PERTURBATIONS OF POSITIVE SEMIGROUPS

MMATH PROJECT
CHARLIE COLLIER
SUPERVISOR: KARSTEN MATTHIES

ABSTRACT. In this paper, we discuss the theory of semigroups of bounded operators and how they characterise the solution of the differential equation $u' = Au$. The fundamental results are given in the form of the Hille-Yosida and Lumer-Phillips theorems, which characterise the generator of such semigroups. Further, we discuss perturbations of semigroups and positivity of elements and linear operators, and how this theory relates to the particular example of a birth-and-death process in an analytic framework.

CONTENTS

Date: April 5, 2019.

1. Introduction

The aim of this paper is to study the theory of semigroups and how they can be used to solve differential equations of the form $u'(t) = Au(t)$ for $t > 0$ for some initial condition in a Banach space X. It is known that for *nice* operators A, for example matrices or bounded linear operators in X, the solution is characterised by $u(t) = e^{tA}u_0$, however when A is not necessarily a bounded operator then it is not possible to understand the matrix exponential e^{tA}. Thus we need a tool to generalise the exponential and find such solutions in this case.

We begin in chapter 2 by recalling some facts about bounded and unbounded operators, and study the idea of closedness and density of an operator. These concepts are in fact necessary conditions for the operator A to generate a *semigroup*. We also recall the uniform boundedness theorem for a collection of bounded operators and the closed graph theorem, as uniform bounds and closedness are of importance when studying semigroups. We also introduce the notion of the resolvent and spectrum of a closed and densely defined operator, proving some characteristics of such objects, in particular that the resolvent is always bounded for closed operators (of which we mostly only consider later on).

In chapter 3 we discuss the notion of *positivity*. In some spaces, for example the reals $\mathbb{R}$, we are used to considering the natural order $\leq$ which is reflexive, anti-symmetric and transitive, and this ordering gives the notion of a positive element (namely, $x \geq 0$). However in a general Banach space X, such a notion may not be explicitly obvious, and so we introduce a partial ordering on X to deal with this. With such constructions, we can generalise a lot of properties which hold in $(\mathbb{R}, |\cdot|)$, for example the triangle inequality. Using positivity of a space allows us to define what it means for an operator to be positive in a natural way.

In chapter 4, we introduce the theory of strongly continuous semigroups. As discussed above, this will allow us to define the solution of differential equations $u' = Au$ when A is an unbounded operator, provided that our initial condition is in the domain of A. After proving some preliminary results about semigroups, we discuss the Hille-Yosida and Lumer-Phillips generation theorems. These are theorems which give necessary and sufficient conditions for a linear operator to be the generator of a semigroup, so in particular if A is not at least closed and densely defined, then it doesn't stand a chance of generating a semigroup. The Lumer-Phillips theorem introduces the notion of a dissipative operator in relation to semigroups of contractions.

In chapter 5, we consider a different approach to solving $u' = Au$ using the perturbation of semigroups. Indeed, it may be possible to deconstruct A as the sum of two operators, of which one may clearly generate a strongly continuous semigroup, and so the perturbation theory will allow us to see if the perturbing operator is nice enough so that it preserves properties from the semigroup generated by A. We state and prove an important spectral criterion which is used throughout the chapter. We consider some particular perturbations, in fact bounded perturbations and *Miyadera* perturbations. Further, we state a theorem relating to the perturbation of a positive semigroup of contractions, which is of use in chapter 6.

In chapter 6, we apply some of the above theory to the particular example of a birth and death process. Using standard modelling relating to birth and death rates, we obtain the equation $u' = \mathcal{A}u + \mathcal{B}u$ for some operators $\mathcal{A}$ and $\mathcal{B}$. From this, we can use the theory of semigroups of contractions and perturbations to characterise the solution in some sense.

2. Linear Operators

In order to discuss the notion of a positive operator, we first recall some facts about operators and some fundamental theorems, namely the Banach-Steinhaus, and closed graph theorem. We also recall the spectrum and resolvent of a closed, densely defined operator since this will be of importance when discussing semigroups and their perturbations.

For further discussions of these concepts, see [1, §2.1.2].

2.1. **Definitions.**

In the following, X and Y denote Banach spaces over the field $\mathbb{R}$ or $\mathbb{C}$.

The respective norms are denoted by $\|\cdot\|_X$ or $\|\cdot\|_Y$, or just $\|\cdot\|$ if the ambient space is clear.

Definition 2.1. An *operator* from X to Y is a linear mapping $A : \mathscr{D}(A) \to Y$ where $\mathscr{D}(A)$ is a linear subspace of the space X, called the *domain* of A. We often denote an operator with its domain as $(A, \mathscr{D}(A))$. The space of all such operators is the set $L(X, Y)$.

If $\mathscr{D}(A) = X$, then we say that A is an *everywhere defined* operator.

If $\overline{\mathscr{D}(A)} = X$, then we say that A is a *densely defined* operator.

Remark 2.2. If $Y = \mathbb{F}$ is a field of scalars, then elements of $L(X, Y)$ are called *functionals.*

Definition 2.3. An everywhere defined operator A is *bounded* if there exists a constant $M \geq 0$ such that $\|Ax\| \leq M\|x\|$ for all $x \in X$. In such a case, we write $A \in \mathcal{L}(X, Y)$.

We write $\mathcal{L}(X)$ for $\mathcal{L}(X, X)$, and note that $\mathcal{L}(X, Y) \subseteq L(X, Y)$.

For any operator A in $\mathcal{L}(X, Y)$, its *operator norm* is defined by

$$\|A\| := \sup_{\substack{x\in X\\ \|x\|\leq 1}} \|Ax\|.$$

We also have the following equivalent definitions for the operator norm,

$$\|A\| = \sup_{\substack{x\in X\\ \|x\|=1}} \|Ax\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Ax\|}{\|x\|}.$$

In fact, the operator norm is the minimum value of M for which $\|Ax\| \leq M\|x\|$ for all $x \in X$.

With the operator norm, $\mathcal{L}(X, Y)$ becomes a Banach space.

Definition 2.4. Let $(A, \mathscr{D}(A))$ be an operator in X and $Y \subseteq X$.

Then the *part of* A *in* Y is the function A_Y defined by $A_Y y := Ay$ with domain

$$\mathscr{D}(A_Y) = \{x \in \mathscr{D}(A) \cap Y : Ax \in Y\}.$$

Definition 2.5. Let $(A, \mathscr{D}(A))$ and $(B, \mathscr{D}(B))$ be two operators in $L(X, Y)$. Then we say that B is an *extension of* A, denoted $A \subseteq B$, if $\mathscr{D}(A) \subseteq \mathscr{D}(B)$ and $Ax = Bx$ for all $x \in \mathscr{D}(A)$.

We now consider *unbounded* operators $(A, \mathscr{D}(A))$, so operators in $L(X, Y)\backslash\mathcal{L}(X, Y)$.

Definition 2.6. The *graph norm* of A is the norm $\|\cdot\|_A$ on $\mathscr{D}(A)$ defined by

$$\|x\|_A := \|x\| + \|Ax\|, \quad x \in \mathscr{D}(A).$$

It is easy to see that $(\mathscr{D}(A), \|\cdot\|_A)$ is indeed a normed space.

Definition 2.7. The *graph* of an operator A is the set $\mathscr{G}(A) := \{(x, Ax) \in X \times Y : x \in \mathscr{D}(A)\}$.

Definition 2.8. An operator A is *closed* if $\mathscr{G}(A)$ is a closed subset of $X \times Y$.

Here, closedness is with respect to the norm $\|(x,y)\|_{X\times Y} = \|x\|_X + \|y\|_Y$ on $X \times Y$.

If X and Y are both Banach spaces, then $X \times Y$ is Banach with this norm.

Proposition 2.9. *A is a closed operator if and only if for all sequences $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A)$ such that $x_n \to x$ in X and $Ax_n \to y$ in X as $n \to \infty$, we have that $x \in \mathscr{D}(A)$ and $Ax = y$.*

Proof. This follows since A is closed iff $\mathscr{G}(A) = \overline{\mathscr{G}(A)}$, that is $\mathscr{G}(A)$ contains its limit points. □

Lemma 2.10. *If $A \in L(X,Y)$ is a closed operator, then so is $-A$ and $A - \lambda I$ for any $\lambda \in \mathbb{F}$. Further, if $X = Y$ and A is invertible then A^{-1} is a closed operator.*

Proof. We note that $\mathscr{D}(A) = \mathscr{D}(-A) = \mathscr{D}(A-\lambda)$. If A is closed, consider a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A)$ such that $x_n \to x$ in X and $(-A)x_n \to y$ in Y as $n \to \infty$. Then we have that $Ax_n \to -y$, and so $x \in \mathscr{D}(A)$ with $Ax = -y$ since A is closed, or equivalently $(-A)x = y$, so that $-A$ is a closed operator. A similar proof follows to show that $A - \lambda$ is closed. Now if $A : \mathscr{D}(A) \to X$ is invertible, A^{-1} exists. Then since A is a closed operator, $\mathscr{G}(A)$ is closed in $X \times X$ and any set in $\mathscr{G}(A)$ is of the form (x, Ax) for $x \in \mathscr{D}(A)$. Because A^{-1} is onto, there exists a $y \in X$ with $A^{-1}y = x$, or equivalently $y = Ax$. Hence $\mathscr{G}(A) = \{(A^{-1}y, y) : y \in X\}$ is closed, and changing the order of the coordinates gives $\mathscr{G}(A^{-1})$, thus A^{-1} is closed. □

Definition 2.11. An operator A in X is *closable* if $\overline{\mathscr{G}(A)}$ is itself the graph of some operator. That is, if $(0, y) \in \overline{\mathscr{G}(A)}$ then $y = 0$, or equivalently if $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A)$ with $x_n \to 0$ in X and $Ax_n \to y$ in Y, then $y = 0$. We denote by $\overline{A}$ the operator whose graph is the set $\overline{\mathscr{G}(A)}$, where

$$\mathscr{D}(\overline{A}) = \{x \in X : \exists (x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A), y \in X \text{ such that } \|x_n - x\| \to 0 \text{ and } \|Ax_n - y\| \to 0\},$$

with $\overline{A}x = y = \lim_{n\to\infty} Ax_n$.

Example 2.12. Differential operators give examples of unbounded operators which are closed. Let $X = C^0([0,1])$ with the supremum norm and define $D : \mathscr{D}(D) \to X$ by $Dx(t) = x'(t)$ where $x \in \mathscr{D}(D) = C^1([0,1])$. This operator is unbounded since if $x_n(t) = t^n$ for $n \in \mathbb{N}$, then $\|x_n\| = 1$ for all n and also $Dx_n(t) = nt^{n-1}$ so that $\|Dx_n\| = n$, hence $\|D\| \geq n$. Further, this operator is closed. For this, let $C^1([0,1]) \ni x_n \to x$ and $Dx_n = x_n' \to y$ as $n \to \infty$. Then

$$\int_0^t y(\tau)\,\mathrm{d}\tau = \lim_{n\to\infty} \int_0^t x_n'(\tau)\,\mathrm{d}\tau = \lim_{n\to\infty} (x_n(t) - x_n(0)) = x(t) - x(0). \tag{2.1}$$

Here the first, second and third equalities following respectively by the uniform convergence of $x_n' \to y$, the fundamental theorem of calculus and the pointwise convergence of $x_n \to x$ (which is implied by the uniform convergence). Then by (2.1), since y is continuous, $x \in C^1([0,1])$ and also $x' = y$, proving that D is a closed operator.

2.2. **Fundamental Theorems.** We recall the following theorems from functional analysis [2].

Theorem 2.13 (Banach-Steinhaus/Uniform boundedness principle)**.** *Let X be a Banach space and Y a normed space. Let $\mathcal{T} \subseteq \mathcal{L}(X,Y)$ be a collection of bounded linear operators form X to Y such that $\{Tx : T \in \mathcal{T}\}$ is bounded in Y for all $x \in X$. Then $\mathcal{T}$ is bounded in $\mathcal{L}(X,Y)$.*

Consequently, if $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(X,Y)$ is such that $(A_nx)_{n\in\mathbb{N}}$ is bounded for every $x \in X$, then there exists a unique operator $A \in \mathcal{L}(X,Y)$ such that $Ax = \lim_{n\to\infty} A_nx$ for $x \in X$.

Theorem 2.14. *Let X, Y be Banach spaces and $A \in \mathcal{L}(X,Y)$ bijective. Then $A^{-1} \in \mathcal{L}(Y,X)$.*

Theorem 2.15. *Let X and Y be normed spaces and $A \in L(X,Y)$ bounded with domain $\mathscr{D}(A)$.*

(i) If $\mathscr{D}(A)$ is closed in X, then A is a closed operator; and
(ii) If A is a closed operator and Y is complete, then $\mathscr{D}(A)$ is closed in X.

Theorem 2.16 (Closed graph theorem)**.** *Let X and Y be Banach spaces and $(A, \mathscr{D}(A))$ a closed linear operator $\mathscr{D}(A) \to Y$ defined on the closed linear subspace $\mathscr{D}(A)$. Then A is bounded.*

Remark 2.17. If A is a linear operator between two Banach spaces X and Y such that $\mathscr{D}(A) = X$, that is A is everywhere defined, then by the closed graph theorem A is a bounded operator, so $A \in \mathcal{L}(X,Y)$. Conversely if $A \in \mathcal{L}(X,Y)$, then $\mathscr{D}(A) = X$ is closed in X and so A is a closed.

2.3. **Resolvent and Spectrum.** Throughout this subsection, any operator $(A, \mathscr{D}(A))$ in X is assumed to be densely defined, that is $\overline{\mathscr{D}(A)} = X$, and closed. When talking about the spectrum and resolvent of an operator, we need our spaces to be over the field of complex numbers, $\mathbb{C}$. However if X is a space over $\mathbb{R}$, then we can *complexify* the space; see [1, §2.2.5].

Definition 2.18. The *resolvent set* of A, denoted by $\varrho(A)$, is the set of complex scalars $\lambda \in \mathbb{C}$ for which the operator $\lambda I - A : \mathscr{D}(A) \to X$ is an invertible operator. For such λ, we define the *resolvent of* A as the operator $R(\lambda, A) := (\lambda I - A)^{-1}$ which maps X to $\mathscr{D}(A)$.

By definition, we have that $\mathscr{D}(A) = R(\lambda, A)X$ for any $\lambda \in \varrho(A)$.

Definition 2.19. The *spectrum* of A, denoted by $\sigma(A)$, is the complement of the resolvent set. That is, $\sigma(A) = \mathbb{C} \backslash \varrho(A)$.

Remark 2.20. The spectrum of an operator A is usually divided into three subsets:

(i) the *point spectrum*, denoted by $\sigma_p(A)$, which is the set of $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is not one-to-one (the *eigenvalues* of A);
(ii) the *continuous spectrum*, denoted by $\sigma_c(A)$, which is the set of all $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is one-to-one with range which is dense in, but not equal to, X; and
(iii) the *residual spectrum*, denoted by $\sigma_r(A)$, which is the set of all $\lambda \in \sigma(A)$ for which the operator $\lambda I - A$ is one-to-one with range which is not dense in X.

With such constructions, we have that $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$ (as a disjoint union).

The following result gives an insight into why we assume that A is closed in this subsection.

Lemma 2.21. *Let $\lambda \in \varrho(A)$. Then the resolvent $R(\lambda, A)$ is bounded if and only if A is closed.*

Proof. The resolvent $R(\lambda, A) = (\lambda I - A)^{-1} : X \to \mathscr{D}(A) \subseteq X$ is an everywhere defined operator, so by the closed graph theorem (in fact by remark 2.17) the resolvent is bounded if and only if its graph $\mathscr{G}(R(\lambda, A)) = \{(x, (\lambda I - A)^{-1}x) : x \in X\}$ is a closed subspace of $X \times X$. If $y = R(\lambda, A)x$ for some $x \in X$, then $y \in \mathscr{D}(A)$ and also $(\lambda I - A)y = x$. Therefore

$$\mathscr{G}(R(\lambda, A)) = \{((\lambda I - A)y, y) : y \in \mathscr{D}(A)\}$$

is closed. Since a reordering of the coordinate components does not affect the closedness of this set, we have that $\{(y, (\lambda I - A)y) : y \in \mathscr{D}(A)\} = \mathscr{G}(\lambda I - A)$ is closed. In other words, $\lambda I - A$ is a closed operator, or equivalently A is closed by lemma 2.10. □

Proposition 2.22 (Resolvent identity)**.** *For any* $\lambda, \mu \in \varrho(A)$*,*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A).$$

Proof. Since $R(\lambda, A)(\lambda I - A) = (\lambda I - A) R(\lambda, A) = I$ and similarly for $R(\mu, A)$, we have that

$$\begin{aligned} R(\lambda, A) - R(\mu, A) &= R(\lambda, A)(\mu I - A) R(\mu, A) - R(\lambda, A)(\lambda I - A) R(\mu, A) \\ &= (\mu - \lambda) R(\lambda, A) R(\mu, A), \end{aligned}$$

as required. □

Remark 2.23. Consequently, the resolvents $R(\lambda, A)$ and $R(\mu, A)$ commute for any $\lambda, \mu \in \varrho(A)$.

