

A mathematical model of biological evolution

九大 理 石井一成

松田博嗣

理研 萩田直史

§0. はじめに. 最近の分子生物学の進歩には著しいものがあるが, 進化に関しても相同タンパク質やDNAの比較生物学的研究が行なわれ, 進化速度の定量的データが過去10数年間に蓄積されてきた。その結果によると, 相同タンパク質の進化は年あたりほぼ一定の速さで起ったと考えられる。この速さは, 相同タンパク質の種類毎に固有のものであるが, 同一の相同タンパク質については系統樹のどの経路を考へるかに依らず一定であるとされている(実験データの解説については, 例えは[1]を参照)。

進化速度の決定機構として, ネオ・ダーウィニズムは環境の変動を重視し, 進化速度は環境の持続時間の逆数に等しいと考へている。他方, タンパク質やDNAなどの分子レベルでは自然淘汰が働かないとする分子進化の中立説によれば,

進化速度は突然変異率に等しい(中立説の最近の発展については, [1] を参照)。しかし, これまで進化速度の理論集団遺伝学的研究は極めて不十分で, 例えは, 変動環境下での自然淘汰の影響が統一的に研究されたことは殆んどなかった。このため, せうかく分子進化速度の定量的データを手にしなから, その集団遺伝学的決定機構に関して現在の所決定的なことは何も言えない状態に留まっている。

分子進化の一つの側面である集団の多様性については, 電気泳動法によって得られた実験データが持続型揺動淘汰モデルによって首尾よく理解できることが分っている(分子進化の環境ゆらぎ説については, [2] を参照)。この環境ゆらぎ説の下での分子進化速度を調べようとした所, 従来用いられてきた手法(固定確率理論)では十分正しく扱えないことが分った。

そこで, 我々は進化速度 v が突然変異率 μ , 揺動淘汰の変動幅 s , その持続時間 τ , として集団の有効サイズ N にどのように依存するかを統一的に調べることを目的にし, 理想化されたモデルを考え, 研究することにした。前回のシンポジウムでの報告([3])に引き続き, その後得られた結果を以下で紹介したい。ここでの紹介では, モデルの説明とその数学的挙動についての結果に重点を置き, 得られた結果の生物学

的含義の討論については、他の機会に譲ることにしたい。

§ 1. 進化のモデル. 我々の考えたモデルは Wright-Fisher

モデルの特別な場合である。すなわち、 $\mathcal{G} \equiv \{(N_0, N_1, N_2, \dots)\}$;

$N_n \in \mathbb{N}$ ($n=0, 1, 2, \dots$), $\sum_{n=0}^{\infty} N_n = N$ } 上の離散時間マルコフ過程

X_t で、推移確率が

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

で与えられるものを考える。ただし、

$$z_n(t) \equiv \omega_n(t) y_n / \sum_{n=0}^{\infty} \omega_n(t) y_n \quad (2)$$

$$y_n \equiv \mu x_{n-1} + (1-\mu) x_n \quad (3)$$

$$x_n \equiv N_n / N \quad (n=0, 1, 2, \dots), \quad x_{-1} \equiv 0 \quad (4)$$

である。モデルは突然変異率 μ , 適応度の時系列 $\{\omega_n(t); n, t \in \mathbb{N}\}$ および集団のサイズ N によつて特徴づけられる。

n は遺伝子のスリッパ数と呼ばれ、(3) から分るように、 $t=0$ 以来来た突然変異の回数を表わしている。集団の平均スリッパ数

\bar{n} は

$$\bar{n} \equiv \sum_{n=0}^{\infty} n N_n / N \quad (5)$$

の時点 t における期待値を $\bar{n}(t)$ と表わすとき、このモデルの

進化速度 v は

$$v \equiv \lim_{t \rightarrow \infty} \bar{n}(t) / t \quad (6)$$

と定義される ($t=0$ での初期分布は $X_0 = (N, 0, 0, \dots)$ とし

おく)。

上のモデルでサイズ効果がない ($N \rightarrow \infty$) の場合は, (2)~(4) で $x_n = z_n(t-1)$ とした決定論的過程により記述される。解析の便利さのため, 更に, このモデルの連続時間版として

