$e^x \text{ は連続関数であり } \lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1.$$

さらに指数と対数についての次の関係も事実として利用する。

$$p = e^{\log_e p} = e^{\log p} = \exp(\log_e p) = \exp(\log p),$$

$$p = \log_e e^p = \log e^p.$$

対数の底が記述されない場合は、底が  $e$  であるとするのが通例である。したがって冗長的ではあるが両方併記した。

## 1.6 導関数の導出

### 1.6.1 導関数の導出

微分係数に対して、関数概念を適用したものが「導関数」である。言い換えれば、関数概念の適用によって微分係数から演繹されたものが「導関数」である。ゆえに、「導関数」は単なる定義とみなすべきではないだろう。微分係数というものを定義し、それを関数化したものとして導出されるもの、と見立てた方がよい。これが近代解析の流儀であると思われる。

関数  $f(x)$  に対して、実変数  $x$  がとりうるおのおのの実数点  $a$  にそれぞれ個別に微分係数  $f'(a)$  を対応させれば、それは「実数点に数を対応させる」という写像であり、関数そのものである。この関数を  $f'(x)$  とあらわし、「導関数」とよぶ。すなわち、

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}. \quad (1.6.1)$$

### 1.6.2 典型的な関数の導関数

#### 1.6.2.1 巾乗の導関数

まずはじめに、 $f(x) = x^2$  の導関数  $f'(x)$  をもとめてみよう。(1.6.1) から、

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \rightarrow 0} \frac{2hx + h^2}{h} = \lim_{h \rightarrow 0} \{2x + h\} = 2x.$$

同様に、 $f(x) = x^3$  の導関数は、

$$f'(x) = \lim_{h \rightarrow 0} \frac{(x+h)^3 - x^3}{h} = \lim_{h \rightarrow 0} \frac{3x^2h + 3xh^2 + h^3}{h} = \lim_{h \rightarrow 0} \{3x^2 + 3xh + h^2\} = 3x^2.$$

<sup>6</sup>三角関数  $\sin x$  は、 $\sin$  の図形的意味と単位円上での振る舞いをを基盤にすれば、その連続性はイメージしやすい。しかしながら、指数関数はそう簡単にはいかない。指数が整数や有理数の場合ならば、掛け算の繰り返し（巾乗）のイメージで足りるけれども、無理数の場合（例えば、 $e^x$  など）はどう考えればいいのか？「掛け算の繰り返し」といっても  $\pi$  回繰り返すというのはどういうことなのか？これらは本章の初等的という範囲からは、はみでている。

この形を、二項定理をもちいて一般化する． $n$  を自然数として二項定理から、

$$(x+h)^n = \sum_{i=0}^n {}_n C_i x^{n-i} h^i = \sum_{i=0}^n \binom{n}{i} x^{n-i} h^i = x^n + \binom{n}{1} x^{n-1} h + \binom{n}{2} x^{n-2} h^2 + \cdots + \binom{n}{n-1} x h^{n-1} + h^n$$

ともとまるので、 $f(x) = x^n$  の導関数は

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h} \\ &= \lim_{h \rightarrow 0} \frac{x^n + \binom{n}{1} x^{n-1} h + \binom{n}{2} x^{n-2} h^2 + \cdots + \binom{n}{n-1} x h^{n-1} + h^n - x^n}{h} \\ &= \lim_{h \rightarrow 0} \left\{ \binom{n}{1} x^{n-1} + \binom{n}{2} x^{n-2} h + \cdots + \binom{n}{n-1} x h^{n-2} + h^{n-1} \right\} \\ &= \binom{n}{1} x^{n-1} = n x^{n-1}. \end{aligned}$$

### 1.6.2.2 三角関数の導関数

次に、 $\sin x$  についてみてみよう．計算の途中で三角関数の加法定理を利用し、最後に  $\lim_{h \rightarrow 0} (\sin h/h) = 1$  という数学的事実を用いれば、

$$\begin{aligned} (\sin x)' &= \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} = \lim_{h \rightarrow 0} \frac{\sin x (\cos h - 1) + \cos x \sin h}{h} \\ &= \sin x \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} + \cos x \lim_{h \rightarrow 0} \frac{\sin h}{h} = \sin x \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} + \cos x \end{aligned}$$

とはなる．そしてここで  $\lim_{h \rightarrow 0} \{(\cos h - 1)/h\} = 1$  の取り扱いに困り、行き詰まる．この時に使えるテクニックとして  $h = 2k$  と置くという方法がある． $h \rightarrow 0$  のとき  $k \rightarrow 0$  であるから、加法定理と倍角性質などの三角関数の操作技術を駆使して、

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{k \rightarrow 0} \frac{\sin(x+2k) - \sin x}{2k} = \lim_{k \rightarrow 0} \frac{\sin x \cos 2k + \cos x \sin 2k - \sin x}{2k} \\ &= \lim_{k \rightarrow 0} \frac{\sin x (1 - 2 \sin^2 k) + 2 \cos x \sin k \cos k - \sin x}{2k} \\ &= \lim_{k \rightarrow 0} \frac{-2 \sin x \sin^2 k + 2 \cos x \sin k \cos k}{2k} \\ &= \lim_{k \rightarrow 0} \frac{2 \sin k (\cos x \cos k - \sin x \sin k)}{2k} \\ &= \lim_{k \rightarrow 0} \frac{\sin k}{k} \cos(x+k) \end{aligned}$$

をひとまず得る．ここで  $\lim_{k \rightarrow 0} (\sin k/k)$  と  $\lim_{k \rightarrow 0} \cos(x+k)$  はおのおの 1 と  $\cos x$  となる極限をもつので、極限の算術から

$$\lim_{k \rightarrow 0} \frac{\sin k}{k} \cos(x+k) = \lim_{k \rightarrow 0} \frac{\sin k}{k} \cdot \lim_{k \rightarrow 0} \cos(x+k) = \cos x$$

となる．つまり  $(\sin x)' = \cos x$ .

同様に、 $\cos x$  もみてみると

$$\begin{aligned}
 (\cos x)' &= \lim_{k \rightarrow 0} \frac{\cos(x+2k) - \cos x}{2k} \\
 &= \lim_{k \rightarrow 0} \frac{\cos x \cos 2k - \sin x \sin 2k - \cos x}{2k} \\
 &= \lim_{k \rightarrow 0} \frac{\cos x(1 - 2\sin^2 k) - 2\sin x \sin k \cos k - \cos x}{2k} \\
 &= \lim_{k \rightarrow 0} \frac{-2\cos x \sin^2 k - 2\sin x \sin k \cos k}{2k} \\
 &= \lim_{k \rightarrow 0} \frac{-2\sin k(\cos x \sin k + \sin x \cos k)}{2k} \\
 &= -\lim_{k \rightarrow 0} \frac{\sin k}{k} \sin(x+k) \\
 &= -\sin x
 \end{aligned}$$

となり、 $(\cos x)' = -\sin x$ .

### 1.6.2.3 指数関数の導関数

$e^x$  の導関数は、 $\lim_{h \rightarrow 0} \{(e^h - 1)/h\} = 1$  という数学的事実を用いて

$$(e^x)' = \lim_{h \rightarrow 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \rightarrow 0} \frac{e^x e^h - e^x}{h} = \lim_{h \rightarrow 0} \frac{e^x(e^h - 1)}{h} = e^x \cdot \lim_{h \rightarrow 0} \frac{e^h - 1}{h} = e^x$$

となる。

## 1.7 導関数関連の定理

導関数の計算をより一般の関数に拡大するために、ここで導関数にまつわる性質を述べた定理を証明しておこう。対象の関数は拡大されるが、留意する点は、様々な関数形が出てきても、その総体としての関数と独立変数の変化が導関数を規定するということである。例えば2つの関数の積  $f(x)g(x)$  の導関数は、 $F(x) = f(x)g(x)$  として

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h}$$

を考えるのであり、合成関数の場合には  $F(x) = f(g(x))$  として関数の総体を捉えて

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{f(g(x+h)) - f(g(x))}{h}$$

を考えるということである。

### 1.7.1 基本的な定理

次の定理はもっとも基本的な定理である：

$$\begin{aligned}
 f(x) = C &\implies f'(x) = 0 \quad (C \text{ は定数}), \\
 \{\alpha \cdot f(x) + \beta \cdot g(x)\}' &= \alpha \cdot f'(x) + \beta \cdot g'(x) \quad (\text{線形性. } \alpha, \beta \text{ は定数}), \\
 \{f(x) \cdot g(x)\}' &= f'(x) \cdot g(x) + f(x) \cdot g'(x).
 \end{aligned}$$

以下順次証明をしていく。

定理 1.7.1.  $f(x) = C \implies f'(x) = 0$  ( $C$  は定数)

証明. 導関数のそもそもの表式から始めて,

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{C - C}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0.$$

□

定理 1.7.2.  $\{\alpha \cdot f(x) + \beta \cdot g(x)\}' = \alpha \cdot f'(x) + \beta \cdot g'(x)$

証明. これも導関数のそもそもの表式から始めればよい.

$$\begin{aligned} \{\alpha \cdot f(x) + \beta \cdot g(x)\}' &= \lim_{h \rightarrow 0} \frac{\alpha \cdot f(x+h) + \beta \cdot g(x+h) - \{\alpha \cdot f(x) + \beta \cdot g(x)\}}{h} \\ &= \lim_{h \rightarrow 0} \left\{ \alpha \cdot \frac{f(x+h) - f(x)}{h} + \beta \cdot \frac{g(x+h) - g(x)}{h} \right\} \\ &= \alpha \cdot \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} + \beta \cdot \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} \\ &= \alpha \cdot f'(x) + \beta \cdot g'(x) \end{aligned}$$

□

定理 1.7.3.  $\{f(x) \cdot g(x)\}' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

証明. これも導関数のそもそもの表式から始めればよいのだが, 途中で  $f(x)g(x+h)$  を引いて加える (結果変化なし) というところが技巧的である.

$$\begin{aligned} \{f(x) \cdot g(x)\}' &= \lim_{h \rightarrow 0} \frac{f(x+h) \cdot g(x+h) - f(x) \cdot g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h) \cdot g(x+h) - f(x) \cdot g(x+h) + f(x) \cdot g(x+h) - f(x) \cdot g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{(f(x+h) - f(x)) \cdot g(x+h) + f(x) \cdot (g(x+h) - g(x))}{h} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{f(x+h) - f(x)}{h} \cdot g(x+h) + f(x) \cdot \frac{g(x+h) - g(x)}{h} \right\} \\ &= \lim_{h \rightarrow 0} \left\{ \frac{f(x+h) - f(x)}{h} \cdot g(x+h) \right\} + \lim_{h \rightarrow 0} \left\{ f(x) \cdot \frac{g(x+h) - g(x)}{h} \right\} \\ &= f'(x) \cdot g(x) + f(x) \cdot g'(x). \end{aligned}$$

□

定理 1.7.3 はいろいろと便利である. さきに二項定理で求めた巾乗関数の導関数も, この定理から導くことができる (若干頭の体操的).  $f(x) = x, g(x) = x$  とすると,  $x^2 = f(x)g(x)$  であり,  $f'(x) = g'(x) = (x)' = 1$  であるから,

$$(x^2)' = \{f(x)g(x)\}' = f'(x)g(x) + f(x)g'(x) = 1 \cdot x + x \cdot 1 = 2x.$$

次に,  $h(x) = x^2$  とすれば,  $x^3 = f(x)h(x)$  であり,  $h'(x) = 2x$  であるので

$$(x^3)' = \{f(x)h(x)\}' = f'(x)h(x) + f(x)h'(x) = x^2 + x \cdot 2x = 3x^2.$$

以下数珠繋ぎで、 $(x^n)' = nx^{n-1}$  となる。

「商の導関数」と言われるものについても同様にこの定理 1.7.3 からもとめることができる（もちろん基本通りにやってもできるだろうが、計算が面倒になりそうだ）。今、 $h(x) = f(x)/g(x)$ （ただし、 $g(x) \neq 0$ ）としよう。すると、 $f(x) = h(x)g(x)$  である。この両辺それぞれの導関数を求めると、

$$\text{(左辺)} = f'(x),$$

$$\text{(右辺)} = h'(x) \cdot g(x) + h(x) \cdot g'(x),$$

であり、これらは等しくなるべきなのだから、

$$h'(x) \cdot g(x) + h(x) \cdot g'(x) = f'(x).$$

ここで、 $h(x)$  をそもそもの  $f(x)/g(x)$  に戻す。すると、

$$\begin{aligned} h'(x) \cdot g(x) + h(x) \cdot g'(x) = f'(x) &\iff h'(x) \cdot g(x) + \frac{f(x)}{g(x)} \cdot g'(x) = f'(x) \\ &\iff h'(x) \cdot g(x) = f'(x) - \frac{f(x)}{g(x)} \cdot g'(x) \\ &\iff h'(x) \cdot g(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)} \\ &\iff h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{\{g(x)\}^2}. \end{aligned}$$

これも定理としておこう。

$$\text{定理 1.7.4. } \left(\frac{f(x)}{g(x)}\right)' = \frac{\{f'(x)g(x) - f(x)g'(x)\}}{\{g(x)\}^2}$$

証明. 上述のとおり。 □

特に、 $f(x) = 1$  とすれば、分数関数の導関数として、

$$\left(\frac{1}{g(x)}\right)' = -\frac{g'(x)}{\{g(x)\}^2}$$

を得る。ここで再び巾乗関数について応用してみよう。 $x^{-n} = 1/x^n$  ( $n$  は自然数) であったから、

$$(x^{-n})' = \left(\frac{1}{x^n}\right)' = -\frac{(x^n)'}{(x^n)^2} = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1}$$

となる。 $m = -n$  とすれば、これは負の整数  $m$  に対しても  $(x^m)' = mx^{m-1}$  が成り立つことを示している。いままでの  $n$  の場合と合わせれば、正の整数（自然数）と負の整数をあらわす  $l$  で  $(x^l)' = lx^{l-1}$  が成り立つことが導出されたことになる。ではひとつ抜け落ちていた  $l = 0$  の場合はどうか？ $l = 0$  の時は、 $(x^0)' = (1)' = 0$  であり、かつ、 $lx^{l-1}$  も  $0$  であるから、問題はない。よって全ての整数  $l$  において  $(x^l)' = lx^{l-1}$  となるのである。

## 1.7.2 合成関数の導関数の定理

定理を先にのべると、合成関数の導関数の定理は  $(f(g(x)))' = f'(g(x))g'(x)$  というものである。これを証明するには、すこしの工夫が必要になる。導関数の極限の形式による表現に対して、 $h$  の変化による振る舞いを観察して作用をあらわに書き出すという工夫である。その理路をみてみよう。



関数  $g(x)$  の導関数は

$$g'(x) = \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h}$$

であった。これは、

$h$  をどんどん小さくして 0 に近づけていけば、 $g'(x)$  と  $\frac{g(x+h) - g(x)}{h}$  の差はどんどん小さくなって 0 に近づく

という動き意味している<sup>7</sup>。上記の「差」は、 $h$  の値に応じて決まることは明らかだろう。その「差」を  $\epsilon(h)$  と記せば、

$$\frac{g(x+h) - g(x)}{h} - g'(x) = \epsilon(h) \iff g(x+h) - g(x) = g'(x)h + \epsilon(h)h.$$

きちんと導関数になるには、 $h \rightarrow 0$  の時に  $\epsilon(h) \rightarrow 0$  とならねばならない ( $\lim_{h \rightarrow 0} \epsilon(h) = 0$ )。逆に言えば、 $h \rightarrow 0$  の時 0 となるような関数  $\epsilon(h)$  であると諒解しておく。

同様なことを、今度は (変数名を変えて)  $f(u)$  にも適用する。紛らわしいので、 $h$  の代わりに  $t$  を用いる。また、差分を表す関数は同じ関数である必要はないので、 $\eta(t)$  とする (ただし、 $\lim_{t \rightarrow 0} \eta(t) = 0$  という性質を持たねばならないことは同じである)。すなわち、

$$\frac{f(u+t) - f(u)}{t} - f'(u) = \eta(t) \iff f(u+t) - f(u) = f'(u)t + \eta(t)t.$$

ここで、 $u$  を  $g(x)$  として本来の合成関数をみると、まず、

$$f(g(x)+t) - f(g(x)) = f'(g(x))t + \eta(t)t$$

である。技巧的だが、ここで、 $t$  に  $g(x+h) - g(x)$  を示していた  $g'(x)h + \epsilon(h)h$  を用いる、つまり  $t = g'(x)h + \epsilon(h)h$  とすると (見通しが悪くなるので、 $\eta(t)$  はそのまま)、

$$f(g(x) + g'(x)h + \epsilon(h)h) - f(g(x)) = f'(g(x))(g'(x)h + \epsilon(h)h) + \eta(t)(g'(x)h + \epsilon(h)h)$$

であり、左辺は  $f(g(x+h)) - f(g(x))$  であるから結局

$$f(g(x+h)) - f(g(x)) = f'(g(x))(g'(x)h + \epsilon(h)h) + \eta(t)(g'(x)h + \epsilon(h)h)$$

となる。さらに両辺を  $h$  で割れば、

$$\frac{f(g(x+h)) - f(g(x))}{h} = f'(g(x))(g'(x) + \epsilon(h)) + \eta(t)(g'(x) + \epsilon(h))$$

である。 $h \rightarrow 0$  の極限をとれば、

$$(\text{左辺}) = \lim_{h \rightarrow 0} \frac{f(g(x+h)) - f(g(x))}{h} = (f(g(x)))' \quad (\text{合成関数の導関数の定義そのもの}),$$

$$(\text{右辺}) = \lim_{h \rightarrow 0} \{f'(g(x))(g'(x) + \epsilon(h)) + \eta(t)(g'(x) + \epsilon(h))\} = f'(g(x))g'(x)$$

となる。なぜならば、 $\lim_{h \rightarrow 0} \epsilon(h) = 0$  と決めたものであり、 $\lim_{h \rightarrow 0} t = \lim_{h \rightarrow 0} \{g'(x)h + \epsilon(h)h\} = 0$  であるので  $h \rightarrow 0$  のとき  $t \rightarrow 0$ 、ゆえに  $\lim_{h \rightarrow 0} \eta(t) = \lim_{t \rightarrow 0} \eta(t) = 0$  となるからである<sup>8</sup>。以上から、

<sup>7</sup>これは、無限小量を用いる時によく用いられる言葉使いである。

<sup>8</sup>この極限の議論は少し甘い。付録 A でもうすこし詳細にみておいた。

定理 1.7.5.  $(f(g(x)))' = f'(g(x))g'(x)$

証明. 上述のとおり. □

例をひとつ見てみよう.  $e^{x^2}$  の導関数を計算してみる. 今,  $f(y) = e^y$ ,  $y = g(x) = x^2$  とすれば,  $e^{x^2} = f(g(x))$  という合成関数とみなせる. ここで,  $f'(y) = (e^y)' = e^y$  であるので  $f'(g(x)) = e^{g(x)}$  ( $y$  を置き換えただけ). また  $g'(x) = 2x$  である. したがって,

$$(e^{x^2})' = (f(g(x)))' = f'(g(x))g'(x) = e^{g(x)} \cdot 2x = e^{x^2} \cdot 2x.$$

$y = g(x) = x \log a$  ( $a > 0$ ) の場合はとても面白い結果を得る.  $g'(x) = \log a$  であるから,

$$(e^{x \log a})' = (f(g(x)))' = f'(g(x))g'(x) = e^{g(x)} \cdot \log a = e^{x \log a} \cdot \log a$$

となる.  $e^{x \log a} = (e^{\log a})^x = a^x$  であることを利用すれば

$$(a^x)' = a^x \log a.$$

$a = e$  とすれば, お馴染みの  $(e^x)' = e^x$  でもある<sup>9</sup>.

### 1.7.3 逆関数の導関数の定理

逆関数の基本的な例をみてみよう<sup>10</sup>. 関数  $f$  を,  $x$  を  $x \geq 0$  なる実数として

$$x \xrightarrow{f} x^2$$

<sup>9</sup>そうではあるけれども, そもそも  $(e^x)' = e^x$  であることを利用して (前提として)  $(a^x)' = \log a \cdot a^x$  が導き出されたのであるから, これは当然のことを確認しているにすぎない.

<sup>10</sup>時に混乱することがあるので, 整理しておこう. 混乱の元は, 関数, 逆関数における独立変数と, グラフ用紙へ描く時の  $x$  軸,  $y$  軸という記号にその原因があるのではないかと推察している (個人的に).

