「核酸および蛋白質合成反応の情報理論」

大阪大学理学部生物学教室　　塚　本　吉　彦

（現在の勤務先：東北大学医学部生理学第二教室）

内　　容

塚本吉彦

要　約：

核酸および蛋白質合成反応の情報理論が作られた。これは鋳型符号が基質符号を相補的対応性によって選択する過程を，速度論の立場から取り扱うものである。一つの鋳型符号に対して相補的な基質符号が出現するまでの基質符号の系列が基質符号語である。基質符号語の長さは，それを構成する基質符号の個数で計られ，幾何分布をなし，平均値を有する。与えられた鋳型のすべての符号に関して平均された，一つの鋳型符号あたりに必要な基質符号語の平均長を表わす関数（F）が導かれる。この関数の最小値として，鋳型符号の鎖状の系列の不確定性および媒液の基質符号の衝突過程の不確定性を表わす関数が求められた。鋳型の不確定性を表わす関数（G）によって，鋳型の貯える遺伝的な情報量が定義される。その情報量の単位をコット（cot）と名付けた。媒液の不確定性を表わす関数（M）は，合成反応系に供給することのできる基質の実効的な種類の数を意味する。また，それは異種分子の混合過程に含まれる，熱力学的な不確定性（エントロピー）だけでは表わすことのできない，情報的な性質を表わす。それらの関数（G，M）は確率変数がすべて等しい時に最大値をとること，符号を結合して高次の符号を新たに作る時に乗法性を有することを基本的性質とする。合成の情報理論は従来の通信の情報理論と同じ数理計画法上の構成をもっている。しかし前者の不確定性を表わす関数は，後者のそれが加法性を有することと比べて，乗法性を有する。この顕著な相異点は両者の識別機構の違いにもとずく。

1. 序　論

核酸（DNA または RNA）は，基質（ヌクレオチド），鋳型（DNA または RNA）合成酵素，およびその他の因子からなる化学反応系で合成される（例えば，Cold Spr Harb. Symp. Quant. Biol. 1968，1970）。そのとき相補的に対応する塩基の一対が，２あるいは３対の水素結合によって作られる（Watson & Crick，1953 a，b）。それらの塩基はシングレット符号として鋳型符号系と基質符号系を構成する。

タンパク質は，基質（アミノアシル tRNA），鋳型（mRNA），合成酵素（リボゾーム），およびその他の多くの因子からなる化学反応系で合成される（例えば，Cold Spr. Harb. Symp. Quant. Biol. 1969）。 そのとき相補的に対応する塩基のトリプレットの一対が塩基間水素結合によって作られる（Nirenberg，et al，1966)。それらの塩基はトリプレット符号として鋳型符号（コードン）系と基質符号（アンチコ

ードン）系を構成する。

鋳型符号の鎖状の系列の特徴は，パタンを構成するすべての符号が同時にしかも常に並存することである。パタンの安定にとって符号間の結合が必須であり，それはリン酸エステル結合に依っている。符号間結合が変わらない限り，いわば鋳型符号のパタンは時間的に伝達（遺伝）される。

基質符号の時間的な系列の特徴は，パタンを構成する符号が特定の空間部位に一つずつ継起することである。パタンの形成にとって，各符号が順次に更新することが必須であり，それは熱的自由運動に依っている。基質の熱的運動がある限り，パタンは更新過程として展開される。基質符号のパタンの展開は時間経過を必要とする過程である。

核酸および蛋白質の合成反応は鎖の開始，伸長，終端の三相に大別される。開始および終端の問題は，この論文では取り扱わない。われわれはこれから，鎖の伸長相とりわけ鋳型符号による相補的な基質符号の選択過程に着目しよう。

鎖の伸長相において，合成部位は鋳型鎖に沿って鋳型符号を一つずつ走査する。それぞれの鋳型符号による基質符号の選択は，基質符号の時間的な系列にたいして行なわれる。したがって，選択時間経過を必要とする過程であり，それを対象にした速度論が問題となりうる。

シャノンの通信に関する情報理論は，通報の伝送速度に関係する理論である（Shannon & Weaver，1949；Khinchin，1953；Fano，1963）。例として確率的に独立な通報からなる情報源の電気通信を考えよう。通報は二元符号の時系列に変換して伝送される。個々の二元符号は二者択一の選択操作によって決定される。一つの通報あたりに必要な平均符号個数（符号語の平均の長さ）が定義される。符号語の平均の長さが短かければ短かいほど伝送速度は大きくなる。出現確率の大きな通報に比較的短かな符号語を対応させ，出現確率の小さな通報に比較的長い符号語を対応させることによって，符号語の平均の長さを最小にできる。最適な完全符号語が作られたとき，二元符号の時系列の符号あたりの不確定性あるいは個々の二元符号の担う情報量は１（ビット）である。符号語の平均の長さの最小値はビット単位で計られた情報源の事象あたりの不確定性あるいは符号語の担う情報量に等しい。

情報は不確定な状態を確定した状態へ転換させる媒介物である。情報量は確定した状態に転換した不確定な状態の量である。不確定な状態の量はそれを確定した状態へ転換させるために必要な標準化された選択操作の平均回数で計られる。選択操作は具体的な

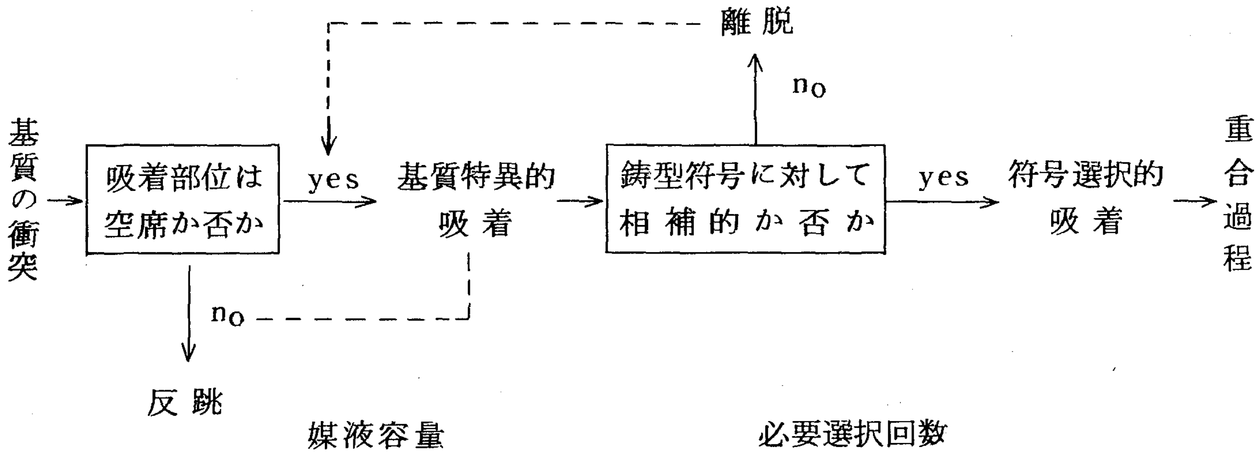
識別機構に依存する。

鋳型を介する合成反応の場合，鋳型符号の基質符号に対する相補的か非相補的かの判別が二者択一の選択操作である。その識別機構は電気通信の場合と同じではない。しかし，この場合にも，一つの鋳型符号あたりに必要な基質符号の平均個数が定義されるのではないだろうか。もしそれが定義されるならば，その値は反応速度に関係する。またその最小値としては鋳型符号の鎖状の系列の不確定性あるいはそれの貯える情報量が計られるであろう。

2．鎖伸長の反応速度論

媒液には鋳型符号に対応する基質符号が溶けているとする。合成によって消費された分を補給するホメオスタシス機構（負のフィードバック制御）を仮定すれば，それらの濃度は一定と考えられる。熱力学的な平衡状態，すなわち媒液中の基質の速度，混合比，濃度の均一性を仮定する。媒液の基質符号のプールから鋳型符号に相補的な基質符号が選択される過程を，以下のように設定する（図1）。

図1　基質の流れ図（相補的な基質符号の選択過程）



媒液容量

$$V = \frac{1}{\frac{1}{k_2} + \frac{1}{k_1 C}}$$

必要選択回数

$$F = \sum_{i=1}^{n} \frac{p_i}{q_i}$$

まず合成酵素の吸着部位に基質符号が熱運動によって衝突する。もしも吸着部位が空席ならば，塩基符号以外の部位における弱い結合によって，それは基質特異的に，しか

も塩基符号に対しては非選択的に吸着される。逆に，他の基質が吸着部位を占有しているならば，その基質は反跳する。いわば，類似物による可逆的拮抗阻害が起る。単位時間あたりの基質特異的吸着の平均回数が媒液容量（V）である。これは基質の全濃度に依存する関数である。

基質符号の時間的な系列に対して鋳型符号による選択が行なわれる。合成部位に出現した鋳型符号に対して，基質符号が相補的な場合には，それを塩基符号間の水素結合によって符号選択的に吸着し，反応は後続の重合過程へと進む。他方，基質符号が非相補的な場合には，それを吸着部位から離脱し，吸着部位は空席にもどる。このような操作が，相補的な基質符号が吸着するまでくり返される。これは，鋳型符号が基質符号に対して相補的か非相補的かを判別し，より分ける二者択一の選択過程である。この過程に対して，のちに述べるように，一つの鋳型符号あたりに必要な基質符号の平均個数（F）が定義される。これは各種の基質符号の相対濃度に依存する関数である。

媒液容量をV（1／sec），必要な基質符号の平均個数をFとすれば，一個の相補的な基質符号を選択するために必要な時間は，F／V（sec）である。選択された一個の基質符号を重合するために必要な時間をT（sec）とする。一個の基質を付加するために必要な時間は，選択時間と重合時間の和である。それは鎖伸長の合成反応速度v（1／sec）の逆数に等しい。したがって，

$$1/v = F/V + T$$

すなわち，Vは大きければ大きいほど，Fは小さければ小さいほど反応速度は大きくなる（図2）。

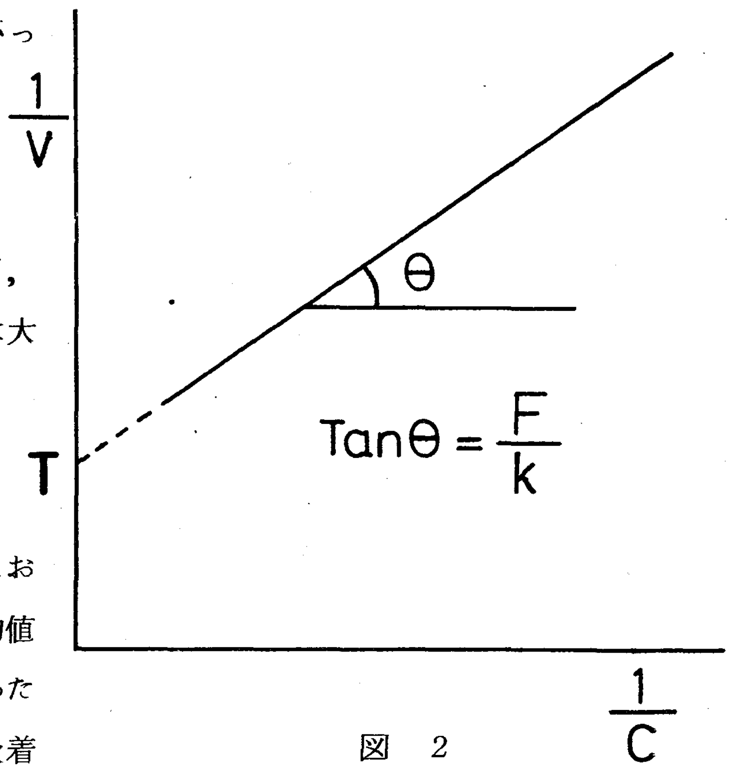


図　2

3．媒液容量

合成酵素の基質特異的な吸着部位における単位時間あたりの吸着回数の平均値が媒液容量である。それを求めるにあたり，まず時間〔0，t〕に加算される吸着

回数の総数Nt を求める。この問題は型 I の計数器に関する更新過程と数学的に同じ内容のものである（Cox，1962）。

基質の衝突過程は単位時間の平均衝突回数 ρ をパラメーターとするポアソン過程であると仮定する（図3）。

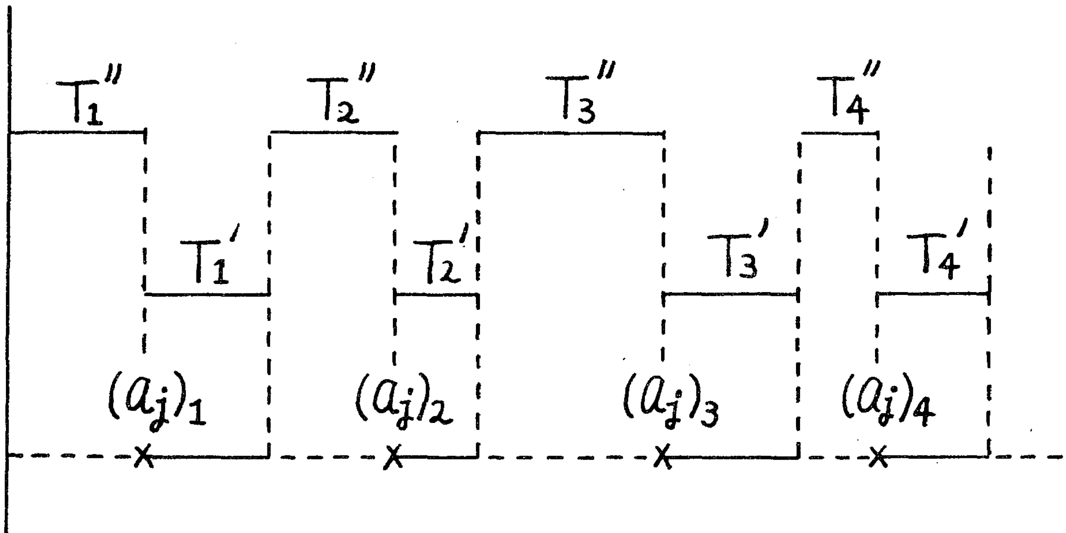


図　3

はじめに吸着部位は空席であるが，T_1'' 後に基質 $(a_j)_1$ が衝突したとする。基質 $(a_j)_1$ は基質特異的に吸着し，時間 T_1' の間，部位を占有する。このとき，別の基質が衝突しても，吸着できないので反跳する。$(a_j)_1$ が脱着すると部位は空席にもどる。T_2'' 後に次の基質 $(a_j)_2$ が衝突すると，それは吸着する。このような繰り返しによって，空席状態の時間間隔の系列 $\{T_1'', T_2'', \cdots\}$ がえられる。これは，パラメーター ρ で指数的にたがいに独立に分布する。

