Positivity in the applied sciences

J. Banasiak, School of Mathematical Sciences,
University of KwaZulu-Natal, Durban 4041, South Africa

Banach Centre/CIME COURSE and WORKSHOP
”FROM A MICROSCOPIC TO A MACROSCOPIC
DESCRIPTION OF COMPLEX SYSTEMS”

Contents

1 Introduction

Laws of physics and, increasingly, also those of other sciences are in many cases expressed in terms of differential or integro–differential equations. If one models systems evolving with time, then the variable describing time plays a special role, as the equations are built by balancing the change of the system in time against its 'spatial' behaviour. In mathematics such equations are called *evolution equations*.

Equations of the applied sciences are usually formulated pointwise; that is, all the operations, such as differentiation and integration, are understood in the classical (calculus) sense and the equation itself is supposed to be satisfied for all values of the independent variables in the relevant domain:

$$\begin{aligned} \frac{\partial}{\partial t}u(t,x) &= [\mathcal{A}u(t,\cdot)](x), \quad x\in\Omega \\ u(t,0) &= \overset{\circ}{u}, \end{aligned} \tag{1}$$

where $\mathcal{A}$ is a certain expression, differential, integral, or functional, that can be evaluated at any point $x \in \Omega$ for all functions from a certain subset S.

When we are trying to solve (1), we change its meaning by imposing various *a priori* restrictions on the solution to make it amenable to particular techniques. Quite often (1) does not provide a complete description of the dynamics even if it looks complete from the modelling point of view. Then the obtained solution maybe be not what we have been looking for. This becomes particularly important if we cannot get our hands on the actual solution but use 'soft analysis' to find important properties of it. These lecture notes are devoted predominantly to one particular way of looking at the evolution of a system in which we describe time changes as transitions from one state to another; that is, the evolution is described by a family of operators $(G(t))_{t\geq 0}$, parameterised by time, that map an initial state of the system to all subsequent states in the evolution; that is solutions are represented as

$$u(t) = G(t)u_0, \tag{2}$$

where $(G(t))_{t\geq 0}$ is the semigroup and u_0 is an initial state.

In this case we place everything in some abstract space X which is chosen partially for the relevance to the problem and partially for mathematical convenience. For example, if (1) describes the evolution of an ensemble of particles, then u is the particle density function and the natural space seems to be $L_1(\Omega)$ as in this case the norm of a nonnegative u, that is, the integral over Ω, gives the total number of particles in the ensemble. It is important to note that this choice is not unique but is rather a mathematical intervention into the model, which could change it in a quite dramatic way. For instance, in this case we could choose the space of measures on Ω with the same interpretation of the norm, but also, if we are interested in controlling the maximal concentration of particles, a more proper choice would be some reasonable space with a supremum norm, such as, for example, the space of bounded continuous functions on Ω, $C_b(\Omega)$. Once we select our space, the right-hand side can be interpreted as an operator $A : D(A) \to X$ (we hope) defined on some subset $D(A)$ of X (not necessarily equal to X) such that $x \to [\mathcal{A}u](x) \in X$. With this, (1) can be written as an ordinary differential equation in X:

$$\begin{aligned} u_t &= Au, \qquad t > 0, \\ u(0) &= u_0 \in X. \end{aligned} \tag{3}$$

The domain $D(A)$ is also not uniquely defined by the model. Clearly, we would like to choose it in such a way that the solution originating from $D(A)$ could be differentiated and belong to $D(A)$ so that both sides of the equation make sense. As we shall see, semigroup theory in some sense forces $D(A)$ upon us, although it is not necessarily the optimal choice from a modelling point of view. Although throughout the lectures we assume that the underlying space is given, the choice of $D(A)$, on which we define the realisation A of the expression $\mathcal{A}$, is a more complicated thing and has major implications as to whether we are getting from the model what we bargained for.

Though we also discuss a general theory, we focus on models preserving some notion of *positivity*: non-negative inputs should give non-negative outputs (in a suitable sense of the word).

1.1 What can go wrong?

Dishonesty. Models are based on certain laws coming from the applied sciences and we expect the solutions to equations of these models to return these laws. However, this is not always true: we will see models built on the basis of population conservation principles, solutions of which, for certain classes of parameters, do not preserve populations. Such models are called *dishonest*. Dishonesty could be a sign of a phase transition happening in the model, or simply indicate limits of validity of the model.

Multiple solutions. Even if all side conditions relevant to a physical process seems to have been built into the model, we may find that the model does not provide full description of the dynamics; while for some classes of parameters the model gives uniquely determined solutions, for others there exist multiple solutions.

We will see that methods based on positivity methods provided a comprehensive explanation of these two 'pathological' phenomena.

1.2 And if everything seems to be fine?

If we make sure that the abstract model (3) gives as a reasonable description of the phenomena at hand, we can analyse its further properties. One of the most frequent questions asked by practitioners is stability and long time behaviour of solutions. In particular, in population theory an important problem is the existence of dominating long time pattern of evolution. More precisely, we can pose the following questions:

1. Does there exists a special solution u^* of (3) of the form $u^*(t) = e^{\lambda^* t}u_0^*$ for some real λ^* and an element $u_0^* \in X$ such that for any other solution there is a constant C such that

$$u(t) = Ce^{\lambda^* t}u_0^* + O(\exp{(\lambda^* - \epsilon)t}) \quad (4)$$

for some $\epsilon > 0$ (independent of u)? An added bonus would be if u_0^* could be selected positive.

2. If this is impossible, may be there is a finite dimensional projection P, which commutes with the semigroup $G(t)$ and such that

$$e^{-\lambda^* t}G(t) - P \to 0, \quad \text{exponentially fast.} \quad (5)$$

3. More generally, we may ask whether there exists a finite dimensional projection P, which commutes with the semigroup $G(t)$ and such that $G(t)|_{PX}$

can be extended to a group of operators of the form e^{tM} with all eigenvalues of M satisfying $\Re\lambda = \lambda^*$ and

$$(I-P)G(t) = O(e^{-(\lambda^*-\epsilon)t}), \quad \text{as } t \to \infty. \tag{6}$$

for some $\epsilon > 0$.

Those familiar with the finite dimensional population theory will recognize that in the first case we have primitive irreducible transition matrix while in the second and third the matrix is only irreducible with different properties of the largest eigenvalue.

We say that the semigroup $(G(t))_{t\geq 0}$ has *asynchronous exponential growth* (AEG) if (4) is verified (and positive AEG if u^* is positive). If only (5) is satisfied, then we say that $(G(t))_{t\geq 0}$ has *multiple asynchronous growth* (MAEG) and, finally, if (6) holds, then we say that $(G(t))_{t\geq 0}$ has *extended asynchronous growth* (EAEG).

The name 'asynchronous exponential growth' comes precisely from the population biology when it is observed that in many cases initially synchronized populations lose synchrony after just a few generations. It reflects the fact that whatever distribution was observed at an initial times, the population evolves towards an asymptotic distribution, where the proportion of individuals in a given stage is constant.

Here the interplay of compactness and positivity techniques can produce in infinite dimension results which are very close to the classical Frobenius-Perron theory.

However, unlike in finite dimension, some models which behave perfectly well for some classes of parameters, can degenerate into chaotic behaviour for others. We shall demonstrate this on two examples. One is taken from classical birth-and-death type problems, the other in a variant of the age structured population model.

It is worthwhile to note that phase transitions and chaos usually are associated with nonlinear phenomena. Here we will see that they can occur in linear ones but for this the latter must be infinite dimensional.

2 Spectral properties of operators

The considerations below will be carried in an arbitrary Banach space. However, most applications in the present lectures are restricted to the Banach spaces which are commonly used in the population theory due to their natural interpretation. It is worthwhile to understand that, in applications, working in a particular Banach space means simply that the functions we are working must satisfy a numerical restriction which is important in the modelling process. In population theory usually we are interested in the evolution of an ensemble of elements the state of which is described by a function $n(t,x)$ representing either a number of elements in a given state (if the number of states is finite or countable) or the density of particles in the state x, if x is a continuous variable. In many cases we are interested in tracking the total number of elements of the population which, for a time t, is given by

$$\sum_{x\in\Omega} n(t,x),$$

where Ω is the state space, if Ω is countable and

$$\int_{x\in\Omega} n(t,x)dx,$$

if Ω is a continuum. To make a full use of the tools of the functional analysis, we must allow entries of arbitrary sign, so it is not surprising that using this point of view we are working either in

$$l_1 := \{(n_i)_{i\in\mathbb{N}};\ \sum_{i=1}^{\infty} |n_i| < \infty\}$$

or

$$L_1(\Omega) := \{x \to n(x);\ \int\limits_{\Omega} |n(x)|dx < \infty\},$$

where in the first case we took $\Omega = \mathbb{N}$.

Of course, as noted in Introduction, if we are more interested in maximal concentration of elements, we should rather work in spaces of functions with supremum norm.

If uncomfortable with abstract notions, one can substitute one of the spaces described above for a general X to get a better understanding of the main ideas of the lectures.

2.1 Operators

Let X, Y be real or complex Banach spaces with the norm denoted by $\|\cdot\|$ or $\|\cdot\|_X$.

An *operator* from X to Y is a linear rule $A : D(A) \to Y$, where $D(A)$ is a linear subspace of X, called the *domain* of A. We use the notation $(A, D(A))$ to denote the operator A with domain $D(A)$.

By $\mathcal{L}(X, Y)$, we denote the space of all bounded operators between X and Y; that is, the operators for which

$$\|A\| := \sup_{\|x\|\leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| < +\infty. \tag{7}$$

The above defines a norm which turns $\mathcal{L}(X, Y)$ into a Banach space. by introducing the norm $\mathcal{L}(X, X)$ is abbreviated as $\mathcal{L}(X)$.

If $Y \subset X$ is a linear space, then the *part* of A in Y is defined as

$$A_Y y = Ay, \quad D(A_Y) = \{x \in D(A) \cap Y;\ Ax \in Y\}. \tag{8}$$

A *restriction of operator* of $(A, D(A))$ to $D \subset D(A)$ is denoted by $A|_D$.

For $A, B \in L(X, Y)$, we write $A \subset B$ if $D(A) \subset D(B)$ and $B|_{D(A)} = A$.

Two operators $A, B \in \mathcal{L}(X)$ are said to commute if $AB = BA$. An arbitrary operator A is said to *commute* with $B \in \mathcal{L}(X)$ if

$$BA \subset AB. \tag{9}$$

This means that for any $x \in D(A)$, $Bx \in D(A)$ and $BAx = ABx$.

We define the *image* of A by

$$Im\ A = \{y \in Y;\ y = Ax \text{ for some } x \in D(A)\}$$

and the *kernel* of A by

$$Ker\ A = \{x \in D(A);\ Ax = 0\}.$$

Furthermore, the *graph* of A is defined as

$$G(A) = \{(x, y) \in X \times Y;\ x \in D(A), y = Ax\}. \tag{10}$$

We say that the operator A is *closed* if $G(A)$ is a closed subspace of $X \times Y$. Equivalently, A is closed if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$, if $\lim_{n\to\infty} x_n = x$ in X and $\lim_{n\to\infty} Ax_n = y$ in Y, then $x \in D(A)$ and $y = Ax$.

An operator A in X is *closable* if the closure of its graph $\overline{G(A)}$ is itself a graph of an operator, that is, if $(0, y) \in \overline{G(A)}$ implies $y = 0$. Equivalently, A is closable if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$, if $\lim_{n\to\infty} x_n = 0$ in X and $\lim_{n\to\infty} Ax_n = y$ in Y, then $y = 0$. In such a case the operator whose graph is $\overline{G(A)}$ is called the *closure* of A and denoted by $\overline{A}$.

2.1.1 Compact operators

Let us recall that an operator $K \in \mathcal{L}(X, Y)$, X, Y-Banach spaces, is compact (resp. weakly compact) if the image of the unit ball in X is a relatively compact (resp. weakly compact) subset of Y.

Most relevant properties of compact operators are preserved if the operator $K \in X$ is power compact; that is, if K^m is compact for some $m \in \mathbb{N}$.

Importance of power compact operators stems, in particular, from the fact that in certain spaces ($C(\Omega)$, $L_1(\Omega)$) the square of a weakly compact operator is compact.

In applications it is often needed that AK be power compact for any $A \in \mathcal{L}(X)$. Such operators are called *strictly power compact*. Since the the space of weakly compact (and also compact, for that matter) operators is a two sided ideal in $\mathcal{L}(X)$, weakly compact operators in $C(\Omega), L_1(\Omega)$ are strictly power compact.

Example 1 Consider the integral operator given formally by

$$Tf(x) = \int\limits_{\Omega} k(x,y)f(y)dy,$$

where $\Omega \subseteq \mathbb{R}^n$. The operator T is compact from $L_p(\Omega)$ to $L_p(\Omega)$ if $k \in L_{p,q}(\Omega \times \Omega)$, where $1/p + 1/q = 1$, provided $p > 1$. For $p = 1$, the assumption corresponding assumption

$$k \in L_{1,\infty}(\Omega \times \Omega) \tag{11}$$

is not sufficient for compactness. For such K to be compact, we require e.g. $k \in C(\Omega, L_\infty(\Omega))$ (see [36, p.53]). However, under assumption (11), the operator K is weakly compact and thus strictly power compact ([24]).

2.2 Spectral properties of a single operator

Let A be any operator in X. The *resolvent set* of A is defined as

$$\rho(A) = \{\lambda \in \mathbb{C};\ \lambda I - A : D(A) \to X \text{ is invertible}\}. \tag{12}$$

We call $(\lambda I - A)^{-1}$ the resolvent of A and denote it by

$$R(\lambda, A) = (\lambda I - A)^{-1}, \quad \lambda \in \rho(A).$$

The complement of $\rho(A)$ in $\mathbb{C}$ is called the *spectrum* of A and denoted by $\sigma(A)$. In general, it is possible that either $\rho(A)$ or $\sigma(A)$ is empty. The spectrum is usually subdivided into several subsets. We follow the approach of [38, 26] which, though being not the most common, is very suitable for the description of asymptotics of semigroups. The most important is

- *Point spectrum* $\sigma_p(A)$ is the set of $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is not one-to-one. In other words, $\sigma_p(A)$ is the set of all eigenvalues of A.

A generalization of the point spectrum which will play an important role later is the approximate spectrum:

- *Approximate spectrum* $\sigma_a(A)$ is the set of $\lambda \in \sigma(A)$ for which either the operator $\lambda I - A$ is not one-to-one or the range $Im\ A$ is not closed.

The name *approximate spectrum* comes from the following property which is often used to as a definition.

Lemma 1 *If $(A, D(A))$ is a closed operator in X, then $\lambda \in \sigma_a(A)$ if and only if there is a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ such that $\|x_n\| = 1$, $n \in \mathbb{N}$, and* $\lim_{n\to\infty} \|Ax_n - \lambda x_n\| = 0$.

The last part of the spectrum is

- *Residual spectrum* $\sigma_r(A)$ is the set of $\lambda \in \sigma(A)$ for which $Im\ (\lambda I - A)$ is not dense in X.

Clearly, the σ_p, σ_a and σ_r are not disjoint (in particular, $\sigma_p \subset \sigma_a$) but we clearly have

$$\sigma(A) = \sigma_a(A) \cup \sigma_r(A).$$

Moreover, $\sigma_r(A) = \sigma_p(A^*)$ (A^* denotes the adjoint of A) and the topological boundary of $\sigma(A)$ satisfies

$$\partial\sigma(A) \subset \sigma_a(A) \tag{13}$$

Remark 1 Typically, $\sigma(A)$ is divided into $\sigma_p(A)$ (defined as above), the continuous spectrum $\sigma_c(A)$ which is the set of $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is one-to-one and its range is dense in X but not equal to X and the residual spectrum is defined as the set of $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is one-to-one and its range is not dense in X. Clearly, $\sigma_c(A) \subset \sigma_a(A)$ but we shall not explore further relations between these two definitions. However, the continuous spectrum will come in handy in e.g. Theorem 45.

The resolvent of any operator A satisfies the *resolvent identity*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A), \qquad \lambda, \mu \in \rho(A), \tag{14}$$

from which it follows, in particular, that $R(\lambda, A)$ and $R(\mu, A)$ commute. It follows that $\rho(A)$ is an open set and $R(\lambda, A)$ is an analytic function of $\lambda \in \rho(A)$ which can be written as the power series

$$R(\lambda, A) = \sum_{n=0}^{\infty}(\mu - \lambda)^n R(\mu, A)^{n+1} \tag{15}$$

for $|\mu - \lambda| < \|R(\mu, A)\|^{-1}$. For any bounded operator the spectrum is a compact subset of $\mathbb{C}$ so that $\rho(A) \neq \emptyset$. If A is bounded, then the limit

$$r(A) = \lim_{n\to\infty} \sqrt[n]{\|A^n\|} \tag{16}$$

exists and is called *the spectral radius*. Clearly, $r(A) \leq \|A\|$.

Theorem 2 *The spectral radius of A has the following properties.*

(i) We have

$$R(\lambda, A) = \sum_{n=0}^{\infty} \lambda^{-(n+1)} A^n, \tag{17}$$

where the series converges in the operator norm for $|\lambda| > r(A)$.

(ii) For $|\lambda| < r(A)$ the series in (17) diverges (in the operator norm).

(iii)

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda|. \tag{18}$$

To show that a point $\lambda \in \mathbb{C}$ belongs to the spectrum we often use the following result.

Theorem 3 *Let A be a closed operator. If $\lambda \in \rho(A)$, then*

$$dist(\lambda, \sigma(A)) = \frac{1}{r(R(\lambda, A))} \geq \frac{1}{\|R(\lambda, A)\|}. \tag{19}$$

In particular, if $\lambda_n \to \lambda$, $\lambda_n \in \rho(A)$, then $\lambda \in \sigma(A)$ if and only if $\{\|R(\lambda_n, A)\|\}_{n\in\mathbb{N}}$ is unbounded.

The concept of the spectral radius allows to introduce another frequently used part of the spectrum. The *peripheral spectrum* of a bounded operator A is the set

$$\sigma_{per,r(A)} = \{\lambda \in \sigma(A);\ |\lambda| = r(A)\}. \tag{20}$$

Clearly, $\sigma_{per,r(A)}(A)$ is compact and, by (18), non-empty.

For an unbounded operator A the role of the spectral radius often is played by the *spectral bound* $s(A)$ defined as

$$s(A) = \sup\{\Re\lambda;\ \lambda \in \sigma(A)\}, \tag{21}$$

and the peripheral spectrum of A in this case is correspondingly defined as

$$\sigma_{per,s(A)} = \{\lambda \in \sigma(A);\ Re\lambda = s(A)\}. \tag{22}$$

An important role in analysis is played by the Spectral Mapping Theorem.

Suppose $A \in \mathcal{L}(X)$ and $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is an analytic function in a disc containing $\sigma(A)$. Then we can define a function $f(A)$ by

$$f(A) = \sum_{i=0}^{\infty} a_n A^n$$

where the series is convergent as $\sigma(A)$ is contained in a circle with radius $r(A)$. An alternative definition can be obtained by the Dunford integral

$$f(A) = (2\pi i)^{-1} \int_{\gamma} f(\lambda) R(\lambda, A) d\lambda,$$

where γ is a closed contour surrounding $\sigma(A)$.

Spectra of A and $f(A)$ are related by the Spectral Mapping Formula

$$\sigma(f(A)) = f(\sigma(A)). \tag{23}$$

2.2.1 Decomposition of the spectrum

Let A be a closed operator. An important case occurs if $\sigma(A)$ can be decomposed into two disjoint parts, one of which is compact and the other closed. We shall focus on the case when the compact part consists of an isolated point λ_0 of $\sigma(A)$. This means that the resolvent can be expanded into a Laurent series

$$R(\lambda, A) = \sum_{n=-\infty}^{\infty} (\lambda - \lambda_0)^n B_n \tag{24}$$

for $0 < |\lambda - \lambda_0| < \delta$ for sufficiently small δ. The coefficients B_n are bounded operators given by the formula

$$B_n = \frac{1}{2\pi i} \int_\gamma (\lambda - \lambda_0)^{-n-1} R(\lambda, A) d\lambda, \quad n \in \mathbb{Z} \tag{25}$$

where γ is a positively oriented simple curve surrounding λ_0 in $\rho(A)$. Application of the Cauchy integral formula gives

$$B_{-n} B_{-k} = B_{-n-k+1}, \quad n, k \in \mathbb{N} \tag{26}$$

The coefficient $P = B_{-1}$ is called the *residue* of A. If there exists k such that $B_{-k} \neq 0$ while B_{-n}, $n > k$, then λ_0 is called the *pole* of $R(\lambda, A)$ *of order* k. We have

$$B_k = \lim_{\lambda \to \lambda_0} (\lambda - \lambda_0)^k R(\lambda, A).$$

The following properties can be found in e.g. [32, 47].

Theorem 4 *1. The operator B_{-1} is a projection on X with $Im\ B_{-1}$ and $Im\ (I - B_{-1})$ closed.*

2. The restriction of A to $Im\ B_{-1}$ is bounded and has spectrum $\{\lambda_0\}$.

3. If $dim\ Im\ B_{-1} < \infty$, then λ_0 is a pole of $R(\lambda, A)$.

4. If λ_0 is a pole of $R(\lambda, A)$ of order k, then it is an eigenvalue of A and, for $j \geq 0$,

$$\begin{aligned} Im\ B_{-1} &= Ker\ (\lambda_0 I - A)^k = Ker\ (\lambda_0 I - A)^{k+j}, \\ Im\ (I - B_{-1}) &= Im\ (\lambda_0 I - A)^k = Im\ (\lambda_0 I - A)^{k+j}, \end{aligned} \tag{27}$$

and

$$X = Ker\ (\lambda_0 I - A)^k \bigoplus Im\ (\lambda_0 I - A)^k.$$

Let us prove the first part of (4), which frequently occurs in applications. Multiplying (24) by $(\lambda I - A)$ we obtain

$$\begin{aligned} I &= ((\lambda-\lambda_0)I + (\lambda_0 I - A)) \sum_{n=-\infty}^{\infty} (\lambda-\lambda_0)^n B_n \\ &= \sum_{n=-\infty}^{\infty} (\lambda-\lambda_0)^{n+1} B_n + \sum_{n=-\infty}^{\infty} (\lambda-\lambda_0)^n (\lambda_0 I - A) B_n \end{aligned}$$

so that

$$(\lambda_0 I - A)B_{-n} = -B_{-(n+1)}.$$

Since n is the order of the pole, $B_{-(n+1)} = 0$. On the other hand, since $B_{-n} \neq 0$, there is f such that $x = B_{-n}f \neq 0$ is an eigenvector corresponding to λ_0. $\square$

We define

$$Ker_\infty(\lambda_0 I - A) = \bigcup_{k\geq 0} Ker\ (\lambda_0 I - A)^k;$$

Ker_∞ is called the generalized eigenspace of A corresponding to the eigenvalue λ_0. $dim\ \ Im\ \ P$ is called the ***algebraic multiplicity*** of λ_0, denoted m_a, while $m_g = dim\ Ker\ (\lambda_0 I - A)$ is called the ***geometric multiplicity***. If $m_a = 1$, then λ_0 is called an ***algebraically simple*** pole. If k is the order of the pole ($k = \infty$ if λ_0 is an essential singularity), then

$$m_g + k - 1 \leq m_a \leq m_g k$$

$(0 \cdot \infty := \infty)$. Thus, $m_a < \infty$ if and only if λ_0 is a pole with $m_g < \infty$.

If A is closed with $\rho(A) \neq \emptyset$, then λ_0 is an isolated point of $\sigma(A)$ if and only if $(\lambda - \lambda_0)^{-1}$ is isolated in $\sigma(R(\lambda, A))$ and the residues and orders of the respective poles coincide.

In particular, if A has compact resolvent, then $\sigma(A)$ consists only of poles of finite algebraic multiplicity.

2.2.2 Turning approximate eigenvalues into eigenvalues

There is a very useful construction extending a given Banach space, called an *ultrapower* of X ([1]) or F*-product* ([38]). Here we shall discuss it in a restricted setting. Let $l_\infty(X)$ (resp. $c_0(X)$) be the vector space of bounded (resp. converging to 0) sequences $(x_n)_{n\in\mathbb{N}} \subset X$. We denote

$$\hat{X} = l_\infty(X)/c_0(X)$$

with the classes of equivalence denoted by

$$\hat{x} = (x_n)_{n\in\mathbb{N}} + c_0(X).$$

The space $\hat{X}$ becomes a Banach space under the norm

$$\|\hat{x}\| = \limsup_{n\to\infty} \|x_n\|.$$

There is a natural embedding $X \ni x \to (x, x, \ldots) + c_0(X) \in \hat{X}$ so that X can be identified with a closed subspace of $\hat{X}$.

Bounded operators on X give rise to bounded operators on $\hat{X}$: for $A \in \mathcal{L}(X)$ and $\hat{x} = (x_n)_{n\in\mathbb{N}} + c_0(X)$ we have

$$\hat{A}\hat{x} = (Ax_1, \widehat{Ax_2}, \ldots)$$

and it can be proved that $\|A\| = \|\hat{A}\|$.

If $(x_n)_{n\in\mathbb{N}}$ is approximate eigenvector of A with approximate eigenvalue λ, then $\|Ax_n - \lambda x_n\| \to 0$ as $n \to \infty$. But this is the same as saying that $\hat{x} = (x_1, x_2, \ldots) + c_0(X)$ is an eigenvector of $\hat{A}$ with the same eigenvalue. Actually, even more is true.

Theorem 5 *[21, p.290] Let* $A \in \mathcal{L}(X)$. *Then*

1. $\sigma(A) = \sigma(\hat{A})$;

2. $\sigma_a(A) = \sigma_a(\hat{A}) = \sigma_p(\hat{A})$;

3. $\widehat{R(\lambda, A)} = R(\lambda, \hat{A})$ *for* $\lambda \in \rho(A) = \rho(\hat{A})$;

4. $\lambda_0 \in \sigma(A)$ *is a pole of* $R(\lambda, A)$ *of order* p *if and only if* $\lambda_0 \in \sigma(\hat{A})$ *is a pole of* $R(\lambda, \hat{A})$ *of order* p.

Unfortunately, for unbounded operators and semigroups the situation becomes more complicated and we shall return to this topic later.

2.2.3 Spectrum of compact and power compact operators

The main results, summarizing the spectral properties of compact and power compact operators, are given in the following theorem.

Theorem 6 *If* K *is compact (or power compact), then*

(i) The spectrum of K *is at most countable and contains* $\{0\}$ *if* $dimX = \infty$;

(ii) If $\sigma(K)$ *is infinite and* $\{\lambda_1, \lambda_2, \cdots\}$ *is any enumeration of it, then* $\lambda_n \to 0$ *as* $n \to \infty$.

(iii) Every non-zero point of spectrum is a pole of the resolvent and thus is an eigenvalue.

Proof. For compact operators this result is known as the Fisher-Riesz theory. Extension to power compact operators is possible due to the Spectral Mapping Theorem which gives $\sigma(K^n) = \{\lambda^n;\ \lambda \in \sigma(K)\}$ and proves the assertions about the spectrum. To prove (iii) we consider the restriction of K to the projection $X_\lambda = B_{-1}X$ associated with $0 \neq \lambda \in \sigma(K)$. X_λ is invariant with respect to K and thus $K|_{X_\lambda}$ is power compact with $\sigma(K|_{X_\lambda}) = \{\lambda\}$ by Theorem 4(2). If $dim X_\lambda = \infty$, then this contradicts point (i) of the present theorem. Thus $dim X_\lambda < \infty$ and the result follows by Theorem 4 (4). □

2.2.4 Essential spectrum

As we have seen above, it is important to separate 'good' points of spectrum from 'bad' ones. The concept of *essential spectrum* have been introduced with this idea in mind.

Definition 1 *The essential spectrum of A, denoted by $\sigma_e(A)$ is the set of $\lambda \in \sigma(A)$ which satisfy at least one of the following conditions*

(i) $Im\ (\lambda I - A)$ is not closed;

(ii) $dim\ K_\infty(\lambda I - A) = \infty$;

(iii) λ is an accumulation point of $\sigma(A)$.

Essential spectrum is closely related to the concept of Fredholm points of A. We say that λ is a Fredholm point of A, and write $\lambda \in \rho_\Phi(A)$, if $Ker\ (\lambda I - A)$ is finite dimensional and $Im\ (\lambda I - A)$ is closed of finite codimension. The Fredholm spectrum of A, denoted $\sigma_\Phi(A)$, is the set of $\lambda \in \mathbb{C}$ which are not Fredholm points of A. Clearly,

$$\sigma_\phi(A) \subset \sigma_e(A),$$

but, in general, these sets are different (e.g., there may exist non-isolated Fredholm points of A.

Remark 2 Several authors (see e.g. [38, 26]) define the essential spectrum as the Fredholm spectrum. It has the additional advantage that it coincides with the normal spectrum of the canonical image of A in the quotient space $\mathcal{L}(X)/\mathcal{K}(X)$, where $\mathcal{K}(X)$ is the ideal of compact operators in X. This allows to define the Fredholm norm of A as

$$\|A\|_\Phi = dist(A, \mathcal{K}(X)) = \inf\{\|A - K\|,\ K \in \mathcal{K}(X)\}$$

As we shall see later, for the purpose of these lectures, the difference between both definitions are not significant.

We mention that there are also other, non-equivalent, definitions of essential spectrum.

We have the following result [20, 21].

Theorem 7 *Suppose $\lambda_0 \in \sigma(A)$ and $dim\ Ker\ (\lambda_0 I - A) < +\infty$. Then $\lambda_0 \in \sigma(A) \setminus \sigma_e(A)$ if and only if $R(\lambda, A)$ is analytic in a neighbourhood of λ_0 and has a pole at λ_0.*

Without assumption that $dim\ Ker\ (\lambda_0 I - A) < +\infty$ we can prove only that if $\lambda_0 \in \sigma(A) \setminus \sigma_e(A)$, then λ_0 is a pole of $R(\lambda, A)$.

In particular, if λ_0 is a non-essential point of $\sigma(A)$, then $Im\ (\lambda_0 I - A)$ is of finite codimension (see (27)) and thus $\lambda_0 \in \rho_\Phi(A)$.

We note some properties of the spectrum, [5]:

(a) $\mathrm{int}\sigma \subset \sigma_e$;

(b) $\partial\sigma_e \subset \sigma_\Phi$.

We can use characterization (18) of the spectral radius of a bounded operator to define analogous concepts related to the essential and Fredholm spectra of A:

$$\begin{aligned} r_e(A) &= \sup_{\lambda\in\sigma_e(A)} |\lambda|, \\ r_\Phi(A) &= \sup_{\lambda\in\sigma_\Phi(A)} |\lambda|. \end{aligned} \tag{28}$$

Clearly, we have $r_\Phi(A) \leq r_e(A)$. On the other hand, since $\sigma_e(A)$ is a compact set (for A bounded), there is $\lambda \in \sigma_e(A)$ with $|\lambda| = r_e(A)$. Such λ is in $\partial\sigma_e(A)$, hence, by (b) above, it is in $\sigma_\Phi(A)$. Therefore $r_\Phi(A) \geq r_e(A)$ and

$$r_\Phi(A) = r_e(A). \tag{29}$$

Remark 3 Since $\|\cdot\|_\Phi$ is a norm and $\sigma_\Phi(A)$ is coincides with the spectrum of the canonical image $\tilde{A}$ of A in $\mathcal{L}(X)/\mathcal{K}(X)$, we have also

$$r_\Phi(A) = r(\tilde{A}) = \lim_{n\to\infty} \sqrt[n]{\|\tilde{A}^n\|_\Phi}.$$

Using the above discussion, the essential radius can be characterized as follows

> $r_e(A)$ is the smallest $r \in \mathbb{R}_+$ such that every $\lambda \in \sigma(A)$ satisfying $|\lambda| > r$ is an isolated pole of finite algebraic multiplicity. For any $r > r_e(A)$, the set $\{\lambda \in \sigma(A);\ |\lambda| \geq r\}$ is finite.

The last statement follows from the fact that the spectrum of a bounded operator is compact and any accumulation point of $\sigma(A)$ belongs to $\sigma_e(A)$.

3 Banach Lattices and Positive Operators

In many processes in the natural sciences only nonnegative solutions are meaningful. This is the case when the solution is a probability, a density function, the absolute temperature, and so on. Thus, mathematical models of such processes should have the property that nonnegative data yield nonnegative solutions. If we work in concrete spaces of functions, then the notion of positivity is natural: either pointwise for continuous functions or almost everywhere in the spaces of measurable functions. However, in a general setting we have to find an abstract notion generalizing the pointwise concepts of positivity.

3.1 Defining Order

In a given vector space X an order can be introduced either geometrically, by defining the so-called *positive cone* (in other words, what it means to be a *positive element* of X), or through the axiomatic definition. We follow the second approach.

Definition 2 *Let X be an arbitrary set. A partial order (or simply, an order) on X is a binary relation, denoted here by '$\geq$', which is reflexive, transitive, and antisymmetric, that is,*

(1) $x \geq x$ for each $x \in X$;

(2) $x \geq y$ and $y \geq x$ imply $x = y$ for any $x, y \in X$;

(3) $x \geq y$ and $y \geq z$ imply $x \geq z$ for any $x, y, z \in X$.

We need a number of related conventions and definitions. The notation $x \leq y$ means $y \geq x$. $x > y$ means $x \geq y$ and $x \neq y$. An *upper bound* for a set $S \subset X$ is an element $x \in X$ satisfying $x \geq y$ for all $y \in S$. An element $x \in S$ is said to be *maximal* if there is no $S \ni y \neq x$ for which $y \geq x$. A *lower bound* for S and a *minimal element* are defined analogously. A *greatest element* (respectively, a *least element*) of S is an $x \in S$ satisfying $x \geq y$ (respectively, $x \leq y$) for all $y \in S$.

We note here that in an ordered space in general there are elements that cannot be compared and hence the distinction between maximal and greatest elements is important. A maximal element is the 'largest' amongst all comparable elements in S, whereas a greatest element is the 'largest' amongst all elements in S. If a greatest (or least) element exists, it must be unique by axiom (2).

The *supremum* of a set is its least upper bound and the *infimum* is the greatest lower bound. The supremum and infimum of a set need not exist. It is worthwhile to emphasize that an element s, which is an upper bound of S, is a supremum of the set S if, for any upper bound y of S, we have $s \leq y$.

Let $x, y \in X$ and $x \leq y$. The *order interval* $[x, y]$ is defined by

$$[x, y] := \{z \in X;\ x \leq z \leq y\}.$$

For a two-point set $\{x, y\}$ we write $x \wedge y$ or $\inf\{x, y\}$ to denote its infimum and $x \vee y$ or $\sup\{x, y\}$ to denote supremum. We say that X is a *lattice* if every pair of elements (and so every finite collection of them) has both supremum and infimum. From now on, unless stated otherwise, any vector space X is real.

Definition 3 *An ordered vector space is a vector space X equipped with partial order which is compatible with its vector structure in the sense that*

(4) $x \geq y$ implies $x + z \geq y + z$ for all $x, y, z \in X$;

(5) $x \geq y$ implies $\alpha x \geq \alpha y$ for any $x, y \in X$ and $\alpha \geq 0$.

The set $X_+ = \{x \in X;\ x \geq 0\}$ is referred to as the positive cone of X.

If the ordered vector space X is also a lattice, then it is called a vector lattice or a Riesz space.

Typical examples of Riesz spaces are provided by *function spaces*. If X is a vector space of real-valued functions on a set Ω, then we can introduce a pointwise order in X by saying that $f \leq g$ in X if $f(x) \leq g(x)$ for any $x \in S$. Equipped with such an order, X becomes an ordered vector space. Let us define on $X \times X$ the operations $f \vee g$ and $f \wedge g$ by taking pointwise maxima and minima; that is, for any $f, g \in X$,

$$\begin{aligned}(f \vee g)(x) &:= \max\{f(x), g(x)\},\\ (f \wedge g)(x) &:= \min\{f(x), g(x)\}.\end{aligned}$$

We say that X is a *function space* if $f \vee g, f \wedge g \in X$, whenever $f, g \in X$. Clearly, a function space with pointwise ordering is a Riesz space. Examples of function spaces are offered by the spaces of all real functions $\mathbb{R}^{\Omega}$ or all real bounded functions $M(\Omega)$ on a set Ω, and by, defined earlier, spaces $C(\Omega)$, $C(\overline{\Omega})$, or l_p, $1 \leq p \leq \infty$.

If Ω is a measure space, then all above considerations are valid when the pointwise order is replaced by $f \leq g$ if $f(x) \leq g(x)$ almost everywhere. With this understanding, $L_0(\Omega)$ and $L_p(\Omega)$ spaces with $1 \leq p \leq \infty$ become function spaces and are thus Riesz spaces.

We only consider Archimedean spaces; that is, spaces havig the property that if $\inf_{n\in\mathbb{N}}\{n^{-1}x\} = 0$ holds for any $x \in X_+$.

The operations of taking supremum or infimum in a Riesz space have several useful properties which make them similar to the numerical case. In particular, we can define the positive and negative part of $x \in X$, and its absolute value, respectively, by

$$x_+ = \sup\{x, 0\}, \quad x_- = \sup\{-x, 0\}, \quad |x| = \sup\{x, -x\}.$$

The functions $(x, y) \to \sup\{x, y\}, (x, y) \to \inf\{x, y\}, x \to x_\pm$ and $x \to |x|$ are collectively referred to as the *lattice operations* of a Riesz space. They are related by

$$x = x_+ - x_-, \qquad |x| = x_+ + x_-. \tag{30}$$

The absolute value has a number of useful properties that are reminiscent of the properties of the scalar absolute value; that is, for example, $|x| = 0$ if and only if $x = 0$, $|\alpha x| = |\alpha||x|$ for any $x \in X$ and any scalar α.

For a subset S of a Riesz space we write

$$\begin{aligned} \sup\{x, S\} &= x \vee S := \{\sup\{x, s\};\ s \in S\}, \\ \inf\{x, S\} &= x \wedge S := \{\inf\{x, s\};\ s \in S\}. \end{aligned}$$

The following infinite distributive laws are used later.

Proposition 1 *[3, Theorem 1.5] and [34, Theorem 2.13.1] Let S be a nonempty subset of a Riesz space X. If $\sup S$ exists, then $\sup\{\inf\{x, S\}\}$ and $\sup\{\sup\{x, S\}\}$ exist for each $x \in X$ and*

$$\begin{aligned} \sup\{\inf\{x, S\}\} &= \inf\{x, \sup S\}, \\ \sup\{\sup\{x, S\}\} &= \sup\{x, \sup S\}. \end{aligned} \tag{31}$$

Similarly, if $\inf S$ exists, then $\inf\{\sup\{x, S\}\}$, $\inf\{\inf\{x, S\}\}$ exist for each $x \in X$ and

$$\begin{aligned} \inf\{\sup\{x, S\}\} &= \sup\{x, \inf S\}, \\ \inf\{\inf\{x, S\}\} &= \inf\{x, \inf S\}. \end{aligned} \tag{32}$$

The existence of suprema or infima of finite sets, ensured by the definition of a Riesz space, does not extend to infinite sets. This warrants introducing a more restrictive class of spaces.

Definition 4 *We say that a Riesz space X is Dedekind (or order) complete if every nonempty and bounded from above subset of X has a least upper bound in X. X is said to be σ-Dedekind or (σ-order) complete, if every bounded from above nonempty countable subset of X has a least upper bound X.*

Example 8 The space $C([0,1])$ is not σ-order complete (and thus also not order complete). To see this, consider the sequence of functions given by

$$f_n(x) = \begin{cases} 1 & \text{for} \quad 0 \leq x \leq \frac{1}{2} - \frac{1}{n}, \\ n\left(\frac{1}{2} - x\right) & \text{for} \quad \frac{1}{2} - \frac{1}{n} < x \leq \frac{1}{2}, \\ 0 & \text{for} \quad \frac{1}{2} < x < 1. \end{cases}$$

This is clearly an increasing sequence bounded from above by $g(x) \equiv 1$. However, it converges pointwise to a discontinuous function $f(x) = 1$ for $x \in [0, 1/2)$ and $f(x) = 0$ for $x \in [1/2, 0]$. In general, spaces $C(\Omega)$ are not σ-order complete unless Ω consists of isolated points. On the other hand, the spaces l_p, $1 \leq p \leq \infty$, are clearly order complete, as taking the coordinatewise suprema of sequences bounded from above by an l_p sequence produces a sequence which is in l_p.

