DOES ASSORTATIVE MATING LEAD TO A POLYMORPHIC POPULATION? A TOY MODEL JUSTIFICATION

RYSZARD RUDNICKI
Institute of Mathematics, Polish Academy of Sciences
Bankowa 14, 40-007 Katowice, Poland

RADOSŁAW WIECZOREK
Institute of Mathematics, University of Silesia
Bankowa 14, 40-007 Katowice, Poland

Abstract. We consider a model of phenotypic evolution in populations with assortative mating of individuals. The model is given by a nonlinear operator acting on the space of probability measures and describes the relation between parental and offspring trait distributions. We study long-time behavior of trait distribution and show that it converges to a combination of Dirac measures. This result means that assortative mating can lead to a polymorphic population and sympatric speciation.

1. Introduction. Quantitative investigation of the inheritance of phenotypic traits has a long tradition, going back to the seminal paper [9]. Extensive information and references about classical models of inheritance and trait evolution can be found in [3]. Speciation is one of the most important topics of evolutionary biology. It is clear that one of the reasons of speciation is geographical isolation (allopatric, peripatric and parapatric speciation) but evolutionary biologists consider also more controversial sympatric speciation — the process through which new species evolve from a single ancestral species while inhabiting the same geographic region [13]. Such speciation is a product of reproductive isolation and the reasons of it can be disruptive selection (a natural selection which favors both extremes of the population) or assortative mating. Assortative mating appears when individuals with similar traits mate more often than they would if they were choosing a partner randomly. In this paper we present a simple model of the evolution of phenotypic traits with assortative mating and show mathematical results which can be used as an argument for sympatric speciation. The traits are assumed to remain unchanged during lifetime, and their examples are skin color, shape of leaves and shell pattern. The trait of an individual is described by a real number. We consider a hermaphroditic population, i.e., a population where every individual has both male and female reproductive system. An individual chooses a partner according to a preference function [8, 10, 14, 18, 22, 23]. We allow also for self-fertilization or vegetative reproduction in the case when an individual has not found a proper partner.

2010 Mathematics Subject Classification. Primary: 47H30; Secondary: 92D15.
Key words and phrases. Assortative mating, phenotypic evolution, polymorphic population, nonlinear operator of trait inheritance, convergence to multimodal distribution.
* Corresponding author.
Self-fertilization or vegetative reproduction occurs in most flowering plants, numerous protozoans, and many invertebrates. As a result of mating or self-fertilization, a new individual is born. In the main model we assume that the offspring trait is a random variable with values between the parental traits. This assumption seems to be artificial but it allows us to give a strict mathematical proof of the convergence of trait distribution to a combination of Dirac measures. In a real situation the range of the offspring trait can be wider which can lead to a multimodal limit distribution.

We consider here a discrete time population model, i.e., a model with a fixed interval between generations. Such models are appropriate when species have no overlap between successive generations. The model is described by an operator defined on the space of probability measures and this operator describes the relation between the distributions of traits in consecutive generations. Continuous time models of phenotypic evolution were considered in [4, 8, 21, 24].

Although the biologists agree that assortative mating constitutes a key component of the speciation process (see [15], [19], references therein), it is rather difficult to give a precise proof of this conjecture. Some numerical results presented in the paper [8] suggest that in the case of assortative mating phenotypic profiles converges to multimodal limit distributions, which means that the population becomes polymorphic. The main goal of our paper is to show that the presented model has a similar property, and the speciation is achieved without mutation or external selective pressure. Precisely, we prove that the distribution of traits converges to a combination of Dirac measures. A similar behavior was observed only in a few biological models, e.g., in a Rubinow maturity-structured type equation [16], in an asexual phenotype-structured models [12, 17], and in a nonlinear age-structured model of semelparous species [20].

The scheme of the paper is the following. In Section 2 we introduce our model. Section 3 contains a theorem on the convergence to a combination of Dirac measures. The next two sections are devoted to the answer to the following question. Does there exist an initial measure $\mu$ supported on some interval such that the sequence of $(P^n\mu)$ converges to a combination of at least two Dirac measures? In Section 4 we present an example which precisely answer this question. The numerical simulations suggesting that such behavior is rather typical are presented in Section 5. Also in this section we present some simulations when the offspring trait is distributed around parental traits and we show that the limit distribution can be multimodal. In the last section we comment our results and discuss problems for future investigation.

