Nonlinear expectations and a semigroup approach to fully nonlinear PDEs

Dissertation zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Max Nendel

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Mathematik und Statistik

Konstanz, 2017

Tag der mündlichen Prüfung: 20. Dezember 2017

1. Referent: Prof. Dr. Robert Denk
2. Referent: Prof. Dr. Michael Kupper

Für meine Mutter Susanne.

“Amor es la distancia más corta entre dos humanos.”

Danksagung

Zunächst möchte ich mich bei meiner Mutter Susanne Nendel bedanken. Ohne dich wäre ich nicht, wo ich heute bin. Du bist ein wahres Vorbild und man könnte sich keine bessere Mutter wünschen. Weiterhin möchte ich meinen Pateneltern Martina Fritz und Wolfgang Messerschmidt für die vielen guten Ratschläge und Werte, die sie mir mitgegeben haben, danken. Auch ihr habt mein Leben maßlos beeinflusst und ich danke euch für eure liebevolle Unterstützung. Ebenso möchte ich gegenüber Herrn Luigi Blason und Frau Hannelore Blason meinen tiefsten Dank ausdrücken. Ich denke stets mit Freude an meine schöne Kindheit in Ihrem Haus als Teil Ihrer Familie zurück und danke Ihnen für die Werte, die Sie mir mitgegeben haben.

Mein tiefster Dank gilt natürlich auch meinen beiden Betreuern Prof. Dr. Robert Denk und Prof. Dr. Michael Kupper. Eure Leidenschaft zur Mathematik hat mich sehr inspiriert und ich danke euch für eure intensive und stets sehr persönliche Betreuung. Robert, vielen Dank für deine zahlreichen Ratschläge, sowohl mathematischer als auch privater Natur, für all die schönen Dinge, die ich dank deiner Unterstützung erleben durfte und für deinen herzlichen Umgang. Michael, auch dir möchte ich für unzählige schöne Momente, deine vielen guten Ratschläge sowie deine bedingungslose und aufopfernde Art der Betreuung danken.

Zudem bin ich Prof. Dr. Reinhard Racke für die vielen angenehmen Gespräche, die vielen guten Ratschläge und das Teilen vieler Erfahrungen sehr dankbar. Meinen Freunden und (ehemaligen) Arbeitskollegen Sarah Kroschewski, Felix Kammerlander, Felix Hummel, Daniel Bartl, Stefan Banholzer, David Palosch, Karsten Herth, Lisa Fischer, Dr. Asgar Jamneshan, Dr. Markus Kunze, Dr. Johannes Schnur, Dr. Tim Seger, Dr. Martin Saal, Dr. Alexander Schöwe, Dr. Marco Ritter, Dr. Karin Borgmeyer, Dr. Michael Pokojovy, Dr. Tobias Nau, Dr. Mario Kaip, Gerda Baumann und Brigitte Wehle danke ich für die schöne Zeit, ihre Unterstützung sowie das stets angenehme und humorvolle Arbeitsklima. Mein tiefster Dank gilt auch meinem Freund und (fast immer) Büropartner Felix Kammerlander für das Korrekturlesen dieser Arbeit und die vielen gemeinsam verbrachten Stunden. Des Weiteren möchte ich mich bei Herrn Joachim Hahn, Dr. Bienvenido Barraza Martínez, Dr. Jairo Hernández Monzón und allen Mitarbeitern des Fachbereichs für Mathematik und Statistik der Universidad del Norte für ihre Unterstützung und die anregenden Gespräche während meines fünfmonatigen Aufenthalts in Barranquilla bedanken. Außerdem danke ich Frau Yureicy Bolaño Bonett von ganzem Herzen für ihren liebevollen Umgang, ihre motivierende und inspirierende Art sowie die schöne gemeinsam verbrachte Zeit.

Abschließend möchte ich sagen, dass ich auf meine Promotion als die bisher schönste Zeit meines Lebens zurückblicke und das habe ich allen Menschen zu verdanken, mit denen ich innerhalb und außerhalb der Universität Konstanz Kontakt hatte. Bitte seiet euch meines tiefsten Dankes sicher und verzeiht mir, wenn ich euch an dieser Stelle nicht alle erwähnen kann.

Deutsche Zusammenfassung

In vielen Bereichen, wie zum Beispiel Physik, Biologie, Medizin oder Wirtschaftswissenschaften, beschreiben partielle Differentialgleichungen die Veränderung dynamischer Systeme in Ort und Zeit. Zum Beispiel werden die Bewegung einer Welle, die Ausbreitung von Wärme, das Schwingen einer Saite, die Konzentration von gewissen Substanzen in der Blutlaufbahn oder Preise in Finanzmärkten, wie beim Bachelier Modell oder beim Black-Scholes Modell, durch partielle Differentialgleichungen beschrieben. Im Gegensatz zu gewöhnlichen Differentialgleichungen ist es selbst bei autonomen linearen partiellen Differentialgleichungen nur in den seltensten Fällen möglich, die Lösung explizit anzugeben. Daher bedient man sich häufig abstrakter Methoden, um die Existenz und Eindeutigkeit einer Lösung zu beweisen. Dabei spielen, insbesondere im Zusammenhang mit autonomen linearen partiellen Differentialgleichungen, Halbgruppen eine wichtige Rolle. Die wesentliche Idee dieser Theorie ist es, die Orts- und Zeitvariablen getrennt zu betrachten und die partielle Differentialgleichung als eine Banachraum-wertige gewöhnliche Differentialgleichung in der Zeit aufzufassen. Dies führt zu einem abstrakten Cauchy-Problem der Form

$$u'(t) = Au(t), \quad t \geq 0, \tag{0.1}$$

$$u(0) = u_0. \tag{0.2}$$

Hierbei ist A in den Anwendungen zumeist ein linearer Differentialoperator beziehungsweise ein Integro-Differentialoperator in den Ortsvariablen und u_0 ein Element eines entsprechenden Funktionenraumes. In der vorliegenden Arbeit wird der auftretende Funktionenraum meist der Raum $\mathrm{BUC}(G)$ aller beschränkten, gleichmäßig stetigen Funktionen $G \to \mathbb{R}$ auf einem metrischen Grundraum G sein. Das Ziel der Halbgruppentheorie ist es, für eine derartige Differentialgleichung die Existenz einer eindeutigen Lösung nachzuweisen. Die Lösung soll zusätzlich stetig von den Anfangsdaten abhängen. In diesem Fall spricht man davon, dass das abstrakte Cauchy-Problem (0.1) - (0.2) wohlgestellt ist. Ein Hauptresultat der Halbgruppentheorie ist, dass das Cauchy-Problem (0.1) - (0.2) genau dann wohlgestellt ist, wenn A der Generator einer stark stetigen Halbgruppe $(S(t))_{t\geq 0}$ ist. Hierbei kann $S(t)$ für $t \geq 0$ als eine abstrakte Variante des Exponentials e^{tA} gesehen weden. Die eindeutige Lösung u des obigen Anfangswertproblems ist dann durch $u(t) := S(t)u_0$ für alle $t \geq 0$ gegeben. Des Weiteren werden die Generatoren von stark stetigen Halbgruppen durch den Satz von Hille-Yosida genau charakterisiert. Für eine detaillierte Diskussion über Operatorhalbgruppen verweisen wir an dieser Stelle auf Pazy [63] oder Engel-Nagel [34], [35].

In der vorliegenden Arbeit betrachten wir insbesondere Differentialgleichungen, deren Lösungen durch Preisprozesse in einem Finanzmarkt gegeben sind. Dies führt zu gewissen Bedingungen an den Generator A, wie zum Beispiel, dass A ein positives Maximumprinzip erfüllt und alle konstanten Funktionen im Kern von A liegen. Ein Beispiel für einen derartigen Operator im Fall

$G = \mathbb{R}$ ist $A := \frac{\sigma^2}{2}\partial_{xx}$ mit $\sigma > 0$, wobei ∂_{xx} die zweite partielle Ableitung nach der Ortsvariablen bezeichnet. Dies führt zu der Wärmeleitungsgleichung im Ganzraum

$$\begin{aligned} \partial_t u(t,x) &= \frac{\sigma^2}{2}\partial_{xx}u(t,x), \quad t \geq 0,\ x \in \mathbb{R}, \\ u(0,x) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

Es ist wohlbekannt (vgl. Bauer [6]), dass die Lösung dieser Gleichung für entsprechende Anfangsdaten u_0 durch

$$u(t,x) := \int_\Omega u_0(x + X_t)\,\mathrm{d}\mathbb{P} \quad (t \geq 0,\ x \in \mathbb{R}) \tag{0.3}$$

gegeben ist, wobei $(X_t)_{t\geq 0}$ eine 1-dimensionale Brownsche Bewegung mit Volatilitätsparameter σ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathscr{F}, \mathbb{P})$ ist. Allgemein ist der hier auftretende stochastische Prozess $(X_t)_{t\geq 0}$ ein Markov-Prozess oder, unter zusätzlichen Annahmen an den Grundraum G, ein Lévy-Prozess. Falls die Lösung u des Cauchy-Problems (0.1) - (0.2) eine Darstellung der Form (0.3) mit einem Lévy-Prozess $(X_t)_{t\geq 0}$ besitzt, so wird A auch als Generator dieses Lévy-Prozesses bezeichnet.

Nehmen wir nun an, dass die Volatilität auf Grund von Parameterunsicherheit im zugrundeliegenden Finanzmarkt nicht genau bekannt ist, so führt die Bestimmung von Preisen zu der sogenannten G-Wärmeleitungsgleichung

$$\partial_t u(t,x) = \sup_{\sigma \in [\sigma_\ell, \sigma_h]} \frac{\sigma^2}{2}\partial_{xx}u(t,x), \quad t \geq 0,\ x \in \mathbb{R}, \tag{0.4}$$

$$u(0,x) = u_0(x), \quad x \in \mathbb{R} \tag{0.5}$$

mit $0 < \sigma_\ell < \sigma_h$. Diese Gleichung ist eine voll nichtlineare partielle Differentialgleichung und gehört zur Klasse der Hamilton-Jacobi-Bellman-Gleichungen, welche auch im Zusammenhang mit stochastischen Kontrollproblemen auftreten, vgl. Krylov [56], Fleming-Soner [38] oder Yong-Zhou [83]. Um derartige Gleichungen auf Lösbarkeit untersuchen zu können, wurde das Konzept der Viskositätslösung entwickelt, welches eine, wenn auch sehr schwache, Lösung für eine sehr große Anzahl an Gleichungen liefert, siehe z.B. Ishii-Lions [52], Crandall-Ishii-Lions [21] oder Caffarelli-Cabré [14]. Da es sich, wie bereits erwähnt, bei der G-Wärmeleitungsgleichung um eine nichtlineare Gleichung handelt, können die Lösungen nicht mithilfe eines Integrals bezüglich eines Wahrscheinlichkeitsmaßes dargestellt werden.

Zu Beginn des 21. Jahrhunderts wurden nichtlineare Erwartungen als Erweiterungen von Erwartungswerten bezüglich Wahrscheinlichkeitsmaßen eingeführt, siehe Coquet et al. [20] oder Peng [65]. Nichtlineare Erwartungen sind im sublinearen Fall eng verbunden mit kohärenten monetären Risikomaßen, welche 1999 von Artzner et al. [4] eingeführt wurden und in den folgenden Jahren unter anderem von Delbaen [23], [24], Fritteli-Rosazza Gianin [45], [44], Föllmer-Schied [41], [42], [43] und Schied [75] zu einer kompletten Theorie entwickelt wurden. Mithilfe des Konzepts einer nichtlinearen Erwartung erhält man für die eindeutige Viskositätslösung von (0.4) - (0.5) eine ähnliche Darstellung wie in (0.3) mit einer sublinearen Erwartung anstelle eines Integrals. Die Lösung u ist dann durch

$$u(t,x) = \mathcal{E}(u_0(x + X_t)) \quad (t \geq 0,\ x \in \mathbb{R}) \tag{0.6}$$

gegeben, wobei $(X_t)_{t\geq 0}$ eine G-Brownsche Bewegung unter der nichtlinearen Erwartung $\mathcal{E}$ ist. Letztere wird häufig auch als G-Erwartung bezeichnet und geht auf Peng [66], [67] zurück. Wir verweisen weiterhin auf Cheridito et al. [18] sowie Soner et al. [76], [77] für den Zusammenhang zwischen der G-Erwartung und voll nichtlinearen partiellen Differentialgleichungen sowie stochastischen Rückwärtsgleichungen zweiter Ordnung, sogenannten 2BSDEs. Wählt man

$$\Lambda := [\sigma_\ell, \sigma_h] \quad \text{und} \quad A_\lambda := \frac{\lambda^2}{2}\partial_{xx} \quad (\lambda \in \Lambda),$$

so erhält man die Banachraum-wertige Differentialgleichung

$$u'(t) = \sup_{\lambda\in\Lambda} A_\lambda u(t), \quad t \geq 0, \tag{0.7}$$

$$u(0) = u_0 \tag{0.8}$$

als abstrakte Formulierung der G-Wärmeleitungsgleichung, wobei das auftretende Supremum punktweise in der Ortsvariablen zu verstehen ist.

In der vorliegenden Arbeit betrachten wir Gleichungen der Form (0.7) - (0.8) mit einer nichtleeren Indexmenge Λ und einer Familie $(A_\lambda)_{\lambda\in\lambda}$ von Generatoren von Lévy-Prozessen als abstrakte Versionen von Hamilton-Jacobi-Bellman-Gleichungen. Derartige Gleichungen untersuchen wir im Hinblick auf die Existenz von Viskositätslösungen und stellen die Lösung unter Berücksichtigung der zu den $(A_\lambda)_{\lambda\in\Lambda}$ gehörigen Halbgruppen ebenfalls durch eine Halbgruppe dar. Der hierbei betrachtete Zugang ist durch Nisio [60] inspiriert und führt zu sogenannten Nisio-Halbgruppen. Die wesentliche Idee dieses Ansatzes ist es, auf einer immer feineren Zeitpartition über die zu den $(A_\lambda)_{\lambda\in\Lambda}$ gehörigen Halbgruppen zu optimieren und in den Grenzwert zu gehen. Schlussendlich werden wir zeigen, dass die Lösung von (0.7) - (0.8) sich ebenfalls in der Form (0.6) darstellen lässt. Hierbei ist der auftretende stochastische Prozess ein Lévy-Prozess unter einer sublinearen Erwartung $\mathcal{E}$, welche in einem gewissen Sinne als Supremum über Verteilungen von stochastischen Integralen über die zu den $(A_\lambda)_{\lambda\in\Lambda}$ gehörigen Halbgruppen definiert wird. Wir verweisen auf Neufeld-Nutz [58] und Hu-Peng [50] für ähnliche Ergebnisse mit anderen Methoden unter teils leicht stärkeren Annahmen.

Als Grundlage, um für die Lösung von (0.7) - (0.8) die Darstellung (0.6) zu erlangen, diskutieren wir nichtlineare Erwartungen. In Einklang mit Peng [65] definieren wir eine nichtlineare Erwartung als ein monotones, konstantenerhaltendes Funktional auf dem Raum $\mathscr{L}^\infty(\Omega, \mathscr{F})$ aller beschränkten messbaren Funktionen $\Omega \to \mathbb{R}$, wobei $(\Omega, \mathscr{F})$ ein beliebiger Messraum ist. Da der topologische Dualraum von $\mathscr{L}^\infty(\Omega, \mathscr{F})$ mit dem Raum $\mathrm{ba}(\Omega, \mathscr{F})$ aller endlich-additiven, signierten Maße mit beschränkter Variation identifiziert werden kann, entsprechen lineare Erwartungen den Integralen von endlich additiven Wahrscheinlichkeitsmaßen. Um ein nichtlineares Analogon von Integralen σ-additiver Wahrscheinlichkeitsmaße zu erhalten, müssen zusätzliche Stetigkeitsannahmen an die nichtlineare Erwartung getroffen werden. Hierbei muss im Wesentlichen zwischen drei Arten der Stetigkeit unterschieden werden, die im linearen Fall alle äquivalent zur σ-Additivität des zugehörigen Wahrscheinlichkeitsmaßes sind. Bereits für sublineare Erwartungen jedoch hat jede einzelne dieser drei Stetigkeitsarten ganz unterschiedliche Konsequenzen für die Erwartung. Unter gewissen Stetigkeitsannahmen an die Erwartung werden wir schließlich die Existenz von Lévy-Prozessen unter konvexen beziehungsweise sublinearen

Erwartungen diskutieren, wobei wir eine nichtlineare Version des Satzes von Kolmogorov verwenden. Damit erhalten wir eine natürliche Erweiterung der entsprechenden wohlbekannten Beziehung zwischen sogenannten Markovschen Faltungshalbgruppen und Lévy-Prozessen (vgl. Applebaum [2] oder Sato [72]).

Die vorliegende Arbeit ist wie folgt gegliedert:

In **Kapitel 1** beschäftigen wir uns mit der Erweiterung von sogennanten Prä-Erwartungen auf einem Unterraum von $\mathscr{L}^\infty(\Omega, \mathscr{F})$ zu nichtlinearen Erwartungen. Dabei werden wir ähnlich wie Cerreia-Vioglio et al. [15] zunächst maximale Erweiterungen von Prä-Erwartungen ohne zusätzliche Stetigkeitsannahmen betrachen und diese Erweiterungen auf verschiedene Arten charakterisieren. Im konvexen Fall leiten wir mithilfe von konvexer Dualitätstheorie eine duale Erweiterungsmethode her, welche im späteren Verlauf von Nutzen sein wird, um Erweiterungen unter zusätzlichen Stetigkeitsannahmen herzuleiten, siehe Theorem 1.61. Dies ist das Hauptresultat des Kapitels und kann als eine konvexe Version des Satzes von Daniell-Stone gesehen werden, wobei wir im Beweis eine Version von Choquets Kapazibilitätstheorem verwenden. Wir werden sehen, dass für konvexe Prä-Erwartungen die beiden oben genannten Erweiterungen jeweils zu konvexen Erwartungen führen und auf duale Darstellungen der jeweiligen Erweiterung eingehen.

Die oben erwähnten Erweiterungsverfahren werden wir in **Kapitel 2** verwenden, um eine nichtlineare Version des Satzes von Kolmogorov über die Existenz von stochastischen Prozessen zu beweisen. Hierbei ist die Fragestellung, ob zu einer Familie von endlich-dimensionalen Randverteilungen ein stochastischer Prozess mit diesen Randverteilungen existiert. Im linearen Fall ist die wesentliche Bedingung hierfür, dass die Randverteilungen eine gewisse Konsistenzbedingung erfüllen müssen, welche die Wohldefiniertheit eines σ-additiven Inhalts auf einer Algebra sicherstellt. Im nichtlinearen Fall fordern wir eine ähnliche Konsistenzbedingung an eine Familie von endlich-dimensionalen Randverteilungen, in diesem Fall nichtlineare Erwartungen. Dies stellt die Wohldefiniertheit einer Prä-Erwartung auf einem Unterraum von $\mathscr{L}^\infty(\Omega, \mathscr{F})$ sicher, wobei $(\Omega, \mathscr{F})$ hier der kanonische Pfadraum ist. Unter dieser Konsistenzbedingung beweisen wir dann eine nichtlineare Version des oben genannten Satzes von Kolmogorov, siehe Theorem 2.10. Im sublinearen Fall stellt die duale Version dieses Ergebnisses eine robuste Version des Satzes von Kolmogorov dar, siehe Theorem 2.13. Diese Ergebnisse ermöglichen es uns, mit Hilfe von nichtlinearen Kernen, Markov- und Lévy-Prozesse unter nichtlinearen Erwartungen zu konstruieren, siehe Theorem 2.28 und Theorem 2.32. Hierbei müssen die Familien von Übergangsverteilungen (nichtlineare Kerne) die sogenannten Chapman-Kolmogorov-Gleichungen erfüllen. Diese Gleichungen entsprechen der Zeitkonsistenz von dynamischen monetären Risikomaßen und implizieren die Konsistenz der resultierenden endlich-dimensionalen Randverteilungen. Wir verweisen auf Delbaen [25], Delbaen et al. [26], Cheridito et al. [16], Föllmer-Penner [40] oder Bartl [5] für eine Diskussion über die Zeitkonsistenz von dynamischen monetären Risikomaßen.

In **Kapitel 3** konstruieren wir mithilfe von Nisio-Halbgruppen zu Familien $(A_\lambda)_{\lambda \in \Lambda}$ von Generatoren explizit Lévy-Prozesse, deren Preisprozesse unter sublinearen Erwartungen Lösungen von Differentialgleichungen der Form (0.7) liefern. Hierzu verwenden wir einen Halbgruppentheoretischen Zugang, welcher eine von der klassischen Literatur abweichende Definition einer

Viskositätslösung benötigt. Diese Definition einer Viskositätslösung führt jedoch in den meisten Fällen sogar auf eine etwas größere Klasse von Testfunktionen, wodurch der Lösungsbegriff leicht stärker ist. Des Weiteren zeigen wir in Theorem 3.9 und Theorem 3.16, dass für Gleichungen der Form (0.7) - (0.8), bei denen alle Operatoren A_λ orts- und zeitunabhängige Koeffizienten haben, die Lösungen von der Form (0.6) sind, wobei $(X_t)_{t\geq 0}$ ein Lévy-Prozess unter einer sublinearen Erwartung $\mathcal{E}$ ist. Weiterhin zeigen wir, dass diese Erwartung sich als Supremum von Integralen über Wahrscheinlichkeitsmaße schreiben lässt. Genauer sind diese Wahrscheinlichkeitsmaße durch Halbgruppen-theoretisch definierte Verteilungen von stochastischen Integralen mit endlicher Orts-Zeit-Partition gegeben. Wir wenden diese Ergebnisse dann auf allgemeine partielle Integro-Differentialgleichungen im Ganzraum sowie auf dem Torus an, siehe Beispiel 3.19 beziehungsweise Beispiel 3.23. Abschließend betrachten wir zeithomogene Markov-Ketten mit endlichem Zustandsraum in stetiger Zeit unter sublinearen Erwartungen. Hierbei verallgemeinern wir den Begriff einer Q-Matrix, siehe Norris [61], zu einem Q-Operator und erhalten zu jedem sublinearen Q-Operator $\mathcal{Q}$ eine zeithomogene Markov-Kette, deren Übergangsverteilungen Lösungen gewöhnlicher Differentialgleichungen der Form

$$\begin{aligned} u'(t) &= \mathcal{Q}u(t), \quad t \geq 0, \\ u(0) &= u_0 \end{aligned}$$

mit $u_0 \in \mathbb{R}^d$ sind. Hierbei bezeichnet d die Mächtigkeit des Zustandsraumes. Es wird sich zudem herausstellen, dass die Übergangsverteilungen der zugehörigen Markov-Kette durch eine Nisio-Halbgruppe zu einer Familie von Q-Matrizen gegeben sind, welche den sublinearen Q-Operator $\mathcal{Q}$ dual darstellt.

Teile dieser Arbeit sind in gemeinsamer Zusammenarbeit mit Robert Denk und Michael Kupper entstanden, siehe [28] und [29].

Contents

Introduction and main results

In many areas, such as physics, biology, medicine or economics, partial differential equations describe the evolution of dynamic systems in space and time. For example, the movement of a wave, the propagation of heat, the vibration of a string, the concentration of certain substances in the bloodstream or prices in financial markets, such as the Bachelier model or the Black-Scholes model, are being described by partial differential equations. In contrast to ordinary differential equations, even for linear autonomous partial differential equations it is only in rare cases possible to explicitly compute the solution. Therefore, in most cases, abstract methods are being used in order to prove the existence and uniqueness of a solution. In particular, in the analysis of linear autonomous partial differential equations semigroups play a fundamental role. The main idea of this theory is to consider the space and time variables separately and to read the partial differential equation as a Banach space valued ordinary differential equation. This leads to an abstract Cauchy problem of the form

$$u'(t) = Au(t), \quad t \geq 0, \tag{0.9}$$

$$u(0) = u_0. \tag{0.10}$$

Here, A is in most applications a linear differential or integro-differential operator in the space variables and u_0 is an element of a suitable function space. In the present work, the considered function space will always be the space $\mathrm{BUC}(G)$ of all bounded uniformly continuous functions $G \to \mathbb{R}$ for some metric space G. The aim of semigroup theory is to prove the existence of a unique solution for such a differential equation. Moreover, the solution should depend continuously on the initial data. In this case, one says that the Cauchy problem (0.9) - (0.10) is well-posed. One of the main results of semigroup theory is that the abstract Cauchy problem (0.9) - (0.10) is well-posed if and only if A is the generator of a strongly continuous semigroup $(S(t))_{t\geq 0}$. Here, $S(t)$ for $t \geq 0$ can be interpreted as an abstract version of the exponential e^{tA} and the unique solution u of the above initial value problem is given by $u(t) := S(t)u_0$ for all $t \geq 0$. Another main result of semigroup theory is the Hille-Yosida Theorem, which fully characterizes the generators of strongly continuous semigroups. For a detailed discussion on operator semigroups we refer to Pazy [63] or Engel-Nagel [34], [35].

In the present work, we particularly consider differential equations whose solutions describe price processes in financial markets. This leads to certain conditions on the generator A, e.g. that A satisfies a positive maximum principle and that all constant functions lie in the kernel of A. An example for such an operator in the case $G = \mathbb{R}$ is given by $A := \frac{\sigma^2}{2}\partial_{xx}$ with $\sigma > 0$, where ∂_{xx} denotes the second partial derivative with respect to the space variable. This leads to the heat equation in the whole space

$$\begin{aligned} \partial_t u(t,x) &= \frac{\sigma^2}{2}\partial_{xx}u(t,x), \quad t \geq 0,\ x \in \mathbb{R}, \\ u(0,x) &= u_0(x), \quad x \in \mathbb{R}. \end{aligned}$$

It is well known (cf. Bauer [6]) that, for suitable initial data u_0, the solution to this equation is given by

$$u(t,x) := \int_\Omega u_0(x+X_t)\,\mathrm{d}\mathbb{P} \quad (t \geq 0,\ x \in \mathbb{R}), \tag{0.11}$$

where $(X_t)_{t\geq 0}$ is a 1-dimensional Brownian Motion with volatility parameter σ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. In general, the appearing stochastic process $(X_t)_{t\geq 0}$ is a Markov process or, under additional assumptions on the state space G, a Lévy process. If the solution u to the Cauchy problem (0.9) - (0.10) admits a representation of the form (0.11) with a Lévy process $(X_t)_{t\geq 0}$, one says that A is the generator of this Lévy process.

If we assume that, due to parameter uncertainty in the underlying financial market, the volatility is not exactly known, the pricing of certain financial assets leads to the so called G-heat equation

$$\partial_t u(t,x) = \sup_{\sigma\in[\sigma_\ell,\sigma_h]} \frac{\sigma^2}{2}\partial_{xx}u(t,x), \quad t \geq 0,\ x \in \mathbb{R}, \tag{0.12}$$

$$u(0,x) = u_0(x), \quad x \in \mathbb{R} \tag{0.13}$$

with $0 < \sigma_\ell < \sigma_h$. This equation is a fully nonlinear partial differential equation and belongs to the class of Hamilton-Jacobi-Bellman equations, which also appear in the context of stochastic optimal control, cf. Krylov [56], Fleming-Soner [38] or Yong-Zhou [83]. In order to investigate the solvability of such equations, the concept of viscosity solutions has been developed, cf. Ishii-Lions [52], Crandall-Ishii-Lions [21] or Caffarelli-Cabré [14]. Although this is a rather weak notion of a solution, it guarantees the solvability for a large amount of equations. Since, as mentioned above, the G-heat equation is a nonlinear equation, solutions cannot be represented via an integral with respect to a probability measure.

At the beginning of the 21st century, nonlinear expectations have been introduced as generalizations of expected values under probability measures, cf. Coquet et al. [20] or Peng [65]. In the sublinear case, nonlinear expectations are closely related to coherent monetary risk measures as introduced by Artzner et al. [4] and further developed by Delbaen [23], [24], Fritteli-Rosazza Gianin [45], [44], Föllmer-Schied [41], [42], [43] and Schied [75] amongst others. Using the concept of a nonlinear expectation, the unique viscosity solution to (0.12) - (0.13) admits a similar representation as in (0.11) with a sublinear expectation instead of an integral. The solution u is then given by

$$u(t,x) = \mathcal{E}(u_0(x+X_t)) \quad (t \geq 0,\ x \in \mathbb{R}), \tag{0.14}$$

where $(X_t)_{t\geq 0}$ is a G-Brownian Motion under the nonlinear expectation $\mathcal{E}$. The latter is often being referred to as G-expectation and is due to Peng [66], [67]. Further, we refer to Cheridito et al. [18] as well as Soner et al. [76], [77] for the connection between the G-expectation, fully nonlinear partial differential equations and second order backward stochastic differential equations, so-called 2BSDEs. Choosing

$$\Lambda := [\sigma_\ell, \sigma_h] \quad \text{and} \quad A_\lambda := \frac{\lambda^2}{2}\partial_{xx} \quad (\lambda \in \Lambda),$$

we obtain the Banach space valued differential equation

$$u'(t) = \sup_{\lambda\in\Lambda} A_\lambda u(t), \quad t \geq 0, \tag{0.15}$$

$$u(0) = u_0 \tag{0.16}$$

as an abstract formulation of the G-heat equation, where the supremum is taken pointwise in the space variable.

In the present work, we consider equations of the form (0.15) - (0.16) with a nonempty index set Λ and a family $(A_\lambda)_{\lambda\in\Lambda}$ of generators of Lévy processes as abstract versions of Hamilton-Jacobi-Bellman equations. We investigate such equations in view of the existence of viscosity solutions and represent the solution as a semigroup using the semigroups belonging to the family $(A_\lambda)_{\lambda\in\Lambda}$. The ansatz we consider here, in order to obtain a solution, is inspired by Nisio [60] and leads to so-called Nisio semigroups. The main idea of this approach is to optimize over the semigroups belonging to the $(A_\lambda)_{\lambda\in\Lambda}$ on a finer and finer time partition and then pass to the limit. Finally, we show that the solution to (0.15) - (0.16) admits a representation of the form (0.11). Here, the appearing stochastic process is a Lévy process under a sublinear expectation, which is in some sense the supremum over distributions of stochastic integrals, which are being defined by the semigroups belonging to the family $(A_\lambda)_{\lambda\in\Lambda}$. We refer to Neufeld-Nutz [58] and Hu-Peng [50] for similar results with different methods under slightly more restrictive assumptions.

In order to obtain the representation (0.11) for solutions to abstract initial value problems of the form (0.15) - (0.16), we discuss nonlinear expectations. Following Peng [65], we define a nonlinear expectation to be a monotone, constant preserving functional on the space $\mathscr{L}^\infty(\Omega,\mathscr{F})$ of all bounded measurable functions $\Omega \to \mathbb{R}$, where $(\Omega,\mathscr{F})$ is an arbitrary measurable space. Since the topological dual space of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ can be identified with the space $\mathrm{ba}(\Omega,\mathscr{F})$ of all finitely additive signed measures with bounded variation, linear expectations correspond to integrals with respect to finitely additive probability measures. In order to obtain a nonlinear analogon of integrals with respect to σ-additive probability measures, additional continuity assumptions have to be imposed on the nonlinear expectation. Here, one basically has to distinguish between three continuity assumptions which, in the linear case, are all equivalent to the σ-additivity of the respective probability measure. However, already for sublinear expectations each one of these three types of continuity leads to quite different consequences for the expectation. Under certain continuity assumptions, we then discuss the existence of Lévy processes under convex or sublinear expectations using a nonlinear version of Kolmogorov's theorem. This provides a natural extension of the well-known relation between so-called Markovian convolution semigroups and Lévy processes (cf. Applebaum [2] or Sato [72]).

The present thesis is organized as follows:

Chapter 1 is dedicated to the extension of so-called pre-expectations on a subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ to nonlinear expectations. In a similar way as Cerreia-Vioglio et al. [15], we first consider maximal extensions of pre-expectations without assuming any additional continuity properties. We further characterize this maximal extension in various ways. In the convex case, we derive a dual extension procedure using tools from convex analysis and duality theory. Later, this will be useful in the course of deriving extensions under additional continuity assumptions, see Theorem 1.61. The latter is the main result of this chapter and can be seen as a convex version of the Daniell-Stone Theorem, where we use a variant of Choquet's Capacitability Theorem in the proof. We will see that for convex pre-expectations both extension procedures mentioned above lead to convex expectations and we further discuss dual representations of the respective extension.

We use the extension methods mentioned above in **Chapter 2** in order to derive a nonlinear version of Kolmogorov's theorem on the existence of stochastic processes. That is, given a family of finite-dimensional marginal distributions, we are looking for a stochastic process which has exactly these marginals. In the linear case, the essential condition for the existence of such a process is that the family of marginal distributions satisfies a certain consistency condition, which guarantees the well-definedness of some pre-measure on a certain algebra. In the nonlinear case, we will impose a similar consistency condition on the family of finite-dimensional marginal distributions, which are nonlinear expectations in this case. This condition will ensure that a certain pre-expectation is well-defined on a subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$, where $(\Omega,\mathscr{F})$ is the canonical path space. Under the above mentioned consistency condition, we then derive a nonlinear version of Kolmogorov's extension theorem, see Theorem 2.10. In the sublinear case, the dual version of the latter yields a robust version of the Kolmogorov Theorem, see Theorem 2.13. These results allow us to construct nonlinear Markov and Lévy processes under nonlinear expectations by means of nonlinear kernels, see Theorem 2.28 and Theorem 2.32, respectively. Here, the family of transition probabilites (nonlinear kernels) have to satisfy the so-called Chapman-Kolmogorov equations. These equations correspond to the time consistency of dynamic monetary risk measures and imply the consistency of the resulting family of finite-dimensional marginal distributions. We refer to Delbaen [25], Delbaen et al. [26], Cheridito et al. [16], Föllmer-Penner [40] or Bartl [5] for a discussion on time consistency of dynamic monetary risk measures.

