

FROM CONVERGENCE OF OPERATOR SEMIGROUPS TO GENE EXPRESSION, AND BACK AGAIN

ADAM BOBROWSKI

*Institute of Mathematics, Polish Academy of Sciences
Katowice branch, Bankowa 14, 40-007 Katowice, Poland
on leave from*

*Department of Mathematics
Faculty of Electrical Engineering and Computer Science
Lublin University of Technology
Nadbystrzycka 38A, 20-618 Lublin, Poland
E-mail: a.bobrowski@pollub.pl*

To the memory of Professor Andrzej Lasota

Abstract. The subject of the paper is reciprocal influence of pure mathematics and applied sciences. We illustrate the idea by giving a review of mathematical results obtained recently, related to the model of stochastic gene expression due to Lipniacki et al. [38]. In this model, featuring mRNA and protein levels, and gene activity, the stochastic part of processes involved in gene expression is distinguished from the part that seems to be mostly deterministic, and the dynamics is expressed by means of a piece-wise deterministic Markov process. Mathematical results pertain to asymptotic behavior of the process in time as well as limit behavior when certain parameters may be assumed to be large. These results are but an inspiration to considering the ways applied sciences influence pure mathematics by supplying fresh ideas and providing

2000 *Mathematics Subject Classification:* 47D07, 60J25, 60J35, 60J55.

Key words and phrases: contraction semigroup, generator, convex combination, approximation formulae, Trotter-Kato theorem, degenerate convergence, Feller process, weak convergence of processes, time-changed process, Feynman-Kac formula, Volkonskii's formula, gene expression, Fokker-Planck equation, asymptotic stability.

I would like to thank all those who in various ways contributed to this research. These include (in random order) T. Lipniacki, M. Kimmel, A. B. Brasier, A. Marciniak-Czochra, P. Paszek, J. Kiszyński, J. Zabczyk, H. Leszczyński, R. Rudnicki, K. Pichór, K. Bogdan, T. Komorowski and W. Rzymowski. This work was partly supported by the Polish Government, grant number 4T 07A 001 30, and by TODEQ.

The paper is in final form and no version of it will be published elsewhere.

new challenges. On the other hand, they may also be seen as an exemplification of the fact that statements that seem to be almost obvious and are often taken for granted in applied sciences may require mathematical scrutiny and non-standard proofs.

1. Introduction. As the title is meant to suggest [51], this paper may be seen as a report from scientific adventure voyage I undertook around 1997 to come back home some 10 years later. I did not have any dragons to fight against, as was the case with the poor hobbit, but my story seems to be of some interest as it shows how close the two apparently distant words – these of pure, abstract mathematics and of applied sciences – really are. And how mightily they may influence each other, if only we allow a two-way traffic.

I did not make the trip by myself, and am not a pioneer in any way – I was helped by fearless guides living in the high mountains and dreadful deserts of the Lands Between, and it is to them I owe thanks for the safe passage there and back again. But the way is open, and many may go, as I did, to render help to inhabitants of the far lands and bring back the treasures of the old Smaug.

2. Home at Bag End: mathematical background

2.1. Convergence of Laplace transforms, and the Trotter–Kato–Neveu theorem. Let \mathbb{X} be a Banach space, $f_n : [0, \infty) \rightarrow \mathbb{X}, n \geq 1$ be equibounded measurable functions in that there exists an $M \geq 0$ such that $\|f_n(t)\| \leq M$ for all n and almost all $t \in [0, \infty)$, and let $r_n : (0, \infty) \rightarrow \mathbb{X}, n \geq 1$, be the Laplace transforms of f_n :

$$r_n(\lambda) = \int_0^\infty e^{-\lambda t} f_n(t) dt.$$

By the Lebesgue Dominated Convergence Theorem, convergence of f_n implies convergence of r_n . However, as an example of $f_n(t) = e^{\iota nt}$ (where $\iota = \sqrt{-1}$) shows, the converse statement is in general not true, and some additional assumptions have to be made on f_n to guarantee its validity. A common assumption of such a type is that of equicontinuity of f_n at each point $t \in [0, \infty)$. (An elegant proof of sufficiency of equicontinuity, due to Kiszyński [3, 7], is based on the fact that in such a set up, $[0, \infty) \ni t \mapsto F(t) = (f_n(t))_{n \geq 1}$ is a continuous function in the space of bounded \mathbb{X} -valued sequences (equipped with the supremum norm). Since the Laplace transform of F (exists and) has values in the subspace of convergent sequences, subspace separation theorems imply that the values of F must be in the same subspace, giving the desired result.) Additionally, under such assumption, the limit function is continuous, and, hence, convergence is uniform on compact subintervals of $[0, \infty)$ – a mode of convergence known as *almost uniform*. Moreover, if we restrict our attention to almost uniform convergence, equicontinuity turns out to be necessary: continuous f_n converge almost uniformly on compact subintervals iff they are equicontinuous and their Laplace transforms converge [3] Thm 1.7.5. – but, certainly, it is the previous result that is of more practical value.

In this context, the Trotter–Kato–Neveu theorem, one of the corner stones of the theory of one-parameter operator semigroups, is at least surprising. It says that a sequence of equibounded C_0 semigroups $\{T_n(t), t \geq 0\}, n \geq 1$ converges strongly (*i.e.* in the

sense of the norm in \mathbb{X}) to a limit C_0 semigroup $\{T(t), t \geq 0\}$ iff the Laplace transforms $R_{\lambda,n} = \int_0^\infty e^{-\lambda t} T_n(t) dt$ converge to $R_\lambda = \int_0^\infty e^{-\lambda t} T(t) dt$, and convergence of semigroups is almost uniform in $[0, \infty)$ – apparently, no assumption of equicontinuity is needed here. (Equiboundedness means that there exists a constant $M > 0$ such that $\|T_n(t)\| \leq M$ for all $t \geq 0$ and $n \geq 1$.) But of course, as we shall see soon, this is one of the cases of “hidden equicontinuity”, where equicontinuity is implicit in other assumptions, like equiboundedness of derivatives.

Let us recall that a C_0 semigroup (or: a strongly continuous semigroup) of operators is a family of bounded operators $\{T(t), t \geq 0\}$ in \mathbb{X} such that:

- $T(t)T(s) = T(t + s), s, t \geq 0,$
- $T(0) = I_{\mathbb{X}},$
- $\lim_{t \rightarrow 0^+} T(t) = I_{\mathbb{X}}$ (strongly).

As developed in the well-known and beautiful theory (see e.g. [3, 20, 22, 27, 44, 52] etc.) such families have very nice properties that make them of importance and great applicability [3, 8, 19, 20, 44] etc; we name three of them. The first is that for all $x \in \mathbb{X}$, the map $[0, \infty) \ni t \mapsto T(t)x$, called the trajectory of the semigroup, is strongly continuous. The second is that the set $D(A)$ of x for which the right-hand derivative at $t = 0$ of related trajectory exists, is dense in \mathbb{X} . Moreover, the operator of this derivative,

$$Ax = \lim_{t \rightarrow 0} \frac{1}{t} (T(t)x - x)$$

defined on $D(A)$ is closed, and for $x \in D(A)$, the related trajectories are differentiable with

$$\frac{dT(t)x}{dt} = AT(t)x = T(t)Ax. \tag{1}$$

A is called the generator of the semigroup. In other words, trajectories of the semigroup starting at elements of $D(A)$ are solutions to the Cauchy problem for its generator. Conversely, if a Cauchy problem for a densely defined operator A is well-posed in that solutions exist for initial conditions $x \in D(A)$ and depend continuously on the initial conditions, then A is the generator of a semigroup. (The Cauchy problem for a, say, linear operator A and initial condition x_0 is that of existence and uniqueness of solutions to equation $x(t) = Ax(t), t \geq 0$ with $x(0) = x_0$.)

