連載（講義）

電子光学入門
－　電子分光装置の理解のために　－
(第３回)

嘉藤　誠
日本電子 (株) 〒 196-8558 東京都昭島市武蔵野 3-1-2
kato@jeol.co.jp
(2004 年 10 月 7 日受理)

幾何光学の理論は，光線を決定するフェルマーの原理をもとにした定式化が可能であり，光線に付随する波面，いわゆるアイコナールの概念をもちいて光線束の振舞いが研究されます．この理論形式を粒子の運動に適用することで，力学における変分原理が得られ，電子ビームがみたすべき種々の制約がこの定式化を通じて導かれます．今回は，光学理論からはじめて，力学における変分原理への一般化の過程を示します．

Introduction to Electron Optics for the Study of Energy Analyzing Systems (3)

M. Kato
JEOL Ltd., 3-1-2 Musashino, Akishima, Tokyo 196-8558.
kato@jeol.co.jp
(Received: October 7, 2004)

A study of the laws of geometrical optics is greatly facilitated by a variational approach. The behaviour of light rays may be derived from Fermat's principle, and this yields the concept of eikonal, i.e., wavefronts accompanying light rays. Variational principles in classical mechanics are the generalization of the optical counterparts, and several restrictions imposed on electron beams are obtained through this formulation. In this chapter, we begin with the derivation of the laws of geometirical optics, and genelarize from them the vatiational principles of mechanics.

3　幾何光学と力学の原理

3.1　はじめに

前回までの議論（第１章「電子光学概論」，第２章「電磁場中の電子の運動」）において，最小作用の原理とよばれるものの重要性を述べました．この原理は，電子光学の定式化を見通しよくするだけではなく，電子の集団としての振舞いを研究するための有力な道具です．本章では，最小作用の原理の導出を行うとともに，変分原理から導かれる力学法則について述べていきます．

電子光学でもちいられる最小作用の原理は，解析力学のなかでも難解なもののひとつです．この原理の導出過程を示すためには，まず解析力学の基本的な理解がなければなりません．しかし，解析力学の教科書はすでに多く存在します．それらと同じ内容をたどるかわりに，ここでは電子光学の講座らしく，光学の原理を出発点としたアプローチを行うことにします．もともと解析力学は，幾何光学の理論を単純に一般化したものであり，歴史を考えてもこのような扱いは正統的なものです．

出発点となる光学の原理は，すでに第１章で紹介したフェルマーの原理です．これは，空間の２点を結ぶ曲線のうち光路長が最短のものが，現実に光線の

たどる経路となるということを言っています．光路長とは，曲線に沿っての実際の長さに，その場所の屈折率をかけながら積分したものです．もし空間に物質が存在せず，屈折率が一様に1であるなら，光路長とは通常の意味での長さのことですから，２点を結ぶ最短の経路としての直線が光線経路として選ばれることになります．

少し一般化して考えましょう．数学的に用意されたn次元空間があって，その空間に「長さ」の概念が定義されているとします．すると，空間内の2点を結ぶいろいろな曲線にたいしてその長さを求めて比較すれば，2点をむすぶ最短経路というものが決定できるでしょう．そして，その最短経路の長さを2点間の「距離」と呼ぶのは自然です．このときの最短経路自身は，数学では一般に**測地線** (geodesic) とよばれます．

測地線の概念は，もともと地球上の2点を結ぶ最短航路にたいして名づけられたもので，2点間の距離を測るための経路という意味でもちいられたものです．この場合には，地球の表面，つまり球面という「曲がった」2次元空間が長さの定義された空間であり，球面上の2点を通るピンと張ったロープのかたちが測地線となるわけです．

この測地線の概念をもちいるなら，フェルマーの原理によって決定される光線の経路とは，光路長に関しての測地線であるわけです．つまり，通常のユークリッド的な長さとは別の，光路長として定義された「長さ」に関しての測地線を，現実の光線がたどるということです．

このように，2点を結ぶ曲線のうちのどれが測地線になるかは，その空間に定義された「長さ」がどのようなものであるかに依存します．逆に，2点を結ぶ曲線のうちのある特定のものが選ばれるように,「長さ」をうまく定義する問題も考えることができるでしょう．

ニュートン以後の古典力学がたどった道のりのひとつは，まさにこのようなものです．すなわち，ある質点の運動，あるいはある力学系のたどる時間発展を，その系の自由度にたいして定義される空間（時間はその空間の次元のひとつとなります）における測地線としてとらえるという定式化です．そしてその発想の起源は，光学にあるわけです．

力学を，光学理論に沿って定式化することの意義は何でしょうか?その一つは，測地線という概念が幾何学的な意味をもち，よって特定の座標系を想定する必要がないということです．そのような形式で表現された運動法則は，具体的に書き下す際には任意の座標系をもちいることができます．すなわち，任意の曲線座標系，あるいは非慣性系（たとえば前章で述べたラーモア座標系）でさえも許されます．

別のメリットとして，変分原理を軸とした定式化によって，力学系に課される本質的な制限というものが見出されます．そのような制限は，いわゆる正準変換論として展開されるものですが，これは電子光学系における電子ビームの振舞いを研究するうえで，本質的な意味をもちます．

本章では，まずフェルマーの原理の本質を理解し，幾何光学の原理を示した上で，力学における変分原理へと向かうことにします．なお，次回以降では，ここで導いた結果のいくつかをもちいるだけですので，今回は気楽な読み物として目を通していただければ結構です．

3.2　幾何光学の原理

3.2.1　直線とはなにか

何もない空間，つまり真空の領域において，ある一点からレーザービームを発射すれば，その光線の経路はまっすぐな直線になります．しかし，この「まっすぐ」とはなにを意味しているのでしょうか?これをきちんと説明するのはさほど簡単ではありません．

もし空間に適当な座標系を設置すれば，その座標をもちいて直線の方程式を示すことができます．しかし，ここで考えているのは直線という幾何学的な概念であり，われわれがどんな座標を空間に張りめぐらすかとは無関係に，意味をもつものであるはずです．

直線を座標をもちいないで定義する方法のひとつは,「直線とは2点をむすぶ最短経路のことである」とすることです．この定義が可能であるためには「長さ」という概念が必要ですが，これはわれわれの空間に最初からそなわっています．そこで，空間の2点を任意に指定したとき，それを結ぶすべての曲線のうちで最小の長さをもつものを直線とよべばよいわけです．

これで，直線の意味，よってまた「真空中の光線の経路を決定する原理」が幾何学的な言葉で表現されたことになります．しかし，もちろんわれわれは座標をもちいた議論に慣れていて，定量的になにかを決定するには座標の助けが必要です．そこで，上で述べたような幾何学的な直線の定義を，座標によって表現するとどうなるかを考えてみます．

まず，空間のなかにデカルト座標 xyz をとります．空間内に２点 A，B を考え，この２点を通る任意の曲線を指定します．この曲線の長さ l を，とりあえず形式的に次のように書いておきます．

$$l = \int_A^B ds \tag{1}$$

ここで ds は線素，すなわち曲線上の微小区間の長さを表していて，その区間の両端の座標を (x, y, z), $(x + dx, y + dy, z + dz)$ とすれば，その長さの２乗は，

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2}$$

で与えられます．微小な長さ ds を曲線に沿って加えていったものが曲線の長さとなります．

(1) を実際に計算するには，曲線の形状を指定する必要があります．光学では z 軸を光軸として選ぶことが多いので，z を独立変数としてもちいて，

$$x = x(z),\ y = y(z) \quad z_A \leq z \leq z_B \tag{3}$$

と表します．

あとで力学を論じる際は，時間 t を独立変数として，粒子の運動を $(x(t), y(t), z(t))$ のように表します．よって，光学における z は，力学での時間の役目を果たす変数です．われわれは無意識のうちに xyz 空間の３つの次元をひとまとまりにして考えてしまいがちです．しかし光学の定式化においては，xy 平面上の光線の軌跡を z をパラメタとして記述する，という認識をもたなければなりません．

もし光学において xyz 空間における概念が登場したら，それは力学においては，$xyzt$ という４次元の時空間における概念を考えていることに相当します．そのような,「物理空間と独立変数を合わせた空間」を考えることはむしろ特殊な立場であり，そしてまた，非常に重要な見方でもあります．(2) で定義される長さが，すでにそのような量になっています．

さて，曲線が (3) のように与えられると，z の微小区間 $[z, z + dz]$ における x 座標の変化は $dx = x'(z)dz$, y 座標の変化は $dy = y'(z)dz$ となります．よって，この区間における曲線の長さの２乗は，

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= (x'dz)^2 + (y'dz)^2 + dz^2 \\ &= (x'^2 + y'^2 + 1)\ dz^2 \end{aligned} \tag{4}$$

となります．これから，(3) のように与えられた曲線の長さを具体的に計算する式として，

$$l = \int_{z_A}^{z_B} \sqrt{x'(z)^2 + y'(z)^2 + 1}\ dz \tag{5}$$

が得られます．

さて，２点 A，B を結ぶ曲線のうちで，長さの最小のものが直線であると定義したわけです．この定義は，次のような変分原理の形式で与えることができます．

$$\delta l = \delta \int_A^B ds = 0 \tag{6}$$

上式における δ は，曲線の形状を２点 A，B を通るという条件のもとでわずかに変化させる操作を表していて,「曲線に変分を与える」という言い方をします．上式は，そのような変分にたいして長さに変化がないということを要求しています．２点を結ぶすべての曲線から直線を選び出すために，長さが停留（極小）になれば十分であり，最小であることを要求する必要はありません．一般に，このような変分原理によって決定される曲線を測地線とよぶことはすでに述べましたが，これを**停留曲線** (stationary curve) ということもあります．

上の原理 (6) は，最初に空間内の２点が指定されて，その２点を通る曲線のなかの特別な一つを選び出すはたらきをしています．しかしこれを，真空中での光線の経路を決定する原理として見るとき，先に始点と終点を指定するということが不自然に思えてしまいます．(力学における変分原理を最初に習ったとき，なぜ両端だけが特別扱いなのか？と悩んだ人は少なくないはずです.)

しかし，今考えている原理は長さの概念にもとづいていて，２点を結ぶ最短線を選び出そうとしているわけです．比較の際に両端を動かしていいのであれば，長さが最短になるのは２点がくっつく場合であるという結果が導かれるだけです．つまり，曲線のうちの特別なひとつを選ぶ原理にはなりえません．２点を固定したうえで比較を行うということが，長さの概念をもちいて直線を定義するために必然的なことなのです．

さて，(5) は確かに曲線の長さを与える式ですが，z の関数 $\sqrt{x'(z)^2 + y'(z)^2 + 1}$ を z に関して積分する，つまり,「z 軸上に分布した値を足し合わせていく」という形式になっています．しかし，曲線の長さというものは本来,「曲線に沿って分布する量を曲線に沿って足しあげていく」というかたちで理解すべきものでしょう．

そのような「曲線に沿っての積分」のうち，これからの議論において有用なのはベクトルの線積分としてのかたちです．つまり，空間全体，あるいは少なく

とも曲線上でベクトル $\boldsymbol{P}$ が定義されていれば，

$$\int_A^B \boldsymbol{P}\cdot d\boldsymbol{X} \tag{7}$$

という積分を考えることができます．ここで $d\boldsymbol{X} = (dx, dy, dz)$ は，曲線に沿った微小な接ベクトルです．曲線の長さとはこのベクトル $d\boldsymbol{X}$ の長さを加えていったものですから，曲線に沿った単位接ベクトルを $\boldsymbol{U}$ とすれば，これを $\boldsymbol{P}$ として採用すればよいことがわかります．すなわち，

$$l = \int_A^B \boldsymbol{U}\cdot d\boldsymbol{X} \tag{8}$$

となります．この様子を Fig.1 に示します．

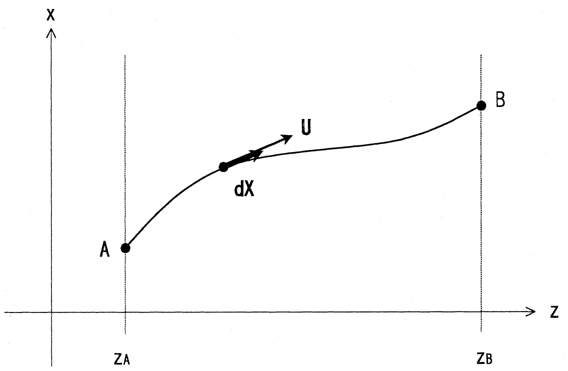

Fig. 1: Unit tangential vector $\boldsymbol{U}$ of a curve joining two points A and B. Integration of $\boldsymbol{U}$ along a curve yields length l of the curve.

単位接ベクトル $\boldsymbol{U}$ の成分は，$d\boldsymbol{X} = (x', y', 1)\ dz$, $|d\boldsymbol{X}| = \sqrt{x'^2 + y'^2 + 1}\ dz$ をもちいれば次のように与えられます．

$$\begin{aligned}\boldsymbol{U} &= \frac{d\boldsymbol{X}}{|d\boldsymbol{X}|}\\ &= \left(\frac{x'}{\sqrt{x'^2+y'^2+1}}, \frac{y'}{\sqrt{x'^2+y'^2+1}},\right.\\ &\qquad\left.\frac{1}{\sqrt{x'^2+y'^2+1}}\right)\end{aligned} \tag{9}$$

よって，$\boldsymbol{U} = (u_x, u_y, u_z)$ をもちいて曲線の長さを与える公式として，次式が得られます．

$$\begin{aligned}l &= \int_A^B u_x dx + u_y dy + u_z dz\\ &= \int_A^B \frac{x'}{\sqrt{x'^2+y'^2+1}}\, dx + \frac{x'}{\sqrt{x'^2+y'^2+1}}\, dy\\ &\qquad + \frac{1}{\sqrt{x'^2+y'^2+1}}\, dz\end{aligned} \tag{10}$$

これを具体的に計算する際は $x' = x'(z)$, $dx = x'(z)dz$ などを代入することになるので，結局は (5) を計算することになります．しかし，これから展開する光学と力学の定式化のために，曲線の長さを (10) のかたちに表しておくことが重要になるのです．

3.2.2　フェルマーの原理

屈折率の分布 $n = n(x, y, z)$ をもつ媒質中での光線経路を決定するのが，次のフェルマーの原理です．

$$\delta l = \delta \int_A^B n ds = 0 \tag{11}$$

積分 l は光路長とよばれ，この光路長という長さに関しての測地線が現実の光線経路であることを意味しています．(上式は前節と異なる「長さ」の定義ですが，同じ記号 l をもちいます.)