2. Model. We consider a hermaphroditic population with sexual and asexual reproduction. We start with an individual-based model. We assume that every individual is described by a phenotypic trait $x \in F$, where $F$ is a closed and bounded interval with non-empty interior. Consider a population which consists of $n$ individuals with traits $x_1, \ldots, x_n$. The process of sexual reproduction depends on a preference function $\psi: F \times F \to [0, 1]$, which is any measurable function. We suppose that if individuals with traits $x$ and $y$ encounter, then they mate with probability $\psi(x, y)$ assuming that the first one plays the role of male and the second female. We also assume that different individuals encounter randomly, thus, individuals of traits $x_i, x_j, i \neq j$, mate with rate

$$m(x_i, x_j) = \frac{\psi(x_i, x_j) + \psi(x_j, x_i)}{n - 1} \quad (1)$$
and

\[ m(x_i, x_i) = 1 - \frac{1}{2} \sum_{j \neq i} m(x_i, x_j) \]  

(2)

is a self-fertilization or vegetative reproduction rate. After mating/self-fertilization an offspring is born with probability 1, it survives till the next reproduction. We also assume that there is no overlap between successive generations. The above assumptions seem to be strong because in each generation, each adult individual is replaced by an offspring and the size of the population remains constant. We neglect here questions concerning control of the population size and intraspecific competition to make the model as simple as possible to be mathematically tractable. We discuss more advanced models including these questions in the last section.

Consider the case of sexual reproduction and let \( x \) and \( y \) be parental traits. We assume that the distribution of trait of the offspring is the random variable \( \xi_{x,y} = \frac{x + y}{2} + \frac{|x - y|}{2} \xi_{x,y} \),

(3)

where \( \xi_{x,y} \) is a random variable with values in the interval \([-1, 1]\) such that

\[ \mathbb{E}\xi_{x,y} = 0 \quad \text{and} \quad \mathbb{E}\xi_{x,y}^2 \leq c < 1, \]

(4)

where \( c \) is a constant independent of \( x \) and \( y \). It means that the trait of the offspring is distributed between traits of parents, the expected offspring’s trait is the parental mean trait and the offspring can have a different trait than parents. By \( K(x, y, dz) \) we denote the distribution of the random variable \( \xi_{x,y} \). In the case of self-fertilization the offspring has the same trait as the parent, which means that there is no mutation.

By increasing the number of individuals to infinity, we obtain a nonlinear operator which describes the evolution of phenotypic distribution. Denote by \( \mathcal{M} \) the set of all probability Borel measures on \( F \). Let the measure \( \mu \) be the distribution of traits in one generation. We denote by \( P\mu \) the distribution of traits in the next generation. The operator \( P: \mathcal{M} \to \mathcal{M} \) is given by the formula

\[
P\mu(dz) = \int_F \int_F \psi(x, y)K(x, y, dz) \mu(dx) \mu(dy) \]

\[ + \int_F \int_F \left( \frac{1}{2} \delta_x(dz) + \frac{1}{2} \delta_y(dz) \right) \left( 1 - \psi(x, y) \right) \mu(dx) \mu(dy). \]

(5)

In (5) the term \( \int_F \int_F \psi(x, y)K(x, y, dz) \mu(dx) \mu(dy) \) denotes the distribution of newborn individuals in the process of sexual reproduction while the last term is related to vegetative reproduction. The operator \( P \) can be also written in the following way

\[
P\mu(dz) = \mu(dz) + \int_F \int_F \left( K(x, y, dz) - \frac{1}{2} \delta_x(dz) - \frac{1}{2} \delta_y(dz) \right) \psi(x, y) \mu(dx) \mu(dy).
\]

3. Convergence to a combination of Dirac measures. Denote by \( \mathcal{M}_{\delta,d} \) the set of probability Borel measures on \( F \) which are convex combinations of Dirac measures \( \delta_i \) supported on sets of isolated points with distance at least \( d \), i.e.

\[
\mathcal{M}_{\delta,d} = \{ \nu \in \mathcal{M}: \nu = c_1 \delta_{x_1} + \cdots + c_n \delta_{x_n}, |x_i - x_j| \geq d \text{ if } i \neq j, \ n \in \mathbb{N} \}.
\]

**Theorem 3.1.** Assume that \( \psi(x, y) > 0 \) if \( |x - y| < d \). Then for any measure \( \mu \in \mathcal{M} \) there exists a measure \( \nu \in \mathcal{M}_{\delta,d} \) such that the sequence \( (P^n\mu) \) converges weakly to \( \nu \). Moreover \( \int_F x \mu(dx) = \int_F x \nu(dx) \).
The proof is based on the method of moments (see e.g. Section 30 of [2]). The idea of the proof is similar to that used in [11, Theorem 4.2].