The third property is that for any semigroup there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that $\|T(t)\| \leq Me^{\omega t}$, and, consequently, the Laplace transform $R_\lambda x = \int_0^\infty e^{-\lambda t} T(t)x dt$ exists for all $\lambda > \omega$. It turns out that all $\lambda > \omega$ belong to the resolvent set of A and that $R_\lambda = (\lambda - A)^{-1}$.

It is these properties that force equicontinuity in the set up of the Trotter–Kato–Neveu theorem. To see this, let us be more precise as to what is assumed there. We assume namely that $R_{\lambda,n}$ converge and that the limit operators R_λ are of quite a special form: they are to be equal to the resolvent of the generator of the limit semigroup, and hence, in particular they must have a common range, equal to $D(A)$, which is dense in \mathbb{X} . How this comes into play is seen from the following proof of the Trotter–Kato–Neveu theorem, which goes back to T.G. Kurtz [35]. Let us take $x \in D(A)$ and $\lambda > 0$. Then, there exists a $y \in \mathbb{X}$ such that $x = (\lambda - A)^{-1}y$. Hence, $x_n = R_{\lambda,n}y = (\lambda - A_n)^{-1}y$ belongs to the domain

$D(A_n)$ of the infinitesimal generator A_n of the n -th semigroup, and $\lim_{n \rightarrow \infty} x_n = x$, by assumption. On the other hand, $A_n x_n = \lambda(\lambda - A_n)^{-1}y - y = \lambda x_n - y$ converges to $\lambda x - y$. Therefore, $\sup_n \|A_n x_n\| < \infty$. Consequently, by (1), the derivatives of $[0, \infty) \ni t \mapsto T_n(t)x_n$ are equibounded, and so these maps are equicontinuous, even more, they are Lipschitz continuous with the same constant. Since their Laplace transforms converge to the Laplace transform of $[0, \infty) \ni t \mapsto T(t)x$, the functions must converge to this function (almost uniformly). Since x_n converges to x and the semigroups are equibounded, the same is true of $[0, \infty) \ni t \mapsto T_n(t)x$. Using equicontinuity of the semigroups, $D(A)$ being dense in \mathbb{X} , we obtain convergence for all $x \in \mathbb{X}$.

2.2. Trouble with generalizations of the Trotter–Kato–Neveu theorem: degenerate convergence. A serious drawback of the Trotter–Kato–Neveu theorem is that it assumes a priori existence of the limit semigroup $\{T(t), t \geq 0\}$. In a more natural set up the limit semigroup is not given and we know merely that $R_{\lambda,n}$ converge:

$$\lim_{n \rightarrow \infty} R_{\lambda,n} =: R_\lambda, \quad \lambda > 0.$$

Since the $R_{\lambda,n}$ are resolvents of generators A_n of $\{T_n(t), t \geq 0\}$, they satisfy the so-called Hilbert equation: $(\lambda - \mu)R_{\lambda,n}R_{\mu,n} = R_{\mu,n} - R_{\lambda,n}$ for $\lambda, \mu > 0$, and so must R_λ :

$$(\lambda - \mu)R_\lambda R_\mu = R_\mu - R_\lambda, \quad \lambda, \mu > 0.$$

This is to say that $R_\lambda, \lambda > 0$ is a pseudoresolvent; in particular, the operators R_λ have common range and kernel.

In general, as it will be clear from examples to be presented below, R_λ is not a resolvent of a single operator. However, if the range of R_λ is dense in \mathbb{X} , we may construct a semigroup $\{T(t), t \geq 0\}$ with generator A satisfying $\int_0^\infty e^{-\lambda t} T(t) dt = (\lambda - A)^{-1} = R_\lambda$. This semigroup then turns out to be the limit semigroup of $\{T_n(t), t \geq 0\}$. If the range of R_λ is not dense in \mathbb{X} , the semigroup as above may be constructed merely on

$$\mathbb{X}_0 = cl(\text{Range } R_\lambda).$$

More specifically, in \mathbb{X}_0 there exists an operator A such that all $\lambda > 0$ belong to its resolvent set and $(\lambda - A)^{-1}x = R_\lambda x$, for $x \in \mathbb{X}_0$. This operator may be shown to generate a strongly continuous semigroup $\{T(t), t \geq 0\}$ such that

$$\lim_{n \rightarrow \infty} T_n(t)x = T(t)x, \quad x \in \mathbb{X}_0, t \geq 0,$$

almost uniformly in $[0, \infty)$. It is worth stressing that \mathbb{X}_0 is *the* subspace where the semigroups $\{T_n(t), t \geq 0\}$ converge almost uniformly in $[0, \infty)$. In other words, outside of \mathbb{X}_0 the semigroups either do not converge at all, or do converge but not almost uniformly in $[0, \infty)$.

By the way, let us note that the proof of the Trotter–Kato–Neveu theorem presented in the previous section was based on the fact that we were able to find a large class of $x \in \mathbb{X}$ such that there existed $x_n \in D(A_n)$ with $\lim_{n \rightarrow \infty} x_n = x$ and $A_n x_n$ convergent. It turns out that all information concerning almost uniform convergence of semigroups $\{T_n(t), t \geq 0\}$ is hidden in the relation $A_{ex} \subset \mathbb{X} \times \mathbb{X}$ defined as $(x, y) \in A_{ex}$ iff there exists $x_n \in D(A_n)$ with $\lim_{n \rightarrow \infty} x_n = x$ and $\lim_{n \rightarrow \infty} A_n x_n = y$, often termed *extended limit* of A_n . (In general this is not an operator, since given x there may be many y such that

$(x, y) \in A_{ex}$.) A close inspection of Kurtz’s proof of the Trotter–Kato–Neveu theorem (see the previous subsection) shows two facts. First of all, $D(A_{ex})$, defined as the set of $x \in \mathbb{X}$ such that $(x, y) \in A_{ex}$ for some y , is contained in \mathbb{X}_0 . Secondly, if A is the generator of the limit semigroup defined in \mathbb{X}_0 , then $D(A) \subset D(A_{ex})$. Therefore, \mathbb{X}_0 is the closure of $D(A_{ex})$. Moreover, it may be shown that the resolvents $R_{\lambda, n}$ converge iff for some $\lambda > 0$ the set of elements of the form $\lambda x - y$ where $(x, y) \in A_{ex}$ is dense in \mathbb{X} . Finally, the part of A_{ex} in \mathbb{X}_0 , i.e. the relation $A_{ex} \cap (\mathbb{X}_0 \times \mathbb{X}_0)$ is the graph of an operator in \mathbb{X}_0 – this operator is the generator of the limit semigroup $\{T(t), t \geq 0\}$ in \mathbb{X}_0 . All the above facts constitute the Sova-Kurtz version of the Trotter–Kato–Neveu theorem – see [8, 19, 35, 49].

