APPLICATIONES MATHEMATICAE
27,1 (2000), pp. 81–101

A. MARCINIAK (Warszawa)

DIFFUSION LIMIT FOR THE PHENOMENON OF RANDOM GENETIC DRIFT

Abstract. The paper deals with mathematical modelling of population genetics processes. The formulated model describes the random genetic drift. The fluctuations of gene frequency in consecutive generations are described in terms of a random walk. The position of a moving particle is interpreted as the state of the population expressed as the frequency of appearance of a specific gene. This leads to a continuous model on the microscopic level in the form of two first order differential equations (known as the telegraph equations). Applying the modified Chapman–Enskog procedure we show the transition from this system to a macroscopic model which is a diffusion type equation. Finally, the error of approximation is estimated.

1. Introduction. This work is devoted to mathematical modelling of processes in population genetics. Population genetics involves investigations of gene distribution in populations and analysis of changes of specific genes proportions in consecutive generations. Population genetics, as a scientific discipline, has an importance for medicine due to the consequences of genes dynamics for human being evolution [4], [5].

Basing on some geneticists' papers [8], [6] and Wright–Moran's mathematical model [7], [13] I assume that fluctuations of gene frequency can be treated as a "random walk" phenomenon. In this model, the position of a randomly moving particle is equivalent to the state of the population expressed in terms of the frequency of appearance of the gene considered. The state space is therefore the real interval $[0,1]$. Further, I assume that there exists a positive correlation between two consecutive steps of the random walk. Under these assumptions the probability density of the population being in a particular state is given by a system of first order differential

2000 *Mathematics Subject Classification*: Primary 35B25.
Key words and phrases: genetic drift, random walk, singularly perturbed systems, telegraph equations, diffusion equation.

equations, known as the telegraph equations. Then I show that the solution of that system can be approximated by the solution of an appropriate diffusion type equation.

The formulated model describes the phenomenon of genetic drift, which still intrigues the geneticists. I study mathematical connections between various descriptions of this phenomenon including the random walk description, the telegraph equations description, and the description by an appropriate diffusion equation.

2. Mathematical models in population genetics

2.1. *Genetic basics.* We confine ourselves to the case of the simplest population consisting of diploid individuals. In every diploid cell, both chromosomes and particular genes are coupled. A gene located at a specific position of a chromosome has its counterpart (allele) in another homologous chromosome. Thus, there is always a pair of genes responsible for any feature or disease with Mendel's type of heritage.

As is customary, I will mark dominant genes with capital letters, and recessive ones with small letters. In the simplest case of only two alleles, A and a, gametes can produce three different genotypes (pairs of genes): AA, aa, Aa. The difference between Aa and aA configurations is of no importance because we cannot distinguish between the genes in homologous pairs. AA and aa are homozygotes while Aa is a heterozygote. A diploid cell can give the gamete (a reproductive cell) only one allele from a pair. The genotypes of progeny are created in a random way. The Mendelian segregation law states that every parental gene is transmitted to an offspring genotype with probability 1/2. For example, the possible genotypes formed from the crossing $Aa \times Aa$ are AA, Aa and aa, with probabilities respectively: 1/4, 1/2, 1/4, but the crossing $AA \times aa$ results only in a hybrid offspring Aa.

The rule of random crossing assumes that all processes of parental pairs formation are stochastically independent and equally probable. This is certainly valid only in very large populations and in the absence of any selection. The fundamental mechanism of evolution is based on the changes in the frequency of genes.

The simplest mathematical approach to the problem is to treat the process as a deterministic one [6]. This kind of description is acceptable for very large populations and special mating conditions.

The Hardy–Weinberg law says that under such circumstances, random mating results in unchanging probability distribution of genotypes and genes in each consecutive generation [5]. This law is commonly used to calculate the frequencies of appearance of recessive features carriers, for example phenylketonuria, mucoviscidosis, albinism, inborn deafness, and others [4].

In practice there are, however, additional important factors influencing the genetic variety of a population. Apart from nonrandom mating (sib mating for example) there are also: mutation, migration and selection.

2.2. *Phenomenon of genetic drift.* Genetic drift as a factor of evolution is still a controversial issue among geneticists. It can be defined as a process of gene frequency variation in a natural finite population due to the random nature of transmitting alleles during reproduction and other accidental random events that influence the allele frequency [8]. This process of random fluctuations is repeated from generation to generation. The changes in subsequent generations do not depend on the initial state ("lack of genetic memory").

Genetic drift can result in falling away of one of alleles and creation of a homozygotic population. It can be the result of population finiteness; it is then called a pure genetic drift. In such a population, in the case of non-numerous offspring, it can happen that gene frequencies are not exactly reproduced in future generations. This can be interpreted as a sampling error. It is worth noticing that falling away of both alleles, a and A, is equally probable. Actually, quite often those probabilities might not be the same as a result of some selections. So, in our further consideration of genetic drift the existence of selections is assumed.

According to some geneticists, genetic drift is one of the most important mechanisms of evolution [8].

The basic model of genetic drift was formulated in terms of finite Markov chains by Wright and Malecot [13]. This model, after modification by Moran [7], can be described as a "random walk with an absorbing boundary". Investigations using Markov chains were carried out by Feller, Karlin, and McGregor. An overview of the known results can be found in [7].

Kimura [8] and Ludwig [9] proposed a continuous stochastic process for modelling the gene frequency changes in time. They assumed that the changes of gene frequency exhibited the lack of memory and therefore could be modelled by a Markov process. Their model has the form of diffusion equations known as the Kolmogorov prospective and retrospective equations, but it is not strictly justified. Nevertheless, using diffusion approximation, a range of genetic problems have been solved ([8], [9]).

While analysing these results, it is of interest to show that the solution of a diffusion equation can be a good approximation of the probability density of finding the population in a given set of states. The state of the population is defined as the ratio of the number of alleles A to the number of all alleles $(2N)$, so the state variable takes values in $[0, 1]$.

3. Setting the problem. Following the above mentioned idea of Moran to conceptualize variations of gene frequency in terms of random walk, I selected for further investigation the model of random walk proposed by Taylor and Goldstein and the model of random walk with correlations investigated by Banasiak and Mika [3] just in the aspect of diffusion approximation. Contrary to the model analysed by Banasiak and Mika, the model formulated in the present paper takes into account the phenomenon of drift. In [3], the diffusion equation did not describe a drift, as the drift coefficient (the coefficient of the first derivative of the space variable x) was equal to 0. That was a consequence of their prior assumption that the probability of taking a specific direction by a moving particle does not depend on its previous direction. In the case of genetic applications a substantial difference can be seen between the probability that the number of alleles A will continue to increase if it has been increasing earlier, and the probability that the number of alleles A will continue to decrease if it has been decreasing earlier. This difference induces some kind of selection.

In the present work these probabilities are taken as different functions. As a result, in the system of telegraph equations there appears an additional term, and in the diffusion equation a nonzero drift coefficient. It can be either positive or negative, which reflects the possibility of genetic drift in both directions.

When describing variations of gene frequency in the language of random walk, the position of a randomly walking particle is identified with the concentration of alleles A in the population. That concentration (frequency) can range from 0, which corresponds to the homozygotic population aa, to 1, which corresponds to the homozygotic population AA. Thus, in the model considered, the state space is the real interval $[0,1]$, where 0 and 1 are absorbing states, because when a population becomes homozygotic, the change of its state is possible only through mutation.

The main objective of the present work is to show that the solution of the diffusion equation with drift, with an appropriately modified initial value, complemented with initial layer equations, approximates the solution of a singularly perturbed telegraph system with drift and that this system is an adequate description of a random walk with correlations.

4. Correlated random walk with drift

4.1. *Formulation of the model.* We consider the problem of correlated random walk. We assume that a particle moves along the interval [0,1]. The location x of the randomly moving particle corresponds to the ratio of the number $X(t)$ of alleles A to the number $2N$ of all alleles in the population. We consider the general case when there exists a positive correlation between two consecutive steps of the random walk that approaches value 1 when the time step approaches 0. This correlation reflects the tendency of particles to move in a given direction with a finite velocity.

