Irregular convergence of mild solutions of semi-linear equations

Adam Bobrowski and Markus Kunze

Abstract. We prove that even irregular convergence of semigroups of operators implies similar convergence of mild solutions of the related semi-linear equations with Lipschitz continuous nonlinearity. This result is then applied to three models originating from mathematical biology: shadow systems, diffusions on thin layers, and dynamics of neurotransmitters.

Mathematics Subject Classification (2010). 35K57, 47D06, 35B25, 35K58.

Keywords. semigroups of operators, semi-linear equations, irregular convergence, singular perturbations, boundary conditions, shadow systems, thin layers, signaling pathways.

1. Introduction

A number of phenomena of mathematical biology and mathematical physics, such as reaction-diffusion equations [34], can be modelled by a semilinear equation in a Banach space, i.e., an equation of the form

\[ \frac{du(t)}{dt} = Au(t) + F(t, u(t)), \quad t \geq 0, \]  

(1.1)

where \( A \) is the generator of a strongly continuous semigroup \( (e^{tA})_{t \geq 0} \) of operators in a Banach space \( \mathbb{X} \), and \( F : \mathbb{R}^+ \times \mathbb{X} \to \mathbb{X} \) is a jointly continuous map that is globally Lipschitz continuous in the second variable: there is an \( L > 0 \) such that for \( x, y \in \mathbb{X} \) and \( t \geq 0 \)

\[ \|F(t, x) - F(t, y)\| \leq L\|x - y\|. \]  

(1.2)

(This assumption may be a bit relaxed, see in particular our examples in Section 3 and consult [31] and/or [10]). We should point out that it is the nonlinearity \( F \) which is responsible for many characteristic phenomena which cannot occur in a linear equation. As an example, we mention bistability of
the system and the existence of homo- and hetero-clinical waves, which are critical phenomena for signalling pathways in living cells [13][20].

This paper is devoted to the question of convergence of mild solutions of a sequence of semilinear equations. Thus, given semigroups \((e^{tA_n})_{t \geq 0}\) and \(x \in \mathcal{X}\) we are interested in the question of whether the mild solutions to the equation

\[
\frac{du_n(t)}{dt} = A_n u(t) + F(t, u_n(t)), \quad u_n(0) = x, t \geq 0,
\]

(1.3)

converge or not. We recall that a mild solution of (1.3) is a function \(u_n\) satisfying

\[
u_n(t) = e^{tA_n}x + \int_0^t e^{(t-s)A_n}F(s, u_n(s)) \, ds, \quad t \geq 0,
\]

(1.4)

and it can be proved using Banach’s Fixed-Point Theorem that for each \(x\) such a solution exists and is unique (see e.g. [10], Chapter 36).

In addressing such a question it is sometimes a good strategy to prove convergence of the semigroups \((e^{tA_n})_{t \geq 0}\) first, and only then start to worry about the second term in (1.4). Here the classical convergence theorem for semigroups of operators, often called the Trotter–Kato–Neveu Theorem [10][17][19][31], comes in handy. This theorem asserts that a sequence of semigroups \((e^{tA_n})_{t \geq 0}, n \geq 0\), converges strongly (i.e., for each \(x\) in norm of the Banach space \(\mathcal{X}\)) to a strongly continuous semigroup if and only if:

(i) There exist constants \(C\) and \(\omega\) such that

\[
\|e^{tA_n}\| \leq Ce^{\omega t}, \quad t \geq 0, n \geq 1.
\]

(ii) For \(\lambda > \omega\), the resolvents converge (strongly), i.e., there exists the limit

\[
R_\lambda = \lim_{n \to \infty} (\lambda - A_n)^{-1}.
\]

(iii) The closure of the range of the operator \(R_\lambda\) is the entire \(\mathcal{X}\) for some (and hence for all) \(\lambda > \omega\).

In this case, the operators \(R_\lambda, \lambda > \omega\) form the resolvent of an operator \(A\), i.e., \(R_\lambda = (\lambda - A)^{-1}\) for \(\lambda > \omega\), and \(A\) is the generator of the limit semigroup. We should point out that, if the conditions of the Trotter–Kato–Neveu Theorem are met, the semigroups converge uniformly on the compact subsets of \([0, \infty)\). In this situation, a straightforward use of the Gronwall lemma implies that the mild solutions of (1.3) converge uniformly on compact subsets of \([0, \infty)\) to mild solutions of (1.1).

However, this is not the entire story [10], for quite often, in singular perturbations for example, the third condition in the Trotter–Kato–Neveu Theorem fails. Then, the limit semigroup acts on the regularity space \(\mathcal{X}_0\) defined as the closure of the (common) range of \(R_\lambda\):

\[
\mathcal{X}_0 = \text{cl} \text{Range}R_\lambda
\]

(1.5)

and for \(x \in \mathcal{X}_0\) convergence is still uniform on the compact subsets of \([0, \infty)\); such convergence is termed regular, and therefore \(\mathcal{X}_0\) is called the regularity
Irregular convergence 3

space. There are examples showing that for $x \not\in \mathcal{X}_0$, $\lim_{n \to \infty} e^{tA_n}x$ in general does not exists (see [6] or [10], Chapter 7). Nevertheless, in many important cases it does, except that convergence is not uniform on compact subsets of $(0, \infty)$, but merely uniform on compact subsets of $(0, \infty)$ which is to say that it is uniform in each interval of the form $[\tau^{-1}, \tau], \tau > 1$ (see [7] or [10], Chapter 28). Such convergence is termed irregular.

Thus, in this paper, motivated in particular by biological models presented in our Section 3 we consider the following situation.

- We assume that the semigroups $(e^{tA_n})_{t \geq 0}$ are equibounded, i.e., that condition (i) in the Trotter–Kato–Neveu Theorem holds with $\omega = 0$. This is a customary simplification in semigroup theory, since the general case may be deduced from this one by considering $A'_n = A_n + \omega I$.
- We assume that there is a closed subspace $\mathcal{X}_0$ (the regularity space) and a strongly continuous semigroup $(e^{tA})_{t \geq 0}$ in $\mathcal{X}_0$, such that for $x \in \mathcal{X}_0$,

$$\lim_{n \to \infty} e^{tA_n}x = e^{tA}x$$

uniformly on the compact subsets of $[0, \infty)$. (In fact, $\mathcal{X}_0$ is given by (1.5).)
- There is also a projection $P : \mathcal{X} \to \mathcal{X}_0$ such that for $x \not\in \mathcal{X}_0$,

$$\lim_{n \to \infty} e^{tA_n}x = e^{tA}P x, \quad t > 0$$

uniformly on the compact subsets of $(0, \infty)$.

We note that in choosing these assumptions, besides motivations of biological origin named above, we were guided by T.G. Kurtz’s singular perturbation theorem [10, 18, 24, 25], where assumptions of similar nature are made. Our main theorem reads:

**Theorem 1.1.** Under the stated assumptions, the mild solutions of (1.3) with initial value $x \in \mathcal{X}$ converge to that of

$$\frac{du(t)}{dt} = Au(t) + PF(t, u(t)), \quad t \geq 0,$$

(note the change in the non-linear term) with initial value $P x$. As in the case of semigroups, for $x \in \mathcal{X}_0$, the convergence is uniform on compact subsets of $[0, \infty)$, and for $x \not\in \mathcal{X}_0$ it is uniform on compact subsets of $(0, \infty)$.

It should perhaps be remarked here that the theorem applies in particular to the situation where $F(t, x) = B(t)x$ and $t \mapsto B(t)$ is a strongly continuous family of equibounded operators. In the yet more concrete case where $B$ does not depend on time $t$, this theorem has been obtained in [10] Chapter 29 and its implications for the so-called Stein model of neuronal activity were discussed in Chapter 30 there.

The formulation of our main result presented above is best suited for the motivating models but, remarkably, our proof never uses the fact that the semigroups involved are strongly continuous at $t = 0$. We can thus consider equibounded, but perhaps not continuous at $t = 0$, semigroups $(T_n(t))_{t \geq 0}$.
In order to generalize the variation of constants formula to this setting, strong Bochner measurability of the orbits seems a minimal assumption. As is well known (see [21, Theorem 10.2.3]) this already implies strong continuity on \((0, \infty)\). We will thus assume that all semigroups involved have orbits which are continuous on \((0, \infty)\). As we shall see, for each \(x \in X, n \geq 1\) there exists precisely one continuous function \(u_n : (0, \infty) \to X\) of at most exponential growth, bounded on \((0, 1]\), and such that

\[
u_n(t) = T_n(t)x + \int_0^t T_n(t-s)F(s, u_n(s)) \, ds, \quad t > 0, \tag{1.7}
\]

and of course the same may be said of existence and uniqueness of \(u\) satisfying

\[
u(t) = T(t)x + \int_0^t T(t-s)F(s, u(s)) \, ds, \quad t > 0. \tag{1.8}
\]

In this context our main theorem may be rephrased as follows.