まず,  $f, g$  などを関数とし,  $s, r$  を実数をあらわすものとしよう. そして, 次のような関数の働きの図式が成り立つ場合を考える.

$$s \xrightarrow{f} r \xrightarrow{g} s$$

この図式を, 数式で記述すると,

$$r = f(s), \quad s = g(r), \quad \text{まとめて書くと} \quad s = g(f(s))$$

である. これを成立させる  $g$  を  $f$  の逆関数と呼び, 場合によっては  $f^{-1}$  と書く. ここまでは混乱は少ない. さてこの関数  $f, g$  をグラフ用紙に書こうとすると, 通常は関数の変数 (独立変数) を  $x$  として  $x$  軸に, 関数の値 (従属変数) を  $y$  として  $y$  軸に書くことになる. ゆえに,

$$y = f(x), \\ y = g(x) = f^{-1}(x)$$

としてプロットすることになるが, この立場では, 上で述べた関係  $s = g(f(s))$  がおもてにみえてこない. なかんずく, さきの図式の  $s, r$  の役割をそれぞれ  $x, y$  で表記し直してみると,

$$x \xrightarrow{f} y \xrightarrow{g} x \quad (y = f(x), \quad x = g(y), \quad x = g(f(x)))$$

となる.  $g$  に関しては  $x = g(y)$  なのである. このグラフ用紙での書き方 (つまり, 独立変数を  $x$  とする習慣) と逆関数関係を示すときの独立変数の扱いの違いが混乱のもとである. グラフ化の際は, 逆関数  $g$  において,  $x$  と  $y$  を入れ替えていると考えてもいい. 関数を比較する場合には, このグラフ用紙に書く場合の立場をとって, 独立変数を同じもの (記号) に揃えるのである. そこに留意さえすれば, 混乱は取れる (背後には, 関数の変数をあらわす文字はなんでも構わないという基本が横たわっている).

また実際にグラフを書いてみればわかる事柄であるが,  $y = f(x)$  と  $x = f^{-1}(y)$  は同じグラフになる. そして,  $y = f(x)$  と  $y = f^{-1}(x)$  は  $y = x$  で対象なグラフになる.

とする関数であるとする。  $x \geq 0$  と取り決めたので、単射である。したがって、その逆関数を  $f^{-1}$  とすれば、

$$x \xrightarrow{f} x^2 \xrightarrow{f^{-1}} x$$

となる。関数をあらわに書けば、  $f(x) = x^2, f^{-1}(x) = \sqrt{x}$  となる。また指数関数と対数関数については、

$$x \xrightarrow{f} e^x \xrightarrow{f^{-1}} x$$

であり、  $f(x) = e^x, f^{-1}(x) = \log x$  なのである。もちろん  $f^{-1}(e^x) = x$ 。このような、ある関数  $f$  の逆関数  $f^{-1}$  の導関数にたいして、次の定理が成立する。

定理 1.7.6.  $(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}$

証明.  $f, f^{-1}$  は逆関数の関係であるので、変数を  $x, y$  とすれば次の関係が成立している<sup>11</sup>。

$$y \xrightarrow{f} x \xrightarrow{f^{-1}} y \quad (x = f(y), \quad y = f^{-1}(x))$$

$h, t$  を実数として、  $h = f(y+t) - f(y)$  とする。明らかに、  $f(y+t) = f(y) + h = x + h$ 。この時、逆関数の関係から、

$$y + t = f^{-1}(f(y+t)) = f^{-1}(x + h)$$

が成り立つ。  $f^{-1}(x)$  の導関数にこの結果をあてはめていけば、

$$\begin{aligned} (f^{-1}(x))' &= \lim_{h \rightarrow 0} \frac{f^{-1}(x+h) - f^{-1}(x)}{h} = \lim_{h \rightarrow 0} \frac{y+t-y}{f(y+t) - f(y)} = \lim_{h \rightarrow 0} \frac{t}{f(y+t) - f(y)} \\ &= \lim_{h \rightarrow 0} \frac{1}{\frac{f(y+t) - f(y)}{t}} \end{aligned}$$

$h$  を  $h = f(y+t) - f(y)$  と決めたのだから、  $h \rightarrow 0 \iff t \rightarrow 0$ 、したがって、

$$(f^{-1}(x))' = \lim_{t \rightarrow 0} \frac{1}{\frac{f(y+t) - f(y)}{t}} = \frac{1}{f'(y)} = \frac{1}{f'(f^{-1}(x))}.$$

□

この定理 1.7.6 をもちいて対数関数の導関数をもとめてみよう。ここで

$$p = e^{\log p} = \exp(\log p), \quad q = \log e^q$$

という数学的事実を再確認しておく。さて今、便宜上、指数関数を  $f$ 、その逆関数である対数関数を  $f^{-1}$  としておくと、定理から

$$(f^{-1}(x))' = (\log x)' = \frac{1}{f'(\log x)} = \frac{1}{\exp(\log x)} = \frac{1}{x}$$

<sup>11</sup>ここで、 $x$  を真ん中においたのは、逆関数の独立変数として  $x$  を使いたかったからである。前の脚註で述べた「グラフ用紙書き（言い換えれば数学的慣習）」のゆえである。 $y$  はなんでも良かったが、 $x$  の相棒といえば  $y$  であろう。

となる。さらに、合成関数の導関数についての定理 1.7.5 と組み合わせれば、

$$(\log f(x))' = \frac{f'(x)}{f(x)} \quad (1.7.1)$$

が得られる<sup>12</sup>。

再度巾乗関数について、有理数乗を考えてみよう。まず  $x^{1/m}$  について考える<sup>13</sup>。これは  $m$  乗の逆関数であるとみて、逆関数の図式を書いてみると、

$$y \xrightarrow{f: m \text{ 乗}} x \xrightarrow{f^{-1}: 1/m \text{ 乗}} y \quad (x = f(y) = y^m, \quad y = f^{-1}(x) = x^{1/m})$$

である。 $f'(y) = my^{m-1}$  であることに留意して、逆関数の導関数の定理を使うと、

$$(x^{1/m})' = \frac{1}{f'(f^{-1}(x))} = \frac{1}{f'(x^{1/m})} = \frac{1}{m(x^{1/m})^{m-1}} = \frac{1}{mx^{1-(1/m)}} = \frac{1}{m}x^{(1/m)-1}.$$

次に、 $x^{n/m}$  の導関数を計算しよう。これは  $f(y) = y^n$ ,  $y = g(x) = x^{1/m}$  という合成関数と見立てられるので、合成関数の導関数の定理を使う。すなわち、

$$(x^{n/m})' = f'(g(x))g'(x) = ny^{n-1} \frac{1}{m}x^{(1/m)-1} = \frac{n}{m}(x^{1/m})^{n-1}x^{(1/m)-1} = \frac{n}{m}x^{(n-1)/m}x^{(1/m)-1} = \frac{n}{m}x^{(n/m)-1}.$$

ゆえに、今までの巾乗についての結果をまとめれば、 $\alpha$  が自然数、整数、有理数のどの場合でも、

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

となるのである<sup>14</sup>。

巾乗関数の導関数の導出の締めくくりとして、無理数乗の場合を考察しよう。対数関数の導関数が重要な役割を果たす。 $q$  を無理数として、 $x^q$  を考えるのだが、そもそも「無理数乗」とは何なのか？有理数乗までの導出においては、勘定できる数  $n$  をつかって「 $n$  回掛ける」ということを巾乗の基本として対処してきた。しかしここでは無理数である。 $2^{\sqrt{2}}$  とは「何回」になるのか？回数の勘定は諦めて、対数による表現を使おう。 $x > 0$  の時、 $x = e^{\log x}$  であった。したがって、

$$x^q = (e^{\log x})^q = e^{q \log x}$$

と変形できる。これらが指数関数の性質を満たしていることを確認しよう。 $q, r$  を無理数として計算すると、

$$\begin{aligned} q = 0 &\implies x^0 = e^0 = 1, \\ x^q x^r &= e^{q \log x} e^{r \log x} = e^{q \log x + r \log x} = e^{(q+r) \log x} = x^{q+r}, \\ (x^q)^r &= (e^{q \log x})^r = e^{qr \log x} = x^{qr}, \end{aligned}$$

が得られ、指数関数の性質が満たされていることが見てとれる。さらに  $x^q = e^{q \log x}$  に対して、両辺対数をとると、先に確認した  $y = \log e^y$  という数学的事実から

$$\log x^q = \log e^{q \log x} = q \log x$$

<sup>12</sup>念のために簡単に復習をしておこう。対数関数を  $p(x) = \log x$  と表せば  $\log f(x) = p(f(x))$  である。さらに  $p'(x) = 1/x$ 。これらを用いて合成関数の導関数の定理にあてはめれば  $(p(f(x)))' = p'(f(x))f'(x) = f'(x)/f(x)$ 。

<sup>13</sup>有理数乗を考える時は、 $x > 0$  である。 $x^{1/2} = \sqrt{x}$  の考慮の結果である。

<sup>14</sup>ただし、 $x$  の定義域は異なる（関数が定義できない場合を除くため）。 $\alpha$  が自然数の場合は  $x$  に制限はない。整数の場合は、 $x^{-2} = 1/x^2$  などが出てくる為、 $x \neq 0$  である。有理数の場合は先に脚註したように  $x > 0$ 。

を得る。指数対数の関係も充分である。

(1.7.1)を用いる。  $f(x) = x^q$  とすれば、

$$(\log f(x))' = \frac{f'(x)}{f(x)} = \frac{f'(x)}{x^q}$$

であり、一方

$$(\log f(x))' = (\log x^q)' = (q \log x)' = \frac{q}{x}$$

であるので、

$$\frac{f'(x)}{x^q} = \frac{q}{x} \iff f'(x) = qx^{q-1}$$

となる。

以上より、無理数でもこの形式が成り立つことが分かった。いままでの自然数整数、有理数での結果を全部合わせて、実数  $\alpha$  にたいして

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

となるのである。

## 1.8 高階導関数

関数  $f(x)$  が微分可能であった時、その導関数を  $f'(x)$  とした。この  $f'(x)$  も関数であるので、その導関数を考えることができる。それを  $f''(x)$  と書く。さらにこれも導関数なので — というように階層をどんどんと深くしていくことが可能になる。階層を深くした導関数を「高階導関数」とよぶ。' を並べて書くことにも限界があるので、 $(n)$  と記す。呼び方と書き方を列挙すれば、

$$1 \text{ 階の導関数: } f'(x) =: f^{(1)}(x)$$

$$2 \text{ 階の導関数: } f''(x) =: f^{(2)}(x)$$

$$3 \text{ 階の導関数: } f'''(x) =: f^{(3)}(x)$$

.....

$$n \text{ 階の導関数: } f^{(n)}(x) =: f^{(n)}(x)$$

である。

## 1.9 多変数関数と偏微分記号

関数  $f$  が2つの独立変数  $x, y$  から構成される関数の場合、それを「2変数関数」と言い、 $f(x, y)$  と書く。独立変数の個数を  $n$  個にまで広げれば、それは「 $n$ 変数関数」とよばれ、たとえば、 $f(x_1, x_2, \dots, x_m, \dots, x_n)$  のように書く。

多変数関数  $f$  において、ひとつの独立変数  $x_m$  に注目する。その際、その他の独立変数は、一旦「定数」であると見なし（すなわち、固定され変動しないものと見なし）、 $x_m = a + h$  と  $x_m = a$  の場合の差をとって

$$\lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, a+h, \dots, x_n) - f(x_1, x_2, \dots, a, \dots, x_n)}{h}$$

なる極限を考える。この極限が存在する場合、これを偏微分係数と呼び、変数  $x_m$  に関して実数点  $a$  において偏微分可能であるという。ここから、 $x_m$  がとりうるおのおのの実数点  $a$  にこの偏微分係数を対応させればそれが関数になることは、1変数の場合と同じである。したがって、新しい記号をもちいて、この偏導関数を次の様にあらわす：

$$\frac{\partial f}{\partial x_m} := \lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, x_m + h, \dots, x_n) - f(x_1, x_2, \dots, x_m, \dots, x_n)}{h} .$$

## 第2章 古典解析（無限小解析）流儀での微分法と導関数の導出

その昔、 $y = f(x)$  としてその導関数を  $f'(x)$  と書くのはもちろんのこと、かつ、 $dy/dx$  とも書く、と習った。そしてこれは「分数ではない」。

積分が微分の単なる逆演算であるならばその書き方は  $f(x) = \int f'(x)$  でもよかったはずだ。しかし、 $f(x) = \int f'(x) dx$  と  $dx$  がつく。とにかく積分では  $dx$  が単独で出てくる。微分方程式もしかり。であれば、 $dy, dx$  をそのまま扱える体系があってもいい。

歴史的には、「無限小解析」と呼ばれる方法論があった。天真爛漫にもちいると、すくなからぬ綻びが散見されるけれども、なにより直観に訴えてくるものである。この綻びを何とかしようとして、先人たちは多大な洞察を重ね、解析学を厳密にしてきた。しかしながら、その厳密な解析学は取り付きやすいものではない。綻びを綻びとして諒解しておけば、無限小解析と言われた微分量操作の方法論は、理解しやすい方法論であると思う。操作性も高い。ゆえに 응용数学方面（物理や工学など）では重宝する。そして綻びがあっても、結果は整合する。微分量は深淵である。

そしてこの微分量にもとづく理路は、導関数というものを通じて、 $dx$  が定量であるという事実を持って前章で述べた近代解析流儀と整合するのである。

### 2.1 微分量とその計算ルール（公理系）

#### 2.1.1 変化量と微分量

あらためて、「変化する量（単純に変化量と呼ぶことにする）」というものを考えてみる。具体的な例としては、変数が増えたり減ったりするし、変数の変化に伴って変化する関数が増えたり減ったりしよう。今ここでは変数や関数という属性は一旦忘れて、「変化する」という動きを持つものそれら全体を変化量と名づけ、 $u, v$  などで代表させることにする。

変化量  $u, v$  が変化した場合、当然そこには、その変化に対応する「変化分の量」が考えられる。そのうち微小なものを「微分量」と呼ぶことにし、それぞれ  $du, dv$  とあらわすことにする<sup>1</sup>。さらに、この微分量を求める（計算する）ことを『微分する』とも言う。

#### 2.1.2 微分量に関する公理

微分量に対して次の3個の公理を要請する。

公理 2.1.1.  $a$  を変化しない量（定量または定数と呼ぶ）とするとき、 $da = 0$ 。

公理 2.1.2.  $d(u + v) = du + dv$ 。

<sup>1</sup>この微小な変化分の量  $du$  は、変化量の「変化後と変化前」のそれぞれの量から  $du = (u + du) - u$  というようにあらわされうるが、この式はトートロジーであってここから特段の実りは現出しない。

公理 2.1.3.  $d(uv) = du \cdot v + u \cdot dv$ .

公理 2.1.1 は、定量は変化しないということを言明したものである。2.1.2, 2.1.3 は微分量の世界の構造を与えるものである。

### 2.1.3 独立変化量と、従属変化量（関数）

何者にも依存せずに、独自に変化する、変化できる変化量を独立変化量という。また、独立変化量の変化に応じて変化する量を従属変化量ということにする。慣れ親しんで来た「関数」は、変数の変化に応じて変化するものだから従属変化量である。

従属変化量も上の公理を満たすように構築したい。そこで従属変化量の微分量、すなわち関数の微分量を次のように定める（ただし、 $x$  は独立変化量<sup>2</sup>）：

$$d(f(x)) = f(x+dx) - f(x).$$

次の記述はすべて同じものである。表記の簡便性、時と場合と対象の区別の明確化などの理由で使い分けられる。

$$d(f(x)) = df(x) = df = f(x+dx) - f(x).$$

このように定義された関数（従属変化量）の微分量が、微分量の公理を満たすことを確認しておこう。 $h(x) := f(x) + g(x)$  とすれば、関数の微分量の定義から

$$\begin{aligned} dh(x) &= h(x+dx) - h(x) = (f(x+dx) + g(x+dx)) - (f(x) + g(x)) \\ &= (f(x+dx) - f(x)) + (g(x+dx) - g(x)) \\ &= df(x) + dg(x) \end{aligned}$$

すなわち  $d(f(x) + g(x)) = df(x) + dg(x)$  となり、公理 2.1.2 が成立していることがわかる。

同様に  $h(x) := f(x)g(x)$  とすると、

$$\begin{aligned} dh(x) &= h(x+dx) - h(x) \\ &= f(x+dx)g(x+dx) - f(x)g(x) \\ &= (df(x) + f(x))(dg(x) + g(x)) - f(x)g(x) \\ &= df(x) \cdot g(x) + f(x) \cdot dg(x) + df(x) \cdot dg(x) \end{aligned}$$

となり、最後に余分なものが出来てしまう。公理を満たすためにはこの項は 0 でなければならない。つまり公理 2.1.3 は  $df(x) \cdot dg(x) = 0$  を要請するものなのである<sup>3</sup>。そしてその時  $d(f(x)g(x)) = df(x) \cdot g(x) + f(x) \cdot dg(x)$ 。

### 2.1.4 基本的な定理

まず線形性と商の微分量の定理を証明しよう。

<sup>2</sup>では独立変化量と従属変化量の違いは何なのか、という疑問が当然起こる。第 2.5 節「高階導関数と独立変化量（独立変数）」での議論において、変化量が独立であるとはどういうことか、が明らかになる。

<sup>3</sup>と言い切ってしまうが、素直には納得しがたい。さらに、第 2.5 節「高階導関数と独立変化量（独立変数）」のところで、論理の綻びをみる。



定理 2.1.1.  $\alpha, \beta$  を定数とすると,  $d(\alpha u + \beta v) = \alpha \cdot du + \beta \cdot dv$  (線形性).

証明. まず,  $d(\alpha u) = d\alpha \cdot u + \alpha \cdot du$  であり,  $\alpha$  は定数であるから,  $d\alpha = 0$ . したがって,  $d(\alpha u) = \alpha \cdot du$ .  
同様に,  $\beta$  も定数であるから,  $d(\beta v) = \beta \cdot dv$ . ゆえに,

$$d(\alpha u + \beta v) = d(\alpha u) + d(\beta v) = \alpha \cdot du + \beta \cdot dv.$$

□

定理 2.1.2.  $v \neq 0$  として,  $d\left(\frac{u}{v}\right) = \frac{du \cdot v - u \cdot dv}{v^2}$  (商の微分量).

証明.  $w = u/v$  すなわち  $wv = u$  として両辺を微分すると

$$\begin{aligned} d(wv) = d(u) &\iff dw \cdot v + w \cdot dv = du &\iff dw = \frac{du - w \cdot dv}{v} \\ &\iff dw = \frac{du - \frac{u}{v} \cdot dv}{v} &\iff dw = \frac{du \cdot v - u \cdot dv}{v^2}. \end{aligned}$$

□

次に, 合成関数の微分量 (合成関数の微分) の定理である.

定理 2.1.3.  $y = f(u)$ ,  $u = g(x)$  なる関係で  $y$  と  $x$  が  $u$  を介して結ばれている場合,  $dy = \frac{dy}{du} \frac{du}{dx} dx$ .

証明. 証明の基本は, 微分量  $du$ ,  $dx$  について,  $1 = du/du$  という式を用いて掛けたり割ったりするだけである.

$$dy = dy \cdot 1 = dy \frac{du}{du} = dy \frac{du}{du} \cdot 1 = dy \frac{du}{du} \frac{dx}{dx} = \frac{dy}{du} \frac{du}{dx} dx.$$

この時, 本章のはじめに述べたように, 書き方のバリエーションとして,  $dy = d(f(x)) = df$  もありうるので, この定理は次のように書くこともある:

$$dy = \frac{dy}{du} \frac{du}{dx} dx = \frac{df}{dg} \frac{dg}{dx} dx.$$

□

最後に逆関数の微分量 (逆関数の微分) の定理である.

定理 2.1.4.  $x = f(y)$  の逆関数を  $y = f^{-1}(x)$  と書くと,  $df^{-1} = \frac{1}{df} dx$ .

証明.

$$dx = dx \frac{dy}{dy} = \frac{dx}{dy} dy \iff dy = \frac{1}{\frac{dx}{dy}} dx.$$

$dx = df$ ,  $dy = df^{-1}$  でもあるから,

$$df^{-1} = \frac{1}{df} dx.$$

□

## 2.2 微分量の公理系を支える数学的事実

微小である微分量と、超越関数に対しては、次の事実をあらかじめ認めておく<sup>4</sup>.

$$e^{(dx)} - 1 = dx,$$

$$\sin(dx) = dx,$$

$$\cos(dx) = 1.$$

## 2.3 微分量の具体的な計算例

- 自然数による巾乗

$$d(x^2) = d(xx) = dx \cdot x + x \cdot dx = 2x dx$$

これを数珠つなぎしていけば、 $n$  を自然数として、

$$d(x^n) = nx^{n-1} dx$$

を得る<sup>5</sup>.