We will define two probability density functions $\alpha(x,t)$ and $\beta(x,t)$ corresponding to the particle movement to the right and left, respectively. Denote by $p^+(x)$ the probability that a particle continues to move in the same direction while moving to the right, $q^+(x)$ the probability that a particle changes its direction while moving to the right, and $p^-(x)$, $q^-(x)$ similar probabilities for the initial movement to the left.

The assumption that $p^-(x) = p^+(x)$ and $q^-(x) = q^+(x)$ leads to diffusion approximation with zero drift coefficient ([2], [3]). This corresponds to pure genetic drift without selections.

The probability that a particle "rests" at a given step is independent of the earlier movement direction and equals $1-\sigma(x)$, where $\sigma(x) = p^+(x) + q^+(x) = p^-(x) + q^-(x)$.

Assume that σ is a Lipschitz continuous function and

$$0 < \sigma_0 \le \sigma(x) \le \sigma_1 < \infty, \qquad \forall x \in [0,1]. \tag{4.1}$$

Under the above assumptions we have the following system of difference equations for the functions $\alpha(x,t)$ and $\beta(x,t)$:

$$\begin{aligned} \alpha(x,t+\theta) &= [1-p^+(x)-q^+(x)]\alpha(x,t) + p^+(x-\delta)\alpha(x-\delta,t) \\ &\quad + q^-(x-\delta)\beta(x-\delta,t), \end{aligned} \tag{4.2}$$

$$\begin{aligned} \beta(x,t+\theta) &= [1-p^-(x)-q^-(x)]\beta(x,t) + p^-(x+\delta)\beta(x+\delta,t) \\ &\quad + q^+(x+\delta)\alpha(x+\delta,t). \end{aligned} \tag{4.3}$$

Assuming the presence of positive correlations we expect that as $\theta \to 0$, the probabilities $p^+(x)$ and $p^-(x)$ approach $\sigma(x)$ and the probabilities $q^+(x)$ and $q^-(x)$ of direction reversal approach 0. Thus, it is reasonable to take those probabilities in the form

$$\begin{aligned} p^+(x) &= \sigma(x) - \lambda^+(x)\theta + O(\theta^2), \\ p^-(x) &= \sigma(x) - \lambda^-(x)\theta + O(\theta^2), \end{aligned} \tag{4.4}$$

$$q^+(x) = \lambda^+(x)\theta + O(\theta^2), \qquad q^-(x) = \lambda^-(x)\theta + O(\theta^2), \tag{4.5}$$

where $\lambda^+(x)$ and $\lambda^-(x)$ are the rates of direction reversal. These are related to the strength of correlations in the system. These equations characterize the nature of the correlated process. We assume that $\lambda^+(x)$ and $\lambda^-(x)$ are Lipschitz continuous functions satisfying

$$\begin{aligned} &0 < \lambda_0^+ \le \lambda^+(x) \le \lambda_1^+ < \infty, \\ &0 < \lambda_0^- \le \lambda^-(x) \le \lambda_1^- < \infty, \qquad \forall x \in [0,1]. \end{aligned} \tag{4.6}$$

Expanding the functions α and β in Taylor series, using (4.4) and (4.5) and letting δ, $\theta \to 0$ in such a way that $\gamma = \lim_{\delta\to 0,\, \theta\to 0} \delta/\theta$ remains finite, we obtain the system of equations

$$\partial_t\alpha(x,t) = -\gamma\partial_x[\sigma(x)\alpha(x,t)] - \lambda^+(x)\alpha(x,t) + \lambda^-(x)\beta(x,t), \tag{4.7}$$

$$\partial_t\beta(x,t) = \gamma\partial_x[\sigma(x)\beta(x,t)] - \lambda^-(x)\beta(x,t) + \lambda^+(x)\alpha(x,t). \tag{4.8}$$

For further transformations we assume that the functions $\lambda^+(x)$ and $\lambda^-(x)$ have the form

$$\lambda^+(x) = \lambda(x) - \varphi(x), \qquad \lambda^-(x) = \lambda(x) + \varphi(x), \tag{4.9}$$

where $0 < \lambda_0 \le \lambda(x) \le \lambda_1 < \infty$, and the function φ is bounded.

We introduce two new functions:

$$v(x,t) = \alpha(x,t) + \beta(x,t), \qquad w(x,t) = \alpha(x,t) - \beta(x,t), \tag{4.10}$$

where v is the density and w the net flux of particles to the right.

Finally we get the following system of equations (already referred to as the *telegraph system*):

$$\begin{aligned} \partial_t v + \gamma\partial_x(\sigma w) &= 0,\\ \partial_t w + \gamma\partial_x(\sigma v) + 2\lambda w - 2\varphi v &= 0, \end{aligned} \tag{4.11}$$

with initial conditions

$$v(x,0) = \mathring{v}(x), \qquad w(x,0) = \mathring{w}(x). \tag{4.12}$$

4.2. *Existence and uniqueness of solution of the telegraph equations for the correlated random walk.* The system (4.11) is considered on the interval $[0,1]$ with Dirichlet's homogeneous boundary conditions

$$v(0,t) = v(1,t) = 0, \qquad t > 0. \tag{4.13}$$

The conditions correspond to the absorbing border and hence are appropriate for the model considered as shown in [14]. The solution must satisfy the initial conditions (4.12).

Let us rewrite (4.11) as the following evolution system in the Hilbert space $W^{0,2}_\sigma$:

$$\partial_t \begin{bmatrix} v\\ w\end{bmatrix} = \gamma S\begin{bmatrix} v\\ w\end{bmatrix} + 2\lambda C\begin{bmatrix} v\\ w\end{bmatrix} + 2\varphi B\begin{bmatrix} v\\ w\end{bmatrix}, \tag{4.14}$$

where

$$S = \begin{bmatrix} 0 & -\partial_x\sigma\\ -\partial_x\sigma & 0\end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 0\\ 0 & -1\end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0\\ 1 & 0\end{bmatrix}, \tag{4.15}$$

and $W^{0,2}_\sigma = L^2([0,1]) \times L^2([0,1])$ with the norm

$$\|(v,w)\|^2_\sigma = \int_0^1 |v(x)|^2\sigma(x)\,dx + \int_0^1 |w(x)|^2\sigma(x)\,dx.$$

This norm is equivalent to the standard norm in $L^2([0,1]^2)$ by (4.1).

It will be shown that the operator S is maximal dissipative and hence it generates a semigroup of contractions. Since the operators $2\lambda C$ and $2\varphi B$ are bounded, the operator $K = \gamma S + 2\lambda C + 2\varphi B$ also generates a semigroup. Hence the Cauchy problem related to (4.14) is solvable.

Further, some concepts and theorems from the theory of semigroups of bounded linear operators [12] are used.

LEMMA 4.1. *The operator S defined on $D(S) = W^{1,2}_{0,\sigma}([0,1]) \times W^{1,2}_{0,\sigma}([0,1])$ is dissipative.*

Proof. We have

$$\begin{aligned}
\mathrm{Re}\left(S\begin{bmatrix} v \\ w \end{bmatrix}, \begin{bmatrix} v \\ w \end{bmatrix}\right) &= \int_0^1 \begin{bmatrix} -\partial_x \sigma w \\ -\partial_x \sigma v \end{bmatrix}^T \begin{bmatrix} \overline{v} \\ \overline{w} \end{bmatrix} \sigma\, dx \\
&= -\int_0^1 (\partial_x(\sigma w)\overline{v}\sigma + \partial_x(\sigma v)\overline{w}\sigma)\, dx \\
&= \int_0^1 (\partial_x(\sigma w)\overline{v}\sigma + \partial_x(v\sigma)\overline{w}\sigma)\, dx \\
&= -\mathrm{Re}\left(\begin{bmatrix} v \\ w \end{bmatrix}, S\begin{bmatrix} v \\ w \end{bmatrix}\right).
\end{aligned}$$

So, we obtain the equality $\mathrm{Re}\, z = -\mathrm{Re}\, \overline{z}$. Therefore $\mathrm{Re}\, z = 0$ and the dissipativity of S follows. ∎

LEMMA 4.2. *$R(I - S) = X$.*

Proof. Since S is dissipative we have

$$\|(\lambda I - S)u\| \ge \lambda\|u\| \quad \forall u \in D(S) \text{ and } \lambda > 0.$$

This implies that the range $R(\lambda I - S)$ is closed. Indeed, take a sequence $f_n \in R(\lambda I - S)$ such that $f_n \to f$. Then there is u_n such that $f_n = (\lambda I - S)u_n$. Since $\|f_n\| = \|(\lambda I - S)u_n\| \ge \lambda\|u_n\|$, we have $u_n \to u$, $u \in D(\lambda I - S)$ and $(\lambda I - S)u = f$. This means that $f \in X$. So, it follows that $R(\lambda I - S)$ is closed for all $\lambda > 0$, in particular for $\lambda = 1$.