$$f(T'') = \rho e^{-\rho T''}$$

また，吸着による占有状態の時間間隔の系列 $\{T_1', T_2', \cdots\}$ がえられる。この系列は確率密度関数（p.d.f.）b(t)を有し，たがいに独立に分布すると仮定する。系列 $\{T_1'', T_2'', \cdots\}$ と $\{T_1', T_2', \cdots\}$ とはたがいに独立であると仮定する。吸着する基質符号の系列は $\{(a_j)_1, (a_j)_2, \cdots\}$ である。ここに，j は種類を表わす系列 1, …, n の中の数である。相続く吸着（吸着の終止点をとる）の時間間隔，すなわち

$$T_1 = T_1'' + T_2'',\quad T_2 = T_2'' + T_1'',\ \cdots\cdots$$

によって系列 $\{T_1, T_2, \cdots\}$ がつくられる。これは p.d.f. として $b(t)$ とパラメーター ρ の指数分布との重ね合わせ(convolution)をもつ，通常の更新過程である。μ と σ^2 を占有時間の平均値と分散としよう。そうすれば，吸着の時間間隔の平均値と分散は，$\mu + 1/\rho$ および $\sigma^2 + 1/\rho^2$ である。時間 t における吸着回数 N_t は，漸近的に正規分布になり，次の平均値と分散をそれぞれもつ。

$$\overline{N}_t = \rho t / (1 + \mu\rho)$$

と，

$$(1 + \sigma^2 \rho^2)\rho t / (1 + \mu\rho)^3$$

(証明は補遺1および Feller，1950 を参照)。単位時間($t=1$)における吸着回数の平均値が媒液容量であるから，

$$V = \frac{1}{\mu + 1/\rho}$$

がえられる。

次に別途として，吸着平衡の現象論によって媒液容量を求めよう。

$$c + (a - x) \underset{k_2}{\overset{k_1}{\rightleftarrows}} x$$

吸着部位の全数 a のうち x に基質が吸着しているとする。基質の濃度 c に比例して基質の吸着部位に対する衝突が起るから，その速度定数を k_1 とする。到着速度は k_1c である。到着したとき吸着部位の空席状態の数($a-x$)あるいは占有状態の数 x に比例して吸着あるいは反跳がおこる。ここに吸着速度は $k_1c(a-x)$ である。占有状態の数 x に比例して基質の脱着がおこる。その速度定数を k_2 とすれば，脱着速度は k_2x である。平衡状態では吸着速度と脱着速度が等しいから，

$$k_1 c(a - x) = k_2 x$$

これより，

$$x = \frac{a}{1 + k_2/k_1 c}$$

塚本吉彦

平衡状態における吸着（または脱着）速度は，

$$\frac{a}{1/k_2 + 1/k_1 c}$$

一個の部位における平衡状態での吸着速度が媒液容量であるから，

$$V = \frac{1}{1/k_2 + 1/k_1 c}$$

のように前述と同様の式がえられる。到着速度に関しては $k_1 c = \rho$，脱着の速度定数 k_2 と占有時間 μ の間に，$k_2 = 1/\mu$ の関係がある。一例として，占有時間 $\mu = 0.1$ (sec)，濃度 ρ（任意の単位）のときの媒液容量を図 4 に示す。

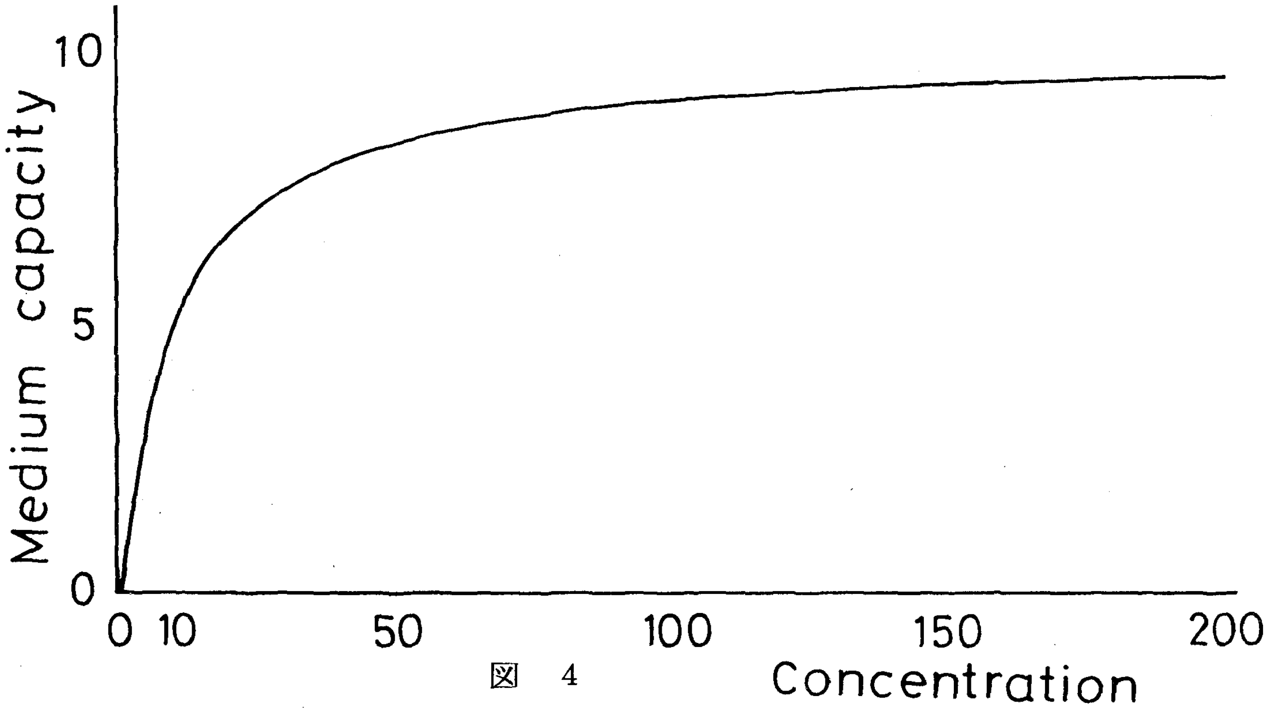


図　4

4．鋳型と基質の符号系の構成

鋳型符号系 $X = \{x_1, \cdots, x_n\}$ と基質符号系 $A = \{a_1, \cdots, a_n\}$ とを以下のように構成する。

鋳型は L 個の鋳型符号 x_j を順に配列した鎖である。

$$(x_j)_1 \cdot (x_j)_2 \cdot \cdots \cdot (x_j)_i \cdot \cdots \cdot (x_j)_L$$

ここに $i = 1, \cdots, L$，および j は種類を表わす系列 $1, \cdots, n$ の中の数である。第 j 種の鋳型符号 x_j が含まれる個数を L_j とする。x_j の出現確率 p_j は相対頻度で与えられる。

$$p_j = L_j / L \quad (j = 1, \cdots, n)$$

ここに，$\sum_{j=1}^{n} p_j = 1$ である。

基質符号 a_j（$j = 1, \cdots, n$）は合成部位に対して吸着をくりかえし，時間的な系列をつくる。

$(a_j)_1$，$(a_j)_2$，…，$(a_j)_i$，… ここに，$i = 1, 2, \cdots$。基質符号の系列 $\{(a_j)_i\}$ は確率的に相互に独立であると仮定する。媒液における全基質の濃度を〔S〕，それに含まれる第 j 種の基質の濃度を〔S_j〕とする。第 j 種の基質符号 a_j の出現確率 q_j は第 j 種の基質の相対濃度によって与えられる。

$$q_j = [S_j]/[S] \qquad (j = 1, \cdots, n)$$

ここに $\sum_{j=i}^{n} q_j = 1$ である。この条件がなりたつとき，この媒液を完全系と呼ぶ。第 j 種のそれぞれの符号 x_j と a_j は相補的に対応する。基質特異的な吸着は行なうけれども，鋳型符号に対応する基質符号をもたない基質類似物が存在する場合に，それらを第 n ＋ 1 種としてまとめると，

$$\sum_{j=1}^{n+1} q_j = 1 \qquad \text{したがって} \qquad \sum_{j=1}^{n} q_j < 1$$

である。このような，いわば“冗長な基質”を含む媒液を不完全系と呼ぶ。

5．選択過程と不確定性の関数

基質符号の時間的な系列において，第 j 種の基質符号が m 回目（$m \geqq 1$）にはじめて出現する確率は $q_j(1-q_j)^{m-1}$ である。この分布の母関数は，

$$\begin{aligned}\Phi(s) &= \sum_{m=1}^{\infty} q_j(1-q_j)^{m-1} s^m \\ &= q_j^2 / (1-q_j)\{(q_j-1)s+1\}。\end{aligned}$$

第 j 種の基質符号が一回出現するまでに必要な，種類を問わない基質符号の平均個数は，

$$E(x_j) = \Phi'(1) = 1/q_j$$

その分散は，

$$V_{ar}(x_j) = \Phi''(1) + \Phi'(1) - \Phi'(1)^2$$

$$= (1 - q_j) / q_j^{\,2} \quad 。$$

第 j 種の鋳型符号に対して相補的な第 j 種の基質符号の系列を第 j 種の基質符号語と定義する。基質符号語の長さは，それを構成する基質符号の個数で計られ，幾何分布をなし，平均値を有する。第 j 種の基質符号語の平均長は $1/q_j$ である。

与えられた鋳型によって使用される基質符号語の平均長を，それぞれの種類の使用頻度にしたがって平均すれば，

$$F(q_1, \cdots, q_n) = \sum_{j=1}^{n} (q_j / q_j)$$

$$= p_1 / q_1 + \cdots + p_n / q_n$$

$$\left(\sum_{j=1}^{n} p_j = 1 , \quad \sum_{j=1}^{n} q_j \leqq 1 \right)$$

これは一つの鋳型符号あたりに必要な基質符号語の平均長である。数理計画法において，F は目標変数，p_j は計画係数，q_j は計画変数，$\sum_j p_j = 1$ と $\sum_j q_j \leqq 1$ はそれぞれ係数と変数に関する条件である。F は基質符号の出現確率 q_j ($j = 1, \cdots, n$) の関数である。完全系 ($\sum_j q_j = 1$) の場合に，F 関数は，

$$q_j^{*} = \sqrt{p_j / G} \qquad (j = 1, \cdots, n)$$

のとき，最小値

$$G(p_1, \cdots, p_n) = \left(\sum_{j=1}^{n} \sqrt{p_j} \right)^2$$

をとる (証明は補遺 2 を参照)。すなわち，最適媒液の基質符号語の平均長は，

$$1 / q_j^{*} = \sqrt{G / p_j} \qquad (j = 1, \cdots, n)$$

である。この結果は，確率の大きな鋳型符号の種類については，対応する基質符号語の平均長を比較的短かくし，確率の小さな鋳型符号の種類については，対応する基質符号語の平均長を比較的長くするように，媒液中の基質符号の相対濃度を与えることによって，F の最小値が実現されることを示す。

この最小値Gは，通信の情報理論の場合と同様に，鋳型の不確定性あるいはそれの担う情報量を表わすと考えられる。鋳型に貯えられた情報は合成反応を通じて読み出されるものである。Gは各種の鋳型符号の出現確率 p_j（$j=1, \cdots, n$）の関数である。この関数によって鋳型に貯えられた平均情報量が定義される。その単位を”コット”（cot：collision unit）と名付ける。

$$G \leqq F \quad \text{すなわち} \quad G/F \leqq 1$$

鋳型の読み出しに必要な一つの鋳型符号あたりの基質符号語の平均長は，必らず鋳型の平均情報量より大きい。等号は媒液が最適のときなりたつ。したがって，一個の基質符号は最大にして1コット，一般にはG／Fコットの情報量を移送する。

$X=\{x_j \mid j=1, \cdots, n\}$ の要素 x_j に付与された確率 p_j（$j=1, \cdots, n$）の関数として定義される，鋳型の不確定性あるいは平均情報量は，次の三つの表わし方をとりうる。

$$G(X) = \left(\sum_{j=1}^{n} \sqrt{p_j}\right)^2 = \sum_{j=1}^{n} p_j \cdot g_j(X)$$

$$= \frac{1}{\sum_{j=1}^{n} \{1/g_j(X)^2\}}$$

ここで，

$$g_j(X) = \sum_{i=1}^{n} \sqrt{p_i}/\sqrt{p_j}$$

は，第 j 種の鋳型符号の担う情報量というべきものである。いま，最適媒液の第 j 種の基質符号の出現確率を q_j^*（$j=1, \cdots, n$）とすれば，

$$g_j(X) = 1/q_j^*$$

である。したがって，最適媒液のとき，基質符号語はその平均長に等しい $g_j(X)$ コットの情報量を移送する。この関係を上式に代入して，

$$G(X) = 1/\sum_{j=1}^{n} g_j^{*2}$$

がえられる。Fの最小値Gは，最適媒液の基質符号の出現確率の関数としても与えられる。したがって，Gは最適媒液の不確定性をも表わすように見える。しかし，この関係式は，媒液の不確定性が媒液における基質符号の出現確率の関数として独自に定義され，その特殊な場合において，媒液の不確定性と鋳型の不確定性が一致することを示すと解釈される。