- 整数による巾乗

$n$  を自然数とすれば、 $x^{-n} = 1/x^n$  である。まずこれを諒解しておく。さて、今、 $y = 1/x$  とすれば、 $xy = 1$  であるから、これを両辺微分すると、

$$\begin{aligned} d(xy) = dx \cdot y + x \cdot dy = 0 &\iff dy = -\frac{1}{x^2} dx = -x^{-2} dx \\ &\iff d(x^{-1}) = -x^{-2} dx. \end{aligned}$$

よってこれも同様に数珠つなぎで、

$$d(x^{-n}) = -nx^{-n-1} dx.$$

結果、上の自然数の例と合わせて、 $m$  を整数として、

$$d(x^m) = mx^{m-1} dx.$$

<sup>4</sup>いまここで立てている理路の枠組みの中では、これらは導出できない。Taylor 展開は導出すべきものであるとしているから、それも利用できないのはもちろんである。微分量が微小であることを利用して、極限の方で使った事実  $\lim_{h \rightarrow 0} (e^h - 1)/h = 1$  や  $\lim_{h \rightarrow 0} \sinh/h = 1$  を形をかえて利用することにした。これでいいのか？他に方法はあるのか？

<sup>5</sup>数学的帰納法の練習問題に適しているよう。

- 三角関数

$$\begin{aligned} d(\sin x) &= \sin(x+dx) - \sin(x) \\ &= \sin x \cos(dx) + \cos x \sin(dx) - \sin x \\ &= \sin x + \cos x \cdot dx - \sin x \\ &= \cos x \cdot dx. \end{aligned}$$

$$\begin{aligned} d(\cos x) &= \cos(x+dx) - \cos x \\ &= \cos x \cos(dx) - \sin x \sin(dx) - \cos x \\ &= \cos x - \sin x \cdot dx - \cos x \\ &= -\sin x \cdot dx. \end{aligned}$$

- 指数関数

$$d(e^x) = e^{x+dx} - e^x = e^x(e^{dx} - 1) = e^x dx$$

- 対数関数

$y = \log x$  は  $x = e^y$  の逆関数である。したがって、

$$dy = \frac{1}{dx} dx = \frac{1}{e^y dy} dx = \frac{1}{e^y} dx = \frac{1}{x} dx \iff d(\log x) = \frac{1}{x} dx$$

- 実数による巾乗

すべての実数  $q$  において、 $x^q = e^{q \log x}$  であることを、再度利用する。いま、 $y = x^q$  とし、かつ、 $u = q \log x$  とすれば、 $y = e^u$  であるので、

$$dy = \frac{dy}{du} \frac{du}{dx} dx = e^u q \frac{1}{x} dx = qx^q \frac{1}{x} dx = qx^{q-1} dx$$

となる。

## 2.4 導関数

導関数の導入の目的は、近代解析流儀との対応を見るためであり、かつ、整合させるためである。

$f(x)$  の微分量は、 $df = f(x+dx) - f(x)$  であった。この時、両辺を微分量  $dx$  で割ったものを導関数と名付け、 $f'(x)$  と記す。すなわち、

$$\frac{f(x+dx) - f(x)}{dx} =: f'(x).$$

この導関数を利用すると

$$df = f'(x) dx \iff \frac{df}{dx} = f'(x)$$

となることはあきらか<sup>6</sup>。導関数もそれ自身変化量であるので、公理を適用し、導関数に対する基本的な計算方法を導いておこう。

<sup>6</sup>単なる書きかたの変更にすぎないけれども、近代解析流儀での導関数と古典解析流儀での微分量との対応がここから導かれる。

- 定量の公理 (公理 2.1.1)

定量  $a$  の微分量は  $da=0$  であった。ここで両辺を  $dx$  で割れば、

$$\frac{da}{dx} = 0$$

である。そして、左辺は  $a$  の導関数の定義に他ならないので、定量の導関数は  $0$  という事実を得る。

- 和の公理 (公理 2.1.2)

変化量  $x$  の関数を  $f, g$  とする。和の微分量は  $d(f+g)=df+dg$  であった。ここで両辺を  $dx$  で割れば、

$$\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx} = f' + g'$$

である。そして、左辺は  $f+g$  の導関数の定義に他ならないので、次を得る：

$$(f+g)' = f' + g'.$$

- 積の公理 (公理 2.1.3)

積の微分量は  $d(fg)=df \cdot g + f \cdot dg$  であった。上と同様に両辺を  $dx$  で割れば、

$$\frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx} = f' \cdot g + f \cdot g'$$

であり、左辺は  $fg$  の導関数の定義に他ならないので、次を得る：

$$(fg)' = f' \cdot g + f \cdot g'.$$

- 導関数の線形性 (定理 2.1.1)

$\alpha, \beta$  は定数であるとする。この時、 $d(\alpha f + \beta g) = \alpha \cdot df + \beta \cdot dg$  であった。上と同様な論法で

$$\frac{d(\alpha f + \beta g)}{dx} = \alpha \cdot \frac{df}{dx} + \beta \cdot \frac{dg}{dx} = \alpha \cdot f' + \beta \cdot g' \iff (\alpha f + \beta g)' = \alpha \cdot f' + \beta \cdot g'.$$

- 商の導関数 (定理 2.1.2)

商の微分量は  $d(f/g) = (df \cdot g - f \cdot dg)/g^2$  であった。同様な論法で、

$$\frac{d\left(\frac{f}{g}\right)}{dx} = \frac{1}{g^2} \left( \frac{df}{dx} \cdot g - f \cdot \frac{dg}{dx} \right) \iff \left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}.$$

- 合成関数の導関数 (定理 2.1.3)

$y = f(u)$ ,  $u = g(x)$  の時、 $dy = \frac{dy}{du} \frac{du}{dx} dx$  であった。したがって、

$$\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = f'(u)g'(x) = f'(g(x))g'(x).$$

- 逆関数の導関数 (定理 2.1.4)

$x = f(y)$ ,  $y = f^{-1}(x)$  の時、 $df^{-1} = \frac{1}{df} dx$  であった。したがって、 $\frac{df^{-1}}{dx} = \frac{1}{df}$  となる。この場合、 $\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$

の方が若干見通しは良い<sup>7</sup>。

<sup>7</sup>ここまで見てきたように、微分量  $dy, dx$  などの扱いはほとんど「数」と同じで縦横無尽である。高校時代に「それは分数ではない」と言われた  $\frac{dy}{dx}$  は、導関数としての  $\frac{dy}{dx}$  であったわけなのだろう。しかし、それでは積分の時に単独で出てきた  $dx$  を説明できないのではないか？もちろんいろんな制約があったのだろうとは思えけれども。

## 2.5 高階導関数と独立変化量（独立変数）

### 2.5.1 高階導関数

$f(x)$  の導関数とは,

$$df = f'(x)dx, \quad f'(x) = \frac{f(x+dx) - f(x)}{dx}$$

となる関数  $f'(x)$  のことであつた。ここで、焦点をこの  $f'(x)$  に移す。  $f'(x)$  も変化量である。関数に別名を与えて混乱をさけよう。  $g(x) := f'(x)$  と表せば、この  $g$  の微分量は

$$dg = g'(x)dx, \quad g'(x) = \frac{g(x+dx) - g(x)}{dx} = \frac{f'(x+dx) - f'(x)}{dx}$$

となる。この  $g'(x)$  は、  $f'(x)$  の導関数であり、  $f'(x)$  は  $f(x)$  の導関数であつた。これをまとめて、  $g'(x)$  は  $f(x)$  の「2階導関数」と呼ぶことにする。階級を追加するのである。そしてそれを  $f''(x)$  と ' の個数で表現したり、  $f^{(2)}(x)$  のように表現したりする。こんな感じである。

$$f'(x) = f^{(1)}(x) = \frac{f(x+dx) - f(x)}{dx},$$
$$f''(x) = f^{(2)}(x) = \frac{f'(x+dx) - f'(x)}{dx}.$$

「3階導関数」や一般の「 $n$ 階導関数」も推して知るべしである。

### 2.5.2 独立変化量（独立変数）の導入

導関数の導入は混乱なく素直に受け入れられる。これが2階導関数になつた場合、もとの関数と2階導関数のあいだにあらたな事態が生じてくる。その事態を、  $f(x) = x^2$  という関数でみてみよう。微分量の計算から1階導関数、2階導関数をもとめれば、

$$df = 2x dx \implies \frac{df}{dx} = f'(x) = 2x$$
$$df' = 2 dx \implies \frac{df'}{dx} = f''(x) = 2$$

となる。さて今ここで、  $df$  の微分量を求めてみる。実際に計算すれば、

$$d(df) = d(2x dx) = d(2x) \cdot dx + 2x \cdot d(dx) = 2 dx dx + 2x \cdot d(dx)$$

となる。もうひとつ変形を進めると、

$$\frac{d(df)}{dx dx} = 2 + 2x \cdot \frac{d(dx)}{dx dx}$$

である。一方、  $f''(x) = 2$  であつた。したがつて、

$$\frac{d(df)}{dx dx} = f''(x) + 2x \cdot \frac{d(dx)}{dx dx}$$

と表せる。ゆえに、  $2x \cdot \frac{d(dx)}{dx dx} = 0$  ならば、

$$\frac{d(df)}{dx dx} = f''(x)$$

ということが言えるようになる。この条件において、 $x$  は変数なので様々な値をとりうるから、恒等的に 0 になるための候補は  $d(dx)=0$  しかない。繰り返すと、

$$d(dx)=0 \implies \frac{d(df)}{dx dx} = f''(x)$$

となるのである。

以上の事態を一般化してみよう。見通しをよくするため、 $f(x), f'(x)$  などの  $(x)$  は省いて記載する。

導関数  $f'$  は  $df = f' dx$  であった。これが  $f'$  と  $dx$  の積であることから公理 2.1.3 (微分量の積の公理) を利用して  $df$  の微分量を求めると、

$$d(df) = d(f' \cdot dx) = df' \cdot dx + f' \cdot d(dx)$$

が得られる。また  $f'$  の導関数  $f''$  をもちいれば、 $f'$  の微分量は、 $df' = f'' \cdot dx$  であるのでそれを代入すれば、

$$d(df) = f'' \cdot dx dx + f' \cdot d(dx) \iff \frac{d(df)}{dx dx} = f'' + f' \cdot \frac{d(dx)}{dx dx} \quad (2.5.1)$$

となる。これが一般的に導かれる形である。

そしてこの時、 $d(dx)=0$  であれば、 $\frac{d(df)}{dx dx} = f''$  が成立する。

さて、 $d(dx)=0$  とはどういうことか？これは  $dx$  の微分量が 0 であるということ、つまり、 $dx$  は定量である (変化量ではない) ということを物語っている。このように独立変化量 (独立変数) においては、微分量が定量にならざるを得ないのである<sup>8</sup>。逆にいえば、 $x$  が独立変化量 (独立変数) であれば、 $d(dx)=0$  であり、 $\frac{d(df)}{dx dx} = f''$  が成り立つのである。

### 2.5.3 微分記号 $d$ の記法の拡張

$d(df)$  なる微分量は「微分の微分」、すなわち「2回微分した」という行為を表明しているので、この括弧省いた形として、あたかも  $d$  を普通の数のように扱うことを認める。すなわち、

$$d(df) = dd f = d^2 f.$$

つまり、 $dd \dots d$  と  $n$  個並んでいる場合には  $dd \dots d = d^n$  を認めるのである。

つぎに、ちょっと無理矢理感が否めないが、 $dx dx = (dx)^2 = dx^2$  と書くことを認める。一般には、 $(dx)^n = dx^n$  であるとする。高階導関数の記法の見通しをよくするための処置である。これらの記法を採用すれば、先の式 (2.5.1) は、

$$\frac{d^2 f}{dx^2} = f'' + f' \cdot \frac{d^2 x}{dx^2}$$

である。これからわかるように、この古典解析流儀では、一般には  $\frac{d^2 f}{dx^2} = f''$  は成り立たない。余分な項がついてまわるのである。これを導関数と等しくするためには、 $d(dx)=0$  でなければならない。つまり、 $x$  を「独立変数」というものと見なして進んでいかなくてはならない。独立変数であるので、 $d^2 x = d(dx)=0$  となり (というか、これは独立変数というものの定義である)、

$$\frac{d^2 f}{dx^2} = f''$$

<sup>8</sup>微分すると定量になる変化量とは何か？その素性は何なのか？にわかには認識しがたく、微分量の深淵が伺われるものである。「無限小」として始まった微分量の扱いにおいて、様々な格闘が先人達にはあったようだ。そしてその格闘に巻き込まれる時我々は混乱する。有意義な混乱であるけれども、一抹の不安は常に尽きない。高瀬正仁著『 $dx$  と  $dy$  の解析学—オイラーに学ぶ』[1] はこの不安に対して心強い味方になってくれる本である。

というお馴染みの式が出てくるのである。この背後には  $dx$  は定数という神秘が隠れているのであった。これをもって、近代的流儀との整合性が保たれるのである。

さて、ここで困ったことがひとつおこる。2.1.3 節で、 $df(x) \cdot dg(x)$  は 0 であると考えねばならなかった。では今出てきた  $dx^2 = (dx)^2 = dx dx$  は 0 ではないのか？ たしかに、毎回毎回の微分の操作（導関数導出の操作）においては、 $dx$  が単独で出てきて、 $dx dx$  は出てこない。しかしながら  $n$  回微分や  $n$  階導関数を拡張した記法で書くと  $dx dx = dx^2$  や  $dx^3$  などが出てくる。ここが形式的な綻びが出現する場面なのである。

## 2.6 多変数関数と全微分

独立変数  $n$  個からなる関数を「 $n$  変数関数」とよぶことは近代解析流儀と同様で変わらない。今それを  $f(x_1, x_2, \dots, x_m, \dots, x_n)$  のように書く。この関数  $f$  において、ひとつの独立変数  $x_m$  に注目する。その際、その他の独立変数は、一旦「定数」と見なす（すなわち、固定され変動しない）。 $f$  をそのように見立てた場合を  $f_m$  と記すことにすれば、

$$df_m = f(x_1, x_2, \dots, x_m + dx_m, \dots, x_n) - f(x_1, x_2, \dots, x_m, \dots, x_n)$$

であり、この時の導関数は

$$df_m = f'_m dx_m \iff f'_m = \frac{df_m}{dx_m} = \frac{f(x_1, x_2, \dots, x_m + dx_m, \dots, x_n) - f(x_1, x_2, \dots, x_m, \dots, x_n)}{dx_m}$$

であらわせる。この導関数  $f'_m$  を、「見立て」を尊重して  $\partial f / \partial x_m$  と書く。偏導関数と呼ばれることもある。すなわち、

$$df_m = \frac{\partial f}{\partial x_m} dx_m$$

さらに、 $x_m$  は独立変数としたのだから、 $d(dx_m) = 0$  という条件が成立し、その元で  $n$  階の偏導関数を、 $\frac{\partial^n f}{\partial x_m^n}$  と記す。

$f$  の微分は、全ての独立変数において偏導関数を計算し、それを加えたものとして定義される（多変数に拡張したので、定義するしかないのだ<sup>9</sup>）。これを全微分と呼ぶこともある。次のような表式となる：

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n.$$

<sup>9</sup>ベクトルの力を借りればそうでもないかもしれない。ベクトルを変数とする関数  $f(\vec{x})$  の微分を  $df = f(\vec{x} + d\vec{x}) - f(\vec{x})$  とし、これに同じ公理を持ち込んで出発すれば、理論が構築できるような気もする（ベクトル解析の基礎づけ？）。しかし、「独立ベクトルの微分が定数」なんて言えるのだろうか？

## 第3章 Taylor 展開

関数を巾乗で近似するという思考の飛躍をもって、近似的な方法で、Maclaurin 展開（マクローリン展開）を導出し、それを平行移動して Taylor 展開（テイラー展開）を導く<sup>1</sup>。

### 3.1 近代解析的導出

#### 3.1.1 Maclaurin 展開

関数  $f(x)$  は  $x$  の巾乗の多項式とほぼ等しいのではないか、という思考の飛躍に基づいた仮説を立ててみる。つまり、

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 + a_7x^7 + \dots$$

1 階、2 階、3 階の導関数をそれぞれ求めてみると

$$f^{(1)}(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4 + \dots$$

$$f^{(2)}(x) = 2a_2 + 6a_3x + 12a_4x^2 + 20a_5x^3 + 30a_6x^4 \dots$$

$$f^{(3)}(x) = 6a_3 + 24a_4x + 60a_5x^2 + 120a_6x^3 + 210a_7x^4 \dots$$

となる。ここで  $x=0$  とすれば、順次

$$a_0 = f(0), \quad a_1 = f^{(1)}(0), \quad a_2 = \frac{1}{2}f^{(2)}(0), \quad a_3 = \frac{1}{6}f^{(3)}(0)$$

と求まる。みて分かるように、定数項のものが値を決定する。順次やっていくのは計算機に任せて、すこし頭を使ってみよう。もともとの  $f(x)$  には  $a_nx^n$  があった。これを定数項にもっていきたい。1 回の微分で、巾の肩が降り、巾の肩は 1 減るので、 $n$  回微分すると

$$(a_nx^n)^{(n)} = a_n n(n-1)(n-2)\dots 2 \cdot 1 = a_n n!$$

という定数項になる。したがって、

$$f^{(n)}(x) = a_n n! + Ax + Bx^2 + \dots$$

という形となることがわかる ( $A, B$  は適当な数)。それゆえ、

$$a_n = \frac{1}{n!} f^{(n)}(0)$$

と求まる。この結果から、

$$f(x) = f(0) + \frac{f^{(1)}(0)}{1!}x + \frac{f^{(2)}(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots$$

となる。最初の仮説のもとで、微分係数を用いて得られたこの式を、 $f(x)$  の Maclaurin 展開という。 $n \rightarrow \infty$  でこれが成立するかどうかは吟味が必要であるが、形式的にはこのようにもとまる。

<sup>1</sup>厳密証明はどうするか？ Lagrange, Rolle, Cauchy と進まねばなるまいが、気が重い。



### 3.1.2 Taylor 展開

Maclaurin 展開が成立する場合に、そこから Taylor 展開を導出してみよう。今、 $g(x) = f(x+a)$  なる関数  $g(x)$  を考える ( $f(x)$  を  $a$  平行移動した関数と考えても良い)。この  $g(x)$  を Maclaurin 展開すれば、

$$g(x) = g(0) + \frac{g^{(1)}(0)}{1!}x + \frac{g^{(2)}(0)}{2!}x^2 + \frac{g^{(3)}(0)}{3!}x^3 + \dots + \frac{g^{(n)}(0)}{n!}x^n + \dots$$

である。一方、関数が等しいのだから導関数も等しく、

$$g'(x) = f'(x+a)$$

という関係になる<sup>2</sup>。この関係は  $n$  階導関数でも同様。したがって、

$$g(0) = f(a), \quad g'(0) = f'(a), \quad \dots, \quad g^{(n)}(0) = f^{(n)}(a), \quad \dots$$

となるので、 $g(x)$  の Maclaurin 展開は

$$g(x) = f(a) + \frac{f^{(1)}(a)}{1!}x + \frac{f^{(2)}(a)}{2!}x^2 + \frac{f^{(3)}(a)}{3!}x^3 + \dots + \frac{f^{(n)}(a)}{n!}x^n + \dots$$

そもそも  $g(x) = f(x+a)$  であったから、

$$f(x+a) = f(a) + \frac{f^{(1)}(a)}{1!}x + \frac{f^{(2)}(a)}{2!}x^2 + \frac{f^{(3)}(a)}{3!}x^3 + \dots + \frac{f^{(n)}(a)}{n!}x^n + \dots$$

ここで、 $u = x+a$  とすると、もちろん  $x = u-a$  であるので、

$$f(u) = f(a) + \frac{f^{(1)}(a)}{1!}(u-a) + \frac{f^{(2)}(a)}{2!}(u-a)^2 + \frac{f^{(3)}(a)}{3!}(u-a)^3 + \dots + \frac{f^{(n)}(a)}{n!}(u-a)^n + \dots$$

これを  $f(u)$  の  $a$  のまわりの Taylor 展開という。

さらに、 $u-a = \Delta a$  とすれば、

$$f(a + \Delta a) = f(a) + \frac{f^{(1)}(a)}{1!}\Delta a + \frac{f^{(2)}(a)}{2!}(\Delta a)^2 + \frac{f^{(3)}(a)}{3!}(\Delta a)^3 + \dots + \frac{f^{(n)}(a)}{n!}(\Delta a)^n + \dots$$

となる。 $a$  を関数の変数と見た場合の Taylor 展開のもう一つの表式である。 $\lim_{\Delta a \rightarrow 0}$  で導関数そのものになっていることが見て取れよう。

## 3.2 古典解析的導出

古典解析流儀の微分量の扱いの操作性の高さをまじまじと実感させてくれる方法をなぞってみよう<sup>3</sup>。二項係数の近似計算結果をそのまま利用するというところに躊躇と混乱、議論はあるのだが。

変化量を  $x$  として、 $y = f(x)$  を考える。すると  $dy$  を計算することによって、

$$\begin{aligned} dy = f(x+dx) - f(x) &\iff f(x+dx) = f(x) + dy \\ &\iff f(x+dx) = y + dy \end{aligned}$$

<sup>2</sup> $f$  の方の導関数は、精密に言えば、合成関数の導関数をもとめる chain rule を使っている。

<sup>3</sup>この方法は、高瀬 [1] から学んだ。

が得られる。いまここで、一時的に  $g(x) := f(x+dx)$  と置いてみよう。そして  $g$  の微分量を計算すると、

$$d(g(x)) = g(x+dx) - g(x) \iff g(x+dx) = g(x) + d(g(x)).$$

さて、 $f$  と  $g$  の関係をおさらいすると、 $g(x) = f(x+dx)$  であったから、 $g(x+dx) = f(x+2dx)$  であり、 $d(g(x) = d(f(x+dx))$  である。結果、

$$\begin{aligned} g(x+dx) = g(x) + d(g(x)) &\iff f(x+2dx) = f(x+dx) + d(f(x+dx)) \\ &\iff f(x+2dx) = y + dy + d(y+dy) = y + 2dy + d^2y \end{aligned}$$