Coming back to our main thought: outside of \mathbb{X}_0 the semigroups may not converge, and there are plenty of examples of such behavior. For instance, the semigroups in $\mathbb{X} = \mathbb{C}$, given by $T_n(t)x = e^{tn}x$ we have encountered at the beginning of this paper. Also, the translation semigroups $T_n(t)x(\tau) = x(\tau - nt)$ for $\tau \geq nt$ and $= 0$ for $\tau < nt$, in the space of continuous functions with $x(0) = \lim_{\tau \rightarrow \infty} x(\tau) = 0$. In both cases, $\mathbb{X}_0 = 0$ and outside of this space the semigroups do not converge (see [6] or [8] for details). Still another example is given by the Yosida approximation of the operator $x \mapsto -x'$ (the derivative) defined in the space of continuous functions on $[0, \infty]$, with domain composed of x from this space such that $x(0) = 0$ and x' belongs to this space – see [8] Example 8.2.10.

More interestingly, the semigroups $\{T_n(t), t \geq 0\}$ may converge outside of \mathbb{X}_0 , but certainly not almost uniformly in $[0, \infty)$. And in fact, as we will yet see, quite often they do. (To be sure, many examples of convergence of semigroups known in the literature, both in pure and applied mathematics, are instances of such convergence, this is particularly so with the singular perturbations and parabolic approximations to hyperbolic problems – see e.g. [2, 4, 5, 28, 40].) However, there are few general theorems that can be used to prove such convergence in practice.

To see where the difficulty lies, note that if the semigroups converge in this way, in the limit we still obtain a semigroup, an extension of the semigroup $\{T(t), t \geq 0\}$ acting in \mathbb{X}_0 , but one that is not of class C_0 . In fact, denoting it still by $\{T(t), t \geq 0\}$, by definition of \mathbb{X}_0 , we have $\lim_{t \rightarrow 0} T(t)x = x$ merely for $x \in \mathbb{X}_0$, and so the third point of the definition of C_0 semigroup fails. As a consequence, the limit semigroup does not have nice properties enjoyed by C_0 semigroups. Yes, that is true that being measurable (as a limit of measurable semigroups) it must be strongly continuous for $t > 0$, by [27] Thm 3.5.4, but still is not of class C_0 . In particular, the set of x for which the related trajectory is differentiable is not dense in \mathbb{X} , we cannot do the trick with equiboundedness of derivatives leading to equicontinuity, and need to look for more subtle methods.

To be sure, convergence for $x \notin \mathbb{X}_0$ cannot – by nature – be proved using the Trotter–Kato–Neveu theorem, and \mathbb{X}_0 may be pretty small or insignificant – see [11].

2.3. Examples. Let us turn to examples. For the first of these, given a semigroup $\{S(t), t \geq 0\}$ with $\|S(t)\| \leq M, t \geq 0$ for some $M > 0$, we define $T_n(t) = e^{-nt}S(t)$. Then, $\{T_n(t), t \geq 0\}, n \geq 1$ are equibounded and certainly $\lim_{n \rightarrow \infty} T_n(t) = 0, t > 0$. On the other hand, we have $R_{\lambda, n} = \int_0^\infty e^{-(\lambda+n)t}S(t) dt \rightarrow 0$, as $n \rightarrow \infty$. Hence, $\mathbb{X}_0 = \{0\}$

and all the Trotter–Kato–Neveu theorem shows is that $0 = e^{-nt}S(t)0$ converges to 0. Without realization that this theorem deals merely with convergence that is almost uniform in $[0, \infty)$ it could be found really amazing that convergence that is completely obvious by inspection cannot be captured by means of it.

For a next example, consider the most general time-continuous two-state Markov chain. Its intensity matrix is given by $\begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$ where $a, b \geq 0$ are given constants. Recall that such a chain starting from state 1, waits for an exponential time of parameter a to jump to state 2. While at 2, it waits for an exponential time of parameter b , before jumping to state 1. Matrix $A_n = \begin{pmatrix} -a & a \\ b & -b-n \end{pmatrix}$ describes a modified chain, which at 1 behaves as before, but at state 2 it waits for exponential time with parameter $b+n$ and jumps to state 1 with probability $\frac{b}{b+n}$ or disappears from the state-space with probability $\frac{n}{b+n}$. Hence, putting $T_n(t) = \exp A_n t$, we should have

$$\lim_{n \rightarrow \infty} T_n(t) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^{-at}x \\ 0 \end{pmatrix}, \quad x, y \in \mathbb{R}, t > 0. \quad (2)$$

($T_n(t)$ are contractions in \mathbb{R}^2 equipped with the norm $\|(x, y)\| = |x| + |y|$.) To prove this, calculating $R_{\lambda, n} = \frac{1}{\lambda(a+b)+n(\lambda+a)} \begin{pmatrix} b+n & a \\ b-\lambda & \lambda+a \end{pmatrix}$ we see that it converges to $\begin{pmatrix} \frac{1}{\lambda+a} & 0 \\ 0 & 0 \end{pmatrix}$. Hence, $\mathbb{X}_0 = \mathbb{R} \times \{0\}$ and the Trotter–Kato–Neveu Theorem gives (2) for $y = 0$. In particular, it does not and may not provide any information for $y \neq 0$. Fortunately, the full result may be obtained from the explicit form of $T_n(t)$ given e.g. in [42], see [11] for details.

For our third example, suppose we are given $\epsilon > 0$ and the generator A of a bounded (and strongly continuous) cosine operator function. (Cosine operator functions are related to the second order Cauchy problem $x''(t) = Ax(t)$, $x(0) = y$, $x'(0) = 0$ in the same way as the semigroups are related to the first order Cauchy problem. An operator A generates a strongly continuous cosine operator function iff the related second order Cauchy problem is well-posed.) Consider the abstract telegraph equation with small parameter:

$$\epsilon x''_{\epsilon}(t) + x'_{\epsilon}(t) = Ax_{\epsilon}(t), \quad x_{\epsilon}(0) = y, x'_{\epsilon}(0) = z. \quad (3)$$

Formally, letting $\epsilon \rightarrow 0$, we obtain the abstract heat equation:

$$x'_0(t) = Ax_0(t), \quad x_0(0) = y,$$

and so, our intuition is that $\lim_{\epsilon \rightarrow 0} x_{\epsilon}(t) = x_0(t)$. When we turn to derivatives, we may still expect that $\lim_{\epsilon \rightarrow 0} x'_{\epsilon}(t) = x'_0(t)$ for $t > 0$, but we cannot obtain convergence at $t = 0$ unless $z = Ay$, for we have $x'_{\epsilon}(0) = z$ and $x'_0(0) = Ay$.

Using the main theorem of [34], telegraph equation (3) may be transformed to an equivalent first-order system of equations, and then the claim concerning convergence of solutions and derivatives with initial conditions $x_{\epsilon}(0) = y$, $x'_{\epsilon}(0) = Ay$ follows from the Trotter–Kato–Neveu theorem – see [6]. The proof for $x'_{\epsilon}(0) \neq Ay$ requires more subtle methods. In [6] we use an explicit representation of solutions of (3) in terms of related stochastic processes, and the continuity theorem. In [11] we show the desired result by comparing the related semigroups with semigroups that are easier to show to be

convergent. Still another method is presented in [33]. (Of course, diffusion approximation to telegraph equation has been studied in many contexts by a number of authors – see *e.g.* the bibliography in [5], [6] and [28]. Here, we merely stress its semigroup-theoretic implications.)

