

生物進化モデルの数理

九大・理 石井一成, 松田博嗣  
理研 萩田直史

§1. はじめに 分子遺伝学を中心に分子生物学が花々しく展開される今日、生物進化も蛋白質やDNAなど情報高分子の比較生化学や集団遺伝学として活発に研究されている。こうした分子進化のデータには、進化速度や生物集団の多型に関して、かなり一般的な経験則が認められている。しかしながら、その進化機構については、1968年に木村資生博士(国立遺伝研)が分子進化の中立説を提唱されて以来<sup>1)</sup>、いわゆる中立説・適応説論争が未だに継続中で、定説がないのが現状である。私たちは、分子進化ではある継続時間をもつて正負に変動する自然淘汰の効果が大切であろうとの判断に基づき、分子進化の環境ゆらぎ説を提唱し<sup>2)</sup>、その理論集団遺伝学的検討を行っている。題目に掲げた「生物進化モデル」も、実は、分子進化の理論的混沌を(主に進化速度の側面において)解消することを目指して導入され、研究されているものである。今日は、今までに得られた厳密な結果のあらましと、未解決の問題のいくつかを紹介したい。なるべく多くの方々に生物進化に関心をもって頂き、それぞれの観

点から進化モデルの研究に貢献して頂きたいからである。そこで、モデルの導入とその性質の紹介に入る前準備として、分子進化速度の定量的データのあらましと、その決定機構を巡る理論的問題点を簡単に見ておこう。

§ 2. 分子進化速度のデータ § 2. 比較生化学の発展により、相同蛋白質（共通の祖先に由来し、似た機能をもつ蛋白質。例えばヒトとコウのヘモグロビンは互に相同な蛋白質である）の1次構造（アミノ酸配列）の生物種間比較が数多く行われた。相同蛋白質の1次構造の差違は、共通の祖先蛋白質から現存する蛋白質に至る系統上のどこかでアミノ酸置換がいくつか起ったことによつて生じたものである。相同蛋白質の種間比較と古生物学で得られた生物種の分岐年代を用いると、祖先から現存する蛋白質分子に至る系統線上での年当りアミノ酸置換率が推定される。この置換率を当該蛋白質の分子進化速度と呼んでいる。分子進化速度は相同でない蛋白質間では異なるが、互に相同な蛋白質同志では、系統線の違ひによらず、年当りほぼ一定であることが分つてゐる。その実例をいくつか表1に示す。

分子進化速度に見られる経験的通則として、上に挙げた (i) 年当りの近似的定速性の他に、(ii) 機能上の制約の強い相同蛋白質（部分）程その進化速度は小さいようである。

こうした分子進化の特徴を説明する仮説として、中立説が提唱されている<sup>1), 3)</sup>。これは、分子レベルではほとんどの進化上の変化が適応的に中立な突然変異の(生物集団サイズの有限性に由る)「ランダム・ドリフト」によって起るとの仮説である。特に、分子進化では変異体に働く正の自然選択が無視されるという意味で、非ダーウィニズムと呼ばれている。中立説では自然選択を無視できるため、仮定の理論的帰結をさまざまな進化状況の下で定量的に調べるのが容易である。この解析のため、遺伝子頻度のランダム・ドリフトを拡散過程として調べる(木村資生博士によって長年の間に開発され発展させられた)拡散モデルの手法が活用された<sup>4), 5)</sup>。

中立説の下では、進化速度は中立突然変異率に等しいことが導かれる。したがって、通則(ii)は、機能上の制約の強い相同蛋白質(部分)程その中立突然変異率は小さいと考えさえすれば、一応理解できる。しかし、突然変異率というものは総じて年当り一定であるよりはむしろ世代当り一定であるようなので、通則(i)は中立説の下で理解し難い。さらに、中立説の下での蛋白質多型の特徴は集団の有効サイズに強く依存するが、実測は余り種に依存しない。これらの困難は中立説の提案者自身が認める所で、このための付加的な仮定を導入することによって、問題の解消が試みられている<sup>3)</sup>。しかし、

それらの仮定の妥当性は現在の所確認されてない。

では、適応説はどうかであろうか。一般に、正の自然選択を認めるといつても、選択様式としては無限の可能性が考えられる。したがって、ここでの問題は、分子進化の通則を説明できる自然選択の様式は何かということになる。この際、分子進化の通則自身から生じる枠枷があることに注意しよう。

まず第1に、電気泳動法で始めて同定できるようなアイソタイム間の選択係数は非常に小さい( $10^{-3}$ を越さないとの評価もある)。次に、表1に与えたような分子進化速度から、脊椎動物など高等動物のゲノム(卵または精子のような1倍体細胞に含まれるDNAの完全な組)当りの塩基置換率を推定してみよう。ゲノムの塩基対数が $10^9$ かそれ以上のオーダーであり、アミノ酸を指定するコドンが3個の塩基から成ることを考えると、毎世代ゲノム当たり1のオーダーの塩基置換が起ることになる<sup>1)</sup>。これらの置換過程は時間的に互に重なったものとなりなければならない。なぜなら、一つの置換過程が終了した後には次のものが始まると過程すると、1世代のオーダーの短期間に野生型と変異型との完全な入れ替りが起ることになり、野生型の短期間絶滅による集団個体数の減少を、変異型が高等動物としてありそうもない異常に高い増殖率で穴埋めしていかなければ、集団は全体として絶滅してしまう

ことになるからである。このような観点から、分子進化においては、ほぼ何時も集団中には3種類以上の異なるゲノムが同時点で存在していることになる。

以上の注意を念頭において、尚、比較的簡単に理論解析できる自然選択様式として、ゲノムの適応度が各塩基座位からの寄与分の相乗積として定まるものがある。この場合の遺伝子頻度変化は自然選択が働く2アレル拡散モデルによって解析できる。この自然選択様式は、実は、中立論者自身によって、中立説の対照仮説として徹底的に解析され、理論結果が経験通則と比較された。例えば、 $\mu$ を淘汰有利度 $s (> 0)$ をもつ変異型への突然変異率とし、集団は2倍体生物集団とすると、変異体が集団中に現れる確率は、 $N\mu \ll 1$ ならば、単位時間当たりほぼ $2N\mu$ である。そこで、進化速度はほぼ $v = 2N\mu = 4N\mu s \dots (1)$ となる。ただし、ここで $\mu$ は変異体の固定確率で、2アレル拡散モデルによる結果 $\mu = 2s$ を用いた。したがって正の自然選択( $s > 0$ )が分子進化に利くとすると、 $v$ は $Ns$ に比例することになり、これは現実において生物種ごとに $N, s$ 特に $N$ は大きく違うから、進化速度の一定性は理解し難い<sup>3)</sup>。

次に、適応説の一つとして筆者らが提唱した環境ゆらぎ説について考えてみよう<sup>2),6)</sup>。この仮説では、あるゲノムの適

応度がその塩基配列全体によって決定される様式は、上記の  
 ような単純な相乗積型選択様式ではないと考える。(実際、既に  
 多くの選択論者によって注意されたように、理論解析の容  
 易化をもたらす魅力を別とすれば、特に相乗積型選択様式を  
 想定すべき経験的根拠は何もない) さらに、上述したように  
 分子進化において同時点に3種類以上の異なるゲノムが集団  
 中に共存していることは、それらのゲノムが他のゲノムと異  
 なって、当面する環境に適応した特別なゲノムであることを  
 意味していると考えられる。これら特別なゲノムの適応度につ  
 いては、アイソガイム間の選択係数が微弱であるらしい事実を  
 も考慮に入れ、互に独立同等な確率過程として良く記述で  
 きると考える。ただし、この際、微弱な自然選択の累積効果  
を保証するような自己相関を確率過程はもっているとする。