**Theorem 1.2.** Suppose

\[
\lim_{n \to \infty} T_n(t)x = T(t)x, \quad x \in X, t > 0. \tag{1.9}
\]

Then \(u_n\) converge to \(u\) uniformly on the compact subsets of \((0, \infty)\). Additionally, if \(T(0)\) and \(T_n(0)\) are defined for all \(n\) and \(x\) is such that (1.9) is also true for \(t = 0\) and the limit is uniform on the compact subsets of \([0, \infty)\), then also is the corresponding mild solutions \(u_n\) converge uniformly on the compact subsets of \([0, \infty)\) to \(u\).

We then obtain Theorem 1.1 as a direct consequence of Theorem 1.2, setting \(T_n(t) = e^{tA_n}\), and \(T(t) := e^{tA}P\). The latter theorem will be proved in Section 2. The remaining part of the paper is devoted to applications, including shadow systems, diffusion in a thin layer between two parallel planes, and modeling activity of fast neurotransmitters: these are presented in Section 3.

### 2. Proof of Theorem 1.2

Whenever we have a sequence of equations that can be solved by means of the Banach Fixed-Point Theorem, there is a ‘folklore recipe’ to establish convergence of the solutions. All one needs to establish is

(i) a uniform bound on the Lipschitz constants of the fixed-point maps less than 1 and

(ii) pointwise convergence of the fixed-point maps;

see Section 2.2 Proposition 2.1 below for a more precise statement. While this general principle does not appear so often in the study of deterministic equations (where it seems that the use of Gronwall’s Lemma is preferred), it is used quite often in the context of stochastic equations, see [15][23][29].

As so often, the main difficulty in following the path described above is the choice of the ‘right’ underlying space. This is also the case in our situation. While, under assumption of strong continuity of all approximating semigroups, it is possible to use a space of continuous functions on \([0, \infty)\)
to solve a single equation, there is no hope that this space will be suitable for establishing a convergence result, as the maps involved do not converge (uniformly on compact subsets of \([0, \infty)\)). Thus, to prove Theorem 1.2 we proceed in two steps. First, rather than a space of continuous functions, we consider a space of integrable functions as our underlying space and establish convergence of our solutions in the \(L^1\)-sense. It is only afterwards that this result is refined to uniform convergence on compact subsets of \((0, \infty)\). In Section 2.3 we will see in fact that while convergence of \(T_n(t)x\) to \(T(t)x\) may perhaps be non-uniform in the vicinity of \(t = 0\), the other term in (1.7) converges uniformly in any compact subset of \([0, \infty)\).

2.1. Existence of the functions \(u\) and \(u_n\)

In this section, given \(x \in X\), we establish existence and uniqueness of \(u\) satisfying (1.8) (in the sense specified below); this argument certainly applies for \(u_n\), as well. We focus on global existence and for this we assume additionally that there are non-negative constants \(C_0\) and \(\lambda_0\) such that

\[
\int_0^t \|F(s, 0)\| \, ds \leq C_0 e^{\lambda_0 t}, \quad t > 0; \tag{2.1}
\]

we stress that for local existence, i.e., for existence on finite intervals this assumption is not necessary and that it is automatically satisfied if \(F\) does not depend on \(t\) (which is the case in all our examples). Recall that besides Lipschitz continuity described in (1.2) we assume that for certain \(C\),

\[
\|T_n(t)\| \leq C, \quad t > 0 \tag{2.2}
\]

and thus also \(\|T(t)\| \leq C\) for \(t > 0\).

For \(\lambda \geq \lambda_0\), let \(BC_\lambda(\mathbb{R}_0^+, X)\) be the space of continuous functions \(v\) on \(\mathbb{R}_0^+\) with values in \(X\), which are bounded on \((0, 1]\) and are of at most exponential growth \(e^{\lambda t}\), i.e., satisfy

\[
\|v\|_\lambda := \sup_{t > 0} e^{-\lambda t} \|v(t)\| < \infty. \tag{2.3}
\]

When equipped with the norm \(\| \cdot \|_\lambda\), \(BC_\lambda(\mathbb{R}_0^+, X)\) is a Banach space. (This type of norm has been first introduced by Adam Bielecki, see [4][16].) Next, for \(v \in BC_\lambda(\mathbb{R}_0^+, X)\), we introduce

\[
[\Phi(v)](t) = T(t)x + \int_0^t T(t-s)F(s, v(s)) \, ds, \quad t \geq 0.
\]

It is clear that \(\Phi(v)\) is a continuous function on \((0, \infty)\) which on \((0, 1]\) is bounded by

\[
C\|x\| + \int_0^t \|T(t-s)[F(s, v(s)) - F(s, 0)]\| \, ds + \int_0^t \|T(t-s)F(s, 0)\| \, ds
\]

\[
\leq C\|x\| + C \int_0^t \|F(s, v(s)) - F(s, 0)\| \, ds + C \int_0^t \|F(s, 0)\| \, ds
\]

\[
\leq C\|x\| + CL \max_{s \in (0,1]} \|v(s)\| + C \int_0^1 \|F(s, 0)\| \, ds < \infty;
\]
the last integral being finite since $F$ is assumed to be jointly continuous. Also, taking another element of $BC_\lambda(\mathbb{R}_0^+, \mathcal{X})$, say $w$, we have for $t > 0$,

$$e^{-\lambda t} \| [\Phi(v)](t) - [\Phi(w)](t) \| \leq e^{-\lambda t} \int_0^t \| T(t-s)[F(s, v(s)) - F(s, w(s))] \| \, ds$$

$$\leq C L \int_0^t e^{-\lambda (t-s)} e^{-\lambda s} \| v(s) - w(s) \| \, ds$$

$$\leq C L \int_0^t e^{-\lambda (t-s)} \| v-w \|_\lambda \, ds$$

$$< \frac{CL}{\lambda} \| v-w \|_\lambda.$$ (2.4)

For $w = 0$, we have

$$\| [\Phi(w)](t) \| \leq C \| x \| + C \int_0^t \| F(s,0) \| \, ds,$$

proving, by assumption (2.1), that $\Phi(0)$ belongs to all $BC_\lambda(\mathbb{R}_0^+, \mathcal{X}), \lambda \geq \lambda_0$. Therefore, (2.4) with $w = 0$ shows that $\Phi(v)$ belongs to $BC_\lambda(\mathbb{R}_0^+, \mathcal{X})$ as well: $\Phi$ maps $BC_\lambda(\mathbb{R}_0^+, \mathcal{X})$ into itself. Moreover, taking the supremum over $t > 0$ we obtain

$$\| \Phi(v) - \Phi(w) \|_\lambda \leq \frac{CL}{\lambda} \| v-w \|_\lambda.$$ (2.5)

It follows that in the spaces $BC_\lambda(\mathbb{R}_0^+, \mathcal{X})$ with $\lambda > CL$ (and $\lambda \geq \lambda_0$), $\Phi$ is a contraction. The Banach Fixed-Point Theorem shows now existence and uniqueness of a fixed-point of $\Phi$, i.e., existence and uniqueness of a $u$ satisfying (1.8).

One may wonder if the solution thus obtained is also locally unique, but such uniqueness can be ascertained by the same calculation in the space $BC( (0, \tau], \mathcal{X}), \tau > 0$ of bounded, continuous functions on $(0, \tau]$ with values in $\mathcal{X}$ equipped with the Bielecki-type norm

$$\| v \|_\lambda = \sup_{t \in (0, \tau]} e^{-\lambda t} \| v(t) \|.$$ (2.6)

We stress again that for existence (and uniqueness) of local solutions assumption (2.1) is not necessary.

### 2.2. Convergence in the $L^1$ sense

Our goal in this subsection is to prove that for any $\tau > 0$

$$\lim_{n \to \infty} \int_0^\tau \| u_n(s) - u(s) \| \, ds.$$ (2.5)

We will use the following ‘folk wisdom’ theorem.

