第九章　汎関数と変分法

この章では物理学の学習には欠かすことができない「**汎関数**」の意味と、それを用いた「**変分法**」について学ぶ。

§1. 汎関数と変分法

【汎関数とは】

まず「**汎関数**」の意味を与えることから始める。変数xの値を与え、それに対してfで指定される一連の操作を施し、値$f(x)$を返す操作が「**関数**」$f(x)$の意味であった。同じ様に、xの関数$f(x)$を与え、それに対してJで指定される一連の捜査を施した後に出来る関数を返す操作を「**汎関数**」といい、$J[f]$と書く。簡単に言えば「汎関数は関数を変数とする関数である」。関数と汎関数の違いを並べて示すと下表のようになる。

	関数$f(x)$の場合	汎関数$J[f]$の場合
変数	x	$f(x)$
xに施す操作の内容	f	J
操作の結果	$f(x)$	$J[f]$

初学者にはなかなかわかりにくいので、具体的かつ簡単な汎関数の例を以下に与える。

【汎関数の例】

変数xの関数を$f(x)$とするとき、この関数を$x=0$から$x=1$まで積分した値を与える操作$\int_0^1 f(x)dx$を考える。すなわち、ここでの$f(x)$に対する「操作の内容」は「$x=0$から$x=1$までの積分」である。といってもこの結果がどのような値になるか分からない。$f(x)$が何か分からないので積分が実行できないからである。もし$f(x)$が与えられれば、むずかしいかやさしいかは別にして積分を実行することができ、その結果を得ることができる。すなわち$\int_0^1 f(x)dx$の値は関数fを与えることによって初めて定まる。これが「$\int_0^1 f(x)dx$は関数fを変数とする汎関数」の意味である。そしてそれを

$$\langle 9\text{-}1\rangle \quad J[f]=\int_0^1 f(x)dx$$

と書くのである。$[\ldots]$の意味は[…]内の関数の形を変えると結果が変わることを表している。たとえば、もし$f(x)=x$であれば<9-2>$\left(J=\int_0^1 xdx=\left[\frac{1}{2}x^2\right]_0^1=\frac{1}{2}\right)$であり、もし$f(x)=x^2$であれば<9-3>$\left(J=\int_0^1 x^2dx=\left[\frac{1}{3}x^3\right]_0^1=\frac{1}{3}\right)$であり、もし$f(x)=x^3$であれば<9-4>$\left(J=\int_0^1 x^3dx=\left[\frac{1}{4}x^4\right]_0^1=\frac{1}{4}\right)$である。

これをもう一段階複雑にした汎関数が物理では重要になる。つまり$f(x)$を与えて直ちにJが決まるのではなく、関数$f(x)$を与えるとまず$f(x)$の汎関数$F[f(x)]$が決まり、さらにxの関数であるFを積分して得られるJを問題にすることがある。したがってこの場合は、$f(x)$が与えられてもFがわからなければ$J[f]$が決まらない。たとえば、$F[f(x)]$を$x=0$から$x=1$まで積分した

$$\langle 9\text{-}5\rangle \quad J[f]=\int_0^1 F[f(x)]dx$$

はその例であるが、これはFの形を与えたとしてもfを与えない限り決まらないので、結局はfの汎関数である。例えば上の場合で$F[f(x)]$が関数$f(x)$に1を加える操作

<9-6> $F[f(x)] = f(x) + 1$

であるとする。この時はもし$f(x) = x$であれば$F = x + 1$であるから、<9-7>
$\left(J = \int_0^1 (x+1)dx = \left[\frac{1}{2}x^2 + x\right]_0^1 = \frac{3}{2}\right)$であり、もし $f(x) = x^2$であれば<9-8>
$\left(J = \int_0^1 (x^2+1)dx = \left[\frac{1}{3}x^3 + x\right]_0^1 = \frac{4}{3}\right)$であり、もし$f(x) = x^3$であれば<9-9>
$\left(J = \int_0^1 (x^3+1)dx = \left[\frac{1}{4}x^4 + x\right]_0^1 = \frac{5}{4}\right)$である。

物理では「関数$f(x)$を与えた時に決まる $J[f]$を最も小さくする関数 fを知りたい」というような場合がある。それに使われるのが「**変分法**」という方法である。

【変分法】

「汎関数は関数を変数とする関数」であることと「変分法の基本的な考え方は関数の微分計算と同じ」を合わせて、「**変分は関数による汎関数の微分である**」と簡単に言っても良い。「汎関数は関数を変数と考えた関数である」ことを理解し、「変分法の基本的な考え方は関数の微分計算と同じである」ことを理解するために、上に与えた汎関数の例を使って具体的に説明しよう。例ににあるように関数 $f(x)$の汎関数 $F[f(x)]$を積分の形で含む汎関数、すなわち