これらの基本的見解に立つと、環境ゆらぎ説は等価Kアレ  
 ルモデルを蛋白質多型を説明するモデルとして提出した<sup>2)</sup>。  
 これは、サイズ効果によるのとは異なると、遺伝子頻度のラ  
 ンダム・ドリフトを引き起し、やはり一種の拡散モデルで解  
 析できることが示された<sup>3)</sup>。多型のパターンは突然変異とラ  
 ンダム・ドリフトの釣合で決定されるが、中立説の場合と違  
 って、パラメタ依存性が強くない。多型の経験通則も首尾よ  
 く説明することが分っている<sup>4)</sup>。

では、環境ゆらぎ説の下での分子進化速度はどうなるのか。この場合、従来の理論集団遺伝学で活用されてきた、(1)のような進化速度の固定確率表式は有効な解析方法でなくなる。なぜならば、非相乗積型選択様式の下にある異なったゲノムが3種類以上同時点で共存しつつ進行する進化状況では、固定確率 $\mu$ を計算する簡単な方法がないからである。もちろん、計算機シミュレーションの手法を用いれば、変異体の置換率としての進化速度の性質を調べることは困難ではない。ただし、この際にも、計算機実験で示唆され、発見される性質が、解析的手法によっても確認できれば申し合ない。

生物進化モデルは、上述したような進化速度に関する新しい理論的問題状況を解明するために考案された。このように全く新しい問題では、本質的な要因は保持した上で、可能な限り簡単化した場合をまず調べるのが賢明である。また、最初から特定の場合作りに分析を限定するのは望ましくない。そこで、私たちは、進化速度 $\mu$ が突然変異率 $\mu$ 、自然選択の強さ $s$ 、その継続時間で、集団のサイズ $N$ にどう依存するかを、包括的に調べることを目標とした。現在の所、研究は未だ環境ゆらぎ説の基本見解全部を取り入れた場合に言及できるまでには進んでいないが、進化のありうるパターンについて興味ある結果がいくつか得られていり。以下では、これに関

その研究の一端を紹介しよう。

§3. 生物進化モデル 一般に、ある期間生存して自己複製の単位となるものを広くレプリコン (replicon) と呼ぶことにする。考察の対象とする時間の始まり、すなわち時点  $t=0$  をアダムと呼ぶ。アダム以外のレプリコンは、に生存したレプリコンはすべて唯一つのレプリコンを親として 生存を始め、やがて死ぬ。生き残ったレプリコンは一般に子レプリコンを生む可能性をもつ。各レプリコンはそれぞれ定まったレプリコンは変異体と呼ばれる。上の定義より、各 遺伝的状態をもつ。親と違った状態をもつ子レプリコンから親子関係を通じ祖先に遡れば、それは唯一つのアダムに至る系図をもつことになる。その系図上に現れる変異体の数  $n$  を、当該レプリコンのステップ数 と呼ぶことにする。このようにもななく、アダムのステップ数は 0 である。時点  $t$  に生存するすべてのレプリコンの集団中、ステップ数  $n$  のレプリコンの頻度を  $x_n(t)$  とする。集団の平均ステップ数  $\bar{n}(t) \equiv \sum_{n=0}^{\infty} n x_n(t) \cdots (2)$  の増大率として、進化速度  $v(t) \equiv d\bar{n}(t)/dt \cdots (3)$  が定義される。頻度  $x_n(t)$  の時間変化は

$$dx_n(t)/dt = \{m_n(t) - \bar{m}(t) - \mu\} x_n(t) + \mu x_{n-1}(t) \quad (4)$$

$$x_n(0) = \delta_{n0}, \quad x_{n-1}(t) \equiv 0 \quad (n = 0, 1, 2, \dots)$$

により与えられるとする。ただし、 $m_n(t)$  はステップ数  $n$  のレプリコンの時点  $t$  での純増殖率 (マルサス径数)、正定数  $\mu$  は突然変異率である。  $\bar{m}(t) \equiv \sum_{n=0}^{\infty} m_n(t) x_n(t) \cdots (5)$  は



時点  $t$  での集団の平均マルサス径数である。(4)では集団のサイズ  $N$  が十分大きいとし、サイズ効果を無視してある。サイズ効果の影響は計算機実験で調べるが、これについては §6 で述べる。

生物進化モデルの性質を解析するためには、(4)と等価な線型連立微分方程式

$$dN_n(t)/dt = \{m_n(t) - \mu\} N_n(t) + \mu N_{n-1}(t) \quad (6)$$

$$N_n(0) = f_{n0}, \quad N_{-1}(t) \equiv 0 \quad (n=0, 1, 2, \dots)$$

を考える方が簡単である。(4)と(6)は  $x_n(t) = N_n(t)/N(t) \dots (7)$

により対応づけられる。ただし、ここで、 $N(t) \equiv \sum_{n=0}^{\infty} N_n(t)$

$\dots (8)$  は時点  $t$  に生存するレプリコンの総数である。(7)を

用いると、(6)は  $\bar{n}(t) = M(t)/N(t) \dots (9)$  となる。ただし、

ここで  $M(t) \equiv \sum_{n=0}^{\infty} n N_n(t) \dots (10)$  は集団のステップ総数である。したがって、進化速度の長時間平均

$$\bar{v}_{\infty} \equiv \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t v(t') dt' = \lim_{t \rightarrow \infty} \bar{n}(t)/t \dots (11)$$

を調べるためには、 $N_n(t)$  が(6)に従うときの  $N(t)$ ,  $M(t)$  の  $t \rightarrow \infty$  での漸近挙動を調べれば良い。この際、(6)の解が  $N_0(t) = U_0(t, 0)$ ,  $N_n(t) =$

$$\mu \int_0^t U_n(t, t') N_{n-1}(t') dt' \quad (n=1, 2, 3, \dots) \dots (12)$$

の形に与えられることに注意しよう。ただし、

$$U_n(t, t') \equiv \exp \left\{ \int_{t'}^t \{m_n(t'') - \mu\} dt'' \right\} \quad (n=0, 1, 2, \dots) \quad (13)$$

である。

§ 4. 進化モデルの厳密な結果 進化モデルの性質は、自然選択の様式を表わす  $\{m_n(t), n=0, 1, 2, \dots\}$  と突然変異率  $\mu$  に依存する。以下に、厳密かつ興味ある結果が得られた2つの場合を紹介しよう。

A. 定数マルサス径数モデル これは、 $m_n(t)$  が  $n$  には依ってもよいが、時点  $t$  には依らない定数  $m_n$  である場合である。特に、 $\{m_n\}$  が最初の有限個を除いて周期列である場合、すなわち、ある自然数  $n_0$  と  $\nu$  が存在して  $m_{n+\nu} = m_n$  ( $n = n_0, n_0+1, n_0+2, \dots$ )  $\dots$  (14) である場合には、 $\{m_n\}$  の上限を  $\hat{m}$  とし、 $\mu_c \equiv \exp \left[ \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\alpha=0}^n \log(\hat{m} - m_\alpha + \varepsilon) \right] \dots$  (15) とすると、 $\mu_c$  は臨界突然変異率であって、 $\mu < \mu_c$  ならば、 $v_\infty = 0$ 。一方  $\mu > \mu_c$  ならば  $v_\infty = \left[ \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\alpha=0}^n 1/(P^* + \mu - m_\alpha) \right]^{-1} \leq \mu \dots$  (16) が成り立つ。ただし、 $P^*$  は  $\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{\alpha=0}^n \log(P^* + \mu - m_\alpha) = \log \mu \dots$  (17) の  $[\hat{m} - \mu, \hat{m}]$  内の実根である。表式(15)~(17)は非周期列  $\{m_n\}$  に対しても意味をもつのであるが、上の結果はもっと一般の定数マルサス径数モデルについても成り立つと予想される。この予想の下に、 $\{m_n\}$  の列の分布が  $[-s, s]$  上の一様分布で  $\hat{m} = s$  であるときの臨界突然変異率を(15)により求めると、 $\mu_c = 2s/e \dots$  (18) となり、 $v_\infty/\mu$  は図1に見られるように、 $\mu/s$  の単調増加関数である。