In **Chapter 3** we use Nisio semigroups to given families $(A_\lambda)_{\lambda\in\Lambda}$ of generators in order to explicitly construct Lévy processes under sublinear expectations such that their price processes yield solutions to differential equations of the form (0.15). For this, we use a semigroup theoretic approach which requires a slightly different definition of a viscosity solution than in the classical literature. However, in most cases, this definition of a viscosity solution leads to an even slightly larger class of test functions and we therefore end up with a slightly stronger notion of a solution. Moreover, in Theorem 3.9 and Theorem 3.16 we show that for equations of the form (0.15) - (0.16), where all operators A_λ have space and time independent coefficients, the solutions are of the form (0.14) with a Lévy process $(X_t)_{t\geq 0}$ under a sublinear expectation $\mathcal{E}$. Furthermore, we show that this expectation has a representation as a supremum over integrals with respect to probability measures. More precisely, these probability measures are distributions of stochastic integrals with a finite space-time partition, which are defined in a semigroup theoretic way. We then apply these results to general partial integro-differential equations in the whole space and on the torus, see Example 3.19 and Example 3.23, respectively. We close this chapter by considering time homogeneous Markov chains with finite state space in continuous time under sublinear expectations. Here, we first generalize the notion of a Q-matrix, see for example Norris [61], to a Q-operator and derive a dual representation of such operators by means of Q-matrices in the sublinear case. For every sublinear Q-operator $\mathcal{Q}$ we then construct a time homogeneous Markov chain, whose transition probabilities yield solutions to ordinary differential equations of the form

$$\begin{aligned} u'(t) &= \mathcal{Q}u(t), \quad t \geq 0, \\ u(0) &= u_0 \end{aligned}$$

with $u_0 \in \mathbb{R}^d$. Here d denotes the cardinality of the state space. Moreover, it turns out that the transition probabilites of the corresponding Markov chain are given by a Nisio semigroup to a family of Q-matrices which represents the sublinear Q-operator $\mathcal{Q}$.

Parts of this thesis are based on joint work with Robert Denk and Michael Kupper, see [28] and [29].

Chapter 1

Nonlinear expectations

1.1 Extension of nonlinear pre-expectations

In this section, we consider extensions of nonlinear pre-expectations without assuming any continuity properties. In the linear case, this leads to finitely additive probability measures. We will specialize to the continuous case (leading to σ-additive measures) in Section 1.4. We start by defining nonlinear expectations and showing some of their basic properties.

Throughout this section, let Ω be a nonempty set and $\mathscr{F} \subset 2^\Omega$ an arbitrary σ-algebra on Ω, where 2^S denotes the power set of a set S. We want to emphasize that, throughout this section, $\mathscr{F} = 2^\Omega$ is a possible choice for $\mathscr{F}$. We denote by $\mathscr{L}^\infty(\Omega, \mathscr{F})$ the space of all bounded $\mathscr{F}$-$\mathscr{B}(\mathbb{R})$-measurable random variables $X\colon \Omega \to \mathbb{R}$, where $\mathscr{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. We write $\mathrm{ba}(\Omega, \mathscr{F})$ for the space of all real-valued and finitely additive measures on $(\Omega, \mathscr{F})$ with finite total variation and $\mathrm{ca}(\Omega, \mathscr{F})$ for the subspace of all σ-additive signed measures on $(\Omega, \mathscr{F})$. The subset $\mathrm{ba}_+(\Omega, \mathscr{F})$ stands for all positive elements in $\mathrm{ba}(\Omega, \mathscr{F})$, and we write $\mathrm{ba}_+^1(\Omega, \mathscr{F})$ for the set of all $\mu \in \mathrm{ba}_+(\Omega, \mathscr{F})$ with $\mu(\Omega) = 1$. Analogously, we define $\mathrm{ca}_+(\Omega, \mathscr{F})$ and $\mathrm{ca}_+^1(\Omega, \mathscr{F})$.

The space $\mathscr{L}^\infty(\Omega, \mathscr{F})$ and subspaces thereof will be endowed with the standard norm $\|\cdot\|_\infty$, defined by

$$\|X\|_\infty := \sup_{\omega \in \Omega} |X(\omega)| \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

For $\alpha \in \mathbb{R}$ we will make use of the notation $\alpha := \alpha 1_\Omega$, and for $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ we will write $\mathbb{R} \subset M$ instead of $\{\alpha 1_\Omega \colon \alpha \in \mathbb{R}\} \subset M$. Here, 1_A stands for the indicator function of $A \subset \Omega$.

On the topological dual space M' of a normed space M, we always consider the weak* topology and on subsets of M' we take the trace topology of the weak* topology.

Using the identification $\mathrm{ba}(\Omega, \mathscr{F}) = (\mathscr{L}^\infty(\Omega, \mathscr{F}))'$ via $\mu X := \int_\Omega X \,\mathrm{d}\mu$ for $\mu \in \mathrm{ba}(\Omega, \mathscr{F})$ and $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ (cf. [33, p. 258]), every monotone linear functional $\mathbb{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ with $\mathbb{E}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$ is the expectation of a finitely additive probability measure $\mu \in \mathrm{ba}_+^1(\Omega, \mathscr{F})$. This motivates the following definition, which is due to Peng [65].

1.1 Definition. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$. A *(nonlinear) pre-expectation* $\mathcal{E}$ on M is a functional $\mathcal{E}\colon M \to \mathbb{R}$ with the following properties:

(i) Monotonicity: $\mathcal{E}(X) \le \mathcal{E}(Y)$ for all $X, Y \in M$ with $X \le Y$.

(ii) Constant preserving: $\mathcal{E}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$.

A pre-expectation on $\mathscr{L}^\infty(\Omega, \mathscr{F})$ is called an *expectation*.

In the following remark we give a collection of some basic properties and terminologies in the context of nonlinear expectations.

1.2 Remark. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation on M.

a) For all $X \in M$ it holds $|\mathcal{E}(X)| \leq \|X\|_\infty$, and therefore, $\mathcal{E}$ is continuous at 0. Indeed,

$$-\|X\|_\infty = \mathcal{E}(-\|X\|_\infty) \leq \mathcal{E}(X) \leq \mathcal{E}(\|X\|_\infty) = \|X\|_\infty$$

for all $X \in M$.

b) Assume that $M + \mathbb{R} := \{X + \alpha\colon X \in M,\ \alpha \in \mathbb{R}\} \subset M$ and that $\mathcal{E}(X + \alpha) = \mathcal{E}(X) + \alpha$ for all $X \in M$ and $\alpha \in \mathbb{R}$. This property is often referred to as *translation invariance* (cf. [24]) or *cash additivity* (cf. [43]). Then,

$$\mathcal{E}(X) - \mathcal{E}(Y) \leq \mathcal{E}(Y + \|X - Y\|_\infty) - \mathcal{E}(Y) = \|X - Y\|_\infty$$

for all $X, Y \in M$. Due to symmetry, we thus obtain that $\mathcal{E}\colon M \to \mathbb{R}$ is 1-*Lipschitz*, i.e. Lipschitz continuous with Lipschitz constant 1.

c) Assume that $M + M := \{X + Y\colon X, Y \in M\} \subset M$ and that $\mathcal{E}\colon M \to \mathbb{R}$ is *subadditive*, i.e. $\mathcal{E}(X + Y) \leq \mathcal{E}(X) + \mathcal{E}(Y)$ for all $X, Y \in M$. Then, for all $\alpha \in \mathbb{R}$ and all $X \in M$ we have that

$$\mathcal{E}(X + \alpha) \leq \mathcal{E}(X) + \mathcal{E}(\alpha) = \mathcal{E}(X) + \alpha \leq \mathcal{E}(X + \alpha) + \mathcal{E}(-\alpha) + \alpha = \mathcal{E}(X + \alpha),$$

i.e. $\mathcal{E}(X + \alpha) = \mathcal{E}(X) + \alpha$. Therefore, $\mathcal{E}$ is cash additive and, in particular, 1-Lipschitz.

d) Assume that M is a linear subspace and that $\mathcal{E}$ is cash additive. Then, $\rho(X) := \mathcal{E}(-X)$ for $X \in M$ defines a *monetary risk measure* on M. For a detailed discussion of monetary risk measures we refer to [43] and the references therein.

e) The set $\mathcal{A}_\mathcal{E} := \{X \in M\colon \mathcal{E}(X) \geq 0\}$ is called the *acceptance set* of $\mathcal{E}$. If $M + \mathbb{R} \subset M$ and $\mathcal{E}$ is cash additive, we get that

$$\mathcal{E}(X) = \sup\{\alpha \in \mathbb{R}\colon \mathcal{E}(X - \alpha) \geq 0\} = \sup\{\alpha \in \mathbb{R}\colon X - \alpha \in \mathcal{A}_\mathcal{E}\}$$

for all $X \in M$.

f) Assume $M + \mathbb{R} \subset M$. Then, the following three statements are equivalent:
   (i) $\mathcal{E}$ is cash additive,
   (ii) $\mathcal{E}(X + \alpha) \leq \mathcal{E}(X) + \alpha$ for all $X \in M$ and all $\alpha \in \mathbb{R}$,
   (iii) $\mathcal{E}(X + \alpha) \geq \mathcal{E}(X) + \alpha$ for all $X \in M$ and all $\alpha \in \mathbb{R}$.

   In fact, first assume that $\mathcal{E}(X + \alpha) \leq \mathcal{E}(X) + \alpha$ for all $X \in M$ and all $\alpha \in \mathbb{R}$. Then,

$$\mathcal{E}(X) = \mathcal{E}\big((X + \alpha) - \alpha\big) \leq \mathcal{E}(X + \alpha) - \alpha.$$

   Now, assume that $\mathcal{E}(X + \alpha) \geq \mathcal{E}(X) + \alpha$ for all $X \in M$ and all $\alpha \in \mathbb{R}$. Then,

$$\mathcal{E}(X + \alpha) - \alpha \leq \mathcal{E}\big((X + \alpha) - \alpha\big) = \mathcal{E}(X).$$

g) Assume that M be a *convex cone*, i.e. M is convex and $\lambda M := \{\lambda X\colon X \in M\} \subset M$ for all $\lambda > 0$. Then, $X + Y = 2\big(\frac{1}{2}X + \frac{1}{2}Y\big) \in M$ for all $X, Y \in M$, and therefore, $M + M \subset M$. Since $0 \in M$, any two of the following three conditions imply the remaining third:

   (i) $\mathcal{E}$ is convex,

   (ii) $\mathcal{E}$ is positive homogeneous (of degree 1),

   (iii) $\mathcal{E}$ is subadditive.

   If two (and therefore all three) of the above conditions are fulfilled, we say that $\mathcal{E}$ is *sublinear*.

h) Assume that M be a linear subspace of $\mathscr{L}^\infty(\Omega, \mathscr{F})$ and assume that $\mathcal{E}$ is sublinear. Let $\|\cdot\|_{\mathcal{E}}\colon M \to \mathbb{R}$ be defined by

$$\|X\|_{\mathcal{E}} := \mathcal{E}(|X|) \quad (X \in M).$$

   Then, $\|\cdot\|_{\mathcal{E}}$ is a seminorm on M with $\|X\|_{\mathcal{E}} \leq \|X\|_\infty$. Indeed, we have that $\|0\|_{\mathcal{E}} = \mathcal{E}(0) = 0$,

$$\|X + Y\|_{\mathcal{E}} = \mathcal{E}(|X + Y|) \leq \mathcal{E}(|X| + |Y|) \leq \mathcal{E}(|X|) + \mathcal{E}(|Y|) = \|X\|_{\mathcal{E}} + \|Y\|_{\mathcal{E}}$$

   for all $X, Y \in M$ and $\|\alpha X\|_{\mathcal{E}} = \mathcal{E}(|\alpha||X|) = |\alpha|\mathcal{E}(|X|) = |\alpha|\|X\|_{\mathcal{E}}$ for all $X \in M$ and $\alpha \in \mathbb{R}$. Note that if $M = \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $\mathcal{E} = \mu \in \mathrm{ba}(\Omega, \mathscr{F})$ is linear, then $\|X\|_{\mathcal{E}} = \int_\Omega |X|\,\mathrm{d}\mu$ for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$.

i) Let $S \neq \emptyset$ be a set and $T\colon \Omega \to S$ a mapping. Then, $N := \{Y \in \mathscr{L}^\infty(S, 2^S)\colon Y \circ T \in M\}$ contains all constant functions $S \to \mathbb{R}$, and one readily verifies that

$$\mathcal{E} \circ T^{-1}\colon N \to \mathbb{R}, \quad Y \mapsto \mathcal{E}(Y \circ T)$$

   defines a nonlinear pre-expectation on N. We call $\mathcal{E} \circ T^{-1}$ the *distribution* of T under $\mathcal{E}$. Note that if $M = \mathscr{L}^\infty(\Omega, \mathscr{F})$, then $N = \mathscr{L}^\infty(S, \mathscr{S})$, where $\mathscr{S} := \{B \in 2^S\colon T^{-1}(B) \in \mathscr{F}\}$. In particular, $N \subset \mathscr{L}^\infty(S, \mathscr{S})$ for all $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$.

1.3 Example. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$. A set $\mathcal{A} \subset M$ is called an *acceptance set* of M if

(i) $\inf\{\alpha \in \mathbb{R}\colon \alpha \in \mathcal{A}\} = 0$,

(ii) For all $X \in \mathcal{A}$ and all $Y \in M$ with $Y \geq X$ we have that $Y \in \mathcal{A}$.

Assume that $M + \mathbb{R} \subset M$ and that $\mathcal{A} \subset M$ is an acceptance set. Then,

$$\mathcal{E}(X) := \sup\{\alpha \in \mathbb{R}\colon X - \alpha \in \mathcal{A}\} \quad (X \in M)$$

defines a cash additive pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$. In this case, we have that $\mathcal{A}_{\mathcal{E}} = \mathcal{A}$ (see Remark 1.2 e)). Therefore, the mapping $\mathcal{E} \mapsto \mathcal{A}_{\mathcal{E}}$ is a bijection between the set of all cash additive pre-expectations on M and the set of all acceptance sets of M.

1.4 Example. A typical example for an expectation $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ is given by the following construction: Let $\mathscr{P} \subset \mathrm{ba}^1_+(\Omega, \mathscr{F})$ be nonempty, and define $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ by

$$\mathcal{E}(X) := \sup_{\mu \in \mathscr{P}} \mu X \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

Then, it is immediately seen that $\mathcal{E}$ is an expectation, which is even sublinear.

The following lemma shows that a sublinear expectation (cf. Remark 1.2 g)) that coincides with a finitely additive measure on all $\mathscr{F}$-measurable sets is already linear.

1.5 Lemma. *Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a sublinear expectation and $\mu\in \mathrm{ba}^1_+(\Omega,\mathscr{F})$ with $\mathcal{E}(1_A)\le\mu(A)$ for all $A\in\mathscr{F}$. Then, $\mathcal{E}(X)=\mu X$ for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ and therefore, $\mathcal{E}$ is a linear expectation.*

Proof. Let $D:=\operatorname{span}\{1_A\colon A\in\mathscr{F}\}$ and $X\in D$ with $X\ge 0$. Then, there exists some $n\in\mathbb{N}$, some $A_1,\ldots,A_n\in\mathscr{F}$ and some $c_1,\ldots c_n>0$, such that $X=\sum_{k=1}^n c_k 1_{A_k}$. Hence, as $\mathcal{E}$ is sublinear,

$$\mathcal{E}(X)\le\sum_{k=1}^n c_k\mathcal{E}(1_{A_k})\le\sum_{k=1}^n c_k\mu(A_k)=\mu X.$$

Now let $X\in D$ be arbitrary. Then, $X+\|X\|_\infty\in D$ with $X+\|X\|_\infty\ge 0$. Thus,

$$\mathcal{E}(X)=\mathcal{E}(X+\|X\|_\infty)-\|X\|_\infty\le\mu(X+\|X\|_\infty)-\|X\|_\infty=\mu X.$$

Since D is dense in $\mathscr{L}^\infty(\Omega,\mathscr{F})$, and since, by Remark 1.2 b) and c), $\mathcal{E}\colon\mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ and $\mu\colon\mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ are both continuous, we get that $\mathcal{E}(X)\le\mu X$ for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. As $\mathcal{E}$ is subadditive with $\mathcal{E}(0)=0$, we get that

$$0=\mathcal{E}(0)=\mathcal{E}(X-X)\le\mathcal{E}(X)+\mathcal{E}(-X)$$

for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. Therefore,

$$\mathcal{E}(X)\ge-\mathcal{E}(-X)\ge-\mu(-X)=\mu X$$

for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. This shows that $\mathcal{E}(X)=\mu X$ for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. □

Let $M\subset\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $\mathbb{R}\subset M$. Given a pre-expectation $\mathcal{E}\colon M\to\mathbb{R}$, we are looking for extensions of $\mathcal{E}$ to an expectation on $\mathscr{L}^\infty(\Omega,\mathscr{F})$. Hereby, we are interested in the existence and uniqueness of such extensions. We start with the extension of linear pre-expectations. Here, the main challenge is to maintain the monotonicity. However, as the following remark indicates, already in the linear case, there is no hope for uniqueness without additional continuity assumptions.

1.6 Remark. Let $M\subset\mathscr{L}^\infty(\Omega,\mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $1\in M$. We then denote by $\mathrm{ba}^1_+(M)$ the space of all linear pre-expectations on M. A natural question, that arises in this context, is, if the mapping

$$\mathrm{ba}^1_+(\Omega,\mathscr{F})\to\mathrm{ba}^1_+(M),\quad \nu\mapsto\nu|_M \tag{1.1}$$

is bijective. The following theorem by Kantorovich shows that this mapping is surjective, i.e. any linear pre-expectation on M can be extended to a linear expectation on $\mathscr{L}^\infty(\Omega,\mathscr{F})$. For the reader's convenience, we state this theorem and provide a short sketch of the proof. For more details we refer to [80, p. 277]. However, in Example 1.69 we will see that the mapping in (1.1) is not neccessarily injective, not even if $\mathscr{F}=\sigma(M)$, i.e. already if $\mathscr{F}=\sigma(M)$, a linear pre-expectation on M, in general, admits various extensions to an expectation on $\mathscr{L}^\infty(\Omega,\mathscr{F})$.

1.7 Theorem (Kantorovich)**.** *Let $M \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $1 \in M$ and $\mu\colon M \to \mathbb{R}$ be a linear pre-expectation on M. Then, there exists a linear expectation $\nu\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ with $\nu|_M = \mu$.*

Proof. Let
$$\hat\mu(X) := \inf\{\mu X_0\colon X_0 \in M,\ X_0 \geq X\}$$
for all $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. Then, $\hat\mu\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ is a sublinear expectation with $\hat\mu|_M = \mu$. By the extension theorem of Hahn-Banach (see e.g. [79, Theorem 18.1]), there exists a linear functional $\nu\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ with $\nu|_M = \mu$ and $\nu X \leq \hat\mu(X)$ for all $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. Thus,
$$\nu X - \nu Y = \nu(X-Y) \leq \hat\mu(X-Y) \leq \hat\mu(0) = 0$$
for all $X, Y \in \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X \leq Y$. □

In the proof of the previous theorem, before applying the Hahn-Banach Theorem, the linear pre-expectation $\mu\colon M \to \mathbb{R}$ is extended to a sublinear expectation $\hat\mu\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ via
$$\hat\mu(X) := \inf\{\mu X_0\colon X_0 \in M,\ X_0 \geq X\} \quad (X \in \mathscr{L}^\infty(\Omega,\mathscr{F})).$$
The idea for the first extension procedure therefore is, to extend a pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$ on $M \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $\mathbb{R} \subset M$ via
$$\hat{\mathcal{E}}(X) := \inf\{\mathcal{E}(X_0)\colon X_0 \in M,\ X_0 \geq X\} \quad (X \in \mathscr{L}^\infty(\Omega,\mathscr{F})).$$
The following proposition shows that $\hat{\mathcal{E}}$ is an expectation with $\hat{\mathcal{E}}|_M = \mathcal{E}$. Moreover, if M is assumed to be convex or a convex cone, then convexity or sublinearity of $\mathcal{E}$ carry over to the extension $\hat{\mathcal{E}}$, respectively. For related extension results on niveloids we refer to Maccheroni et al. [15].

1.8 Proposition. *Let $M \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation on M. Further, let*
$$\hat{\mathcal{E}}(X) := \inf\{\mathcal{E}(X_0)\colon X_0 \in M,\ X_0 \geq X\}$$
for all $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. Then, the following assertions hold:

a) $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ is an expectation with $\hat{\mathcal{E}}|_M = \mathcal{E}$.

b) If M is convex and $\mathcal{E}$ is convex, then $\hat{\mathcal{E}}$ is convex.

c) If M is a convex cone and $\mathcal{E}$ is sublinear, then $\hat{\mathcal{E}}$ is sublinear.

Proof. a) Let $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. As $\mathbb{R} \subset M$, we have that $\|X\|_\infty \in M$ with $\|X\|_\infty \geq X$. Thus, the set $\{\mathcal{E}(X_0)\colon X_0 \in M,\ X_0 \geq X\}$ is nonempty. Since $X_0 \geq -\|X\|_\infty$ for all $X_0 \in M$ with $X_0 \geq X$, we obtain that
$$\mathcal{E}(X_0) \geq \mathcal{E}(-\|X\|_\infty) = -\|X\|_\infty$$
for all $X_0 \in M$ with $X_0 \geq X$. Hence, $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ is well-defined. Further, if $X \in M$ we have that $\mathcal{E}(X) \leq \mathcal{E}(X_0)$ for all $X_0 \in M$ with $X_0 \geq X$ and therefore, $\hat{\mathcal{E}}(X) = \mathcal{E}(X)$. Since $\mathbb{R} \subset M$, we thus obtain that $\hat{\mathcal{E}}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$. Now, let $X, Y \in \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X \leq Y$. Then, $Y_0 \geq X$ for all $Y_0 \in M$ with $Y_0 \geq Y$ and therefore, $\hat{\mathcal{E}}(X) \leq \hat{\mathcal{E}}(Y)$.

b) Assume that M is convex and that $\mathcal{E}$ is convex. Let $X, Y \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $\lambda \in [0,1]$. Moreover, let $X_0, Y_0 \in M$ with $X_0 \geq X$ and $Y_0 \geq Y$. Since M is convex, $\lambda X_0 + (1-\lambda)Y_0 \in M$ with

$$\lambda X_0 + (1-\lambda)Y_0 \geq \lambda X + (1-\lambda)Y.$$

Due to convexity of $\mathcal{E}$, we thus obtain that

$$\hat{\mathcal{E}}(\lambda X + (1-\lambda)Y) \leq \mathcal{E}(\lambda X_0 + (1-\lambda)Y_0) \leq \lambda \mathcal{E}(X_0) + (1-\lambda)\mathcal{E}(Y_0).$$

Taking the infimum over all $X_0, Y_0 \in M$ with $X_0 \geq X$ and $Y_0 \geq Y$ we get that

$$\hat{\mathcal{E}}(\lambda X + (1-\lambda)Y) \leq \lambda \hat{\mathcal{E}}(X) + (1-\lambda)\hat{\mathcal{E}}(Y).$$

c) Now assume that M is a convex cone and that $\mathcal{E}$ is sublinear. Then, $\mathcal{E}$ is convex and part b) yields that $\hat{\mathcal{E}}$ is convex as well. Moreover, as $\lambda X_0 \in M$ for all $X_0 \in M$ and $\lambda > 0$ we have that

$$\hat{\mathcal{E}}(\lambda X) = \inf\{\mathcal{E}(\lambda X_0) \colon X_0 \in M,\ X_0 \geq X\} = \inf\{\lambda \mathcal{E}(X_0) \colon X_0 \in M,\ X_0 \geq X\} = \lambda \hat{\mathcal{E}}(X)$$

for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ and all $\lambda > 0$. Hence, $\hat{\mathcal{E}}$ is convex and positive homogeneous. Therefore, $\hat{\mathcal{E}}$ is sublinear by Remark 1.2 g).

□

1.9 Remark. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation on M. Let $\tilde{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be an expectation with $\tilde{\mathcal{E}}|_M = \mathcal{E}$ and $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. Then,

$$\widetilde{\mathcal{E}}(X) \leq \widetilde{\mathcal{E}}(X_0) = \mathcal{E}(X_0)$$

for all $X_0 \in M$ with $X_0 \geq X$. Taking the infimum over all $X_0 \in M$ with $X_0 \geq X$, we see that $\widetilde{\mathcal{E}}(X) \leq \hat{\mathcal{E}}(X)$. That is, $\hat{\mathcal{E}}$ is the largest expectation, which extends $\mathcal{E}$.

1.10 Remark. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E} \colon M \to \mathbb{R}$ be a pre-expectation on M. For $X \in M$ let

$$\check{\mathcal{E}}(X) := \sup\{\mathcal{E}(X_0) \colon X_0 \in M,\ X_0 \leq X\}.$$

Then, one readily verifies that $\check{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ is the smallest expectation, which extends $\mathcal{E}$. However, convexity of $\mathcal{E}$ usually does not carry over to $\check{\mathcal{E}}$.

1.11 Remark. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ and $M + \mathbb{R} \subset M$. Further, let $\mathcal{E} \colon M \to \mathbb{R}$ be a cash additive pre-expectation on M.

a) Then, $\hat{\mathcal{E}}$ is cash additive as well. In fact, let $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $\alpha \in \mathbb{R}$. Then, for all $X_0 \in M$ with $X_0 \geq X + \alpha$ we have that $X_0 - \alpha \geq X$ and therefore,

$$\mathcal{E}(X_0) = \mathcal{E}(X_0 - \alpha) + \alpha \geq \hat{\mathcal{E}}(X) + \alpha.$$

Taking the infimum over all $X_0 \in M$ with $X_0 \geq X + \alpha$, we get that $\hat{\mathcal{E}}(X + \alpha) \geq \hat{\mathcal{E}}(X) + \alpha$. By Remark 1.2 f), we thus obtain that $\hat{\mathcal{E}}$ is cash additive.

b) Let $\widehat{\mathcal{A}}_{\mathcal{E}} := \{X \in \mathscr{L}^\infty(\Omega, \mathscr{F})\colon X_0 \in \mathcal{A}_{\mathcal{E}} \text{ for all } X_0 \in M \text{ with } X_0 \geq X\}$. Then, $\widehat{\mathcal{A}}_{\mathcal{E}} = \mathcal{A}_{\hat{\mathcal{E}}}$, i.e. $\widehat{\mathcal{A}}_{\mathcal{E}}$ is the acceptance set of $\hat{\mathcal{E}}$, and therefore,

$$\hat{\mathcal{E}}(X) = \sup\big\{\alpha \in \mathbb{R}\colon X - \alpha \in \widehat{\mathcal{A}}_{\mathcal{E}}\big\} \tag{1.2}$$

for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. Thus, (1.2) provides a second extension procedure for $\mathcal{E}$, which is extending $\mathcal{E}$ via its acceptable positions $\mathcal{A}_{\mathcal{E}}$.

Throughout the remainder of this section, let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega, \mathscr{F})$ with $1 \in M$. For a convex pre-expectation $\mathcal{E}$ on M we want to derive other descriptions of $\hat{\mathcal{E}}$ in terms of $\mathcal{E}$ using tools from convex analysis and duality theory. For a convex function $\mathcal{E}\colon M \to \mathbb{R}$ we write $\mathcal{E}^*$ for its *conjugate function*, i.e. we define

$$\mathcal{E}^*(\mu) := \sup_{X \in M} \mu X - \mathcal{E}(X)$$

for all linear functionals $\mu\colon M \to \mathbb{R}$. Note that the conjugate function $\mathcal{E}^*$ may also take the value $+\infty$. We will see that for a convex pre-expectation $\mathcal{E}$ on M its conjugate function $\mathcal{E}^*$ is concentrated on the class of linear pre-expectations on M. That is, $\mathcal{E}^*$ is finite only for linear pre-expectations on M. As every linear pre-expectation on M is continuous, we therefore obtain the representation

$$\mathcal{E}(X) = \sup_{\mu \in M'} \mu X - \mathcal{E}^*(\mu) \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F}))$$

for all convex pre-expectations $\mathcal{E}$ on M. This allows us to derive a dual representation of $\hat{\mathcal{E}}$ in terms of the convex pre-expectation $\mathcal{E}$. We start with a collection of several well-known facts from convex analysis and duality theory, which we will use frequently throughout the remainder of this chapter.

1.12 Lemma. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on M.*

a) Every linear functional $\mu\colon M \to \mathbb{R}$ with $\mathcal{E}^*(\mu) < \infty$ is a linear pre-expectation on M and therefore, $\mu \in M'$ with $\|\mu\|_{M'} = 1$, where $\|\cdot\|_{M'}$ denotes the operator norm on M'.

b) For all $c \in \mathbb{R}$ the level set

$$\mathscr{P}_c := \{\mu \in M'\colon \mathcal{E}^*(\mu) \leq c\}$$

is a convex compact subset of M'.

c) For all $X \in M$ and all $c \geq \|X\|_\infty - \mathcal{E}(X)$ we have that

$$\mathcal{E}(X) = \max_{\mu \in \mathscr{P}_c} \mu X - \mathcal{E}^*(\mu).$$

Moreover, $\mathscr{P}_c \neq \emptyset$ if and only if $c \geq 0$.

Proof. a) Let $\mu\colon M \to \mathbb{R}$ be linear with $\mathcal{E}^*(\mu) < \infty$. Then, for all $\lambda > 0$ it holds

$$\begin{aligned} 1 - \lambda^{-1}\mathcal{E}^*(\mu) &= -\lambda^{-1}\big(\mathcal{E}(-\lambda) + \mathcal{E}^*(\mu)\big) \\ &\leq -\lambda^{-1}\mu(-\lambda) = \mu 1 = \lambda^{-1}\mu(\lambda) \\ &\leq \lambda^{-1}\big(\mathcal{E}(\lambda) + \mathcal{E}^*(\mu)\big) = 1 + \lambda^{-1}\mathcal{E}^*(\mu). \end{aligned}$$

Letting $\lambda \to \infty$, we obtain that $\mu 1 = 1$. Moreover, for $\lambda > 0$ and all $X, Y \in M$ with $X \leq Y$ we have that

$$\mu(X - Y) = \lambda^{-1}\mu\big(\lambda(X - Y)\big) \leq \lambda^{-1}\big[\mathcal{E}\big(\lambda(X - Y)\big) + \mathcal{E}^*(\mu)\big] \leq \lambda^{-1}\mathcal{E}^*(\mu) \to 0, \quad \lambda \to \infty.$$

This shows that $\mu\colon M \to \mathbb{R}$ is a linear pre-expectation on M. Therefore, Remark 1.2 b) and c) imply that $\mu \in M'$ with $\|\mu\|_{M'} \leq 1$. Since $\mu 1 = 1$, it follows that $\|\mu\|_{M'} = 1$.

b) Let $c \in \mathbb{R}$. Since

$$\mathscr{P}_c = \bigcap_{X \in M} \{\mu \in M' \colon \mu X \leq \mathcal{E}(X) + \alpha\}$$

is closed, and $\mathscr{P}_c \subset \{\mu \in M' \colon \|\mu\|_{M'} \leq 1\}$, we obtain that $\mathscr{P}_c$ is compact by the Banach-Alaoglu Theorem (see e.g. [69, Theorem 3.15, p. 66]). Now, let $\mu, \nu \in \mathscr{P}_c$ and $\lambda \in [0,1]$. Then, for all $X \in M$,

$$\begin{aligned}\lambda\mu X + (1-\lambda)\nu X - \mathcal{E}(X) &= \lambda\big(\mu X - \mathcal{E}(X)\big) + (1-\lambda)\big(\nu X - \mathcal{E}(X)\big)\\ &\leq \lambda\mathcal{E}^*(\mu) + (1-\lambda)\mathcal{E}^*(\nu)\\ &\leq c\end{aligned}$$

and therefore, $\lambda\mu + (1-\lambda)\nu \in \mathscr{P}_c$.

c) Let $X \in M$ and $c \geq \|X\|_\infty - \mathcal{E}(X)$. By definition of the conjugate function $\mathcal{E}^*$, we have that

$$\mathcal{E}(X) = \mu X - \big(\mu X - \mathcal{E}(X)\big) \geq \mu X - \mathcal{E}^*(\mu)$$

for all $\mu \in \mathscr{P}_c$. Let $\mathcal{E}_0(Y) := \mathcal{E}(X + Y) - \mathcal{E}(X)$ for all $Y \in M$. Then, $\mathcal{E}_0\colon M \to \mathbb{R}$ is convex and $\mathcal{E}_0(0) = 0$. By the extension theorem of Hahn-Banach (cf. [79, Theorem 18.1]), there exists a linear functional $\mu\colon M \to \mathbb{R}$ with $\mu Y \leq \mathcal{E}_0(Y)$ for all $Y \in M$. That is,

$$\mu Y - \mathcal{E}(Y) \leq \mu X - \mathcal{E}(X) =: c_0$$

for all $Y \in M$. As $\mu X - \mathcal{E}(X) = c_0$, we get that $\mathcal{E}^*(\mu) = c_0$. By part a), we thus obtain that $\mu \in M'$ with $\|\mu\|_{M'} = 1$ and therefore,

$$\mathcal{E}^*(\mu) = c_0 = \mu X - \mathcal{E}(X) \leq \|X\|_\infty - \mathcal{E}(X) \leq c.$$

Hence, $\mu \in \mathscr{P}_c$ with $\mathcal{E}(X) = \mu X - \mathcal{E}^*(\mu)$. As $0 \in M$ with $\|0\|_\infty - \mathcal{E}(0) = 0$, we thus obtain that $\mathscr{P}_0 \neq \emptyset$ and therefore $\mathscr{P}_c \neq \emptyset$ for all $c \geq 0$. Again, as $0 \in M$ with $\mathcal{E}(0) = 0$, it follows that $\mathcal{E}^*(\mu) \geq 0$ for all linear functionals $\mu\colon M \to \mathbb{R}$.

