RANDOM COLLISION MODELS IN ORIENTED GRAPHS

YOSHIAKI ITOH,* *Institute of Statistical Mathematics, Tokyo*

Abstract

We investigate a random collision model for competition between types of individuals in a population. There are dominance relations defined for each pair of types such that if two individuals of different types collide then after the collision both are of the dominant type. These dominance relations are represented by an oriented graph, called a tournament. It is shown that tournaments having a particular form are relatively stable, while other tournaments are relatively unstable. A measure of the stability of the stable tournaments is given in the main theorem.

RANDOM COLLISION; ORIENTED GRAPHS; TOURNAMENTS; COMPETING SPECIES; MARTINGALES; ASYMPTOTIC PROBABILITY OF COEXISTENCE

1. Introduction

The problem of competing species has been considered by many authors since Lotka (1925) and Volterra (1931). Ehrenfest's urn model was discussed in detail by Kac (1959), and another urn model for a problem in genetics was treated by Moran (1958). Random collision models for competing species are considered by Itoh (1971), (1973), (1975), and in Itoh (1973), several qualitative properties for species coexistence were given with a computer simulation. Kimura (1955) calculated the asymptotic probabilities for Wright's model in genetics, which can be interpreted as a random collision model without competition. Kimura (1958) and Mather (1969) have considered deterministic versions of the stochastic model investigated in this paper.

We investigate a random collision model for competition between types of individuals in a population. Consider a population consisting of individuals, each of which is one of p types. The types may represent species, genotypes, types of consumer or other classifications. We let $N_i(t)$ denote the number of individuals of type i present in the population at time t, and let $N(t) = (N_1(t), N_2(t), \cdots, N_p(t))$.

For each pair of types, a dominance relation is defined: either type i

---

dominates type j or vice versa. We then consider random collisions between individuals which are assumed to occur at the rate of one per unit time. If two individuals of different types collide then two individuals of the dominant type result from the collision. It should be noted that the term 'dominate' is not used here in the sense in which it is used in genetics: see Moon (1968).

An oriented graph is a convenient means of representing the dominance relations. An oriented graph is a collection of points, called nodes, representing the types, and directed lines between the points, representing the dominance relations. An oriented graph representing a complete set of dominance relations is called a tournament.

The tournament $[T_r]$ consists of a set of $2r + 1$ nodes such that each node is dominated by r nodes and dominates r nodes. This symmetry gives such a tournament a significant degree of stability, and the tournaments $[T_r]$, $r = 0, 1, 2, \cdots$, play an important role in the investigation of the random collision model. As the types successively become extinct, the tournament of the non-extinct types is altered. The main result of this paper indicates that tournaments isomorphic to $[T_r]$, $r = 0, 1, 2, \cdots$, are relatively stable, while other tournaments are relatively unstable. A measure of the stability of a tournament isomorphic to $[T_r]$ is given by the main theorem which indicates that

$$\Pr([R(t)] \cong [T_r]) \sim C\theta_r^t, \qquad \text{as } t \to \infty,$$

where $[R(t)]$ denotes the tournament of non-extinct types at time t, $\theta_r = 1 - 2\binom{2r+1}{2}/n(n-1)$ and C is a constant. The proof uses a family of martingales, $\{(K_r(t), \mathcal{F}_t), t = 0, 1, 2, \cdots\}$, $r = 0, 1, 2, \cdots$, where $\mathcal{F}_t$ denotes the σ-field generated by $\{N(u), u = 0, 1, 2, \cdots, t\}$.

To illustrate the behaviour of the model, we consider an example with tournament $[V]$ given in the Figure 1(a). We note that $[V]$ is isomorphic to $[T_2]$. A computer simulation of the process $\{N(t), t = 0, 1, 2, \cdots\}$ is given in Figure 2. At time 0 there are five types each with 20 individuals, and from time 0 to time 1392 the five types coexist with tournament $[V]$. At time 1392 type 5 becomes
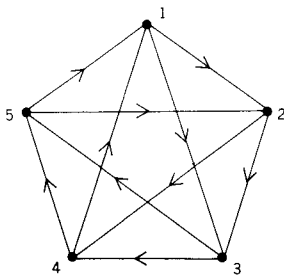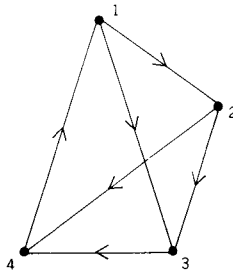