$$dx_n(t)/dt = \{m_n(t) - \bar{m}(t) - \mu\} x_n(t) + \mu x_{n-1}(t) \quad (7)$$

を考えよう。ここでヌルサス径数 $m_n(t)$ は適応度 $w_n(t)$ に対応し, $\bar{m}(t)$ は集団の平均ヌルサス径数

$$\bar{m}(t) \equiv \sum_{n=0}^{\infty} m_n(t) x_n(t) \quad (8)$$

である。(7)は変数変換

$$x_n(t) \equiv N_n(t) / N(t) \quad (9)$$

$$N(t) \equiv \sum_{n=0}^{\infty} N_n(t) \quad (10)$$

により, 2線型常微分方程式系

$$dN_n(t)/dt = \{m_n(t) - \mu\} N_n(t) + \mu N_{n-1}(t) \quad (11)$$

と等価である。この連続時間モデルでは

$$\bar{n}(t) \equiv M(t) / N(t) \quad (12)$$

$$M(t) \equiv \sum_{n=0}^{\infty} n N_n(t) \quad (13)$$

が集団の平均ヌルサス径数を与え, 進化速度 σ はこれを用いた (6)により定義される。

§ 2. 固定確率理論 (従来の方法). 一般の Wright-Fisher

モデルを考えよう。遺伝子の状態をある時点 (例えば $t=0$ と

しよう) に存在する N 個の遺伝子のどれの子孫 (突然変異を伴う複製) であるかによつて分類することになると、確率 1 で有限時間の後に集団はどれか 1 個の遺伝子の子孫のみになることが分る。それが第 i 番目の遺伝子の子孫である確率を u_i と表わすと、どの遺伝子の子孫が固定するかはお互いに背反事象であるから

$$\sum_{i=1}^N u_i = 1 \quad (14)$$

である。さて、固定確率理論によれば進化速度 v は

$$v = \sum_{i=1}^N \mu \cdot u_i' = \mu \sum_{i=1}^N u_i' \quad (15)$$

と表わされる。ここで μ は一世代あたりの突然変異率、 u_i' は第 i 番目の遺伝子に $t=0$ から $t=1$ の間に突然変異が起った結果生じる遺伝子の固定確率である。(15) は

$$v = N\mu u, \quad u \equiv \frac{\sum_{i=1}^N u_i'}{N} \quad (16)$$

と表わされる。ここで、 u は突然変異を起った各遺伝子の平均固定確率である。

中立突然変異の場合、変異が起ったことは遺伝子の運命に影響を及ぼさないから

$$u_i' = u_i \quad (17)$$

である。したがって、(14)、(16) より

$$u = N^{-1}, \quad v = \mu \quad (18)$$

となる。

ある遺伝子に突然変異が起る。それからそれが固定するまでの間に集団中に突然変異が起こらない場合には、自然淘汰が働いている場合に対しても、平均固定確率 u の計算を拡散過程モデルを用いて行なうことが可能である (2Pレルモデル)。しかし、上の条件が破れると、一般に同時点の集団中に3種類以上のPレルが存在することになり、自然淘汰が働いている場合の平均固定確率 u の計算は、拡散過程モデルを用いても非常に困難になる。

§3. 定数適応度モデル. サイズ効果のない連続時間モデル (7), (11) について調べよう。本節では、 $m_n(t)$ が n には依存しても良いが t には依存しない定数 m_n である場合を考える。系列 $\{m_n; n=0, 1, 2, \dots\}$ は上に有界

$$M \equiv \sup_n m_n < \infty \quad (19)$$

で、かつ一意な極限分布

$$p([a, b]) \equiv \lim_{n \rightarrow \infty} \# \{m_\alpha; a \leq m_\alpha \leq b, \alpha=0, 1, 2, \dots, n\} / (n+1) \quad (20)$$

をもつと仮定しよう。またサイズ系列の特性関数 $\lambda(p)$ を

$$\lambda(p) \equiv \int_{-\infty}^{\hat{m}} \log(p-m) p(dm) \quad (p > m) \quad (21)$$