まず，光路長の定義は前節の長さの定義に屈折率 n をかけたものなので，線素の表式は (2) に対応して，

$$dl^2 = n^2(x, y, z)\ (dx^2 + dy^2 + dz^2) \tag{12}$$

となります．よって光路長を具体的に計算する式として，前節の (5) に対応して，

$$l = \int_{z_A}^{z_B} n(x, y, z)\sqrt{x'^2 + y'^2 + 1}\ dz \tag{13}$$

が得られます．

上の (12)，あるいはその特別な場合の (2) のように，$dl^2 =$「座標の微分 dx, dy, dz の 2 次形式」のかたちで空間に長さが定義されるとき，「(リーマン) 計量が導入される」という言い方をします．しかし，力学では「非リーマン的な」長さもでてくるので，本稿では計量という用語はもちいないことにします．また，測地線というよびかたはリーマン計量にたいしてもちいることが多いのですが，本稿では積分を停留にする曲線という意味で，広義にもちいます．

さて，前節と同様に，光路長をベクトルの線積分のかたちで与えることができます．いまの場合には，曲線に沿って積分すべきベクトルは n の因子だけの違いで，

$$\boldsymbol{P} = n\boldsymbol{U} \tag{14}$$

となります．

このベクトル $\boldsymbol{P}$ は，これを曲線に沿って積分することで光路長を与えるので，以下では**測地ベクトル** (geodesic vector) とよぶことにします．このベクトル

は，曲線に沿って光路長が増加する方向を向き，大きさがその増加率を与えるようなものです．

(14) で定義される測地ベクトル $\boldsymbol{P}$ は，光学では**光線ベクトル** (ray vector) とよぶのが普通ですが，力学を考える際も同様のはたらきをする量を導入できます．$\boldsymbol{P}$ は xyz 空間における概念であり，(13) の積分のように独立変数 z を特別扱いするものではないことに注意しましょう．

重要なことは，$\boldsymbol{P}$ は１点 (x, y, z) を指定して定まるベクトルではなく，その点を通る曲線の傾き (x', y') も同時に指定してはじめて決まるということです．すなわち，一本の曲線を指定したときに，その曲線に沿ってだけ $\boldsymbol{P}$ の分布が定まります．この測地ベクトルのような「曲線に沿って定まる量」というものが，光学において重要な量となります．(力学においてこれに対応するのは，運動に沿って定まる量，たとえば運動量とかエネルギーです！)

測地ベクトル $\boldsymbol{P} = (p_x, p_y, p_z)$ をもちいて光路長を表す式は，(10) に対応して，

$$\begin{aligned} l &= \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} \\ &= \int_A^B p_x dx + p_y dy + p_z dz \\ &= \int_A^B \frac{nx'}{\sqrt{x'^2 + y'^2 + 1}}\, dx + \frac{ny'}{\sqrt{x'^2 + y'^2 + 1}}\, dy \\ &\qquad + \frac{n}{\sqrt{x'^2 + y'^2 + 1}}\, dz \end{aligned} \tag{15}$$

となります．測地ベクトルをもちいれば，フェルマーの原理 (11) は次のように表すことができます．

$$\delta \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} = 0 \tag{16}$$

さて，$\boldsymbol{P}$ は任意の１点で任意の傾きをもつ曲線にたいして定義されますが，$\boldsymbol{P}$ の各成分は全く勝手な値をとることはなく，必ずみたすべき一つの関係式が存在します．それは，$\boldsymbol{P} = n\boldsymbol{U}$ の長さが n でなければならないこと，すなわち，

$$|\boldsymbol{P}| = n \tag{17}$$

であることからの帰結で，$\boldsymbol{P}$ の成分間には，

$$p_x^2 + p_y^2 + p_z^2 = n^2 \tag{18}$$

という関係が存在します．

さて，光路長の定義と測地ベクトルの性質を述べてきましたが，フェルマーの原理によって決定される光線がどのような性質をもつものかはまだ議論していません．一般に，フェルマーの原理のような変分原理には，それと等価な微分方程式が必ず存在します．それを，変分問題にたいしての**オイラーの方程式** (Euler's equation) とよびます．

変分法に関しての詳しいことは§3.4 を見ていただくことにして，ここでは結果だけ述べます．フェルマーの原理に対応するオイラーの方程式は，次のように与えられます．

$$\frac{d\boldsymbol{P}}{ds} = \nabla n \tag{19}$$

上式において，左辺は光線に沿っての測地ベクトルの変化率です．この方程式の意味するところは，∇n の方向，すなわち屈折率の大きくなる方向に光線が曲げられていくということです．

3.2.3　波としての光

現在われわれは，光は電磁波であり，日常的な意味で粒子とよぶようなものではないことを知っています．フェルマーの原理が決定するものは光線の経路であって，光の波としての伝播法則を与えるものではありません．しかし，光の本性が波であるとするならば，フェルマーの原理は波としての振舞いの何らかの反映であるはずです．

まず単純な例から考えましょう．Fig.2 において，屈折率が図の上下方向のみに変化していて，水平方向には一定であるとします．光線が点 A に左手から水平に入射したときの，その後の光線の経路を考えてみます．

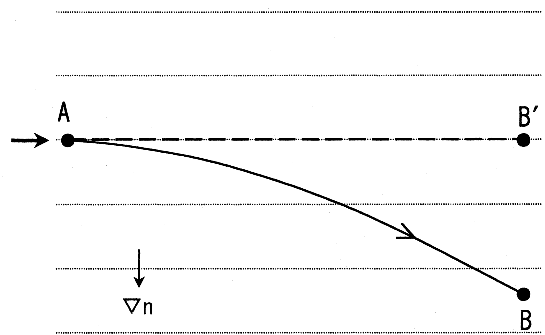

Fig. 2: Light ray incident on point A is deflected toward the gradient of refractive index n. Straight path from A to B' is impossible, though n along the line is constant.

前節の最後で示した光線の方程式 (19) によれば，光線は屈折率の増加する方向，すなわち図の下方に向

かって曲げられて，$A \to B$ のような経路をたどるはずです．しかし仮に，光線が曲げられずに $A \to B'$ という直線経路を進んだとしてみましょう．直線 AB' に沿っては屈折率は定数であり，光線はこの直線上しか進まないとすれば，全空間の屈折率が一定であるのと何も変わりがないはずです．

つまり，光線が $A \to B'$ という経路を進むことに何の異論もないはずです．現実の光線がこの経路を選ばないということを，どのように説明すればいいのでしょうか？この経路を通ってはいけないということを，光線自身がどうやって知るのでしょうか？

この問題を考えるまえに，フェルマーの原理をもちいて A と B' を通る光線を決めるなら，どんな経路になるかを考えておきましょう．この場合，まっすぐ A から B' に進むのではなく，屈折率の小さい側（すなわち図の上方）に少し回り道をすることで，経路の実際の長さは増えても光路長は最短になるようにするでしょう．すなわち，A と B' を通る現実の経路は，この 2 点を通って少し上方に反った弧になります．これは結局，光線は屈折率の大きい側に向かって曲げられながら進むという，光線の方程式 (19) と同じ結果を与えるわけです．

ところが，光線が自分の進路を決定するのに，このようにいろいろな経路の光路長をあらかじめ調べておくことはできないはずです．先に述べた問題，すなわちなぜ $A \to B'$ という経路が不可能であるかにたいしての一つの説明は，屈折率 n は値そのものではなく勾配 ∇n のみに物理的な意味があるとすることです．光線が感じるのはつねに今いる場所の ∇n であり，光線はその方向に向かって曲げられるという解釈です．つまり (19) の右辺を，質点に働く力のようなものとして見るわけです．

とすれば，光線の方程式 (19) のもとになったフェルマーの原理は次のような説明が可能でしょう．フェルマーの原理においては，現実の光線を決定するのに，それとわずかに違う経路の光線と光路長の比較をしなければなりません．その際に，屈折率 n そのものではなくてその変化の様子，すなわち勾配 ∇n が効いてくるはずです．つまり，n の勾配をさぐる働きを，変分をとる操作によって行っているということです．もしそのように解釈するなら，光線の経路を決定する法則はあくまで (19) であり，フェルマーの原理は，単にその数学的な言い換えであると見なすべきです．

しかしながら，屈折率という量は媒質を直接特徴づけるものであることを考えると，勾配だけに意味をもたせることには無理があるでしょう．しかしそれ以外に，$A \to B'$ という経路が不可能であるということを説明する方法があるでしょうか？

このように考えていくなら，光は幅をもたない光線という存在ではなく，何らかの意味で拡がった存在であるべきことに思い至るでしょう．もちろん，われわれは光が波であることを知っているので，そのような考えに行きつくわけです．光を拡がった波として考えた場合に，光線経路をどのように説明できるかを考えてみましょう．

Fig.2 の場合に，点 A に向かって入射してくる光線にたいして，ある波長をもった有限の幅の平面波を付随させて考えてみます．この光線が $A \to B$ のように曲げられるなら，付随する波は，Fig.3 のように変形されていかなければならないでしょう．すなわち，屈折率の大きな場所ほど波長が短くなると考えれば，光線が曲がる方向を説明できます．

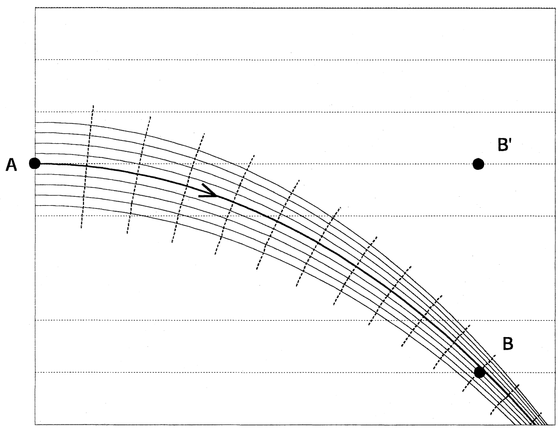

Fig. 3: Deflection of light rays is explained if the rays are accompanied with wavefronts. Wavelength λ is assumed to be a function of refractive index n.

われわれが光線とよぶものには波が付随していて，その波長が屈折率の関数であると仮定すれば，光線経路が説明できるわけです．このようにすれば，屈折率の勾配ではなく屈折率の値自身に，物理的な意味を持たせることができます．

では，この立場においては，フェルマーの原理はどのように解釈すべきでしょうか？やはり単なる数学的な言いかえでしょうか．それとも，物理的な意味を付与することができるでしょうか．次節において，まず光を波として記述する方法を定式化したのち，フェルマーの原理の解釈を再考することにします．

3.2.4　アイコナール

前節では，光の波長が屈折率の関数であると考えれば，光線経路を説明できることを述べました．ここでは，まず屈折率と波長の関係を定量的に決定しましょう．

Fig.4 のように，光線が空気と水の境界で屈折される状況を考えます．光線に平面波を付随させて考えて，水中のほうが波長が短くなると考えれば，屈折の方向が説明できます．定量的には，屈折の法則，いわゆるスネルの法則に一致しなければなりません．結果だけを示すと次のようになります．真空中（空気中）の波長を λ_0，屈折率 n の媒質中の波長を λ としたとき，

$$\lambda = \frac{\lambda_0}{n} \tag{20}$$

という関係があれば，スネルの法則が説明されます．

一方，スネルの法則はフェルマーの原理による結果と一致します．つまり，Fig.4 における 2 点 A，B をむすぶ経路のうち，光路長が最短になるものが現実の光線経路となることを示すことができます．よって，媒質の境界面における屈折の現象に関するかぎり，(20) もフェルマーの原理も，現実の光線経路を決定するという意味で対等であるわけです．

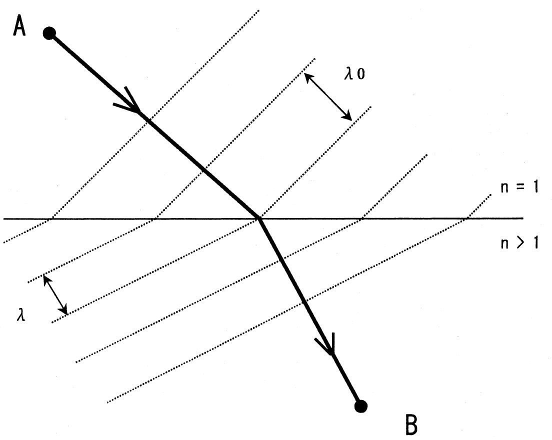

Fig. 4: Refraction of a plane wave. Snell's law is understood if a light ray is accompanied with a plane wave whose wavelength λ is inversely proportional to refractive index n.

この関係を，Fig.4 のような不連続な屈折面だけでなく，連続的な屈折率分布 $n(x,y,z)$ にたいしても要求することは自然でしょう．その際は，波長が場所の関数として与えられることになります．すなわち，

$$\lambda(x,y,z) = \frac{\lambda_0}{n(x,y,z)} \tag{21}$$

です．

もし適当に境界条件を設定するなら，上の関係をもちいて空間全体での波面の分布を決定することができるはずです．境界面での屈折現象との類推で考えるなら，そのように決定された波面に直交する曲線，すなわち光線は，フェルマーの原理から決定されるものと一致することが期待できるでしょう．

以下では，実際に波面を決定する方程式を導き，その方程式とフェルマーの原理が等価であることの証明を行います．そのために，まずいくつか準備をしておきます．

最初に，光路長を波の言葉で表しておきます．真空中では，波の進行方向に λ_0 進むごとに位相が 2π 変化します．屈折率が n の領域では，$\lambda = \lambda_0/n$ 進むごとに位相が 2π 変わります．実際の長さに n をかけたものが光路長の定義であることを思い出すと，n の値とは関係なく，光路長が λ_0 増えるごとに位相が 2π 変化することになります．この意味で，光路長とは真空中に換算した長さであるわけです．

次に，波面の分布を表現するための関数を定義します．波面とは，同じ位相をもつ点をつなげてできる面，すなわち等位相面のことでした．空間における位相分布を $\phi(x,y,z)$ という関数で与えるなら，その ϕ の等高面が波面であり，ϕ が 2π 増えるごとに次の山がきます．そこで，

$$S(x,y,z) = \frac{\lambda_0}{2\pi}\phi(x,y,z) \tag{22}$$

によって関数 S を定義するなら，位相が 2π 増えるごとに S が λ_0 増えるので，S は位相分布を光路長に換算した関数となります．この S を，**アイコナール** (eikonal) とよびます．Fig.5 に S の分布の例を示します．

アイコナール S は位相分布と実質的に同じものですが，S の定義はすぐにはピンとこないはずですので，性質を箇条書きにしておきます．

（1）S は，波の進行に沿っての光路長を与える．すなわち，ds 進むごとに nds だけ増加するような関数である．

（2）S が λ_0 増えるごとに，位相は 2π 増える．S の等高面を λ_0 ごとに描くなら，それは位相 2π ごとの波面である．

（3）λ_0 ごとに描いた S の等高面の間隔は，その場所での波長 λ を表す．

次に，アイコナール S と光線との関係を考えます．波面に直交する曲線が光線ですから，ある 1 点における S の勾配ベクトル ∇S は，その点を通る光線の方

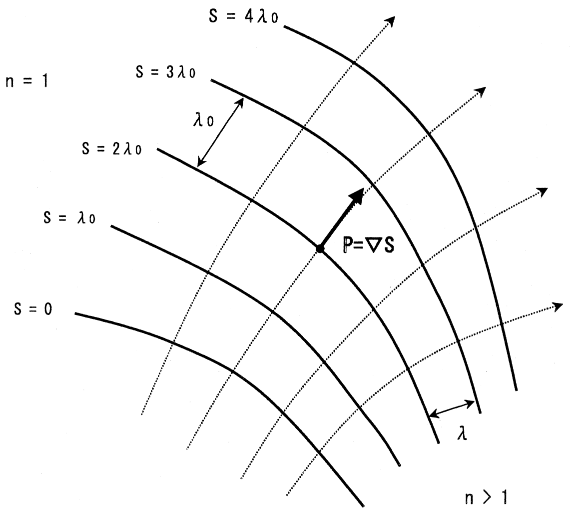

Fig. 5: Contour surfaces of eikonal $S(x, y, z)$. Geometrical light rays are defined as the orthogonal trajectories to the surfaces. The gradient of eikonal S is equal to geodesic vector $\boldsymbol{P}$ along a ray.