を得る<sup>4</sup>。さらに続けて、先ほどの  $g(x)$  は忘れて、再度一時的に  $g(x) := f(x+2dx)$  と置いてみよう。この新しい  $g$  の微分量は、 $d(g(x)) = g(x+dx) - g(x)$  で変わらないが、 $f$  と  $g$  の関係を上記の結果も用いてあらわすと、

$$\begin{aligned} g(x) &= f(x+2dx) = y + dy + d(y+dy) = y + 2dy + d^2y, \\ g(x+dx) &= f(x+3dx), \\ d(g(x)) &= d(f(x+2dx)) = d(y + dy + d(y+dy)) = dy + 2d^2y + d^3y \end{aligned}$$

であり、結果

$$\begin{aligned} d(g(x)) = g(x+dx) - g(x) &\iff g(x+dx) = g(x) + d(g(x)) \\ &\iff f(x+3dx) = f(x+2dx) + d(f(x+2dx)) \\ &\iff f(x+3dx) = y + 3dy + 3d^2y + d^3y \end{aligned}$$

を得る。この手順を  $n$  まで繰り返せば、

$$f(x+ndx) = y + \binom{n}{1}dy + \binom{n}{2}d^2y + \dots + \binom{n}{k}d^ky + \dots + \binom{n}{n-1}d^{n-1}y + d^ny.$$

さてここでひとつズルをする。ズルの前に、二項係数についての次の変形を見ておくと、

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} = \frac{n^k}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right)$$

であるから、 $n$  が  $k$  に比べて非常に大きい時には

$$\binom{n}{k} \sim \frac{n^k}{k!}$$

と近似できる。これを利用してしまふというところがズルのもとである。

次に、ある有限量（変化量ではない定量）を  $\Delta x$  として、

$$n =: \frac{\Delta x}{dx} \quad (\text{i.e. } ndx = \Delta x)$$

と定める。すると微分量  $dx$  の微小性から、この  $n$  は非常に大きな数になる。したがって、上の近似を適用<sup>5</sup>し、 $n$  の巨大性は和を無限大まで拡張することと考えて良いから（すこし精密さに欠ける議論ではある）、

$$\begin{aligned} f(x+\Delta x) &= y + \frac{n}{1!}dy + \frac{n^2}{2!}d^2y + \dots + \frac{n^k}{k!}d^ky + \dots \\ &= y + \frac{1}{1!} \left(\frac{\Delta x}{dx}\right) dy + \frac{1}{2!} \left(\frac{\Delta x}{dx}\right)^2 d^2y + \dots + \frac{1}{k!} \left(\frac{\Delta x}{dx}\right)^k d^ky + \dots \\ &= y + \frac{1}{1!} \frac{dy}{dx} (\Delta x) + \frac{1}{2!} \frac{d^2y}{dx^2} (\Delta x)^2 + \dots + \frac{1}{k!} \frac{d^ky}{dx^k} (\Delta x)^k + \dots \end{aligned}$$

<sup>4</sup> $d(dy) = d^2y$ ,  $d(d^2y) = d^3y$  等の表記を利用している。

<sup>5</sup> $(dx)^k = dx^k$  という表記を用いている。

となる。

$x$  が独立変化量であった場合（すなわち、 $y$  は従属変化量の場合）、微分量の商は導関数と等しくなるので

$$\frac{d^k y}{dx^k} = f^{(k)}(x)$$

とできる。したがって、この場合には、

$$f(x + \Delta x) = f(x) + \frac{f^{(1)}(x)}{1!}(\Delta x) + \frac{f^{(2)}(x)}{2!}(\Delta x)^2 + \cdots + \frac{f^{(k)}(x)}{k!}(\Delta x)^k + \cdots$$

となる。ここでグラフ的に右へ  $\Delta x$  ずらす、つまり、 $x \mapsto x - \Delta x$  と変換を行えば、

$$f(x) = f(x - \Delta x) + \frac{f^{(1)}(x - \Delta x)}{1!}(\Delta x) + \frac{f^{(2)}(x - \Delta x)}{2!}(\Delta x)^2 + \cdots + \frac{f^{(k)}(x - \Delta x)}{k!}(\Delta x)^k + \cdots$$

であり、さらに  $x - \Delta x = a$ （すなわち、 $\Delta x = x - a$ ）とおけばお馴染みの

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x - a) + \frac{f^{(2)}(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x - a)^k + \cdots$$

を得る。

最後に、 $a = 0$  として、次の Maclaurin 展開の式が出てくる。

$$f(x) = f(0) + \frac{f^{(1)}(0)}{1!}x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(k)}(0)}{k!}x^k + \cdots$$

## 第4章 積分

変化量という存在をもとに微分量を考え、幾つかの公理を要請してその公理のもとでなりたつ世界を見てきたのが古典的流儀であった。立場を変えて、出発点を微分量とその世界とし、そこから元になる変化量について考察するというのが積分である。まず微分量の逆演算として不定積分というものを導入し、その計算規則を見る。さらに、関数概念を前面にだして、変数の動く領域を限定することにより、定積分という考え方に至る。そして最後に定積分と不定積分の関係を見る。「微分量の寄せ集め」の力強さを知ることになる。

### 4.1 不定積分

#### 4.1.1 不定性

なにをもって「不定」というのか？はじめに、そこを明らかにしていこう。微分量からもとの変化量を求めることを、「不定積分  $\int$ 」ということにする。微分の「逆演算」として積分というものを考えようという姿勢である。変化量  $u$  と微分量  $du$  に対しての働きは、それぞれの演算を  $d$ ,  $\int$  とあらわすことにして、

$$u \xrightarrow{d} du \xrightarrow{\int} u$$

という感じになる。そして、微分量が  $du$  となる変化量のことを

$$\int du$$

と記述し、「不定積分」となづけることにする。

ここで、 $C$  を任意の定数であるとする。定数の微分は  $0$  であるという大切な公理があった。ゆえに  $d(u+C) = du$  である。そしてこの結果から、 $\int du = u$  でもあり、 $\int du = u + C$  でもあることがわかる。 $C$  は定数であればなんでも構わないので、不定積分で示される変化量は無数に存在することになる。しかしながら、無数に存在するとはいえ、違いは加算する定数の部分だけだ。これが「不定」という単語で言いたいことなのである。

いま、ここで、「 $a$  と  $b$  の違いは加算する定数の分だけである」ということを加味した「等号」を「 $\doteq$ 」と記すことにしよう<sup>1</sup>。すなわち、 $C$  を任意の定数として、

$$a \doteq b \iff a = b + C$$

という意味合いのものであるとする。これを用いれば、不定積分に関しては、

$$\int du \doteq u$$

と言えるし、書ける。不定積分に関しては、これをじゃんじゃん使っていく。

<sup>1</sup>似たようなものとして、モジュロ演算 (mod) と等号が思い浮かぶ。

### 4.1.2 計算規則を与える定理

微分量の公理および定理から、不定積分に対して、計算規則を与える次の定理が導出される。

**定理 4.1.1.**  $\alpha$  を定数として、 $\int \alpha du \doteq \alpha \int du$  (定数倍に関して)。

**証明.** まずすなおに、不定積分の記述方法から

$$\int d(\alpha u) \doteq \alpha u .$$

一方、微分量の公理から、 $d(\alpha u) = \alpha du$  なので、

$$\int d(\alpha u) \doteq \int \alpha du .$$

したがって、

$$\int \alpha du \doteq \alpha u .$$

一方、 $u$  を不定積分で書き表すと  $u \doteq \int du$  であるから、 $\alpha u \doteq \alpha \int du$  となり、最終的に

$$\int d(\alpha u) \doteq \int \alpha du \doteq \alpha \int du \doteq \alpha u .$$

□

**定理 4.1.2.**  $\int (du + dv) \doteq \int du + \int dv$  (和の分解)。

**証明.** まず、 $\int d(u+v) \doteq u+v$  であり、 $u \doteq \int du$ 、 $v \doteq \int dv$  であるから、

$$\int d(u+v) \doteq \int du + \int dv .$$

さらに、 $d(u+v) = du + dv$  であるので、 $\int d(u+v) \doteq \int (du + dv)$ 。したがって、

$$\int (du + dv) \doteq \int du + \int dv .$$

合わせて、 $u+v \doteq \int du + \int dv$  もわかる。

□

**定理 4.1.3.**  $\alpha, \beta$  を定数とすると、 $\int (\alpha du + \beta dv) \doteq \alpha \int du + \beta \int dv$  (線形性)。

**証明.** 定理 (4.1.1), (4.1.2) より明らかである。

□

**定理 4.1.4.**  $\int d(uv) \doteq \int du \cdot v + \int u \cdot dv$  (積の分解, 部分積分)。

証明. 微分量の公理と, 上記の和の分解の定理 (4.1.2) からそのまま導出できる.

$$\int d(uv) \doteq \int (du \cdot v + u \cdot dv) \doteq \int du \cdot v + \int u \cdot dv.$$

$\int d(uv) \doteq uv$  と合わせて,

$$\begin{aligned} uv \doteq \int du \cdot v + \int u \cdot dv &\iff \int du \cdot v \doteq uv - \int u \cdot dv \\ &\iff \int u \cdot dv \doteq uv - \int du \cdot v \end{aligned}$$

が「部分積分」でお馴染みのものである. □

系 4.1.1.  $\int d\left(\frac{u}{v}\right) \doteq \int du \frac{1}{v} - \int \frac{u}{v^2} dv$  (商の積分).

証明. 微分量に関する定理 (2.1.2)

$$d\left(\frac{u}{v}\right) = \frac{du \cdot v - u \cdot dv}{v^2}$$

より明らか. □

定理 4.1.5.  $y = f(u)$ ,  $u = g(x)$  なる関係で  $y$  と  $x$  が  $u$  を介して結ばれている場合,  $y \doteq \int \frac{dy}{du} \frac{du}{dx} dx$ .

証明. 微分量に関する定理 (2.1.3) 文字通りのままで,

$$dy = \frac{dy}{du} \frac{du}{dx} dx$$

より明らか. □

### 4.1.3 導関数での計算規則

関数概念を前面に押し出して, 導関数をもちいての計算規則を考えよう. 今,  $x$  を独立変数 (つまり, 独立変化量) とすると, 関数  $f(x)$  とその導関数  $f'(x)$  の間には,

$$df(x) = f'(x)dx$$

の関係があった.  $f(x)$  も  $x$  も変化量であるから, 上記計算規則が使えて,

$$f(x) \doteq \int df(x) \doteq \int f'(x)dx$$

が得られる. この時,  $f(x)$  を「原始関数」ともいう. 微分すると  $f'(x)$  となる関数の総称である (導関数の「逆」の意味合い).

部分積分の表現も導出しよう。  $f(x)$  に加えて  $g(x)$  も仲間に入れる。計算規則から

$$\begin{aligned} \int d(f(x)g(x)) &\doteq \int df(x) \cdot g(x) + \int f(x) \cdot dg(x) \doteq \int f'(x)g(x)dx + \int f(x)g'(x)dx \\ &\iff \\ f(x)g(x) &\doteq \int f'(x)g(x)dx + \int f(x)g'(x)dx \end{aligned}$$

とお馴染みのものが求まる。

$y = f(u)$ ,  $u = g(x)$  なる関係で  $y$  と  $x$  が  $u$  を介して結ばれている合成関数の場合を考えると、

$$dy = f'(u)du, \quad du = g'(x)dx$$

であるから、

$$y \doteq \int \frac{dy}{du} \frac{du}{dx} dx \doteq \int f'(u)g'(x)dx \doteq \int f'(g(x))g'(x)dx$$

となる<sup>2</sup>。

## 4.2 定積分

不定積分は、導関数と原始関数との関係をあらわすもの、言い換えれば、関数と関数の対応関係をあらわすものであった。そこでその関係を利用して、関数と数値の対応を積分の世界に導入する。汎関数的な見方である。この数値との対応は、独立変数の定義域をごとに定まるものである。

### 4.2.1 定義か定理か

今、 $f(x)$  の不定積分  $\int f(x)dx$  において、 $x$  の変化する範囲つまり  $x$  の定義域を  $a \leq x \leq b$  と定めてみよう (ただし、 $a \leq b$  とする)。そして面倒な事柄をさけるため、この範囲で  $f(x)$  は微分可能 (したがって連続でもある) とする。この定義域での積分を

$$\int_a^b f(x)dx$$

と記すことにし、これを「定積分」と名づけよう。

さてここで、驚くべき事実に遭遇する。今、 $f(x)$  の原始関数が  $F(x)$  であるとする<sup>3</sup>。この原始関数を用いれば、上の定積分は

$$\int_a^b f(x)dx = F(b) - F(a)$$

<sup>2</sup>もちろん、順次代入して

$$y \doteq \int f'(u)du \doteq \int f'(u)g'(x)dx \doteq \int f'(g(x))g'(x)dx$$

と変形してもよい (備忘録)。

<sup>3</sup>逆に言えば、 $F(x)$  の導関数が  $f(x)$  ということである。すなわち、 $dF(x) = f(x)dx$  (または、 $F'(x) = f(x)$ )。

となるのである。これは定義だろうか？それとも何かから演繹される定理だろうか？一般的には、次の節で考える「Riemann 和の極限」で証明するという手段がとられる<sup>4</sup>。そしてこの関係を「微分積分学の基本定理」と呼ぶ場合もある。

簡略記法としては、次のものが用意されている。

$$\int_a^b f(x)dx = \left[ F(x) \right]_{x=a}^{x=b} = \left[ F(x) \right]_a^b = F(b) - F(a).$$

さて、なにゆえに「定」積分なのか？不定積分では、例によって  $C$  を任意の定数として

$$\int f(x)dx = F(x) + C$$

と原始関数に対して  $C$  だけの任意性があった。いまここで

$$G_1(x) := F(x) + C_1, \quad G_2(x) := F(x) + C_2 \quad (\text{もちろん } C_1 \neq C_2)$$

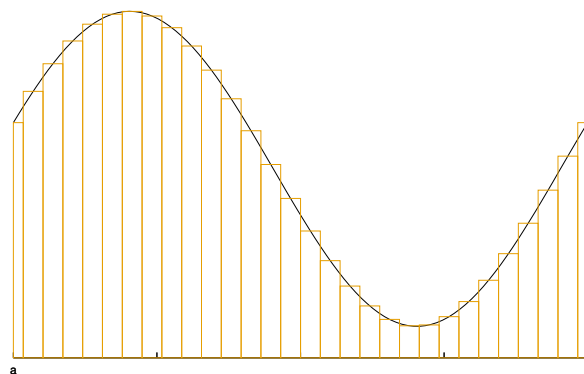
とおいてみよう。  $G_1(x)$ ,  $G_2(x)$  が  $f(x)$  の原始関数であることは明らかだ。すると定積分は、

$$\begin{aligned} \int_a^b f(x)dx &= G_1(b) - G_1(a) = (F(b) + C_1) - (F(a) + C_1) = F(b) - F(a) \\ \int_a^b f(x)dx &= G_2(b) - G_2(a) = (F(b) + C_2) - (F(a) + C_2) = F(b) - F(a) \end{aligned}$$

となり、不定積分について回った定数の任意性が消えてしまう。ゆえに「定積分」なのである。

## 4.2.2 Riemann 和の極限 ( $dx$ の寄せ集めのイメージ)

図形的なイメージからみてみよう。右の図は、ある連続でなめらかな関数（すなわち、微分可能な関数）  $f(x)$  を  $x=a$  から  $x=b$  まで描き、それを幅  $\Delta x$  の長方形で刻んだものである。おおよそのところ、各長方形は、幅  $\Delta x$  で高さ  $f(x)$  であるとみて差し支えないだろう。したがって、この長方形の面積の和は  $f(x)\Delta x$  の和であり、ほぼそれがこの関数が囲む図形の面積と近いだろうと予測できる。



ここで、区間を  $a = x_0 < x_1 < x_2 < \dots < x_n = b$  と  $n$  分割してみる（等幅でなくても良い）。さらに、  $\Delta x_i = x_{i+1} - x_i$  と表記することにする。上の説明では、高さがだいたい  $f(x)$  であるとしたけれども、その「だいたい」をなくすのが目的であるので、  $\Delta x_i$  の区間に対応する高さを  $h_i$  とひとまず定める。まだなんだかわからないがとにかく  $h_i$  とするのである。その時の長方形の和は、  $\sum \Delta x_i h_i$  である。

そして、  $f(x)$  の原始関数  $F(x)$  をもってくる。これは不定積分から持ってくるのが可能である。天下りの的に平均値の定理を使おう。平均値の定理によれば、

$$\frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} = F'(c_i)$$

<sup>4</sup>その証明は「平均値の定理」がもちいられ、平均値の定理は「ロルの定理を前提として」、というなにやらの積分学総動員の趣きのあるものである。



となる  $c_i$  が区間  $\Delta x_i$  の中に存在することが保証される。原始関数であることを利用すれば、

$$\frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} = F'(c_i) \iff \frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} = f(c_i) \iff F(x_{i+1}) - F(x_i) = f(c_i)(x_{i+1} - x_i)$$

である。  $c_i$  が区間内に存在するのでそれを高さに採用、つまり、  $h_i = f(c_i)$  とすることにする。

ここで、  $S = \sum_{i=0}^n f(c_i)\Delta x_i$  という和を考える。これが、Riemann 和と呼ばれるものである。計算をすすめると、

$$\begin{aligned} S &= \sum_{i=0}^n f(c_i)\Delta x_i = \sum_{i=0}^n f(c_i)(x_{i+1} - x_i) \\ &= \sum_{i=0}^n [F(x_{i+1}) - F(x_i)] = F(x_n) - F(x_0) \\ &= F(b) - F(a) \end{aligned}$$

この結果は区間の分割の仕方によらない。区間の個数や幅がなんであっても成立するものである。そして、定積分というものを Riemann 和の分割の極限と定義すると、

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} S = F(b) - F(a)$$

が証明されてくる。証明されてくるとは言っても、すこし戸惑いもある。まず Riemann 和が平均値の定理によって、どう区間を分割しても  $F(b) - F(a)$  になることは良い。その分割の極限を定積分と定義しただけではないのか、という疑問はぬぐいきれない。なんとなくすっきりしないものが残る<sup>5</sup>。

補足として、古典解析的に考えてみよう。  $dx$  は無限小という不思議な量であった。つまり、それは定量であるけれども、どこまででも 0 に近づくことができるものであった。長方形の分割の幅を  $dx$  とすれば、極限的には、長方形のごつごつがなめらかなつながりになるだろうと考えられる気がする。そしてその長方形の和をなめらかにした、と言う意味で、この「和」を

$$\int_a^b f(x)dx$$

と記すのである。定積分が、『 $f(x)$  と  $x$  軸によって囲まれる面積』であると認識できるのはこのイメージのおかげである。ただし、この古典的解析的な見方では、  $F(b) - F(a)$  は出てこない。やはりすっきりはしない<sup>6</sup>。

### 4.2.3 定積分の性質

$\int_a^b f(x)dx = F(b) - F(a)$  であることを諒解してきた。ここから、まず、

$$\int_a^a f(x)dx = F(a) - F(a) = 0$$

が導かれる。さらに、

$$-\int_a^b f(x)dx = -(F(b) - F(a)) = F(a) - F(b) = \int_b^a f(x)dx$$

も導ける。そしてさらに、

$$\int_a^c f(x)dx + \int_c^b f(x)dx = F(c) - F(a) + F(b) - F(c) = F(b) - F(a) = \int_a^b f(x)dx$$

<sup>5</sup>直接の関係はないが、小島 [4]70 ページの「青空ゼミナール」というコラムで、解析学の定理総動員による厳密性の獲得について書いてある。しかし、本当に「平均値の定理」から実数の公理が導き出せるのだろうか（もはや公理ではないよなあ、それ）。

<sup>6</sup>高瀬 [1] の 152 ページに「微分積分学の基本定理」の説明がある。この説明は示唆に富んでいる。

であることもわかる。

あとは、実際に原始関数を求めることに帰着する。原始関数が求まれば、定積分の値が計算できるようになる。

### 4.3 不定積分と定積分をつなぐもの

積分範囲に独立変数をおくことにより、不定積分と定積分をつなぎ合わせることができる。  $f(x)$  の原始関数を  $F(x)$  とすれば、

$$\int_a^x f(t)dt = [F(t)]_a^x = F(x) - F(a)$$

である。一方、不定積分においては、  $C$  を適当な（任意の）定数として、

$$\int f(x)dx = F(x) + C$$

となる。この2式から、

$$\int_a^x f(t)dt = F(x) - F(a) = \int f(x)dx + C - F(a)$$

が導かれ、  $C - F(a) =: D$  は定数であることに注目すれば、

$$\int_a^x f(t)dt + D = \int f(x)dx \iff \int_a^x f(t)dt \doteq \int f(x)dx$$

という関係が得られる。

### 4.4 積分の例題

具体的な例題は、高校の教科書・参考書・問題集に任せよう。そこにはいろいろな事例がのっている。特に修練が必要なものは部分積分と置換積分であろう。部分積分は、導関数のもととなる原始関数のあたりの付け方が重要で、そののちは慎重に計算を進めていけばよい。ややこしいのは置換積分で、ある種判じ物めいた様相が漂う。なにをどう置換するかは、修練するしかない。時には名人芸もある。一般的な目安としては、

$$\begin{aligned} (1) \quad (ax+b)^n &\implies t := ax+b \\ (2) \quad \sqrt{ax+b} &\implies t := \sqrt{ax+b} \\ (3) \quad \sqrt{a^2-x^2} &\implies x := a \sin \theta \left( -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2} \right) \\ (4) \quad (a^2+x^2)^n &\implies x := a \tan \theta \left( -\frac{\pi}{2} < \theta < \frac{\pi}{2} \right) \quad \text{or} \quad x := a \cdot \frac{e^t - e^{-t}}{2} = a \cdot \sinh t \end{aligned}$$