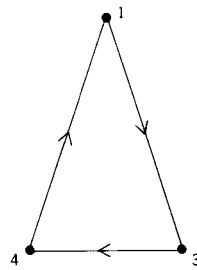
Figure 1 (a)   Figure 1 (b)   Figure 1 (c)
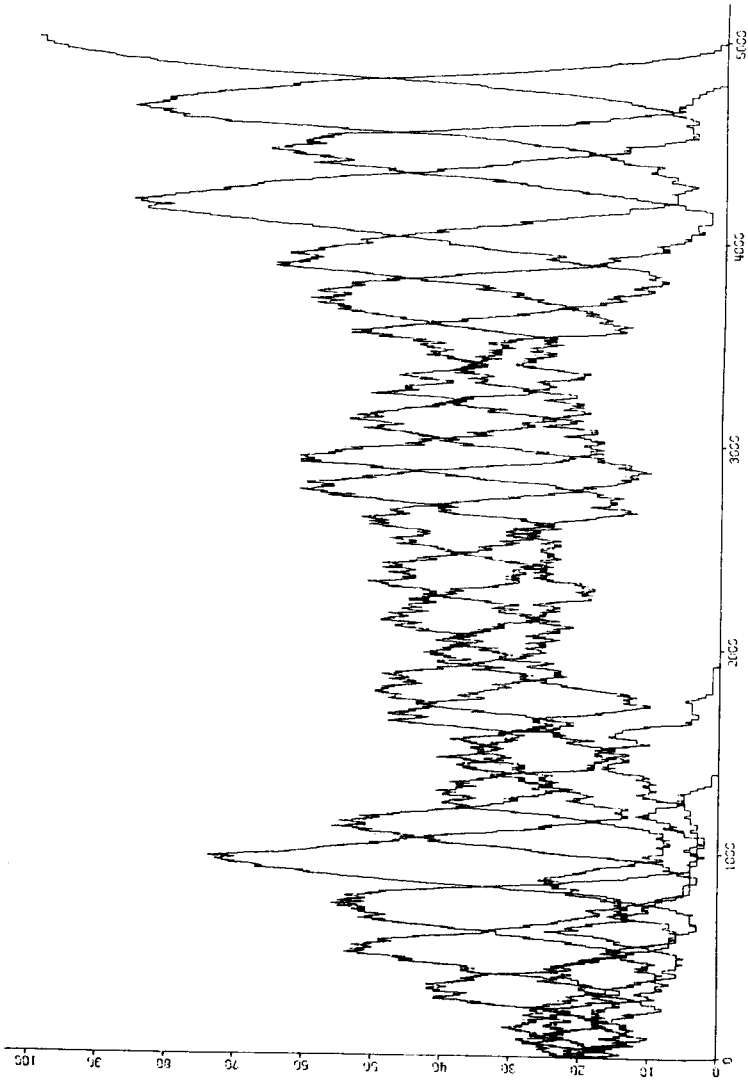
Figure 2

extinct. From time 1393 to time 1921 there are four types having the tournament obtained from [V] by deleting the node corresponding to type 5 and lines to it, as given in Figure 1(b). At time 1921 type 2 becomes extinct, and from time 1922 to time 4786, there are three types having a tournament isomorphic to [T_1], as given in Figure 1(c). The relative stability of the tournaments [T_2] and [T_1] is indicated by this realization.

2. Tournaments

A tournament [T] consists of a set of nodes T, and oriented lines joining each pair of nodes, such that each pair of nodes is joined by one and only one oriented line. Thus, if $i, j \in T$ then either $\overrightarrow{ij}$ or $\overrightarrow{ji}$ is in [T]. If the line $\overrightarrow{ji}$ is in [T], then we say that i dominates j, and write $i > j$, or $j < i$.

Two tournaments [S] and [T] are isomorphic, and we write [S] $\cong$ [T], if there exists a one-to-one dominance preserving mapping between their nodes.