B. 周期的マルサス径数モデル      これは、マルサス径数

$$m_n(t) \text{ が } m_n(t) = (-1)^{n+[\cdot]t/\tau} \cdot s \quad (n=0, 1, 2, \dots, t \geq 0) \quad (19)$$

で与えられる場合である。s,  $\tau$  は正定数で、自然選択の強さと継続時間を表わす。[ $\cdot$ ] は Gauss 記号で、 $\cdot$  を越えない最大整数を表わす。

$N(t), M(t)$  の時間変化を調べるため、母関数  $g(x, t) \equiv \sum_{n=0}^{\infty} x^n N_n(t) \cdots (20)$  を導入すると、この時間変化は(16)より

$$\partial g(x, t) / \partial t = (-1)^{[t/\tau]} s g(-x, t) + \mu (x-1) g(x, t) \quad (21)$$

となる。 $g(\pm x, t)$  について閉じた方程式(21)を解いて  $g(x, t)$  を求めれば、 $N(t) = g(1, t), M(t) = g_x(1, t)$  の  $t \rightarrow \infty$  での漸近挙動を調べることもができる。

平均進化速度  $v_{\infty}$  は、 $\mu^* \equiv \mu/s, v^* \equiv s\tau$  とし?

$$v^* \equiv v_{\infty}/s = \mu^* (C\mu^{*2} + S/\alpha) / \sqrt{(1+\mu^{*2})(1+C^2\mu^{*2})} \quad (22)$$

によって与えられる。ただし、 $C \equiv \cosh \alpha, S \equiv \sinh \alpha,$

$\alpha \equiv \tau^* \sqrt{1+\mu^{*2}}$  である。次に、集団の平均マルサス径数  $\bar{m}(t)$

の長時間平均  $\bar{m}_{\infty}$  は  $\bar{m}^* \equiv \bar{m}_{\infty}/s = \frac{1}{\tau^*} \log \left[ (C S \mu^* + \sqrt{1+C^2\mu^{*2}}) / \sqrt{(1+\mu^{*2})(1+C^2\mu^{*2})} \right] \cdots (23)$  によって与えられる。図2は、

これらの結果のパラメタ依存性を等高線図として表わしている。

§5. 揺動マルサス径数モデル      これは、マルサス径数

$m_n(t)$  が定常確率過程であった、 $n$  が異なるものは互に独立

かつ同等である場合である。これは揺動環境下での自然選択のモデル化である。中でも特に興味ある環境変動は、ある平均継続時間をもちて揺動するようなものである。

今の場合、 $\{m_n(t)\}$ が確率過程であるから、進化速度 $v_\infty$ も確率変数となる。その平均 $\langle v_\infty \rangle$ などを求めることは、残念ながら未だできていない。しかし、 $\langle v_\infty \rangle$ の近似としての  $\tilde{v}_\infty \equiv \lim_{t \rightarrow \infty} \langle M(t) \rangle / t \langle N(t) \rangle \dots (24)$  の値を求めることは割合にやさしい。このためには $\langle N_n(t) \rangle$ の時間変化を調べれば済むからである。その結果は  $\tilde{v}_\infty = -1/\mu \langle \tilde{U}'_0(P^*) \rangle \leq \mu \dots (25)$  である。ただし、ここで $P^*$ は

$1/\mu = \langle \tilde{U}_0(P^*) \rangle \equiv \int_0^\infty e^{-Pt} \langle U_0(t, 0) \rangle dt \dots (26)$  の実数部最大根であり、 $\langle \tilde{U}'_0(P^*) \rangle$ は $\langle \tilde{U}_0(P^*) \rangle$ の $P^*$ に関する導関数である。

(13)を用いると、 $\langle U_0(t, 0) \rangle$ のラプラス変換 $\langle \tilde{U}_0(P) \rangle$ は  $\langle \tilde{U}_0(P) \rangle = \langle u(m_0) \rangle$ ,  $u(m) \equiv \langle \int_0^\infty e^{-Pt} \exp \left[ \int_0^t \{m_0(t') - \mu\} dt' \right] dt \rangle_{m_0(0)=m} \dots (27)$  の形に表わされる。ただし、ここで  $\langle \cdot \rangle_{m_0(0)=m}$  は条件 $m_0(0)=m$ の下での $\cdot$ の平均である。 $M_n(t)$ がマルコフ過程である場合には、条件付平均 $u(m)$ を求めるために、ファインマン-カッツの公式  $(A+m-\mu-P)u(m) = -1 \dots (28)$  を利用することができる。ただし、ここで $A$ はマルコフ過程 $m_0(t)$ の生成作用素である。例えば、マルサス経数

$M_n(t)$  が  $\pm S$  の値をとる定常マルコフ過程であった。

$M_n(0) = m$  から出発したとき  $M_n(t) = m'$  となる推移確率

$P(m, m', t)$  が

$$P(m, m', t) = e^{-t/\tau} \delta_{mm'} + \frac{1}{2} (1 - e^{-t/\tau}) \quad (29)$$

である場合に、上の手順で  $\tilde{v}_\infty$  を求めると、 $Se \equiv (1 + 1/\mu\tau)^{-1} S$

として

$$\tilde{v}_\infty = \frac{1}{2} \mu (1 + 1/\sqrt{1 + 4S_0^2}) \leq \mu \quad (30)$$

の結果が得られる (図3)。

### § 6. Wright-Fisher 型進化モデル 以上 §§ 3 ~ 5 で調

べた進化モデルでは、サイズ効果を完全に無視した。これを取り入れるためには、例えば、(4)の右辺に適切な揺動項を追加してもよい。しかし、そのような確率微分方程式系による研究によっても、進化速度について未だ見るべき結果が得られていない。ここからは、サイズ効果を取り入れる今一つの方法を紹介しよう。これは、計算機実験を行って速度  $v_\infty$  のパラメータ依存性を調べるのに適した、離散時間モデルである。

これは、 $\mathcal{G} \equiv \{(N_0, N_1, N_2, \dots); N_n \in \mathbb{N} (n=0, 1, 2, \dots)\}$ ,

$\sum_{n=0}^{\infty} N_n = N$  } 上のマルコフ連鎖  $X_t$  で、推移確率が

$$P\{X_{t+1} = (N'_0, N'_1, N'_2, \dots) \mid X_t = (N_0, N_1, N_2, \dots)\} = N! \prod_{n=0}^{\infty} \{z_n(t)^{N'_n} / N'_n!\} \quad (31)$$

で与えられるものである。ただし、

$$z_n(t) \equiv w_n(t) y_n / \sum_{n=0}^{\infty} w_n(t) y_n, \quad y_n \equiv \mu x_{n-1} + (1-\mu)x_n, \quad x_n \equiv N_n/N$$

$$x_{n+1} \equiv 0 \quad (n=0, 1, 2, \dots) \quad (32)$$

であつて、モデルは突然変異率  $\mu$ 、適応度の時系列  $\{w_n(t); n, t \in \mathbb{N}\}$  および集団のサイズ  $N$  によつて特徴づけられる ( $N \rightarrow \infty$  の極限では、(31) の確率的变化は消失し、(32) で  $x_n = z_n(t-1)$  としての決定論的過程になる。この定差方程式系の連続時間版が、(4) の微分方程式系である。後者を特徴づけるマルサス径数  $\{m_n(t)\}$  と  $\{w_n(t)\}$  とを、 $w_n(t) = \exp\{m_n(t)\}$  によつて対応させると、自然選択、突然変異による一世代間の変化が小さい場合には、離散時間過程と連続時間過程の挙動は一致するようになる)。