It is now sufficient to show that $R(I - S)$ is dense, e.g. by showing that

$$\forall \begin{bmatrix} f \\ g \end{bmatrix} \in C_0^\infty \times C_0^\infty\ \exists \begin{bmatrix} v \\ w \end{bmatrix} : \quad (I - S)\begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}.$$

We have

$$\begin{gathered}
v + \partial_x(\sigma w) = f, \qquad w + \partial_x(\sigma v) = g, \qquad w = g - \partial_x(\sigma v), \\
v + \partial_x(\sigma g) - \partial_x(\sigma \partial_x(\sigma v)) = f, \\
v - \partial_x(\sigma \partial_x(\sigma v)) = f - \partial_x(\sigma g).
\end{gathered}$$

So, we get

$$Au = h \quad \text{in } C_0^\infty, \tag{4.16}$$

where

$$h = f - \partial_x(\sigma g), \qquad A = I - \partial_x \sigma \partial_x \sigma. \tag{4.17}$$

Take $\phi \in C_0^\infty$. Multiplying (4.16) by $\phi\sigma$ and integrating by parts we get

$$(Au, \phi) = (h, \phi), \tag{4.18}$$

where

$$\begin{aligned} (Au, \phi) &= -\int_0^1 \partial_x(\sigma\partial_x(\sigma u))\phi\sigma\, dx + \int_0^1 u\phi\sigma\, dx \\ &= \int_0^1 \partial_x(\sigma u)\partial_x(\sigma\phi)\sigma\, dx - \partial_x(\sigma u)\sigma\phi\sigma|_0^1 + \int_0^1 u\phi\sigma\, dx. \end{aligned} \tag{4.19}$$

Now $\partial_x(\sigma u)\sigma\phi\sigma|_0^1$ vanishes because $\phi \in C_0^\infty$. Therefore we have

$$(Au, \phi) = \int_0^1 \partial_x(\sigma u)\partial_x(\sigma\phi)\sigma\, dx + \int_0^1 u\phi\sigma\, dx. \tag{4.20}$$

Now, we show the uniqueness of the solution of the above problem in the space $W_{0,\sigma}^{1,2}$, which is the closure of C_0^∞ under the norm $W_{,\sigma}^{1,2}$.

Let $a(u, \phi) = (Au, \phi)$. From Hölder's inequality and boundedness of σ, it follows that $a(u, \phi)$ is a bounded bilinear form with respect to u, ϕ, and (h, ϕ) is a linear functional. Moreover, $a(\cdot, \cdot)$ is a coercive form, i.e. $\mathrm{Re}(Au, u) \ge c\|u\|^2_{W_{0,\sigma}^{1,2}}$, which follows from the properties of σ. So, the assumptions of the Lax–Milgram theorem are satisfied. Therefore the problem considered has a unique solution. This implies that $R(I - S)$ is dense in X. ■

Remark 4.1. From Lemmas 4.1 and 4.2 it follows that the operator S is m-dissipative and hence it is maximal dissipative.

Lemma 4.3. $\overline{D(S)} = X$.

Proof. The operator S is m-dissipative and $W_{0,\sigma}^{1,2}$ is a reflexive space. Therefore the domain of S is dense in X. ■

Lemma 4.4. *S is the generator of a C_0-semigroup of contractions.*

Proof. This follows immediately from the Lumer–Phillips theorem. The assumptions of that theorem are satisfied by Lemmas 4.1–4.3. ■

Theorem 4.1. *For each pair $(\mathring{v}, \mathring{w}) \in W_{0,\sigma}^{1,2}([0,1]) \times W_{,\sigma}^{1,2}([0,1])$ the Cauchy problem for* (4.14) *has a unique classical solution on* $[0, \infty)$.

Proof. The operator $K = S + 2\lambda C + 2\varphi B$ is the generator of a semigroup denoted by $\{G_K(t)\}_{t\ge 0}$. This results from the fact that S generates a semigroup of contractions and the operators $2\lambda C$ and $2\varphi B$ are bounded (by the form of the matrices C and B and the assumptions on φ). Therefore the Cauchy problem for (4.14) has a unique solution. ■

Now, we show that the semigroup $\{G_K(t)\}_{t\geq 0}$ is bounded, i.e. that there is a constant $M>0$ such that $\|G_K(t)\|\leq M,\ t\geq 0$.

We introduce the notation $A=2\lambda C+2\varphi B$. The operator A is then of the form

$$A=\begin{bmatrix} 0 & 0\\ 2\varphi & -2\lambda\end{bmatrix}. \tag{4.21}$$

For abbreviation we set $\overline{S}=\gamma S$. Equation (4.14) can be rewritten as follows:

$$\partial_t\begin{bmatrix} v\\ w\end{bmatrix}=\overline{S}\begin{bmatrix} v\\ w\end{bmatrix}+A\begin{bmatrix} v\\ w\end{bmatrix}. \tag{4.22}$$

Consider the problem (4.22) for the following functions:

$$\widehat{v}=ve^{-D(x)},\qquad \widehat{w}=we^{-D(x)}, \tag{4.23}$$

where D is a function of x.

Substituting (4.23) into (4.22) we obtain, after some transformations,

$$\partial_t\begin{bmatrix} \widehat{v}\\ \widehat{w}\end{bmatrix}=\overline{S}\begin{bmatrix} \widehat{v}\\ \widehat{w}\end{bmatrix}+\overline{A}\begin{bmatrix} \widehat{v}\\ \widehat{w}\end{bmatrix}, \tag{4.24}$$

where

$$\overline{A}=\begin{bmatrix} 0 & -\gamma D'\sigma\\ -\gamma D'\sigma+2\sqrt{\lambda}\varphi & -2\lambda\end{bmatrix} \tag{4.25}$$

and D' is the first derivative of D.

By Lemma 4.4, S is the generator of a C_0-semigroup of contractions. Since γ is a constant, it is easy to see that $\overline{S}$ is also the generator of a C_0-semigroup of contractions. We will show that there exists D such that $\overline{A}$ is the generator of a C_0-semigroup of contractions.

LEMMA 4.5. *There exists D such that the operator $\overline{A}$ defined on $D(\overline{A})=W^{1,2}_{0,\sigma}([0,1])\times W^{1,2}_{0,\sigma}([0,1])$ is dissipative.*

Proof. We have

$$\begin{aligned}\operatorname{Re}\left(\overline{A}\begin{bmatrix} \widehat{v}\\ \widehat{w}\end{bmatrix},\begin{bmatrix} \widehat{v}\\ \widehat{w}\end{bmatrix}\right)&=\int_0^1\begin{bmatrix} -\gamma D'\sigma\widehat{w}\\ (-\gamma D'\sigma+2\varphi)\widehat{v}-2\lambda\widehat{w}\end{bmatrix}^T\begin{bmatrix} \overline{\widehat{v}}\\ \overline{\widehat{w}}\end{bmatrix}\sigma\,dx\\ &=\int_0^1(-2\gamma D'\sigma\widehat{w}\widehat{v}\sigma+2\varphi\widehat{v}\widehat{w}\sigma-2\lambda\widehat{w}\widehat{w}\sigma)\,dx\leq 0.\end{aligned}$$

The above inequality holds if $-2\gamma D'\sigma^2 + 2\varphi\sigma = 0$, thus if $D'(x) = \varphi_{(x)}/(\gamma\sigma(x))$. Therefore, $\overline{A}$ is dissipative for

$$D(x) = \int_0^x \frac{\varphi_{(s)}}{\gamma\sigma(s)}\,ds + \frac{\varphi_{(0)}}{\gamma\sigma(0)}. \quad\blacksquare \tag{4.26}$$

Further, the operator $\overline{A}$ is considered with the D given by (4.26).