向を向いています．そして，その大きさ $|\nabla S|$ は，単位長さ進んだときの S の増加量です．よって，上の性質（1）からただちに，$|\nabla S|$ はその場所の n であることがわかります．すなわち，

$$|\nabla S(x, y, z)| = n(x, y, z) \tag{23}$$

です．

結局，∇S は光線の方向を向いていて大きさが n ですから，光線に沿った測地ベクトル $\boldsymbol{P} = n\boldsymbol{U}$ に他なりません．∇S はこの空間に定義されるベクトル場なので，対応する測地ベクトル $\boldsymbol{P}$ もまたベクトル場になっています．すなわち，

$$\boldsymbol{P}(x, y, z) = \nabla S(x, y, z) \tag{24}$$

です．

このように，測地ベクトル $\boldsymbol{P}$ がベクトル場として決定されるということが重要です．前節では，$\boldsymbol{P}$ は曲線に沿って定義される量として導入されました．しかし，ここではまず波面の分布としての S を考え，この波面に直交する曲線としての光線の集合がつくられたわけです．そして，それらの一つ一つの光線に沿って定義される $\boldsymbol{P}$ を集めたものが，ベクトル場としての $\boldsymbol{P}$ です．

さて次に，波面分布を決定する問題に戻りましょう．われわれは結局，アイコナール $S(x, y, z)$ を決定できればいいわけです．上で導いた (23) は，S に課すべき条件のひとつです．しかしよく考えれば，空間のある 1 点において，S という関数がみたすべき条件は (23) 以外にはないことがわかるでしょう．

すなわち，∇S が向く方向はつまりその 1 点を通る光線の方向ですが，これはどんな方向でも可能であり，なんの制約も存在しません．各点で (23) をみたし，かつその条件が大域的にうまくつながって $S(x, y, z)$ という関数がつくられるとき，それは現実に存在しうるアイコナールであるはずです．

このように見たとき，(23) は**アイコナール方程式** (eikonal equation) とよばれます．これは 1 階の偏微分方程式であり，成分で書くなら，

$$\left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2 = n^2(x, y, z) \tag{25}$$

となります．適当な境界条件，たとえば $S = 0$ の面をどこかに与えるなら，そこから等光路長の面を次々につくっていくことでアイコナール方程式の解を決定することができます．(この手続きを具体的に示したものが，ホイヘンスの原理といわれるものです.)

アイコナール方程式を導く手順を，測地ベクトル $\boldsymbol{P}$ を主役にして述べるなら次のようになります．$\boldsymbol{P}$ が局所的にみたすべき関係式として，(18) がありました．これと同時に，大域的な条件としての $\boldsymbol{P} = \nabla S$，すなわち $\boldsymbol{P}$ が S の勾配として生成されることを要求します．具体的には，(18) において，$p_x = \partial S/\partial x$，$p_y = \partial S/\partial y$，$p_z = \partial S/\partial z$ とおきかえればよいわけです．こうして得られるものがアイコナール方程式です．

アイコナール S は，(22) が示すように，空間の位相分布に対応するものであって波動そのものではありません．したがって，アイコナール方程式 (25) は波動方程式としての意味をもつものではありません．(25) は 1 階の偏微分方程式なので，解は一般には空間的に単調に変化していくようなものです．S に対応する波動関数（単に，波動を表す関数という意味です）を書くなら，角振動数を ω として，

$$\Psi(x, y, z, t) = A(x, y, z, t)\; e^{\,i\left[\dfrac{2\pi}{\lambda_0} S(x, y, z) - \omega t\right]} \tag{26}$$

というかたちになります．

なお，今までは波動の時間依存性のことを何も考えていませんでしたので，少し解説しておきます．出発点となったフェルマーの原理には時間の概念がなく，光線経路の道すじだけが考慮されます．そこで，この原理で決定される光線になんらかの波動を対応させるとすれば，時間的に実質的に変化のないもの，

すなわち空間のすべての点が同一の振動数をもった定常状態を考えるべきでしょう．そのような定常状態において，空間において同じ位相で振動している点をつないでいけば，一般にはある曲面が構成されます．この等位相面を波面とよび，この分布を記述するものが今考えているアイコナールであるわけです．

歴史的には，アイコナール方程式は電磁波の方程式から位相の部分だけを分離することで導かれたものです．一般に，波動においては振幅と位相は複雑に関係し合い，これは波動方程式の種類には依りません．よって，位相部分だけがみたす方程式というものは存在しません．しかし，屈折率が波長にくらべてゆっくり変化する，言い換えれば波長が 0 に近づいた極限を考えることで，分離が可能になるのです．

この極限操作によって，波としての特有の現象である回折が無視されてしまいます．しかしその見返りとして，光線という概念によって記述できるような，すなわち幾何光学が適用される状況がつくられます．この近似が，いわゆる**幾何光学近似** (geometirical optics approximation) です．アイコナールとは，そのような極限の状況における「波面」であり，波動方程式の厳密な解の位相を表すのではありません．その意味で，アイコナールの等高面を**幾何光学的波面** (geometirical wavefront) ということがあります．

さて次に，アイコナール方程式とフェルマーの原理が等価であることを示します．アイコナール方程式の解 S が存在すれば，その等高面に直交する曲線が光線であるわけですから，その曲線がフェルマーの原理をみたす，すなわち測地線であることを示せばよいわけです．

Fig.6 のように，アイコナール S の等高面に直交する曲線のひとつを c として，その上の 2 点 A，B をとります．この 2 点を通る別の曲線 c' をとると，c' の光路長は c の光路長より必ず大きくなることを示せばよいわけです．これが成り立つことは「図から明らか」で済ませてもよさそうに思えますが，よく考えるとそれほど明らかでもありません．

正式な証明は次のようになります．まず曲線 c の光路長は，c に沿ってのベクトル場 $\boldsymbol{P} = \nabla S$ の線積分で与えられます．しかし，一般に勾配ベクトル場の線積分は経路によらないので，このベクトル場 $\boldsymbol{P}$ を曲線 c' に沿って積分しても値は同じです．式で書けば，

$$l_c = \int_c \boldsymbol{P} \cdot d\boldsymbol{X} = \int_{c'} \boldsymbol{P} \cdot d\boldsymbol{X} \tag{27}$$

です．

一方，曲線 c' の光路長は，この曲線に沿って定義される測地ベクトルを積分すれば得られますが，この測地ベクトルは，ベクトル場としての $\boldsymbol{P}$ とは異なります．そこで，曲線 c' に沿っての測地ベクトルを $\bar{\boldsymbol{P}}$ と書けば，曲線 c' の光路長は次式で与えられます．

$$l_{c'} = \int_{c'} \bar{\boldsymbol{P}} \cdot d\boldsymbol{X} \tag{28}$$

ここで，測地ベクトルの長さはその点の n だけで決まるので ((17) 式)，曲線 c' 上のすべての点で $|\boldsymbol{P}| = |\bar{\boldsymbol{P}}|$ です．そこで，曲線 c' に沿った $d\boldsymbol{X}$ は $\bar{\boldsymbol{P}}$ の方向を向いていることを考えると，曲線 c' 上のすべての点で，

$$\bar{\boldsymbol{P}} \cdot d\boldsymbol{X} \geq \boldsymbol{P} \cdot d\boldsymbol{X} \tag{29}$$

であることがわかります．これと (27)，(28) から，$l_{c'} \geq l_c$，すなわち，曲線 c' の光路長は曲線 c の光路長よりも必ず大きくなることがわかります．これで，曲線 c は測地線であることが示されました．

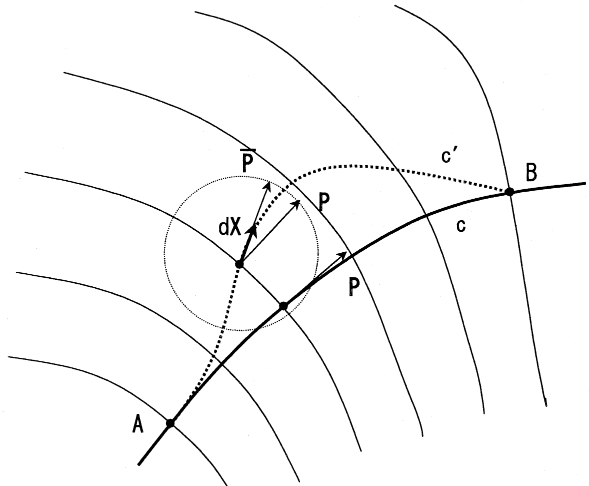

Fig. 6: Illustrating the proof of Fermat's principle. Orthogonal curve c to the contour surfaces of eikonal S turns out to be a geodesic.

以上によって，光線経路を決定するフェルマーの原理と，波としての振舞いを決めるアイコナール方程式が等価であることが示されました．光線とは「光路長に関しての最短線」であり，そしてアイコナールの等高面は「光路長に関しての等距離面」です．これらは，空間に与えられた「長さ」としての光路長が決定する幾何学的な概念であるわけです．

屈折率の与えられた空間は，感覚的な言い方をするなら，「曲がった」空間としてのイメージが可能です．波はその曲がった空間をまっすぐ進もうとして，

われわれの目から見れば歪んだ波面形状になります．波は空間的に拡がった存在なので，空間の曲がりを感じながら進んで行くことができます．

一方，光線はいくぶん人工的な概念であって，幅をもたない曲線として扱われます．1次元的な拡がりしかないために，自分の周辺の空間がどのように曲がっているのかを十分に知ることができません．そのような曲線の概念を用い，しかも本性としては波としての存在である光の行動を決定するものとして，フェルマーの原理を見ることができるでしょう．

ここまでの議論によれば，変分原理で記述されることがそのまま波動性を意味するという印象をもたれるかも知れません．しかし§3.2.3の最初に戻って考えるなら，屈折率 n 自身ではなくその勾配 ∇n に物理的意味を持たせることで，粒子的な描像に固執することは可能です．

本節では，光が決まった波長をもつ波であることを前提に定式化を始めました．しかし，最終的に導かれたアイコナール方程式には波長という量は含まれていません．アイコナールは，あくまで光線経路に直交するという意味での波面を表すものであり，回折や干渉をおこすような真の波動性に必然的に結びつくものではありません．

3.2.5　ハミルトンの特性関数

アイコナール方程式の解が見出されると，それには光線束が対応します．光線の方程式を解くのとは違って，アイコナール方程式を解くことは，無数の光線を一度に決定することです．しかし，アイコナール方程式を実際に解くのは大変です．たとえ数値計算に頼るにしても，一般には光線を一本一本決定していくほうがはるかに現実的です．

アイコナール S を考えることの理由は別のところにあります．すなわち，S を考えることで，光線の集団的な振舞いにたいしての議論が可能になるということです．ただし，ひとつの関数 S に対応する光線束はかなり限られた素性のものです．つまり，空間の1点で光線が交わることのない，たとえば点光源から光線が放出されているような状況しか考慮することができません．

現実の光線の場はもちろん，もっと複雑です．たとえば光学顕微鏡の場合を考えると，試料を照らすための光源はかならず有限の広がりをもち，その光源の各点から放出される光線束が重なり合って試料まで到達します．そのような現実的な光の場を考えるべきであり，その際にどんな法則が成立するか，あるいはどんな制約が存在するのかがわれわれの知りたいことです．

この目的のために，アイコナールの概念を少し拡張して，次のような関数を考えます．

$$S(\boldsymbol{X}_A, \boldsymbol{X}_B) = \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} \tag{30}$$

この S は前節のようなスカラー場ではなく，2点 A, B の座標 $\boldsymbol{X}_A = (x_A, y_A, z_A)$，$\boldsymbol{X}_B = (x_B, y_B, z_B)$ の関数であり，この2点を通る測地線（つまり現実の光線）に沿った光路長として定義されます．すなわち，2点 A，B が指定されたときに，その2点を通る測地線をまず求め，その測地線の経路に沿って $\boldsymbol{P}$ を積分した値が $S(\boldsymbol{X}_A, \boldsymbol{X}_B)$ です．

この関数 S は，**ハミルトンの特性関数** (characteristic function of Hamilton) とよばれます．この関数は前節のアイコナールと本質的に異なるものではないので，ここでは同じ S という名前にしています．

なお，一般に測地ベクトル $\boldsymbol{P}$ は任意の曲線に沿って定義されますが，特性関数 S の定義 (30) における測地ベクトルは測地線に沿ったものでなければなりません．慣れるまでは，この式の $\boldsymbol{P}$ はたとえば $\tilde{\boldsymbol{P}}$ のように区別して書くほうが無難です．

特性関数 S は2点の座標の関数なので，前節のアイコナールのように等高面を図示するようなことはできません．しかし，たとえば始点 A を固定して終点 B のみを動かして考えるなら，$\boldsymbol{X}_B$ の関数としての S を図示することができます．たとえばFig.7のようになるでしょう．

これは，点 A に点光源を置いたときに空間につくられるアイコナールにほかなりません．この S は,「点 A から測った距離の場」としての意味をもつことがわかるでしょう．もちろんこの場合の「距離」とは，測地線に沿って測った光路長という意味です．

いま考えている $\boldsymbol{X}_B$ の関数としての S にたいして，点 B における勾配ベクトルを考えることができます．これを ∇S_B と書くことにします．これは，2点の関数としての $S(\boldsymbol{X}_A, \boldsymbol{X}_B)$ の $\boldsymbol{X}_B$ に関しての偏微分にほかなりません．そこで，$\nabla S_B = \partial S(\boldsymbol{X}_A, \boldsymbol{X}_B)/\partial \boldsymbol{X}_B$ と書くこともできます．

さて，この $\boldsymbol{X}_B$ の関数としての S をアイコナールとして見るなら，∇S_B は，2点 A，B を通る測地線の B における測地ベクトルでもあるわけです．その測地ベクトルを $\boldsymbol{P}_B$ と書けば，$\boldsymbol{P}_B = \nabla S_B$ という関係式が得られたことになります．

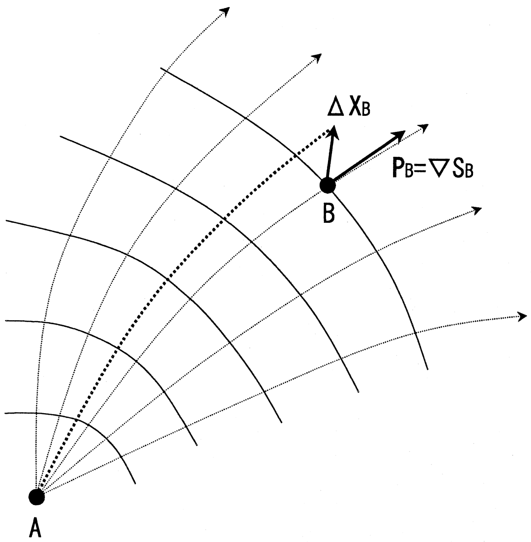

Fig. 7: Contour surfaces of Hamilton's characteristic function $S(\boldsymbol{X}_A, \boldsymbol{X}_B)$ when point A is fixed.