The spaces $L_p(\Omega), p \in \{0\} \cup [1, \infty]$ are also order complete but the proof is much more delicate, see [9, Example 2.52].

3.2 Banach Lattices

As the next step, we investigate the relation between the lattice structure and the norm when X is both a normed and an ordered vector space.

Definition 5 *A norm on a vector lattice X is called a lattice norm if*

$$|x| \leq |y| \quad \text{implies} \quad \|x\| \leq \|y\|. \tag{33}$$

A Riesz space X complete under the lattice norm is called a Banach lattice.

Property (33) gives the important identity:

$$\|x\| = \||x|\|, \qquad x \in X. \tag{34}$$

If X is a normed lattice, then all lattice operations are uniformly continuous in the norm of X with respect to all variables involved.

Positive operators will be discussed in more detail below. However, we need some terminology related to operators at this instance. An operator A defined on X is said to be positive if $Ax \geq 0$ for $x \geq 0$. A positive operator A is said to be a *lattice homomorphism* if $A(x \vee y) = Ax \vee Ay$. It can be proved that this is equivalent to A preserving all other lattice operations (e.g. $|Ax| = |x|$, $(Ax)^+ = Ax^+$, etc). If A is a one-to-one lattice homomorphism, it is called a *lattice isomorphism* and if, additionally, A is an isometry, then it is called a *lattice isometry*.

Bounded positive functionals form a convex cone in X^* and thus define a natural ordering of X^*. It can be proved, [3, Theorem 12.1], that the normed dual of a normed Riesz space is a Banach lattice under this order. In addition, the evaluation map $X \to X^{**}$ is a lattice isometry so that X becomes a Riesz subspace of X^{**}.

3.2.1 Sublattices, ideals, bands, etc

A vector subspace X_0 of a vector lattice X, which is ordered by the order inherited from X, may fail to be a vector sublattice of X in the sense that X_0 may be not closed under lattice operations. For instance, the subspace

$$X_0 := \{f \in L_1(\mathbb{R});\ \int_{-\infty}^{\infty} f(t)dt = 0\}$$

does not contain any nontrivial nonnegative function, and thus it is not closed under the operations of taking $f_\pm$ or $|f|$.

Accordingly, we call X_0 a *vector sublattice* or a *Riesz subspace* if X_0 is closed under lattice operations.

A subset S of a vector lattice is called solid if for any $x, y \in X$ from $y \in S$ and $|x| \leq |y|$ it follows that $x \in S$. A solid linear subspace is called *ideal*; ideals are automatically Riesz subspaces. A *band* in X is an ideal that contains suprema of all its subsets. Any subset $S \subset X$ uniquely determines the smallest (in the inclusion sense) Riesz subspace (respectively, ideal, band) in X containing S, called the *Riesz subspace (respectively, ideal, band) generated by* S.

Example 9 Closed ideals can be used to construct new useful Banach lattices by taking quotients. Let X be a Banach lattice and E a closed ideal in X. Then the quotient space X/E is a Banach space. We can define an order in X/E through the following relation. For $X/E \ni \tilde{x}, \tilde{y}$ we say that $\tilde{x} \leq \tilde{y}$ if there are $x_1 \in \tilde{x}$ and $y_1 \in \tilde{y}$ such that $x_1 \leq y_1$ in X and one can prove that X/E with this order and the canonical quotient norm is a Banach lattice.

Consider, in particular, the F-product discussed in Subsection 2.2.2. If X is a Banach lattice, then the absolute value on $l_\infty(X)$ is given by

$$|(x_n)_{n\in\mathbb{N}}| = (|x_n|)_{n\in\mathbb{N}}.$$

Since $c_0(X)$ is a closed ideal in $l_\infty(X)$, then $\hat{X}$ is a lattice with the canonical injection becoming a lattice homomorphism.

Example 10 Closed ideals in $X = L_p(\Omega)$, which are not equal to X, are precisely the sets of the form

$$I = \{f \in X;\ \exists_{\Omega' \subset \Omega, \mu(\Omega')}\ f|_{\Omega'} = 0\ a.e.\}.$$

I clearly is a closed ideal.On the other hand, let $f(x) > 0$ a.e. on Ω and $g \in X_+$. Consider sets $\Omega_n = \{x \in \Omega;\ f(x) \geq 1/n\}$ and define $g_n(x) = 0$ on Ω_n, $g_n = \min g, n$ otherwise. We have $0 \leq g_n \leq n^2 f$, hence $g_n \in X$ and, clearly $g_n \to g$ in $L_p(\Omega)$ since $\mu(\Omega_n) \to 0$ as $n \to \infty$.

In the theory developed later a particularly important part is played by ideals generated by a single point, say $\{x\}$. Such an ideal, called the *principal ideal* generated by x, is given by

$$E_x = \{y \in X;\ \text{there exists } \lambda \geq 0 \text{ such that } |y| \leq \lambda |x|\}.$$

If for some vector $e \in X$ we have $E_e = X$, then e is called an *order unit*.

A *principal band* generated by $x \in X$ is given by

$$B_x = \{y \in X;\ \sup_{n \in \mathbb{N}}\{|y| \wedge n|x|\} = |y|\}.$$

An element $e \in X$ is said to be a *weak unit* if $B_e = X$. It follows that, in a vector lattice, $e > 0$ is a weak unit if and only if, for any $x \in X$, $|x| \wedge e = 0$ implies $x = 0$. Every order unit is a weak unit. If $X = C(\Omega)$, where Ω is compact, then any strictly positive function is an order unit. On the other hand, L_p and l_1 spaces, $1 \leq p < +\infty$, will not typically have order units (L_p include functions that could be unbounded, for l_p one can always find a sequence converging to 0 at a slower rate than a given one). However, any strictly positive a.e. L_p function is a weak order unit.

An intermediate notion between order unit and weak order unit is played by *quasi-interior points*. We say that $0 \neq u \in X_+$ is a quasi-interior point of X if $\overline{E_u} = X$. We have

Lemma 2 *[1, Lemma 4.15] For $0 \neq u \in X_+$ the following are equivalent.*

(a) u is a quasi-interior point of X;

(b) For each $x \in X_+$ *we have* $\lim_{n\to\infty} \|x \wedge nu - x\| = 0$*;*

(b) If $0 < x^* \in X_+^*$*, then* $<x^*, u> > 0$*.*

The name ’quasi-interior point’ comes from the fact that a unit is an interior point of a positive cone. Thus, we have

$$\text{order unit} \Rightarrow \text{quasi} - \text{interior point} \Rightarrow \text{weak order unit}$$

and, in general, the implications cannot be reversed. The importance of quasi interior points will become more clear when we will discuss AM-spaces.

3.2.2 AM- and AL-spaces

Two important classes of Banach lattices that play a significant role later are provided by the AL- and AM- spaces.

Definition 6 *We say that a Banach lattice* X *is*

(i) an AL-space if $\|x + y\| = \|x\| + \|y\|$ *for all* $x, y \in X_+$*,*

(ii) an AM-space if $\|x \vee y\| = \max\{\|x\|, \|y\|\}$ *for all* $x, y \in X_+$*.*

Example 11 Standard examples of AM-spaces are offered by the spaces $C(\overline{\Omega})$, where $\overline{\Omega}$ is either a bounded subset of $\mathbb{R}^n$, or in general, a compact topological space. Also the space $L_\infty(\Omega)$ is an AM-space. On the other hand, most known examples of AL-spaces are the spaces $L_1(\Omega)$. We observe later that these examples exhaust all (up to a lattice isometry) cases of AM- and AL-spaces. However, particular representations of these spaces can be very different.

It can be proved, [3, Theorem 12.22] and [1, Theorem 3.3], that a Banach lattice X is an AL-space (respectively, AM-space) if and only if its dual X^* is an AM-space (respectively, AL-space). Moreover, if X is an AL-space, then X^* is a Dedekind complete AM-space with unit e^* defined by

$$X^* \ni e^*(x) = \|x_+\| - \|x_-\|$$

for $x \in X$ (thus e^* coincides with the norm of x on the positive cone). Moreover, if X is an AM-space with unit e, then X^{**} is also an AM-space with unit e.

Any AM-space X with unit e can be equivalently normed by

$$\|x\|_\infty = \inf\{\lambda > 0;\ |x| \leq \lambda e\}$$

(see, e.g., [3, p. 188]). In this norm the unit ball of X coincides with the order interval $[-e, e]$. On the other hand, any Banach lattice contains AM-spaces with unit. Precisely speaking, [3, Theorem 12.20], the principal ideal generated by any element $u \in X$ with the norm

$$\|x\|_\infty = \inf\{\lambda > 0;\ |x| \leq \lambda |u|\}, \tag{35}$$

becomes an AM-space with unit $|u|$, whose closed unit ball coincides with the order interval $[-|u|, |u|]$.

The following results give the full characterisation of AL- and AM- spaces.

Theorem 12 *[3, Theorem 12.26] A Banach lattice is an AL-space if and only if it is lattice isometric to an $L_1(\Omega)$ space.*

Theorem 13 *[3, Theorem 12.28] A Banach lattice X is an AM-space with unit if and only if it is lattice isometric to some $C(\Omega)$ for a unique (up to a homeomorphism) compact Hausdorff space Ω. In particular, X is an AM-space if and only if it is lattice isometric to a closed vector sublattice of a $C(\Omega)$ space.*

We provide a brief information about the main parts of the proof of the latter theorem.

Proof. The compact space Ω turns out to be

$$\begin{aligned}\Omega &= \{x^* \in B^*_{1,+};\ x^* \text{ extr. p. of } B^*_1 \text{ with } \|x^*\| = \|x^*(e)\| = 1\} \\ &= \{x^* \in B^*_{1,+};\ x^* \text{ lat. hom. with } \|x^*\| = \|x^*(e)\| = 1\}.\end{aligned}$$

Here, B^*_1 is the unit ball in the dual space and extreme points of a set are understood as points which do not belong to any proper segment with endpoints in this set. Establishing this equality is a difficult part of the proof. It follows that Ω is non-empty (by Krein-Milman theorem) and weakly* compact. Thus, Ω equipped with the weak* topology will be our compact topological space. For $x \in X$ we define the mapping

$$(Tx)(x^*) = <x^*, x>, \quad x^* \in \Omega.$$

It can be proved that T is a norm preserving lattice isomorphism from X into $C(\Omega)$. Since $(Te)(x^*) = x^*(e) = 1$ for all $x^* \in \Omega$, $T(E)$ is closed and separates points of Ω, it follows from the Stone-Weierstrass theorem that $T(E) = C(\Omega)$. □

Using the last theorem, we see that each Banach lattice 'locally' is a lattice isomorphic to $C(\Omega)$. More precisely, given $0 < u \in X$ we take the principal ideal E_u which can be converted into an AM-space normed by (35). This norm is not equivalent to the norm in X. However, if we have a bounded operator defined on X, then the transferred operator on $C(\Omega)$ will be again a positive everywhere defined operator and thus bounded (by Theorem 15). Conversely, operators specific to $C(\Omega)$, such as multiplication or composition operators, can be transferred to bounded operators on X_u. If u is a quasi-internal point and the given operator happen to be bounded in the original norm, then it can be extended by density to the whole Banach lattice. We shall use this construction later to define the modulus of an element of a complex Banach lattice and the signum operator.

3.3 Positive Operators

Definition 7 *A linear operator A from a Banach lattice X into a Banach lattice Y is called positive, denoted by $A \geq 0$, if $Ax \geq 0$ for any $x \geq 0$.*

An operator A is positive if and only if $|Ax| \leq A|x|$. This follows easily from $-|x| \leq x \leq |x|$ so, if A is positive, then $-A|x| \leq Ax \leq A|x|$. Conversely, taking $x \geq 0$, we obtain $0 \leq |Ax| \leq A|x| = Ax$.

Positive operators are fully determined by their behaviour on the positive cone. Precisely speaking, we have the following theorem (eg. [9, Theorem 2.64]).

Theorem 14 *If $A : X_+ \to Y_+$ is additive, then A extends uniquely to a positive linear operator from X to Y. Keeping the notation A for the extension, we have, for each $x \in X$,*

$$Ax = Ax_+ - Ax_-. \tag{36}$$

Another frequently used property of positive operators is given in the following theorem.

Theorem 15 *If A is an everywhere defined positive operator from a Banach lattice to a normed Riesz space, then A is bounded.*

Proof. If A were not bounded, then we would have a sequence $(x_n)_{n\in\mathbb{N}}$ satisfying $\|x_n\| = 1$ and $\|Ax_n\| \geq n^3$, $n \in \mathbb{N}$. Because X is a Banach space, $x := \sum_{n=1}^{\infty} n^{-2}|x_n| \in X$. Because $0 \leq |x_n|/n^2 \leq x$, we have $\infty > \|Ax\| \geq \|A(|x_n|/n^2)\| \geq \|A(x_n/n^2)\| \geq n$ for all n, which is a contradiction. □

A striking consequence of this fact is that all norms, under which X is a Banach lattice, are equivalent as the identity map must be continuously invertible, [3, Corollary 12.4].

Example 16 The assumption that X in Theorem 15 is a complete space is essential. Indeed, let X be a space of all real sequences which have only a finite number of nonzero terms. It is a normed Riesz space under the norm $\|\mathbf{x}\| = \sup_n\{|x_n|\}$, where $\mathbf{x} = (x_n)_{n\in\mathbb{N}}$. Consider the functional

$$f(\mathbf{x}) = \sum_{n=1}^{\infty} x_n.$$

It is a positive everywhere defined linear functional. However, taking the sequence of elements $\mathbf{x}_n = (1, 1, \ldots, 1, 0, 0, \ldots)$, where 0 appears starting from the $n + 1$st place, we see that $\|\mathbf{x}_n\| = 1$ and $f(\mathbf{x}_n) = n$ for each $n \in \mathbb{N}$ so that f is not bounded.

The set of all positive operators from a Banach lattice X to another Banach lattice Y is a convex cone in the space $\mathcal{L}(X, Y)$, thus it generates a natural order: $A \leq B$ whenever $Ax \leq Bx$ for all $x \in X_+$. This cone, however, in general does not generate $\mathcal{L}(X, Y)$ (e.g., [3, Example 1.11]). The norm of a positive operator can be evaluated by

$$\|A\| = \sup_{x\geq 0,\, \|x\|\leq 1} \|Ax\|.$$

As a consequence, we note that if $0 \leq A \leq B$, then $\|A\| \leq \|B\|$. Moreover, it is worthwhile to emphasize that if there exists K such that $\|Ax\| \leq K\|x\|$ for $x \geq 0$, then this inequality holds for any $x \in X$.

Irreducible operators. An important class of positive operators are *irreducible operators*. We say that an operator A on a Banach lattice X is irreducible if $\{0\}$ and X are the only invariant ideals under A. We say that A is *strongly irreducible* if Au is quasi-interior point for any $u > 0$. Strongly irreducible operators are irreducible. Indeed, any closed ideal $E \neq \{0\}$ contains a positive point u so that $Au \in AE \subset E$ provided E is invariant. Since Au is quasi-interior, this implies $E = X$ We shall return to this concept in Subsection 6.2.1.

Example 17 An important role in the following considerations is played by the multiplication by the *signum operator*. In function spaces the definition is obvious: given $u \neq 0$ and $f \in C(\Omega)$, we define $S_u f = u|u|^{-1} f$. Clearly, in this setting S_u is a linear isometry satisfying $|S_u f| = |f|$; its inverse is $S_{\bar{u}}$, where $\bar{u}$ is the complex conjugate of u.

In general situation, we restrict our attention to u such that $|u|$ is quasi-interior point of X. In this case we define this operator on $E_{|u|}$ by passing to the representation $C(\Omega)$ and transferring back the signum operator defined above to X. We note that in this setting S_h is still invertible and has the same properties as in $C(\Omega)$. By $|S_u f| = |f|$ we can extend S_u by density to $X = \overline{E_{|u|}}$ preserving invertibility.

It is possible to extend this definition to the case when $|u|$ is no longer a quasi-interior point but it will not be needed in what follows (see e.g. [38, p. 245].

3.4 Relation Between Order and Norm

Existence of an order in some set X allows us to introduce in a natural way the notion of convergence. However, in general, sequences are not sufficient to properly describe all related phenomena and thus we have to resort to nets.

We say that an ordered set Δ is *directed* if any pair of elements has an upper bound. Then, by a *net* $(x_\alpha)_{\alpha\in\Delta}$ in a set X, we understand a function from the *index set* Δ into X.

By a *subnet* we understand a net $(y_\beta)_{\beta\in B}$ such that for any $\alpha \in \Delta$ there is $\beta \in B$ such that for each $B \ni \beta' \geq \beta$ there is $\alpha' \geq \alpha$ such that $y_{\beta'} = x_{\alpha'}$. A net $(x_\alpha)_{\alpha\in\Delta}$ in a normed space X converges to some point $x \in X$ if for any $\epsilon > 0$ there is $\alpha_0 \in \Delta$ such that for any $\alpha \geq \alpha_0$ we have $\|x_\alpha - x\| \leq \epsilon$. We write this as $x_\alpha \xrightarrow{n} x$ or explicitly $\lim_{\alpha\in\Delta} x_\alpha = x$ in norm.

A net $(x_\alpha)_{\alpha\in\Delta}$ in an ordered set X is said to be *decreasing* (in symbols $x_\alpha\downarrow$) if for any $\alpha_1, \alpha_2 \in \Delta$ with $\alpha_1 \geq \alpha_2$ we have $x_{\alpha_1} \leq x_{\alpha_2}$. The notation $x_\alpha \downarrow x$ means that $x_\alpha\downarrow$ and $\inf\{x_\alpha;\ \alpha \in \Delta\} = x$. Furthermore, we write $x_\alpha \downarrow\geq x$ if the net is decreasing and $x_\alpha \geq x$ for all $\alpha \in \Delta$.

Symbols $x_\alpha\uparrow$, $x_\alpha \uparrow\ x$, and $x_\alpha \uparrow\leq\ x$ have analogous meaning.

Using these definitions we can analyse convergence of increasing and decreasing nets, where the limit is, respectively, the supremum or infimum of the net. If $(x_\alpha)_{\alpha\in\Delta}$ is a net of arbitrary elements of X, then we say that it is *order convergent* to x if there are nets $(y_\beta)_{\beta\in B}$ and $(z_\gamma)_{\gamma\in\Gamma}$ such that $y_\beta\uparrow x$, $z_\gamma \downarrow x$ and such that for any $\beta \in B$ and $\gamma \in \Gamma$ there is $\alpha \in \Delta$ such that $y_\beta \leq x_\alpha \leq z_\gamma$. We write this as $x_\alpha \xrightarrow{o} x$. It can be proved, [1, p. 17], that we can take the sets B and Γ to be equal.

We note that a net in a partially ordered space can have at most one order limit. Furthermore, if either $x_\alpha\uparrow x$ or $x_\alpha \downarrow x$, then $x_\alpha \xrightarrow{o} x$. Conversely, if $x_\alpha\uparrow$ (resp., $x_\alpha \downarrow$) and $x_\alpha \xrightarrow{o} x$, then $x_\alpha\uparrow x$ (resp., $x_\alpha \downarrow x$). The proofs can be found in [9, Examples 2.71 and 2.72]. One of the basic results here is

Proposition 2 *Let X be a normed lattice. Then:*

(1) The positive cone X_+ is closed.

(2) If $X \ni x_\alpha\uparrow$ and $\lim_{\alpha\in\Delta} x_\alpha = x$ in the norm of X, then

$$x = \sup\{x_\alpha;\ \alpha \in \Delta\}.$$

(3) If $X \ni x_\alpha \downarrow$ *and* $\lim_{\alpha\in\Delta} x_\alpha = x$ *in the norm of* X*, then*

$$x = \inf\{x_\alpha;\ \alpha \in \Delta\}.$$

Proof. (1) Because $X_+ = \{x \in X;\ x_- = 0\}$ and lattice operation $X \ni x \to x_- \in X$ is continuous we see that X_+ is closed.

(2) For any fixed $\alpha \in \Delta$ we have

$$\lim_{\beta\in\Delta}(x_\beta - x_\alpha) = x - x_\alpha$$

in norm and $x_\beta - x_\alpha \in X_+$ for $\beta \geq \alpha$ so that $x - x_\alpha \in X_+$ for any $\alpha \in \Delta$ by (1). Thus x is an upper bound for the net $\{x_\alpha\}_{\alpha\in\Delta}$. On the other hand, if $x_\alpha \leq y$ for all α, then $0 \leq y - x_\alpha \xrightarrow{n} y - x$ so that, again by (1), we have $y \geq x$ and hence $x = \sup\{x_\alpha;\ \alpha \in \Delta\}$.

The proof of (3) is analogous. □

Example 18 The converse of Proposition 2(2) is false; that is, we may have $x_\alpha \uparrow x$ but $(x_\alpha)_{\alpha\in\Delta}$ does not converge in norm. Indeed, consider $\mathbf{x}_n = (1, 1, 1\ldots, 1, 0, 0, \ldots) \in l_\infty$, where 1 occupies only the n first positions. Clearly, $\sup_{n\in\mathbb{N}} \mathbf{x}_n = \mathbf{x} := (1, 1, \ldots, 1, \ldots)$ but $\|\mathbf{x}_n - \mathbf{x}\|_\infty = 1$.

This example justifies introducing a special class of Banach lattices.

Definition 8 *We say that a Banach lattice* X *has order continuous norm if for any net* $(x_\alpha)_{\alpha\in\Delta}$, $x_\alpha \downarrow 0$ *implies* $\|x_\alpha\| \downarrow 0$.

Before we give examples of Banach lattices with order continuous norm, we state and prove basic properties of them.

Theorem 19 *[3, Theorem 12.9] For a Banach lattice X, the statements below are equivalent.*

(1) X has order continuous norm;

(2) If $0 \leq x_n \uparrow \leq x$ holds in X, then $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence;

(3) X is σ-order complete and $x_n \downarrow 0$ implies $\|x_n\| \to 0$;

(4) X is an ideal in X^{**};

(5) For every $a, b \in X$, the order interval $\{x;\ a \leq x \leq y\}$ is weakly compact.

Moreover, every Banach lattice with order continuous norm is order complete.

Example 20 For $1 \leq p < \infty$, the Banach lattice $L_p(\Omega)$ has order continuous norm. Indeed, let $f_n \downarrow 0$ almost everywhere. Then $\|f_n\|^p = \int_\Omega f_n^p d\mu \to 0$ from the dominated convergence theorem and the statement follows from Theorem 19(3) as $L_p(\Omega)$ is σ-order complete by Example 8.

Incidentally, this gives an independent proof that $L_p(\Omega), 1 \leq p < \infty$ are order complete.

On the other hand, $L_\infty(\Omega)$ is order complete by Example 8 but its norm is not order continuous. To see this, consider the σ-algebra Σ of measurable subsets of Ω and let Δ be the subset of Σ containing the sets which differ from Ω by sets of positive measure, directed by the relation of inclusion. Finally, take the net $(\chi_\alpha)_{\alpha \in \Delta}$ of characteristic functions of sets from Δ. Then $\chi_\Omega - \chi_\alpha \downarrow 0$ but $\|\chi_\Omega - \chi_\alpha\| = 1$ for all $\alpha \in \Delta$.

The importance of Banach lattices with order continuous norm stems mainly from property 2 of Theorem 19 which states that increasing sequences dominated in the order sense must necessarily converge in norm. There is an important subset of this class of Banach lattices with a stronger property that increasing and norm bounded sequences are norm convergent.

Definition 9 *We say that a Banach lattice X is a KB-*space *(Kantorovič–Banach space) if every increasing norm bounded sequence of elements of X_+ converges in norm in X.*

Example 21 We observe that if $x_n \uparrow x$, then $\|x_n\| \leq \|x\|$ for all $n \in \mathbb{N}$ and thus any KB-space has order continuous norm by Theorem 19. Hence, spaces which do not have order continuous norm cannot be KB-spaces. This rules out the spaces of continuous functions, l_∞ and $L_\infty(\Omega)$.

To see that the KB-class is indeed strictly smaller, let us consider the space c_0. First we prove that it has order continuous norm. It is clearly σ-order complete. Let the sequence $(\mathbf{x}_n)_{n\in\mathbb{N}}$, given by $\mathbf{x}_n = (x_k^n)_{k\in\mathbb{N}}$, satisfy $\mathbf{x}_n \downarrow 0$. For a given $\epsilon > 0$, we find k_0 such that $|x_k^1| < \epsilon$ for all $k \geq k_0$. Because $(\mathbf{x}_n)_{n\in\mathbb{N}}$ is decreasing, we also have $|x_k^n| < \epsilon$ for all $k \geq k_0$ and $n \geq 1$. Then, we find n_0 such that $|x_k^n| < \epsilon$ for all $n \geq n_0$ and $1 \leq k \leq k_0$ and combining these estimates we see that $\|\mathbf{x}_n\| < \epsilon$ for all $n \geq n_0$ so $\|\mathbf{x}_n\| \to 0$.

On the other hand, let us again take the sequence $\mathbf{x}_n = (1, 1, \ldots, 1, 0, 0, \ldots)$ where 1 occupies n first positions. It is clearly norm bounded and increasing, but it does not converge in norm to any element of c_0. Hence, c_0 has not got an order continuous norm.

The next theorems characterize the KB-spaces which appear in applications.

Theorem 22 *[9, Theorem 2.82] Assume that X is a weakly sequentially complete Banach lattice. If $(x_n)_{n\in\mathbb{N}}$ is increasing and $(\|x_n\|)_{n\in\mathbb{N}}$ is bounded, then there is $x \in X$ such that $\lim_{n\to\infty} x_n = x$ in X.*

The next result shows the same property for AL-spaces.

Theorem 23 *Any AL-space is a KB-space.*

Proof. If $(x_n)_{n\in\mathbb{N}}$ is an increasing and norm bounded sequence, then for $0 \le x_n \le x_m$, we have

$$\|x_m\| = \|x_m - x_n\| + \|x_n\|$$

as $x_m - x_n \ge 0$ so that

$$\|x_m - x_n\| = \|x_m\| - \|x_n\| = |\|x_m\| - \|x_n\|| \,.$$

By assumption, $(\|x_n\|)_{n\in\mathbb{N}}$ is monotonic and bounded, and hence convergent, we see that $(x_n)_{n\in\mathbb{N}}$ is Cauchy. □

3.5 Complexification

Our main interest is in real operators on real Banach spaces. However, in some cases, especially when we want to use spectral theory, we need to move the problem to a complex space. This is done by the procedure called *complexification*.

Definition 10 *Let X be a real vector lattice. The complexification X_C of X is the set of pairs $(x, y) \in X \times X$ where, following the scalar convention, we write $(x, y) = x + iy$. Vector operations are defined as in scalar case*

$$\begin{aligned} x_1 + iy_1 + x_2 + iy_2 &= x_1 + x_2 + i(y_1 + y_2), \\ (\alpha + i\beta)(x + iy) &= \alpha x - \beta y + i(\beta x + \alpha y). \end{aligned}$$

The partial order in X_C is defined by

$$x_0 + iy_0 \le x_1 + iy_1 \quad \text{if and only if} \quad x_0 \le x_1 \text{ and } y_0 = y_1. \tag{37}$$

The operations of the complex adjoint, real part, and imaginary part of $z = x+iy$ *are defined through:*

$$\begin{aligned}\bar{z} &= \overline{x+iy} = x - iy,\\ \Re z &= \frac{z+\bar{z}}{2} = x,\\ \Im z &= \frac{z-\bar{z}}{2i} = y.\end{aligned}$$

Remark 4 Note, that from the definition, it follows that $x \geq 0$ in X_C is equivalent to $x \in X$ and $x \geq 0$ in X. In particular, X_C with partial order (37) is not a lattice.

It is a more complicated task to introduce a norm on X_C because standard product norms, in general, fail to preserve the homogeneity of the norm.

First we introduce the modulus on X_C. In the scalar case we obviously have

$$\sup_{\theta\in[0,2\pi]} (\alpha\cos\theta + \beta\sin\theta) = |\alpha + i\beta|. \tag{38}$$

Mimicking this, for $x + iy \in X_C$ we define

$$|x+iy| = \sup_{\theta\in[0,2\pi]} (x\cos\theta + y\sin\theta).$$

It can be proved that this element exists. This follows because elements over which we take the supremum belong to the principal ideal generated by $|x| + |y|$ and, as we noted when discussing AM-spaces, such an ideal is an AM-space with unit $|x|+|y|$ and thus it is lattice isometric to some $C(\Omega)$. For $C(\Omega)$ the existence of $|x + iy|$ is proved pointwise by the argument leading to (38).

Such a defined modulus has all standard properties of the scalar complex modulus, [2, Problem 3.2.2]: for any $z, z_1, z_2 \in X_C$ and $\lambda \in C$,

(a) $|z| \geq 0$ and $|z| = 0$ if and only if $z = 0$,

(b) $|\lambda z| = |\lambda||z|$,

(c) $|z_1 + z_2| \leq |z_1| + |z_2|$ (triangle inequality),

and thus one can define a norm on the complexification X_C by

$$\|z\|_c = \|x + iy\|_c = \||x + iy|\|. \tag{39}$$

As the norm $\| \cdot \|$ is a lattice norm, we have $\|z_1\|_c \leq \|z_2\|_c$, whenever $|z_1| \leq |z_2|$, and $\| \cdot \|_c$ becomes a lattice norm on X_C.

Definition 11 *A complex Banach lattice is an ordered complex Banach space X_C that arises as the complexification of a real Banach lattice X, according to Definition 10, equipped with the norm (39).*

We extend A to X_C according to the formula

$$A_C(x + iy) = Ax + iAy,$$

and observe that if A is a positive operator between real Banach lattices X and Y then, for $z = x + iy \in X_C$, we have

$$(Ax)\cos\theta + (Ay)\sin\theta = A(x\cos\theta + y\sin\theta) \leq A|z|.$$

therefore $|A_C z| \leq A|z|$. Hence for positive operators

$$\|A_C\|_c = \|A\|. \tag{40}$$

There are examples, where $\|A\| < \|A_C\|_c$.

Note that the standard $L_p(\Omega)$ and $C(\Omega)$ norms are of the type (39). These spaces have a nice property of preserving the operator norm even for operators which are not necessarily positive, see [9, p. 63].

Example 24 Any positive linear operator A on X_C is a real operator; that is, $A : X \to X$. In fact, let $X_C \ni x = x_+ - x_-$. By definition, $Ax_+ \geq 0$ and $Ax_- \geq 0$ so $Ax_+, Ax_- \in X$ and thus $Ax = Ax_+ - Ax_- \in X$.

Remark 5 If for a linear operator A we prove that it generates a semigroup of say, contractions, in X, then this semigroup will be also a semigroup of contractions on X_C, hence, in particular, A is a dissipative operator in the complex setting. Due to this observation we confine ourselves to real operators in real spaces.

3.6 Series of Positive Elements in Banach Lattices

In this subsection we discuss two results which are series counterparts of the dominated and monotone convergence theorems in Banach lattices.

Theorem 25 *Let $(x_n(t))_{n\in\mathbb{N}}$ be family of nonnegative sequences in a Banach lattice X, parameterized by a parameter $t \in T \subset \mathbb{R}$, and let $t_0 \in \overline{T}$.*

(i) If for each $n \in \mathbb{N}$ the function $t \to x_n(t)$ is non-decreasing and $\lim\limits_{t\nearrow t_0} x_n(t) = x_n$ in norm, then

$$\lim_{t\nearrow t_0} \sum_{n=0}^{\infty} x_n(t) = \sum_{n=0}^{\infty} x_n, \tag{41}$$

irrespective of whether the right hand side exists in X or $\|\sum\limits_{n=0}^{\infty} x_n\| := \sup\{\|\sum\limits_{n=0}^{N} x_n\|;\ N \in \mathbb{N}\} = \infty$. In the latter case the equality should be understood as the norms of both sides being infinite.

(ii) If $\lim\limits_{t\to t_0} x_n(t) = x_n$ in norm for each $n \in \mathbb{N}$ and there exists $(a_n)_{n\in\mathbb{N}}$ such that $x_n(t) \leq a_n$ for any $t \in T, n \in \mathbb{N}$ with $\sum\limits_{n=0}^{\infty} \|a_n\| < \infty$, then (41) *holds as well.*

Remark 6 Note that if X is a KB-space, then $\lim_{t\nearrow t_0}\sum\limits_{n=0}^{\infty} x_n(t) \in X$ implies convergence of $\sum_{n=0}^{\infty} x_n$. In fact, since $x_n \geq 0$ (by closedness of the positive cone), $N \to \sum_{n=0}^{N} x_n$ is non-decreasing, and hence either $\sum_{n=0}^{\infty} x_n \in X$, or $\|\sum_{n=0}^{\infty} x_n\| = \infty$, and in the latter case we have $\left\|\lim_{t\nearrow t_0}\sum\limits_{n=0}^{\infty} x_n(t)\right\| = \infty$.

3.7 Spectral Radius of Positive Operators

Let $A \in \mathcal{L}(X)$. First we note that that the peripheral spectrum $\sigma_{per,r(A)}$, see (22), is non-empty. Also, $r(A) \in \{|\lambda|;\ \lambda \in \sigma(A)\}$. This follows from the compactness of $\sigma(A)$.

As a more serious application of the theory of Banach lattices, here we prove that if A is a positive operator, then its spectral radius is an element of the spectrum of A; that is, $r(A) \in \sigma(A)$. This, and related, results are usually referred to as the Frobenius-Perron theorem, after the authors of the matrix versions of them.

First we note that we can carry the considerations in the complexification of X, if necessary. Since all operators are positive, the operator norms in the real lattice and its complexification are equal, see (40) and we shall not distinguish them in the proofs.

Theorem 26 *Let $r(A)$ be the spectral radius of a positive operator A on a Banach lattice X. Then $r(A) \in \sigma(A)$.*

Proof. Let $\lambda_n = r(A) + 1/n$, then $\lambda_n \in \rho(A)$ for any n. Since $\lambda_n \to r(A)$. To show that $r(A) \in \sigma(A)$, it suffices, by Theorem 3, to show $\lim_{n\to\infty} \|R(\lambda_n, A)\| = \infty$.

Since the peripheral spectrum is non-empty, let $\alpha \in \sigma(A)$ with $|\alpha| = r(A)$ and define $\mu_n = \alpha\lambda_n/|\alpha|$. We have $\mu_n \in \rho(A)$ and $\mu_n \to \alpha$ so that, invoking

Theorem 3 again, $\lim_{n\to\infty} \|R(\mu_n, A)\| = \infty$. Next, for each n we pick a unit vector z_n satisfying

$$\|R(\mu_n, A)z_n\| \geq \frac{1}{2}\|R(\mu_n, A)\|$$

Using the series representation of the resolvent (17) we easily infer

$$|R(\lambda, A)z| \leq R(|\lambda|, A)|z|$$

so that $|R(\mu_n, A)z_n| \leq R(\lambda_n, A)|z_n|$ and consequently

$$\|R(\lambda_n, A)\| \geq \|R(\lambda_n, A)|z_n|\| \geq \|R(\mu_n, A)z_n\| \geq \frac{1}{2}\|R(\mu_n, A)\|$$

which proves the thesis. □

Theorem 27 *If $A : X \to X$ is a compact positive operator on a Banach lattice X with $r(A) > 0$, then $r(A)$ is an eigenvalue with positive eigenvector.*

Proof. Since $r(A) > 0$, by Theorems 26 and 6 it is an eigenvalue. As in the proof of the previous theorem, we put $\lambda_n = r(A) + 1/n$ so that $\lambda_n \downarrow r(A)$ and $\|R(\lambda_n, A)\| \to \infty$ as $n \to \infty$. Furthermore, for each n there is z_n with $\|z_n\| = 1$ satisfying

$$\|R(\lambda_n, A)z_n\| \geq \frac{1}{2}\|R(\lambda_n, A)\|.$$

We define

$$x_n = \frac{R(\lambda_n, A)z_n}{\|R(\lambda_n, A)z_n\|}$$

and note that x_n is a positive unit vector.

From

$$\begin{aligned} Ax_n - r(A)x_n &= (\lambda_n - r(A))x_n + Ax_n - \lambda_n x_n \\ &= \frac{x_n}{n} - \frac{z_n}{\|R(\lambda_n, A)z_n\|} \end{aligned}$$

we obtain

$$\|Ax_n - r(A)x_n\| \to 0, \quad n \to \infty.$$

Since A is compact, the sequence $(Ax_n)_{n\in\mathbb{N}}$ has a convergent subsequence which we denote by $(Ax_n)_{n\in\mathbb{N}}$ again. Using $r(A) > 0$ and $\|x_n\| = 1$, we find from the above that x_n converges to a positive (unit) vector x. This vector satisfies $Ax = r(A)x$. □

Corollary 1 *The thesis of Theorem 27 remains valid if the positive operator A is only power compact.*

Proof. If $r = r(A) > 0$ and A is power compact, then from the Spectral Mapping Theorem we have $A^k x = r^k x$ for some $x > 0$. Putting $y = \sum_{i=0}^{k-1} r^i A^{k-1-i} x$ we find that $y > 0$ (from positivity of A,x and r) and

$$Ay - ry = A^k x - r^k x = 0.$$

Remark 7 The assumption $r(A) > 0$ is crucial in infinite dimensional case (in finite dimension convergence of a subsequence of $(x_n)_{n\in\mathbb{N}}$ is obvious. Possibly the best result ensuring this was given by de Pagter, [42] and [1, p.359]. It reads that an irreducible power compact positive operator has a positive spectral radius.

4 First semigroups

The semigroup theory is concerned with methods of finding solutions of the Cauchy problem.

Definition 12 *Given a complex Banach space and a linear operator $\mathcal{A}$ with $D(\mathcal{A})$, $Im\mathcal{A} \subset X$ and given $u_0 \in X$, find a function $u(t) = u(t, u_0)$ such that*

1. $u \in C^0([0,\infty)) \cap C^1((0,\infty))$,

2. *for each* $t > 0$, $u(t) \in D(\mathcal{A})$ *and*

$$u'(t) = \mathcal{A}u(t), \quad t > 0, \tag{42}$$

3.

$$\lim_{t \to 0^+} u(t) = u_0 \tag{43}$$

in the norm of X.

A function satisfying all conditions above is called the classical (or strict) solution of (42), (43).

If the solution to (42), (43) is unique, then we can introduce a family of operators $(G(t))_{t\geq 0}$ such that $u(t, u_0) = G(t)u_0$. Ideally, $G(t)$ should be defined on the whole space for each $t > 0$, and the function $t \to G(t)u_0$ should be continuous for each $u_0 \in X$, leading to well-posedness of (42), (43). Moreover, uniqueness and linearity of $\mathcal{A}$ imply that $G(t)$ are linear operators. A fine-tuning of these requirements leads to the following definition.

Definition 13 *A family* $(G(t))_{t\geq 0}$ *of bounded linear operators on* X *is called a* C_0*-semigroup, or a strongly continuous semigroup, if*

(i) $G(0) = I$;

(ii) $G(t + s) = G(t)G(s)$ *for all* $t, s \geq 0$;

(iii) $\lim_{t\to 0^+} G(t)x = x$ *for any* $x \in X$.

A linear operator A *is called the (infinitesimal) generator of* $(G(t))_{t\geq 0}$ *if*

$$Ax = \lim_{h \to 0^+} \frac{G(h)x - x}{h}, \tag{44}$$

with $D(A)$ *defined as the set of all* $x \in X$ *for which this limit exists. Typically the semigroup generated by* A *is denoted by* $(G_A(t))_{t\geq 0}$.