**Proposition 2.1.** Let $(S, d)$ be a complete metric space and suppose that maps $\Phi_n : S \to S$, $n \geq 1$ are Lipschitz continuous with the same Lipschitz constant $q \in (0, 1)$, i.e., for all $s, s' \in S$ and $n \geq 1$,

$$d(\Phi_n(s), \Phi_n(s')) \leq q d(s, s').$$ (2.6)
Assume also that for each \( s \in S \) the limit
\[
\Phi(s) := \lim_{n \to \infty} \Phi_n(s)
\] (2.7)
exists. Then the unique fixed-points \( s^*_n \) of maps \( \Phi_n \) (which exist by the Banach Fixed-Point Theorem) converge to an \( s^* \in S \) which is the unique fixed-point of the map \( \Phi \).

We present two proofs of this result; the first is by straightforward calculation, and the second is almost by inspection. The advantage of the latter, though a bit longer, argument is that it shows that the proposition is a direct consequence of the Banach Fixed-Point Theorem.

**Proof by calculation.** Combing (2.6) and (2.7) we see that \( \Phi \) is Lipschitz continuous with Lipschitz constant \( q \), and thus possesses precisely one fixed-point. Call it \( s^* \). We have
\[
d(s^*, s^*_n) \leq d(s^*, \Phi_n(s^*)) + d(\Phi_n(s^*), s^*_n)
\]
\[
= d(\Phi(s^*), \Phi_n(s^*)) + d(\Phi_n(s^*), \Phi_n(s^*_n))
\]
\[
\leq d(\Phi(s^*), \Phi_n(s^*)) + q d(s^*, s^*_n).
\]
Thus,
\[
d(s^*, s^*_n) \leq \frac{1}{1-q} d(\Phi(s^*), \Phi_n(s^*))
\]
and the right-hand side converges to 0, by assumption. \( \square \)

**Proof by inspection.** Let \( c(S) \) be the space of convergent sequences of elements of \( S \); this is a complete metric space with metric given by
\[
d_{\sup}((s_n)_{n \geq 1}, (s'_n)_{n \geq 1}) = \sup_{n \geq 1} d(s_n, s'_n).
\]
If \((s_n)_{n \geq 1}\) converges in \( S \) then, by (2.6) and (2.7), so does \((\Phi_n(s_n))_{n \geq 1}\) and we have
\[
\lim_{n \to \infty} \Phi_n(s_n) = \Phi(\lim_{n \to \infty} s_n).
\] (2.8)
Therefore, \((s_n)_{n \geq 1} \mapsto (\Phi_n(s_n))_{n \geq 1}\) defines a map, call it \( \Psi \), in \( c(S) \), and (2.6) shows that \( \Psi \) is Lipschitz continuous with the constant \( q \). The Banach Fixed-Point Theorem asserts now that \( \Psi \) has a unique fixed-point. By the very definition it is clear that the \( n \)-th coordinate of this fixed-point is \( s^*_n \), i.e., we have
\[
(\Phi_n(s^*_n))_{n \geq 1} = \Psi((s^*_n)_{n \geq 1}) = (s^*_n)_{n \geq 1}.
\]
The limit \( s^* = \lim_{n \to \infty} s^*_n \) exists since the fixed-point belongs to \( c(S) \). Using (2.8) we obtain \( \Phi(s^*) = s^* \), proving that \( s^* \) is a fixed-point of \( \Phi \). On the other hand, combing (2.6) and (2.7) we see that \( \Phi \) is Lipschitz continuous with Lipschitz constant \( q \), and thus possesses precisely one fixed-point. \( \square \)

Returning to the proof of (2.5), we fix \( \tau > 0 \) and consider the space \( L^1((0, \tau), \mathcal{X}) \) of (equivalence classes of) Bochner measurable functions on
(0, \tau) with values in \mathbb{X} that are integrable on this interval. This is a Banach space when equipped with the norm
\[ \|v\|_{L^1} = \int_0^\tau \|v(s)\| \, ds \]
but for us it will be convenient to work with the equivalent Bielecki-type norm
\[ \|v\|_{\lambda} = \sup_{t \in (0, \tau]} e^{-\lambda t} \int_0^t \|v(s)\| \, ds \] (2.9)
where \( \lambda > CL \). For \( v \in L^1((0, \tau), \mathbb{X}) \) let
\[ \Phi_n(v)(t) = T_n(t)x + \int_0^t T_n(t-s)F(s,v(s)) \, ds, \quad t \in (0, \tau); \]
\( \Phi_n(v) \) is a measurable function. Since
\[ \|\Phi_n(0)(t)\| \leq C\|x\| + C \int_0^t \|F(s,0)\| \, ds \leq C\|x\| + C \int_0^\tau \|F(s,0)\| \, ds, \]
\( \Phi_n(0) \) is a bounded function, and thus a member of \( L^1((0, \tau), \mathbb{X}) \). Also, for \( w \in L^1((0, \tau), \mathbb{X}) \) and \( t > 0 \),
\[ \int_0^t \|\Phi_n(v)(s) - \Phi_n(w)(s)\| \, ds \leq \int_0^t \int_0^s \|T_n(s-u)[F(u,v(u)) - F(u,w(u))]\| \, du \, ds \]
\[ \leq CL \int_0^t \int_0^s \|v(u) - w(u)\| \, du \, ds. \] (2.10)
Since for \( w = 0 \) the right-hand side does not exceed \( CL\tau\|v\|_{L^1} \), \( \Phi_n(v) \) is a member of \( L^1((0, \tau), \mathbb{X}) \). Thus \( \Phi_n \) maps \( L^1((0, \tau), \mathbb{X}) \) into itself. Multiplying the right-most and left-most members of (2.10) by \( e^{-\lambda t} \) we see also that
\[ \|\Phi_n(v) - \Phi_n(w)\|_{\lambda} \leq CL \sup_{t \in (0, \tau]} \int_0^t e^{-\lambda(t-s)} \, ds \|v - w\|_{\lambda} < \frac{CL}{\lambda}\|v - w\|_{\lambda}. \]
It follows that the maps \( \Phi_n \) are Lipschitz continuous with the same Lipschitz constant \( \frac{CL}{\lambda} \in (0, 1) \) as in our Proposition 2.1. In particular, each of them has a unique fixed-point. This fix point, say \( u_n \), satisfies (1.7) almost surely for \( t \in (0, \tau] \). Since from the previous subsection we know that solutions of (1.7) are unique in \( BC((0, \tau], \mathbb{X}) \) and since members of \( BC((0, \tau], \mathbb{X}) \) belong to \( L^1((0, \tau), \mathbb{X}) \) as well, the so-found solutions of (1.7) coincide with those found earlier. Hence, by Proposition 2.1 (2.5) will be proved once we show that
\[ \lim_{n \to \infty} \Phi_n(v) = \Phi(v) \]
in the sense of \( \| \cdot \|_{\lambda} \) norm or, equivalently, in the norm \( \| \cdot \|_{L^1} \), where, of course,
\[ \Phi(v)(t) = T(t)x + \int_0^t T(t-s)F(s,v(s)) \, ds. \]
Irregular convergence

That \( \lim_{n \to \infty} \int_0^\tau \|T_n(t)x - T(t)x\| \, ds = 0 \) is clear by the Dominated Convergence Theorem, because the integrands here converge pointwise and are dominated by \( 2C\|x\|. \) Hence, it suffices to show that

\[
\lim_{n \to \infty} \int_0^\tau \int_0^t \|T_n(t-s)F(s,v(s)) - T(t-s)F(s,v(s))\| \, ds \, dt = 0. \tag{2.11}
\]

The integrands in

\[
\int_0^t \|T_n(t-s)F(s,v(s)) - T(t-s)F(s,v(s))\| \, ds \tag{2.12}
\]

converge pointwise to 0 and are bounded by

\[
2C\|F(s,v(s))\| \leq 2C\|F(s,v(s)) - F(s,0)\| + 2C\|F(s,0)\|
\leq 2CL\|v(s)\| + 2C\sup_{0 \leq s \leq \tau} \|F(s,0)\|
=: 2CL\|v(s)\| + C_1. \tag{2.13}
\]

Since \( s \mapsto 2CL\|v(s)\| + C_1 \) is integrable on \((0,\tau)\), invoking the Dominated Convergence Theorem again we see that the expression in \((2.11)\) converges to 0 for all \( t \in (0,\tau) \). Moreover, \( \int_0^t \|T_n(t-s)F(s,v(s)) - T(t-s)F(s,v(s))\| \, ds \) does not exceed

\[
2CL \int_0^t \|v(s)\| \, ds + 2C_1t \leq 2CL\|v\|_{L^1} + 2C\tau
\]

so that the same argument proves \((2.11)\).