A tournament [S] is a subtournament of a tournament [T] if there exists a one-to-one dominance preserving mapping between S and a subset of T. The subtournament of [T] generated by $S \subset T$ is the tournament with S as its set of nodes and with dominance relations as in [T], and we write [S] $\subset$ [T].

We define the dominance indicator for a tournament [T] as

$$a_{ij} = a_{ij}(T) = \begin{cases} 1, & \text{if } i > j, \\ 0, & \text{if } i = j, \\ -1, & \text{if } i < j. \end{cases}$$

Note that $a_{ij} = -a_{ji}$.

For a tournament [T], $S \subset T$ and a node $k \in T \setminus S$, we define the dominance relation as follows:

$$k > S \Leftrightarrow \sum_{i \in S} a_{ik} < 0$$

$$k < S \Leftrightarrow \sum_{i \in S} a_{ik} > 0.$$

The tournament [T_r] consists of a set of nodes $T_r = \{0, 1, 2, \cdots, 2r\}$ and dominance relations defined by

$$i > j \quad \text{if } i - j \equiv 1, 2, \cdots, r \quad (\text{mod } 2r + 1).$$

Thus, each node dominates (and is dominated by) r other nodes. In dealing with the tournament [T_r], it is convenient to define node i, for an integer $i \notin T_r$, as node i_r where i_r denotes the residue of i modulus $2r + 1$. Thus node $2r + 1$ is node 0, and so on.

3. Definition of the model

We consider a population of individuals of p types, $1, 2, \cdots, p$, whose numbers are $N_1(t), N_2(t), \cdots, N_p(t)$ respectively at time t. We define $N(t) =$

$(N_1(t), N_2(t), \cdots, N_p(t))$. It is assumed that initially $N(t) = \alpha$, that in each unit of time one collision occurs, and that all possible collisions are equiprobable. There exist dominance relations between the types as defined by the tournament $[V]$, with nodes $V = \{1, 2, \cdots, p\}$ corresponding to the types, which is assumed to be a subtournament of $[U] \cong [T_s]$. Further, it is assumed that after a collision between individuals of different types, both are of the dominant type, while a collision between individuals of the same type has no effect.

These assumptions imply that $\{N(t), t = 0, 1, 2, \cdots\}$ is a Markov chain with transition probabilities defined by

$$
\pi_{ij} = \Pr[N(t+1) = n_{ij} \mid N(t) = n] = 2n_i n_j / n(n-1) \qquad i, j = 1, 2, \cdots, p, \ i \neq j;
$$
(1)
$$
\pi_{ii} = \Pr[N(t+1) = n \mid N(t) = n] = \sum_{i=1}^{p} n_i (n_i - 1)/n(n-1);
$$

where $n = (n_1, n_2, \cdots, n_p)$, $n = \sum_{i=1}^{p} n_i$, and n_{ij} is a $p \times 1$ vector with ith component $n_i + a_{ij}$, jth component $n_j + a_{ji}$, and all other components equal to those of n.

If we let $f_i(\tau) = N_i(n\tau)/n$, then we are led to the deterministic approximation, $f(\tau)$ which is such that

$$
(2) \qquad\qquad f_i'(\tau) = f_i(\tau) \sum_{j \in V} a_{ij} f_j(\tau), \qquad i \in V,
$$

with $f_i(\tau) > 0$ and $\sum_{i \in V} f_i(\tau) = 1$.

We denote by $\mathscr{F}_t$ the σ-field generated by $\{N(u), u = 0, 1, 2, \cdots, t\}$. We denote the tournament of non-extinct types at time t by $[R(t)]$, so that $[R(0)] = [V]$; and we define $m = m(V) = \max\{r : [X] \subset [V], [X] \cong [T_r]\}$. We also define

$$
g_X(n) = \prod_{i \in X} n_i, \quad \text{and} \quad H_{V, r}(n) = \sum_{X \in A_r} g_X(n),
$$

where $A_r = \{X : [X] \subset [V], [X] \cong [T_r]\}$. Note that $H_{V, r}(n)$ is equal to the number of possible tournaments of individuals isomorphic to $[T_r]$ when $N(t) = n$, and so the process $H_{V, r}(N(t))$ can be thought of as a measure of $[T_r]$-ness of the population at time t.