によって定義し、臨界突然変異率 μ_c を

$$\mu_c \equiv \exp\{\lambda(\hat{m}+0)\} \quad (22)$$

と定義しよう。すると、進化速度 v についての2命題が証

明される。

命題1. マルカス径数列 $\{m_n; n=0, 1, 2, \dots\}$ が本質的に週期的, すなわち, 2つの自然数 n_0, ν が存在して

$$m_{n+\nu} = m_n \quad (n = n_0, n_0+1, n_0+2, \dots) \quad (23)$$

である場合, 突然変異率 μ と臨界突然変異率 μ_c との大小関係に応じて, 次の3つの性質がある。

(I) $\mu > \mu_c \geq 0$ の場合. (10) の $N(t)$, (13) の $M(t)$ は

$$N(t) \sim A e^{p^* t}, \quad M(t) \sim A \nu t e^{p^* t} \quad (t \rightarrow \infty) \quad (24)$$

なる漸近挙動を示す. ここで, A は正数で, p^* は

$$\lambda(p^* + \mu) = \log \mu \quad (25)$$

の区間 $(\hat{m} - \mu, \hat{m})$ 中に一意に存在する根である. ν は進化速度で,

$$\nu = 1/\lambda'(p^* + \mu) = \left\{ \int_{-\infty}^{\hat{m}} \varphi(dm) / (p^* + \mu - m) \right\}^{-1} \quad (26)$$

により与えられる. この ν は不等式

$$0 < \nu \leq \mu \quad (27)$$

を満足する。

(II) $\mu = \mu_c > 0$ の場合.

$$N(t) \sim A t^{\hat{n}} e^{(\hat{m} - \mu)t}, \quad M(t) \sim A \nu t^{\hat{n}+1} e^{(\hat{m} - \mu)t} \quad (t \rightarrow \infty). \quad (28)$$

ただし, A は正数, \hat{n} は \hat{m} に等しい m_n の個数である. 進化速度 ν は

$$\nu = 1/(\hat{n}+1)\lambda'(\hat{m}+0) = \left\{ (\hat{n}+1) \int_{-\infty}^{\hat{m}} \varphi(dm) / (\hat{m} - m) \right\}^{-1} \quad (29)$$

によ、 z とえられ、不等式

$$0 < v \leq \mu / (\hat{n} + 1) \quad (30)$$

を満足する。

(Ⅳ) $0 < \mu < \mu_c$ の場合。

$$N(t) \sim A t^{\hat{n}-1} e^{(\hat{n}-\mu)t}, \quad M(t) \sim B t^{\hat{n}-1} e^{(\hat{n}-\mu)t} \quad (t \rightarrow \infty). \quad (31)$$

ただし、 A, B は正数、 \hat{n} は \hat{m} に等しい m_n の個数である。

この場合の進化速度は

$$v = 0 \quad (32)$$

である。 ■

命題 2. (一般のマルカス径数の場合) $\mu > \mu_c$ に対し $N(t), M(t)$ が $t \rightarrow \infty$ で (24) の型の漸近挙動を示すと仮定すれば、 P^* は (25) の区間 $(\hat{m}-\mu, \hat{m})$ 内の一意根であり、 v は (26) によ、 z とえられる。 ■

§ 4. 周期的適応度モデル。 やはりサイズ効果のない連

続時間モデル (7), (11) について調べよう。本節では、マルカス径数の時間変動の効果を検討するため

$$m_n(t) = (-1)^{[t/2]+n} \cdot s \quad (n=0, 1, 2, \dots, 0 \leq t < \infty) \quad (33)$$

と仮定しよう。ここで、 $[\cdot]$ はガウス記号である。(33) に

よれば、 $m_n(t)$ は z 時間同一の値 ($+s$ 又は $-s$) を保持し、

次いで突然符号を逆転し、又 z 時間同一の値を保持する。

値の変動幅は $2s$ である。

この周期的適応度モデルの進化速度 v は

$$v = \mu (\mu^2 C + s^2 S / \xi) / \sqrt{(\mu^2 + s^2)(\mu^2 C^2 + S^2)} \quad (34)$$