Remark 2.24 (Pseudoresolvents)**.** More generally, the set $\{R(\lambda, A)\}_{\lambda \in \varrho(A)}$ is a family of *pseudo resolvents*. In fact, if $\Delta \subseteq \mathbb{C}$, a family $\{J(\lambda)\}_{\lambda \in \Delta}$ of bounded linear operators on X that satisfies $J(\lambda) - J(\mu) = (\mu - \lambda) J(\lambda) J(\mu)$ for all $\lambda, \mu \in \Delta$ is called a pseudoresolvent on Δ. Consequently, $J(\lambda)$ and $J(\mu)$ commute for any λ, μ. It is easy to see that $\operatorname{im} J(\lambda)$ is independent of $\lambda \in \Delta$ as

$$J(\lambda) = J(\mu)(I + (\mu - \lambda) J(\lambda)),$$

and so if $y \in \operatorname{im} J(\lambda)$, then $y = J(\lambda)x$ for some $x \in X$ and hence from the above

$$y = J(\mu)(I + (\mu - \lambda) J(\lambda))x = J(\mu)(x + (\mu - \lambda) y),$$

so $y \in \operatorname{im} J(\mu)$. By symmetry, we obtain that $\operatorname{im} J(\lambda) = \operatorname{im} J(\mu)$ for any λ, μ. Further, $\ker J(\lambda)$ is also independent of $\lambda \in \Delta$ since $J(\lambda) = (I + (\mu - \lambda) J(\lambda)) J(\mu)$, so clearly $\ker J(\mu) \subseteq \ker J(\lambda)$ and by symmetry we obtain the equality. It also follows that $J(\lambda)$ is the resolvent of a unique densely defined closed operator A if and only if $\ker J(\lambda) = \{0\}$ and $\operatorname{im} J(\lambda)$ is dense in X; see [1, Theorem 3.41] for a proof of this.

Theorem 2.25. *The resolvent set* $\varrho(A)$ *is an open subset of* $\mathbb{C}$*.*

Proof. Let $\lambda \in \varrho(A)$. Then for any $\mu \in \mathbb{C}$ we have that

$$\mu I - A = \lambda I - A + (\mu - \lambda) I = (\lambda I - A)[I + (\mu - \lambda) R(\lambda, A)].$$

Therefore if $|\mu - \lambda| < \|R(\lambda, A)\|^{-1}$, then $I + (\mu - \lambda) R(\lambda, A)$ is an invertible operator using the Neumann series, and thus $\mu I - A$ is invertible as it is the composition of two invertible operators. In other words, $B(\lambda, \|R(\lambda, A)\|^{-1}) \subseteq \varrho(A)$ and so $\varrho(A)$ is an open subset of $\mathbb{C}$. □

Corollary 2.26. $\lambda \mapsto R(\lambda, A)$ *is an analytic function such that for any* $n \in \mathbb{N}_0$*,*

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} R(\lambda, A) = (-1)^n n! R(\lambda, A)^{n+1}. \tag{2.2}$$

Proof. Let $\lambda, \mu \in \varrho(A)$ with $|\mu - \lambda| < \|R(\mu, A)\|^{-1}$. Then from the resolvent identity,

$$R(\lambda, A)[I - (\mu - \lambda) R(\mu, A)] = R(\mu, A).$$

Therefore from the assumption on $|\mu - \lambda|$, we have the Neumann series

$$R(\lambda, A) = \sum_{k=0}^{\infty} (\mu - \lambda)^k R(\mu, A)^{k+1}, \tag{2.3}$$

thus $R(\lambda, A)$ is analytic in λ. As the function is analytic, we can differentiate (2.3) term-by-term n-times and hence obtain the result (2.2). More formally, clearly (2.2) holds for $n = 0$ and also for $n = 1$ since from (2.3) we see that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} R(\lambda, A) = \sum_{k=1}^{\infty} (-k)(\mu - \lambda)^{k-1} R(\mu, A)^{k+1},$$

so taking $\mu = \lambda$ the right-hand side becomes $-R(\lambda, A)^2$ with convention $0^0 = 1$. Now assume the result holds for all $0 \leq m \leq n$ for some $n \geq 1$. Then from the Neumann series we have that

$$\frac{\mathrm{d}^{n+1}}{\mathrm{d}\lambda^{n+1}} R(\lambda, A) = \frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\mathrm{d}^n}{\mathrm{d}\lambda^n} R(\lambda, A) = (-1)^n n! \frac{\mathrm{d}}{\mathrm{d}\lambda} R(\lambda, A)^{n+1}$$

Then using that the result holds for $m = 1$, we obtain that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} R(\lambda, A)^{n+1} = (n+1) R(\lambda, A)^n \frac{\mathrm{d}}{\mathrm{d}\lambda} R(\lambda, A) = -(n+1) R(\lambda, A)^{n+2},$$

thus the result follows by induction. □

Definition 2.27. If A is a bounded operator in X, then its *spectral radius* is the value

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

If $(A, \mathscr{D}(A))$ is an unbounded operator, then its *spectral bound* is $s(A) := \sup\{\Re[\lambda] : \lambda \in \sigma(A)\}$.

Theorem 2.28. *For any bounded operator A in X, we have that*

$$r(A) = \lim_{n \to \infty} \|A^n\|^{1/n}.$$

Consequently, $r(A) \leq \|A\|$. Now if $\|A\| < 1$ for any bounded operator A, then $1 \in \varrho(A)$ and

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n \text{ where } A^0 = I. \tag{2.4}$$

The proof follows since the above series defines a Cauchy sequence as $\|A\| < 1$, which is hence convergent as X is Banach. The following proposition gives more general conditions for $1 \in \varrho(A)$.

Proposition 2.29. *Let $A \in \mathcal{L}(X)$. If $r(A) < 1$, then $1 \in \varrho(A)$ and (2.4) holds.*

Proof. If $r(A) < 1$, then we have that $|\lambda| < 1$ for all $\lambda \in \varrho(A)$. Hence $1 \in \varrho(A)$, as if not then $1 \in \sigma(A)$, giving rise to a contradiction. Therefore $R(1, A) = (I - A)^{-1}$ exists. Now using the equivalent characterisation of $r(A)$ from theorem 2.28, if q is such that $r(A) < q < 1$, then there exists an $N \in \mathbb{N}$ such that $\|A^n\|^{1/n} < q$ for all $n \geq N$ by inertia, that is $\|A^n\| < q^n$. Then

$$\sum_{k=0}^{\infty} \|A^k\| \leq \sum_{k=0}^{N-1} \|A^k\| + \sum_{k=N}^{\infty} q^n < \infty,$$

so the Neumann series is absolutely convergent. Then it can be shown that $(I - A) S_n = I - A^{n+1}$ and where S_n denote the partial sums of $\sum_{k=0}^{\infty} A^k$, and taking $n \to \infty$ we obtain that

$$(I - A) \sum_{k=0}^{\infty} A^k = I,$$

since $\|A^n x\| \leq \|A^n\| \|x\| \leq q^n \|x\| \to 0$ as $n \to \infty$, so $A^{n+1} \to 0$ as $n \to \infty$. It can similarly be shown that $\sum_{k=0}^{\infty} A^k (I - A) = I$, and so the Neumann series (2.4) holds, as required. □

3. Banach Spaces and Order

For a more detailed description of the following concepts, see [1, §2.2].

3.1. **Ordered Banach Spaces.**

Before discussing what it means for an operator to be *positive*, we consider what it actually means for an element of some arbitrary vector space to be positive or negative. In $\mathbb{R}$, there is a natural ordering given by $\leq$, and this binary operation is reflexive, anti-symmetric and transitive. In other words, for all $x, y, z \in \mathbb{R}$ we have that (i) $x \leq x$, (ii) $x \leq y$ and $y \leq x$ implies $x = y$ and (iii) if $x \leq y$ and $y \leq z$, then $x \leq z$, respectively. Since for any $x, y \in \mathbb{R}$ we know that either $x \leq y$ or $y \leq x$, any elements in $\mathbb{R}$ are *comparable* with respect to $\leq$. An ordering with the above such properties is called a *total ordering*. Similar orderings can be defined on $\mathbb{R}^n$ (componentwise) and even on abstract function spaces such as $L^1(\Omega)$ for a measure space Ω, where the ordering is defined pointwise almost everywhere; in other words, $f \leq g$ if and only if $f(x) \leq g(x)$ for almost every $x \in \Omega$. These are natural orderings, but for a general vector space there may not exist such an obvious ordering.

We now introduce some basic theory to introduce partial orders and ordered vector spaces.

Definition 3.1. Let X be any set. A *partial order* (or simply *order*) on X is a binary relation, denoted by $\preceq$, which is reflexive, anti-symmetric and transitive. Respectively, that is

(O1) $x \preceq x$ for any $x \in X$;

(O2) for any $x, y \in X$, if $x \preceq y$ and $y \preceq x$, then $x = y$; and

(O3) for any $x, y, z \in X$, if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

We also write $y \succeq x$ to mean $x \preceq y$.

Definition 3.2. Let $S \subseteq X$ be any subset.

(i) An element $x \in X$ is an *upper bound* for S if $s \preceq x$ for all $s \in S$.

(ii) An element $x \in S$ is said to be *maximal* if there is no $s \in S$, $s \neq x$ such that $x \preceq s$.

(iii) An element $x \in S$ is the *greatest element* of S if $s \preceq x$ for all $s \in S$.

We define a *lower bound*, *minimal element* and *least element* analogously.

Remark 3.3. Note the distinction between maximal elements and greatest elements. In fact, $x \in S$ is maximal if it is the largest amongst all *comparable* elements of S, whereas a greatest element is the largest amongst *all elements* in S.

Now as in $\mathbb{R}$, we can define the supremum and infimum of any set $S \subseteq X$.

Definition 3.4. Let $S \subseteq X$ be any subset.

(i) We call $x \in X$ the *supremum* of S, denoted $x = \sup S$, if it is the least upper bound of S. That is, $s \preceq x$ for all $s \in S$, and any other upper bound $y \in X$ of S has $x \preceq y$.

(ii) We call $x \in X$ the *infimum* of S, denoted $x = \inf S$, if it is the greatest lower bound of S.

We denote the supremum and infimum of the two-point set $\{x, y\}$ by $x \vee y$ and $x \wedge y$, respectively.

Definition 3.5. Let X be a set with partial ordering $\preceq$.

We call X a *lattice* if every pair of elements has supremum and infimum.

We can equip a partial order to any vector space which respects the vector space structure. We assume that all vector spaces X are over the field $\mathbb{R}$ unless otherwise stated.

To use spectral theory, we need to work over $\mathbb{C}$, but this can be done by complexification of the space X. For a more detailed description of this concept, see [1, §2.2.5].

Definition 3.6. An *ordered vector space* is a vector space X equipped with a partial order $\preceq$ which is compatible with the vector space structure in the sense that

(O4) for any $x, y, z \in X$, if $x \preceq y$ then $x + z \preceq y + z$; and
(O5) for any $x, y \in X$ and scalar $\alpha \geq 0$, if $x \preceq y$ then $\alpha \cdot x \preceq \alpha \cdot y$,

where $+, \cdot$ are the vector space operations.

Moreover, if X is also a *lattice*, then it is called a *vector lattice* (or *Riesz space*).

Example 3.7. Function spaces are typical examples of Riesz spaces. If $X = \mathbb{R}^\Omega$ for some set Ω, then we can introduce a *pointwise* ordering on X by $f \preceq g$ if and only if $f(x) \leq g(x)$ for all $x \in \Omega$ (or $f(x) \leq g(x)$ for almost everywhere x if Ω is a measure space). Now we define the operations $f \vee g$ and $f \wedge g$ on $X \times X$ by $(f \vee g)(x) := \max\{f(x), g(x)\}$ and $(f \wedge g)(x) := \min\{f(x), g(x)\}$. Then we call X a *function space* if $f \vee g, f \wedge g \in X$ for all $f, g \in X$.

Definition 3.8. The *positive cone* of an ordered vector space X is the set

$$X_+ := \{x \in X : x \succeq 0\}.$$

Remark 3.9. More generally, a *convex cone* in a vector space X is a set C with the following characterising properties: $C + C \subseteq C$, $\alpha C \subseteq C$ for any $\alpha \geq 0$ and $C \cap (-C) = \{0\}$. It is not hard to see that X_+ is in fact a convex cone in X using the compatibility of $\preceq$ with the vector space operations. On the other hand, given a convex cone C, the relation (on X) given by $x \preceq y$ if and only if $y - x \in C$ makes X an ordered vector space with $X_+ = C$ since

$$X_+ = \{x \in X : x \succeq 0\} = \{x \in X : x - 0 = x \in C\} = C.$$

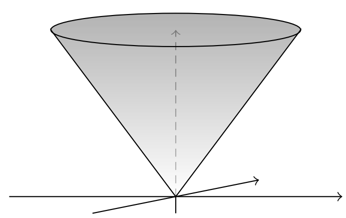

FIGURE 1. Visualisation of a positive cone in $\mathbb{R}^3$.

The convex cone C of X is called *generating* if $X = C - C$, so every vector $x \in X$ can be expressed as the difference of two 'positive' vectors, that is $x = c_1 - c_2$ for some $c_1, c_2 \succeq 0$.

Remark 3.10. In the real numbers $\mathbb{R}$, there are no infinitely large or small numbers. In other words, for any $r \in \mathbb{R}_+$ we have that $\lim_{n\to\infty} n^{-1} r = 0$. This is called the *Archimedean* property of the reals, and this motivates an analogous definition in any Riesz space.

Definition 3.11. A Riesz space is *Archimedean* if $\inf_{n\in\mathbb{N}}\{n^{-1}x\} = 0$ for any $x \in X_+$.

3.2. **Generalising Properties of** $(\mathbb{R}, |\cdot|)$**.** The following concepts are generalisations of results from the reals $\mathbb{R}$ with ordering $\leq$. In particular, the modulus or absolute value of an element and the triangle inequality. In order to do this, we need to define the absolute value in a general Riesz space which is the analogue of $|x| = \max\{x, -x\}$ in $\mathbb{R}$. We first consider some properties of suprema and infima which will be used to prove later results.

Proposition 3.12. *Let X be a Riesz space and let $x, y, z \in X$. Then*

(i) $x + y = \sup\{x, y\} + \inf\{x, y\}$*;*
(ii) $x + \sup\{y, z\} = \sup\{x + y, x + z\}$ *and* $x + \inf\{y, z\} = \inf\{x + y, x + z\}$*;*
(iii) $\sup\{x, y\} = -\inf\{-x, -y\}$ *and* $\inf\{x, y\} = -\sup\{-x, -y\}$*; and*
(iv) $\alpha \sup\{x, y\} = \sup\{\alpha x, \alpha y\}$ *and* $\alpha \inf\{x, y\} = \inf\{\alpha x, \alpha y\}$ *for* $\alpha \geq 0$.

Proof. (i) Since $\inf\{x, y\} \preceq y$ we have that $x + \inf\{x, y\} \preceq x + y$ by (O4). Therefore

$$x \preceq x + y - \inf\{x, y\},$$

and similarly $y \preceq x + y - \inf\{x, y\}$. Hence $\sup\{x, y\} \preceq x + y - \inf\{x, y\}$. Now on the other hand, $y \preceq \sup\{x, y\}$ and so $x + y - \sup\{x, y\} \preceq x$ and similarly $x + y - \sup\{x, y\} \preceq y$. So we obtain that $x + y - \sup\{x, y\} \preceq \inf\{x, y\}$, and hence the result follows by (O2).

(ii) Clearly $x + y \preceq x + \sup\{y, z\}$ and $x + z \preceq x + \sup\{y, z\}$ by (O4), therefore we deduce that

$$\sup\{x + y, x + z\} \preceq x + \sup\{y, z\}.$$

On the other hand, $y = -x + (x + y) \preceq -x + \sup\{x + y, x + z\}$ and similarly for z. Hence $\sup\{y, z\} \preceq -x + \sup\{x + y, x + z\}$, as required. The equality for infima is proven similarly.

(iii) Since $x, y \preceq \sup\{x, y\}$, we have $-\sup\{x, y\} \preceq -x$ and $-\sup\{x, y\} \preceq -y$ and therefore we obtain $-\sup\{x, y\} \preceq \inf\{-x, -y\}$. Now suppose that z is any lower bound of $\{-x, -y\}$, then we have $z \preceq -x$ and $z \preceq -y$ and so $-z \succeq x, y$. Hence $-z \succeq \sup\{x, y\}$, or equivalently $z \preceq -\sup\{x, y\}$, showing that $-\sup\{x, y\}$ is the greatest lower bound of the set $\{-x, -y\}$. That is, $-\sup\{x, y\} = \inf\{-x, -y\}$, and the other equality is proven similarly.

(iv) If $\alpha = 0$ the result is trivial. Otherwise, it is easy to see that $\sup\{\alpha x, \alpha y\} \preceq \alpha \sup\{x, y\}$. Now suppose that $z \succeq \alpha x, \alpha y$ (so that z is an upper bound of $\{\alpha x, \alpha y\}$), then $\alpha^{-1} z \succeq x, y$ and so $\alpha^{-1} z \succeq \sup\{x, y\}$, or equivalently $z \succeq \alpha \sup\{x, y\}$ using (O5). Note here that α^{-1} exists since α is a non-zero element of the field $\mathbb{R}$. Therefore $\alpha \sup\{x, y\}$ is the least upper bound of the set $\{\alpha x, \alpha y\}$, and the result for infima is analogous. □

Definition 3.13. Let X be a Riesz space and $x \in X$.

(i) The *positive part* of x is $x_+ = \sup\{x, 0\}$.
(ii) The *negative part* of x is $x_- = \sup\{-x, 0\} = -\inf\{x, 0\}$.
(iii) The *absolute value* of x is $|x| = \sup\{x, -x\}$.

With these notions, we have the *lattice operations* of a Riesz space, given by

$$(x, y) \mapsto x \vee y, \ (x, y) \mapsto x \wedge y, \ x \mapsto x_\pm \text{ and } x \mapsto |x|. \tag{3.1}$$

Notice that $x_+, x_-, |x| \succeq 0$, so these lattice operations really map $X \to X_+ \subseteq X$.