Proof. Without lost of generality we can assume that $F \subseteq [0, \infty)$. Let $\mu \in \mathcal{M}$ and $h: F \to \mathbb{R}$ be a continuous function. Then

$$\int_F h(z) \, P_\mu(dz) \leq \int_F h(z) \, \mu(dz) + \int_F \int_F h_\nu(x, y) \psi(x, y) \, \mu(dx) \, \mu(dy), \quad (6)$$

where

$$h_\nu(x, y) = \int_F h(z) K(x, y, dz) - \frac{1}{2} h(x) - \frac{1}{2} h(y). \quad (7)$$

In particular, if $h(x) = x$, then $h_\nu(x, y) = 0$ and, consequently, the measures $\mu$ and $P_\mu$ have the same first moment. If $h(x) = x^2$, then

$$h_\nu(x, y) = \frac{(x + y)^2}{4} + \frac{(x - y)^2}{4} \mathbb{E}_{x, y} z^2 - \frac{x^2 + y^2}{2} \leq \frac{(x - y)^2 (c - 1)}{4} \quad (8)$$

and from (6) it follows that

$$\int_F z^2 \, P_\mu(dz) \leq \int_F z^2 \, \mu(dz) - \int_F \int_F \frac{1}{2} (x - y)^2 (1 - c) \psi(x, y) \, \mu(dx) \, \mu(dy), \quad (9)$$

If $h$ is a convex function, then

$$\mathbb{E} h(\zeta_{x, y}) \leq \frac{1}{2} (h(x) + h(y)). \quad (10)$$

Indeed, since $\zeta_{x, y}$ has values between $x$ and $y$ we have

$$h(\zeta_{x, y}) \leq \frac{\zeta_{x, y} - y}{x - y} \, h(x) + \frac{\zeta_{x, y} - x}{y - x} \, h(y),$$

and therefore

$$\mathbb{E} h(\zeta_{x, y}) \leq \frac{\mathbb{E} \zeta_{x, y} - y}{x - y} \, h(x) + \frac{\mathbb{E} \zeta_{x, y} - x}{y - x} \, h(y) = \frac{1}{2} (h(x) + h(y))$$

because $\mathbb{E} \zeta_{x, y} = (x + y)/2$. From (7) and (10) it follows that $h_\nu(x, y) \leq 0$ for all $x, y \in F$, and consequently, (6) gives

$$\int_F h(z) \, P_\mu(dz) \leq \int_F h(z) \, \mu(dz). \quad (11)$$

For a positive integer $k$, we denote by $m_k(\mu)$ the $k$-th moment of the measure $\mu \in \mathcal{M}$, i.e. $m_k(\mu) = \int_F x^k \, \mu(dx)$. Since $F \subseteq [0, \infty)$, the function $h(x) = x^k$ is convex and, therefore, $m_k(P_\mu) \leq m_k(\mu)$ for $\mu \in \mathcal{M}$. We now take a $\mu \in \mathcal{M}$ and consider the sequence $\mu_n = P_n \mu$, $n \geq 0$. Since $F$ is a compact set, this sequence is tight and, therefore, for every subsequence $(\mu_{n_k})$ there exists a further subsequence $(\mu_{n_{k(i)}})$ convergent weakly to some measure $\nu \in \mathcal{M}$. We check that $\nu \in \mathcal{M}_{\delta, d}$. Indeed, from (9) it follows that

$$0 \leq m_2(P^{m+1} \mu) \leq m_2(\mu) - \frac{1}{4} (1 - c) \sum_{n=0}^m a_n,$$

where

$$a_n = \int_F \int_F (x - y)^2 \psi(x, y) \, \mu_n(dx) \, \mu_n(dy).$$
From the above it follows that the series $\sum_{n=0}^{\infty} a_n$ is convergent and, therefore, $\lim_{n \to \infty} a_n = 0$. Since the measure $\nu$ is the limit of some subsequence of $(\mu_n)$ we deduce that

$$\int_F \int_F (x - y)^2 \psi(x, y) \nu(dx) \nu(dy) = 0.$$  \hspace{1cm} (12)

We claim that (12) and the assumption $\psi(x, y) > 0$ if $|x - y| < d$ imply that $\nu \notin M_{\delta,d}$. Suppose, contrary to our claim, that $\nu \notin M_{\delta,d}$. Then there exist $\varepsilon > 0$ and points $x_0, x_1 \in F$ such that $2\varepsilon < |x_0 - x_1| < d - 2\varepsilon$, and $x_0, x_1 \in \text{supp } \nu$. Then $\nu((x_i - \varepsilon, x_i + \varepsilon)) > 0$ for $i = 0, 1$. But since for $x \in (x_0 - \varepsilon, x_0 + \varepsilon)$ and $y \in (x_1 - \varepsilon, x_1 + \varepsilon)$ we have $0 < |x - y| < d$, the integral in (12) is positive, a contradiction. It remains to check that $\nu$ is unique. Since the sequence $m_k(\mu_n)$, $n \geq 0$, is decreasing and bounded from below by zero, each moment is convergent to some nonnegative number. Now, the uniqueness of the measure $\nu$ can be checked by using Theorem 30.1 of [2] which gives conditions under which a distribution is uniquely determined by its moments, but to make the paper self-contained we use another simpler argument. If two measures $\nu_1$ and $\nu_2$ from the set $M_{\delta,d}$ have the same moments, then all moments the signed measure $\bar{\nu} = \nu_1 - \nu_2$ are zero. Since $\bar{\nu} = \sum_{i=1}^{n} c_i \delta_{x_i}$, where $c_i \in \mathbb{R}$ and $x_1, \ldots, x_n$ are different real numbers, then the equations for moments $j = 0, 1, \ldots, n - 1$ are following

$$c_1 x_1^j + \cdots + c_n x_n^j = 0, \quad \text{for } j = 0, 1, \ldots, n - 1.$$  \hspace{1cm} (13)