LEMMA 4.6. $R(I - \overline{A}) = X$.

Proof. Since $\overline{A}$ is dissipative we have $\|(\lambda I - \overline{A})u\| \ge \lambda\|u\|$. From this inequality it follows that $R(\lambda I - \overline{A})$ is closed for each $\lambda > 0$, and in particular for $\lambda = 1$. It is sufficient to show that $R(I - \overline{A})$ is dense, e.g. by showing the existence and uniqueness of solution of the problem

$$\begin{bmatrix} 1 & \gamma D'\sigma \\ \gamma D'\sigma - 2\varphi & 1 + 2\lambda \end{bmatrix} \begin{bmatrix} \widehat{v} \\ \widehat{w} \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \quad \text{in } C_0^\infty. \tag{4.27}$$

The above problem has a unique solution if and only if $\det(I - \overline{A}) \ne 0$. But

$$\det(I - \overline{A}) = 1 + 2\lambda - \gamma D'\sigma(\gamma D'\sigma - 2\varphi) = 1 + 2\lambda + \varphi^2 > 0 \quad \forall x. \ \blacksquare$$

LEMMA 4.7. *$\overline{A}$ is the generator of a C_0-semigroup of contractions.*

Proof. From Lemmas 4.5 and 4.6 it follows that $\overline{A}$ is m-dissipative, and hence it is maximal dissipative. Moreover, $\overline{D(\overline{A})} = X$, as in the case of S. Thus, by the Lumer–Phillips theorem, for D given by (4.26), $\overline{A}$ is the generator of a C_0-semigroup of contractions.

THEOREM 4.2. *The semigroup $\{G_K(t)\}_{t\ge0}$ is bounded, so there exists a constant $M > 0$ such that*

$$\|G_K(t)\| \le M, \quad t \ge 0.$$

Proof. By Lemma 4.4, S is the generator of a C_0-semigroup of contractions, thus $\overline{S}$ is the generator of a semigroup of contractions. From Lemma 4.7 it follows that so is $\overline{A}$. Using these facts and the Trotter formula ([12]), we will show that $\overline{K} = \overline{S} + \overline{A}$ is also the generator of a C_0-semigroup of contractions. We have

$$\|(G_{\overline{S}}(t)G_{\overline{A}}(t))^n\| \le \|G_{\overline{S}}(t)\|^n \|G_{\overline{A}}(t)\|^n \le 1. \tag{4.28}$$

It is easy to check that the assumptions of the Trotter formula are satisfied, i.e. that there exists $\mu > 0$ such that $R(\mu I - (\overline{S} + \overline{A}))$ is dense in X. Therefore $\overline{K} = \overline{S} + \overline{A}$ is the generator of a semigroup of contractions, thus $\|G_{\overline{K}}(t)\| \le 1$, and hence

$$\|G_K(t)\| \le M,$$

where M is a positive constant. This follows from the fact that $e^{D(x)}$ is bounded for D given by (4.26) and $x \in [0, 1]$. $\blacksquare$

5. Analysis of the singularly perturbed evolution equation. In the case of weak correlations, i.e. large λ, one can expect that the correlated process of random walk is similar to an uncorrelated one described by diffusion equations.

Applying the theory of singularly perturbed systems we will show that the solution of the diffusion equation with drift approximates the solution of the system of correlated random walk with drift and we will estimate the approximation error.

In the spirit of that theory we now perform an appropriate scaling of the coefficients in (4.14) by introducing a small positive parameter $\varepsilon > 0$, independent of x.

Various scalings are possible. For the scaling

$$\gamma = \widehat{\gamma}/\varepsilon, \quad \lambda = \widehat{\lambda}/\varepsilon^2, \quad \varphi = \widehat{\varphi}/\varepsilon, \tag{5.1}$$

equation (4.14) takes the form

$$\partial_t \begin{bmatrix} v \\ w \end{bmatrix} = \frac{\widehat{\gamma}}{\varepsilon} S \begin{bmatrix} v \\ w \end{bmatrix} + \frac{2\widehat{\lambda}}{\varepsilon^2} C \begin{bmatrix} v \\ w \end{bmatrix} + \frac{2\widehat{\varphi}}{\varepsilon} B \begin{bmatrix} v \\ w \end{bmatrix}, \tag{5.2}$$

whereas for the scaling

$$\lambda = \widehat{\lambda}/\varepsilon, \quad \varphi = \widehat{\varphi}/\varepsilon, \tag{5.3}$$

it takes the form

$$\partial_t \begin{bmatrix} v \\ w \end{bmatrix} = \widehat{\gamma} S \begin{bmatrix} v \\ w \end{bmatrix} + \frac{2\widehat{\lambda}}{\varepsilon} C \begin{bmatrix} v \\ w \end{bmatrix} + \frac{2\widehat{\varphi}}{\varepsilon} B \begin{bmatrix} v \\ w \end{bmatrix}. \tag{5.4}$$

In both cases $\widehat{\gamma}$, $\widehat{\lambda}$ and $\widehat{\varphi}$ are independent.

Further we will use the scaling (5.1) and equation (5.2). There is no need to analyse both scalings separately since the substitution $t' = \varepsilon t$ transforms (5.4) into (5.2). Thus the scaling (5.3) also leads to a diffusion equation with the only difference that the equation involves the parameter ε and the diffusion coefficient is of order $1/\lambda$.

5.1. *Analysis of a singularly perturbed telegraph system.* Now, consider the singularly perturbed problem (5.2):

$$\begin{gathered} \partial_t v + \frac{\widehat{\gamma}}{\varepsilon} \partial_x(\sigma w) = 0, \\ \partial_t w + \frac{\widehat{\gamma}}{\varepsilon} \partial_x(\sigma v) + \frac{2\widehat{\lambda}}{\varepsilon^2} w - \frac{2\widehat{\varphi}}{\varepsilon} v = 0, \end{gathered} \tag{5.5}$$

with initial conditions

$$v(0,x) = \mathring{v}, \quad w(0,x) = \mathring{w}.$$

By the results already obtained we know that the above system has a unique solution.

In order to obtain "diffusion approximation" we apply a mathematical version of the Chapman–Enskog method proposed by Banasiak and Mika ([2], [3]).

That method relies on representing the functions v and w in the following form:

$$v(t) = \overline{v}(t) + \widetilde{v}(\tau), \tag{5.6}$$

$$w(t) = \overline{w}(t) + \widetilde{w}(\tau), \tag{5.7}$$

where $\tau = t/\varepsilon^2$. This representation is needed for the description of the initial layer phenomena.

The function $\overline{w}$ depends on t only through its dependence on $\overline{v}$, i.e. there is a function W such that $\overline{w}(t) = W(\overline{v}(t))$. This is the essence of Chapman–Enskog's procedure.

The functions $\overline{w}$, $\widetilde{v}$, $\widetilde{w}$ are expanded into asymptotic series in ε:

$$\begin{aligned} \overline{w}(t) &= \overline{w}_0(t) + \varepsilon\overline{w}_1(t) + O(\varepsilon^2) \\ &= W_0(\overline{v}(t)) + \varepsilon W_1(\overline{v}(t)) + O(\varepsilon^2), \end{aligned} \tag{5.8}$$

$$\widetilde{v}(\tau) = \widetilde{v}_0(\tau) + \varepsilon\widetilde{v}_1(\tau) + O(\varepsilon^2), \tag{5.9}$$

$$\widetilde{w}(\tau) = \widetilde{w}_0(\tau) + \varepsilon\widetilde{w}_1(\tau) + O(\varepsilon^2), \tag{5.10}$$

The clue of the method is that the function v remains unexpanded and it is treated as a function of order $O(1)$ at all levels of approximation, and the function $\overline{w}$ depends on t only through its dependence on $\overline{v}$.

The Chapman–Enskog procedure consists in inserting the expanded functions (5.8)–(5.10) into (5.5) and comparing the coefficients of the same powers of ε.