4. Results

Lemma 1. If $[X] \subset [V]$, $[X] \cong [T_r]$ then

$$
\sum_{i \in X} a_{ij} = \begin{cases} 1, & \text{if } j < X \\ 0, & \text{if } j \in X \\ -1, & \text{if } j > X. \end{cases}
$$

Proof. This follows by straightforward enumeration of cases.

Lemma 2.   If $[X] \subset [V]$, $[X] \cong [T_r]$, and $X > i$ $(X < i)$ for every $i \in V \setminus X$, then $[X]$ is uniquely determined.

Proof.   This follows from the uniqueness of the solution of (2).

Lemma 3.   Let $[X] \subset [V]$, $[X] \cong [T_r]$ and $k \in V \setminus X$, such that $x_i < k < x_{i+1}$. If $X < k$ $(X > k)$, and X^*, k^* are such that $X^* \cup \{k^*\} = X \cup \{k\}$, $[X^*] \cong [T_r]$, and $X^* > k^*$ $(X^* < k^*)$, then X^* and k^* are uniquely determined by $k^* = x_i$ $(k^* = x_{i+1})$, and $X^* = (X \setminus \{k^*\}) \cup \{k\}$. We write $X^* = X(k)$ and $k^* = k(X)$.

Proof.   From the assumptions we have, for nodes x of X,

$$k < x_{i+j}, \qquad j = 1, 2, \cdots, r;$$

$$k > x_{i+j}, \qquad j = 0, -1, -2, \cdots, -r.$$

It follows that

$$\{x : k > x, x \in (X \setminus \{x_i\})\} = \{x : x_i > x, x \in X\}$$

$$\{ \quad \{x : k < x, x \in (X \setminus \{x_i\})\} = \{x : x_i < x, x \in X\}.$$

Thus we see that X^* and k^* are as specified above. The uniqueness follows from Lemma 2.

Lemma 4.   Let $[X] \subset [V] \subset [U]$ and $[X] \cong [T_m]$, where $m = \max\{r: [X] \subset [V], [X] \cong [T_r]\}$. Suppose that $X < k$ for all $k \in V \setminus X$, and that $l < X$ for some $l \in U \setminus V$; then one of the following two statements holds:
  (i) $\max\{r : [X] \subset [V \cup \{l\}], [X] \cong [T_r]\} = m + 1$,
  (ii) $X(l) < k$ for all $k \in (V \cup \{l\}) \setminus X(l)$.

Proof.   Let $Y = X \cup \{i, j\} \subset U$, $[X] \cong [T_r]$, $i > X$, $j < X$ and $X(j) > i$, then it follows from Lemma 3 that $[Y] \cong [T_{r+1}]$. Therefore, if (ii) does not hold, then there exists $h \in V \setminus X$ such that $X(l) > h$, so (i) must hold.

Lemma 5.   Let $m = \max\{r : [X] \subset [V], [X] \cong [T_r]\}$. If $[W]$ is such that

$$[W] \cong [T_m], \quad i > W \quad \text{for all} \quad i \in V \setminus W, \quad \text{and} \quad [W] \subset [V],$$

then $[W]$ is uniquely determined. Further, for the tournament $[W]$,

$$E(g_W(N(t+1)) \mid \mathscr{F}_t) = \left[1 - \left(2 \binom{2m+1}{2} + \sum_{i \in V \setminus W} N_i(t) \right) \Big/ n(n-1) \right] g_W(N(t)).$$

Proof.   We construct $[W]$ from an arbitrary $[X] \cong [T_m]$. From Lemma 3, if $k \in V \setminus X$ then either (i) $X < k$, or (ii) $X(k) < k(X)$. Since $\max\{r : [X] \subset [V], [X] \cong [T_r]\} = m$, Lemma 4 assures us that W can be constructed recursively. Using Lemma 1, we find that

$$\frac{1}{2} \sum_{i \in W} \sum_{j \in W} (g_W(\boldsymbol{n}_{ij}) - g_W(\boldsymbol{n})) \pi_{ij} = -2 \binom{2m+1}{2} g_W(\boldsymbol{n})/n(n-1),$$

$$\sum_{i \in W} \sum_{j \in V \setminus W} (g_W(\boldsymbol{n}_{ij}) - g_W(\boldsymbol{n})) \pi_{ij} = -\left(\sum_{i \in V \setminus W} n_i \right) g_W(\boldsymbol{n})/n(n-1).$$

These relations lead directly to the given result.