The matrix $X = [x_i^j]$ is a Vandermonde matrix and it has a nonzero determinant. Thus the system (13) has the unique solution $c_1 = \cdots = c_n = 0$. It means that $\nu_1 = \nu_2$ and $\nu$ is uniquely determined by its moments.

\textbf{Remark 1.} An essential role in the proof of Theorem 3.1 is played by the assumption that the offspring trait has values between the parental traits – see assumption (3). Without this assumption we expect that the sequence $(P^n \mu)$ converges weakly to some measure $\nu$ but $\nu$ does not belong to $M_{\delta,d}$. Some numerical simulations suggest that the limit distributions is multimodal.

Now we study a special case of operator $P$ which is related to assortative mating. Let

$$\psi(x, y) = \varphi(|x - y|)$$

where $\varphi: [0, \infty) \to [0, \infty)$ is a non-increasing function such that $\varphi(r) > 0$ for $r \in [0, d)$ and $\varphi(r) = 0$ for $r \geq d$. From Theorem 3.1 it follows immediately that if the length of the interval $F$ is less than $d$, then the sequence of $(P^n \mu)$ converges weakly to $\delta_a$, where $a = \int_F x \mu(dx)$. Assume that the interval $F$ is longer than $d$ and the initial measure $\mu$ is of the form $\mu = \mu_1 + \cdots + \mu_k$ with $\text{diam} \text{supp } \mu_i < d$ for all $i$, and $\text{dist} \text{supp } \mu_i, \text{supp } \mu_j \geq d$ for $i \neq j$, where $\text{supp } \mu$ is the topological support of $\mu$ and $\text{dist}(A, B)$ denotes the distance between sets $A$ and $B$. Then the sequence of $(P^n \mu)$ converges weakly to $\nu = c_1 \delta_{a_1} + \cdots + c_k \delta_{a_k}$, where $c_i = \mu_i(F)$ and $a_i = \int_F x \mu_i(dx)$.

\textbf{4. Example.} Having in mind results of the previous section one can ask the following question: Does there exist an initial measure $\mu$ supported on some interval such that the sequence of $(P^n \mu)$ converges to a combination of at least two Dirac measures? Though numerical simulations suggest that such behavior is rather typical (see Section 5) it is not easy to give a precise mathematical proof.
In order to construct such an example we consider the operator $P$ with $\psi(x, y) = \varphi(|x - y|)$ and $K(x, y, dz) = \delta_{x+y}(dz)$, i.e., the offspring’s trait is the parental mean trait’s. We restrict the operator $P$ to absolutely continuous measures with respect to the Lebesgue measure. If $\mu(dx) = f(x) dx$ then the measure $P\mu$ has also a density given by

$$Pf(x) = f(x) + \int_{-\infty}^{\infty} (f(x + \frac{r}{2})f(x - \frac{r}{2}) - f(x)f(x-r))\varphi(|r|) \, dr,$$

where we formally assume that $f(x) = 0$ for $x \notin F$.

We consider further the case when $F = [-1, 1]$ and $\varphi(r) = 1_{[0,1]}(r)$. Then from (14) it follows that

$$Pf(x) = f(x) \left(1 - \int_{-1}^{x+1} f(r) \, dr\right) + 2 \int_{-1/2}^{1/2} f(x-r)f(x+r) \, dr$$

$$= f(x) \left(1 - \int_{-1}^{x+1} f(r) \, dr\right) + 4 \int_{x}^{x+1/2} f(2x-y)f(y) \, dy.$$ (15)

Let $\alpha_0$, $\eta$, and $\bar{x}$ be positive constants such that $\eta \geq 10$, $\alpha_0 e^{\eta} \leq \frac{1}{10}$, and $\bar{x} < 1$ be such that $1 - \bar{x} \leq \frac{1}{4} e^{-\eta}$. Define a functional $I: L^1([-1, 1]) \to \mathbb{R}$ by

$$I(f) = \int_{-\bar{x}}^{\bar{x}} e^{\eta (x-y)} f(y) \, dy.$$ 

Let a density $f_0$ be an even function such that

$$f_0(x) \leq \alpha_0 e^{\eta|x|} \text{ for } |x| \leq \bar{x} \text{ and } I(f_0) \leq \frac{1}{32}.$$ 

Using inductive arguments we show that

$$P^n f_0(x) \leq \alpha_n e^{\eta|x|} \text{ for } |x| \leq \bar{x},$$ (16)

where the sequence $(\alpha_n)$ converges to zero as $n \to \infty$ and, in consequence, the sequence $(\mu_n)$ of measures with densities $P^n f_0$ converges weakly to a combination of two Dirac measures. Since $P^n f_0$ are also even functions, it is sufficient to check (16) for $x \in [0, \bar{x}]$.