Proposition 3.14. *If X is a Riesz space and $x \in X$, then $x = x_+ - x_-$ and $|x| = x_+ + x_-$.*

Proof. Using proposition 3.12, we obtain that

$$x = x + 0 = \sup\{x, 0\} + \inf\{x, 0\} = \sup\{x, 0\} - \sup\{-x, 0\} = x_+ - x_-, \text{ and}$$
$$|x| = \sup\{x, -x\} = \sup\{2x, 0\} - x = 2x_+ - (x_+ - x_-) = x_+ + x_-,$$

as required. □

Remark 3.15. In $\mathbb{R}$, the absolute value $|x| = \max\{x, -x\}$ defines a norm, and we have a similar result in a Riesz space X. Indeed, if $x = 0$ then clearly $|x| = \sup\{0, 0\} = 0$ and if $|x| = 0$, then $\sup\{x, -x\} = 0$ and since the supremum is an upper bound, $x \preceq 0$ and $-x \preceq 0$, that is $x \succeq 0$. Then by (O2) we obtain that $x = 0$. Also $|\alpha x| = |\alpha||x|$ for any scalar α and $x \in X$. Indeed,

$$|\alpha x| = (\alpha x)_+ + (\alpha x)_- = \sup\{\alpha x, 0\} + \sup\{-\alpha x, 0\}$$

If $\alpha \geq 0$, then using proposition 3.12 we have that $\sup\{\alpha x, 0\} = \alpha x_+$. Also, for the second term we have $\sup\{-\alpha x, 0\} = -\inf\{\alpha x, 0\} = -\alpha \inf\{x, 0\} = \alpha \sup\{-x, 0\} = \alpha x_-$. Since $|\alpha| = \alpha$ for $\alpha \geq 0$, we obtain that $|\alpha x| = |\alpha|(x_+ + x_-) = |\alpha||x|$, and similarly if $\alpha \leq 0$.

The generalisation of the triangle property in $(\mathbb{R}, |\cdot|)$ is given in the below proposition.

Lemma 3.16. *Let X be a Riesz space and $x, y \in X$. Then $(x + y)_+ \preceq x_+ + y_+$ and $x_+ \preceq |x|$.*

Proof. Clearly $0 \preceq \sup\{x, 0\} + \sup\{y, 0\} = x_+ + y_+$ and also $x + y \preceq x_+ + y_+$. Thus

$$(x + y)_+ = \sup\{x + y, 0\} \preceq x_+ + y_+.$$

For the latter equation, by (O4) we have that $|x| = x_+ + x_- \succeq x_+ + 0 = x_+$ since $x_- \succeq 0$. □

Proposition 3.17 (Triangle inequality)**.** *Let X be a Riesz space and $x, y \in X$. Then*

$$||x| - |y|| \preceq |x + y| \preceq |x| + |y|. \tag{3.2}$$

Further, $|\sup\{x, z\} - \sup\{y, z\}| \preceq |x - y|$ and $|\inf\{x, z\} - \inf\{y, z\}| \preceq |x - y|$.

Proof. Clearly $x \preceq |x|$ and $y \preceq |y|$, therefore by (O4) we have that

$$x + y \preceq |x| + y \preceq |x| + |y|.$$

Similarly, $-x \preceq |x|$ and $-y \preceq |y|$ and so $-x - y \preceq |x| + |y|$. Therefore we obtain that

$$|x + y| = \sup\{x + y, -x - y\} \preceq |x| + |y|.$$

Then using the above, we have $|x| = |x + y - y| \preceq |x + y| + |-y| = |x + y| + |y|$ so that

$$|x| - |y| \preceq |x + y|,$$

and the symmetry of x and y gives (3.2). Now by proposition 3.12(ii), we have that

$$\sup\{x, z\} = \sup\{(x - z) + z, (z - z) + z\} = \sup\{x - z, 0\} + z$$
$$\Rightarrow \quad \sup\{x, z\} - \sup\{y, z\} = (\sup\{x - z, 0\} + z) - (\sup\{y - z, 0\} + z)$$
$$= (x - z)_+ - (y - z)_+ = ((x - y) + (y - z))_+ - (y - z)_+.$$

Then by lemma 3.16 we see that

$$\sup\{x, z\} - \sup\{y, z\} \preceq (x - y)_+ + (y - z)_+ - (y - z)_+ = (x - y)_+ \preceq |x - y|,$$

and similarly it can be shown that $\sup\{y, z\} - \sup\{x, z\} \preceq |x - y|$.

The equation involving infima is proven in an analogous way. □

3.3. **Banach Lattices.** Now since we are interested in studying operators defined on Banach spaces, and in particular ordered Banach spaces, we also have a norm $\|\cdot\|$ to consider. In fact, we consider a particular type of norm on X which, in a way, *respects* the partial ordering of the space X. This norm is called a lattice norm, as defined below.

Definition 3.18. A norm $\|\cdot\|$ on a Riesz space X is called a *lattice norm* if

$$\forall x, y \in X,\ |x| \preceq |y| \ \Rightarrow\ \|x\| \leq \|y\|.$$

A *Banach lattice* is a Riesz space which is complete when equipped with the lattice norm.

Example 3.19. Let $n \in \mathbb{N}$ and consider the Riesz space $\mathbb{R}^n$ with componentwise ordering (in other words, the *lexicographic* ordering). Then any $x \in \mathbb{R}^n$ can be expressed as $x = (x_1, \dots, x_n)$ and we have that $|x| = (|x_1|, \dots, |x_n|)$. Thus if $|x| \preceq |y|$, then $|x_i| \leq |y_i|$ in $\mathbb{R}$ for all $i = 1, \dots, n$ and thus clearly $\|x\| \leq \|y\|$ where $\|\cdot\|$ is the standard Euclidean norm, so $\|\cdot\|$ is a lattice norm. Since $\mathbb{R}^n$ is complete with norm, we have that $(\mathbb{R}, \|\cdot\|, \preceq)$ is a Banach lattice.

Lemma 3.20. *If X is a Banach lattice, then $\|x\| = \||x|\|$ for any $x \in X$.*

Proof. Since $x \preceq |x|$ for any $x \in X$, we have that $|x| \preceq |(|x|)|$. But since $|x| \succeq 0$, it follows that $|(|x|)| = |x|$ and thus $|x| \preceq |(|x|)| \preceq |x|$ for any $x \in X$. Thus by definition of the lattice norm, we have that $\|x\| \leq \||x|\| \leq \|x\|$, as required. ☐

Proposition 3.21. *If X is a normed lattice, then the lattice operations are uniformly continuous in X with respect to any of the variables involved.*

Proof. We denote the lattice operations from (3.1) by the maps $f_1, \dots, f_5$, respectively. In other words, $f_1(x,y) = x \vee y$, $f_2(x,y) = x \wedge y$, $f_3(x) = x_+$, $f_4(x) = x_-$ and $f_5(x) = |x|$. Now we have for f_4 that $|f_4(x) - f_4(y)| = |\sup\{x, 0\} - \sup\{y, 0\}| \preceq |x - y|$ and thus $\|f_4(x) - f_4(y)\| \leq \|x - y\|$, proving uniform continuity of f_4 and similarly for f_3. Since $f_5 = f_3 + f_4$, it follows immediately that f_5 is uniformly continuous. Now for f_1, we using proposition 3.17 to deduce that

$$|\sup\{x, z\} - \sup\{y, v\}| = |\sup\{x, z\} - \sup\{y, z\} + \sup\{y, z\} - \sup\{y, v\}| \preceq |x - y| + |z - v|.$$

Therefore using lemma 3.20 and the definition of the lattice norm, we have that

$$\|f_1(x,z) - f_1(y,v)\| \leq \||x - y| + |z - v|\| \leq \|x - y\| + \|z - v\|$$

Hence f_1 is uniformly continuous in both of its variables, and similarly so is f_2. ☐

Definition 3.22. We say that a Banach lattice X is

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

Examples 3.23. The space of continuous functions $X = C^0([0,1], \mathbb{R})$ with supremum norm and ordering as in example 3.7 is an AM-space. First note that X is complete under this norm, and it is a lattice norm since if $|f| \preceq |g|$, then $|f(x)| \leq |g(x)|$ for any $x \in [0,1]$, and thus taking the supremum yields $\|f\| \leq \|g\|$. An example of an AL-space is $X = L^1(\Omega)$, where X has the natural ordering given by $f \preceq g$ if and only if $f(x) \leq g(x)$ for almost every $x \in \Omega$.

For further discussions of such spaces, see [1, Remark 2.58 – 2.61].

3.4. **Order and Norm.** Existence of an order in some set X introduces a natural way to define convergence in the Banach space. Now any sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ can equivalently be thought of as an element of $\mathbb{R}^{\mathbb{N}}$, that is a function $\mathbb{N} \to \mathbb{R}$ with $n \mapsto x_n$, and this motivates the definition of a net in an arbitrary set X. Since $\mathbb{R}$ is a metric space, we have the notion of distances and thus convergence, but for general ordered spaces X we can not do this, hence we need to define what it means for a net $(x_\alpha)_{\alpha\in\Delta}$ to converge. In fact, many of the results in $(\mathbb{R}, |\cdot|)$ hold more generally, for example unique limits and converging and increasing sequences $(x_n)_{n\in\mathbb{N}}$ have limit $\sup\{x_n : n \in \mathbb{N}\}$, amongst other things.

Definition 3.24. Let Δ be a *directed* ordered set (so any pair of elements has an upper bound). Then a *net* $(x_\alpha)_{\alpha\in\Delta}$ in a set X is a function from the *index set* Δ to X, with $\alpha \mapsto x_\alpha$.

A *subnet* of $(x_\alpha)_{\alpha\in\Delta}$ is a net $(y_\beta)_{\beta\in B}$ such that

$$\forall \alpha \in \Delta\, \exists \beta \in B \text{ such that } \forall B \ni \beta' \succeq \beta\, \exists \alpha' \succeq \alpha \text{ such that } y_{\beta'} = x_{\alpha'}.$$

Example 3.25. Sequences and subsequences are examples of nets and subnets.

Definition 3.26. A net $(x_\alpha)_{\alpha\in\Delta}$ in a normed space X converges to some $x \in X$ if

$$\forall \varepsilon > 0\, \exists \alpha_0 \in \Delta \text{ such that } \|x_\alpha - x\| \leq \varepsilon\, \forall \alpha \succeq \alpha_0.$$

This is denoted as $x_\alpha \xrightarrow{n} x$ or $x = \lim_{\alpha\in\Delta} x_\alpha$.

Definition 3.27. Let X be a Riesz space and $(x_\alpha)_{\alpha\in\Delta}$ a net in X, and let $x \in X$.

(i) The net is *decreasing*, denoted $x_\alpha\downarrow$, if for any $\alpha_1, \alpha_2 \in \Delta$, if $\alpha_1 \preceq \alpha_2$ then $x_{\alpha_1} \succeq x_{\alpha_2}$.
(ii) We write $x_\alpha \downarrow x$ if the net is decreasing and $x = \inf\{x_\alpha : \alpha \in \Delta\}$.
(iii) We write $x_\alpha \downarrow\succeq x$ if the net is decreasing and bounded below by x, that is $x_\alpha \succeq x\, \forall \alpha \in \Delta$.

We can define $x_\alpha\uparrow$, $x_\alpha \uparrow x$ and $x_\alpha \uparrow\preceq x$ in an analogous way.

Definition 3.28. The net $(x_\alpha)_{\alpha\in\Delta}$ in X is *order convergent* to $x \in X$ if there exist nets $(y_\beta)_{\beta\in B}$ and $(z_\gamma)_{\gamma\in\Gamma}$ such that $y_\beta \uparrow x$ and $z_\gamma \downarrow x$ and for any $(\beta, \gamma) \in B \times \Gamma$, there exists an $\alpha \in \Delta$ such that $y_\beta \preceq x_\alpha \preceq z_\gamma$. We denote this by $x_\alpha \xrightarrow{o} x$.

Proposition 3.29. *Let X be an ordered set and $(x_\alpha)_{\alpha\in\Delta}$ a net in X.*

(i) If either $x_\alpha \uparrow x$ or $x_\alpha \downarrow x$ for some $x \in X$, then $x_\alpha \xrightarrow{o} x$.
(ii) Conversely, if $x_\alpha\uparrow$ (respectively $x_\alpha\downarrow$) and $x_\alpha \xrightarrow{o} x$, then $x_\alpha \uparrow x$ (respectively $x_\alpha \downarrow x$).

Proof. (i) Suppose that $x_\alpha \downarrow x$, that is the net is decreasing and $x = \inf\{x_\alpha : \alpha \in \Delta\}$. Now we consider the nets $(y_\alpha)_{\alpha\in\Delta}$ and $(z_\alpha)_{\alpha\in\Delta}$ defined by $y_\alpha = x_\alpha$ and $z_\alpha = x$ for each $\alpha \in \Delta$. Then we have that $y_\alpha \preceq x_\alpha \preceq z_\alpha$ for all $\alpha \in \Delta$ and hence $x_\alpha \xrightarrow{o} x$.

(ii) Suppose that $x_\alpha\uparrow$ and $x_\alpha \xrightarrow{o} x$ and let $(y_\beta)_{\beta\in B}$ and $(z_\gamma)_{\gamma\in\Gamma}$ be the nets which define the convergence $x_\alpha \xrightarrow{o} x$. Then for each (β, γ) there exists an $\alpha_{\beta,\gamma}$ such that $y_\beta \preceq x_\alpha \preceq z_\gamma$ for any $\alpha \succeq \alpha_{\beta,\gamma}$. Let α, β, γ be fixed, then for $\alpha' \succeq \sup\{\alpha_{\beta,\gamma}, \alpha\}$ we have that $x_\alpha \preceq x_{\alpha'} \preceq z_\gamma$ and so $x_\alpha \preceq z_\gamma$ for any α, γ. Since $z_\gamma \downarrow x$, taking the infimum yields $x_\alpha \preceq x$ for any α, so x is an upper bound. Now let y be any other upper bound for $(x_\alpha)_{\alpha\in\Delta}$, then $x_\alpha \preceq y$ and for any $\beta \in B$, as above, there exists an α with $y_\beta \preceq x_\alpha \preceq y$ and thus $y_\beta \preceq y$ for all $\beta \in B$. Then as $y_\beta \uparrow x$, we have that $x \preceq y$ and so indeed x is the least upper bound of the net $(x_\alpha)_{\alpha\in\Delta}$, as required. □

Proposition 3.30. *Let X be a normed lattice and $(x_\alpha)_{\alpha\in\Delta}$ a net in X. Then*

(i) the positive cone X_+ is closed in X;
(ii) if $x_\alpha\uparrow$ and $\lim_{\alpha\in\Delta} x_\alpha = x$, then $x = \sup\{x_\alpha : \alpha \in \Delta\}$; and
(iii) if $x_\alpha\downarrow$ and $\lim_{\alpha\in\Delta} x_\alpha = x$, then $x = \inf\{x_\alpha : \alpha \in \Delta\}$.

Proof. (i) By definition, $X_+ = \{x \in X : x_- = 0\}$ and the lattice operation $f_4 : X \to X$ with $x \mapsto x_-$ is continuous by proposition 3.21. Hence X_+ is closed as it is the preimage of a closed set under a continuous map, that is $X_+ = f_4^{-1}(\{0\})$. Now $\{0\}$ is closed if and only if $X\backslash\{0\}$ is open, and for any $x \neq 0$ we have that $B(x, \|x\|) \subseteq X\backslash\{0\}$, so $\{0\}$ is closed.

(ii) For fixed $\alpha \in \Delta$ we have that $\lim_{\beta\in\Delta}(x_\beta - x_\alpha) = x - x_\alpha$ and since $(x_\alpha)_{\alpha\in\Delta}$ is increasing we see that $x_\beta - x_\alpha \in X_+$ for $\beta \succeq \alpha$. Thus $x - x_\alpha \in X_+$ since X_+ is closed, hence x is an upper bound for the net. Now let y be any other upper bound, then $x_\alpha \preceq y$ for all $\alpha \in \Delta$. Then $0 \preceq y - x_\alpha \xrightarrow{n} y - x$ since $x_\alpha \xrightarrow{n} x$ by assumption. Hence again since X_+ is closed, we obtain that $y - x \succeq 0$ and so indeed $x = \sup\{x_\alpha : \alpha \in \Delta\}$.

(iii) This follows analogously to (ii). □

Definition 3.31. A Banach lattice X is a KB-space if every increasing norm-bounded sequence of elements in X_+ converges in norm in X. That is, if $(x_n)_{n\in\mathbb{N}}$ is a sequence of elements in X_+ with $x_n \preceq x_m$ for $n \leq m$ and $\|x_n\| \leq M$ for some $M > 0$ and all $n \in \mathbb{N}$, then $(x_n)_{n\in\mathbb{N}}$ converges.

Theorem 3.32. *If X is an AL-space, then it is a KB-space.*

Proof. Suppose that $(x_n)_{n\in\mathbb{N}}$ is an increasing and normed-bounded sequence in X_+. Then we have that $0 \preceq x_n \preceq x_m$ for any $n \leq m$, and so $|x_n| \preceq |x_m|$, thus $\|x_n\| \leq \|x_m\|$ as X is a Banach lattice with lattice norm. Therefore $(\|x_n\|)_{n\in\mathbb{N}}$ is an increasing and bounded sequence, and thus Cauchy as it is convergent since $\mathbb{R}$ is complete. Now for $0 \preceq x_n \preceq x_m$, we have that

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

since X is an AL-space. Thus $(x_n)_{n\in\mathbb{N}}$ is Cauchy and hence convergent, as required. □

3.5. **Positive Operators.** We are now in a position to define positive operators.