こんどは，点 B を固定して点 A のみを動かします．$\nabla S_A = \partial S(\boldsymbol{X}_A, \boldsymbol{X}_B)/\partial \boldsymbol{X}_A$ を考えると，これはやはり2点 A と B を通る測地線の A における測地ベクトルとなります．ただし，A における測地ベクトルは S が減少する方向に向かうので，符号だけが異なって $\boldsymbol{P}_A = -\nabla S_A$ となります．

まとめると，$S(\boldsymbol{X}_A, \boldsymbol{X}_B)$ が知れているとき，2点 A, B における測地ベクトルを与える公式として次のものが得られました．

$$\begin{cases} \boldsymbol{P}_A = -\nabla S_A \\ \boldsymbol{P}_B = \quad \nabla S_B \end{cases} \tag{31}$$

上式は，A から B にいたる光線がみたす法則を導くために，中心的な役割をはたします．その議論を行うまえに，上式をもちいてフェルマーの原理を別のかたちに言い換えることができるので，それを先に述べておきます．

まず，偏微分係数 $\nabla S_B = \partial S(\boldsymbol{X}_A, \boldsymbol{X}_B)/\partial \boldsymbol{X}_B$ は，点 B をわずかに変化させたときの，点 A からの距離の変化率を与えます．そのときの，点 B の座標の変化を $\boldsymbol{X}_B \to \boldsymbol{X}_B + \Delta\boldsymbol{X}_B$ とすれば，距離の変化は $\Delta S = \nabla S_B \cdot \Delta\boldsymbol{X}_B$ で与えられます．(Fig.7 を見てください.)

点 A のみを動かしたときも同様なので，両方を同時に動かしたときの S の変化は，(31) をもちいれば，

$$\Delta S(\boldsymbol{X}_A, \boldsymbol{X}_B) = -\boldsymbol{P}_A \cdot \Delta\boldsymbol{X}_A + \boldsymbol{P}_B \cdot \Delta\boldsymbol{X}_B \tag{32}$$

となります．あるいは (30) をもちいれば，

$$\Delta \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} = [\boldsymbol{P} \cdot \Delta\boldsymbol{X}]_A^B \tag{33}$$

と書くことができます．上式は，変分を与えるまえの積分路が測地線であるときだけ成り立つ式ですが，逆に，積分路が測地線であるための条件式と見ることもできます．

われわれは，測地線を与える変分原理として二つの表現を得たことになります．並べて書けば，

$$\delta \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} = 0 \tag{34}$$

$$\Delta \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} = [\boldsymbol{P} \cdot \Delta\boldsymbol{X}]_A^B \tag{35}$$

です．このうち (34) すなわちフェルマーの原理は，両端を固定した「δ 変分」，一方 (35) は両端を動かすことを許した「Δ 変分」によって与えられます．後者の特別な場合が前者ですので，後者のほうがより広い応用をもつことになります．この公式 (35) は，力学の最小作用の原理を導く際にもちいます．

さて，(31) に戻りましょう．測地ベクトル $\boldsymbol{P}$ の3つの成分は独立ではなく，長さがその場所での n に等しい，すなわち (17) という条件があったことを思い出します．そこで，ある1点で p_x と p_y が決まれば，p_z も決まってしまいます．よって (31) においては，p_x と p_y に関しての式だけが本質的です．

そこで，まず測地ベクトルを x, y 成分と z 成分に分離して，$\boldsymbol{P} = (p_x, p_y, p_z) = (\boldsymbol{p}, p_z)$ のように書きます．また空間の位置ベクトルも $\boldsymbol{X} = (x, y, z) = (\boldsymbol{x}, z)$ のように表します．簡単のために，以下では z_A, z_B を固定して考えて，変数として表示しないことにします．すると，(31) で x, y 成分だけを書いた式は，

$$\begin{cases} \boldsymbol{p}_A = -\dfrac{\partial S(\boldsymbol{x}_A, \boldsymbol{x}_B)}{\partial \boldsymbol{x}_A} \\ \boldsymbol{p}_B = \quad \dfrac{\partial S(\boldsymbol{x}_A, \boldsymbol{x}_B)}{\partial \boldsymbol{x}_B} \end{cases} \tag{36}$$

となります．これらを $\boldsymbol{x}_B$ と $\boldsymbol{p}_B$ について解けば（解けたとすれば），下のような関係式が得られることになります．

$$\begin{cases} \boldsymbol{x}_B = \boldsymbol{x}_B(\boldsymbol{x}_A,\ \boldsymbol{p}_A) \\ \boldsymbol{p}_B = \boldsymbol{p}_B(\boldsymbol{x}_A,\ \boldsymbol{p}_A) \end{cases} \tag{37}$$

上式は，$z = z_A$ における $(\boldsymbol{x}_A, \boldsymbol{p}_A)$ が与えられれば，$z = z_B$ での $(\boldsymbol{x}_B, \boldsymbol{p}_B)$ が知れるということを言っています．この関係式を具体的に得るためには，まず特性関数 S を決定しておく必要があります．しかしながら，アイコナール方程式を解くことでさえ十分難しいのですから，それを行うのは一般には無理です．

ではなぜ特性関数を考えるかといえば，この関数が存在するという事実だけから，S がどんな関数形に

なるかとは無関係に，光線の場がみたすべき法則が導かれるからです．(37) は，光線の始状態を与えたときに，その後の光線の行き先を決定するものです．その決まり方を決めているのは，(36) における関数 S ただ一つです．このことから，始状態 $(\boldsymbol{x}_A, \boldsymbol{p}_A)$ と終状態 $(\boldsymbol{x}_B, \boldsymbol{p}_B)$ とが全く勝手な関係をもつことは許されなくなるのです．

この法則は，ある一本の光線にたいしての制約を与えるものではありません．先に S の $(\boldsymbol{x}_A, \boldsymbol{x}_B)$ の近傍における値を自由に与えるなら，$(\boldsymbol{x}_A, \boldsymbol{p}_A)$ と $(\boldsymbol{x}_B, \boldsymbol{p}_B)$ のほとんど好き勝手な組み合わせをつくることができます．しかし，それとわずかに始状態の異なる光線を考えると，その光線の終状態も S によって決まってしまいます．すなわち S の存在は，一本の光線にたいしてではなく，光線の集合体にたいしての制限を与えるということです．

これを考えるための図を，Fig.8 に示します．一般に $F = F(x, y)$ という関数を (x, y, F) という座標をとって曲面として図示したとき，法線ベクトルは $(-\partial F/\partial x, -\partial F/\partial y, 1)$ となります．そこで，いまの $S(\boldsymbol{x}_A, \boldsymbol{x}_B)$ を $(\boldsymbol{x}_A, \boldsymbol{x}_B, S)$ という座標で表したとすれば，(36) によって，法線ベクトルは $(\boldsymbol{p}_A, -\boldsymbol{p}_B, 1)$ で与えられます．

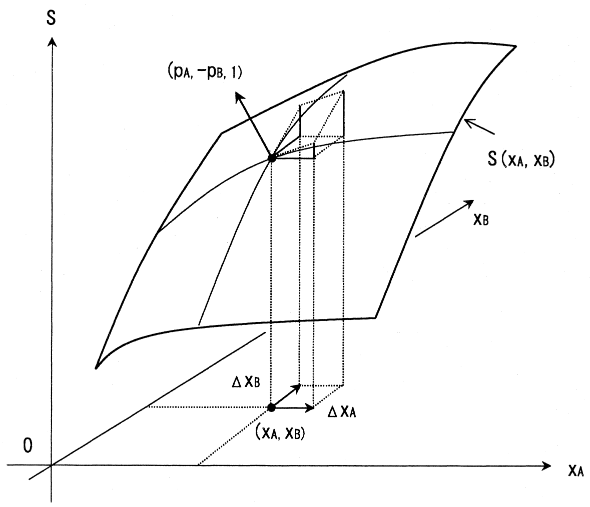

Fig. 8: Restriction imposed on light rays is derived from the presence of Hamilton's characteristic function $S(\boldsymbol{x}_A, \boldsymbol{x}_B)$. The derivatives of S with respect to $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ give deodesic vectors $\boldsymbol{p}_A$ and $\boldsymbol{p}_B$ at those points.

具体的に，光線の場にたいしてどのような制約が生じるのかが，まだわかりにくいと思います．ここでは，簡単に導くことのできる関係式を一つ示しておきます．(36) において，$\partial^2 S/(\partial \boldsymbol{x}_A \partial \boldsymbol{x}_B) = \partial^2 S/(\partial \boldsymbol{x}_B \partial \boldsymbol{x}_A)$ であることをもちいます．すなわち，(36) の第１式を $\boldsymbol{x}_B$ で微分して符号を変えたものと，第２式を $\boldsymbol{x}_A$ で微分したものを等しくおけば，

$$-\frac{\partial \boldsymbol{p}_A}{\partial \boldsymbol{x}_B} = \frac{\partial \boldsymbol{p}_B}{\partial \boldsymbol{x}_A} \tag{38}$$

となります．そこで，始状態 $(\boldsymbol{x}_A, \boldsymbol{p}_A)$ のまわりにある幅 $(\Delta\boldsymbol{x}_A, \Delta\boldsymbol{p}_A)$ をもたせ（符号は考えず），その領域に対応する終状態を考えるなら，

$$\Delta\boldsymbol{x}_A \cdot \Delta\boldsymbol{p}_A = \Delta\boldsymbol{x}_B \cdot \Delta\boldsymbol{p}_B \tag{39}$$

という関係が得られます．これは，$z = z_A$ を物面，$z = z_B$ を像面として考えるなら，像を縮小すれば $(\Delta\boldsymbol{x}_A > \Delta\boldsymbol{x}_B)$，ビームの開き角は大きくなる $(\Delta\boldsymbol{p}_A < \Delta\boldsymbol{p}_B)$ という，輝度不変則に他なりません．(説明が足りませんが，詳しくは後の章で述べます.)

力学においても，変分原理による定式化が可能であるので，全く同様の状況が生じます．力学での言葉づかいのほうが聞きなれているはずですので，上に現れてきた量を力学における用語で言うなら次のようになります．

まず $\boldsymbol{p}$ は**正準運動量** (canonical momentum)，$\boldsymbol{x}$ と $\boldsymbol{p}$ で張られる空間は**相空間** (phase space) といいます．この相空間における変換 $(\boldsymbol{x}_A, \boldsymbol{p}_A) \to (\boldsymbol{x}_B, \boldsymbol{p}_B)$ が関数 S によって (36) のように与えられるとき，**正準変換** (canonical transformation) とよびます．そして，関数 S 自身は，正準変換の**母関数** (generating function) とよばれます．

なお，phase space は古い文献では位相空間と訳されていますが，数学の集合論における位相空間 (topological space) と区別するために，最近は相空間といいます．解析力学では，相空間の位相（topology）が議論されるので，混乱を避けなければなりません．

正準変換のもつ特性を研究することは，光学と力学に共通したテーマの一つです．光学系においての輝度不変則やアッベの正弦則とよばれるもの，あるいは力学におけるリュービルの定理などは，すべて正準変換の性質として導かれるものです．これらに関しては，後章で電子分光の感度を議論する際に，詳しく述べることにします．

3.3　力学の原理

3.3.1　時空間における長さ

力学における粒子の運動は，空間と時間で構成される４次元の「時空間」における曲線と見なすことができます．時間を刻々と追いかけながら現象を眺めるというわれわれの日常の感覚とは違って，粒子の運動は「時空間における静的な曲線」と見なされます．そして，そのような曲線のうちの特別なものを選び出すものとして，力学における変分原理を確立したいわけです．

光学との類推からすれば，時空間に「長さ」の概念を導入することから始めればよいでしょう．3次元空間においては，長さとよぶべき量が最初からそなわっていました．しかし，時空間においてそのような量を考えるということの意味自体が，そう簡単ではありません．問題は，「どんな量を長さとよぶべきなのか」ということです．

光学を考えた際には，曲線の長さを具体的に表すために，空間にデカルト座標をとりました．この際に，いろいろな方向を向いたデカルト座標のうちのどれを選ぶかは問題にはなりません．すべてのデカルト座標は同等であり，これは「空間は等方的である」ということの反映であるわけです．そこで，ある特定のデカルト座標系において近接した2点をむすぶ長さ ds が $ds^2 = dx^2 + dy^2 + dz^2$ で定義されるなら，それは別のデカルト座標系 $x'y'z'$ でも $ds^2 = dx'^2 + dy'^2 + dz'^2$ でなければなりません．

もし二つのデカルト座標で同一の２点間の長さを示す２つの表式が異なったとすれば，二つの座標系が区別できることになってしまいます．逆に，同等であるすべてのデカルト座標系で長さの表式が変わらないという事実は，長さというものが幾何学的な量であり，座標とは独立した意味をもっていることを意味しています．物理法則というものは，そのような幾何学的な量によって表現されなければなりません．

そこで，時空間における「長さ」にたいしても，座標と独立した意味をもつことを要求すべきです．まず，上で述べた空間の等方性はもちろんみたす必要があります．しかし，時間軸が加わったことによって別の対称性が要求されます．それは，「互いに一定の相対速度で等速運動するデカルト座標系は，互いに対等で区別ができない」という，経験的な事実にもとづくものです．そのような座標系はいわゆる**慣性系** (inertial system) であり，すべての慣性系が対等であるという，時空間における対称性をみたす必要があるということです．

この話の続きを詳しく述べるには，マックスウェル方程式の変換性やエーテルの話などをしなければなりません．しかしこれに関しては特殊相対論の教科書にゆずり，ここでは結論だけを述べることにします．

特殊相対論は，すべての慣性系が対等であり，かつ光速度 c が慣性系によらない定数であるという経験事実を基本原理として据えるものです．複数の慣性系を考慮するには，それらの間の座標変換式が必要です．相対論以前の，すなわちニュートン力学においては，いわゆる**ガリレイ変換** (Galilei transformation) がそのような役割を果たします．二つの慣性系の相対速度 $\boldsymbol{V}$ の方向に x 軸と x' 軸をとり，$t = 0$ で原点が一致しているとすれば，ガリレイ変換は次のように与えられます．

$$x' = x - Vt,\ y' = y,\ z' = z\ ,t' = t \tag{40}$$

この変換によって，ニュートンの運動方程式は不変に保たれます．

しかしこの変換によって，ある特定の速度が一定値になることは不可能です．光速が不変であるということは，原点から発せられた球面波の波面の方程式，

$$x^2 + y^2 + z^2 = c^2t^2 \tag{41}$$

の表式が不変でなければなりません．この条件をみたす座標変換が，次の**ローレンツ変換** (Lorentz transformation) です．

$$x' = \frac{x - Vt}{\sqrt{1 - V^2/c^2}},\ y' = y,\ z' = z,\ t' = \frac{t - Vx/c^2}{\sqrt{1 - V^2/c^2}} \tag{42}$$

(42) にたいして，時空間の近接した２点 (t, x, y, z) と $(t + dt, x + dx, y + dy, z + dz)$ に関しての次の不変量が導かれます．

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 \tag{43}$$

すなわち，すべての慣性系が互いにローレンツ変換で結ばれるなら，光速は不変になり，(43) は慣性系によらない幾何学的な量になるということです．

さて，上式はデカルト座標で表した長さの式によく似ていて，しかもそれは慣性系によらない不変量です．負号が混じっているのが気になるところですが，もし線素 ds を粒子の運動に沿って考えるなら必ず正となります．なぜなら，粒子の速度 v は c を超えない