一般に，媒液の不確定性Mは，基質符号の出現確率 q_j ($j=1, \cdots, n$) の関数として，同じ表式で定義されるものとする。

$$M(q_1, \cdots, q_n) = 1 \Big/ \sum_{j=1}^{n} q_j^2$$

$A=\{a_j \mid j=1, \cdots, n\}$ の要素 a_j に付与された確率 q_j ($j=1, \cdots, n$) の関数として定義される媒液の不確定性は，次の三つの表わし方をとりうる。

$$M(A) = 1 \Big/ \sum_{j=1}^{n} q_j^2 = \sum_{j=1}^{n} \left(1/\mu_j(A)\, q_j\right)$$

$$= \left\{ \sum_{j=1}^{n} \left(1/\sqrt{\mu_j(A)}\right)\right\}^2$$

ここで，

$$\mu_j(A) = \sum_{i=1}^{n} q_i^2 / q_j^2$$

である。最適媒液のときには，第 j 種の基質符号語が使用される確率について，

$$p_j = 1/\mu_j(A)$$

がなりたつ。

6. 媒液の不確定性の数え方

媒液の不確定性は次のような仕方で数えられる。熱的平衡状態にある媒液の中で，基質が特定の部位に対して衝突をくりかえし，基質符号の時間的な系列をつくるとする。基質符号を２個ずつ重ならないように区切り，長さ２の連をつくる。

$|(a_j)_1,\ (a_j)_2|(a_j)_3,\ (a_j)_4|$ ……

$$|(a_j)_i,\ (a_j)_{i+1}| \ \cdots\cdots$$

ここで, $i=1, 2, \cdots$, および j は種類を表わす系列 $1, 2, \cdots, n$ の中の数である。連が同種の基質符号から構成されるときにホモ（homo）の連と呼び，異種の基質符号から構成されるときにヘテロ（hetero）の連と呼ぶ。第 j 種のホモの連の出現確率は q_j^2 である。種類を問わないホモの連の出現確率は $\sum_j q_j^2$ である。もとの連の系列にたいして，ホモの連を選択する過程を考える。$M(A) = 1/\sum_j q_j^2$ は，ホモの連が1個出現するまでの連の平均個数である。そのような選択によってホモの連の系列がえられる。$\mu_j(A) = \sum_i q_i^2/q_j^2$ は第 j 種のホモの連が1個出現するまでのホモの連の平均個数である。そのような選択によって第 j 種のホモの連のみがえられる。

7. 関数M, G, Fの特徴

（i）n を固定したとき，条件 $\sum_{j=1}^{n} q_j=1$ のもとで，$M(q_1, \cdots, q_n)$ は $q_j = 1/n$（$j=1, \cdots, n$）のとき最大値 n をとる（証明は補遺4）。

$$M(q_1, \cdots, q_2) \leqq M(1/n, \cdots, 1/n) = n$$

図5に，2元符号系におけるMの値が基質符号の確率の関数として示されている。$q=1/2 \pm 1/2\sqrt{3}$ で変曲点をもつ。両端で最小値として単位元をとる。等確率すなわち中央で最大値2をとる。

（ii）n を固定したとき，条件 $\sum_{j=1}^{n} p_j=1$ のもとで，$G(p_1, \cdots, p_n)$ は $p_j = 1/n$（$j=1, \cdots, n$）のとき最大値 n をとる（証明は補遺4）。

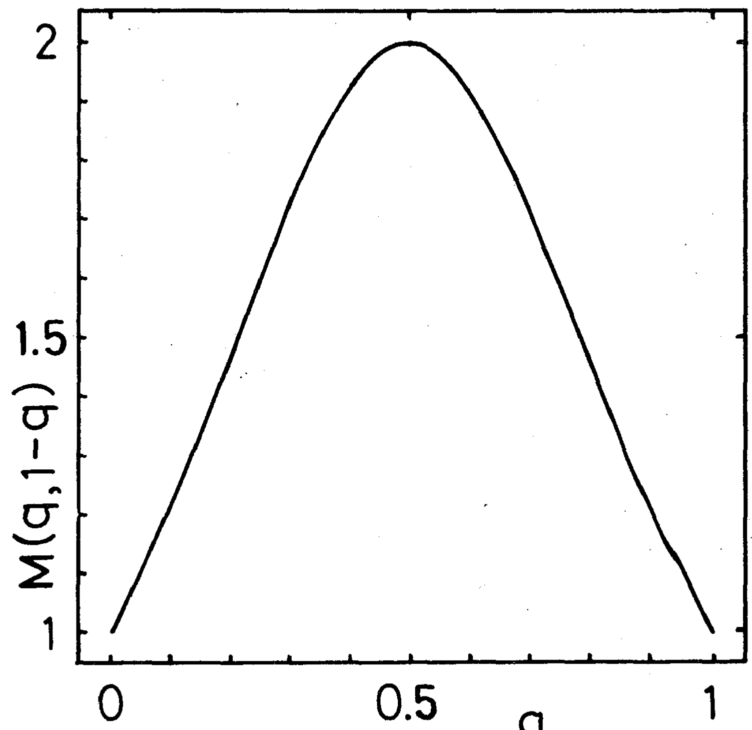


図 5

$$G(p_1,\ \cdots,\ p_n) \leqq G(1/n,\ \cdots,\ 1/n) = n$$

図６に，２元符号系におけるGの値が鋳型符号の確率の関数として示されている。両端で最小値として単位元をとる。等確率すなわち中央で最大値２をとる。

（iii）　$F(q_1,\ \cdots,\ q_n)$ は，$q_j = 1/n\ (j=1,\ \cdots,\ n)$ のとき

$$F = n$$

また，$q_j = p_j\ (j=1,\ \cdots,\ n)$ のとき

$$F = n$$

をとる。図７に，２元符号系における $F = p/q + (1-p)/(1-q)$ の値が q の関数として示されている。（p，1−p）の値は，1.（0.1，0.9）；2.（0.3，0.7）；3.（0.5，0.5）；4.（0.7，0.3）；5.（0.9，0.1）である。中央付近では q が変化しても，Fはそれほど変化しない。$q = 1-q = 1/2$ および $q = p$，$q = 1-p$ のとき $F = 2$ である。最小値は両点の中間にあって２より小さい。

８．媒液の不確定性と異種分子の混合過程

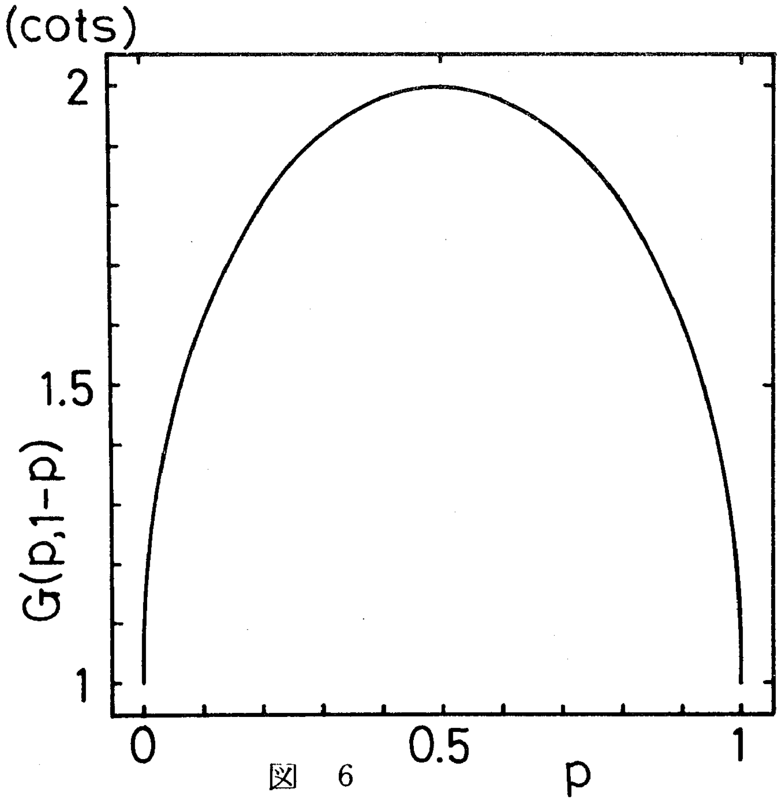


図　6

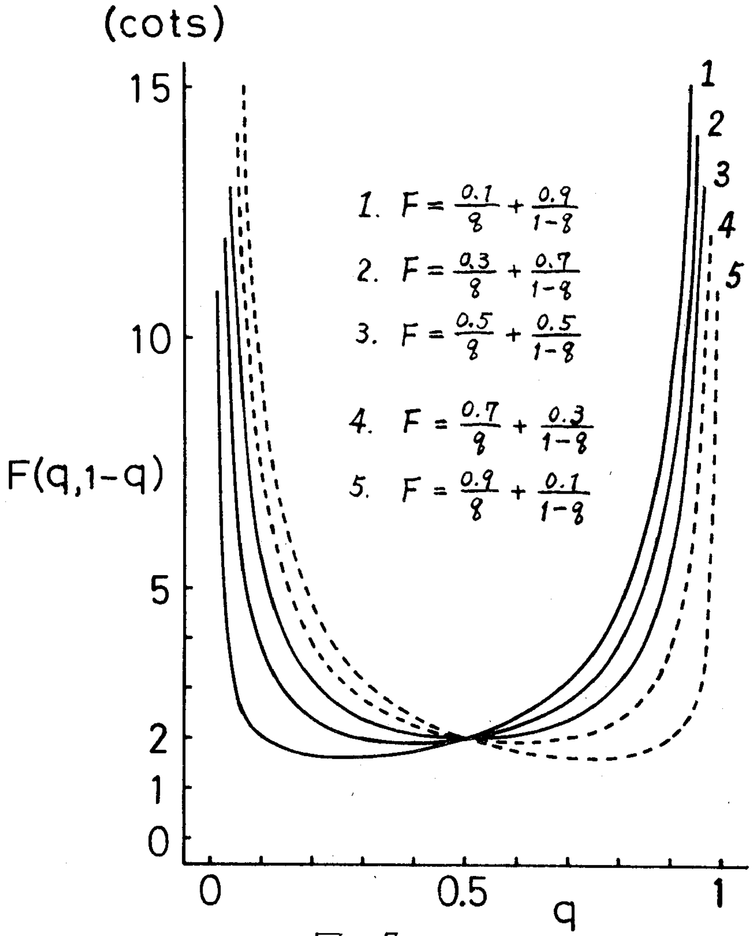


図　7

媒液は溶媒の中に各種の基質が溶け込んだ混合溶液である。各基質は細胞の各所にそれぞれの合成反応系によって生成され，溶媒の中を拡散し，たがいに混合する。この混合過程は自発的であり，熱力学的なエントロピー（不確定性）は増大する。このような媒液の生成過程に媒液の不確定性Mはどのように関係するだろうか。この点を明らかにするためには，溶液における溶質分子の混合を運動学的に調べる必要がある。実際には計算の簡単のために，異種の気体分子の混合過程を量的に検討する。気体の場合には気体分子は真空中を運動するから，溶媒の影響を考慮しなくともよい。われわれが問題にする点について，以下で導かれる結果は溶液の場合にも同様になりたつことがのちに明らかになるだろう。

熱力学にしたがって，2種類の気体分子の混合過程を考える（久保ら，1961；中村1967）。隔壁によって分けられた二つの部屋1，2があるとする（図8参照）。

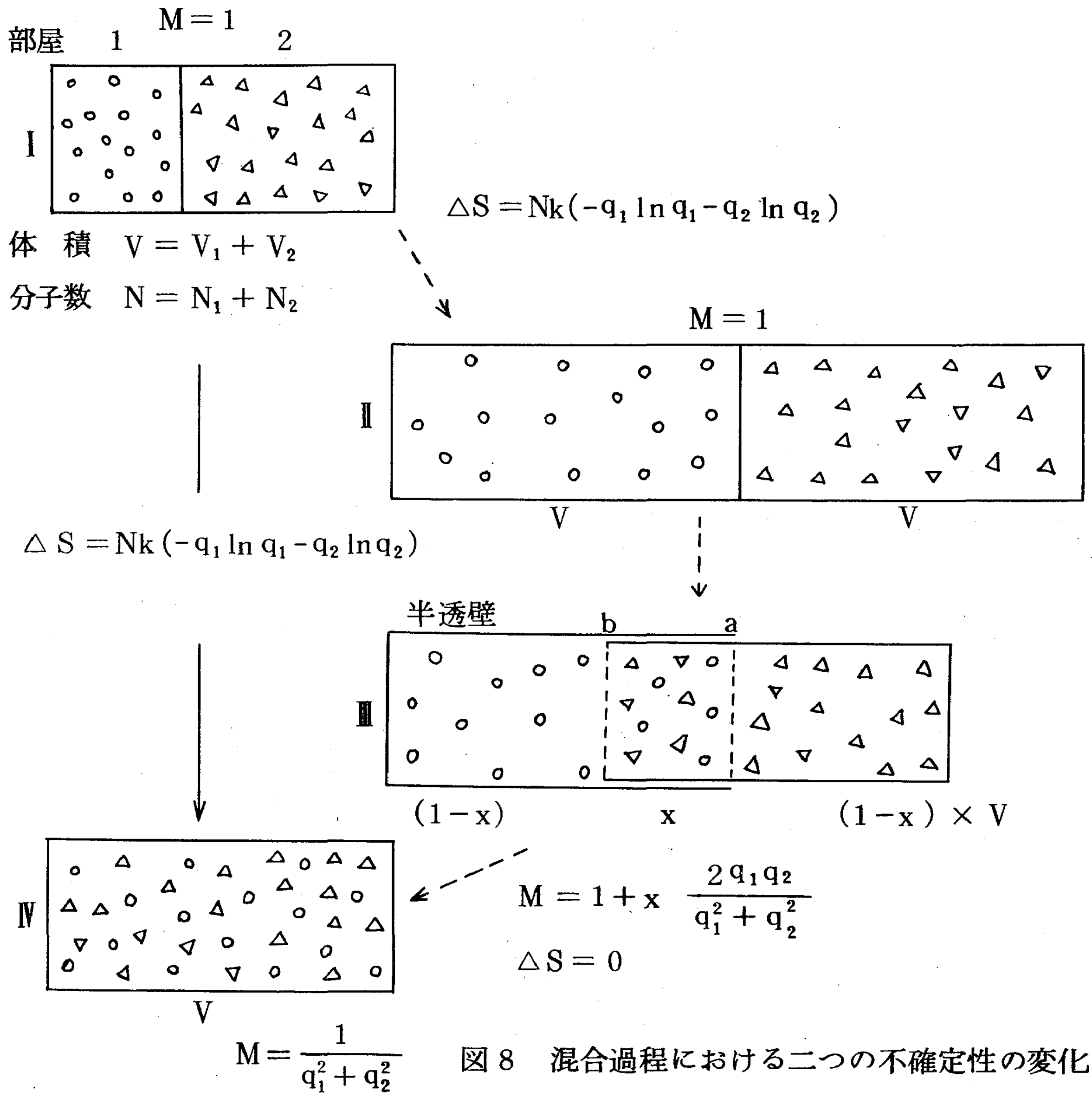


図８　混合過程における二つの不確定性の変化

塚本吉彦

それぞれの部屋に温度 T も圧力 P もたがいに等しい 2 種類の分子 1，2 が入っている。部屋の体積をそれぞれ V_1，V_2 また分子の数を N_1，N_2 とする。全体積は $V=V_1+V_2$，全分子数は $N=N_1+N_2$ である。分子密度が小さいとき，気体の状態は理想気体の状態方程式で近似的に表わされる。