Our final example, also taken from [6], concerns convergence of Brownian motions in $[0, \infty)$ with changing behavior at the screen $\tau = 0$. Namely, we consider $\mathbb{X} = C[0, \infty]$ and the operators $A_\epsilon, \epsilon > 0$ defined as $A_\epsilon = \frac{d^2}{d\tau^2}$ with $\mathcal{D}(A_\epsilon)$ composed of $x \in \mathbb{X}$ such that $\frac{d^2x}{d\tau^2}$ exists and belongs to \mathbb{X} , and $x(0) = \epsilon x'(0)$. These operators are generators of Feller semigroups $\{T_\epsilon(t), t \geq 0\}$ in $C[0, \infty]$ related to Brownian motions that upon touching the boundary where $\tau = 0$ are partly reflected and partly disappear from the space – see [29, 41, 46]. The $\epsilon > 0$ is a coefficient of elasticity; it decides what happens to the Brownian particle touching the screen. The smaller the ϵ the larger is the fraction of particles that are being annihilated at 0; the remaining ones are reflected at the screen and continue their motion. From this description, and from the form of the generators, it is intuitively clear that as $\epsilon \rightarrow 0$, these processes approach the minimal Brownian motion, where at the screen $\tau = 0$ the particle is instantly annihilated and disappears from the space; the semigroup $\{T_0(t), t \geq 0\}$ related to the minimal Brownian motion is defined in the subspace \mathbb{X}_0 of \mathbb{X} where $x(0) = 0$, and its infinitesimal generator is $A_0x = \frac{1}{2}x''$, on the natural (maximal) domain.

The fact that $\lim_{\epsilon \rightarrow 0} T_\epsilon(t)f = T_0(t)f$ for $f \in \mathbb{X}_0$ may be proved using the Trotter–Kato–Neveu theorem. This theorem, however, does not give information on $\lim_{\epsilon \rightarrow 0} T_\epsilon(t)f$ for $f \notin \mathbb{X}_0$. To treat the latter case, we consider $C[0, \infty]$ as the space of complex functions and note that the semigroups $\{T_\epsilon(t), t \geq 0\}$ are uniformly analytic, and express them with the help of the Dunford integral – see [6] for details.

3. From pure to applied mathematics: gene expression. In 1994, when I defended my PhD thesis with Prof. J. Kisiński as the thesis advisor, examples such as presented above called for a general treatment, but none seemed to be available. For some time I enjoyed finding connections of these problems with the theory of non-densely defined operators, in particular integrated semigroups, but gradually got persuaded I do not have much fresh inspiration to deal with them. The subject seemed to be more and more dry and uninteresting, and I decided I need to start doing things that are related to “real life” and lead to “real applications”, little though I knew about them. I became interested in biological models.

As a result of my pursuit and a punishment for my folly, I was thrown into the dungeon (actually: the second floor) of a building where there was no mathematician to be found: just geneticists. For a number of years I became involved with studying mathematical models of one of the main population genetics forces, called genetic drift, and in particular became familiar with the beautiful process of Kingman’s coalescence – see [32]. Of course, being a mathematician, I tended to use mathematical language I was familiar with to describe the objects from biology, and used mathematical tools to deal with the problems coming from this field – to be sure, it is fun. I never thought, though, I would ever come back to “degenerate convergence of semigroups.”

But then, in May 2005 I was asked for my opinion on the then to-be-published model of stochastic gene expression due to Lipniacki et al. [38], see also [39] and [25].

To explain the model, let us recall that gene expression is the process in which gene's DNA sequence is converted into the structures and functions of a cell (most often: proteins or RNA). The idea behind this notion is apparently that the gene itself is invisible, or at least *was* invisible at the time when the term was coined, but expresses its existence via its visible products, such as proteins. The process itself is quite complicated and there are significant differences in gene expression between eukaryote and bacteria. Even if we restrict ourselves to eukaryote, the process will still depend on a kind of cell, phase of its development, and on metabolic and physiological state.

Fortunately, there are some stages of gene expression which are more or less common to all eukaryote, which in an extremely simplified way may be described as follows. First, we must recall that (almost) all genetic information is contained in the nucleus of a cell where cell's DNA is, and the process of gene expression does not start until for this or that reason the gene becomes active. At this point, information contained in DNA is transcribed into so-called messenger RNA (mRNA), a medium of transportation. Then, mRNA molecules carry the information to the cytoplasm, where it is translated into protein sequence. After translation, the resulting protein may still be modified and then travels to its destination (target) in the cell. The last stages are natural degradation of mRNA and degradation of mRNA.

It should be noted here that typically there are several copies of a single gene in the genome. (Eukaryotic cells are diploid, i.e. typically have two copies of each gene in their genome.) It seems reasonable to assume that they become active and inactive independently of each other. However, it is biologically established that the level of proteins influences intensities of activation and inactivation of the copies.

Authors of [38] argue that while the process of transcription, transportation and translation are more or less deterministic, the process of gene activation (inactivation) is of stochastic nature, basically due to a limited number reactants. This, by the way, is one of the main subjects and important input of the paper. Moreover, they propose a mathematical model where the state of the process is described by three variables (x, y, γ) where x is the mRNA level, y is the protein level, and γ is the number of active gene copies. After some simplifying and re-scaling procedure, the dynamics is expressed in the form of the following system of stochastic ordinary differential equations:

$$\begin{aligned} \frac{dx}{dt} &= \gamma - x, \\ \frac{dy}{dt} &= r(x - y), \end{aligned} \tag{4}$$

where $r > 0$ is a given constant, and γ is an integer-valued random process with intensity of increase by 1 equal to $a\alpha(x, y)$ and intensity of decrease by 1 equal to $b\beta(x, y)$ for some non-negative functions α and β and positive constants a and b . (Biologically, it is clear that α and β do depend on y , but not on x . Introducing assumption on dependence on x to the model does not, however, lead to additional mathematical difficulties.) Such a model captures all the important features of gene expression. More specifically, the first

equation shows that when all gene copies are inactive, *i.e.* when $\gamma = 0$, the number of mRNA decreases exponentially, and that an increase in γ results in the gradual growth of x . On the other hand, by the second equation, a burst (or a decrease in number) of mRNA molecules produces a burst (or a decrease in number) of protein molecules, with protein level y “catching up” with mRNA level with efficiency r . Finally, the level of mRNA and protein influences probability of the change in the number of active gene copies.

I need to confess, the example offered by this model is a very bad one, pedagogically. For years I was told, and got used to the fact that in applied mathematics, for each problem you face, an appropriate mathematical language with appropriate tools must be chosen. That, in other words, you should not stick to your favorite theory, but always learn new ones to match the problem you are to solve. But for me, mathematically, this model was exactly what I specialized in and what I was interested in for years. Of course, I had the advantage of working on a model that was already prepared by somebody else. But still, it seemed to suit my interests too well.

Actually the number of mathematical connections of the model I became aware of is quite surprising. First of all, it is clear that were it not for the fact that the intensities of jumps of γ depend on y (and possibly on x), the process (x, y, γ) would have been an example of *random evolution*, *i.e.* a member of a class of stochastic processes introduced by Griego and Hersh [23, 24] after Kac’s famous work on stochastic solutions of the telegraph equation [30]. After I consulted with prof. J. Zabczyk, the process turned out to be an example of a piece-wise deterministic process of M.H.A. Davis [14, 15, 16], who prophesied in 1984: “These (processes) were only isolated rather recently but seen general enough to include as special cases practically all the non-diffusion continuous-time processes of applied probability” ([15], p. i).