<9-10> (9.1.1) $J[f] = \int_A^B F[f(x)]dx$

から始める。汎関数の微分を理解するには、汎関数の意味をしっかりと理解していないといけない。そのために(9.1.1)式を使ってもう一度、汎関数の意味を確かめておく:

- fはxの関数であるから、それとFの具体的な形が与えられれば、xの関数として$F[f(x)]$が決まり、積分を実行することができる。それが(9.1.1)式右辺の意味である。積分を実行した結果のJはfを変えると変わる。それが左辺の$J[f]$の意味である。

物理学で重要になるのは、fを少し変えたときその変化が Fを通して左辺のJに与える影響である。すなわち、$f(x)$に小さな関数$\Delta(x)$をつけ加え、$f(x)$を $(f(x) + \Delta(x))$とする。必要なのはfに Δを加えたことによってJがどれだけ変わったか、すなわち

<9-11> (9.1.2)
$$J[f+\Delta] - J[f] = \int_A^B F[f(x) + \Delta(x)]dx - \int_A^B F[f(x)]dx$$

を知ることである。物理学を学ぶときにはそれを知ることがなぜ重要であるかを理解することは必要であるが、今は数学の問題としてこの式を考えよう。(9.1.2)式の両辺に対して以下の二つの考察を行なう。

1. いま$f(x)$と $\Delta(x)$が積分を行う区間 ($A \le x \le B$)で突然大きな変化をすることがないとする。そうすると、$\Delta(x)$を限りなく小さくすれば(9.1.2)式右辺の引き算の結果もやはり小さくなり、左辺の量も小さくなる。その意味を表すために左辺を $\delta J[f]$と書く。δはfにとても小さなΔを加えたことによって生じる Jの変化が「とても小さい」ということを表す記号である。
2. 少し面倒な計算をすると、同じ区間$(A \le x \le B)$で「xのある関数」を積分した量として(9.1.2)式の右辺を表すことができる。汎関数$F[f(x)]$についてそれを実行した時の「xのある関数」を $\left(\frac{\partial F}{\partial f}\delta f\right)$と書く。不思議な記号を使うが、「$x$のある関数」が持つ意味を正確に伝えるためには仕方がない。このなかのδfは 「Fに含まれるfが僅かに変わったとき」ということ

を忘れないための記号である。そして$\frac{\partial F}{\partial f}$は$F$のなかにある関数$f$を普通の変数であるかのように考えて、それで$F$を微分して得られる微分係数である。

以上より、結局(9.1.2)式を

〈9-12〉 (9.1.3) $$\delta J[f] = \int_A^B \left(\frac{\partial F}{\partial f}\delta f\right)dx$$

と書き表すことができる。

「物理学」でしばしば、(9.1.1)式のJに最も小さな値を与える関数fを求めることが必要になる。いま関数$f(x)$を与えたときにJが最も小さくなってJ_0となったとする。その値を与える$f(x)$をわずかに変えるとJはJ_0から必ず大きくなるので、(9.1.3)式の$\delta J[f]$は0でない値を与えるはずである。したがってJにJ_0の値を与える関数fは(9.1.3)式の$\delta J[f]$を0とするfであるということになる。もし(9.1.3)式右辺の積分の中にある$\frac{\partial F}{\partial f}$を$0$にするような$f$があれば、それは必ず$\delta J[f]$を$0$にする。したがって、もし$\frac{\partial F}{\partial f}=0$を満足する$f$を探し出すことができれば、それが求める$f$を与えることになる。

物理で生じる実際の状況はもう一つ複雑になる。それはFがfの微係数$\frac{df}{dx}\equiv f'$を含む$F[f(x),f'(x)]$の形をし、その結果Jが$J[f,f']$となる時、それに最も小さな値を与える関数$f(x)$を求めることが必要になる。この時には、(9.1.3)式以下で与えたと同じ議論をていねいに行えば、Fの中にあるfとf'が関係式

〈9-13〉 (9.1.4) $$\frac{d}{dx}\frac{\partial F}{\partial f'}-\frac{\partial F}{\partial f}=0$$

を満たせば、$F[f(x),f'(x)]$が最小となることが分かる。この方程式は一般に「オイラー方程式」と呼ばれ、物理学では「最小作用原理」を表す「オイラー-ラグランジの方程式」として現われる。

$\frac{\partial F}{\partial f}$は$F$のなかにある関数$f$を普通の変数と考えて偏微分を実行して得られる偏微分係数といった。同様に$\frac{\partial F}{\partial f'}$は$F$のなかにある関数$f'$を普通の変数と考えて偏微分を実行して得られる偏微分係数である。したがって、$J[f]$に最も小さな値を与える関数fは、fとf'によるFの偏微分を実行して得られた偏微分係数を(9.1.4)式に代入し、その方程式を満足するfを求める問題に帰着する。どのようにこれが実行されるかを理解するため、簡単な練習問題をひとつあげる。