厳密解が得られたモデル A, B では、サイズ効果と、 $m_n(t)$  の時間変化が一般的に非周期的に起るための効果、および  $m_n(t)$  のとる値が 2 値に限らない等の効果は全く無視されている。これらの効果により、上で得られた結果がどの程度質的影響を受けるかについては、上で紹介した W-F モデルに対して計算機実験を行うことにより、系統的に調べることができる。

図 4 は § 4 の周期的マルサス径数モデルに対してサイズ効果を取り入れた結果である。これを図 2-b と比較すると、サイズ効果は有効集団サイズ  $N$  と  $\mu$  の積が 1 よりかなり小さくならない限り、少くとも  $v_\infty/\mu$  に対しては影響はない

ようである。

図5は§5の揺動マルサス径数モデルに対するサイズ効果を取り入れた結果である。サイズ効果が無視できる場合の厳密な結果が分っていないので、周期的マルサス径数モデルの場合程確定的に言えないが、それぞれ  $N\mu = \frac{1}{4}$ , 4の場合に対する図5-a, 5-bを比較すると、このモデルでも  $\tilde{v}_\infty/\mu$  に対するサイズ効果は、 $N\mu$  が1より大きければ、さ程でないようである。さらに、図2-bと比較すると、周期的マルサス径数モデルと揺動マルサス径数モデルは、進化速度のパラメータ依存性について似通った点が多いことがわかる。最後に、図5と図3を比較すると、せっかく解析的に求められはしたが、§5の  $\tilde{v}_\infty$  は、残念ながら、真の  $\langle v_\infty \rangle$  とかなり相違しているように思われる。

§7. おわりに 上の紹介で明らかのように、生物進化モデルの数理については、解明すべき問題が多い。純粹に数学的性質に限らずとも、例えば、(1)定数マルサス径数モデルに対する表式(5)-(7)型性質の妥当する径数列  $\{m_n\}$  の範囲は何か。(2)揺動マルサス径数モデルの  $\langle v_\infty \rangle$  をいかに解析するか、がある。後者に関しては、揺動淘汰による確率論的変動を拡散過程として近似する着想があるが、<sup>(5), (6)</sup> 非可逆的な突然変異を伴う進化モデルに対する有効性については、前節の W-F

モデルの計算機実験を通して、批判的検討が進められている。

次に、上で紹介した進化モデルの研究結果の生物学的意義について簡単に述べよう。サイズ効果を無視した進化モデルを分子進化に適用する際、分子の機能上の制約の強さの大小はパラメタ  $s$ 、この大小に対応すると考えよう。図 1 より  $v^* \equiv v_\infty / \mu$  は  $s$  の減少関数であり、図 2-b から分かるように、 $v^* < 1$  なるパラメタ領域では、 $v^*$  は  $s$  の減少関数であり、いずれも分子進化の通則 (ii) を与える。ただし、 $v^* > 1$  なる領域では、 $v^*$  は  $s$  の増加関数であり、 $v^*$  については最大値が存在し、通則 (ii) を与えない。

それでは、現実の分子進化が、通則 (ii) の成り立つような  $v^* < 1$  なるパラメタ領域に対応する理由があるであろうか。図 2-b と図 2-c を比較すれば分かるように、 $s$ 、 $\tau$  を固定して  $\mu$  を変化させると、 $v^* < 1$  なる領域では  $\partial \bar{m}_\infty / \partial \mu < 0$ 、 $v^* > 1$  なる領域では  $\partial \bar{m}_\infty / \partial \mu > 0$  となる。図 2 は周期的マルサス径数モデルの結果であるが、モデル (4) において  $\bar{m}_n(t)$  が  $\mu$  に依存しないとする限り、全く一般的に

$$\partial \bar{m}(t) / \partial \mu = v(t) / \mu - 1 \quad (33)$$

の関係が成り立つことが証明される。

実際には、突然変異率を下げるためには自己複製に費やす



時間と自由エネルギーを多く要するであろうから、

$\partial m_n(t)/\partial \mu = c(\mu) > 0$  ( $c(\mu)$ は $\mu$ だけの関数)と仮定すると、

(33)は

$$\partial \bar{m}(t)/\partial \mu = v(t)/\mu - 1 + c(\mu) \quad (34)$$

となる。一方、突然変異率の大小は遺伝的形質であろうから、突然変異率に対する自然選択の結果、 $\mu$ は $\partial \bar{m}_\infty/\partial \mu = 0$ をみたすように最適化されているとすると、(34)の長時間平均より

$$v^* = 1 - c(\mu) < 1 \quad (35)$$

となる。以上で論じてきたのは、生物進化モデルの中でも特別に簡単化された場合の性質であるが、これだけからでも既に、分子進化速度の通則(ii)は、中立説だけでなく、分子レベルに働く自然選択にある種の定常な変動を仮定しさえすれば導出できることが示唆されていると思われる。

§5で与えた揺動マルサス径数モデルの $\tilde{v}_\infty$ については、本来求めるべき $\langle v_\infty \rangle$ の近似に過ぎないというところで、上の議論では言及を控えた。ただ、最後に一言、§5で求めた $\tilde{v}_\infty$ は以下に述べるモデルの真の $\langle v_\infty \rangle$ を与えることを注意しておこう。そのモデルでは、ある遺伝的状態 $\sigma$ に突然変異が起ると、これまで存在しなかった全く新しい状態 $\sigma'$ が生じると考え、 $\sigma$ 状態レガリコンの時点 $t$ でのマルサス径数 $m_\sigma(t)$ が、 $\sigma$ が異なると互に独立・同等な、定常確率過程であると

仮定するのである。このモデルに対しても、(4)が形式的に妥当であるが、実質的な時間変化を解析するには、もちろん、時点  $t$  での  $\sigma$  状態レプリコンの頻度  $x_{\sigma}(t)$  に対する方程式から出発しなければならない。ここで、 $\tilde{v}_{\infty}$  に関する上述の注意も、 $N_{\sigma}(t)$  と  $N_n(t)$  の関係の検討から示されるのである。

なお、生物進化モデルについては、現在までの所下記<sup>(5), (6), (8), (9)</sup>に詳しい報告が行われている。

#### 文 献

- 1) M. Kimura. Nature 217 (1968), 624.
- 2) H. Matsuda and T. Gojobori. Adv. Biophys. 12 (1979), 53.
- 3) 木村資生, サイエンス 10 (1980), 32.
- 4) J. F. Crow and M. Kimura. An Introduction to Population Genetics Theory. Harper & Row, New York, 1970.
- 5) 福島正俊, 石井一成, 自然現象と確率過程, 入門 | 現代の数学 (10). 日本評論社 (1980).
- 6) 松田博嗣, 石井一成. 生物集団と進化の数理. 岩波書店 (1980).
- 7) N. Takahata, K. Ishii and H. Matsuda, Proc. Natl. Acad. Sci. USA 72 (1975) 4541-4545.
- 8) K. Ishii, H. Matsuda and N. Ogita. Proc. International Symposium on

Mathematical Topics in Biology, Kyoto, Sept. 1978, 24-33.