Secondly, the authors of [38] were interested in the the Fokker-Planck-type equation for the process $(x(t), y(t), \gamma(t)), t \geq 0$, which we write here in the case of a single gene:

$$\begin{aligned} \frac{\partial f_0}{\partial t} + \frac{\partial}{\partial x}(-x f_0) + r \frac{\partial}{\partial y}((x - y) f_0) &= b\beta f_1 - \alpha\alpha f_0, \\ \frac{\partial f_1}{\partial t} + \frac{\partial}{\partial x}((1 - x) f_1) + r \frac{\partial}{\partial y}((x - y) f_1) &= \alpha\alpha f_0 - b\beta f_1; \end{aligned} \tag{5}$$

here $f_i(t, \cdot, \cdot), i = 0, 1$ are parts of a density of the process at time $t \geq 0$ related to $\gamma = 0, 1$, respectively. (Let us note in passing that were the intensities α and β equal to zero, the above system would have been uncoupled and would describe two independent motions along trajectories of solutions of (4) with fixed γ equal either 0 or 1, by means of a family of Frobenius–Perron operators in a related L^1 space – see [36]; the non-zero right-hand side describes the rules of jumps between states $\gamma = 0$ and $\gamma = 1$.) In particular one of the questions was that of existence (and form) of densities that would be invariant for the equation. Certainly, being A. Lasota’s student, I was familiar with the theory of Markov semigroups in L^1 that provides the means for approaching such questions – roughly speaking, Markov semigroups are all about dynamics of densities of Markov processes, being composed of operators that to an initial density of a Markov process assign its densities at times $t > 0$; see [36]. As a result of another student of Lasota’s,

namely R. Rudnicki's expertise in this field, existence of such a density and convergence of all other densities to this one as $t \rightarrow \infty$ was proved in [12].

But these were two hypotheses concerning limit behavior of (4) as a, b and r tend to infinity, that really captured my attention, as they proved to be hypotheses on degenerate convergence of semigroups. Parenthetically, let us recall what Einar Hille wrote in 1948 [26]: "I hail a semigroup when I see one and I seem to see them everywhere. (Friends have observed, however, that there are mathematical objects which are not semigroups.)" In a sense, thus, it is not surprising that I also see them even in a model of gene expression. Nevertheless, I am still amazed that questions from abstract mathematics may be of relevance in applied sciences.

The first hypothesis (not discussed in the published version of Lipniacki's paper) concerned the limit behavior of (4) when $r \rightarrow \infty$. Heuristically, as $r \rightarrow \infty$, the efficiency with which y chases x , increases infinitely, and so in the limit, we should expect that y catches up with x immediately after any change. In other words, we should have $x = y$ all the time and (4) should reduce to a single stochastic equation

$$\frac{dx}{dt} = \gamma - x, \quad (6)$$

where γ is a random process with values in $\{0, 1\}$, with activation intensity $a\tilde{\alpha}(x) = a\alpha(x, x)$ and inactivation intensity $b\tilde{\beta}(x) = b\beta(x, x)$. Biologically, r is defined as the ratio of protein and mRNA degradation rates. Hence, large r corresponds to the case where mRNA is much more stable than the corresponding protein. In such a scenario, the mRNA's level dictates the protein's level: compare [45] Fig. 7A ($R=0.78$) and Fig. 7D ($R=0.92$). This is the case with protein $I\kappa B\alpha$ in the $\text{NF}\kappa\text{B}$ module (in the absence of A20 protein) – see [37], and with protein p53 – see [13].

The second hypothesis concerned the limit behavior of (4) as a and b tend to infinity in such way that a/b tends to a positive constant c . Lipniacki argues that in such a case the jumps occur infinitely often and γ reaches statistical equilibrium. Hence, in (4) it may be replaced by its conditional expected value (given x and y). In other words, in the limit, (4) becomes

$$\begin{aligned} \frac{dx}{dt} &= \frac{c\alpha}{c\alpha + \beta} - x, \\ \frac{dy}{dt} &= r(x - y). \end{aligned} \quad (7)$$

To explain it in more detail: the paper [38] was concerned mainly with the situation where transitions between active and inactive states of the gene are relatively rare, so that to explain, in particular, bursts in production of mRNA reported in [21]. Passage to the limit as a and b tend to infinity corresponds to the case of frequent transitions studied already in the 1980' – see [1, 48], the case that can be called classical today. See [50] where frequent transitions are assumed tacitly, especially Fig. 3 where the transition rate is very similar to the term $\frac{c\alpha}{c\alpha + \beta}$ appearing in (7); for an analogous formulae for transition rate in a slightly different scenario see [13]. Similar passages to the limit may be found in [31]; in particular, passage from (B1) to (B20), from (34) to (39), and from (C1)-(C3) to (46)-(47). See also [43].

When put in the context of weak convergence of related stochastic processes, these hypotheses are statements on convergence of certain conditional expected values, as explained below. First, let us note that in the case of a single gene copy, *i.e.* the case where γ belongs to $\{0, 1\}$, it suffices to consider (4) for x and y in $[0, 1]$. Indeed, all solutions to these equations starting from outside of the unit square $\mathcal{Q} = [0, 1]^2$ tend to this square, and upon reaching it stay there for ever. Hence, we may consider $(x(t), y(t), \gamma(t)), t \geq 0$ as a process with values in two copies of \mathcal{Q} (one corresponding to $\gamma = 0$, the other to $\gamma = 1$). A continuous function on the state space may therefore be identified with two functions, say f_0 and g_0 , on \mathcal{S} . For $t \geq 0$, we consider

$$\begin{aligned} f(t, x, y) &= \mathbb{E}_{x,y,0} \{f_0(x(t), y(t))1_{\{\gamma(t)=0\}} + g_0(x(t), y(t))1_{\{\gamma(t)=1\}}\} \\ g(t, x, y) &= \mathbb{E}_{x,y,1} \{f_0(x(t), y(t))1_{\{\gamma(t)=0\}} + g_0(x(t), y(t))1_{\{\gamma(t)=1\}}\}, \end{aligned} \tag{8}$$

where $\mathbb{E}_{x,y,i}, i = 0, 1$ is the conditional expectation, given that $x(0) = x, y(0) = y$ and $\gamma(0) = i$. Since the process $(x(t), y(t), \gamma(t)), t \geq 0$ may be shown to be Feller (see Section 4 for definition), these functions are continuous, and if f_0 and g_0 are sufficiently regular (for example of class C^1), they satisfy the so-called Kolmogorov equation:

$$\begin{aligned} \frac{\partial f}{\partial t} &= -x \frac{\partial f}{\partial x} + r(x-y) \frac{\partial f}{\partial y} - a\alpha f + a\alpha g, \\ \frac{\partial g}{\partial t} &= (1-x) \frac{\partial g}{\partial x} + r(x-y) \frac{\partial g}{\partial y} - b\beta g + b\beta f, \end{aligned} \tag{9}$$

with initial conditions $f(0, x, y) = f_0(x, y)$ and $g(0, x, y) = g_0(x, y)$, in a sense dual to (5).