If $(G(t))_{t\geq 0}$ is a C_0-semigroup, then the local boundedness and (ii) lead to the existence of constants $M > 0$ and ω such that for all $t \geq 0$

$$\|G(t)\|_X \leq Me^{\omega t}. \tag{45}$$

We say that $A \in \mathcal{G}(M, \omega)$ if it generates $(G(t))_{t\geq 0}$ satisfying (45). The *type*, or *uniform growth bound,* $\omega_0(G)$ of $(G(t))_{t\geq 0}$ is defined as

$$\omega_0(G) = \inf\{\omega;\ \text{there is } M \text{ such that (45) holds}\}. \tag{46}$$

From (44) and the condition (iii) of Definition 13 we see that if A is the generator of $(G(t))_{t\geq 0}$, then for $x \in D(A)$ the function $t \to G(t)x$ is a classical solution of the following Cauchy problem,

$$\begin{aligned} \partial_t u(t) &= A(u(t)), \quad t > 0, &(47)\\ \lim_{t\to 0^+} u(t) &= x. &(48) \end{aligned}$$

We note that ideally the generator A should coincide with $\mathcal{A}$ but in reality very often it is not so. In fact, a large part of the theory discussed here is concerned with finding a relation between $\mathcal{A}$ and its realisation A which generates a semigroup. Such problems are addressed later. However, for most of this section we are concerned with solvability of (47), (48); that is, with the case when $\mathcal{A}$ of (42) is the generator of a semigroup.

We noted above that for $x \in D(A)$ the function $u(t) = G(t)x$ is a classical solution to (47), (48). For $x \in X \setminus D(A)$, however, the function $u(t) = G(t)x$ is continuous but, in general, not differentiable, nor $D(A)$-valued, and, therefore, not a classical solution. Nevertheless, it follows that the integral $v(t) = \int_0^t u(s)ds \in D(A)$ and therefore it is a strict solution of the integrated version of (47), (48):

$$\begin{aligned} \partial_t v &= Av + x, \quad t > 0\\ v(0) &= 0, \end{aligned} \tag{49}$$

or equivalently,

$$u(t) = A\int_0^t u(s)ds + x. \tag{50}$$

We say that a function u satisfying (49) (or, equivalently, (50)) is a *mild solution* or *integral solution* of (47), (48).

Proposition 3 *Let $(G(t))_{t\geq 0}$ be the semigroup generated by $(A, D(A))$. Then $t \to G(t)x$, $x \in D(A)$, is the only solution of (47), (48) taking values in $D(A)$. Similarly, for $x \in X$, the function $t \to G(t)x$ is the only mild solution to (47), (48).*

Thus, if we have a semigroup, we can identify the Cauchy problem of which it is a solution. Usually, however, we are interested in the reverse question, that is, in finding the semigroup for a given equation. The answer is given by the Hille–Yoshida theorem (or, more properly, the Feller–Miyadera–Hille–Phillips–Yosida theorem).

4.1 Around the Hille–Yosida Theorem

Theorem 28 *$A \in \mathcal{G}(M, \omega)$ if and only if*

(a) A is closed and densely defined,

(b) there exist $M > 0, \omega \in \mathbb{R}$ such that $(\omega, \infty) \subset \rho(A)$ and for all $n \geq 1, \lambda > \omega$,

$$\|(\lambda I - A)^{-n}\| \leq \frac{M}{(\lambda - \omega)^n}. \tag{51}$$

If A is the generator of $(G(t))_{t\geq 0}$, then properties (i) and (ii) follow from the formula relating $(G(t))_{t\geq 0}$ with $R(\lambda, A)$: for $\lambda > \omega_0(G)$, where $\omega_0(G)$ is defined by (45), then $\lambda \in \rho(A)$ and

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t} G(t)x dt \tag{52}$$

is valid for all $x \in X$.

Another widely used formula relating A with $(G(t))_{t\geq 0}$ is:

$$G(t)x = \lim_{n\to\infty}\left(I - \frac{t}{n}A\right)^{-n} x = \lim_{n\to\infty}\left(\frac{n}{t}R\left(\frac{n}{t}, A\right)\right)^{n} x \tag{53}$$

for any $x \in X$, and the limit is uniform in t on bounded intervals.

As we noticed earlier, a given operator $(A, D(A))$ can generate at most one C_0-semigroup. Using the Hille–Yosida theorem we can prove a stronger result which is useful later.

Proposition 4 *Assume that the closure $(\overline{A}, D(\overline{A}))$ of an operator (A, D) generates a C_0-semigroup in X. If $(B, D(B))$ is also a generator such that $B|_D = A$, then $(B, D(B)) = (\overline{A}, D(\overline{A}))$.*

Without the assumption that the closure of A is a generator there may be infinitely many extensions of a given operator which generate a semigroup: consider the semigroups generated by the realizations of the Laplacian subject to Dirichlet, Neumann, or mixed boundary conditions – all the generators coincide if restricted to the space of C_0^∞ functions.

Example 29 Let $X = L_p(I)$, where I is either $\mathbb{R}$ or $\mathbb{R}_+$. In both cases we can define a *(left) translation semigroup* by

$$(G(t)f)(s) := f(t+s), \quad f \in X, \text{ and } s, t \in I. \tag{54}$$

The semigroup property is obvious. Next, for each $t \geq 0$, we have

$$\|G(t)f\|_p^p = \int_I |f(t+s)|^p ds \leq \int_I |f(r)|^p dr = \|f\|_p^p,$$

where, in the case $I = \mathbb{R}$, we have the equality. Hence $(G(t))_{t\geq 0}$ satisfies

$$\|G(t)\| \leq 1, \tag{55}$$

and so $(G(t))_{t\geq 0}$ is a semigroup of contractions.

To prove that $(G(t))_{t\geq 0}$ is strongly continuous, we use an approximation approach. First let $\phi \in C_0^\infty(I)$. It is uniformly continuous (having compact support) hence for any $\epsilon > 0$ there is $\delta > 0$ such that for any $s \in I$ and $0 < t < \delta$,

$$|\phi(t+s) - \phi(s)| < \epsilon.$$

Thus,

$$\int_I |\phi(t+s) - \phi(s)|^p ds \leq M_\phi \epsilon^p,$$

where M_ϕ is the measure of some fixed neighbourhood of the support of ϕ containing supports of all $s \to \phi(t+s)$ with $0 < t < \delta$. Because $C_0^\infty(I)$ is dense in $L_p(I)$ for $1 \leq p < \infty$, (55) allows us to use Banach-Steinhaus thorem to claim that $(G(t))_{t\geq 0}$ is a strongly continuous semigroup.

It follows that there is a measurable representation $(t,s) \to [G(t)f](s)$ of $G(t)f$ which is measurable on $\mathbb{R}_+ \times I$ and such that the Riemann integral of $t \to G(t)f$ coincides for almost every $s \in I$ with the Lebesgue integral of $[G(t)f](s)$ with respect to t. Note that in this case it follows directly as the composition of a measurable function with $(t,s) \to t+s$ is measurable, but in general it is not that obvious. Hence, from now on we do not distinguish between a vector-valued function and its measurable representation.

Let us denote by $(A, D(A))$ the generator of $(G(t))_{t\geq 0}$ and let $g := Af \in L_p(I)$. Thus, $\Delta_h f := h^{-1}(G(h)f - f) \to g$ in $L_p(I)$. Taking a compact interval $[a,b] \subset I$, we have

$$\left|\int_a^b (\Delta_h f(s) - g(s))ds\right| \leq \int_a^b |\Delta_h f(s) - g(s)|ds$$

$$\leq \quad |b-a|^{1/q}\|\Delta_h f - g\|_{L_p(I)},$$

so

$$\lim_{h\to 0^+} \int_a^b h^{-1}(f(s+h)-f(s))ds = \int_a^b g(s)ds.$$

On the other hand, we can write

$$\int_a^b h^{-1}(f(s+h)-f(s))ds = h^{-1}\int_b^{b+h} f(s)ds - h^{-1}\int_a^{a+h} f(s)ds,$$

where the terms are the difference quotients of the function $\int_{t_0}^t f(s)ds$ at $t=a$ and $t=b$, respectively. Because f is integrable on compact intervals, $\int_{t_0}^t f(s)ds \in AC(I)$ (absolutely continuous) and its derivative is almost everywhere given by the integrand f. By redefining f on a set of measure zero, we can write

$$f(x) = f(a) + \int_a^x g(s)ds, \quad x \in I.$$

Thus, we see that $A \subset T$, where T is the maximal differential operator on $L_p(I)$. Since T is invertible, similarly to Proposition 4 we obtain $A = T$.

We note that the identification of the generator of the translation semigroup can be done by finding the resolvent through the Laplace transform (52).

4.2 Dissipative Operators

Let X be a Banach space (real or complex) and X^* be its dual. From the Hahn–Banach theorem, for every $x \in X$ there exists $x^* \in X^*$ satisfying

$$<x^*, x>= \|x\|^2 = \|x^*\|^2.$$

Therefore the *duality set*

$$\mathcal{J}(x) = \{x^* \in X^*;\ <x^*, x>= \|x\|^2 = \|x^*\|^2\} \qquad (56)$$

is nonempty for every $x \in X$.

Definition 14 *We say that an operator $(A, D(A))$ is* dissipative *if for every $x \in D(A)$ there is $x^* \in \mathcal{J}(x)$ such that*

$$\Re < x^*, Ax > \leq 0. \tag{57}$$

An important equivalent characterisation of dissipative operators, [43, Theorem 1.4.2], is that A is dissipative if and only if for all $\lambda > 0$ and $x \in D(A)$,

$$\|(\lambda I - A)x\| \geq \lambda \|x\|. \tag{58}$$

We note some important properties of dissipative operators.

Proposition 5 *[26] If $(A, D(A))$ is dissipative, then*

(i) $Im(\lambda I - A) = X$ for some $\lambda > 0$ if and only if $Im(\lambda I - A) = X$ for all $\lambda > 0$.

(ii) A is closed if and only if $Im(\lambda I - A)$ is closed for some (and hence all) $\lambda > 0$.

(iii) If A is densely defined, then A is closable and $\overline{A}$ is dissipative. Moreover, $\overline{Im(\lambda I - A)} = Im(\lambda I - \overline{A})$.

Combination of the Hille–Yosida theorem with the above properties gives a generation theorem for dissipative operators, known as the Lumer–Phillips theorem ([43, Theorem 1.43] or [26, Theorem II.3.15]).

Theorem 30 *For a densely defined dissipative operator $(A, D(A))$ on a Banach space X, the following statements are equivalent.*

(a) The closure $\overline{A}$ generates a semigroup of contractions.

(b) $\overline{Im(\lambda I - A)} = X$ for some (and hence all) $\lambda > 0$.

If either condition is satisfied, then A satisfies (57) for any $x^* \in \mathcal{J}(x)$.

In particular, if we know that A is closed then the density of $Im(\lambda I - A)$ is sufficient for A to be a generator. On the other hand, if we do not know a priori that A is closed then $Im(\lambda I - A) = X$ yields A being closed and consequently that it is the generator.

Example 31 If $(A, D(A))$ is a densely defined operator in X and both A and its adjoint A^* are dissipative, then $\overline{A}$ generates a semigroup of contractions in X. In fact, because $\overline{A}$ is dissipative and closed, $Im(I-\overline{A})$ is closed. If $Im(I-\overline{A}) \neq X$, then for some $0 \neq x^* \in X^*$ we have

$$0 = <x^*, x - \overline{A}x> = <x^* - \overline{A}^* x^*, x>$$

for all $x \in D(\overline{A})$. Because $\overline{A}$ is densely defined, $x^* - \overline{A}^* x^* = 0$ and because $\overline{A}^*$ is dissipative, $x^* = 0$. Hence $Im(I - \overline{A}) = X$ and $\overline{A}$ is the generator of a dissipative semigroup by Theorem 30. In particular, dissipative self-adjoint operators on Hilbert spaces are always generators.

4.3 Nonhomogeneous Problems

Consider the problem of finding the solution to:

$$\begin{aligned} \frac{du}{dt}(t) &= Au(t) + f(t), \quad 0 < t < T \\ u(0) &= u_0, \end{aligned} \tag{59}$$

where $0 < T \leq \infty$, A is the generator of a semigroup, and $f : (0, T) \to X$ is a known function. For u to be a continuous solution, f must be continuous.

However, this condition proves to be insufficient. We observe that if u is a classical solution of (59), then it must be given by

$$u(t) = G(t)u_0 + \int_0^t G(t-s)f(s)ds. \tag{60}$$

The integral is well defined even if $f \in L_1([0,T],X)$ and $u_0 \in X$. We call u defined by (60) the *mild solution* of (59). For an integrable f such u is continuous but not necessarily differentiable, and therefore it may be not a solution to (59). The following theorem gives sufficient conditions for a mild solution to be a classical solution (see, e.g., [43, Corollary 4.2.5 and 4.2.6]).

Theorem 32 *Let A be the generator of a C_0-semigroup $(G(t))_{t\geq 0}$ and $x \in D(A)$. Then (60) is a classical solution of (59) if either*

(i) $f \in C^1([0,T],X)$, or

(ii) $f \in C([0,T],X) \cap L_1([0,T],D(A))$.

The assumptions of this theorem are often too restrictive for applications. On the other hand, it is not clear exactly what the mild solutions solve. We present here a result from [26, p. 451] which is particularly suitable for the applications.

Proposition 6 *A function $u \in C(\mathbb{R}_+,X)$ is a mild solution to (59) with $f \in L_1(\mathbb{R}_+,X)$ in the sense of (60) if and only if $\int_0^t u(s)ds \in D(A)$ and*

$$u(t) = u_0 + A\int_0^t u(s)ds + \int_0^t f(s)ds, \qquad t \geq 0. \tag{61}$$

4.4 Long time behaviour of semigroups

It is important to note that the Hille–Yosida theorem is valid in both real and complex Banach spaces with the same formulation. Thus if A is an operator in a real Banach space X, generating a semigroup $(G(t))_{t\geq 0}$, then its complexification will generate a complex semigroup of the same type in the complexification X_C of X. This allows us to extend (52) to complex values of λ. Precisely, the integral in (52) is absolutely convergent for $\Re\lambda > \omega_0(A)$. Moreover, iterations of the resolvent give the following formula,

$$\begin{aligned} R(\lambda, A)^n x &= \frac{(-1)^{n-1}}{(n-1)!}\frac{d^{n-1}}{d\lambda^{n-1}}R(\lambda, A) \\ &= \frac{1}{(n-1)!}\int\limits_0^\infty t^{n-1}e^{-\lambda t}G(t)xdt, \end{aligned} \tag{62}$$

valid for all $x \in X$.

4.4.1 Story of four numbers

Formula (62) yields the estimate

$$\|R(\lambda, A)^n\| \leq \frac{M}{(\Re\lambda - \omega_0(G))^n}, \qquad \Re\lambda > \omega_0(G). \tag{63}$$

An immediate consequence of the above considerations is that the spectrum of a semigroup generator is always contained in a left half-plane, given by the spectral bound

$$s(A) = \sup\{\Re\lambda;\ \lambda \in \sigma(A)\}, \tag{64}$$

defined in (21). For semigroups generated by bounded operators and, in particular, by matrices, Liapunov's theorem, see e.g. [26, Theorem I.2.10], states that the type $\omega_0(G)$ of the semigroup is equal to $s(A)$. This is no longer true for strongly continuous semigroups in general; see for example, [43, Example 4.4.2] or [38, Example A-III.1.3], where it is shown that the translation semigroup $[G(t)f](s) = f(t+s)$ on the space $X = L_p(\mathbb{R}_+) \cap E$, where E is the weighted space $E := \{f \in L_p(\mathbb{R}_+), e^s ds\}$, whose generator A is the differentiation operator, satisfies $\omega_0(G) = 0$ and $s(A) = -1$.

That the type $\omega_0(G)$ might be a rather crude estimate of $s(A)$ can be expected because the former is determined by the absolute convergence of the Laplace integral and the Laplace integral may converge as an improper integral in a possibly larger half-plane $\Re\lambda > abs(G)$, where by $abs(G)$ we denoted the abscissa of convergence (of the Laplace integral treated as an improper integral). That, $abs(G) = \inf\{\lambda \in \mathbb{C}\}$ for which

$$B_\lambda x := \lim_{\tau\to\infty} \int\limits_0^\tau e^{-\lambda t} G(t) x dt \tag{65}$$

exists for all $x \in X$. Moreover, any such λ satisfies $\lambda \in \rho(A)$ and $B_\lambda x = R(\lambda, A)x$ for all $x \in X$.

Thus at this moment we only have the obvious estimate

$$s(A) \leq \omega_0(G) < +\infty. \tag{66}$$

We can prove, however, that $abs(G)$ controls the growth of classical solutions of (47), (48), that is, of the solutions emanating from $x \in D(A)$. To make this concept precise, we define the *growth bound* $\omega_1(G)$ by

$$\begin{aligned} \omega_1(G) &= \inf\{\omega; \text{there is } M \text{ such that } \|G(t)x\| \leq Me^{\omega t}\|x\|_{D(A)}, \\ & \quad x \in D(A), t \geq 0\}. \end{aligned} \tag{67}$$

Clearly, $\omega_1(G) \leq \omega_0(G)$. The following result is true.

Proposition 7 *For a semigroup* $(G(t))_{t\geq 0}$ *we have*

$$\omega_1(G) = abs(G). \tag{68}$$

4.4.2 Fine structure of the spectrum of A and long time behaviour of $(G(t))_{t\geq 0}$

One of the most important questions in the theory of strongly continuous semigroups is to determine the long time behaviour of a semigroup through spectral properties of its generator.

Spectral Mapping Theorem for semigroups If $(G(t))_{t\geq 0}$ is generated by a bounded operator A, then $G(t) = \exp tA$ and the the Spectral Mapping Theorem (23) gives

$$\sigma(G(t)) = e^{t\sigma(A)}. \tag{69}$$

Hence

$$e^{t\omega(G)} = r(G(t)) = e^{ts(A)}$$

and thus, in particular, (69) yields the Lyapunov theorem for dynamical systems generated by bounded operators. However, we have seen that for C_0-semigroups the spectrum of the generator does not fully determine the spectrum of the semigroup; that is, the Spectral Mapping Theorem (23) fails in this case.

Note that while the number zero can be in the spectrum of a semigroup $(G(t))_{t\geq 0}$ (e.g. for eventually compact semigroups), it cannot be obtained from any finite spectral value of A through (69). Thus, we shall restrict our considerations to $\sigma(G(t)) \setminus \{0\}$. Furthermore, validity of (69) for a given $\lambda \in \sigma(G(t))$ means that there exist $k \in \mathbb{Z}$ such that

$$\mu + 2k\pi i/t \in \sigma(A) \quad \text{with} \quad \lambda = e^{t\mu}. \tag{70}$$

We note the following general result, [26, Theorems 6.2 and 6.3]

Theorem 33 *Let $(G(t))_{t\geq 0}$ be the strongly continuous semigroup generated by A. Then*

1. $e^{t\sigma(A)} \subset \sigma(G(t))$*;*

2. $e^{t\sigma_p(A)} = \sigma_p(G(t)) \setminus \{0\}$*;*

3. $e^{t\sigma_r(A)} = \sigma_r(G(t)) \setminus \{0\}$*;*

4. $e^{t\sigma_a(A)} \subset \sigma_a(G(t))$

Ultrapowers in the context of semigroups As we noted above, the main obstacle for validity of the Spectral Mapping Theorem is caused by the approximate spectrum. We have introduced a method of converting the approximate spectrum into the point spectrum in Paragraph 2.2.2 and it is natural to ask whether it can be used to alleviate the encountered problems. Let $(G(t))_{t\geq 0}$ be a strongly continuous semigroup. As noted in Par. 2.2.2, for each $t \geq 0$, the bounded operator $G(t)$ extends to $\widehat{G(t)}$ on X preserving norms, spectra etc. Unfortunately, the family $(\widehat{G(t)})_{t\geq 0}$ is strongly continuous if and only if the generator A of $(G(t))_{t\geq 0}$ is bounded. The problem is created at the first step of construction as the extension of $(G(t))_{t\geq 0}$ to $l_\infty(X)$, denoted by $(\tilde{G}(t))_{t\geq 0}$,

$$\tilde{G}(t)[(x_n)_{n\in\mathbb{N}}] := (G(t)x_n)_{n\in\mathbb{N}}$$

is not strongly continuous.

To get around this difficulty, we proceed as in the definition of the *sun-dual* and first define the subspace of $l_\infty(X)$ by

$$l_\infty^G(X) := \{(x_n)_{n\in\mathbb{N}} \in l_\infty(X);\ \lim_{t\to 0^+} \|G(t)x_n - x_n\| = 0, \quad \text{uniformly in } n\}$$

Clearly $l_\infty^G(X)$ is $(\tilde{G}(t))_{t\geq 0}$ invariant and it turns out that the restriction of $(\tilde{G}(t))_{t\geq 0}$ to this subspace is strongly continuous. Moreover, since a strongly continuous semigroup is uniformly continuous on compact subsets, we see that $c_0(X) \subset l_\infty^G(X)$. Then, instead of $\hat{X}$, we consider the quotient space

$$\hat{X}^G = l_\infty^G(X)/c_0(X) \tag{71}$$

and define the semigroup $(\hat{G}(t)(t))_{t\geq 0}$ as the canonical projection of $(\tilde{G}(t)(t))_{t\geq 0}$ to $\hat{X}^G$:

$$\hat{G}(t)[(x_n)_{n\in\mathbb{N}} + c_0(X)] := (G(t)x_n)_{n\in\mathbb{N}} + c_0(X) \tag{72}$$

for $(x_n)_{n\in\mathbb{N}} \in l_\infty^G(X)$. Again, with the canonical injection $X \ni x \to (x, x, \ldots) \in \hat{X}^G$ the operators $\hat{G}(t)$ become extensions of $G(t)$ for any $t \geq 0$ and restrictions of $\widehat{G(t)}$ defined on $\hat{X}$. Using standard results for quotient semigroups, we find that the generator $\hat{A}$ of $(\hat{G}(t)(t))_{t\geq 0}$ on $\hat{X}^G$ is given by

$$\begin{aligned}
&\hat{A}[(x_n)_{n\in\mathbb{N}} + c_0(X)] = (Ax_n)_{n\in\mathbb{N}} + c_0(X) \quad \text{on} \\
&D(\hat{A}) = \{(x_n)_{n\in\mathbb{N}} + c_0(X);\ (x_n)_{n\in\mathbb{N}} \in D(A), \\
&\qquad (x_n)_{n\in\mathbb{N}}, (Ax_n)_{n\in\mathbb{N}} \in \hat{X}^G\}
\end{aligned}$$

Unfortunately, there is a price to pay: in general it is not true that $\sigma(G(t)) = \sigma(\hat{G}(t))$. This apparent contradiction with Theorem 5 is explained by the observation that the later theorem would refer to $\widehat{G(t)}$ and not to $\hat{G}(t)$. For instance, an approximate eigenvector for $G(t)$ may fail to satisfy the condition defining $l_\infty^G(X)$ and thus fail to be an approximate eigenvector of $\hat{G}(t)$. Of course, if an approximate eigenvector $(x_n)_{n\in\mathbb{N}}$ satisfies $(x_n)_{n\in\mathbb{N}} \subset l_\infty^G(X)$, then $\hat{x} = (x_n)_{n\in\mathbb{N}} + c_0(X)$ it is an eigenvector of $\hat{G}(t)$. We will see one way of getting around this difficulty below.

Eventually uniformly continuous semigroups If a semigroup $(G(t))_{t\geq 0}$ is continuous in the uniform operator topology for $t \geq 0$, then is generator is bounded and we can use classical Lyapunov theorem. However, if $(G(t))_{t\geq 0}$ is uniformly continuous for $t > 0$ (*immediately uniformly continuous*) or even for $t \geq t_0$ for some $t_0 > 0$ (*eventually uniformly continuous*), then the situation becomes nontrivial. We note that analytic semigroups and eventually compact semigroups are eventually uniformly continuous.

To prove the latter statement assume that $T(t_0)$ is compact and let $t, s \geq t_0$. Since $t \to G(t)x$ is uniformly continuous for x in compact sets (Banach-Steinhaus theorem) and $G(t_0)B_1$ is relatively compact (B_1 is the unit ball), we have that

$$\lim_{t\to s}(G(t)x - G(s)x) = \lim_{t\to s}(G(t-t_0) - G(s-t_0))G(t_0)x$$

converges uniformly to zero for $x \in B_1$ giving uniform continuity of $(G(t))_{t\geq 0}$ for $t \geq t_0$.

Theorem 34 *If $(G(t))_{t\geq 0}$ is an eventually uniformly continuous semigroup with generator A, then*

$$\sigma(G(t)) \setminus \{0\} = e^{t\sigma(A)}.$$

Proof. For the proof it suffices to show that $\sigma_a(G(t))\setminus\{0\} \subset e^{t\sigma_a(A)}$. Furthermore, it is enough to consider $1 \in \sigma_a(G(t_1))$ for some $t_1 > 0$. In fact, any other λ and t_2 can be reduced to this situation by considering the rescaled semigroup

$$(S(t)(t))_{t\geq 0} = (e^{-t\ln\lambda/t_1}G(tt_2/t_1))_{t\geq 0}$$

with generator $B = (t_2A - \ln\lambda)/t_1$. The spectral properties of 1 for $S(t_1)$ are the same as of λ for $G(t_2)$ as

$$G(t_2) - \lambda I = \lambda(S(t_1) - I).$$

Take $(f_n)_{n\in\mathbb{N}} \in \sigma_a(G(t_1))$; that is $\|f\| = 1$ with

$$\lim_{n\to\infty} \|G(t_1)f_n - f_n\| = 0.$$

Let $(G(t))_{t\geq 0}$ be uniformly continuous for $t \geq t_0$. We choose $k \in \mathbb{N}$ such that $kt_1 > t_0$ and define $g_n = G(kt_1)f_n$. Then we have

$$\lim_{n\to\infty} \|g_n\| = \lim_{n\to\infty} \|[G(t_1)]^n f_n\| = \lim_{n\to\infty} \|f_n\| = 1$$

as well as

$$\lim_{n\to\infty} \|G(t_1)g_n - g_n\| \leq \lim_{n\to\infty} \|G(t_1)\|^k \|G(t_1)f_n - f_n\| = 0,$$

so $(g_n)_{n\in\mathbb{N}}$ also is an approximate eigenvector with approximate eigenvalue 1. However, $(G(t))_{t\geq 0}$ is uniformly continuous on sets of the form $G(t_0)U$ where U is a bounded set. In particular, $(G(t))_{t\geq 0}$ is uniformly continuous on $(g_n)_{n\in\mathbb{N}}$ and hence $\hat{g} = (g_n)_{n\in\mathbb{N}}$ is an element in the semigroup ultrapower $\hat{X}^G$.

By comments at the end of Example 4.4.2, $\hat{g}$ is an eigenvector of $(\hat{G}(t))_{t\geq 0}$ with eigenvalue 1 hence, by the Spectral Mapping Theorem for the point spectrum, there is an eigenvalue $2\pi i n/t_1$ of $\hat{A}$ for some $n \in \mathbb{Z}$. Since $\sigma(A) = \sigma(\hat{A})$, we obtain the thesis. $\square$

Another theorem which plays an important role in analysis of long time behaviour of semigroups is

Theorem 35 *If $(A, D(A))$ is the generator of an eventually uniformly continuous semigroup $(G(t))_{t\geq 0}$, then, for every $b \in \mathbb{R}$, the set $\{\lambda \in \sigma(A);\ \Re\lambda \geq b\}$ is bounded.*

Proof. Fix arbitrary $a > \omega_0(G)$. The proof consists in showing that for every $\gamma > 0$ there exist $r_0 \geq 0$ such that for any $r > r_0$ we have $dist(a + ir, \sigma(A)) \geq \gamma$. Indeed, if we assume the contrary, then there exists γ such that for any r_0 there is $r > r_0$ with $dist(a + ir, \sigma(A)) < \gamma$ which, in turn, shows that $\sigma(A)$ extends to infinity in the strip $a - \gamma =: b < \Re\lambda < a$, showing its unboundeness. Further, using (18), we find

$$dist(a + ir, \sigma(A)) = \frac{1}{r(R(\lambda + ir, A))} \geq \|R(\lambda + ir, A)^n\|^{-1/n}$$

so we have to prove that for any $\epsilon > 0$ there is r_0 and n such that for all $r > r_0$ we have $\|R(\lambda + ir, A)^n\|^{1/n} < \epsilon$.

The proof uses the representation

$$R(a + ir, A)^{n+1}x = \frac{1}{n!}\int_0^\infty e^{-(a+ir)t}t^n G(t)xdt.$$

We use the fact that $t \to \|G(t)\|$ is measurable. If $(G(t))_{t\geq 0}$ is uniformly continuous for $t \geq t_1$, then the domain of integration is split into $[0, t_1]$, $[t_1, t_2]$ and $[t_2, \infty)$. The first integral can be made uniformly small by sufficiently large n, the last by sufficiently large t_2 and the for the integral over $[t_1, t_2]$ for fixed t_2 we use uniform continuity of the integrand and the Riemann-Lebesgue lemma to show that it is small for sufficiently large r. □

4.4.3 Bad spactrum – chaos

Though our main interest lies with linear dynamical systems, the general framework discussed here applies to a much larger class of dynamical systems.

Let the space (X, d) be a complete metric space and $(G(t))_{t\geq 0}$ be a continuous dynamical system on X with generator A. By $O(p) = \{G(t)p\}_{t\geq 0}$ we denote the *orbit* of $(G(t))_{t\geq 0}$ originating from p.

We say that $(G(t))_{t\geq 0}$ is *topologically transitive* if for any two non-empty open sets $U, V \subset X$ there is $t_0 \geq 0$ such that $G(t_0)U \cap V \neq \emptyset$.

A *periodic point* of $(G(t))_{t\geq 0}$ is any point $p \in X$ satisfying $G(\tau)p = p$ for some $\tau > 0$.

Definition 15 *[23] Let X be a metric space. A dynamical system $(G(t))_{t\geq 0}$ in X is said to be (topologically)* **chaotic** *in X if it is transitive and its set of periodic points is dense in X.*

Devaney's definition is related to the property called hypercyclicity: a dynamical system $(G(t))_{t\geq 0}$ is called *hypercyclic* if for some $x \in X$ we have

$$\overline{\{G(t)x\}_{t\geq 0}} = X;$$

that is, $(G(t))_{t\geq 0}$ has a dense orbit in X.

Hypercyclicity is equivalent to topological transitivity. Thus, Devaney's definition means that $(G(t))_{t\geq 0}$ is chaotic if it has an orbit dense in X and its set of periodic points is dense.

Remark. Hypercyclic (and thus chaotic) dynamical systems can only occur in separable spaces.

Positive criteria The classical criterion for chaoticity of linear semigroups is given in the following theorem.

Theorem 36 *[22] Let X be a separable Banach space and let A be the generator of a semigroup $(G(t))_{t\geq 0}$ on X. Suppose that*

1. The point spectrum of A, $\sigma_p(A)$, contains an open connected set U such that $U \cap i\mathbb{R} \neq \emptyset$;
2. There exists a selection $U \ni \lambda \to x(\lambda)$ of eigenvectors of A, that is analytic in U;
3. $\overline{Span\{x(\lambda), \lambda \in U\}} = X$.
Then $(G(t))_{t\geq 0}$ is chaotic.

The proof uses the observation that $(G(t))_{t\geq 0}$ is hypercyclic if

$$\begin{aligned} X_0 &= \{x \in X;\ \lim_{t\to\infty} G(t)x = 0\} \\ X_\infty &= \{w \in X;\ \forall_{\epsilon>0}\exists_{x\in X, t>0}\|x\| < \epsilon \\ &\qquad \text{and } \|G(t)x - w\| < \epsilon\} \end{aligned}$$

are dense in X. Thus, if also the set of periodic points X_p is dense, then $(G(t))_{t\geq 0}$ is chaotic. Condition 3. is used through the following argument. If $U' \subset U$ with an accumulation point in U and $\Phi \in X^*$ satisfy $< \Phi, x(\lambda) >= 0$ for $\lambda \in U'$, then from the principle of isolated zeros $F_\Phi(\lambda) =< \Phi, x(\lambda) >= 0$ in U which by Condition 3. is possible only if $\Phi = 0$. This in turn shows that $\overline{Span\{x(\lambda), \lambda \in U'\}} = X$. Now, it is easy to see that the sets $U_- = U \cap \{\lambda, \Re\lambda < 0\}, U_+ = U \cap \{\lambda, \Re\lambda > 0\}, U_0 = U \cap \{\lambda, \Re\lambda = 0, \Im\lambda \text{ is rational}\}$ have accumulation points in U. Moreover $Span\{x(\lambda), \lambda \in U_-\} \subset X_0$, by $x(\lambda) = G(t)e^{-\lambda t}x(\lambda)$ we see that $Span\{x(\lambda), \lambda \in U_+\} \subset X_\infty$ and $Span\{x(\lambda), \lambda \in U_0\} \subset X_p$ so that if Condition 3 is satisfied, X_0, X_∞ and X_p are dense in X and therefore $(G(t))_{t\geq 0}$ is chaotic.□

Proposition 8 *If A is a closed operator in X and for some function $x(\lambda)$ that is analytic in an open connected set U we have*

$$Ax(\lambda) = \lambda x(\lambda), \tag{73}$$

then, denoting by a_{n,λ_0} the n-th coefficient of Taylor's expansion of $x(\lambda)$ at $\lambda_0 \in U$, we have

$$Z = Z_{\lambda_0} = \overline{Span\{a_{n,\lambda_0}, n \in \mathbb{N}_0\}}$$

is independent of λ_0. Moreover, for any $U' \subset U$ having an accumulation point in U we have

$$Z = \overline{Span\{x(\lambda), \lambda \in U'\}} = \overline{Span\{x(\lambda), \lambda \in U\}}.$$

The proof is an essay about the identity

$$0 =< \Phi, x(\lambda) >= \sum_{n=0}^{\infty} < \Phi, a_{n,\lambda_0} > (\lambda - \lambda_0)^n$$

$\lambda, \lambda_0 \in U'$, $\Phi \in X^*$, and the principle of isolated zeros. □

Theorem 37 [16] *Suppose that conditions 1. and 2 of Theorem 36 are satisfied. Then there exists an infinite dimensional closed subspace $Y \subseteq X$ that is invariant for $(G(t))_{t\geq 0}$ such that $(G|_Y(t))_{t\geq 0}$ is chaotic.*

The proof of this result uses the previous proposition to show that closed linear spans of eigenvectors with $\Re\lambda > 0, \Re\lambda < 0$ and $\Re\lambda = 0$ are the same. Since a closed linear span of eigenvectors of the generator is invariant w.r.t. the semigroup, the theorem follows. $\square$

The previous result justifies the following definition.

Definition 16 *Suppose $(G(t))_{t\geq 0}$ is a continuous dynamical system on X. If there exists a closed subspace Y which is invariant for $(G(t))_{t\geq 0}$ such that*

1. $\overline{\{G(t)x\}_{t\geq 0}} = Y$ *for some $x \in Y$, then we say that $(G(t))_{t\geq 0}$ is* **sub-hypercyclic***;*
2. $(G(t))_{t\geq 0}$ *is chaotic in Y, then we say that $(G(t))_{t\geq 0}$ is* **sub-chaotic***.*

The subspace Y is called, respectively, the hypercyclicity and chaoticity subspace for $(G(t))_{t\geq 0}$.

Negative criteria It is important to distinguish cases when the dynamical system cannot be chaotic, even in a subspace.

For sets $M \subset X$ and $N \subset X^*$ denote

$$\begin{aligned} M^{\perp} &= \{f \in X^*; < f, x >= 0, \forall x \in M\} \\ {}^{\perp}N &= \{x \in X; < f, x >= 0, \forall f \in N\}. \end{aligned}$$

Theorem 38 *Let $(G(t))_{t\geq 0}$ be a continuous linear dynamical generated by A in a Banach space X, having an orbit dense in some subspace $X_{ch} \subset X$. Then the adjoint A^* of A and the dual dynamical system $(G^*(t))_{t\geq 0}$ have the following properties:*

(i) Let $0 \neq \phi \in X^*$. If $\{G^*(t)\phi\}_{t\geq 0}$ is bounded, then $\phi \in X_{ch}^{\perp}$,

(ii) If ϕ is an eigenvector of A^*, then $\phi \in X_{ch}^{\perp}$.

In particular, $(G(t))_{t\geq 0}$ cannot be chaotic if

$$\sigma_p(A^*) = \emptyset.$$

The proof is based on the following observation. Let $0 \neq \Phi \in X^*$ be such that $\|G^*(t)\Phi\|$ is bounded. Consider

$$<G^*(t)\Phi, x>=<\Phi, G(t)x> .$$

Along a dense trajectory $\{G(t)x_0\}_{t\geq 0}$ (for a fixed x_0) we can find $x = T(t_\epsilon)x_0$ for which $\|x\| < \epsilon$ and so the right hand side can be made arbitrarily small. This shows (modulo some limiting argument) that Φ is orthogonal to the span of $\{G(t)x_0\}_{t\geq 0}$. Similar argument works for (ii). □

Corollary 2 *Let $E(\lambda)$ be the eigenspace corresponding to λ and*

$$E_* = \overline{\bigoplus_{\lambda\in\sigma(A^*)} E_\lambda}.$$

Then

$$X_{ch} \subseteq {}^{\perp}E_*.$$

Consequently, if

$$codim\ {}^{\perp}E_* < +\infty,$$

then there is no subspace of X in which $(G(t))_{t\geq 0}$ is chaotic.

These results can be used to rule out important classes of semigroups from being hypercyclic.

Corollary 3 *[22] Let $(G(t))_{t\geq 0}$ be a strongly continuous semigroup generated by A in a Banach space X. Assume that $(G(t))_{t\geq 0}$ is eventually uniformly continuous and that the resolvent of A, $R(\lambda, A)$ is compact. Then $(G(t))_{t\geq 0}$ is not hypercyclic.*

Indeed, a hypercyclic semigroup must have positive growth bound $\omega_0(G)$. Since it is eventually uniformly continuous, $s(A) = \omega_0(A) > -\infty$ by Theorem 34. Since $R(\lambda, A)$ is compact, $R(\lambda, A^*)$ is also compact and, since $s(A) > -\infty$, the spectrum of A is not empty and consists solely of eigenvalues. □

For instance, the diffusion semigroup on a bounded domain is analytic with compact resolvent and thus cannot be chaotic.

Recent criteria

Theorem 39 *[25] Let A be the generator of a strongly continuous semigroup $(G(t))_{t\geq 0}$ on a separable Banach space X. Assume that there is $\Omega := (\omega_1, \omega_2) \subset \mathbb{R}$ with $\mu(\Omega) > 0$ and a strongly measurable $x : \Omega \to X$ such that $Ax(\lambda) = i\lambda x(\lambda)$ for almost any $\lambda \in \Omega$ and*

$$\overline{Span\{x(\lambda);\ \lambda \in \Omega \setminus \Omega'\}} = X \tag{74}$$

for any $\Omega' \subset \mathbb{R}$ with $\mu(\Omega') = 0$. Then $(G(t))_{t\geq 0}$ is hypercyclic in X.

Remark. If x is continuous, then (74) can be replaced by

$$\overline{Span\{x(\lambda);\ \lambda \in \Omega\}} = X \tag{75}$$

and one obtains automatically that $(G(t))_{t\geq 0}$ is chaotic in X.

Proof. Note that: (a) x is a non-zero function, and (b) using a scalar multiplier we can assume that x is (Bochner) integrable.