Lemma 6. Let $\theta_r = 1 - 2\binom{2r+1}{2}/n(n-1)$, and define

$$K_r(t) = \theta_r^{-t} H_{V,r}(N(t)),$$

then $\{(K_r(t), \mathscr{F}_t), t = 0, 1, 2, \cdots\}$ is a martingale for $r = 0, 1, 2, \cdots$.

Proof.

$$E(H_{V,r}(N(t+1)) \mid \mathscr{F}_t) - H_{V,r}(N(t)) = \frac{1}{2} \sum_{X \in A} \sum_{i \in V} \sum_{j \in V} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij}.$$

Now,

$$\sum_{i \in V} \sum_{j \in V} = \sum_{i \in X} \sum_{j \in X} + 2 \sum_{i \in X} \sum_{j \in Y} + \sum_{i \in Y} \sum_{j \in Y}$$

where $Y = V \setminus X$. We have, using Lemma 1,

$$\sum_{i \in X} \sum_{j \in X} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij} = -4 \binom{2r+1}{2} g_X(\boldsymbol{n})/n(n-1).$$

Also, since $g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n}) = 0$ for $i, j \in Y$, we have

$$\sum_{i \in Y} \sum_{j \in Y} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij} = 0.$$

Finally, we have

$$\sum_{X \in A} \sum_{j \in Y} \sum_{i \in X} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij}$$

$$= \sum_{X \in A} \sum_{j \in Y} \sum_{i \in X(j)} [g_{X(j)}(\boldsymbol{n}_{ij(X)}) - g_{X(j)}(\boldsymbol{n})] \pi_{ij(X)}$$

by definition of $X(j)$ and $j(X)$. Further, using Lemma 3 and Lemma 1, we find that

$$\sum_{i \in X} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij} + \sum_{i \in X(j)} [g_{X(j)}(\boldsymbol{n}_{ij(X)}) - g_{X(j)}(\boldsymbol{n})] \pi_{ij(X)} = 0.$$

It follows that

$$\sum_{X \in A} \sum_{i \in X} \sum_{j \in Y} [g_X(\boldsymbol{n}_{ij}) - g_X(\boldsymbol{n})] \pi_{ij} = 0.$$

Therefore, we have

$$E(H_{V,r}(N(t+1))\,|\,\mathscr{F}_t) = \theta_r H_{V,r}(N(t))$$

from which the result follows.

Theorem.

$$\lim_{t\to\infty} \frac{\Pr\{[R(t)] \cong [T_r]\,|\,N(0)=\alpha\}}{H_{V,r}(\alpha)Q_r^{-1}\theta_r^t} = 1, \qquad r = 0, 1, \cdots, m\,;$$

where $Q_r = (1/L_r)\sum_{n\in E_r} n_0 n_1 \cdots n_{2r}$, $E_r = \{(n_0, n_1, \cdots, n_{2r}): n_i$ are positive integers such that $\sum_{i=0}^{2r} n_i = n\}$, and L_r denotes the number of elements in the set E_r.

Proof. Suppose first that $[V] \cong [T_m]$. The process $\{N(t), t = 0, 1, 2, \cdots\}$ is a Markov chain on the state space $E_m \cup \{0\}$: the set of states for which $N_1 N_2 \cdots N_p = 0$ is an absorbing set which we denote as state 0. The transition probability matrix of this Markov chain has unique second largest eigenvalue $\lambda_1 = 1 - 2\binom{2m+1}{2}/n(n-1)$, and corresponding left eigenvector $x_1 = (d, L_m^{-1}, L_m^{-1}, \cdots, L_m^{-1})$ for an appropriate d.