- 9) K. Ishii, H. Matsuda and N. Ogita. A Mathematical Model of Biological Evolution. J. Math. Biol. 投稿中

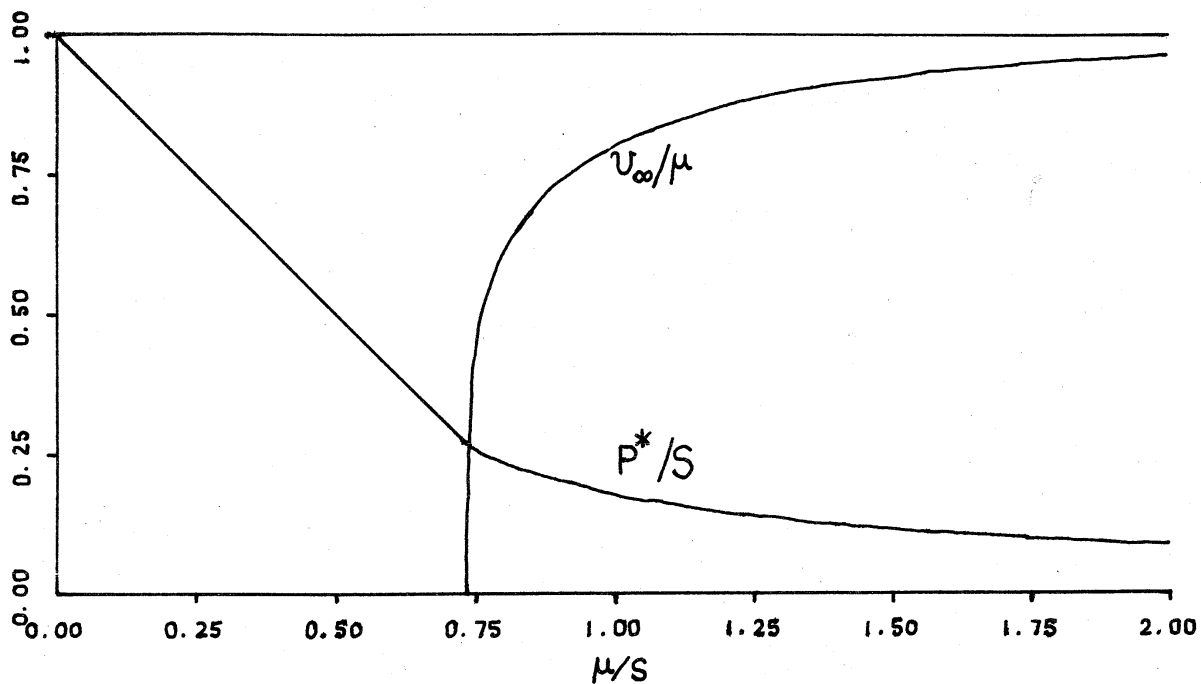


図1 区間 $[-s, s]$ 上に一様分布する定数マルサス径数モデルの平均進化速度  $v_{\infty}$  と  $P^*$

蛋白質	$10^9$ 年当りアミノ酸座位当り置換率
フィブリノゲン	9.0
ヘモグロビン	1.4
ミオグロビン	1.1
チトクロムc	0.21
ヒストンIV	0.009

表1 蛋白質の進化速度

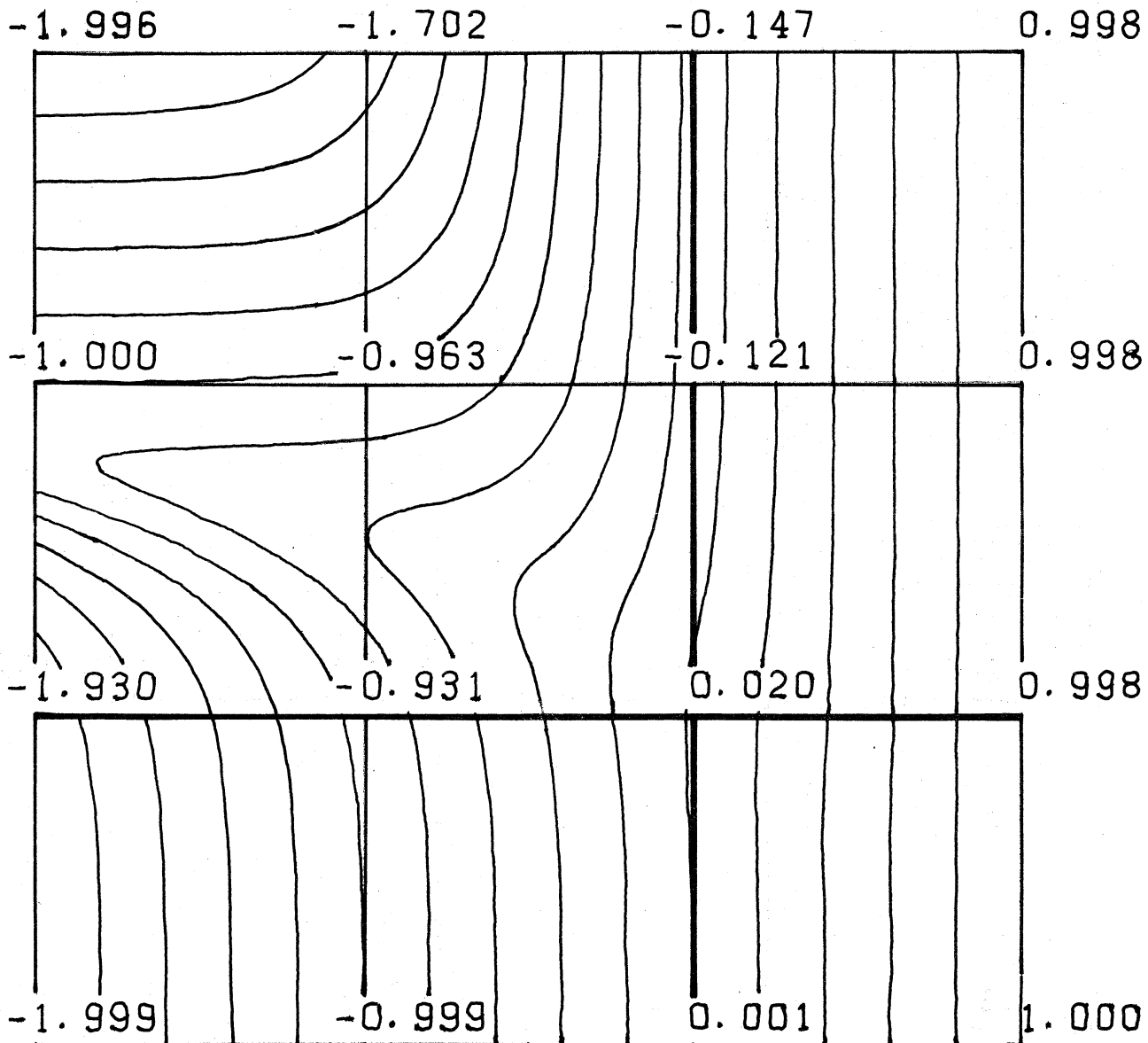


図 2-a 周期的マルサス径数モデルの  $\log_{10} v^*$  の等高線図。等高線間隔は 0.2。パラメータ領域の横軸は  $\log_{10} \mu^*$ 、縦軸は  $\log_{10} z^*$ 。図は、 $-2 \leq \log_{10} \mu^* \leq 1$ 、 $-1 \leq \log_{10} z^* \leq 2$  なる領域に対するもの。座標軸は太線で表示。