ので，(43) から

$$ds^2 = c^2dt^2\left(1 - \frac{v^2}{c^2}\right) \geq 0 \tag{44}$$

となります．

これはわれわれの求めている，時空間における幾何学的な量です．(43) はどんな慣性系の座標をもちいて計算しても同一の値になります．これを「長さ」とよぶなら，この長さが停留になるという性質は座標とは無関係であり，ある慣性系から見た停留曲線は，ほかのどんな慣性系から見ても停留曲線であるわけです．

以上のように，光速が不変であるという基本原理にもとづいて，(43) という幾何学的な概念がつくられました．「速度の次元をもつ普遍定数 c が存在する」という事実は，時間と空間を一体化する役割を果たし，これによって時空間というものが幾何学の対象となります．相対論がなければ，時間と空間は切り離された存在であって，そのような幾何学的な見方は不可能です．

3.3.2　自由粒子

前節において，時空間において長さとよぶべき量が見いだされました．この長さに関しての停留条件が何を意味するかを調べてみます．

まず，時空間における曲線の長さの表式を求めてみます．時空間の 2 点 $A(t_A, x_A, y_A, z_A)$, $B(t_B, x_B, y_B, z_B)$ を通る曲線が，

$$x = x(t),\ y = y(t),\ z = z(t) \quad t_A \leq t \leq t_B \tag{45}$$

のかたちで与えられているとします．これは時空間における静的な「曲線」と見るべきですが，われわれの感覚からすれば，時間を独立変数とした「運動」であるわけです．以下では，曲線と言ったり運動と言ったりしますが，実質的には違いはありません．

さて (43) から，曲線の長さは，

$$\begin{aligned} l &= \int_A^B ds \\ &= \int_A^B \sqrt{c^2dt^2 - dx^2 - dy^2 - dz^2} \\ &= c\int_{t_A}^{t_B} \sqrt{1 - v^2/c^2}\ dt \end{aligned} \tag{46}$$

と表されます．

この曲線に沿って実際に運動する人を考え，その人がはめている腕時計が示す時間を考えます．これをその人の**固有時間** (proper time) といいます．点 A から B まで運動する際の固有時間の進み $\Delta\tau$ は，次式で与えられます．

$$\Delta\tau = \int_{t_A}^{t_B} \sqrt{1 - v^2/c^2}\ dt \tag{47}$$

この値は，観測している人にとっての経過時間 $\Delta t = t_B - t_A$ より小さい値になります．すなわち「動いている時計は遅れる」わけです．(46) と (47) の違いは c の因子だけなので，曲線の長さと言っても固有時間と言っても，実質的には同じことです．

さて，(46) で定義される長さに関しての停留曲線は，時空間におけるまっすぐな線，すなわち等速直線運動であることを示すことができます．なぜそうなるかは，次のように考えるとわかりやすいでしょう．

まず，時空間における曲線の長さの計算値は，慣性系に依らず同一になります．(そのように長さを定義したわけです.) そこで，2 点 A, B が同一の空間位置になるような座標系に移って考えることができます．その座標系においては，(46) の最後の式における v の項の寄与によって，A から B までじっとして動かない場合が l が一番長くなり，途中で動いて最後に B にもどるような運動をすると，l はそれより小さくなります．

すなわち，ここで定義されている長さは，直線のときに最短線ではなく最長線となります．しかし停留性だけを考えるなら問題ありません．測地線という言い方は最短線にたいしてふさわしいものですが，以下では停留曲線の意味で，測地線という用語をもちいます．

結局，時空間の 2 点 A, B を通るような自由粒子の運動は，次の変分原理によって決定されることがわかりました．

$$\delta l = \delta\int_A^B ds = 0 \tag{48}$$

なお，一か所にじっとしているときに l が最長になるということは，固有時間におきかえて考えるなら，そうしているときが時間が一番多く進むということです．たとえば，宇宙旅行に行ってきた双子の兄より，地球で待っていた弟のほうが多く歳をとるわけです．

さて，ある慣性系をとり，その原点から「等距離」の点をつなげてみると Fig.9 のようになります．ユークリッド的な長さで考えるなら等距離面は球面ですが，いま考えている時空間における長さに関しての等距離面は，$ct - x$ 平面上で描くなら図のような直角双曲線になります．

ここで考えている等距離面は，光学におけるアイコナールを時空間で考えたものとなります．ここでのアイコナールは，$S = S(t, x, y, z)$ というかたちの関数です．アイコナールの等高面に直交する曲線が測地線である，と言いたいところですが，図の上ではこれらは直交していません．しかし，われわれは時空間のベクトルとか内積をまだ定義していないことに注意しましょう．現時点では,「直交する」という言葉をもちいることができないわけです．

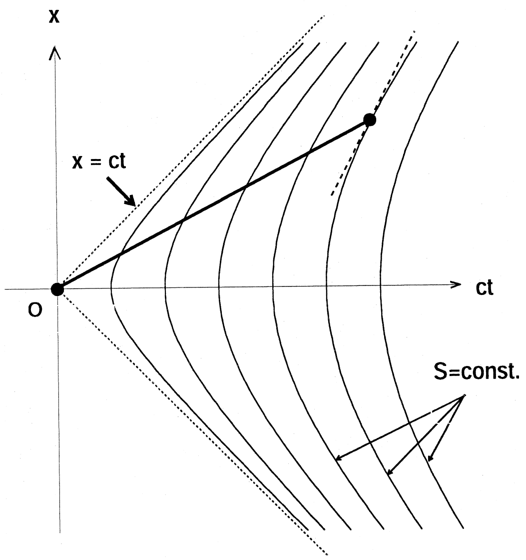

Fig. 9: Contour surfaces of the distance measured from origin O in four-dimensional space-time. $\boldsymbol{x}$ symbolizes three space directions.

3.3.3 時空間における測地ベクトル

時空間でのベクトルの概念を導入するために，まず3次元空間におけるベクトルの定義から考えておきます．3次元ベクトルは，デカルト座標の変換にたいして座標の変換と同じ変換をうける3成分の量として定義されます．近接する2点の座標の差 $d\boldsymbol{X} = (dx, dy, dz)$ はそのような量です．

4次元の時空間においては，座標 (ct, x, y, z) と同じ変換をうける量が **4元ベクトル** (4-vector) とよばれます．一般に4元ベクトルは，

$$A_\mu = (A_0, \boldsymbol{A}) = (A_0, A_x, A_y, A_z) \tag{49}$$

のように表します．ここで A_0 は A_μ の時間成分，$\boldsymbol{A}$ は空間成分とよばれます．たとえば (cdt, dx, dy, dz) は4元ベクトルで，これを $dx_\mu = (cdt, d\boldsymbol{x}) = (cdt, dx, dy, dz)$ のように記します．

長さの定義(43)は，4元ベクトル dx_μ の長さの2乗と見ることができますが，これを4元ベクトルの内積として，$dx_\mu dx_\mu = |dx_\mu|^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$ のように書きます．一般に，2つの4元ベクトル A_μ, B_μ があったとき，それらの内積は，

$$\begin{aligned} A_\mu B_\mu &= A_0B_0 - \boldsymbol{A}\cdot\boldsymbol{B} \\ &= A_0B_0 - A_xB_x - A_yB_y - A_zB_z \end{aligned} \tag{50}$$

で定義されます．

この内積が0になるとき，A_μ と B_μ は「直交する」といいます．もし時空間を $ct - x$ 平面として2次元的に見るなら，A_μ と B_μ の始点を原点に移したとき，直線 $x = ct$ にたいして互いに折り返した角度の関係にあるときに直交します．

また，3次元の勾配演算子 ∇ に対応するものは，

$$\nabla_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right) \tag{51}$$

で定義されます．この勾配ベクトルは，普通のユークリッド的な意味での勾配にたいして，空間成分の符号を反転させたものです．Fig.9において，アイコナール S の等高面は直線と直交し，S の勾配 $\nabla_\mu S$ は直線の方向を向きます．

なお，A_μ のように書くのはテンソル解析の指標記法です．本来は添え字の上つき（反変成分）と下つき（共変成分）を区別すべきですが，ここではその区別をする代わりに，内積の定義における符号を調整することで済ませています．

次に，時空間における測地ベクトルを求めることを考えます．光学の場合には，自由空間では曲線の単位接ベクトル $\boldsymbol{U}$ が測地ベクトルであったので，それに対応する単位4元ベクトル u_μ を求めればよいことになります．もちろん，単位ベクトルといっても4元ベクトルとしての長さが1ということで，(50)で定義された内積に関して $u_\mu u_\mu = 1$ となるような4元ベクトルのことです．

もう少していねいに言えば，まず光学で3次元の単位接ベクトル $\boldsymbol{U}$ をもちいた理由は，線素が $|d\boldsymbol{X}| = \boldsymbol{U}\cdot d\boldsymbol{X}$ として与えられたからです．$\boldsymbol{U}$ 自身は $\boldsymbol{U} = d\boldsymbol{X}/|d\boldsymbol{X}|$ として求められます．いま必要な u_μ は $|dx_\mu| = u_\mu dx_\mu$ となるべきもので，成分は次のように求められます．

$$\begin{aligned} u_\mu &= \frac{dx_\mu}{|dx_\mu|} = \frac{(cdt,\ d\boldsymbol{x})}{\sqrt{c^2dt^2 - d\boldsymbol{x}^2}} \\ &= \left(\frac{1}{\sqrt{1 - v^2/c^2}}, \frac{\boldsymbol{v}}{c\sqrt{1 - v^2/c^2}}\right) \end{aligned} \tag{52}$$

これは，光学においてもちいた，3 次元空間の単位ベクトルの成分 (9) に対応するものです．

この u_μ を測地ベクトルとすればよいわけですが，実はこの成分は，定数を除けばよく知られた量です．すなわち，今考えている粒子の質量を m として，上式に mc をかけたものを p_μ と定義すれば，

$$p_\mu = mcu_\mu = \left(\frac{mc}{\sqrt{1-v^2/c^2}}, \frac{m\boldsymbol{v}}{\sqrt{1-v^2/c^2}} \right) \quad (53)$$

となります．この時間成分は粒子の相対論的なエネルギー E を c でわったもの，空間成分は運動量 $\boldsymbol{p}$ となっています．すなわち，

$$\begin{cases} E = \dfrac{mc^2}{\sqrt{1-v^2/c^2}} \\ \boldsymbol{p} = \dfrac{m\boldsymbol{v}}{\sqrt{1-v^2/c^2}} \end{cases} \quad (54)$$

です．よって (53) は，

$$p_\mu = \left(\frac{E}{c}, \, \boldsymbol{p} \right) \quad (55)$$

と書けることがわかります．p_μ は **4 元運動量** (4-momentum) とよばれるものです．

この p_μ は，単位ベクトル u_μ から $p_\mu = mcu_\mu$ として定義されたので，p_μ の長さは mc です．このことを成分で表すなら，

$$\frac{E^2}{c^2} - \boldsymbol{p}^2 = m^2c^2 \quad (56)$$

となります．これは，4 元運動量の成分，すなわち運動量とエネルギーが独立な量ではないことを直接に示しています．

時空間の単位ベクトル u_μ よりも，それと定数だけ違う 4 元運動量 p_μ をもちいるほうが，物理的な意味が明確になります．そこで以下では，(46) に mc をかけたものを改めて曲線の長さ l と定義します．これによって，l はふつう長さの次元をもつ量ではなく，作用とよばれる量の次元（エネルギー × 時間）をもつことになります．この約束のもとで，測地ベクトルは p_μ となり，曲線の長さが次のように与えられます．

$$l = \int_A^B p_\mu dx_\mu = \int_A^B Edt - \boldsymbol{p} \cdot d\boldsymbol{x} \quad (57)$$

さて，上の議論を逆に辿って考えるなら，エネルギーと運動量がいっしょになって 4 元ベクトル p_μ をつくることを証明したことになります．粒子のエネルギーと運動量は力学において重要な役割を演じますが，これらは単位ベクトル u_μ の成分と定数の違いしかないわけです．つまり，エネルギーも運動量も，時空間の曲線のかたちを示すための幾何学的な量であるということです．

考えてみれば，ニュートン力学における運動エネルギー $T = mv^2/2$ と運動量 $\boldsymbol{p} = m\boldsymbol{v}$ は，m を除いて考えるなら残りは $\boldsymbol{v} = d\boldsymbol{x}/dt$ だけで決まり，この $\boldsymbol{v}$ は時空間における曲線の傾きです．つまり，曲線の形状に関しての幾何学的な量です．

3 次元空間における通常の意味での長さと，時空間の曲線の長さの表式を並べて書いてみましょう．3 次元空間の長さは，測地ベクトル $\boldsymbol{P}$ $(= \boldsymbol{U})$ をもちいて，

$$\begin{aligned} l_3 &= \int_A^B \boldsymbol{P} \cdot d\boldsymbol{X} \\ &= \int_A^B p_x dx + p_y dy + p_z dz \end{aligned} \quad (58)$$

でした．一方，ここで考えてきた 4 次元の時空間の長さは，

$$\begin{aligned} l_4 &= \int_A^B p_\mu dx_\mu \\ &= \int_A^B Edt - p_x dx - p_y dy - p_z dz \end{aligned} \quad (59)$$

となります．(58) での A，B は 3 次元空間の 2 点であり，(59) では時空間における 2 点ですが，区別しないで書いています．3 次元空間の測地ベクトルの成分と 4 元運動量の空間成分が同じ名前 p になってしまいましたが，むしろこの対応を考えて，光学における測地ベクトルを最初から $\boldsymbol{P}$ と書いたわけです．(58) と (59) の違いは，それぞれの空間における長さの定義の違いに対応しています．

さて，光学の場合には測地ベクトルの成分間に (18) という関係がありましたが，それに対応するここでの関係式は (56) です．この条件式において，ベクトル場としての p_μ がアイコナールの勾配 $\nabla_\mu S$ として生成されることを要求すれば，それがそのまま S にたいしてのアイコナール方程式となります．

すなわち，光学における (24) に対応して $p_\mu = -\nabla_\mu S$ とおきます．これを成分ごとに書けば，$E = -\partial S/\partial t$，$\boldsymbol{p} = \nabla S$ となります．(空間成分を光学の場合と同じにするために，$\nabla_\mu S$ に負号をつけました．) これらを (56) に代入すれば，

$$\frac{1}{c^2}\left(\frac{\partial S}{\partial t}\right)^2 - (\nabla S)^2 = m^2c^2 \quad (60)$$

が得られます．これが自由粒子にたいしてのアイコ

ナール方程式であり，力学では**ハミルトン・ヤコビ方程式** (Hamilton-Jacobi equation) とばれます．

3.3.4　電磁場中の荷電粒子

光学の場合，空間にもとから備わっている長さに屈折率という場 $n(x,y,z)$ をかけたものを光路長として定義すれば，現実の光線は光路長に関しての測地線となるのでした．

同様に，粒子に力が作用する場合にも，時空間における長さの定義をうまく定義し直して，粒子の運動が測地線になるようにしたいわけです．その場合の長さの定義は，すべての慣性系で表式が不変に保たれることをとりあえず要求すべきでしょう．しかし，そのような制限のもとでも，自由空間における長さの定義 (43) を一般化する方法は一つではありません．

