Dokl. Akad. Nauk SSSR
Tom 209 (1973), No. 2

Soviet Math. Dokl.
Vol. 14 (1973), No. 2

ON THE BEHAVIOR OF THE TRAJECTORIES OF A CLASS OF GENETIC SYSTEMS

UDC 519.9+575.1

V. M. KIRŽNER

In the present note, we will follow the definitions and notation of the work of Ju. I. Ljubič [1]. A series of works have been devoted to the study of the behavior of trajectories of an infinite population under the condition of panmixia and the absence of mutations, migration, and selection. A theorem about convergence for haploid populations of gametes (and consequently diploid populations of zygotes) was first obtained by H. Geiringer in [2]. For the description of evolution O. Reiersøl [3] proposed a special method of differential operators, which he applied to the proof of a theorem about the convergence of trajectories for any number of loci rigidly linked with sexual X-chromosomes. The trajectory for one partially sex-linked locus was investigated by O. Kempthorne in [4].

The method of differential operators was developed by Ljubič in [1] and further results about the behavior of the trajectories of autosomal (isopolyploid) populations were obtained. In the current note, these results are extended to general diploid genetic systems with sexual differentiation, including partial sex linkage. The number of loci with this or another type of sex linkage may be arbitrary. The number of alleles at each locus is arbitrary. For this class of systems, a theorem about the convergence of trajectories under panmixia and the absence of selection (mutations, however, are admitted) is obtained, the set of equilibrium states is described and the limiting states of the trajectories are calculated in terms of the initial states. Further, explicit expressions for evolutionary roots are obtained and on this basis the speed of convergence to equilibrium is investigated. In conclusion, we write out an explicit evolution formula in the first approximation with respect to measures of nonequilibrium. The complete formula is not cited due to its extreme unwieldiness.

The set of loci $L = \{1, 2, \cdots, l\}$ will be divided into subsets of the following form: in the set L_X, loci representing only X-chromosomes occur; in the set L_Y are loci represented only at Y-chromosomes; in the set L_{XY} are loci represented at both sex chromosomes; in the set L_A are the loci of the autosomes. We put

$$L_{♀} = L_X \cup L_{XY} \cup L_A, \quad L_{♂} = L_Y \cup L_{XY} \cup L_A. \qquad (a)$$

Let $n_i \geq 2$ be the number of alleles at the ith locus. We denote by x_k^i, $i = 1, 2, \cdots, l$; $k = 1, 2, \cdots, n_i$, the kth gene of the ith locus. We define female and male gametes as monomials having the forms

---

AMS (MOS) subject classifications (1970). Primary 92A10.

Copyright © 1973, American Mathematical Society

$$g_f = \prod_{i \in L_f} \varepsilon^1_{h_i}, \quad g_m = \prod_{i \in L_m} \varepsilon^1_{h_i},$$

respectively.

The sets of female and male gametes we denote by Γ_f and Γ_m, respectively. We denote the real linear spaces with bases Γ_f and Γ_m by $\mathfrak{R}_f$ and $\mathfrak{R}_m$, while σ_f and σ_m are the standard simplexes of these spaces. The state of the population under consideration will be identified with the elements G of the set $\sigma_f \times \sigma_f \times \sigma_m$, which is contained in the linear space $\mathfrak{R} = \mathfrak{R}_f \times \mathfrak{R}_f \times \mathfrak{R}_m$. Each state of the population G is a collection $(G(1), G(2), G(3))$ of states of different parts of the population: $G(1)$ is the state of the female gametes of maternal origin, $G(2)$ is the state of female gametes of paternal origin, $G(3)$ is the state of the male gametes.

Crossovers in a separate sex system are described by two linkage distributions $\mathfrak{P}_f(L)$ and $\mathfrak{P}_m(L)$, given on the sets L_f and L_m, respectively. We note that the linkage distribution $\mathfrak{P}_m(L)$ is not symmetric, i.e. the probabilities $p_m(u \mid v)$ and $p_m(v \mid u)$ are, generally speaking, unequal.