The proof uses the Fourier transform

$$\phi(r) = \int_{-\infty}^{\infty} e^{irs} x(s) ds$$

with $x(s)$ extended by zero outside Ω, if necessary. Denote

$$Y_x = \overline{Span\{\phi(\mathbb{R})\}} = \overline{Span\{\phi(r);\ r \in \mathbb{R}\}}. \tag{76}$$

By Riemann-Lebesgue theorem (see e.g. [26, Lemma C.8]), $\phi \in C_0(\mathbb{R}, X)$; that is, $\lim_{|r|\to\infty} \phi(r) = 0$. Let us fix $r \in \mathbb{R}$. Since

$$[G(t)\phi](r) = \int_{-\infty}^{\infty} e^{i(t+r)s} x(s) ds,$$

we see that

$$\lim_{t\to\infty} [G(t)\phi](r) = 0.$$

Thus, $Span\{\phi(\mathbb{R})\} \subset X_0$. Similarly,

$$\phi(r) = G(t) \int_{-\infty}^{\infty} e^{i(-t+r)s} x(s) ds =: [G(t)\psi](r)$$

where $\|\psi(r)\|$ can be made as small as we wish. Hence, $Span\{\phi(\mathbb{R})\} \subset X_\infty$. The last assumption is used to show that $Y_x = X$. Assume that $\Phi \in X^*$ annihilates $Span\{\phi(\mathbb{R})\}$, then for any r

$$0 =< \Phi, \phi(r) >= \int_{-\infty}^{\infty} e^{irs} < \Phi, x(s) > ds$$

which, by uniqueness of the Fourier transform means that $s \to< \Phi, x(s) >$ is zero almost everywhere. Now, since $x(s)$ is only defined almost everywhere, to assert $\Phi = 0$ we must assume that the property that $\overline{Span\{x(\Omega)\}} = X$ is stable under changes of $\mathbf{x}$ on sets of measure zero. With this assumption we obtain, in particular, that

$$\overline{Span\{\phi(\mathbb{R})\}} = X. \tag{77}$$

Some generalizations A closer look at the proof above shows that actually we have a stronger result:

Corollary 4 *Let X be an arbitrary (not necessarily separable) Banach space. Let all assumptions of Theorem 39 except (74) be satisfied and $s \to x(s)$ is a non-zero function. Then $(G(t))_{t\geq 0}$ is hypercyclic in Y_x.*

Let X be an arbitrary Banach space, (Ω, μ) be a measure space, and $f : \Omega \to X$ be a strongly measurable function. For any measurable $U \subset \Omega$ we define the *essential image* of U through f defined as

$$f(U)_{ess} := \{x \in X;\ \mu(\{s \in U : \|f(s) - x\| < \epsilon\}) \neq 0, \forall \epsilon > 0\},$$

Lemma 3 *Let U be a measurable subset of Ω. The essential image has the following properties:*

(a) If $\mu(U) > 0$, then $f(U)_{ess} \cap f(U) \neq \emptyset$. Consequently, the set Z of elements $x \in U$ such that $f(x) \notin f(U)_{ess}$ satisfies $\mu(Z) = 0$;

(b) If $\mu(U \setminus U') = 0$, and $f(U') \subset \overline{Span\{f(U)_{ess}\}}$, then $\overline{Span\{f(U)_{ess}\}} = \overline{Span\{f(U')\}}$.

(c) $\overline{Span\{f(U)_{ess}\}}$ is separable.

Theorem 40 *Let $(G(t))_{t\geq 0}$ be a C_0-semigroup generated by the operator A on an arbitrary Banach space X. Assume that $\sigma_p(A) \cap i\mathbb{R} =: i\Omega \neq \emptyset$, where $\Omega \subset \mathbb{R}$*

is measurable with $\mu(\Omega) > 0$, and that there is a (strongly) measurable function $x : \Omega \to X$ such that $0 \neq x(\lambda) \in ker(i\lambda - A)$ for any $\lambda \in \Omega$. Then $(G(t))_{t\geq 0}$ is sub-hypercyclic, with the hypercyclicity space $X_x := \overline{Span\{x(\Omega)_{ess}\}}$.

Corollary 5

$$X_x = \overline{Span\{\mathcal{F}[x(\cdot)](r),\ r \in \mathbb{R}\}}$$

where $\mathcal{F}$ is the Fourier transform of $\lambda \to x(\lambda)$.

Corollary 6 *If there is an interval $I \subset \Omega$ such that $x(I) \subset x(\Omega)_{ess}$, then $(G(t))_{t\geq 0}$ is subchaotic (with chaoticity space possibly smaller that X_x).*

Corollary 7 *Under notation of Theorem 40, if $\Omega = [a, b]$ and $x(\lambda)$ is weakly (sequentially) continuous on Ω, then $(G(t))_{t\geq 0}$ is chaotic in $X_f = \overline{Span\{x(\Omega)\}}$.*

A counterexample It is often suggested that sufficiently many periodic solutions leads to chaos. For linear systems, periodic solutions are the solutions corresponding to imaginary eigenvalues, thus Theorem 40 seems to be a step in right direction. However, we have:

Example 41 Consider $X = C_b(\mathbb{R})$, the space of bounded continuous functions with $\sup$ norm and translation semigroup $(G(t))_{t\geq 0}$ on X:

$$(G(t)f)(x) = f(t + x). \tag{78}$$

Clearly,

$$\|G(t)f\| = \sup_{x\in\mathbb{R}} |f(t + x)| = \sup_{x\in\mathbb{R}} |f(x)| = \|f\|$$

for any $f \in X$, thus it is a semigroup of isometries but it is not a C_0-semigroup on X.

Consider, however, $Y = \overline{Span\{f_\gamma;\ \gamma \in \mathbb{R}\}}$, where $f_\gamma(x) = e^{i\gamma x}$. Then $Y \subset X$, $G(t)Y \subset Y$ and $(G(t))_{t\geq 0}$ is a strongly continuous semigroup on Y. Moreover,

$$Af_\gamma = i\gamma f_\gamma$$

hence $i\mathbb{R} \subset \sigma_p(A)$ with corresponding eigenvectors $f_\gamma(x) = e^{i\gamma x}$. Thus we have an example of a strongly continuous semigroup on a (non-separable) Banach space which is not sub-hypercyclic and therefore the richness of the imaginary point spectrum is not sufficient for chaos.

4.5 Positive Semigroups

Definition 17 *Let X be a Banach lattice. We say that the semigroup $(G(t))_{t\geq 0}$ on X is positive if for any $x \in X_+$ and $t \geq 0$,*

$$G(t)x \geq 0.$$

We say that an operator $(A, D(A))$ is resolvent positive if there is ω such that $(\omega, \infty) \subset \rho(A)$ and $R(\lambda, A) \geq 0$ for all $\lambda > \omega$.

A strongly continuous semigroup is positive if and only if its generator is resolvent positive. In fact, the positivity of the resolvent for $\lambda > \omega$ follows from (52) and closedness of the positive cone; see Proposition 2. Conversely, the latter with the exponential formula (53) shows that resolvent positive generators generate positive semigroups.

A number of spectral results for semigroups can be substantially improved if the semigroup in question is positive. The following theorem holds, [39, Theorem 1.4.1].

Theorem 42 *Let $(G(t))_{t\geq 0}$ be a positive semigroup on a Banach lattice, with* *generator A. Then*

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t} G(t)x dt \tag{79}$$

for all $\lambda \in \mathbb{C}$ with $\Re\lambda > s(A)$. Furthermore,

(i) Either $s(A) = -\infty$ or $s(A) \in \sigma(A)$ and

$$s(A) = \omega_1(G);$$

(ii) For a given $\lambda \in \rho(A)$, we have $R(\lambda, A) \geq 0$ if and only if $\lambda > s(A)$;

(iii) For all $\Re\lambda > s(A)$ and $x \in X$, we have $|R(\lambda, A)x| \leq R(\Re\lambda, A)|x|$.

From Theorem 42 we see that the spectral bound of the generator of a positive semigroup controls the growth rate of all classical solutions. However, the strict inequality $s(A) < \omega_0(G)$ can still occur, as was shown by Arendt; see [39, Example 1.4.4]. In this example $X = L_p([1, \infty)) \cap L_q([1, \infty))$, $1 \leq p < q < \infty$, and the semigroup in question is $(G(t)f)(s) := f(se^t)$, $s > 1, t > 0$. Its generator is $(Af)(s) = sf'(s)$ on the maximal domain and it can be proved that $s(A) = -1/p < -1/q = \omega_0(G)$. Interestingly enough, $s(A) = \omega_0(G)$ holds for positive semigroups on L^p-spaces. This was proved a few years ago by L. Weis, see the proof in, say, [39, Section 3.5]. However, for the case $p = 1$, which is most relevant for the applications described in this book, it can be proved with much less effort.

Theorem 43 *Let $(G(t))_{t\geq 0}$ be a positive semigroup on an AL-space and let A be its generator. Then $s(A) = \omega_0(G)$.*

The theorem is a corollary of a general result known as the Datko theorem.

Theorem 44 *Let A be the generator of a semigroup $(G(t))_{t\geq 0}$. If, for some $p \in [1, \infty)$,*

$$\int_0^\infty \|G(t)x\|^p dt < \infty, \tag{80}$$

for all $x \in X$, then $\omega_0(G) < 0$.

Proof of Theorem 43. Defining $<f, x>:= \|x\|$ for $x \in X_+$ we obtain a positive additive functional which can be extended to a bounded positive linear functional by Theorems 14 and 15. Let $\omega > abs(G) = s(A)$ (see Theorem 42). Then for $x \geq 0$ and $\tau > 0$, we have

$$\int_0^\tau e^{-\omega t}\|G(t)x\|dt = \left\langle f, \int_0^\tau e^{-\omega t}G(t)xdt \right\rangle \leq <f, R(\omega, A)x> .$$

Therefore

$$\int_0^\infty e^{-\omega t}\|G(t)x\|dt < +\infty$$

for all $x \in X_+$ and hence for all $x \in X$. Theorem 44 then implies $\|G(t)\| \leq Me^{(\omega-\mu)t}$ for some $\mu > 0$, hence $\omega_0(G) < \omega$ which yields $\omega_0(G) \leq s(A)$ and consequently $s(A) = \omega_0(G)$. □

4.6 Generation through Perturbation

Verifying conditions of the Hille–Yosida, or even the Lumer–Phillips, theorems for a concrete problem is quite often a formidable task. On the other hand, in many cases the operator appearing in the evolution equation at hand is built as a combination of much simpler operators that are relatively easy to analyse. The question now is to what extent the properties of these simpler operators are inherited by the full equation. More precisely, we are interested in the problem:

Problem P. *Let $(A, D(A))$ be a generator of a C_0-semigroup on a Banach space X and $(B, D(B))$ be another operator in X. Under what conditions does $A + B$ generate a C_0-semigroup on X?*

Before attempting to address this problem we point out a difficulty that arises immediately from the above formulation. As A and B are unbounded operators, we have to realize that the sum $A + B$ is, at this moment, defined only as $(A + B)x = Ax + Bx$ on $D(A + B) = D(A) \cap D(B)$, where the latter can reduce in some cases to $\{0\}$. Also, the sum of two closed operators is not necessarily closed: a trivial example is offered by $B = -A$ and $A + B = 0$, defined on $D(A)$, is not a closed operator. Thus, $A + B$ with $B = -A$ does not generate a semigroup. On the other hand, the closure of $A + B$ that is the zero operator defined on the whole space is the generator of a constant uniformly bounded semigroup. This situation happens quite often and suggests that the formulation of Problem P is too restrictive and we often restrict ourselves to the following weaker formulation of it.

Problem P′. *Let $(A, D(A))$ be a generator of a C_0-semigroup on a Banach space X and $(B, D(B))$ be another operator in X. Find conditions that ensure that there is an extension K of $A + B$ that generates a C_0-semigroup on X and characterise this extension.*

The characterisation of extensions of $A+B$ that generate a semigroup (in general, there can be many extensions having this property) provides essential information on the properties of the semigroup and plays a role of the regularity theorems in the theory of differential equations. The best situation is when $K = A + B$ or $K = \overline{A + B}$, as there is then a close link between K and A and B. However, there are cases where K is an unspecified extension of $A + B$ in which case the semigroup can display features that are rather impossible to deduct from the properties of A and B alone.

4.6.1 A Spectral Criterion

Usually the first step in establishing whether $A+B$ or some of its extensions generates a semigroup is to find if $\lambda I-(A+B)$ (or its extension) is invertible for all sufficiently large λ.

In all cases discussed here we have the generator $(A, D(A))$ of a semigroup and a perturbing operator $(B, D(B))$ with $D(A) \subseteq D(B)$.

We note that B is A-bounded; that is, for some $a, b \geq 0$ we have

$$\|Bx\| \leq a\|Ax\| + b\|x\|, \qquad x \in D(A) \tag{81}$$

if and only if $BR(\lambda, A) \in \mathcal{L}(X)$ for $\lambda \in \rho(A)$.

In what follows we denote by K an extension of $A+B$. We now present an elegant result relating the invertibility properties of $\lambda I - K$ to the properties of 1 as an element of the spectrum of BL_λ, first derived in [28].

Theorem 45 *Assume that* $\Lambda = \rho(A) \cap \rho(K) \neq \emptyset$*.*

(a) $1 \notin \sigma_p(BR(\lambda, A))$ *for any* $\lambda \in \Lambda$*;*

(b) $1 \in \rho(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *if and only if* $D(K) = D(A)$ *and* $K = A + B$*;*

(c) $1 \in \sigma_c(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *if and only if* $D(A) \subsetneq D(K)$ *and* $K = \overline{A+B}$*;*

(d) $1 \in \sigma_r(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *if and only if* $K \supsetneq \overline{A+B}$*.*

Corollary 8 *Under the assumptions of Theorem 45,* $K = A + B$ *if one of the following criteria is satisfied: for some* $\lambda \in \rho(A)$ *either*

(i) $BR(\lambda, A)$ *is compact (or, if* $X = L_1(\Omega, d\mu)$*, weakly compact), or*

(ii) the spectral radius $r(BR(\lambda, A)) < 1$.

Proof If (ii) holds, then obviously $I - BR(\lambda, A)$ is invertible by the Neumann series theorem:

$$(I - BR(\lambda, A))^{-1} = \sum_{n=0}^{\infty} (BR(\lambda, A))^n, \tag{82}$$

giving the thesis by Proposition 45 (b). Additionally, we obtain

$$\begin{aligned} R(\lambda, A + B) &= R(\lambda, A)(I - BR(\lambda, A))^{-1} \\ &= R(\lambda, A) \sum_{n=0}^{\infty} (BR(\lambda, A))^n. \end{aligned} \tag{83}$$

If (i) holds, then either $BR(\lambda, A)$ is compact or, in L_1 setting, $(BR(\lambda, A))^2$ is compact, [24, p. 510], and therefore, if $I - BR(\lambda, A)$ is not invertible, then 1 must be an eigenvalue, which is impossible by Theorem 45(c). □

If we write the resolvent equation

$$(\lambda I - (A + B))x = y, \quad y \in X, \tag{84}$$

in the (formally) equivalent form

$$x - R(\lambda, A)Bx = R(\lambda, A)y, \tag{85}$$

then we see that we can hope to recover x provided the Neumann series

$$R(\lambda) := \sum_{n=0}^{\infty} (R(\lambda, A)B)^n R(\lambda, A)y = \sum_{n=0}^{\infty} R(\lambda, A)(BR(\lambda, A))^n y. \tag{86}$$

is convergent. Clearly, if (82) converges, then we can factor out $R(\lambda, A)$ from the series above getting again (83). However, $R(\lambda, A)$ inside acts as a regularising factor and (86) converges under weaker assumptions than (82) and this fact is frequently used to construct the resolvent of an extension of $A + B$ (see, e.g., Theorem 50, Theorem 58 or Section 4.7).

4.6.2 Bounded Perturbation Theorem and Related Results

Theorem 46 *Let* $(A, D(A)) \in \mathcal{G}(M, \omega)$ *for some* $\omega \in \mathbb{R}, M \geq 1$. *If* $B \in \mathcal{L}(X)$, *then* $(K, D(K)) = (A + B, D(A)) \in \mathcal{G}(M, \omega + M\|B\|)$.

Moreover, the semigroup $(G_{A+B}(t))_{t\geq 0}$ *generated by* $A + B$ *satisfies either Duhamel equation:*

$$G_{A+B}(t)x = G_A(t)x + \int_0^t G_A(t-s)BG_{A+B}(s)xds, \quad t \geq 0, x \in X \qquad (87)$$

and

$$G_{A+B}(t)x = G_A(t)x + \int_0^t G_{A+B}(t-s)BG_A(s)xds, \quad t \geq 0, x \in X, \qquad (88)$$

where the integrals are defined in the strong operator topology.

Moreover, $(G_{A+B}(t))_{t\geq 0}$ *is given by the Dyson–Phillips series obtained by iterating (87):*

$$G_{A+B}(t) = \sum_{n=0}^{\infty} G_n(t), \qquad (89)$$

where $G_0(t) = G_A(t)$ *and*

$$G_{n+1}(t)x = \int_0^t G_A(t-s)BG_n(s)xds. \qquad t \geq 0, x \in X. \qquad (90)$$

The series converges in the operator norm of $\mathcal{L}(X)$ *and uniformly for* t *in bounded intervals.*

Proof. First, the problem is reduced to one with $\omega = 0$ by shifting the generator, and then with $M = 1$ by renorming the space using the equivalent norm

$$|||x||| := \sup_{t\geq 0} \|G_A(t)x\|. \qquad (91)$$

Next, because any bounded operator is A-bounded (with constant $a = 0$), by Theorem 45(b) we see that $\lambda \in \rho(A+B)$ if and only if $I - BR(\lambda, A)$ is invertible in $\mathcal{L}(X)$. By the Hille–Yosida theorem this can be achieved if $\Re\lambda > \|B\|$ as then $r(BR(\lambda, A)) \leq \|BR(\lambda, A)\| < 1$ in which case the Neumann series (83) gives the estimate

$$\|R(\lambda, A+B)\| \leq \frac{1}{\Re\lambda} \cdot \frac{1}{1 - \frac{\|B\|}{\Re\lambda}} = \frac{1}{\Re\lambda - \|B\|} \tag{92}$$

yielding the generation result. The Duhamel formula (87) is obtained by considering the function $\phi_x(s) = G_A(t-s)G_{A+B}(s)x$, $x \in D(A)$, and $s \in [0, t]$. Because $G_{A+B}(s)x$ is in $D(A) = D(A+B)$, ϕ_x is differentiable with

$$\frac{d}{ds}\phi_x(s) = G_A(t-s)BG_{A+B}(s)x$$

yielding (87) by integration and extension by density to X, which is justified as all the operators are bounded. The other Duhamel formula follows by considering the function $\psi_x(s) = G_{A+B}(t-s)G_A(s)x$.

Finally, the Dyson–Phillips expansion (89) follows by solving (87) by iterations, as for a scalar Volterra equation.

4.6.3 Perturbations of Dissipative Operators

Theorem 47 *Let A and B be linear operators in X with $D(A) \subseteq D(B)$ and $A + tB$ is dissipative for all $0 \leq t \leq 1$. If*

$$\|Bx\| \leq a\|Ax\| + b\|x\|, \tag{93}$$

for all $x \in D(A)$ with $0 \leq a < 1$ and for some $t_0 \in [0, 1]$ the operator $(A + t_0B, D(A))$ generates a semigroup (of contractions), then $A + tB$ generates a semigroup of contractions for every $t \in [0, 1]$.

Proof. The proof consists in showing, by using Neumann expansion, that if $I - (A + t_0 B)$ is invertible, then $I - (A + tB)$ is invertible provided $|t - t_0| < 1 - a/(2a+b)$. Since the length of the interval on which $I-(A+tB)$ is invertible is independent of the starting point t_0, by using finitely many successive steps, we can cover the whole interval $[0, 1]$. Thus $(A + tB, D(A))$ is a dissipative operator such that $I - (A + tB)$ is surjective for all $t \in [0, 1]$. It is also densely defined because $D(A)$ is dense and so $(A + tB, D(A))$ generates a semigroup of contractions. □

The fact that $a < 1$ in the previous theorem is crucial and a lot of work has been done to change $<$ to $=$. One result, in general setting, is given below. Some others, employing positivity, are discussed further on.

Theorem 48 *Let A be the generator of a semigroup of contractions and B, with $D(A) \subset D(B)$, is such that $A + tB$ is dissipative for all $t \in [0, 1]$. If*

$$\|Bx\| \leq \|Ax\| + b\|x\|, \tag{94}$$

for $x \in D(A)$ and B^* is densely defined, then $\overline{A + B}$ is the generator of a contractive semigroup.

Remark 8 If B is closable and X reflexive, then B^* is automatically densely defined.

We complete this part with a quick glance at possibly the most general perturbation theorem for general operators, called the Miyadera perturbation theorem.

We say that an operator B is a *Miyadera perturbation* of A if B is A-bounded and there exist numbers α and γ with $0 < \alpha < \infty$, $0 \leq \gamma < 1$ such that

$$\int_0^\alpha \|BG_A(t)x\| dt \leq \gamma \|x\| \tag{95}$$

for all $x \in D(A)$.

Theorem 49 *If B is a Miyadera perturbation of A, then $(A+B, D(A))$ is the generator of a C_0-semigroup $(G(t))_{t\geq 0}$.*

4.7 Positive Perturbations of Positive Semigroups

In most perturbation theorems of the previous chapter an essential role was played by a strict inequality in some condition comparing A and B (or $(G_A(t))_{t\geq 0}$ and B). This provided some link between the generator and both operators A and B, and ensured that the semigroup was generated by $A+B$ or, at worst, by $\overline{A+B}$. In many cases of practical importance, however, this inequality becomes a weak inequality or even an equality. We show that in such a case we can still get existence of a semigroup albeit we usually lose control over its generator that can turn to be a larger extension of $A+B$ than $\overline{A+B}$. In such a case the resulting semigroup has properties that are not 'contained' in A and B alone; these are discussed in the next chapter. Here we provide the generation theorem, obtained in [17], which is a generalisation of Kato's result from 1954, [31], as well as some of its consequences.

Theorem 50 *Let X be a KB-space. Let us assume that we have two operators $(A, D(A))$ and $(B, D(B))$ satisfying:*

(A1) A *generates a positive semigroup of contractions* $(G_A(t))_{t\geq 0}$,

(A2) $r(BR(\lambda, A)) \leq 1$ *for some* $\lambda > 0 (= s(A))$,

(A3) $Bx \geq 0$ *for* $x \in D(A)_+$,

(A4) $<x^*, (A+B)x> \leq 0$ *for any* $x \in D(A)_+$, *where* $<x^*, x> = \|x\|$, $x^* \geq 0$.

Then there is an extension $(K, D(K))$ of $(A+B, D(A))$ generating a C_0-semigroup of contractions, say, $(G_K(t))_{t\geq 0}$. The generator K satisfies, for $\lambda > 0$,

$$\begin{aligned} R(\lambda, K)x &= \lim_{n\to\infty} R(\lambda, A)\sum_{k=0}^{n}(BR(\lambda, A))^k x \\ &= \sum_{k=0}^{\infty} R(\lambda, A)(BR(\lambda, A))^k x. \end{aligned} \tag{96}$$

Remark 9 If $-A$ is a positive operator, then assumption (A2) can be replaced by the simpler one:

(*A2'*) $$\|Bx\| \leq \|Ax\|, \qquad x \in D(A)_+.$$

Proof of Theorem 50. We define operators K_r, $0 \leq r < 1$ by $K_r = A + rB$, $D(K_r) = D(A)$. We see that, as by (A2) the spectral radius of $rBR(\lambda, A)$ does not exceed $r < 1$, the resolvent $(\lambda I - (A + rB))^{-1}$ exists and is given by

$$R(\lambda, K_r) := (\lambda I - (A + rB))^{-1} = R(\lambda, A)\sum_{n=0}^{\infty} r^n \left(BR(\lambda, A)\right)^n, \tag{97}$$

where the series converges absolutely and each term is positive. Hence, it follows that

$$\|R(\lambda, K_r)y\| \leq \lambda^{-1}\|y\| \tag{98}$$

for all $y \in X$. Therefore, by the Lumer–Phillips theorem, for each $0 \leq r < 1$, $(K_r, D(A))$ generates a contraction semigroup which we denote $(G_r(t))_{t\geq 0}$. The net $(R(\lambda, K_r)x)_{0\leq r<1}$ is increasing as $r \uparrow 1$ for each $x \in X_+$ and $\{\|R(\lambda, K_r)x\|\}_{0\leq r<1}$ is bounded, so by assumption that X is a KB-space, there is an element $y_{\lambda,x} \in X_+$ such that

$$\lim_{r\to 1^-} R(\lambda, K_r)x = y_{\lambda,x}$$

in X. By the Banach–Steinhaus theorem we obtain the existence of a bounded positive operator on X, denoted by $R(\lambda)$, such that $R(\lambda)x = y_{\lambda,x}$. We use the Trotter–Kato theorem to obtain that $R(\lambda)$ is defined for all $\lambda > 0$ and it is the resolvent of a densely defined closed operator K which generates a semigroup of contractions $(G_K(t))_{t\geq 0}$. Moreover, for any $x \in X$,

$$\lim_{r\to 1^-} G_r(t)x = G_K(t)x, \tag{99}$$

and the limit is uniform in t on bounded intervals and, provided $x \geq 0$, monotone as $r \uparrow 1$. By the monotone convergence theorem, Theorem 25, we have

$$R(\lambda, K)x = \sum_{k=0}^{\infty} R(\lambda, A)(BR(\lambda, A))^k x, \qquad x \in X \tag{100}$$

and we can prove that

$$R(\lambda, K)(\lambda I - (A + B))x = x$$

which shows that $K \supseteq A + B$. □

The semigroup $(G_K(t))_{t\geq 0}$ obtained in Theorem 50 is the smallest in the following sense.

Proposition 9 *Let D be a core of A. If $(G(t))_{t\geq 0}$ is another positive semigroup generated by an extension of $(A + B, D)$, then $G(t) \geq G_K(t)$.*

The assumption (A2) of Theorem 50 is stronger than the assumption that B is A-bounded, used in Theorem 48. Thus, it is worthwhile to compare Theorem 50 with Theorems 48 and 47.

Proposition 10 *Let $(G(t))_{t\geq 0}$ be the semigroup generated by $A + B$ or $\overline{A + B}$ under conditions of Theorems 47 or 48, respectively. If A is a resolvent positive operator and B is positive, then $(G(t))_{t\geq 0}$ is positive.*

Proof. The first part follows as in the proof of Theorem 47 as the extensions are done via the Neumann series which preserves positivity. Consider now the case $a = 1$. All the semigroups $(G_r(t))_{t\geq 0}$ generated by $A + rB$ are positive semigroups of contractions. Moreover, for each $x \in D(A)$ we have

$$\lim_{r\to 1^-}(A + rB)x = (A + B)x$$

and it follows that the semigroup $(G(t))_{t\geq 0}$, generated by $\overline{A + B}$, is the limit of semigroups $(G_r(t))_{t\geq 0}$ as $r \to 1$, that are positive. Hence $(G(t))_{t\geq 0}$ is also positive. □

Thus, if X is reflexive and B is closable, then Theorem 48 is evidently stronger than Theorem 50 as the former requires positivity of neither $(G_A(t))_{t\geq 0}$ nor of B. Moreover, in Theorem 48, we obtain the full characterisation of the generator as the closure of $A + B$. However, checking the closability of the operator B in particular applications is often difficult, whereas the positivity is often obvious. Also, there is a large class of nonclosable operators which can nevertheless be positive, for example, finite-rank operators (in particular, functionals) are closable if and only if they are bounded, [32, p.166]. Moreover, Theorem 50 gives a constructive formula (96) for the resolvent of the generator, which seems to be unavailable in general case, and this, in turn, allows other representation results that are discussed below. Also, what is possibly the most important fact, in nonreflexive spaces Theorem 50 refers to a substantially different class of phenomena because, as we show in the next chapter, in many cases covered by this theorem the generator does not coincide with the closure of $A + B$. Arguments used in the proof of Theorem 50 are very powerful and can be generalized in many ways. We present here a theorem in which the sign of the perturbation is reversed; some other with yet more general perturbations are given below.

Theorem 51 *Let $(A_0, D(A_0))$ be the generator of a positive semigroup of contractions on a KB-space X and $(N, D(N))$ be a positive operator. Assume that there exists an increasing sequence $((N_n, D(N_n)))_{n\in\mathbb{N}}$ of positive operators satisfying*

1. *$D(A_0) \cap D(N)$ is dense in X,*

2. *$D(N_n) \supset D(N)$,*

3. *There is a dense set $D \subset D(A_0) \cap D(N)$ such that $\lim_{n\to\infty} N_n y = Ny$ for $y \in D$,*

4. *$(A_0 - N_n, D(A_0) \cap D(N_n))$ generates a positive semigroup of contractions for $n = 1, 2 \ldots$.*

Then there is an extension $(\mathcal{A}, D(\mathcal{K}))$ of $(A_0 - N, D)$ which generates a semigroup of contractions.

The next result is known as the Desch perturbation theorem.

Theorem 52 *Let A be the generator of a positive C_0-semigroup in $X = L_1(\Omega)$ and let $B \in \mathcal{L}(D(A), X)$ be a positive operator. If for some $\lambda > s(A)$ the operator $\lambda I - A - B$ is resolvent positive, then $(A + B, D(A))$ generates a positive C_0-semigroup on X.*

We note that the Desch theorem, Theorem 52, is in fact equivalent to the Miyadera theorem. This is due to the fact that, for any operator C with $r(C) < 1$, we can introduce an equivalent norm on $X = L_1(\Omega)$ for which $\|C\| < 1$ and, under such norm, the assumptions of the Desch theorem become equivalent to the ones for the Miyadera theorem. This, in particular, yields

Corollary 9 *Let $(G(t))_{t\geq 0}$ be the semigroup generated by $(A + B, D(A))$ (according to Theorem 52). Then $(G(t))_{t\geq 0}$ satisfies the Duhamel equation (88) and is given by the Dyson–Phillips expansion (89).*

Theorem 50 in L_1-setting reads:

Corollary 10 *Let $X = L_1(\Omega)$ and suppose that the operators A and B satisfy*

1. $(A, D(A))$ generates a substochastic semigroup $(G_A(t))_{t\geq 0}$;

2. $D(B) \supset D(A)$ and $Bu \geq 0$ for $u \in D(B)_+$;

3. for all $u \in D(A)_+$

$$\int_\Omega (Au + Bu) d\mu \leq 0. \tag{101}$$

Then the assumptions of Theorem 50 are satisfied.

Proof. First, assumption (101) gives us assumption (A4), that is, dissipativity on the positive cone. Next, let us take $u = R(\lambda, A)x = (\lambda I - A)^{-1}x$ for $x \in X_+$ so that $u \in D(A)_+$. Because $R(\lambda, A)$ is a surjection from X onto $D(A)$, by

$$(A + B)u = (A + B)R(\lambda, A)x = -x + BR(\lambda, A)x + \lambda R(\lambda, A)x,$$

we have

$$-\int_\Omega x\, d\mu + \int_\Omega BR(\lambda, A)x\, d\mu + \lambda \int_\Omega R(\lambda, A)x\, d\mu \leq 0. \tag{102}$$

Rewriting the above in terms of the norm, we obtain

$$\lambda \|R(\lambda, A)x\| + \|BR(\lambda, A)x\| - \|x\| \leq 0, \qquad x \in X_+, \tag{103}$$

from which $\|BR(\lambda, A)\| \leq 1$; that is, assumption (A2) is satisfied. □

The following extension of the above result could be proved by techniques of Theorem 50.

Corollary 11 *Assume that A is the generator of a positive C_0-semigroup of contractions in $X = L_1(\Omega)$ and let $B = B_+ - B_-$ be such that $B_\pm \geq 0$, $D(B_\pm) \supset D(A)$ and there exists $C \geq 0$ with $D(A) \subset D(C)$ such that $B_+ + B_- \leq C$ and for all $x \in D(A)_+$,*

$$\int_\Omega (Ax + Cx) d\mu \leq 0. \tag{104}$$

Then there is an extension K_B of $A+B$ which generates a semigroup of contractions.

The Dyson–Phillips expansion seems to be unavailable for semigroups generated under the assumptions of Theorem 50 in general KB-spaces. However, it can be proved in the L_1 case. The following theorem is a consequence of Theorem 50 but can be also proved from scratch.

Theorem 53 *Under the adopted assumptions, the Dyson–Phillips expansion*

$$G_K(t)f = \sum_{n=0}^{\infty} S_n(t)f, \quad f \in X, \tag{105}$$

where the iterates $S_n(t)$ are defined through

$$\begin{aligned} S_0(t)f &= G_A(t)f, \\ S_n(t)f &= \int_0^t S_{n-1}(t-s)BG_A(s)fds, \quad n > 0, \end{aligned} \tag{106}$$

for $f \in D(A)$ and $t \geq 0$, converges uniformly in t on bounded intervals to a positive semigroup of contractions $(G'(t))_{t\geq 0}$.

This semigroup satisfies the integral equation

$$G'(t)f = G_A(t)f + \int_0^t G'(t-s)BG_A(s)fds \tag{107}$$

for any $f \in D(A)$ and $t \geq 0$. The generator K' of $(G'(t))_{t\geq 0}$ is given by

$$(I-K')^{-1}f = \sum_{n=0}^{\infty} (I-A)^{-1}(B(I-A)^{-1})^n f, \tag{108}$$

and hence $(G'(t))_{t\geq 0}$=$(G_K(t))_{t\geq 0}$, where $(G_K(t))_{t\geq 0}$ is the semigroup obtained in Corollary 10.

5 What can go wrong?

5.1 An overview

Let us consider the classical birth-and-death process that describes the evolution of a population whose size k at any time t may increase to $k+1$ or decrease to $k-1$ owing to a 'birth' or 'death' of an individual; the probability that a birth or death occurs in time interval Δt being $b_k\Delta t + o(\Delta t)$ and $d_k\Delta t + o(\Delta t)$, respectively. If we denote by $u_k(t)$ the probability that the population is of size k at time t, then the corresponding (so-called forward) Kolmogorov system takes the form:

$$\begin{array}{rcl} u_0' &=& -b_0u_0 + d_1u_1, \\ &\vdots& \\ u_n' &=& -(b_n+d_n)u_n + d_{n+1}u_{n+1} + b_{n-1}u_{n-1}, \\ &\vdots& . \end{array} \tag{109}$$

We use the convention that boldface letters denote sequences; for example, $\mathbf{u} = (u_0, u_1, \ldots, u_n, \ldots)$. We also put $b_{-1} = d_0 = 0$ and, to avoid technicalities, we assume that $b_n, d_n > 0$ for all other indices.

System (109) is considered in the Banach space $X = l^1$; this choice is dictated by the fact that if u_k is the probability, then $u_k \geq 0$ and

$$\|\mathbf{u}\| = \sum_{k=0}^{\infty} u_k = 1$$

so that the norm of X should be preserved in the evolution.

First we introduce formal mappings of sequences. Remembering the convention $b_{-1} = d_0 = 0$, we let $\mathbf{w} = \mathcal{A}\mathbf{u} = -\{(b_n + d_n)u_n\}_{n\in\mathbb{N}_0}$. By $\mathcal{B}$ we denote the mapping $\mathbf{v} = \mathcal{B}\mathbf{u}$, where $\mathbf{v} = \{d_{n+1}u_{n+1} + b_{n-1}u_{n-1}\}_{n\in\mathbb{N}_0}$.

The formal mappings $\mathcal{A}$ and $\mathcal{B}$ can define various operators in X. As a basic choice, we define the operator A in X as the restriction of $\mathcal{A}$ to the domain $D(A) = \{\mathbf{u} \in X;\ \mathcal{A}\mathbf{u} \in X\}$. In particular, if $\mathbf{u} \in D(A)_+$, then $\mathbf{v} = \mathcal{B}\mathbf{u} \in X_+$ with

$$\sum_{n=0}^{\infty}(v_n + w_n) = 0. \tag{110}$$

This allows us to define a positive operator B as the restriction of $\mathcal{B}$ to $D(A)$. It follows then that for $\mathbf{u} \in D(A)$ we have

$$\|B\mathbf{u}\| \leq \|A\mathbf{u}\|. \tag{111}$$

As we said earlier, mathematical equations of the applied sciences are built by combining various conservation and constitutive laws. They are also formulated and understood pointwise.

This means that all the operations, such as differentiation, summation, or integration, are meant in the classical 'calculus' sense, and the equation itself is supposed to be satisfied for all reasonable values of the independent variables. Thus the birth-and-death system (109) is basically understood as

$$\mathbf{u}' = \mathcal{A}\mathbf{u} + \mathcal{B}\mathbf{u}, \tag{112}$$

where the system, taken row by row, should be satisfied for all $\mathbf{u}$ for which the expression above makes sense. The modelling interpretation suggests that one should have $u_n(t) \geq 0$ for all $n \in \mathbb{N}_0$ and $t \geq 0$, and

$$\sum_{n=0}^{\infty} u_n(t) = \sum_{n=0}^{\infty} u_n(0) < +\infty, \qquad t > 0.$$

However, if we prove the existence of a semigroup 'solving' (112), then what we really obtain is a solution to a particular reformulation of the original problem in which on the right-hand side stands the generator K of this semigroup. This generator may be quite different from $\mathcal{A} + \mathcal{B}$ and only a detailed characterization of its domain can reveal whether the constructed semigroup gives the full picture of the dynamics described by Eq. (112). As we show, the generator K is between the minimal operator $K_{\min} = A + B$ (defined on $D(A)$) and the maximal operator $K_{\max} = \mathcal{A} + \mathcal{B}$ defined on

$$D_{\max} = \{\mathbf{u} \in X;\ \mathcal{A}\mathbf{u} + \mathcal{B}\mathbf{u} \in X\};$$

that is, $K_{\min} \subset K \subset K_{\max}$. Where K is situated on this scale determines the well-posedness of the problem (112). The following situations are possible

1. $K_{\min} = K = K_{\max}$,
2. $K_{\min} \subsetneq K = \overline{K_{\min}} = K_{\max}$,
3. $K_{\min} = K \subsetneq K_{\max}$,
4. $K_{\min} \subsetneq K = \overline{K_{\min}} \subsetneq K_{\max}$,
5. $\overline{K_{\min}} \subsetneq K \subsetneq K_{\max}$,

and each of them has its own specific interpretation in the model.

In all cases where $K \subsetneq K_{\max}$ we don't have uniqueness; that is, there are differentiable X-valued solutions to (112) emanating from zero and therefore they are not described by the constructed dynamical system: 'there is more to life, than meets the semigroup' [12]. To achieve uniqueness here, one has to impose additional constraints on the solution.

If $\overline{K_{\min}} \subsetneq K$, then despite the fact that the model is formally conservative, (110), the solutions are not; the described quantity leaks out from the system and the mechanism of this leakage is not present in the model. In the Markov processes such a case is referred to as dishonesty of the transition function, [4].

Finally, as b_n, d_n are the rates of change of states in the population, for any solution $\mathbf{u}(t)$, the quantity

$$\Delta t \sum_{n=0}^{\infty} (b_n + d_n) u_n(t) \tag{113}$$

describes the total number of state changes in the time interval Δt. Thus condition $\mathbf{u}(t) \in D(A)$ for any t, equivalent to (113) being finite, reflects the realistic property of a finite total number of 'switches' at any time. Thus, if $K \neq K_{\min}$, then an infinite number of state changes in a finite time interval may occur.

Therefore, strictly speaking, only problems with $K = K_{\min} = K_{\max}$ can be physically realistic. However, in many applications, the last condition is disregarded and the case $K = \overline{K_{\min}} = K_{\max}$ is considered to be 'optimal'.

5.2 The mathematics behind it

5.2.1 Dishonesty

Equations describing the evolution of u are typically constructed by balancing, for any state x, the loss of $u(x,t)$ that is due to the transfer of a part of the population to other states x', and the gain due to the transfer of parts of the population from other states x' to the state x. A general form of such equations is as follows,

$$\partial_t u = T_0 u + Au + Bu, \tag{114}$$

where A is the loss operator, B is the gain operator, and T_0 may describe some transport in the state space (e.g., free streaming or diffusion). The very nature of the modelling process sketched above requires that the described quantity should be preserved; that is, u should add up (or integrate) to a constant independent of t, for instance to 1 if u is the probability density, or to the initial number of particles in the second example mentioned above. If this is the case, then the semigroup describing the evolution is conservative for positive initial data and is called a *stochastic semigroup*.