【練習問題】$(x-y)$平面上の原点Oと平面上の座標(a,b)にあるP点を結ぶ最も短い曲線を求めよ。答えはもちろんOとP点を結ぶ直線、$y=\frac{b}{a}x$である[1]。

【解】まずこの問題を汎関数を使った変分法の問題として表す。そのため、OとPを結ぶ曲線は非常に短い直線の連続で出来ていると考える。一般に$(x-y)$平面上の二点(x_1,y_1)と(x_2,y_2)間の直線距離sを与える公式は〈9-14〉$\left(s=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2}\right)$である。これを使うと、$O$と$P$を結ぶ曲線上の一点$(x,y)$と、それから少し離れた点$(x+\Delta x, y+\Delta y)$間の短い直線距離$\Delta s$は

〈9-15〉
$$\Delta s = \sqrt{(\Delta x)^2+(\Delta y)^2}$$
$$= \sqrt{1+\left(\frac{\Delta y}{\Delta x}\right)^2}\,\Delta x$$

である。したがってOからPまでの距離は、Oから短い直線距離Δsを加えながらPにたどり着いたときのΔsの総和に等しい。この総和はΔs(したがってΔxとΔy)を短くすればするほど考えている曲線の距離を正確に表すことになる。数学的にこ

れを実行してOからPに至る曲線の距離Lを与えると、Lは

〈9-16〉 (9.1.5) $$L=\int_0^a \sqrt{1+y'^2}\,dx$$

となる。ここで、OからPに至る曲線上のx座標と y座標の関係を表す与える式（曲線の方程式）を $y(x)$と書いたとき、$y'=\frac{dy}{dx}$は xにおけるその曲線の傾きを表すyの微分係数である。したがって、もし(9.1.4)式のLを最小にする右辺の中にあるy'がわかれば、それは OとPを最短でつなぐ曲線のx座標とy座標の関係を表す微分方程式を与えることになり、それを解けば曲線がわかる。

(9.1.5)式右辺で積分される関数はy'のみを含むので、それを(9.1.1)式の Fと考え、$F[y'(x)]$と書くことができる。よってFのyに関する微分は 0である。したがってFをy'で微分したものを(9.1.4)式のオイラー方程式に代入すれば、$y(x)$が満たす微分方程式がただちに得られる。Fのy'に関する微分を実行すると

〈9-17〉 (9.1.6) $$\frac{dF[y'(x)]}{dy'}=\frac{d\sqrt{1+y'^2}}{dy'}=\frac{y'}{\sqrt{1+y'^2}}$$

となる。これと$\frac{dF}{dy}=0$を(9.1.4)式に代入すれば、Oと Pを最短でつなぐ曲線を表す関数が満足するオイラー方程式は

〈9-18〉 (9.1.7) $$\frac{d}{dx}\frac{y'}{\sqrt{1+y'^2}}=0$$

となる。したがって$\frac{y'}{\sqrt{1+y'^2}}$は定数であるから、それをCと置きy'を求めれば $y'=\frac{C}{\sqrt{1-C^2}}$となる。右辺は定数であるから、それを改めて Aとすれば

〈9-19〉 (9.1.8) $$y'=A$$

である。xに関する一階の微分係数が定数 Aとなるyを一般に

〈9-20〉 (9.1.9) $$y=Ax+B$$

と書くことができる。ここで、Bはもう一つの定数である。これは直線を表す方程式であるが、この直線は原点$O(0,0)$と$P(a,b)$点を通るから

〈9-21〉 (9.1.10) $$\begin{cases}0=B\\ b=Aa\end{cases}$$

であるから、よって二つの定数は

〈9-22〉 (9.1.11) $$\begin{cases}B=0\\ A=\frac{b}{a}\end{cases}$$

でなければならない。したがって原点とP点をつなぐ最短曲線は

<9-23> (9.1.12) $y = \frac{b}{a} x$

で与えられる直線になることが変分の考え方を使って証明できた。

上で与えたのは自明でつまらない例のように思うかも知れないが、「相対性理論」で光が通る道筋が四次元空間内の直線になることは基本的に同じやり方で示すことができる。さらには、自然界の全てを支配する「最小作用の原理」から、ニュートン方程式を初めとする自然界の諸方程式を導くのも基本的にはオイラー方程式を使って同じやり方で行う。そのような、物理学独特の考え方を理解するためには、ここで与えた簡単な例をしっかりと理解しておくことが大切である。

[1] これが物理で現れたときは、空間のある点からある点まで光がたどる経路を求める問題になる。その時には、もし空間が真空であれば答えはこの練習問題の結論のように直線になるが、もし空間に質量を持つ物質があれば、光の経路は物質の存在によって曲げられて、最短経路は曲線になる。そのような例は宇宙物理に現れる。