$$PV_1 = N_1 kT, \qquad PV_2 = N_2 kT$$

ここで，k はボルツマン定数である。状態方程式から

$$V_1 : V_2 = N_1 : N_2$$

の関係がある。隔壁を除くと気体は混ざり合う。状態 I が状態 II へ変化する際のエントロピー変化は，

$$\Delta S = -N_1 k \ell n (V_1/V) - N_2 k \ell n (V_2/V)$$

である。この過程は思考実験によって状態 II，III を通るものとして考えることができる。

（i） I → II の変化： 隔壁のあるまま反対側の壁を拡げて，部屋 1 と 2 の体積を別々に V_1 から V に，V_2 から V に膨張させる。それぞれのエントロピー変化は，

$$\Delta S_1 = - N_1 k \ell n (V_1/V)$$

$$\Delta S_2 = - N_2 k \ell n (V_1/V)$$

である。系全体のエントロピー変化は

$$\Delta S = \Delta S_1 + \Delta S_2$$

すなわち，混合の熱力学的なエントロピー増加は，分子の各成分が同一の体積 V をそれぞれ単独に占めたときのエントロピー増加の和に等しい。気体 1，2 の相対濃度 q_1，q_2

$$q_1 = N_1/N, \qquad q_2 = N_2/N$$

を使って，書き直せば，

$$\Delta S = Nk(-q_1 \ell n\, q_1 - q_2 \ell n\, q_2)$$

である。

(ii) II → III → IV の変化：部屋1と2を，隔壁の代わりにそれぞれ半透壁 a，b をもつ，たがいに入れ子になった二つのシリンダーでしきる。半透壁 a は分子1を通さないが，分子2を通す。半透壁 b は分子2を通さないが，分子1を通す。シリンダー1，2の中には圧力がそれぞれ P_1，P_2 の気体が入っている。シリンダー1を固定しておいて，シリンダー2をさし込んでいく。シリンダー1と2の共通部分には，気体1と2が混在する。この過程で，半透膜にはそれを通りえない気体の分圧が実際の圧力として働く。したがって気体2は，部屋2の半透壁 b には右側から，その反対側の壁には左側から，同一の圧力 P_2 をそれぞれ及ぼす。変位するシリンダーの受ける正味の力はゼロである。この混合過程は外部から熱も仕事も加えられることなく進行する。半透壁 b がシリンダー1の壁に到着したときに気体分子1，2は完全に混合される。状態 II と IV の間の変化は可逆であり，熱力学的なエントロピー変化はない。熱力学的なエントロピーは状態 I から状態 II への変化を表わすことができるけれども，状態 II から状態 IV への変化を表わすことができない。熱力学的な混合のエントロピーが，異種分子がたがいに混ざらないとき（I → II）に変化し，それらがたがいに混ざり合うとき（II → IV）に変化しないことは一つのパラドックスである。混合された無秩序な状態から分離された秩序のある状態への逆の過程（IV → II）では，何らかの仕事のなされていることが現象的に明白である。このような仕事は熱力学的に把えられ得ないのである。

われわれは，上で仮定した二つの半透壁 a，b をたんに分子の大きさの相違を利用して作ることは不可能であることに注意する必要がある。たとえ孔の小さい方の半透壁が一種類の分子しか通さなくとも，孔の大きい方の半透壁が二種類の分子をともに通すからである。われわれは，Maxwell の魔物（分子の“速度”を測定する弁の作用）や，Szilard の模型（分子の“位置”を検出するしきりの作用）と同様の問題に遭遇している（Brillouin，L.，1956 参照）。半透壁の孔は分子の“種類”を判別する弁の作用を備える必要がある。先の例で言えば，孔の大きな方の半透壁は大きさの小さい方の分子が向かって来たとき，それを判別して孔を閉じればよいだろう。もしも分子の質量や大きさが等しいならば，分子の種類は表面の光学的性質や立体特異性などの相違を利用して判別されることもあるだろう。分子の種類に関する情報を得るために必要な最小限熱雑音に打ち克つエネルギー移動によって，一定の熱力学的なエントロピーの増大が不可避であることは，Maxwell の魔物と同様の事情である。しかし，その議論は半透壁 a，b の仮定の根拠に関するものである。したがって，それは半透壁が仮定された

上での上述の混合過程の状態変化の議論には直接関係しない。

前述の２種類の分子１と２が混合した状態Ⅳを考える。２種類の分子は質量（μ），直径（σ）がたがいに等しい球状の粒子とする。両者は分子表面の光学的性質や立体特異性等によって区別されるものとする。熱的平衡にある気体において，非常に多数の分子が熱運動によってたがいにランダムに衝突をくりかえしているとする。単位時間に行なわれるすべての衝突をたがいに衝突し合う２個の分子の種類によって類別する。この場合の衝突は既述した２個の基質符号から構成される一つの連に相当する。衝突する２個の分子が同種のときにホモと呼び，異種のときにヘテロと呼ぶ。

単位体積について単位時間の分子間の衝突数 z は，気体分子運動論より，

$$z = K(N/V)^2 \qquad \text{ただし} \qquad K = \sigma^2 [4\pi k T/\mu]^{1/2}$$

全体積 V についての衝突数は単位体積についてのものに体積 V を乗ずれば得られる。

$$Z = KN^2/V$$

単位体積について，分子１および分子２のホモの衝突数 z_{11}，z_{22} はそれぞれ，

$$z_{11} = K(N_1/V)^2 \ , \qquad z_{22} = K(N_2/V)^2$$

また，分子１と分子２のヘテロの衝突数 z_{12} は，

$$z_{12} = 2K(N_1/V)(N_2/V)$$

衝突数 z は各種のホモの衝突数 z_{11}, z_{22} とヘテロの衝突数 z_{12} の和である。

$$z = z_{11} + z_{22} + z_{12}$$

衝突数 z のホモの衝突数 $z_{11} + z_{22}$ に対する倍率を計算すると，

$$z/(z_{11} + z_{22}) = N^2/(N_1^2 + N_2^2) = 1/(q_1^2 + q_2^2)$$
$$= M(q_1, q_2)$$

である。媒液の不確定性は，全衝突数のホモの衝突数に対する倍率を表わす。

部屋１と２の全体を一つの系として，その状態を媒液の不確定性によって調べる。状

態Ⅰでは分子密度に関して,

$$N_1/V_1 = N_2/V_2 = N/V$$

より,単位体積および全体積について単位時間の分子間の衝突数は状態Ⅳに等しい。すべての衝突はホモの衝突であるから,

$$M = 1$$

である。状態Ⅱでは分子密度が減少するから,単位体積について単位時間の分子間の衝突数は減少する。部屋1と2のそれぞれについて,

$$z(1) = K(N_1/V)^2 \quad , \quad z(2) = K(N_2/V)^2$$

全体積については

$$Z = K(N_1^2 + N_2^2)/V$$

しかし,すべての衝突はホモの衝突であるから,

$$M = 1$$

である。状態Ⅲでは,シリンダー1と2の共通部分において,ホモの衝突にヘテロの衝突が新たに加わる。単位体積について単位時間の衝突数は,シリンダー1,2のそれぞれの単独部分に関しては状態Ⅱの場合と,シリンダーの共通部分に関しては状態Ⅳの場合と同じである。共通部分の比率を x($0 \leqq x \leqq 1$)とする。全体積についての単位時間の全衝突数は,

$$Z = K\{(1-x)(N_1^2 + N_2^2)/V + x\,N^2/V\}$$

である。全体積についての全衝突数のホモの衝突数に対する倍率は,

$$M = 1 + x \cdot 2q_1 q_2/(q_1^2 + q_2^2) \qquad (0 \leqq x \leqq 1)$$

である。ここで,$x=0$ は状態Ⅱに,$x=1$ は状態Ⅳに相当している。状態Ⅱから状態Ⅳへの変化は,媒液の不確定性Mによって表わされる。以上のように,媒液の不確定性が全衝突数のホモの衝突数にたいする倍率を意味することは,稀薄気体の場合に運動学

的に証明された。衝突数を決める因子のうち分子密度だけがこの証明に直接関係する。したがって，稀薄溶液における溶質分子間の衝突過程にも同様の結果があてはまることは明らかである。

上述した思考実験によって混合過程に含まれる二つの不確定性の変化を分離して計算できた。実際の混合過程では二つの不確定性はともに変化する。媒液の不確定性の変化は，現実の異種分子の混合過程に含まれる，熱力学的なエントロピーだけでは表わすことのできない，情報的な側面の量的変化を表わす。

一般にｎ種類の分子の場合にも，上述の結論が２種類の分子の場合と同様に証明される。媒液の不確定性は細胞の化学反応系にとって，どんな意義をもつのだろうか。媒液は基質特異的な反応に必要な基質分子を供給する。分子の種類の数ｎが大きくなるにともなって，不確定性Ｍの値は増大する。それは媒液に溶ける基質の種類が多いほど，基質を供給することのできる反応の種類も多くなりうること，いわば種類の容量が大きくなることを意味する。また，各種の分子の濃度が一様化するにともなって，不確定性Ｍの値も増大する。ｎ種の基質の濃度がたがいに等しいとき，媒液の不確定性Ｍは種類の数ｎに等しい。それは，各種の基質の濃度がたがいに等しいとき，それぞれの基質は等頻度で供給されることに対応する。もしも濃度が種類によって偏っているとき，Ｍの値は種類の数より小さくなる。例えば，ℓ（$\ell < n$）種の基質がほとんど供給されないほど極端に低濃度である場合を考える。そのとき，他の（$n-\ell$）種の基質がたがいに等濃度であれば，Ｍは（$n-\ell$）に近ずくであろう。媒液の不確定性は，実際の反応過程で供給される基質の実効的な種類の数を表わす。それは，各種の基質の選択的な供給という事象に関して，どれだけ多くの種類が実効的に可能かという“事象あたりの媒液容量”であると言える。

９．ダブレット符号系の構成と性質

二つのシングレット符号系が結合したダブレット符号系，すなわち鋳型符号系 $XY = x_1y_1, \cdots, x_jy_k, \cdots, x_ny_m\}$ と基質符号系 $AB = \{a_1b_1, \cdots, a_jb_k, \cdots, a_nb_m\}$ を考える。第 jk 種のそれぞれの符号 x_jy_k と a_jb_k は相補的に対応する。

鋳型はＬ個の鋳型符号 x_jy_k（$j = 1, \cdots, n$；$k = 1, \cdots, m$）が順次に配列した鎖である。

$$(x_j y_k)_1 \cdot (x_j y_k)_2 \cdot \cdots \cdot (x_j y_k)_i \cdot \cdots \cdot (x_j y_k)_L$$

第 j k 種の鋳型符号 $x_j y_k$ の含まれる個数を L_{jk} とする。 $x_j y_k$ の出現確率 π_{jk} は相対頻度で与えられる。

$$\pi_{jk} = L_{jk} / L \quad (j = 1, \cdots, n\ ;\ k = 1, \cdots, m)$$

ここに, $\sum_j \sum_k \pi_{jk} = 1$ 。ダブレットの鋳型符号を構成する符号系 X で符号 x_j が出現する確率 p_j は,

$$p_j = \sum_{k=1}^{m} \pi_{jk} \qquad (j = 1, \cdots, n)$$

符号系 Y で符号 y_k が出現する確率 r_k は,

$$r_k = \sum_{j=1}^{n} \pi_{jk} \qquad (k = 1, \cdots, m)$$

で与えられる。ここに, $\sum_j p_j = 1$, $\sum_k r_k = 1$ 。 鋳型符号を構成する上で, 二つの符号系 X と Y とが確率的に独立である場合には,

$$\pi_{jk} = p_j r_k$$

である。次に, 符号系 X で符号 x_j が出現した条件のもとで, 符号系 Y で符号 y_k が出現する条件つき確率を r_{jk} とすれば,

$$\pi_{jk} = p_j r_{jk}$$

の関係がある。ここに, $\sum_k r_{jk} = 1$ 。

基質符号 $a_j b_k$ ($j = 1, \cdots, n\ ;\ k = 1, \cdots, m$) は, 特定の部位に吸着をくりかえして時間的な系列をつくる。

$$(a_j b_k)_1, \ (a_j b_k)_2, \ \cdots, \ (a_j b_k)_i, \ \cdots$$

基質符号の系列 $\{(a_j b_k)_i\}$ は確率的に相互に独立であると仮定する。 媒液における全基質の濃度を〔S〕, それに含まれる第 j k 種の基質の濃度を〔S_{jk}〕とする。第 j k 種の基質符号 $a_j b_k$ の出現確率 ρ_{jk} は第 j k 種の基質の相対濃度で与えられる。

$$\rho_{jk} = [S_{jk}] / [S] \qquad (j=1, \cdots, n \,;\, k=1, \cdots, m)$$

ここに，$\sum_j \sum_k \rho_{jk} = 1$。基質符号を構成している符号系Aで符号 a_j が出現する確率 q_j は，

$$q_j = \sum_{k=1}^{m} \rho_{jk} \qquad (j=1, \cdots, n)$$

符号系Bで符号 b_k が出現する確率 d_k は，

$$d_k = \sum_{j=1}^{n} \rho_{jk} \qquad (k=1, \cdots, m)$$

で与えられる。ここに，$\sum_j q_j = 1$，$\sum_k d_k = 1$。 基質符号を構成する上で，二つの符号系AとBとが確率的に独立である場合には，

$$\rho_{jk} = q_j d_k$$

である。次に，符号系Aで符号 a_j が出現した条件のもとで，符号系Bで符号 b_k が出現する条件つき確率を d_{jk} とすれば，

$$\rho_{jk} = q_j d_{jk}$$

の関係がある。ここに，$\sum_k d_{jk} = 1$。

鋳型符号系XYの不確定性は，

$$G(XY) = \left(\sum_{j=1}^{m} \sum_{k=1}^{m} \sqrt{\pi_{jk}} \right)^2$$

$$= \frac{1}{\sum_{j=1}^{n} \sum_{k=1}^{m} \{ 1 / g_{jk}(XY)^2 \}}$$

と表わされる。ここに，

$$g_{jk}(XY) = \sum_{i=1}^{n} \sum_{h=1}^{m} \sqrt{\pi_{ih}} / \sqrt{\pi_{jk}}$$

「鋳型符号系の不確定性を表わす関数Gは，二つの鋳型符号系XとYとを結合して，

新しい鋳型符号系XY をつくるとき，乗法性を有する。」

（ⅰ） X と Y とが独立のとき，

$$G(XY) = G(X)G(Y)$$

$$G(X) = \left(\sum_{j=1}^{n} \sqrt{p_j}\right)^2, \qquad G(Y) = \left(\sum_{k=1}^{m} \sqrt{r_k}\right)^2$$