In many cases, however, the semigroup turns out not to be conservative even though the modelled physical system should have this property. Markov processes exhibiting the latter property are well known in probability theory and are referred to as *dishonest*, or *explosive*, Markov processes. In such cases we have a leakage of the described quantity out of the system that is not accounted for in the modelling processes. This in turn indicates a possibility of the phase transition during evolution and shows that the model does not provide an adequate description of the full process. It seems, however, that this phenomenon is much less understood from the functional-analytic point of view and though a number of scattered results, often limited to a particular application, can be found in earlier literature, [31, 45, 6, 7, 27, **?**, 35, 30], a systematic study has been initiated only recently in a series of papers, [11, 12, 14, 28, 15], and has yielded strong results.

In many cases, however, in the modelling process a mechanism appears that allows the amount of the described quantity to decrease. It could be an absorbing or permeable boundary, or some reaction removing a portion of the quantity from the system. In such a case we say that the semigroup describing the evolution is *strictly substochastic*; that is, the substochasticity of it is not caused by a dishonesty of the process. The theory of Markov processes deals with such a case by introducing an additional state that accounts for the loss, and redefines the process so that the resulting process is Markovian. However, the loss-functional defining the leakage from the system carries important information about the evolution, for example, in the fragmentation models it describes the rate of mass loss due to internal reactions and therefore plays a special role in the description of the process. It is thus important that we do not amalgamate it with other states so that we can keep track of mass loss in the evolution.

Moreover, also for strictly substochastic processes, we can have an analogue of dishonesty; that is, the described quantity can leak out from the system faster than predicted by the loss-functional and thus it is important to separate these two causes of leakage in the model.

Property of honesty/dishonesty of a semigroup is closely related to the characterisation of the generator of the semigroup. To explain why, let us look at a simplified situation when (114) with $T_0 = 0$ is supposed to model a conservative system in $X = L_1(\Omega, d\mu)$; that is, for sufficiently regular u, say $u \in D(A)$,

$$\int_{\Omega} (A + B)ud\mu = 0$$

(the total gain is equal to the total loss, according to our terminology from the beginning of this section). If A generates a substochastic semigroup and B is positive, then by Corollary 10, there is an extension K of $A + B$ generating a semigroup of contractions, say $(G_K(t))_{t\geq 0}$.

Assume now that the semigroup $(G_K(t))_{t\geq 0}$ is generated by $(K, D(K)) = (A+B, D(A))$. Then the solution $u(t) = G_K(t)u_0$, emanating from $u_0 \in D(K)_+$, satisfies $u(t) \in D(A)_+$ and, therefore, because

$$\frac{d}{dt}u(t) = Ku(t) = Au(t) + Bu(t),$$

we obtain that for any $t \geq 0$

$$\frac{d}{dt}\|u(t)\| = \int_{\Omega} \frac{du(t)}{dt}d\mu = \int_{\Omega} (Au(t) + Bu(t))d\mu = 0, \tag{115}$$

so that $\|u(t)\| = \|u_0\|$ for any $t \geq 0$ and the solutions are indeed conservative.

If $K = \overline{A + B}$, then for $u \in D(K)$ there exists a sequence $(u_n)_{n\in\mathbb{N}}$ of elements of $D(A)$ such that $u_n \to u$ and $(A + B)u_n \to Ku$ in X as $n \to \infty$, thus

$$\int_{\Omega} Kud\mu = \lim_{n\to\infty} \int_{\Omega} (A + B)u_n d\mu = 0. \tag{116}$$

This in turn shows again that if $u_0 \in D(K)_+$, then $u(t) = G(t)u \in D(K)_+$ for any $t \geq 0$ and (115) takes the form

$$\frac{d}{dt}\|u(t)\| = \int_\Omega \frac{du(t)}{dt} d\mu = \int_\Omega K u(t) d\mu = 0,$$

and the solutions are conservative as well. That $K = \overline{A+B}$ is also the necessary condition is not that clear but can be proved, see Theorem 56.

To make the above terminology precise, the semigroup $(G(t))_{t\geq 0}$ is said to be a *substochastic semigroup* if for any $t \geq 0$ and $x \geq 0$, $G(t)x \geq 0$ and $\|G(t)x\| \leq \|x\|$, and a *stochastic semigroup* if additionally $\|G(t)f\| = \|f\|$ for $f \in X_+$.

We consider linear operators in $X = L_1(\Omega, d\mu)$: $T \subset T_0 + A$ with $D(T) \subset D(T_0) \cap D(A)$, and B, that satisfy the assumptions of Corollary 10; that is,

1. $(T, D(T))$ generates a substochastic semigroup $(G_T(t))_{t\geq 0}$;
2. $D(B) \supset D(T)$ and $Bf \geq 0$ for $f \in D(B)_+$;
3. for all $f \in D(T)_+$,

$$\int_\Omega (Tf + Bf)\, d\mu = -c(f) \leq 0. \tag{117}$$

c is an integral functional; that is, for some $\varsigma > 0$

$$c(u) = \int_\Omega \varsigma(x) u(x)\, d\mu'_x. \tag{118}$$

Under these assumptions, Corollary 10, Theorem 50, and other results of the previous chapter give the existence of a smallest substochastic semigroup $(G_K(t))_{t\geq 0}$ generated by an extension K of the operator $T + B$. This semigroup, for arbitrary $f \in D(K)$ and $t > 0$, satisfies

$$\frac{d}{dt} G_K(t) f = K G_K(t) f. \tag{119}$$

The semigroup $(G_K(t))_{t\geq 0}$ can be obtained as the strong limit in X of semigroups $(G_r(t))_{t\geq 0}$ generated by $(T + rB, D(T))$ as $r \uparrow 1^-$; if $f \in X_+$, then the limit is monotonic. It is also given as the solution to the Duhamel equation (87) and by the Dyson–Phillips expansion (89). Moreover, the generator K of $(G_K(t))_{t\geq 0}$ is characterised by

$$(\lambda I - K)^{-1} f = \sum_{n=0}^{\infty} (\lambda I - T)^{-1} [B(\lambda I - T)^{-1}]^n f, \ f \in X, \ \lambda > 0. \tag{120}$$

It is important to distinguish the class of semigroups corresponding to $c \neq 0$, as such semigroups cannot be stochastic but their substochasticity is built into the model and not caused by the dishonesty of it.

Definition 18 *A positive semigroup $(G_K(t))_{t\geq 0}$ generated by an extension K of the operator $T + B$ is said to be strictly substochastic if (117) holds with $c \neq 0$.*

Next we extend the concept of honesty to strictly substochastic semigroups.

Definition 19 *We say that a positive semigroup $(G_K(t))_{t\geq 0}$ (generated by an extension K of the operator $T + B$) is honest if c extends to $D(K)$ and for any $0 \leq \overset{\circ}{f} \in D(K)$ the solution $u(t) = G_K(t)\overset{\circ}{f}$ of (119) satisfies*

$$\frac{d}{dt} \int\limits_{\Omega} u(t)\, d\mu = \frac{d}{dt} \|u(t)\| = -c\,(u(t))\,. \tag{121}$$

It can be proved that (121) is equivalent to it 'integrated' version:

Proposition 11 *$(G_K(t))_{t\geq 0}$ is honest if and only if for any $f \in X_+$ and $t \geq 0$,*

$$\|G_K(t)f\| = \|f\| - c\left(\int\limits_0^t G_K(s) f ds\right). \tag{122}$$

This result allows for the introduction of the defect function

$$\eta_f(t) = \|G_K(t)f\| - \|f\| + \int\limits_0^t c\,(G_K(s)f)\,ds \tag{123}$$

for $f \in X_+$ and $t \geq 0$. It follows that η_f is a nonpositive and nonincreasing function for $t \geq 0$. For $\lambda > 0$ we define $L_\lambda = R(\lambda, T) = (\lambda I - T)^{-1}$. Arguing as in (103) we obtain that condition (117) is equivalent to

$$-c(L_\lambda f) = \lambda\|L_\lambda f\| + \|BL_\lambda f\| - \|f\|, \qquad f \in X_+. \tag{124}$$

The following theorem is fundamental for analysing honesty of substochastic semigroups.

Theorem 54 *For any fixed $\lambda > 0$, there is $0 \leq \beta_\lambda \in X^*$ with $\|\beta_\lambda\| \leq 1$ such that for any $f \in X_+$,*

$$\lambda\|R(\lambda, K)f\| = \|f\| - <\beta_\lambda, f> - c\,(R(\lambda, K)f)\,. \tag{125}$$

In particular, c extends to a nonnegative continuous linear functional on $D(K)$, given again by (118).

Proof. Let us fix $f \in X_+$. From (120) and nonnegativity we obtain

$$\lambda\|(\lambda I - K)^{-1}f\| = \lim_{N\to\infty}\sum_{n=0}^{N}\lambda\|L_\lambda(BL_\lambda)^n f\|.$$

By (124) we get

$$\sum_{n=0}^{N}\lambda\|L_\lambda(BL_\lambda)^n f\| = \|f\| - \|(BL_\lambda)^{N+1}f\| - c\left(\sum_{n=0}^{N}L_\lambda(BL_\lambda)^n f\right).$$

By non-negativity, the monotone convergence theorem gives

$$\lim_{N\to\infty} c\left(\sum_{n=0}^{N}L_\lambda(BL_\lambda)^n f\right) = c(R(\lambda, K)f) < +\infty.$$

This shows that c extend to a finite functional on $D(K)$ which is continuous in the graph topology. Returning to (**??**) we see also that $\|(BL_\lambda)^{N+1}f\|$ converges to some $\beta_\lambda(f) \geq 0$ and, by a similar argument, β_λ extends to a continuous linear functional on X with the norm not exceeding 1. $\square$

By taking the Laplace transform of η_f, we obtain

$$<\beta_\lambda, f>= -\lambda \int\limits_0^\infty e^{-\lambda t}\eta_f(t)dt$$

for $f \in X_+$and hence the following result is true

Theorem 55 *$(G_K(t))_{t\geq 0}$ is honest if and only if $\beta_\lambda \equiv 0$ for any (some) $\lambda > 0$.*

In particular

Corollary 12 *If $(G_K(t))_{t\geq 0}$ is dishonest, then there is $f \in X_+$ such that $\|G_K(t)f\| < \|f\| - \int_0^t c\,(G_K(s)f)\,ds$ for any $t > 0$.*

A central result on the characterization of honesty is:

Theorem 56 *[8] The semigroup $(G_K(t))_{t\geq 0}$ is honest if and only if one of the following holds:*
(a) $K = \overline{T+B}$.
(b) $\int\limits_\Omega Ku\,d\mu \geq -c(u), \quad u \in D(K)_+$.

Proof. (a) implies honesty as in (116) - properties of the functional c allow passage to the limit.

Conversely, if $(G_K(t))_{t\geq 0}$ is honest, then $\beta_\lambda \equiv 0$ for any $\lambda > 0$, which means, by the proof of Theorem 54, that

$$\lim_{n\to\infty}(BL_\lambda)^n f = 0.$$

Hence the series in (120) converges to $R(\lambda, \overline{T+B})$.

If $(G_K(t))_{t\geq 0}$ is honest, then first part gives (b) with the equality sign. Conversely, for $u = R(\lambda, K)f$, $f \in X_+$, we have

$$\int\limits_{\Omega} Ku\, d\mu = -\|f\| + \lambda\|R(\lambda, K)f\| = -c(u) - <\beta_\lambda, f>,$$

which implies $<\beta_\lambda, f> \leq 0$ for all $f \in X_+$, thus $\beta_\lambda = 0$. □

Unfortunately, typically we do not know K and thus condition (c) has a limited practical value. There are two important theorems providing conditions for honesty and dishonesty in terms of known operators. The first is based on Theorem 45 which, combined with Theorem 56, shows that $(G_K(t))_{t\geq 0}$ is honest if and only if

$$1 \notin \sigma_p((BL_\lambda)^*).$$

In particular, using the definition of β_λ we see that

$$<\beta_\lambda, BR(\lambda, T)f> = \lim_{n\to\infty} \|(BR(\lambda, T))^{n+1} f\| = <\beta_\lambda, f>,$$

$f \in X_+$, so that

$$(BR(\lambda, T))^* \beta_\lambda = \beta_\lambda. \qquad (126)$$

The other set of results is based on the fact that we know at least one extension of the generator K, namely K_{max}. Let $\mathcal{K}$ be any extension of K.

Theorem 57 *[8]*
(a) If $\int_\Omega \mathcal{K}u\, d\mu \geq -c(u)$ for all $u \in D(\mathcal{K})_+$ then the semigroup is honest.
(b) If there exists $u \in D(\mathcal{K})_+ \cap X$ such that for some $\lambda > 0$
(i) $\lambda u(x) - [\mathcal{K}u](x) = g(x) \geq 0, \quad a.e.,$
(ii) $c(u)$ is finite and

$$\int\limits_{\Omega} \mathcal{K}u\, d\mu < -c(u), \qquad (127)$$

then the semigroup $(G_K(t))_{t\geq 0}$ is not honest.

Proof. The statement (a) is obvious from Theorem 56(c) as $\mathcal{K}$ contains K. In practice, however, we are interested to use the smallest possible extension since taking a too large one could spoil the inequality, [7]. Similarly, (b) uses Theorem 56(c) but here the function $u \in D(\mathcal{K})$, satisfying (127) may fail to belong to $D(K)$; the other two conditions allow one to prove that there is an element of $D(K)$ satisfying (127), thus proving dishonesty of $(G_K(t))_{t\geq 0}$. □

Extension Techniques For further reference we briefly sketch a particularly effective extension technique. We embed $X = L_1(\Omega, d\mu)$ in the set of μ-measurable functions that are defined on Ω and take values in the extended set of real numbers, denoted by E; by E_f we denote the subspace of E consisting of functions that are finite almost everywhere. E is a lattice with respect to the usual relation: '$\leq$ almost everywhere', $X \subset \mathsf{E}_f \subset \mathsf{E}$ with X and E_f being sublattices of E.

Let $\mathsf{F} \subset \mathsf{E}$ be defined by the condition: $f \in \mathsf{F}$ if and only if for any nonnegative and nondecreasing sequence $(f_n)_{n\in\mathbb{N}}$ satisfying $\sup_{n\in\mathbb{N}} f_n = |f|$ we have $\sup_{n\in\mathbb{N}}(I-T)^{-1}f_n \in X$.

We define mapping $\mathsf{L} : \mathsf{F}_+ \to X_+$ by

$$\mathsf{L}f := \sup_{n\in\mathbb{N}} R(1,T)f_n, \qquad f \in \mathsf{F}_+,$$

where $0 \leq f_n \leq f_{n+1}$ for any $n \in \mathbb{N}$, and $\sup_{n\in\mathbb{N}} f_n = f$and extend it to a positive linear operators on the whole $D(\mathsf{B})$ and F, respectively, Theorem 14.

In most applications $(I-T)^{-1}$ is an integral operator with positive kernel so that, by monotone convergence theorem, F coincides is the set of $L_1(\Omega)$ functions for which the integral exists. In the same way we define B on $D(\mathsf{B})$. It turns out that L is one-to-one therefore we can define the operator T with $D(\mathsf{T}) = \mathsf{LF} \subset X$

by

$$\mathsf{T}u = u - \mathsf{L}^{-1}u, \tag{128}$$

so that T is an extension of T. The central theorem of this paragraph reads:

Theorem 58 *If $(T, D(T))$ and $(B, D(B))$ are operators in X such that $(T, D(T))$ generates a substochastic semigroup $(G_T(t))_{t\geq 0}$ on X, $D(B) \supset D(T)$, $Bu \geq 0$ for $u \in D(B)_+$, and*

$$\int_\Omega (Tu + Bu)\, d\mu \leq 0, \tag{129}$$

for all $u \in D(T)_+$, then the extension K of $A + B$, that generates a substochastic semigroup on X by Corollary 10, is given by

$$Ku = \mathsf{T}u + \mathsf{B}u, \tag{130}$$

with

$$\begin{aligned} D(K) &= \{u \in D(\mathsf{T}) \cap D(\mathsf{B}) : \mathsf{T}u + \mathsf{B}u \in X, \\ &\quad \text{and} \lim_{n\to+\infty} ||(\mathsf{LB})^n u|| = 0\}. \end{aligned}$$

Using the current notation, we can give a more focused version of Theorem 57(a).

Theorem 59 *If for any $g \in \mathsf{F}_+$ such that $-g + \mathsf{BL}g \in X$, and $c(\mathsf{L}g)$ exists,*

$$\int_\Omega \mathsf{L}g\, d\mu + \int_\Omega (-g + \mathsf{BL}g)\; d\mu \geq -c\,(\mathsf{L}g)\,, \tag{131}$$

then $K = \overline{T + B}$.

Remark 10 It is worthwhile to reflect on the nature of dishonesty. By definition, $(G_K(t))_{t\geq 0}$ is dishonest if it is not honest and therefore for $(G_K(t))_{t\geq 0}$ to be dishonest, it is enough that (122) does not hold for just one $f \in X_+$ at one moment of time $t > 0$. Hence it makes sense to consider 'pointwise in space' honesty and say that $(G_K(t))_{t\geq 0}$ is *honest along the trajectory* $\{G_K(t)f\}_{t\geq 0}$ if (122) holds for this particular f and for all $t \geq 0$. Accordingly, such a trajectory is called an *honest trajectory*. Thus $(G_K(t))_{t\geq 0}$ is honest if and only if each trajectory $\{G_K(t)f\}_{t\geq 0}$ is honest. Moreover, honesty can also be considered to be a 'pointwise in time' phenomenon. Indeed, if $u(t_0) \in D(\overline{T+B})$ for some $t_0 > 0$ then, by (116),

$$\frac{d}{dt}\|u\|\bigg|_{t=t_0} = -c(u(t_0))$$

and therefore we can say that the trajectory $\{G_K(t)f\}_{t\geq 0}$ is honest over a time interval I if and only if $G_K(t)f \in D(\overline{T+B})$ for $t \in I$.

In other words, $(G_K(t))_{t\geq 0}$ is dishonest along the whole trajectory $\{G_K(t)f\}_{t\geq 0}$ if and only if this trajectory, starting from $f \in D(\overline{T+B})$, leaves $D(\overline{T+B})$ immediately and stays in $D(K) \setminus D(\overline{T+B})$ for all $t > 0$.

In general, our theory cannot determine, in general, whether a given system $(G_K(t))_{t\geq 0}$ can be dishonest along some trajectories and honest along the others. Using specific properties of birth-and-death and fragmentation models, however, we can show that dishonesty in these models is spatially universal. That is, if it occurs along one trajectory, it must occur along any other; see Theorem 65.

Unfortunately, much less can be said about how dishonest trajectories behave in time. One of the reasons for this is that our theory is based on the Laplace transform approach which gives, in some sense, time averages of solutions which provide little information about the properties which are local in time.

5.2.2 Multiple solutions

Let us return to the general Cauchy problem (42), (43). If $\mathcal{A}$ is the generator of a semigroup, then the problem is always uniquely solvable. Hence, multiple solution only can occur if the original operator is not a generator.

Assume that, for a given u_0, (42), (43) has two solutions. Then their difference is again a solution of (42) but corresponding to the null initial condition – it is called a *nul-solution*.

Theorem 60 *[29, Theorem 23.7.1] If $\mathcal{A}$ is a closed operator whose point spectrum is not dense in any right half-plane, then for each $u_0 \in X$ the Cauchy problem of Definition 12 has at most one exponentially bounded solution.*

Proof. The proof essentially follows by taking the Laplace transform of both sides of (42) and some careful manipulation to ensure convergence. □

A useful reformulation of the previous theorem reads as follows:

Theorem 61 *[29, Theorem 23.7.2] Let $\mathcal{A}$ be a closed operator. The Cauchy problem (42), (43) has an exponentially bounded nul-solution of type $\leq \omega$ if and only if the eigenvalue problem*

$$\mathcal{A}y(\lambda) = \lambda y(\lambda) \tag{132}$$

has a solution $y(\lambda) \neq 0$ that is a bounded and holomorphic function of λ in each half-plane $\Re\lambda \geq \omega + \epsilon$, $\epsilon > 0$.

Now we investigate a relation between Cauchy problems (42), (43) and (47), (48). Let $(A, D(A))$ be the generator of a C_0-semigroup $(G(t))_{t\geq 0}$ on a Banach space X. To simplify notation we assume that $(G(t))_{t\geq 0}$ is a semigroup of contractions, hence $\{\lambda;\ Re\lambda > 0\} \subset \rho(A)$. Let us further assume that there exists an extension $\mathcal{A}$ of A defined on the domain $D(\mathcal{A})$. We have the following basic result.

Lemma 4 *Under the above assumptions, for any λ with $Re\lambda > 0$,*

$$D(\mathcal{A}) = D(A) \oplus Ker(\lambda I - \mathcal{A}). \tag{133}$$

The next corollary links Theorem 61 with the above lemma.

Corollary 13 *If $D(\mathcal{A})\backslash D(A) \neq \emptyset$, then $\sigma_p(\mathcal{A}) \supseteq \{\lambda \in \mathbb{C};\ Re\lambda > 0\}$. Moreover, there exists a holomorphic (in the norm of X) function $\{\lambda \in \mathbb{C};\ Re\lambda > 0\} \ni \lambda \to e_\lambda$ such that for any λ with $Re\,\lambda > 0$, $e_\lambda \in Ker(\lambda I - \mathcal{A})$, which is also bounded in any closed half-plane, $\{\lambda \in \mathbb{C};\ Re\lambda \geq \gamma > 0\}$.*

An important observation is that analogous considerations can be carried also for mild (or integral) solutions of (42), (43), defined as for the semigroup: We say that u is a mild solution of (42), (43) if $u \in C([0,\infty), X)$, $\int_0^t u(s)ds \in D(\mathcal{A})$ for any $t > 0$, and

$$u(t) = \overset{\circ}{u} + \mathcal{A}\int_0^t u(s)ds, \quad t > 0. \tag{134}$$

For mild solution we have the following counterpart of Theorem 61.

Corollary 14 *Let $\mathcal{A}$ be a closed operator. If (42), (43) has a mild nul-solution of type $\leq \omega$, then the characteristic equation*

$$\mathcal{A}y(\lambda) = \lambda y(\lambda) \tag{135}$$

has a solution $y(\lambda) \neq 0$, which is a bounded and holomorphic function of λ in each half-plane $Re\lambda \geq \omega + \epsilon$, $\epsilon > 0$. Again, $y(\lambda)$ in (135) can be taken as

$$y(\lambda) = \int_0^\infty e^{-\lambda t}u(t)dt. \tag{136}$$

This observation allows to check uniqueness only for continuous solutions of the integral version of the problem, which is technically simpler.

5.3 Applications to birth-and-death type problems

Let us recall that here we deal with the system

$$\begin{aligned}
u_0' &= -a_0u_0 + d_1u_1, \\
&\vdots \\
u_n' &= -a_nu_n + d_{n+1}u_{n+1} + b_{n-1}u_{n-1}, \\
&\vdots \quad .
\end{aligned} \tag{137}$$

We assume that the rates of change are given and are denoted by d_n and b_n for changes $n \to n-1$ and $n \to n+1$, respectively. In general, we can also include a mechanism that changes a number of objects at the state n by, for example, removing them from the environment or, otherwise, introducing them. The rate of this mechanism is denoted by $\mathbf{c} = (c_n)_{n\in\mathbb{N}}$ and in such a case we have $c_n = b_n + d_n - a_n$. The classical application of this system comes from population theory, where it is a particular case of a Kolmogorov system; in this case u_n is the probability that the described population consists of n individuals and its state can change by either the death or birth of an individual thus moving the population to the state $n-1$ or $n+1$, respectively, hence the name birth-and-death system. The classical birth-and-death system is formally conservative; this is equivalent to $a_n = d_n + b_n$. However, recently a number of other important applications have emerged. For example, [33], we can consider an ensemble of cancer cells structured by the number of copies of a drug-resistant gene they contain. Here, the number of cells with n copies of the gene can change due to mutations, but the cells also undergo division without changing the number of genes in their offspring which is modelled by a nonzero vector $\mathbf{c}$. Finally, system (137) can be thought of as a simplified kinetic system consisting of particles labelled by internal energy n and interacting inelastically with the surrounding matter where in each interaction they can either gain or lose a unit (quantum) of energy. Some particles can decay without a trace or be removed from the system leading again to a nonzero $\mathbf{c}$.

The most common setting for birth-and-death problems is the space l_1. Here we extend it to other l_p spaces to demonstrate the applicability of Theorem 50. The existence results of this section for $p > 1$ can also be proved using Proposition 10; see [19].

5.3.1 Existence Results

Let us recall that the boldface letters denote sequences, for example, $\mathbf{u} = (u_0, u_1, \ldots)$. We assume that the sequences $\mathbf{d}$, $\mathbf{b}$, and $\mathbf{a}$ are nonnegative with $b_{-1} = d_0$.

By $\mathcal{K}$ we denote the matrix of coefficients of the right-hand side of (137) and, without causing any misunderstanding, the formal operator in the space l of all sequences, acting as

$$(\mathcal{K}\mathbf{u})_n = b_{n-1}u_{n-1} - a_n u_n + d_{n+1}u_{n+1}.$$

In the same way, we define $\mathcal{A}$ and $\mathcal{B}$ as $(\mathcal{A}\mathbf{u})_n = -a_n u_n$ and $(\mathcal{B}\mathbf{u})_n = b_{n-1}u_{n-1} + d_{n+1}u_{n+1}$, respectively.

By $\mathcal{K}_p$ we denote the maximal realization of $\mathcal{K}$ in l_p, $p \in [1, \infty)$; that is,

$$\mathcal{K}_p\mathbf{u} = \mathcal{K}\mathbf{u}$$

on

$$D(\mathcal{K}_p) = \{\mathbf{u} \in l_p;\ \mathcal{K}\mathbf{u} \in l_p\}. \tag{138}$$

It is easy to check that he maximal operator $\mathcal{K}_p$ is closed for any $p \in [1, \infty)$. Next, define the operator A_p by restricting $\mathcal{A}$ to

$$D(A_p) = \{\mathbf{u} \in l_p;\ \mathcal{A}\mathbf{u} \in l_p\} = \{\mathbf{u} \in l_p;\ \sum_{n=0}^{\infty} a_n^p |u_n|^p < +\infty\}.$$

Again, it is standard that $(A_p, D(A_p))$ generates a semigroup of contractions in l_p. Using Theorem 50 we can prove the following result.

Theorem 62 *Assume that sequences* $\mathbf{b}$ *and* $\mathbf{d}$ *are nondecreasing and there is* $\alpha \in [0,1]$ *such that for all* n,

$$0 \leq b_n \leq \alpha a_n, \qquad 0 \leq d_{n+1} \leq (1-\alpha)a_n. \tag{139}$$

Then there is an extension K_p *of the operator* $(A_p + B_p, D(A_p))$, *where* $B_p = \mathcal{B}|_{D(A_p)}$, *which generates a positive semigroup of contractions in* l_p, $p \in (1, \infty)$.

Furthermore, we can prove that $\mathcal{B}$ is closed and thus B_p is closable. Then Theorem 48 implies

$$K_p = \overline{A_p + B_p}$$

provided $p \in (1, \infty)$. The situation in l_1 is completely different.

Corollary 15 *Let* $p = 1$. *Assume that sequences* $\mathbf{b}$ *and* $\mathbf{d}$ *are nonnegative and*

$$a_n \geq (b_n + d_n). \tag{140}$$

Then there is an extension K_1 *of the operator* $(A_1 + B_1, D(A_1))$, *where* $B_1 = \mathcal{B}|_{D(A_1)}$, *which generates a positive semigroup of contractions in* l_1.

Proof. Using the definition of $D(A_1)$ we see, from (140), that $0 \leq b_n \leq a_n$ and $0 \leq d_n \leq a_n$ for $n \in \mathbb{N}$. Hence, A_1 is well defined and condition (117) takes the form

$$\begin{aligned}\sum_{n=0}^{\infty}((A_1 + B_1)\mathbf{u})_n &= -\sum_{n=0}^{\infty} a_n u_n + \sum_{n=0}^{\infty} b_{n-1}u_{n-1} + \sum_{n=0}^{\infty} d_{n+1}u_{n+1} \\ &= -\sum_{n=0}^{\infty} a_n u_n + \sum_{n=0}^{\infty} b_n u_n + \sum_{n=0}^{\infty} d_n u_n \leq 0,\end{aligned}$$

where we used the convention $b_{-1} = d_0 = 0$. □

We have also the following result.

Theorem 63 *For any* $p \in [1, \infty)$ *we have* $K_p \subset \mathcal{K}_p$.

For $p = 1$, it is immediate consequence of Theorem 58.

5.3.2 Birth-and-death problem – honesty results

We now find whether the constructed semigroup is honest (conservative) or dishonest by means of the extension techniques of Subsection 5.2.1. In the case of matrix operators it is particularly easy to give explicit descriptions of the extended operators and related spaces. In particular, $\mathsf{E}_f = m$, the set of bounded sequences and, for instance,

$$\mathsf{L}\mathbf{u} = \left(\frac{u_n}{1 + b_n + d_n}\right)_{n\in\mathbb{N}}$$

on $\mathsf{F} = \{\mathbf{u} \in m;\ \mathsf{L}\mathbf{u} \in l_1\}$, $\mathsf{A}\mathbf{u} = ((b_n + d_n)u_n)_{n\in\mathbb{N}}$ on $D(\mathsf{A}) = \mathsf{LF}$, and similarly for the other operators and spaces introduced in Subsection 5.2.1.

Recall that by $\mathcal{K}$ we denoted the matrix of coefficients and, at the same time, the formal operator acting on m given by multiplication by $\mathcal{K}$. It is easy to see that the maximal operator $\mathcal{K}_1$ (see (138)) is precisely

$$\mathcal{K}_1 = \mathsf{K} = \mathsf{A} + \mathsf{B}. \tag{141}$$

Note too that for $\mathbf{u} \in D(\mathsf{K})$, the integral $\int_\Omega \mathsf{K}\mathbf{u}d\mu$, which plays an essential role in a number of theorems (e.g., Theorems 56 and 57), is given here by

$$\begin{aligned} & \sum_{n=0}^{\infty}(-(b_n + d_n)u_n + b_{n-1}u_{n-1} + d_{n+1}u_{n+1}) \\ = \quad & \lim_{n\to+\infty}\sum_{k=0}^{n}(-(b_k + d_k)u_k + b_{k-1}u_{k-1} + d_{k+1}u_{k+1}) \\ = \quad & \lim_{n\to+\infty}(-b_nu_n + d_{n+1}u_{n+1}), \end{aligned} \tag{142}$$

where the limit exists as $\mathbf{u} \in D(\mathsf{K})$ yields the convergence of the series.

In the theorems concerning honesty and maximality we assume, to avoid technicalities, that $b_n > 0$ for $n \geq 0$ and $d_n > 0$ for $n \geq 1$.

Theorem 64 $K = \overline{A+B}$ *if and only if*

$$\sum_{n=0}^{\infty} \frac{1}{b_n} \left(\sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{d_{n+j}}{b_{n+j}} \right) = +\infty \tag{143}$$

(where we put $\prod_{j=1}^{0} = 1$*).*

Proof. To prove honesty, we use Theorem 59. Thus, by (142) it suffices to prove that for any $\mathbf{u} \in D(\mathsf{K})_+$

$$\lim_{n\to+\infty} (-b_n u_n + d_{n+1} u_{n+1}) \geq 0,$$

where we know that the sequence above converges. If we assume the contrary, that for some $0 \leq \mathbf{u} \in D(\mathsf{K})$, then limit in (142) is negative so that there exists $b > 0$ such that

$$-b_n u_n + d_{n+1} u_{n+1} \leq -b \tag{144}$$

for all $n \geq n_0$ with large enough n_0. Using (144) as a recurrence we get

$$u_n \geq \frac{b}{b_n} \left(\sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{d_{n+j}}{b_{n+j}} \right)$$

and, if the assumption (143) is satisfied, we obtain $\sum_{n=0}^{\infty} u_n = +\infty$ which contradicts the assumption of the summability of $(u_n)_{n\in\mathbb{N}}$.

The proof of necessity is an application of Theorem 57. If the series in (143) is convergent, then, by some algebra,

$$u_n = \frac{b}{b_0} \prod_{i=0}^{n-1} \frac{b_i}{d_{i+1}} \left(\sum_{l=n}^{\infty} \prod_{i=1}^{l} \frac{d_i}{b_i} \right).$$

is summable and satisfies

$$-b = -b_n u_n + d_{n+1} u_{n+1}, \quad n \geq 0$$

so that assumption (iii) of Theorem 57 is satisfied. By construction, $\mathcal{A}\mathbf{u}+\mathcal{B}\mathbf{u} \in l_1$, so that $\mathbf{u} \in D(\mathsf{K})$. We must show that $\mathbf{g} = \mathbf{u} - (\mathcal{A}\mathbf{u} + \mathcal{B}\mathbf{u}) \geq 0$. By direct calculations, we obtain $g_0 = u_0 + b_0 u_0 - d_1 u_1 = u_0 + b$ and for $n > 0$,

$$g_n = u_n + b_n u_n + d_n u_n - b_{n-1} u_{n-1} - d_{n+1} u_{n+1} = u_n,$$

so that $0 \leq \mathbf{g} \in l_1$. Hence assumption (i) of Theorem 57 is satisfied. □

5.3.3 Universality of Dishonesty

Theorem 65 *If $(G_K(t))_{t\geq 0}$ is dishonest, that is, if*

$$\sum_{n=0}^{\infty} \frac{1}{b_n} \left(\sum_{i=0}^{\infty} \prod_{j=1}^{i} \frac{d_{n+j}}{b_{n+j}} \right) < +\infty, \tag{145}$$

then for each $\mathbf{u}_0 \in X_+$ there is $t_0 \geq 0$ such that $\|G_K(t)\mathbf{u}_0\| < \|\mathbf{u}_0\|$ for all $t > t_0$.

Proof. By Theorem 55, $(G_K(t))_{t\geq 0}$ is dishonest if and only if the functional β_λ, defined in Theorem 54, is not identically zero. The defect function along the trajectory originating at $\mathbf{u}_0$, which in our case is given by $\eta_{\mathbf{u}_0}(t) = \|G_K(t)\mathbf{u}_0\| - \|\mathbf{u}_0\|$, is related to β_λ by

$$\int_0^{\infty} e^{-\lambda t} \eta_{\mathbf{u}_0}(t) dt = -\frac{1}{\lambda} <\beta_\lambda, \mathbf{u}_0> .$$

Clearly, λ is inessential. Putting $\beta_\lambda = \beta = (\beta_n)_{n\in\mathbb{N}}$ with $\beta_n \geq 0$, we see that for universality of dishonesty we must have $\beta_n > 0$ for any $n \geq 0$. On the other hand, by (126), β_λ is an eigenvector of $(BR(\lambda, A))^*$. Any eigenvector $(\phi_n)_{n\in\mathbb{N}}$ satisfies

$$\begin{aligned} \frac{b_0}{1+b_0}\phi_1 &= \phi_0, \\ &\vdots \quad , \end{aligned}$$

$$\begin{aligned}\frac{d_n}{1+b_n+d_n}\phi_{n-1}+\frac{b_n}{1+b_n+d_n}\phi_{n+1} &= \phi_n,\\ &\vdots\ ,\end{aligned}$$

and, because $b_0/(1+b_0) < 1$, we have $\phi_1 > \phi_0$. Rearranging the terms in nth equation,

$$\phi_{n+1} = \left(1+\frac{1}{b_n}\right)\phi_n + \frac{d_n}{b_n}(\phi_n - \phi_{n-1}).$$

Hence $\phi_{n+1} > \phi_n$ whenever $\phi_n \geq \phi_{n-1}$ we end the proof by induction. □

5.3.4 Maximality of the Generator

Let us recall that the relation between the generator K and its extensions K and $\mathcal{K}$ is given in (141). In particular, K is the maximal operator.

Proposition 12 *If $(G_K(t))_{t\geq 0}$ is a substochastic semigroup generated by K and for some $0 \leq \mathbf{h} \in D(K)$,*

$$\int_\Omega \mathsf{K}\mathbf{h}d\mu > 0, \tag{146}$$

then $K \neq \mathsf{K}$; that is, the generator is not maximal.

Conversely, assume that if $0 \neq \mathbf{u} \in l$ solves the formal equation

$$\mathcal{K}\mathbf{u} = \lambda\mathbf{u}, \quad \lambda > 0, \tag{147}$$

then either $\mathbf{u} \geq 0$ or $\mathbf{u} \leq 0$, and

$$\int_\Omega \mathsf{K}\mathbf{h}d\mu = 0, \tag{148}$$

for any $\mathbf{h} \in D(\mathsf{K})$. Then $\mathsf{K} = K$; that is, the generator is the maximal operator.

Proof. It follows that if $\mathbf{h} \in D(K)$, then $\int_\Omega K\mathbf{h}d\mu = 0$. Because $K \subset \mathsf{K}$, (146) shows that $\mathbf{h} \notin D(K)$.

If $\mathsf{K} \neq K$ then, by Corollary 13, we have $N(\lambda I - \mathsf{K})_+ \neq \emptyset$. Because (147) is linear, then the assumption ascertains the existence of $0 \neq \mathbf{h} \in N(\lambda I - \mathsf{K})_+$ and for such an $\mathbf{h}$

$$\int\limits_\Omega \mathsf{K}\mathbf{h}d\mu = \lambda \int\limits_\Omega \mathbf{h}d\mu \neq 0, \tag{149}$$

contradicting (148). □

To be able to use this result, we have the following lemma.

Lemma 5 *Let $\lambda > 0$ be fixed. Any solution to (147) is either nonnegative or nonpositive.*

On the basis of the above lemma, we have:

Theorem 66 $K \neq \mathsf{K}$ *if and only if*

$$\sum_{n=1}^{\infty} \frac{1}{d_n} \prod_{j=1}^{n-1} \frac{b_j}{d_j} \left(\sum_{i=0}^{n-1} \prod_{j=1}^{i} \frac{d_i}{b_i} \right) < +\infty, \tag{150}$$

where, as before, $\prod_{j=1}^{0} = 1$.

Proof. By Lemma 5 and Proposition 12, $K \neq \mathsf{K}$ if and only if for each $0 \leq (u_n)_{n\in\mathbb{N}} \in l_1$, such that $(-(b_n + d_n)u_n + b_{n-1}u_{n-1} + d_{n+1}u_{n+1_n})_{n\in\mathbb{N}} \in l_1$, we have

$$I = \sum_{n=0}^{\infty} (-(b_n + d_n)u_n + b_{n-1}u_{n-1} + d_{n+1}u_{n+1}) > 0.$$

and, similarly to the proof of Theorem 64 and (144) we need to investigate the behaviour of the sequence $(r_n)_{n\in\mathbb{N}}$ defined as

$$r_n = -b_n u_n + d_{n+1}u_{n+1}, \quad n \geq 0, \tag{151}$$

or,solving for u_n, for $n \geq 1$,

$$u_n = \frac{1}{d_n}\sum_{i=0}^{n-1}\left(r_i \prod_{j=1}^{n-1-i}\frac{b_{n-j}}{d_{n-j}}\right) + \frac{u_0 b_0}{d_n}\prod_{j=1}^{n-1}\frac{b_j}{d_j}. \tag{152}$$

If $K \neq \mathsf{K}$, then there is a nonnegative $(u_n)_{n\in\mathbb{N}} \in l_1$ for which $I = \lim_{n\to\infty} r_n > 0$ and, by some algebra, it is enough to consider a nonnegative sequence $(u_n)_{n\in\mathbb{N}} \in D(\mathsf{K})$ with the associated sequence $(r_n)_{n\in\mathbb{N}}$ satisfying $\inf_{n\in\mathbb{N}} r_n = r > 0$. Then it can be proved that the series in (150) is convergent.

To prove the converse, define u_n by (151) with arbitrary $(r_n)_{n\in\mathbb{N}}$ converging to $I > 0$ (e.g., we may take $r_n = r$ for all n for a constant positive r). By (150) $(u_n)_{n\in\mathbb{N}} \in l_1$, so that $(u_n)_{n\in\mathbb{N}} \in D(\mathsf{K})$ and because $I > 0$, the thesis follows by (146). □

5.3.5 Examples

We provide a few examples showing that all possible cases of relations between the generator and maximal and minimal operators can be realized.