力の存在のもとでの長さの定義を決定する作業は結局，われわれがすでに知っている運動方程式を正しく導くようなものとして決めるしかありません．電磁場中の荷電粒子の相対論的な運動方程式は，前章ですでに示したように，

$$\frac{d}{dt}\left(\frac{m\boldsymbol{v}}{\sqrt{1-v^2/c^2}}\right) = q\left(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B}\right) \tag{61}$$

です．電子の場合は，上式で $q=-e$ $(e>0)$ とおきます．この運動方程式が正しく導かれるような変分原理をつくりだすことができればよいわけです．（変分原理と等価な微分方程式に関しては，変分法の節 §3.4 を見てください．）これは単なる試行錯誤の作業なので，あとは結果だけ示します．

まず電磁場は，一つの 4 元ベクトル，

$$A_\mu = \left(\frac{\Phi}{c},\ \boldsymbol{A}\right) \tag{62}$$

で完全に記述されることが知られています．ここで $\Phi(t,x,y,z)$ はスカラーポテンシャル，$\boldsymbol{A}(t,x,y,z)$ はベクトルポテンシャルで，A_μ は**4元ポテンシャル** (4-potential) とよばれます．電磁場との関係は次次のように定義されます．

$$\begin{cases} \boldsymbol{E} = -\nabla\Phi - \dfrac{\partial \boldsymbol{A}}{\partial t} \\ \boldsymbol{B} = \nabla\times\boldsymbol{A} \end{cases} \tag{63}$$

この 4 元ポテンシャルをもちいて，時空間の曲線に沿って積分できる量をつくるとすれば，一番簡単なものは $A_\mu dx_\mu$ です．自由空間における長さの式 (57) にこの積分を加えれば，新しい長さが定義できます．

この際に，$A_\mu dx_\mu$ にかける定数を自由に選ぶことができますが，これを電荷 q に選ぶと，対応する微分方程式は正しく電磁場中の電子の運動方程式を与えることが確認されます．すなわち，時空間における曲線の長さを与える表式は次のようになります．

$$\begin{aligned} l &= \int_A^B (p_\mu + qA_\mu)\,dx_\mu \\ &= \int_A^B (E+q\Phi)dt - (\boldsymbol{p}+q\boldsymbol{A})\cdot d\boldsymbol{x} \end{aligned} \tag{64}$$

上式の変分を 0 とおいたもの，すなわち，

$$\delta\int_A^B (E+q\Phi)dt - (\boldsymbol{p}+q\boldsymbol{A})\cdot d\boldsymbol{x} = 0 \tag{65}$$

が，電磁場中の荷電粒子の運動を与える変分原理となります．

さて，4 元ポテンシャルには，ある不定性が存在します．すなわち，もとの A_μ のかわりに，$\chi(t,x,y,z)$ を任意関数として $A'_\mu = A_\mu - \nabla_\mu\chi$ でおきかえても，それから導かれる電磁場は不変です．これは**ゲージ変換** (gauge transformation) とよばれるもので，成分に分けて書けば，

$$\begin{cases} \Phi' = \Phi - \dfrac{\partial\chi}{\partial t} \\ \boldsymbol{A}' = \boldsymbol{A} + \nabla\chi \end{cases} \tag{66}$$

となります．

(64) の積分にたいしてこのゲージ変換をほどこすと，新たに付け加わる量は $\int (d\chi/dt)dt$ となり，これは両端の座標だけで決まるので変分には寄与しません．すなわち，変分原理はポテンシャルで表されていますが，決定される荷電粒子の運動はゲージ変換に影響されません．これは，運動方程式が電磁場だけで表されていることに正しく対応しています．

光学においてフェルマーの原理を考えた際に，屈折率 n とその勾配 ∇n のどちらが物理的意味をもつかという問題がありました．力学の問題においてこれに対応するのは，ポテンシャルと，その微分である電磁場のどちらが物理的な実体であるかということです．

上で見たように，粒子の運動を決定するのは電磁場です．しかし，もし光学のときのように，測地ベクトルの線積分 (64) を位相変化にむすびつけるなら，積分値を決定するポテンシャル自身が意味をもたなくてはなりません．

ポテンシャルの実在性はすでに実証されていて，(64) の積分を位相に対応させることで量子力学を構築することが可能です．これは，経路積分の方法とよ

ばれています．(ただし，粒子にたいするアイコナール理論がそのまま量子力学を導くのではありません．章末の文献を参照してください．)

さてここで，本質的なことではありませんが，定義の変更と記号の約束をしておきます．まず，(64) は相対論的な要請から自然にかたちが定められたものですが，ふつうの教科書では，非相対論的な変分原理から出発する都合で，(64) の符号を変えたものが変分原理にもちいられます．変分原理では停留条件のみを問題にするので，符号はどうでもいいわけです．

そこで，以下では慣例に合わせる意味で，l を改めて次のように定義します．

$$l = \int_A^B \boldsymbol{p}_c \cdot d\boldsymbol{x} - H dt \tag{67}$$

ここで，

$$\begin{cases} \boldsymbol{p}_c = \boldsymbol{p} + q\boldsymbol{A} = \dfrac{m\boldsymbol{v}}{\sqrt{1 - v^2/c^2}} + q\boldsymbol{A} \\ H = E + q\Phi = \dfrac{mc^2}{\sqrt{1 - v^2/c^2}} + q\Phi \end{cases} \tag{68}$$

です．

上式のうち，まず $\boldsymbol{p}_c$ は§3.2.5 での言葉づかいで正準運動量とよぶべきもので，力学的な運動量 $\boldsymbol{p}$ に，ベクトルポテンシャル $\boldsymbol{A}$ による項が付け加わっています．$\boldsymbol{p}$ とは異なり，$\boldsymbol{p}_c$ は $\boldsymbol{A}$ の項を含むために，一般には粒子の運動方向を向いていないことに注意しましょう．H はエネルギーであり，静止エネルギーと運動エネルギー，スカラーポテンシャル Φ による位置エネルギーを合わせたものです．

さて，変分原理と等価な微分方程式を求める際には，まず変分の対象となる積分を時間積分のかたちで表しておく必要があります．(67) にたいしてその書き換えを行うと次のようになります．

$$\begin{aligned} l &= \int_A^B (\boldsymbol{p}_c \cdot \boldsymbol{v} - H) dt \\ &= \int_{t_A}^{t_B} L dt \end{aligned} \tag{69}$$

ここで L は，

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} - q(\Phi - \boldsymbol{A} \cdot \boldsymbol{v}) \tag{70}$$

で与えられます．

一般に，力学の変分原理が時間積分によって表されているとき，その被積分関数は**ラグランジアン** (Lagrangian)，変分原理は**ハミルトンの原理** (Hamilton's principle) とよばれます．よって，いまの場合のラグランジアンは (70) であり，荷電粒子の運動を決定するハミルトンの原理が次式で与えられるわけです．

$$\delta \int_{t_A}^{t_B} L dt = 0 \tag{71}$$

3.3.5　非相対論的極限

いままで見てきたように，変分原理にもとづいて力学を定式化するためには，相対論的な扱いが不可欠です．しかし，だからといって非相対論的な力学，すなわちニュートン力学を与える変分原理が存在しないというわけではありません．

ラグランジアン (70) において，電磁場の項を落としたもの，すなわち第 1 項は自由粒子のラグランジアンです．これを，v/c が十分小さいとしてべき展開すれば，

$$L = -mc^2 \left(1 - \frac{v^2}{2c^2} + \cdots \right) = -mc^2 + \frac{1}{2} m v^2 + \cdots \tag{72}$$

となります．変分に寄与しない定数は無視できるので，ラグランジアンは非相対論的な極限で v^2 であることになります．この時間積分は，

$$\begin{aligned} \int_{t_A}^{t_B} v^2 dt &= \int_{t_A}^{t_B} \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 \right] dt \\ &= \int_A^B \frac{dx^2 + dy^2 + dz^2}{dt} \end{aligned} \tag{73}$$

よって，これがもし曲線の長さとしての $\int ds$ というかたちの積分であるなら，線素 ds は，

$$ds = \frac{dx^2 + dy^2 + dz^2}{dt} \tag{74}$$

で与えられることになります．これは今までのような「$ds^2 =$ 座標の微分の 2 次形式」というかたちとは異なりますが，問題は，これがガリレイ変換 (40) で不変であるかということです．

それを見るには (40) の逆変換を上式に代入してみればよいわけです．しかし，v^2 の時間積分がガリレイ変換で不変でないことを示すだけなら，そのような計算をするまでもありません．すなわち，ある慣性系 K から見た粒子の速度 v と，別の慣性系 K' から見た速度 v' が異なるのは当然です．

しかし，K にとって v^2 の時間積分が停留になるような運動が等速直線運動であるのなら，K' にとって v'^2 の積分が停留になるのもやはり等速直線運動です．

すなわち，積分の値そのものは変わってしまっても，それが停留になるという条件は正しく運動を決定するわけです．

そこで，われわれは非相対論的な力学を変分原理の枠組みに取り込むために，定義の拡張を迫られることになります．すなわち，停留となるべき積分はすべての慣性系での不変量であるという要求は，捨てなければなりません．そのかわり，たとえ積分値そのものは変わってしまっても，それが停留性に影響しなければよい，つまり最終的に決定される運動が同一であればよいということです．

電磁場中でのラグランジアン (70) の非相対論的な表式は，

$$L = \frac{1}{2}mv^2 - q(\Phi - \boldsymbol{A}\cdot\boldsymbol{v}) \tag{75}$$

となります．もし磁場が存在しなければ，

$$L = \frac{1}{2}mv^2 - q\Phi \tag{76}$$

となって，$L = T - U =$（運動エネルギー）－（位置エネルギー）という，解析力学の教科書でおなじみの公式となるわけです．一方，相対論的に正しい L すなわち (70) は，たとえ磁場がなくても $L = T - U$ のかたちではありません．

3.3.6　最小作用の原理

ハミルトンの原理は，電磁場中の電子の運動を $(x(t), y(t), z(t))$ のかたちで得るためのものです．しかし前回（第2章）ですでに述べたように，電子光学で問題となるのは空間中の道すじとしての電子の「軌道」であり，時間の概念は必要がありません．そこで運動方程式から時間を消去して，電子軌道をたとえば $(x(z), y(z))$ のかたちで求めるための方程式を導いたわけです．

ここでは，ハミルトンの原理から時間を消去して，軌道を決定する変分原理を導きます．最終的に導かれるものが，最小作用の原理とよばれるものです．

まず，どんな場合でも時間を消去できるわけではないことに注意する必要があります．これができるためには，時間座標の値そのものが本質的ではなく，その相対値だけが問題となるような系でなければなりません．具体的には，ラグランジアンが時間を陽に含まないということであり，そのような場合にだけエネルギーが保存されます．(§3.4.4 で証明を述べます.) 第2章では，運動方程式から時間を消去するためにエネルギー保存則をもちいましたが，その必然性がここで理解されます．

出発点となる変分原理を，(67) によって次のように書いておきます．

$$\delta\int_A^B \boldsymbol{p}_c\cdot d\boldsymbol{x} - H dt = 0 \tag{77}$$

上式は，時空間における2点 A と B をむすぶ曲線のうちの一つを選び出すものですが，これを何らかの意味で xyz 空間へ射影して，時間をのぞいた空間における軌道を決める原理として書き直すことを考えてみます．

つまり Fig.10 のように，2点 A と B を空間方向に射影した2点を $\bar{A}$, $\bar{B}$ として，この2点を通る xyz 空間上の曲線のうちから，現実に電子がたどる軌道を選び出したいわけです．(この図では，空間を xy 平面として，時間を加えた時空間が3次元になるように描いています.)

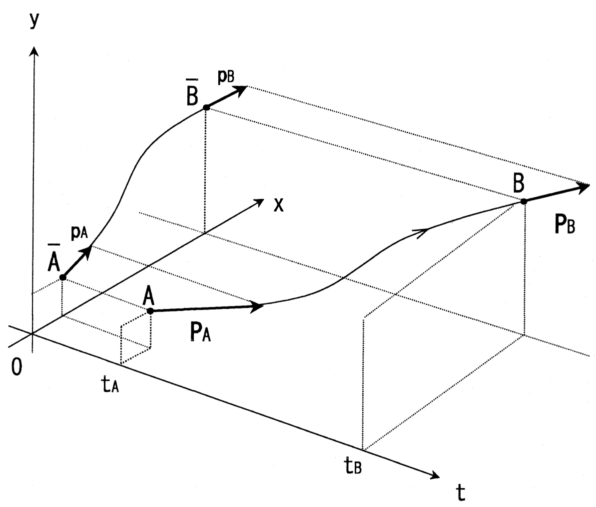

Fig. 10: Illustrating the projection of the motion joining two points A and B in space-time onto the trajectory joining corresponding points $\bar{A}$ and $\bar{B}$ in space. The principle of least action is derived from Hamilton's principle using this projection.

運動方程式から時間を消去したときと同様に，(77) においてエネルギー保存則を適用することを考えてみます．すなわち，(77) の試行関数として，ある一定のエネルギーをもつものに限定するということです．現実の運動はエネルギー保存則をみたすので，試行関数をそのように限定しても正しい解が得られるはずです．

このとき，(77) における H に関しての積分は定数になり，したがって変分には寄与しません．つまり H を含む項はおとすことができて，(77) の原理は次の

ように簡略化されます．

$$\delta\int_A^B \boldsymbol{p}_c \cdot d\boldsymbol{x} = 0 \tag{78}$$

この段階ではまだ $\boldsymbol{p}_c$ は時間の概念を含んでいて，積分路は (77) と同じ時空間におけるものです．しかし第2章で示したように，エネルギー保存則をもちいると $\boldsymbol{p}_c$ は場所の関数として決まってしまうのでしたから，上式の積分は，2点 $\bar{A}$，$\bar{B}$ を通る曲線を積分経路として，

$$\delta\int_{\bar{A}}^{\bar{B}} \boldsymbol{p}_c(\boldsymbol{x}) \cdot d\boldsymbol{x} = 0 \tag{79}$$

でおきかえることができます．これで，時間がすべて追い出されたことになります．

そこで，粒子の軌道を決定するための変分原理として，この (79) をもちいればよいように思えます．実は，この結果そのものは正しいのですが，途中の推論は誤っています．しかし，厳密な議論をしようとするとかなりやっかいなので，上のような説明だけで済ましている教科書も少なくありません．

最小作用の原理は，もともとモーペルテューイが目的論的な思索から「思いついた」ものであり，なにか理論的な根拠があったわけではありません．結果的には正しかったわけですが，いざきちんと証明をしようとすると大変なのです．以下では，この原理の正確な証明を与えるのではなく，その概略だけを述べておきます．

われわれの導きたい変分原理は，2点 $\bar{A}$，$\bar{B}$ を通るすべての曲線のなかから，あるエネルギー値 $H = H_0$ に対応する現実の軌道を選び出すものです．2点 $\bar{A}$，$\bar{B}$ を通る曲線を，エネルギー値が H_0 であることをもちいて時空間における曲線に変換することはいつでも可能です．しかし，そのときの始点 A をたとえ共通に選んだとしても，終点 B の時刻 t_B は，曲線ごとに同じになる保障はありません．

つまり，われわれが試行関数としたい，2点 $\bar{A}$，$\bar{B}$ を通る曲線は，時空間においては両端の時刻はまちまちであるということです．したがって，そのような曲線のなかから現実のものを選び出すために，(77) の原理をそのまま適用することはできません．