（ⅱ） X と Y とが関連しているとき，

$$G(XY) = G(X|Y)\,G_X(Y)$$

第1符号系Xに関する不確定性 $G(X|Y)$ は，第2符号系Yによって遡及的に影響される。

$$G(X|Y) = \frac{1}{\sum_{j=1}^{n}\{1/g_j(X|Y)^2\}}$$

$$g_j(X|Y) = \sum_{i=1}^{n}\sum_{h=1}^{m}\sqrt{\pi_{ih}} \Big/ \sum_{k=1}^{m}\sqrt{\pi_{jk}}$$

$$= \sum_{i=1}^{n}\sqrt{p_i}\,G_i(Y) \Big/ \sqrt{p_j}\,G_j(Y)$$

第1符号系Xが出現した条件のもとでの第2符号系Yに関する条件つき不確定性 $G_X(Y)$ は，

$$G_X(Y) = \sum_{j=1}^{n} p_j\,G_j(Y), \qquad G_j(Y) = \left(\sum_{k=1}^{m}\sqrt{r_{jk}}\right)^2$$

$G_j(Y)$ は第1符号系Xで第j種の符号 x_j が出現した条件のもとでの第2符号系 Y に関する条件つき不確定性である。

$$G_j(Y) = \frac{1}{\sum_{k=1}^{m}\{1/g_{jk}(Y)^2\}} = \sum_{k=1}^{m} r_{jk}\,g_{jk}(Y)$$

$$g_{jk}(Y) = \sum_{h=1}^{m}\sqrt{r_{jh}} \Big/ \sqrt{r_{jk}}$$

基質符号系ＡＢの不確定性は，

$$M(AB) = 1 / \sum_{j=1}^{n} \sum_{k=1}^{m} \rho_{jk}^2$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} \{ 1/\mu_{jk}(AB)\, \rho_{jk} \}$$

と表わされる。ここに，

$$\mu_{jk}(AB) = \sum_{i=1}^{n} \sum_{h=1}^{m} \rho_{ih}^2 / \rho_{jk}^2$$

「基質符号系ABの不確定性を表わす関数Ｍは，二つの基質符号系ＡとＢとを結合して，新しい基質符号系ＡＢをつくるとき，乗法性を有する。」

（i）ＡとＢとが独立のとき

$$M(AB) = M(A)\, M(B)$$

$$M(A) = 1/ \sum_{j=1}^{n} q_j^2, \quad M(B) = 1/ \sum_{k=1}^{m} d_k^2$$

（ii）ＡとＢとが関連しているとき

$$M(AB) = M(A)\, M_A(B)$$

第１符号系Ａに関する不確定性M(A)は，独立の場合と同じである。

$$M(A) = 1/ \sum_{j=1}^{n} q_j^2$$

第１符号系Ａが出現した条件のもとでの，第２符号系Ｂに関する条件つき不確定性 $M_A(B)$は，

$$M_A(B) = \sum_{j=1}^{n} \{ M_j(B) / \mu_j(A|B) \}$$

$$\mu_j(A|B) = \sum_{i=1}^{n} \sum_{h=1}^{m} \rho_{ih}^2 / \sum_{k=1}^{m} \rho_{jk}^2 = \frac{\sum_{i=1}^{n} q_i^2 / M_i(B)}{q_j^2 / M_j(B)}$$

$$M_j(B) = q_j^2 / \sum_{k=1}^{m} \rho_{jk}^2 = 1 / \sum_{k=1}^{m} d_{jk}^2$$

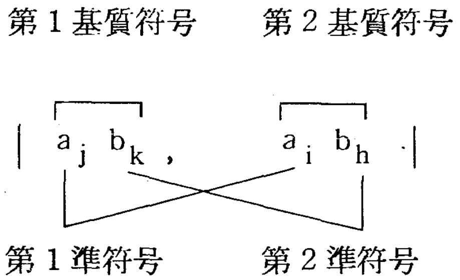
$M_j(B)$ は第1符号系Aで第 j 種の符号 a_j が出現した条件のもとでの第2符号系 B に関する条件つき不確定性である。

$$M_j(B) = \sum_{k=1}^{m} \{ 1 / \mu_{jk}(B) \; d_{jk} \}$$

$$\mu_{jk}(B) = \sum_{h=1}^{m} d_{jh}^{2} / d_{jk}^{2}$$

10．ダブレット符号の基質が溶けた媒液の不確定性の数え方

ダブレット符号の基質が溶けた熱平衡状態にある媒液を考える。基質が特定の部位に衝突をくりかえし，基質符号の時間的な系列をつくるとする。基質符号を2個ずつ重ならないように区切り，長さ2の連をつくる。連は第1基質符号と第2基質符号の2つの基質符号から構成される。基質符号は第1準符号と第2準符号の2つの準符号から構成される。



(j , i＝1, …, n;
k, h＝1, …, m)

連をつくる一対の基質符号の第1準符号がたがいに同種である場合に，第1準ホモと呼ぶ。それ以外の場合に第1ヘテロと呼ぶ。同様にそれらの第2準符号がたがいに同種である場合に，第2準ホモと呼ぶ。それ以外の場合に第2準ヘテロと呼ぶ。一対の基質符号の第1および第2準符号がともにたがいに同種の場合に，それらはホモであると呼ぶ。それ以外の場合にヘテロと呼ぶ。

第 jk 種のホモの連の出現確率は ρ_{jk}^{2} である。全部で nm 種類のホモの連があるが，何らかのホモの連の出現確率は $\sum_{j} \sum_{k} \rho_{jk}^{2}$ である。もとの連の系列に対して，ホモの連を選択する過程を考える。$M(AB) = 1 / \sum_{i} \sum_{h} \rho_{ih}^{2}$ は，もとの連の系列においてホモの連が1個出現するまでの連の平均個数である。そのような選択によってホモの連の系列が得られる。$\mu_{jk}(AB) = \sum_{i} \sum_{h} \rho_{ih}^{2} / \rho_{jk}^{2}$ は，ホモの連の系列において，第

jk 種のホモの連が１個出現するまでの連の平均個数である。そのような選択によって第 jk 種のホモの連のみの系列が得られる。

もとの連の系列にたいして，第１準ホモの連の選択する過程を考える。第 j 種の第１準ホモの連の出現確率は q_j^2 である。全部で n 種類の第１準ホモの連があり，何らかの第１準ホモの連の出現確率は $\sum_j q_j^2$ である。$M(A) = 1/\sum_i q_i^2$ は，もとの連の系列において，第１準符号のホモの連が１個出現するまでの連の平均個数である。そのような選択によって，第１準ホモの系列が得られる。第１準ホモの連の系列に対して，第 j 種の第１準ホモの連を選択する過程を考える。$\mu_j(A) = \sum_i q_i^2 / q_j^2$ は，第１準ホモの連の系列において，第 j 種の第１準ホモの連が１個出現するまでの連の平均個数である。そのような選択によって，第 j 種の第１準ホモの出現に条件づけられた連の系列が得られる。第 j 種の第１準ホモの連の系列に対して，第２準ホモの連を選択する過程を考える。第 j 種の第１準ホモの連の出現に条件づけられた，第 k 種の第２準ホモの連の出現確率は d_{jk}^2 である。全部で m 種類の第２準ホモの連があり，何らかの第２準ホモの連の出現確率は $\sum_k d_{jk}^2$ である。$M_j(B) = 1/\sum_k d_{jk}^2$ は，第 j 種の第１準ホモの出現に条件づけられた連の系列において，第２準ホモの連が１個出現するまでの連の平均個数である。そのような選択によって，第 j 種の第１準ホモの出現に条件づけられたホモの連の系列が得られる。

他方，第１準ホモの連の系列に対して，まず第２準ホモの連を選択する過程を考える。$M_A(B) = \sum_j q_j^2 / \sum_i \sum_h \rho_{ih}^2$ は，第１準ホモの連の系列において，第２準ホモの連が１個出現するまでの連の平均個数である。そのような選択によって，第１準符号と第２準符号がともに同種であるホモの連の系列が得られる。ホモの連の系列に対して，第 j 種の第１準ホモの連を選択する過程を考える。$\mu_j(A|B) = \sum_i \sum_h \rho_{ih}^2 / \sum_k \rho_{jk}^2$ は，ホモの連の系列において，第 j 種の第１準ホモの連が１個出現するまでの平均個数である。そのような選択によって，第 j 種の第１準ホモの出現に条件づけられたホモの連の系列が得られる。

第 j 種の第１準ホモの出現に条件づけられたホモの連の系列に対して，第 k 種の第２準ホモの連を選択する過程を考える。$\mu_{jk}(B) = \sum_h d_{jh}^2 / d_{jk}^2$ は，第 j 種の第１準ホモの出現に条件づけられたホモの連の系列において，第 k 種の第２準ホモの連が１個

出現するまでの連の平均個数である。そのような選択によって，第 j k 種のホモの連のみの系列が得られる。

11．ダブレット符号系の選択過程

与えられた鋳型に最適な媒液における基質符号の系列で，任意の第 i k 種の鋳型符 $(x_j\, y_k)$ に相補的に対応する第 j k 種の基質符号 $(a_j\, b_k)$ が選択される過程を調べる。

基質符号のもとの系列において，第 j k 種の基質符号語の平均長は，

$$g_{jk}(XY) = 1/\rho_{jk}^{*}$$

基質符号語の平均長を n m 種の符号に関して平均すれば，

$$G(XY) = M(AB)$$

このようなもとの系列から基質符号が選択される過程を，二つの段階に分けて描出するこことができる。

第一の段階は，第 1 準符号系 X と A との間における選択である。基質符号のもとの系列において，第 j 種の第 1 準符号が出現するまでの基質符号の平均個数，すなわち第 j 種の第 1 準符号語の平均長は，

$$g_j(X|Y) = 1/q_j^{*}$$

各種の第 1 準符号語の平均長を第 1 準符号系の n 種の符号に関して平均すれば.

$$G(X|Y) = M(A)$$

そのような選択によって，各種の第 1 準符号の出現に条件づけられた，基質符号の n 種類の準系列が得られる。

第二の段階は，第 2 準符号系 Y と B との間における選択である。第 j 種の第 1 準符号の出現に条件づけられた基質符号の準系列において，第 k 種の第 2 準符号が出現するまでの基質符号の平均個数，すなわち第 j 種の第 1 準符号に条件づけられた第 k 種の第 2 準符号語の平均長は，

$$g_{jk}(Y) = 1/d_{jk}^{*}$$

各種の第２準符号語の平均長を第２準符号系のｍ種類の符号に関して平均すれば，

$$G_j(Y) = M_j(B)$$

第２準符号語の平均長を，さらに第１準符号系のｎ種類の符号に関して平均すれば，

$$G_X(Y) = M_A(B)$$

以上の二つの段階の選択過程は乗法的に結合されている。第ｊｋ種の符号について，

$$g_{jk}(XY) = g_j(X|Y)\,g_{jk}(Y)$$

$$1/\rho_{jk}^{*} = 1/q_j^{*}\cdot d_{jk}^{*}$$

ｎｍ種類の符号に関する平均について，

$$G(XY) = G(X|Y)\,G_X(Y)$$

$$M(AB) = M(A)\,M_A(B)$$

実際の選択は二つの段階に分けて行なわれるのではない。たとえば，基質符号の準系列は取り出された形で存在するのではない。準系列はもとの基質符号の系列に備わる準符号に関する性質として認められるものである。

最適媒液では第ｊｋ種の基質符号語が使用される確率について，

$$\pi_{jk} = 1/\mu_{jk}(AB)$$

第２準符号系に遡及的に影響されながら，第ｊ種の第１準符号語が使用される確率に関して，

$$p_j = 1/\mu_j(A|B)$$

第ｊ種の第１準符号に条件づけられた，第ｋ種の第２準符号語が使用される確定に関して，

$$r_{jk} = 1/\mu_{jk}(B)$$

がなりたつ。

12．不確定性の関数の基本的性質と定理

上述した基質符号系 A と B とは以下の事象系を構成している。二つの完全事象系 A と B とを次のように与える。

$$A = \bigcup_{j=1}^{n} a_j \ , \quad a_j \cap a_i = \phi \qquad (j \neq i)$$

$$\mathrm{Prob}(a_j) = q_j \geqq 0 \qquad (j = 1, \cdots, n)$$

$$B = \bigcup_{k=1}^{m} b_k \ , \quad b_k \cap b_\ell = \phi \qquad (k \neq \ell)$$

$$\mathrm{Prob}(b_k) = d_k \geqq 0 \qquad (k = 1, \cdots, m)$$

ここに，$\sum_j q_j = 1$，$\sum_k d_k = 1$。

二つの事象 a_j と b_k とがともに起る確率を ρ_{jk} で表わせば，事象 $a_j b_k$ とその確率 ρ_{jk}（$j=1, \cdots, n$；$k=1, \cdots, m$）の全体は新しい完全事象系をつくる。　ここに，$\sum_j \sum_k \rho_{jk} = 1$。これを事象系 A と B との結合とよび，AB で表わそう。二つの事象系 A と B とが独立である場合には，

$$\rho_{jk} = q_j \, d_k$$

つぎに，事象系 A と B とがたがいに関連している場合，事象系 A で事象 a_j が起った条件のもとで，事象系 B で事象 b_k が起る確率を $d_{jk} \geqq 0$ で表わすと，

$$\rho_{jk} = q_j \, d_{jk}$$

ここに，$\sum_k d_{jk} = 1$。

A，B，AB に対して次の関数がそれぞれ定義される。

$$M(A) = M(q_1, \cdots, q_n)$$

$$M(B) = M(d_1, \cdots, d_m)$$

$$M(AB) = M(\rho_{11}, \cdots, \rho_{nm})$$

また，Ａで a_j が起った条件のもとで，Ｂに対して次の関数が定義される。

$$M_j(B) = M(d_{j1}, \cdots, d_{jm})$$

Ａが起った条件のもとで，Ｂに対して次の関数が定義される。

$$M_A(B) = \sum_{j=1}^{n} q_j M_j(B)$$

$M_A(B)$ はその特殊な場合の関数形として，ＡとＢとが独立である場合の $M(B)$ を含んでいる。

関数Ｍは次の二つの基本的性質を有していることが，これまでに明らかとなった。

（1） 任意の固定された n のもとで，$\sum_{j=1}^{n} q_j = 1$ ならば，$M(q_1, \cdots, q_n)$ は $q_j = 1/n\ (j=1, \cdots, n)$ のとき最大値をとる。