があげられる<sup>7</sup>。

<sup>7</sup>この諸法は、山本他共編『数学コメンタール』[3]から引用した。ちなみに  $\sinh t = -i \sin(it)$  である ( $\cosh t = \cos(it)$ )。

## 第5章 変分法

関数によって定まるあるものを「汎関数」と名付け、関数の微小な変化と汎関数の変化の関係を考察するものが変分法である。写像的には、汎関数は、関数に数値を対応させるものとして捉えられる。この変分の世界は、『関数の変化の量が微小であるときには古典解析と同様な公理構造を持つ』ものとして、構成される。変分法の応用で顕著なものは、汎関数が停留（極大、極小ももちろん含む）するための関数の条件を求めることである。これは微分法が、関数が停留するための変数の条件を求めていたことによく似ている。

ここで本章の議論をはっきりさせるために、次のことに留意しておこう。変数  $x$ ,  $y$ , などと言った時には、これは単なる変数であって、 $x$  が何か別の変数の関数であるということはない。 $x$  が関数である場合には、必ず「関数  $x$ 」と明示される<sup>1</sup>。

### 5.1 変分の世界の構築

#### 5.1.1 汎関数と変分量

変化のもととなるものが、変数  $x$  の場合、それに対応する変化量は関数  $f(x)$  であった。そしてそれらの変化量の微小なものを微分量と呼んだ。これを敷衍し、変化のもととなるものを関数  $f(x)$  として、それに対応する変化量を考えてみる。それを「汎関数」と呼ぶことにしよう。そしてそれらの変化量の微小なものを「変分量」と呼ぶことにする。これらの関数、汎関数、変分量をもとに、微分とはすこし様子の違うもうひとつの別の世界を構築されてゆく。

さてこの「汎関数」とは何か？ある意味素直に、「関数の関数」を指すものとして考えても差し支えはないと思う。つまり、関数を受け取り、その関数ごとにある値（数）を対応させる働き（写像である）を持つものと諒解して良いだろう<sup>2</sup>。さらに付け加えれば、関数を受け取るということは、その関数の定義域すべての関数値を受けとることでもあるので、関数の変数が実数ならばその関数値は「定義域の実数の個数分」あることになる。実数の個数などというものはひとつふたつと勘定できるものではない。無限個存在する。つまり、「無限個の変数をとる関数」とも見立てることが可能なのである。（もちろん、無限個でなく、有限個である場合があってもよい。離散的な汎関数とでも言えようか。第 6.2.2 節の「離散確率版」はその離散的な例である）。

ここではとりあえずこのようなものとして汎関数をとらえておこう<sup>3</sup>。

書き方は、 $x$  を変数としてもつ関数を  $f(x)$  として、汎関数を  $F[f]$ ,  $F[f(x)]$ ,  $F[f(*)]$  などと書く。最後の記法は目新しい<sup>4</sup>。

<sup>1</sup>こころ辺がわりとおおらか、というか、ルーズな書物が多いような気がする。学習者は時に、このおおらかさによって、混乱する。

<sup>2</sup>簡単な汎関数の例としては、定積分があげられる。 $\int_a^b f(x)dx$  は与えられた関数  $f(x)$  それぞれに（定積分の計算結果で導き出される）「対応する数」が決まる。

<sup>3</sup>これについての雑感を第 5.5 節に書いてみた。

<sup>4</sup>前野 [5] に倣った。この記法は、汎関数のもうひとつの見方としてある、変数が無限個の多変数関数という立場を尊重したものだ。実際、 $F[f(x)]$  においては、 $x$  はひとつに固定されているものではない。なかんずく、 $x$  を決めれば  $F[f(x)]$  が決まるというものでもない。 $f(x_0), f(x_1), \dots$  などのすべてのものが  $f(x)$  に集約されている、いいかえれば定義域の範囲のすべての  $x$  に対する  $f(x)$  を変数としてもつので、変数の個数が無限個の関数という見立てが成立するのだ。それを象徴的に書いた記号である。

## 5.1.2 微分と変分

微小な変化量として微分量を対象にして考察を進めたのが微分の世界であり、その微分量の対象は独立変数と関数であった。そして微分量には  $d$  なる記号が利用された。これと同様にして、微小な変化量として変分量を対象にし、変分量の対象を関数と汎関数にしたものが変分の世界である。変分量に対しては、 $\delta$  なる記号を使うのが慣習である。

変分世界の構築にあたっては、微分の世界で用いた同一の公理構造をこの変分世界にも適用する<sup>5</sup>。公理の適用の詳細、実際の計算方法などについては、第 5.2 節以降で述べる。

上記を踏まえて、微分と変分の世界の対応関係を図式化すれば、

$$\begin{array}{ccc}
 f : x & \longrightarrow & f(x) & \text{微分 } (dx, df) \\
 \vdots & & \vdots & \\
 F : f(x) & \longrightarrow & F[f(*)] & \text{変分 } (\delta f, \delta F)
 \end{array}$$

となる。関数がどちらの世界にも顔を出している。これからわかるように、微分と変分の世界の橋渡しとなるものが関数であり、関数が微分変分概念の理解の複雑さを少なからず招くことも否めない<sup>6</sup>。

## 5.2 変分量とその計算ルール（公理系）

### 5.2.1 変分量の記述

まず、関数  $f(x)$  を変化させるという事柄を考えてみよう ( $x$  は独立変数)。いま、 $f(x)$  が何らかの理由で変化したその関数を  $g(x)$  とする。これはすべての点  $x$  で  $f \rightarrow g$  となったこと、つまり、各点を部分的に取り出せば、

$$\begin{array}{c}
 \dots\dots \\
 f(x_0) \rightarrow g(x_0) \\
 f(x_1) \rightarrow g(x_1) \\
 f(x_2) \rightarrow g(x_2) \\
 \dots\dots
 \end{array}$$

となったということである。これが  $x$  に対して連続的に連なっている場合を想定しているのである。それぞれの各点  $x$  での変化分は  $g(x) - f(x)$  であることは諒解できよう。したがって、この変化分が微小な時  $\delta f(x)$  とあらわすことにすれば、

$$g(x) - f(x) = \delta f(x) \iff g(x) = f(x) + \delta f(x)$$

となる。そして、 $\delta f(x)$  は各点  $x$  ごとそれぞれに決まるのだから、 $x$  の関数であることも見て取れる。この  $\delta f(x)$  が関数の変分量である。

<sup>5</sup>第 2 章で導入した公理群は、いわば、微量に対する公理とも言えるもので、それを微分量に適用したのが「微分」、変分量に適用したのが「変分」と捉えて良いはずである。

<sup>6</sup>付け加えて、関数に対しては、微分に加えて積分も考察してきた。ならば、汎関数の変分に対応する「積分（汎関数積分）」というものも考えられるだろう。経路積分はその一例であるらしい。しかしながらそれは流石にこのノートのレベルを越えている。

これにより、汎関数の変分量は、関数の変分量もちいて

$$\delta F[f] = F[f + \delta f] - F[f]$$

と表されることになる。

## 5.2.2 変分の公理系

変分量に対して、微分の世界と同一の公理構造を用いて理路を構築する。記号を変えて、今一度公理を記してみよう。

公理 5.2.1.  $a$  を変化しない量 (定量) とするとき,  $\delta a = 0$ .

公理 5.2.2.  $\delta(f + g) = \delta f + \delta g$ .

公理 5.2.3.  $\delta(fg) = \delta f \cdot g + f \cdot \delta g$ .

汎関数の変分に対して、この公理の適用を試みよう。汎関数の変分量は、

$$\delta F[f] = F[f + \delta f] - F[f]$$

であった。したがって、

$$\begin{aligned} \delta(F[f] + G[f]) &= (F[f + \delta f] + G[f + \delta f]) - (F[f] + G[f]) \\ &= F[f + \delta f] - F[f] + G[f + \delta f] - G[f] \\ &= \delta F[f] + \delta G[f] \end{aligned}$$

となる。積の公理については、 $H[f] = F[f] \cdot G[f]$  とおき、変分量の定義から求まる  $F[f + \delta f] = \delta F[f] + F[f]$  などを用いれば、

$$\begin{aligned} \delta H &= H[f + \delta f] - H[f] \\ &= F[f + \delta f] \cdot G[f + \delta f] - F[f] \cdot G[f] \\ &= (\delta F[f] + F[f]) \cdot (\delta G[f] + G[f]) - F[f] \cdot G[f] \\ &= \delta F[f] \cdot G[f] + \delta F[f] \cdot G[f] + \delta F[f] \cdot \delta G[f] \end{aligned}$$

となり、変分量どうしの積は 0 である<sup>7</sup>として整理すると、

$$\delta(F[f] \cdot G[f]) = \delta F[f] \cdot G[f] + F[f] \cdot \delta G[f]$$

となって、汎関数の変分量も公理を満たすことがわかる。

---

<sup>7</sup>微分の時と同様 (2.1.3 節), この結果は公理からの要請である。

### 5.2.3 基本的な定理

まず素直にもとまる定理から.

定理 5.2.1.  $\alpha, \beta$  を定数とすると,

$$\delta(\alpha f + \beta g) = \alpha \cdot \delta f + \beta \cdot \delta g \quad (\text{線形性}).$$

証明.

$$\delta(\alpha f + \beta g) = \delta(\alpha f) + \delta(\beta g) = \delta\alpha \cdot f + \alpha \cdot \delta f + \delta\beta \cdot g + \beta \cdot \delta g = \alpha \cdot \delta f + \beta \cdot \delta g$$

□

定理 5.2.2.  $g \neq 0$  として,

$$\delta\left(\frac{f}{g}\right) = \frac{\delta f \cdot g - f \cdot \delta g}{g^2} \quad (\text{商の変分}).$$

証明.  $w = f/g$  すなわち  $wg = f$  として両辺を変分すると

$$\begin{aligned} \delta(wg) = \delta(f) &\iff \delta w \cdot g + w \cdot \delta g = \delta f &\iff \delta w = \frac{\delta f - w \cdot \delta g}{g} \\ &\iff \delta w = \frac{\delta f - \frac{f}{g} \cdot \delta g}{g} &\iff \delta w = \frac{\delta f \cdot g - f \cdot \delta g}{g^2}. \end{aligned}$$

□

次に合成関数についての変分を考えよう.  $f = f(u(x))$  という形の関数があるとする.  $f$  は関数であり.  $u$  ももちろん関数である. まず  $f$  自身が変化した変分量は,  $\delta f = (f + \delta f) - f$  であることは定義からも明らかだ. 次に,  $u \rightarrow u + \delta u$  と変化したときの  $f$  の変分はどうか, ということを考えてみる. この  $u$  の変化によって引き起こされる  $f$  の変分量も

$$\delta f = f(u + \delta u) - f(u) \tag{5.2.1}$$

とするのは素直であると思う. これは発見的な論であるが, いまここでこれを認めてしまうことにする. さらに.  $\delta u$  は変分量であるとはいえ「量」に変わりはないので,  $\delta u / \delta u = 1$  が認められる. したがって

$$\delta f = f(u + \delta u) - f(u) = (f(u + \delta u) - f(u)) \cdot 1 = (f(u + \delta u) - f(u)) \cdot \frac{\delta u}{\delta u} = \frac{f(u + \delta u) - f(u)}{\delta u} \cdot \delta u$$

となる. これは簡略化して, 次の様にも記述される (再度 (5.2.1) を利用)<sup>8</sup>.

$$\delta f = \frac{f(u + \delta u) - f(u)}{\delta u} \cdot \delta u = \frac{\delta f}{\delta u} \cdot \delta u.$$

この結果を得て,

$$\delta f = \delta f \cdot 1 = \delta f \cdot \frac{\delta u}{\delta u} = \frac{\delta f}{\delta u} \cdot \delta u$$

という操作的な見立ても認められる.

以上を前提として, 次の定理が成り立つ.

<sup>8</sup>しごく当り前の変形である. 約分すればトートロジー:-)

定理 5.2.3.  $f = f(u)$ ,  $u = u(v)$ ,  $v = v(x)$  なる関係の時,  $v \rightarrow v + \delta v$  と変化した時の  $f$  の変分は

$$\delta f = \frac{\delta f}{\delta u} \frac{\delta u}{\delta v} \delta v \quad (\text{合成関数の変分}).$$

証明. まず  $f$  に対して  $u \rightarrow u + \delta u$ , 次に,  $u$  に対して  $v \rightarrow v + \delta v$  という段取りで行くと,

$$\begin{aligned} \delta f &= f(u + \delta u) - f(u) = \frac{f(u + \delta u) - f(u)}{\delta u} \cdot \delta u = \frac{f(u + \delta u) - f(u)}{\delta u} \cdot (u(v + \delta v) - u(v)) \\ &= \frac{f(u + \delta u) - f(u)}{\delta u} \cdot \frac{u(v + \delta v) - u(v)}{\delta v} \cdot \delta v \\ &= \frac{\delta f}{\delta u} \frac{\delta u}{\delta v} \delta v. \end{aligned}$$

□

逆関数の変分に応用しよう.

定理 5.2.4.  $u(x) = f(v(x))$  の逆関数を  $v(x) = f^{-1}(u(x))$  と書くと ( $x$  は変数),  $u \rightarrow u + \delta u$  と変化した時の  $f^{-1}$  の変分は,

$$\delta f^{-1} = \frac{1}{\frac{\delta f}{\delta v}} \delta u \quad (\text{逆関数の変分}).$$

証明.  $v \rightarrow v + \delta v$  としたときの  $u$  の変分は,

$$\delta u = f(v + \delta v) - f(v) = \frac{f(v + \delta v) - f(v)}{\delta v} \delta v = \frac{\delta f}{\delta v} \delta v \quad \iff \quad \delta v = \frac{1}{\frac{\delta f}{\delta v}} \delta u$$

または一気に

$$\delta u = \delta f = \delta f \cdot 1 = \delta f \cdot \frac{\delta v}{\delta v} = \frac{\delta f}{\delta v} \delta v \quad \iff \quad \delta v = \frac{1}{\frac{\delta f}{\delta v}} \delta u$$

としてもいいだろう. 結果,  $\delta v = \delta f^{-1}$  だから

$$\delta f^{-1} = \frac{1}{\frac{\delta f}{\delta v}} \delta u$$

□

## 5.2.4 変分量の公理系を支える数学的事実

微小である変分量に対して, 次の事実をあらかじめ認めておく

$$e^{(\delta f)} - 1 = \delta f,$$

$$\sin(\delta f) = \delta f,$$

$$\cos(\delta f) = 1.$$

## 5.2.5 基本的な関数の変分の例

- 自然数による関数の巾乗

$$\delta(f^2) = \delta(ff) = \delta f \cdot f + f \cdot \delta f = 2f \delta f$$

これを数珠つなぎしていけば,  $n$  を自然数として,

$$\delta(f^n) = n f^{n-1} \delta f$$

を得る.

- 関数の三角関数

先に導入した, 変分における数学的事実を用いる.

$$\begin{aligned}\delta(\sin(f)) &= \sin(f + \delta f) - \sin(f) \\ &= \sin(f) \cos(\delta f) + \cos(f) \sin(\delta f) - \sin(f) \\ &= \sin(f) + \cos(f) \cdot \delta f - \sin(f) \\ &= \cos(f) \cdot \delta f.\end{aligned}$$

$$\begin{aligned}\delta(\cos(f)) &= \cos(f + \delta f) - \cos(f) \\ &= \cos(f) \cos(\delta f) - \sin(f) \sin(\delta f) - \cos(f) \\ &= \cos(f) - \sin(f) \cdot \delta f - \cos(f) \\ &= -\sin(f) \cdot \delta f.\end{aligned}$$

- 関数の指数関数

これも先に導入した, 変分における数学的事実を用いる.

$$\delta(e^f) = e^{f+\delta f} - e^f = e^f (e^{\delta f} - 1) = e^f \delta f.$$

- 関数の対数関数

$u = e^v$ ,  $v = \log u$  とすれば,  $u, v$  は逆関数の関係にある ( $u(v) = v(u) = 1$ ). したがって, 逆関数の変分の定理から,

$$\delta v = \frac{1}{\frac{\delta u}{\delta v}} \delta u = \frac{1}{\frac{e^v \delta v}{\delta v}} \delta u = \frac{1}{e^v} \delta u = \frac{1}{u} \delta u \iff \delta(\log u) = \frac{1}{u} \delta u.$$

- 実数による関数の巾乗

$\alpha$  を任意の実数とする.  $f^\alpha = \exp(\alpha \log f)$  であるから,

$$\delta(f^\alpha) = \delta(e^{\alpha \log f}) = e^{\alpha \log f} \cdot \delta(\alpha \log f) = e^{\alpha \log f} \cdot \alpha \cdot \frac{1}{f} \cdot \delta f.$$

再び  $f^\alpha = \exp(\alpha \log f)$  を使えば,

$$\delta(f^\alpha) = e^{\alpha \log f} \cdot \alpha \cdot \frac{1}{f} \cdot \delta f = f^\alpha \cdot \alpha \cdot \frac{1}{f} \cdot \delta f = \alpha f^{\alpha-1} \delta f.$$



## 5.3 微分と変分の関係

前節までで見たように、変分の演算は微分と非常に似ている、というか、その構造は同一であるといっても差し支えない。これは、まず第一に、基礎とおくものを「変化量」と抽象したこと、第二に、その変化量に対する公理は同一のものを採用していること（微量量についての公理として抽象化）、による。したがって、その公理から導出される論理構造は変わらない。ただ、対象とする数学的実体が、《数（変数）と関数》であるか、《関数と汎関数》であるか、が違うだけなのである。とはいえ、この2つは区別したい、それゆえ異なる記号  $d$  と  $\delta$  で区別をした。

しかしながら、どちらにも「関数」が関わっている。これを糊にして、微分と変分の世界をつなぐことを考えてみる。

### 5.3.1 形式と操作

#### § 変量がひとつの場合

まず変量がひとつの場合の  $L(q) = \exp(q) + q^2$  という記号列を元にして考えよう<sup>9</sup>。

最初にこの  $q$  が単なる変数である場合を想定する。変数  $q$  が  $q \rightarrow q + dq$  と変化した時、 $L$  の微分  $dL$  が考えられて、

$$dL = e^q \cdot dq + 2q \cdot dq \iff \frac{dL}{dq} = e^q + 2q$$

となった。

つぎにこの  $q$  が  $q = q(x)$  ( $x$  はただの変数) という関数である場合を想定する。関数  $q(x)$  が  $q(x) \rightarrow q(x) + \delta q(x)$  と変化した時、 $L$  の変分が考えられて、

$$\delta L = e^q \cdot \delta q + 2q \cdot \delta q \iff \frac{\delta L}{\delta q} = e^q + 2q$$

となった。

これより、微分演算操作、変分演算操作のどちらの場合でも、その結果の実態は形式的には変わらないことが見て取れる。対象の違い、演算操作の違いによって、記号  $d$  と  $\delta$  を使い分けているに過ぎないともいえる。もちろん微分と変分は概念的には似て非なるものであるが、「操作的」な結果の形式は微分も変分も何ら変わりはない、ということが言えてくるのである。

#### § 多変数の場合

次に、変量が2つの場合もみてみよう。変数  $q$  と  $p$  は互いに何の関係もないもの（つまり、各々独立）であるとして、 $L(q, p) = \exp(q) + q^2 + \log p$  なる記号列を考える。

$q, p$  が単なる変数である場合には  $q \rightarrow q + dq, p \rightarrow p + dp$  の変化のもとで全微分が考えられ

$$dL = \frac{\partial L}{\partial q} \cdot dq + \frac{\partial L}{\partial p} \cdot dp = (e^q + 2q) \cdot dq + \frac{1}{p} \cdot dp \quad \left( \text{i.e.} \quad \frac{\partial L}{\partial q} = e^q + 2q, \quad \frac{\partial L}{\partial p} = \frac{1}{p} \right).$$

<sup>9</sup>この節では、いままでのように  $f$  や  $g$  を使うのではなく、 $q, p$  という記号を使う。 $f, g$  にはどうも関数の意味合いが色濃く染み付いている気がするからである。

つぎにこの  $q$  が  $q = q(x)$ ,  $p$  が  $p = p(x)$  (やはり  $x$  はただの変数) という関数である場合を想定する. 関数  $q(x)$  と  $p(x)$  が何の関係もないことは同じである. そして  $q(x) \rightarrow q(x) + \delta q(x)$  と  $q(x)$  のみが変化した時の変分は,

$$\begin{aligned} L(q + \delta q, p) - L(q, p) &= \left( e^{(q + \delta q)} + (q + \delta q)^2 + \log p \right) - (e^q + q^2 + \log p) \\ &= \left( e^{(q + \delta q)} - e^q \right) + ((q + \delta q)^2 - q^2) \\ &= e^q \cdot \delta q + 2q \cdot \delta q = (e^q + 2q) \cdot \delta q \end{aligned}$$

ともとまる. この結果を, 暫定的な記号をもちいて, 次の様に記すことにしよう:

$$\frac{\hat{\partial} L}{\hat{\partial} q} := \frac{L(q + \delta q, p) - L(q, p)}{\delta q} = e^q + 2q.$$

同様に,  $p(x) = p(x) + \delta p(x)$  と  $p(x)$  のみが変化した時の変分は,

$$\begin{aligned} L(q, p + \delta p) - L(q, p) &= (e^q + q^2 + \log(p + \delta p)) - (e^q + q^2 + \log p) \\ &= \log(p + \delta p) - \log p \\ &= \frac{1}{p} \cdot \delta p \end{aligned}$$

となり, 先の暫定的な記号を用いれば,

$$\frac{\hat{\partial} L}{\hat{\partial} p} := \frac{L(q + \delta q, p) - L(q, p)}{\delta p} = \frac{1}{p}$$

とあらわせることになる.