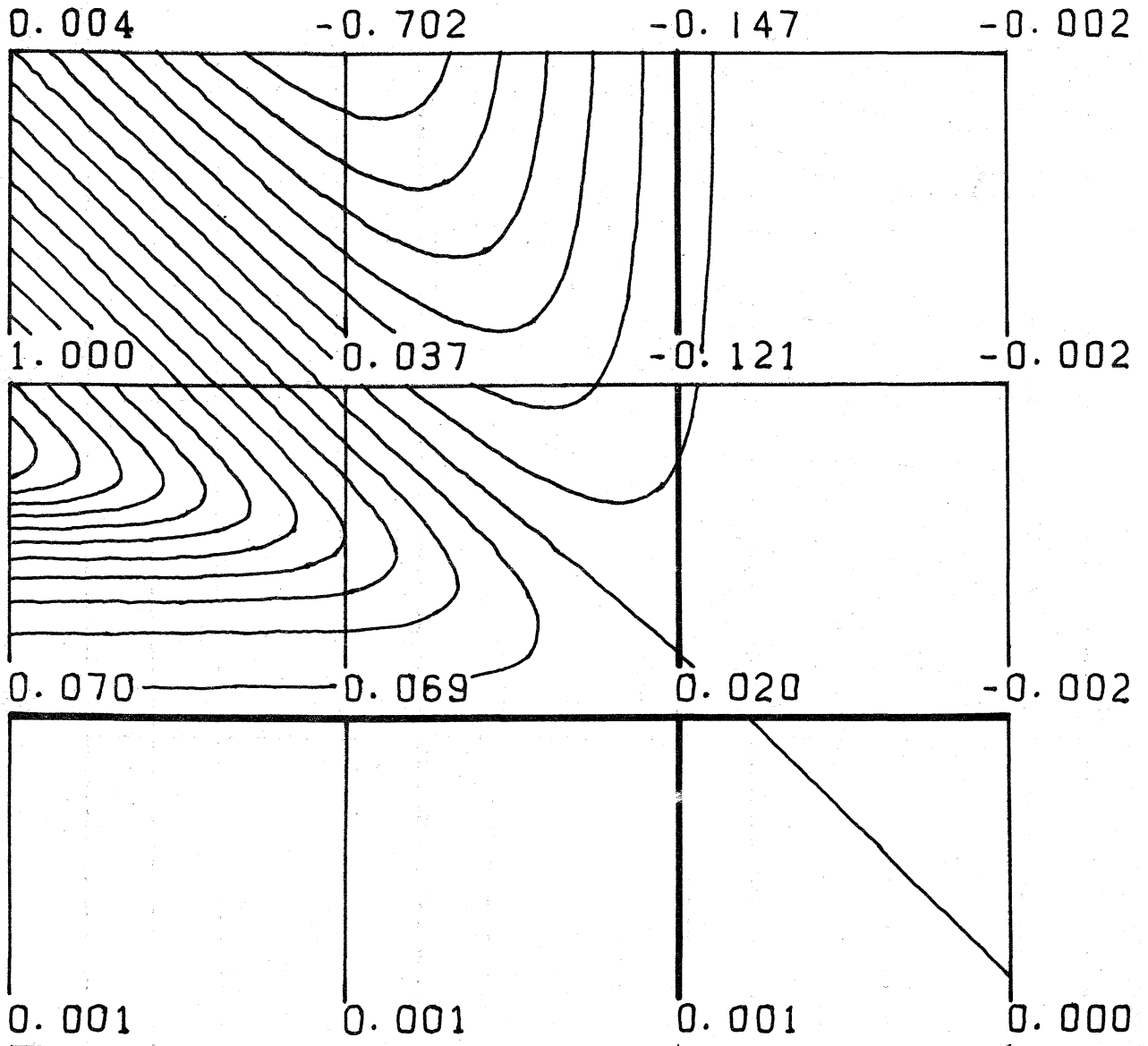


図2-b 周期的マルサス径数モデルの  $\log_{10}(v_0/\mu)$  の等高線図.

等高線間隔は0.1. パラメタ領域は図2-aと同じ.