□

1.13 Corollary. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on M. Then,*

$$\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$$

is convex and every $\mu \in \mathscr{P}$ is a linear pre-expectation on M. Moreover,

$$\mathcal{E}(X) = \max_{\mu \in \mathscr{P}} \mu X - \mathcal{E}^*(\mu)$$

for all $X \in M$.

1.14 Lemma. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation and $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$.*

a) $\mathcal{E}$ is sublinear if and only if $\mathscr{P} = \{\mu \in M' \colon \mathcal{E}^*(\mu) = 0\}$.

b) $\mathcal{E}$ is linear if and only if $\#\mathscr{P} = 1$. In this case, we have that $\mathscr{P} = \{\mathcal{E}\}$.

Proof. a) First, assume that $\mathcal{E}$ is sublinear. Let $\mu \in \mathscr{P}$ and $X \in M$ be arbitrary. Then,

$$\lambda\big(\mu X - \mathcal{E}(X)\big) = \mu(\lambda X) - \mathcal{E}(\lambda X) \leq \mathcal{E}^*(\mu) < \infty$$

for all $\lambda > 0$, and therefore, $\mu X - \mathcal{E}(X) \leq 0$. Taking the supremum over all $X \in M$, we obtain that $\mathcal{E}^*(\mu) \leq 0$. By Lemma 1.12 c), this implies that $\mathcal{E}^*(\mu) = 0$.

Now assume that $\mathscr{P} = \{\mu \in M' \colon \mathcal{E}^*(\mu) = 0\}$. Then, for all $X \in M$ and all $\lambda > 0$ we have that

$$\mathcal{E}(\lambda X) = \max_{\mu \in \mathscr{P}} \mu(\lambda X) = \lambda \max_{\mu \in \mathscr{P}} \mu X = \lambda \mathcal{E}(X).$$

By Remark 1.2 g), we thus obtain that $\mathcal{E}$ is sublinear.

b) Assume that $\mathcal{E}$ is linear and let $\mu \in \mathscr{P}$. Then, by part a) we have that $\mathcal{E}^*(\mu) = 0$ and therefore, $\mu X \leq \mathcal{E}(X)$ for all $X \in M$. As $\mathcal{E}$ is linear, we thus obtain that $\mu = \mathcal{E}$. This shows that $\mathscr{P} = \{\mathcal{E}\}$.

Now let $\#\mathscr{P} = 1$ and let $\mu \in \mathscr{P}$. Then, by Corollary 1.13, we have that

$$\mathcal{E}(X) = \max_{\nu \in \mathscr{P}} \nu X - \mathcal{E}^*(\nu) = \mu X - \mathcal{E}^*(\mu)$$

for all $X \in M$. As $\mathcal{E}(0) = \mu 0 = 0$, we get that $\mathcal{E}^*(\mu) = 0$, i.e. $\mathcal{E} = \mu$ is linear.

□

The previous two lemmas have a series of consequences, which we collect in the following remark.

1.15 Remark. Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation and $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$.

a) By Lemma 1.12 b), all level sets of $\mathcal{E}^*$ are convex and therefore, $\mathcal{E}^*$ is convex. Moreover, by Lemma 1.12 c), $\mathcal{E}^*(\mu) \geq 0$ for all linear functionals $\mu\colon M \to \mathbb{R}$.

b) The mapping $\mathscr{P} \to \mathbb{R}$, $\mu \mapsto \mathcal{E}^*(\mu)$ is convex and lower semicontinuous. Indeed, as $\mathscr{P} \subset M'$ is convex, we have that, for fixed $X \in M$, the mapping

$$\mathscr{P} \to \mathbb{R}, \quad \mu \mapsto \mu X - \mathcal{E}(X)$$

is convex and continuous. Taking the pointwise supremum over all $X \in M$, we see that the mapping $\mathscr{P} \to \mathbb{R}$, $\mu \mapsto \sup_{X \in M} \mu X - \mathcal{E}(X)$ is convex and lower semicontinuous.

c) By Corollary 1.13, we have the dual representation

$$\mathcal{E}(X) = \max_{\mu \in \mathscr{P}} \mu X - \mathcal{E}^*(\mu),$$

where the set $\mathscr{P}$ is independent of X. However, in general the set $\mathscr{P}$ is not compact, which often leads to technical problems. Although the choice of the level $c \in \mathbb{R}$ in Lemma 1.12 c) heavily depends on X, it is oftentimes very useful to have a representation in terms of the convex compact set $\mathscr{P}_c$ with $c \geq 0$.

d) Assume that $\mathcal{E}$ is sublinear. Then, by Lemma 1.12 c) and Lemma 1.14 a), $\mathscr{P} = \mathscr{P}_0$ is compact. Moreover,
$$\mathcal{E}(X) = \max_{\mu \in \mathscr{P}} \mu X$$
for all $X \in M$.

e) Assume that $\mathcal{E}$ is sublinear and let $\mathcal{Q} \subset M'$ be convex and compact with
$$\mathcal{E}(X) = \sup_{\mu \in \mathcal{Q}} \mu X$$
for all $X \in M$. Then, we already have that $\mathscr{P} = \mathcal{Q}$. In fact, by definition of $\mathcal{E}^*$, we have that $\mathcal{Q} \subset \mathscr{P}$. In order to prove the other implication, let $\nu \in M' \setminus \mathcal{Q}$. Then, by the separation theorem of Hahn-Banach (cf. [79, Proposition 18.2]), there exists some $X \in M$ with
$$\mathcal{E}(X) = \sup_{\mu \in \mathscr{P}} \mu X < \nu X,$$
where we used the fact that the topological dual of M' (endowed with the weak* topology) is M. Hence, $\mathcal{E}^*(\nu) > 0$ and therefore, by Lemma 1.14 a), $\nu \notin \mathscr{P}$.

f) The previous remark and Example 1.4 imply that there exists a one to one correspondence betweeen the set of all convex compact subsets of $\mathrm{ba}^1_+(\Omega, \mathscr{F})$ and the set of all sublinear expectations. Note that, by the Banach-Alaoglu Theorem (see [69, Theorem 3.15, p. 66]), subsets of $\mathrm{ba}^1_+(\Omega, \mathscr{F})$ are compact if and only if they are closed.

g) By Lemma 1.12 a) and Corollary 1.13, we have that
$$\begin{aligned}\mathcal{E}(X+\alpha) &= \max_{\mu \in \mathscr{P}} \mu(X+\alpha) - \mathcal{E}^*(\mu) = \max_{\mu \in \mathscr{P}} \mu X + \mu\alpha - \mathcal{E}^*(\mu)\\ &= \max_{\mu \in \mathscr{P}} \mu X + \alpha - \mathcal{E}^*(\mu) = \mathcal{E}(X) + \alpha\end{aligned}$$
for all $X \in M$ and $\alpha \in \mathbb{R}$. This shows that $\mathcal{E}$ is cash additive and therefore 1-Lipschitz. In particular, $\rho(X) := \mathcal{E}(-X)$ defines a convex monetary risk measure on M (see Remark 1.2 d)). If, in addition, $\mathcal{E}$ is positive homogeneous, then ρ is a coherent monetary risk measure as introduced by Artzner et al. [4], see also Delbaen [23], [24].

h) Let $S \neq \emptyset$ be a set, $T\colon \Omega \to S$ an arbitrary mapping and assume that $\mathcal{E}$ is sublinear. Moreover, let $N := \{Y \in \mathscr{L}^\infty(S, 2^S)\colon Y \circ T \in M\}$ and $\mathcal{E} \circ T^{-1}\colon N \to \mathbb{R}$, $Y \mapsto \mathcal{E}(Y \circ T)$ the distribution of T under $\mathcal{E}$ (see Remark 1.2 i)). Then, N is a subspace of $\mathscr{L}^\infty(S, 2^S)$ with $1 \in N$ and $\mathcal{E} \circ T^{-1}$ is sublinear with
$$\big\{\nu \in N'\colon (\mathcal{E} \circ T^{-1})^*(\nu) < \infty\big\} = \{\mu \circ T^{-1}\colon \mu \in \mathscr{P}\} =: \mathscr{P} \circ T^{-1}.$$
In fact, as the mapping $M' \to N'$, $\mu \mapsto \mu \circ T^{-1}$ is continuous, part d) implies that $\mathscr{P} \circ T^{-1}$ is compact. Moreover, for all $Y \in N$ we have that
$$(\mathcal{E} \circ T^{-1})(Y) = \mathcal{E}(Y \circ T) = \max_{\mu \in \mathscr{P}} \mu(Y \circ T) = \max_{\mu \in \mathscr{P}} (\mu \circ T^{-1})Y = \max_{\nu \in \mathscr{P} \circ T^{-1}} \nu Y.$$
By part e), the assertion follows.

1.16 Proposition. *Let* $\mathcal{E}\colon M \to \mathbb{R}$ *be a convex pre-expectation and* $c > 0$. *Then, for all* $\lambda > 0$ *and all* $X, Y \in M$ *with* $2\|X\|_\infty \le c$ *and* $2\|Y\|_\infty \le c$ *it holds*

$$|\mathcal{E}(X) - \mathcal{E}(Y)| \le \frac{1}{\lambda}\mathcal{E}(\lambda|X-Y|) + \frac{c}{\lambda}.$$

Proof. Let $\lambda > 0$ and $X, Y \in M$ with $2\|X\|_\infty \le c$ and $2\|Y\|_\infty \le c$. Due to symmetry, it suffices to show that

$$\mathcal{E}(X) - \mathcal{E}(Y) \le \frac{1}{\lambda}\mathcal{E}(\lambda|X-Y|) + \frac{c}{\lambda}.$$

By Lemma 1.12 c), there exists some $\mu \in M'$ with $\mathcal{E}^*(\mu) \le \|X\|_\infty - \mathcal{E}(X) \le 2\|X\|_\infty \le c$ and $\mathcal{E}(X) = \mu X - \mathcal{E}^*(\mu)$. Hence, we obtain that

$$\begin{aligned}
\mathcal{E}(X) - \mathcal{E}(Y) &= \mu X - \mathcal{E}^*(\mu) - \mathcal{E}(Y) \le \mu X - \mu Y = \mu(X-Y) \\
&\le \mu(|X-Y|) = \frac{1}{\lambda}\mu(\lambda|X-Y|) \\
&= \frac{1}{\lambda}\big(\mu(\lambda|X-Y|) - \mathcal{E}^*(\mu)\big) + \frac{1}{\lambda}\mathcal{E}^*(\mu) \\
&\le \frac{1}{\lambda}\mathcal{E}(\lambda|X-Y|) + \frac{1}{\lambda}\mathcal{E}^*(\mu) \\
&\le \frac{1}{\lambda}\mathcal{E}(\lambda|X-Y|) + \frac{c}{\lambda}.
\end{aligned}$$

□

1.17 Lemma. *Let* $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ *be a convex expectation and* $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. *Then,*

$$|\mathcal{E}(X)| \le \mathcal{E}(|X|).$$

Proof. Since $X \le |X|$, we obtain that $\mathcal{E}(X) \le \mathcal{E}(|X|)$. By Corollary 1.13, there exists some $\mu \in \mathrm{ba}^1_+(\Omega, \mathscr{F})$ with $0 = -\mathcal{E}(0) = \mathcal{E}^*(\mu) - \mu 0 = \mathcal{E}^*(\mu)$. It follows that

$$-\mathcal{E}(X) \le -\mu X = \mu(-X) \le \mu|X| \le \mathcal{E}(|X|).$$

□

Although Proposition 1.8 implies the existence of an extension $\hat{\mathcal{E}}$ for every pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$, this extension is not neccessarily unique. However, as convex pre-expectations are 1-Lipschitz by Remark 1.15 g), the extension is uniquely determined on the closure $\overline{M}$ of M.

1.18 Proposition. *Let* $\mathcal{E}\colon M \to \mathbb{R}$ *be a convex pre-expectation. Then, there exists exactly one convex pre-expectation* $\hat{\mathcal{E}}\colon \overline{M} \to \mathbb{R}$ *with* $\hat{\mathcal{E}}|_M = \mathcal{E}$. *Here,* $\overline{M}$ *denotes the closure of* M *as a subset of* $\mathscr{L}^\infty(\Omega, \mathscr{F})$ *and* $\hat{\mathcal{E}}$ *is given as in Proposition 1.8.*

Proof. By Proposition 1.8, there exists a convex pre-expectation $\hat{\mathcal{E}}\colon \overline{M} \to \mathbb{R}$ with $\hat{\mathcal{E}}|_M = \mathcal{E}$. As $\overline{M}$ is again a vector space with $1 \in \overline{M}$, by Remark 1.15 g), every convex pre-expectation $\widetilde{\mathcal{E}}\colon \overline{M} \to \mathbb{R}$ is 1-Lipschitz and therefore uniquely determined by its values on M. □

1.19 Lemma. *Let* $N \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ *be a subspace with* $M \subset N$ *and* $\mathcal{E}\colon N \to \mathbb{R}$ *a convex pre-expectation. Then, we have that*

$$\big\{\mu \in M' \colon (\mathcal{E}|_M)^*(\mu) < \infty\big\} = \big\{\nu|_M \colon \nu \in N',\ \mathcal{E}^*(\nu) < \infty\big\}.$$

Proof. It is clear that

$$\{\mu \in M' \colon (\mathcal{E}|_M)^*(\mu) < \infty\} \subset \{\nu|_M \colon \nu \in N',\ \mathcal{E}^*(\nu) < \infty\}.$$

Therefore, let $\mu \in M'$ with $(\mathcal{E}|_M)^*(\mu) < \infty$. Then, we have that

$$\mu X \le \mathcal{E}(X) - \mathcal{E}^*(\mu)$$

for all $X \in M$. Hence, by the extension theorem of Hahn-Banach (cf. [79, Theorem 18.1]), there exists a linear functional $\nu \colon N \to \mathbb{R}$ with $\nu|_M = \mu$ and

$$\nu X \le \mathcal{E}(X) + \mathcal{E}^*(\mu)$$

for all $X \in N$. Therefore, $\mathcal{E}^*(\nu) < \infty$, which, by Lemma 1.12 a), implies that $\nu \in N'$. □

We apply Lemma 1.19 to the case $N = \mathscr{L}^\infty(\Omega, \mathscr{F})$ and obtain the following corollary.

1.20 Corollary. *Let $\mathcal{E} \colon M \to \mathbb{R}$ be a convex pre-expectation and $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$. Moreover, let $\widetilde{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be a convex expectation, which extends $\mathcal{E}$. Then, we have that*

$$\mathscr{P} = \{\nu|_M \colon \nu \in \mathrm{ba}^1_+(\Omega, \mathscr{F}),\ \widetilde{\mathcal{E}}^*(\nu) < \infty\}.$$

In view of Corollary 1.13 and Corollary 1.20, a natural approach to extend a convex pre-expectation $\mathcal{E}$ on M would be to consider $\hat{\mathscr{P}} := \{\nu \in \mathrm{ba}^1_+(\Omega, \mathscr{F}) \colon \nu|_M \in \mathscr{P}\}$ and define $\widetilde{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ by

$$\widetilde{\mathcal{E}}(X) := \sup_{\nu \in \hat{\mathscr{P}}} \nu X - \mathcal{E}^*(\nu|_M) \tag{1.3}$$

for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. By definition, $\widetilde{\mathcal{E}}|_M = \mathcal{E}$ and therefore, $\widetilde{\mathcal{E}}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$. Moreover, as $\hat{\mathscr{P}} \subset \mathrm{ba}_+(\Omega, \mathscr{F})$, it holds

$$\widetilde{\mathcal{E}}(X) = \sup_{\nu \in \hat{\mathscr{P}}} \nu X - \mathcal{E}^*(\nu|_M) \le \sup_{\nu \in \hat{\mathscr{P}}} \nu Y - \mathcal{E}^*(\nu|_M) = \widetilde{\mathcal{E}}(Y)$$

for all $X, Y \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $X \le Y$. Hence, $\widetilde{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ is an expectation with $\widetilde{\mathcal{E}}|_M = \mathcal{E}$. In the following theorem, we will prove that $\hat{\mathcal{E}} = \widetilde{\mathcal{E}}$ and that the supremum in (1.3) is attained.

1.21 Theorem. *Let $\mathcal{E} \colon M \to \mathbb{R}$ be a convex pre-expectation. Further, let*

$$\hat{\mathscr{P}} := \{\nu \in \mathrm{ba}^1_+(\Omega, \mathscr{F}) \colon \nu|_M \in \mathscr{P}\},$$

where $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$. Then, we have that $\hat{\mathscr{P}} = \{\nu \in \mathrm{ba}(\Omega, \mathscr{F}) \colon \hat{\mathcal{E}}^*(\nu) < \infty\}$. Moreover, for all $\nu \in \hat{\mathscr{P}}$ it holds $\hat{\mathcal{E}}^*(\nu) = \mathcal{E}^*(\nu|_M)$ and therefore,

$$\hat{\mathcal{E}}(X) = \max_{\nu \in \hat{\mathscr{P}}} \nu X - \mathcal{E}^*(\nu|_M) \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

Proof. Let
$$\mathcal{Q} := \{\nu \in \mathrm{ba}(\Omega, \mathscr{F})\colon \hat{\mathcal{E}}^*(\nu) < \infty\}$$
and $\widetilde{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be given by (1.3). Then, as we showed above, $\widetilde{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ is an expectation with $\widetilde{\mathcal{E}}|_M = \mathcal{E}$. By Remark 1.9, we thus obtain that

$$\nu X - \mathcal{E}^*(\nu|_M) \le \widetilde{\mathcal{E}}(X) \le \hat{\mathcal{E}}(X) \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F}))$$

for all $\nu \in \hat{\mathscr{P}}$. Therefore, $\nu \in \mathcal{Q}$ with

$$\hat{\mathcal{E}}^*(\nu) = \sup_{X \in \mathscr{L}^\infty(\Omega, \mathscr{F})} \nu X - \hat{\mathcal{E}}(X) \le \mathcal{E}^*(\nu|_M).$$

By Lemma 1.12 a), for all $\nu \in \mathcal{Q}$ it holds $\nu \in \mathrm{ba}_+^1(\Omega, \mathscr{F})$ and

$$\mathcal{E}^*(\nu|_M) = \sup_{X \in M} \nu X - \mathcal{E}(X) = \sup_{X \in M} \nu X - \hat{\mathcal{E}}(X) \le \hat{\mathcal{E}}^*(\nu),$$

i.e. $\nu \in \hat{\mathscr{P}}$ with $\mathcal{E}^*(\nu|_M) \le \hat{\mathcal{E}}^*(\nu)$. Corollary 1.13 thus implies that

$$\hat{\mathcal{E}}(X) = \max_{\nu \in \hat{\mathscr{P}}} \nu X - \mathcal{E}^*(\nu|_M) \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

□

1.22 Corollary. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a sublinear pre-expectation. Further, let*

$$\hat{\mathscr{P}} := \{\nu \in \mathrm{ba}_+^1(\Omega, \mathscr{F})\colon \nu|_M \in \mathscr{P}\},$$

where $\mathscr{P} := \{\mu \in M'\colon \mathcal{E}^*(\mu) = 0\}$. Then, we have that $\hat{\mathscr{P}} = \{\nu \in \mathrm{ba}(\Omega, \mathscr{F})\colon \hat{\mathcal{E}}^*(\nu) = 0\}$ and therefore,

$$\hat{\mathcal{E}}(X) = \max_{\nu \in \hat{\mathscr{P}}} \nu X \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

1.23 Corollary. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a sublinear pre-expectation. Then, there exists a convex compact set $\mathscr{P} \subset \mathrm{ba}_+^1(\Omega, \mathscr{F})$ such that*

$$\mathcal{E}(X) = \max_{\mu \in \mathscr{P}} \mu X$$

for all $X \in M$.

We return to the setting of Theorem 1.7. For a given linear pre-expectation $\mu\colon M \to \mathbb{R}$ on M, we consider the sublinear expectation $\hat{\mu}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ which extends μ. Then, for every $\nu \in M'$ we have that $\mu^*(\nu) = 0$ if and only if $\nu = \mu$. With the previous results, we therefore obtain the following corollary, which states that $\hat{\mu}$ is the pointwise maximum of all linear expectations $\nu \in \mathrm{ba}_+^1(\Omega, \mathscr{F})$ that extend μ.

1.24 Corollary. *Let $\mu\colon M \to \mathbb{R}$ be a linear pre-expectation on M and*

$$\hat{\mathscr{P}} := \{\nu \in \mathrm{ba}_+^1(\Omega, \mathscr{F})\colon \nu|_M = \mu\}.$$

Then, we have that $\hat{\mathscr{P}} = \{\nu \in \mathrm{ba}(\Omega, \mathscr{F})\colon \hat{\mu}^*(\nu) = 0\}$ and therefore,

$$\hat{\mu}(X) = \max_{\nu \in \hat{\mathscr{P}}} \nu X \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation and $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$. Considering the sublinear expectation $\hat{\mu}$, which extends $\mu \in \mathscr{P}$, one could also think of

$$\widetilde{\mathcal{E}}(X) := \sup_{\mu \in \mathscr{P}} \hat{\mu}(X) - \mathcal{E}^*(\mu) \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})) \tag{1.4}$$

as another possible extension of $\mathcal{E}$. Clearly, we have that $\widetilde{\mathcal{E}}|_M = \mathcal{E}$ and therefore, $\widetilde{\mathcal{E}}(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$. Moreover, as $\hat{\mu}$ is monotone for all $\mu \in \mathscr{P}$, we also have that $\widetilde{\mathcal{E}}$ is monotone. Hence, $\widetilde{\mathcal{E}}$ is an expectation which extends $\mathcal{E}$. In the following theorem, we will use Corollary 1.24 to show that the expectation $\widetilde{\mathcal{E}}$ coincides with $\hat{\mathcal{E}}$ and that the supremum in (1.4) is attained.

1.25 Theorem. *Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on M and*

$$\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}.$$

Then, for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ it holds

$$\hat{\mathcal{E}}(X) = \max_{\mu \in \mathscr{P}} \hat{\mu}(X) - \mathcal{E}^*(\mu).$$

Proof. Let $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $\widetilde{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be given by (1.4). Above, we have seen that $\widetilde{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ is an expectation that extends $\mathcal{E}$. Remark 1.9 thus yields that $\widetilde{\mathcal{E}}(X) \leq \hat{\mathcal{E}}(X)$. By Theorem 1.21, there exists a linear expectation $\nu \in \mathrm{ba}^1_+(\Omega, \mathscr{F})$ which satisfies $\mu := \nu|_M \in \mathscr{P}$ and $\hat{\mathcal{E}}(X) = \nu X - \mathcal{E}^*(\mu)$. Hence, by Corollary 1.24, we get that

$$\hat{\mathcal{E}}(X) = \nu X - \mathcal{E}^*(\mu) \leq \hat{\mu}(X) - \mathcal{E}^*(\mu) \leq \widetilde{\mathcal{E}}(X) \leq \hat{\mathcal{E}}(X).$$

This shows that $\hat{\mathcal{E}}(X) = \nu X - \mathcal{E}^*(\mu) = \hat{\mu}(X) - \mathcal{E}^*(\mu)$ and we therefore obtain the assertion. $\square$

1.26 Remark. Let $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on M. In Remark 1.11 b), Proposition 1.8, Theorem 1.21 and Theorem 1.25 we showed that the following four extension procedures all lead to the same (maximal) expectation extending $\mathcal{E}$.

(i) $\mathcal{E} \mapsto \hat{\mathcal{E}}$,

(ii) $\mathcal{E} \mapsto \mathcal{A}_{\mathcal{E}} \mapsto \widehat{\mathcal{A}}_{\mathcal{E}} \mapsto \Big[X \mapsto \sup\big\{\alpha \in \mathbb{R} \colon X - \alpha \in \widehat{\mathcal{A}}_{\mathcal{E}}\big\}\Big]$,

(iii) $\mathcal{E} \mapsto (\mathscr{P}, \mathcal{E}^*) \mapsto \big(\hat{\mathscr{P}}, [\nu \mapsto \mathcal{E}^*(\nu|_M)]\big) \mapsto \big[X \mapsto \max_{\nu \in \hat{\mathscr{P}}} \nu X - \mathcal{E}^*(\nu|_M)\big]$,

(iv) $\mathcal{E} \mapsto (\mathscr{P}, \mathcal{E}^*) \mapsto \big(\{\hat{\mu} \colon \mu \in \mathscr{P}\}, \mathcal{E}^*\big) \mapsto \big[X \mapsto \max_{\mu \in \mathscr{P}} \hat{\mu} X - \mathcal{E}^*(\mu)\big]$.

1.2 Continuity of expectations

In this section, we introduce continuity concepts for nonlinear expectations, which are a nonlinear analogon of the σ-additivity of measures. We start with the most important concepts, which are continuity from above and continuity from below, before we consider a third continuity concept which is called ca-weak lower semicontinuity. Although these concepts are equivalent for probability measures, in the nonlinear case, they lead to different properties. We will see that for convex expectations, continuity from above implies continuity from below. However, the inverse implication is not true, in general. This phenomenon is due to the direction of the inequality in the definition of convexity. Considering concave expectations, continuity from below would be the stronger continuity property.

Again, let Ω be a nonempty set and $\mathscr{F}$ a σ-algebra on Ω. For a sequence $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ and $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ we write $X_n \searrow X$ or $X_n \nearrow X$ as $n \to \infty$ if $X_n \geq X_{n+1}$ or $X_n \leq X_{n+1}$ and $X_n(\omega) \to X(\omega)$ as $n \to \infty$ for all $\omega \in \Omega$, respectively. Analogously, we write $a_n \searrow a$ or $a_n \nearrow a$ as $n \to \infty$ for a sequence $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ and $a \in \mathbb{R}$.

1.27 Definition. Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation.

a) We say that $\mathcal{E}$ is *continuous from above*, if

$$\mathcal{E}(X_n) \searrow \mathcal{E}(X)$$

as $n \to \infty$ for all $(X_n)_{n\in\mathbb{N}} \subset M$ and $X \in M$ with $X_n \searrow X$ as $n \to \infty$.

b) We say that $\mathcal{E}$ is *continuous from below*, if

$$\mathcal{E}(X_n) \nearrow \mathcal{E}(X)$$

as $n \to \infty$ for all $(X_n)_{n\in\mathbb{N}} \subset M$ and $X \in M$ with $X_n \nearrow X$ as $n \to \infty$.

1.28 Remark. a) Let Ω be a compact topological space, $\mathscr{F}$ be the Borel σ-algebra on Ω and $C_b(\Omega)$ denote the space of all continuous bounded functions $\Omega \to \mathbb{R}$. Then, by Proposition 1.16 and Dini's lemma, every convex pre-expectation $\mathcal{E}\colon C_b(\Omega) \to \mathbb{R}$ is continuous from above.

b) Let $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega, \mathscr{F})$ with $1 \in M$. Then, for any linear pre-expectation μ on M the following four statements are equivalent (cp. [7, Theorem 3.2]):

(i) μ is continuous from below,

(ii) μ is continuous from above,

(iii) $\mu X_n \searrow 0$ as $n \to \infty$ for every sequence $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow 0$ as $n \to \infty$,

(iv) μ is σ-additive, i.e. for all $(X_n)_{n\in\mathbb{N}}$ with $X_n \geq 0$ for all $n \in \mathbb{N}$ and $\sum_{n\in\mathbb{N}} X_n \in M$ it holds

$$\mu\left(\sum_{n\in\mathbb{N}} X_n\right) = \sum_{n\in\mathbb{N}} \mu X_n,$$

where $\sum_{n\in\mathbb{N}} X_n$ is to be understood pointwise.

We denote by $\mathrm{ca}_+^1(M)$ the space of all linear pre-expectations on M which are continuous from above. Then, the Daniell-Stone theorem (cf. [12, Theorem 7.8.1] or [32, Theorem 4.5.2]) implies that the mapping

$$\mathrm{ca}_+^1(\Omega,\sigma(M)) \to \mathrm{ca}_+^1(M), \quad \nu \mapsto \nu|_M$$

is bijective.

c) Let $M \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $1 \in M$, $\mathcal{E}\colon M \to \mathbb{R}$ be a convex pre-expectation on M and $\mathscr{P} := \{\mu \in M'\colon \mathcal{E}^*(\mu) < \infty\}$. Assume that there exists a set $\mathcal{Q} \subset \mathscr{P}$, such that every linear pre-expectation $\mu \in \mathcal{Q}$ is continuous from above and

$$\mathcal{E}(X) = \sup_{\mu\in\mathcal{Q}} \mu X - \mathcal{E}^*(\mu)$$

for all $X \in M$. Then, $\mathcal{E}$ is continuous from below. In fact, let $(X_n)_{n\in\mathbb{N}} \subset M$ and $X \in M$ with $X_n \nearrow X$ as $n \to \infty$. Then,

$$\mathcal{E}(X) = \sup_{\mu\in\mathcal{Q}} \mu X - \mathcal{E}^*(\mu) = \sup_{\mu\in\mathcal{Q}}\sup_{n\in\mathbb{N}} \mu X_n - \mathcal{E}^*(\mu) = \sup_{n\in\mathbb{N}}\sup_{\mu\in\mathcal{Q}} \mu X_n - \mathcal{E}^*(\mu) = \sup_{n\in\mathbb{N}} \mathcal{E}(X_n).$$

Although continuity from above and continuity from below are equivalent in the linear case, already in the sublinear case there is a difference between these concepts as the following example illustrates.

1.29 Example. Let $\Omega := [0,1]$, $\mathscr{F} := \mathscr{B}([0,1])$ be the Borel σ-algebra on $[0,1]$ and

$$\mathcal{E}(X) := \sup_{\omega\in\Omega} X(\omega) = \sup_{\omega\in\Omega} \delta_\omega X$$

for all $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$, where $\delta_\omega \in \mathrm{ca}_+^1(\Omega,\mathscr{F})$ is given by

$$\delta_\omega(A) := \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A \end{cases} \qquad (A \in \mathscr{F})$$

for all $\omega \in \Omega$. Then, by Remark 1.28 c), the sublinear expectation $\mathcal{E}$ is continuous from below. Let $X_n := 1_{A_n}$ with $A_n := \left\{\frac{1}{k}\colon k \in \mathbb{N},\ k \geq n\right\}$ for all $n \in \mathbb{N}$. Then, we have that $X_n \searrow 0$ as $n \to \infty$. Since $\mathcal{E}(0) = 0$ and

$$\mathcal{E}(X_n) = 1$$

for all $n \in \mathbb{N}$, we see that $\mathcal{E}$ is not continuous from above.

1.30 Remark. The previous example shows that an expectation which is continuous from below, in general, needs not be continuous from above. However, every convex expectation, which is continuous from above, is already continuous from below, as the following lemma shows. In the proof, we apply a minimax theorem due to Fan [37]. In [37], a function $f\colon E \times F \to \mathbb{R}$ defined on arbitrary sets E and F is said to be convex on F, if for all $y_1, y_2 \in F$ and $\lambda \in [0,1]$ there exists an element $y_0 \in F$ such that

$$f(x,y_0) \leq \lambda f(x,y_1) + (1-\lambda) f(x,y_2) \quad (x \in E).$$

Analogously, concavity on E is defined. The minimax theorem by Fan ([37, Theorem 2]) states that

$$\max_{x\in E}\inf_{y\in F} f(x,y)=\inf_{y\in F}\max_{x\in E} f(x,y)$$

holds if E is a compact Hausdorff space, $f(\,\cdot\,,y)$ is upper semicontinuous (u.s.c.) on E for each $y\in F$, and f is convex on F and concave on E.