と厳密に解析的に求められる。また、集団の平均エルサス径

数 $\bar{m}(t)$ の長時間平均 \bar{m}_∞ は

$$\bar{m}_\infty \equiv \lim_{t \rightarrow \infty} \int_0^t \bar{m}(t') dt' / t = z^{-1} \log \left\{ \frac{\sqrt{s^2 + \mu^2 C^2} + \mu S}{\sqrt{s^2 + \mu^2}} \right\} - \mu \quad (35)$$

となる。ただし、ここで

$$\xi \equiv z \sqrt{\mu^2 + s^2}, \quad C \equiv \cosh \xi, \quad S \equiv \sinh \xi \quad (36)$$

とおいた。

(34) によれば $v^* \equiv v/s$ は $\mu^* \equiv \mu/s$ と $z^* \equiv sz$ のみの関数である。図1は $(\log \mu^*, \log z^*)$ 平面上に描いた $\log v^*$ の等高線地図である。これより、 $\log \mu^* \times \log z^*$ 平面は進化速度 v の $\log \mu^* \times \log z^*$ 依存性に関して、(I) $v \equiv 1/2$ なる“ z 領域” (第2象限内, 横軸に平行な等高線), (II) $v \equiv \mu$ なる“ μ 領域” (第1, 3, 4象限内, 縦軸に平行な等高線), と(III)両領域間の遷移域, に大別されることが分る。

§5. サイズ効果の影響。

サイズ効果が進化速度に及ぼす影響については Wright-Fisher モデル (1) ~ (4) で調べてみた。

この場合の解析的な結果はまだ得られていない。周期的適応度モデルの場合, 適応度として

$$w_n(t) = \exp \{ (-1)^{[t/2] + n} \cdot s \} \quad (37)$$

を与え、マルコフ過程の標本過程をモンテ・カルロ法によつて発生させ、平均ステップ数の増大率などを調べた。図2は $N\mu = 4$ ($N=200, \mu=0.02$) の場合の結果を図1と同様に表示した等高線図である。計算機実験はパラメータ領域 $-1 \leq \log \mu^*$, $\log z^* \leq 1$ に $5 \times 5 = 25$ 個の格子点を取り、各点で $10^4/\mu = 5 \times 10^5$ 世代にわたる実験を行なった。これら25個の格子点で得られた進化速度の値は、 $N = \infty$ のときの(34)による値と非常に良い一致を示した。また、これら25個の格子点でのデータに適当な内挿操作を施して描かせた図2の等高線図から、 $N\mu = 4$ の場合にも μ 領域の存在することが $N = \infty$ の場合と同様に見取れる。

§6. 揺動適応度モデル。 図3は、 n に関して $w_n(t)$ が

$$P\{\log w_n(t+1) = m' \mid \log w_n(t) = m\} = (1 - e^{-1/2}) \delta(m - m') + (2s)^{|m - m'|} e^{-1/2} \quad (38)$$

$$(-s \leq m, m' \leq s)$$

であるような適応度系列の標本過程で特徴づけられる Wright-Fisher モデル(1)~(4)を、やはりモンテ・カルロ法により計算機実験した結果を示している。 s は揺動幅、 z は平均持続世代数であつて、図の表示法は図1、図2と同様である。図

3は $N\mu = 4$ の場合の結果である。図2, 図3は少くとも定性的には良く似ており, どちらも μ 領域, τ 領域の存在を示している。この結果は, 週期的な適度モデルがかなり極端に簡単化されたモデルであるにもかかわらず, 進化速度パラメータ依存性の本質的特徴に関しては, 相当広い範囲のより現実的な変動淘汰モデルの典型例となつてゐる可能性を示唆している。

サイズ効果を無視した揺動適度モデルの連続時間版については, まだ厳密な解析的結果は求められていない。

参考文献

- [1] M. Kimura, The neutral theory of molecular evolution. *American Scientist* 241 (1979) No. 5, pp. 94-104.
- [2] H. Matsuda and T. Gojobori, Protein polymorphism and fluctuation of environments. *Adv. Biophys.* 12 (1979) pp. 53-99.
- [3] K. Ishii, H. Matsuda and N. Ogita, Molecular evolution by stochastic selection. *Proc. Intern. Symposium on Mathematical Topics in Biology, Kyoto, Sept. 1978*, pp. 24-33.