これらから,  $q(x)$  と  $p(x)$  を同時に変化させた場合について,

$$\delta L = \frac{\hat{\partial} L}{\hat{\partial} q} \cdot \delta q + \frac{\hat{\partial} L}{\hat{\partial} p} \cdot \delta p$$

であるということを認める. 何か他の事実から演繹されるものではないけれども, 個別個別の変化を内在しているので, 十分に説得力のある事柄であろう<sup>10</sup>.

以上から, 変数が2つの場合でも, 微分演算操作, 変分演算操作の結果の実態は形式的には変わらないことが見て取れる. この理路では,  $\hat{\partial} L / \hat{\partial} q$  という記号を, 偏微分の意味合いと同様に, 『 $q(x)$  のみを変量とする, 言い換えれば,  $q(x)$  以外のものを定数とみなして変分をとる』という操作を示すものであるとした. 微分と変分は概念が異なるものであるから, この様に記号をきちんと区別していけばいいと思うのだけれども, 世の中そう几帳面にはできておらず, ここに  $\partial$  をそのまま用いる慣習があるようだ<sup>11</sup>. これが少なからず混乱をもたらす. それはそれとしてとにかく, 慣習に則れば, 変分が

$$\delta L = \frac{\partial L}{\partial q} \cdot \delta q + \frac{\partial L}{\partial p} \cdot \delta p$$

とあらわされることになるのである<sup>12</sup>.

<sup>10</sup>例えば,  $q(x)$  のみが変化するのであれば,  $p(x)$  は変化しないので,  $\delta p = 0$  である. ゆえに,  $\delta L(q, p) = \frac{\hat{\partial} L}{\hat{\partial} q} \cdot \delta q$  となり整合する. 問題はないだろう.

<sup>11</sup>単なる記号の節約と言えるのかもしれない. ただ学習者にとっては親切な慣習であるとは言えまい. 最初は, なかなかこのように割りきれないものだ.

<sup>12</sup>多変量の場合について, 微分の時と同様な言葉使いで, 全変分, 偏変分と名づけても良いのではと思うけれど, そういう名称はあまり使われない様だ.

もうひとつ、 $L$  に変数  $x$  があらわにあらわれる場合も考察しておこう。

$$L = L(q(x), p(x), x)$$

という形式である。この変数  $x$  についての全微分は（合成関数の微分ルールを適用して）

$$dL = \frac{\partial L}{\partial q} \cdot \frac{dq}{dx} \cdot dx + \frac{\partial L}{\partial p} \cdot \frac{dp}{dx} \cdot dx + \frac{\partial L}{\partial x} \cdot dx \quad \left( \Rightarrow \frac{dL}{dx} = \frac{\partial L}{\partial q} \cdot \frac{dq}{dx} + \frac{\partial L}{\partial p} \cdot \frac{dp}{dx} + \frac{\partial L}{\partial x} \right)$$

と求まる。一方、 $q(x) \rightarrow q(x) + \delta q(x)$ ,  $p(x) = p(x) + \delta p(x)$  と変化した時の変分は（変数  $x$  については変分はないことに留意）,

$$\delta L = \frac{\partial L}{\partial q} \cdot \delta q + \frac{\partial L}{\partial p} \cdot \delta p$$

となる。この形の  $L$  には、微分にはあって変分にはない項が存在するのである。

最後に、変数が  $n$  個の場合も（簡単に想像がつくであろうが）記しておく。

$$\delta L(q_1(x), q_2(x), \dots, q_n(x)) = \sum_{i=1}^n \frac{\partial L}{\partial q_i} \cdot \delta q_i = \sum_{i=1}^n \frac{\partial L}{\partial q_i} \cdot \delta q_i$$

### 5.3.2 微分と変分をつなぐもの

まず、直観的なイメージからとらえていこう。 $f(x)$  をその変分  $\delta f(x)$  だけ変化させたとする。つまり、

$$f(x) \mapsto f(x) + \delta f(x)$$

と対応させる。であれば、両辺の導関数をとってもその対応関係は保存するだろうと考えることには理があるはずである。微分の世界の導関数を示す記号に ' をもちいることにすると、

$$f'(x) \mapsto f'(x) + (\delta f(x))'$$

一方、先に  $f(x)$  の導関数を考え、それに、その変分  $\delta f'(x)$  だけ変化させた対応

$$f'(x) \mapsto f'(x) + \delta f'(x)$$

を考える。この両者は等しくなってほしい。すなわち、

$$(\delta f(x))' = \delta f'(x)$$

がなりたってほしい。そして実際にこれは成り立つ（成り立たせる、といった方が事実に近い）のだが、そう一筋縄ではいかない。

微分量  $df(x)$  の変分をとると、

$$\delta(df(x)) = \delta(f(x+dx) - f(x)) = \delta f(x+dx) - \delta f(x) = d(\delta f(x))$$

と求まる（途中で変分操作の和の分解の公理、微分量の定義を使った）。つまり、微分と変分の操作は順序交換可能なのである。これを簡略化して、 $\delta df(x) = d\delta f(x)$  と書く。つぎに、積の公理を用いてみよう。まず微分の世界では、 $df = f' dx$  であったから、

$$\delta df = \delta(f' dx) = \delta f' \cdot dx + f' \cdot \delta dx.$$

両辺を  $dx$  で割って

$$\frac{\delta df}{dx} = \delta f' + f' \frac{\delta dx}{dx}.$$

$\delta df = d\delta f$  であったから置き換えて,

$$\frac{d\delta f}{dx} = \delta f' + f' \frac{\delta dx}{dx}.$$

左辺の形式は導関数であることを示している. すなわち

$$(\delta f)' = \delta f' + f' \frac{\delta dx}{dx}.$$

さて,  $x$  が独立変数であるならば,  $dx$  は定数であった. したがってその場合,  $\delta dx = 0$  となり,  $(\delta f)' = \delta f'$  となる. また, 微分と変分の順序は入れ替えて良いのであったから,  $d\delta x = 0$  でもある. まとめると,

- 変化量にたいしては, 微分と変分の順序は変更可能である. つまり,  $\delta du = d\delta u$ .
- $x$  が独立変数である場合には,  $\delta dx = d\delta x = 0$ .
- 関数  $f(x)$  の変数  $x$  が独立変数であるならば,  $(\delta f(x))' = \delta f'(x)$ . 言い換えれば, 関数の変数が独立変数の時, 導関数の変分は変分の導関数に等しい.

となるだろう. これが微分と変分を橋渡す事実なのである.

## 5.4 積分と変分の関係

汎関数の例では, 積分を経由したものによく出会う (積分汎関数といわれる) ので, 積分と変分の関係 — 順序変更の関係 — をここで見ておこう. といっても実質は簡単である<sup>13</sup>.

簡単な場合から入ろう. 汎関数  $F$  として,

$$F[f] = \int f(x) dx$$

というものを考える. 変分を定義どおりに行えば,

$$\begin{aligned} \delta F[f] &= F[f + \delta f] - F[f] \\ &= \int \{f(x) + \delta f(x)\} dx - \int f(x) dx \\ &= \int \delta f(x) dx \end{aligned}$$

である. つまり,

$$\delta F[f] = \delta \int f(x) dx = \int \delta f(x) dx$$

であり, これから積分記号と変分記号の順序が交換できることがわかる.

<sup>13</sup>積分は足し算の「妖怪」であると考えれば, 和の分解の公理から, ある程度自明な事柄である, と言えるかもしれない. 何が「妖怪」なのかはうまく言えないけれど.

$f(x)$  をもとに形作られる関数（汎関数ではない）を  $L(f(x))$  と書こう<sup>14</sup>。そして汎関数として、 $F[f] = \int L(f(x)) dx$  を考えてみる。変分をとれば、

$$\begin{aligned}\delta F[f] &= F[f + \delta f] - F[f] \\ &= \int \{L(f(x) + \delta f(x))\} dx - \int L(f(x)) dx \\ &= \int \{L(f(x) + \delta f(x)) - L(f(x))\} dx\end{aligned}$$

であり、この積分の中身は  $L(f(x))$  の変分であるので、

$$\delta F[f] = \delta \int L(f(x)) dx = \int \delta L(f(x)) dx$$

となる。ここでも積分記号と変分記号の順序が交換できることが見て取れる。

複数の関数を変化量とする場合も同様である。独立な2個の関数  $f(x), g(x)$  を変化量とする  $L(f, g)$  に対する汎関数  $F[f, g] = \int L(f, g) dx$  の変分は

$$\delta F[f, g] = \delta \int L(f, g) dx = \int \delta L(f, g) dx = \int \left\{ \frac{\partial L}{\partial f} \delta f + \frac{\partial L}{\partial g} \delta g \right\} dx$$

となる。 $n$  個であれば、

$$\delta F[f_1, f_2, \dots, f_n] = \delta \int L(f_1, f_2, \dots, f_n) dx = \int \delta L(f_1, f_2, \dots, f_n) dx = \int \left\{ \sum_{i=1}^n \frac{\partial L}{\partial f_i} \delta f_i \right\} dx$$

である。これらの計算方法は Euler-Lagrange の方程式を求める際に大活躍する。

## 5.5 閑話休題 — 関数概念と汎関数 —

本章では、関数、汎関数を、

変化のもととなるものが変数  $x$  の場合、それに対応する変化量を関数  $f(x)$ 、  
変化のもととなるものが関数  $f$  であった場合にそれに対応する変化量を汎関数  $F[f]$

という意味合いで使ってきた。ここで、関数とは何ぞや、というところに思いを馳せてみる。

関数  $f$  とは何であったか。ことさら大げさに構えるつもりはないけれども、今一度この馴染みの深い言葉と、それを学んできた過程をさらってみる。ここをはっきりさせておくと、変分や汎関数が見通し良くなると思うからである。

はじめて関数にふれた頃は、説明として、「変化させる働き・変化させる機能」という比喩が使われていたように覚えている。数の「3」に働いて（機能して）「6」に変化させる、「4」に働きかけた結果は「8」である、というような現象がまず説明されて、さてその規則性は？と流れてゆき、この「働き・機能」は「2倍する」と言うものであろうと予測した。そこから、『では2倍するという働きを数式化してみましよう』というところで「変数」としての  $x$  が登場し、『この働きは  $2x$  である』と展開していったと記憶している。そしてこの  $x$  なるものはその中身はなんでもよく、逆に言えば、全ての数を代表して  $x$  としているにすぎないのである、と諒解した。「独立変数」との出会いのころである。

<sup>14</sup>例えば、 $L = (f(x))^2$  とか、 $L = -f(x) \log_e f(x)$  とか。

また、入力と出力という観点に注目して、「関数はブラックボックス」という比喩も使われていた。ブラックボックスの機能は「2倍する」ことであり、3を入力すれば6が出てきて、5を入力すれば10が出てくる、これを「任意の入力  $x$  に対して、出力を  $2x$  とするブラックボックス」とよび、それを関数ととらえていた。

数に対する「働き・機能」というものをもとにして、「関数」というものに馴染んで来た、そのついでに、入力の抽象化を  $x$  という文字で表現してきた、というのが今感じられる実感である。

次に、この「働き・機能」というものがだんだんと薄れてきて、数と数の対応関係に焦点が移ってきた様に記憶する。対応関係を一般化抽象化したものとしての「写像<sup>15</sup>」が登場し、関数は、

関数：数から、数への写像

とみなされてゆく。そして  $n$  変数関数（多変数関数）は、

関数： $n$  個の数の組から、数への写像

となってゆく。

多変数関数になると、変数が独立ということにも注目しなくならなくなる。 $n$  個の変数はそれぞれ独立である、ということは、各変数どうしにはなんの関係も存在しない、依存関係もなにもない、ということである。

原像となる「 $n$  個の数の組」を幾何学的に考えてみると、これは  $n$  次元空間の点（座標点）と同一視できる。ゆえに、

関数：点（座標点）から、数への写像

とも考えられる。実際、実数は数直線上のひとつの点であらわされるので、1変数関数もこの見立てにかなう。 $n$  変数関数は、 $n$  次元空間の点から数への写像となる。「点（座標点）」という言葉には、その点を構成するのはおのおの独立である、というニュアンスがこもっていることが多いので、便利でもある。

<sup>15</sup>ここで、写像について、定義と言葉を簡単にまとめておく。この手の数学的概念を規定する場合には、「集合論」の言葉は実に便利かつ有用である（「集合」とは何か、ということも本来必要なことではあると思うけれども、話は別次元になってしまうと思われるので、ここでは割愛）。

- 写像の定義  
空集合ではない集合  $A, B$  があって、 $A$  の「各要素すべて」に  $B$  の要素が「ただひとつ」対応している時、この対応を  $A$  から  $B$  への「写像 (mapping)」と名付ける。図式的には、写像を  $m$  と記して、次のように書いたりもする：

$$m: A \rightarrow B.$$

- 原像と像  
上記写像において、 $a \in A$  とした時、この  $a$  を「原像」と名付ける。 $a$  に対応する  $B$  の要素を「像」と名付ける。この  $a$  に対応する像を  $b_a (\in B)$  と記すことにすれば、図式的には、

$$m: a \rightarrow b_a, \quad m: a \rightarrow m(a), \quad b_a = m(a)$$

などとわりと自由に書いたりもする。

註1: 複数の原像が同じ像に対応しても構わない。写像の定義はそれを禁止していない。

- 定義域、値域  
定義域とは、集合  $A$  に他ならない。値域とは、像となるものすべてを集めた集合である。すなわち、値域を  $V$  とすれば、 $V = \{b_a \mid b_a = m(a), \forall a \in A\}$ 。

註2: 集合  $B$  の中には、像にならない要素があっても構わない。写像の定義はそれを禁止していない。したがって、 $V \subseteq B$ 。

- 全射（または、上への写像）  
像にならない  $b$  が存在しないこと。言い換えれば、すべての  $b (\in B)$  が像であること。つまり、 $V = B$ 。
- 単射（または、1対1写像）  
同じ像になる原像はひとつしかないこと。つまり、 $a_1 \neq a_2 \Rightarrow m(a_1) \neq m(a_2)$ （対偶をとれば  $m(a_1) = m(a_2) \Rightarrow a_1 = a_2$  である）。
- 全単射  
全射、かつ、単射。

像のバリエーションもいろいろ出てきた。上に述べてきたような「数」を像とするものは、スカラー関数と言われることがある。さらに、点にベクトルを対応させるつまり、像をベクトルとするものは物理でお馴染みであり、ベクトル場と言われたりもする。そういう多様な関数概念に出会ってきた。

少し戻る。

関数を「点から数への写像」とみなす所まで来た。ここで「点」を見直す。座標点には物理のイメージがついて回るが、それも忘れる。そして、関数がうじゃうじゃと存在している世界を考え、それを「関数空間」と名づけることにしよう。そして、この関数空間のなかのひとつの「点」がある関数をあらわすものとする。さらに想像を飛躍させ、関数空間の中の点は連続であるとする。この関数空間の中の各関数が、実数軸の点のように、連続にはびこっていると考えるのである<sup>16</sup>。そしてこの関数空間の点（つまり関数）から数への写像を考える。これは「関数」という概念そのものであるが、同じ名前を使うと紛らわしいので、「汎関数」と呼ぶことにするのである。すなわち、

汎関数：関数空間内の点（関数）から実数への写像

である。この世界での解析学が変分法に他ならない、と考えてよいのではないかと思う。

---

<sup>16</sup>この考えを成立させるためには、おそらく、関数と関数との「距離（間隔）」というものを明かにしなければならないだろう。位相構造の導入である。そこら辺をきちんとやるのが関数解析という分野なのだと思う。良い教科書は何だろうか？

## 第6章 微分と変分の実践

### 6.1 長方形の等周問題

長方形の周の長さを  $2L$  に固定したとき、その長方形の面積が最大になるのはどういう場合か、という「長方形の等周問題」を通して、いろいろな解法について考えてみる。なお以下、長方形の辺を  $x, y$  とし、面積を  $S$  とする。もちろん  $S = xy$  である<sup>1</sup>。そして  $x, y$  はただの変数なので、微分が使われる。

#### 6.1.1 2次関数

等周の条件は、 $2x + 2y = 2L$  であるから、 $y = L - x$ 。したがって面積  $S$  は

$$S = xy = x(L - x) = -x^2 + Lx = -\left(x - \frac{L}{2}\right)^2 + \frac{L^2}{4}$$

となり、この完全平方形から、 $x = L/2$  の時最大値  $L^2/4$  となる。

#### 6.1.2 微分の応用

面積を与える関数は  $S(x) = -x^2 + Lx$  である。これを微分して導関数を求めると、

$$\frac{dS}{dx} = -2x + L$$

となる。増減表を作れば、

$x$	...	$L/2$	...
$dS/dx$	+	0	-
$S$	↗	$L^2/4$	↘

なので、 $x = L/2$  の時最大値  $L^2/4$  となる。

#### 6.1.3 ベクトルの利用（未定乗数法一歩手前）

$S$  が最大値をとるとすれば、それは停留値でもあるのだから、全微分は  $0$ 、すなわち、

$$dS = \frac{\partial S}{\partial x} dx + \frac{\partial S}{\partial y} dy = 0 \iff \begin{pmatrix} \frac{\partial S}{\partial x} & \frac{\partial S}{\partial y} \end{pmatrix} \cdot \overrightarrow{(dx, dy)} = 0$$

<sup>1</sup>この証明は必要ですかねえ。というか、ここしばらく、面積という言葉には敏感になっているので、つい、不要ないらぬ心配をしてしまう。Riemann 和の定積分で証明するべきなのだろうか:-)



一方、等周条件を関数化して  $g(x,y)=x+y-L=0$  とすると、この全微分も

$$dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy = 0 \iff \overrightarrow{\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)} \cdot \overrightarrow{(dx, dy)} = 0$$

どちらも  $\overrightarrow{(dx, dy)}$  と直交するので、2次元の空間を舞台にすれば、

$$\overrightarrow{\left(\frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}\right)} = \lambda \cdot \overrightarrow{\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)}$$

が成立する。したがって、偏微分の実行結果を一緒に書けば、

$$\begin{aligned} \frac{\partial S}{\partial x} - \lambda \frac{\partial g}{\partial x} &= 0 \iff y - \lambda = 0 \\ \frac{\partial S}{\partial y} - \lambda \frac{\partial g}{\partial y} &= 0 \iff x - \lambda = 0. \end{aligned}$$

したがって、 $x=y=\lambda$  であり、これを等周条件  $2x+2y=2L$  にいれれば、 $\lambda=L/2$  となる。この時  $S$  が停留する。しかしながら、これだけでは最大値であることはまだ言えない。それは別方面からの検討が必要である。

#### 6.1.4 未定乗数法微分版

今、 $F(x,y)=S(x,y)-\lambda g(x,y)$  とし、 $F$  の全微分をとる。この未定乗数法使用時の留意点は、 $x,y,\lambda$  それぞれ全てを独立変数とみなすことである。すると、

$$\begin{aligned} dF &= dS - d(\lambda g) \\ &= \frac{\partial S}{\partial x} dx + \frac{\partial S}{\partial y} dy - d\lambda \cdot g - \lambda dg \\ &= \frac{\partial S}{\partial x} dx + \frac{\partial S}{\partial y} dy - d\lambda \cdot g - \lambda \frac{\partial g}{\partial x} dx - \lambda \frac{\partial g}{\partial y} dy \\ &= \left(\frac{\partial S}{\partial x} - \lambda \frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial S}{\partial y} - \lambda \frac{\partial g}{\partial y}\right) dy - d\lambda \cdot g \end{aligned}$$

となる。 $F$  が停留値をとる条件は  $dF=0$ 。それが任意の  $dx, dy, d\lambda$  で成り立つと考えるのだから、それぞれの係数は0でなければならない。すなわち、

$$\begin{aligned} \frac{\partial S}{\partial x} - \lambda \frac{\partial g}{\partial x} &= 0, \\ \frac{\partial S}{\partial y} - \lambda \frac{\partial g}{\partial y} &= 0, \\ g &= 0. \end{aligned}$$