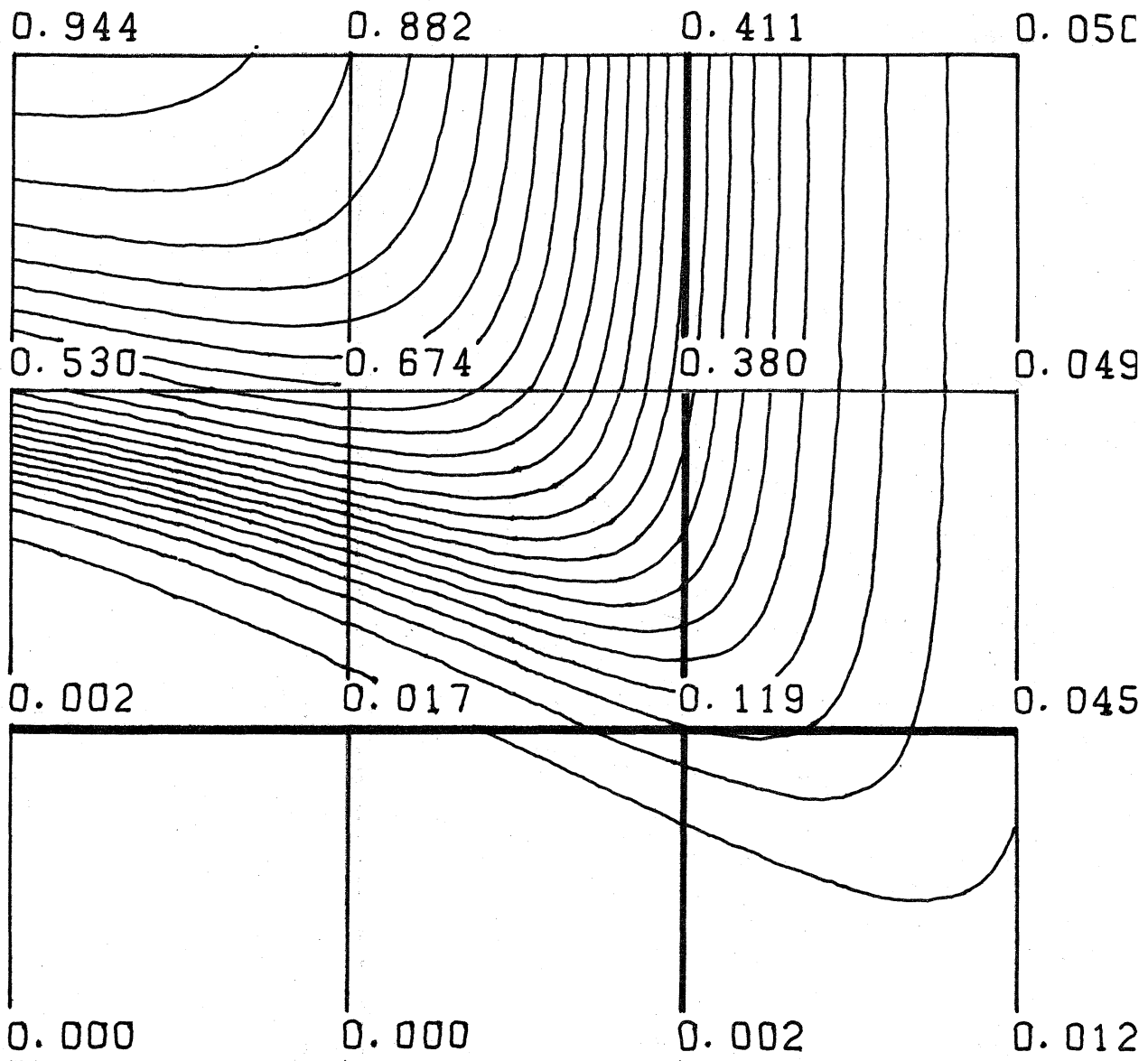


図 2-c 周期的マルサス径数モデルの  $\bar{m}^* \equiv \bar{m}_0/s$  の等高線図。

等高線間隔は 0.04. パラメタ領域は図 2-a と同じ。

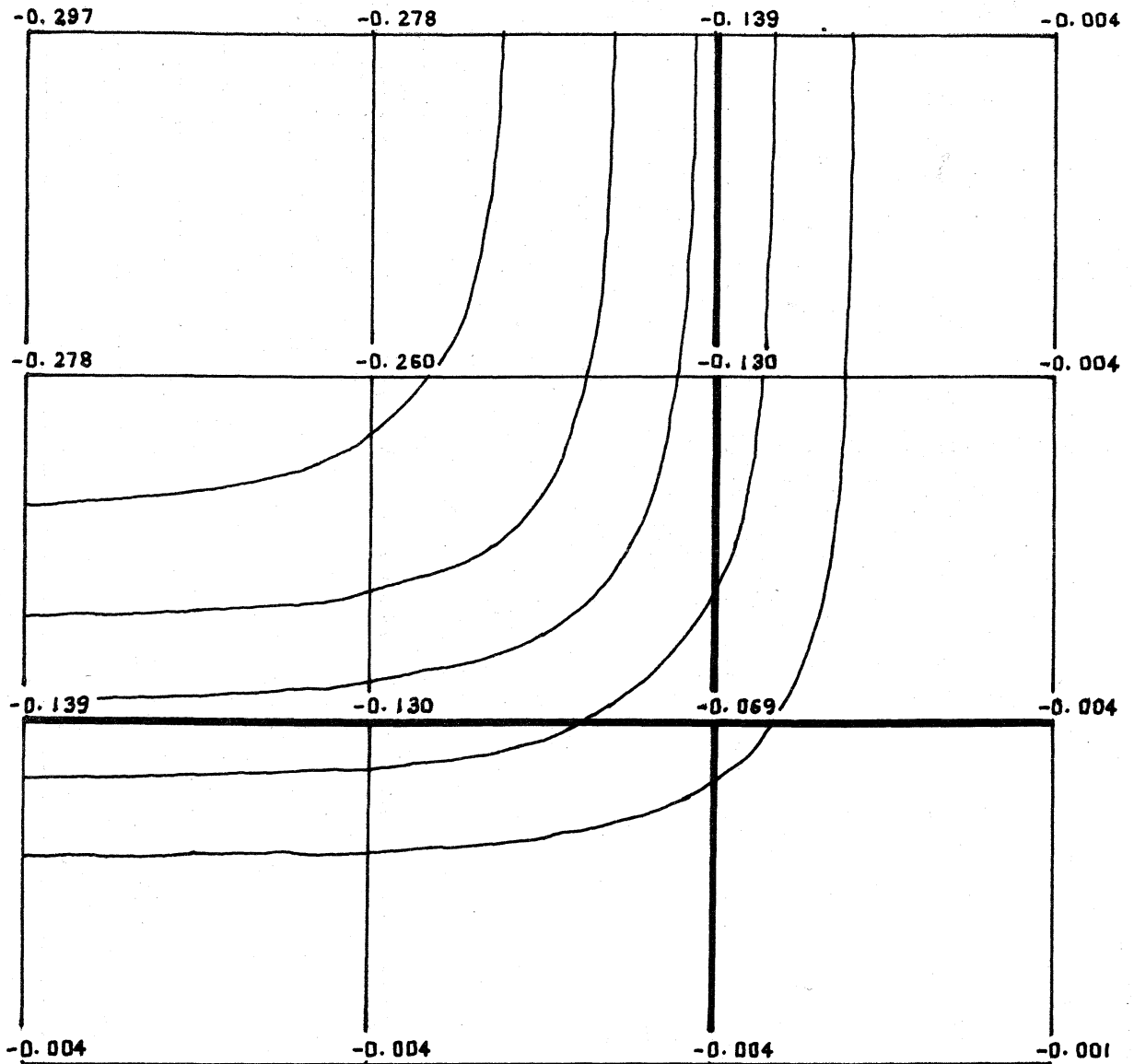


図3 揺動マルサス径数モデルの  $\log_{10}(\bar{v}_{\infty}/\mu)$  の等高線図 (29) の推移確率をもったマルサス径数の場合)。等高線間隔は0.05。パラメタ領域は図2と同じ。



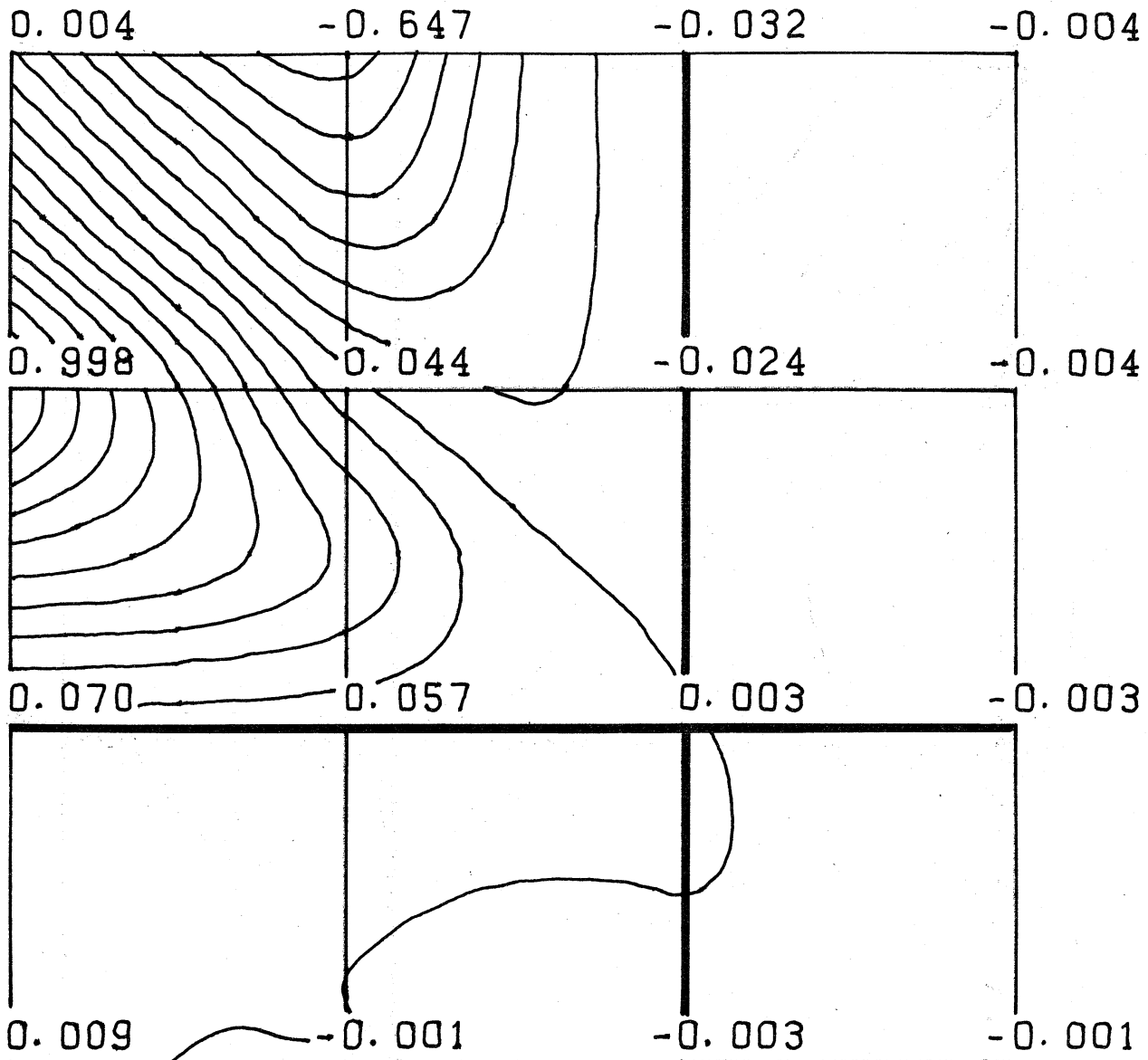


図4 サイズ効果を取り入れた周期的マルサス径数モデルの  $\log_{10}(v_0/\mu)$  の等高線図 ( $N\mu = 1/4$  の場合)。等高線間隔は 0.1。パラメタ領域は図2と同じ。領域上の  $7 \times 7$  のパラメタ点での計算機実験の結果に基づいた図。