Definition 3.33. Let X and Y be Banach lattices with partial orders $\preceq_1$ and $\preceq_2$, respectively. Then a linear operator $A \in L(X, Y)$ is *positive* if $Ax \succeq_2 0$ for any $x \in X_+$, that is $x \succeq_1 0$.

We denote this by $A \succeq 0$.

Lemma 3.34. *$A \in L(X, Y)$ is a positive operator between Banach lattices if and only if*

$$|Ax| \preceq A|x|, \quad \forall x \in X. \tag{3.3}$$

Proof. Suppose that A is a positive operator and let $x \in X$. Then $-|x| \preceq x \preceq |x|$, so $|x| - x \succeq 0$ and thus $A(|x| - x) = A|x| - Ax \succeq 0$. Therefore $Ax \preceq A|x|$ and similarly $-A|x| \preceq Ax$, proving (3.3). Conversely let $x \in X_+$ and suppose that (3.3) holds. Then since $|x| = x$, it follows that $|Ax| \preceq A|x| = Ax$ and thus $Ax \succeq 0$, that is A is a positive operator. □

Example 3.35. Consider the space $X = L^1(\Omega)$. If $k \geq 0$ is measurable on Ω, then

$$(Af)(x) = \int_{\Omega} k(x,y) f(y) \, \mathrm{d}y$$

defines a positive linear operator A on X.

Example 3.36. Consider the space $\mathcal{L}(X,Y)$ of linear operators where X, Y are Banach lattices. The space $\mathcal{C}$ of positive operators in $\mathcal{L}(X,Y)$ is a convex cone. To see this, consider $A := A_1 + A_2$ for some $A_1, A_2 \succeq 0$. With pointwise operations, for any $x \in X_+$, we have $Ax = A_1x + A_2x \succeq 0$ since $A_ix \succeq 0$ for $i = 1, 2$ and therefore $A \succeq 0$, proving that $\mathcal{C} + \mathcal{C} \subseteq \mathcal{C}$. It is clear that $\alpha\mathcal{C} \subseteq \mathcal{C}$ for any $\alpha \geq 0$ because $(\alpha A)x = \alpha(Ax) \succeq \alpha \cdot 0 = 0$ for $x \in X_+$. Finally, if $A \in \mathcal{C} \cap (-\mathcal{C})$, then $Ax \succeq 0$ and $(-A)x \succeq 0$ for any $x \in X_+$. Thus $Ax \succeq 0$ and $-(Ax) \succeq 0$, and therefore $Ax = 0$ from (O2). Since A is a linear operator, it follows that $A = \mathbf{0}$ and so $\mathcal{C} \cap (-\mathcal{C}) = \{\mathbf{0}\}$, where $\mathbf{0}$ denotes the zero operator. Since $\mathcal{C}$ is a convex cone in $\mathcal{L}(X,Y)$, it generates a natural ordering for $\mathcal{L}(X,Y)$. In particular, $A \preceq B$ if and only if $B - A \in \mathcal{C}$, or in other words $Ax \preceq Bx$ for all $x \in X_+$. However, this cone does not generate $\mathcal{L}(X,Y)$.

Linear operators between finite-dimensional vector spaces are uniquely determined by their behaviour on any basis of the domain. Analogously, positive operators are uniquely determined by their action on the positive cone of their domain. More precisely, we have the following:

Theorem 3.37. *Suppose that X and Y are Banach lattices and $A : X_+ \to Y_+$ is additive, that is $A(x+y) = Ax + Ay$ for any $x, y \in X_+$. Then A extends uniquely to a positive linear operator $\tilde{A} : X \to Y$. Moreover, we have that $\tilde{A}x = Ax_+ - Ax_-$ for any $x \in X$.*

Proof. See [1, Theorem 2.64]. □

4. Semigroups

We begin with an example which motivates the use of semigroups. The initial value problem $u'(t) = Au(t)$ for $t > 0$ with $u(0) = u_0$ has a well-known solution when either $A \in M_{n,n}(\mathbb{C})$ or $A \in \mathcal{L}(X)$ for some Banach space X, namely $u(t) = e^{tA}u_0$ in both cases. Here we define

$$e^{tA} := \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n, \tag{4.1}$$

the *matrix exponential*. When $A \in \mathcal{L}(X)$, this operator is well-defined as its partial sums define a Cauchy sequence in X and $\|e^{tA}\| \leq e^{t\|A\|}$. However, when A is an unbounded operator defined on some domain $\mathscr{D}(A) \subseteq X$, it is not possible to immediately make sense of (4.1). Despite this, we would still like to define the solution when A is unbounded, and it turns out that the theory of *strongly continuous semigroups* makes this possible.

Definition 4.1. Let X be a Banach space. Then a one-parameter family $\{T(t)\}_{t\geq 0} \subseteq \mathcal{L}(X)$ is a *semigroup* if $T(0) = I$ and $T(t+s) = T(t)T(s)$ for every $t, s \geq 0$. A semigroup is said to be *uniformly continuous* or *strongly continuous*, respectively, if

$$\lim_{t\to 0^+} \|T(t) - I\| = 0 \text{ or } \lim_{t\to 0^+} \|T(t)x - x\| = 0 \; \forall x \in X.$$

To every semigroup $T = \{T(t)\}_{t\geq 0}$ we can define its *infinitesimal generator* as the operator

$$Ax = \lim_{t\to 0^+} \frac{T(t)x - x}{t} \text{ for } \mathscr{D}(A) := \Big\{x \in X : \exists \lim_{t\to 0^+} \frac{T(t)x - x}{t}\Big\},$$

where $\mathscr{D}(A)$ is the domain of A.

Remark 4.2. A strongly continuous semigroup is sometimes referred to as a C_0-semigroup. Our main interest lies with C_0-semigroups since we are considering unbounded operators on X, but for completion we state some properties of *uniformly continuous* semigroups (see [3, §1.1]). An operator A generates (i.e. is the infinitesimal generator for) a uniformly continuous semigroup if and only if $A \in \mathcal{L}(X)$, in which case the semigroup is given by $\{e^{tA}\}_{t\geq 0}$. Every semigroup has a unique infinitesimal generator by definition, and on the other hand every bounded linear operator A gives rise to a unique uniformly continuous semigroup. Further, for any uniformly continuous semigroup $\{T(t)\}_{t\geq 0}$, there exists $\omega \geq 0$ with $\|T(t)\| \leq e^{\omega t}$ for all $t \geq 0$ and $t \mapsto T(t)$ is differentiable in norm with derivative $AT(t) = T(t)A$.

Proposition 4.3. *Let $\{T(t)\}_{t\geq 0}$ be a C_0-semigroup. Then there exists $\omega \geq 0$ and $M \geq 1$ with*

$$\|T(t)\| \leq Me^{\omega t}, \quad t \geq 0.$$

Proof. We claim first that $\|T(t)\|$ is bounded on $[0, \eta]$ for some $\eta > 0$. Suppose not, then there exists a null sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}_+$ with $\|T(t_n)\| \geq n$. From the Banach-Steinhaus theorem, for some $x \in X$, $T(t_n)x$ is unbounded but this contradicts the strong continuity of the semigroup since $T(t_n)x \to x$ in norm. Therefore there exist $M, \eta \geq 0$ such that $\|T(t)\| \leq M$ for $t \in [0, \eta]$, where in fact $M \geq 1$ since $T(0) = I$ and $\|I\| = 1$. Let $\omega = \eta^{-1} \log M \geq 0$, then for any $t \geq 0$ we have that $t = n\eta + \delta$ for some $0 \leq \delta < \eta$, hence $\|T(t)\| = \|T(\delta)T(\eta)^n\| \leq MM^n = Me^{\omega t}$. □

We write $A \in \mathcal{G}(M, \omega)$ if A generates a C_0-semigroup satisfying $\|T(t)\| \leq Me^{\omega t}$, as above.

Definition 4.4. The *type* or *uniform growth bound* of the semigroup $\{T(t)\}_{t\geq 0}$ is defined by

$$\omega_0 = \omega_0(T) := \inf\{\omega : \exists M > 0 \text{ such that } \|T(t)\| \leq Me^{\omega t}\ \forall t \geq 0\}.$$

Now, the one-sided continuity of $t \mapsto T(t)x$ as $t \to 0^+$ is equivalent to full continuity.

Lemma 4.5. *If $\{T(t)\}_{t\geq 0}$ is a C_0-semigroup, then $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is continuous.*

Proof. Let $t \geq 0$, then for any $h \geq 0$ we have, using proposition 4.3, that

$$\|T(t+h)x - T(t)x\| \leq \|T(t)\|\|T(h)x - x\| \leq Me^{\omega t}\|T(h)x - x\|,$$

and for $t \geq h \geq 0$, we have $\|T(t-h)x - T(t)x\| \leq \|T(t-h)\|\|x - T(h)x\| \leq Me^{\omega t}\|x - T(h)x\|$, hence continuity follows by taking $h \to 0$ and using the strong continuity of the semigroup. □

Corollary 4.6. *For any C_0-semigroup $\{T(t)\}_{t\geq 0}$, for any $x \in X$ we have that*

$$\lim_{h\to 0^+} \frac{1}{h}\int_t^{t+h} T(s)x \,\mathrm{d}s = T(t)x.$$

Remark 4.7. Let I be a closed interval of the form $[a, b]$ for some $a, b \in \mathbb{R}$ with $a < b$ and also let $f : I \to X$ be a continuous function. If $(A, \mathscr{D}(A))$ is a closed operator on X and $f(t)$ is any $\mathscr{D}(A)$-valued function such that both $f(t)$ and $Af(t)$ are continuous on X, then it follows that

$$\int_a^b Af(t)\,\mathrm{d}t = A\int_a^b f(t)\,\mathrm{d}t. \tag{4.2}$$

Consider the case where $X = \mathbb{R}$, and define a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ in the following way. Since f is continuous, it is Riemann integrable, thus we consider a uniform partition of $[a, b]$ into $n+1$ subintervals; that is, define $\{u_0, \dots, u_n\}$ by $u_i = a+(b-a)i/n$. Then we can define the integral of f over $[a, b]$ as the limit of Riemann-Darboux sums, that is

$$x := \int_a^b f(t)\,\mathrm{d}t = \lim_{n\to\infty}\sum_{i=1}^n f(u_{i-1})\frac{1}{n} =: \lim_{n\to\infty} x_n.$$

Then since A is a linear operator on $\mathbb{R}$, it is continuous and therefore

$$\lim_{n\to\infty} Ax_n = \lim_{n\to\infty}\sum_{i=1}^n Af(u_{i-1})\frac{1}{n} = \int_a^b Af(t)\,\mathrm{d}t =: y.$$

Then using that A is closed, $Ax = y$ and $x \in \mathscr{D}(A)$, and this is exactly (4.2). The above ideas can be generalised to prove the result for any space X using the *Bochner integral* [1, §2.1.5]. In particular, the equality (4.2) also holds for elements of a semigroup $\{T(t)\}_{t\geq 0}$ because they are bounded, everywhere defined operators on X, and so closed by the closed graph theorem.

We now prove some elementary properties relating semigroups and their generators.

Proposition 4.8. *Let $\{T(t)\}_{t\geq 0}$ be a C_0-semigroup on X with infinitesimal generator $(A, \mathscr{D}(A))$.*

(i) If $x \in X$, then $\int_0^t T(s)x\,\mathrm{d}s \in \mathscr{D}(A)$ and

$$A\int_0^t T(s)x\,\mathrm{d}s = T(t)x - x.$$

(ii) If $x \in \mathscr{D}(A)$, then $T(t)x \in \mathscr{D}(A)$ and $\mathbb{R}_+ \ni t \mapsto T(t)x \in X$ is differentiable with

$$\frac{\mathrm{d}}{\mathrm{d}t} T(t)x = AT(t)x = T(t)Ax. \tag{4.3}$$

Consequently, for any $t, s \geq 0$, we obtain that

$$T(t)x - T(s)x = \int_s^t T(\tau)Ax \,\mathrm{d}\tau = \int_s^t AT(\tau)x \,\mathrm{d}\tau. \tag{4.4}$$

Proof. Let $x \in X$ and $h > 0$. Then using (4.2), we obtain that

$$\frac{T(h) - I}{h} \int_0^t T(s)x \,\mathrm{d}s = \frac{1}{h} \int_0^t (T(s+h)x - T(s)x) \,\mathrm{d}s = \frac{1}{h} \int_t^{t+h} T(s)x \,\mathrm{d}s - \frac{1}{h} \int_0^h T(s)x \,\mathrm{d}s,$$

and so taking h to zero and using corollary 4.6 we obtain the result in (i). If $x \in \mathscr{D}(A)$, then

$$\frac{T(h) - I}{h} T(t)x = T(t) \left(\frac{T(h) - I}{h} \right) x,$$

and so taking h to zero proves that $T(t)x \in \mathscr{D}(A)$ and also $AT(t)x = T(t)Ax$. Further, this also implies that $t \mapsto T(t)x$ is differentiable with the required derivative. Finally, (4.4) follows by integrating (4.3) over the interval $[s, t]$. □

Remark 4.9. As a result, the differential equation in X given by

$$\frac{\partial u(t)}{\partial t} = Au(t) \text{ for } t > 0 \text{ with } \lim_{t \to 0^+} u(t) = x$$

has solution given by $u(t) = T(t)x$, provided that $x \in \mathscr{D}(A)$.

Theorem 4.10. *The generator A of a C_0-semigroup $\{T(t)\}_{t \geq 0}$ has dense domain and is closed.*

Proof. For $x \in X$, let $x_t := t^{-1} \int_0^t T(s)x \,\mathrm{d}s$. Then we have $x_t \in \mathscr{D}(A)$ for $t > 0$ by proposition 4.8(i) and $x_t \to x$ as $t \to 0^+$ by continuity, so $\mathscr{D}(A)$ is dense in X. For closedness, let $\mathscr{D}(A) \ni x_n \to x$ and $Ax_n \to y$ in X as $n \to \infty$. Then by (4.4),

$$T(t)x_n - x_n = \int_0^t T(s)Ax_n \,\mathrm{d}s.$$

The integrand converges uniformly on bounded intervals to $T(s)y$, so taking $n \to \infty$ gives

$$T(t)x - x = \int_0^t T(s)y \,\mathrm{d}s.$$

Diving by $t > 0$ and taking $t \to 0^+$ it is clear that $x \in \mathscr{D}(A)$ and $Ax = y$, so A is closed. □

Lemma 4.11. *A generates the semigroup $\{T(t)\}_{t \geq 0}$ if and only if $A - \omega I$ generates $\{e^{-\omega t} T(t)\}_{t \geq 0}$.*

Proof. Suppose that A generates the semigroup $\{T(t)\}_{t \geq 0}$ and define $S(t) := e^{-\omega t} T(t)$. Clearly $S(t+s) = S(t)S(s)$ for any $t, s \geq 0$ and $S(0) = e^{-\omega 0} T(0) = I$. Also, we have that

$$\|S(t)x - x\| = \|e^{-\omega t} T(t)x - x\| \leq e^{-\omega t} \|T(t)x - x\| + |e^{-\omega t} - 1| \|x\|,$$

and so taking $t \to 0^+$ we see that $\{S(t)\}_{t \geq 0}$ generates a C_0-semigroup with generator $A - \omega I$ as

$$\lim_{t \to 0^+} \frac{S(t)x - x}{t} = \lim_{t \to 0^+} \left[e^{-\omega t} \frac{T(t)x - x}{t} + \frac{e^{-\omega t} - 1}{t} x \right] = Ax - \omega x,$$

which holds for any $x \in \mathscr{D}(A)$. Conversely suppose that $\{S(t)\}_{t\geq 0}$ is a semigroup with generator $A-\omega I$. Then for any $t,s\geq 0$, $S(t+s)=S(t)S(s)$ implies that $T(t+s)=T(t)T(s)$ by division by $e^{-\omega(t+s)}$, and $I=S(0)=T(0)$. Moreover for any $x\in\mathscr{D}(A-\omega I)=\mathscr{D}(A)$,

$$\lim_{t\to 0^+}\frac{T(t)x-x}{t}=\lim_{t\to 0^+}\left[e^{\omega t}\frac{e^{-\omega t}T(t)x-x}{t}+\frac{e^{\omega t}-1}{t}x\right]=(A-\omega I)x+\omega x=Ax,$$

and so A is the generator of the semigroup $\{T(t)\}_{t\geq 0}$, as required. □

Example 4.12 (Translation semigroup)**.** Consider the Banach space $X=\mathrm{BUC}(\mathbb{R})$ containing the uniformly continuous and bounded functions on $\mathbb{R}$ with the supremum norm. Then define

$$(T(t)f)(s):=f(t+s)\text{ for any }f\in X\text{ and fixed }t\geq 0.$$

This defines a C_0-semigroup (of *contractions*) $\{T(t)\}_{t\geq 0}$. To see this, we have the following.

(i) It is clear that $T(0)=I$ since $(T(0)f)(s)=f(s)$.

(ii) For any $t,t'\geq 0$, by definition $(T(t+t')f)(s)=f(t+t'+s)$. Now defining the function $f_{t'}$ by $f_{t'}(s)=f(t'+s)$ for $s\in\mathbb{R}$, we obtain that $(T(t)T(t')f)(s)=(T(t)f_{t'})(s)=f_{t'}(t+s)$, and hence $T(t+t')=T(t)T(t')$ for any $t,t'\geq 0$.

(iii) Let $f\in X$ and let $\varepsilon>0$ and consider the following,

$$\|T(t)f-f\|=\sup_{s\in\mathbb{R}}|(T(t)f)(s)-f(s)|=\sup_{s\in\mathbb{R}}|f(t+s)-f(s)|. \tag{4.5}$$

Since f is uniformly continuous, there exists a $\delta>0$ such that for all $t',s'\in\mathbb{R}$, if $|t'-s'|<\delta$ then $|f(t')-f(s')|<\varepsilon$. Since we are interested in taking $t\to 0^+$ in (4.5), we can assume that $|(t+s)-s|=|t|<\delta$. Then $\|T(t)f-f\|<\varepsilon$, proving strong continuity.