From (15) it follows that

$$Pf(x) \leq \frac{1}{2} f(x) + 2 \int_{-1/2}^{1/2} f(x-r)f(x+r) \, dr$$ (17)

for any even density $f$. Now let $f$ be an even density such that $f(x) \leq \alpha e^{\eta|x|}$ for $|x| \leq \bar{x}$, $I(f) \leq \frac{1}{10}$, and $\alpha > 0$ satisfies the inequality $\alpha e^{\eta} \leq \frac{1}{10}$. We estimate the function $Pf$ from above in the interval $[0, \bar{x}]$. We have

$$2 \int_{-1/2}^{1/2} f(x-r)f(x+r) \, dr \leq 4\alpha^2 \int_0^{1/2} e^{\eta(|x-r| + |x+r|)} \, dr + Sf(x),$$ (18)

where $Sf(x) = 0$ for $x \in [0, \bar{x} - 1/2]$ and $Sf(x) = 2 \int_{\bar{x}-x}^{1/2} f(x-r)f(x+r) \, dr$ for $x \in (\bar{x} - 1/2, \bar{x}]$. Since $\alpha \leq \frac{1}{16} e^{-\eta}$ and

$$\int_0^{1/2} e^{\eta(|x-r| + |x+r|)} \, dr \leq \frac{1}{2} e^{\eta(1+x)} \text{ for } \eta \geq 2 \text{ and } x \in [0, 1],$$
the first term on the right hand side of inequality (18) is bounded from above by 
$\alpha e^{\eta x}/8$. Now we estimate $Sf(x)$. First we consider the case $x \in [\bar{x}-1/2, 1/2]$. Then 
$|x - r| \leq 1/8$ if $r \in [\bar{x} - x, 1/2]$. Hence

$$Sf(x) \leq \alpha e^{\eta/8} \leq \frac{1}{8} \alpha e^{3\eta/8} \leq \frac{1}{8} \alpha e^{\eta x}.$$ 

If $x \geq 1/2$, then

$$Sf(x) \leq 2f(x) \int_{\bar{x} - x}^{1/2} e^{-\eta r} f(x + r) \, dr = 2f(x) \int_{\bar{x}}^{1} e^{\eta(x-y)} f(y) \, dy \leq 2f(x)I(f).$$

Since $I(f) \leq 1/16$ we have $Sf(x) \leq \frac{1}{4} f(x) \leq \frac{1}{8} \alpha e^{\eta|x|}$ for $x \in [0, \bar{x}]$. Thus

$$Pf(x) \leq \frac{3}{4} \alpha e^{\eta|x|} \text{ for } x \in [0, \bar{x}] .$$

Now we show that

$$I(Pf) \leq (1 + \alpha)I(f). \ (20)$$

Indeed,

$$I(Pf) - I(f) = \int_{\bar{x}}^{1} e^{\eta(x-y)} \left( \int_{\bar{x}}^{1} 4f(2y-z) f(y) \, dz - f(y) \int_{y-1}^{1} f(r) \, dr \right) \, dy$$

$$= 2J_1 + 2J_2 - J_3 - J_4,$$

where

$$J_1 = \int_{\bar{x}}^{1} \int_{2x-t}^{\bar{x}} e^{\eta(\bar{x}-(s+t)/2)} f(s)f(t) \, ds \, dt, \quad J_3 = \int_{\bar{x}}^{1} \int_{1-t}^{\bar{x}} e^{\eta(\bar{x}-t)} f(s)f(t) \, ds \, dt,$$

$$J_2 = \int_{\bar{x}}^{1} \int_{2x-t}^{\bar{x}} e^{\eta(\bar{x}-(s+t)/2)} f(s)f(t) \, ds \, dt, \quad J_4 = \int_{\bar{x}}^{1} \int_{\bar{x}}^{1} e^{\eta(\bar{x}-t)} f(s)f(t) \, ds \, dt.$$

We have

$$J_4 = \int_{\bar{x}}^{1} \int_{\bar{x}}^{1} \frac{1}{2} \left( e^{\eta(\bar{x}-t)} + e^{\eta(\bar{x}-s)} \right) f(s)f(t) \, ds \, dt$$

$$= 2 \int_{\bar{x}}^{1} \int_{\bar{x}}^{t} \frac{1}{2} \left( e^{\eta(\bar{x}-t)} + e^{\eta(\bar{x}-s)} \right) f(s)f(t) \, ds \, dt$$

$$\geq 2 \int_{\bar{x}}^{1} \int_{\bar{x}}^{t} e^{\eta(\bar{x}-(s+t)/2)} f(s)f(t) \, ds \, dt = 2J_2.$$