2.3. Uniform convergence on compact sets

To complete the proof of Theorem 1.2 it suffices to show that

\[
\lim_{n \to \infty} \int_0^\tau T_n(t-s)F(s,u_n(s)) \, ds = \int_0^\tau T(t-s)F(s,u(s)) \, ds
\]

uniformly on \([0,\tau], \tau > 0\). However, the norm of the difference between the two integrals featuring here does not exceed

\[
\int_0^\tau \|T_n(t-s)[F(s,u_n(s)) - F(s,u(s))]\| \, ds + I_n(t)
\leq CL \int_0^\tau \|u_n(s) - u(s)\| \, ds + I_n(t),
\]

where \( I_n(t) \) is the integral appearing in \((2.12)\) with \( v \) replaced by \( u \). Hence, by \((2.5)\), we are left with proving uniform convergence of \( I_n(t), t \in [0,\tau] \) to 0.

We begin by recalling that by assumption for any \( y \in \mathcal{F} \), \( T_n(r)y \) converges to \( T(r)y \) uniformly on compact subsets of \((0,\infty)\). In other words, for each \( y \) and each \( \delta > 0 \),

\[
\sup_{r \in [\delta,\tau]} \|T_n(r)y - T(r)y\| < \epsilon.
\]
may be made arbitrarily small by choosing sufficiently large \( n \). A familiar 3-epsilon argument (using the assumption of equiboundedness (2.2)) shows that for any compact set \( K \subset \mathcal{X} \), the same is true of
\[
\sup_{y \in K} \sup_{r \in [\delta, \tau]} \| T_n(r) y - T(r) y \|.
\]
Now, since \( u \) is bounded on \( (0, \tau] \), our estimate (2.13) (again, with \( v \) replaced by \( u \)) shows that the integrands in \( I_n(t) \) are dominated by a certain common constant. Therefore, given \( \varepsilon > 0 \) one may find a \( \delta > 0 \) such that the norm of the sum of parts of \( I_n(t) \) resulting from integration over \( [0, \delta] \) and \( [t - \delta, t] \) is smaller than \( \frac{\varepsilon}{2} \), regardless of the choice of \( \tau \geq t > 2\delta \). (If \( t < 2\delta \), \( \| I_n(t) \| < \varepsilon \).) Finally, the remaining integral, the one resulting from integration over \( [\delta, t - \delta] \), can by estimated by
\[
\int_{\delta}^{t - \delta} \sup_{y \in K_{\delta}} \| T_n(t - s) y - T_n(t - s) y \| \, ds
\]
\[
\leq (t - 2\delta) \sup_{y \in K_{\delta}} \sup_{r \in [\delta, \tau]} \| T_n(r)y - T(r)y \|,
\]
where \( K_{\delta} \) is the image of \( [\delta, \tau] \) via \( s \mapsto F(s, u(s)) \). Since the latter function is continuous (by continuity of \( u \) and joint continuity of \( F \)), \( K_{\delta} \) is a compact set, and so the last expression may be made smaller than \( \frac{\varepsilon}{2} \) by taking \( n \) sufficiently large. This completes the proof.

3. Examples
3.1. Shadow systems
Let \( S \) be a compact, Hausdorff space, \( N \) and \( N_0 < N \) be two integers, and \( B_i, i = 1, \ldots, N \) be the generators of conservative Feller semigroups in \( C(S) \). Given \( \kappa_n, n \geq 1 \), such that \( \lim_{n \to \infty} \kappa_n = \infty \), we consider the semigroups in \( \mathcal{X} = (C(S))^N \), the Cartesian product of \( N \) copies of \( C(S) \), given by
\[
T_n(t) (f_i)_{i=1,\ldots,N} = \left( e^{tB_1} f_1, \ldots, e^{tB_0} f_{N_0}, e^{t\kappa_n B_{N_0+1}} f_{N_0+1}, \ldots, e^{t\kappa_n B_N} f_N \right),
\]
so that the last \( N - N_0 \) ‘variables’ are ‘fast’, and the first \( N_0 \) are ‘slow’. We assume also that the (strong) limits
\[
\lim_{t \to \infty} e^{tB_i} f =: P_i f, \quad f \in C(S),
\]
exist for each \( N_0 + 1 \leq i \leq N \). Then, for each \( n \geq 1 \), we consider the following system of equations for an unknown function \( t \mapsto u_n(t) = (u_{n,i}(t))_{i=1,\ldots,N} \) with values in \( \mathcal{X} \):
\[
\begin{align*}
    u_{n,i}(t) &= B_i u_{n,i}(t) + F_i(u_n), & t \geq 0, i = 1, \ldots, N_0, \\
    u_{n,i}(t) &= \kappa_n B_i u_{n,i}(t) + F_i(u_n), & t \geq 0, i = N_0 + 1, \ldots, N,
\end{align*}
\]
where $F_i : \mathcal{X} \rightarrow C(S), i = 1, \ldots, N$, are Lipschitz continuous maps. We will show that as $n \to \infty$, mild solutions to these equations with initial condition $u_n(0) = (f_i)_{i=1,\ldots,N} \in \mathcal{X}$ converge to those of the Shadow System:

$$
\begin{align*}
  u_i'(t) &= B_i u_i(t) + F_i(u), & t \geq 0, & i = 1, \ldots, N_0, \\
  u_i'(t) &= P_i F_i(u), & t \geq 0, & i = N_0 + 1, \ldots, N,
\end{align*}
$$

(3.3)

for an unknown function $t \mapsto u(t) = (u_i(t))_{i=1,\ldots,N}$, with initial condition

$$
u(0) = \mathcal{P} (f_i)_{i=1,\ldots,N} := (f_1, \ldots, f_{N_0}, P_{N_0+1} f_{N_0+1}, \ldots, P_N f_N).$$

Indeed, this is a direct consequence of our Theorem 1.2, because we have

$$
\lim_{n \to \infty} T_n(t) (f_i)_{i=1,\ldots,N} = (e^{tB_1} f_1, \ldots, e^{tB_{N_0}} f_{N_0}, P_{N_0+1} f_{N_0+1}, \ldots, P_N f_N),
$$

where $\mathcal{P}$ is defined in (3.4), and $(e^{tA})_{t \geq 0}$ is the strongly continuous semigroup in the Cartesian product

$$
\mathcal{X}_0 := [C(S)]^{N_0} \times \prod_{i=N_0+1}^N \text{Range } P_i
$$
given by

$$
e^{tA} (f_i)_{i=1,\ldots,N} = (e^{tB_1} f_1, \ldots, e^{tB_{N_0}} f_{N_0}, f_{N_0+1}, \ldots, f_N).
$$

In looking for a prototype of this scheme, we go back to the paper by J.P. Keener [22], who studied a model of activator-inhibitor reaction system of the form:

$$
\begin{align*}
  \frac{\partial a}{\partial t} (t, x) &= d_a \frac{\partial^2 a}{\partial x^2} (t, x) + F_1(a(t, x), h(t, x)), \\
  \frac{\partial h}{\partial t} (t, x) &= d_h \frac{\partial^2 h}{\partial x^2} (t, x) + F_2(a(t, x), h(t, x)),
\end{align*}
$$

(3.5)

where both activator $a$ and inhibitor $h$ are distributed on $[0, 1]$ and $F_1$ and $F_2$ and certain Lipschitz continuous functions. Under additional assumptions on $F_1$ and $F_2$, and under Neumann boundary conditions, Keener studies this system in the case where the inhibitor diffusion coefficient $d_h$ is much larger than that for the activator (i.e., $d_a$); he seems to be also responsible for introducing the term Shadow System. Since the Laplace operator with Neumann boundary conditions is known to have property (3.1) with the limit projection mapping an $f \in C([0,1])$ to the constant function being the average of $f$ over $[0,1]$, Keener’s example falls into our scheme, and the Shadow System for (3.5), obtained by letting $d_h \to \infty$, is of the form

$$
\begin{align*}
  \frac{\partial a}{\partial t} (t, x) &= d_a \frac{\partial^2 a}{\partial x^2} (t, x) + F_1(a(t, x), h(t, x)), \\
  \frac{\partial h}{\partial t} (t, x) &= \int_0^1 F_2(a(t, x), h(t, x)) \, dx;
\end{align*}
$$
with initial conditions $a(0, x) = a_0(x), h(0, x) = \int_0^1 h_0(x) \, dx$ where $a_0$ and $h_0$ are initial conditions of (3.3) (see Keener’s equations (5.7a)–(5.7c)).