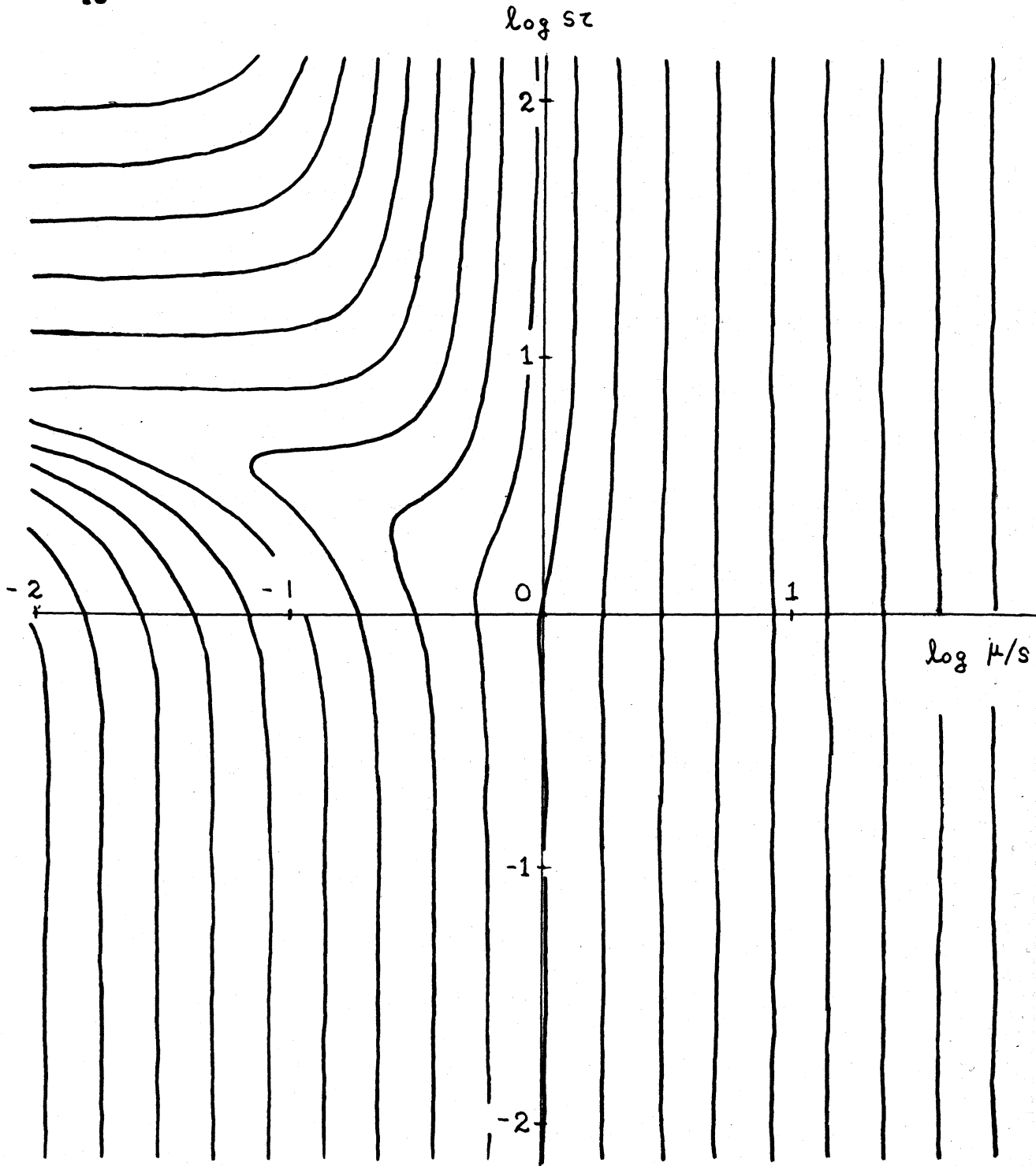


図1. 週期的適応度モデルの進化速度 ($N = \infty$ の場合)