Hence $\{T(t)\}_{t\geq 0}$ is indeed a C_0-semigroup. To see that it is a semigroup of contractions, consider

$$\|T(t)\|=\sup_{\|f\|\leq 1}\|T(t)f\|=\sup_{\|f\|\leq 1}\|f\|\leq 1,\quad \forall t\geq 0.$$

The infinitesimal generator of this semigroup is defined by $(Af)(s)=f'(s)$ for $f\in\mathscr{D}(A)$ since

$$(Af)(s):=\lim_{h\to 0^+}\frac{T(h)f-f}{h}(s)=\lim_{h\to 0^+}\frac{f(s+h)-f(s)}{h},$$

and the domain $\mathscr{D}(A)$ consists of all $f\in X$ for which f' exists and $f'\in X$.

4.1. **Hille-Yosida Generation Theorem.** Given a C_0-semigroup $\{T(t)\}_{t\geq 0}$ with infinitesimal generator A, we know that the solution of the initial value problem $\partial_t u(t)=Au(t)$ on $t>0$ with $u(0)=u_0$ is given by $u(t)=T(t)u_0$ if $u_0\in\mathscr{D}(A)$, however with most problems the operator A will be given and it is not immediately obvious what semigroup it generates (it might not even generate a semigroup). In turns out that there are necessary and sufficient conditions for a linear operator to generate a C_0-semigroup, and this is realised in the theorem of Einar Hille and Kōsaku Yosida, or more generally by William Feller, Isao Miyadera and Ralph Phillips.

Theorem 4.13 (Feller-Miyadera-Phillips)**.** *$A\in\mathcal{G}(M,\omega)$ if and only if*

(i) A is closed and densely defined; and

(ii) $(\omega,\infty)\subseteq\varrho(A)$ and for all $n\geq 1$ and $\lambda>\omega$,

$$\|(\lambda I-A)^{-n}\|\leq\frac{M}{(\lambda-\omega)^n}. \tag{4.6}$$

Proof. Suppose that $A \in \mathcal{G}(M,\omega)$, then A is closed and densely defined by theorem 4.10, and by assumption we have that A generates a C_0-semigroup $\{T(t)\}_{t\geq 0}$ which satisfies $\|T(t)\| \leq Me^{\omega t}$ for some $M \geq 1$ and $\omega \in \mathbb{R}$. Now for $\lambda > \omega$, we define

$$R(\lambda)x := \int_0^\infty e^{-\lambda t} T(t)x \,\mathrm{d}t. \tag{4.7}$$

This is well-defined since $\|e^{-\lambda t}T(t)x\| \leq e^{-\lambda t}\|T(t)\|\|x\| \leq Me^{(\omega-\lambda)t}\|x\|$ and therefore

$$\int_0^\infty \|e^{-\lambda t}T(t)x\| \,\mathrm{d}t \leq M\|x\| \int_0^\infty e^{(\omega-\lambda)t} \,\mathrm{d}t = \frac{M\|x\|}{\lambda-\omega} < \infty.$$

Consequently, $R(\lambda)$ defines a bounded linear operator on X. Now for $h > 0$,

$$\begin{aligned}\frac{T(h)-I}{h}R(\lambda)x &= \frac{1}{h}\int_0^\infty e^{-\lambda t}(T(t+h)-T(t))x \,\mathrm{d}t \\ &= \frac{e^{\lambda h}}{h}\int_h^\infty e^{-\lambda\tau}T(\tau)x \,\mathrm{d}\tau - \frac{1}{h}\int_0^\infty e^{-\lambda t}T(t)x \,\mathrm{d}t \text{ where } \tau := t+h \\ &= \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}T(t)x \,\mathrm{d}t - \frac{1}{h}\int_0^h e^{-\lambda t}T(t)x \,\mathrm{d}t\end{aligned}$$

Then taking $h \to 0^+$ we obtain that $R(\lambda)x \in \mathscr{D}(A)$ for any $x \in X$ and $AR(\lambda)x = \lambda R(\lambda)x - x$. Equivalently, $(\lambda I - A)R(\lambda)x = x$ for any $x \in X$ so that $(\lambda I - A)R(\lambda) = I$. Also for $x \in \mathscr{D}(A)$,

$$R(\lambda)Ax = \int_0^\infty e^{-\lambda t}T(t)Ax \,\mathrm{d}t = \int_0^\infty e^{-\lambda t}AT(t)x \,\mathrm{d}t = AR(\lambda)x,$$

using (4.4) and also (4.2) since A is a closed operator. Therefore also $R(\lambda)Ax = \lambda R(\lambda)x - x$ for any $x \in \mathscr{D}(A)$ so that $R(\lambda)(\lambda I - A)x = x$. Since $\mathscr{D}(A)$ is dense in X, this also holds for any $x \in X$. Consequently we obtain that $R(\lambda) = R(\lambda, A)$ and so necessarily $(\omega,\infty) \subseteq \varrho(A)$ because (4.7) was well-defined for any $\mathbb{R} \ni \lambda > \omega$. Now for the estimate (4.6), we have for $\lambda > \omega$ that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}R(\lambda,A)x = \frac{\mathrm{d}}{\mathrm{d}\lambda}\int_0^\infty e^{-\lambda t}T(t)x \,\mathrm{d}t = -\int_0^\infty te^{-\lambda t}T(t)x \,\mathrm{d}t,$$

and therefore by induction, for any $n \geq 1$,

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}R(\lambda,A)x = (-1)^n\int_0^\infty t^n e^{-\lambda t}T(t)x \,\mathrm{d}t. \tag{4.8}$$

On the other hand, as a result of the resolvent identity, we have from (2.2) that

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}R(\lambda,A) = (-1)^n n! R(\lambda,A)^{n+1}.$$

Therefore by comparing this and (4.8) we see that for $n \geq 1$,

$$R(\lambda,A)^n x = \frac{1}{(n-1)!}\int_0^\infty t^{n-1}e^{-\lambda t}T(t)x \,\mathrm{d}t.$$

As a result, taking norms and using that $\|T(t)\| \leq Me^{\omega t}$ we obtain that

$$\|(\lambda I - A)^{-n}x\| = \|R(\lambda,A)^n x\| \leq \frac{M\|x\|}{(n-1)!}\int_0^\infty t^{n-1}e^{-(\lambda-\omega)t} \,\mathrm{d}t = \frac{M\|x\|}{(\lambda-\omega)^n}.$$

The converse relies on the *Yosida approximations* of A, namely the bounded operators

$$A_\lambda = \lambda AR(\lambda,A) = \lambda^2 R(\lambda,A) - \lambda I.$$

For a proof, refer to [3, §1.3–§1.5]. ☐

When $M = 1$, the conditions of the generation theorem greatly simplify to the following.

Corollary 4.14 (Hille-Yosida). *$A \in \mathcal{G}(1, \omega)$ if and only if*

(i) A is closed and densely defined; and

(ii) $(\omega, \infty) \subseteq \varrho(A)$ and for any $\lambda > \omega$, we have $\|(\lambda I - A)^{-1}\| \leq (\lambda - \omega)^{-1}$.

Proof. ($\Rightarrow$) follows by theorem 4.13 and ($\Leftarrow$) follows again using the Yosida approximations. □

Remark 4.15. Taking $\omega = 0$ in corollary 4.14 gives conditions for A to generate a semigroup of contractions, that is a semigroup satisfying $\|T(t)\| \leq 1$ for all $t \geq 0$. In fact, we can recover the Hille-Yosida theorem from this contraction semigroup case. Indeed, suppose that $\{T(t)\}_{t \geq 0}$ is a C_0-semigroup satisfying $\|T(t)\| \leq e^{\omega t}$ for $t \geq 0$ and some $\omega \in \mathbb{R}$. Then consider $S(t) = e^{-\omega t} T(t)$. Clearly $\{S(t)\}_{t \geq 0}$ is a semigroup of contractions, with generator $A - \omega I$ by lemma 4.11, where A is the generator of $\{T(t)\}_{t \geq 0}$. Then $A - \omega I$ is closed and densely defined, hence A is closed by lemma 2.10 and densely defined as $\mathscr{D}(A) = \mathscr{D}(A - \omega I)$. Further, $(0, \infty) \subseteq \varrho(A - \omega I)$ and for any $\lambda > 0$ we have that $\|R(\lambda, A - \omega I)\| \leq \lambda^{-1}$. Now if $\lambda > \omega$, then $\lambda - \omega > 0$ and the operator $(\lambda - \omega)I - (A - \omega I) = \lambda I - A$ is invertible, hence $(\omega, \infty) \subseteq \varrho(A)$. Moreover, we have that

$$\|R(\lambda - \omega, A - \omega I)\| \leq \frac{1}{\lambda - \omega},$$

but in fact $R(\lambda - \omega, A - \omega I) = R(\lambda, A)$. Hence if we only assume that the Hille-Yosida theorem works for contractions semigroups, that is semigroups with $\|T(t)\| \leq 1$ for all $t \geq 0$, then we can deduce a generation theorem for semigroups with $\|T(t)\| \leq e^{\omega t}$ for some $\omega \in \mathbb{R}$.

Remark 4.16. The resolvent formula in (4.7) also holds for any $\lambda \in \mathbb{C}$ with $\Re[\lambda] > \omega_0(T)$, so

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda t} T(t)x \, \mathrm{d}t. \tag{4.9}$$

This result can be proven in a similar way to that in the proof of the Hille-Yosida theorem.

4.2. **Lumer-Phillips Generation Theorem.** We now consider the *Lumer-Phillips* generation theorem, which relates to dissipative operators, but first we recall the Hanh-Banach theorem. Note that for a normed space X, the dual space $X^* = \mathcal{L}(X, \mathbb{F})$ is the space of all continuous (or bounded) linear functionals on X, and we usually denote by $\langle x^*, x \rangle$ the value of the functional $x^* \in X^*$ at the point $x \in X$.

Theorem 4.17 (Hanh-Banach). *Let X be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $p : X \to \mathbb{R}$ be a seminorm, that is a function which is subadditive and homogeneous. Suppose that $Z \subseteq X$ is a linear subspace and $f : Z \to \mathbb{F}$ is a linear functional satisfy $|f(z)| \leq p(z)$ for all $z \in Z$. Then there exists a linear functional $F : X \to \mathbb{F}$ such that $F|_Z = f$ and $|F(x)| \leq p(x)$ for all $x \in X$.*

Corollary 4.18. *Let X be a normed space and $x \neq 0$.*

Then there exists an $x^* \in X^*$ with $\langle x^*, x \rangle = \|x\|^2 = \|x^*\|^2$.

Proof. Let $f \in V^*$ for some linear subspace $V \subseteq X$. The function $p(x) = \|f\| \|x\|$ for $x \in X$ is a seminorm which satisfies $|f(v)| \leq p(v)$ for all $v \in V$. Therefore by the Hanh-Banach theorem, there exists an extension F with $F|_V = f$ and $|F(x)| \leq \|f\| \|x\|$ for all $x \in X$, thus $\|F\| \leq \|f\|$. In fact $\|F\| = \|f\|$ as $F|_Z = f$. Now consider the subspace $Z = \mathbb{F}x$ and define $g \in L(Z, \mathbb{F})$ by

$g(\alpha x) = \alpha\|x\|$. Now any $\alpha x \in Z$ has norm one if and only if $|\alpha| = 1/\|x\|$ and thus $\|g\| = 1$, that is $g \in Z^*$. Thus from the above, there exists a $G \in X^*$ with $G|_Z = g$ and also $\|G\| = \|g\| = 1$. So taking $x^* = \|x\|G$ we have $\langle x^*, x\rangle = \|x\|G(x) = \|x\|^2$ as $x \in Z$ and $\|x^*\| = \|x\|\|G\| = \|x\|$. □

Definition 4.19. Let $x \in X$. Then the *duality set* $F(x) \subseteq X^*$ of x is the set

$$F(x) = \{x^* : x^* \in X^* \text{ and } \langle x^*, x\rangle = \|x\|^2 = \|x^*\|^2\}.$$

This set is non-empty for all $x \in X\backslash\{0\}$ by corollary 4.18 and since $0 \in F(0)$.

Since the duality set is non-empty, we are now in a position to define dissipative operators.

Definition 4.20. A linear operator $(A, \mathscr{D}(A))$ in X is called *dissipative* if

$$\forall x \in \mathscr{D}(A)\, \exists x^* \in F(x) \text{ such that } \Re\langle x^*, Ax\rangle \le 0.$$

The following proposition gives a useful characterisation of dissipative operators.

Proposition 4.21. *A linear operator $(A, \mathscr{D}(A))$ in X is dissipative if and only if*

$$\|(\lambda I - A)x\| \ge \lambda\|x\| \text{ for all } x \in \mathscr{D}(A) \text{ and } \lambda > 0. \tag{4.10}$$

Proof. Suppose that A is dissipative and that $x \in \mathscr{D}(A)$ and $\lambda > 0$ are arbitrary. If $x = 0$, then (4.10) holds immediately, so assume now that $x \neq 0$. Now because A is dissipative, there exists an $x^* \in F(x)$, so $\|x^*\| = \|x\|$, such that $\Re\langle x^*, Ax\rangle \le 0$ and therefore we obtain that

$$\begin{aligned}\|(\lambda I - A)x\|\|x\| &= \|(\lambda I - A)x\|\|x^*\| \ge |\langle x^*, (\lambda I - A)x\rangle| \\ &\ge \Re\langle x^*, (\lambda I - A)x\rangle = \Re[\lambda\langle x^*, x\rangle - \langle x^*, Ax\rangle] \ge \lambda\|x\|^2.\end{aligned}$$

The converse statement follows by the Banach-Alaoglu theorem, which states that the unit ball in X^* is compact in the weak-star topology. For a proof of this, see [3, Theorem 1.4.2]. □

Lemma 4.22. *Suppose that A is a dissipative operator in X and $\lambda_0 > 0$.*
If $\operatorname{im}(\lambda_0 I - A) = X$*, then* $\operatorname{im}(\lambda I - A) = X$ *for all* $\lambda > 0$*.*

Proof. Let $\Delta \subseteq \mathbb{C}$ be the set of $\lambda \in \mathbb{C}$ with $\lambda \in (0, \infty)$ and $\operatorname{im}(\lambda I - A) = X$. Now if $\Delta = (0, \infty)$, then the result holds, so we show that Δ is both open and closed in $[0, \infty]$ and non-empty (since then the connectedness of the interval $(0, \infty)$ implies the only open and closed sets in $(0, \infty)$ are either $\emptyset$ or $(0, \infty)$). To show that Δ is open, let $\lambda \in \Delta$. Now by definition, $\operatorname{im}(\lambda I - A) = X$ so that $\lambda I - A$ is onto, and if $x \in \mathscr{D}(A)$ is such that $(\lambda I - A)x = 0$, then by (4.10) we have that $\lambda\|x\| = 0$ and so $\lambda I - A$ is one-to-one. Therefore $(\lambda I - A)^{-1}$ exists and so $\lambda \in \varrho(A)$. Now by theorem 2.25, the resolvent set is open and so there exists a neighbourhood U of λ in $\varrho(A)$. Then $U \cap (0, \infty) \subseteq \Delta$ and so Δ is open in $(0, \infty)$. To show that Δ is closed in $(0, \infty)$, consider a sequence $(\lambda_n)_{n\in\mathbb{N}} \subseteq \Delta$ with $\lambda_n \to \lambda > 0$. Then $\lambda \in \Delta$ if and only if $\operatorname{im}(\lambda I - A) = X$. Clearly $\operatorname{im}(\lambda I - A) \subseteq X$, so let $y \in X$. Now for all $n \in \mathbb{N}$ there is an $x_n \in \mathscr{D}(A)$ with $(\lambda_n I - A)x_n = y$, and so by (4.10) we have that $\|x_n\| \le \lambda_n^{-1}\|y\|$ and as $\lambda_n^{-1} \to \lambda^{-1}$ is convergent, it is bounded, so in fact $\|x_n\| \le C$ for some $C > 0$. Hence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence since

$$\begin{aligned}\lambda_m\|x_n - x_m\| &\le \|\lambda_m(x_n - x_m) - A(x_n - x_m)\| = \|(\lambda_m I - A)x_n - (\lambda_m I - A)x_m\| \\ &= \|(\lambda_m I - A)x_n - (\lambda_n I - A)x_n\| = |\lambda_n - \lambda_m|\|x_n\| \le C|\lambda_n - \lambda_m|.\end{aligned}$$

Since X is complete, $x_n \to x$ for some $x \in X$ and $Ax_n = \lambda_n x_n - y \to \lambda x - y$. Now as in the proof of theorem 4.23(i), A is a closed operator as it is dissipative, and thus $x \in \mathscr{D}(A)$ and also $Ax = \lambda x - y$, that is $y \in \operatorname{im}(\lambda I - A)$. Therefore we have that $\operatorname{im}(\lambda I - A) = X$ and hence $\lambda \in \Delta$. Now since $\lambda_0 \in \Delta$, we have that $\Delta \neq \emptyset$ and so $\Delta = (0, \infty)$, as required. □

With the above results, we can now state and prove the Lumer-Phillips theorem.