Thus

$$I(Pf) - I(f) \leq 2J_1 = 2 \int_{\bar{x}}^{1} \int_{2x-t}^{\bar{x}} e^{\eta(\bar{x}-(s+t)/2)} f(s)f(t) \, ds \, dt.$$

Since $f(s) \leq \alpha e^{\eta s}$ for $s \in [0, \bar{x}]$ we have

$I(Pf) - I(f) \leq 2\alpha e^{\eta \bar{x}} \int_{\bar{x}}^{1} f(t) \int_{2x-t}^{\bar{x}} e^{\eta(s-t)/2} \, ds \, dt$

$$= 2\alpha e^{\eta \bar{x}} \int_{\bar{x}}^{1} f(t) 2\eta^{-1} \left( e^{\eta(\bar{x}-t)/2} - e^{\eta(\bar{x}-t)} \right) \, dt$$

$$= \alpha \int_{\bar{x}}^{1} e^{\eta(\bar{x}-t)} f(t) \left[ 4\eta^{-1} \left( e^{\eta(t+\bar{x})/2} - e^{\eta^2} \right) \right] \, dt$$

From inequality $1 - \bar{x} \leq \frac{1}{4} e^{-\eta}$ it follows that the expression in the square brackets is less than 1, which implies that (20) holds. Now let $\alpha_n = \left(3\frac{1}{2}\right)^n \alpha_0$. Since $I(f_0) \leq \frac{1}{32}$ and $\prod_{n=0}^{\infty} (1 + \alpha_n) < 2$ from (19) and (20) it follows that $P^n f_0(x) \leq \alpha_n e^{\eta|x|}$ for
Figure 1. Evolution of trait distribution with $K(x, y, dz) = \delta_{x+y}(dz)$ and the preference function $\varphi(r) = 1$, for $r < 1$ and 0 otherwise. The initial function is similar to $f_0$ in Example.

$x \in [-\bar{x}, \bar{x}]$ and for all $n$. Therefore, the sequence $(\mu_n)$ of measures with densities $P^nf_0$ converges weakly to a combination of two Dirac measures.

**Remark 2.** The proof of convergence to a combination of two Dirac measures is quite technical. The assumptions on the initial function $f_0$ are rather restrictive in order to have precise and reasonably simple proof, but it is clear that other similar functions, which have two peaks more distant then 1 and less distant then two, will have the same behavior. To illustrate the shape of $f_0$, in Figure 4, we show a simulation of evolution with the initial density given by $f_0(x) = C e^{10|x|}$, where $C$ is a normalization constant.

5. **Numerical simulations.** In this section we present some interesting results of numerical simulations. The aim of these simulations is mainly to illustrate the fact, that for many initial distributions the sequence of iterations converges to a multimodal measure.

5.1. Model with the offspring trait between the parental traits. We assume, as in Section 4, that $\psi(x, y) = \varphi(|x - y|)$ and $K(x, y, dz) = \delta_{x+y}(dz)$ and we start from a measure that is absolutely continuous with respect to the Lebesgue measure with some density $f_0$. Then the operator on densities has the form (14).

In Figure 2, we consider three different preference functions: $\varphi_1(r) = 1$, for $r < 1$ and 0 otherwise, which means that the individuals mate always if they are within the range of 1, and self-fertilization or vegetative reproduction occurs only for those which are more distant then 1; the second example is given by the so called beta distribution function $\varphi_2(r) = (r-1)^2(r+1)^2$ for $r < 1$, which resembles a truncated Gaussian function; the last one decreases faster: $\varphi_3(r) = (1-r)^3$ for $r < 1$. In all three cases the initial population’s trait is uniformly distributed on the interval $[-1.5, 1.5]$. In the first case the limit measure is a single Dirac measure in the middle, whereas for the next two functions the limit consists of two separated Dirac measures.

In Figure 3, we can see four sequences of graphs showing evolution with the same preference function $\varphi_2(r) = (r-1)^2(r+1)^2$ for initial trait uniformly distributed on different intervals.