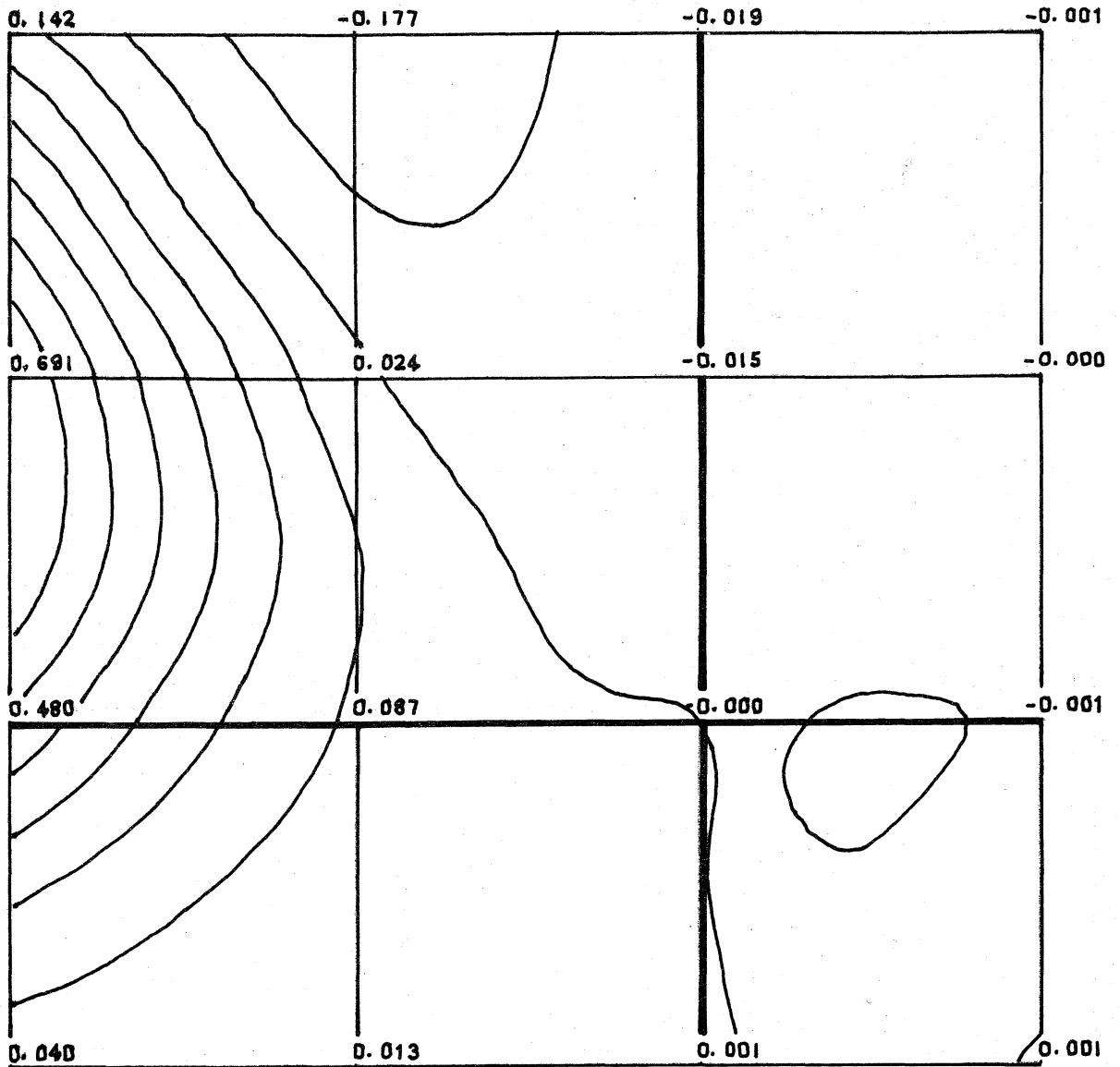


図5-a サイズ効果を取り入れた揺動マルサス径数モデルの  $\log_{10}(v_0/\mu)$  の等高線図 ( $N\mu = 1/4$  の場合)。等高線間隔は0.1。パラメタ領域は図2と同じ。領域上の7×7のパラメタ点での計算機実験の結果に基づいた図。

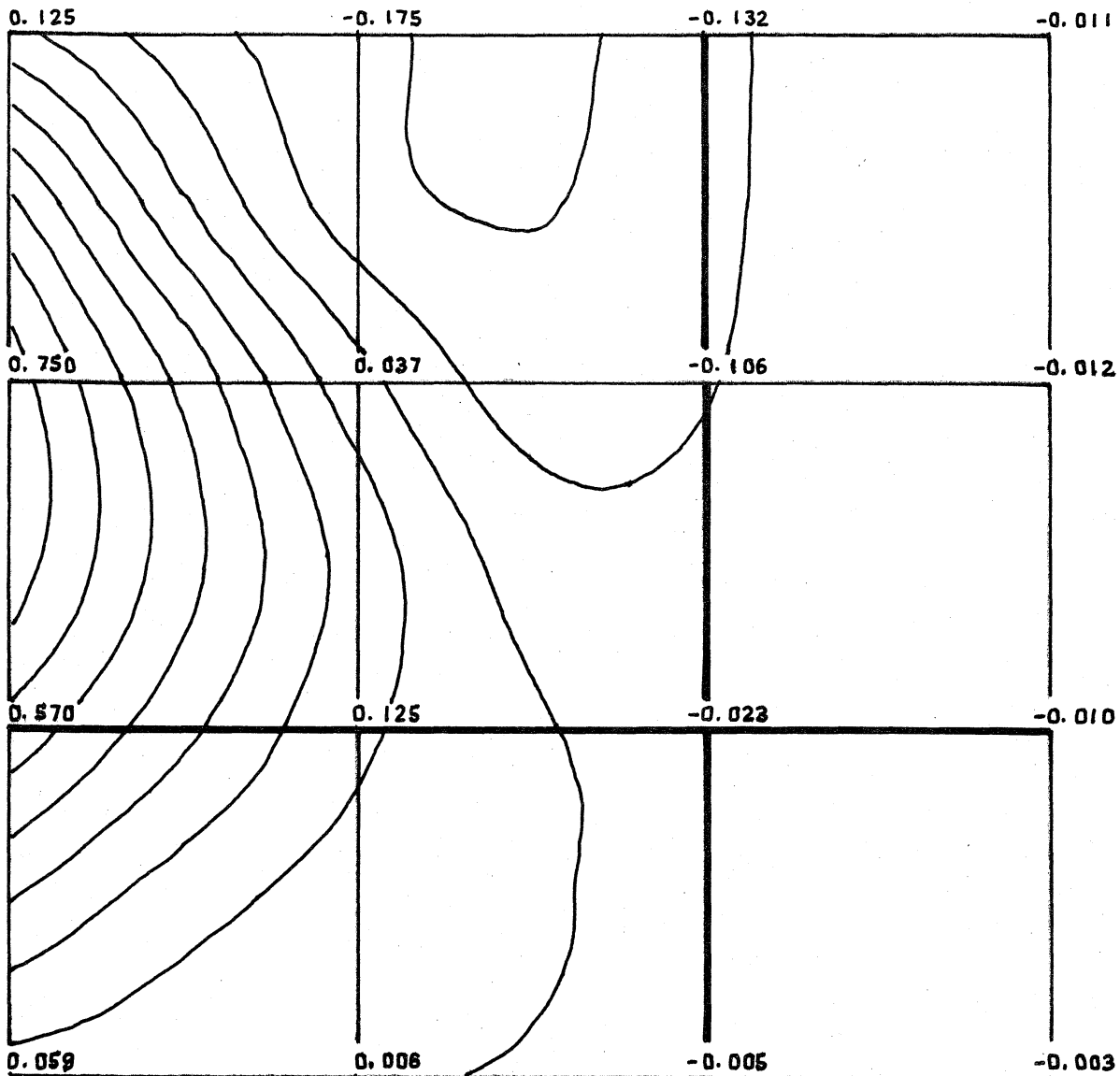


図5-6 サイズ効果を取り入れた揺動マルサス径数モデルの  $\log_{10}(u_0/\mu)$  の等高線図 ( $N\mu=4$  の場合)。等高線間隔は 0.1。パラメータ領域は図2と同じ。領域上の  $7 \times 7$  のパラメータ点での計算機実験の結果に基づいた図。