われわれの目的のためには，(77) において両端を動かしたときの積分値の変化を与える公式が必要です．これは光学において (35) で与えられているもので，これを今の場合に書くと，

$$\Delta\int_A^B \boldsymbol{p}_c \cdot d\boldsymbol{x} - H dt = \left[\boldsymbol{p}_c \cdot \Delta\boldsymbol{x} - H\Delta t\right]_A^B \tag{80}$$

となります．ここで，Δ は両端の移動までを含めた変分を表します．積分値の変化が上式で与えられることと，もとの積分路が測地線であることが同値です．

上式において，まず $H = H_0$ とすれば，両端における Δt のずれの寄与は両辺で同じになります．つまり，両端の時刻がどのようにずれても，両辺でうまく打ち消し合ってくれるわけです．さらに，$\bar{A}$ と $\bar{B}$ は固定なので $\Delta\boldsymbol{x}_A = \Delta\boldsymbol{x}_B = 0$ であり，これで (78) が得られます．

ここでさらに $\boldsymbol{p}_c$ を場所の関数としてしまえば，両端 $\bar{A}$，$\bar{B}$ を固定した原理と見なすことができます．これで結局 (79) となります．このようにして正しく導かれる (79) が，モーペルテューイの最小作用の原理であるわけです．この原理は，これからの電子光学の定式化のための出発点になります．

3.4　変分法

本章のテーマは変分原理ですが，なるべく変分法の知識なしで済ませてきました．ここでは変分法の立場から，前節までの内容をまとめなおします．

3.4.1　変分問題の一般的な扱い

光学におけるフェルマーの原理は，変分問題の典型です．これは次式で与えられていました．

$$\delta l = \delta\int_{z_A}^{z_B} n(x,y,z)\sqrt{x'^2 + y'^2 + 1}\, dz = 0 \tag{81}$$

この積分路は，次のかたちで与えられる空間曲線です．

$$x = x(z),\ y = y(z) \quad z_A \le z \le z_B \tag{82}$$

このような変分問題を一般化して考えるために，以下では，

$$\delta I = \delta\int_{z_A}^{z_B} F(x,y,x',y';z)\, dz = 0 \tag{83}$$

のかたちで考えていきます．歴史的に最初に研究された変分問題は，このタイプのものです．たとえば，ベルヌーイの最速降下線の問題 (1696) が有名です．

上式における積分 I は，曲線にたいしての一般化された「長さ」を定義するものと見なすことができます．そして，両端を固定したときに I が停留になるような曲線を，停留曲線あるいは測地線とよびます．(83) のような積分は，関数（曲線）にたいして一つの数値を対応させる規則と見て，**汎関数** (functional)

とよぶことがあります．これは，「関数を独立変数とする関数」という意味です．

光学の場合には，(83) における x，y，z は通常のユークリッド空間のデカルト座標です．しかし，もし力学を考えるなら，曲線のパラメタとしての z は実は時間 t であり，xy 平面は xyz 空間として読みなおすことになります．力学の場合は F は L と書いて，それをラグランジアンとよぶのでした．

一粒子の力学ではなく，もっと多自由度の力学系にたいしても，(83) のかたちの変分原理が存在します．そのような場合でも，ただ一つの関数 F によって系が記述されるのは同じであり，この事実が変分原理をもちいることの大きなメリットです．変数がいくつあっても本質は同じなので，数学的な定式化は (83) にたいして行えば十分です．

さて，(83) の被積分関数 F の引数を $F(x,y,x',y';z)$ のように書いているのは，積分路 (82) に沿った座標だけでなく，その傾きまで指定しないと値が定まらないことを主張しています．すなわち，ある z における座標 (x,y) と傾き (x',y') が必要ということです．このような量の身分を明確にするために，いくつか言葉の定義をしておきます．

まず，z は純粋に曲線を表示するためのパラメタと見なし，位置座標 (x,y) をもつ2次元空間を**配位空間** (configuration space) とよびます．そして，その各点に傾き x',y' の自由度をくくりつけた，座標 (x,y,x',y') をもつ4次元空間を**状態空間** (state space) とよびます．ある z において (x,y,x',y') が指定されることで，その z における曲線の状態が定まります．すると，$(x,y,x',y';z)$ という引数をもつ量は，スカラーであれベクトルであれ，状態空間上で定義された量，あるいは簡単に「状態空間の量」とよぶことができます．

(83) のかたちの変分原理には，それと等価な常微分方程式が存在します．その導出の過程を以下に示します．まず，簡単のために $\boldsymbol{x}=(x,y)$，$\boldsymbol{x}'=(x',y')$ とおいて，(83) を，

$$\delta I=\delta\int_{z_A}^{z_B}F(\boldsymbol{x},\boldsymbol{x}';z)dz=0 \tag{84}$$

と書きます．この積分路 $\boldsymbol{x}(z)$ にたいして小さな変分 $\delta\boldsymbol{x}(z)$ を与えて，曲線を，

$$\boldsymbol{x}(z)\to\boldsymbol{x}(z)+\delta\boldsymbol{x}(z) \tag{85}$$

のように変化させます．ただし，両端は動かさないという約束のもとで，

$$\delta\boldsymbol{x}(z_A)=\delta\boldsymbol{x}(z_B)=0 \tag{86}$$

という条件をつけます．

さて，(84) の積分路がこのように変化したときの積分値の変化は，$\delta\boldsymbol{x}$ に関して1次の項のみを残すと次のようになります．

$$\begin{aligned}\delta I &= \delta\int_{z_A}^{z_B}F(\boldsymbol{x},\boldsymbol{x}';z)dz\\ &= \int_{z_A}^{z_B}\left[F(\boldsymbol{x}+\delta\boldsymbol{x},\boldsymbol{x}'+\delta\boldsymbol{x}';z)-F(\boldsymbol{x},\boldsymbol{x}';z)\right]dz\\ &= \int_{z_A}^{z_B}\left(\frac{\partial F}{\partial\boldsymbol{x}}\cdot\delta\boldsymbol{x}+\frac{\partial F}{\partial\boldsymbol{x}'}\cdot\delta\boldsymbol{x}'\right)dz\\ &= \int_{z_A}^{z_B}\left(\frac{\partial F}{\partial\boldsymbol{x}}-\frac{d}{dz}\frac{\partial F}{\partial\boldsymbol{x}'}\right)\cdot\delta\boldsymbol{x}dz\\ &\qquad +\left[\frac{\partial F}{\partial\boldsymbol{x}'}\cdot\delta\boldsymbol{x}\right]_{z_A}^{z_B}\end{aligned} \tag{87}$$

この変形の最後で，部分積分をもちいています．上式で与えられる1次の変分が0になるような曲線が，停留曲線とよばれます．

上式の最後の式において，まず (86) の条件から，第2項（境界項）は0となります．第1項において，$\delta\boldsymbol{x}(z)$ は任意関数なので，任意の $\delta\boldsymbol{x}(z)$ にたいして第1項が0になるためには，次式が成り立つ必要があります．

$$\frac{d}{dz}\frac{\partial F}{\partial\boldsymbol{x}'}-\frac{\partial F}{\partial\boldsymbol{x}}=0 \tag{88}$$

数学書では，この証明がかならずつきますが，ほぼ明らかなので省略します．上式は成分ごとに書けば，

$$\left\{\begin{aligned}\frac{d}{dz}\frac{\partial F}{\partial x'}-\frac{\partial F}{\partial x}=0\\ \frac{d}{dz}\frac{\partial F}{\partial y'}-\frac{\partial F}{\partial y}=0\end{aligned}\right. \tag{89}$$

となります．

上式が変分問題 (84) に対応するオイラーの方程式であり，力学においてこれに対応するのがラグランジュの運動方程式です．一般に，変分問題はこのような常微分方程式系と等価であるわけです．

さて，光学の定式化においては，測地ベクトル $\boldsymbol{P}$ が本質的な役割をはたしました．ここでも，(84) の積分を測地ベクトルをもちいた表式に書きなおすことを考えます．すなわち，

$$\int_{z_A}^{z_B}Fdz=\int_A^B\boldsymbol{P}\cdot d\boldsymbol{X} \tag{90}$$

となるように，$\boldsymbol{P}=(p_x,p_y,p_z)$ の表式を決定することです．

光学の場合には，幾何学的な意味を考えて $\boldsymbol{P}$ を決めてしまいましたが，一般の F によって積分が定義されているとき，それを上式のように表す方法はすぐにはわかりません．そもそも，上式では $\boldsymbol{P}$ の $d\boldsymbol{X}$ 方向の射影成分だけが問題になるので，この条件だけで $\boldsymbol{P}$ を決めることはできません．

光学において，$\boldsymbol{P}$ がどのような役割を果たしたかを思い出しましょう．2点間の距離の関数としてのハミルトンの特性関数 S を考えたとき，$\boldsymbol{P}$ はその2点における S の勾配を与えました（(31) 式）．一般の変分問題でも，$\boldsymbol{P}$ はそのような意味をもつものとして定義されるべきでしょう．

具体的には，(31) と等価な (35)，すなわち次式を要求します．

$$\Delta\int_{z_A}^{z_B} F dz = [\boldsymbol{P}\cdot\Delta\boldsymbol{X}]_A^B \tag{91}$$

これと，(87) の最後の式で，変分を与えるまえの積分路が測地線の場合，すなわち，

$$\delta\int_{z_A}^{z_B} F dz = \left[\frac{\partial F}{\partial \boldsymbol{x}'}\cdot\delta\boldsymbol{x}\right]_{z_A}^{z_B} \tag{92}$$

を見比べます．

(92) における変分 δ は，端点で任意の変位を許す Δ 変分とは違いますが，この式の段階ではまだ両端固定の条件を使っていません．そこで，(91) と (92) の右辺において，x と y の変分に関しての変化率を等しくおくことができます．これにより，

$$p_x = \frac{\partial F}{\partial x'},\ p_y = \frac{\partial F}{\partial y'} \tag{93}$$

が得られます．

あとは p_z を求めればいいわけですが，それは (90) から，$Fdz = p_x dx + p_y dy + p_z dz$ となるべきことから得られます．この結果を (93) といっしょにして書けば，

$$\int_{z_A}^{z_B} F dz = \int_A^B p_x dx + p_y dy + p_z dz \tag{94}$$

ここで，

$$\begin{cases} p_x = \dfrac{\partial F}{\partial x'} \\ p_y = \dfrac{\partial F}{\partial y'} \\ p_z = F - \dfrac{\partial F}{\partial x'}x' - \dfrac{\partial F}{\partial y'}y' \end{cases} \tag{95}$$

となります．(95) の3つの量はすべて状態空間の量，すなわち曲線を指定してはじめて，その曲線に沿って決まる量です．すなわち，$p_x = p_x(x, y, x', y'; z)$ などです．

さて，光学における測地ベクトル $\boldsymbol{P}$ の成分は互いに独立ではなく，代数的関係式 (18) をみたすのでした．ここで議論している一般の場合でも，上で見たように p_x と p_y から p_z が決まってしまいます．この関係式を具体的に表示するには，(95) の最初の2式を x' と y' に関して解いてそれを最後の式に代入します．これによって，次のようなかたちの関係式が導かれます．

$$p_z = p_z(x, y, p_x, p_y; z) \tag{96}$$

光学の場合と同様に，$\boldsymbol{P}$ をベクトル場と考え，それが関数 $S = S(x, y, z)$ から $\boldsymbol{P} = \nabla S$ として生成されることを要求すれば，S のみたすべき方程式，すなわちアイコナール方程式が得られます．具体的には，上式で $p_x = \partial S/\partial x$，$p_y = \partial S/\partial y$，$p_z = \partial S/\partial z$ とおきかえればよいわけです．

光学におけるアイコナール方程式の解 S が見出されると，それに対応するベクトル場 $\boldsymbol{P}$ が測地線の場を与えることを §3.2.4 で示しました．これは，一般の場合でも証明することができます．

3.4.2　ハミルトン形式

力学において，(83) のかたちで変分原理を与え，それに対応するオイラーの方程式，すなわちラグランジュの運動方程式をもちいて系を研究する手法を，**ラグランジュ形式** (Lagrangian formulation) といいます．このラグランジュ形式から出発して，これと等価な，**ハミルトン形式** (Hamiltonian formulation) とよばれるものに移行することができます．

ハミルトン形式における概念の多くはすでに登場していて，あとはよびかたの問題です．まず，測地ベクトル $\boldsymbol{P}$ の成分のうち，$\boldsymbol{p} = (p_x, p_y)$ を正準運動量とよび，$p_z = -H$ とおいて H を**ハミルトニアン** (Hamiltonian) とよびます．

力学においては H はエネルギーとなるので，そのようにおくと都合がよいのです．光学の場合には，すでに §3.2.2 で見たように $p_z > 0$ なので，H は負の量になってしまいます．しかしそのような場合でも，考えている系を「ハミルトン系」とよんで，力学と同じ言葉で議論を行うことができます．

ハミルトニアン H ともとの F との関係は，数学的にはいわゆる**ルジャンドル変換** (Legendre transformation) となります．これを説明するために，まず

(94) と (95) を H によって書き直しておくと,

$$\int_{z_A}^{z_B} F dz = \int_A^B p_x dx + p_y dy - H dz \tag{97}$$

$$\begin{cases} p_x = \dfrac{\partial F}{\partial x'} \\ p_y = \dfrac{\partial F}{\partial y'} \\ H = \dfrac{\partial F}{\partial x'}x' + \dfrac{\partial F}{\partial y'}y' - F \end{cases} \tag{98}$$

となります．力学においてこれに対応するものが (67) です．

上式の段階では，p_x，p_y，H は状態空間の量であり，F から導かれた量としてこれら 3 つは対等です．つまり，単に測地ベクトルの 3 つの成分ということです．ここで，(98) の最初の 2 式を x' と y' で解いて，それを最後の式に代入すれば，H は p_x と p_y の関数として定まります．これはもちろん (96) と同じことで，

$$H = H(x, y, p_x, p_y; z) \tag{99}$$

のかたちになります．狭義では，H がこのかたちで表された段階ではじめて，ハミルトニアンとよばれます．

以上の手順において，(98) の最初の 2 式を，F をもちいて p_x と p_y という変数を定義する式と見ることができます．そして最後の式を，p_x と p_y の関数としての H（すなわち (99)）を定義する式と見ます．この，x' と y' の関数としての F から，p_x と p_y の関数としての H への変換が，数学でルジャンドル変換とよばれるものに一致します．

ハミルトニアン H が (99) のかたちで与えられれば，アイコナール方程式を直ちに書き下すことができます．すなわち，(99) において $p_x = \partial S/\partial x$，$p_y = \partial S/\partial y$，$H = -\partial S/\partial z$ とおくことで，

$$\frac{\partial S}{\partial z} = -H\left(x, y, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}; z\right) \tag{100}$$

が得られます．これは力学では，ハミルトン・ヤコビ方程式とよばれます．このように，もともとは運動に沿ってのエネルギーを表す量であったハミルトニアンは，(99) のかたちで与えられることで，系の運動を支配する関数として昇格されます．

もっと直接的に，ラグランジュの運動方程式 (89) を F ではなく H によって表すことができます．結果のみを記すと，(89) と等価な運動方程式として次式が得られます．

$$\begin{cases} x' = \dfrac{\partial H}{\partial p_x},\ y' = \dfrac{\partial H}{\partial p_y} \\ p'_x = -\dfrac{\partial H}{\partial x},\ p'_y = -\dfrac{\partial H}{\partial y} \end{cases} \tag{101}$$