The question arises: how the limit measure depends on the spread of trait in the initial population. As a partial answer, in Figure 4 we present a kind of “bifurcation graph” in which we show the dependence of positions of Dirac measures in limit distributions on the width of the support of initial measure.
5.2 Model with a wider distribution of the offspring trait. Although in Section 2 we assumed that the trait of offspring is between the parental traits, the simulations show the interesting asymptotic behavior without this assumption. We consider the case when the trait of the offspring after sexual mating is distributed around the average of parents' traits with some fixed symmetric distribution, that do not depend on the parental traits. Let this distribution have a density \( \kappa \), so \( K(x, y, dz) = \kappa \left( z - \frac{x + y}{2} \right) dz \). In this case, as before, if \( \mu \) has a density \( f \) then the density of \( P\mu \) is

\[
Pf(z) = \int \varphi(|x - y|)\kappa \left( z - \frac{x + y}{2} \right) f(x)f(y) \, dx \, dy,
\]

\[+ \int (1 - \varphi(|z - y|))f(z)f(y) \, dy,
\]

In all the above simulations, the offspring trait is exactly the average of parental traits, which case is just the simplest one to simulate. Apparently, we could use any other distribution of offspring trait that satisfies the assumptions of Theorem 3.1. Actually, we have run some simulations with the offspring trait distributed uniformly between parental traits and the results were very similar, although the convergence to singular measures is a little slower and one need a bit wider initial distributions to obtain the same number of separate subpopulations.
Figure 3. Evolution of trait distribution with \( K(x, y, dz) = \delta_{x+y} (dz) \) and \( \varphi(r) = (r - 1)^2 (r + 1)^2 \), for \( r < 1 \) and 0 otherwise. The initial trait is uniformly distributed on the intervals of length 2.5, 3, 4.3, and 6 in subsequent rows.

Figure 4. A “bifurcation graph” of positions of the limit Dirac measures with respect to the size of the support of initial function.
ASSORTATIVE MATING 469

ASSORTATIVE MATING 469

ASSORTATIVE MATING 469

ASSORTATIVE MATING 469

Figure 5. Evolution of trait distribution with $K(x, y, dz) = \kappa(z - \frac{x + y}{2}) dz$ with probability distribution $\kappa(r) = C_a(r - a)^2(r + a)^2$, for $|r| \leq a$, where $a = 0.125$, $a = 0.25$, $a = 0.5$ in subsequent rows.

The evolution of the trait distribution for this operator is shown in Figure 5. In all three cases, the preference function is the same, $\varphi(r) = (r - 1)^2(r + 1)^2$ for $r < 1$, and the initial trait is uniformly distributed on an interval of length 4. The variation of offspring’s trait from the parent’s average has the beta distribution $\kappa(r) = C_a(r - a)^2(r + a)^2$, for $|r| \leq a$, where $C_a$ is a normalizing constant. We can see that if the spread of offspring trait is much narrower than the maximal mating distance, then the limit distribution is multimodal.

Finally, let us assume that the trait of all offspring, even those originating from self-fertilization, is distributed with some probability density. This may be caused by mutations. Such an assumption implies the following form of the operator on densities:

$$P f(z) = \int \int \varphi(|x - y|) \kappa \left(z - \frac{x + y}{2}\right) f(x)f(y) \, dx \, dy,$$

$$+ \int \int (1 - \varphi(|x - y|)) \kappa(z - y) f(x)f(y) \, dx \, dy.$$

Figure 6 suggests that in this case the width of offspring’s trait distribution should be very small to obtain a multimodal limit distribution.

6. Discussion. In the paper we considered a very simple phenotype structured population model with assortative mating. The model is given by a nonlinear operator acting on the space of probability measures and describes the relation between parental and offspring trait distributions. We prove that if the offspring trait has values between the parental traits than the limit distribution of trait converges to a combination of Dirac measures and such a combination can be non-trivial (i.e. it is a combination of at least two Dirac measures). This result means that assortative mating can lead to a polymorphic population and sympatric speciation. We also present some numerical simulations which show how a uniformly distributed
population evolves to become distinct species. We also provide some simulations in the case when the range of the offspring trait is wider the interval between parental traits. Then the limit distribution can be multimodal.

Our model is rather simple and does not contain many important components. It will be not difficult to make it more realistic by adding birth and death rates which depend on the traits of mating individuals and the total size of the population $m_n = \mu_n(F)$ or by considering a model with a competition kernel (see e.g. [5, 6, 8, 12]). In the case when birth and death rates depend only on the total size, the properties of the model remain the same if we replace $\mu_n$ by the trait profile $\mu_n/m_n$. In the case when birth and death rates depend also on the trait of an individual, the trait distribution will concentrate around the local maximum of the reproduction rate. Assortative mating will sustain this phenomenon. It means that if we consider a model with both disruptive or competition selection and with assortative mating it will be difficult to decide which factor is responsible for sympatric speciation. If we consider a model only with assortative mating, then all traits are equally favorable and one can expect that the limit distribution will be uniform or will concentrate at the Dirac measures located in the mean of the initial distribution, but we show the picture is more complex. Moreover, one of the main biological arguments that assortative mating can have influence on the speciation process is that rare phenotypes experience less competition than common phenotypes [7]. On the other hand such mating can reduce the number of individuals with rare traits because they can have a problem finding partners, and therefore it seems sensible to consider a simplified model neglecting important ecological aspects as competition for resources and partners.