In the space $\mathfrak{R}$, we define a multiplication $G \times H = E$ by the formulas

$$\begin{gathered} E(1) = \sum_{u \mid v} \frac{1}{2} p_f(u \mid v) \left(D_u G(1) D_v H(2) + D_u H(2) D_v G(1) \right), \\ E(2) = \sum_{u \mid v} p_m(u \mid v) D_u G(1) D_v D_{L_f} H(3), \\ E(3) = \sum_{u \mid v} p_m(u \mid v) D_u H(3) D_v D_{L_m} G(1). \end{gathered} \tag{1}$$

We denote the algebra thus obtained by $\mathfrak{A}(L; \mathfrak{P}_f(L), \mathfrak{P}_m(L))$. Using (1), we may write the evolution equation without regard to mutations in the form

$$G_{n+1} = G_n \times G_n. \tag{2}$$

In the case of one locus, mutations are described by a column stochastic transition matrix $M = \|\mu_{ij}\|$, where μ_{ij} is the probability of passage of the jth gene into the ith. We assume that mutations take place at each locus independently of the others. In this case, it is possible to describe them by means of the matrix

$$M = \bigotimes_{j=1}^{l} M_j,$$

where M_j is the matrix of mutations at the jth locus, and the symbol $\otimes$ denotes the tensor (Kronecker) product. The evolution equation, taking account of mutations, has the form

$$\tilde{G}_{n+1} = M(\tilde{G}_n \times \tilde{G}_n), \tag{3}$$

where $\{\tilde{G}_n\}$ is the trajectory of the population with mutations. We denote by $\mathfrak{P}_f(K)$ and $\mathfrak{P}_m(K)$ the induced linkage distributions on the sets of loci K_f and K_m, respectively.

Lemma 1. *The operator* $\tilde{D}_K$ *is a homomorphism of the algebra* $\mathfrak{A}(L; \mathfrak{P}_f(L), \mathfrak{P}_m(L))$ *into the algebra* $\mathfrak{A}(K; \mathfrak{P}_f(K), \mathfrak{P}_m(K))$.

Lemma 2. *The linear operator* $\mathfrak{M}$ *is a homomorphism of the algebra* $\mathfrak{A}(L; \mathfrak{P}_f(L), \mathfrak{P}_m(L))$.

Theorem 1. *If the matrices* M_i, $i = 1, 2, \cdots$, *have no characteristic values of unit modulus different from* $\lambda = 1$, *then all trajectories converge.*

Theorem 2. *In order that the state* G *be an equilibrium state, it is necessary and sufficient that it have the form*

$$G(1) = G(2) = M^{\infty} \prod_{i \in L_Y} H_i, \quad G(3) = M^{\infty} \prod_{i \in L_m} H_i, \tag{**}$$

where H_i *is an arbitrary state of the ith locus.*

Corollary. *The trajectory defined by the initial state* G_0 *converges to the equilibrium state* G_∞:

$$G_\infty(1) = G_\infty(2) = M^{\infty} \prod_{i \in L_Y \setminus L_m} \left\{ \frac{1}{2} \tilde{D}_i G_0(1) + \frac{1}{4} \tilde{D}_i G_0(2) \right. \tag{***}$$

$$\left. + \frac{1}{4} \tilde{D}_i G_0(3) \right\} \prod_{i \in L_X} \left\{ \frac{2}{3} \tilde{D}_i G_0(1) + \frac{1}{3} \tilde{D}_i G_0(2) \right\},$$

$$G_\infty(3) = M^{\infty} \{ \tilde{D}_{L_Y} G_0(3) \} \prod_{i \in L_Y \setminus L_m} \left\{ \frac{1}{2} \tilde{D}_i G_0(1) + \frac{1}{4} \tilde{D}_i G_0(2) + \frac{1}{4} \tilde{D}_i G_0(3) \right\}.$$

We now pass to a description of the evolution spectrum in the system L_f. Analogous formulas for the system L_m will not be cited. We denote by $Q_2(L)$ the matrix

$$Q_2(L) = \begin{pmatrix} \frac{p_f(L)}{2} & \frac{p_f(L)}{2} & 0 \\ p_m(L \mid \emptyset) & 0 & p_m(\emptyset \mid L) \\ p_m(\emptyset \mid L) & 0 & p_m(L \mid \emptyset) \end{pmatrix},$$

while

$$Q_1(L) = \begin{pmatrix} \frac{p_f(L)}{2} & \frac{p_f(L)}{2} \\ p_m(L \mid \emptyset) & 0 \end{pmatrix}.$$

With each subsystem $K \subset L_f$ we associate some part of the spectrum of the matrix $Q_2(K)$ or $Q_1(K)$. Namely, if $K_X = K_Y = \emptyset$, then we denote by $\pi^{\alpha}(K)$, $\alpha \le 3$, the characteristic values of the matrix $Q_2(K)$ differing from 0 or 1. If $K_X \cup K_Y \ne \emptyset$, then we denote by $\pi^{\alpha}(K)$, $\alpha \le 2$, the characteristic values of the matrix $Q_1(K)$, different from 0 or 1. We will assume that in all cases $\pi^1(K) = \max_\alpha \pi^\alpha(K)$.