Weak convergence of processes involved in the first hypothesis is then equivalent to convergence of solutions of (9) to solutions of the Kolmogorov equation related to (6), which reads

$$\begin{aligned} \frac{\partial f}{\partial t} &= -x \frac{\partial f}{\partial x} - a\tilde{\alpha}f + a\tilde{\alpha}g, \\ \frac{\partial g}{\partial t} &= (1-x) \frac{\partial g}{\partial x} - b\tilde{\beta}g + b\tilde{\beta}f. \end{aligned} \tag{10}$$

Here, $f(t, \cdot)$ and $g(t, \cdot)$ are functions of one variable $x \in [0, 1]$, and it should be noted that such functions may be identified with continuous functions on \mathcal{S} that do not depend on y . In other words, we have another, and quite natural, example of a family of C_0 semigroups in $C(\mathcal{Q}) \times C(\mathcal{Q})$ approximating a C_0 semigroup acting only in a subspace of this space – the subspace that is isometrically isomorphic to $C[0, 1] \times C[0, 1]$. And again, what we may prove with the help of the Trotter–Kato theorem is that if we choose initial conditions of (9) not to depend on y , then in the limit as $r \rightarrow \infty$ we will obtain solutions to (10), with the same initial condition. This, however, does not prove weak convergence of related processes. To establish weak convergence, we need to prove a similar result for any initial conditions. In [9] we show that if we start from any $(f_0, g_0) \in C(\mathcal{Q}) \times C(\mathcal{Q})$, then for all $t > 0$ in the limit we obtain solutions of (10) starting from $x \mapsto f_0(x, x)$ and $x \mapsto g_0(x, x)$ – the proof is based on the Phillips perturbation theorem.

The situation is similar with the second hypothesis. In terms of weak convergence of related processes it states that as a and b tend to infinity in the prescribed way, the

solutions to (9) converge to the solutions of the equation

$$\frac{\partial f}{\partial t} = \left(\frac{c\alpha}{c\alpha + \beta} - x \right) \frac{\partial f}{\partial x} + r(x - y) \frac{\partial f}{\partial y}. \quad (11)$$

It is worth noting that the limit process described by (7) is deterministic: (7) is a system of autonomous ordinary differential equations. Solutions to this system exist under suitable, classic assumptions on α and β , and form a semi-flow, and solutions to (11) may be conveniently expressed through this semi-flow. What is important is that again, (11) is about dynamics of a single function form $C(\mathcal{Q})$, while (9) is about dynamics of pairs of functions from this space. In other words, if we identify elements of $f \in C(\mathcal{Q})$ with pairs (f, f) in $C(\mathcal{Q}) \times C(\mathcal{Q})$, we have another example of a family of C_0 semigroups in $C(\mathcal{Q}) \times C(\mathcal{Q})$ approximating a semigroup that acts merely on the subspace of $C(\mathcal{Q}) \times C(\mathcal{Q})$; this subspace is isometrically isomorphic to $C(\mathcal{Q})$. What the Trotter–Kato–Neveu theorem shows is that if we choose initial conditions of (9) in such a way that $f_0 = g_0$, then both f and g in this equation converge to the solution of (11) starting from f_0 – see [9]. But much more is true, and this cannot be proved with the help of the Trotter–Kato–Neveu theorem. Specifically, in [9] we also prove that taking any f_0 and g_0 , both parts of the solution of (9) at $t > 0$ tend to the solution of (11) starting from $\frac{\beta}{c\alpha + \beta}f + \frac{c\alpha}{c\alpha + \beta}g$.

4. From applied to pure mathematics. An afterthought to the story described above is as follows. Even if we treat applied mathematics as a tool to prove hypotheses of non-mathematicians, who gains more by taking part in cooperation, a non-mathematician who’s hypothesis is substantiated, or a mathematician who deals with a non-trivial problem (with a “real” application)? To be sure, I am not easily thrilled by the word “applied”. I am persuaded that there are as many worthless papers published in applied sciences as there are in pure mathematics. I agree also that mathematics behind many biological problems, even those really relevant, is simply trivial – not all problems interesting for biologists or physicists are interesting for mathematicians. On the other hand, however, intuitions born outside of mathematics may become an invaluable source of inspiration to mathematicians, and an incentive to develop their theories to cover cases that are relevant for other sciences.

Just to say the least: the proof of asymptotic stability of (5) required a reconsideration of the methods available for such problems and lead to their further improvement and development – especially dependence of α and β on y (and possibly on x), natural from biological point of view, constituted a substantial difficulty – see [12]. Similarly with examples of degenerate convergence of semigroups: for several months they were a substantial challenge and lead me to a further research on limitations of the Trotter–Kato–Neveu theorem – see [11].

But I would like to conclude with another example of the way intuitions of non-mathematicians may lead to establishing new theorems of pure mathematics. This example concerns convex combinations of Feller generators, and is another by-product of Lipniacki’s second hypothesis.

To explain the idea, let us recall that a Markov process $p(t), t \geq 0$ with values in a locally compact space \mathcal{S} is said to be a Feller process ([8, 19, 20, 46]) if the corresponding

transition semigroup

$$T(t)f(p) = \mathbb{E}_p f(p(t)), \quad p \in \mathcal{S}, t \geq 0, f \in C_0(\mathcal{S}), \tag{12}$$

where \mathbb{E}_p is the conditional expectation conditional on $p(0) = p$, is a strongly continuous semigroup in $C_0(\mathcal{S})$; in particular, we assume that $T(t)f$ belongs to $C_0(\mathcal{S})$ for $f \in C_0(\mathcal{S})$ and $t \geq 0$ – the reader will notice that (8) is a particular case of (12). Isolated and studied first by Feller himself, Feller semigroups (i.e. semigroups related to Feller processes) are comparatively easily treatable analytically and are best described (see references given above) – in fact, we could claim their theory is by all means classical.

It is clear that all information on finite-dimensional distributions of the process is then hidden in the related semigroup. On the other hand, such a semigroup is fully characterized by its infinitesimal generator, or by the related Cauchy problem. Now, (9) is a Cauchy problem for such a process with quite “compound” generator. In particular it involves operators, say $B_i = (i - x)\frac{\partial}{\partial x} + r(x - y)\frac{\partial}{\partial y}$ which are generators of processes of deterministic motions along integral curves of ordinary differential equations:

$$\begin{aligned} \frac{dx}{dt} &= i - x, \\ \frac{dy}{dt} &= r(x - y), \end{aligned} \tag{13}$$

and its remaining part is responsible for jumps between two copies of \mathcal{Q} . It is interesting to note that equation (11) may be obtained from (9) by letting $f = g$, and adding the first equation multiplied by $\frac{\beta}{c\alpha + \beta}$ to the second equation multiplied by $\frac{c\alpha}{c\alpha + \beta}$. Moreover, $\frac{\beta}{c\alpha + \beta}$ and $\frac{c\alpha}{c\alpha + \beta}$ are clearly the probabilities that the gene activity process γ being at statistical equilibrium, equals 0 and 1, respectively. In other words the generator of the limit, deterministic process of movement along integral curves of (7) is a convex combination of operators B_i , and that process may be seen as a result of averaging motions along integral curves of (13), with weights $\frac{\beta}{c\alpha + \beta}$ and $\frac{c\alpha}{c\alpha + \beta}$.

This suggests the following, more abstract result: let, as above, \mathcal{S} be a locally compact Hausdorff space, and let α and β be two non-negative continuous functions on \mathcal{S} with $\alpha + \beta = 1$. Furthermore, let A, B and C be generators of Feller semigroups $\{T_A(t), t \geq 0\}$, $\{T_B(t), t \geq 0\}$, and $\{T_C(t), t \geq 0\}$ in $C_0(\mathcal{S})$ with related Feller processes $\{X_A(t), t \geq 0\}$, $\{X_B(t), t \geq 0\}$ and $\{X_C(t), t \geq 0\}$. Finally, suppose that $\mathcal{D}(A) \cap \mathcal{D}(B)$ is a core for C and that $Cf = \alpha Af + \beta Bf$ for $f \in \mathcal{D}(A) \cap \mathcal{D}(B)$. Then, $\{X_C(t), t \geq 0\}$ may be described as follows: conditional on being at $p \in \mathcal{S}$, $X_C(t)$ behaves like $X_A(t)$ or $X_B(t)$ with probabilities $\alpha(p)$ and $\beta(p)$, respectively.