（2） 事象系ＡとＢとを結合して新しい事象系ABをつくるとき，乗法性を有する。

$$M(AB) = M(A)\, M_A(B)$$

さらに，次の二つの性質をつけ加えることができる。

（3） $M(A) = M(q_1, \cdots, q_n)$ は，全定義域で連続である。

（4） 事象系Ａに不可能な事象 a_{n+1}, $\mathrm{Prob}(a_{n+1}) = 0$ を加えても，関数値は不変である。

$$M(q_1, \cdots, q_n, 0) = M(q_1, \cdots, q_n)$$

〔定理〕 $M(q_1, \cdots, q_n)$ を，任意の自然数 n と，$q_j \geqq 0\ (j=1, \cdots, n)$，$\sum_{j=1}^{n} q_j = 1$ を満たすすべての $q_1, \cdots, q_n$ にたいして定義された関数とする。もしも，この関数が上述の性質（1），（2），（3），（4）を満たすならば，

$$M(q_1, \cdots, q_n) = 1 \Big/ \sum_{j=1}^{n} q_j^{1+\lambda}$$

ここに，λ は正の定数である。（証明は補遺３を参照）。

関数

$$M(q_1, \cdots, q_n) = 1 \Big/ \sum_{j=1}^{n} q_j^{1+\lambda}$$

に変数変換

$$q_j = p_j^{1/(1+\lambda)} \Big/ \sum_{i=1}^{n} p_i^{1/(1+\lambda)}$$

を行なえば，

$$G(p_1, \cdots, p_n) = \left(\sum_{i=1}^{n} p_i^{1/(1+\lambda)}\right)^{1+\lambda}$$

が得られる。これらの関数について，

$$M(q_1, \cdots, q_n) \leqq M(1/n, \cdots, 1/n) = n^{\lambda}$$

および

$$G(p_1, \cdots, p_n) \leqq G(1/n, \cdots, 1/n) = n^{\lambda}$$

の関係がなりたつ（証明は補遺4を参照）。

13．コードン系とアンチコードン系の構成

4種類の符号から構成される符号系が三重に結合した鋳型符号系XYZ（コードン：codon）と基質符号系ABC（アンチコードン：anti−codon）を考える。4種類の符号はmRNAおよびtRNAを構成する塩基，すなわち，Adenine，Uracil，Guanine，Cytosineである。鋳型符号の構成を次のように決める。

$$X = \{x_j \mid j = 1, 2, 3, 4\}$$

$$= \{x_1, x_2, x_3, x_4\} = \{A, G, U, C\}$$

$$Y = \{y_k \mid k = 1, 2, 3, 4\}$$

$$= \{y_1, y_2, y_3, y_4\} = \{A, G, U, C\}$$

$$Z = \{z_\ell \mid \ell = 1, 2, 3, 4\}$$

$$= \{z_1,\ z_2,\ z_3,\ z_4\} = \{A,\ G,\ U,\ C\}$$

準符号系X，Y，Z を結合した符号系XYZについて，

$$XYZ = \{x_j\ y_k\ z_\ell \mid j,\ k,\ \ell = 1,\ 2,\ 3,\ 4\}$$

$$= \{x_1\ y_1\ z_1,\ \cdots,\ x_1\ y_4\ z_3,\ \cdots,\ x_4\ y_4\ z_4\}$$

$$= \{AAA,\ \cdots,\ ACU,\ \cdots,\ CCC\}$$

mRNA の鋳型は，L個の鋳型符号 $x_j\ y_k\ z_\ell$（$j,\ k,\ \ell = 1,\ 2,\ 3,\ 4$）が順次に配列した鎖であるとする。

$$(x_j\ y_k\ z_\ell)_1 \cdot (x_j\ y_k\ z_\ell)_2 \cdot \cdots \cdot$$

$$(x_j\ y_k\ z_\ell)_i \cdot \cdots \cdot (x_j\ y_k\ z_\ell)_L$$

第 $jk\ell$ 種の符号に含まれる個数を $L_{jk\ell}$ 個とする。第 $jk\ell$ 種の鋳型符号 $x_j\ y_k\ z_\ell$ の出現確率 $\pi_{jk\ell}$ は相対頻度で与えられる。

$$\pi_{jk\ell} = L_{jk\ell} / L \quad (j,\ k,\ \ell = 1,\ 2,\ 3,\ 4)$$

準符号系X，Y，Z で準符号 x_j，y_k，z_ℓ がそれぞれ出現する確率 p_j，r_k，s_ℓ は，

$$p_j = \sum_k \sum_\ell \pi_{jk\ell},\quad r_k = \sum_j \sum_\ell \pi_{jk\ell},\quad s_\ell = \sum_j \sum_k \pi_{jk\ell}$$

で与えられる。

鋳型符号を構成する上で，三つの準符号系 X，Y，Z が独立の場合には，

$$\pi_{jk\ell} = p_j\ r_k\ s_\ell$$

の関係がある。また準符号系X，Y，Z が隣接するものの間で関連する場合，準符号 x_j が出現した条件のもとで準符号 y_k が出現する条件つき確率を r_{jk}，同様に準符号 y_k が出現した条件のもとで準符号 z_ℓ が出現する条件つき確率を $s_{k\ell}$ とすれば，

$$\pi_{jk\ell} = p_j\ r_{jk}\ s_{k\ell}$$

の関係がある。ここに，$\sum_{k} r_{jk}=1$，$\sum_{\ell} s_{k\ell}=1$ である。

基質符号の構成を次のように決める。

$$A=\{a_j \mid j=1,2,3,4\}$$
$$=\{a_1, a_2, a_3, a_4\}=\{U, C, A, G\}$$
$$B=\{b_k \mid k=1,2,3,4\}$$
$$=\{b_1, b_2, b_3, b_4\}=\{U, C, A, G\}$$
$$C=\{c_\ell \mid \ell=1,2,3,4\}$$
$$=\{c_1, c_2, c_3, c_4\}=\{U, C, A, G\}$$

準符号系A，B，Cを結合した符号系ABCについて，

$$ABC=\{a_j\, b_k\, c_\ell \mid j, k, \ell=1,2,3,4\}$$
$$=\{a_1 b_1 c_1, \cdots, a_1 b_4 c_3, \cdots, a_4 b_4 c_4\}$$
$$=\{UUU, \cdots, UGA, \cdots, GGG\}$$

基質符号 $a_j\, b_k\, c_\ell$（$j, k, \ell=1,2,3,4$）は，特定の部位に吸着をくりかえし，時間的な系列をつくる。

$$(a_j\, b_k\, c_\ell)_1,\ (a_j\, b_k\, c_\ell)_2,\ \cdots,$$
$$(a_j\, b_k\, c_\ell)_i,\ \cdots$$

基質符号の系列 $\{(a_j\, b_k\, c_\ell)_i\}$ は確率的に相互に独立であると仮定する。媒液における全基質の濃度を〔S〕，それに含まれる第 $jk\ell$ 種の基質の濃度を〔$S_{jk\ell}$〕とする。第 $jk\ell$ 種の基質符号 $a_{jk\ell}$ の出現確率 $\rho_{jk\ell}$ は，第 $jk\ell$ 種の基質の相対濃度によって与えられる。

$$\rho_{jk\ell}=[S_{jk\ell}]/[S] \quad (j, k, \ell=1,2,3,4)$$

準符号系 A，B，C で準符号 a_j，b_k，c_ℓ がそれぞれ出現する確定 q_j，d_k，t_ℓ は，

$$q_j = \sum_k \sum_\ell \rho_{jk\ell},\quad d_k = \sum_j \sum_\ell \rho_{jk\ell},\quad t_\ell = \sum_j \sum_k \rho_{jk\ell}$$

で与えられる。

基質符号を構成する上で，三つの準符号系 A，B，C が独立の場合には，

$$\rho_{jk\ell} = q_j\, d_k\, t_\ell$$

の関係がある。また準符号系 A，B，C が隣接するものの間で関連する場合，準符号 a_j が出現した条件のもとで準符号 b_k が出現する条件つき確率を d_{jk}，同様に準符号 b_k が出現した条件のもとで準符号 c_ℓ が出現する条件つき確率を $t_{k\ell}$ とすれば，

$$\rho_{jk\ell} = q_j\, d_{jk}\, t_{k\ell}$$

の関係がある。

14．コードンによるアンチコードンの選択過程

任意の一例としてコードン ACU に相補的なアンチコードン UGA が最適媒液のアンチコードンの系列において選択される過程を描出しよう（図9，10参照）。

アンチコードンのもとの系列における UGA 種の基質符号語の平均長は，

$$g_{ACU}(XYZ) = 1/\rho_{UGA}^{*}$$

各種の基質符号語の平均長を，64種の符号に関して平均すれば，

$$G(XYZ) = M(ABC)$$

このような，もとの系列から1個のアンチコードンが選択される過程を，三つの段階に分けて描出することができる。

第一の段階は，第1準符号系 X と A との間における選択である。もとの3次の系列において，第1準符号が U であるようなアンチコードン U b_k c_ℓ が選択されるまでに必要なアンチコードンの平均個数，すなわち U 種の第1準符号語の平均長は，

$$g_A(X \mid YZ) = 1/q_U^{*}$$

第1準符号系の4種の準符号に関する，この値の平均は，

$$G(X \mid YZ) = M(A)$$

そのような選択の結果，第1準符号の出現に条件づけられた4種の2次の準系列が得られる。U種の2次の準系列において，UGAが出現するまでに必要なアンチコードンの平均個数は，

$$g_{ACU}(YZ) = 1 / d^{*}_{UG}\ t^{*}_{GA}$$

図9　Anti－codonの系列におけるCodon ACUに相補的なAnti－codon UGAの選択過程

〔anti－codonのもとの3次の系列〕

UAA　CUC　CAG　UCC　UGA　GCU　ACU　CAG　CUU　AGU
CUA　GCC　AAG　UUG　ACU　GGC　UAG　AUC　UAC　GAA
…………………

↙U　　　C↘　A↘　G↘

〔第1準符号Uに条件づけられた，2次の系列〕

他　3

UAA　UCC　UGA　UUG　UAG　UAC
UCA　UGU　UUA　UAA　UGC　UCG
…………………

↙G　　　U↘　C↘　A↘

〔第2準符号Gに条件づけられた，1次の系列〕

他　15

UGA　UGU　UGC　UGA　UGG
UGC　UGU　UGC　UGG　UGA
…………………

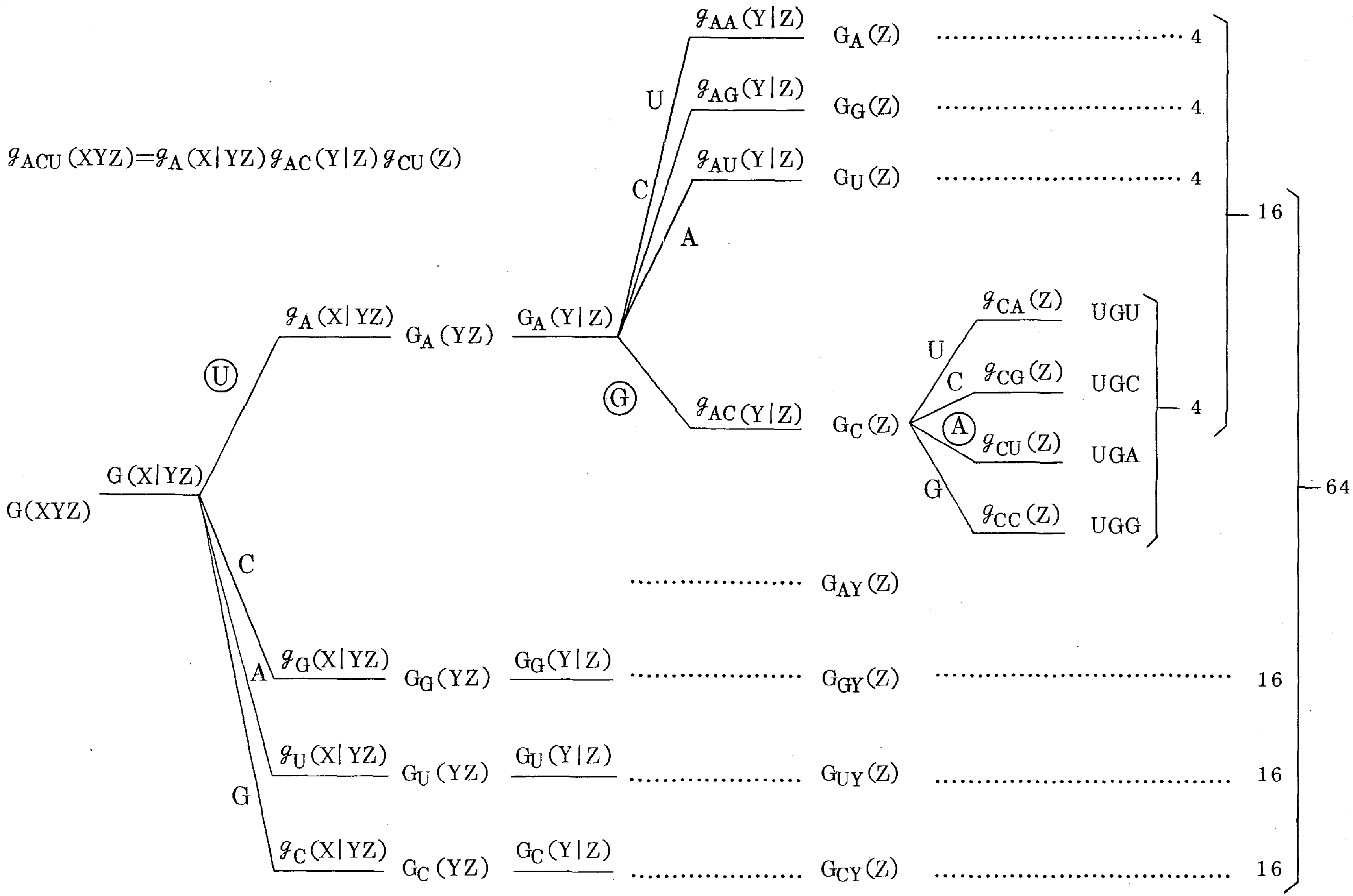
↙A　　　U↘　C↘　G↘

〔第3準符号Aによって，一意に選択されたanti－codon〕

他　63

UGA

図10　ACUに相補的なUGAの選択過程



第2，3準符号系の16種の準符号に関する，この値の平均は，

$$G_A(YZ) = M_U(BC)$$

さらに，第1準符号系の4種の準符号に関する，この値の平均は，

$$G_X(YZ) = M_A(BC)$$

第二の段階は，第2準符号系YとBとの間における選択である。U種の2次の準系列において，第2準符号がGであるようなアンチコードンUGc_ℓが選択されるまでに必要なアンチコードンの平均個数，すなわちU種の第1準符号に条件づけられた，G種の第2準符号語の平均長は，