< $\log v/s$ の等高線地図 >

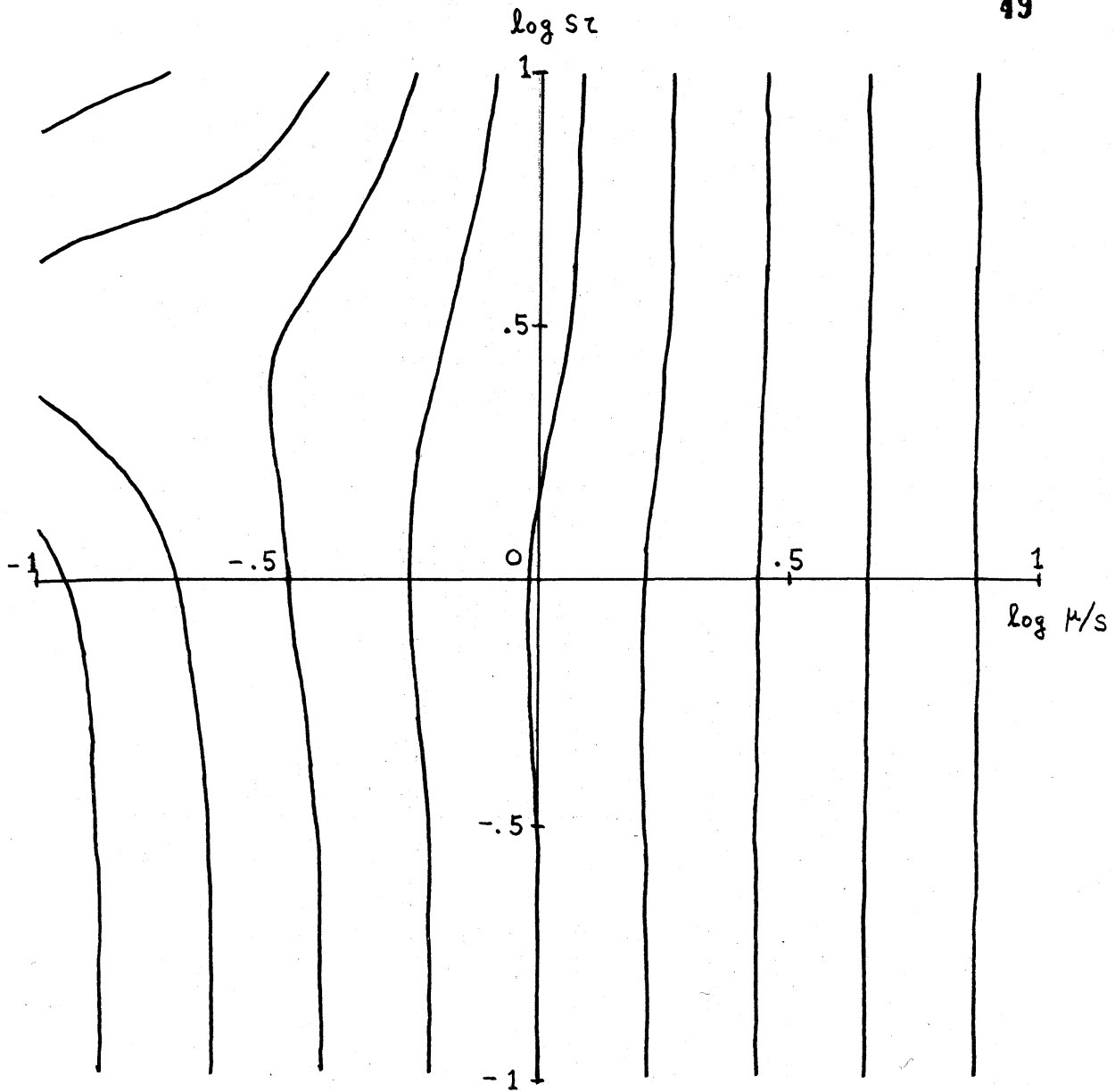


図2. 週期的適応度モデルの進化速度 ($N\mu = 4$ の場合)