あとは、6.1.3 節と同様である。

## 6.2 確率変数のエントロピー

エントロピーが最大となる確率分布を求めることが目的である。離散版と連続版両方を実践しよう。

## 6.2.1 離散確率版

離散確率変数  $p_i$  ( $i = 1, \dots, n$ ) のエントロピーは、 $-\sum_{i=1}^n p_i \log_2 p_i$  とあらわされることが知られている。ここでは、この確率変数  $p_i$  を何がしかの関数であると見立てる<sup>2</sup>。そしてこのエントロピー  $H$  を  $p_i$  の汎関数にとらえ、それが最大となる場合には確率変数の各々の分布はどうなるのか、という問題を考える。最大であるということは、少なくとも停留するはずだ。さらに、確率変数の大前提である  $\sum_{i=1}^n p_i = 1$  という条件はつねに成立する。これを拘束条件として、変分法と未定乗数法<sup>3</sup> を使って実践してみる。

汎関数として捉えられるエントロピー  $H$  を

$$H[p_*] = -\sum_{i=1}^n p_i \log_2 p_i$$

と書くことにしよう。さてここから、汎関数  $F$  として、拘束条件に未定乗数  $\lambda$  をかけたもの

$$F[p_*, \lambda] = H[p_*] - \lambda \left( \sum_{i=1}^n p_i - 1 \right) = -\sum_{i=1}^n p_i \log_2 p_i - \lambda \left( \sum_{i=1}^n p_i - 1 \right)$$

を考えていく。解析的計算をスムーズにするために、すこし式を操作しよう。底が 2 の対数は解析には向いていないので、対数の底変換を施し、 $\log_2 p_i = \log_e p_i / \log_e 2 =: A \log_e p_i$  とする ( $A = 1/\log_e 2$  とした)。したがって、

$$F[p_*, \lambda] = -\sum_{i=1}^n A p_i \log_e p_i - \lambda \left( \sum_{i=1}^n p_i - 1 \right)$$

であり、変分は、

$$\delta F = \frac{\partial F}{\partial p_1} \delta p_1 + \dots + \frac{\partial F}{\partial p_n} \delta p_n + \frac{\partial F}{\partial \lambda} \delta \lambda$$

となる。計算しよう。まず、

$$\frac{\partial F}{\partial \lambda} = \sum_{i=1}^n p_i - 1$$

であることはすぐに分かる。次に、 $\partial F / \partial p_j$  を計算する。それぞれの  $p_i$  は独立である所に留意して、

$$\begin{aligned} \frac{\partial F}{\partial p_j} &= \frac{\partial}{\partial p_j} \left\{ -\sum_{i=1}^n A p_i \log_e p_i - \lambda \left( \sum_{i=1}^n p_i - 1 \right) \right\} \\ &= -\sum_{i=1}^n A \frac{\partial}{\partial p_j} \{ p_i \log_e p_i \} - \lambda \sum_{i=1}^n \frac{\partial}{\partial p_j} p_i \\ &= -A \frac{\partial}{\partial p_j} \{ p_j \log_e p_j \} - \lambda \frac{\partial}{\partial p_j} p_j \\ &= -A (\log_e p_j + 1) - \lambda \end{aligned}$$

ともとまる。以上から、変分の最終形は

$$\delta F = \sum_{i=1}^n \{ -A (\log_e p_i + 1) - \lambda \} \delta p_i + \left( \sum_{i=1}^n p_i - 1 \right) \delta \lambda$$

<sup>2</sup>この見立てによって変分法を用いるのであるが、この問題は、 $p_i$  を単なる変数とみなして微分を用いることによっても、同じように解くことができる。

<sup>3</sup>Lagrange の未定乗数法。

となる。停留の必要条件は  $\delta F = 0$ 。これが任意の  $\delta p_i$ ,  $\delta \lambda$  で成立するには、

$$\begin{aligned} -A(\log_e p_i + 1) - \lambda = 0 &\iff \log_e p_i = -\lambda \log_e 2 - 1 \quad (\text{for all } i = 1, \dots, n), \\ \sum_{i=1}^n p_i - 1 = 0. \end{aligned}$$

(6.2.1) は、すべての  $p_i$  が同じ値を持つことを意味している。その値を  $p$  として拘束条件である (6.2.1) にあてはめれば、

$$\sum_{i=1}^n p_i - 1 = \sum_{i=1}^n p - 1 = p \sum_{i=1}^n 1 - 1 = np - 1 = 0 \iff p = \frac{1}{n}$$

すなわち、すべて等確率の一様分布がエントロピー最大の確率分布となるのである。

「変分法の実践」であるから、丁寧に  $\delta/\delta p_i$  をとらないで、一挙に変分をとる方法も練習してみよう。

$$F = -\sum_{i=1}^n A p_i \log_e p_i - \lambda \left( \sum_{i=1}^n p_i - 1 \right)$$

であったから、変分は

$$\delta F = -\delta \left( \sum_{i=1}^n A p_i \log_e p_i \right) - \delta \left( \lambda \left( \sum_{i=1}^n p_i - 1 \right) \right)$$

となる。細かい計算を先にしておこう。まず変分に対しての和の分解と積の展開から、

$$\begin{aligned} \delta \left( \sum_{i=1}^n A p_i \log_e p_i \right) &= \sum_{i=1}^n \delta (A p_i \log_e p_i) = \sum_{i=1}^n A (\delta p_i \cdot \log_e p_i + p_i \cdot \delta(\log_e p_i)) \\ &= \sum_{i=1}^n A (\delta p_i \cdot \log_e p_i + \delta p_i) = \sum_{i=1}^n A (\log_e p_i + 1) \delta p_i \end{aligned}$$

となる。もう一つの項の変分は、未定乗数法を用いているので  $\lambda$  も独立変数として、

$$\delta \left( \lambda \left( \sum_{i=1}^n p_i - 1 \right) \right) = \delta \lambda \cdot \left( \sum_{i=1}^n p_i - 1 \right) + \lambda \cdot \delta \left( \sum_{i=1}^n p_i - 1 \right) = \delta \lambda \cdot \left( \sum_{i=1}^n p_i - 1 \right) + \lambda \cdot \sum_{i=1}^n \delta p_i.$$

まとめると、

$$\begin{aligned} \delta F &= -\sum_{i=1}^n A (\log_e p_i + 1) \delta p_i - \delta \lambda \cdot \left( \sum_{i=1}^n p_i - 1 \right) - \lambda \cdot \sum_{i=1}^n \delta p_i \\ &= -\sum_{i=1}^n \{A (\log_e p_i + 1) + \lambda\} \delta p_i - \delta \lambda \cdot \left( \sum_{i=1}^n p_i - 1 \right) \end{aligned}$$

となりこれは、(6.2.1) に他ならない。

## 6.2.2 連続確率版

確率変数が連続であるときは、その確率変数を  $p(x)$  として、エントロピーは

$$H[p(*)] = -\int p(x) \log_2 p(x) dx$$

で与えられる。拘束条件は,

$$\int p(x)dx = 1$$

である。この問題は、変量が  $p(x)$  であるので、変分を適用するしかない。未定乗数法を使う所は離散確率の場合と同じであり、計算も  $\Sigma$  が  $\int$  に変わるくらいで実質は一緒である。留意すべきは、変分  $\delta$  が積分  $\int$  の中に入り込めることである。

離散確率のときと同様に  $A := 1/\log_e 2$  とすれば  $H[p(*)] = -\int A \cdot p(x) \log p(x) dx$  である。ここで、未定乗数を  $\lambda$  として、

$$F[p(*), \lambda] = H[p(*)] - \lambda \left( \int p(x) dx - 1 \right)$$

を考え、変分を計算すると

$$\begin{aligned} \delta F[p(*), \lambda] &= \delta H[p(*)] - \delta \left\{ \lambda \left( \int p(x) dx - 1 \right) \right\} \\ &= -\delta \int A \cdot p(x) \log p(x) dx - \delta \lambda \cdot \left( \int p(x) dx - 1 \right) - \lambda \cdot \delta \int p(x) dx \\ &= -\int A \cdot \delta \left( p(x) \log p(x) \right) dx - \delta \lambda \cdot \left( \int p(x) dx - 1 \right) - \lambda \int \delta p(x) dx \\ &= -\int A \cdot \left\{ \delta p(x) \cdot \log p(x) + p(x) \cdot \frac{1}{p(x)} \delta p(x) \right\} dx - \lambda \int \delta p(x) dx - \delta \lambda \cdot \left( \int p(x) dx - 1 \right) \\ &= -\int A \cdot \left\{ \delta p(x) \cdot \log p(x) + \delta p(x) + \frac{\lambda}{A} \delta p(x) \right\} dx - \delta \lambda \cdot \left( \int p(x) dx - 1 \right) \\ &= -\int A \cdot \left\{ \log p(x) + 1 + \frac{\lambda}{A} \right\} \delta p(x) dx - \delta \lambda \cdot \left( \int p(x) dx - 1 \right) \end{aligned}$$

となる。停留値の場合には、 $\delta F = 0$  であり、それが任意の  $\delta p(x)$ ,  $\delta \lambda$  に対して、成り立つには、

$$\begin{aligned} \log p(x) + 1 + \frac{\lambda}{A} = 0 &\iff \log p(x) = -\frac{\lambda}{A} - 1, \\ \int p(x) dx - 1 = 0, \end{aligned}$$

でなければならないことがわかる。結果、 $p(x)$  が  $x$  によらず同じ値を持つことが示され、ここでも  $p(x)$  が一様分布となることがわかった。

### 6.3 2点間の最短経路

2次元平面上において、2点間を結ぶ曲線でその曲線の長さが最小になるものを求める。とはいっても、それが直線であることはあきらかであると習ってきた。そのあきらかさを変分法で確認してみる。

2点を  $P(x_p, y_p)$ ,  $Q(x_q, y_q)$  としこの2点を結ぶ曲線を対象にすることにしよう (右図 6.1 参照)。さらに、その曲線の線素を  $ds$  とする。 $ds$  と  $dx$ ,  $dy$  で構成される直角三角形 (いふなれば、無限小直角三角形) においては、直角三角形であることからピタゴラスの定理、

$$(ds)^2 = (dx)^2 + (dy)^2$$

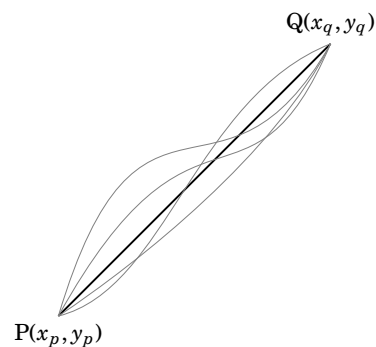


図 6.1: 2点間を結ぶ経路

が成り立つ（2次元平面はユークリッド空間であることを前提としているのでこれは正当化されよう）。ただし、ここでは  $ds, dx, dy$  すべて長さと考えて、0 より大きい量であるとする。微分量にたいしてもこういうことをしているのか、と問われたら、「小さいだけで、数であることに変わりはない<sup>4</sup>」ので問題はない（という立場を取っている）。すると曲線  $PQ$  の長さ  $l$  は

$$l = \int_P^Q ds$$

となる。道具立てはここまで。あとは計算するのみである。

まず、(6.3) から、

$$(ds)^2 = (dx)^2 + (dy)^2 \iff ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.$$

曲線を、 $y(x)$  と表すことにして、諸々表記を簡単にする。すなわち、 $dy/dx = y'(x) = y'$  として、

$$ds = \sqrt{1 + (y')^2} dx$$

であり、積分範囲は  $(P \rightarrow Q) \implies (x_p \rightarrow x_q)$  と移るので

$$l = \int_P^Q ds = \int_{x_p}^{x_q} \sqrt{1 + (y')^2} dx$$

となる。 $y(x) \rightarrow y(x) + \delta y(x)$  と微小量ずらした時の変分を計算しよう。 $y$  の変分は  $\delta y$  であり、点  $P, Q$  は固定して考えるので、 $\delta y(x_p) = 0, \delta y(x_q) = 0$  であることが必要である。その上で  $l$  の変分は、

$$\delta l = \delta \int_{x_p}^{x_q} \sqrt{1 + (y')^2} dx = \int_{x_p}^{x_q} \delta \left( \sqrt{1 + (y')^2} \right) dx$$

と変形できる。

ここでいきなり積分内の関数について変分計算を行ってもいいのだけれども、それはまあ熟練者コース。足元から行ってみよう。問題は  $y$  が変化した時の  $y'$  の変化を変分でどう表現できるかにある。 $y \rightarrow y + \delta y$  の時、 $(y')^2$  はどう変化するかみてみよう。変分の定義から、

$$\begin{aligned} (y')^2 &\longrightarrow ((y + \delta y)')^2 - (y')^2 = (y' + (\delta y)')^2 - (y')^2 \\ &= (y')^2 + 2y'(\delta y)' + ((\delta y)')^2 - (y')^2 \\ &= 2y'(\delta y)' + ((\delta y)')^2 \end{aligned}$$

$\delta$  の2乗の分は無視する。ゆえに、 $(y')^2$  の変分は、微分と変分の順序は変更できるので、

$$\delta (y')^2 = 2y'(\delta y)' = 2y'\delta y'$$

となる。この事実の裏打ちのもと、熟練者は、 $y'$  をあたかも変分の直接対象の関数であるとして、ただの変分操作のみで、

$$\delta (y')^2 = 2y'\delta y'$$

と書き下すのである。これはなにも  $(y')^2$  に限らず、一般の  $f(y')$  なる関数に対しても適用できる熟練法、すなわち、 $y$  を変分させるときにおける、 $y'$  の変分をとる操作法なのである。

<sup>4</sup>違いは独立変化量であること、すなわち、 $d(dx) = 0, d(dy) = 0$  であることのみと言い切って良いだろう。

本題である (6.3) の計算にもどる。以下の変分計算においては、考えている変分は  $y \rightarrow y + \delta y$  であるけれども、今見たように、 $y'$  を直接変分の対象として進めて問題はない。実行しよう。後半で部分積分を用いていることに留意し、点  $P, Q$  固定の要請  $\delta y(x_p) = 0, \delta y(x_q) = 0$  をもちいて

$$\begin{aligned} \delta l &= \int_{x_p}^{x_q} \delta \left( \sqrt{1+(y')^2} \right) dx = \int_{x_p}^{x_q} \delta \left( (1+(y')^2)^{\frac{1}{2}} \right) dx \\ &= \int_{x_p}^{x_q} \frac{1}{2} (1+(y')^2)^{-\frac{1}{2}} \delta ((y')^2) dx = \int_{x_p}^{x_q} \frac{1}{2} (1+(y')^2)^{-\frac{1}{2}} 2y' \delta y' dx \\ &= \int_{x_p}^{x_q} \frac{y'}{\sqrt{1+(y')^2}} \delta y' dx = \int_{x_p}^{x_q} \frac{y'}{\sqrt{1+(y')^2}} \cdot \delta \frac{dy}{dx} \cdot dx \\ &= \int_{x_p}^{x_q} \frac{y'}{\sqrt{1+(y')^2}} \cdot \frac{d}{dx} \delta y \cdot dx \\ &= \left[ \frac{y'}{\sqrt{1+(y')^2}} \delta y \right]_{x_p}^{x_q} - \int_{x_p}^{x_q} \frac{d}{dx} \left( \frac{y'}{\sqrt{1+(y')^2}} \right) \delta y dx \\ &= - \int_{x_p}^{x_q} \frac{d}{dx} \left( \frac{y'}{\sqrt{1+(y')^2}} \right) \delta y dx \end{aligned}$$

を得る。求めたいものは、曲線の長さ  $l$  の最小値であった。最小値であれば少なくともそれは極値であるので、 $\delta l = 0$  であることが必要。従って

$$\forall \delta y, \delta l = 0 \iff \frac{d}{dx} \left( \frac{y'}{\sqrt{1+(y')^2}} \right) = 0 \iff \frac{y'}{\sqrt{1+(y')^2}} = Const.$$

となる ( $Const.$  は定数という事実を示す)。これは  $y' = Const.$  という事にほかならない<sup>5</sup>。つまり、 $y(x)$  は導関数が定数、すなわち、 $x$  の一次式つまり「直線」なのである。

$\delta l = 0$  の時、 $y' = Const.$  と求まったのだが、さてこれは今のところ極値であるのみだ。本当に最小値だろうか？というところで、もとの  $l$  を見直すと、

$$l = \int_{x_p}^{x_q} \sqrt{1+(y')^2} dx$$

であったので、定数として  $C$  をもちいておくと、

$$l = \int_{x_p}^{x_q} \sqrt{1+C^2} dx$$

である。一方、 $y \rightarrow y + \delta y$  の時、 $y' \rightarrow y' + \delta y'$  だったから、それは  $C^2$  に何かしらの量を加えたものになるはずだ。それを  $C^2 + D$  としてみる。すると  $y$  を変化した後の長さは、

$$l + \delta l = \int_{x_p}^{x_q} \sqrt{1+C^2+D} dx$$

で、積分される関数は大きくなっている。つまりずらせば  $l$  は大きくなるので、ずらすまえは最小ということになって、めでたしである。

<sup>5</sup>直感的に明らか、としたいが、最初は戸惑った。この式は、 $y' = C\sqrt{1+(y')^2}$  であるから、両辺2乗すれば、

$$(y')^2 = C^2(1+(y')^2) \iff (1-C^2) \cdot (y')^2 = C^2 \iff (y')^2 = \frac{C^2}{1-C^2}.$$

これが定数でなくてなんであろうか？

## 6.4 Fermat の原理 (幾何光学)

「光は、最短時間となる経路を通過して走る」というのが、Fermat の原理の大意である。もし光の速さが一定であれば、「光は、最短の経路を通過して走る」と同義になる。そして光の速さは、空間の様相<sup>6</sup>が同一ならば、どんな慣性座標系でも一定らしい<sup>7</sup>。物理に深入りするのはやめよう。Fermat の原理を認めると、6.3 節で見たように、2次元平面上において2点間を結ぶ最小の曲線は直線であったので、その2点間の光の経路は直線であるということがわかる。それをもとにして、光の反射と屈折における経路についての法則を、変分を用いて導出してみる。

### 6.4.1 反射の法則

次の図 6.2 のように、点 P から  $x$  軸上の点 R で反射して、点 Q に至る光の経路を考えてみよう。Fermat

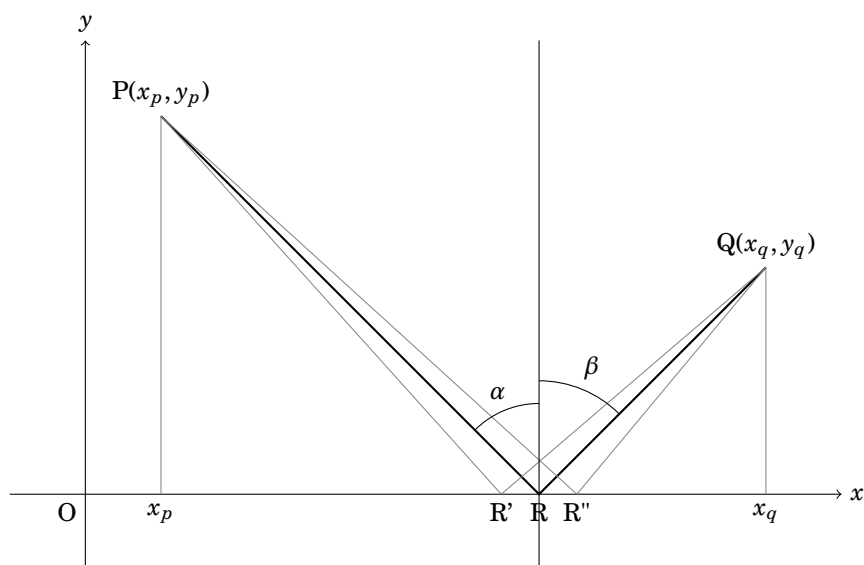


図 6.2: 光の反射の図

の原理から、数ある経路のうち最短のものが選択されることになる。つまり、直線 PR の長さと同様に直線 RQ の長さの和が最短の経路が選ばれる。われわれはその最短のものを探したい。最短を求めるのであるからこれは極値問題である。変分法が利用できる。

各点の座標を、 $P(x_p, y_p)$ 、 $Q(x_q, y_q)$ 、 $R(x, 0)$  とする。変分法の経路は、

PRQ の経路が最短だと仮定する。すると、ちよつとずれた経路、たとえば、PR'Q や PR''Q はそれよりは長いはずである。PRQ が極値であることが示せば、この仮定は成り立つ。ゆえに、その極値となる条件を変分法で求める。

<sup>6</sup>波動的には「媒質」という。ただ、我々はすでに光が単なる波でないこと、「エーテル」は不必要なことを知っている。なのでここでは「空間の様相」と書いた。たとえば、真空中、宇宙空間、大気中と水中ではそれぞれの空間様相がことなる、そういうことを言いたいのである。