1.31 Lemma. *Let $M\subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $1\in M$, $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation and $\mathscr{P}:=\{\mu\in M'\colon \mathcal{E}^*(\mu)<\infty\}$. Then every $\mu\in\mathscr{P}$ is a linear pre-expectation on M and $\mathcal{E}$ is continuous from above if and only if every $\mu\in\mathscr{P}$ is continuous from above. In this case, $\mathcal{E}$ is already continuous from below.*

Proof. By Lemma 1.12 a) and Remark 1.28 c), it remains to show that $\mathcal{E}$ is continuous from above if and only if every $\mu\in\mathscr{P}$ is continuous from above. First, assume that $\mathcal{E}$ is continuous from above and let $\mu\in\mathscr{P}$. Let $(X_n)_{n\in\mathbb{N}}\subset M$ and $X\in M$ with $X_n\searrow X$ as $n\to\infty$. Then,

$$\begin{aligned}0&\le \mu X_n-\mu X=\mu(X_n-X)=\lambda^{-1}\mu(\lambda(X_n-X))\\&\le \lambda^{-1}\mathcal{E}(\lambda(X_n-X))+\lambda^{-1}\mathcal{E}^*(\mu)\searrow\lambda^{-1}\mathcal{E}^*(\mu),\quad n\to\infty\end{aligned}$$

for all $\lambda>0$. Letting $\lambda\to\infty$, we thus obtain that $\inf_{n\in\mathbb{N}}\mu X_n=\mu X$.

Now, assume that every $\mu\in\mathscr{P}$ is continuous from above, and let $(X_n)_{n\in\mathbb{N}}\subset M$ and $X\in M$ with $X_n\searrow X$ as $n\to\infty$. Let $c:=2\max\{\|X\|_\infty,\|X_1\|_\infty\}$ and

$$\mathscr{P}_c:=\{\mu\in M'\colon \mathcal{E}^*(\mu)\le c\}.$$

Then, we have that $\|X\|_\infty-\mathcal{E}(X)\le c$ and, as $X\le X_n\le X_1$ for all $n\in\mathbb{N}$, we have that $\|X_n\|_\infty-\mathcal{E}(X_n)\le 2\|X_n\|_\infty\le c$ for all $n\in\mathbb{N}$. Therefore, by Lemma 1.12 c), we have that

$$\mathcal{E}(X)=\max_{\mu\in\mathscr{P}_c}\mu X-\mathcal{E}^*(\mu)\quad\text{and}\quad \mathcal{E}(X_n)=\max_{\mu\in\mathscr{P}_c}\mu X_n-\mathcal{E}^*(\mu)\tag{1.5}$$

for all $n\in\mathbb{N}$. Consider the mapping $f\colon\mathscr{P}_c\times\mathbb{N}\to\mathbb{R}$, $(\mu,n)\mapsto\mu X_n-\mathcal{E}^*(\mu)$. Then, by Remark 1.15 b), for every $n\in\mathbb{N}$ the mapping $f(\,\cdot\,,n)$ is u.s.c. on $\mathscr{P}_c$. As $\mathscr{P}_c$ is convex, f is concave on $\mathscr{P}_c$ by Remark 1.15 b) and, as every $\mu\in\mathscr{P}_c$ is monotone, f is convex on $\mathbb{N}$. As M' is a Hausdorff space and $\mathscr{P}_c\subset M'$ is compact by Lemma 1.12 c), Fan's minimax theorem together with (1.5) yields that

$$\begin{aligned}\mathcal{E}(X)&=\max_{\mu\in\mathscr{P}_c}\mu X-\mathcal{E}^*(\mu)=\max_{\mu\in\mathscr{P}_c}\inf_{n\in\mathbb{N}}\mu X_n-\mathcal{E}^*(\mu)\\&=\inf_{n\in\mathbb{N}}\max_{\mu\in\mathscr{P}_c}\mu X_n-\mathcal{E}^*(\mu)=\inf_{n\in\mathbb{N}}\mathcal{E}(X_n).\end{aligned}$$

□

1.32 Corollary. *Let $M\subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ be a linear subspace of $\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $1\in M$ and $\mathcal{E}\colon M\to\mathbb{R}$ be a convex pre-expectation on M. Then, $\mathcal{E}$ is continuous from above if and only if $\mathcal{E}(X_n)\searrow 0$ as $n\to\infty$ for every sequence $(X_n)_{n\in\mathbb{N}}\subset M$ with $X_n\searrow 0$ as $n\to\infty$.*

Proof. Assume that $\mathcal{E}(X_n) \searrow 0$ as $n \to \infty$ for every sequence $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow 0$ as $n \to \infty$. Let $(X_n)_{n\in\mathbb{N}} \subset M$ be such a sequence. Further, let $\mu \in M'$ with $\mathcal{E}^*(\mu) < \infty$ and $\lambda > 0$. Then, we have that

$$0 \leq \mu X_n = \lambda^{-1}\big(\mu(\lambda X_n) - \mathcal{E}^*(\mu)\big) + \lambda^{-1}\mathcal{E}^*(\mu) \leq \lambda^{-1}\mathcal{E}(\lambda X_n) + \lambda^{-1}\mathcal{E}^*(\mu) \searrow \lambda^{-1}\mathcal{E}^*(\mu)$$

as $n \to \infty$. Letting $\lambda \to \infty$, we obtain that $\mu X_n \searrow 0$ as $n \to \infty$. Hence, by Remark 1.28 b), μ is continuous from above and therefore, we obtain that $\mathcal{E}$ is continuous from above by Lemma 1.31. The other implication is trivial. □

Part a) and b) of the following remark are the nonlinear analogon of Fatou's lemma and Lebesgue's dominated convergence theorem, respectively. Both statements are well-known facts for convex monetary risk measures (cf. [43, Theorem 4.31]) and carry over to expectations.

1.33 Remark. a) Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ be an expectation. Then, $\mathcal{E}$ is continuous from below if and only if $\mathcal{E}$ satisfies the following *Fatou property*: If $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ is a bounded sequence which converges pointwise to $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$, then

$$\mathcal{E}(X) \leq \liminf_{n\to\infty} \mathcal{E}(X_n).$$

In fact, let $\mathcal{E}$ be continuous from below and $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ a bounded sequence which converges pointwise to $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. Then, as $\mathcal{E}$ is continuous from below, we have that

$$\mathcal{E}(X) = \mathcal{E}\left(\sup_{n\in\mathbb{N}}\inf_{k\geq n} X_k\right) = \sup_{n\in\mathbb{N}}\mathcal{E}\left(\inf_{k\geq n} X_k\right) \leq \sup_{n\in\mathbb{N}}\inf_{k\geq n}\mathcal{E}(X_k) = \liminf_{n\to\infty}\mathcal{E}(X_n).$$

Conversely, assume that $\mathcal{E}$ satisfies the Fatou property and let $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ and $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X_n \nearrow X$ as $n \to \infty$. Then, $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ is bounded and converges pointwise to X. Therefore, by monotonicity of $\mathcal{E}$, we obtain that

$$\mathcal{E}(X) \leq \liminf_{n\to\infty}\mathcal{E}(X_n) = \lim_{n\to\infty}\mathcal{E}(X_n) \leq \mathcal{E}(X),$$

i.e. $\mathcal{E}(X) = \lim_{n\to\infty}\mathcal{E}(X_n)$.

b) Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ be a convex expectation. Then, $\mathcal{E}$ is continuous from above if and only if $\mathcal{E}$ satisfies the following *Lebesgue property*: If $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ is a bounded sequence which converges pointwise to $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$, then

$$\mathcal{E}(X) = \lim_{n\to\infty}\mathcal{E}(X_n).$$

In fact, let $\mathcal{E}$ be continuous from above and $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ a bounded sequence which converges pointwise to $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. By Lemma 1.31, $\mathcal{E}$ is continuous from below and therefore satisfies the Fatou property. Hence,

$$\mathcal{E}(X) \leq \liminf_{n\to\infty}\mathcal{E}(X_n).$$

On the other hand, as $\mathcal{E}$ is continuous from above,

$$\mathcal{E}(X) = \mathcal{E}\left(\inf_{n\in\mathbb{N}}\sup_{k\geq n} X_k\right) = \inf_{n\in\mathbb{N}}\mathcal{E}\left(\sup_{k\geq n} X_k\right) \geq \inf_{n\in\mathbb{N}}\sup_{k\geq n}\mathcal{E}(X_k) = \limsup_{n\to\infty}\mathcal{E}(X_n).$$

This shows that $\mathcal{E}(X) = \lim_{n\to\infty} \mathcal{E}(X)$.

Now, assume that $\mathcal{E}$ satisfies the Lebesgue property and let $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $X_n \searrow X$ as $n \to \infty$. Then, $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ is bounded and converges pointwise to X. Therefore,

$$\mathcal{E}(X) = \lim_{n\to\infty} \mathcal{E}(X_n).$$

In Remark 1.33, the Fatou property and the Lebesgue property both involve the pointwise convergence of bounded sequences of $\mathscr{F}$-$\mathscr{B}(\mathbb{R})$-measurable random variables. We now characterize this convergence by means of a topological convergence, which provides a topological interpretation of the Fatou property and the Lebesgue property. We start with a short remark on weak topologies.

1.34 Remark. Given a vector space E over the field of real numbers $\mathbb{R}$, we denote by E^* the algebraic dual space, i.e. the space of all linear functionals $E \to \mathbb{R}$. For a subset $F \subset E^*$ we denote by $\sigma(E, F)$ the smallest topology on E such that all linear functionals $f \in F$ are continuous. The topology $\sigma(E, F)$ is called the *F-weak topology on E*. For a more detailed discussion of weak topologies we refer to [49, Section 62], [74, Chapter II, Section 5, p. 52] or [81, Definition VIII.3.2]. We focus on the following weak topologies:

- $\sigma(\mathscr{L}^\infty(\Omega, \mathscr{F}), \mathrm{ca}(\Omega, \mathscr{F}))$: The smallest topology, such that for all $\mu \in \mathrm{ca}(\Omega, \mathscr{F})$ the mapping
$$\mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}, \quad X \mapsto \int_\Omega X \,\mathrm{d}\mu$$
is continuous.

- $\sigma(\mathrm{ca}(\Omega, \mathscr{F}), \mathscr{L}^\infty(\Omega, \mathscr{F}))$: The smallest topology, such that for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ the mapping
$$\mathrm{ca}(\Omega, \mathscr{F}) \to \mathbb{R}, \quad \mu \mapsto \int_\Omega X \,\mathrm{d}\mu$$
is continuous. Note that this is the trace topology of the weak* topology on $\mathrm{ba}(\Omega, \mathscr{F})$.

- $\sigma(\mathrm{ca}(\Omega, \mathscr{F}), C_b(\Omega))$ for a topological space Ω with Borel σ-algebra $\mathscr{F}$: The smallest topology, such that for all $X \in C_b(\Omega)$ the mapping
$$\mathrm{ca}(\Omega, \mathscr{F}) \to \mathbb{R}, \quad \mu \mapsto \int_\Omega X \,\mathrm{d}\mu$$
is continuous. Here, $C_b(\Omega)$ denotes the space of all bounded continuous functions $\Omega \to \mathbb{R}$. Note that this is the topology that induces the weak convergence of measures.

1.35 Lemma. *Let $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ and $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. Then, the following statements are equivalent:*

(i) $(X_n)_{n\in\mathbb{N}}$ is a bounded sequence that converges pointwise to X,

(ii) $\mu X_n \to \mu X$ as $n \to \infty$ for all $\mu \in \mathrm{ca}(\Omega, \mathscr{F})$, i.e. $(X_n)_{n\in\mathbb{N}}$ converges to X in the topology $\sigma(\mathscr{L}^\infty(\Omega, \mathscr{F}), \mathrm{ca}(\Omega, \mathscr{F}))$.

Proof. First, assume that $(X_n)_{n\in\mathbb{N}}$ is a bounded sequence which converges pointwise to X. Then, by the dominated convergence theorem, we get that $\mu X_n \to \mu X$ as $n\to\infty$ for all $\mu \in \mathrm{ca}(\Omega,\mathscr{F})$.

Now, assume that $\mu X_n \to \mu X$ as $n\to\infty$ for all $\mu \in \mathrm{ca}(\Omega,\mathscr{F})$. Then, we have that

$$X_n(\omega) = \delta_\omega X_n \to \delta_\omega X = X(\omega), \quad n\to\infty$$

for all $\omega\in\Omega$, i.e. $(X_n)_{n\in\mathbb{N}}$ converges pointwise to X. Moreover, we have that $(\mu X_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ is bounded for every $\mu \in \mathrm{ca}(\Omega,\mathscr{F})$. As $\mathrm{ca}(\Omega,\mathscr{F})$ is a Banach space and

$$\sup_{\mu\in\mathrm{ca}(\Omega,\mathscr{F})} \frac{|\mu X_n|}{\|\mu\|_{\mathrm{ba}}} = \|X_n\|_\infty$$

for all $n\in\mathbb{N}$, the theorem of Banach-Steinhaus (see [71, Theorem 5.8]) implies that

$$\sup_{n\in\mathbb{N}} \|X_n\|_\infty < \infty,$$

i.e. $(X_n)_{n\in\mathbb{N}}$ is bounded. Here, $\|\mu\|_{\mathrm{ba}}$ denotes the total variation of $\mu\in\mathrm{ba}(\Omega,\mathscr{F})$. □

1.36 Corollary. *Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be an expectation. Then, the following statements are equivalent:*

(i) $\mathcal{E}$ is continuous from below,

(ii) $\mathcal{E}$ satisfies the Fatou property,

(iii) $\mathcal{E}$ is $\sigma(\mathscr{L}^\infty(\Omega,\mathscr{F}),\mathrm{ca}(\Omega,\mathscr{F}))$-sequentially lower semicontinuous.

1.37 Corollary. *Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a convex expectation. Then, the following statements are equivalent:*

(i) $\mathcal{E}$ is continuous from above,

(ii) $\mathcal{E}$ satisfies the Lebesgue property,

(iii) $\mathcal{E}$ is $\sigma(\mathscr{L}^\infty(\Omega,\mathscr{F}),\mathrm{ca}(\Omega,\mathscr{F}))$-sequentially continuous.

The following lemma is inspired by [19, Extensions du théorème 1, 1) a)]. It states that, once an expectation $\mathcal{E}$ is continuous from below, it can be extended to $\mathscr{L}^\infty(\Omega,2^\Omega)$ via $\hat{\mathcal{E}}$ (see Proposition 1.8), maintaining the continuity from below.

1.38 Lemma. *Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be an expectation which is continuous from below. Then, $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega,2^\Omega)\to\mathbb{R}$ is continuous from below, as well.*

Proof. Let $X\in\mathscr{L}^\infty(\Omega,2^\Omega)$ and $(X_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{L}^\infty(\Omega,2^\Omega)$ with $X_n \nearrow X$ as $n\to\infty$. Fix $\varepsilon>0$. Then, for every $n\in\mathbb{N}$, there exists $X_0^n\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X_n\le X_0^n\le\|X\|_\infty$ and

$$\mathcal{E}(X_0^n)\le\hat{\mathcal{E}}(X_n)+\varepsilon$$

for all $n\in\mathbb{N}$. Define $Y_n := \inf_{k\ge n} X_0^k$. Then, $Y_n\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X_n\le Y_n\le Y_{n+1}\le\|X\|_\infty$ and

$$\mathcal{E}(Y_n)\le\mathcal{E}(X_0^n)\le\hat{\mathcal{E}}(X_n)+\varepsilon$$

for all $n \in \mathbb{N}$. As $X_n \leq Y_n \leq \|X\|_\infty$ for all $n \in \mathbb{N}$, we get that $Y := \sup_{n\in\mathbb{N}} Y_n \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $X = \sup_{n\in\mathbb{N}} X_n \leq \sup_{n\in\mathbb{N}} Y_n = Y$ and $Y_n \nearrow Y$. Since $\mathcal{E}$ is continuous from below, we obtain that

$$\hat{\mathcal{E}}(X) \leq \mathcal{E}(Y) = \lim_{n\to\infty} \mathcal{E}(Y_n) \leq \lim_{n\to\infty} \hat{\mathcal{E}}(X_n) + \varepsilon.$$

Letting $\varepsilon \searrow 0$, we obtain that $\hat{\mathcal{E}}(X) \leq \lim_{n\to\infty} \hat{\mathcal{E}}(X_n)$, and therefore, $\hat{\mathcal{E}}(X) = \lim_{n\to\infty} \hat{\mathcal{E}}(X_n)$. □

In Remark 1.28 c), we saw that a sublinear expectation $\mathcal{E}$ that admits a representation in terms of probability measures is continuous from below. A natural question therefore is, if the inverse implication also holds, i.e. if every sublinear expectation $\mathcal{E}$, which is continuous from below, admits a representation by probability measures. By Corollary 1.36, an equivalent formulation of this question is, if sequential lower semicontinuity implies lower semicontinuity in the topology $\sigma(\mathscr{L}^\infty(\Omega, \mathscr{F}), \mathrm{ca}(\Omega, \mathscr{F}))$. If $\mathcal{E}$ is continuous from below and dominated by some probability measure ν, i.e. $\mathcal{E}(X) = \mathcal{E}(Y)$ for all $X, Y \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $X = Y$ ν-almost surely, then, it is well-known that $\mathcal{E}$ can even be represented by probability measures, which are absolutely continuous w.r.t. ν (cf. [43], Theorem 4.33). The following example shows that, without a dominating probability measure, such a representation does not neccesarily exist. We will see that for a sublinear expectation $\mathcal{E}$, which is continuous from below, in general there exists not even one probability measure μ with $\mathcal{E}^*(\mu) < \infty$. In particular, not every sequentially closed convex cone, which contains 0, is closed in the topology $\sigma(\mathscr{L}^\infty(\Omega, \mathscr{F}), \mathrm{ca}(\Omega, \mathscr{F}))$.

1.39 Example. Let Ω be a set of cardinality $|\Omega| = \aleph_1$ (see e.g. [13, p. 242]). Let

$$\mathcal{A} := \{A \in 2^\Omega \colon |A| \leq \aleph_0 \text{ or } |\Omega \setminus A| \leq \aleph_0\}$$

and

$$\lambda(A) := \begin{cases} 0, & |A| \leq \aleph_0, \\ 1, & |\Omega \setminus A| \leq \aleph_0 \end{cases}$$

for all $A \in \mathcal{A}$. Then, $(\Omega, \mathcal{A}, \lambda)$ is a probability space and by Proposition 1.8 and Lemma 1.38, $\hat{\lambda} \colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ is a sublinear expectation which is continuous from below and extends λ. However, by a result due to Bierlein [9, Satz 1C], there exists no $\mu \in \mathrm{ca}_+^1(\Omega, 2^\Omega)$ with $\mu|_\mathcal{A} = \lambda$. Hence, by Corollary 1.24, there exists no $\mu \in \mathrm{ca}_+^1(\Omega, 2^\Omega)$ with $\hat{\lambda}^*(\mu) < \infty$. Assuming the continuum hypothesis, we may choose $\Omega = [0, 1]$. In this case $\lambda \colon \mathcal{A} \to \mathbb{R}$ is the restriction of the Lebesgue measure to $\mathcal{A}$.

For two probability measures $\mu, \nu \in \mathrm{ca}_+^1(\Omega, \mathscr{F})$ one says that μ is *absolutely continuous* w.r.t. ν if every ν-null set is a μ-null set. We denote by $\mathrm{ca}_+^1(\Omega, \mathscr{F}, \nu)$ the set of all probability measures, which are absolutely continuous w.r.t. $\nu \in \mathrm{ca}_+^1(\Omega, \mathscr{F})$. It is well-known that μ is absolutely continuous w.r.t. ν if and only if for all $\varepsilon > 0$ there exists some $\delta > 0$, such that for all $A \in \mathscr{F}$ with $\nu(A) \leq \delta$ we have that $\mu(A) \leq \varepsilon$. This motivates the following definition.

1.40 Definition. Let $\mathcal{E} \colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be a convex pre-expectation and $\nu \in \mathrm{ca}_+^1(\Omega, \mathscr{F})$ be a probability measure. Then, we say that $\mathcal{E}$ is *absolutely continuous* w.r.t. ν if for all $c \geq 0$ and all $\varepsilon > 0$ there exists some $\delta > 0$ such that for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $0 \leq X \leq c$ and $\nu X \leq \delta$ we have that $\mathcal{E}(X) \leq \varepsilon$.

1.41 Lemma. *Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a convex expectation, $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})$ be a probability measure and $\mathscr{P}:=\{\mu\in \mathrm{ba}(\Omega,\mathscr{F})\colon \mathcal{E}^*(\mu)<\infty\}$. Moreover, let $\mathscr{P}_c:=\{\mu\in M'\colon \mathcal{E}^*(\mu)\le c\}$ for all $c\ge 0$. Then, the following statements are equivalent:*

(i) $\mathcal{E}$ is absolutely continuous w.r.t. ν,

(ii) For all $c\ge 0$ and all $\varepsilon>0$ there exists some $\delta>0$, such that $\sup_{\mu\in\mathscr{P}_c}\mu(A)\le\varepsilon$ for all $A\in\mathscr{F}$ with $\nu(A)\le\delta$,

(iii) $\mathscr{P}\subset \mathrm{ca}_+^1(\Omega,\mathscr{F},\nu)$ and for all $c\ge 0$ the set $\big\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\colon \mu\in\mathscr{P}_c\big\}$ is uniformly integrable w.r.t. ν.

As $\mathscr{P}_c\subset\mathscr{P}_d$ for all $0\le c\le d$, one may replace $c\ge 0$ by $c\in\mathbb{N}$ in (ii) and (iii).

Proof. First assume that $\mathcal{E}$ is absolutely continuous w.r.t. ν and let $c\ge 0$. Let $\varepsilon>0$ and $n\in\mathbb{N}$ such that $\frac{c}{n}\le\frac{\varepsilon}{2}$. Then, there exists some $\delta_n>0$ such that for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $0\le X\le n$ and $\nu(X)\le\delta_n$ we have that $\mathcal{E}(X)\le\frac{n\varepsilon}{2}$. Let $\delta:=\frac{\delta_n}{n}$ and $A\in\mathscr{F}$ with $\nu(A)\le\delta$. Then, we have that $\nu(n1_A)=n\nu(A)\le\delta_n$ and therefore, $\mathcal{E}(n1_A)\le\frac{n\varepsilon}{2}$. It follows that

$$\mu(A)=\frac{1}{n}\mu(n1_A)\le\frac{1}{n}\big(\mathcal{E}(n1_A)+\mathcal{E}^*(\mu)\big)\le\frac{\varepsilon}{2}+\frac{c}{n}\le\varepsilon$$

for all $\mu\in\mathscr{P}_c$.

Now, assume that for all $c\ge 0$ and all $\varepsilon>0$ there exists some $\delta>0$ such that $\sup_{\mu\in\mathscr{P}_c}\mu(A)\le\varepsilon$ for all $A\in\mathscr{F}$ with $\nu(A)\le\delta$. Then, by [11, Proposition 4.5.3, p. 267], $\mathscr{P}\subset\mathrm{ca}_+^1(\Omega,\mathscr{F},\nu)$ and $\big\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\colon\mu\in\mathscr{P}_c\big\}$ is uniformly integrable w.r.t. ν for all $c\ge 0$.

Finally, assume that $\mathscr{P}\subset\mathrm{ca}_+^1(\Omega,\mathscr{F},\nu)$ and that $\big\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\colon\mu\in\mathscr{P}_c\big\}$ is uniformly integrable w.r.t. ν for all $c\ge 0$. Let $c\ge 0$ and $\varepsilon>0$. Then, there exists some $N\in\mathbb{N}$ such that

$$\sup_{\mu\in\mathscr{P}_c}\mu\left(\left\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}>N\right\}\right)\le\frac{\varepsilon}{2}.$$

Let $\delta:=\frac{\varepsilon}{2N}$. Thus, by Lemma 1.12 c), for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $0\le X\le c$ and $\nu X\le\delta$ we have that

$$\mathcal{E}(X)=\sup_{\mu\in\mathscr{P}_c}\mu X-\mathcal{E}^*(\mu)\le\sup_{\mu\in\mathscr{P}_c}\mu X\le N\nu X+\sup_{\mu\in\mathscr{P}_c}\mu\left(\left\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}>N\right\}\right)\le N\nu X+\frac{\varepsilon}{2}\le\varepsilon.$$

□

Using Lemma 1.41, we can now give a full characterization of the continuity from above.

1.42 Proposition. *Let $\mathcal{E}\colon\mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a convex expectation and*

$$\mathscr{P}:=\{\mu\in\mathrm{ba}(\Omega,\mathscr{F})\colon\mathcal{E}^*(\mu)<\infty\}.$$

Then, the following statements are equivalent:

(i) $\mathcal{E}$ is continuous from above,

(ii) For all $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X_n \searrow 0$ as $n\to\infty$ we have that $\mathcal{E}(X_n)\searrow 0$ as $n\to\infty$,

(iii) $\mathcal{E}$ satisfies the Lebesgue property,

(iv) $\mathcal{E}$ is $\sigma(\mathscr{L}^\infty(\Omega,\mathscr{F}), \mathrm{ca}(\Omega,\mathscr{F}))$-sequentially continuous,

(v) There exists some $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})\cap\mathscr{P}$ such that $\mathcal{E}$ is absolutely continuous w.r.t. ν,

(vi) There exists some $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})$ such that $\mathcal{E}$ is absolutely continuous w.r.t. ν.

Proof. By Corollary 1.32 and Corollary 1.37, we already have the equivalence of the statements (i) - (iv). Assume that $\mathcal{E}$ is continuous from above. Then, Lemma 1.12 and Lemma 1.31 imply that

$$\mathscr{P}_n := \{\mu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})\colon \mathcal{E}^*(\mu)\le n\}$$

is a convex compact subset of $\mathrm{ba}(\Omega,\mathscr{F})$ and therefore, $\mathscr{P}_n$ is countably convex for all $n\in\mathbb{N}$. Hence, for all $n\in\mathbb{N}$ there exists some probability measure $\nu_n\in \mathrm{ca}_+^1(\Omega,\mathscr{F})\cap\mathscr{P}_n$ such that all $\mu\in\mathscr{P}_n$ are ν_n-continuous and the family $\big\{\frac{\mathrm{d}\mu}{\mathrm{d}\nu_n}\colon \mu\in\mathscr{P}_n\big\}$ is uniformly integrable w.r.t. the probability measure ν_n (cf. [11, Theorem 4.7.25, p. 291]). Let $\nu := \sum_{n=1}^\infty 2^{-n}\nu_n \in \mathrm{ca}_+^1(\Omega,\mathscr{F})$. As $\nu_n\in\mathscr{P}_n$ for all $n\in\mathbb{N}$, we have that

$$\nu X - \mathcal{E}(X) = \sum_{n=1}^\infty 2^{-n}\big(\nu_n X - \mathcal{E}(X)\big) \le \sum_{n=1}^\infty 2^{-n} n = 4 < \infty$$

for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. This shows that $\nu\in\mathscr{P}$. It remains to show that $\mathcal{E}$ is absolutely continuous w.r.t. ν. Let $n\in\mathbb{N}$ and $\varepsilon>0$. Then, again by [11, Proposition 4.5.3, p. 267], there exists some $\delta>0$ such that $\sup_{\mu\in\mathscr{P}_n}\mu(A)\le\varepsilon$ for all $A\in\mathscr{F}$ with $\nu_n(A)\le\delta$. Let $A\in\mathscr{F}$ with $\nu(A)\le 2^{-n}\delta$. Then, we have that $\nu_n(A)\le 2^n\nu(A)\le\delta$ and therefore,

$$\sup_{\mu\in\mathscr{P}_n}\mu(A)\le\varepsilon.$$

By Lemma 1.41, we thus obtain that $\mathcal{E}$ is absolutely continuous w.r.t. ν.

It remains to show that statement (vi) implies statement (ii). To show this, let $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})$ such that $\mathcal{E}$ is absolutely continuous w.r.t. ν. Let $(X_n)_{n\in\mathbb{N}}\subset\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X_n\searrow 0$ as $n\to\infty$. Then, we have that $0\le X_n\le \|X_1\|_\infty =: c$ for all $n\in\mathbb{N}$. Moreover, as $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})$, we have that $\nu X_n\searrow 0$ as $n\to\infty$. Let $\varepsilon>0$. As $\mathcal{E}$ is absolutely continuous w.r.t. ν, there exists some $\delta>0$ such that $\mathcal{E}(X)\le\varepsilon$ for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $0\le X\le c$ and $\nu X\le\delta$. Since $\nu X_n\searrow 0$ as $n\to\infty$, there exists some $n_0\in\mathbb{N}$ such that $\nu X_n\le\delta$ and therefore, $0\le\mathcal{E}(X_n)\le\varepsilon$ for all $n\in\mathbb{N}$ with $n\ge n_0$. Thus, we have that $\mathcal{E}(X_n)\searrow 0$ as $n\to\infty$. □

1.43 Remark. Let $\mathcal{E}\colon\mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a convex expectation which is continuous from above. Then, by Theorem 1.42, there exists a probability measure $\nu\in \mathrm{ca}_+^1(\Omega,\mathscr{F})$ such that $\mathscr{P}\subset \mathrm{ca}_+^1(\Omega,\mathscr{F},\nu)$. In particular, by Corollary 1.13, we get that

$$\mathcal{E}(X) = \max_{\mu\in\mathscr{P}}\mu X - \mathcal{E}^*(\mu) = \sup_{\mu\in\mathscr{P}}\mu Y - \mathcal{E}^*(\mu) = \mathcal{E}(Y)$$

for all $X,Y\in\mathscr{L}^\infty(\Omega,\mathscr{F})$ with $X=Y$ ν-a.s. That is, $\mathcal{E}$ is *dominated* by ν.

By Remark 1.28 c), every convex expectation $\mathcal{E}$, which is continuous from above, admits a representation in terms of probability measures, which in turn implies that $\mathcal{E}$ is continuous from below. However, as Example 1.39 shows, a convex expectation, which is continuous from below, in general does not admit a representation in terms of probability measures, i.e. a convex expectation, which is continuous from below, is not neccesarily lower semicontinuous in the topology $\sigma(\mathscr{L}^\infty(\Omega,\mathscr{F}),\mathrm{ca}(\Omega,\mathscr{F}))$. Looking for an analogon of σ-additivity for convex expectations, regarding Lemma 1.38, continuity from below seems to be too weak as it automatically carries over to $\mathscr{L}^\infty(\Omega,2^\Omega)$. On the other hand, in view of Remark 1.43, continuity from above seems to be too strong as it automatically implies that the expectation $\mathcal{E}$ is dominated by some probability measure, which is too restrivtive in many applications. This motivates the following definition.

1.44 Definition. Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be an expectation. Then, we say that $\mathcal{E}$ is ca-*weakly lower semicontinuous* if $\mathcal{E}$ is lower semicontinuous w.r.t. the topology $\sigma(\mathscr{L}^\infty(\Omega,\mathscr{F}),\mathrm{ca}(\Omega,\mathscr{F}))$. In this case, we say that $(\Omega,\mathscr{F},\mathcal{E})$ is a *(nonlinear) expectation space*. We say that $(\Omega,\mathscr{F},\mathcal{E})$ is *convex* or *sublinear* if $\mathcal{E}$ is convex or sublinear, respectively.

1.45 Remark. Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be an expectation.

a) Let $(\Omega,\mathscr{F},\mathcal{E})$ be an expectation space. Then, by Corollary 1.36, $\mathcal{E}$ is continuous from below.

b) Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a convex expectation. Then, $\mathcal{E}$ is ca-weakly lower semicontinuous if and only if
$$\mathcal{E}(X)=\sup_{\mu\in\mathrm{ca}_+^1(\Omega,\mathscr{F})}\mu X-\mathcal{E}^*(\mu)$$
for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$. In this case, we have that
$$\mathcal{E}(X)=\sup_{\mu\in\mathcal{Q}}\mu X-\mathcal{E}^*(\mu)\quad (X\in\mathscr{L}^\infty(\Omega,\mathscr{F})),$$
where $\mathcal{Q}:=\{\mu\in\mathrm{ca}_+^1(\Omega,\mathscr{F})\colon \mathcal{E}^*(\mu)<\infty\}$.

c) Let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a sublinear expectation. Then $\mathcal{E}$ is ca-weakly lower semicontinuous if and only if there exists a set $\mathcal{Q}\subset\mathrm{ca}_+^1(\Omega,\mathscr{F})$ such that
$$\mathcal{E}(X)=\sup_{\mu\in\mathcal{Q}}\mu X$$
for all $X\in\mathscr{L}^\infty(\Omega,\mathscr{F})$.

Since, by Remark 1.43, continuity from above on $\mathscr{L}^\infty(\Omega,\mathscr{F})$ of a convex expectation $\mathcal{E}$ already implies that $\mathcal{E}$ is dominated by some reference measure, this assumption is too strong in many applications. The following proposition is a standard result which shows that, in a topological space Ω, tightness is sufficient to at least obtain continuity from above on the space $C_b(\Omega)$ of all bounded continuous functions. For the reader's convenience, we provide a proof of this statement.