As a second example, we consider the following (rescaled) model of early lung cancer due to Marciniak–Czochra and Kimmel, see [27, 28]:

$$\frac{\partial c}{\partial t}(t, x) = ((2p - 1)a(b(t, x), c(t, x)) - d_c)c(t, x),$$
$$\frac{\partial b}{\partial t}(t, x) = \alpha(c(t, x))g(t, x) - d_b b(t, x) - db(t, x),$$
$$\frac{\partial g}{\partial t}(t, x) = \frac{1}{\gamma} \frac{\partial^2 g}{\partial x^2} - \alpha(c(t, x))g(t, x) - d_g g(t, x) + \kappa(c(t, x)) + db(t, x), \quad (3.6)$$

where $c, b$ and $g$ are densities of cells, and of bound and free growth factor molecules, respectively, distributed across the unit interval $[0, 1]$ (we keep the notations of [27]). Without going into details on the meaning of functions and coefficients in this system (see also Chapter 36 in [10]), we simply state that as $\gamma \to 0$, i.e., in the case where the growth factor molecules are diffusing very quickly, the solutions to this system converge to those of

$$\frac{\partial c}{\partial t}(t, x) = ((2p - 1)a(b(t, x), c(t, x)) - d_c)c(t, x),$$
$$\frac{\partial b}{\partial t}(t, x) = \alpha(c(t, x))g(t, x) - d_b b(t, x) - db(t, x),$$
$$\frac{\partial g}{\partial t}(t, x) = -\int_0^1 \left[ \alpha(c(t, x))g(t, x) + d_g g(t, x) - \kappa(c(t, x)) \right] dx. \quad (3.7)$$

This is a nearly direct consequence of our main theorem coupled with the fact that the third equation in the model is supplied with Neumann boundary conditions. The only difficulty lies in the fact that the $F_1, F_2$ and $F_3$ functions involved here are not globally Lipschitz continuous: they are merely locally Lipschitz continuous (see e.g. [10, 31] for appropriate definitions). However, it may be shown (see Chapter 36 in [10]) that so-called Müller conditions are satisfied: there are constants $C_1, C_2$ and $C_3$ such that at the boundaries of the region $\mathcal{R}$ where $c \in [0, C_1], b \in [0, C_2]$ and $g \in [0, C_3]$ the vector field formed by $F_1, F_2$ and $F_3$ points towards the region. Then, the remedy is to extend the $F_i$’s to globally Lipschitz continuous maps on the whole of $C([0, 1])$ with the Lipschitz constant suitable for $\mathcal{R}$. Then, the theory already developed asserts existence and uniqueness of solutions to (3.6) with modified $F_i$’s. But the force of Müller conditions is that solutions starting in $\mathcal{R}$ never leave $\mathcal{R}$, and so they are in fact solutions for (3.6) with original $F_i$’s. Hence, a more precise statement should read: mild solutions of (3.6) converge to those of (3.7) provided they start in $\mathcal{R}$.

This connection between (3.6) and (3.7) has been made in [26]. The latter paper also gives more delicate information on the speed of convergence based on heat semigroup estimates to be found e.g., in [32, p. 25] or [35, Lemma 1.3]. However, this result still seemed to suggest that the formation of the Shadow System is somehow related to special properties of the Neumann Laplacian. It was in [9] that we argued that the phenomenon occurs for a
much larger class of Feller generators \( B_i \), namely for those with the property that there exists a projection \( P_i \) such that for certain constants \( \varepsilon \) and \( K \),
\[
\| e^{t B_i} - P_i \| \leq K e^{-\varepsilon t}, \quad t > 0, \quad i = N_0 + 1, \ldots, N,
\]
but the reasoning presented in [9] seems to be flawed. Fortunately, our Theorem 1.1 solves the problem, by proving, as we have seen, our conjecture even under less stringent condition (3.1).

3.2. Diffusion in the thin layer between two parallel planes

In modeling biological phenomena via reaction-diffusion equations one sometimes needs to account for the fact that diffusion occurs in a thin layer between two boundaries. For example, nuclei of so-called B-lymphocytes are so large that it is sensible to think of signal-transmitting kinases as diffusing on a two-dimensional sphere modeling the cell membrane rather than in a three-dimensional region between the nucleus and the membrane (see [12, 13] for more details). A crucial issue related to such approximations is the question of how boundary/transmission conditions featuring in the three-dimensional model are transferred to the two-dimensional model.

Here we consider a somewhat idealized situation in which the boundaries are two parallel planes lying close to each other, and show that in the limit the boundary conditions become legitimate, integral parts of the master equation. As discussed in [13], failure to include these parts in the limit reaction term may render the limit equation incapable of capturing the true nature of biological processes.

More specifically, given \( \varepsilon > 0 \) we consider the thin layer
\[
\Omega_{\varepsilon} := \{(x, y, z) \in \mathbb{R}^3 : 0 \leq z \leq \varepsilon\};
\]
and are concerned with the reaction-diffusion equation in \( \Omega_{\varepsilon} \):
\[
\partial_t u(t, x, y, z) = \partial_x^2 u(t, x, y, z) + \partial_y^2 u(t, x, y, z) + \partial_z^2 u(t, x, y) + F(u(t, x, y, z)), \quad (x, y, z) \in \Omega_{\varepsilon}, t > 0
\]
with the reaction term \( F : \mathbb{R} \to \mathbb{R} \) assumed to be (globally) Lipschitz continuous. On the upper and lower planes \( \{(x, y, \varepsilon) : x, y \in \mathbb{R} \} \) and \( \{(x, y, 0) : x, y \in \mathbb{R} \} \) of \( \Omega_{\varepsilon} \) we impose Robin boundary conditions of the form
\[
\partial_z u(t, x, y, \varepsilon) = -\varepsilon c(x, y) u(t, x, y, \varepsilon), \\
\partial_z u(t, x, y, 0) = \varepsilon d(x, y) u(t, x, y, 0),
\]
where \( c, d : \mathbb{R}^2 \to [0, \infty) \) are given, essentially bounded functions. As explained in [12, 13], the scaling factor (i.e., \( \varepsilon \)) is needed in these boundary conditions; otherwise the limit discussed below will be uninteresting. These boundary conditions describe a stochastic mechanism of removing some of the diffusing particles touching the boundaries.

We want to study the behavior of the solutions to the problem (3.8)–(3.9) as the parameter \( \varepsilon \to 0 \), i.e. when the thin layer collapses to a two-dimensional plane. We will argue that as \( \varepsilon \to 0 \) solutions to our system become more and more ‘flat in the \( z \)-direction’, and thus may be thought of
as functions of two variables. To see this effect more clearly, we look at $\Omega_\varepsilon$ through a magnifying glass, by introducing the change of variables, $\tilde{z} = \varepsilon^{-1}z$, which transforms $\Omega_\varepsilon$ into

$$\Omega := \Omega_1.$$ We write $\tilde{u}(t, x, y, \tilde{z}) = u(t, x, y, \varepsilon^{-1}z)$. A short computation shows that the reaction-diffusion equation (3.8) transforms to

$$\begin{align*}
\partial_t \tilde{u}(t, x, y, \tilde{z}) &= \partial_x^2 \tilde{u}(t, x, y, \tilde{z}) + \partial_y^2 \tilde{u}(t, x, y, \tilde{z}) + \varepsilon^{-2} \partial_{\tilde{z}}^2 \tilde{u}(t, x, y, \tilde{z}) \\
&+ F(\tilde{u}(t, x, y, \tilde{z})), \quad (x, y, \tilde{z}) \in \Omega, t > 0.
\end{align*}$$

(3.10)

while the boundary conditions (3.9) are transformed to

$$\begin{align*}
\partial_x \tilde{u}(t, x, y, 1) &= -\varepsilon^2 c(x, y)\tilde{u}(t, x, y, 1) \\
\partial_{\tilde{z}} \tilde{u}(t, x, y, 0) &= \varepsilon^2 d(x, y)\tilde{u}(t, x, y, 0) \quad x, y \in \mathbb{R}, t > 0.
\end{align*}$$