Substituting (5.8) into (5.5) and neglecting terms of order $O(\varepsilon)$, we obtain

$$\partial_t\overline{v} = -\frac{1}{\varepsilon}\widehat{\gamma}\partial_x(\sigma\overline{w}_0) - \widehat{\gamma}\partial_x(\sigma\overline{w}_1). \tag{5.11}$$

Setting $W_j' := dW_j/d\overline{v}$ for $j = 0, 1, \ldots$ we have

$$\begin{aligned} W_0'\partial_t\overline{v} + \varepsilon W_1'\partial_t\overline{v} + \ldots + \frac{1}{\varepsilon}\widehat{\gamma}\partial_x(\sigma\overline{v}) \\ +2\widehat{\lambda}\left(\frac{1}{\varepsilon^2}W_0(\overline{v}) + \frac{1}{\varepsilon}W_1(\overline{v}) - \ldots\right) - \frac{2\widehat{\varphi}}{\varepsilon}\overline{v} = 0. \end{aligned} \tag{5.12}$$

Substituting the expansion (5.11) of $\partial_t\overline{v}$ into (5.12) and comparing coefficients of the same powers of ε we get

$$W_0 \equiv 0, \tag{5.13}$$

$$W_1 = -\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x(\sigma\overline{v}) + \frac{1}{\widehat{\lambda}}\widehat{\varphi}\overline{v}, \tag{5.14}$$

$$W_2 \equiv 0. \tag{5.15}$$

Thus we obtain the following equation for $\overline{v}^{(1)}$, the first order approximation of $\overline{v}$:

$$\partial_t \overline{v}^{(1)} = \frac{\widehat{\gamma}^2}{2}\partial_x\left(\frac{\sigma}{\widehat{\lambda}}\partial_x(\sigma\overline{v}^{(1)})\right) - \widehat{\gamma}\partial_x\left(\frac{\sigma}{\widehat{\lambda}}\widehat{\varphi}\overline{v}^{(1)}\right). \tag{5.16}$$

It is a diffusion type equation, but neither a prospective Kolmogorov equation nor a strictly retrospective one.

Now, using the fact that $\widehat{\gamma}^2/\widehat{\lambda} = \gamma^2/\lambda$ and $\widehat{\varphi}/\widehat{\lambda} = \varphi/\lambda$ we rewrite the diffusion equation with original coefficients. We get

$$\partial_t \overline{v}^{(1)} = \frac{\gamma^2}{2}\partial_x\left(\frac{\sigma}{\lambda}\partial_x(\sigma\overline{v}^{(1)})\right) - \gamma\partial_x\left(\frac{\sigma}{\lambda}\varphi\overline{v}^{(1)}\right). \tag{5.17}$$

To obtain the equations of the initial layer for (5.16) we substitute (5.9) and (5.10) into (5.5) and compare the coefficients of the same powers of ε. We get

$$\partial_\tau \widetilde{w}_0 = -2\widehat{\lambda}\widetilde{w}_0, \tag{5.18}$$

$$\partial_\tau \widetilde{v}_0 = 0, \tag{5.19}$$

$$\partial_\tau \widetilde{w}_1 = -2\widehat{\lambda}\widetilde{w}_1, \tag{5.20}$$

$$\partial_\tau \widetilde{v}_1 = -\widehat{\gamma}\partial_x(\sigma\widetilde{w}_0). \tag{5.21}$$

Since we are looking for solutions which represent the initial layer, they must decay exponentially at infinity. Thus we obtain

$$\widetilde{v}_0(\tau) \equiv 0, \tag{5.22}$$

$$\widetilde{w}_0(\tau) = \widetilde{w}_0(0)e^{-2\widehat{\lambda}\tau}, \tag{5.23}$$

$$\widetilde{v}_1(\tau) = \widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}e^{-2\widehat{\lambda}\tau}\widetilde{w}_0(0)\right), \tag{5.24}$$

$$\widetilde{w}_1(\tau) = \widetilde{w}_1(0)e^{-2\widehat{\lambda}\tau}. \tag{5.25}$$

To establish initial values for (5.16), (5.18)–(5.21) take

$$\mathring{v} = \overline{v}^{(1)}(0) + \varepsilon\widetilde{v}_1(0), \tag{5.26}$$

$$\mathring{w} = \widetilde{w}_0(0) + \varepsilon(\overline{w}_1(0) + \widetilde{w}_1(0)). \tag{5.27}$$

Equation (5.26) defines a solution $\overline{v}^{(1)}(0)$. From (5.23)–(5.25) and (5.27) we get

$$\widetilde{w}_0(0) = \mathring{w}, \tag{5.28}$$

$$\widetilde{v}_1(0) = \widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\right), \tag{5.29}$$

$$\widetilde{w}_1(0) = -\overline{w}_1(0) = \frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x(\sigma\overline{v}^{(1)}(0)) - \frac{1}{\widehat{\lambda}}\widehat{\varphi}\overline{v}^{(1)}(0). \tag{5.30}$$

And for $\overline{v}^{(1)}(0)$ we have

$$\overline{v}^{(1)}(0) = \mathring{v} - \varepsilon\widetilde{v}_1(0), \tag{5.31}$$

$$\overline{v}^{(1)}(0) = \mathring{v} - \varepsilon\widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\right). \tag{5.32}$$

The function $\overline{v}^{(1)}$ must satisfy the same boundary conditions as v, thus

$$\overline{v}^{(1)}(0,t) = \overline{v}^{(1)}(1,t) = 0. \tag{5.33}$$

5.2. *Estimation of the approximation error.* The approximation error is defined as follows:

$$y(t) = v(t) - \overline{v}^{(1)}(t) - \varepsilon\widetilde{v}_1\left(\frac{t}{\varepsilon^2}\right), \tag{5.34}$$

$$z(t) = w(t) - \widetilde{w}_0\left(\frac{t}{\varepsilon^2}\right) - \varepsilon\left(\overline{w}_1 + \widetilde{w}_1\left(\frac{t}{\varepsilon^2}\right)\right). \tag{5.35}$$

To determine it, we substitute (formally) v and w defined by (5.34) and (5.35) into equation (5.5). After algebraic transformations we obtain

$$\partial_t y + \frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma z) = -\widehat{\gamma}\partial_x\sigma\widetilde{w}_1, \tag{5.36}$$

$$\partial_t z + \frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma y) + \frac{2\widehat{\lambda}}{\varepsilon^2}z - \frac{2}{\varepsilon}\widehat{\varphi}y = -\varepsilon\partial_t\overline{w}_1 - \widehat{\gamma}\partial_x\sigma\widetilde{v}_1 + 2\widehat{\varphi}\widetilde{v}_1. \tag{5.37}$$

The initial values for the above system are

$$y(0) = 0, \tag{5.38}$$

$$z(0) = 0, \tag{5.39}$$

whereas the boundary values for y are

$$y(t)\Big|_{\substack{x=0\\x=1}} = -\varepsilon\widetilde{v}_1(\tau)\Big|_{\substack{x=0\\x=1}} = -\varepsilon\widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}e^{-2\widehat{\lambda}\tau}\mathring{w}\right)\Big|_{\substack{x=0\\x=1}}. \tag{5.40}$$

This follows from (5.34), (5.25), (5.33) and the absorbing boundary conditions for v.

Analysing (5.40) shows that the boundary conditions can be written as follows:

$$\begin{aligned} y(t,0) &= \varepsilon Q^0(t/\varepsilon^2) = \varepsilon e^{-2\widehat{\lambda}(0)t/\varepsilon^2}q^0(t/\varepsilon^2),\\ y(t,1) &= \varepsilon Q^1(t/\varepsilon^2) = \varepsilon e^{-2\widehat{\lambda}(1)t/\varepsilon^2}q^1(t/\varepsilon^2), \end{aligned} \tag{5.41}$$

where q^0, q^1 are polynomials with respect to $\tau = t/\varepsilon^2$ with coefficients depending on $\mathring{w}$, $\widehat{\lambda}$, σ, γ and their derivatives at $x = 0$ and $x = 1$ respectively.

We will show that under appropriate assumptions concerning the initial values the approximation error is of the order of $O(\varepsilon^2)$.