1.46 Proposition. *Let Ω be a topological space and $\mathscr{F}$ be the Borel σ-algebra on Ω. Moreover, let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega,\mathscr{F})\to\mathbb{R}$ be a sublinear expectation that is* ca*-weakly lower semicontinuous, and assume that $\mathcal{Q}:=\{\mu\in\mathrm{ca}_+^1(\Omega,\mathscr{F})\colon \mathcal{E}^*(\mu)<\infty\}$ is tight (see e.g. [10]). Then, $\mathcal{E}$ is continuous from above on $C_b(\Omega)$, i.e. the pre-expectation $\mathcal{E}|_{C_b(\Omega)}$ is continuous from above.*

Proof. Let $(X_n)_{n\in\mathbb{N}} \subset C_b(\Omega)$ with $X_n \searrow 0$ as $n \to \infty$ and $\varepsilon > 0$. As $\mathcal{Q}$ is tight, there exists a compact set $K \subset \Omega$ such that

$$\|X_1\|_\infty \cdot \sup_{\mu\in\mathcal{Q}_c} (\Omega \setminus K) \leq \varepsilon.$$

By Dini's lemma, we have that $\|X_n 1_K\|_\infty \searrow 0$ as $n \to \infty$. As $\mathcal{E}$ is ca-weakly lower semicontinuous, we thus get that

$$\begin{aligned}\mathcal{E}(X_n) &= \sup_{\mu\in\mathcal{Q}} \mu X_n - \mathcal{E}^*(\mu) = \sup_{\mu\in\mathcal{Q}} \big(\mu(X_n 1_K) - \mathcal{E}^*(\mu)\big) + \mu(X_n 1_{\Omega\setminus K})\\ &\leq \mathcal{E}(X_n 1_K) + \|X_1\|_\infty \cdot \sup_{\mu\in\mathcal{Q}} \mu(\Omega \setminus K)\\ &\leq \|X_n 1_K\|_\infty + \|X_1\|_\infty \cdot \sup_{\mu\in\mathcal{Q}} \mu(\Omega \setminus K)\\ &\leq \|X_n 1_K\|_\infty + \varepsilon \searrow \varepsilon, \quad n \to \infty.\end{aligned}$$

Letting $\varepsilon \searrow 0$, we obtain that $\lim_{n\to\infty} \mathcal{E}(X_n) = 0$, and therefore, $\mathcal{E}$ is continuous from above at 0. Now the assertion follows from Corollary 1.32. □

1.47 Remark. In the situation of Proposition 1.46, if Ω is a Polish space (cf. [78, Section 2.2]), Prohorov's theorem (cf. [12, Theorem 8.6.2, p. 202]), Remark 1.15 e) and Lemma 1.31 imply that $\mathcal{E}$ is continuous from above on $C_b(\Omega)$ if and only if $\mathcal{Q}$ is tight. Therefore, the following three statements are equivalent:

(i) $\mathcal{E}$ is continuous from above on $C_b(\Omega)$,

(ii) $\mathcal{Q}$ is tight,

(iii) $\mathcal{E}$ is *tight*, i.e. for all $\varepsilon > 0$ there exists a compact set $K \subset \Omega$ with $\mathcal{E}(1_{\Omega\setminus K}) \leq \varepsilon$.

In fact, (iii) implies (i) for any convex expectation $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$, as the following lemma shows.

1.48 Lemma. *Let Ω be a topological space and $\mathscr{F}$ the Borel σ-algebra on Ω. Moreover, let $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$ be an expectation which is convex and tight in the sense of Remark 1.47. Then, $\mathcal{E}$ is continuous from above on $C_b(\Omega)$.*

Proof. Let $(X_n)_{n\in\mathbb{N}} \subset C_b(\Omega)$ with $X_n \searrow 0$ as $n \to \infty$ and $\varepsilon > 0$. Then, there exists a compact set $K \subset \Omega$ with $\mathcal{E}(1_{\Omega\setminus K}) \leq \varepsilon$. As $\mathcal{E}$ is 1-Lipschitz, using Dini's lemma, we get that

$$\begin{aligned}\mathcal{E}(X_n) &\leq \|X_n 1_K\|_\infty + \mathcal{E}(X_n 1_{\Omega\setminus K}) \leq \|X_n 1_K\|_\infty + \|X_1\|_\infty \mathcal{E}(1_{\Omega\setminus K})\\ &\leq \|X_n 1_K\|_\infty + \varepsilon \searrow \varepsilon, \quad n \to \infty.\end{aligned}$$

Letting $\varepsilon \searrow 0$, we obtain that $\mathcal{E}(X_n) \searrow 0$ as $n \to \infty$. By Corollary 1.32, the assertion follows. □

1.3 Some notes on Choquet's Capacitability Theorem

In this section, we provide a proof of a variant of the famous Capacitability Theorem by Choquet [19, Théorème 1], see also [27, Chapter III, Theorem 28]. We basically follow the proof, given in [19], as indicated in [19, Extensions du théorème 1, 2)].

Throughout this section, we consider an arbitrary nonempty set Ω. Suprema and infima, which are taken over a family of functions, are always to be understood in a pointwise sense. For $X, Y \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ let $(X \wedge Y)(\omega) := \min\{X(\omega), Y(\omega)\}$ and $(X \vee Y)(\omega) := \max\{X(\omega), Y(\omega)\}$ for all $\omega \in \Omega$. We say that a set $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ is *directed upwards* or *directed downwards* if $M = M \vee M$ or $M \wedge M = M$, respectively. Here, $M \wedge M$ and $M \vee M$ denote the sets of all $X \wedge Y$ and $X \vee Y$ with $X, Y \in M$, respectively. Moreover, for any set $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ of bounded functions let M_δ and M_σ denote the sets of all bounded functions $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$, for which there exists a sequence $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow X$ and $X_n \nearrow X$ as $n \to \infty$, respectively. We say that $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ is a *Riesz subspace* of $\mathscr{L}^\infty(\Omega, 2^\Omega)$, if M is a subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ which is directed upwards or, equivalently, if M is a subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ which is directed downwards.

We say that a set $\mathscr{H} \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ of bounded functions is a *monotone class* if $\mathscr{H} = \mathscr{H}_\sigma = \mathscr{H}_\delta$. The following theorem is a generalized version of the Monotone Class Theorem and can be found in [27, Chapter I, (22.3)]. For the reader's convenience, we provide a proof of this theorem.

1.49 Theorem (Monotone Class Theorem). *Let $\mathscr{H} \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a monotone class and $M \subset \mathscr{H}$ a Riesz subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ with $1 \in M$. Then, we have that $\mathscr{L}^\infty(\Omega, \sigma(M)) \subset \mathscr{H}$.*

Proof. Let $\mathscr{M}$ denote the set of all subspaces $\mathscr{L}$ of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ with $M \subset \mathscr{L} \subset \mathscr{H}$. By Zorn's lemma, there exists a maximal element $\mathscr{H}^* \in \mathscr{M}$. First, we show that $\mathscr{H}^*$ is a monotone class. Let $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ and $(X_n)_{n\in\mathbb{N}} \subset \mathscr{H}^*$ with $X_n \nearrow X$ as $n \to \infty$. Then,

$$\mathscr{L} := \{Y + \alpha X \colon Y \in \mathscr{H}^*, \alpha \in \mathbb{R}\}$$

is a subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$. Let $Y \in \mathscr{H}^*$ and $\alpha \in \mathbb{R}$. Then, as $\mathscr{H}^*$ is a subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$, we have that $Y + \alpha X_n \in \mathscr{H}^* \subset \mathscr{H}$ for all $n \in \mathbb{N}$. Hence, we get that $Y + \alpha X \in \mathscr{H}_\sigma$ or $Y + \alpha X \in \mathscr{H}_\delta$ if $\alpha \geq 0$ or $\alpha \leq 0$, respectively. Since $\mathscr{H}$ is a monotone class, in any case, we get that $X + \alpha Y \in \mathscr{H}$. This shows that $\mathscr{L} \in \mathscr{M}$ with $\mathscr{H}^* \subset \mathscr{L}$ and we may conclude that $\mathscr{H}^* = \mathscr{L}$ as $\mathscr{H}^*$ is a maximal element of $\mathscr{M}$. In particular, $X \in \mathscr{H}^*$, showing that $\mathscr{H}^*_\sigma = \mathscr{H}^*$. As $\mathscr{H}^*$ is a vector space, this already implies that $\mathscr{H}^*$ is a monotone class. Now, let

$$\mathcal{A} := \left\{ \bigcap_{k=1}^m \{X^k > \alpha_k\} \colon m \in \mathbb{N},\ \alpha_1, \ldots, \alpha_m \in \mathbb{R},\ X^1, \ldots, X^m \in M \right\}.$$

Then, $\mathcal{A}$ is stable under finite intersections. Let $m \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and $X^1, \ldots, X^m \in M$. For $n \in \mathbb{N}$ and $k = 1, \ldots, m$ let

$$X_n^k := n\big(0 \vee (X^k - \alpha_k) \wedge n^{-1}\big).$$

Since M is a Riesz subspace of $\mathscr{L}^\infty(\Omega, \mathscr{F})$ with $1 \in M$, we obtain that $X_n^k \in M$ for all $n \in \mathbb{N}$ and $k = 1, \ldots, m$. Moreover, we have that $X_n^k \nearrow 1_{\{X^k > \alpha_k\}}$ as $n \to \infty$ for $k = 1, \ldots, m$. As M

is a Riesz subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$, we obtain that $X_n := X_n^1 \wedge \ldots \wedge X_n^m \in M$ for all $n \in \mathbb{N}$ with

$$X_n \nearrow 1_{\bigcap_{k=1}^m \{X^k > \alpha_k\}}$$

as $n \to \infty$. Since $X_n \in M \subset \mathscr{H}^*$ for all $n \in \mathbb{N}$, we get that $1_{\bigcap_{k=1}^m \{X^k > \alpha_k\}} \in \mathscr{H}^*$. This shows that

$$\{1_A \colon A \in \mathcal{A}\} \subset \mathscr{H}^*.$$

As $\mathscr{H}^*$ is a vector space with $1 \in \mathscr{H}^*$ and $(\mathscr{H}^*)_\sigma = \mathscr{H}^*$, we get that $\mathcal{D} := \{D \in 2^\Omega \colon 1_D \in \mathscr{H}\}$ is a Dynkin system with $\mathcal{A} \subset \mathcal{D}$. The Dynkin lemma implies that $\sigma(M) = \sigma(\mathcal{A}) \subset \mathcal{D}$. Due to the fact that $\mathscr{H}^*$ is a subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$, we get that $\operatorname{span}\{1_A \colon A \in \sigma(M)\} \subset \mathscr{H}^*$. Since $(\mathscr{H}^*)_\sigma = \mathscr{H}^* = (\mathscr{H}^*)_\delta$, we finally obtain that $\mathscr{L}^\infty(\Omega, \sigma(M)) \subset \mathscr{H}^* \subset \mathscr{H}$. □

Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a set of bounded functions, $S := \bigcup_{n \in \mathbb{N}} \mathbb{N}^n$ be the set of all finite sequences of positive integers and $\Sigma := \mathbb{N}^\mathbb{N}$ denote the set of all infinite sequences of positive integers. A mapping $H \colon S \to M$ is called a *Souslin scheme* if there exists a constant $C > 0$ such that $\|H(s)\|_\infty \leq C$ for all $s \in S$. Let H be a Souslin scheme and $C > 0$ such that $\|H(s)\|_\infty \leq C$ for all $s \in S$. Then, we have that

$$H(\sigma) := \inf_{n \in \mathbb{N}} H(\sigma_1, \ldots, \sigma_n)$$

is a bounded function with $\|H(\sigma)\|_\infty \leq C$ for all $\sigma \in \Sigma$. Hence, the *nucleus*

$$N(H) := \sup_{\sigma \in \Sigma} H(\sigma) = \sup_{\sigma \in \Sigma} \inf_{n \in \mathbb{N}} H(\sigma_1, \ldots, \sigma_n)$$

of the Souslin scheme H is again a bounded function with $\|N(H)\|_\infty \leq C$. We denote by $\mathscr{S}(M)$ the set of all nuclei $N(H)$ of Souslin schemes $H \colon S \to M$.

The above construction of M-valued Souslin schemes and their nuclei is inspired by [19, Extensions du théorème 1, 2)] and basically follows the definition of set-valued Souslin schemes and their nuclei (cf. [8, Definition 7.15, p. 157]).

1.50 Remark. Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards and let $X \in \mathscr{S}(M)$.

a) Let $X_0, X_\infty \in M$ with $X_0 \leq X \leq X_\infty$. Then, there exists a Souslin scheme $H \colon S \to M$ with

$$X_0 \leq H(s) \leq X_\infty$$

for all $s \in S$ and $X = N(H)$. In fact, let $K \colon S \to M$ be a Souslin scheme for X and

$$H \colon S \to M, \quad s \mapsto X_0 \vee K(s) \wedge X_\infty.$$

Then, we have that

$$X = X_0 \vee X \wedge X_\infty = \sup_{\sigma \in \Sigma} \big(X_0 \vee K(\sigma) \wedge X_\infty\big) = \sup_{\sigma \in \Sigma} \inf_{n \in \mathbb{N}} \big(X_0 \vee K(\sigma_1, \ldots, \sigma_n) \wedge X_\infty\big).$$

b) Let $K\colon S \to M$ be a Souslin scheme for X. Then, there exists a Souslin scheme L for X with

$$L(\sigma_1 \dots, \sigma_n) \searrow K(\sigma)$$

as $n \to \infty$ for all $\sigma \in \Sigma$. In particular, $K(\sigma) = L(\sigma) \in M_\delta$. Let $C > 0$ such that $\|K(s)\|_\infty \le C$ for all $s \in S$. For each $n \in \mathbb{N}$ and $s \in \mathbb{N}^n$ let

$$L(s) := K(s_1) \wedge \ldots \wedge K(s_1, \dots, s_n).$$

Since M is directed downwards, $L(s) \in M$ with $\|L(s)\|_\infty \le C$ for all $s \in S$. Hence, $L\colon S \to M$ is a Souslin scheme for X with $L(\sigma_1, \dots, \sigma_n) \searrow K(\sigma)$ as $n \to \infty$ for all $\sigma \in \Sigma$.

c) For $s, t \in S$ we write $t \le s$ if $s, t \in \mathbb{N}^n$ for some $n \in \mathbb{N}$ and $t_i \le s_i$ for all $i \in \{1, \dots, n\}$. Then, there exists a Souslin scheme H for X with $H(t) \le H(s)$ for all $s, t \in S$ with $t \le s$ and

$$H(\sigma_1, \dots, \sigma_n) \searrow H(\sigma) \in M_\delta$$

as $n \to \infty$ for all $\sigma \in \Sigma$. In fact, by part b), there exists a Souslin scheme $L\colon S \to M$ for X with $L(\sigma_1, \dots, \sigma_n) \searrow L(\sigma)$ as $n \to \infty$ for all $\sigma \in \Sigma$. Define

$$H\colon S \to M, \quad s \mapsto \sup\{L(t)\colon t \in S,\ t \le s\}.$$

Note that $\{L(t)\colon t \in S,\ t \le s\}$ is finite and nonempty for all $s \in S$. Then, $H\colon S \to M$ is a Souslin scheme with $H(t) \le H(s)$ for all $s, t \in S$ with $t \le s$. Let $\sigma \in \Sigma$. As

$$L(s,k) \le L(s)$$

for all $s \in S$ and $k \in \mathbb{N}$, we further have that $H(\sigma_1, \dots, \sigma_n) \searrow H(\sigma)$ as $n \to \infty$. For $\tau \in \Sigma$ we write $\tau \le \sigma$ if $\tau_i \le \sigma_i$ for all $i \in \mathbb{N}$. Then, by definition of H, we have that

$$\sup\{L(\tau)\colon \tau \in \Sigma,\ \tau \le \sigma\} \le H(\sigma). \tag{1.6}$$

We now show that in (1.6) equality holds, which implies that $X = N(H)$. Let $\omega \in \Omega$ be fixed. We define

$$A_n := \big\{\tau \in \Sigma\colon \tau \le \sigma \text{ and } \big(H(\sigma)\big)(\omega) \le \big(L(\tau_1, \dots, \tau_n)\big)(\omega)\big\}$$

for all $n \in \mathbb{N}$, and equip $\Sigma = \mathbb{N}^\mathbb{N}$ with the product topology. Then, $A_n \subset \Sigma$ is closed for all $n \in \mathbb{N}$ as the mapping

$$\Sigma \to \mathbb{R}, \quad \tau \mapsto \big(L(\tau_1, \dots, \tau_n)\big)(\omega)$$

is continuous, and $K := \{\tau \in \Sigma\colon \tau \le \sigma\} \subset \Sigma$ is compact by Tychonoff's theorem (see e.g. [55, Chapter X]). Since

$$L(s,k) \le L(s)$$

for all $s \in S$ and $k \in \mathbb{N}$, we have that $A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$. Moreover, as

$$H(\sigma) \le H(\sigma_1, \dots, \sigma_n)$$

for all $n \in \mathbb{N}$, we have that $A_n \neq \emptyset$ for all $n \in \mathbb{N}$. Due to $A_n \subset K$ for all $n \in \mathbb{N}$, we get that $\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$, i.e. there exists some $\tau \in \Sigma$ with $\tau \le \sigma$ and

$$\big(H(\sigma)\big)(\omega) \le \big(L(\tau_1, \dots, \tau_n)\big)(\omega)$$

for all $n \in \mathbb{N}$. This shows that $\big(H(\sigma)\big)(\omega) \le \big(H(\tau)\big)(\omega)$ and therefore, in (1.6) equality holds and the supremum can even be replaced by a maximum.

The proof of the following lemma is a variant of the proof of [8, Proposition 7.35].

1.51 Lemma. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards with $\mathbb{R} \subset M$. Then, $\mathscr{S}(M)$ is a monotone class with $M \subset \mathscr{S}(M)$.*

Proof. Let $X \in M$ and define $H(s) := X$ for all $s \in S$. Then, $H\colon S \to M$ is a Souslin scheme with $X = N(H) \in \mathscr{S}(M)$. This shows that $M \subset \mathscr{S}(M)$.

Now, let $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ and $(X_k)_{k\in\mathbb{N}} \subset \mathscr{S}(M)$ with $X_k \nearrow X$ as $k \to \infty$. Then, we have that $-\|X_1\|_\infty \leq X_k \leq \|X\|_\infty$ for all $k \in \mathbb{N}$. Let $C := \max\{\|X\|_\infty, \|X_1\|_\infty\}$. As $\mathbb{R} \subset M$, by Remark 1.50 a), for each $k \in \mathbb{N}$ there exists a Souslin scheme $H_k\colon S \to M$ with $X_k = N(H_k)$ and $\|H_k(s)\|_\infty \leq C$ for all $s \in S$. Define $H\colon S \to M$ by

$$H(k) := C \quad \text{and} \quad H(k, s') := H_k(s')$$

for all $k \in \mathbb{N}$ and $s' \in S$. By definition of H, it holds $\|H(s)\|_\infty \leq C$ for all $s \in S$ and therefore, H is a Souslin scheme with

$$H(k, \sigma') = \inf_{n\in\mathbb{N}} H(k, \sigma_1', \ldots, \sigma_n') = \inf_{n\in\mathbb{N}} H_k(\sigma_1', \ldots, \sigma_n') = H_k(\sigma').$$

for all $k \in \mathbb{N}$ and $\sigma' \in \Sigma$. Hence, we get that

$$X = \sup_{k\in\mathbb{N}} X_k = \sup_{k\in\mathbb{N}} N(H_k) = \sup_{k\in\mathbb{N}} \sup_{\sigma'\in\Sigma} H_k(\sigma') = \sup_{k\in\mathbb{N}} \sup_{\sigma'\in\Sigma} H(k, \sigma') = \sup_{\sigma\in\Sigma} H(\sigma).$$

Finally, in order to show that $\mathscr{S}(M)_\delta = \mathscr{S}(M)$, let $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ and $(X_k)_{k\in\mathbb{N}} \subset \mathscr{S}(M)$ with $X_k \searrow X$ as $k \to \infty$. Then, we have that $-\|X\|_\infty \leq X_k \leq \|X_1\|_\infty$ for all $k \in \mathbb{N}$. Again, let $C := \max\{\|X\|_\infty, \|X_1\|_\infty\}$. As $\mathbb{R} \subset M$, by Remark 1.50 a), for each $k \in \mathbb{N}$ there exists a Souslin scheme $H_k\colon S \to M$ with $X_k = N(H_k)$ and $\|H_k(s)\|_\infty \leq C$ for all $s \in S$. Let $\pi\colon \mathbb{N}^2 \to \mathbb{N}$ a bijection between $\mathbb{N}^2$ and $\mathbb{N}$, e.g.

$$\pi\colon \mathbb{N}^2 \to \mathbb{N}, \quad (k, m) \mapsto (2m-1)2^{k-1}.$$

Let $n \in \mathbb{N}$ and $s \in \mathbb{N}^n$. Then, there exists exactly one pair $(k, m) \in \mathbb{N}^2$ with $n = \pi(k, m)$, and we define

$$H(s) := H_k(s_{\pi(k,1)}, \ldots, s_{\pi(k,m)}).$$

Then, $\|H(s)\|_\infty \leq C$ for all $s \in S$, and therefore, $H\colon S \to M$ defines a Souslin scheme. It remains to show that $X = N(H)$. Let $\sigma \in \Sigma$, and $\sigma^k \in \Sigma$ be given by

$$\sigma_m^k := \sigma_{\pi(k,m)}$$

for all $k, m \in \mathbb{N}$, where σ_m^k denotes the m-th coordinate of σ^k. Then, we have that

$$H(\sigma) = \inf_{n\in\mathbb{N}} H(\sigma_1, \ldots, \sigma_n) = \inf_{k\in\mathbb{N}} \inf_{m\in\mathbb{N}} H_k(\sigma_{\pi(k,1)}, \ldots, \sigma_{\pi(k,m)}) = \inf_{k\in\mathbb{N}} H_k(\sigma^k), \tag{1.7}$$

where the first equality holds since $\pi\colon \mathbb{N}^2 \to \mathbb{N}$ is a bijection. Hence, we get that

$$H(\sigma) = \inf_{k\in\mathbb{N}} H_k(\sigma^k) \leq \inf_{k\in\mathbb{N}} N(H_k) = \inf_{k\in\mathbb{N}} X_k = X$$

for all $\sigma \in \Sigma$. Taking the supremum over all $\sigma \in \Sigma$, we get that $N(H) \leq X$. In order to show that $X \leq N(H)$, let $\omega \in \Omega$ and $\varepsilon > 0$. Then, for every $k \in \mathbb{N}$ there exists some $\sigma^k \in \Sigma$ such that

$$\big(N(H_k)\big)(\omega) \leq \big(H_k(\sigma^k)\big)(\omega) + \varepsilon.$$

Let $n \in \mathbb{N}$. As π is a bijection between $\mathbb{N}^2$ and $\mathbb{N}$,

$$\sigma_{\pi(k,m)} := \sigma_m^k \quad (k, m \in \mathbb{N})$$

defines a sequence $\sigma \in \Sigma$ and (1.7) yields that

$$X(\omega) = \inf_{k \in \mathbb{N}} \big(N(H_k)\big)(\omega) \leq \inf_{k \in \mathbb{N}} \big(H(\sigma^k)\big)(\omega) + \varepsilon = \big(H(\sigma)\big)(\omega) + \varepsilon \leq \big(N(H)\big)(\omega) + \varepsilon.$$

Letting $\varepsilon \searrow 0$, we obtain that $X(\omega) \leq N(H)(\omega)$ and therefore $X = N(H)$. □

Combining the Monotone Class Theorem 1.49 with Lemma 1.51, we obtain the following corollary.

1.52 Corollary. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a Riesz subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ with $1 \in M$. Then, $\mathscr{L}^\infty(\Omega, \sigma(M)) \subset \mathscr{S}(M)$.*

We continue with a functional version of Choquet's Capacitability Theorem (cf. [19, Théorème 1]), which yields the uniqueness of certain extensions of pre-expectations.

1.53 Theorem (Choquet's Capacitability Theorem)**.** *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a Riesz subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ with $1 \in M$ and $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ be an expectation. We assume that $\mathcal{E}$ is continuous from below and that $\mathcal{E}|_{M_\delta}$ is continuous from above. Then, for all $X \in \mathscr{S}(M)$ we have that*

$$\mathcal{E}(X) = \sup\big\{\mathcal{E}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\},$$

i.e. $\mathcal{E}$ is uniquely determined by its values on M_δ.

Proof. Let $X \in \mathscr{S}(M)$. By Remark 1.50 c), there exists a Souslin scheme $H\colon S \to M$ for X with $H(\sigma_1, \ldots, \sigma_n) \searrow H(\sigma)$ as $n \to \infty$ for all $\sigma \in \Sigma$ and $H(t) \leq H(s)$ for all $s, t \in S$ with $t \leq s$. For all $s \in S$ we set

$$X(s) := \sup\big\{H(\sigma)\colon \sigma \in \Sigma,\ (\sigma_1, \ldots, \sigma_n) \leq s\big\} \leq H(s). \tag{1.8}$$

Let $\varepsilon > 0$. Note that $X(k) \nearrow X$ and $X(s,k) \nearrow X(s)$ as $k \to \infty$ for all $s \in S$. Hence, as $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ is continuous from below, there exists some $\sigma \in \Sigma$ with

$$\mathcal{E}(X) \leq \mathcal{E}\big(X(\sigma_1)\big) + \frac{\varepsilon}{2}$$

and

$$\mathcal{E}\big(X(\sigma_1, \ldots, \sigma_{n-1})\big) \leq \mathcal{E}\big(X(\sigma_1, \ldots, \sigma_n)\big) + \frac{\varepsilon}{2^n}$$

for all $n \in \mathbb{N}$ with $n \geq 2$. By (1.8), we thus get that

$$\mathcal{E}(X) \leq \mathcal{E}\big(X(\sigma_1, \ldots, \sigma_n)\big) + \sum_{i=1}^n \frac{\varepsilon}{2^n} \leq \mathcal{E}\big(X(\sigma_1, \ldots, \sigma_n)\big) + \varepsilon \leq \mathcal{E}\big(H(\sigma_1, \ldots, \sigma_n)\big) + \varepsilon \tag{1.9}$$

for all $n \in \mathbb{N}$. Since $H(\sigma_1, \ldots, \sigma_n) \searrow H(\sigma)$ as $n \to \infty$ and $\mathcal{E}|_{M_\delta}$ is continuous from above, equality (1.9) implies that

$$\mathcal{E}(X) \leq \mathcal{E}\big(H(\sigma)\big) + \varepsilon \leq \sup\big\{\mathcal{E}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\} + \varepsilon.$$

Letting $\varepsilon \searrow 0$, we obtain that $\mathcal{E}(X) \leq \sup\big\{\mathcal{E}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\}$ and the assertion then follows from Remark 1.10. □

A combination of Lemma 1.38, Corollary 1.52 and Theorem 1.53 yields the following uniqueness result for nonlinear expectations on $\mathscr{L}^\infty(\Omega, \sigma(M))$, where $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ is a Riesz subspace with $1 \in M$.

1.54 Corollary. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a Riesz subspace of $\mathscr{L}^\infty(\Omega, 2^\Omega)$ with $1 \in M$ and $\mathcal{E}\colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$ be an expectation. We assume that $\mathcal{E}$ is continuous from below and that $\mathcal{E}|_{M_\delta}$ is continuous from above. Then, for all $X \in \mathscr{L}^\infty(\Omega, \sigma(M))$ we have that*

$$\mathcal{E}(X) = \sup\big\{\mathcal{E}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\},$$

i.e. $\mathcal{E}$ is uniquely determined by its values on M_δ.

Proof. By Lemma 1.38, $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ is continuous from below with $\hat{\mathcal{E}}|_{\mathscr{L}^\infty(\Omega,\sigma(M))} = \mathcal{E}$. In particular, $\hat{\mathcal{E}}$ is continuous from above on M_δ. By Theorem 1.53, we thus have that

$$\hat{\mathcal{E}}(X) = \sup\big\{\hat{\mathcal{E}}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\} = \sup\big\{\mathcal{E}(X_0)\colon X_0 \in M_\delta,\, X_0 \leq X\big\}$$

for all $X \in \mathscr{S}(M)$. The assertion now follows from Corollary 1.52. □

1.4 Extension of continuous pre-expectations

In Section 1.1, we extended a given pre-expectation $\mathcal{E}$ on $M \subset \mathscr{L}^\infty(\Omega, \mathscr{F})$ with $\mathbb{R} \subset M$ via

$$\hat{\mathcal{E}}(X) := \inf\{\mathcal{E}(X_0)\colon X_0 \in M, X_0 \geq X\} \quad (X \in \mathscr{L}^\infty(\Omega, \mathscr{F})).$$

We saw that, if M is a linear subspace of $\mathscr{L}^\infty(\Omega, \mathscr{F})$ and $\mathcal{E}$ is convex, both $\mathcal{E}$ and $\hat{\mathcal{E}}$ have a representation in terms of finitely additve probability measures. However, even if the pre-expectation $\mathcal{E}$ has a representation in terms of (σ-additive) probability measures, we cannot expect our maximal extension $\hat{\mathcal{E}}$ to be represented by probability measures as well.

Given that M is a Riesz subspace (cf. Section 1.3) with $1 \in M$, by the Daniell-Stone Theorem, for every linear pre-expectation $\mu\colon M \to \mathbb{R}$ which is continuous from above, there exists a unique linear expectation $\nu \in \mathrm{ca}^1_+(\Omega, \sigma(M))$ which is continuous from above and extends μ, i.e. $\mu X = \int X d\nu$ for all $X \in M$ (see [12, Theorem 7.8.1] or [32, Theorem 4.5.2]). However, already in the sublinear case, a similar statement does not hold as illustrated by the following example.

1.55 Example. Let $\Omega := [0, 1]$ and $\mathcal{E}(X) := \max_{\omega\in\Omega} X(\omega)$ for all $X \in M := C_b(\Omega)$. By Dini's lemma, $\mathcal{E}\colon M \to \mathbb{R}$ is continuous from above and has the representation

$$\mathcal{E}(X) = \max_{\mu\in \mathrm{ca}^1_+(\Omega,\mathscr{F})} \mu X \quad \text{for all } X \in M,$$

where $\mathscr{F}$ denotes the Borel σ-algebra on Ω. Notice that $\mathrm{ca}^1_+(\Omega, \mathscr{F})$ is a compact subset of $\mathrm{ca}(\Omega, \mathscr{F}) = C_b(\Omega)'$ or, equivalently, tight by Prohorov's theorem. However, it is not compact in $\mathrm{ba}(\Omega, \mathscr{F}) = \mathscr{L}^\infty(\Omega, \mathscr{F})'$. Suppose there existed an expectation $\widetilde{\mathcal{E}} : \mathscr{L}^\infty(\Omega, \mathscr{F}) \to \mathbb{R}$, which extends $\mathcal{E}$ and is continuous from above. Approximating the upper semicontinuous indicator function $1_{\{\omega\}}$ with continuous functions from above implies that $\widetilde{\mathcal{E}}(1_{\{\omega\}}) \geq 1$ for all $\omega \in \Omega$. Hence, for every sequence $(A_n)_{n\in\mathbb{N}} \subset \mathscr{F}$ with $A_n \neq \emptyset$ and $1_{A_n} \searrow 0$, one has $\widetilde{\mathcal{E}}(1_{A_n}) \geq 1$ for all $n \in \mathbb{N}$ and $\widetilde{\mathcal{E}}(0) = 0$.

Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a Riesz subspace with $1 \in M$ and $\mathcal{E}\colon M \to \mathbb{R}$ a convex pre-expectation, which is continuous from above. Then, the previous example indicates that $\mathcal{E}$ cannot be extended to an expectation which is continuous from above. The aim of this section is to derive an extension of $\mathcal{E}$ which is ca-weakly lower semicontinuous and coincides with the extension given by the Daniell-Stone Theorem in the linear case. Given that M is a Riesz subspace with $1 \in M$, the main theorem of this section, Theorem 1.61, states that for every convex pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$, which is continuous from above, there exists exactly one expectation $\overline{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$ which is continuous from below on $\mathscr{L}^\infty(\Omega, \sigma(M))$ and continuous from above on M_δ. Moreover, $\overline{\mathcal{E}}$ is convex and ca-weakly lower semicontinuous and coincides with the Daniell-Stone Extension in the linear case. This theorem can therefore be seen as a convex version of the Daniell-Stone Theorem.

We start by extending a pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$, which is continuous from above, to a pre-expectation on M_δ (cp. Section 1.3), which is continuous from above. Note that, by definition of M_δ, such an extension is automatically unique.