<sup>7</sup>相対論で名高い「光速不変の原理」である。「原理」であることに注意。自然はそうになっている。

というものだ。今の場合、変化量は  $x$  のみであり、それはただの変数であるので微分を使う。経路の長さを  $l(x)$  として、 $x$  の変化に対して停留するものを微分で求める。計算を始めよう。

$$l(x) = PR + RQ = \sqrt{(x-x_p)^2 + y_p^2} + \sqrt{(x_q-x)^2 + y_q^2}$$

であるから、微分をとれば、

$$dl(x) = \frac{x-x_p}{\sqrt{(x-x_p)^2 + y_p^2}} \cdot dx - \frac{x_q-x}{\sqrt{(x_q-x)^2 + y_q^2}} \cdot dx = \left( \frac{x-x_p}{\sqrt{PR}} - \frac{x_q-x}{\sqrt{RQ}} \right) \cdot dx.$$

$dl(x)/dx = 0$  が極値の条件であった。したがって、

$$\frac{x-x_p}{\sqrt{PR}} - \frac{x_q-x}{\sqrt{RQ}} = 0 \iff \frac{x-x_p}{\sqrt{PR}} = \frac{x_q-x}{\sqrt{RQ}}$$

でなければならないことが要請される。さて、これは何を意味しているか？これは、直角三角形  $PRx_p$  と  $RQx_q$  について、斜辺ともう一つの辺の比がそれぞれ等しいことを示している。その場合、2つの直角三角形は相似になる<sup>8</sup>。その相似関係から  $\angle PRx_p = \angle QRx_q$  となり、したがって、図中の角度  $\alpha, \beta$  も  $\alpha = \beta$  となる。一般にこの  $\alpha$  を「入射角」、 $\beta$  を「反射角」という。まとめると、Fermat の原理を認めれば『光の反射の際の経路では、入射角と反射角は等しい』ということになる。

蛇足的補足をひとつ。いまの問題で、微分を求める際に平方根が入っているのは面倒だからといって、この微分の計算の対象となる量を

$$m(x) := \left( (x-x_p)^2 + y_p^2 \right) + \left( (x_q-x)^2 + y_q^2 \right)$$

としてはならない。これは  $PR$  を一辺とする正方形と  $RQ$  を一辺とする正方形の面積の和を表している（次元的にも、長さの2乗である）。したがって、微分をとると、その面積の和が最小値となる  $R$  がもとまるのである。最短経路ではない。実際に計算をすすめてみると、

$$dm(x) = 2x \cdot dx - 2x_p \cdot dx + 2x \cdot dx - 2x_q \cdot dx = (2x - x_p - x_q) \cdot 2dx$$

で、極値の条件  $dm(x)/dx = 0$  が成り立つためには、

$$2x - x_p - x_q = 0 \iff x = \frac{x_q - x_p}{2}$$

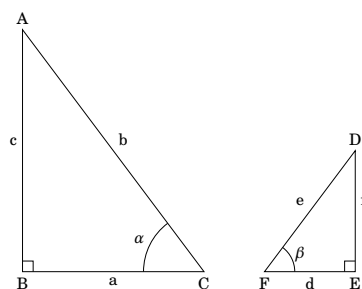
となる。つまり、 $R$  は  $x_p$  と  $x_q$  の中点になり、その時2つの正方形の面積の和が最小となるのである。横着は往々にして、問題を変更してしまう。

8

この相似関係は次のようにして確認できる。今、右の2つの直角三角形  $ABC$  と  $DEF$  において、斜辺と底辺の間の長さの関係として、 $\frac{a}{b} = \frac{d}{e}$  が成り立っているとす。そして図のように角  $\alpha$  と  $\beta$  を決める。直角三角形と言えば三角関数である。三角関数の定義から、

$$\cos \alpha = \frac{a}{b}, \quad \cos \beta = \frac{d}{e}$$

である。一方両式の右辺は、等しいのであった。したがって、 $\cos \alpha = \cos \beta$ 。三角形の内角は  $\pi$  より小さいという制限があるので、この場合  $\alpha = \beta$  となる。となれば、三角形  $ABC$  と  $DEF$  は全ての角が等しいことになり、相似の関係にあることが言える。直角三角形の場合は、このように辺の比に着目すれば相似であることがわかるのである。





もうひとつ余分な蛇足を<sup>9</sup>。  $x$  は変数であるけれども、これを無理やりに変分の対象として計算してみる。すなわち、  $l(x)$  において  $x$  での変分をとれば、

$$\delta l(x) = \frac{x-x_p}{\sqrt{(x-x_p)^2+y_p^2}} \cdot \delta x - \frac{x_q-x}{\sqrt{(x_q-x)^2+y_q^2}} \cdot \delta x = \left( \frac{x-x_p}{\sqrt{PR}} - \frac{x_q-x}{\sqrt{RQ}} \right) \cdot \delta x.$$

$\delta l(x) = 0$  が極値の条件であった。それが任意の  $\delta x$  で成り立つためには、

$$\frac{x-x_p}{\sqrt{PR}} - \frac{x_q-x}{\sqrt{RQ}} = 0 \iff \frac{x-x_p}{\sqrt{PR}} = \frac{x_q-x}{\sqrt{RQ}}$$

でなければならないことになる。結果は同じである。これは、物理の世界で「仮想変位を  $\delta x$  とする」などと言う場合によく用いられる方法である。本来は変数であるので変分を使うべきではないと思うけれども、とにかく、こういう流儀も存在する。物理の世界はおおらかである。

### 6.4.2 屈折の法則

屈折の法則とよばれているものは、光が、様相の異なる2つの空間（波動的に言えば、異なる媒質で満たされている2つの空間）を通る場合に成り立つ法則のことである。そして、物理的には、様相のことなる空間では光の速さも異なる、という事実がある。この事実と Fermat の原理から、屈折の法則を導いてみよう。

この問題を図表化したものが図 6.3。  $y = 0$  を境界として、  $y > 0$  の空間は様相が均一でそこでの光の速さを  $c$  とする。同様に、  $y < 0$  の空間も様相は均一でそこでの光の速さを  $v$  とする。Fermat の原理は光は最短時間の経路を通るというものであった<sup>10</sup>。  $PR$  を光が進む時間を  $t_{PR}$ 、  $RQ$  のそれを  $t_{RQ}$  としよう。各々は、

$$t_{PR} = \frac{\sqrt{(x-x_p)^2+y_p^2}}{c}, \quad t_{RQ} = \frac{\sqrt{(x_q-x)^2+y_q^2}}{v}$$

となることは明らか。したがって、この  $PRQ$  と進む時の合計時間を  $t(x)$  とすれば、

$$t(x) = t_{PR} + t_{RQ} = \frac{\sqrt{(x-x_p)^2+y_p^2}}{c} + \frac{\sqrt{(x_q-x)^2+y_q^2}}{v}$$

とあらあせる。変数  $x$  に対する微分をとれば、

$$dt(x) = \frac{1}{c} \frac{x-x_p}{\sqrt{(x-x_p)^2+y_p^2}} \cdot dx - \frac{1}{v} \frac{x_q-x}{\sqrt{(x_q-x)^2+y_q^2}} \cdot dx = \left( \frac{1}{c} \frac{x-x_p}{\sqrt{(x-x_p)^2+y_p^2}} - \frac{1}{v} \frac{x_q-x}{\sqrt{(x_q-x)^2+y_q^2}} \right) \cdot dx$$

であり、  $dt(x)/dx = 0$  が成り立つためには、

$$\frac{1}{c} \frac{x-x_p}{\sqrt{(x-x_p)^2+y_p^2}} - \frac{1}{v} \frac{x_q-x}{\sqrt{(x_q-x)^2+y_q^2}} = 0$$

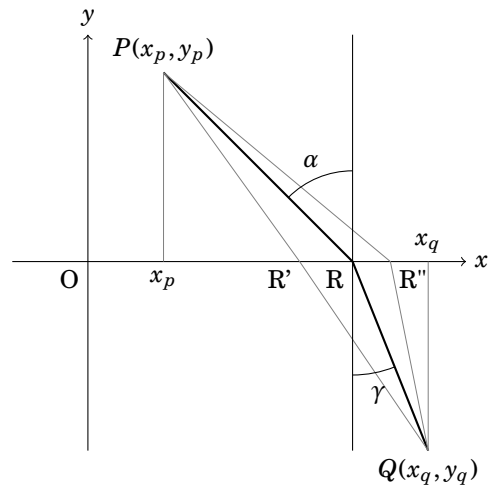


図 6.3: 光の屈折の図

<sup>9</sup>蛇足とは余分なものであるから、この日本語はおそらくおかしい。

<sup>10</sup>今の場合は、2つの空間で光の速度が異なるので、経路の長さが最小すなわち「最短経路」で考えてもうまくはいかない。

でなければならないことが分かる<sup>11</sup>。理解のために、図 6.3 の記号をもちいると、上式は

$$\frac{1}{c} \cdot \frac{x-x_p}{\sqrt{PR}} = \frac{1}{v} \cdot \frac{x_q-x}{\sqrt{RQ}} \iff \frac{1}{c} \cdot \sin \alpha = \frac{1}{v} \cdot \sin \gamma \iff \frac{\sin \alpha}{\sin \gamma} = \frac{c}{v}$$

となる。この関係をみたす点  $R(x,0)$  が最短時間を与えることになる。おのおのの空間での速さ  $c, v$  に依存していることに注意。  $\alpha$  を入射角、  $\gamma$  を屈折角と呼ぶ。またこの関係式には「スネルの法則」という名称がついている。

反射の法則の導出のときと同じように、変数  $x$  を変分の対象の関数として  $t(x)$  の変分を計算してみると（計算の操作は全く同様で）

$$\delta t(x) = \frac{1}{c} \frac{x-x_p}{\sqrt{(x-x_p)^2 + y_p^2}} \cdot \delta x - \frac{1}{v} \frac{x_q-x}{\sqrt{(x_q-x)^2 + y_q^2}} \cdot \delta x$$

となる。変位  $\delta x$  を用いる（物理屋的）流儀である。

### 6.4.3 補足

Fermat 原理に基づいての、反射の経路や屈折の経路の探索方法では、微分でも（無理矢理感はあるけれども）変分でも同様の結果を得た。その議論をフォローするとわかるように、対象の変化量は  $x$  を点とみてるか、関数と見立てるか、だけが違って、計算操作には大きな差異はない。物理の場合は、このような問題、つまり『点をちょっとだけずらす、仮想的にずらす』という操作を考える問題に対しては、どうも  $\delta$  を好む傾向があるような気がする。

<sup>11</sup>6.4.1 節の反射の法則の様に、「仮想変位」を導入して  $\delta x$  での変分を用いても同様の結果を得る。

## 付録A 合成関数の導関数の導出はなぜこんなにややこしいのか

1.7.3 節で示した合成関数の導関数はなぜあのようにややこしいのか、すこし考えてみる。結論をさきに言ってしまうと、精密であろうとする数学的意志の反映がややこしさを生み出している、といえるとおもう。数学のある種の宿命なのかもしれない。

いま、 $y = f(u)$ 、 $u = g(x)$  という関係にある合成関数を考えよう。この導関数の導出については、

$x \rightarrow x + \Delta x$  とした時に  $u \rightarrow u + \Delta u$  となったとする。この  $u$  の変動に対しての  $y$  の変動も  $y \rightarrow y + \Delta y$  とあらわすことにする。ここで  $\Delta y$  と  $\Delta x$  の比にたいして、

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \frac{\Delta u}{\Delta x}$$

という関係が導ける（単に分母分子に  $\Delta u$  を掛けただけ）。さてこの時、導関数の定義から

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = g'(x)$$

となる。さらに最初の  $\Delta u$  の導入の形をみれば、 $\Delta x \rightarrow 0$  のとき、 $\Delta u$  も  $\Delta u \rightarrow 0$  となるから、

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta u} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} = f'(u)$$

したがって、

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = f'(u)g'(x) = f'(g(x))g'(x).$$

と説明されることがある。もうすこし簡単化して、

$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \frac{\Delta u}{\Delta x}$  であるから、

$$\Delta x \rightarrow 0 \text{ のとき } \frac{\Delta u}{\Delta x} \rightarrow \frac{du}{dx}$$

であり、 $\Delta x \rightarrow 0$  のとき、 $\Delta u \rightarrow 0$  となるから、

$$\Delta u \rightarrow 0 \text{ のとき } \frac{\Delta y}{\Delta u} \rightarrow \frac{dy}{du}$$

となる。したがって、 $\Delta x \rightarrow 0$  のとき、

$$\frac{\Delta y}{\Delta x} \rightarrow \frac{dy}{du} \frac{du}{dx}.$$

というもある。

合成関数の導関数として、結論はあっている。細かいことを気にしない「入門的」な説明、というか、導関数作成の操作習得のみが目的ならば、これでもかまわないかもしれないが、数学魂がそれを拒む<sup>1</sup>。

問題は、この合成関数においては、独立変数は  $x$  であり、 $u$  はその変化に応じて変化する従属変数であることにある。したがって、 $x$  の操作は自由に行って良いが、その結果として変化する  $u$  には慎重になるべきなのである。かいつまんで言えば、 $\Delta x$  の取り方によっては  $\Delta u = 0$  の場合があるので、それを無邪気に分母に持ってきたり、お気楽に極限操作に使っては駄目なのだ。それを避けるための工夫が、1.7.3 節のややこしさだったのである。

回避の工夫は、 $\Delta u = 0$  になる場合を考慮して、無条件に分母にしないことにある。その工夫を、導関数の定義を利用して実現してみよう。

いま  $\Delta u := g(x + \Delta x) - g(x)$  とすれば、 $u = g(x)$  の導関数の定義から、 $\Delta x \neq 0$  として

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = g'(x)$$

であることは問題はない。ここで極限を取り外せば

$$\frac{\Delta u}{\Delta x} = g'(x) + \epsilon(\Delta x) \tag{A.0.1}$$

という形に書きあらわすことができる ( $\epsilon$  は  $\Delta x$  の取り方に依存するかもしれないので  $\Delta x$  の関数であるとしておこう<sup>2</sup>)。この  $\epsilon(\Delta x)$  に課せられる条件は  $\lim_{\Delta x \rightarrow 0} \epsilon(\Delta x) = 0$  でありさえすれば良い。そのような  $\epsilon(\Delta x)$  であれば、(A.0.1) から

$$\Delta u = g'(x)\Delta x + \epsilon(\Delta x)\Delta x$$

であるから、必ず

$$\lim_{\Delta x \rightarrow 0} \Delta u = \lim_{\Delta x \rightarrow 0} \{g'(x)\Delta x + \epsilon(\Delta x)\Delta x\} = \lim_{\Delta x \rightarrow 0} \{g'(x)\Delta x\} + \lim_{\Delta x \rightarrow 0} \{\epsilon(\Delta x)\Delta x\} = 0 \tag{A.0.2}$$

となる<sup>3</sup>。

$y = f(u)$  についても同様のことを考えると、 $\Delta u \neq 0$  であれば

$$\begin{aligned} \Delta y &= f(u + \Delta u) - f(u) = f'(u)\Delta u + \eta(\Delta u)\Delta u, \\ \frac{\Delta y}{\Delta u} &= f'(u) + \eta(\Delta u), \\ \lim_{\Delta u \rightarrow 0} \eta(\Delta u) &= 0 \end{aligned}$$

が導き出される。

さて、ここで再び、 $\Delta x \rightarrow 0$  の時の  $\Delta u$  を考えよう。 $(\Delta x \rightarrow 0) \implies (\Delta u \rightarrow 0)$  であることは (A.0.2) で導き出されている。しかしながら、 $\Delta x \neq 0$  であっても  $\Delta u = 0$  となる可能性はある。それゆえ、 $\Delta u = 0$  の場合の考慮をしておくことが必要になる。そのために、 $\lim_{\Delta u \rightarrow 0} \eta(\Delta u) = 0$  を自然に拡張して、 $\eta(0) = 0$  という条件を課すことにする。逆に言えば、そういう関数である  $\eta(\Delta u)$  を選ぶのである。その結果、 $\Delta u = 0$  である場合でも  $\eta(\Delta u) = 0$  が成り立つことになるので、すべてをひっくるめると、

$$\lim_{\Delta x \rightarrow 0} \eta(\Delta u) = 0$$

<sup>1</sup>筆者に常にその魂があるかどうかは別問題である。操作性の習得のみで満足している分野のいかに多いことか。

<sup>2</sup>ここはなかなかストレートに納得しがたい。高木貞治の解析概論 [2] 37 ページでは、 $x$  と  $\Delta x$  の関数としている。この嗅覚はどこからくるのか？

<sup>3</sup> $\epsilon(\Delta x)$  の形式によっては、一般には  $\lim_{\Delta x \rightarrow 0} \{\epsilon(\Delta x)\Delta x\} = 0$  となるとは限らない。それゆえその前に  $\lim_{\Delta x \rightarrow 0} \epsilon(\Delta x) = 0$  を導いておいたのである。これによって  $\Delta x \rightarrow 0$  の極限が 0 になることが保証される。

と言えることになる。以上の結果をもちいると、

$$\begin{aligned}\Delta y &= f'(u)\Delta u + \eta(\Delta u)\Delta u = (f'(u) + \eta(\Delta u))\Delta u \\ &= (f'(u) + \eta(\Delta u))(g'(x) + \epsilon(\Delta x))\Delta x \\ &= f'(u)g'(x)\Delta x + \{f'(u)\epsilon(\Delta x) + g'(x)\eta(\Delta u) + \eta(\Delta u)\epsilon(\Delta x)\}\Delta x\end{aligned}$$

であり、 $\Delta x$  に関しては、 $\Delta x \neq 0$  として構わなかったので

$$\frac{\Delta y}{\Delta x} = f'(u)g'(x) + \{f'(u)\epsilon(\Delta x) + g'(x)\eta(\Delta u) + \eta(\Delta u)\epsilon(\Delta x)\}$$

となり、最後の項は

$$\lim_{\Delta x \rightarrow 0} \{f'(u)\epsilon(\Delta x) + g'(x)\eta(\Delta u) + \eta(\Delta u)\epsilon(\Delta x)\} = \lim_{\Delta x \rightarrow 0} \{f'(u)\epsilon(\Delta x)\} + \lim_{\Delta x \rightarrow 0} \{g'(x)\eta(\Delta u)\} + \lim_{\Delta x \rightarrow 0} \{\eta(\Delta u)\epsilon(\Delta x)\} = 0$$

であるから

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(u)g'(x)$$

となる。

まとめてみよう。 $\epsilon(\Delta x)$  は  $\lim_{\Delta x \rightarrow 0} \epsilon(x)$  を満たす関数であればなんでもよかった。一方、 $\eta(\Delta u)$  は、まず少なくとも  $\lim_{\Delta u \rightarrow 0} \eta(\Delta u)$  を満たさねばならない。そのような  $\eta(\Delta u)$  のうちで、 $\eta(0) = 0$  が成り立つものを採用しているのである。それらが全て出揃ったところで、合成関数の微分が導出されてくる。これが「ややこしさ」の中身だったのである。

1.7.3 節では、 $\Delta x$  の代わりに  $h$ 、 $\Delta u$  の代わりに  $t$  を用いた。その背景にはここに述べたような事態があったのである。

## エピローグ

ここまで書いてきたこのノートについて思うことは、

- 高木の解析概論 [2] のほんの数章，本の一部のみにしかならない。
- 対象としている関数は，概ね連続関数であるという仮定に頼っている．いくなれば桃源郷の世界での記述である。
- 偏微分についての操作技術は書いていない。
- 積分と微分の順序交換，いわゆる積分下の微分を省いてしまった。
- 数学的事実としてアприオリに導入した事柄は，有理式の微積分からの導出，つまり，微積分の世界のなかで導出することが可能であるけれど，その導出の方法を省いた。

などなどである（当然のことながら，複素関数については何も触れなかった）．これらについては，いつか機会があれば勉強してまとめて見たいと考えている．

小さな自分の部屋で，Mac での  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X} 2_{\epsilon}$  でこのノートを執筆している時に，友人の尾島由郎君の *Une Collection des Chainons I, II* を繰り返し繰り返し聴いた．解析の世界，なかんづく，微小量の不思議な世界を覗くとき，そして理解が進まず悪戦苦闘しているとき，この音楽とこの音の塊に元気づけられ，励まされた．微小量の玄妙さとの相乗効果は，筆者をえもいわれぬ気分にもしてくれた．ありがたかった．尾島くんどうもね．

## 関連図書

- [1] 高瀬 正仁. 『 $dx$  と  $dy$  の解析学—オイラーに学ぶ』. 日本評論社, 増補版, 2015. (増補版第一刷を参照した).
- [2] 高木 貞治. 『解析概論』. 岩波書店, 改訂第三版, 1961. (第 19 刷を参照した).
- [3] 山本茂年・小川徹夫・阿部雄二・荒木重蔵 共編. 『数学コメンタール〈数学 I・II・III・A・B・C〉』. 駿台文庫株式会社, 初版, 1996. (初版第 12 刷 (2004) を参照した).
- [4] 小島 寛之. 『ゼロから学ぶ微分積分』. 講談社, 2001. (第 16 刷 (2014) を参照した).
- [5] 前野 昌弘. 『よくわかる解析力学』. 東京図書株式会社, 2013. (第 2 刷 (2014) を参照した).