We will consider all possible partitions $K_1 | \cdots | K_l$ of all possible subsystems $K \subset L_f$, satisfying the condition $|K_i| \ge 1$, $i = 1, 2, \cdots, l$, where the equality $|K_i| = 1$ is allowed only if $K_i \subset L_{XY}$. We shall call such partitions admissible. With an admissible partition $K_1 | \cdots | K_l$, we associate the set of all possible collections $\{\pi^{\alpha_i}(K_i)\}_{i=1}^{l}$. Each such collection is determined by some collection of indices $\alpha = \{\alpha_i\}_{i=1}^{l}$, which we shall also call admissible. For ease of writing, we set $\pi^{\alpha}(K_i) = \pi^{\alpha_i}(K_i)$, $\alpha = \{\alpha_i\}_{i=1}^{l}$.

Theorem 3. *The evolution spectrum in the system* L_f *consists of the values*

$$\lambda_{K_1 | \cdots | K_l}(\nu, \alpha) = \left(-\frac{1}{2} \right)^{\nu} \prod \pi^{\alpha}(K_i), \quad \lambda_0 = 1, \tag{4}$$

where $K_1|\dots|K_r$ *is any admissible partition,* α *is any admissible collection of indices, and* $\nu \le |L_X \setminus \bigcup K_i|$.

The speed of convergence to equilibrium is characterized by the coefficient of stabilization $\varkappa$. For a general linkage distribution we have the following

Theorem 4. *The coefficient of stabilization is*

$$\varkappa = \max_{|K|\le 2} \varkappa^1(K), \tag{5}$$

where the maximum is taken over all K *for which* $|K| = 2$, *if* $K \not\subset L_{XY} \cup L_X$, *and* $|K| \le 2$ *if* $K \subset L_{XY} \cup L_X$.

Using (5), it is possible to obtain a bound on the coefficient of stabilization, depending only on the number of loci.

Theorem 5. *We have the inequality*(1)

$$\varkappa \geqslant \left[\frac{l-1}{2}\right]\left(2\left[\frac{l-1}{2}\right]+1\right)^{-1}.$$

We now determine a linear approximation of the trajectory in the system L_Y. For this, it is convenient to introduce the point of reference from the generation G_2. We denote by $\overline{E^{\alpha}_{Q}(G_2)}$ measures of nonequilibrium of the first degree

$$\overline{E^{\alpha}_{Q}(G_2)} = (\bar{D}_Q G_2(\alpha) - \bar{D}_Q G_\infty(\alpha))\, D_Q G_\infty(\alpha).$$

By $A^Q_K(\iota, \alpha, s)$ we denote the solution of the following system of equations:

$$\sum_{\alpha} A^Q_K(\iota, \alpha, s) = (-1)^{|K \setminus Q|}\delta_{\iota, 1},$$

$$[Q(K) - \lambda_K(\alpha)]\, A^Q_K(\alpha, s) = 0,$$

where $Q(K) = Q_0(K)$, if $K \cap L_X = \varnothing$, and $Q(K) = Q_1(K)$, if $K \cap L_X \ne \varnothing$; $A^Q_K(\alpha, s) = \|A^Q_K(\iota, \alpha, s)\|_\iota$.

Theorem 6. *The trajectory, linearized relative to measures of nonequilibrium of the first degree, has the form*

$$G_n(s) = G_\infty(s) + \left(-\frac{1}{2}\right)^{n-2} \sum_{\max|g-\tilde{g}|=1} \overline{E^1_{\tilde g}(G_2)} + \sum_{K, s} [\varkappa^s(K)]^{n-2}\left[\sum_{Q \subset K} A^Q_K(\iota, \alpha, s)\, \overline{E^\alpha_Q(G_2)}\right].$$

The author expresses thanks to Ju. I. Ljubič for guidance of this work.

Physical Technical Institute of Low Temperatures
Academy of Sciences of the USSR

Received 10/MAY/72

BIBLIOGRAPHY

1. Ju. I. Ljubič, *Basic concepts and theorems of evolution genetics of free populations*, Uspehi Mat. Nauk 26 (1971), no. 5 (161), 51–116 = Russian Math. Surveys 26 (1971), no. 5, 51–123.

2. H. Geiringer, *On the probability theory of linkage in Mendelian heredity*, Ann. Math. Statist. 15 (1944), 25–57. MR 6, 11.

---

(1) The square brackets denote the integral part of a number.

3. O. Reiersøl, *Genetic algebras studied recursively and by means of differential operators*, Math. Scand. 10 (1962), 25–64. MR 25 #1189.

4. O. Kempthorne, *An introduction to genetic statistics*, Wiley, New York; Chapman and Hall, London, 1957. MR 19, 622.

Translated by J. CASTI