1.56 Lemma. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards (cp. Section 1.3) with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ a pre-expectation, which is continuous from above. Then, there exists a unique pre-expectation $\mathcal{E}_\delta\colon M_\delta \to \mathbb{R}$, which is continuous from above and extends $\mathcal{E}$. Moreover, $\mathcal{E}_\delta = \hat{\mathcal{E}}|_{M_\delta}$, i.e. $\mathcal{E}_\delta$ is the largest pre-expectation $\widetilde{\mathcal{E}}\colon M_\delta \to \mathbb{R}$ with $\widetilde{\mathcal{E}}|_M = \mathcal{E}$.*

Proof. Let $X \in M_\sigma$ and $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow X$ as $n \to \infty$. Then,

$$\|X\|_\infty \le \mathcal{E}(X_n) \le \mathcal{E}(X_1)$$

for all $n \in \mathbb{N}$ and therefore,

$$\mathcal{E}_\delta(X) := \lim_{n\to\infty} \mathcal{E}(X_n) \in \mathbb{R}.$$

First, we show that the definition of $\mathcal{E}_\delta(X)$ is independent of the sequence $(X_n)_{n\in\mathbb{N}} \subset M$ and that $\mathcal{E}_\delta$ is monotone. Let $Y \in M_\delta$ with $X \le Y$ and $(Y_n)_{n\in\mathbb{N}} \subset M$ with $Y_n \searrow Y$ as $n \to \infty$. Then, $Z_k^n := X_k \vee Y_n \in M$ for all $k, n \in \mathbb{N}$ and $Z_k^n \searrow Y_n$ as $k \to \infty$ for all $n \in \mathbb{N}$, and we obtain that

$$\lim_{k\to\infty} \mathcal{E}(X_k) \le \lim_{k\to\infty} \mathcal{E}(Z_k^n) = \mathcal{E}(Y_n)$$

for all $n \in \mathbb{N}$ since $\mathcal{E}$ is continuous from above. Letting $n \to \infty$, we get that

$$\lim_{n\to\infty} \mathcal{E}(X_n) \le \lim_{n\to\infty} \mathcal{E}(Y_n). \tag{1.10}$$

Taking $Y = X$, by a symmetry argument, we obtain that $\lim_{n\to\infty} \mathcal{E}(Y_n) = \lim_{n\to\infty} \mathcal{E}(X_n)$, which shows that $\mathcal{E}_\delta$ is well-defined. Moreover, (1.10) implies that $\mathcal{E}_\delta$ is monotone. Choosing $X_n = X$ for $n \in \mathbb{N}$ and $X \in M$, we obtain that $\mathcal{E}_\delta|_M = \mathcal{E}$. In particular, $\mathcal{E}_\delta(\alpha) = \alpha$ for all $\alpha \in \mathbb{R}$ as $\mathbb{R} \subset M$. Therefore, $\mathcal{E}_\delta$ defines a pre-expectation on M_δ with $\mathcal{E}_\delta|_M = \mathcal{E}$.

Now, let $X \in M_\delta$ and $(X_n)_{n\in\mathbb{N}} \subset M_\delta$ with $X_n \searrow X$ as $n \to \infty$. For all $n \in \mathbb{N}$ let $(X_k^n)_{k\in\mathbb{N}} \subset M$ with $X_k^n \searrow X_n$ as $k \to \infty$. Define

$$Y_n := X_n^1 \wedge \ldots \wedge X_n^n$$

for all $n \in \mathbb{N}$. Then, as M is directed downwards, we have that $Y_n \in M$ with $Y_n \ge Y_{n+1}$ for all $n \in \mathbb{N}$. Moreover, $Y_n \ge X_1 \wedge \ldots \wedge X_n = X_n$ for all $n \in \mathbb{N}$ and $X_k^n \ge X_k^k \ge Y_k$ for all $k, n \in \mathbb{N}$ with $k \ge n$. Hence,

$$X_n = \lim_{k\to\infty} X_k^n \ge \lim_{k\to\infty} Y_k \ge \lim_{k\to\infty} X_k = X$$

for all $n \in \mathbb{N}$. Letting $n \to \infty$, we obtain that $X = \lim_{n\to\infty} Y_n$. Altogether, we thus have that $Y_n \searrow X$ as $n \to \infty$ with $Y_n \ge X_n$ for all $n \in \mathbb{N}$ and therefore,

$$\mathcal{E}_\sigma(X) = \lim_{n\to\infty} \mathcal{E}(Y_n) \ge \lim_{n\to\infty} \mathcal{E}_\delta(X_n).$$

As $\mathcal{E}_\delta(X_n) \ge \mathcal{E}_\delta(X)$ for all $n \in \mathbb{N}$, we get that $\mathcal{E}_\delta(X) = \lim_{n\to\infty} \mathcal{E}_\sigma(X_n)$.

It remains to show that $\mathcal{E}_\delta = \hat{\mathcal{E}}|_{M_\delta}$. To this end, let $X \in M_\delta$ and $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow X$ as $n \to \infty$. Then, it holds that

$$\mathcal{E}_\delta(X) = \lim_{n\to\infty} \mathcal{E}(X_n) = \lim_{n\to\infty} \hat{\mathcal{E}}(X_n) \ge \hat{\mathcal{E}}(X)$$

and, by Remark 1.9, we thus get that $\mathcal{E}_\delta(X) = \hat{\mathcal{E}}(X)$. □

1.57 Remark. A similar result as in Lemma 1.56 also holds for expectations which are continuous from below. Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ a pre-expectation which is continuous from below. Then, there exists a unique pre-expectation $\mathcal{E}_\sigma\colon M_\sigma \to \mathbb{R}$ which is continuous from below and extends $\mathcal{E}$. This can be seen by considering the pre-expectation

$$N \to \mathbb{R}, \quad X \mapsto -\mathcal{E}(-X)$$

on $N := \{X \in \mathscr{L}^\infty(\Omega, 2^\Omega)\colon -X \in M\}$. Note that this pre-expectation is then continuous from above and therefore Lemma 1.56 can be applied. As $\mathcal{E}_\sigma$ is continuous from below, $\mathcal{E}_\sigma$ is given by

$$\mathcal{E}_\sigma(X) = \lim_{n\to\infty} \mathcal{E}(X_n)$$

for $X \in M_\sigma$ and $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \nearrow X$ as $n \to \infty$. Therefore, $\mathcal{E}_\sigma$ is the smallest pre-expectation $\widetilde{\mathcal{E}}\colon M_\sigma \to \mathbb{R}$, which extends $\mathcal{E}$.

The following lemma is a direct consequence of the fact that $\mathcal{E}_\delta(X) = \lim_{n\to\infty} \mathcal{E}(X_n)$ for $X \in M_\delta$ and $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_n \searrow X$ as $n \to \infty$. A similar result holds for $\mathcal{E}_\sigma$.

1.58 Lemma. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation which is continuous from above.*

(i) If M is convex and $\mathcal{E}$ is convex, then M_δ is convex and $\mathcal{E}_\delta$ is convex.

(ii) If M is a convex cone and $\mathcal{E}$ is sublinear, then M_δ is a convex cone and $\mathcal{E}_\delta$ is sublinear.

1.59 Remark. Let Ω be a metric space with metric d and $\mathrm{BUC}(\Omega)$ the space of all bounded uniformly continuous functions $\Omega \to \mathbb{R}$.

a) It holds that $\mathrm{BUC}(\Omega)_\sigma = \mathrm{LSC}_b(\Omega)$, where $\mathrm{LSC}_b(\Omega)$ denotes the space of all bounded lower semicontinuous functions $\Omega \to \mathbb{R}$.

   In fact, it is well-known that $\mathrm{BUC}(\Omega)_\sigma \subset \mathrm{LSC}_b(\Omega)$. In order to show the inverse implication, let $X \in \mathrm{LSC}_b(\Omega)$. W.l.o.g. we may assume that $X \geq 0$ (otherwise consider $X + \|X\|_\infty$). For $k, n \in \mathbb{N}_0$ let
   $$U_k^n := \big\{\omega \in \Omega\colon X(\omega) > 2^{-n}k\big\}.$$
   As U_n^k is open and Ω is a metric space, we have that $2^{-n}k1_{U_k^n} \in \mathrm{BUC}(\Omega)_\sigma$ for all $k, n \in \mathbb{N}_0$. Note that
   $$n\big(d(\omega, U^c) \wedge n^{-1}\big) \nearrow 1_U(\omega)$$
   as $n \to \infty$ for all $\omega \in \Omega$ and any open set $U \subset \Omega$. Finally, for all $n \in \mathbb{N}_0$ let
   $$X_n := \sup_{k\in\mathbb{N}_0} 2^{-n}k1_{U_k^n}.$$
   Then, we have that $X_n \in \mathrm{BUC}(\Omega)_\sigma$ with $X_n \leq X_{n+1} \leq X$ and $\|X - X_n\|_\infty \leq 2^{-n}$ for all $n \in \mathbb{N}_0$. In particular, $X_n \nearrow X$ as $n \to \infty$, and therefore, $X \in \mathrm{BUC}(\Omega)_\sigma$.

   As $\mathrm{BUC}(\Omega)$ is a vector space, we thus have that $\mathrm{BUC}(\Omega)_\delta = \mathrm{USC}_b(\Omega)$, where $\mathrm{USC}_b(\Omega)$ denotes the space of all bounded upper semicontinuous functions $\Omega \to \mathbb{R}$.

b) Let $\mathcal{E}\colon \mathrm{BUC}(\Omega) \to \mathbb{R}$ be a convex pre-expectation which is continuous from above. Then, there exists exactly one pre-expectation on $C_b(\Omega)$ which is continuous from above and extends $\mathcal{E}$. The extension is given by $\mathcal{E}_\delta|_{C_b(\Omega)}$. That is, there is a one to one correspondence between the set of all pre-expectations which are continuous from above on $\mathrm{BUC}(\Omega)$ and the set of all pre-expectations which are continuous from above on $C_b(\Omega)$ (cp. Remark 1.47).

Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation which is continuous from above. By Lemma 1.56, we can extend $\mathcal{E}$ in a unique way to a pre-expectation which is continuous from above on M_δ. Looking for an extension of $\mathcal{E}$ to an expectation which is continuous from below, Choquet's Capacitability Theorem indicates the following Ansatz, which is a nonlinear analogon of the concept of an inner measure.

1.60 Proposition. *Let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be directed upwards and downwards with $\mathbb{R} \subset M$ and $\mathcal{E}\colon M \to \mathbb{R}$ be a pre-expectation which is continuous from above. For $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ let*

$$\overline{\mathcal{E}}(X) := \sup\left\{ \inf_{n\in\mathbb{N}} \mathcal{E}(X_n) \colon (X_n)_{n\in\mathbb{N}} \subset M,\ X_{n+1} \le X_n\ (n \in \mathbb{N}),\ \inf_{n\in\mathbb{N}} X_n \le X \right\}. \tag{1.11}$$

Then,

$$\overline{\mathcal{E}}(X) = \sup\big\{\mathcal{E}_\delta(X_0)\colon X_0 \in M_\delta,\ X_0 \le X\big\} = (\mathcal{E}_\delta)^\vee(X)$$

for all $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ (cf. Remark 1.10). In particular, $\overline{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ is an expectation with $\overline{\mathcal{E}}|_M = \mathcal{E}$, and $\overline{\mathcal{E}}|_{M_\delta}$ is continuous from above.

Proof. Let $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$ and

$$B := \left\{ \inf_{n\in\mathbb{N}} \mathcal{E}(X_n) \colon (X_n)_{n\in\mathbb{N}} \subset M,\ X_{n+1} \le X_n\ (n \in \mathbb{N}),\ \inf_{n\in\mathbb{N}} X_n \le X \right\}.$$

As $B \supset \big\{\mathcal{E}_\delta(X_0)\colon X_0 \in M_\delta,\ X_0 \le X\big\}$, we have that

$$\overline{\mathcal{E}}(X) \ge \sup\big\{\mathcal{E}_\delta(X_0)\colon X_0 \in M_\delta,\ X_0 \le X\big\}. \tag{1.12}$$

It remains to show that equality holds in (1.12). Let $(X_n)_{n\in\mathbb{N}} \subset M$ with $X_{n+1} \le X_n$ for all $n \in \mathbb{N}$ and $\inf_{n\in\mathbb{N}} X_n \le X$. As X is bounded and $\mathbb{R} \subset M$, we have that

$$\widetilde{X}_n := X_n \vee (-\|X\|_\infty) \in M$$

for all $n \in \mathbb{N}$ with $\widetilde{X}_{n+1} \le \widetilde{X}_n$ for all $n \in \mathbb{N}$ and $\inf_{n\in\mathbb{N}} \widetilde{X}_n \le X$. Since $X_n \le \widetilde{X}_n$ for all $n \in \mathbb{N}$, we get that

$$\inf_{n\in\mathbb{N}} \mathcal{E}(X_n) \le \inf_{n\in\mathbb{N}} \mathcal{E}(\widetilde{X}_n)$$

and the assertion follows. □

We now focus on the convex case and want to apply the uniqueness result (Corollary 1.54) obtained in Section 1.3 in order prove that the extension, described in (1.11), is in some sense unique. Moreover, we want to derive a dual representation in terms of probability measures for (1.11), showing that this extension procedure leads to a convex expectation space (see Definition 1.44). Throughout the rest of this section, let $M \subset \mathscr{L}^\infty(\Omega, 2^\Omega)$ be a Riesz subspace with $1 \in M$. Typical examples for M are:

(i) The space $\operatorname{span}\{1_A \colon A \in \mathcal{A}\}$ of all $\mathcal{A}$-step functions, where $\mathcal{A} \subset 2^\Omega$ is an algebra. In this case, $\sigma(M) = \sigma(\mathcal{A})$,

(ii) The space $C_b(\Omega)$ of all bounded continuous functions $\Omega \to \mathbb{R}$, provided that Ω is a topological space. In this case $\sigma(M)$ is the Borel σ-algebra on Ω,

(iii) The space $\mathrm{BUC}(\Omega)$ of all bounded uniformly continuous functions $\Omega \to \mathbb{R}$, given that Ω is a metric space. In this case $\sigma(M)$ is the Borel σ-algebra on Ω,

(iv) The space $\mathrm{Lip}_b(\Omega)$ of all bounded Lipschitz continuous functions $\Omega \to \mathbb{R}$, where Ω is assumed to be a metric space. In this case $\sigma(M)$ is the Borel σ-algebra on Ω,

(v) The space $C_b^\theta(\Omega)$ of all bounded Hölder continuous functions $\Omega \to \mathbb{R}$ with Hölder exponent $\theta \in (0,1)$, given that Ω is a metric space. In this case $\sigma(M)$ is the Borel σ-algebra on Ω.

Let $\mathcal{E} \colon M \to \mathbb{R}$ be a convex pre-expectation which is continuous from above. Then, by Lemma 1.31, $\mathcal{E}$ has a dual representation in terms of linear pre-expectations which are continuous from above. One can therefore extend $\mathcal{E}$ to an expectation by extending these linear pre-expectations via the Daniell-Stone Theorem (see e.g. Cheridito et al. [17] or Maccheroni et al. [15]). This extension procedure leads to an expectation which is ca-weakly lower semicontinuous (see Definition 1.44). Note that the continuity from above on $\mathscr{L}^\infty(\Omega, \sigma(M))$ of such an extension would already imply that this extension is dominated by some probability measure (see Remark 1.43). Using Choquet's Capacitability Theorem (more precisely Corollary 1.54), we will show that, on $\mathscr{L}^\infty(\Omega, \sigma(M))$, this construction coincides with the extension $\overline{\mathcal{E}}$, given in Proposition 1.60.

1.61 Theorem. *Let $\mathcal{E} \colon M \to \mathbb{R}$ be a convex pre-expectation which is continuous from above. Then, $\overline{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$, given by* (1.11)*, is the only expectation which is continuous from below on $\mathscr{L}^\infty(\Omega, \sigma(M))$, continuous from above on M_δ and extends $\mathcal{E}$. Moreover, $\overline{\mathcal{E}}$ is convex with the dual representation*

$$\overline{\mathcal{E}}(X) = \sup_{\nu \in \mathrm{ca}^1_+(\Omega, \sigma(M))} \nu X - \mathcal{E}^*(\nu|_M)$$

for all $X \in \mathscr{L}^\infty(\Omega, \sigma(M))$. In particular, $\overline{\mathcal{E}}$ is ca*-weakly lower semicontinuous, i.e. $\big(\Omega, \sigma(M), \overline{\mathcal{E}}\big)$ is a convex expectation space.*

Proof. Let $\widetilde{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$ be given by

$$\widetilde{\mathcal{E}}(X) := \sup_{\nu \in \mathrm{ca}^1_+(\Omega, \sigma(M))} \nu X - \mathcal{E}^*(\nu|_M) \quad (X \in \mathscr{L}^\infty(\Omega, \sigma(M))).$$

By the theorem of Daniell-Stone, it follows that $\widetilde{\mathcal{E}}$ is a convex expectation which is ca-weakly lower semicontinuous and extends $\mathcal{E}$. Moreover, $\widetilde{\mathcal{E}}$ is continuous from above on M_δ. Indeed, let $(X_n)_{n \in \mathbb{N}} \in M$ with $X_n \searrow X$ as $n \to \infty$ for some $X \in M_\delta$. Let

$$\mathcal{Q} := \big\{\mu \in M' \,:\, \mathcal{E}^*(\mu) \le \|X_1\|_\infty + \|X\|_\infty\big\}$$

and

$$f \colon \mathcal{Q} \times \mathbb{N} \to \mathbb{R}, \quad (\mu, n) \mapsto \mu X_n - \mathcal{E}^*(\mu).$$

Then, by Lemma 1.12 c), $\mathcal{Q}$ is convex and compact. Moreover, f is concave on $\mathcal{Q}$ and convex on $\mathbb{N}$ in the sense of [37] (see Remark 1.30), and $f(\,\cdot\,,n)$ is upper semicontinuous for all $n\in\mathbb{N}$. As $\mathcal{E}$ is continuous from above, Lemma 1.31 implies that every $\mu\in\mathcal{Q}$ is continuous from above. By Lemma 1.12, Fan's minimax theorem and the Daniell-Stone Theorem, we thus obtain that

$$\begin{aligned}\inf_{n\in\mathbb{N}}\widetilde{\mathcal{E}}(X_n)&=\inf_{n\in\mathbb{N}}\max_{\mu\in\mathcal{Q}}\mu X_n-\mathcal{E}^*(\mu)=\max_{\mu\in\mathcal{Q}}\inf_{n\in\mathbb{N}}\mu X_n-\mathcal{E}^*(\mu)\\&\le\max_{\nu\in\mathrm{ca}^1_+(\Omega,\sigma(M))}\inf_{n\in\mathbb{N}}\nu X_n-\mathcal{E}^*(\nu|_M)=\max_{\nu\in\mathrm{ca}^1_+(\Omega,\sigma(M))}\nu X-\mathcal{E}^*(\nu|_M)\\&=\widetilde{\mathcal{E}}(X).\end{aligned}$$

Hence, $\widetilde{\mathcal{E}}(X)=\inf_{n\in\mathbb{N}}\widetilde{\mathcal{E}}(X_n)$, so that $\widetilde{\mathcal{E}}$ is continuous from above on M_δ, which implies that $\widetilde{\mathcal{E}}|_{M_\delta}=\mathcal{E}_\delta$. The assertion now follows from Corollary 1.54 and Proposition 1.60. □

We apply Theorem 1.61 to the linear case and obtain the following corollary.

1.62 Corollary. *Let $\mu\colon M\to\mathbb{R}$ be a linear pre-expectation which is continuous from above and $\nu\in\mathrm{ca}^1_+(\Omega,\mathscr{F})$ with $\nu|_M=\mu$. Then, ν is given by*

$$\nu X=\sup\left\{\inf_{n\in\mathbb{N}}\mu X_n\colon (X_n)_{n\in\mathbb{N}}\subset M,\ X_{n+1}\le X_n\ (n\in\mathbb{N}),\ \inf_{n\in\mathbb{N}}X_n\le X\right\}=\overline{\mu}X$$

for all $X\in\mathscr{L}^\infty(\Omega,\sigma(M))$. In particular, $\overline{\mu}\colon\mathscr{L}^\infty(\Omega,\sigma(M))\to\mathbb{R}$ is linear.

1.63 Remark. We consider the situation of Theorem 1.61.

a) Let $\nu\in\mathrm{ca}^1_+(\Omega,\sigma(M))$ and $\mu:=\nu|_M$. Then,

$$\overline{\mathcal{E}}^*(\nu)=\mathcal{E}^*(\mu),$$

where $\overline{\mathcal{E}}^*$ denotes the conjugate function of $\overline{\mathcal{E}}$. In particular, $\overline{\mathcal{E}}$ is sublinear if $\mathcal{E}$ is sublinear and $\overline{\mathcal{E}}$ is linear if $\mathcal{E}$ is linear.

In fact, by definition, we have that $\mathcal{E}^*(\mu)\le\overline{\mathcal{E}}^*(\nu)$. In order to show the inverse inequality, let $X\in\mathscr{L}^\infty(\Omega,\sigma(M))$ and $\varepsilon>0$. Then, by Corollary 1.62, there exists a sequence $(X_n)_{n\in\mathbb{N}}\subset M$ with $X_{n+1}\le X_n$ for all $n\in\mathbb{N}$, $\inf_{n\in\mathbb{N}}X_n\le X$ and $\nu X\le\inf_{n\in\mathbb{N}}\mu X_n+\varepsilon$. Further, there exists some $n_0\in\mathbb{N}$ such that $\mathcal{E}(X_{n_0})-\varepsilon\le\inf_{n\in\mathbb{N}}\mathcal{E}(X_n)$. We thus obtain that

$$\begin{aligned}\nu X-\overline{\mathcal{E}}(X)&\le\inf_{n\in\mathbb{N}}\mu X_n-\overline{\mathcal{E}}(X)+\varepsilon\le\mu X_{n_0}-\overline{\mathcal{E}}(X)+\varepsilon\\&\le\mu X_{n_0}-\inf_{n\in\mathbb{N}}\mathcal{E}(X_n)+\varepsilon\le\mu X_{n_0}-\mathcal{E}(X_{n_0})+2\varepsilon\\&\le\mathcal{E}^*(\mu)+2\varepsilon.\end{aligned}$$

Letting $\varepsilon\searrow 0$, we get that $\nu X-\overline{\mathcal{E}}(X)\le\mathcal{E}^*(\mu)$ and therefore $\overline{\mathcal{E}}^*(\nu)\le\mathcal{E}^*(\mu)$.

b) By Theorem 1.61, there exists at most one expectation on $\mathscr{L}^\infty(\Omega,\sigma(M))$ which is continuous from above and extends $\mathcal{E}$. If there exists such an extension, it is given by $\overline{\mathcal{E}}$.

c) Let $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) < \infty\}$. Then, by Theorem 1.61 and Corollary 1.62, we have that

$$\overline{\mathcal{E}}(X) = \sup_{\mu \in \mathscr{P}} \overline{\mu} X - \mathcal{E}^*(\mu)$$

for all $X \in \mathscr{L}^\infty(\Omega, \mathscr{F})$. Although, in part a), we have proved that

$$\{\overline{\mu} \colon \mu \in \mathscr{P}\} \subset \{\nu \in \mathrm{ba}(\Omega, \sigma(M)) \colon \overline{\mathcal{E}}^*(\nu) < \infty\}, \tag{1.13}$$

in general, equality does not hold in (1.13). Note that in (1.13) equality holds if and only if $\overline{\mathcal{E}}$ is continuous from above (cp. Lemma 1.41 and Proposition 1.42).

1.64 Corollary. *Let $\mathcal{E} \colon M \to \mathbb{R}$ be a sublinear pre-expectation which is continuous from above and $\mathscr{P} := \{\mu \in M' \colon \mathcal{E}^*(\mu) = 0\}$. Then, $\overline{\mathcal{E}} \colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$, given by* (1.11)*, is the only expectation which is continuous from below on $\mathscr{L}^\infty(\Omega, \sigma(M))$, continuous from above on M_δ and extends $\mathcal{E}$. Moreover, we have that*

$$\overline{\mathcal{E}}(X) = \sup_{\mu \in \mathscr{P}} \overline{\mu} X$$

for all $X \in \mathscr{L}^\infty(\Omega, \sigma(M))$ and therefore, $(\Omega, \sigma(M), \overline{\mathcal{E}})$ is a sublinear expectation space.

1.65 Remark. Consider the situation of Corollary 1.62. Since $\nu \colon \mathscr{L}^\infty(\Omega, \sigma(M)) \to \mathbb{R}$ is linear, we have that

$$\begin{aligned} \nu X &= -\nu(-X) \\ &= -\sup \left\{ \inf_{n \in \mathbb{N}} \mu Y_n \colon (Y_n)_{n \in \mathbb{N}} \subset M,\ Y_{n+1} \le Y_n \ (n \in \mathbb{N}),\ \inf_{n \in \mathbb{N}} Y_n \le -X \right\} \\ &= \inf \left\{ -\inf_{n \in \mathbb{N}} \mu Y_n \colon (Y_n)_{n \in \mathbb{N}} \subset M,\ Y_{n+1} \le Y_n \ (n \in \mathbb{N}),\ -\inf_{n \in \mathbb{N}} Y_n \ge X \right\} \\ &= \inf \left\{ \sup_{n \in \mathbb{N}} \mu X_n \colon (X_n)_{n \in \mathbb{N}} \subset M,\ X_n \le X_{n+1} \ (n \in \mathbb{N}),\ X \le \sup_{n \in \mathbb{N}} X_n \right\} \end{aligned}$$

for all $X \in \mathscr{L}^\infty(\Omega, \sigma(M))$. Let $\mu^* \colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ (in this case μ^* does not refer to the convex conjugate of μ) be defined by

$$\mu^*(X) := \inf \left\{ \sup_{n \in \mathbb{N}} \mu X_n \colon (X_n)_{n \in \mathbb{N}} \subset M,\ X_n \le X_{n+1} \ (n \in \mathbb{N}),\ X \le \sup_{n \in \mathbb{N}} X_n \right\}$$

for $X \in \mathscr{L}^\infty(\Omega, 2^\Omega)$. As $\mu^*|_{\mathscr{L}^\infty(\Omega, \sigma(M))} = \nu$, we obtain that $\mu^* = \hat{\nu}$ is sublinear and continuous from below by Lemma 1.38. Therefore, the mapping

$$2^\Omega \to [0,1], \quad A \mapsto \mu^*(1_A)$$

defines an outer measure on Ω (cf. [7, Definition 5.2]). That is, the Daniell-Stone integral of μ can be obtained from a functional version of an outer measure to μ. Moreover, by the theorem of Carathéodory (cf. [7, Theorem 5.3]),

$$\sigma(\mu^*) := \{A \in 2^\Omega \colon \mu^*(1_B) = \mu^*(1_{B \cap A}) + \mu^*(1_{B \setminus A}) \text{ for all } B \in 2^\Omega\}$$

defines a σ-algebra on Ω with $\sigma(M) \subset \sigma(\mu^*)$ and $\mu^*|_{\mathscr{L}^\infty(\Omega,\sigma(\mu^*))}$ is the expectation of a complete probability measure. Hence, we can extend any linear pre-expectation on M which is continuous from above, to a probability measure on the σ-algebra

$$\overline{\sigma}(M) := \bigcap \big\{\sigma(\mu^*)\colon \mu \text{ is a linear pre-expectation on } M \text{ which is continuous from below}\big\}$$

of all *universally M-measureable* sets. Note that $\sigma(M) \subset \overline{\sigma}(M)$. Therefore, we may even extend a convex pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$ which is continuous from above to a convex expectation on $\mathscr{L}^\infty(\Omega, \overline{\sigma}(M))$ which is ca-weakly lower semicontinuous.

By Theorem 1.61 we have that $\overline{\mathcal{E}}$ is the only expectation which is continuous from below on $\mathscr{L}^\infty(\Omega,\mathscr{F})$, continuous from above on M_δ and extends $\mathcal{E}$. However, there may exist infinitely many expectations which are continuous from below (even ca-weakly lower semicontinuous) and extend $\mathcal{E}$, as the following example shows.

1.66 Example. Let $\Omega := [0,1]$, $\mathscr{F}$ the Borel σ-algebra on $[0,1]$ and

$$\mathcal{E}(X) := \max_{\omega\in\Omega} X(\omega) = \max_{\mu\in \mathrm{ca}^1_+(\Omega,\mathscr{F})} \mu X = \max_{\omega\in\Omega} \delta_\omega X$$

for all $X \in M := C_b(\Omega)$, where $\delta_\omega \in \mathrm{ca}^1_+(\Omega,\mathscr{F})$ is given by

$$\delta_\omega(A) := \begin{cases} 1, & \omega \in A,\\ 0, & \omega \notin A \end{cases}$$

for all $A \in \mathscr{F}$ and $\omega \in \Omega$. Then, $\overline{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$ is given by

$$\overline{\mathcal{E}}(X) = \sup_{\mu\in \mathrm{ca}^1_+(\Omega,\mathscr{F})} \mu X = \sup_{\omega\in\Omega} \delta_\omega X = \sup_{\omega\in\Omega} X(\omega)$$

for all $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$. However, for every $\omega_0 \in \Omega$, we have that $\mathcal{E}_0\colon \mathscr{L}^\infty(\Omega,\mathscr{F}) \to \mathbb{R}$, given by

$$\mathcal{E}_0(X) := \sup_{\omega\in\Omega\setminus\{\omega_0\}} X(\omega) = \sup_{\omega\in\Omega\setminus\{\omega_0\}} \delta_\omega X$$

for $X \in \mathscr{L}^\infty(\Omega,\mathscr{F})$, is an expectation which is ca-weakly lower semicontinuous and extends $\mathcal{E}$.

1.67 Example. Let $P([0,1])$ denote the space of all polynomials $[0,1] \to \mathbb{R}$ with real coefficients and consider a sequence $(m_n)_{n\in\mathbb{N}_0} \subset \mathbb{R}$. The Hausdorff Moment Problem deals with the question, under which conditions this sequence is a sequence of moments of a probability measure on the Borel σ-algebra $\mathscr{B}([0,1])$. In [47] and [48] Hausdorff answered this question by proving that $(m_n)_{n\in\mathbb{N}_0}$ is a sequence of moments of a probability measure on $\mathscr{B}([0,1])$ if and only if $m_0 = 1$ and $(m_n)_{n\in\mathbb{N}_0}$ is completely monotonic, i.e.

$$(-1)^k(\Delta^k m)_n \geq 0$$

for all $k, n \in \mathbb{N}_0$. Here, $(\Delta m)_n := m_{n+1} - m_n$ for all $n \in \mathbb{N}_0$. We briefly illustrate this proof. Let $X \in P([0,1])$ be given by $X(\omega) = \sum_{k=0}^n a_k\omega^k$ for all $\omega \in [0,1]$ with $a_1, \ldots, a_n \in \mathbb{R}$ and $n \in \mathbb{N}_0$. We define the linear functional $\mu\colon P([0,1]) \to \mathbb{R}$ by setting

$$\mu X := \sum_{k=0}^n a_k m_k.$$

For $k, n \in \mathbb{N}_0$ let $X_{k,n} \in P([0,1])$ be given by $X_{k,n}(\omega) := \omega^n(1-\omega)^k$ for all $\omega \in [0,1]$. Then,

$$\mu X_{k,n} = \sum_{l=0}^{k} \binom{k}{l} (-1)^l m_{n+l} = (-1)^k \sum_{l=0}^{k} \binom{k}{l} (-1)^{k-l} m_{n+l} = (-1)^k (\Delta^k m)_n$$

for all $k, n \in \mathbb{N}_0$. Moreover, for all $X \in P([0,1])$ with $X \geq 0$ there exist $k_0, n_0 \in \mathbb{N}_0$ and $a_{k,n} \in [0,\infty)$ for all $1 \leq k \leq k_0$ and $1 \leq n \leq n_0$ such that

$$X = \sum_{k=1}^{k_0} \sum_{n=1}^{n_0} a_{k,n} X_{k,n}.$$

Hence, μ defines a linear pre-expectation on $P([0,1])$ if and only if $m_0 = 1$ and $(m_n)_{n\in\mathbb{N}_0}$ is completely monotonic. Therefore, the Hausdorff Moment Problem is equivalent to the extension of a linear pre-expectation on $P([0,1])$. Using Theorem 1.61, we can generalize Hausdorff's moment problem in the following sense. Let $\mathcal{E}\colon P([0,1]) \to \mathbb{R}$ be a convex pre-expectation on $P([0,1])$. As $\overline{P([0,1])} = C([0,1])$ by the Weierstrass approximation theorem, Proposition 1.18 yields that there exists a unique convex pre-expectation on $C([0,1])$ which extends $\mathcal{E}$, and will again be denoted by $\mathcal{E}$. By Remark 1.28 a), $\mathcal{E}\colon C([0,1]) \to \mathbb{R}$ is continuous from above. Hence, by Theorem 1.61, there exists a convex expectation $\overline{\mathcal{E}}\colon \mathscr{L}^\infty([0,1], \mathscr{B}([0,1])) \to \mathbb{R}$ which is ca-weakly lower semicontinuous and extends $\mathcal{E}$.

1.68 Remark. Let $\mathcal{Q}_1, \mathcal{Q}_2 \subset \mathrm{ca}^1_+(\Omega, \mathscr{F})$ with $\mathcal{Q}_1|_M = \mathcal{Q}_2|_M$. Then, it already holds that $\mathcal{Q}_1 = \mathcal{Q}_2$. In fact, let $\nu_1 \in \mathcal{Q}_1$. Then, we have that $\nu_1|_M \in \mathcal{Q}_2|_M$. Hence, there exists some $\nu_2 \in \mathcal{Q}_2$ with $\nu_2|_M = \nu_1|M$ and therefore, $\nu_1 = \nu_2 \in \mathcal{Q}_2$ by the Daniell-Stone Theorem. By a symmetry argument, we obtain that $\mathcal{Q}_1 = \mathcal{Q}_2$.