Proposition 13 *If both sequences $(b_n^{-1})_{n\in\mathbb{N}}, (d_n^{-1})_{n\in\mathbb{N}} \notin l_1$, then $K = \overline{A+B} = \mathsf{K}$. In particular, this is true for the standard birth-and-death problem of population theory where the coefficients are affine functions of n.*

Proof. Expanding (150) we get, for a fixed n,

$$\frac{1}{d_n}\left(1 + \frac{b_{n-1}}{d_{n-1}} + \cdots + \frac{b_{n-1}\ldots b_1}{d_{n-1}\ldots d_1}\right) \geq \frac{1}{d_n}.$$

Similarly, expanding (143), we get

$$\frac{1}{b_n}\left(1 + \frac{d_{n+1}}{b_{n+1}}\cdots +\right)$$

which gives divergence of both series. $\square$

The proofs of the following results are obtained in a similar way.

Proposition 14 *If* $(d_n^{-1})_{n\in\mathbb{N}} \in l_1$ *and*

$$\lim_{n\to\infty} \frac{b_n}{d_n} = q < 1, \tag{153}$$

then $K = \overline{A+B} \neq \mathsf{K}$.

Proposition 15 *If the sequence* $(d_n)_{n\in\mathbb{N}}$ *is of polynomial growth:* $d_n = O(n^\beta)$ *for some* β *as* $n \to \infty$, $(b_n^{-1})_{n\in\mathbb{N}} \in l_1$ *and*

$$\lim_{n\to\infty} \frac{b_n}{d_n} = q > 1, \tag{154}$$

then $\overline{A+B} \subsetneq K = \mathsf{K}$.

Proposition 16 *There are sequences* $(b_n)_{n\in\mathbb{N}}$ *and* $(d_n)_{n\in\mathbb{N}}$ *for which* $\overline{A+B} \subsetneq K \subsetneq \mathsf{K}$.

Proof. Take $b_n = 2\cdot 3^n$ and $d_n = 3^n$. $\square$

5.4 Chaos in population theory

We consider a population of cancer cells characterized by different levels of drug resistance. The cells belonging to 0–th subpopulation are sensitive to antineoplastic drugs. The cells of n–th subpopulation, $n > 0$, are resistant with the level of resistance increasing with n. Each subpopulation contains cells characterized by a number of copies of a drug resistance gene. The more copies of the gene exist, the more resistant the cell, with the understanding that it can survive under higher concentration of the drug. Since the number of gene copies can be very large, we use a model with an infinite number of cell subpopulations. We consider a gene amplification – deamplification process characterized by two components: the conservative one and the proliferative one.

The conservative component of the process describes the mutations of cells modelled as in standard birth-and-death process. The proliferative component of the process is related to the assumption that the moment of death represents the moment of cell division and that the average life–span is given by the coefficient λ_n for the n–subpopulation ($n \geq 0$). This model leads to the infinite system of ordinary differential equations

$$\begin{aligned} \frac{df_0}{dt} &= -a_0 f_0 + d_1 f_1, \\ \vdots \quad & \quad \vdots \\ \frac{df_n}{dt} &= -a_n f_n + b_{n-1} f_{n-1} + d_{n+1} f_{n+1}, \; n \geq 1, \end{aligned} \tag{155}$$

where we denoted $a_0 = -\lambda_0 + b_0$ and $a_n = -\lambda_n + b_n + d_n$ for $n \in \mathbb{N}$. We denote by $\mathbf{f} = \{f_n(t)\}_{n\geq 0}$ the distribution function and by L the infinite matrix of the coefficients on the right-hand side of (155). The proper Banach space for the process defined by Eq. (155) is l^1, where the norm

$$\|\mathbf{f}\|_1 = \sum_{n=0}^{\infty} f_n \,, \tag{156}$$

of any element $\mathbf{f}$ in the positive cone l^1_+: $l^1_+ = \{\mathbf{f} \in l^1 \;;\quad f_n \geq 0 \quad n = 0, 1, 2 \ldots\}$ represents the total number of cells. For the sake of completeness we shall consider also the Banach spaces l^p, $1 \leq p < \infty$, and c_0 (the space sequences converging to 0), with natural norms.

In [13], Eqn. (155) is considered under the assumption that the coefficients a_n, b_n (for $n \in \mathbb{N}_0$), d_n (for $n \in \mathbb{N}$) are nonnegative and

(A1) for some $a \geq 0$, $a_n = a + \alpha_n, \; n \in \mathbb{N}_0$, with $\lim\limits_{n\to\infty} \alpha_n = 0$, (A2) for some $d > 0$ $\lim_{n\to\infty} d_n = d$

(A3) $\limsup_{n\to\infty} b_n = 0$.

Let $\mathcal{L}_p$, $p \in [1, \infty[\cup \{0\}$ denote the operator realization of L in l^p and c_0, respectively. The operators $\mathcal{L}_p$ are bounded, hence they generate dynamical systems $(T_p(t))_{t\geq 0}$ in l^p and c_0, respectively.

Theorem 67 *Let the assumptions (A1), (A2) and (A3) be satisfied. There is $q > 0$ such that if $|\alpha_n| \leq dq^{n+1}$, $|b_n d_{n-1}| \leq d^2 q^{2n+4}$ and $a < d$, then the semigroup generated by $\mathcal{L}_p$ is chaotic in any l^p, $1 \leq p < \infty$, and in c_0.*

Consider the system transposed to (155)

$$\begin{aligned} \frac{df_0}{dt} &= -a_0 f_0 + b_0 f_1, \\ \frac{df_n}{dt} &= -a_n f_n + d_n f_{n-1} + b_n f_{n+1}, \quad n \in \mathbb{N}. \end{aligned} \tag{157}$$

Using the fact that $c_0^* = l^1$ and $(l^p)^* = l^r$, $1/p + 1/r = 1$, by Theorem 38, if (157) was chaotic in any subspace, then the codimension of the span of all eigenvectors of the operator in (155) in respective space would be finite. Since this is not true, we have

Corollary 16 *Suppose that the sequences (a_n), (b_n) and (d_n) are as in Theorem 67. Then the semigroup generated by (157) is chaotic in no subspace of l^p, $1 \leq p < \infty$, or of c_0.*

Theorem 67 ensures the topological chaos for large deamplification ("death") rates and small amplification ("birth") rates, i.e. for the process which is subcritical. On the contrary, chaos will not appear in processes with small deamplification rates and possibly large amplification rates. The assumptions of Theorem 67 are often not realistic – in most standard applications the coefficients depend in an affine way on n. This creates numerous problems starting from the generation of the semigroup through the construction of eigenvectors to their density in l^p.

We adopt the following assumption.

Assumption AC *There exists* $N_0 \geq 1$ *such that*

$$\left.\begin{array}{l} a_n = an + \alpha\,, \\ d_{n+1} = dn + \delta\,, \\ b_{n-1} = bn + \beta\,, \quad n \geq N_0, \end{array}\right\} \tag{158}$$

with $a = -(b + d)\,, b\,,\ d \geq 0\,, \alpha, \beta, \delta \in \mathbb{R}$.

In this case the proliferation rate does not depend on n for large n, and equals

$$\gamma = \alpha + \beta + \delta + b - d.$$

Recall that L is the infinite matrix of coefficients. We define

$$D(\mathcal{L}_{max}) = \{f \in l^p;\ Lf \in l^p\}$$

and $\mathcal{L}_{max} = L|_{D(\mathcal{L}_{max})}$.

Theorem 68 *[18] Suppose that Assumption* AC *is satisfied and* $p \in [1; +\infty)$. *Then* $\mathcal{L}_{max}$ *is a unique realization of* L *that generates a* C_0*-semigroup in* l^p.

The importance of the identification of $\mathcal{L}_{max}$ as the generator stems from the fact that l^p-solutions of the infinite system

$$\begin{array}{rcl} \lambda f_0 &=& -a_0 f_0 + d_1 f_1, \\ \vdots && \vdots \\ \lambda f_n &=& -a_n f_n + b_{n-1} f_{n-1} + d_{n+1} f_{n+1},\ n \geq 1, \end{array} \tag{159}$$

are the eigenvectors of the generator.

Proposition 17 *Let Assumption AC be satisfied,* $d > b$*,* $N_0' := \max\{n \geq 0 : d_n = 0\}$*. For any* $\lambda \in \mathbb{C}$ *there exists a unique sequence* $f(\lambda) = (f_n(\lambda))_{n\geq 0}$ *satisfying (159) and the initial conditions* $f_n(\lambda) = 0$ *for* $n < N_0'$*,* $f_{N_0'}(\lambda) = 1$*. Moreover,*

(i) $f_n(\lambda)$ *is a polynomial in* λ *of degree* $n - N_0'$ *for* $n \geq N_0'$*;*

(ii) *for any* $\lambda_0 \in \mathbb{C}$ *and* $\epsilon > 0$*, there exists* $K > 0$ *such that if* $|\lambda - \lambda_0| < \epsilon$ *and* $n \geq N_0' + 1$*, then*

$$|f_n(\lambda)| \leq Kn^{-\frac{\alpha+\beta+\delta-\Re\lambda}{d-b}}. \tag{160}$$

Denote by $\Pi_p(b, d, \alpha, \beta, \delta)$ the open left half-plane defined by

$$\Pi_p(b, d, \alpha, \beta, \delta) = \{\lambda \in \mathbb{C} \;:\; \Re\lambda < \gamma_p\}, \tag{161}$$

where

$$\gamma_p = \alpha + \beta + \delta - \frac{d-b}{p}. \tag{162}$$

Corollary 17 *Consider the operator* $\mathcal{L}_{max}$ *acting in the space* l^p*,* $1 \leq p < \infty$*. If Assumption* AC *holds with* $d > b$*, then* $\Pi_p(b, d, \alpha, \beta, \delta) \subset \sigma_p(\mathcal{L}_{max})$*. Moreover, for any* $\lambda \in \Pi_p(b, d, \alpha, \beta, \delta)$ *the sequence* $f(\lambda)$*, given by Proposition 17, is an eigenvector of* $\mathcal{L}_{max}$ *for the eigenvalue* λ*, and the vector-valued function* $\Pi_p(b, d, \alpha, \beta, \delta) \ni \lambda \to f(\lambda) \in l^p$ *is analytic.*

Theorem 69 *Suppose that* $1 \leq p < \infty$ *and that Assumption* AC *holds with* $d > b$ *and* $\gamma_p > 0$*. Then the* C_0*-semigroup generated by* $\mathcal{L}_{max}$ *in* l^p *is sub-chaotic.*

Theorem 70 *Suppose that Assumption AC is satisfied,* $p \in [1; +\infty)$*, and either of two cases hold:*

(i) $b > d$*,*

(ii) $d_{m_0} = 0$*,*

for some $m_0 \geq 1$*. Then the* C_0*-semigroup generated by* $\mathcal{L}_{max}$ *is not topologically chaotic.*

5.4.1 Meaning of chaos

In most biological applications only non-negative solutions make sense (the solution should stay in the non–negative cone of l_1) and it is only fair to note that the chaotic properties discussed here cannot occur for such solutions. In fact, for systems with strictly positive proliferation, the l_1 norm of a solution may only grow and hence the solution cannot wander.

On the other hand, as we are dealing with linear systems we may wish to consider the differences between two physical (i.e. non–negative) solutions and such a difference certainly need not be non–negative. In fact, we have

Proposition 18 *Let X be a Banach lattice. If $(G(t))_{t\geq 0}$ is chaotic (subchaotic) in X, then for any $\epsilon > 0$ there exist $x_1, x_2 \in X_+$ such that $\|x_1 - x_2\| < \epsilon$ and $t \to G(t)x_1 - G(t)x_2$ is dense in the space of chaoticity of $(G(t))_{t\geq 0}$.*

Proof. Let X_{ch} be a space of chaoticity of $(G(t))_{t\geq 0}$. There is a dense trajectory in X_{ch} so, in particular, for any $\epsilon > 0$ there is $z \in X_{ch}$ such that $\|z\| < \epsilon$ and $\{G(t)z\}_{t\geq 0}$ is dense in X_{ch} (since any tail of a dense trajectory is dense). Since the positive cone in a Banach lattice is generating, there are $x_1, x_2 \in X_+$ such that $z = x_1 - x_2$. From linearity, $G(t)z = G(t)x_1 - G(t)x_2$. □

6 Asynchronous growth

The analysis in Subsection 4.4 gives some information about how fast a semigroup can grow but does not yield any clue as to whether there are any long term patterns of the behaviour of the semigroup. Some such patterns were discussed in Subsection 1.2. In many cases, as in the finite dimensional case, such patterns are associated with the eigenvalues with largest real value. This section is devoted to existence of such eigenvalues.

The first step in this direction is to ensure existence decomposition of the spectrum of the semigroup into isolated point part and the rest which, hopefully, will have real parts smaller that the point part. This of course occurs if the semigroup is compact (or even eventually compact) and, more generally, when its essential spectrum radius is strictly smaller that the spectral radius.

6.1 Essential growth bound

The concept of essential spectrum provides more insight into the long time behaviour of semigroups. Since $r_\Phi(G(t))$ is defined through the norm in the quotient space, we can define the Fredholm growth rate of the semigroup using the Fredholm norm $\|\cdot\|_\Phi$ and prove in the same way as for the growth rate that

$$\omega_\Phi(A) = \lim_{t\to\infty} \frac{1}{t} \log \|G(t)\|_\phi \tag{163}$$

and

$$e^{t\omega_\Phi(A)} = r_\Phi(G(t))$$

However, using (29) we can replace r_Φ by r_{ess} and call ω_Φ the essential growth rate and denote it by ω_e. This shows that at the level of spectral radii and growth bounds the distinction between Fredholm and essential spectra (and thus between approaches of [38, 26] and [21, 5]) disappears.

Clearly, $\omega_e(G) \leq \omega_0(G)$. If $\omega_e(G) < \omega_0(G)$, then there is an eigenvalue of $(G(t))_{t\geq 0}$ satisfying

$$|\lambda| = e^{\omega_0(G)t}$$

hence, by Theorem 23(2), there is $\lambda_1 \in \sigma_p(A)$ such that $\Re\lambda_1 = \omega_0$. Since $s(A) \leq \omega_0(G)$ we obtain the important result

$$\omega_0(G) = \max\{\omega_e(G), s(A)\} \tag{164}$$

We shall look more closely into implications of $\omega_e(G) < \omega_0(G)$. In this case $s(A) = \omega_0(G)$.

Theorem 71 *Suppose* $\omega_e(G) < \omega_0(G)$. *Then* $\sigma_{per,s(A)} \neq \emptyset$ *and is finite. Moreover,* X *can be decomposed in a unique way into a sum* $N \bigoplus S$ *of two closed* $G(t)$*-invariant subspaces with one of them (say* N*) of finite dimension. Furthermore,* $\sigma(A|_N) = \sigma_{per,s(A)}$ *and* $\omega_0(G|_S) < \omega_0(G)$

Proof. First we note that, by (164) and the definition of $s(A)$, for any $\gamma \in (\omega_e(G), \omega_0(G)]$ there is $\lambda \in \sigma(A) \setminus \sigma_e(A)$ (and also $\lambda \in \sigma(A) \setminus \sigma_\Phi(A)$ with $\gamma \leq \Re\lambda \leq \omega_0(G)$. We will show that for any $\gamma > \omega_e(G)$ there are only finitely many λ satisfying $\Re\lambda \geq \gamma$. To the contrary, assume that there is an infinite sequence $(\lambda_n)_{n\in\mathbb{N}}$ satisfying $\omega_e(G) < \gamma \leq \Re\lambda_n \leq \omega_0(G)$. Since each $\lambda_n \in \sigma_p(A)$ thus, by the Spectral Mapping Theorem for the point spectrum, $\mu_n := e^{t\lambda_n} \in \sigma_p(G(t))$ for any $t \geq 0$. Assume that for some $t_0 > 0$ the sequence $(\mu_n)_{n\in\mathbb{N}}$ has an accumulation point. By the definition of the essential spectrum, this implies $r_e(G(t_0)) \geq e^{\gamma t_0}$ but then $\omega_e(G) \geq \gamma$, which is a contradiction. So, none of the sequences $(e^{t\lambda_n})_{n\in\mathbb{N}}$ has an accumulation point hence, being bounded, must be finite. Fix again $t > 0$. There may be an infinite sequence of λ_n (denoted again by $(\lambda_n)_{n\in\mathbb{N}}$) satisfying $\mu = e^{t\lambda_n}$ for each n, see (70). The eigenspaces of A corresponding to distinct λ_n are linearly independent. But then their direct sum is infinite dimensional and corresponds to the eigenspace of $G(t)$ corresponding to μ contradicting again the definition of the essential spectrum. The first two statements of the lemma follow now by specifying $\gamma = s(A)$. The other two can be obtained by defining N as the sum of $Ker_\infty(\lambda I - A)$ over $\lambda \in \sigma_{per,s(A)}$ and S as the intersection of $Im\ ((\lambda I - A)^k)$ over $k \in \mathbb{N}$ and $\lambda \in \sigma_{per,s(A)}$ (or equivalently, using the fact that $\sigma_{per,s(A)}$ is isolated in $\sigma(A)$ and compact, by taking the spectral projection corresponding to $\sigma_{per,s(A)}$ and its complement. $\square$

Remark 11 Using the terminology of Subsection 1.2, we see that EAEG holds if $\omega_e(G) < \omega_0(G)$. Then, MAEG holds if, moreover, $\sigma_{per,s(A)}$ consists of a single eigenvalue. Finally, AEG holds if, in addition, this eigenvalue has multiplicity one. In fact, these conditions are necessary and sufficient, [5].

It seems that to make a progress towards MAEG and AEG, we have to assume that the semigroup at hand is positive.

6.2 Peripheral spectrum of positive semigroups

The main result of this subsection is

Theorem 72 *If $(G(t))_{t\geq 0}$ is a positive semigroup on a Banach lattice X generated by A such that $s(A) > -\infty$ is a pole of the resolvent $R(\lambda, A)$. Then $\sigma_{per,s(A)}$ is additively cyclic.*

The proof of this result is quite technical and draws on numerous results from the theory of positive operators on Banach lattices and we shall refrain from giving it in detail. It is, however, worthwhile to discuss a few salient point of the proof which use the relations between Banach lattices and the space of continuous function $C(K)$, given by the Kakutani-Krein theorem (Theorem 13).

We recall the signum operator S_u (see Example 17)) and define

$$u^k = S_u^k |u|, \qquad k \in \mathbb{Z}$$

where $u^{-1} := \bar{u}$.

The crucial result is the following lemma.

Lemma 6 *If T and R are two bounded operators satisfying $|Rx| \leq T|x|$, $x \in X$, and $Ru = u$ and $T|u| = |u|$ for $u \in X$ such that $|u|$ is a quasi-interior point (see Lemma 2), then $T = S_u^{-1} R S_u$.*

Proof. The proof uses the fact that, as u is quasi-interior, X can be identified with a space of continuous functions. Details are given in [21, Lemmas 8.8-9].

To illustrate this result we shall discuss two other results which are related to Theorem 72 but which are relatively simpler.

Proposition 19 *If L is a positive operator on a Banach lattice X (and thus bounded) and suppose that $Lu = \alpha u$ and $L|u| = |u|$ for some $u \in X$ and $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Then, for every $k \in \mathbb{Z}$ we have $Lu^k = \alpha^k u^k$.*

Proof. Since L is bounded, it leaves $\overline{X_u}$ invariant. Define $T = L|_{\overline{X_u}}$ and $R = \alpha^{-1}L$. Then $Tu = Ru$, $T|u| = L|u| = |u|$, and $|Tx| = |Lx| = |Rx|$ for $x \in \overline{X_{|u|}}$. Hence, $T = S_u^{-1}RS_u = \alpha^{-1}S_u^{-1}TS_u$. Iteration yields $T = \alpha^{-k}S_u^{-k}TS_u^k$. Hence

$$Tu^{-k} = TS_u^{-k}|u| = \alpha^{-k}S_u^{-k}T|u| = \alpha^{-k}S_u^{-k}|u| = \alpha^{-k}u^{-k},\ k \in \mathbb{Z},$$

which gives the thesis. □

Corollary 18 *Let $(G(t))_{t\geq 0}$ be a positive semigroup on a Banach lattice generated by A, and suppose that for some $u \in X$ and $\alpha, \beta \in \mathbb{R}$ we have*

$$Au = (\alpha + i\beta)u, \quad A|u| = \alpha|u| \tag{165}$$

Then

$$Au^n = (\alpha + in\beta)u^n. \quad n \in \mathbb{Z} \tag{166}$$

Furthermore, if $|u|$ is a quasi-interior point of X, then $S_uD(A) = D(A)$ and $A + i\beta I = S_u^{-1}AS_u$.

Proof. We may assume $\alpha = 0$. Eq. (165) implies $G(t)|u| = |u|$ and $G(t)u = e^{i\beta t}u$, $t \geq 0$, by Theorem 23 (2). Thus, by Proposition 19, we have $G(t)u^n = e^{in\beta}u^n$ which, again by Theorem 23, is equivalent to $Au^n = i\beta nu^n$.

If $|u|$ is a quasi-interior point of X_+ then in the proof of Proposition 19 we have $X = \overline{X_{|u|}}$ so that $T = L = G(t)$ and $R = e^{-i\beta t}G(t)$. Thus yields $e^{i\beta t}G(t) = S_u^{-1}G(t)S_u$ for all $t \geq 0$ which implies $S_uD(A) = D(A)$ (by $S_ue^{i\beta t}G(t) = G(t)S_u$) and $A + i\beta I = S_u^{-1}AS_u$. □

The above result allows to give a simple proof a theorem on cyclicity of point spectrum in Banach lattices with strictly monotonic norm: $0 \leq f < g$ implies $\|f\| < \|g\|$. In particular, the spaces L_p have this property.

Theorem 73 *Suppose X is a Banach lattice with strictly monotone norm. If $(G(t))_{t\geq 0}$ is a positive contractive semigroup with $s(A) = 0$, then $\sigma_{per,s(A)} \cap \sigma_p(A)$ is imaginary additively cyclic.*

Proof. Suppose that $Au = i\beta u$ for some $\beta \in \mathbb{R}, u \in X$. Then $G(t)u = e^{i\beta t}u$ and $|u| \leq G(t)|u|$. Hence $\|u\| \leq \|G(t)|u|\| \leq \|u\|$ since $(G(t))_{t\geq 0}$ is contractive. Hence $\|G(t)|u|\| = \|u\|$ and, by strict monotonicity of the norm, $G(t)|u| = |0|$, which implies $A|u| = 0$. Using Corollary 18 we obtain the thesis. □

Corollary 19 *If assumptions of Theorem 73 are satisfied and $\omega_e(G) < \omega_0(G)$, then $\sigma_{per,s(A)} = \{s(A)\}$. Thus, $(G(t))_{t\geq 0}$ has MAEG.*

Proof. Since $\omega_e(G) < \omega_0(G)$, the peripheral spectrum $\sigma_{per,s(A)}$ is the point spectrum and, by Theorem 71, must be finite and non-empty. The only way for it to be additively cyclic is to consist of one point. □

In general case this result will follow from Theorem 72 whose proof is much more involved. After this interlude let us return to the discussion of Theorem 72.

Proof. We may assume $s(A) = 0$. Let us start with the case that $s(A) = 0$ is a pole of $R(\lambda, A)$ of order 1. Assume that $i\nu \in \sigma_{per,s(A)}$. Using the Spectral

Mapping Theorem for the resolvent and (13) we find that for any $\lambda \in \rho(A)$ we have

$$(\lambda - i\nu)^{-1} \in \partial\sigma(R(\lambda, A)) \subset \sigma_a(R(\lambda, A)).$$

The problem we face is that $(\lambda - i\nu)^{-1}$ is an approximate eigenvalue. Using considerations of Paragraph 2.2.2, we embed the problem into the ultrapower $\hat{X}$ so that the approximate eigenvalues become eigenvalues of the extended operator. The snag is that the extended resolvent $\hat{R}(\lambda) := \widehat{R(\lambda, A)}$ is no longer the resolvent of a densely defined operator if A is unbounded. It is however, a pseudoresolvent with the same domain of definition, which satisfies the same estimate as the resolvent of the generator of a positive semigroup (Theorem 42 (iii)):

$$|\hat{R}(\lambda)\hat{x}| \leq \hat{R}(\Re\lambda)|\hat{x}|, \qquad x \in X, \Re\lambda > 0. \tag{167}$$

By Theorem 5, $(\lambda - i\nu)^{-1}$ is an eigenvalue and a pole of $\hat{R}(\lambda)$ of the same order. Fixing $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$, there is $\hat{u}$ satisfying

$$\hat{R}(\lambda)\hat{u} = (\lambda - i\nu)^{-1}\hat{u}$$

and, by properties of pseudo-resolvents, the above equation is satisfied for all λ with $\Re\lambda > 0$. Summarizing, we have

$$\begin{aligned} \hat{R}(\lambda)\hat{u} &= (\lambda - i\nu)^{-1}\hat{u}, \quad \Re\lambda > 0, \\ \lambda\hat{R}(\lambda)|\hat{u}| &\geq |\hat{u}|, \quad \lambda > 0, \end{aligned} \tag{168}$$

where the second relation follows from the first and (167). Now, if $\hat{u}$ were a quasi-interior point and if we had equality in the second relation, then we would be in the position to use Lemma 6 with $R = \hat{R}(\lambda)$ and $T = R(\Re\lambda)$ to get

$$\lambda\hat{R}(\lambda) = S_u^{-1}\lambda\hat{R}(\lambda + i\nu)S_u$$

for all λ, $\Re\lambda > 0$ (by analytic continuation).

Replacing λ by $\lambda + i\nu$ on the right hand side and iterating we obtain

$$\lambda\hat{R}(\lambda) = S_u^{-k}\lambda\hat{R}(\lambda + ik\nu)S_u^k$$

and, applying this to $|\hat{u}|$, we obtain

$$
\begin{aligned}
S_u^{-k}\lambda\hat{R}(\lambda+ik\nu)S_u^k|u| = S_u^{-k}\lambda\hat{R}(\lambda+ik\nu)u^k \\
= \lambda\hat{R}(\lambda)|u| = |u|, \quad \Re\lambda > 0, k \in \mathbb{Z},
\end{aligned}
$$

or

$$\lambda\hat{R}(\lambda+ik\nu)u^k = u^k$$

so that the peripheral spectrum of $\hat{R}(\lambda)$ and thus of the generator would be cyclic.

As we noted, there are two snags. One is that $\hat{u}$ is not necessarily a quasi-interior point. This, however, can be remedied by restricting considerations to the closed ideal $\overline{X_u}$. The second, that we have inequality rather than equality in the second relation of (168). This is a more serious problem, though a solution is similar if much more technically involved. Since $|\hat{u}| > 0$, there is a positive functional such that $< |\hat{u}|, \phi >$.

As we noted, since $\lambda = 0$ is a first order pole of $R(\lambda, A)$, it is also a first order pole for $\hat{R}(\lambda)$, which means that $\{\lambda\hat{R}(\lambda);\ 0 < \lambda < 1\}$ and so $\{\lambda\hat{R}(\lambda)^*\phi;\ 0 < \lambda < 1\}$ are norm bounded. Let $(\lambda_n)_{n\in\mathbb{N}}$ converge to zero. Then $\{\lambda_n\hat{R}(\lambda_n)^*\phi\}_n$ is weak-$*$ relatively compact and thus have a weak-$*$ accumulation point, say, ϕ. It can be proved that

$$< x, \lambda\hat{R}(\lambda)\phi >=< x, \phi >, \quad x \in X,$$

and by properties of pseudo-resolvents, this extends to all λ with $\Re\lambda > 0$. In other words, $\phi = \lambda\hat{R}(\lambda)\phi$.

Furthermore,

$$< |\hat{u}|, \phi ><< |\hat{u}|, \phi >\leq< \lambda\hat{R}(\lambda)|\hat{u}|, \phi >=< |\hat{u}|, \lambda\hat{R}(\lambda)^*\phi >$$

and, since $< |\hat{u}|, \phi >$ is independent of λ, we obtain $< |\hat{u}|, \phi >$.

Next, for arbitrary $\hat{x} \in \hat{X}$ we have

$$
\begin{aligned}
< |\hat{R}(\lambda)\hat{x}|, \phi > &\leq < \hat{R}(\Re\lambda)|\hat{x}|, \phi > \\
=< |\hat{x}|, \hat{R}(\Re\lambda)^*\phi > &= \Re\lambda^{-1} < |\hat{x}, \phi >
\end{aligned}
$$

which means that the ideal $I = \{\hat{x} \in \hat{X};\ < |x|, \phi >= 0\}$ is $\hat{R}(\lambda)$ invariant. Denote by $\tilde{R}(\lambda)$ the canonical image of $\hat{R}(\lambda)$ in the quotient space $\hat{X}/I$; in the same way $\tilde{x}$ denotes the canonical image of $\hat{x}$. Then, since the canonical injection is a lattice homeomorphism (see Example 9), we obtain

$$|\tilde{R}(\lambda)\tilde{x}| \leq \tilde{R}(\Re\lambda)|\tilde{x}|, \qquad \tilde{x} \in \hat{X}/I, \Re\lambda > 0$$

and, obviously

$$\tilde{R}(\lambda)\tilde{u} = (\lambda - i\nu)^{-1}\tilde{u}.$$

Moreover, $< |\hat{u}|, \phi >> 0$ which means $\tilde{u} \neq 0$ in X/I and

$$< \hat{R}(\lambda)|\hat{u}| - |\hat{u}|, \phi >=$$

which is the same as $\tilde{R}(\lambda)|\tilde{u}| = |\tilde{u}|$ which means that we can use the argument above to obtain

$$\lambda\tilde{R}(\lambda + ik\nu)\tilde{u}^k = \tilde{u}^k, \quad k \in \mathbb{Z}, \Re\lambda > 0.$$

This means that $ik\nu \in \sigma(A)$ and the statement is proved for $s(A)$ being a first order pole.

Let $\lambda = 0$ be a pole of order $p \geq 2$ and suppose we can prove the result for poles of any order less than p. Define

$$Q = \lim_{\lambda \to 0} \lambda^p R(\lambda, A).$$

Then $Q > 0$ is a bounded operator which, moreover satisfies $QA = AQ = 0$. The ideal

$$I = \{x \in X;\ Q|x| = 0\}$$

is $(G(t))_{t\geq 0}$-invariant and thus we can consider the problem in the quotient lattice $\tilde{X} = X/I$ on which the canonical image $\tilde{Q}$ is zero. This implies that $\lambda = 0$ is a pole of order less than p of the canonical image of the resolvent and the prove is finished by induction. □

As we said earlier, Theorem 72 yields Corollary 19 in full generality.

Corollary 20 *Let $(G(t))_{t\geq 0}$ be a positive semigroup satisfying $\omega_e(G) < \omega_0(G)$. Then $\sigma_{per,s(A)} = \{s(A)\}$. Thus, $(G(t))_{t\geq 0}$ has MAEG.*

Note that assumption $\omega_e(G) < \omega_0(G)$ ensures that $s(A)$ is a pole of $R(\lambda, A)$.

The next step towards AEG requires irreducibility of the semigroup.

6.2.1 Peripheral spectrum of irreducible semigroups

Let $(G(t))_{t\geq 0}$ be a positive semigroup on a Banach lattice X generated by A. Recall that a closed ideal $E \subset X$ is said to be invariant under $(G(t))_{t\geq 0}$ (or $\{G(t)\}$-invariant) if it is $G(t)$-invariant for any $t \geq 0$. The semigroup $(G(t))_{t\geq 0}$ is called irreducible if $\{0\}$ and X are the only $\{G(t)\}$-invariant ideals of X. Furthermore, $(G(t))_{t\geq 0}$ is called strongly irreducible if $G(t)$ is a strongly irreducible operator for any $t \geq 0$ (that is, if $G(t)u$ is a quasi-interior point for any $0 < u \in X$). Clearly, strongly irreducible semigroup is irreducible (see Paragraph 3.3). We have the following characterization of irreducible semigroups.

Proposition 20 *For a positive semigroups $(G(t))_{t\geq 0}$ on a Banach lattice X, the following are equivalent:*

(i) $(G(t))_{t\geq 0}$ *is irreducible;*

(ii) For every $0 < x \in X$ and $0 < \phi \in X^*$, there exists $t \geq 0$ such that $<\phi, G(t)x>> 0$;

(iii) $R(\lambda, A)$ *is strongly irreducible for all (some)* $\lambda > s(A)$*;*

(iv) $R(\lambda, A)$ *is irreducible for all (some)* $\lambda > s(A)$.

Proof. $(i) \Rightarrow (ii)$ For $0 < \phi \in X^*$ we define

$$E = \{x \in X; <\phi, G(t)|x|>>= 0 \text{ for all } t \geq 0\}.$$

This is a closed $\{G(t)\}$-invariant ideal in X. Since $(G(t))_{t\geq 0}$ is irreducible and $E \neq X$, we have $E = \{0\}$ which gives (ii). $(ii) \Rightarrow (iii)$ Take $0 < u \in X$ and $\lambda > s(A)$. From (79) we have

$$<\phi, R(\lambda, A)u>= \int_0^\infty e^{-\lambda t} <\phi, G(t)u> dt > 0$$

which shows that $R(\lambda, A)u$ is a quasi-interior point for any λ.

$(iii) \Rightarrow (iv)$ Obvious.

$(iv) \Rightarrow (i)$ Using again (79) we see that the closed linear span of $R(\lambda, A)E$ is contained in the closed linear span of $\{G(t)E;\ t \geq 0\}$. If E is $\{G(t)\}$-invariant, then $R(\lambda, E) \subseteq E$. $\square$ Much more information about the spectrum can be obtained if $(G(t))_{t\geq 0}$ is an irreducible semigroup (see Paragraph **??**). We have the following main theorem.

Theorem 74 *Let $(G(t))_{t\geq 0}$ be a positive irreducible semigroup generated by A and let $s(A) > -\infty$ be a pole of the resolvent $R(\lambda, A)$. Then:*

1. *$s(A)$ is a first-order pole with geometric multiplicity 1; moreover there exists a quasi-interior point $x_0 \in X_+$ satisfying*

   $$Ax_0 = s(A)x_0,$$

   and a strictly positive $x_0^* \in X_+^*$ such that

   $$A^*x_0 = s(A)x_0^*.$$

2. *$\sigma_{per,s(A)} = s(A) + i\nu\mathbb{Z}$ for some $\nu \geq 0$ and all elements of $\sigma_{per,s(A)}$ are first-order poles of $R(\lambda, A)$ with algebraic multiplicity 1.*

Proof. As usual assume $s(A) = 0$ and let p be the order of the pole 0 of $R(\lambda, A)$. Define

$$Q = \lim_{\lambda \to 0} \lambda^p R(\lambda, A).$$

If $p > 1$ then, by (26), $Q^2 = B_{-p}^2 = B_{-2p+1} = 0$ as $-2p + 1 < -p$. On the other hand, consider again the $(G(t))_{t \geq 0}$ invariant ideal

$$I = \{x \in X;\ Q|x| = 0\}.$$

By irreducibility, $I = \{0\}$ (as it cannot be X due to $Q \neq 0$). Thus, $Q^2 \neq 0$ and this contradiction proves $p = 1$.

The operator Q is thus a positive projection on $Ker\ A$. Let $x \in X_+$ be such that $Qx = x_0 \neq 0$. Since $AQ = QA$, we have $Ax_0 = 0$ and, by the Spectral Mapping Theorem for point spectrum, $G(t)x_0 = x_0$ for any $t \geq 0$ and $\bar{X}_{x_0}$ is a $(G(t))_{t \geq 0}$ invariant ideal yielding, by irreducibility, $X = \bar{X}_{x_0}$. Hence, x_0 is a quasi-interior point.

Since Q^* is a positive projection on $Ker\ A^*$, let us consider $x_0^* = Q^* x^*$; then $A^* x_0^* = 0$ and $G(t)^* x_0^* = x_0^*$. Consequently,

$$J = \{x \in X;\ < |x|, x_0 >= 0\}$$

is a $(G(t))_{t \geq 0}$ invariant closed ideal and thus $J = \{0\}$. This means that x_0^* is strictly positive. We can normalize it so that $< x_0, x_0^* >= 1$.

To prove that 0 is simple, first let us consider $x > 0$ satisfying $Ax = 0$ and normalized to $< x, x_0^* >= 1$. Since we have

$$G(t)|x - x_0| \geq |G(t)(x - x_0)| = |x - x_0|$$

hence $A|x - x_0| = 0$. If $G(t)|x - x_0| > |x - x_0|$, then, by strict positivity of x_0^*, $< G(t)|x - x_0|, x_0^* >>< |x - x_0|, x_0^* >$. This is, however, impossible, as $< G(t)|x - x_0|, x_0^* >=< G(t)|x - x_0|, G(t)^* x_0^* >=< |x - x_0|, x_0^* >$ Thus $G(t)|x - x_0| = |x - x_0|$ and consequently $A|x - x_0| = |x - x_0|$. Define $u = |x -$

$x_0| + (x - x_0) = 2\sup\{(x - x_0), 0\} = 2(x - x_0)^+$ and $v == |x - x_0| - (x - x_0) = 2\inf\{(x - x_0), 0\} = 2(x - x_0)^-$. Thus, $u, v \in X_+$ and $Au = Av = 0$. By the above, u, v are quasi-interior points, or 0. If both were non-zero, then both would be weak units. On the other hand, they are disjoint, hence either $u = 0$ or $v = 0$, so one of them must be 0. Assume $v = 0$. Then $|x - x_0| = (x - x_0)$ and thus $< |x - x_0|, x_0^* >=< x, x_0^* > - < x_0, x_0^* >= 1 - 1 = 0$, which yields $x = x_0$. The case $u = 0$ is analogous.

The next step is taking arbitrary $y \in X$ satisfying $Ay = 0$. We write $y = y^+ - y^-$. Since $G(t)y = y$, we have $|y| = |G(t)y)| \leq G(t)|y|$ and, by the argument of the previous paragraph, we find $G(t)|y| = |y|$. Thus we have $G(t)(y^+ - y^-) = y^+ - y^-$ and $G(t)(y^+ + y^-) = y^+ + y^-$, yielding $G(t)y^{\pm} = y^{\pm}$. Using again the argument of the previous paragraph, we find $y^{\pm} =< y^{\pm}, x_0^* > x_0$ which gives $y =< y_0, x_0^* > x_0$ and proves geometric simplicity of $s(A) = 0$.

To prove the second statement, we note that elements of $\sigma_{per,s(A)}$ belong to $\partial\sigma(A)$ and thus can be converted into eigenvalues by embedding the problem into the ultrapower $\hat{X}$. Details are, however, quite involved and we omit them here, see [38, p. 314]. Assume then $\sigma_{per,s(A)} = \sigma_p(A)$. From Theorem 72 we know that $\sigma_{per,s(A)}$ is cyclic and since 0 is an isolated point, it follows that $\sigma_{per,s(A)} = i\nu\mathbb{Z}$ for some $\nu \geq 0$. Let $i\nu \in \sigma_{per,s(A)}$ so that $Au = i\nu u$ for some $0 \neq u \in X$. Then $G(t)u = e^{i\nu t}u$ and thus $G(t)|u| \geq |G(t)u| = |u|$ and, as in the first part of the proof, we find $G(t)|u| = |u|$ or $A|u| = 0$. This gives $\lambda R(\lambda, A)|u| = |u|$ and $\lambda R(\lambda + i\nu, A)u = u$ for $\lambda \in \mathbb{C}$ with $\Re\lambda > 0$. As in the discussion of the proof of Theorem 72, we see $\lambda R(\lambda + ik\nu, A)u^k = u^k$ or, equivalently, $Au^k = ik\nu u^k$. Since u is a quasi-interior point of X we can claim, as in Lemma 6, that actually

$$R(\lambda, A) = S_u^{-k} R(\lambda + ik\nu, A) S_u^k.$$

Since $s(A) = 0$ is a first-order pole, we see that every element of $\sigma_{per,s(A)}$ has the same property. $\square$

Hence we can state the final result in our quest for asynchronous exponential growth.