Theorem 4.23 (Lumer-Phillips). *Let $(A, \mathscr{D}(A))$ be a linear operator in X with dense domain.*

- *(i) If A is dissipative and there is a $\lambda_0 > 0$ for which $\operatorname{im}(\lambda_0 I - A) = X$, then A is the generator of a C_0-semigroup of contractions in X.*
- *(ii) If A is the generator of a C_0-semigroup of contractions in X, then $\operatorname{im}(\lambda I - A) = X$ for all $\lambda > 0$ and A is dissipative. Moreover, $\Re\langle x^*, Ax\rangle \leq 0$ for all $x \in \mathscr{D}(A)$ and $x^* \in F(x)$.*

Proof. (i) By lemma 4.22 we have that $\operatorname{im}(\lambda I - A) = X$ for all $\lambda \in (0, \infty)$. Now let $\lambda \in (0, \infty)$, then from (4.10) we have that $\lambda I - A$ is one-to-one. Indeed, if $(\lambda I - A)x = 0$ then $\lambda\|x\| = 0$, that is $x = 0$. Hence $(\lambda I - A)^{-1}$ exists as an operator $X \to \mathscr{D}(A)$, and it is bounded as if $x \in X$, then $x = (\lambda I - A)y$ for some $y \in \mathscr{D}(A)$ and so (4.10) gives

$$\|(\lambda I - A)^{-1}x\| = \|y\| \leq \frac{1}{\lambda}\|(\lambda I - A)y\| = \frac{1}{\lambda}\|x\|. \tag{4.11}$$

As the operator $(\lambda I - A)^{-1}$ is bounded, it is closed by the closed graph theorem. Then by lemma 2.10, $\lambda I - A$ is closed and hence so is A. Moreover, (4.11) gives that $\|R(\lambda, A)\| \leq \lambda^{-1}$ and clearly $\lambda \in \varrho(A)$ from the above. Then by Hille-Yosida (corollary 4.14) we have that A generates a semigroup of contractions in X.

(ii) We have that $A \in \mathcal{G}(1, 0)$ as A generates a semigroup of contractions, and so $(0, \infty) \subseteq \varrho(A)$ from the Hille-Yosida theorem (in particular, corollary 4.14). Then for any $\lambda \in (0, \infty)$ we have that $\lambda I - A$ is an invertible operator $\mathscr{D}(A) \to X$, and so $\operatorname{im}(\lambda I - A) = X$. Now let $x \in \mathscr{D}(A)$ and $x^* \in F(x)$ be arbitrary. Then it follows that

$$\begin{aligned} |\langle x^*, T(t)x\rangle| &\leq \|x^*\|\|T(t)x\| \leq \|T(t)\|\|x\|\|x^*\| = \|x\|^2 \\ \Rightarrow \Re\langle x^*, T(t)x - x\rangle &= \Re\langle x^*, T(t)x\rangle - \|x\|^2 \leq 0. \end{aligned}$$

Then diving the above by $t > 0$ and taking the limit as $t \to 0^+$ we obtain that

$$\Re\langle x^*, Ax\rangle = \Re\left\langle x^*, \lim_{t\to 0^+} \frac{T(t)x - x}{t}\right\rangle \leq 0,$$

using the continuity of x^* and $\Re(\cdot)$. In particular, A is a dissipative operator. □

4.3. **Positive Semigroups.** A positive semigroup is itself a semigroup but which consists of a family of positive operators, as defined in definition 3.33.

Definition 4.24. Let X be a Banach lattice. Then the semigroup $\{T(t)\}_{t\geq 0}$ on X is *positive* if for any $x \in X_+$ and $t \geq 0$, we have that $T(t)x \succeq 0$. Further, an operator $(A, \mathscr{D}(A))$ is said to be *resolvent positive* if there exists $\omega \in \mathbb{R}$ with $(\omega, \infty) \subseteq \varrho(A)$ and $R(\lambda, A) \succeq 0$ for all $\lambda > \omega$.

If $(A, \mathscr{D}(A))$ is a bounded operator in X, then it generates a C_0-semigroup and the semigroup is defined by $T(t) = e^{tA}$ for $t \geq 0$. Now when A is unbounded, provided it satisfies the conditions of the Hille-Yosida or Lumer-Phillips generation theorems, it generates a C_0 semigroup

$\{T(t)\}_{t\geq 0}$ but we can not make sense of e^{tA} for unbounded A. Despite this, C_0-semigroups are generalisations of the exponential and thus we would expect some properties to carry over.

It is a well-known fact that the scalar exponential e^{at} can be expressed as the following limit,

$$\lim_{n\to\infty}\Big(1+\frac{ta}{n}\Big)^n = \lim_{n\to\infty}\Big(1-\frac{ta}{n}\Big)^{-n} = e^{at},$$

and the following is an analogous result for semigroups.

Theorem 4.25 (Exponential formula)**.** *Let $\{T(t)\}_{t\geq 0}$ be a C_0-semigroup on X.*

If A is the infinitesimal generator of $\{T(t)\}_{t\geq 0}$, then for any $x \in X$,

$$T(t)x = \lim_{n\to\infty}\Big(I-\frac{t}{n}A\Big)^{-n}x = \lim_{n\to\infty}\Big[\frac{n}{t}R\Big(\frac{n}{t},A\Big)\Big]^n x,$$

and the limit is uniform in t on any bounded interval.

Proof. See [3, Theorem 1.8.3]. □

The proof relies on the resolvent formula (4.9) and also (2.2).

Proposition 4.26. *A C_0-semigroup is positive if and only if its generator is resolvent positive.*

Proof. Let $\{T(t)\}_{t\geq 0}$ be the semigroup with infinitesimal generator A. If A is resolvent positive, then from the exponential formula in theorem 4.25 we see that $T(t)x \succeq 0$ for any $x \in X_+$. Now if the semigroup is positive, then as in remark 4.16 we have for any $x \in X$ that

$$R(\lambda,A)x = \int_0^\infty e^{-\lambda t}T(t)x\,\mathrm{d}t, \quad \Re[\lambda] > \omega_0(T).$$

Therefore for $\mathbb{R} \ni \lambda > \omega_0(T)$ we have that $R(\lambda,A)x \succeq 0$ for all $x \in X_+$. □

5. Perturbations

Let X be a Banach space and consider the differential equation in X given by

$$\frac{\partial u(t)}{\partial t} = Au(t),\ t > 0 \text{ and } u(0) = u_0 \in X. \tag{5.1}$$

As discussed in the previous chapter, the issue of finding a semigroup $\{T(t)\}_{t\geq 0}$ with generator A in order to find an explicit solution to (5.1) is given by the Hille-Yosida theorems. If such a semigroup where to exist, it is necessary that the operator A is closed and densely defined and satisfies some operator norm bounds on the resolvent $R(\lambda, A)$ for sufficiently large λ. We now consider a different approach, whereby we deconstruct A into the sum of two operators, namely $A = A_1 + A_2$. If, without loss of generality, A_1 is the infinitesimal generator of some semigroup $\{T_1(t)\}_{t\geq 0}$, then we consider whether $A_1 + A_2$ generates a C_0-semigroup as a perturbation of the semigroup $\{T_1(t)\}_{t\geq 0}$. The advantage here is that the operator A_1 may be a relatively nicer operator than A itself, and a generator may be more easily found. However, there is a bit of work involved on the perturbation side. So more formally, we consider the following problem.

Problem. If $(A, \mathscr{D}(A))$ generates a C_0-semigroup on a Banach space X and $(B, \mathscr{D}(B))$ is another operator in X, then under what conditions does $A + B$ generate a C_0-semigroup.

Such a problem already has some potential errors. As an example, the operator $A + B$ is only defined on $\mathscr{D}(A) \cap \mathscr{D}(B)$, which may be empty. Thus we make the slightly stronger assumption that $\mathscr{D}(A) \subseteq \mathscr{D}(B)$, and also that B is A-bounded. This means that there exists $a, b \geq 0$ with

$$\|Bx\| \leq a\|Ax\| + b\|x\|, \quad \forall x \in \mathscr{D}(A),$$

and $\mathscr{D}(A) \subseteq \mathscr{D}(B)$. This assumption is needed as having $\mathscr{D}(A) \subseteq \mathscr{D}(B)$ alone is too restrictive. We also generalise further, to asking if there exists an extension K of $A + B$ which generates a C_0-semigroup. In the following, we assume that $(A, \mathscr{D}(A))$ generates a C_0-semigroup, which we often denote by $\{T_A(t)\}_{t\geq 0}$, and $(B, \mathscr{D}(B))$ is the *perturbing* operator.

5.1. A Spectral Criterion.

The following theorem gives a nice characterisation for when K, an extension of the operator $A+B$, generates a semigroup. The importance of this theorem can not be understated, and it will be used when discussing bounded and Miyadera perturbations. Recall the definitions of the continuous, point and residual spectrum and the resolvent set.

Theorem 5.1. *Assume that $\Lambda := \varrho(A) \cap \varrho(K) \neq \emptyset$. Then*

(i) $1 \notin \sigma_p(BR(\lambda, A))$ *for any* $\lambda \in \Lambda$;
(ii) $1 \in \varrho(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *iff* $\mathscr{D}(K) = \mathscr{D}(A)$ *and* $K = A + B$;
(iii) $1 \in \sigma_c(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *iff* $\mathscr{D}(A) \subsetneq \mathscr{D}(K)$ *and* $K = \overline{A + B}$; *and*
(iv) $1 \in \sigma_r(BR(\lambda, A))$ *for some/all* $\lambda \in \Lambda$ *iff* $K \supsetneq \overline{A + B}$.

In order to prove this result, we need some preliminary lemmas.

Lemma 5.2. *Suppose that $\varrho(A) \neq \emptyset$. Then B is A-bounded if and only if*

$$BR(\lambda, A) \in \mathcal{L}(X) \text{ for any } \lambda \in \varrho(A). \tag{5.2}$$

Proof. Suppose that B is A-bounded and $\lambda \in \varrho(A)$, then there exist $a, b \geq 0$ such that $\|Bx\| \leq a\|Ax\| + b\|x\|$ for all $x \in \mathscr{D}(A)$. Now for any $y \in X$, we have $R(\lambda, A)y \in \mathscr{D}(A)$ since $R(\lambda, A)X = \mathscr{D}(A)$. Therefore, as $AR(\lambda, A) = -I + \lambda R(\lambda, A)$, we have that

$$\|BR(\lambda, A)y\| \leq a\|AR(\lambda, A)y\| + b\|R(\lambda, A)y\| \leq M\|y\|,$$

for some $M \geq 0$ and so $BR(\lambda, A) \in \mathcal{L}(X)$. Conversely, suppose (5.2) holds and let $x \in \mathscr{D}(A)$. Then there exists a $y \in X$ such that $x = R(\lambda, A)y$ and thus

$$\|Bx\| = \|BR(\lambda, A)y\| \leq M\|y\| = M\|(\lambda I - A)x\| \leq M\|Ax\| + \lambda M\|x\|,$$

proving that B is A-bounded, as required. □

Lemma 5.3. *Let $\lambda \in \Lambda$ and $f \in X$. Then $R(\lambda, K)f \in \mathscr{D}(A)$ if and only if $f \in (I - BR(\lambda, A))X$. In particular, for all $x \in \mathscr{D}(A)$ there exists a $g \in X$ such that $x = R(\lambda, K)(I - BR(\lambda, A))g$.*

Proof. Suppose that $x = R(\lambda, K)f \in \mathscr{D}(A)$. Then since K is an extension of $A + B$, we have

$$f = (\lambda I - K)x = (\lambda I - A - B)x.$$

Therefore with $g = (\lambda I - A)x \in X$ we see that $f = g - Bx = g - BR(\lambda, A)g = (I - BR(\lambda, A))g$. Conversely, suppose that $f = (I - BR(\lambda, A))g$ for some $g \in X$. Then with $x = R(\lambda, A)g$,

$$f = ((\lambda I - A) - B)R(\lambda, A)g = (\lambda I - A - B)x.$$

As K is an extension of $A + B$, we have that

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

and by definition $x \in \mathscr{D}(A)$, proving the equality. Now if $x \in \mathscr{D}(A)$, then as $R(\lambda, A)X = \mathscr{D}(A)$ and the resolvent operator defines a bijective map, there exists a $g \in X$ such that $x = R(\lambda, A)g$. Then defining $f = (I - BR(\lambda, A))g$, from the above we see that $x = R(\lambda, K)f$, as required. □

For the following lemma, we recall the definition of the closure of an operator.

If $(A, \mathscr{D}(A))$ is a linear unbounded operator which is not necessarily closed, then $\overline{A}$ is the linear operator whose graph is the closure of A, that is $\mathscr{G}(\overline{A}) = \overline{\mathscr{G}(A)}$. In fact, we have the more formal definition where $\mathscr{D}(\overline{A})$ is the set of all $x \in X$ for which there exists $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A)$ and a $y \in X$ such that $x_n \to x$ in X and $Ax_n \to y$ in Y as $n \to \infty$ and for any $x \in \mathscr{D}(\overline{A})$,

$$\overline{A}x = y = \lim_{n\to\infty} Ax_n.$$

Lemma 5.4. *For any $\lambda \in \Lambda$, we have that $\mathscr{D}(\overline{A + B}) = R(\lambda, K)\overline{(I - BR(\lambda, A))X}$.*

Proof. Suppose first that $x \in \mathscr{D}(\overline{A + B})$. Now our underlying assumption is that $\mathscr{D}(A) \subseteq \mathscr{D}(B)$, so $A + B$ is an operator defined on $\mathscr{D}(A)$. Thus there exists an $y \in X$ and a sequence $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A)$ with $x_n \to x$ and $(A + B)x_n \to y$ as $n \to \infty$. Then

$$\|(\lambda I - (A + B))x_n - (\lambda x - y)\| \leq |\lambda|\|x_n - x\| + \|(A + B)x_n - y\|,$$

and thus with $f = \lambda x - y$ we have that $(\lambda I - (A + B))x_n \to f$ as $n \to \infty$. Since $R(\lambda, A)X = \mathscr{D}(A)$, for each $n \in \mathbb{N}$ we can write $x_n = R(\lambda, A)g_n$ for some $g_n \in X$ and obtain that

$$f = \lim_{n\to\infty} (\lambda I - (A + B))x_n = \lim_{n\to\infty} (I - BR(\lambda, A))g_n, \tag{5.3}$$

hence $f \in \overline{(I - BR(\lambda, A))X}$. Moreover, by lemma 5.3 we have that

$$x_n - R(\lambda, K)f = R(\lambda, A)g_n - R(\lambda, K)f = R(\lambda, K)((I - BR(\lambda, A))g_n - f).$$

Therefore we see that $\|x_n - R(\lambda, K)f\| \leq \|R(\lambda, K)\|\|(I - BR(\lambda, A))g_n - f\|$, thus by (5.3) we have that $x_n \to R(\lambda, K)f$ as $n \to \infty$ and so by uniqueness of limits,

$$x = R(\lambda, K)f \in R(\lambda, K)\overline{(I - BR(\lambda, A))X},$$

and clearly this holds for all $\lambda \in \Lambda$. Conversely, suppose that $x \in R(\lambda, K)\overline{(I - BR(\lambda, A))X}$, so

$$x = R(\lambda, K)f = \lim_{n\to\infty} R(\lambda, K)f_n =: \lim_{n\to\infty} x_n,$$

for some $(f_n)_{n\in\mathbb{N}} \subseteq (I - BR(\lambda, A))X$ which converges to $f \in \overline{(I - BR(\lambda, A))X}$. Now by lemma 5.3 we have that $x_n = R(\lambda, K)f_n \in \mathscr{D}(A)$, thus $x \in \mathscr{D}(\overline{A + B})$. □

Lemma 5.5. *Let X and Y be Banach spaces and $A, B \in L(X, Y)$.*
If $A \subseteq B$, $\ker B = \{0\}$ and $\operatorname{im} A = Y$, then $A = B$.

Proof. It suffices to show that $\mathscr{D}(A) = \mathscr{D}(B)$, so suppose that $x \in \mathscr{D}(B)\backslash\mathscr{D}(A)$ and let $y = Bx$. Since A is onto, there exists an $x' \in \mathscr{D}(A)$ with $y = Ax'$, and because B is an extension of A, we have that $x' \in \mathscr{D}(B)$ and also $y = Ax' = Bx'$. Now B is one-to-one and $Bx = Bx'$, so $x = x'$, but this contradicts that $x \notin \mathscr{D}(A)$. So $\mathscr{D}(B)\backslash\mathscr{D}(A) = \emptyset$, that is $\mathscr{D}(A) = \mathscr{D}(B)$. □

We are now in a position to prove the spectral criterion theorem.

Proof of theorem 5.1. (i) Let $\lambda \in \Lambda$. Then $1 \notin \sigma_p(BR(\lambda, A))$ if the operator $I - BR(\lambda, A)$ is one-to-one, that is $\ker(I - BR(\lambda, A)) = \{0\}$. Then suppose that $(I - BR(\lambda, A))x = 0$ for some $x \in X$. Now since K is an extension of $A + B$, we have that

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

thus $(\lambda I - K)R(\lambda, A)x = 0$. By assumption we have that $\lambda \in \varrho(K)$, so $(\lambda I - K)$ is one-to-one and thus $R(\lambda, A)x = 0$, that is $x \in \ker R(\lambda, A)$. But since $R(\lambda, A)$ is an invertible, hence one-to-one, operator, we obtain that $x = 0$, proving the result.

(ii) For any $\lambda \in \Lambda \subseteq \varrho(A)$ we have that

$$\lambda I - (A + B) = (I - BR(\lambda, A))(\lambda I - A),$$

and thus the invertibility of $\lambda I - (A + B)$ is equivalent to the invertibility of $I - BR(\lambda, A)$. Now if $K = A + B$, then $I - BR(\lambda, A)$ is invertible since $\lambda \in \varrho(K)$ and so $\lambda - (A + B) = \lambda I - K$ is invertible, that is $1 \in \varrho(BR(\lambda, A))$ for any $\lambda \in \Lambda$. Conversely, if $I - BR(\lambda, A)$ is invertible for some/all $\lambda \in \Lambda$, then $\lambda I - (A + B)$ is invertible and since K is an extension of $A + B$ we have that $\lambda I - (A + B) \subseteq \lambda I - K$ and so $K = A + B$ by lemma 5.5.