The proof of the convergence to a combination of Dirac measures is essentially based on the assumption that the offspring are distributed between the phenotypic values of the parents. It is clear that this assumption is an oversimplification and it should be replaced by, for example, that the offspring distribution is close to normal with a mean coinciding with the parents’ mean. This problem is discussed in details in [1]. Simulation presented in Section 5.2 confirm that we expect in such a case multimodality of phenotype distributions. But in this case a serious mathematical problem appears: how to define the class of multimodal distributions,
which additionally could explain sympatric speciation. It is why we only make a first step in this direction considering a model with such a strong assumption concerning offspring traits, but we hope to find a mathematical method of studying a model with different assumptions. It would be also interesting to investigate a continuous time model described in the last section of [21]. This model is more realistic but also more difficult to study. Till now we have only proved a version of Theorem 3.1 for this model but we are still unable to construct an example similar to that considered in Section 4.

Acknowledgments. This research was partially supported by the National Science Centre (Poland) Grant No. 2014/13/B/ST1/00224.

REFERENCES

[1] N. H Barton, A. M. Etheridge and A. Véber, The infinitesimal model: Definition, derivation, and implications, Theor. Popul Biol., (2017).
[2] P. Billingsley, Probability and Measure, John Wiley and Sons, New York, 1986.
[3] M. G. Bulmer, The Mathematical Theory of Quantitative Genetics, Clarendon Press, Oxford, 1980.
[4] J. A. Cañizo, J. A. Carrillo and S. Cuadrado, Measure solutions for some models in population dynamics, Acta Appl. Math., 123 (2013), 141–156.
[5] L. Desvillettes, P. E. Jabin, S. Mischler and G. Raoul, On selection dynamics for continuous populations, Commun. Math. Sci., 6 (2008), 729–747.
[6] O. Diekmann, P. E. Jabin, S. Mischler and B. Perthame, The dynamics of adaptation: An illuminating example and a Hamilton-Jacobi approach, Theor. Popul Biol., 67 (2005), 257–271.
[7] U. Dieckmann and M. Doebeli, On the origin of species by sympatric speciation, Nature, 400 (1999), 354–357.
[8] M. Doebeli, H. J. Blok, O. Leimar and U. Dieckmann, Multimodal pattern formation in phenotype distributions of sexual populations, Proc. R. Soc. B, 274 (2007), 347–357.
[9] R. A. Fisher, The correlations between relatives on the supposition of Mendelian inheritance, Trans. R. Soc. Edinburgh, 52 (1919), 399–433.
[10] S. Gavrilets and C. R. B. Boake, On the evolution of premating isolation after a founder event, The American Naturalist, 152 (1998), 706–716.
[11] P. Hinow, et al., Analysis of a model for transfer phenomena in biological populations, SIAM J. Appl. Math., 70 (2009), 40–62.
[12] P. Jabin and G. Raoul, On selection dynamics for competitive interactions, J. Math. Biol., 63 (2011), 493–517.
[13] F. A. Kondrashov and A. S. Kondrashov, Interactions among quantitative traits in the course of sympatric speciation, Nature, 400 (1999), 351–354.
[14] C. Matessi, A. Gimelfarb and S. Gavrilets, Long-term buildup of reproductive isolation promoted by disruptive selection: How far does it go?, Selection, 2 (2001), 41–64.
[15] J. Maynard Smith, Sympatric speciation, Am. Nat., 100 (1966), 637–650.
[16] P. Michel, A singular asymptotic behavior of a transport equation, C. R. Acad. Sci. Paris, Ser. I, 346 (2008), 155–159.
[17] B. Perthame, Transport Equations in Biology, Frontiers in Mathematics, Birkhäuser, Basel, 2007.
[18] J. Polechová and N. H. Barton, Speciation through competition: A critical review, Evolution, 59 (2005), 1194–1210.
[19] O. Puebla, E. Berringham and F. Guichard, Pairing dynamics and the origin of species, Proc. Biol. Sci., 279 (2012), 1085–1092.
[20] R. Rudnicki and R. Wieczorek, On a nonlinear age-structured model of semelparous species, Discrete Contin. Dyn. Syst. Ser. B, 19 (2014), 2641–2656.
[21] R. Rudnicki and P. Wołowiński, Model of phenotypic evolution in hermaphroditic populations, J. Math. Biol., 70 (2015), 1295–1321.
[22] K. A. Schneider and R. Bürger, Does competitive divergence occur if assortative mating is costly?, Journal of Evolutionary Biology, 19 (2006), 570–588.
[23] K. A. Schneider and S. Peischl, Evolution of assortative mating in a population expressing dominance, *PLoS ONE*, 6 (2011), e16821.

[24] P. Zwoleński, Trait evolution in two-sex populations, *Math. Mod. Nat. Phenom.*, 10 (2015), 163–181.

Received October 2016; revised February 2017.

E-mail address: rudnicki@us.edu.pl
E-mail address: radoslaw.wieczorek@us.edu.pl