In [10] we support the intuition presented above by showing that $\{T_C(t), t \geq 0\}$ may be approximated by the semigroups $\{\mathcal{T}_n(t), t \geq 0\}$ in $C_0(\mathcal{S}) \times C_0(\mathcal{S})$ generated by operators \mathcal{A}_n with common domain $\mathcal{D}(\mathcal{A}_n) = \mathcal{D}(A) \times \mathcal{D}(B)$, given by $\mathcal{A}_n \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} Af - n\beta f + n\beta g \\ Bg + n\alpha f - n\alpha y g \end{pmatrix}$, for $n \geq 0$. More specifically, and the reader will not be surprised at the remark that this is yet another example of degenerate convergence of semigroups,

$$\lim_{n \rightarrow \infty} \mathcal{T}_n(t) \begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} T_C(t)(\alpha f + \beta g) \\ T_C(t)(\alpha f + \beta g) \end{pmatrix}, \quad f, g \in C_0(\mathcal{S}), t > 0.$$

Here $C_0(\mathcal{S}) \times C_0(\mathcal{S})$ (with the norm $\| \begin{pmatrix} f \\ g \end{pmatrix} \| = \max\{\|f\|_{C_0(\mathcal{S})}, \|g\|_{C_0(\mathcal{S})}\}$) is identified via

isometric isomorphism with $C_0(\mathcal{S}_A \cup \mathcal{S}_B)$ where $\mathcal{S}_A = \mathcal{S} \times \{A\}$ and $\mathcal{S}_B = \mathcal{S} \times \{B\}$ are two copies of \mathcal{S} , and $\{\mathcal{T}_n(t), t \geq 0\}$ are viewed as Feller semigroups related to the following random processes $\{X_n(t), t \geq 0\}$ with values in $\mathcal{S}_A \cup \mathcal{S}_B$ (see Figure 1). While at \mathcal{S}_A , X_n behaves like X_A , in particular, it may jump with X_A , but stays in \mathcal{S}_A until the random time τ when it leaves \mathcal{S}_A to go to \mathcal{S}_B and behave like X_B . At time τ , conditional on being at a point $(p, A) \in \mathcal{S}_A$, the process jumps to the copy (p, B) of this point in \mathcal{S}_B and the intensity of jump varies with $p \in \mathcal{S}_A$, being equal to $n\beta(p)$. While at \mathcal{S}_B , the process behaves like X_B and the intensity of the jump to \mathcal{S}_A is $n\alpha$.

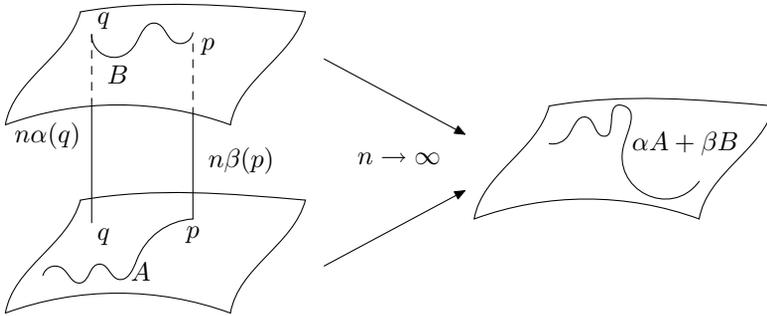


Fig. 1

It is helpful to think of X_n as a mixture of two processes. In the first of them, when starting at (p, A) we always stay at \mathcal{S}_A and move according to the rules of process X_A ; conditional on starting at (p, B) we stay at \mathcal{S}_B all the time and move according to the rules of process X_B . In the second process the only possible transitions are jumps between $(p, A) \in \mathcal{S}_A$ and $(p, B) \in \mathcal{S}_B$. In other words, conditional on starting at (p, A) the second process is a Markov chain with two possible values (p, A) and (p, B) , and with intensities $\begin{pmatrix} -n\beta(p) & n\beta(p) \\ n\alpha(p) & -n\alpha(p) \end{pmatrix}$ or, which is the same, with transition probabilities $\begin{pmatrix} \alpha(p) + \beta(p)e^{-nt} & \beta(p) - \beta(p)e^{-nt} \\ \alpha(p) - \alpha(p)e^{-nt} & \alpha(p)e^{-nt} + \beta(p) \end{pmatrix}, t \geq 0$. Hence, the second process, the process of jumps between \mathcal{S}_A and \mathcal{S}_B , is composed of independent Markov chains indexed by p . As $n \rightarrow \infty$, for each $t > 0$, each of these chains reaches its equilibrium $(\alpha(p), \beta(p))$. Therefore, the limit process of X_n may indeed be thought of as the one behaving at p as X_A with probability $\alpha(p)$ or X_B with probability $\beta(p)$ (see Figure 1).

Apparently, there seems to be no other known approximation formula for the semigroup $\{T_C(t), t \geq 0\}$ that would support this intuition. In particular the Dyson-Phillips perturbation series and the Trotter product formula are of no help here (see [10] for more detailed discussion).

One particular example of a convex combination of two Feller generators is especially worth mentioning here. This is the case where $B = 0$ and α is separated from zero by a constant: $\alpha \geq \alpha_0 > 0$. By the main result of Dorroh [17] (which, by the way, is the precursor of the whole class of “multiplicative perturbation theorems” of semigroup theory – see [18]), αA generates a Feller semigroup. Moreover, B is the generator of the

semigroup composed of identities, which corresponds to no movement at all. Hence, in the light of our theorem, the process related to αA may be described as follows: conditional on being at $p \in \mathcal{S}$, with probability $\alpha(p)$ the process behaves like X_A and with probability $1 - \alpha(p)$ it stays at p . As a result, the process “slows down” as at each point it hesitates whether to move or to stay, and the smaller is α the slower the process moves. This agrees beautifully with the so-called Volkonskii formula [47] of probability theory, which says that multiplying the generator of a Feller process by a function is equivalent to a time-change in the process, a procedure of changing the speed at which the original process runs through its path. In the case under consideration this function is $\alpha \leq 1$, and such a procedure results in “slowing down” of the process, in agreement with our interpretation.

Finally, let us mention that the resolvent of αA may be explicitly calculated in terms of the process X_A and this, via the use of the Feynman-Kac formula, leads to a straightforward proof of the (part of the) result of Dorroh and of the Volkonskii formula – see [10].

Having brushed against the Feynman-Kac formula, I am persuaded I came back home, to pure mathematics with well-grounded intuitions from applied sciences.