(3.11)

For notational simplicity we drop the tildes, and then rewrite this system as an abstract evolution equation on the space $L^2(\Omega)$, as follows:

$$\begin{align*}
\partial_t u_\varepsilon(t) &= A_\varepsilon u_\varepsilon(t) + F(u_\varepsilon(t)), \quad u(0) = u_0
\end{align*}$$

(3.12)

where $u_\varepsilon : [0, \infty) \to L^2(\Omega)$ and $A_\varepsilon$ is a suitable realization of the differential operator $\partial_x^2 + \partial_y^2 + \varepsilon^{-2}\partial_{\tilde{z}}^2$, subject to the boundary conditions (3.11). The reaction term $F$, although denoted by the same letter as the function featuring in (3.8), has a slightly different meaning. Namely, for a $u \in L^2(\Omega)$ we may define

$$F(u)(x, y, z) := F(u(x, y, z))$$

(3.13)

where $F$ on the right-hand side is the function from (3.8). Assuming that $F(0) = 0$ or, more generally, that there is a $u \in L^2(\Omega)$ such that $F(u) \in L^2(\Omega)$, we check, using the existence of a global Lipschitz constant for $F$, that (3.13) defines a globally Lipschitz continuous map $L^2(\Omega) \to L^2(\Omega)$, with the Lipschitz constant inherited from $F$. In (3.14) and in what follows, for simplicity of notation (and to comply with notations of Sections 1 and 2), we do not distinguish between $F$ and $F$. As discussed in Section 1 in dealing with well-posedness and convergence of solutions to (3.12), it is a good strategy to work first with the related problem without the nonlinear term:

$$\begin{align*}
\partial_t u_\varepsilon(t) &= A_\varepsilon u_\varepsilon(t), \quad u(0) = u_0,
\end{align*}$$

(3.14)

and we will follow this path. To this end, we will establish well-posedness of the problem (3.14) making use of the theory of sesquilinear forms. We recall that if $H$ is a complex Hilbert space, a sesquilinear form on $H$ is a mapping $a : D(a) \times D(a) \to \mathbb{C}$ which is linear in the first component and antilinear in the second component. It is called symmetric, if $a[u, v] = a[v, u]$. Note that for a symmetric form $a$ we have $a[u, v] := a[u, u] \in \mathbb{R}$ for every $u \in D(a)$. A symmetric form $a$ is called accretive if $a[u, u] \geq 0$ for all $u \in D(a)$; it is called closed, if $D(a)$ is a Hilbert space with respect to the inner product $[u, v]_a = a[u, v] + [u, v]_H$. A sesquilinear form is called densely defined, if $D(a)$ is dense in $H$. 


Given an accretive, symmetric and closed sesquilinear form $a$ that is densely defined, we can define the associated operator $A$ by setting

$$D(A) = \{ u \in D(a) : \exists f \in H : a[u, v] = [f, v]_H \forall v \in D(a) \}$$

and $Au := f$. We thus have $a[u, v] = [Au, v]_H$ for all $u \in D(A)$ and $v \in D(a)$. It is well known that if $A$ is associated to an accretive, symmetric, densely defined and closed sesquilinear form, then $-A$ is the generator of an analytic contraction semigroup on the space $H$. For more information on sesquilinear forms, we refer to [30].

To employ this theory, we have to find a sesquilinear form $a_\varepsilon$ such that $-A_\varepsilon := \partial_x^2 + \partial_y^2 + \varepsilon^{-2}\partial_z^2$ with boundary conditions (3.11) is the associated operator. To this end, we make a formal computation. We multiply $-A_\varepsilon u$ by a $\bar{v} \in L^2(\Omega)$, integrate over $\Omega$ and integrate by parts. This yields

$$- \int_\Omega A_\varepsilon u \bar{v} \, d(x, y, z) = \int_\Omega \left[ \partial_x u \partial_x \bar{v} + \partial_y u \partial_y \bar{v} + \varepsilon^{-2} \partial_z u \partial_z \bar{v} \right] \, d(x, y, z)$$

$$- \varepsilon^{-2} \int_{R^2} \partial_z u(x, y, 1) \bar{v}(x, y, 1) \, d(x, y)$$

$$+ \varepsilon^{-2} \int_{R^2} \partial_z u(x, y, 0) \bar{v}(x, y, 0) \, d(x, y)$$

$$= \int_\Omega \left[ \partial_x u \partial_x \bar{v} + \partial_y u \partial_y \bar{v} + \varepsilon^{-2} \partial_z u \partial_z \bar{v} \right] \, d(x, y, z)$$

$$+ \int_{R^2} c(x, y) u(x, y, 1) \bar{v}(x, y, 1) \, d(x, y)$$

$$+ \int_{R^2} d(x, y) u(x, y, 0) \bar{v}(x, y, 0) \, d(x, y)$$

$$=: a_\varepsilon[u, v]. \quad (3.15)$$

Here we have used the boundary conditions (3.11) in the second step. Our first result is as follows:

**Lemma 3.1.** Define $a_\varepsilon$ by (3.15) on $D(a) = H^1(\Omega)$. Then $a_\varepsilon$ is an accretive, symmetric, densely defined and closed sesquilinear form on $H = L^2(\Omega)$. Consequently, if we denote the operator associated to $a_\varepsilon$ by $-A_\varepsilon$, then $A_\varepsilon$ generates an analytic contraction semigroup on $L^2(\Omega)$.

**Proof.** Obviously, $a_\varepsilon$ is a symmetric sesquilinear form and $D(a)$ is dense in $L^2(\Omega)$. Moreover, we have

$$[u, u]_{a_\varepsilon} = a_\varepsilon[u] + \|u\|^2_{L^2(\Omega)} \geq \min\{1, \varepsilon^{-2}\} \|u\|^2_{H^1(\Omega)} \geq 0.$$

This shows that $a_\varepsilon$ is accretive and that the norm $\| \cdot \|_{a_\varepsilon}$ can be estimated from below by a multiple of the $H^1$-norm. Using the fact that the trace is a bounded operator from $H^1(\Omega)$ to $L^2(\{(x, y, 1) : x, y \in \mathbb{R}\})$ (and from $H^1(\Omega)$ to $L^2(\{(x, y, 0) : x, y \in \mathbb{R}\})$), see, e.g., Part I Case C of [1], Theorem 4.12, and that $c$ and $d$ are essentially bounded, we see that the $\| \cdot \|_{a_\varepsilon}$-norm is actually equivalent to the $H^1$-norm. This proves that $a_\varepsilon$ is closed. \qed
Remark 3.2. If \( u \in H^2(\Omega) \) satisfies the boundary condition (3.9) in the weak sense and \( v \in H^1(\Omega) \), then the formal computation made above is justified and we find that \( -[A_{\varepsilon} u, v]_{L^2(\Omega)} = a[u, v] \). This shows that the operator associated to the form \( a_{\varepsilon} \) is indeed a suitable realization of the differential operator \( A_{\varepsilon} \).

Next we want to let \( \varepsilon \to 0 \). To that end, we use a convergence theorem for symmetric sesquilinear forms due to Simon [33]. In this theorem, the situation is as follows. We are given an increasing sequence \( a_n \) of accretive, symmetric, densely defined and closed sesquilinear forms on a Hilbert space \( H \). That the sequence is increasing means that \( D(a_{n+1}) \subset D(a_n) \) and \( a_n[u] \leq a_{n+1}[u] \) for all \( u \in D(a_{n+1}) \). Let \( -A_n \) be the associated operator. Then \( A_n \) generates an analytic contraction semigroup \( (e^{tA_n})_{t \geq 0} \) on \( H \). We define the limit form \( A \) by setting
\[
D(A) := \left\{ u \in \bigcap_{n \in \mathbb{N}} D(a_n) : \sup_{n \in \mathbb{N}} a_n[u] < \infty \right\}
\]
and \( a[u] := \sup_{n \in \mathbb{N}} a_n[u] \). For \( u \neq v \) the values of \( a[u, v] \) are defined via polarization. It turns out that \( A \) is closed, accretive and symmetric as well. However, \( a \) need not be densely defined. Let us put \( H_0 := \overline{D(a)} \). Then \( a \) is a closed, densely defined, accretive and symmetric sesquilinear form on \( H_0 \). Let \( -A \) be the associated operator on \( H_0 \). Then \( A \) generates an analytic contraction semigroup on \( H_0 \). We extend this semigroup to all of \( H \), setting it 0 on \( H_0^\perp \). With slight abuse of notation, we write \( e^{tA}P \) for this extension, where \( P \) is the orthogonal projection onto \( H_0 \). Simon’s theorem asserts that in such circumstances \( e^{tA_n} \) converges irregularly to \( e^{tA}P \) as \( n \to \infty \).