(iii) Let $\lambda \in \Lambda$. Then from lemma 5.4 we have that $\mathscr{D}(\overline{A + B}) = \mathscr{D}(K)$ if and only if

$$R(\lambda, K)X = \mathscr{D}(K) = R(\lambda, K)\overline{(I - BR(\lambda, A))X},$$

that is $X = \overline{(I - BR(\lambda, A))X}$ since $R(\lambda, K)$ is a bijective operator. Now from the above, we have that $1 \notin \sigma_p(BR(\lambda, A))$ and so $I - BR(\lambda, A)$ is one-to-one, therefore the statement $X = \overline{(I - BR(\lambda, A))X}$ is equivalent to $1 \in \sigma_c(BR(\lambda, A))$. Since the statement $K = \overline{A + B}$ is independent of λ, it holds for all $\lambda \in \Lambda$ if it holds for some $\lambda \in \Lambda$.

(iv) From the above, all other possibilities have been exhausted. Hence K is a proper extension of $\overline{A+B}$, that is $K \supsetneq \overline{A+B}$ if and only if $1 \in \sigma_r(BR(\lambda, A))$ as the spectrum $\sigma(\cdot)$ is the disjoint union of $\sigma_p(\cdot)$, $\sigma_c(\cdot)$ and $\sigma_r(\cdot)$. □

Corollary 5.6. *Suppose that $\Lambda = \varrho(A) \cap \varrho(K) \neq \emptyset$ and $r(BR(\lambda, A)) < 1$ for some $\lambda \in \varrho(A)$. Then $K = A + B$.*

Proof. If $r(BR(\lambda, A)) < 1$, then by proposition 2.29 the operator $I - BR(\lambda, A)$ is invertible with inverse given by the Neumann series $\sum_{n=0}^{\infty}(BR(\lambda, A))^n$. Therefore $1 \in \varrho(BR(\lambda, A))$ and so by theorem 5.1 we have that $K = A + B$ and $\mathscr{D}(K) = \mathscr{D}(A)$, as required. □

5.2. **Bounded Perturbation Theorem.** If B is a bounded operator, then $A+B$ does generate a C_0-semigroup and we have the following theorem characterising the semigroup it generates. Indeed, this is the best case scenario for a perturbing operator.

Theorem 5.7. *Suppose that $(A, \mathscr{D}(A)) \in \mathcal{G}(M, \omega)$ for some $\omega \in \mathbb{R}$ and $M \geq 1$ and $B \in \mathcal{L}(X)$. Then $K = A+B$ generates a C_0-semigroup, in fact $(A+B, \mathscr{D}(A)) \in \mathcal{G}(M, \omega + M\|B\|)$. Further, the semigroup $\{T_{A+B}(t)\}_{t\geq 0}$ generated by $A+B$ satisfies one of the following Duhamel equations,*

$$T_{A+B}(t)x = T_A(t)x + \int_0^t T_A(t-s)BT_{A+B}(s)x\,\mathrm{d}s \text{ for } t \geq 0 \text{ and } x \in X; \text{ or}$$

$$T_{A+B}(t)x = T_A(t)x + \int_0^t T_{A+B}(t-s)BT_A(s)x\,\mathrm{d}s \text{ for } t \geq 0 \text{ and } x \in X.$$

Moreover, the semigroup $\{T_{A+B}(t)\}_{t\geq 0}$ is given by the Dyson-Phillips series

$$T_{A+B}(t) = \sum_{n=0}^{\infty} T_n(t),$$

where $T_0(t) = T_A(t)$ and the T_n are defined recursively by the formula

$$T_{n+1}(t)x = \int_0^t T_A(t-s)BT_n(s)x\,\mathrm{d}s.$$

The Dyson-Phillips series converges in $\mathcal{L}(X)$ and converges uniformly for t on bounded intervals.

To prove that $(A + B, \mathscr{D}(A)) \in \mathcal{G}(M, \omega + M\|B\|)$, we need the following *re-norming* lemma.

Lemma 5.8. *Let A be a linear operator for which $(0, \infty) \subseteq \varrho(A)$. If for some $M \geq 0$ we have*

$$\|\lambda^n R(\lambda, A)^n\| \leq M \text{ for all } n \geq 1 \text{ and } \lambda > 0.$$

then there exists a norm $|\cdot|$ on X which is equivalent to $\|\cdot\|$ with $\|x\| \leq |x| \leq M\|x\|$ and

$$|\lambda R(\lambda, A)x| \leq |x| \text{ for all } x \in X \text{ and } \lambda > 0.$$

Proof. Let $\mu > 0$, then for any $x \in X$ we define the norm

$$\|x\|_\mu = \sup_{n\geq 0} \|\mu^n R(\mu, A)^n x\|.$$

Then clearly $\|x\| \leq \|x\|_\mu \leq M\|x\|$ for all $x \in X$ and also $\|\mu R(\mu, A)\|_\mu \leq 1$ since

$$\|\mu R(\mu, A)x\|_\mu = \sup_{n\geq 0} \|\mu^{n+1} R(\mu, A)^{n+1} x\| \leq \sup_{n\geq 0} \|\mu^n R(\mu, A)^n x\| = \|x\|_\mu. \tag{5.4}$$

Now define the norm $|\cdot|$ by

$$|x| = \lim_{\mu\to\infty} \|x\|_\mu, \quad x \in X.$$

It is clear that this norm is homogeneous and satisfies the triangle inequality, also if $x = 0$ then $\|x\|_\mu = 0$ for any $\mu > 0$ and so $|x| = 0$. On the other hand if $|x| = 0$, then $\|x\|_\mu \to 0$ as $\mu \to \infty$. Now if $\|x\|_\lambda$ is increasing in λ, then since any norm is non-negative we have that $\|x\|_\mu = 0$ for any $\mu \in (0, \infty)$, and so $x = 0$. So it suffices to prove that $\lambda \mapsto \|x\|_\lambda$ for any $x \in X$.

Now fix $\mu > 0$, then we claim that $\|\lambda R(\lambda, A)\|_\mu \le 1$ for any $0 < \lambda \le \mu$.

For this, let $x \in X$ and set $y = R(\lambda, A)x$. Then $y = R(\mu, A)(x + (\mu - \lambda)y)$ and therefore

$$\|y\|_\mu = \|R(\mu, A)x + (\mu - \lambda)R(\mu, A)y\|_\mu \le \frac{1}{\mu}\|x\|_\mu + \frac{\mu - \lambda}{\mu}\|y\|_\mu,$$

using that $\|\mu R(\mu, A)\|_\mu \le 1$ from the above. Hence $\lambda\|y\|_\mu \le \|x\|_\mu$, as claimed and so

$$\|\lambda^n R(\lambda, A)^n x\| \le \|\lambda^n R(\lambda, A)^n x\|_\mu \le \|x\|_\mu$$

for any λ with $0 < \lambda \le \mu$. Thus taking the supremum over $n \ge 0$ we obtain that $\|x\|_\lambda \le \|x\|_\mu$. Then as $\|x\| \le \|x\|_\mu \le M\|x\|$ for all $\mu > 0$, taking the limit as $\mu \to \infty$ we see that $\|x\| \le |x| \le M\|x\|$ and also since $\|\lambda^n R(\lambda, A)^n x\|_\mu \le \|x\|_\mu$, taking $n = 1$ and then $\mu \to \infty$ we obtain that

$$|\lambda R(\lambda, A)x| \le |x|,$$

as required. □

We are now in a position to prove the first part of the bounded perturbation theorem. Now if $(A, \mathscr{D}(A)) \in \mathcal{G}(M, \omega)$, then A generates a semigroup $\{T(t)\}_{t\ge0}$ with $\|T(t)\| \le Me^{\omega t}$. Then by the Hille-Yosida theorem we have that A is a closed and densely defined operator, $(\omega, \infty) \subseteq \varrho(A)$ and also for any $n \ge 1$ and $\lambda > \omega$, $\|R(\lambda, A)^n\| \le M(\lambda - \omega)^{-n}$. Now if $S(t) := e^{-\omega t}T(t)$, then by lemma 4.11 this defines a semigroup with generator $A' = A - \omega I$. Since $(\omega, \infty) \subseteq \varrho(A)$, we have that $(0, \infty) \subseteq \varrho(A')$ and for any $n \ge 1$ and $\lambda > 0$, it follows that

$$\|R(\lambda, A')^n\| = \|R(\lambda + \omega, A)^n\| \le \frac{M}{\lambda^n},$$

that is $\|\lambda^n R(\lambda, A')^n\| \le M$. Then by lemma 5.8 there exists a norm $|\cdot|$ on X with

$$\|x\| \le |x| \le M\|x\| \text{ and } |\lambda R(\lambda, A')x| \le |x| \text{ for all } x \in X \text{ and } \lambda > 0.$$

Therefore if $\lambda > \omega$, we have that $|(\lambda-\omega)R(\lambda-\omega, A')x| = |(\lambda-\omega)R(\lambda, A)x| \le |x|$ for any $x \in X$.

The above remarks are used in the following proof.

Proof of theorem 5.7. Suppose that $A \in \mathcal{G}(M, \omega)$ generates the semigroup $\{T(t)\}_{t\ge0}$ and that $B \in \mathcal{L}(X)$ is a bounded operator on X. Then there exists a norm $|\cdot|$ on X such that $\|x\| \le |x| \le M\|x\|$ for all $x \in X$ and $|R(\lambda, A)| \le (\lambda - \omega)^{-1}$ for all real $\lambda > \omega$. Therefore for $\lambda > \omega + |B|$, we have that $|BR(\lambda, A)| \le |B||R(\lambda, A)| < 1$ and so $I - BR(\lambda, A)$ is an invertible operator for such λ. Now define the operator

$$R := R(\lambda, A)(I - BR(\lambda, A))^{-1} = \sum_{k=0}^{\infty} R(\lambda, A)(BR(\lambda, A))^k.$$

Then $(\lambda I - A - B)R = (I - BR(\lambda, A))^{-1} - BR(\lambda, A)(I - BR(\lambda, A))^{-1} = I$ and for $x \in \mathscr{D}(A)$,

$$R(\lambda I - A - B)x = R(\lambda, A)(\lambda I - A - B)x + \sum_{k=1}^{\infty} R(\lambda, A)(BR(\lambda, A))^k(\lambda I - A - B)x$$
$$= x - R(\lambda, A)Bx + \sum_{k=1}^{\infty}(R(\lambda, A)B)^k x - \sum_{k=2}^{\infty}(R(\lambda, A)B)^k x = x,$$

and since $\mathscr{D}(A)$ is dense in X by theorem 4.10, we obtain that $R = (\lambda I - A - B)^{-1}$, and so the resolvent of $A + B$ exists for any $\lambda > \omega + |B|$. Furthermore, we have that

$$|(\lambda I - A - B)^{-1}| \leq \sum_{k=0}^{\infty} |R(\lambda, A)||BR(\lambda, A)|^k \leq \frac{(1 - |BR(\lambda, A)|)^{-1}}{\lambda - \omega}.$$

Now $|BR(\lambda, A)| \leq |B|(\lambda - \omega)^{-1}$, hence $1 - |BR(\lambda, A)| \geq (\lambda - \omega - |B|)(\lambda - \omega)^{-1}$ and so

$$|(\lambda I - A - B)^{-1}| \leq \frac{1}{\lambda - \omega - |B|}.$$

Now the operator $A + B$ is densely defined as $\mathscr{D}(A + B) = \mathscr{D}(A) \cap \mathscr{D}(B) = \mathscr{D}(A)$ as $\mathscr{D}(B) = X$ and also closed, since A is closed by theorem 4.10 and B is closed by the closed graph theorem. Therefore by the Hille-Yosida theorem, $A + B$ is the infinitesimal generator of a C_0-semigroup $\{S(t)\}_{t\geq 0}$ which satisfies $|S(t)| \leq e^{(\omega + |B|)t}$ for $t \geq 0$. Then returning to the original norm $\|\cdot\|$ on X we have that $\|S(t)\| \leq Me^{(\omega + M\|B\|)t}$ as $\|x\| \leq M|x|$ and $|B| \leq M\|B\|$, as required.

For a proof of the Duhamel equations and Dyson-Phillips series, see [3, Chapter 3]. □

Remark 5.9. In the above proof we have used that $A + B$ is closed where $(A, \mathscr{D}(A))$ is a closed operator and $B \in \mathcal{L}(X)$. To see why, consider $(x_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}(A + B) = \mathscr{D}(A)$ with $x_n \to x$ and $(A + B)x_n \to y$ as $n \to \infty$. Now since B is bounded, it is continuous and so $Bx_n \to Bx$. Hence we have that $\|Ax_n - (y - Bx)\| \leq \|(A + B)x_n - y\| + \|Bx_n - Bx\|$ and thus $Ax_n \to y - Bx$ as $n \to \infty$. Then because A is closed, we have that $x \in \mathscr{D}(A) = \mathscr{D}(A + B)$ and $Ax = y - Bx$, that is $(A + B)x = y$, so $A + B$ is indeed a closed operator.

5.3. **Perturbations of Dissipative Operators.** As previously discussed, any generator of a C_0-semigroup is closed and densely defined. Moreover, contraction semigroups have dissipative infinitesimal generators. We now consider the following perturbation theorem.

Theorem 5.10. *Let A and B be linear operators in X such that B is A-bounded and $A + tB$ dissipative for all $t \in [0, 1]$. If $(A + t_0B, \mathscr{D}(A))$ generates a semigroup of contractions for some $t_0 \in [0, 1]$, then $A + tB$ generates a semigroup of contractions for every $t \in [0, 1]$.*

Proof. See [1, Theorem 4.11]. The proof relies on the Lumer-Phillips generation theorem. □

5.4. **Miyadera Perturbations.** We now consider another particular type of perturbation. Let $(A, \mathscr{D}(A))$ be the linear operator in X generating the semigroup $\{T_A(t)\}_{t\geq 0}$, and as before we consider B as the perturbing operator and $K \supseteq A + B$ an operator generating a C_0-semigroup.

Definition 5.11. The operator $(B, \mathscr{D}(B))$ is a *Miyadera perturbation* of A if B is A-bounded and there exist constant $\alpha \in (0, \infty)$ and $\gamma \in [0, 1)$ such that for all $x \in \mathscr{D}(A)$,

$$\int_0^{\alpha} \|BT_A(t)x\| \, \mathrm{d}t \leq \gamma\|x\|. \tag{5.5}$$

This definition is for a semigroup $\{T_A(t)\}_{t\geq 0}$ of arbitrary type, namely the value

$$\omega_0 = \inf\{\omega \in \mathbb{R} : \exists M \geq 0 \text{ with } \|T_A(t)\| \leq Me^{\omega t} \text{ for all } t \geq 0\}.$$

Now below we will prove that $A + B$ generates a C_0-semigroup if B is a Miyadera perturbation of A and A generates a C_0-semigroup, and the proof is substantial easier if the semigroup A has negative type, that is $\omega_0 < 0$. Recall by lemma 4.11 that A is the generator of $\{T_A(t)\}_{t\geq 0}$, with type ω_0, if and only if $A - \omega I$ is the generator of $\{e^{-\omega t}T_A(t)\}_{t\geq 0}$, with type $\omega_0 - \omega$. Thus provided that $\omega > \omega_0$, the latter semigroup has negative type. The following lemma allows us to assume that $\{T_A(t)\}_{t\geq 0}$ has negative type in theorem 5.13.

Lemma 5.12. *The operator B is a Miyadera perturbation of A if and only if it is a Miyadera perturbation of $A - \lambda I$ for any $\lambda \in \mathbb{R}$, but possibly with different values of α and γ.*

Proof. It suffices to prove the lemma for $\lambda > 0$. Indeed, if the result is proven for any $\lambda > 0$, then if B is a Miyadera perturbation of A, we have that B is a Miyadera perturbation of $A \pm \lambda I$ for $\lambda > 0$. So addition of λI to A for any $\lambda \in \mathbb{R}$ does not affect the Miyadera perturbation.

Now suppose that B is a Miyadera perturbation of $A - \lambda I$ for some $\lambda > 0$, then

$$\int_0^\alpha \|e^{-\lambda t} BT_A(t)x\| \,\mathrm{d}t \leq \gamma \|x\|$$

for some $\alpha \in (0, \infty)$ and $\gamma \in [0, 1)$ and all $x \in \mathscr{D}(A)$. If $\gamma' = e^{\lambda\alpha} < 1$, then we have that

$$\int_0^\alpha \|BT_A(t)x\| \,\mathrm{d}t \leq e^{\lambda\alpha} \int_0^\alpha e^{-\lambda t}\|BT_A(t)x\| \,\mathrm{d}t \leq e^{\lambda\alpha}\gamma\|x\| = \gamma'\|x\|,$$

and so B is a Miyadera perturbation of A. On the other hand, if $e^{\lambda\alpha}\gamma \geq 1$ then we can find a $\beta \in (0, \alpha)$ such that $\gamma' = e^{\lambda\beta}\gamma < 1$ as $\gamma < 1$. In this case, we see that for any $x \in \mathscr{D}(A)$,

$$\int_0^\beta \|BT_A(t)x\| \,\mathrm{d}t \leq e^{\lambda\beta} \int_0^\alpha e^{-\lambda t}\|BT_A(t)x\| \,\mathrm{d}t \leq e^{\lambda\beta}\gamma\|x\| = \gamma'\|x\|.$$

Conversely, suppose that B is a Miyadera perturbation of A. Then for any $\lambda > 0$, we have

$$\int_0^\alpha e^{-\lambda t}\|BT_A(t)x\| \,\mathrm{d}t \leq \int_0^\alpha \|BT_A(t)x\| \,\mathrm{d}t \leq \gamma\|x\|$$

since $e^a \leq 1$ for $a < 0$, and so B is a Miyadera perturbation of $A - \lambda I$, as required. ☐

Theorem 5.13. *Suppose that B is a Miyadera perturbation of A and $A \in \mathcal{G}(M, \omega')$.*
Then $(A + B, \mathscr{D}(A))$ is the generator of a C_0-semigroup, denoted by $\{T(t)\}_{t\geq 0}$.