$$g_{AC}(Y|Z) = 1/d_{UG}^*$$

第2準符号系の4種の準符号に関する，この値の平均は，

$$G_A(Y|Z) = M_U(B)$$

他の3種の2次の準系列においても同様の選択がなされる。そのような選択の結果，第1，2準符号の出現に条件づけられた16種の1次の準系列が得られる。

第三の段階は，第3準符号系ZとCとの間における選択である。UG種の1次の準系列において，第3準符号がAであるようなアンチコードンUGAが選択されるまでに必要なアンチコードンの平均個数，すなわちG種の第2準符号に条件づけられた，A種の第3準符号語の平均長は，

$$g_{CU}(Z) = 1/t_{GA}^*$$

第3準符号系の4種の準符号に関する，この値の平均は，

$$G_C(Z) = M_G(C)$$

さらに，第2準符号系の4種の準符号に関する，この値の平均は，

$$G_{AY}(Z) = M_{UB}(C)$$

以上の結果として，アンチコードンUGAが選択されるまでの過程において，

$$g_{ACU}(XYZ) = g_{A}(X|YZ)\, g_{AC}(Y|Z)\, g_{CU}(Z)$$

$$1/\rho_{UGA}{}^{*} = 1/q_{U}{}^{*}\ d_{UG}{}^{*}\ t_{GA}{}^{*}$$

がなりたつ。コードン・アンチコードン系の64種の符号に関する，この値の平均について，

$$G(XYZ) = G(X|YZ)\, G_{X}(YZ)$$

$$M(ABC) = M(A)\, M_{A}(BC)$$

および，各種の第1準符号に条件づけられた，第2，3準符号系に関する平均について，

$$G_{A}(YZ) = G_{A}(Y|Z)\, G_{AY}(Z) \qquad \text{他3種}$$

$$M_{U}(BC) = M_{U}(B)\, M_{UB}(C) \qquad \text{他3種}$$

の関係がなりたつ。

15．議　論

遺伝現象の情報的性格については数多く指摘されてきた（Stent，1968）。その分子的機構はDNA → DNA の複製とDNA → mRNA → Protein の形質発現を中心に解明されてきた。しかし遺伝情報は体系的な理論として把握されていなかった。本研究はシャノンの通信理論の数学的形式そのものではなく，そこに含まれている基本的な考え方を遺伝現象に適用したものである。合成の場合と通信の場合とで関数形が異なるのは，対応づけのための識別機構が相違するからである。

本研究は鋳型を介する合成反応における空間と時間の概念を検討することから始まった。それは現象を時空上の物質の運動として明確化するために必要である。鋳型を介する合成反応は分子の並存と継起の相互転換によって進行する，という特質が認められる。分子の並存とは一つの時点に多くの分子が空間的に並んで存在すること（時間の一における空間の多）である。分子の継起とは一つの部位に多くの分子が時間的に順次に生起すること（空間の一における時間の多）である。

鋳型は各時点にそれを構成するすべての鋳型符号が鎖状に並存したものである。鋳型符号間の安定した化学結合（ポテンシャルエネルギー）によって，その系列は時間的に

伝達（遺伝）される。合成酵素が鋳型上を走査運動するにつれて順次に出現する鋳型符号は時間的な系列をなす。この過程は分子の並存から継起への転換である。合成酵素が一段階ずつ位置移動する際に，いかなる種類の鋳型符号が出現するかは，合成酵素にとって偶然的な事象である。ここに鋳型の不確定性が認められる。与えられた鋳型について符号の種類の順序は確定されている。いわば，鋳型の不確定性は確定性を含む不確定性である。

基質符号の系列は一つの吸着部位に媒液を構成する基質符号が一個ずつ時間的に継起したものである。基質符号の熱運動（カイネテイックエネルギー）によって，その系列は順次に独立に更新される。基質符号の時間的な系列から，鋳型を介する合成反応によって，基質が鎖状に結合した系列が作られる。この過程は分子の継起から並存への転換である。個々の衝突によって，いかなる種類の基質符号が基質特異的吸着を行なうかは，吸着部位にとって偶然的な事象である。ここに基質符号の系列（または媒液）の不確定性が認められる。基質符号の時間的な系列はランダムである。基質符号の系列の不確定性はそれ自体としては確定性を含まない不確定性である。基質符号は鋳型を介する合成反応の選択過程を通じて，鋳型の確定性と結びつくことができる。不確定な状態は選択的な対応づけの媒介によって確定した状態へ転換する。その選択的な指示作用をする媒介物が情報である。情報は確定性を結ぶものであり，両方の性質を備えている。

ある物理学者（Delbürck，1949）が，ファージを対象に分子遺伝学を始めたのは，生物における物理法則を探すためだった。生物現象を物理的基礎の上に分析し，理解するために必要な基本的な概念，理論がまだ欠けているように思われた。その法則は，物質の有機的な運動に普遍的に成り立つもので，従来知られている物理法則と相補的な関係にあるものと予想された。生きた細胞を調べていくことによって，明白なパラドックスに出会い，問題の焦点の浮かび上がることが期待された。

本研究の重要な結果の一つは，媒液の不確定性を表わす関数（M）である。

$$M = 1 \Big/ \sum_{i=1}^{n} q_i^{1+\lambda}$$

$$\left(\sum_{i=1}^{n} q_i = 1, \quad \lambda：\text{正の定数} \right)$$

この関数は熱力学および通信理論で用いられるエントロピー関数（H）

$$H = -\lambda \sum_{i=1}^{n} p_i \log p_i$$

$$\left(\sum_{i=1}^{n} p_i = 1, \quad \lambda：\text{正の定数} \right)$$

と比べられる。M関数はH関数のもつ加法性の代わりに乗法性をもつほかは，同様の性質を備えている。それは新しい別のエントロピーと呼ばれるべき資格を有している。異種分子の混合過程を思考実験で調べたとき，われわれは明白なパラドックスにぶつかった。すなわち，熱力学的な混合のエントロピーは，異種分子が混合しないときに変化し，それらが混合するときに変化しない。混合するときの過程は，媒液の不確定性(M)によって表わされることがわかった。熱力学的なエントロピーと媒液の不確定性は相補い合って，異種分子の混合現象を十分に表現する。

生物は物理法則が通用する漠大な数の原子や分子から構成されている。たがいに分離した様々な物質が相互に関連することは，有機体が成立するための不可欠の条件の一つである。まさに“all interrelated and all interdependent”（Delbürck）である。すべての生きた形態は無数の糸で相互に関連しているから，独立した系でなりたち，しかも変わることのない絶対的な現象はないかもしれない。けれどもわれわれは今や相互関連こそが絶対的であることを知る。通信はそのような相互関連の一つの重要な方式である。その意味で通信的過程はすべての有機体にとって普編的な現象である。

DNAからmRNAへの転写においては塩基のシングレット符号が，そしてmRNAからタンパク質への翻訳においてはトリプレット符号が用いられる。mRNAはシングレット符号からトリプレット符号への変換を行なう。トリプレット符号は，すべてのアミノ酸を一義的に指定するに十分な種類の数を有している。シングレット符号からトリプレット符号への変換に際して，1符号あたりの最大不確定性は $G(1/4, \cdots, 1/4) = 4$（コット）から $G(1/64, \cdots, 1/64) = 64$（コット）に増加する。別々に読み出された3個の符号が結合して，高次の1個の符号を形成するとき，3個の符号の加法的な関係（$4+4+4=12$）は，乗法的な関係（$4\times4\times4=64$）へ変わる。一本のポリペプチド鎖に対応するmRNAのコードンの個数をL個とする。それに相当するシングレット符号の個数は3L個である。鋳型全体の不確定性は，一符号あたりの不確定性と符号の個数の積で与えられる。したがって，長さ3L個のmRNAの最大不確

定性は 12L（コット）から 64L（コット）に増加する。すなわち，同じ長さでありながら，64L／12L＝5・1/3 倍になる。これは鋳型の情報容量が符号の解読規則の変更によって増大することを示す。このような事情は不確定性の関数が乗法性を有する場合にのみ起ることで，それが加法性を有する場合には起らない。

ただし，若干考慮すべきことがある。アミノ酸の種類の数（約２０　は，トリプレット符号の種類の数（64）より少ない。mRNA 上のコードンの３番目のＡとＧあるいはＵとＣは区別されないでアンチコードンに読まれるという符号の縮退現象が見つけられている（Eck，1963；Crick，1966；Söll & Khorana et al，1966)。実際に用いられている識別可能な種類の数は，少なくともこの理由によって半減する。これは不必要な鋳型の不確定性を減少させるものである。

結局，鋳型の不確定性の関数のもつ乗法性は，生物の多様性を生み出す機構の一つの特性を示すのであろう。

アミノ酸および核酸塩基の生合成系において，最終産物の負のフィードバックによる制御機構が知られている。基質濃度の制御量を目標値に近づけて定常に保つことが制御のはたらきである。そのような制御は基質符号語（F）のできるだけ小さい適当な値を恒常的に維持するために不可欠である。細胞で知られている実際の制御機構の特性を，このような観点から検討する必要があるだろう。

参 考 文 献

Brillouin, L. "Science and information theory"
Academic Press Inc. New York (1956)

Cold Spring Harbor Symp. Quant. Biol. Vol. 33 (1968)
〃 Vol. 34 (1969)
〃 Vol. 35 (1970)

Cox, D. R. "Renewal theory"
Methuen & Co. Ltd, London (1962)

Crick, F. H. C. J. Mol. Biol. **19**, 548 (1966)

Delbürck, M. "A physicist looks at biology" (1949)
in "Phage and the origins of molecular biology"
Cold Spring Harbor Lab. Quant. Biol. (1966).

Eck, R. V. Science **140**, 477 (1963)

Fano, R. M. "Transmission of information, a statistical theory of communications" (1963)
The M. I. T. Press, Cambridge

Feller, W. "An introduction to probability theory and its applications" Vol. 1
John Wiley & Sons, Inc. New York (1950)

Khinchin, A. I. Uspehi Matematiceskin Nauk **8**, 3 (1953)
国沢，梅垣編訳〝情報理論の進歩”（1965）岩波書店

Shannon, C. E. & Weaver, W. "The mathematical theory of communication" (1964)
The University of Illinois Press, Urbana

Söll, D., H. G. Khorana et al. J. Mol. Biol. **19**, 556 (1966)

Stent, G. "That was the molecular biology, that was"
Science, **160**, 390 (1968) 杉野訳 科学 **38**,（1968）岩波書店

Watson, J. D. & F. H. C. Crick, Nature, **171**, 737 (1953)

Watson, J. D. & F. H. C. Crick, Nature, **171**, 964 (1953)

久保編　「大学演習　熱学・統計力学」　(1961)　裳華房

中村伝　「統計力学」　(1967)　岩波全書２６９

ヤグロム ＆ ヤグロム　「情報理論入門」　(1957)　井関，西田訳，みすず書房

高木貞治　「解析概論」　岩波書店

補遺1．媒液容量Vの導出

時間（0，t）における吸着の総数をN_t，r番目の吸着が終わるまでの時間を$S_r = T_1 + \cdots\cdots + T_r$とおく。$N_t$と$S_r$との確率分布の間には，

$$P(N_t \geqq r) = P(S_r < t) \tag{1}$$

の関係がある。

S_rは，中心極限定理より，充分大きなrに対して，平均$r(\mu+1/\rho)$および分散$r(\sigma^2+1/\rho^2)$をもつ正規分布に漸近的に近ずく。すなわち，任意の定数βに対して．$r\to\infty$のとき，

$$P\left(\frac{S_r - r(\mu+1/\rho)}{\sqrt{r(\sigma^2+1/\rho^2)}} < \beta\right) \longrightarrow \Phi(\beta) \tag{2}$$

が成立する。ここに，$\Phi(\beta)$は標準の正規分布関数である。いま，

$$\frac{t - r(\mu+1/\rho)}{\sqrt{r(\sigma^2+1/\rho^2)}} \to \beta$$

となるように，$t\to\infty$，$r\to\infty$とすれば，(1)と(2)の両式から

$$P(N_t \geqq r) \to \Phi(\beta)$$

書き換えれば，

$$P\left(\frac{N_t - t(\mu+1/\rho)^{-1}}{t^{1/2}(\sigma^2+1/\rho^2)^{1/2}(\mu+1/\rho)^{-3/2}} > \beta\right) \longrightarrow \Phi(\beta)$$