〈 $\log v/s$ の等高線地図 〉

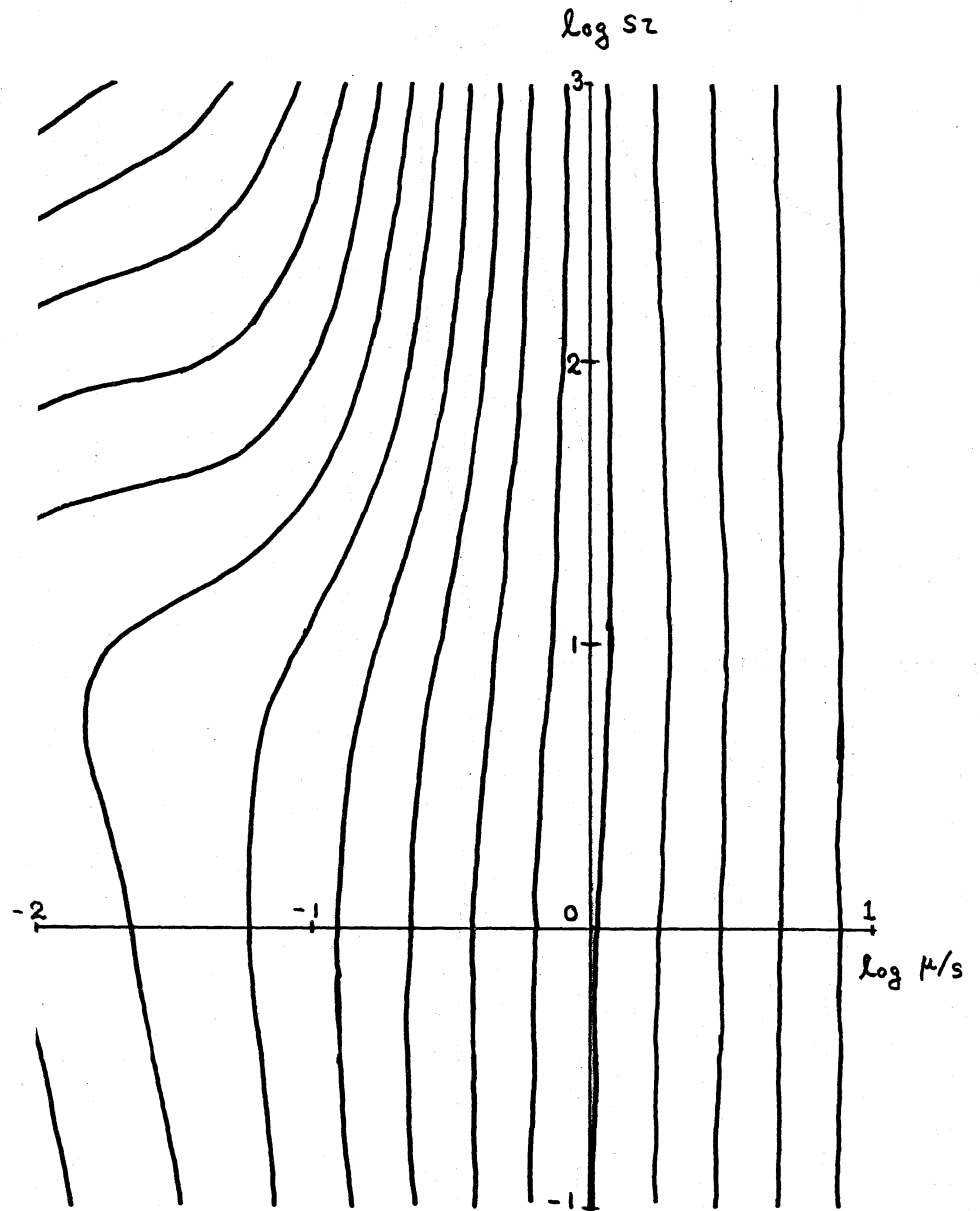


図3. 揺動振幅モデルの進化速度 ($N\mu=4$ の場合)

〈 $\log v/s$ の等高線地図 〉