In the situation of the thin layer, we have the following result:

**Theorem 3.3.** Consider the forms \( a_{\varepsilon} \) defined via (3.15) on \( D(a_{\varepsilon}) = H^1(\Omega) \). This family increases (as \( \varepsilon \) decreases) and the limit form is given as \((a, D(a))\), where
\[
D(a) = \{ u \in H^1(\Omega) \mid \exists u_0 \in H^1(\mathbb{R}^2) : u(x, y, z) = u_0(x, y) \text{ for almost all } x, y, z, 0 < z < 1 \}
\]
and
\[
a[u, v] = \int_{\mathbb{R}^2} [\partial_x u_0(x, y)\partial_x \bar{v}_0(x, y) + \partial_y u_0(x, y)\partial_y \bar{v}_0(x, y)] \, d(x, y)
+ \int_{\mathbb{R}^2} (c(x, y) + d(x, y))u_0(x, y)\bar{v}_0(x, y) \, d(x, y).
\]

**Proof.** Note that \( D(a_{\varepsilon}) \) does not depend on \( \varepsilon \) and that \( a_{\varepsilon}[u] \leq a_{\delta}[u] \) for every \( u \in H^1(\Omega) \) whenever \( \delta \leq \varepsilon \). This shows that \( a_{\varepsilon} \) is increasing as \( \varepsilon \downarrow 0 \). The estimate
\[
a_{\varepsilon}[u] \geq \varepsilon^{-2} \int_{\Omega} |\partial_z u|^2 \, d(x, y, z)
\]
shows that \( \sup_{\varepsilon} a_{\varepsilon}[u] < \infty \) implies that \( \partial_z u = 0 \) almost everywhere. This is equivalent to the fact that \( u \) does not depend on \( z \), i.e. we can find a function
Irregular convergence 17

$u_0 \in H^1(\mathbb{R}^2)$ such that $u(x, y, z) = u_0(x, y)$ for almost every $x, y \in \mathbb{R}$ and $0 < z < 1$. Conversely, it is obvious that for functions with such a representation $\sup_{\varepsilon} a_\varepsilon[u] < \infty$. This proves that the limit form has the claimed domain. Noting that for $u \in D(a)$, $a_\varepsilon[u] = a[u]$ we conclude that also the expression for the limit form itself is clear (by polarization formula).

We note that $\overline{D(a)}$ consists exactly of those functions $f \in L^2(\Omega)$ which do not depend on the $z$-variable. This space is isomorphic to $L^2(\mathbb{R}^2)$. The negative of the operator associated to the limit form $a$ is of the form

$$A = \Delta - (c + d),$$

i.e. it is the Laplacian on $L^2(\mathbb{R}^2)$ perturbed by the potential $c + d$. The Simon theorem now reveals that

$$\lim_{\varepsilon \to 0} e^{tA_\varepsilon} f = e^{tA} P f, \quad t > 0, f \in L^2(\Omega),$$

where $P$, the orthogonal projection on $\overline{D(a)}$, is given by

$$P u(x, y) = \int_0^1 u(x, y, z) \, dz.$$  

To repeat, this result should be interpreted by saying that as $\varepsilon$ tends to zero, mild solutions to (3.14) (which are solutions of (3.8)–(3.10) seen through a magnifying glass) gradually lose dependence on $z$ and become functions of two variables. Their limit dynamics is then governed by the operator $A$. By our main theorem (Theorem 1.1), the same applies to mild solutions to the full reaction-diffusion equations (3.12), and the master equation for the limit dynamics is

$$u'(t) = \Delta u(t) - (c + d) u(t) + PF(u(t)), \quad t \geq 0.$$  

The notable difference between this equation and equation (3.12) (or, as a matter of fact, equation (3.8)) is that, besides the 2D Laplace operator naturally replacing the 3D Laplace operator, the right-hand side here has the additional term, equalling $-(c + d) u(t)$, which is a residue of the boundary conditions (3.11). Let us recall that in the 3D model the latter conditions describe loss of some of the particles touching the boundary. In the 2D limit these conditions disappear along with the boundary and their role is taken over by the term just mentioned; this term also describes loss of some of the diffusing particles, but now this process takes place inside the limiting 2D region.

3.3. Modeling fast neurotransmitters

In this section, we will show that the well-known ODE model of synaptic depression due to Aristizabal and Glavinovič [3] may be obtained as a singular perturbation of the PDE model of Bielecki and Kalita [5], provided that (as discussed in [14] and [8]) in the latter model appropriate transmission conditions are introduced. In our previous paper [11], we showed convergence of the related semigroups of operators, but did not tackle the nonlinearity featuring in the master equation.
In the Bielecki and Kalita model, it is assumed that a terminal bouton $\Omega_0$ is subdivided into 3 subregions $\Omega_1, \Omega_2$ and $\Omega_3$ which correspond to the so-called immediately available (readily releasable), the small (recycling) and the large (resting) pools, respectively (for discussion of alternative names for these pools see [2]), with $\Omega_1$ surrounded by $\Omega_2$ which in turn is surrounded by $\Omega_3$. The master equation for the concentration $u$ of neurotransmitters in $\Omega_0$ is of the form (see eq. (1) in [5])

$$u'(t) = Lu(t) + \beta(u^\sharp - u(t))^+. \tag{3.16}$$

Here, $L$ is a diffusion operator which is complemented with transmission conditions on $\Gamma_{23}$ (the common boundary of $\Omega_2$ and $\Omega_3$) and $\Gamma_{12}$ (the common boundary of $\Omega_1$ and $\Omega_2$), and with Robin boundary conditions on $\Gamma_{01}$ (the outer boundary of $\Omega_0$). The transmission conditions require that the co-normal derivative of the function is proportional to the difference in concentration on both sides of the boundary, and thus allow seeing $\Gamma_{23}$ and $\Gamma_{21}$ as semi-permeable membranes with permeability changing along the boundary; permeability from $\Omega_i$ to $\Omega_j$ and that from $\Omega_j$ and $\Omega_i$ may also differ. Moreover, $\beta : \Omega_0 \to \mathbb{R}$ is a bounded function which is zero on $\Omega_1 \cup \Omega_2$, $u^\sharp$ is a positive constant and $u^+ = \max(u, 0)$. Thus, the non-linearity is a production term: if $u(t)$ is smaller than $u^\sharp$ at a point of the large pool $\Omega_3$, new neurotransmitters are produced there. Certainly, (3.16) is an equation of the form (1.1) with $A = L$ and

$$F(t, u) = F(u) = \beta(u^\sharp - u)^+, \quad u \in L^2(\Omega_0).$$

In [11], we studied the linearized version of (3.16) where the non-linearity $\beta(u^\sharp - u(t))^+$ was replaced by $\beta(u^\sharp - u(t))$. As a particular case of a more general principle of approximating fast diffusions by Markov chains, we proved there that if instead of $L$ we consider the operator $L_\kappa$, obtained by replacing the matrix $D$ of diffusion coefficients of the operator $L$ by $\kappa D$, the associated contraction semigroups $(e^{tL_\kappa})$ converge irregularly, as $\kappa \to \infty$:

$$\lim_{\kappa \to \infty} e^{tL_\kappa} = \Phi e^{tQ}\Phi^{-1}P, \quad t > 0.$$ 

Here $P$ given by

$$Pu = \sum_{i=1}^3 \frac{1}{\lambda(\Omega_i)} \int_{\Omega_i} u \, d\lambda \mathbb{1}_{\Omega_i}, \quad u \in L^2(\Omega_0), \tag{3.17}$$

is the orthogonal projection on the space $\mathcal{X}_0$ of functions $u \in L^2(\Omega_0)$ which are constant on the sets $\Omega_1, \Omega_2$ and $\Omega_3$, $\lambda$ denotes the Lebesgue measure, and $\Phi$ is the isometric isomorphism between $\mathbb{C}^3$, equipped with the norm

$$\|(z_1, z_2, z_3)\| = \sqrt{\sum_{i=1}^3 |z_i|^2 \lambda(\Omega_i)},$$
and \( \mathcal{X}_0 \), given by
\[
\Phi(u_1, u_2, u_3) = \sum_{j=1}^{3} u_j \mathbb{1}_{\Omega_j}, \quad (u_1, u_2, u_3) \in \mathbb{C}^3.
\]

Finally, \( Q \) is a certain intensity (or: Kolmogorov) matrix, describing a Markov chain with three states corresponding to three aggregated regions \( \Omega_1, \Omega_2, \Omega_3 \).