LEMMA 5.1. *If* $\sigma\mathring{w}\partial_x(\widehat{\lambda})\big|_{\substack{x=0\\x=1}} = 0$ *and* $\partial_x\sigma\mathring{w}\big|_{\substack{x=0\\x=1}} = 0$, *then*

$$y(t)\big|_{\substack{x=0\\x=1}} = 0. \tag{5.42}$$

P r o o f. From (5.40) we have

$$\begin{aligned} y(t)\big|_{\substack{x=0\\x=1}} &= -\varepsilon\widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}e^{-2\widehat{\lambda}\tau}\mathring{w}\right)\bigg|_{\substack{x=0\\x=1}} \\ &= \left[-\varepsilon\widehat{\gamma}\left(\frac{\partial_x(\sigma\mathring{w})2\widehat{\lambda}-2\sigma\mathring{w}\partial_x\widehat{\lambda}}{4\widehat{\lambda}^4} - \frac{\sigma\mathring{w}}{2\widehat{\lambda}}2\tau(\partial_x\widehat{\lambda})\right)e^{-2\widehat{\lambda}\tau}\right]\bigg|_{\substack{x=0\\x=1}} = 0, \end{aligned}$$

since $\partial_x(\sigma\mathring{w}) = 0$ and $\sigma\mathring{w}\partial_x\widehat{\lambda} = 0$. ∎

Further we will use the following facts:

PROPOSITION 5.1 ([1]). *The operator* $\mathcal{D}$ *defined by*

$$\mathcal{D}v = \frac{\gamma^2}{2}\partial_x\left(\frac{\sigma}{\lambda}\partial_x(\sigma v)\right) - \gamma\partial_x\left(\frac{\sigma}{\lambda}\varphi v\right) \tag{5.43}$$

with domain

$$D(\mathcal{D}) = W^{2,2}_{,\sigma}([0,1]) \cap W^{1,2}_{0,\sigma}([0,1]) \tag{5.44}$$

is the generator of an analytic semigroup $\{G_{\mathcal{D}}(t)\}$ *in* $L^2([0,1])$.

PROPOSITION 5.2 ([12]). *Let* X *be a reflexive space and let* $G(t)$ *be a* C_0-*semigroup generated by an operator* A. *Then the adjoint semigroup* $G(t)^*$ *is the* C_0-*semigroup generated by* A^*.

PROPOSITION 5.3 ([12]). *If an operator* A *is bounded on* X, *then* A^* *is bounded on* X^* *and* $\|G_A(t)\| = \|G_{A^*}(t)\|$.

PROPOSITION 5.4 ([1], [11]). $D(\mathcal{D}^{1/2}) = W^{1,2}_0([0,1])$, *and for* $k = 3, 4$,

$$\begin{aligned} D(\mathcal{D}^{k/2}) = \Big\{ & u \in W^{k,2}([0,1]) : v\big|_{\substack{x=0\\x=1}} = 0, \\ & \partial_x\left(\frac{\sigma}{\lambda}\partial_x(\sigma v)\right)\bigg|_{\substack{x=0\\x=1}} = 0,\ \gamma\partial_x\left(\frac{\sigma}{\lambda}\varphi v\right)\bigg|_{\substack{x=0\\x=1}} = 0\Big\}. \end{aligned}$$

THEOREM 5.1. *The semigroup generated by the operator defined in* (5.43) *satisfies the condition*

$$\|G_{\mathcal{D}}(t)\| \le Me^{-\omega t} \tag{5.45}$$

for some M, $\omega \ge 0$.

Proof. Consider the operator $\mathcal{D}^*$ adjoint to $\mathcal{D}$, i.e. $(\mathcal{D}v, u)_{W^{1,2}_{0,\sigma}} = (v, \mathcal{D}^* u)_{W^{1,2}_{0,\sigma}}$. Then

$$\mathcal{D}^* u = \frac{\gamma^2}{2}\partial_x\left(\frac{\sigma}{\lambda}\partial_x(\sigma u)\right) + \frac{\gamma\varphi}{\lambda}\partial_x(\sigma u), \tag{5.46}$$

$$u(t,0) = u(t,1) = 0. \tag{5.47}$$

We will show that $\mathcal{D}^*$ is the generator of a semigroup of negative type.

There exists a constant M such that $\|e^{t\mathcal{D}^*}\| \le Me^{-\mu_0 t}$, where $-\mu_0 = \sup\{\operatorname{Re}\mu : \mu \in \sigma(\mathcal{D}^*)\}$ ([1]).

From the maximum principle it follows that $\operatorname{Re}\mu < 0$ for $\mu \in \sigma(\mathcal{D}^*)$. It can be shown that the spectrum of the operator $\mathcal{D}^*$ determined by (5.46) and (5.47) is discrete. By the substitution $u = \phi e^{\xi}$, where $\xi = -\frac{1}{\gamma}\int_0^x \varphi(s)\,ds - \frac{1}{\gamma}\varphi(0)$, the eigenvalue problem $\mathcal{D}^* u = \mu u$ can be reduced to the problem $\widehat{\mathcal{D}}^*\phi = \mu\phi$, where $\widehat{\mathcal{D}}^*$ is a self-adjoint operator. The domain of the latter consists of functions of compact support. Hence $\widehat{\mathcal{D}}^*$ has a discrete spectrum ([10]). This implies that $-\mu_0 = \sup\{\operatorname{Re}\mu : \mu \in \sigma(\mathcal{D}^*)\} < 0$. Therefore $\mathcal{D}^*$ generates a semigroup of negative type.

Propositions 5.2 and 5.3 yield that if $\|e^{t\mathcal{D}^*}\| \le Me^{-\mu_0 t}$, then $\|e^{t\mathcal{D}}\| \le Me^{-\mu_0 t}$, hence $\mathcal{D}$ is the generator of a semigroup of negative type. ■

THEOREM 5.2. *Suppose that*

(i) $\widehat{\lambda} \in C^3([0,1])$, $\sigma \in C^4([0,1])$, $0 < \lambda_0 \le \lambda(x) \le \lambda_1 < \infty$, $0 < \sigma_0 \le \sigma(x) \le \sigma_1 < \infty$,

(ii) $\mathring{v} \in W^{3,2}([0,1])$, $\mathring{v}\big|_{\substack{x=0\\x=1}} = 0$,

$$\partial_x\left(\frac{\sigma}{\lambda}\partial_x(\sigma v)\right)\bigg|_{\substack{x=0\\x=1}} = 0, \qquad \gamma\partial_x\left(\frac{\sigma}{\lambda}\varphi v\right)\bigg|_{\substack{x=0\\x=1}} = 0,$$

(iii) $\mathring{w} \in W^{3,2}([0,1])$, $\partial_x(\sigma\mathring{w})\big|_{\substack{x=0\\x=1}} = 0$, $\sigma\mathring{w}\partial_x(2\widehat{\lambda})\big|_{\substack{x=0\\x=1}} = 0$,

(iv) $\varphi \in C^2([0,1])$, $\varphi \le \varphi_1 < \infty$.

Then there exists a constant C such that the approximation error defined by (5.34)–(5.35) *satisfies the condition*

$$\|\{y(t), z(t)\}\|_{L^2} \le C\varepsilon^2, \tag{5.48}$$

uniformly for $t \in [0,\infty)$.

Proof. Consider the problem (5.36)–(5.40). From assumption (iii) and Lemma 5.1 it follows that $y(t)\big|_{\substack{x=0\\x=1}} = 0$.

Due to the assumptions on the coefficients and initial conditions the error function $\{y, z\}$ defined by (5.34) and (5.35) is differentiable with respect to t and x. Thus, we can substitute it into the system (5.36)–(5.37). Hence $\{y, z\}$ is a solution of the initial-boundary value problem (5.36)–(5.40).