Given an expectation $\mathcal{E}$ which is continuous from below, Lemma 1.38 states that the expectation $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(\Omega, 2^\Omega) \to \mathbb{R}$ is continuous from below and extends $\mathcal{E}$. Therefore, one might think that, extending a pre-expectation $\mathcal{E}\colon M \to \mathbb{R}$, which is continuous from below, the extension $\hat{\mathcal{E}}$ yields an extension, which is continuous from below, as well. However, in general, this is not the case as the following example shows.

1.69 Example. Let $S := \{0,1\}$ be endowed with the topology 2^S. Then, S is a topological space with Borel-σ-algebra $\mathscr{B} := 2^S$. Let $\Omega := S^\mathbb{N}$ be endowed with the product topology $\mathscr{F} := \mathscr{B}^\mathbb{N}$ and $\mathscr{H} := \{J \subset \mathbb{N}\colon |J| \in \mathbb{N}\}$ be the set of all finite nonempty subsets of $\mathbb{N}$. Then, for all $J \in \mathscr{H}$ we have that $|S^J| = 2^{|J|} < \infty$ and therefore, the product σ-algebra $\mathscr{B}^J$ is the power set 2^{S^J} and $\mathscr{L}^\infty(S^J, \mathscr{B}^J)$ consists of all functions $S^J \to \mathbb{R}$. For all $J \in \mathscr{H}$ let

$$\mathrm{pr}_J\colon \Omega \to S^J, \quad (x_n)_{n\in\mathbb{N}} \mapsto (x_n)_{n\in J}.$$

Let $M := \big\{f \circ \mathrm{pr}_J\colon J \in \mathscr{H},\ f \in \mathscr{L}^\infty\big(S^J, 2^{S^J}\big)\big\}$, $y \in \Omega$ be arbitrary and $\delta_y \in \mathrm{ca}^1_+(\Omega, \mathscr{F})$ denote the Dirac measure given by

$$\delta_y(A) := \begin{cases} 1, & y \in A, \\ 0, & y \notin A \end{cases}$$

for all $A \in \mathscr{F}$. For all $n \in \mathbb{N}$ let $S^n := S^{\{1,\ldots,n\}}$ and $\mathrm{pr}_n := \mathrm{pr}_{\{1,\ldots,n\}}$. Let $\mathcal{E}\colon M \to \mathbb{R}$ be given by $\mathcal{E}(g) := \delta_y g$ for $g \in M$ and $f := 1_{S^\mathbb{N}\setminus\{y\}}$. Then, we have that $f \in \mathscr{L}^\infty(S^\mathbb{N})$. For all $n \in \mathbb{N}$ let

$$B_n := \mathrm{pr}_n^{-1}\big(\{(y_1,\ldots,y_n)\}\big) = \big\{x \in S^\mathbb{N}\colon x_i = y_i \text{ for all } i \in \{1,\ldots,n\}\big\} \in \mathscr{B}^\mathbb{N}$$

and $g_n := 1_{S^{\mathbb{N}}\setminus B_n} = 1 - 1_{B_n} \in M$. Then, we have that $g_n \nearrow f$ as $n \to \infty$, i.e. $f \in M_\sigma$. In fact, by definition, we have that $g_n(y) = 0 = f(y)$ for all $n \in \mathbb{N}$. As $y \in B_n$, we have that

$$\hat{\mathcal{E}}(g_n) = \mathcal{E}(g_n) = \delta_y g_n = 0$$

for all $n \in \mathbb{N}$. Let $g \in M$ with $g \geq f$. Then, we have that $g(x) \geq f(x) = 1$ for all $x \in S^{\mathbb{N}} \setminus \{y\}$. On the other hand, there exists some $J \in \mathscr{H}$ and some $f \colon S^J \to \mathbb{R}$ such that $g = f \circ \operatorname{pr}_J$. As $|S| = 2 > 1$, there exists some $x \in S^{\mathbb{N}} \setminus \{y\}$ with $\operatorname{pr}_J(x) = \operatorname{pr}_J(y)$ and therefore,

$$g(y) = f(\operatorname{pr}_J(y)) = f(\operatorname{pr}_J(x)) = g(x) \geq 1.$$

This shows that $g(x) \geq 1$ for all $x \in S^{\mathbb{N}}$. As $1 \geq f$ and $1 \in M$, we thus have that

$$\hat{\mathcal{E}}(f) = 1 \neq 0 = \lim_{n\to\infty} \hat{\mathcal{E}}(g_n).$$

This shows that for a pre-expectation $\mathcal{E}$ which is continuous from above or below, in general, the extension $\hat{\mathcal{E}}$ is not continuous from below, not even on M_σ. Moreover, by Corollary 1.24, we obtain that there exists some $\nu \in \mathrm{ba}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ with $\nu \neq \delta_y$ and $\nu|_M = \delta_y|_M$. In particular, we may deduce that $\mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}) \neq \mathrm{ba}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ although $\mathrm{ca}_+^1(S^n, \mathscr{B}^n) = \mathrm{ba}_+^1(S^n, \mathscr{B}^n)$ for all $n \in \mathbb{N}$.

Chapter 2

Existence of stochastic processes under nonlinear expectations

2.1 A robust version of Kolmogorov's extension theorem

In this section, we apply the extension results of the previous chapter to a Kolmogorov type setting. That is, given a consistent family of finite-dimensional marginal expectations, we want to find an expectation with these marginals. Again, we distinguish between the finitely additive case and the countably additive case. Finally, we will state a robust version of Kolmogorov's extension theorem (cp. [6, Theorem 35.3]).
Throughout this section, let $I \neq \emptyset$ be an index set, $\mathscr{H} := \big\{J \subset I \colon |J| \in \mathbb{N}\big\}$ the set of all finite, nonempty subsets of I and S a Polish space (cf. [78, Section 2.2]) with Borel σ-algebra $\mathscr{B}$. Typical examples for S are the space $\mathbb{R}^d$ with $d \in \mathbb{N}$, any open, halfopen or closed interval, or any separable Banach space. For each $J \in \mathscr{H}$ let $M_J \subset \mathscr{L}^\infty(S^J, \mathscr{B}^J)$ be a linear subspace with $1 \in M_J$, where $\mathscr{B}^J$ is the product σ-algebra on S^J. As before, M_J is always endowed with the norm $\|\cdot\|_\infty$, and on the topological dual M_J' of M_J we consider the weak*-topology for all $J \in \mathscr{H}$. For all $K \subset J \subset I$ let

$$\mathrm{pr}_{JK} \colon S^J \to S^K, \quad (x_i)_{i\in J} \mapsto (x_i)_{i\in K}$$

and $\mathrm{pr}_J := \mathrm{pr}_{IJ}$. Throughout this section, we assume that

$$M_K \circ \mathrm{pr}_{JK} := \big\{f \circ \mathrm{pr}_{JK} \colon f \in M_K\big\} \subset M_J$$

for all $J, K \in \mathscr{H}$ with $K \subset J$. Typical examples for the family $(M_J)_{J\in\mathscr{H}}$ are:

(i) the space $\mathscr{L}^\infty(S^J) := \mathscr{L}^\infty(S^J, \mathscr{B}^J)$ of all bounded $\mathscr{B}^J$-$\mathscr{B}(\mathbb{R})$-measurable functions, where $\mathscr{B}^J$ denotes the product σ-algebra on S^J,

(ii) the space $C_b(S^J)$ of all bounded continuous functions $S^J \to \mathbb{R}$, where S^J is endowed with the product topology,

(iii) the space $\mathrm{BUC}(S^J)$ of all bounded uniformly continuous functions $S^J \to \mathbb{R}$ w.r.t. a fixed metric, which generates the topology on S.

Let $J, K \in \mathscr{H}$ with $K \subset J$. For a pre-expectation $\mathcal{E}_J \colon M_J \to \mathbb{R}$ we then denote by $\mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$ the restriction of the distribution of $\mathcal{E}_J$ under pr_{JK}^{-1} to M_K (cp. Remark 1.2 i)), i.e.

$$\mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1} \colon M_K \to \mathbb{R}, \quad f \mapsto \mathcal{E}_J(f \circ \mathrm{pr}_{JK}).$$

In [65], Peng defines a consistency condition for nonlinear expectations and proves an extension to the subspace

$$M := \big\{f \circ \mathrm{pr}_J \colon J \in \mathscr{H}, f \in \mathscr{L}^\infty(S^J, \mathscr{B}^J)\big\}$$

of $\mathscr{L}^\infty(S^I, \mathscr{B}^I)$. Here $\mathscr{B}^I$ denotes the product σ-algebra of $\mathscr{B}$, i.e. the σ-algebra generated by the sets of the form $\mathrm{pr}_J^{-1}(B_J)$ with $J \in \mathscr{H}$ and $B_J \in \mathscr{B}^J$. In the sequel, we use the same notion of consistency as Peng and apply the extension results from the previous chapter in order to obtain an extension to $\mathscr{L}^\infty(S^I, \mathscr{B}^I)$.

2.1 Definition. For all $J \in \mathscr{H}$ let $\mathcal{E}_J\colon M_J \to \mathbb{R}$ be a pre-expectation. Then, the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is *consistent* if for all $J, K \in \mathscr{H}$ with $K \subset J$ it holds

$$\mathcal{E}_K(f) = \mathcal{E}_J(f \circ \mathrm{pr}_{JK}) \quad (f \in M_K).$$

A family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ of subsets $\mathcal{Q}_J \subset M_J'$ is *consistent* if for all $J, K \in \mathscr{H}$ with $K \subset J$ it holds

$$\mathcal{Q}_K = \big\{\mu \circ \mathrm{pr}_{JK}^{-1}\colon \mu \in \mathcal{Q}_J\big\} =\colon \mathcal{Q}_J \circ \mathrm{pr}_{JK}^{-1}.$$

2.2 Remark. For all $J \in \mathscr{H}$ let $\mathcal{E}_J\colon M_J \to \mathbb{R}$ be a pre-expectation.

a) The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent if and only if

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J, K \in \mathscr{H}$ with $K \subset J$.

b) The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent if and only if

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J, K \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + 1$. In fact, assume that $\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$ for all $J, K \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + 1$. We prove that

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J \in \mathscr{H}$ with $K \subset J$ by induction on $n = |J| - |K| \in \mathbb{N}_0$. For $n = 0$ the statement is trivial. Now, assume that there exists some $n \in \mathbb{N}_0$ such that

$$\mathcal{E}_K = \mathcal{E}_J \circ \mathrm{pr}_{JK}^{-1}$$

for all $J \in \mathscr{H}$ with $K \subset J$ and $|J| = |K| + n$. Let $J \in \mathscr{H}$ with $|J| = |K| + n + 1$, $J' := J \setminus \{i\}$ for some $i \in J \setminus K$ and $f \in M_K$. Then, we have that $g := f \circ \mathrm{pr}_{J'K} \in M_{J'}$ with $g \circ \mathrm{pr}_{JJ'} = f \circ \mathrm{pr}_{JK}$. Therefore, by the induction hypothesis, we have that

$$\mathcal{E}_J(f \circ \mathrm{pr}_{JK}) = \mathcal{E}_J(g \circ \mathrm{pr}_{JJ'}) = \mathcal{E}_{J'}(g) = \mathcal{E}_{J'}(f \circ \mathrm{pr}_{J'K}) = \mathcal{E}_K(f).$$

By Example 1.4 and Remark 1.15 d) and e), for all $J \in \mathscr{H}$ there is a one to one correspondence between the set of all convex compact subsets of M_J' and the set of all sublinear pre-expectations on M_J. Therefore, the consistency of sublinear pre-expectations and convex compact subsets of M_J' is closely related, as the following lemma shows.

2.3 Lemma. *For every $J \in \mathscr{H}$ let $\mathcal{E}_J\colon M_J \to \mathbb{R}$ be a sublinear pre-expectation and*

$$\mathcal{Q}_J := \big\{\mu_J \in M_J'\colon \mu_J f \le \mathcal{E}_J(f) \text{ for all } f \in M_J\big\}.$$

Then, the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent if and only if the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent.

Proof. Suppose that the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent. Then, by Remark 1.15 h) and Lemma 1.19, we obtain that the family $(Q_J)_{J\in\mathscr{H}}$ is consistent as well. Note that the distribution in Remark 1.2 i) and Remark 1.15 h) is a priori defined on a larger space than M_J. We thus have to use Lemma 1.19 in order to restict the elements of the set $\mathcal{Q}_J \circ \mathrm{pr}_{JK}^{-1}$, as defined in Remark 1.15 h), to M_K.

Now suppose that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent and let $J, K \in \mathscr{H}$ with $K \subset J$. Then, by Remark 1.15 d), we get that

$$\begin{aligned}\mathcal{E}_K(f) &= \max_{\mu_K\in\mathcal{Q}_K} \mu_K f = \max_{\mu_K\in\mathcal{Q}_J\circ\mathrm{pr}_{JK}^{-1}} \mu_K f \\ &= \max_{\mu_J\in\mathcal{Q}_J} \mu_J(f\circ\mathrm{pr}_{JK}) = \mathcal{E}_J(f\circ\mathrm{pr}_{JK})\end{aligned}$$

for all $f \in M_K$. □

The following theorem is a finitely additive and nonlinear version of Kolmogorov's extension theorem.

2.4 Theorem. *Let $(\mathcal{E}_J)_{J\in\mathscr{H}}$ be a consistent family of pre-expectations $\mathcal{E}_J\colon M_J \to \mathbb{R}$. Then, there exists an expectation $\hat{\mathcal{E}}\colon \mathscr{L}^\infty(S^I, \mathscr{B}^I) \to \mathbb{R}$ such that*

$$\hat{\mathcal{E}}(f\circ\mathrm{pr}_J) = \mathcal{E}_J(f) \quad \text{for all } J \in \mathscr{H} \text{ and all } f \in M_J.$$

If the pre-expectations $\mathcal{E}_J$ are convex or sublinear for all $J \in \mathscr{H}$, then $\hat{\mathcal{E}}$ is convex or sublinear, respectively.

Proof. Let $M := \{f\circ\mathrm{pr}_J \colon f \in M_J,\ J \in \mathscr{H}\}$. Then M is a linear subspace of $\mathscr{L}^\infty(S^I, \mathscr{B}^I)$ with $1 \in M$. For every $J \in \mathscr{H}$ and $f \in M_J$ let $\mathcal{E}(f\circ\mathrm{pr}_J) := \mathcal{E}_J(f)$. Since the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent, the functional $\mathcal{E}\colon M \to \mathbb{R}$ is well-defined. Moreover, $\mathcal{E}\colon M \to \mathbb{R}$ is a pre-expectation on M. The assertion now follows from Proposition 1.8. □

2.5 Remark. Consider the situation of Theorem 2.4.

a) The proof shows that the statement still holds without the assumption that S is a Polish space. In fact, S could be an arbitrary measurable space.

b) If $\mathcal{E}_J =: \mu_J$ is linear for all $J \in \mathscr{H}$, Corollary 1.24 yields that

$$\hat{\mathcal{E}}(f) = \sup_{\nu\in\mathcal{Q}} \nu f,$$

where $\mathcal{Q} := \{\nu \in \mathrm{ba}_+^1(S^I, \mathscr{B}^I) \colon \nu\circ\mathrm{pr}_J = \mu_J \text{ for all } J \in \mathscr{H}\}$.

c) One readily verifies that the pre-expectation $\mathcal{E}_0\colon M \to \mathbb{R}$ is linear, as soon as $\mathcal{E}_J =: \mu_J$ is linear for all $J \in \mathscr{H}$. Thus, we obtain the following corollary.

2.6 Corollary. *Let $(\mu_J)_{J\in\mathscr{H}}$ be a consistent family of linear pre-expectations $\mu_J\colon M_J \to \mathbb{R}$. Then there exists a linear expectation $\mu\colon \mathscr{L}^\infty(S^I, \mathscr{B}^I) \to \mathbb{R}$ such that*

$$\mu_J f = \mu(f\circ\mathrm{pr}_J)$$

for all $J \in \mathscr{H}$ and all $f \in M_J$.

For all $J \in \mathscr{H}$ we denote by $\mathrm{ba}^1_+(M_J)$ the set of all linear pre-expectations on M_J (cf. Remark 1.6). Using Theorem 2.4 and Lemma 2.3, we obtain the following finitely additive robust version of Kolmogorov's extension theorem. Note that, by Lemma 1.12 a), $\mathrm{ba}^1_+(M_J)$ is a subset of the topological dual M_J' of M_J. For all $J \in \mathscr{H}$ we equip $\mathrm{ba}^1_+(M_J)$ with the trace topology of the weak* topology on M_J'.

2.7 Corollary. *Suppose that $\mathcal{Q}_J \subset \mathrm{ba}^1_+(M_J)$ is convex and closed for all $J \in \mathscr{H}$ and that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists a convex compact set $\mathcal{Q} \subset \mathrm{ba}^1_+(S^I, \mathscr{B}^I)$ with*

$$\mathcal{Q}_J = \big\{\mu \circ \mathrm{pr}_J^{-1} \colon \mu \in \mathcal{Q}\big\} =: \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$

for all $J \in \mathscr{H}$, where $\mu \circ \mathrm{pr}_J^{-1} \colon M_J \to \mathbb{R},\ f \mapsto \mu(f \circ \mathrm{pr}_J)$ for all $\mu \in \mathrm{ba}(S^I, \mathscr{B}^I)$.

Proof. For each $J \in \mathscr{H}$ define the sublinear expectation

$$\mathcal{E}_J(f) := \max_{\mu_J \in \mathcal{Q}_J} \mu_J f \quad \text{for all } f \in M_J.$$

As $\mathcal{Q}_J \subset \mathrm{ba}^1_+(M_J)$ is a closed subset of the unit ball in M_J', the Banach-Alaoglu Theorem (see e.g. [69, Theorem 3.15, p. 66]) implies that $\mathcal{Q}_J$ is compact for all $J \in \mathscr{H}$. Due to Remark 1.15 e) and Lemma 2.3, $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is therefore a consistent family of sublinear expectations. Hence, by Theorem 2.4, there exists a sublinear expectation $\hat{\mathcal{E}} \colon \mathscr{L}^\infty(S^I, \mathscr{B}^I) \to \mathbb{R}$ with

$$\hat{\mathcal{E}}(f \circ \mathrm{pr}_J) = \mathcal{E}_J(f) \quad \text{for all } J \in \mathscr{H} \text{ and all } f \in M_J. \tag{2.1}$$

Using Lemma 1.12, we get that $\hat{\mathcal{E}}(f) = \max_{\mu\in\mathcal{Q}} \mu f$ for all $f \in \mathscr{L}^\infty(S^I, \mathscr{B}^I)$, where

$$\mathcal{Q} := \big\{\mu \in \mathrm{ba}^1_+(S^I, \mathscr{B}^I) \colon \mu f \le \hat{\mathcal{E}}(f) \text{ for all } f \in \mathscr{L}^\infty(S^I, \mathscr{B}^I)\big\}.$$

Remark 1.15 h) and Lemma 1.19 thus imply that $\mathcal{Q} \circ \mathrm{pr}_J^{-1} = \mathcal{Q}_J$ for all $J \in \mathscr{H}$. Note that, as in the proof of Lemma 2.3, we again use Lemma 1.19 since the distribution is a priori defined on a larger space than $\mathscr{L}^\infty(S^J, \mathscr{B}^J)$. □

With a slightly different proof, we can omit the convexity assumption in Corollary 2.7. However, we will no longer be able to use the theory of nonlinear expectations, which makes the proof a little bit more technical.

2.8 Theorem. *For all $J \in \mathscr{H}$ let $\mathcal{Q}_J \subset \mathrm{ba}^1_+(M_J)$ be closed in M_J' and assume that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists a compact set $\mathcal{Q} \subset \mathrm{ba}_+(S^I, \mathscr{B}^I)$ with*

$$\mathcal{Q}_J = \big\{\mu \circ \mathrm{pr}_J^{-1} \colon \mu \in \mathcal{Q}\big\} =: \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$

for all $J \in \mathscr{H}$.

Proof. First, note that, as in the proof of Corolary 2.7, the set $\mathcal{Q}_J \subset \mathrm{ba}^1_+(M_J)$ is compact. Let

$$\mathcal{Q} := \big\{\mu \in \mathrm{ba}^1_+(S^I, \mathscr{B}^I) \colon \mu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J \text{ for all } J \in \mathscr{H}\big\}.$$

As the mapping $\mathrm{ba}(S^I, \mathscr{B}^I) \to M_J,\ \mu \mapsto \mu \circ \mathrm{pr}_J^{-1}$ is continuous, $\mathcal{Q}$ is a closed subset of the unit ball in $\mathrm{ba}(S^I, \mathscr{B}^I)$, which implies the compactness $\mathcal{Q}$. By definition, it holds $\mathcal{Q} \circ \mathrm{pr}_J^{-1} \subset \mathcal{Q}_J$ for all $J \in \mathscr{H}$. In order to show the inverse implication, let $J_0 \in \mathscr{H}$ and $\nu_{J_0} \in \mathcal{Q}_{J_0}$ be fixed. For all $J \in \mathscr{H}$ let

$$A_J := \big\{\mu \in \mathrm{ba}^1_+(S^I, \mathscr{B}^I) \colon \mu \circ \mathrm{pr}_{J_0}^{-1} = \nu_{J_0} \text{ and } \mu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J\big\}.$$

Since $\mathcal{Q}_J \subset \mathrm{ba}^1_+(S^J, \mathscr{B}^J)$ is closed, we obtain that A_J is closed and therefore, by the Banach-Alaoglu Theorem, compact for all $J \in \mathscr{H}$. Let $J \in \mathscr{H}$ with $J_0 \subset J$. As the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent, there exists some $\mu_J \in \mathcal{Q}_J$ with

$$\mu_J \circ \mathrm{pr}_{JJ_0}^{-1} = \nu_{J_0}.$$

Let $M_J := \{f \circ \mathrm{pr}_J \colon f \in M_J\}$ and

$$\mu(f \circ \mathrm{pr}_J) := \mu_J f \quad \text{for all } f \in M_J.$$

Then, by Theorem 1.7, there exists some $\nu \in \mathrm{ba}^1_+(S^I, \mathscr{B}^I)$ with $\nu|_{M_J} = \mu$ and we thus have that $\mu \in A_J$, i.e. $A_J \neq \emptyset$. By assumption, for all $n \in \mathbb{N}$ and $J_1, \ldots, J_n \in \mathscr{H}$, it holds that

$$\bigcap_{i=1}^n A_{J_i} \supset A_{J_0 \cup \ldots \cup J_n} \neq \emptyset$$

and therefore, as A_J is compact for all $J \in \mathscr{H}$, we obtain that $\bigcap_{J\in\mathscr{H}} A_J \neq \emptyset$. That is, there exists some $\nu \in \mathrm{ba}^1_+(S^I, \mathscr{B}^I)$ with $\nu \circ \mathrm{pr}_{J_0}^{-1} = \nu_{J_0}$ and $\nu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J$ for all $J \in \mathscr{H}$. □

We apply Theorem 2.8 with $M_J = \mathscr{L}^\infty(S^J, \mathscr{B}^J)$ for all $J \in \mathscr{H}$ and obtain the following corollary.

2.9 Corollary. *For all $J \in \mathscr{H}$ let $\mathcal{Q}_J \subset \mathrm{ba}^1_+(S^J, \mathscr{B}^J)$ be closed and assume that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists a compact set $\mathcal{Q} \subset \mathrm{ba}_+(S^I, \mathscr{B}^I)$ with*

$$\mathcal{Q}_J = \big\{\mu \circ \mathrm{pr}_J^{-1} \colon \mu \in \mathcal{Q}\big\} =: \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$

for all $J \in \mathscr{H}$.

In Theorem 2.4, we proved the existence of an extension without any continuity properties. In Example 1.69, we saw that this extension does not satisfy any continuity properties and therefore no uniqueness can be obtained. The following theorem is the first main theorem of this section and can be viewed as a continuous version of Theorem 2.4, using Theorem 1.61 in order to obtain an extension to the path space.

2.10 Theorem. *For all $J \in \mathscr{H}$ let $\sigma(M_J) = \mathscr{B}^J$ and $\mathcal{E}_J \colon M_J \to \mathbb{R}$ be a convex pre-expectation, which is continuous from above. Let $M := \big\{f \circ \mathrm{pr}_J \colon f \in M_J,\ J \in \mathscr{H}\big\}$ and assume that the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists exactly one expectation $\bar{\mathcal{E}} \colon \mathscr{L}^\infty(S^I, \mathscr{B}^I) \to \mathbb{R}$ which is continuous from below on $\mathscr{L}^\infty(S^I, \mathscr{B}^I)$, continuous from above on M_δ and satisfies*

$$\mathcal{E}_J(f) = \bar{\mathcal{E}}(f \circ \mathrm{pr}_J) \quad \text{for all } J \in \mathscr{H} \text{ and } f \in M_J.$$

Moreover, $\bar{\mathcal{E}}$ is convex and ca*-weakly lower semicontinuous. If the pre-expectations $\mathcal{E}_J$ are sublinear or linear for all $J \in \mathscr{H}$, then $\bar{\mathcal{E}}$ is sublinear or linear, respectively.*

Proof. Define $\mathcal{E}(f \circ \mathrm{pr}_J) := \mathcal{E}_J(f)$ for all $f \in M_J$ and all $J \in \mathscr{H}$. Since the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent, $\mathcal{E}\colon M \to \mathbb{R}$ defines a convex pre-expectation on M. Let $\mu \in M'$ with $\mathcal{E}^*(\mu) < \infty$. We will first show that $\mu\colon M \to \mathbb{R}$ is continuous from above. Let $\mu_J := \mu \circ \mathrm{pr}_J^{-1}$ for all $J \in \mathscr{H}$. Then, $\mathcal{E}_J^*(\mu_J) \leq \mathcal{E}^*(\mu) < \infty$ and therefore, by Lemma 1.31, $\mu_J\colon M_J \to \mathbb{R}$ is continuous from above. By the theorem of Daniell-Stone, there exists a unique $\nu_J \in \mathrm{ca}_+^1(S^J, \mathscr{B}^J)$ with $\nu_J|_{M_J} = \mu_J$ for all $J \in \mathscr{H}$. Let $J, K \in \mathscr{H}$ with $K \subset J$ and $f \in M_K$. Then,

$$\mu_K f = \mu_J(f \circ \mathrm{pr}_{JK}) = \nu_J(f \circ \mathrm{pr}_{JK}) \quad \text{for all } f \in M_K$$

and therefore,

$$\nu_K f = \nu_J(f \circ \mathrm{pr}_{JK}) \quad \text{for all } f \in \mathscr{L}^\infty(S^K, \mathscr{B}^K)$$

as the extension of μ_K to a probability measure is unique. By Kolmogorov's extension theorem, there exists a unique $\nu \in \mathrm{ca}_+^1(S^I, \mathscr{B}^I)$ with $\nu(f \circ \mathrm{pr}_J) = \nu_J f$ for all $J \in \mathscr{H}$ and $f \in \mathscr{L}^\infty(S^J, \mathscr{B}^J)$. Hence, we get that $\nu|_M = \mu$, which implies that $\mu\colon M \to \mathbb{R}$ is continuous from above as well. By Lemma 1.31, we thus obtain that $\mathcal{E}\colon M \to \mathbb{R}$ is continuous from above.

Next, we show that $\mathscr{B}^I = \sigma(M)$. It is clear that $\mathscr{B}^I \supset \sigma(M)$. In order to show the inverse implication, let $J \in \mathscr{H}$ and $B_J \in \mathscr{B}^J$. Then, by assumption, $B_J \in \sigma(M_J)$ and therefore, we obtain that $\mathrm{pr}_J^{-1}(B_J) \in \sigma(M_J \circ \mathrm{pr}_J) \subset \sigma(M)$.

Finally, since M is a Riesz subspace of $\mathscr{L}^\infty(S^I, \mathscr{B}^I)$ with $1 \in M$ and $\mathscr{B}^I = \sigma(M)$, the assertion follows from Theorem 1.61. □

In the situation of Theorem 2.10, considering the canonical process $(\mathrm{pr}_{\{i\}})_{i\in I}$ on the nonlinear expectation space $\big(S^I, \mathscr{B}^I, \overline{\mathcal{E}}\big)$, we obtain the following two corollaries about the existence of stochastic processes under nonlinear expectations.

2.11 Corollary. *For all $J \in \mathscr{H}$ let $\sigma(M_J) = \mathscr{B}^J$ and $\mathcal{E}_J\colon M_J \to \mathbb{R}$ be a convex pre-expectation which is continuous from above. Then, the following two statements are equivalent:*

(i) The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent,

(ii) There exists a convex expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a stochastic process $(X_i)_{i\in I}$ with

$$\mathcal{E}(f(X_J)) = \mathcal{E}_J(f)$$

for all $J \in \mathscr{H}$ and $f \in M_J$, where $X_J := (X_i)_{i\in J}$.

2.12 Corollary. *For all $J \in \mathscr{H}$ let $\sigma(M_J) = \mathscr{B}^J$ and $\mathcal{E}_J\colon M_J \to \mathbb{R}$ be a sublinear pre-expectation which is continuous from above. Then, the following two statements are equivalent:*

(i) The family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent,

(ii) There exists a sublinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a stochastic process $(X_i)_{i\in I}$ with

$$\mathcal{E}(f(X_J)) = \mathcal{E}_J(f)$$

for all $J \in \mathscr{H}$ and $f \in M_J$, where $X_J := (X_i)_{i\in J}$.

We denote by $\mathrm{ca}_+^1(M_J)$ the set of all linear pre-expectations on M_J which are continuous from above (see Remark 1.28 b)), equipped with the trace topology of the weak* topology on M_J'. We now focus on the extension of sets of probability measures and obtain the following robust version of Kolmogorov's extension theorem. Here, the main task is to make a consistent choice of measures in $(\mathcal{Q}_J)_{J\in\mathscr{H}}$. We use a similar trick as in the proof of Theorem 2.10 in order to make this choice.

2.13 Theorem. *For all $J \in \mathscr{H}$ let $\sigma(M_J) = \mathscr{B}^J$ and $\mathcal{Q}_J \subset \mathrm{ca}_+^1(M_J)$ be compact in M_J'. Moreover, assume that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists a set $\mathcal{Q} \subset \mathrm{ca}_+^1(S^I, \mathscr{B}^I)$ with*

$$\mathcal{Q}_J = \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$

for all $J \in \mathscr{H}$.

Proof. Let

$$\mathcal{Q} := \big\{\mu \in \mathrm{ca}_+^1(S^I, \mathscr{B}^I) \colon \mu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J \text{ for all } J \in \mathscr{H}\big\}.$$

Then, $\mathcal{Q} \circ \mathrm{pr}_J^{-1} \subset \mathcal{Q}_J$ for all $J \in \mathscr{H}$. In order to show the inverse implication, let $J_0 \in \mathscr{H}$ and $\mu_{J_0} \in \mathcal{Q}_{J_0}$ be fixed. By Theorem 2.8, there exists some $\mu \in \mathrm{ba}_+^1(S^I, \mathscr{B}^I)$ with $\mu \circ \mathrm{pr}_{J_0}^{-1} = \mu_{J_0}$ and $\mu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J$ for all $J \in \mathscr{H}$. Let $\mu_J := \mu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J$ for all $J \in \mathscr{H} \setminus \{J_0\}$. Then, the family $(\mu_J)_{J\in\mathscr{H}}$ is consistent and $\mu_J \in \mathcal{Q}_J$ is continuous from above for all $J \in \mathscr{H}$. By the theorem of Daniell-Stone, there exists a unique $\nu_J \in \mathrm{ca}_+^1(S^J, \mathscr{B}^J)$ with $\nu_J|_{M_J} = \mu_J$ for all $J \in \mathscr{H}$. Since

$$\mu_K f = \mu_J(f \circ \mathrm{pr}_{JK}) = \nu_J(f \circ \mathrm{pr}_{JK}) \quad \text{for all } f \in M_K,$$

we obtain that

$$\nu_K f = \nu_J(f \circ \mathrm{pr}_{JK}) \quad \text{for all } f \in \mathscr{L}^\infty(S^J, \mathscr{B}^J),$$

by the uniqueness of ν_K. Hence, by Kolmogorov's extension theorem, there exists a unique $\nu \in \mathrm{ca}_+^1(S^I, \mathscr{B}^I)$ with $\nu \circ \mathrm{pr}_J^{-1} = \nu_J$ for all $J \in \mathscr{H}$. Hence, $\nu \in \mathcal{Q}$ with $\nu \circ \mathrm{pr}_{J_0}^{-1} = \mu_{J_0}$ and $\nu \circ \mathrm{pr}_J^{-1} \in \mathcal{Q}_J$ for all $J \in \mathscr{H}$. □

We apply Theorem 2.13 to the case $M_J = C_b(S^J)$ for all $J \in \mathscr{H}$ and obtain the following corollary, using Prohorov's theorem (cf. [12, Theorem 8.6.2, p. 202]).

2.14 Corollary. *For all $J \in \mathscr{H}$ let $\mathcal{Q}_J \subset \mathrm{ca}_+^1(S^J, \mathscr{B}^J)$ be closed and tight and assume that the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent. Then, there exists a set $\mathcal{Q} \subset \mathrm{ca}_+^1(S^I, \mathscr{B}^I)$ with*

$$\mathcal{Q}_J = \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$

for all $J \in \mathscr{H}$.