As discussed in [11] (see also [11], Section 7.2), \( Q \), whose entries are obtained by integrating permeability coefficients over the separating membranes, may be identified with the matrix governing the ODE model of Aristizabal and Glavinovič.

Making use of our main theorem, we can now tackle the original equation (3.16). Indeed, since \( \beta \) is assumed to be bounded, say by a constant \( M \), we have
\[
\|F(u) - F(v)\|_{L^2(\Omega_0)} \leq M\|u - v\|_{L^2(\Omega_0)}
\]
for any \( u, v \in L^2(\Omega_0) \). This means that the non-linear term is Lipschitz continuous. It follows from Theorem 1.1 that, as \( \kappa \) converges to infinity, solutions to (3.16) (with \( \mathcal{L} \) replaced by \( \mathcal{L}_\kappa \)) converge to an \( \mathcal{X}_0 \)-valued function, which via \( \Phi \) defined above may be identified with the \( \mathbb{C}^3 \)-valued solution \( u(t) \) of the equation
\[
u'(t) = Qu(t) + \Phi^{-1}PF(\Phi(u(t))). \tag{3.18}
\]
Because of (3.17), since \( \beta \) is identically zero on \( \Omega_1 \) and \( \Omega_2 \), we see that \( \Phi^{-1}PF\Phi \) maps a vector \((u_1, u_2, u_3) \in \mathbb{C}^3\) to \((0, 0, v_3)\) where
\[
v_3 = \frac{1}{\lambda(\Omega_3)} \int_{\Omega_3} \beta \, d\lambda (u^\sharp - u_3)^+.
\]

Therefore, also the non-linearity in the limit equation is a production term: the third coordinate of a solution of (3.18) grows if it happens to be below \( u^\sharp \), and the rate of growth is \( \frac{1}{\lambda(\Omega_2)} \int_{\Omega_3} \beta \, d\lambda \).

**Acknowledgment.** This research is supported by National Science Center (Poland) grant 2017/25/B/ST1/01804.

**References**

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Second, Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[2] A. A. Alabi and R.W. Tsien, *Synaptic vesicle pools and dynamics*, Cold Spring Harbor Perspectives in Biology 4 (2012), a013680.

[3] F. Aristizabal and M. I. Glavinovic, *Simulation and parameter estimation of dynamics of synaptic depression*, Biol. Cybern. 90 (2004), 3–18.

[4] A. Bielecki, *Une remarque sur la méthode de Banach–Cacciopoli–Tikhonov*, Bull. Polish Acad. Sci. 4 (1956), 261–268.

[5] A. Bielecki and P. Kalita, *Model of neurotransmitter fast transport in axon terminal of presynaptic neuron*, J. Math. Biol 56 (2008), 559–576.

[6] A. Bobrowski, *Degenerate convergence of semigroups*, Semigroup Forum 49 (1994), no. 3, 303–327.

[7] ______, *A note on convergence of semigroups*, Ann. Polon. Math. 69 (1998), no. 2, 107–127.
From diffusions on graphs to Markov chains via asymptotic state lumping, Ann. Henri Poincare 13 (2012), 1501–1510.

Singular perturbations involving fast diffusion, J. Math. Anal. Appl. 427 (2015), no. 2, 1004–1026.

Convergence of One-parameter Operator Semigroups. In Models of Mathematical Biology and Elsewhere, New Mathematical Monographs, vol. 30, Cambridge University Press, Cambridge, 2016.

A. Bobrowski, B. Kaźmierczak, and M. Kunze, An averaging principle for fast diffusions in domains separated by semi-permeable membranes, Mathematical Models and Methods in Applied Sciences 27 (2017), no. 04, 663–706, available at http://www.worldscientific.com/doi/pdf/10.1142/S0218202517500130.

A. Bobrowski and T. Lipniacki, Singular limit of diffusion equations in 3D domains with thickness converging to zero, Models and Reality: Festschrift For James Robert Thompson, edited by J.A. Dobelman, 2017, pp. 95–116.

Robin-type boundary conditions in transition from reaction-diffusion equations in 3D domains to equations in 2D domains, 2018.

A. Bobrowski and K. Morawwska, From a PDE model to an ODE model of dynamics of synaptic depression, Discr. Cont. Dyn. Syst. B 17 (2012), no. 7, 2313–2327.

Z. Brzézniak, On stochastic convolution in Banach spaces and applications, Stochastics Stochastics Rep. 61 (1997), no. 3-4, 245–295.

R. E. Edwards, Functional Analysis. Theory and Applications, Dover Publications, 1995.

K.-J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Springer, New York, 2000.

S. N. Ethier and T. G. Kurtz, Markov Processes. Characterization and Convergence, Wiley, New York, 1986.

J. A. Goldstein, Semigroups of Linear Operators and Applications, Oxford University Press, New York, 1985.

B. Hat, B. Kaźmierczak, and T. Lipniacki, B cell activation triggered by the formation of the small receptor cluster: a computational study, PLoS Comput Biol. 7(10) (2011 Oct.), e1002197.

E. Hille and R. S. Phillips, Functional Analysis and Semi-Groups, Amer. Math. Soc. Colloq. Publ. 31, Amer. Math. Soc., Providence, R. I., 1957.

J. P. Keener, Activators and inhibitors in pattern formation, Stud. Appl. Math. 59 (1978), no. 1, 1–23.

M. Kunze and J. van Neerven, Approximating the coefficients in semilinear stochastic partial differential equations, J. Evol. Equ. 11 (2011), no. 3, 577–604.

T. G. Kurtz, A limit theorem for perturbed operator semigroups with applications to random evolutions, J. Functional Analysis 12 (1973), 55–67.

Applications of an abstract perturbation theorem to ordinary differential equations, Houston J. Math. 3 (1977), no. 1, 67–82.

A. Marciniak-Czochra, S. Härting, G. Karch, and K. Suzuki, Dynamical spike solutions in a nonlocal model of pattern formation, Nonlinearity 31 (2018), no. 5, 1757–1781.

A. Marciniak-Czochra and M. Kimmel, Modelling of early lung cancer progression: influence of growth factor production and cooperation between partially transformed cells, Math. Mod. Meth. Appl. Sci. 17 (2007), 1693–1719.

Reaction-diffusion model of early carcinogenesis: the effects of influx of mutated cells, Math. Model. Nat. Phenom. 3 (2008), no. 7, 90–114.

C. Marinelli, C. Prévôt, and M. Röckner, Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise, J. Funct. Anal. 258 (2010), no. 2, 616–649.

E. M. Ouhabaz, Analysis of Heat Equations on Domains, Lond. Math. Soc. Monograph Series, vol. 30, Princeton Univ. Press, Princeton, 2005.
REFERENCES

[31] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer, 1983.

[32] F. Rothe, Global Solutions of Reaction-Diffusion Systems, Lecture Notes in Mathematics, vol. 1072, Springer-Verlag, Berlin, 1984.

[33] B. Simon, A canonical decomposition for quadratic forms with applications to monotone convergence theorems, J. Functional Analysis 28 (1978), no. 3, 377–385.

[34] J. A. Smoller, Shock-Waves and Reaction-Diffusion Equations, Springer, 1994.

[35] M. Winkler, Aggregation vs. global diffusive behavior in the higher-dimensional Keller-Segel model, J. Differential Equations 248 (2010), no. 12, 2889–2905.

Adam Bobrowski
Lublin University of Technology
Nadbystrzycka 38A
20-618 Lublin, Poland
e-mail: a.bobrowski@pollub.pl

Markus Kunze
Universität Konstanz,
Fachbereich Mathematik und Statistik,
78457 Konstanz, Germany
e-mail: markus.kunze@uni-konstanz.de