Using (5.14) and (5.15) the term $\partial_t \overline{w}_1$ can be rewritten as follows:

$$
\begin{aligned}
\partial_t \overline{w}_1 &= \mathcal{D}\Big(-\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x(\sigma\overline{v}^{(1)})\Big) + \mathcal{D}\Big(\frac{1}{\widehat{\lambda}}\widehat{\varphi}\overline{v}^{(1)}\Big) \\
&= \mathcal{D}\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x\Big[\sigma\Big(-G_{\mathcal{D}}(t)\mathring{v} + \varepsilon\widehat{\gamma}G_{\mathcal{D}}(t)\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big] \\
&\quad + \mathcal{D}\frac{1}{\widehat{\lambda}}\widehat{\varphi}\Big[G_{\mathcal{D}}(t)\mathring{v} - \varepsilon\Big(\widehat{\gamma}G_{\mathcal{D}}(t)\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big] \\
&= \overline{F}_1(t) + \varepsilon\overline{F}_2(t),
\end{aligned}
\tag{5.49}
$$

because $\mathring{v} - \varepsilon\widehat{\gamma}\partial_x\big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\big) \in D(\mathcal{D})$ by (ii) and (iii). Using Proposition 5.4 and assumptions (ii), (iii), $\|\overline{F}_2(t)\|_{L^2}$ can be estimated as

$$
\begin{aligned}
\|\overline{F}_2(t)\|_{L^2} \le{}& \Big\|\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x\sigma\Big(\mathcal{D}G_{\mathcal{D}}(t)\widehat{\gamma}\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big\|_{L^2} \\
&+ \Big\|\frac{1}{\widehat{\lambda}}\widehat{\varphi}\Big(\mathcal{D}G_{\mathcal{D}}(t)\widehat{\gamma}\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big\|_{L^2}
\end{aligned}
\tag{5.50}
$$

and

$$
\begin{aligned}
&\Big\|\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x\sigma\Big(\mathcal{D}G_{\mathcal{D}}(t)\widehat{\gamma}\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big\|_{L^2} \\
&\qquad \le \Big\|\frac{\widehat{\gamma}}{2\widehat{\lambda}}\mathcal{D}^{3/2}G_{\mathcal{D}}(t)\sigma\widehat{\gamma}\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big\|_{L^2} \\
&\qquad \le \frac{\widehat{\gamma}}{2\widehat{\lambda}_1}\Big\|\mathcal{D}^{1/2}G_{\mathcal{D}}(t)\Big(\mathcal{D}\sigma\widehat{\gamma}\partial_x\Big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\Big)\Big)\Big\|_{L^2} \\
&\qquad \le C\|\mathcal{D}^{1/2}G_{\mathcal{D}}(t)\| \cdot \|\mathring{w}\|_{W^{3,2}},
\end{aligned}
\tag{5.51}
$$

where C is a constant.

To estimate $\|\mathcal{D}^{1/2}G_{\mathcal{D}}(t)\|$ we use the following fact:

Proposition 5.5 ([12]). *If $-A$ is the generator of an analytic semigroup $G_{\mathcal{A}}(t)$ and 0 belongs to the resolvent set $\varrho(A)$ then for each $\alpha \ge 0$ there exists $\beta \ge 0$ such that*

$$\|\mathcal{A}^{\alpha}G_{\mathcal{A}}(t)\| \le M_{\alpha}t^{-\alpha}e^{-\beta t}.$$

Hence we have

$$
\begin{aligned}
\|\mathcal{D}^{1/2}G_{\mathcal{D}}(t)\| &= \|(\mathcal{D} - \delta I + \delta I)^{1/2}G_{\mathcal{D}}(t)\| \\
&\le M\frac{1}{\sqrt{t}}e^{-\beta t} + \delta^{1/2}\|G_{\mathcal{D}}(t)\|.
\end{aligned}
\tag{5.52}
$$

Using (5.52), (5.51) and (5.45) we obtain

$$\left\|\frac{\widehat{\gamma}}{2\widehat{\lambda}}\partial_x\sigma\left(\mathcal{D}G_{\mathcal{D}}(t)\widehat{\gamma}\partial_x\left(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\right)\right)\right\|_{L^2}\le C\left(1+\frac{1}{\sqrt{t}}\right)e^{-\omega t},\tag{5.53}$$

for some constants C and $\omega>0$.

Similarly we can estimate $\big\|\frac{1}{\widehat{\lambda}}\widehat{\varphi}\big(\mathcal{D}G_{\mathcal{D}}(t)\widehat{\gamma}\partial_x\big(\frac{\sigma}{2\widehat{\lambda}}\mathring{w}\big)\big)\big\|_{L^2}$, which finally implies that

$$\|\overline{F}_2(t)\|_{L^2}\le C(1+1/\sqrt{t})e^{-\omega t},\tag{5.54}$$

where C and $\omega>0$ are new constants.

To obtain the desired estimates we consider the following two problems:

- problem $\mathcal{P}_1$:

$$\begin{aligned}&\partial_t y_1+\frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma z_1)=-\widehat{\gamma}\partial_x\sigma\widetilde{w}_1,\\&\partial_t z_1+\frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma y_1)+\frac{2\widehat{\lambda}}{\varepsilon^2}z_1-\frac{2}{\varepsilon}\widehat{\varphi}y_1=-\varepsilon^2\overline{F}_2(t)-\widehat{\gamma}\partial_x\widetilde{v}_1+2\widehat{\varphi}\widetilde{v}_1,\\&z_1(0,x)=0,\quad y_1(0,x)=0,\quad y_1(t,0)=y_1(t,1)=0;\end{aligned}\tag{5.55}$$

- problem $\mathcal{P}_2$:

$$\begin{aligned}&\partial_t y_2+\frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma z_2)=0,\\&\partial_t z_2+\frac{\widehat{\gamma}}{\varepsilon}\partial_x(\sigma y_2)+\frac{2\widehat{\lambda}}{\varepsilon^2}z_2-\frac{2}{\varepsilon}\widehat{\varphi}y_2=-\varepsilon\overline{F}_1(t),\\&z_2(0,x)=0,\quad y_2(0,x)=0,\quad y_2(t,0)=y_2(t,1)=0.\end{aligned}\tag{5.56}$$

We will show that $\|\{y_1,z_1\}\|_{L^2}\le C\varepsilon^2$ and $\|\{y_2,z_2\}\|_{L^2}\le C\varepsilon^2$.

Consider problem $\mathcal{P}_1$. A mild solution $\{y_1,z_1\}$ can be written as

$$\begin{aligned}\begin{bmatrix}y_1\\z_1\end{bmatrix}&=G_{\mathcal{K}_\varepsilon}(t)\begin{bmatrix}y_1(0)\\z_1(0)\end{bmatrix}\\&\quad+\int_0^t G_{\mathcal{K}_\varepsilon}(t-s)\begin{bmatrix}-\widehat{\gamma}\partial_x\sigma\widetilde{w}_1(s/\varepsilon^2)\\-\varepsilon^2\overline{F}_2(s)-\widehat{\gamma}\partial_x\widetilde{v}_1(s/\varepsilon^2)+2\widehat{\varphi}\widetilde{v}_1(s/\varepsilon^2)\end{bmatrix}ds.\end{aligned}\tag{5.57}$$

Here $\{G_{\mathcal{K}_\varepsilon}(t)\}_{t\ge0}$ is the semigroup generated by the operator

$$\mathcal{K}_\varepsilon=\frac{\widehat{\gamma}}{\varepsilon}S+\frac{2\widehat{\lambda}}{\varepsilon^2}C+\frac{2}{\varepsilon}\widehat{\varphi}B.$$

This is a problem similar to that of Theorem 4.1. By Theorem 4.2 and the form of the operator $\mathcal{K}_\varepsilon$, it follows that the semigroup $\{G_{\mathcal{K}_\varepsilon}(t)\}_{t\ge0}$ is bounded. Thus we have $\|G_{K_\varepsilon}(t)\|\le M$, $t\ge0$, for some $M>0$.

Notice that all functions on the right-hand side of problem $\mathcal{P}_1$ are of the form $e^{-2\widehat{\lambda}(t/\varepsilon^2)}p(t/\varepsilon^2)$, where p is a polynomial with respect to t/ε^2, whose coefficients depend on the coefficients of the system (5.56). Using the condition (5.45) and inequality (5.54) we come to the following estimate:

$$\left\| \begin{bmatrix} y_1 \\ z_1 \end{bmatrix} \right\|_{L^2} \le \varepsilon^2 C_1 \int_0^{t/\varepsilon^2} e^{-2\widehat{\lambda}_0 s'} P(s')\, ds' \le C\varepsilon^2, \tag{5.58}$$

where C is a constant depending on the coefficients and initial values but independent of t and ε and P is a polynomial.