これが**ハミルトンの運動方程式** (Hamilton's equations of motion)，あるいは正準運動方程式とよばれるものです．

(89) は x と y に関しての 2 階の微分方程式系ですが，これは x，y，x'，y' に関しての 1 階の微分方程式系と見なすことができます．一方，上式 (101) は x，y，p_x，p_y に関しての 1 階の微分方程式系です．つまり，ラグランジュ形式における (x', y') という変数の組を，(p_x, p_y) でおきかえたものがハミルトン形式であると見ることができます．

われわれは光学を考えた際に，光線に沿っての p_x と p_y をくくりつけた空間，すなわち (x, y, p_x, p_y) という座標をもつ相空間の重要性を見ました．力学において，ラグランジュ形式における運動法則は状態空間上のものですが，ハミルトン形式においては，それが相空間における運動法則に翻訳されるわけです．

ハミルトニアン H によって生成される相空間の運動を正準変換といいます．H の具体形とは無関係に，ある H が存在して運動が (101) のように決定されるとき，その運動は相空間における正準変換です．正準変換にはその母関数とよばれるものが対応し，光学においてはそれがハミルトンの特性関数であるわけです．

このようなメリットによって，最近では光学でも電子光学でも，ハミルトン形式による定式化が主流です．しかし，いいことばかりではありません．たとえば，(x', y') という変数は曲線の形状に直接的に関係しますが，(p_x, p_y) はそうではなく，これが具体的に何を表すかをつねに意識していなければなりません．電磁場中の荷電粒子の場合では，(p_x, p_y) が運動方向を向いているという保障すらありません．

さらに，ラグランジアンの時間積分は相対論的に不変な意味をもちますが，ハミルトニアンは 4 元運動量の成分なのでそうではありません．とは言え，たとえば粒子加速器の設計などでは正準変換性がものをいうので，ハミルトン形式に依らなければ実用的な議論ができません．

3.4.3 フェルマーの原理

前節までに述べた結果を，フェルマーの原理 (81) にたいして適用してみます．ただしラグランジュ形式で考えます．まず，この場合の F は次式で与えられます．

$$F = n(x,y,z)\sqrt{x'^2 + y'^2 + 1} \tag{102}$$

これにたいしてオイラーの方程式 (89) を書き下すと，

$$\begin{cases} \dfrac{d}{dz}\left(\dfrac{nx'}{\sqrt{x'^2+y'^2+1}}\right) = \dfrac{\partial n}{\partial x}\sqrt{x'^2+y'^2+1} \\ \dfrac{d}{dz}\left(\dfrac{ny'}{\sqrt{x'^2+y'^2+1}}\right) = \dfrac{\partial n}{\partial y}\sqrt{x'^2+y'^2+1} \end{cases} \tag{103}$$

となります．これは $x(z)$ と $y(z)$ に関しての連立2階常微分方程式です．

光線経路を決定するためには上式で十分ですが，この方程式の意味を考えやすくするために，少し変形します．まず，$ds^2 = dx^2 + dy^2 + dz^2 = (x'^2+y'^2+1)dz^2$ から，

$$\frac{d}{ds} = \frac{1}{\sqrt{x'^2+y'^2+1}}\frac{d}{dz} \tag{104}$$

となるので，これを (103) にもちいれば，

$$\begin{cases} \dfrac{d}{ds}\left(n\dfrac{dx}{ds}\right) = \dfrac{\partial n}{\partial x} \\ \dfrac{d}{ds}\left(n\dfrac{dy}{ds}\right) = \dfrac{\partial n}{\partial y} \end{cases} \tag{105}$$

と書けます．ここで，

$$\boldsymbol{U} = \frac{d\boldsymbol{X}}{ds} = \left(\frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds}\right) \tag{106}$$

が曲線の単位接ベクトルであることをもちいると，(105) の2式に対応する z に関しての式が得られて，3成分をまとめると次式となります．

$$\frac{d}{ds}(n\boldsymbol{U}) = \nabla n \tag{107}$$

あるいは，測地ベクトル $\boldsymbol{P} = n\boldsymbol{U}$ で表せば，

$$\frac{d\boldsymbol{P}}{ds} = \nabla n \tag{108}$$

となります．これで，§3.2.2 の (19) が導かれたわけです．

なお，(107) の左辺の微分を展開すれば，

$$\frac{d\boldsymbol{U}}{ds} = \frac{1}{n}[\nabla n - (\nabla n \cdot \boldsymbol{U})\boldsymbol{U}] \tag{109}$$

となります．これは前章で導いた，静電ポテンシャル $\Phi(x,y,z)$ の中での電子の軌道方程式において，$n = \sqrt{\Phi}$ とおいた場合の式になります．つまり，静電レンズにたいしては $\sqrt{\Phi}$ を屈折率と考えれば，電子軌道は光線経路と同じになるわけです．

3.4.4 ハミルトンの原理

すでに，力学における定式化に関しても多くを述べましたが，ここで改めて整理しておきます．質点にたいしてのハミルトンの原理は，次のかたちになります．

$$\delta l = \delta\int_{t_A}^{t_B} L(x,y,z,\dot{x},\dot{y},\dot{z};t)dt = 0 \tag{110}$$

この被積分関数がラグランジアンとよばれるわけです．(まぎらわしいですが，ハミルトンの原理はラグランジュ形式に属するものです．ここではラグランジュ形式のみを述べます.)

まず (110) の積分路は，時空間における曲線，すなわち，

$$x = x(t),\ y = y(t),\ z = z(t) \quad t_A \le t \le t_B \tag{111}$$

のかたちで与えられた運動です．ハミルトンの原理に等価な微分方程式は，前節と同様にして導かれて，

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{x}} - \dfrac{\partial L}{\partial x} = 0 \\ \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{y}} - \dfrac{\partial L}{\partial y} = 0 \\ \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{z}} - \dfrac{\partial L}{\partial z} = 0 \end{cases} \tag{112}$$

となります．これがラグランジュの運動方程式です．これが，あらかじめ知られている運動方程式に一致するように L を選ぶことになります．

エネルギー保存則と L の関係を§3.3.6 で触れましたが，これは次のように導かれます．(67) と (69)，(80) を組み合わせると，

$$\Delta\int_{t_A}^{t_B} Ldt = [\boldsymbol{p}_c \cdot \Delta\boldsymbol{x} - H\Delta t]_A^B \tag{113}$$

という関係が得られます．これは，変分を与えるまえの積分路が測地線であるときにだけ成り立つ式であり，ハミルトンの原理を一般の Δ 変分に関するものに拡張したものです．

上式における Δ 変分として，曲線を t 方向に一定値 Δt だけ平行移動する変分を考えます．もし L が t に陽に依存しなければ，この変分に対して (113) の左辺は0となり，一方右辺は，$\Delta\boldsymbol{x} = 0$ なので $-(H_B - H_A)\Delta t$

となります．よって $H_A = H_B$，すなわちエネルギーが不変であることが導かれます．

このエネルギー保存則に限らず，運動量や角運動量などの保存則は，L の対称性のみから導かれます．これが変分原理をもちいることの大きなメリットです．

さて，すでに述べたように，相対論的な L は慣性系間のローレンツ変換で不変であるように定義されます．L の積分が停留になるという要求は幾何学的な意味をもち，したがって座標とは無関係な意味をもちます．そこで，ある座標系で見た測地線は，どんな座標系で見てもやはり測地線であるわけです．

とすれば，L にたいして許される座標変換は，慣性系間のものに限定する必要はないはずです．たとえば空間座標としてデカルト座標以外の曲線座標をもちいたとしても，ハミルトンの原理の要求する内容自身はそれとは無関係に，同じ意味を保っているはずです．そこで，L をたとえば極座標で書き直し，その L にたいしてラグランジュの運動方程式を書き下せば，それは極座標で表した正しい運動方程式であるはずです．

さらに，ハミルトンの原理は時空間の曲線の一つを選び出す働きをするので，その内容において空間座標と時間座標は対等です．そこで，上に述べたような曲線座標だけでなく，時間まで含むような変換，たとえば回転座標系への変換を行ったとしても，停留曲線は正しい解を与えます．

具体的には，もとのデカルト座標と新座標 (q_1, q_2, q_3) の関係を，

$$\begin{cases} x = x(q_1, q_2, q_3, t) \\ y = y(q_1, q_2, q_3, t) \\ z = z(q_1, q_2, q_3, t) \end{cases} \tag{114}$$

とすれば，新座標で表した運動方程式は，

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad i = 1, 2, 3 \tag{115}$$

として与えられます．もちろんこのときの L は，もとの L を新座標で書き直したものです．

このような，一般座標変換にたいしての「共変性」がハミルトンの原理の威力です．これは，変分原理という形式が，幾何学的な性質によって特別な曲線を選び出すものであるということからの帰結です．

ただし，この性質をあまり過大評価してはなりません．われわれは，任意の座標にたいして成り立つような，つまり慣性系とは限らない一般の座標系で成り立つような運動法則を手に入れたわけではありません．(115) が表す運動方程式はあくまで，慣性系で与えられた運動法則を一般の座標で書き直したというだけで，単なる数学的な書き換えにすぎません．つまり物理的には何も付け加わっていないのです．

解析力学の定式化によって任意の座標が許されるというのは，このような意味です．本章で力学を議論した際には，ローレンツ変換で移りあう慣性系の座標しかもちいませんでした．しかし，それによって本質はすべて尽くされるわけです．

3.5　まとめ

力学において，粒子にたいする運動方程式は，単に時間を追って粒子の座標変化を決定するだけのものです．しかし，運動を変分原理で記述するなら，運動に沿って「長さ」が決まり，その等距離面としての「波面」の概念が生じます．運動に沿って定義される測地ベクトルは，その「長さ」の増加率を与えるものです．

粒子的な観点から眺めた測地線には，波面がつねに付随していて，測地線の集団としての振舞いが一定の法則をみたすことになります．これは幾何光学の輝度不変則などの法則に対応するもので，力学においては正準変換の性質として説明されます．

この正準変換性に起因する法則は，粒子の座標と測地ベクトルの成分（正準運動量）を組み合わせた空間，すなわち相空間において有効に研究されることになります．

3.6　文献紹介

まず，解析力学の入門書は数多くありますが，次のものを挙げておきます．

[1] 小出昭一郎，解析力学，岩波書店 (1983)
[2] 高橋康，量子力学を学ぶための解析力学入門，講談社 (1978)
[3] J.V.Jose and E.J.Saletan, Classical Dynamics: A Contemporary Approach, Cambridge(1998)

まず [1] は簡潔によくまとまっており，[2] は解析力学とはどういうものかをてっとりばやく知るのに適しています．

[3] は訳本はありませんが，基本的なことから高度なことまで丁寧に書かれていて，これからじっくり勉強しようとする人にお勧めです．

本格的なものとしては，次のものがあります．

[4] 山本義隆，中村孔一，解析力学 I，II，朝倉書店 (1998)
[5] 野間進　他訳，ゴールドスタイン，古典力学，吉岡書店 (1959); H.Goldstain, Classical Mehanics, 2nd edition, Addison-Wesley(1981)
[6] 高橋康 監訳，ランチョス，解析力学と変分原理，日刊工業新聞社 (1992); C. Lanczos, The variational Principles of Mechanics, 5th edition, Dover(1986)
[7] R.Abraham and J.E.Marsden, Foundations of Mechanics, 2nd edition, Addison-Wesley(1985)

[4] は日本人によって書かれた唯一の本格派で，あらかじめ数学的な素養が要求されます．[5] はよく引用され，要領よくまとまっていますが，いくぶん形式的なので独習用としては適さないように思います．(とは言え，独習用に一冊で済むものはなかなかありません．)

[6] は変分原理を中心に書かれています．(これは名著なのですが，邦訳はミスプリが多いです．) [7] は，極度に抽象化された解析力学の姿が示されています．集合論と多様体の話から始まっていて，数学的にかなり高度です．変分原理に関しては多く書かれていません．

なお，本章で定義した測地ベクトルという名称は，これらの教科書では出てきません．本章では光学と特殊相対論の枠内で議論したので「ベクトル」でよかったのですが，一般的な状態空間上では共変ベクトル，あるいは微分形式として理解すべきものです．[4] や [7] では，微分形式の導入にかなりのページ数を費やしています．本章での測地ベクトルに相当するものは，たいてい「基本1形式」のようなそっけない名前がつけられています．

次のものは，変分原理を出発点として力学法則を導くという一貫したスタイルで書かれています．本章でラグランジアンを決める手順を述べる際は，これらを参考にしました．

[8] 広重徹　他訳，ランダウ，リフシッツ，力学，東京図書 (1974) ; L.Landau, E.Lifshitz, Mechanics, 3rd edition, Pergamon(1976)
[9] 恒藤敏彦　他訳，ランダウ，リフシッツ，場の古典論，東京図書 (1978) ; L.Landau, E.Lifshitz, The Classical Theory of Fields, 4th edition, Pergamon(1975)

力学と電磁気学における4次元的な定式化に関しては，次の文献の記法に合わせました．

[10] 戸田盛和訳，ファインマン物理 IV 電磁波と物性，岩波書店 (1968)
[11] 宮島龍興訳，ファインマン物理学 III 電磁気学，岩波書店 (1969)

[11] ではラグランジアンの積分と波動関数の位相の関係に触れられていて，経路積分の基本的なアイデアを知ることができます．ベクトルポテンシャルの実在性，いわゆるアハロノフ・ボーム効果に関してはこれに説明があります．

次に，数学書を二つ挙げておきます．

[12] ゲリファント，フォーミン，変分法，文一総合出版 (1970)
[13] 犬井鉄郎，偏微分方程式とその応用，コロナ社 (1957)

[12] は変分法に関してのまとまった教科書です．測地線の場からハミルトン・ヤコビ方程式に至るまでの証明の積み重ねの過程が詳しく示されています．

[13] は1階の偏微分方程式に関して詳しくかかれた数少ない本です．ハミルトン・ヤコビ方程式は1階の偏微分方程式ですが，たいていの偏微分方程式の本は2階の方程式が中心になっています．

光学に関しては，第1章でも引用した次のものが代表的です．

[14] 草川徹　他訳，光学の原理（第5版訳）I-III，東海大学出版会 (1974) ; M. Born and E. Wolf, Principles of Optics, 6th ed., Pergamon (1980)

マックスウェル方程式からのアイコナール方程式の導出，フェルマーの原理の証明，ハミルトンの特性関数に関してはこれに示されています．

［前回の訂正］

(58) 式の2行目：$\dfrac{1}{(1+\gamma)\hat{\Phi}} \;\rightarrow\; \dfrac{1}{(1+\gamma)\Phi}$

また，§2.5.5 において，磁界レンズのつくる磁場および走査型電子顕微鏡の収差に関しての記述は不十分でした．後の章で改めて詳しく述べます．

［次回予告］

次章は次の内容を予定しています．収差係数の表式，および収差補正との関連を今回述べる予定でしたが，次回以降の内容とさせていただきます．

1．近軸軌道方程式の導出
2．電子レンズのつくる電磁場と，そのべき展開
3．近軸軌道と電磁場の計算例
4．収差係数の導出