Corollary 21 *If $(G(t))_{t\geq 0}$ is a positive and irreducible semigroup with $\omega_e(G) < \omega_0(G)$, then $\sigma_{per,s(A)} = \{s(A)\}$ and $s(A)$ is a simple eigenvalue admitting a positive eigenvector. Thus, $(G(t))_{t\geq 0}$ has positive AEG.*

The problem is to find working techniques which would allow to determine whether $(G(t))_{t\geq 0}$ satisfies the assumptions of Corollary 21. As for generation, the most fruitful approach seems to be through perturbations. We shall explore several such techniques in the next subsection.

6.3 Compactness, positivity and irreducibility of perturbed semigroups

In Subsection 4.6 we discussed various perturbation theorems ensuring the existence of the semigroup associated with $A + B$. In this section we shall discuss to which extent the asymptotic behaviour of the perturbed semigroup is related to that of the original one. We shall focus on bounded perturbations. Let us recall that, by Theorem 46, in this case the perturbed semigroup $(G_{A+B}(t))_{t\geq 0}$ is related to $(G_A(t))_{t\geq 0}$ by the Duhamel equation:

$$G_{A+B}(t)x = G_A(t)x + \int_0^t G_A(t-s)BG_{A+B}(s)xds, \ t \geq 0, x \in X \qquad (169)$$

where the integral is defined in the strong operator topology. Moreover, $(G_{A+B}(t))_{t\geq 0}$ is given by the Dyson–Phillips series obtained by iterating (87):

$$G_{A+B}(t) = \sum_{n=0}^{\infty} G_n(t), \qquad (170)$$

where $G_0(t) = G_A(t)$ and

$$G_{n+1}(t)x = \int_0^t G_A(t-s)BG_n(s)xds. \qquad t \geq 0, x \in X. \tag{171}$$

The series converges in the operator norm of $\mathcal{L}(X)$ and uniformly for t in bounded intervals.

6.3.1 Compact and weakly compact perturbations

A model result related to the main question discussed here is

Theorem 75 *[26, p. 258] Let $(G_A(t))_{t\geq 0}$ be strongly continuous semigroup on a Banach space X generated by A and let B be a compact operator. If $(G_{A+B}(t))_{t\geq 0}$ is the semigroup generated by $A+B$, then $G_{A+B}(t) - G_A(t)$ is compact for all $t \geq 0$. In particular*

$$\omega_e(A+B) = \omega_e(A). \tag{172}$$

The proof of this results, as well as of the results below, heavily depends on the *convex compactness property*, which we will discuss below.

Theorem 76 *If $\mathcal{B} : \Omega \to \mathcal{L}(X,Y)$ is a bounded and strongly measurable function on a finite measure space $(\Omega, d\mu)$ such that $\mathcal{B}(\omega)$ is a compact operator for each $\omega \in \Omega$, then the integral $\int_\Omega \mathcal{B}(\omega)d\mu_\omega$ is compact as well.*

Proof. There are many proofs of this result. We sketch one of them, specified to our particular case $\mathcal{B}(s) = G_A(t-s)BS(s)$, where S is a strongly continuous function, $s \in [0,t]$ and t is fixed. The function $\mathbb{R}_+ \times X \ni (t,x) \to G_A(t)x$ is jointly continuous. Furthermore, since a strongly continuous function is uniformly continuous on compact sets, the set

$$M = \{G_A(s)Bx \; s \in [0,t], \|x\| \leq c\}$$

is relatively compact in X. Having in mind that the Riemann integral is the norm limit of Riemann sums, we find that $(ct)^{-1}\int\limits_0^t G_A(t-s)BS(s)xds$ is an element of the closed convex hull of M provided $c = \sup\{\|S(s)x\|;\; s \in [0,t], \|x\| \leq 1\}$. Since the closed convex hull of a relatively compact set is compact, the statement follows. □

The assumption of compactness of the perturbing operator is often too restrictive. We mentioned earlier that integral operators with natural kernels in important L_1 spaces are not compact but weakly compact. Fortunately, the convex (weak) compactness property holds in this case as well, though the proof in general case is much more delicate.

Theorem 77 *[44] If $\mathcal{B} : \Omega \to \mathcal{L}(X,Y)$ is a bounded strongly measurable function on a finite measure space $(\Omega, d\mu)$ such that $\mathcal{B}(\omega)$ is a weakly compact operator for each $\omega \in \Omega$, then the integral $\int\limits_\Omega \mathcal{B}(\omega)d\mu_\omega$ is compact as well.*

Proof. We sketch here a simple proof of this fact, from [37] when $X = Y = L_1(\Omega, d\nu)$ with Ω being σ-finite. By Eberlein-Šmulian theorem, weak compactness is equivalent to weak sequential compactness, as so we can restrict our attention to separable X (by considering closed spans of sequences). The crucial ingredient of the proof is the criterion of weak compactness in L_1 (see [24, p. 292]): the set $E \subset L_1(\Omega, d\nu)$ is relatively weakly compact if and only if for any decreasing nested sequence $(\Omega_j)_{j\in\mathbb{N}} \subset \Omega$ of measurable sets satisfying $\bigcap_j \Omega_j = \emptyset$ we have

$$\sup_{f\in E}\int\limits_{\Omega_j} |f(z)|d\nu_z \to 0$$

as $j \to \infty$. Thus, we consider

$$\sup_{\|x\|\le 1} \int_{\Omega_j} \left| [\int_{\Omega} \mathcal{B}(\omega) x d\mu_\omega](z) \right| d\nu_z \le \int_{\Omega} \left(\sup_{\|x\|\le 1} \int_{\Omega_j} |[\mathcal{B}(\omega)x](z)| d\nu_z \right) d\mu_\omega$$

where we used the fact that

$$\Omega \ni \omega \to \sup_{\|x\|\le 1} \int_{\Omega_j} |[\mathcal{B}(\omega)x](z)| d\nu_z$$

is measurable on account of separability of X.

Since $B(\omega)$ is weakly compact, we have

$$\sup_{\|x\|\le 1} \int_{\Omega_j} |[\mathcal{B}(\omega)x](z)| d\nu_z \to 0$$

as $j \to \infty$. Since

$$\sup_{\|x\|\le 1} \int_{\Omega_j} |[\mathcal{B}(\omega)x](z)| d\nu_z \le \sup_{\|x\|\le 1} \|\mathcal{B}(\omega)x\|_{L_1} \le C$$

on account of boundedness of the family $\mathcal{B}(\omega)$, the dominated convergence theorem ends the proof. □

With Theorem 76, the proof of Theorem 75 is immediate, since compact perturbations do not change the essential spectrum.

For weakly compact perturbations the situation is more delicate: clearly we know that the difference $G_{A+B}(t) - G_A(t)$ is weakly compact but this does not yield equality of essential spectral types. Restricting, however, our attention to spaces L_1, we know that the square of a weakly compact operator is compact and we should be able to use Theorem 6 for power compact operators.

Unfortunately, the situation is still not obvious, as the relation between the spectra of A and $A+B$ is determined by properties of $R(\lambda, A)B$ (or $BR(\lambda, A)$) and not of B: for $\lambda \in \rho(A)$

$$\lambda \in \sigma(A+B) \Leftrightarrow 1 \in \sigma(BR(\lambda, A)) \Leftrightarrow 1 \in \sigma(R(\lambda, A)B).$$

This situation prompted Voigt [46] to introduce concepts of T-power compact and strictly power compact operators. C is said to be T power compact on $\Delta \in \rho(T)$ if there is n such that $(CR(\lambda, T))^n$ is compact for $\lambda \in \Delta$. C is strictly power compact if DC is power compact for any bounded D.

We note that Voigt introduces in [46] yet another 'essential spectrum' of an operator C. However, the unbounded component of his essential spectrum coincides with the unbounded component of the set of all Fredholm points of C and thus the essential spectral radii determined by all these definitions coincide. The main result needed here is

Theorem 78 *[46, Corollary 1.4] If C and T are bounded and C is T power compact on the unbounded component of $\rho(A)$, then the unbounded components of the Voigt's essential spectrum of T and $T+C$ coincide.*

By the remark above the theorem, unbounded components of essential spectra of T and $T+C$ coincide and thus

$$r_e(T) = r_e(T+C). \tag{173}$$

The importance of this result here is due to the fact that weakly compact operators form an ideal; that is, if C is weakly compact, then AC and CA are weakly compact for any bounded A. Thus, in any L_1 space, weakly compact operators are strictly power compact with $(AC)^2, (CA)^2$ compact. Hence, arguing as for Theorem 75 with the aid of Theorem 77 we arrive at the following result.

Corollary 22 *If $(G_A(t))_{t\geq 0}$ is a strongly continuous semigroup on a Banach space $X = L_1(\Omega)$ generated by A and let B be a weakly compact operator. If $(G_{A+B}(t))_{t\geq 0}$*

is the semigroup generated by $A + B$, then $G_{A+B}(t) - G_A(t)$ is a strictly power compact for all $t \geq 0$. Hence, in particular

$$\omega_e(G_{A+B}) = \omega_e(G_A). \tag{174}$$

6.3.2 Eventual uniform continuity of perturbed semigroups

If a semigroup is eventually uniformly continuous, then the Spectral Mapping Theorem is valid. Moreover, many compactness results can be proved if the underlying semigroup is eventually uniformly continuous. hence, we shall discuss here a few relevant results for the perturbed semigroup.

If F and G are strongly continuous operator valued functions, then the convolution

$$(F * G)(t)(x) := \int_0^t F(t-s)G(s)xds, \quad t \geq 0, x \in X, \tag{175}$$

is well defined. We have the following basic result:

Lemma 7 *If F and G are strongly continuous, then*

(i) If F is uniformly continuous (resp., compact) on $(0, \infty)$, then $F*G$ and $G*F$ are uniformly continuous (resp., compact) on $(0, \infty)$;

(ii) If F is uniformly continuous (resp., compact) on (α, ∞), and G is uniformly continuous (resp., compact) on (β, ∞), then $F*G$ and $G*F$ are uniformly continuous (resp., compact) on $(\alpha + \beta, \infty)$.

Proof. (i) For $t > 0$,$h > 0$ and $x \in X$ we have

$$\begin{aligned}
&\lim_{h\searrow 0} \|(F*G)(t+h)x - (F*G)(t)x\| \\
&\leq \lim_{h\searrow 0} \| \int_0^t \|F(t+h-s) - F(t-s)\| \sup_{\tau\in[0,t]} \|G(\tau)\|\|x\| ds \\
&+ \lim_{h\searrow 0} \| \int_t^{t+h} \sup_{\tau\in[0,t]} \|F(\tau)\| \sup_{\tau\in[0,t]} \|G(\tau)\|\|x\| ds \\
&= 0,
\end{aligned}$$

uniformly in $\|x\| \leq 1$. This shows continuity from the right. Continuity from the left follows in the same way. The statement for $G * F$ follows by symmetry.

Compactness follows directly from Theorem 76.

The uniform continuity in the statement (ii) follows through similar but more detailed estimates. To prove compactness, we take $t > \alpha+\beta$ and if $0 < s < t-\alpha$, then $t - s > \alpha$ and if $t - \alpha < s < t$, then $s > \beta$. Now

$$(F * G)(t)x = \int_0^{t-\alpha} F(t-s)G(s)x ds + \int_{t-\alpha}^{t} F(t-s)G(s)x ds$$

and the statement follows by applying Theorem 76 to each term. □

Consider the semigroup $(G_A(t))_{t\geq 0}$ and the perturbed semigroup $(G_{A+B}(t))_{t\geq 0}$. From Duhamel formula we have

$$G_{A+B} = G_A + G_A * BG_{A+B} = G_A + G_A B * G_{A+B}. \qquad (176)$$

We define the Volterra operator associated with this problem as

$$\mathcal{V}F = G_A * BF = G_A B * F$$

for any strongly continuous F.

We have the following result.

Theorem 79 *Suppose B is bounded. Then*

(a) If $(G_A(t))_{t\geq 0}$ is immediately uniformly continuous (resp., compact), then the same holds for $(G_{A+B}(t))_{t\geq 0}$;

(b) If $(G_A(t))_{t\geq 0}$ is uniformly continuous (resp., compact) on (α, ∞) and if there exists $k \in \mathbb{N}$ such that $\mathcal{V}^k G_A$ is uniformly continuous (resp., compact) on $(0, \infty)$, then $(G_{A+B}(t))_{t\geq 0}$ is uniformly continuous (resp., compact) on $(k\alpha, \infty)$.

Proof. (i) follows immediately from Lemma 7 (i) as $G_{A+B} = G_A + G_A * BG_{A+B}$.

To prove (ii) we note that by Dyson-Phillips expansion

$$G_{A+B}(t) = \sum_{n=0}^{k} \mathcal{V}^n G_A(t) + \sum_{n=1}^{\infty} \mathcal{V}^n(\mathcal{V}^k G_A(t))$$

where the series converges in uniform operator topology on compact intervals by Theorem 46. The terms in the first part are uniformly continuous (resp., compact) on $(k\alpha, \infty)$ by Lemma 7 (ii). The second part can be written as

$$\sum_{n=1}^{\infty} \mathcal{V}^n(\mathcal{V}^k G_A) = G_A * B(\mathcal{V}^k G_A) + G_A * B(G_A * \mathcal{V}^k G_A) + \dots$$

where each term is uniformly continuous on $(0, \infty)$ by Lemma 7 (i) and converges in uniform operator topology, as above. □

6.3.3 Irreducibility of perturbed semigroups.

Here we assume that $(G_A(t))_{t\geq 0}$ is a positive semigroup and B is aa bounded positive operator. Then $BR(\lambda, A)$ are positive for sufficiently large λ ($\lambda > s(A)$)

and the terms of Dyson-Phillips expansion $\mathcal{V}^n G_A(t)$ are positive operators for $t \geq 0$. The last statement follows from the fact that iterates defining this expansion are positive by (90). As a consequence, we have

$$0 \leq G_A(t) \leq G_{A+B}(t)$$

and consequently

$$\omega_0(A) \leq \omega_0(A+B).$$

The formula (86) shows that for sufficiently large $\lambda \in \mathbb{R}$ we have

$$0 \leq R(\lambda, A) \leq R(\lambda, A+B)$$

so that

$$s(A) \leq s(A+B).$$

This follows, e.g. from the fact that is we approach $s(A)$ then $R(\lambda, A)$ blows up and thus $R(\lambda, A+B)$ must blow up, hence $s(A) \notin \rho(A+B)$.

Theorem 80 *Let X be a Banach lattice, $(G_A(t))_{t\geq 0}$ a positive semigroup and B a positive bounded operator. The perturbed semigroup $(G_{A+B}(t))_{t\geq 0}$ is irreducible if and only if $I = \{0\}$ and $I = X$ are the only closed ideals satisfying*

(a) $G_A(t)I \subseteq I, t \geq 0$,

(b) $BI \subseteq I$.

Proof. Assume that $(G_{A+B}(t))_{t\geq 0}$ is irreducible and let I satisfies (a) and (b). Using Proposition 20, we obtain $R(\lambda, A)I \subseteq I$. Hence $K(\lambda)I = R(\lambda, A)BI \subseteq I$. By (86) we obtain $R(\lambda, A+B)I \subseteq I$ and thus $I = \{0\}$ or $I = X$.

To prove the converse, assume $I = \{0\}$ and $I = X$ are the only closed ideals satisfying (a) and (b). We show that the $R(\lambda, A+B)$ is irreducible which is equivalent, again by Proposition 20, to irreducibility of $(G_{A+B}(t))_{t\geq 0}$. So, let I be a closed ideal satisfying $R(\lambda, A+B)I \subseteq I$ for $\lambda > s(A)$ and $x \in I$. We have

$$|R(\lambda, A)x| \leq R(\lambda, A)|x| \leq R(\lambda, A+B)|x| \in I$$

that is $R(\lambda, A)I \subseteq I$ and, equivalently, $G_A(t)I \subseteq I$. Take again $x \in I$; then for $\lambda > s(A)$

$$\begin{aligned} |R(\lambda, A)BR(\lambda, A+B)x| &\leq R(\lambda, A)BR(\lambda, A+B)|x| \\ = (R(\lambda, A+B) - R(\lambda, A))|x| &\leq R(\lambda, A)|x| \in I \end{aligned}$$

where we used (86). Thus, for $\mu > s(A)$ we have also

$$R(\mu, A)R(\lambda, A)BR(\lambda, A+B)x \in X$$

provided $x \in I$. Using the resolvent equation, we get

$$(R(\mu, A) - R(\lambda, A))BR(\lambda, A+B)x \in I,$$

hence, by linearity of I,

$$R(\mu, A)BR(\lambda, A+B)x \in I,$$

We multiply the above by λ, let $\lambda \to \infty$, use $\lambda R(\lambda, A+B)x \to x$ and closedness of I to obtain $R(\mu, A)Bx \in I$. Multiplying the latter by μ and repeating the argument we obtain $Bx \in I$ hence I is B invariant.

Thus (a) and (b) are satisfied yielding $I = \{0\}$ or $I = X$ and hence $(G_{A+B}(t))_{t\geq 0}$ is irreducible. □

6.3.4 A model of evolution of a blood cell population

We consider a population of blood cells distinguished only by their size and describe the population by the density function $n(t, s)$ of cells having size s in time t. The following processes take place when the time passes:

1. Each cell grows linearly in time;
2. Each cell dies with a probability depending on size;
3. Each cell divides into two daughter cells of equal size with a probability depending on size;

Moreover, we assume that there exists a maximal cell size (here normalized to 1); also there exists a minimal cell size $s = \alpha > 0$ below which no division can occur. As a consequence of the last assumption, if we start with initial population with sizes greater that $\alpha/2$, the size of each cell in the population must satisfy $s > \alpha/2$ and we can assume the boundary condition

$$n(t, \alpha/2) = 0, \qquad t > 0.$$

These assumption lead to the following evolution equation:

$$\begin{aligned} n_t(t,s) &= -n_s(t,s) - \mu(s)n(t,s) - b(s)n(t,s) \\ &\quad +4b(2s)n(t,2s)\chi_{[\alpha/2,1/2]}(s), \quad s > \alpha/2, t > 0 \\ u(0,s) &= n_0(s), \end{aligned} \tag{177}$$

where χ_A is the characteristic function of the set A. We assume that the death rate μ is a positive continuous function on $[\alpha/2, 1]$. The division rate should be continuous with $b(s) > 0$ on $(\alpha, 1)$ and $b(s) = 0$ elsewhere.

We consider this equation as an abstract evolution equation in $X = L_1([\alpha/2, 1], dx)$ and define operators

$$Af = -f' - (\mu + b)f \tag{178}$$

on $D(A) = W^{1,1}([\alpha/2, 1])$ and

$$(Bf)(s) = 4b(2s)n(t,2s)\chi_{[\alpha/2,1/2]}(s) \tag{179}$$

on $D(B) = X$ (since multiplication by 2 is bi-lipschitz, the composition is well-defined) as an operation in X.

Hence, we define

$$K = A + B, \quad D(K) = D(A).$$

The following result is standard.

Lemma 8 *The operator* $(A, D(A))$ *generates a* C_0*-semigroup explicitly given by*

$$G_A(t)f = \begin{cases} e^{-\int\limits_{s-t}^{s}(\mu(\tau)+b(\tau))d\tau} f(s-t) & \text{for} \qquad s-t > \alpha/2 \\ 0 & \text{otherwise.} \end{cases} \tag{180}$$

The spectrum of A *is empty and the resolvent* $R(\lambda, A)$*, given explicitly by*

$$(R(\lambda, A)g)(s) = \int\limits_{\alpha/2}^{s} e^{-\int\limits_{\sigma}^{s}(\mu(\tau)+b(\tau))d\tau} g(\sigma)d\sigma \tag{181}$$

is compact.

We have also

Lemma 9 *The semigroup* $(K(t))_{t\geq 0}$ *is eventually uniformly continuous and eventually compact for* $t > 1 - \alpha/2$*.*

Proof. To prove eventual uniform continuity, we first note that $(G_A(t))_{t\geq 0}$ is zero for $t > 1 - \alpha/2$ and thus uniformly continuous on this interval. Hence, by Theorem 79 (ii), it suffices to prove immediate uniform continuity of some term of the Dyson-Phillips expansion. It turns out that $\mathcal{V}G_A$ is immediately uniform continuous and hence $(K(t))_{t\geq 0}$ is uniformly continuous for $t > 1 - \alpha/2$.

To prove compactness, we note that $R(\lambda, A)$ is compact and, as $R(\lambda, K)$ is given by uniformly converging series (86) of compact operators, $R(\lambda, K)$ is compact as well. Hence, $R(\lambda, A)G_K(t)$ is compact for such t.

It is interesting that this implies compactness of $(G_K(t))_{t\geq 0}$. Indeed, defining $R(t)x = \int\limits_0^t G_K(s)xds$, we have $AR(t)x = G_K(t)x - x$ and

$$R(t) = R(\lambda, A)(I - G_K(t)) + \lambda R(\lambda, A)R(t)x$$

so that, for a fixed $t_0 > 1 - \alpha/2$,

$$\begin{aligned} R(t_0 + h) - R(t_0) &= -R(\lambda, A)[G_K(t_0 + h) - G_K(t_0)] \\ &\quad -\lambda R(\lambda, A)[R(t_0 + h) - R(t_0)]. \end{aligned} \tag{182}$$

Since $(G_K(t))_{t\geq 0}$ is uniformly continuous at t_0,

$$G_K(t_0) = \lim_{h\to 0} h^{-1}(R(t_0 + h) - R(t_0))$$

in the uniform topology. Since the first term on the right-hand side in (182) is compact and the second converges to $\lambda R(\lambda, A)G_K(t_0)$, which is compact, $G_K(t_0)$ is compact. Thus, we get compactness for $t > 1 - \alpha/2$. $\square$

We note that eventual compactness implies $\omega_e(G_K) = -\infty$ and hence clearly $\omega_e(G_K) < \omega_0(G_K)$.

The final step is to establish irreducibility of $(G_K(t))_{t\geq 0}$.

Lemma 10 *The semigroup $(G_K(t))_{t\geq 0}$ is irreducible.*

Proof. Let us analyse how the resolvent

$$R(\lambda, K) = R(\lambda, A) + R(\lambda, A)BR(\lambda, A) + \dots$$

acts on functions with support (a.e) in $(s_0, 1]$ (precisely, $s_0 = \sup\{s\}$ such that $suppf \subset [s, 1]$. Then, by (181), $R(\lambda, A)f > 0$ on $(s_0, 1]$, $BR(\lambda, A)f > 0$ on $(s_0/2, 1/2]$, $R(\lambda, B)BR(\lambda, A)f > 0$ on $(s_0/2, 1]$ and, continuing, $R(\lambda, K)f > 0$ on $[\alpha/2, 1]$. Using the description of ideals in L_1, Example 10, we see that no closed non-trivial ideal can be invariant under $R(\lambda, K)$ and, by Proposition 20, we obtain irreducibility of the semigroup. $\square$

We collect all the information in the following summarising theorem.

Theorem 81 *Let $(G_K(t))_{t\geq 0}$ be the C_0-semigroup corresponding to (186). Then there is a dominant eigenvalue equal to $s(A)$ and the corresponding 1-dimensional positive projection P such that*

$$\|e^{-s(A)t}G_K(t) - P\| \leq Me^{-\epsilon t}$$

for some $M, \epsilon > 0$ and all $t \geq 0$. Thus, $(G_K(t))_{t\geq 0}$ has a positive AEG.

The dominant eigenvalue can be found by solving a scalar equation.

Proposition 21 *Assume $\alpha \geq 1/2$. The spectrum $\sigma(K)$ only consists of eigenvalues which are solutions of the characteristic equation*

$$\xi(\lambda) := -1 + \int_{\alpha/2}^{1/2} 4b(2\sigma)e^{-\int_{\sigma}^{2\sigma}(\mu(\tau)+b(\tau)+\lambda)d\tau} d\sigma = 0. \tag{183}$$

The spectral bound $s(A)$ is the unique $\lambda_0 \in \mathbb{R}$, for which $\xi(\lambda_0) = 0$. The semigroup $(G_K(t))_{t\geq 0}$ is stable if and only if $\xi(0) < 0$.

Proof. The first statement follows from the fact that K has a compact resolvent.

For the second statement we have to solve $\lambda f - Kf = 0$; that is

$$\begin{aligned} &\lambda f(s) + f'(s) + (\mu(s) + b(s))f(s) = 0, \quad \text{for} \quad s \in [1/2, 1], \\ &\lambda f(s) + f'(s) + (\mu(s) + b(s))f(s) - 4b(2s)f(2s) = 0, \\ &\text{for} \quad s \in [\alpha/2, s/2). \end{aligned} \tag{184}$$

Solving the first equation with a normalizing condition $f(1) = 1$ yields

$$f(s) = e^{\int\limits_s^1 (\mu(\sigma)+b(\sigma)+\lambda)d\sigma}$$

for $s \in [1/2, 1]$. Turning to the second equation we see that if $s \in [\alpha/2, 1/2]$, then the argument $g(2s)$ varies between α and 1 and we can substitute

$$f(2s) = e^{\int\limits_{2s}^1 (\mu(\sigma)+b(\sigma)+\lambda)} d\sigma,$$

at least on the interval $[1/4, 1/2]$ and solve the second equation in (184) as a non-homogeneous equation. Taking into account that the solution must be continuous at $s = 1/2$, we obtain

$$f(s) = e^{\int\limits_s^1 (\mu(\sigma)+b(\sigma)+\lambda)d\sigma} \left[1 - 4 \int\limits_s^{1/2} b(2\sigma) e^{-\int\limits_\sigma^{2\sigma} (\mu(\tau)+b(\tau)+\lambda)d\tau} d\sigma \right]$$

This solution must satisfy the boundary condition $f(\alpha/2) = 0$, which gives the desired form of the characteristic function in (183). Next, the function ξ, restricted to $\mathbb{R}$ is continuous, strictly decreasing with $\lim_{\lambda\to-\infty} \xi(\lambda) = +\infty$ and $\lim_{\lambda\to\infty} \xi(\lambda) = -1$ and hence has exactly one real solution λ_0. This solution is negative only if $\xi(0) < 0$. $\square$

6.3.5 Emergence of chaos in the blood cell model

Consider a slightly modified blood cell evolution model (at the beginning, without death and division terms). the transport equation

$$u_t = -xu_x + 0.5u, \quad u(x, 0) = u_0(x), \tag{185}$$

in the space $L_1([0, 1])$. This equation, was analysed in [W] in the space of continuous function and the occurrence of chaos was attributed to the insufficient supply of the most primitive blood cells. It was also investigated in the same space in [Rud, LM] in the statistical framework using the concept of exactness.

Let $u(0,x) = u_0(x), 0 < x < 1$. The explicit solution to (185) is given by $u(t,x) = T(t)u_0(x) = e^{t/2}u_0(xe^{-t})$. We use Theorem 36. The eigenfunctions of the generator are found to be $u_\lambda(x) = x^{-\lambda+1/2}$ provided $\lambda \in U = \{Re\lambda < 3/2\}$. Thus the first assumption of Theorem 36 is satisfied.

Consider now the function

$$F_\lambda[g] = \int\limits_0^1 x^{-\lambda+1/2}g(x)dx$$

where $g \in L_\infty([0,1])$. By $\frac{d}{d\lambda}x^{-\lambda+1/2} = (\ln x)x^{-\lambda+1/2}$ we see that $F_\lambda[g]$ is analytic in U. Changing variable according to $z = -\ln x$ we obtain

$$0 = \int\limits_0^1 x^{-\lambda+1/2}g(x)dx = \int\limits_0^\infty e^{(-\frac{1}{2}+Re\lambda)z}\left(e^{-z}g(e^{-z})\right)e^{iIm\lambda z}dz.$$

Now, the function $F(z) = e^{(-\frac{1}{2}+Re\lambda)z}\left(e^{-z}g(e^{-z})\right)$ is in $L_1([0,\infty])$ for the stipulated range of λ and $Im\lambda$ is not restricted, thus the above integral represents the classical Fourier transform of a function extended by 0 for $z < 0$. Since the transform is zero, $g(x) = 0$ for all x.

Next we consider a more sophisticated version of (185)

$$\begin{aligned} u_t(t,x) &= -xu_x(t,x) + \eta u(t,x) + 4\beta u(t,2x)\chi_{[0,1/2]}(x) \\ u(0,x) &= \phi(x) \end{aligned} \tag{186}$$

in $L_1([0,1],dx)$, where χ_A is the characteristic function of the set A. Change of variables $x = e^{-y}$, $y \geq 0$ gives

$$\begin{aligned} v_t(t,y) &= v_y(t,y) + \eta v(t,x) + 4\beta v(t,y-\ln 2)\chi_{[\ln 2,\infty]}(y) \\ u(0,y) &= \psi(y) = \phi(e^{-y}) \end{aligned} \tag{187}$$

in $X = L_1([0,\infty), e^{-y}dy)$. A nice way to find the eigenvectors (El Mourchid) is to consider the recurrence for $v^n := v|_{[n\ln 2,(n+1)\ln 2)}$

$$v^0 = e^{(\lambda-\eta)}, \qquad y \in [0,\ln 2),$$

$$v_y^n = (\lambda - \eta)v^n - 4\beta v^{n-1}(y - \ln 2),$$
$$y \in [n \ln 2, (n+1)\ln 2), n \geq 1$$

which gives formal eigenvectors as

$$v^n(y) = e^{(\lambda-\eta)y} \sum_{k=0}^{n} \frac{(-4\beta e^{-(\lambda-\eta)\ln 2})^n}{n!}(y - n\ln 2)^n \tag{188}$$

for $n \ln 2 \leq y < (n+1)\ln 2$. Combining and rearranging terms (justified later by absolute convergence), we get

$$v_\lambda(y) = e^{(\lambda-\eta)y} \sum_{n=0}^{\infty} \frac{(-4\beta e^{-(\lambda-\eta)\ln 2})^n}{n!}(y - n\ln 2)^n \chi_{[n\ln 2,\infty)}(y). \tag{189}$$

Estimating the terms of the series in $L_1(\mathbb{R}_+, e^{-y}dy)$ we have

$$\frac{(4\beta e^{-(\Re\lambda-\eta)\ln 2})^n}{n!} \int\limits_{n\ln 2}^{\infty} e^{-(\eta-\Re\lambda+1)y}(y - n\ln 2)^n dy$$
$$= \frac{(2\beta)^n}{n!} \int\limits_0^{\infty} e^{-(\eta-\Re\lambda+1)z} z^n dz$$
$$= \left(\frac{2\beta}{\eta - \Re\lambda + 1}\right)^n$$

and we see that the series for v_λ is convergent in the half-plane $\Re\lambda < \eta + 1 - 2\beta$ and uniformly convergent is any closed half-plane contained in it. It is easy to see that each term of the series is of the form

$$\lambda \to \phi(\lambda) = e^{\lambda(y-a)} f(y)$$

where a is a constant and f is such that the above function is integrable (with weight e^{-y}) for any fixed λ with $\Re\lambda < \eta + 1 - 2\beta$. Thus, taking such λ_1, λ_2 with, say, $\Re\lambda_1 > \Re\lambda_2$, we have

$$\|\phi(\lambda_1) - \phi(\lambda_2)\| \leq \int\limits_0^{\infty} |1 - e^{(\lambda_2-\lambda_1)(y-\alpha)}| e^{\lambda_1(y-\alpha)} f(y) e^{-y} dy$$

we see that the term between the absolute bars is bounded. Hence, by the dominated convergence, ϕ is continuous and, by the uniform convergence of the series, v_λ is a continuous function of λ. Consequently, if $\eta + 1 - 2\beta > 0$, the dynamics generated by (186) is subchaotic.

7 Asymptotic analysis of singularly perturbed dynamical systems

The goal of this section is to give a concise explanation of concepts of asymptotic analysis and, in particular, of one technique of the asymptotic analysis which essentially stems from the Chapman-Enskog procedure.

In order to introduce this asymptotic procedure, let us consider a particular case of singularly perturbed abstract initial value problem

$$\left\{ \begin{array}{rcl} \frac{\partial f_\epsilon}{\partial t} & = & S f_\epsilon + \frac{1}{\epsilon} C f_\epsilon, \\ f_\epsilon(0) & = & f_0, \end{array} \right. \tag{190}$$

where the presence of the small parameter ϵ indicates that the phenomenon modelled by the operator C is more relevant than that modelled by S or, in other words, they act on different time scales.

As elsewhere in these lectures, we are concerned with kinetic type problems and the operator S describes some form of transport, whereas C is an interaction/transition operator describing interstate transfers, e.g., they may be collision operator in the kinetic problems or a transition matrix in the structured population theory.

We are often interested in situations when the transition processes between structure states are dominant. If this is the case, the population quickly becomes homogenised with respect to the structure and starts to behave as an unstructured, governed by s suitable equations (which in analogy with the kinetic theory will be called *hydrodynamic*). These equations should be the limit, or approximating, equation for (190) as $\epsilon \to 0$ (the parameter ϵ in such a case is related to the mean free time between state switches).

To put this in a mathematical framework, we can suppose to have on the right-hand side a family of operators $\{C_\epsilon\}_{\epsilon>0} = \{S + \frac{1}{\epsilon}C\}_{\epsilon>0}$ acting in a suitable Banach space X, and a given initial datum. The classical asymptotic analysis consists in looking for a solution in the form of a truncated power series

$$f_\epsilon^{(n)}(t) = f_0(t) + \epsilon f_1(t) + \epsilon^2 f_2(t) + \cdots + \epsilon^n f_n(t),$$

and builds up an algorithm to determine the coefficients $f_0, f_1, f_2, \ldots, f_n$. Then $f_\epsilon^{(n)}(t)$ is an approximation of order n to the solution $f_\epsilon(t)$ of the original equation in the sense that we should have

$$\|f_\epsilon(t) - f_\epsilon^{(n)}(t)\|_X = o(\epsilon^n)\,, \tag{191}$$

for $0 \le t \le T$, where $T > 0$.

It is important to note that the zeroth-order approximation satisfies

$$Cf_0(t) = 0$$

which is the mathematical expression of the fact that the hydrodynamic approximation should be transition-free and that's why the null-space of the dominant collision operator is called the *hydrodynamic space* of the problem.

Another important observation pertains to the fact that in most cases the limit equation involves less independent variables than the original one. Thus the solution of the former cannot satisfy all boundary and initial conditions of the latter. Such problems are called *singularly perturbed*. If, for example, the approximation (191) does not hold in a neighbourhood of $t = 0$, then it is necessary to introduce an *initial layer* correction by repeating the above procedure with rescaled time to improve the convergence for small t. The original approximation which is valid only away from $t = 0$ is referred to as the *bulk approximation*.

Similarly, there approximation could fail close to the spatial boundary of the domain as well as close to the region where the spatial and temporal boundaries meet. To improve accuracy in such cases one introduces the so-called *boundary* and *corner layer* corrections, but we will not discuss them here.

A first way to look at the problem from the point of view of the approximation theory is to find, in a systematic way, a new (simpler) family of operators, still depending on ϵ, say B_ϵ, generating new evolution problems

$$\frac{\partial \varphi_\epsilon}{\partial t} = B_\epsilon \varphi_\epsilon,$$

supplemented by appropriate initial conditions, such that the solutions $\varphi_\epsilon(t)$ of the new evolution problem satisfy

$$\|f_\epsilon(t) - \varphi_\epsilon(t)\|_X = o(\epsilon^n)\,, \tag{192}$$

for $0 \leq t \leq T$, where $T > 0$ and $n \geq 1$. In this case we say that B_ϵ is a hydrodynamic approximation of C_ϵ to order n.

This approach mathematically produces weaker results than solving system (190) for each ϵ and eventually taking the limit of the solutions as $\epsilon \to 0$. But in real situation, ϵ is small but not zero, and it is interesting to find simpler operators B_ϵ for modelling a particular regime of a physical system of interacting particles.

A slightly different point of view consists in requiring that the limiting equation for the approximate solution does not contain ϵ. In other words, the task is now to find a new (simpler) operator, say B, and a new evolution problem

$$\frac{\partial \varphi}{\partial t} = B\varphi,$$

with an appropriate initial condition, such that the solutions $\varphi(t)$ of the new evolution problem satisfy

$$\|f_\epsilon(t) - \varphi(t)\|_X \to 0, \text{ as } \epsilon \to 0, \tag{193}$$

for $0 \leq t \leq T$, where $T > 0$.

In this case we say that B is the hydrodynamic limit of operators C_ϵ as $\epsilon \to 0$. This approach can be treated as (and in fact is) a particular version of the previous one as very often the operator B is obtained as the first step in the procedure leading eventually to the family $\{B_\epsilon\}_{\epsilon \geq 0}$. For instance, for the nonlinear Boltzmann equation with the original Hilbert scaling, B would correspond to the Euler system, whereas B_ϵ could correspond to the Navier-Stokes system with ϵ-dependent viscosity, or to Burnett equation at yet higher level.

In this review I will focus on the first method which, in some sense, is dictated by particular applications, where the scaling is given. The other may be seen as more mathematical as one is then looking for suitable scalings of independent variables and physical parameters which lead to the limiting equations not depending on ϵ, see [10, 9].

In any case the asymptotic analysis, should consist of two main points:

- determining an algorithm which provides in a systematic way the approximating family B_ϵ (or the limit operator B),
- proving the convergence of f_ϵ in the sense of (192) (or of (193)).

Even if the formal part and the rigorous part of an asymptotic analysis seem not to be related, the formal procedure can be of great help in proving the convergence theorems.

We will focus on the modification of the classical Chapman-Enskog procedure which was adapted to a class of linear evolution equations by J.Mika at the end of the 1970s and later extended to singularly perturbed evolution equations arising in the kinetic theory. The advantage of this procedure is that the projection of the solution to the Boltzmann equation onto the null-space of the collision operator, that is, the hydrodynamic part of the solution, is not expanded in ϵ, and thus the whole information carried by this part is kept together. This is in contrast to the Hilbert type expansions, where, if applicable, only the zero order term of the expansion of the hydrodynamic part is recovered from the limit equation.

The main feature of the modified Chapman-Enskog procedure is that the initial value problem is decomposed into two problems, for the kinetic and hydrodynamic parts of the solution, respectively. This decomposition consists in splitting the unknown function into the part belonging to the null space V of the operator C, which describes the dominant phenomenon, whereas the remaining part belongs to the complementary subspace W.

Thus the first step of the asymptotic procedure is finding the null-space of the dominant collision operator C; then the decomposition is performed using the (spectral) projection $\mathcal{P}$ onto the null-space V by applying $\mathcal{P}$ and the complementary projection $\mathcal{Q} = I - \mathcal{P}$ to equation (190). In this one obtains a system of evolution equations in the subspaces V and W. At this point the kinetic part of the solution is expanded in series of ϵ, but the hydrodynamic part of the solution is left unexpanded. In other words, we keep all orders of approximation of the hydrodynamic part compressed into a single function.

One of the main drawback of the classical approach is that the initial layer contribution is neglected and transitional effects are not taken into account. To overcome this, two time scales are introduced in order to obtain the necessary corrections. In general, the compressed asymptotic algorithm permits to derive in a natural way the hydrodynamic equation, the initial condition to supplement it, and the initial layer corrections.

Summarizing, the original Chapman-Enskog method is improved by the introduction of two new ingredients:

- the projection of the original equation onto the hydrodynamic subspace,
- the analysis of the evolution equations in terms of the theory of semigroups.

Taking these new ingredients into account, we obtain the following main advantages:

- we can build an algorithm listing the steps of the procedure to be followed,
- we are able to establish all the mathematical properties of the full and limit solutions needed for the rigorous convergence proof.

7.1 Compressed expansion

For clarity, we present the compressed method on a simplified model with the small parameter appearing only in one place:

$$\partial_t u = \mathcal{S}u + \frac{1}{\varepsilon}\mathcal{C}u, \tag{194}$$

in a Banach space $\mathcal{X}$. However, the analysis can be extended to more general cases.

The success of the method depends on the spectral properties of the operators $\mathcal{S}$ and $\mathcal{C}$. To be able to start, we must assume that $\lambda = 0$ is the dominant simple eigenvalue of the operator $\mathcal{C}$.