To estimate $\left\| \left[\begin{smallmatrix} y_2 \\ z_2 \end{smallmatrix}\right] \right\|_{L^2}$ we introduce an additional function h defined as the solution of the problem

$$\begin{aligned} \partial_t h + \frac{2\widehat{\lambda}}{\varepsilon^2} h &= -\varepsilon \overline{F}_1, \\ h(0) &= 0. \end{aligned} \tag{5.59}$$

Solving the above problem we obtain

$$h(t,x) = -\varepsilon \int_0^t e^{-2\widehat{\lambda}(x)(t-s)/\varepsilon^2}\, \overline{F}_1(s,x)\, ds. \tag{5.60}$$

Define $z_2^* = z_2 - h$. From this definition problem $\mathcal{P}_2$ is equivalent to the following one:

$$\partial_t y_2 + \frac{\widehat{\gamma}}{\varepsilon} \partial_x(\sigma z_2^*) = -\widehat{\gamma}\partial_x \Big(\sigma \int_0^t e^{-2\widehat{\lambda}(t-s)/\varepsilon^2}\, \overline{F}_1(s)\, ds \Big), \tag{5.61}$$

$$\partial_t z_2^* + \frac{\widehat{\gamma}}{\varepsilon} \partial_x(\sigma y_2) + \frac{2\widehat{\lambda}}{\varepsilon^2} z_2^* - \frac{2}{\varepsilon}\widehat{\varphi} y_2 = 0, \tag{5.62}$$

$$z_2^*(0,x) = 0, \qquad y_2(0,x) = 0, \tag{5.63}$$

$$y_2(t,0) = y_2(t,1) = 0. \tag{5.64}$$

Denote this last problem by $\mathcal{P}_2^*$.

From assumption (ii) and Proposition 5.4 it follows that $\mathring{v} \in D(\mathcal{D}^{3/2})$. Hence similarly to (5.54), the following estimate also holds:

$$\|\partial_x \overline{F}_1(t,\cdot)\|_{L^2} \le C\Big(1 + \frac{1}{\sqrt{t}}\Big) e^{-\omega t}. \tag{5.65}$$

Let us consider a mild solution of $\mathcal{P}_2^*$. It can be written in the form

$$\begin{aligned} \begin{bmatrix} y_2 \\ z_2^* \end{bmatrix} &= G_{\mathcal{K}_\varepsilon}(t) \begin{bmatrix} y_2(0) \\ z_2^*(0) \end{bmatrix} \\ &\quad + \int_0^t G_{\mathcal{K}_\varepsilon}(t-s) \begin{bmatrix} -\widehat{\gamma}\partial_x \Big(\sigma \int_0^t e^{-2\widehat{\lambda}(t-s)/\varepsilon^2}\, \overline{F}_1(s)\, ds \Big) \\ 0 \end{bmatrix} ds. \end{aligned} \tag{5.66}$$

Thus using (5.63) we have the following estimate of the L^2 norm of $\{y_2, z_2^*\}$:

$$\left\| \begin{bmatrix} y_2 \\ z_2^* \end{bmatrix} \right\|_{L^2} \le \int_0^t \|G_{\mathcal{K}_\varepsilon}(t-s)\| \left\| -\widehat{\gamma}\partial_x \left(\sigma \int_0^s e^{-2\widehat{\lambda}(s-s')/\varepsilon^2} \overline{F}_1(s')\, ds' \right) \right\|_{L^2} ds. \tag{5.67}$$

By the properties of the semigroup $\{G_{\mathcal{K}_\varepsilon}(t)\}_{t\ge 0}$ and the estimate (5.65) we obtain

$$\left\| \begin{bmatrix} y_2 \\ z_2^* \end{bmatrix} \right\|_{L^2} \le C \int_0^t \int_0^s e^{-2\widehat{\lambda}(s-s')/\varepsilon^2} \left(1 + \frac{s-s'}{\varepsilon^2}\right) e^{-\omega s'} \left(1 + \frac{1}{\sqrt{s'}}\right) ds'\, ds, \tag{5.68}$$

where C denotes some constant. Hence we obtain

$$\left\| \begin{bmatrix} y_2 \\ z_2^* \end{bmatrix} \right\|_{L^2} \le C \int_0^\infty e^{-\omega s'} \left(1 + \frac{1}{\sqrt{s'}}\right) ds' \int_0^\infty e^{-2\widehat{\lambda} s/\varepsilon^2} \left(1 + \frac{s}{\varepsilon^2}\right) ds. \tag{5.69}$$

Finally we arrive at

$$\left\| \begin{bmatrix} y_2 \\ z_2^* \end{bmatrix} \right\|_{L^2} \le C_1 \varepsilon^2 \tag{5.70}$$

for some constant C_1.

From (5.60), it follows that there exists a constant C_2 such that

$$\|h(t)\|_{L^2} \le C_2 \varepsilon^3. \tag{5.71}$$

From (5.58), (5.70) and (5.71) it follows that the mild solution $\{y_1 + y_2, z_1 + z_2\}$ of the problem (5.36)–(5.40) satisfies

$$\|\{y_1 + y_2, z_1 + z_2\}\|_{L^2} \le C_3 \varepsilon^2. \tag{5.72}$$

for some constant C_3.

Since the problem (5.36)–(5.40) has a classical solution $\{y, z\}$, it equals the mild solution ([14]). Thus, we obtain (5.48). ∎

6. Summary and conclusions. It has been shown how the random walk system with correlations can be modified to improve the model. Starting from the description of genetic drift in terms of random walk we derived the equations of the continuous model having the form of two first order partial differential equations. Then we demonstrated that the long time behaviour of solutions can be aproximated by the appriopriate solutions of an equation of diffusion type. The final diffusion equation involves a nonzero drift coefficient.

It is worth noticing that the parameter ε is introduced in such a way that the strength of correlation is $\lambda = \widehat{\lambda}/\varepsilon^2$. Therefore the approximation error is of order λ^{-1}.

References

[1] H. Amann, *Dual semigroups and second order linear elliptic boundary value problems*, Israel J. Math. 45 (1983), 225–253.

[2] J. Banasiak, *Singularly perturbed linear and semilinear hyperbolic systems: kinetic theory approach to some folk's theorems*, Acta Appl. Math. (in print).

[3] J. Banasiak and J. R. Mika, *Singularly perturbed telegraph equations with application to the random walk theory*, J. Appl. Math. Stochast. Anal. (in print).

[4] J. M. Connor and M. A. Ferguson-Smith, *Essential Medical Genetics*, Blackwell Sci. Publ., Oxford, 1987.

[5] J. Friedman, *Genetics*, William & Wilkins, Baltimore, 1996.

[6] J. B. S. Haldane, *Suggestions as to quantitative measurement of rates of evolution*, Evolution 3 (1949), 51–56.

[7] M. Iosifescu, *Finite Markov Processes and Their Applications*, Wiley, 1980.

[8] M. Kimura, *Diffusion Models in Population Genetics*, Harper&Row, 1970.

[9] D. Ludwig, *Stochastic Population Theories*, Lecture Notes in Math. 3, Springer, Berlin, 1974.

[10] B. P. Mikhailov, *Partial Differential Equations*, Nauka, Moscow, 1976 (in Russian).

[11] J. R. Mika and J. Banasiak, *Singularly Perturbed Evolution Equations with Applications to Kinetic Theory*, World Sci., Singapore, 1995.

[12] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York, 1982.

[13] S. Wright, *Evolution in Mendelian population*, Genetics 16 (1931), 97–159.

[14] E. Zauderer, *Partial Differential Equations of Applied Mathematics*, Wiley, 1988.

Anna Marciniak
Institute of Applied Mathematics and Mechanics
Warsaw University
Banacha 2
02-097 Warszawa, Poland
E-mail: anna.marciniak@iwr.uni-heidelberg.de

Received on 15.1.1999;
revised version on 9.11.1999