References

- [1] G. K. Ackers, A. D. Johnson and M. A. Shea, *Quantitative model for gene regulation by λ phage repressor*, Proc. Natl. Acad. Sci. USA 79 (1982), 1129–1133.
- [2] D. R. Akhmetov and R. Spigler, *Uniform and optimal estimates for solutions to singularly perturbed parabolic equations*, J. Evol. Equations 7 (2007), 347–372.
- [3] W. Arendt, C. J. K. Batty, M. Hieber and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Birkhäuser, 2001.
- [4] O. Arino, E. Sánchez, R. Bravo De La Parra and P. Auger, *A singular perturbation in an age-structured population model*, SIAM J. Appl. Math. 50 (1999), 408–436.
- [5] J. Banasiak, *Singularly perturbed linear and semilinear hyperbolic systems: kinetic theory approach to some folk’s theorems*, Acta Applicandae Mathematicae 49 (1997), 199–228.
- [6] A. Bobrowski, *Degenerate convergence of semigroups*, Semigroup Forum 49 (1994), 303–327.
- [7] A. Bobrowski, *The Widder-Arendt theorem on inverting of the Laplace transform and its relationship with the theory of semigroups of operators*, Methods Funct. Anal. Topology 3 (1997), 1–39.
- [8] A. Bobrowski, *Functional Analysis for Probability and Stochastic Processes*, Cambridge University Press, 2005.
- [9] A. Bobrowski, *Degenerate convergence of semigroups related to a model of stochastic gene expression*, Semigroup Forum 73 (2006), 345–366.
- [10] A. Bobrowski, *On a semigroup generated by a convex combination of two Feller generators*, J. Evol. Equ. 7 (2007), 555–565.
- [11] A. Bobrowski, *On limitations and insufficiency of the Trotter–Kato theorem*, Semigroup Forum 75 (2007), 317–336.
- [12] A. Bobrowski, T. Lipniacki, K. Pichór and R. Rudnicki, *Asymptotic behavior of distributions of mRNA and protein levels in a model of stochastic gene expression*, J. Math. Anal. Appl. 133 (2007), 753–769.

- [13] A. Ciliberto, B. Novak and J. J. Tyson, *Steady states and oscillations in the p53/Mdm2 network*, Cell Cycle 4:3; March 2005, 488–493.
- [14] M. H. A. Davis, *Piece-wise deterministic Markov processes*, J. Royal Statistical Soc. Ser. B 46 (1984) 353–388.
- [15] M. H. A. Davis, *Lectures on Stochastic Control and Nonlinear Filtering*, Springer, 1984.
- [16] M. H. A. Davis, *Markov Processes and Optimization*, Chapman and Hall, 1993.
- [17] B. Dorroh, *Contraction semigroups in a function space*, Pacific J. Math. 19 (1966), 35–38.
- [18] K.-J. Engel and R. Nagel, *One-parameter Semigroups for Linear Evolution Equations*, Springer, 2000.
- [19] S. N. Ethier and T. G. Kurtz, *Markov Processes. Characterization and Convergence*, Wiley, 1986.
- [20] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, (second edition, 1971), Wiley, 1966.
- [21] A. M. Femino, F. S. Fay, K. Fogarty and R. H. Singer, *Visualization of single RNA transcripts in situ*, Science 280 (1998), 585–590.
- [22] J. A. Goldstein, *Semigroups of Linear Operators and Applications*, Oxford Mathematical Monographs, 1985.
- [23] R. J. Griego and R. Hersh, *Random evolutions, Markov chains, and systems of partial differential equations*, Proc. Nat. Acad. Sci. USA 62 (1969), 305–308.
- [24] R. J. Griego and R. Hersh, *Theory of random evolutions with applications to partial differential equations*, Trans. Amer. Math. Soc. 156 (1971), 405–418.
- [25] B. Hat, P. Paszek, M. Kimmel, K. Piechór and T. Lipniacki, *How the number of alleles influences gene expression*, Journal of Statistical Physics 128 (2007), 511–533.
- [26] E. Hille, *Functional Analysis and Semigroups*, Amer. Math. Soc., 1948.
- [27] E. Hille and R. S. Phillips, *Functional Analysis and Semigroups*, Amer. Math. Soc., 1957.
- [28] T. Hillen, *Hyperbolic models for chemosensitive movement*, Math. Models Methods of Appl. Sci. 12 (2002), no. 7 1–28.
- [29] K. Ito and H. P. Mc Kean, Jr., *Diffusion Processes and Their Sample Paths*, Springer, 1965.
- [30] M. Kac, *Some Stochastic Problems in Physics and Mechanics*, Magnolia Petroleum Co. Colloq. Lect. 2, 1956.
- [31] T. B. Kepler and T. C. Elston, *Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations*, Biophysical J. 81 (2001), 3116–3136.
- [32] J. F. C. Kingman, *The coalescent*, Stoch. Proc. Appl. 13 (1982), 235–248.
- [33] J. Kiszyński, *On second order Cauchy's problem in a Banach space*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. et Phys. 18 (1970), 371–374.
- [34] J. Kiszyński, *On cosine operator functions and one-parameter groups of operators*, Studia Math. 44 (1972), 93–105.
- [35] T. G. Kurtz, *Extensions of Trotter's operator semigroup approximation theorems*, J. Funct. Anal. 3 (1969), 354–375.
- [36] A. Lasota and M. C. Mackey, *Chaos, Fractals and Noise*, Springer, 1994.
- [37] T. Lipniacki, P. Paszek, A. R. Brasier, B. Luxon and M. Kimmel, *Mathematical model of NF- κ B regulatory module*, J. Theor. Biol. 228 (2004): 195–215.
- [38] T. Lipniacki, P. Paszek, A. Marciniak-Czochra, A. R. Brasier and M. Kimmel, *Transcriptional stochasticity in gene expression*, J. Theor. Biol. 238 (2006), 348–367.
- [39] T. Lipniacki, P. Paszek, A. R. Brasier, B. Luxon and M. Kimmel, *Stochastic regulation in early immune response*, Biophys. J. 90 (2006), 725–742.

- [40] M. Lisi and S. Totaro, *The Chapman–Enskog procedure for an age-structured population model: initial, boundary and corner layer corrections*, *Mathematical Biosciences* 196 (2005), 153–186.
- [41] P. Mandl, *Analytical Treatment of One-Dimensional Markov Processes*, Springer, 1968.
- [42] R. Nagel (ed.), *One-Parameter Semigroups of Positive Operators*, Springer, 1986.
- [43] P. Paszek, *Modeling stochasticity in gene regulation: characterization in the terms of the underlying distribution function*, *Bulletin of Mathematical Biology* 69 (2007), 1567–1601.
- [44] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, 1983.
- [45] A. Raj, C. S. Peskin, D. Tranchina, D.Y. Vargas and S. Tyagi, *Stochastic mRNA synthesis in mammalian cells*, *PLoS Biol* 4 (10) (2006): e309. DOI: 10.1371/journal.pbio.0040309.
- [46] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, third edition, Springer, 1999.
- [47] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Vol. 1, Foundations*, Cambridge University Press, 2000.
- [48] M. A. Shea and G. K. Ackers, *The Or control system of bacteriophage lambda: a physical-chemical model for gene regulation*, *J. Mol. Biol.* 181 (1985), 211–230.
- [49] M. Sova, *Convergence d’opérations linéaires non bornées*, *Rev. Roum. Math. Pures et App.* 12 (1967), 373–389.
- [50] M. Thattai and A. van Oudenaarden, *Intrinsic noise in gene regulatory networks*, *Proc. Natl. Acad. Sci. USA* 98 (2001), 8614–8619.
- [51] J. R. R. Tolkien, *The Hobbit, or There and Back Again*, Bellantine Books, 1965.
- [52] K. Yosida, *Functional Analysis*, Springer, 1965.