補遺2．関数Fの最小値Gの導出

与えられた係数$p_1, \cdots, p_n$に対して，完全系の媒液の変数$q_1, \cdots, q_n$の関数$F(q_1, \cdots, q_n)$の最小値を求める（高木貞治を参照）。

$$\begin{cases} F(q_1, \cdots, q_n) = \sum_{i=1}^{n} p_i/q_i & (1) \\ \varphi(q_1, \cdots, q_n) = 1 - \sum_{i=1}^{n} q_i = 0 & (2) \end{cases}$$

$$\partial \varphi / \partial q_n = -1 \neq 0 \tag{3}$$

だから，q_n は q_α（$\alpha = 1, \cdots, n-1$）の関数となり，したがって F は q_α の関数となる。

$$F(q_1, \cdots, q_n) = f(q_1, \cdots, q_{n-1}) \tag{4}$$

（i） 極値の必要条件は

$$\partial f / \partial q_\alpha = 0 \qquad (\alpha = 1, \cdots, n-1) \tag{5}$$

すなわち

$$\partial f / \partial q_\alpha = \partial F / \partial q_\alpha + \partial F / \partial q_n \cdot \partial q_n / \partial q_\alpha = 0 \tag{6}$$

(2) 式より

$$\partial \varphi / \partial q_\alpha + \partial \varphi / \partial q_n \cdot \partial q_n / \partial q_\alpha = 0 \tag{7}$$

(6) 式と (7) 式より，$\partial q_n / \partial q_\alpha$ を消去して

$$\partial F / \partial q_\alpha \cdot \partial \varphi / \partial q_n - \partial F / \partial q_n \cdot \partial \varphi / \partial q_\alpha = 0 \tag{8}$$

計算して

$$P_\alpha / q_\alpha^2 = P_n / q_n^2 = G \quad (\text{constant}) \quad (\alpha = 1, \cdots, n-1) \tag{9}$$

これより

$$q_i = \sqrt{p_i / G} \qquad (i = 1, \cdots, n) \tag{10}$$

(2) 式に代入して

$$1 - \sum_{i=1}^{n} \sqrt{p_i / G} = 0$$

これより

$$\sqrt{G} = \sum_{i=1}^{n} \sqrt{p_i} \tag{11}$$

ゆえに

$$q_i = \sqrt{p_i} \Big/ \sum_{j=1}^{n} \sqrt{p_j} \qquad (i = 1, \cdots, n) \tag{12}$$

のとき，$F(q_1, \cdots, q_n)$ は極値をとる。求める極値は，

$$F\left(\sqrt{p_1} \Big/ \sum_{j=1}^{n} \sqrt{p_j}\ , \cdots , \sqrt{p_n} \Big/ \sum_{j=1}^{n} \sqrt{p_j}\right) = \left(\sum_{i=1}^{n} \sqrt{p_i}\right)^2 = G \tag{13}$$

(ii) 極値が最小値であることの証明

点 $S(\sqrt{p_1/G}, \cdots, \sqrt{p_{n-1}/G})$ における f の二次偏導関数として $a_{\alpha\beta}$ を次のように定義する。

$$a_{\alpha\beta} = f_{\alpha\beta}(S) = \frac{\partial}{\partial q_\beta}\left\{\frac{\partial}{\partial q_\alpha} f(S)\right\} \tag{14}$$

$$f_{\alpha\beta} = \frac{\partial}{\partial q_\beta}\left(\frac{\partial f}{\partial q_\alpha}\right) = \frac{\partial}{\partial q_\beta}\left(\frac{\partial F}{\partial q_\alpha} + \frac{\partial F}{\partial q_n} \cdot \frac{\partial q_n}{\partial q_\alpha}\right)$$

$$= \frac{\partial}{\partial q_\beta}\left(\frac{\partial f}{\partial q_\alpha}\right) + \frac{\partial}{\partial q_\beta}\left(\frac{\partial F}{\partial q_n}\right)\frac{\partial q_n}{\partial q_\alpha} + \frac{\partial F}{\partial q_\beta} \cdot \frac{\partial}{\partial q_\beta}\left(\frac{\partial q_n}{\partial q_\alpha}\right)$$

$$= F_{\alpha\beta} + F_{n\beta} \cdot \frac{\partial q_n}{\partial q_\alpha} + F_n \frac{\partial}{\partial q_\beta}\left(\frac{\partial q_n}{\partial q_\alpha}\right) \tag{15}$$

(7) 式より

$$\frac{\partial q_n}{\partial q_\alpha} = -\frac{\partial \varphi / \partial q_\alpha}{\partial \varphi / \partial q_n} = -\frac{\varphi_\alpha}{\varphi_n} \tag{16}$$

また

$$\frac{\partial}{\partial q_\beta}\left(\frac{\partial q_n}{\partial q_\alpha}\right) = \frac{\varphi_\alpha(\partial \varphi_n / \partial q_\beta) - \varphi_n(\partial \varphi_\alpha / \partial q_\beta)}{{\varphi_n}^2}$$

$$= (\varphi_\alpha\ \varphi_{n\beta} - \varphi_n\ \varphi_{\alpha\beta}) / {\varphi_n}^2 \tag{17}$$

よって，

$$f_{\alpha\beta} = F_{\alpha\beta} - F_{n\beta}\ \varphi_{\alpha}/\varphi_{n}$$

$$+ F_{n}\,(\varphi_{n}\ \varphi_{n\beta} - \varphi_{n}\ \varphi_{\alpha\beta})/\varphi_{n}^{2} \tag{18}$$

また,

$$F_{\alpha\beta} = \begin{cases} 2\,q_{\alpha}/q_{\alpha}^{3} & (\alpha=\beta) \\ 0 & (\alpha\neq\beta) \end{cases} \tag{19}$$

$$\varphi_{\alpha} = -1, \qquad \varphi_{\alpha\beta} = 0 \tag{20}$$

以上より, $\alpha\neq\beta$ のとき, $f_{\alpha\beta} = 0$

よって,

$$a_{\alpha\beta} = f_{\alpha\beta}(S) = 0 \tag{21}$$

$\alpha=\beta$ のとき, $f_{\alpha\alpha} = 2\,p_{\alpha}/q_{\alpha}^{3}$

よって,

$$a_{\alpha\alpha} = f_{\alpha\alpha}(S) = 2\,G^{3/2}\ p_{\alpha}^{-1/2} \tag{22}$$

$a_{\alpha\beta}$ の行列式の首座行列式 D_k について,

$$D_k = \begin{vmatrix} a_{11}, & a_{12}, & \cdots, & a_{1k} \\ a_{21}, & a_{22}, & \cdots, & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1}, & a_{k2}, & \cdots, & a_{kk} \end{vmatrix}$$

$$= \begin{vmatrix} 2\,G^{3/2}\ p_{1}^{-1/2}, & 0, & \cdots\cdots, & 0 \\ 0, & 2\,G^{3/2}\ p_{2}^{-1/2}, & \cdots\cdots, & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & \cdots\cdots, & 2\,G^{3/2}\ p_{k}^{-1/2} \end{vmatrix}$$

$$= 2^{k}\ G^{3k/2}\,(p_{1}\ p_{2} \cdots p_{k})^{-1/2} > 0 \tag{23}$$

D_k（$k=1, \cdots, n-1$）はすべて正であるから，求めた極値Gは極小である。

補遺3．不確定性の関数Mの公理的導出
（定理の証明）

$$M(1/n, \cdots, 1/n) = L(n) \qquad n \geqq 2$$

とおく。

$$L(n) = M(1/n, \cdots, 1/n, 0)$$
$$\leqq M(1/(n+1), \cdots, 1/(n+1)) = L(n+1)$$

よって，$L(n)$ は非減少関数である。m，r（$r>1$）を任意の自然数とする。m個のたがいに独立な完全事象系 $S_1, \cdots, S_m$ を考え，それらはそれぞれ r 個の等確率の 事象から成るものとする。

$$M(S_k) = M(1/r, \cdots, 1/r) = L(r) \qquad (k=1, \cdots, m)$$

事象 S_k はたがいに独立だから，

$$M\left(\prod_{k=1}^{m} S_k\right) = \prod_{k=1}^{m} M(S_k) = L(r)^m$$

$S_1 S_2 \cdots S_m$ は r^m 個の等確率事象からなるから，その関数値は $L(r^m)$ に等しい。

$$L(r^m) = L(r)^m \tag{1}$$

よって，任意の自然数の対 n，t（$t>1$）にたいして

$$L(t^n) = L(t)^n \tag{2}$$

$r>1$，$t>1$，n を任意のあたえられた自然数として，m は次の不等式

$$r^m \leqq t^n \leqq r^{m+1} \tag{3}$$

を満足する自然数であるとすると，

$$m \log r \leqq n \log t \leqq (m+1) \log r$$

$\log r > 0$ だから

$$m/n \leqq \log t / \log r < m/n + 1/n \tag{4}$$

$L(n)$ は非減少だから，(3) より

$$L(r^m) \leqq L(t^n) \leqq L(r^{m+1})$$

(1) と (2) より

$$L(r)^m \leqq L(t)^n \leqq L(r)^{m+1}$$

$$m \log L(r) \leqq n \log L(t) \leqq (m+1) \log L(r)$$

$L(r)$ は常に 1 でないから

$$m/n \leqq \log L(t) / \log L(r) \leqq m/n + 1/n \tag{5}$$

(4) と (5) より

$$\left| \frac{\log L(t)}{\log L(r)} - \frac{\log t}{\log r} \right| \leqq \frac{1}{n}$$

n はいくらでも十分に大きくえらべるから，

$$\log L(t) / \log t = \log L(r) / \log r$$

r，t は任意だから，すべての $n(\geqq 2)$ にたいして，

$$\log L(n) = \lambda \log n$$

ゆえに，$L(n) = n^{\lambda}$

$L(n)$ の非減少性より，$\lambda > 0$ である。

事象系Aの第 j 種の事象 a_j の確率 q_j $(j=1, \cdots, n)$ が有理数のとき，適当な自然数 u_j と u により

$$q_j = u_j / u \qquad (j=1, \cdots, n)$$

と表わされる。ここに，$\sum_{j=1}^{n} u_j = u$。

事象系Aと事象系Bとの結合して，新しい事象系ABをつくる。事象系Aで第j種の事象 a_j が起った条件のもとで，事象系Bで第k種の事象 b_k が起る条件つき確率を d_{jk} とする。u_j の中の最大値をm，kを系列1，…，mの中の数とする。そして d_{jk} を，

$$\begin{cases} d_{jk} = 1/u_j & (k=1,\cdots,u_j) \\ d_{jk} = 0 & (k=u_j+1,\cdots,m) \end{cases}$$

になるように定める。

ここに，

$$\sum_{k=1}^{m} d_{jk} = \sum_{k=1}^{u_j} 1/u_j = 1$$

事象系ABの事象 $a_j\ b_k$ の起る確率 ρ_{jk} は，

$$\rho_{jk} = q_j\ d_{jk} \qquad (j=1,\cdots,n\ ;\ k=1,\cdots,m)$$

である。よって，

$$\rho_{jk} = (u_k/u)\cdot(1/u_k) = 1/u$$

事象系Aの第j種の事象 a_j が起った条件のもとでの事象系Bに対する関数は，

$$\begin{aligned} M_j(B) &= M_j(1/u_j,\cdots,1/u_j,0,\cdots,0) \\ &= M_j(1/u_j,\cdots,1/u_j) \\ &= L(u_j) \\ &= u_j^{\lambda} \end{aligned} \tag{6}$$

(6)より，

$$M_A(B) = \sum_{j=1}^{n} q_j\ M_j(B) = \sum_{j=1}^{n} q_j\ u_j^{\lambda} \tag{7}$$

事象系ＡＢに対する関数は,

$$M(AB) = M(1/u, \cdots, 1/u) = u^{\lambda} \tag{8}$$

関係式

$$M(AB) = M(A)\ M_A(B)$$

に (7), (8) を代入すると,

$$u^{\lambda} = M(A) \sum_{j=1}^{n} q_j\ u_j^{\lambda}$$

$$\therefore\ M(A) = \frac{1}{\sum_{j=1}^{n} q_j (u_j / u)^{\lambda}} = \frac{1}{\sum_{j=1}^{n} q_j^{1+\lambda}} \tag{9}$$

有理数 $q_1, \cdots, q_n$ に対して証明された等式(9) は，関数 $M(q_1, \cdots, q_n)$ の連続性の仮定より，これらの変数の実効値に対してもなりたつ。

（証明終り）

補遺４．不確定性の関数Ｍ，Ｇの最大値の等出

$y = f(x)$ を a から b までの区間で凸関数とし，$x_1, \cdots, x_n$ をこの区間内の n 個の値とすれば,

$$\{f(x_1) + \cdots + f(x_n)\}/n \leqq f\{(x_1 + \cdots + x_n)/n\}$$

がなりたつ。ただし，等号は $x_1 = \cdots = x_n$ のときである（ヤグロム & ヤグロム, 1957）。

(i)　$y = -q^{1+\lambda}$　$(\lambda > 0)$ は 0 から 1 までの区間で凸関数であるから，$q_1, \cdots, q_n$ をこの区間内の n 個の値とすれば,

$$-(q_1^{1+\lambda} + \cdots + q_n^{1+\lambda})/n \leqq -\{(q_1 + \cdots + q_n)/n\}^{1+\lambda}$$

がなりたつ。条件 $q_1 + \cdots + q_n = 1$ を使って，変形すれば,

$$(q_1^{1+\lambda} + \cdots + q_n^{1+\lambda})/n \geqq 1/n^{1+\lambda}$$

これより,

$$1/(q_1^{1+\lambda}+\cdots+q_n^{1+\lambda}) \leqq n^{\lambda}$$

が得られる。ただし,等号は $q_1=\cdots=q_n$ のときである。

(ii) $y=p^{1/(1+\lambda)}$ ($\lambda>0$) は0から1までの区間で凸関数であるから,p_1, …, p_2 をこの区間内の n 個の値とすれば,

$$\{p_1^{1/(1+\lambda)}+\cdots+p_n^{1/(1+\lambda)}\}/n \leqq \{(p_1+\cdots+p_n)/n\}^{1/(1+\lambda)}$$

がなりたつ。条件 $p_1+\cdots+p_n=1$ を使って,変形すれば,ただちに

$$\{p_1^{1/(1+\lambda)}+\cdots+p_n^{1/(1+\lambda)}\}^{(1+\lambda)} \leqq n^{\lambda}$$

が得られる。ただし,等号は $p_1=\cdots=p_n$ のときである。