Sketch of the proof. If $\omega' \geq 0$, then we consider $A + B = (A - \lambda I) + (B + \lambda I)$ for some $\lambda > \omega'$. Then $A - \lambda I$ has negative type, and by lemma 5.12 we have that B is a Miyadera perturbation of $A - \lambda I$ and $A + B$ is the generator of a semigroup if and only if $A - \lambda I + B$ is by lemma 4.11. So we can assume without loss of generality that $\{T_A(t)\}_{t\geq 0}$, the semigroup generated by A, as negative type $\omega' < 0$. So in particular, $\|T_A(t)\| \leq Me^{-\omega t}$ for all $t \geq 0$ where $\omega = -\omega' > 0$.

Now for $x \in \mathscr{D}(A)$ we define the operator $U_1(t)$ for $t \geq 0$ by

$$U_1(t)x = \int_0^t T_A(t - s)BT_A(s)x \,\mathrm{d}s$$

Then with $\alpha \in (0,\infty)$ and $\gamma \in [0,1)$ from the definition of B being a Miyadera perturbation of A, we consider t with $n\alpha < t \leq (n+1)\alpha$ for some $n \in \mathbb{N}$ and notice that, for some constant c,

$$\|U_1(t)x\| \leq M \sum_{j=0}^{n} \int_0^\alpha \|BT_A(r)T_A(j\alpha)x\| \,\mathrm{d}r \leq \gamma M \|x\| \sum_{j=0}^{\infty} \|T_A(j\alpha)\| \leq \frac{\gamma M^2 \|x\|}{1-e^{-\omega\alpha}} =: c\|x\|,$$

and as $\mathscr{D}(A)$ is dense in X, $U_1(t)$ extends to a bounded linear operator on X where $\{\|U_1(t)\|\}_{t\geq 0}$ is bounded. Now for $j \in \mathbb{N}$ with $j \geq 2$, we define $\{U_j(t)\}_{t\geq 0}$ inductively by

$$U_j(t)x = \int_0^t U_{j-1}(t-s)BT_A(s)x \,\mathrm{d}s, \quad x \in \mathscr{D}(A).$$

Similarly to the above estimate for $\|U_1(t)x\|$, for each $j \geq 0$, $\{U_j(t)\}_{t\geq 0}$ is a family of bounded operators with norms that are uniformly bounded in t. Now it can be shown by induction that

$$\sum_{j=0}^{n} U_j(t)U_{n-j}(s) = U_n(t+s) \text{ for } n \in \mathbb{N}_0 \text{ and } t,s \geq 0, \tag{5.6}$$

where $U_0(t) = T_A(t)$. Since B is a Miyadera perturbation of A, it can be shown that $t \mapsto U_j(t)x$ is continuous at $t = 0$ for each $x \in \mathscr{D}(A)$. Further, $\{U_j(t)\}_{t\geq 0}$ are strongly continuous at $t = 0$. Now for $t \in [0,\alpha]$ and $x \in \mathscr{D}(A)$, we have that $\|U_1(t)x\| \leq M\gamma\|x\|$, so by density this holds for all $x \in X$ and therefore by induction $\|U_j(t)\| \leq M\gamma^j$. This implies that $\sum_{j=0}^\infty U_j(t)$ converges in $\mathcal{L}(X)$ uniformly on $[0,\alpha]$, and using (5.6) we obtain that $\sum_{n=0}^\infty U_n(2t)$ converges uniformly in $\mathcal{L}(X)$ on $[0,\alpha]$, that is $\sum_{j=0}^\infty U_j(t)$ converges uniformly on $[0,2\alpha]$. Iterating we obtain almost uniform convergence on $[0,\infty)$. Now define $T(t) = \sum_{j=0}^\infty U_j(t)$ for $t \in [0,\infty)$, then $\{T(t)\}_{t\geq 0}$ is a C_0-semigroup on X which satisfies the Duhamel equation

$$T(t)x = T_A(t)x + \int_0^t T(t-s)BT_A(s)x \,\mathrm{d}s, \quad x \in \mathscr{D}(A).$$

Now it can be shown that $\int_0^t T(t-s)BT_A(s)x \,\mathrm{d}s$ is differentiable at $t = 0$ for $x \in \mathscr{D}(A)$ with derivative Bx, and by properties of the semigroup, $T_A(t)x$ is differentiable for $x \in \mathscr{D}(A)$, hence from the Duhamel equation we obtain that $Kx = Ax + Bx$ for $x \in \mathscr{D}(A)$, that is $K \supseteq A + B$. Now $\varrho(A) \cap \varrho(K) \neq \emptyset$ as A and K are both generators, so by theorem 5.1, for $K = A + B$ to hold we need that $I - BR(\lambda, A)$ is invertible for some $\lambda \in \Lambda$. Using (4.9) we can deduce that

$$BR(\lambda, A)x = \int_0^\infty e^{-\lambda t} BT_A(t)x \,\mathrm{d}t$$

for any $\lambda > 0$. Now for sufficiently large λ, we have that $\|BR(\lambda, A)\| < 1$ so that $I - BR(\lambda, A)$ is invertible and so $1 \in \varrho(BR(\lambda, A))$, hence $K = A + B$ and $\mathscr{D}(A) = \mathscr{D}(K)$ by theorem 5.1. □

5.5. **Positive Perturbations and Kato's Theorem.** We state, without proof, the following theorem which relates to perturbing a positive semigroup of contractions.

Theorem 5.14. *Let X be a KB-space and suppose $(A, \mathscr{D}(A))$ and $(B, \mathscr{D}(B))$ satisfy*

(i) A generates a positive semigroup of contractions $\{T_A(t)\}_{t\geq 0}$,
(ii) $r(BR(\lambda, A)) \leq 1$ for some $\lambda > 0$,
(iii) $Bx \succeq 0$ for all $x \in \mathscr{D}(A)_+$, and
(iv) $\langle x^*, (A+B)x\rangle \leq 0$ for any $x \in \mathscr{D}(A)_+$ where $\langle x^*, x\rangle = \|x\|$ and $x^* \succeq 0$.

Then there exists an extension $(K, \mathscr{D}(K))$ *of* $(A+B, \mathscr{D}(A))$ *which generates a* C_0*-semigroup of contractions, denoted* $\{T_K(t)\}_{t\geq 0}$*. The generator* K *satisfies the following equation for* $\lambda > 0$*,*

$$R(\lambda, K)x = \lim_{n\to\infty} R(\lambda, A)\sum_{k=0}^{n}(BR(\lambda, A))^n x = \sum_{k=0}^{\infty} R(\lambda, A)(BR(\lambda, A))^k x.$$

Moreover, the semigroup $\{T_K(t)\}_{t\geq 0}$ *satisfies the Duhamel equation, namely*

$$T_K(t)x = T_A(t)x + \int_0^T T_K(t-s)BT_A(s)x \,\mathrm{d}s, \quad x \in \mathscr{D}(A).$$

Proof. For a proof of this result, see [1, Theorem 5.2 and Corollary 5.8]. □

6. Application to Birth-and-Death Processes

Example 6.1. In this example, we consider a Markov *birth-and-death* process which describes the evolution of a population whose size k at any time t is allowed to increase or decrease by one, as a result of a birth or death of an individual in the population. Now for any time interval Δt, the probability that a birth or death occurs is $b_k\Delta t + o(\Delta t)$ and $d_k\Delta t + o(\Delta t)$, respectively, where b_k and d_k are the instantaneous birth and death rates when the population is of size k.

We denote the probability that the population has size k at time $t \geq 0$ by $u_k(t)$.

Then the corresponding *forward Kolmogorov system* is given by

$$\begin{aligned} u_0' &= -b_0u_0 + d_1u_1, \\ u_n' &= -(b_n + d_n)u_n + d_{n+1}u_{n+1} + b_{n-1}u_{n-1}, \quad n \geq 1. \end{aligned} \tag{6.1}$$

We assume that $b_n, d_n > 0$ for all indices, apart from setting $b_{-1} = d_0 = 0$ for convenience.

The natural setting for such problems is in the Banach space $X = \ell^1$, that is the space

$$\ell^1 = \{x = (x_n)_{n\in\mathbb{N}_0} : \|x\| = \sum_{n=0}^{\infty} |x_n| < \infty\}.$$

We make this choice of space as $u_k \geq 0$ and $\|u\| = \sum_{k=0}^{\infty} u_k = 1$, since the u_k are probabilities. Therefore for this problem, the value of $\|u\|$ should be preserved. Now it is convenient to write the system (6.1) as the sum of two operators $\mathcal{A}$ and $\mathcal{B}$ which are defined for any $u = (u_n)_{n\in\mathbb{N}_0}$ by $(\mathcal{A}u)_n = -(b_n + d_n)u_n$ and $(\mathcal{B}u)_n = d_{n+1}u_{n+1} + b_{n-1}u_{n-1}$. Doing so we obtain that

$$u' = \mathcal{A}u + \mathcal{B}u, \tag{6.2}$$

where $u = (u_n)_{n\in\mathbb{N}_0}$, as an equivalent system to (6.1). Probabilistically we should have that

$$u_n(t) \geq 0 \text{ and } \sum_{k=0}^{\infty} u_k(t) = \sum_{k=0}^{\infty} u_k(0) = 1$$

for all $t \geq 0$ and $n \in \mathbb{N}_0$. The equation (6.2) is now more familiar, considering the perturbation theory of the previous chapter. It turns out that, in the natural ℓ^1 setting, a restriction of $\mathcal{A}$ to a suitable domain generates a contraction semigroup.

In what now follows, we consider the more general setting where $X = \ell^p$ for $p \in [1, \infty)$.

6.1. Existence Results.

Let $u = (u_n)_{n=0}^{\infty}$ where u_n denotes the number of objects in state n. For a probabilistic interpretation, u_n is the probability of observing an object in the state n and so we should have $\|u\|_{\ell^1} = 1$, and we allow any object in the system to change its state by ± 1. Now denote by d_n and b_n the given rates of change for $n \mapsto n-1$ and $n \mapsto n+1$, respectively. In general, we also include a mechanism which allows for the change in the number of objects at the state n (for example, removing objects from the system or introducing them), and we denote the rate of such mechanism by $(c_n)_{n=0}^{\infty}$. Standard modelling yields the following system,

$$\begin{aligned} u_0' &= -a_0u_0 + d_1u_1, \\ u_n' &= -a_nu_n + d_{n+1}u_{n+1} + b_{n-1}u_{n-1}, \quad n \geq 1, \end{aligned} \tag{6.3}$$

where $c_n = b_n + d_n - a_n$. As in example 6.1, when we are discussing a birth-and-death population system, we have that $c_n = 0$, and such a system is called *conservative* as $a_n = b_n + d_n$.

Now we assume that d, b and a are non-negative sequences with $b_{-1} = d_0$.

Define the operator $\mathcal{K}$ by $(\mathcal{K}u)_n = b_{n-1}u_{n-1} - a_n u_n + d_{n+1}u_{n+1}$ and define $\mathcal{A}$ and $\mathcal{B}$ as in example 6.1, that is $(\mathcal{A}u)_n = -a_n u_n$ and $(\mathcal{B}u)_n = b_{n-1}u_{n-1}+d_{n+1}u_{n+1}$ for $n \geq 1$. Then defining the operators in this way, we can think of the system (6.3) as one of the form $u' = \mathcal{A}u + \mathcal{B}u$.

Definition 6.2. We denote by $\mathcal{K}_p$ the *maximal realisation of* $\mathcal{K} \in \ell^p$ for $p \in [1,\infty)$.

That is, $\mathcal{K}_p u = \mathcal{K}u$ with domain $\mathscr{D}(\mathcal{K}_p) = \{u \in \ell^p : \mathcal{K}u \in \ell^p\}$.

Lemma 6.3. *For any $p \in [1,\infty)$, the operator $\mathcal{K}_p$ is closed.*

Proof. Suppose that $u^{(n)} \to u$ and $\mathcal{K}_p u^{(n)} \to v$ in ℓ^p as $n \to \infty$ for some u, v. Then $u_k^{(n)} \to u_k$ for any $k \geq 0$ and therefore for any k, $(\mathcal{K}_p u^{(n)})_k = b_{k-1}u_{k-1}^{(n)} - a_k u_k^{(n)} + d_{k+1}u_{k+1}^{(n)}$ converges to $b_{k-1}u_{k-1} - a_k u_k + d_{k+1}u_{k+1}$, but this is just $(\mathcal{K}_p u)_k$ as $u \in \ell^p$. Then by uniqueness of limits, we obtain that $\mathcal{K}_p u = v$ and so $\mathcal{K}_p$ is closed, as required. □

Definition 6.4. The operator A_p is defined to be the restriction of $\mathcal{A}$ to the domain

$$\mathscr{D}(A_p) = \{u \in \ell^p : \mathcal{A}u \in \ell^p\}.$$

Then considering the definition of $\mathcal{A}$, the condition $\mathcal{A}u \in \ell^p$ is equivalent to $\sum_{n=0}^{\infty} a_n^p |u_n|^p < \infty$.

Lemma 6.5. *The operator $(A_p, \mathscr{D}(A_p))$ generates a semigroup of contractions in ℓ^p.*

Proof. The idea is to use the Hille-Yosida theorem, in particular remark 4.15. Now by a similar proof to that of lemma 6.3, the operator A_p is closed and it is also densely defined. Now for any $y = (y_n)_{n\in\mathbb{N}_0}$, we solve $R(\lambda I - A_p)y = y = (\lambda I - A_p)Ry$ in order to find the resolvent $R(\lambda, A_p)$. Now $(\lambda I - A_p)y = ((\lambda + a_n)y_n)_{n\in\mathbb{N}_0}$ and thus we conjecture that $(Ry)_n = y_n(\lambda + a_n)^{-1}$. Since $a_n \geq 0$, we consider $\lambda > 0$ so that this is well-defined. Now with this, we see that

$$R(\lambda I - A_p) = I = (\lambda I - A_p)R,$$

and thus for any $\lambda > 0$ the resolvent $R(\lambda, A_p)$ exists and so $(0,\infty) \subseteq \varrho(A)$.

Therefore since $A_p u = \mathcal{A}u = (-a_n u_n)_{n\in\mathbb{N}_0}$ and $a_n \geq 0$, we have that

$$\|A_p R(\lambda, A_p)y\|^p = \sum_{n=0}^{\infty} |a_n (R(\lambda, A_p)y)_n|^p = \sum_{n=0}^{\infty} a_n^p \frac{|y_n|^p}{(\lambda + a_n)^p}$$

Now since the function $x \mapsto x^p$ is increasing and well-defined for $x \geq 0$, we obtain that

$$\|A_p R(\lambda, A_p)y\|^p \leq \sum_{n=0}^{\infty} |y_n|^p = \|y\|^p,$$

and thus $\|A_p R(\lambda, A_p)\| \leq 1$. Also, we have that

$$\|R(\lambda, A_p)y\|^p = \sum_{n=0}^{\infty} \frac{1}{(\lambda + a_n)^p} |y_n|^p \leq \frac{1}{\lambda^p} \|y\|^p,$$

and so $\|R(\lambda, A_p)\| \leq 1/\lambda$ for all $\lambda > 0$, as required. □

Theorem 6.6. *Suppose that $b = (b_n)_{n\in\mathbb{N}_0}$ and $d = (d_n)_{n\in\mathbb{N}_0}$ are non-decreasing and there exists an $\alpha \in [0,1]$ such that for all n, we have that*

$$0 \leq b_n \leq \alpha a_n \text{ and } 0 \leq d_{n+1} \leq (1-\alpha)a_n.$$

Then there exists an extension K_p of $(A_p + B_p, \mathscr{D}(A_p))$ which generates a positive semigroups of contractions in ℓ^p with $p \in (1,\infty)$ where B_p is the restriction of $\mathcal{B}$ to $\mathscr{D}(A_p)$.

Proof. This uses theorem 5.14; for a proof, see [1, Theorem 7.3]. □

We now consider again the natural setting of $X = \ell^1$, so with $p = 1$.

Corollary 6.7. *Suppose that* $b = (b_n)_{n\in\mathbb{N}_0}$ *and* $d = (d_n)_{n\in\mathbb{N}_0}$ *are non-negative and*

$$a_n \geq (b_n + d_n). \tag{6.4}$$

Then there exists an extension K_1 *of* $(A_1 + B_1, \mathscr{D}(A_1))$ *which generates a positive semigroup of contractions in* ℓ^1 *where* B_1 *is the restriction of* $\mathcal{B}$ *to* $\mathscr{D}(A_1)$*.*

Proof. For a proof of this result, see [1, Corollary 7.5]. □

List of Figures

References

[1] *Perturbations of Positive Semigroups with Applications*, Jacek Banasiak and Luisa Arlotti, Springer, 2006.
[2] *Functional Analysis*, Walter Rudin, 2nd Edition, 1991.
[3] *Semigroups of Linear Operators and Applications to Partial Differential Equations*, A. Pazy, Springer, 1983.
[4] *Perturbation Theory for Linear Operators*, Tosio Kato, Spring, 1980.