It is easy to see that this requirement amounts to $\mathcal{C}$ being the generator of a semigroup having AEG. The fact that $\lambda = 0$ needs to be dominant ensures an exponential decay of the initial layer. This assumption may, however, be relaxed if we are not that interested in the properties of the layer.

Remark 12 In many cases we have several state variables and the operator $\mathcal{C}$ only acts on some of them. Then the above requirement refers to the action of $\mathcal{C}$ in this restricted space.

The corresponding eigenspace (the hydrodynamic space of $\mathcal{C}$) is thus one dimensional; we denote by $\mathcal{P}$ the spectral projection of the state space onto this space. Let $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ be the complementary projection. Accordingly, by $\mathcal{P}u = v$ we denote the hydrodynamic part of the solution u and by $\mathcal{Q}u = w$ the kinetic part.

Applying these projections on both sides of (194) we get

$$\begin{aligned} \partial_t v &= \mathcal{PSP}v + \mathcal{PSQ}w \\ \varepsilon\partial_t w &= \varepsilon\mathcal{QSQ}w + \varepsilon\mathcal{QSP}v + \mathcal{QCQ}w, \end{aligned} \tag{195}$$

with the initial conditions

$$v(0) = \overset{o}{v},\ w(0) = \overset{o}{w},$$

where $\overset{o}{v} = \mathcal{P}\ \overset{o}{u}$, $\overset{o}{w} = \mathcal{Q}\ \overset{o}{u}$.

We have kept the superfluous symbols $\mathcal{P}v$ and $\mathcal{Q}w$ for the sake of notational symmetry.

The projected operator $\mathcal{PSP}$ vanishes for many types of linear equations and, for simplicity, we perform analysis for such a case. Thus, we obtain the following form of (195)

$$\begin{aligned} \partial_t v &= \mathcal{PSQ}w \\ \partial_t w &= \mathcal{QSP}v + \mathcal{QSQ}w + \mathcal{QAQ}w + \frac{1}{\varepsilon}\mathcal{QCQ}w && (196) \\ v(0) &= \overset{o}{v},\ w(0) = \overset{o}{w}\ . && (197) \end{aligned}$$

We represent the solution of (196) as a sum of the bulk and the initial layer parts:

$$v(t) = \bar{v}(t) + \tilde{v}(\tau), \tag{198}$$

$$w(t) = \bar{w}(t) + \tilde{w}(\tau), \tag{199}$$

where, in this case. the variable τ in the initial layer part is given by $\tau = t/\varepsilon$. Other scalings may require different formulae for τ.

Equations for the bulk part and the initial layer part are sought independently.

The following algorithm describes the main features of the compressed asymptotic procedure are:

Algorithm 1

1. The bulk approximation $\bar{v}$ is not expanded into powers of ε.
2. The bulk approximation $\bar{w}$ is explicitly written in terms of $\bar{v}$ and expanded in powers of ε.
3. The time derivative $\partial_t \bar{v}$ and the initial value $\bar{v}(0)$ are expanded into powers of ε.

Thus

$$\begin{aligned}
\bar{w} &= \bar{w}_0 + \varepsilon \bar{w}_1 + O(\varepsilon^2), \\
\tilde{v} &= \tilde{v}_0 + \varepsilon \tilde{v}_1 + O(\varepsilon^2), \\
\tilde{w} &= \tilde{w}_0 + \varepsilon \tilde{w}_1 + O(\varepsilon^2).
\end{aligned} \tag{200}$$

Substituting the expansion for $\bar{w}$ into (196) and comparing terms of the same powers of ε yield

$$\partial_t \bar{v} = \mathcal{PSQ}(\bar{w}_0 + \varepsilon \bar{w}_1 + O(\varepsilon^2)). \tag{201}$$

and

$$\bar{w}_0 \equiv 0,$$

$$\bar{w}_1 = -(\mathcal{QCQ})^{-1}\mathcal{QSP}\bar{v}.$$

Inserting the expressions for $\bar{w}_0$ and $\bar{w}_1$ into (201) gives the approximate 'diffusion' equation

$$\partial_t \bar{v} = -\varepsilon \mathcal{PSQ}(\mathcal{QCQ})^{-1}\mathcal{QSP}\bar{v} + O(\varepsilon^2). \tag{202}$$

For the initial layer a similar procedure yields

$$\begin{aligned} \tilde{v}_0(\tau) &\equiv 0, \\ \partial_\tau \tilde{w}_0 &= \mathcal{QCQ}\tilde{w}_0, && (203) \\ \partial_\tau \tilde{v}_1 &= \mathcal{PSQ}\tilde{w}_0, && (204) \\ \partial_\tau \tilde{w}_1 &= \mathcal{QCQ}\tilde{w}_1 + \mathcal{QSP}\tilde{v}_0 + \mathcal{QSQ}\tilde{w}_0. && (205) \end{aligned}$$

We observe that, due to $\bar{w}_0 \equiv 0$, the initial condition for $\tilde{w}_0$ is $\tilde{w}_0(0) = \overset{o}{w}$. Solving (203) with this initial value allows to integrate (204) which gives

$$\tilde{v}_1(\tau) = \mathcal{PSQ}(\mathcal{QCQ})^{-1}e^{\tau\mathcal{QCQ}}\overset{o}{w}, \tag{206}$$

upon which $\tilde{v}_1(0) = \mathcal{PSQ}(\mathcal{QCQ})^{-1}\overset{o}{w}$. This in turn allows one to determine the initial condition for the diffusion equation: we have from (199) that $\overset{o}{v} = \bar{v}(0) + \varepsilon\tilde{v}_1(0) + O(\varepsilon^2)$ so that

$$\bar{v}(0) = \overset{o}{v} - \varepsilon\mathcal{PSQ}(\mathcal{QCQ})^{-1}\overset{o}{w} + O(\varepsilon^2). \tag{207}$$

In what follows we adopt a uniform notation valid for all discussed examples. In general, by ρ we shall denote a solution of the "diffusion" equation determined by discarding the $O(\varepsilon^2)$ terms in (202), that is,

$$\partial_t \rho = -\varepsilon\mathcal{PSQ}(\mathcal{QCQ})^{-1}\mathcal{QSP}\rho. \tag{208}$$

Thus, ρ is expected to provide an approximation of $\bar{v}$. By $\hat{\rho}$ we denote the solution of this equation with uncorrected initial condition $\hat{\rho}(0) = \overset{o}{v}$ and by $\bar{\rho}$ the solution with the corrected initial value obtained by discarding the $O(\varepsilon^2)$ terms in (207), that is, $\bar{\rho}(0)$ is given by

$$\bar{\rho}(0) = \overset{o}{v} - \varepsilon \mathcal{PSQ}(\mathcal{QCQ})^{-1} \overset{o}{w} . \tag{209}$$

Specific formulae will be given for each case separately. As we noted earlier, for the the procedure to start, we need $\lambda = 0$ to be a simple eigenvalue which thus admits a spectral projection onto its eigenspace. This condition is satisfied if, in particular, this eigenvalue is isolated. However, the existence of exponentially decaying initial layer requires the operator $\mathcal{QCQ}$ to generate a semigroup of negative type in $\mathcal{QX}$. Since $\mathcal{Q}$ commutes with $\mathcal{C}$, the generation is obvious. However, to ensure the negative type, it is necessary to have $s(\mathcal{QCQ}) < 0$. This condition is equivalent to $(G_{\mathcal{C}}(t))_{t\geq 0}$ having AEG.

7.1.1 Can we prove the convergence?

To this end we need to find an equation satisfied by the error which is defined as

$$\begin{aligned} y(t) &= v(t) - [\bar{v}(t) + \epsilon \tilde{v}_1(t/\epsilon)], \\ z(t) &= w(t) - [\tilde{w}_0(t/\epsilon) + \epsilon(\bar{w}_1(t) + \tilde{w}_1(t/\epsilon))]. \end{aligned} \tag{210}$$

Inserting (formally) the error into (195) we obtain

$$\begin{aligned} \partial_t y &= \mathcal{PSP}y + \mathcal{PSQ}z + \epsilon\mathcal{PSP}\tilde{v}_1 + \epsilon\mathcal{PSQ}\tilde{w}_1, \\ \partial_t z &= \mathcal{QSP}y + \mathcal{QSQ}z + \frac{1}{\epsilon}\mathcal{QCQ}z + \epsilon\mathcal{QSQ}\tilde{w}_1 \\ &\quad +\epsilon\mathcal{QSP}\tilde{v}_1 + \epsilon\mathcal{QSQ}\bar{w}_1 - \epsilon\partial_t\bar{w}_1. \end{aligned} \tag{211}$$

We observe that, denoting the total error $E(t) = y(t) + z(t)$, the error system (211) can be written as

$$\partial_t E = \left(\mathcal{S} + \frac{1}{\epsilon}\mathcal{C}\right) E + \epsilon\bar{F} + \epsilon\tilde{F}$$

Denoting by $(G_\epsilon(t))_{t\geq 0}$ the contractive semigroup generated by $\mathcal{S}+\epsilon^{-1}\mathcal{C}$, we get

$$\|E(t)\| \leq \|E(0)\| + \epsilon \int_0^t \|\bar{F}(s)\|ds + \epsilon \int_0^t \|\tilde{F}(s)\|ds.$$

It can be proved that $E(0) = O(\epsilon^2)$ and so this equation yields the error of approximation to be $O(\epsilon)$, which is not good as we have ϵ order terms in the approximation. A closer look at the term involving $\tilde{F}$ shows that it contains $e^{-t/\epsilon}$ which, upon integration, produces another ϵ so that the initial condition and the initial layer contribution to the error are $O(\epsilon^2)$. The fact that the contribution of $\bar{F}$ is also $O(\epsilon^2)$ is highly nontrivial but can be proved for a large class of problems.

It is important to note that the above considerations show that the presented asymptotic procedure **potentially** produces the convergence of the expected order. Since in most cases we work with unbounded operators, every step must be carefully justified.

7.1.2 How this works in practice: diffusion approximation of the telegraph equation

Here the compressed asymptotic procedure is applied to the telegraph equation

$$\partial_t \begin{bmatrix} v \\ w \end{bmatrix} = \mathcal{S} \begin{bmatrix} v \\ w \end{bmatrix} + \frac{1}{\varepsilon}\mathcal{C} \begin{bmatrix} v \\ w \end{bmatrix}, \tag{212}$$

where

$$\mathcal{S} = \begin{bmatrix} 0 & -b\partial_x \\ -c\partial_x & 0 \end{bmatrix}, \ \mathcal{C} = \begin{bmatrix} 0 & 0 \\ 0 & -d \end{bmatrix}.$$

or

$$\begin{aligned} \partial_t v + b\partial_x w &= 0, \\ \partial_t w + c\partial_x v + \frac{d}{\varepsilon} w &= 0, \end{aligned} \tag{213}$$

with constant coefficients b, c, d and a small parameter $\varepsilon > 0$.
The system (213) is supplemented by the initial conditions

$$v(0) = \overset{o}{v}, \qquad w(0) = \overset{o}{w}, \tag{214}$$

and the homogeneous Dirichlet conditions

$$v(-1,t) = v(1,t) = 0, \ t > 0. \tag{215}$$

To avoid the effect of a boundary layer, we assume that $\overset{o}{v}$ and $\overset{o}{w}$ are three times differentiable and

$$\partial_x \overset{o}{v}(\pm 1) = 0 \ , \ \overset{o}{w}(\pm 1) = 0 \ , \ \partial_{xx} \overset{o}{w} (\pm 1) = 0. \tag{216}$$

This system may describe the voltage and the current in a telegraphic cable, where the a, b, c and d are the loss coefficient, the resistance, the capacity and the self induction respectively or it can be considered as a simplified (two-velocity) linear Boltzmann equation because the relevant spectral properties are similar.

The diffusion approximation can be derived from (208), using the compressed asymptotic procedure, by taking

$$\mathcal{P}\begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix},$$

$$\mathcal{Q}\begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ w \end{bmatrix}.$$

Then

$$\mathcal{QSP}\begin{bmatrix} v \\ 0 \end{bmatrix} = \mathcal{Q}\begin{bmatrix} 0 & -b\partial_x \\ -c\partial_x & 0 \end{bmatrix}\begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -c\partial_x v \end{bmatrix},$$

$$\mathcal{PSQ}\begin{bmatrix} 0 \\ w \end{bmatrix} = \mathcal{P}\begin{bmatrix} 0 & -b\partial_x \\ -c\partial_x & 0 \end{bmatrix}\begin{bmatrix} 0 \\ w \end{bmatrix} = \begin{bmatrix} -b\partial_x w \\ 0 \end{bmatrix},$$

The inverse $(\mathcal{QCQ})^{-1}$ is given by

$$(\mathcal{QCQ})^{-1}\left[\begin{array}{c} 0 \\ w \end{array}\right] = \left[\begin{array}{c} 0 \\ -w/d \end{array}\right].$$

Then

$$\mathcal{PSQ}(\mathcal{QCQ})^{-1}\mathcal{QSP}\left[\begin{array}{c} v \\ 0 \end{array}\right] = \mathcal{PSQ}\left[\begin{array}{c} 0 \\ \frac{c}{d}\partial_x v \end{array}\right] = \left[\begin{array}{c} -\frac{bc}{d}\partial_{xx} v \\ 0 \end{array}\right].$$

Hence the approximation diffusion equation, as given by (208), is

$$\partial_t \rho = \varepsilon \frac{bc}{d}\partial^2_{xx}\rho. \tag{217}$$

The uncorrected initial condition is $\rho(0) = \hat{\rho}(0) = \overset{o}{v}$, whereas the corrected one can be derived from (209) using $\mathcal{PSQ}$ and $(\mathcal{QCQ})^{-1}$ as calculated above, which gives

$$\bar{\rho}(0) = \overset{o}{v} \; -\varepsilon \mathcal{PSQ}(\mathcal{QCQ})^{-1} \overset{o}{w} = \overset{o}{v} - \varepsilon \frac{b}{d}\partial_x \overset{o}{w} . \tag{218}$$

The initial layer is derived from (206) and is given by

$$\tilde{v}_1(\tau) = \mathcal{PSQ}(\mathcal{QCQ})^{-1} e^{\tau \mathcal{QCQ}} \overset{o}{w} = \frac{b}{d} e^{-d\tau}\partial_x \overset{o}{w}, \tag{219}$$

where $\tau = t/\varepsilon$.

Let us denote

$$D_3 = \{u \in W^3_2([-1,1]);\; u|_{x=\pm 1} = 0,\; \partial^2_{xx} u|_{x=\pm 1} = 0\}$$

The following theorem is true.

Theorem 82 *If* $\overset{o}{v}, \overset{o}{w} \in D_3$ *and the compatibility conditions (216) are satisfied. Then there is a constant* C *such that*

$$\|v(t) - \rho(t) - \varepsilon \tilde{v}_1(t/\varepsilon)\| \leq C\varepsilon^2$$

uniformly on $[0, \infty)$

7.1.3 Age structured population model

A seemingly similar system is offered by

$$\partial_t \mathbf{n} = \mathcal{H}\mathbf{n} + \mathcal{M}\mathbf{n} + \frac{1}{\varepsilon}\mathcal{K}\mathbf{n}, \tag{220}$$

where $\mathbf{n} = (n_1, \ldots, N)$, $\mathcal{H} = diag\{-\partial_a, \ldots, -\partial_a\}$, $\mathcal{M} = diag\{-\mu_1, \ldots, -\mu_N\}$, $\mathcal{K} = \{k_{ij}\}_{1\leq i,j\leq N}$. Here n_i is the population density of fish in patch i, a is the age, $\mu_i(a)$ is the mortality rate, and the coefficients k_{ij} represent the migration rates from patch j to patch i, $j \neq i$. The system was introduced to describe evolution of a continuous age-structured population of sole which, however, is further divided into patches (say, egg, larvae, juvenile and adult). The characteristic feature of the population is daily vertical migration provoked by light intensity of which is highly dependent on patches. The small parameter ε corresponds to the fact that the migration processes occur at a much faster time scale than the demographic ones (aging and death). This system must be supplemented by the boundary condition of the McKendrick-Von Forester type

$$\mathbf{n}(0,t) = \int\limits_0^\infty \mathcal{B}(a)n(a,t)da \tag{221}$$

where $\mathcal{B}(a) = diag\{b_1(a), \ldots, b_N(a)\}$ gives the fertility at age a and patches 1 to N. The initial condition is given by

$$\mathbf{n}(a,0) = \Phi(a). \tag{222}$$

The transition matrix $\mathcal{K}$ is a typical transition matrix (of a time-continuous process); that is off-diagonal entries are positive and columns sum up to 0. We further assume that it generates an irreducible (n-dimensional) semigroup. Thus, 0 is the dominant eigenvalue of $\mathcal{K}$ with a positive eigenvector $\mathbf{e}$ which will be fixed to satisfy $\mathbf{1} \cdot \mathbf{e} = 1$, where $\mathbf{1} = (1, 1 \ldots, 1)$. The vector $\mathbf{e} = (e_1, \cdots, e_N)$ represents the stable patch structure; that is, asymptotic distribution of the population into the patches. Thus, it is reasonable to approximate

$$e_i = \frac{n_i}{n}, \quad i = 1, \ldots, N$$

where $n = \sum_{i=1}^{N} n_i$. Adding together equations in (220) and using the above we obtain

$$\partial_t n = -\partial_a n - \mu^*(a)n \tag{223}$$

where $\mu^* = \mathbf{1} \cdot \mathcal{M}\mathbf{e} = \sum_{i=1}^{N} \mu_i e_i$ is the so-called 'aggregate' mortality. This model, supplemented with appropriate averaged boundary condition is called the aggregated model and is expected to provide averaged approximate description of the population.

The assumptions allow to perform the compressed asymptotic analysis. The spectral projections $\mathcal{P}, \mathcal{Q} : \mathbb{R}^n \to \mathbb{R}^n$ are given by

$$\mathcal{P}\mathbf{x} = (\mathbf{1} \cdot \mathbf{x})\mathbf{e}, \quad \mathcal{Q}\mathbf{x} = \mathbf{x} - (\mathbf{1} \cdot \mathbf{x})\mathbf{e}$$

which gives the hydrodynamical space $V := Span\{\mathbf{e}\}$ and the kinetic space

$$W = Im\mathcal{Q} = \{\mathbf{x};\ \mathbf{1} \cdot \mathbf{x} = 0\},$$

as well as the solution decomposition

$$\mathbf{n} = \mathcal{P}\mathbf{n} + \mathcal{Q}\mathbf{n} = v + w = p\mathbf{e} + w$$

where $p = p(a,t)$ is a scalar function. Applying the projections to both sides of (220) we get

$$\begin{aligned} \partial_t v &= \mathcal{P}(\mathcal{H} + \mathcal{M})\mathcal{P}v + \mathcal{P}(\mathcal{H} + \mathcal{M})\mathcal{Q}w \\ \partial_t w &= \mathcal{Q}(\mathcal{H} + \mathcal{M})\mathcal{Q}w + \mathcal{Q}(\mathcal{H} + \mathcal{M})\mathcal{P}v + \frac{1}{\varepsilon}\mathcal{Q}\mathcal{K}\mathcal{Q}w, \end{aligned}$$

with projected initial conditions $v(0) = \overset{o}{v},\ w(0) = \overset{o}{w}$.

Denoting again by $\bar{v}$ and $\bar{w}$ the bulk part of the solution and substituting the expansion for $\bar{w}$ into (224) we obtain as before

$$\bar{w}_0 \equiv 0,$$

$$\bar{w}_1 = -(\mathcal{Q}\mathcal{K}\mathcal{Q})^{-1}\mathcal{Q}(\mathcal{H} + \mathcal{M})\mathcal{P}\bar{v}.$$

Inserting the expressions for $\bar{w}_0$ and $\bar{w}_1$ into the expansion of the hydrodynamic part of the system (224) gives the approximate 'diffusion' equation

$$\partial_t \bar{v} = \mathcal{P}(\mathcal{H}+\mathcal{M})\mathcal{P}\bar{v} - \varepsilon \mathcal{P}(\mathcal{H}+\mathcal{M})\mathcal{Q}(\mathcal{Q}\mathcal{K}\mathcal{Q})^{-1}\mathcal{Q}(\mathcal{H}+\mathcal{M})\mathcal{P}\bar{v}.$$

The explicit expressions for the involved operators can be calculated as

$$\begin{aligned}
\mathcal{P}(\mathcal{H}+\mathcal{M})\mathcal{P}\bar{v} &= -(\partial_a p - p(\mathbf{1}\cdot\mathcal{M}\mathbf{e})\mathbf{e}, \\
\mathcal{Q}(\mathcal{H}+\mathcal{M})\mathcal{P}\bar{v} &= -p(\mathbf{1}\cdot\mathcal{M}\mathbf{e} - \mathcal{M})\mathbf{e}, \\
\mathcal{P}(\mathcal{H}+\mathcal{M})\mathcal{Q}\mathbf{x} &= -(\mathbf{1}\cdot\mathcal{M}\mathbf{x} - \mathbf{1}\cdot\mathcal{M}\mathbf{x})\mathbf{e},
\end{aligned}$$

and, denoting by $\mathbf{h}$ the unique solution in $W = \mathcal{Q}X$ of

$$\mathcal{K}\mathbf{h} = -(\mathbf{1}\cdot\mathcal{M}\mathbf{e} - \mathcal{M})\mathbf{e}$$

we obtain

$$\mathcal{P}(\mathcal{H}+\mathcal{M})\mathcal{Q}(\mathcal{Q}\mathcal{K}\mathcal{Q})^{-1}\mathcal{Q}(\mathcal{H}+\mathcal{M}\mathcal{P}\bar{v} = p(\mathbf{1}\cdot\mathcal{M}\mathbf{h}).$$

Therefore

$$\partial_t p = -\partial_a p + p(\mathbf{1}\cdot\mathcal{M}\mathbf{e} + \epsilon\mathbf{1}\cdot\mathcal{M}\mathbf{h})$$

or, taking into account the form of $\mathcal{M}$, we obtain

$$\partial_t p = -\partial_a p - \mu^* p + \epsilon(\mathbf{1}\cdot\mathcal{M}\mathbf{h})p$$

so that the asymptotic procedure recovers the aggregated model (223) as well as provides its first order correction.

We note that, contrary to the telegraph system, here we haven't obtained a diffusion equation. The (mathematical) reason for this that in the telegraph equation the transport operator appears on the anti-diagonal and thus provides 'mixing' of the hydrodynamic and kinetic parts of the equation. Here the transport occurs only on the diagonal hence, at the transport level, the patches are not mixed and this feature is preserved in the approximating equation.

In this model it is impossible to neglect effects of the boundary conditions and thus one needs to analyse the boundary and corner layers as well as the initial layer. However, we will not discuss them here.

7.1.4 Fokker-Planck equation of Brownian motion

We conclude with a brief discussion of a more complicated example of the Fokker-Planck equation describing n-dimensional Brownian motion. The collision operator $\mathcal{C}$ now is given by the three-dimensional differential operator

$$(\mathcal{C}u)(x,\xi) = \partial_\xi(\xi + \partial_\xi)u(x,\xi), \tag{1}$$

$x, \xi \in \mathbb{R}^n$ and the streaming operator **S** is of the form

$$(\mathcal{S}u)(x,\xi) = \xi\partial_x u(x,\xi). \tag{2}$$

Here u is the particle distribution function in the phase space, x denotes the position and ξ the velocity of the particle.

The Fokker-Planck operator can be transformed to the well-known harmonic oscillator operator: for the function $u(\xi)$ we define $\xi = \sqrt{2}\zeta \in \mathbb{R}^n$ and

$$y(\zeta) = (\mathsf{A}_n u)(\zeta) := (\sqrt{2})^{\frac{n}{2}} e^{\frac{|\zeta|^2}{2}} u(\sqrt{2}\zeta). \tag{3}$$

This is an isometry of the space $L_2(\mathbb{R}^n, e^{\frac{|\xi|^2}{2}} d\xi)$ onto $L_2(\mathbb{R}^n, d\zeta)$ which transforms the Fokker-Planck collision operator $\mathcal{C}$ into

$$\tilde{C}y = \frac{1}{2(\sqrt{2})^{n/2}} e^{-\frac{|\zeta|^2}{2}} \left(\partial_\zeta^2 y - |\zeta|^2 y + ny\right). \tag{4}$$

Dropping the normalizing factor we arrive at the harmonic oscillator operator in $L_2(\mathbb{R}^n)$, denoted hereafter by H,

$$(Hy)(\zeta) = \partial_\zeta^2 y(\zeta) - |\zeta|^2 y(\zeta) + ny(\zeta). \tag{5}$$

To analyse this operator we introduce the sesquilinear form

$$h(\phi,\psi) = \int\limits_{\mathbb{R}^n} \left(\partial_\zeta\phi\partial_\zeta\bar{\psi} + |\zeta|^2\phi\bar{\psi} + \phi\bar{\psi}\right) d\zeta, \tag{6}$$

defined originally on $C_0^\infty(\mathbb{R}^n)$ and the Hilbert space H_1 defined as the closure of $C_0^\infty(\mathbb{R}^n)$ with respect to the norm $\|\phi\|_{H_1} = \sqrt{h(\phi,\phi)}$. Let A_h denote the operator associated with h It follows that the spectrum of A_h consists only of eigenvalues and the operator itself can be expressed in terms of the series of its eigenfunctions. Using the separation of variables and the one-dimensional theory of the harmonic oscillator we obtain the following expression for the eigenfunctions of A_h:

$$H_\alpha^{(n)}(\zeta) = \frac{(-1)^{|\alpha|}}{(2^{|\alpha|}\pi^{n/2}\alpha!)^{1/2}} e^{\frac{|\zeta|^2}{2}} \partial^\alpha e^{|\zeta|^2} = \prod_{i=1}^{n} H_{\alpha_i}^{(1)}(\zeta_i), \tag{7}$$

where $\zeta \in \mathbb{R}^n$ and $\alpha = (\alpha_1, \ldots \alpha_n)$ is a multi-index.

$H_m^{(1)}$ is the normalized one-dimensional Hermite function corresponding to the eigenvalue $\lambda_m = 2m+1$

$$H_m(\zeta) := \frac{(-1)^m}{\sqrt{\sqrt{\pi}2^m m!}} e^{\frac{\zeta^2}{2}} \partial_\zeta^m e^{-\zeta^2}. \tag{8}$$

Let C denote the Fokker-Planck collision operator obtained from A_h by the inverse transformation (3), and thus corresponding to the differential expression (1). For $k = 1, \ldots, n$ and the multi-index $\beta = (\beta_1, \ldots, \beta_k)$ we define

$$\Phi_\beta^{(k)} := \mathsf{A}_k^{-1} H_\beta^{(k)},$$

that is,

$$\Phi_\alpha^{(n)}(\xi) = \frac{(-1)^{|\alpha|}}{(2\pi)^{n/4}\sqrt{\alpha!}} \partial^\alpha e^{-\frac{|\xi|^2}{2}} = \prod_{i=1}^{n} \Phi_{\alpha_i}^{(1)}(\xi_i). \tag{9}$$

Since A_k is an isometric isomorphism, the family $\left\{\Phi_\alpha^{(n)}\right\}_{\alpha\in\mathbb{N}^n}$ forms an orthonormal basis in $L_2(\mathbb{R}^n, e^{\frac{|\xi|^2}{2}} d\xi)$. We have therefore

$$u = \sum_{|\alpha|=0}^{\infty} u_\alpha \Phi_\alpha^{(n)} \tag{10}$$

and

$$\mathsf{C}u = -\sum_{|\alpha|=1}^{\infty} |\alpha| u_\alpha \Phi_\alpha^{(n)}, \tag{11}$$

so that it is clear that C is dissipative and satisfies all assumptions of the general theory.

To conclude we derive the form of the diffusion equation. To this end we express operator $\mathcal{S}$ in terms of eigenfunctions $\Phi_\alpha^{(n)}$. Let us adopt the following convention

$$\alpha(i, \pm 1) = (\alpha_1, \ldots, \alpha_i \pm 1, \ldots, \alpha_n).$$

The Hermite functions satisfy the following recurrence formula for $\Phi_\alpha^{(n)}$. Let $i = 1, \ldots, n$, then

$$\xi_i \Phi_\alpha^{(n)} = \sqrt{\alpha_i + 1}\,\Phi^{(n)}_{\alpha(i,+1)}(\xi) + \sqrt{\alpha_i}\,\Phi^{(n)}_{\alpha(i,-1)}(\xi). \tag{12}$$

If some $\alpha_i = 0$, then naturally the second summand vanishes. By Eq. (12) we obtain formally

$$\mathcal{S}u = -\sum_{k=1}^{n} \partial_k \left(\sum_{|\alpha|=0}^{\infty} \left(\sqrt{\alpha_k}\, u_{\alpha(k,-1)} + \sqrt{\alpha_k + 1}\, u_{\alpha(k,+1)} \right) \Phi_\alpha^{(n)} \right). \tag{13}$$

The hydrodynamic space is clearly spanned by $\Phi_0^{(n)}$. Hence we denote $\bar{v} = \bar{\rho}\Phi_0^{(n)}$ and $\tilde{v}_1 = \tilde{\rho}\Phi_0^{(n)}$. Introducing the notation

$$\mathbf{0}(i; l) = (0, \ldots, l, \ldots, 0)$$

and

$$\mathbf{0}(i, j; k, l) = (0, \ldots, k, \ldots, l, \ldots, 0),$$

where l (resp. (k, l)) are in the i-th (resp. i-th and j-th) place, we get

$$\mathcal{S}\bar{v} = -\sum_{k=1}^{n} \partial_k \Phi^{(n)}_{\mathbf{0}(k;1)} \bar{\rho}$$

and further

$$\begin{aligned}
&\mathcal{SQ}(\mathcal{QCQ})^{-1}\mathcal{QSP}\bar{v} \\
&= -\sum_{k=1}^{n} \partial_k \left(\sum_{l=1, l\neq k}^{n} \partial_l \Phi^{(n)}_{\mathbf{0}(k,l;1,1)} + \partial_k \left(\sqrt{2}\,\Phi^{(n)}_{\mathbf{0}(k,2)} + \Phi_0^{(n)} \right) \right) \bar{\rho}.
\end{aligned}$$

Projecting this onto $\Phi_0^{(n)}$ we get the diffusion operator in the form

$$\mathcal{PSQ}(\mathcal{QCQ})^{-1}\mathcal{QSP}\bar{v} = -\Delta_x \bar{v}.$$

Similarly for the corrector of the initial value we obtain

$$\mathcal{PSQ}(\mathcal{QCQ})^{-1} \overset{\circ}{w} = \Phi_0^{(n)} \sum_{k=1}^{n} \partial_k \overset{\circ}{u}_{\mathbf{0}(k;1)}$$

and the initial layer corrector $\tilde{v}_1$ will have the form

$$\tilde{v}_1\left(\frac{t}{\epsilon}\right) = e^{-\frac{t}{\epsilon}} \Phi_0^{(n)} \sum_{k=1}^{n} \partial_k \overset{\circ}{u}_{\mathbf{0}(k;1)},$$

where $\overset{\circ}{u}_{\mathbf{0}(k;1)}$ is the first moment of the initial value for u.

To formulate the final result of this section we introduce

$$\varrho(t,x) := \int_{\mathbb{R}} u(t,x,\xi) d\xi,$$

where u is the solution of the initial value problem for the Fokker-Planck equation of the Brownian motion. Let $\overset{\circ}{u} \in W_1^3(\mathbb{R}^n, L_2(\mathbb{R}^n, e^{\frac{|\xi|^2}{2}} d\xi))$, then

$$\left\| \varrho(t) - \bar{\rho}(t) - \epsilon \tilde{\rho}\left(\frac{t}{\epsilon}\right) \right\|_{L_2(\mathbb{R}^n \times \mathbb{R}^n, e^{\frac{|\xi|^2}{2}} dx d\xi)} = O(\epsilon^2) \tag{14}$$

uniformly for t in bounded intervals of $[0, \infty[$. Here $\bar{\rho}$ is the solution of the following initial value problem

$$\begin{aligned} \partial_t \bar{\rho} &= \epsilon \partial_x^2 \bar{\rho}, \\ \bar{\rho}(0) &= \overset{\circ}{u}_0 - \epsilon \sum_{k=1}^{n} \partial_{x_k} \overset{\circ}{u}_{\mathbf{0}(k;1)}, \end{aligned}$$

and the function $\tilde{\rho}$ in the initial layer corrector $\tilde{v}_1 = \tilde{\rho}\Phi_0$ is given by

$$\tilde{\rho}\left(\frac{t}{\epsilon}\right) = e^{-t/\epsilon} \sum_{k=1}^{n} \partial_{x_k} \overset{\circ}{u}_{\mathbf{0}(k;1)}.$$

References

[1] Y.A. Abramovich, Ch.D. Aliprantis, *An Invitation to Operator Theory*, AMS, Providence, Rhode Island, 2002.

[2] Y. A. Abramovich, C. D. Aliprantis, *Problems in Operator Theory*, American Mathematical Society, Providence, RI, 2002.

[3] Ch.D. Aliprantis, O. Burkinshaw, *Positive Operators*, Academic Press, Inc., Orlando, 1985.

[4] W. J. Anderson, *Continuous-Time Markov Chains. An Applications-Oriented Approach*, Springer Verlag, New York, 1991.

[5] O. Arino, Some Spectral Properties for the Asymptotic Behaviour of Semigroups Connected to Population Dynamics, *SIAM Review*, **34**(3), 1992, 445-476.

[6] L. Arlotti, The Cauchy problem for the linear Maxwell-Boltzmann equation, *J. Differential Equations*, **69** (2), (1987), 166–184.

[7] L. Arlotti, A perturbation theorem for positive contraction semigroups on L^1-spaces with applications to transport equations and Kolmogorov's differential equations, *Acta Appl. Math.*, **23**(2), (1991), 129–144.

[8] L. Arlotti, J. Banasiak, Strictly substochastic semigroups with application to conservative and shattering solutions to fragmentation equations with mass loss, *J. Math. Anal. Appl.*, **293**(2), (2004), 693–720.

[9] J. Banasiak, L. Arlotti, *Positive perturbations of semigroups with applications*, Springer, London, 2006.

[10] J. Banasiak, G. Frosali, G. Spiga, An interplay between elastic and inelastic scattering in models of extended kinetic theory and their hy-

drodynamic limits - reference manual, *Transport Theory Stat. Phys.*, **31**(3), 2002, 187-248.

[11] J. Banasiak, On an extension of Kato–Voigt perturbation theorem for substochastic semigroups and its applications, *Taiwanese J. Math.*, **5**(1), 2001, 169–191.

[12] J. Banasiak, On non-uniqueness in fragmentation models, *Math. Methods Appl. Sci.*, **25**(7), 2002, 541–556.

[13] . J. Banasiak and M. Lachowicz, Topological chaos for birth–and–death–type models with proliferation, *Math. Models Methods Appl. Sci.*, **12** (6), (2002), 755–775.

[14] J. Banasiak, A complete description of dynamics generated by birth-and-death problems: A semigroup approach, in: *Mathematical Modelling of Population Dynamics*, R. Rudnicki (Ed.), Banach Center Publications, vol. 63, 2004, 165–176.

[15] J. Banasiak, Conservative and shattering solutions for some classes of fragmentation equations, *Math. Models Methods Appl. Sci.*, **14**(4), 2004, 483–501.

[16] J. Banasiak, M. Moszyński, A generalization of Desch-Schappacher-Webb criteria for chaos, *Discr. Cont. Dyn. Sys.-A*, **12**(5), (2005), 959-972,

[17] J. Banasiak, M. Lachowicz, A generalization of Kato's perturbation theorem in Banach lattices, *Studia Mathematica*, (2006), to appear.

[18] . J. Banasiak, M. Lachowicz, M. Moszyński, Chaotic behavior of semigroups related to the process of gene amplification–deamplification with cells' proliferation, *Math. Biosciences*, (2006), in print.

[19] J. Banasiak, M. Lachowicz, M. Moszyński, Semigroups for generalized birth-and-death equations in l^p spaces, *Semigroup Forum*, 2006, to appear.

[20] F. E. Browder, On the spectral theory of elliptic differential operators, *Math. Ann*, **142**, 1961, 22–130.

[21] Ph. Clément, H. J. A. M. Heijmans, S. Angenent, C. J. van Duijn, B. de Pagter, *One-parameter Semigroups*, CWI Monograph, North Holland, Amsterdam, New York, 1987.

[22] . W. Desch, W. Schappacher and G.F. Webb, Hypercyclic and chaotic semigroups of linear operators, *Ergodic Theory Dynam. Systems*, **17**, 1997, 793–819.

[23] R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, 2nd edn., Addison-Wesley, New York, 1989.

[24] N. Dunford, J.T. Schwartz, *Linear Operators I*, Wiley Interscience, New York, 1958.

[25] S. El Mourchid, The imaginary point spectrum and hypercyclicity, *Semigroup Forum*, (2006), to appear.

[26] K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer Verlag, New York, 1999.

[27] G. Frosali, C. van der Mee, S.L. Paveri-Fontana, Conditions for runaway phenomena in the kinetic theory of particle swarms. *J. Math. Phys.*, **30** (5), 1989, 1177–1186.

[28] G. Frosali, C. van der Mee, F. Mugelli, A characterization theorem for the evolution semigroup generated by the sum of two unbounded operators, *Math. Methods Appl. Sci.*, **27**(6), (2004), 669–685.

[29] E. Hille, R. S. Phillips, *Functional Analysis and Semi-groups*, Colloquium Publications, v. 31, American Mathematical Society, Providence, RI, 1957.

[30] I. Jeon, Stochastic fragmentation and some sufficient conditions for shattering transition, *J. Korean Math. Soc.*, **39**(4), 2002, 543–558.

[31] T. Kato, On the semi-groups generated by Kolmogoroff's differential equations, *J. Math. Soc. Jap.*, **6**(1), (1954), 1–15.

[32] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Springer Verlag, Berlin, 1984.

[33] M. Kimmel, A. Świerniak and A. Polański, Infinite–dimensional model of evolution of drug resistance of cancer cells, *J. Math. Systems Estimation Control* **8**(1), 1998, 1–16.

[34] W.A.J. Luxemburg, A.C. Zaanen, *Riesz Spaces*, North-Holland, Amsterdam, 1971.

[35] D. J. McLaughlin, W. Lamb, A. C. McBride, A semigroup approach to fragmentation models, *SIAM J. Math. Anal.*, **28** (5), (1997), 1158–1172.

[36] J. R. Mika, J. Banasiak, *Singularly Perturbed Evolution Equations with Applications in KineticTheory*, World Sci., Singapore, 1995.

[37] M. Mokhtar-Kharroubi, On the convex compactness property for the strong operator topology and related topics, Math. Meth. Appl. Sci., **27**, 2004, 681–701.

[38] R. Nagel (ed.), *One parameter semigroups of positive operators*, LNM 1184, Springer Verlag, Berlin, New York, 1986.

[39] J. van Neerven, *The Asymptotic Behaviour of Semigroups of Linear Operators*, Operator Theory: Advances and Applications 88, Birkhäuser Verlag, Basel, 1996.

[40] J. R. Norris, *Markov Chains*, Cambridge University Press, Cambridge, 1998.

[41] R. Nussbaum, The radius of the essential spectrum, *Duke Math. J.*, **37**, 1970, 473–478.

[42] B. de Pagter, Irreducible compact operators, *Math. Z*, **192**, 1986, 149–153.

[43] A. Pazy *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer Verlag, New York, 1983.

[44] G. Schlüchtermann, On weakly compact operators, *Math. Ann.*, **292**, 1992, 262–266.

[45] J. Voigt, On substochastic C_0-semigroups and their generators, *Transport Theory Statist. Phys.*, **16** (4–6), (1987), 453–466.

[46] J. Voigt, A Perturbation Theorem for the Essential Spectral Radius of Strongly Continuous Semigroups, *Mh. Math.*, **90**, 1980, 153–161.

[47] K. Yosida, *Functional Analysis*, 5th Ed., Springer Verlag, Berlin, 1978.