Thus, as $t \to \infty$, given that $[R(t)] \cong [T_m]$, each of the possible states is equiprobable, so that

$$\lim_{t\to\infty} E(H_{V,m}(N(t))\,|\,[R(t)] \cong [T_m]) = Q_r.$$

Also, from Lemma 6, we have

$$E(H_{V,m}(N(t))\,|\,N(0)=\alpha) = H_{V,m}(\alpha)\theta_m^t.$$

If $[R(t)] \not\cong [T_m]$ then $H_{V,m}(N(t)) = 0$, hence

$$E(H_{V,m}(N(t))\,|\,N(0)=\alpha)$$

$$= E(H_{V,m}(N(t))\,|\,[R(t)] \cong [T_m])\Pr([R(t)] \cong [T_m]\,|\,N(0)=\alpha).$$

Thus the result follows for $r = m$.

If $[T_m] \subset [X]$, then using Lemma 5, we see that

$$\Pr([R(t)] \cong [X]\,|\,N(0)=\alpha) = o(\theta_m^t),$$

thus the time spent in any tournament of which $[T_m]$ is a subtournament is asymptotically negligible compared to the time spent in a tournament isomorphic to $[T_m]$, so the result of the theorem holds for general $[V]$ and $r = m$. The above argument can be applied successively for $r = m - 1, m - 2, \cdots, 1$.

MacArthur (1957) considered from a census of birds that, if $p - 1$ points are chosen at random on a stick, and if the stick is broken at these points then the lengths of the resulting p segments represent the relative abundances of the p species. The result proved above gives justification for MacArthur's hypothesis: it is most likely that $[V] \cong [T_r]$, in which case all possible partitions of the population into $2r + 1$ types are equiprobable.

In many applications a modified time scale is more reasonable. If we assume that each individual collides with one other individual on average in a unit time interval, i.e., there are $\frac{1}{2}n$ collisions per unit time, then the time scale $\tau = 2t/n$, the result of the above theorem gives

$$\Pr([R(\tau)] \cong [T_r] \mid N(0) = \alpha) \sim kH_{v,r}\left(\frac{1}{n}\,\alpha\right)e^{-\binom{2r+1}{2}\tau/n}$$

where $k = \int\int \cdots \int_A f_0 f_1 \cdots f_{2r} \, df_0 df_1 \cdots df_{2r}$, with $A = \{(f_0, f_1, \cdots, f_{2r}): f_0, f_1, \cdots, f_{2r} > 0, \ f_0 + f_1 + \cdots + f_{2r} = 1\}$.

Another possible modification of the model is to assume that if individuals of type i and type j collide then they become two i's with probability $\frac{1}{2} + \delta$ and two j's with probability $\frac{1}{2} - \delta$, for $0 \leq \delta \leq \frac{1}{2}$, if $i > j$. Wright's model is the case $\delta = 0$, for which the asymptotic probabilities of species coexistence have been obtained for an arbitrary combination of species: see Kimura (1955). We have considered the case $\delta = \frac{1}{2}$. Analogous results can be obtained for other values of δ.

Acknowledgement

References

ITOH, Y. (1971) The Boltzmann equation on some algebraic structure concerning struggle for existence. *Proc. Japan Acad.* **47**, 854–858.

ITOH, Y. (1973) On a ruin problem with interaction. *Ann. Inst. Statist. Math.* **25**, 635–641.

ITOH, Y. (1975) An *H*-theorem for a system of competing species. *Proc. Japan Acad.* **51**, 374–379.

KAC, M. (1959) *Probability and Related Topics in Physical Sciences.* Interscience, London.

KIMURA, M. (1955) Random genetic drift in multi-allelic locus. *Evolution* **9**, 419–435.

KIMURA, M. (1958) On the change of population fitness by natural selection. *Heredity* **12**, 145–167.

LOTKA, A. J. (1925) *Elements of Physical Biology.* Williams and Wilkins, Baltimore, Reissued (1956), Dover, New York.

MACARTHUR, R. H. (1957) On the relative abundance of bird species. *Proc. Natn. Acad. Sci. U.S.A.* **43**, 293–295.

MATHER, K. (1969) Selection through competition. *Heredity* **24**, 529–540.

MOON, J. W. (1968) *Topics on Tournaments.* Holt, Rinehart and Winston, New York.

MORAN, P. A. P. (1958) Random processes in genetics. *Proc. Camb. Phil. Soc.* **54**, 60–71.

VOLTERRA, V. (1931) *Leçons sur la théorie mathématique de la lutte pour la vie.* Cahiers Scientifiques VII, Gauthier–Villars, Paris.