We close this section with three examples. The first two examples (Example 2.15 a) and b)) show that the set $\mathcal{Q}$, constructed in Theorem 2.13, is neither compact nor unique. In the third example (Example 2.16), we construct a stochastic process with nonlinear marginal distributions.

2.15 Example. As in Example 1.69, let $S := \{0, 1\}$ be endowed with the topology 2^S. Then, S is a Polish space with Borel-σ-algebra 2^S and for all $J \in \mathscr{H}$ we have that $\#S^J < \infty$. Therefore, the product σ-algebra $\mathscr{B}^J$ is the power set 2^{S^J} and $\mathscr{L}^\infty(S^J, \mathscr{B}^J)$ consists of all functions $S^J \to \mathbb{R}$.

a) Note that the set $\mathcal{Q} \subset \mathrm{ca}_+^1(S^I, \mathscr{B}^I)$ constructed in the proof of Theorem 2.13 is closed in the topology $\sigma(\mathrm{ca}(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}), \mathscr{L}^\infty(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}))$ (see Remark 1.34). However, in general, we may not expect the set $\mathcal{Q}$ to be compact in this topology, not even if $\mathcal{Q}_J$ is convex for all $J \in \mathscr{H}$. In fact, let $\mathcal{Q} := \mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$. Then, we have that $\mathcal{Q}$ is convex and closed in the topology $\sigma(\mathrm{ca}(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}), \mathscr{L}^\infty(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}))$. However, $\mathcal{Q}$ is not a compact subset of $\mathrm{ba}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$, as $\mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}) \neq \mathrm{ba}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ (see Example 1.69). On the other hand, we have that $\mathcal{Q} \circ \mathrm{pr}_J^{-1} = \mathrm{ba}_+^1(S^J, \mathscr{B}^J)$ is a convex and compact for all $J \in \mathscr{H}$.

b) Let $\mathcal{Q} := \mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ and $\mathscr{P} := \{\nu \in \mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}) \,|\, \nu(\{y\}) = 0\}$ for some $y \in S^{\mathbb{N}}$. Then, we have that $\mathscr{P}$ and $\mathcal{Q}$ are both convex and closed in the topology $\sigma(\mathrm{ca}(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}), \mathscr{L}^\infty(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}}))$. Let $J \in \mathscr{H}$. Since $\#S = 2 > 1$ there exists some $x_J \in S^{\mathbb{N}} \setminus \{y\}$ with $\mathrm{pr}_J(x_J) = \mathrm{pr}_J(y)$ and therefore,
$$\mathscr{P} \circ \mathrm{pr}_J^{-1} = \mathrm{ba}_+^1(S^J, \mathscr{B}^J) = \mathcal{Q} \circ \mathrm{pr}_J^{-1}$$
is compact. On the other hand, we have that $\mathscr{P} \neq \mathcal{Q}$. This shows that in Theorem 2.13 uniqueness cannot be obtained. Replacing $\mathrm{ca}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$ by $\mathrm{ba}_+^1(S^{\mathbb{N}}, \mathscr{B}^{\mathbb{N}})$, a similar example shows that also in Theorem 2.7 uniqueness cannot be obtained.

2.16 Example. Let $0 \leq \underline{\sigma} \leq \overline{\sigma}$ and $\underline{\mu} \leq \overline{\mu}$. For fixed $n \in \mathbb{N}$ let $\sigma \in [\underline{\sigma}, \overline{\sigma}]^n$, $\mu \in [\underline{\mu}, \overline{\mu}]^n$ and $0 \leq t_1 < \ldots < t_n$. For $J := \{t_1, \ldots, t_n\}$ and $f \in C_b(\mathbb{R}^n)$ let
$$\mathbb{E}_J^{\mu,\sigma}(f) := \int_{\mathbb{R}^n} f(x_1, x_1 + x_2, \ldots, x_1 + \ldots + x_n)\, \mathrm{d}N_J^{\mu,\sigma}(x_1, \ldots, x_n),$$
where
$$N_J^{\mu,\sigma} := \bigotimes_{k=1}^n N\big(\mu_k(t_k - t_{k-1}), \sigma_k^2(t_k - t_{k-1})\big)$$
is the product of normal distributions with $t_0 := 0$ and $N(0,0) := \delta_0$. Moreover, let
$$\mathcal{E}_J(f) := \sup_{\mu \in [\underline{\mu}, \overline{\mu}]^n, \sigma \in [\underline{\sigma}, \overline{\sigma}]^n} \mathbb{E}_J^{\mu,\sigma}(f)$$
for all $f \in C_b(\mathbb{R}^n)$. We equip $\mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ with the topology $\sigma\big(\mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)), C_b(\mathbb{R}^n)\big)$ (see Remark 1.34). Then, the mapping
$$\mathbb{R}^n \times [0,\infty)^n \to \mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)), \quad (\mu, \sigma) \mapsto N_J^{\mu,\sigma}$$
is continuous by Lévy's continuity theorem (cf. [82, Theorem 18.1]) or by direct computation (note that it suffices to verify sequential continuity as $\mathbb{R}^n \times [0,\infty)^n$ is a metric space). Let $s\colon \mathbb{R}^n \to \mathbb{R}^n$ be given by
$$s(x_1, \ldots, x_n) := (x_1, x_1 + x_2, \ldots, x_1 + \ldots + x_n) \quad \text{for all } x_1, \ldots, x_n \in \mathbb{R}.$$
As $s\colon \mathbb{R}^n \to \mathbb{R}^n$ is continuous, the mapping
$$\mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)) \to \mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)), \quad \nu \mapsto \nu \circ s^{-1}$$
is continuous and therefore, the mapping
$$\mathbb{R}^n \times [0,\infty)^n \to \mathrm{ca}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n)), \quad (\mu, \sigma) \mapsto \mathbb{E}_J^{\mu,\sigma}$$

is continuous. As $[\underline{\mu},\overline{\mu}]^n \times [\underline{\sigma},\overline{\sigma}]^n \subset \mathbb{R}^n \times [0,\infty)^n$ is compact, we thus obtain that the family

$$\mathcal{Q}_J := \big\{\mathbb{E}_J^{\mu,\sigma} \colon \mu \in [\underline{\mu},\overline{\mu}]^n, \sigma \in [\underline{\sigma},\overline{\sigma}]^n\big\} \subset C_b(\mathbb{R}^n)'$$

is compact. Since $\mathcal{Q}_J$ is convex and compact, we get that $\mathcal{E}_J$ is continuous from above with

$$\mathcal{Q}_J = \big\{\mu_J \in C_b(\mathbb{R}^n)' \colon \mu_J f \le \mathcal{E}_J(f) \text{ for all } f \in C_b(\mathbb{R}^n)\big\}$$

for each $J \in \mathscr{H}$. As the family $(\mathcal{Q}_J)_{J\in\mathscr{H}}$ is consistent, we thus get that the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent by Lemma 2.3. Therefore, by Corollary 2.12, there exists a sublinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a stochastic process $(X_t)_{t\ge 0}$ with

$$\mathcal{E}\big(f(X_{t_1},\ldots,X_{t_n})\big) = \mathcal{E}_{\{t_1,\ldots,t_n\}}(f)$$

for all $n \in \mathbb{N}$, $0 \le t_1 < \ldots < t_n$ and $f \in C_b(\mathbb{R}^n)$. Moreover, by Theorem 2.13 (Corollary 2.14, to be precise), there exists a set $\mathcal{Q} \subset \mathrm{ca}_+^1\left(\mathbb{R}^{[0,\infty)}, \mathscr{B}(\mathbb{R})^{[0,\infty)}\right)$ with $\mathcal{Q} \circ \mathrm{pr}_J^{-1} = \mathcal{Q}_J$ for all $J \in \mathscr{H}$. Note that this construction does not lead to the G-Expectation introduced by Peng [66], [67]. In order to construct G-normally distributed processes, we first have to introduce the notion of kernels and Markov processes, specifically Lévy processes, under nonlinear expectations, which will be done in the subsequent sections. The G-Expectation will then be constructed with a semigroup theoretic approach in Example 3.20.

2.2 Markov processes under nonlinear expectations

The aim of this section is to construct nonlinear Markov processes, using the extension results obtained in the previous section. Markov processes are closely related to graph and network theory (see [31], [62]), the theory of semigroups (cf. [34], [35], [53], [63]) and stochastic optimal control (cf. [38], [56], [68], [83]). As in the linear case (cf. [6], [36], [54]), the construction of Markov processes, using Kolmogorov's extension theorem, requires the notion of a kernel. We basically follow the definition of a monetary risk kernel by Föllmer and Klüppelberg [39] and define nonlinear kernels in an analogous way.

Throughout, let S be a Polish space with Borel σ-algebra $\mathscr{B}$. We fix a metric d on S and endow S^n with the metric

$$d_n(x,y) := \bigg(\sum_{i=1}^n d(x_i,y_i)^2\bigg)^{1/2} \quad (x,y\in S^n)$$

for all $n\in\mathbb{N}$.

2.17 Definition. Let $M,N\subset\mathscr{L}^\infty(S,\mathscr{B})$ with $1\in M$ and $1\in N$. A *(nonlinear) kernel* from M to N is a function $\mathcal{E}\colon M\to N$, such that for all $x_0\in S$

$$\mathcal{E}(x_0,\,\cdot\,)\colon M\to\mathbb{R},\quad f\mapsto \mathcal{E}(x_0,f) := \big(\mathcal{E}(f)\big)(x_0)$$

is a pre-expectation on M. We say that a kernel $\mathcal{E}$ from M to N is convex, sublinear or linear if $\mathcal{E}(x_0,\,\cdot\,)$ is convex, sublinear or linear for all $x_0\in S$, respectively. A pre-kernel $\mathcal{E}$ from M to N is said to be continuous from above or below, if $\mathcal{E}(x_0,\,\cdot\,)$ is continuous from above or below for each $x_0\in S$, respectively.

2.18 Remark. Let $M\subset\mathscr{L}^\infty(S,\mathscr{B})$ with $1\in M$. A pre-expectation on M is a kernel from M to $\mathbb{R}$.

We start with a technical lemma, which we will use in Section 3.1 in order to define distributions of stochastic integrals by a semigroup approach.

2.19 Lemma. *Let $\nu\in\mathrm{ca}^1_+(S,\mathscr{B})$ and $\mathcal{E}\colon \mathrm{BUC}(S)\to\mathrm{BUC}(S)$ be a kernel. We assume that there exists a set $\mathcal{Q}$ of linear kernels from $\mathrm{BUC}(S)$ to $\mathrm{BUC}(S)$ such that*

$$\mathcal{E}(x,f) = \sup_{\mu\in\mathcal{Q}} \mu(x,f)$$

for all $f\in\mathrm{BUC}(S)$ and $x\in S$. We denote by $\nu\mathcal{Q}\subset\mathrm{BUC}(S)'$ the set of all pre-expectations of the form

$$\mathrm{BUC}(S)\to\mathbb{R},\quad f\mapsto \nu\bigg(\sum_{k=1}^n 1_{B_k}\mu_k f\bigg)$$

with $n\in\mathbb{N}$, $B_1,\ldots,B_n\in\mathscr{B}$ and $\mu_1,\ldots,\mu_n\in\mathcal{Q}$. Then, for all $f\in\mathrm{BUC}(S)$ it holds that

$$\nu\mathcal{E}(f) = \sup_{\eta\in\nu\mathcal{Q}} \eta f.$$

Proof. Let $f \in \mathrm{BUC}(S)$. Clearly, it holds $\eta f \leq \nu\mathcal{E}(f)$ for all $\eta \in \nu\mathcal{Q}$. It remains to show the inverse inequality. As S is separable, $\mathcal{E}(f) \in \mathrm{BUC}(S)$ and $\mu f \in \mathrm{BUC}(S)$ for all $\mu \in \mathcal{Q}$, there exist $\mu_1^n, \ldots, \mu_n^n \in \mathcal{Q}$ and $B_1^n, \ldots, B_n^n \in \mathscr{B}$ for all $n \in \mathbb{N}$ such that

$$\sum_{k=1}^n 1_{B_k^n} \mu_k^n f \nearrow \mathcal{E}(f)$$

as $n \to \infty$. We therefore consider

$$\eta_n \colon \mathrm{BUC}(S) \to \mathbb{R}, \quad g \mapsto \nu\bigg(\sum_{k=1}^n 1_{B_k^n} \mu_k^n g\bigg)$$

for all $n \in \mathbb{N}$. Since ν is continuous from below, it then follows that

$$\eta_n f = \nu\bigg(\sum_{k=1}^n 1_{B_k^n} \mu_k^n f\bigg) \nearrow \nu\mathcal{E}(f)$$

as $n \to \infty$. □

We illustrate the construction of Markov processes with the following example on Markov chains.

2.20 Example (Markov chains)**.** Let S be a finite state space, which is endowed with the topology 2^S and the discrete metric d. Then, S is a Polish space with Borel-σ-algebra $\mathscr{B} = 2^S$ and $\mathscr{L}^\infty(S, \mathscr{B}) = \mathbb{R}^S$. Let $\mathscr{P} \colon \mathbb{R}^S \to \mathbb{R}^S$ and $\mu_0 \colon \mathbb{R}^S \to \mathbb{R}$. We assume that $\mathscr{P}$ and μ_0 are convex. Moreover, we assume that

$$\mathscr{P}(\alpha) = \alpha$$

for all $\alpha \in \mathbb{R}$, $\mathscr{P}(f) \leq \mathscr{P}(g)$ for all $f, g \in \mathbb{R}^S$,

$$\mu_0(\alpha) = \alpha$$

for all $\alpha \in \mathbb{R}$ and $\mu_0(f) \leq \mu_0(g)$ for all $f, g \in \mathbb{R}^S$. For all $k, l \in \mathbb{N}_0$ with $k \leq l$ we define

$$\mathcal{E}_{k,l}(\,\cdot\,, f) := \mathscr{P}^{l-k}(f)$$

for all $f \in \mathscr{L}^\infty(S, \mathscr{B})$. Then, $\mathcal{E}_{k,l} \colon \mathscr{L}^\infty(S, \mathscr{B}) \to \mathscr{L}^\infty(S, \mathscr{B})$ defines a convex kernel from $\mathscr{L}^\infty(S, \mathscr{B})$ to $\mathscr{L}^\infty(S, \mathscr{B})$ for all $k, l \in \mathbb{N}_0$ with $k \leq l$. Let $\mathscr{H} := \{J \subset \mathbb{N}_0 \colon |J| \in \mathbb{N}\}$ be the set of all finite, nonempty subsets of $\mathbb{N}_0$. For $k \in \mathbb{N}_0$ we define

$$\mathcal{E}_{\{k\}}(f) := \mathcal{E}_{\{0\}}(\mathcal{E}_{0,k}(\,\cdot\,, f)) = \mu_0(\mathscr{P}^k(f)) \quad (f \in \mathscr{L}^\infty(S, \mathscr{B})).$$

For $n \in \mathbb{N}$, $k_1, \ldots, k_{n+1} \in \mathbb{N}_0$ with $k_1 < \ldots < k_{n+1}$ and $f \in \mathscr{L}^\infty(S^{n+1}, \mathscr{B}^{n+1})$ we now define recursively

$$\mathcal{E}_{\{k_1, \ldots, k_{n+1}\}}(f) := \mathcal{E}_{\{k_1, \ldots, k_n\}}(g)$$

with

$$g(x_1, \ldots, x_n) := \mathcal{E}_{k_n, k_{n+1}}(x_n, f(x_1, \ldots, x_n, \,\cdot\,)) \quad (x_1, \ldots, x_n \in S).$$

Since $|S^J| < \infty$, the convex expectation $\mathcal{E}_J \colon \mathscr{L}^\infty(S^J, \mathscr{B}^J) \to \mathbb{R}$ is continuous from above for all $J \in \mathscr{H}$. As

$$\mathcal{E}_{k,m} = \mathcal{E}_{k,l}\mathcal{E}_{l,m}$$

for all $k, l, m \in \mathbb{N}_0$ with $k < l < m$, we obtain that the family $(\mathcal{E}_J)_{J\in\mathscr{H}}$ is consistent. Hence, by Theorem 2.10, there exists a nonlinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a stochastic process $(X_k)_{k\in\mathbb{N}_0}$ with

$$\mathcal{E}\big(f(X_0)\big) = \mu_0(f)$$

for all $f \in \mathscr{L}^\infty(S, \mathscr{B})$ and

$$\mathcal{E}\big(f(X_{k_1}, \ldots, X_{k_n}, X_l)\big) = \mathcal{E}\left(\mathcal{E}_{k,l}\big(X_k, f(X_{k_1}, \ldots, X_{k_n}, \,\cdot\,)\big)\right).$$

for all $n \in \mathbb{N}$, $0 \leq k_1 < \ldots < k_n \leq k < l$ and $f \in \mathscr{L}^\infty(S^{n+1}, \mathscr{B}^{n+1})$. The stochastic process $(X_k)_{k\in\mathbb{N}_0}$ can be seen as a convex time-homogeneous *Markov chain* with *initial distribution* μ_0 and *transition operator* $\mathscr{P}$ on the nonlinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$. If $\mathscr{P}$ is sublinear, the set

$$\big\{\mu \in \mathbb{R}^{S\times S} \colon \mu f \leq \mathscr{P}(f) \text{ for all } f \in \mathbb{R}^S\big\}$$

induces a Markov-set chain, see Hartfiel [46].

In the above example, when constructing $\mathcal{E}_J$, the measurability of g, in general, is not trivial. This motivates the following definition.

2.21 Definition. A *regular kernel* on S is a kernel $\mathcal{E} \colon C_b(S) \to \mathscr{L}^\infty(S, \mathscr{B})$ with the following properties:

(i) $\mathcal{E}$ is continuous from below,

(ii) For all $n \in \mathbb{N}$ and $f \in C_b(S^{n+1})$ the mapping

$$S^{n+1} \to \mathbb{R}, \quad (x_0, \ldots, x_n) \mapsto \mathcal{E}\big(x_0, f(x_1, \ldots, x_n, \,\cdot\,)\big)$$

is $\mathscr{B}^{n+1}$-measurable.

2.22 Remark. a) Let $\mathcal{E} \colon C_b(S) \to \mathrm{USC}_b(S)$ be a convex kernel which is continuous from above. Then, $\mathcal{E}$ is a regular kernel on S. In fact, let $n \in \mathbb{N}$ and $f \in \mathrm{BUC}(S^{n+1})$. Then, for all $\varepsilon > 0$, $x_0 \in S$ and $x' \in S$ there exists some $\delta > 0$ such that

$$\mathcal{E}\big(y_0, f(x', \,\cdot\,)\big) \leq \mathcal{E}\big(x_0, f(x', \,\cdot\,)\big) + \frac{\varepsilon}{2} \quad \text{for all } y_0 \in S \text{ with } d(y_0, x_0) \leq \delta$$

and

$$\sup_{u\in S} \big|f(x', u) - f(y', u)\big| \leq \frac{\varepsilon}{2} \quad \text{for all } y' \in S^n \text{ with } d_n(x', y') \leq \delta.$$

Therefore, we have that

$$\big|\mathcal{E}\big(y_0, f(x', \,\cdot\,)\big) - \mathcal{E}\big(y_0, f(y', \,\cdot\,)\big)\big| \leq \sup_{u\in S} |f(x', u) - f(y', u)| \leq \frac{\varepsilon}{2}$$

for all $y_0 \in S$ and $y' \in S^n$ with $d_n(x', y') \leq \delta$. This implies that

$$\mathcal{E}\big(y_0, f(y', \,\cdot\,)\big) \leq \mathcal{E}\big(y_0, f(x', \,\cdot\,)\big) + \frac{\varepsilon}{2} \leq \mathcal{E}\big(x_0, f(x', \,\cdot\,)\big) + \varepsilon$$

for all $y_0 \in S$ with $d(x_0, y_0) \leq \delta$ and $y' \in S^n$ with $d_n(x', y') \leq \delta$. By Remark 1.59 b), the assertion follows.

b) Let $\mu\colon C_b(S) \to \mathscr{L}^\infty(S)$ be a linear kernel which is continuous from below. Then, μ is a regular kernel on S. First, observe that for all $x \in S$ there exists a unique linear expectation $\overline{\mu}(x,\,\cdot\,)$ on $\mathscr{L}^\infty(S,\mathscr{B})$ which is continuous from below and extends $\mu(x,\,\cdot\,)$. By the Monotone Class Theorem 1.49, the mapping $\overline{\mu}\colon \mathscr{L}^\infty(S,\mathscr{B}) \to \mathscr{L}^\infty(S,\mathscr{B})$ is therefore well defined, i.e. $\overline{\mu}$ is a linear kernel from $\mathscr{L}^\infty(S,\mathscr{B})$ to $\mathscr{L}^\infty(S,\mathscr{B})$ which is continuous from below. Let $\mathscr{H}$ be the set of all $f \in \mathscr{L}^\infty(S^{n+1},\mathscr{B}^{n+1})$ such that the mapping

$$S^{n+1} \to \mathbb{R}, \quad (x_0,\ldots,x_n) \mapsto \overline{\mu}\big(x_0, f(x_1,\ldots,x_n,\,\cdot\,)\big)$$

is $\mathscr{B}^{n+1}$-measurable. As $\overline{\mu}$ is linear and continuous from below, $\mathscr{H}$ is a monotone class. Moreover,

$$\operatorname{span}\big\{1_{A\times B}\colon A \in \mathscr{B}^n, B \in \mathscr{B}\big\} \subset \mathscr{H}.$$

Again, by the Monotone Class Theorem 1.49, we thus obtain that $\mathscr{H} = \mathscr{L}^\infty(S^{n+1},\mathscr{B}^{n+1})$ showing that μ is a regular kernel on S.

2.23 Definition. A *(nonlinear) Markov process* is a quadruple $(\Omega,\mathscr{F},(\mathcal{E}^x)_{x\in S},(X_t)_{t\geq 0})$, where

(i) $X_t\colon \Omega \to S$ is $\mathscr{F}$-$\mathscr{B}$-measurable for all $t \geq 0$,

(ii) $(\Omega,\mathscr{F},\mathcal{E}^x)$ is a nonlinear expectation space with $\mathcal{E}^x \circ X_0^{-1} = \delta_x$ for each $x \in S$,

(iii) For all $0 \leq s < t$ there exists a regular kernel $\mathcal{E}_{s,t}$ on S (cf. Definition 2.21) such that for all $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_n \leq s$ and $f \in C_b(S^{n+1})$ we have that

$$\mathcal{E}^x\big(f(X_{t_1},\ldots,X_{t_n},X_t)\big) = \mathcal{E}^x\left(\mathcal{E}_{s,t}\big(X_s, f(X_{t_1},\ldots,X_{t_n},\,\cdot\,)\big)\right).$$

We say that the Markov process $(\Omega,\mathscr{F},(\mathcal{E}^x)_{x\in S},(X_t)_{t\geq 0})$ is *time-homogeneous* if (iii) is satisfied with

$$\mathcal{E}_{s,t}(x,f) = \mathcal{E}_{0,t-s}(x,f) = \mathcal{E}^x(f(X_{t-s})) \quad (x \in S,\ f \in C_b(S))$$

for all $0 \leq s < t$.

2.24 Remark. a) Let $(\Omega,\mathscr{F},(\mathcal{E}^x)_{x\in S},(X_t)_{t\geq 0})$ be a Markov process. Then, (iii) implies that $\mathcal{E}_{s,t}\mathcal{E}_{t,u} = \mathcal{E}_{s,u}$ for all $0 \leq s < t < u$.

b) Let $(\Omega,\mathscr{F},(\mathbb{E}^x)_{x\in S},(X_t)_{t\geq 0})$ be a linear Markov process in the sense of Definition 2.23 with

$$\mathbb{E}^x\big(f(X_{t_1},\ldots,X_{t_n},X_t)\big) = \mathbb{E}^x\left(\mathbb{E}_{s,t}\big(X_s, f(X_{t_1},\ldots,X_{t_n},\,\cdot\,)\big)\right)$$

for all $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_n \leq s$ and $f \in C_b(S^{n+1})$. Let $\mathscr{F}_t := \sigma(X_s\colon 0 \leq s \leq t)$. Then, (iii) implies that

$$\mathbb{E}^x(f(X_t)|\mathscr{F}_s) = \mathbb{E}_{s,t}(X_s,f)$$

for all $0 \leq s < t$ and $f \in \mathscr{L}^\infty(S)$.

2.25 Definition. Let $M \subset \mathscr{L}^\infty(S,\mathscr{B})$ with $1 \in M$. For all $0 \leq s < t$ let $\mathcal{E}_{s,t}$ be a nonlinear kernel from M to M. Then, we say that the family $(\mathcal{E}_{s,t})_{0\leq s<t}$ satisfies the *Chapman-Kolmogorov equations* if

$$\mathcal{E}_{s,u}(f) = \mathcal{E}_{s,t}(\mathcal{E}_{t,u}(f)) =: (\mathcal{E}_{s,t}\mathcal{E}_{t,u})(f)$$

for all $0 \leq s < t < u$ and $f \in M$.

2.26 Remark. Let $M \subset \mathscr{L}^\infty(S,\mathscr{B})$ with $1 \in M$ and $(\mathcal{E}_{s,t})_{0\le s<t}$ be a family of kernels from M to M with $\mathcal{E}_{s,s+t} = \mathcal{E}_{0,t} =: \mathcal{E}_t$ for all $s \ge 0$ and $t > 0$. Defining $\mathcal{E}_0 := \mathrm{id}_M$, the family $(\mathcal{E}_{s,t})_{0\le s<t}$ satisfies the Chapman-Kolmogorov equations if and only if $(\mathcal{E}_t)_{t\ge 0}$ satisfies the *semigroup property*, i.e.

$$\big(\mathcal{E}_s\mathcal{E}_t\big)(f) = \mathcal{E}_{s+t}(f)$$

for all $s,t \ge 0$ and $f \in M$. Since $\mathcal{E}_0 = \mathrm{id}_M$, this implies that $(\mathcal{E}_t)_{t\ge 0}$ defines a *semigroup* on M.

2.27 Remark. a) Let $\mu\colon \mathscr{L}^\infty(S,\mathscr{B}) \to \mathscr{L}^\infty(S,\mathscr{B})$ be a linear kernel. Then, the following statements are equivalent:

(i) $\mu f \in \mathrm{LSC}_b(S)$ for all $f \in \mathrm{LSC}_b(S)$,

(ii) $\mu f \in \mathrm{USC}_b(S)$ for all $f \in \mathrm{USC}_b(S)$,

(iii) $\mu f \in C_b(S)$ for all $f \in C_b(S)$.

b) Let $\mathcal{E}$ be a convex kernel from $\mathrm{Lip}_b(S)$ to $\mathrm{BUC}(S)$. Then, by Proposition 1.18, there exists exactly one convex kernel $\hat{\mathcal{E}}$ from $\mathrm{BUC}(S)$ to $\mathrm{BUC}(S)$ with $\hat{\mathcal{E}}|_{\mathrm{Lip}_b(S)} = \mathcal{E}$.

c) Let $\mathcal{E}$ be a convex kernel from $\mathrm{BUC}(S)$ to $C_b(S)$ which is continuous from above. Then, there exists exactly one convex kernel $\hat{\mathcal{E}}$ from $C_b(S)$ to $C_b(S)$ which is continuous from above and satisfies $\hat{\mathcal{E}}|_{\mathrm{BUC}(S)} = \mathcal{E}$. In fact, by Remark 1.59 b), there exists a convex kernel $\hat{\mathcal{E}}$ from $C_b(S)$ to $\mathscr{L}^\infty(S)$ which is continuous from above and extends $\mathcal{E}$. As $C_b(S)$ is a vector space, we get that $\hat{\mathcal{E}}$ is continuous from below, as well. Let $f \in C_b(S)$. By Remark 1.59 a), there exist sequences $(\varphi_n)_{n\in\mathbb{N}} \subset \mathrm{BUC}(S)$ and $(\psi_n)_{n\in\mathbb{N}} \subset \mathrm{BUC}(S)$ with $\varphi_n \searrow f$ and $\psi_n \nearrow f$ as $n \to \infty$. Therefore,

$$\inf_{n\in\mathbb{N}} \mathcal{E}(\varphi_n) = \hat{\mathcal{E}}(f) = \sup_{n\in\mathbb{N}} \mathcal{E}(\psi_n),$$

i.e. $\mathcal{E}(f) \in \mathrm{USC}_b(S) \cap \mathrm{LSC}_b(S) = C_b(S)$.

d) Let $\mathcal{E}$ be a convex kernel from $\mathrm{BUC}(S)$ to $\mathrm{USC}_b(S)$ which is continuous from above. Then, by Remark 1.59, there exists a convex kernel $\hat{\mathcal{E}}$ from $\mathrm{USC}_b(S)$ to $\mathrm{USC}_b(S)$ which is continuous from above and extends $\mathcal{E}$.

e) Let $(\mathcal{E}_{s,t})_{0\le s<t}$ be a family of convex kernels from $\mathrm{USC}_b(S)$ to $\mathrm{USC}_b(S)$ which are continuous from above. Then, by the uniqueness obtained in part b), c) and d), the following statements are equivalent:

(i) The family $(\mathcal{E}_{s,t})_{0\le s<t}$ satisfies the Chapman-Kolmogorov equations,

(ii) $\mathcal{E}_{s,u}(f) = \mathcal{E}_{s,t}(\mathcal{E}_{t,u}(f))$ for all $0 \le s < t < u$ and $f \in C_b(S)$,

(iii) $\mathcal{E}_{s,u}(f) = \mathcal{E}_{s,t}(\mathcal{E}_{t,u}(f))$ for all $0 \le s < t < u$ and $f \in \mathrm{BUC}(S)$,

(iv) $\mathcal{E}_{s,u}(f) = \mathcal{E}_{s,t}(\mathcal{E}_{t,u}(f))$ for all $0 \le s < t < u$ and $f \in \mathrm{Lip}_b(S)$.

Hence, for $M \in \{\mathrm{Lip}_b(S), \mathrm{BUC}(S), C_b(S)\}$ the extension of families of convex kernels from M to M, which are continuous from above, is included in the case $M = \mathrm{USC}_b(S)$. We therefore restrict ourselves in the following to the extension of families of convex kernels from $\mathrm{USC}_b(S)$ to $\mathrm{USC}_b(S)$ which are continuous from above.

2.28 Theorem. *For all* $0 \leq s < t$ *let* $\mathcal{E}_{s,t}\colon \mathrm{USC}_b(S) \to \mathrm{USC}_b(S)$ *be a convex kernel which is continuous from above, and assume that the family* $(\mathcal{E}_{s,t})_{0\leq s<t}$ *satisfies the Chapman-Kolmogorov equations. Then,* $\mathcal{E}_{s,t}|_{C_b(S)}$ *is a regular kernel for all* $0 \leq s < t$ *and there exists a convex Markov process* $(\Omega, \mathscr{F}, (\mathcal{E}^x)_{x\in S}, (X_t)_{t\geq 0})$ *with*

$$\mathcal{E}^x\big(f(X_{t_1},\ldots,X_{t_n},X_t)\big) = \mathcal{E}^x\left(\mathcal{E}_{s,t}\big(X_s, f(X_{t_1},\ldots,X_{t_n},\,\cdot\,)\big)\right)$$

for all $0 \leq s < t$, $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_n \leq s$ *and* $f \in C_b(S^{n+1})$. *If* $\mathcal{E}_{s,t}$ *is sublinear for all* $0 \leq s < t$, *the Markov process* $(\Omega, \mathscr{F}, (\mathcal{E}^x)_{x\in S}, (X_t)_{t\geq 0})$ *is sublinear as well.*

Proof. Let $\mathscr{H} := \{J \subset [0,\infty)\colon |J| \in \mathbb{N}\}$ be the set of all finite, nonempty subsets of $[0,\infty)$ and $\mathcal{E}_0^x(f) := f(x)$ for all $f \in C_b(S)$ and $x \in S$. For $0 < t < \infty$ we define

$$\mathcal{E}_{\{t\}}(f) := \mathcal{E}_0^x(\mathcal{E}_{0,t}(f)) = \mathcal{E}_{0,t}(x,f) \quad \text{for all } f \in C_b(S).$$

For $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_{n+1} < \infty$ and $f \in C_b(S^{n+1})$, we recursively define

$$\mathcal{E}^x_{\{t_1,\ldots,t_{n+1}\}}(f) := \mathcal{E}^x_{\{t_1,\ldots,t_n\}}(g),$$

where

$$g(x_1,\ldots,x_n) := \mathcal{E}_{t_n,t_{n+1}}(x_n, f(x_1,\ldots,x_n,\,\cdot\,))$$

for all $x_1,\ldots,x_n \in S$. Note that $g \in \mathscr{L}^\infty(S^n)$ as $\mathcal{E}_{t_n,t_{n+1}}|_{C_b(S)}$ is a regular kernel by Remark 2.22 a). Then, $\mathcal{E}^x_J\colon C_b(S^J) \to \mathbb{R}$ is a convex pre-expectation which is continuous from above for all $J \in \mathscr{H}$. By the Chapman-Kolmogorov equations and Remark 2.2 b), we obtain that the family $(\mathcal{E}^x_J)_{J\in\mathscr{H}}$ is consistent and the statement follows from Theorem 2.10 with $M_J = C_b(S^J)$ for all $J \in \mathscr{H}$ and $X_t = \mathrm{pr}_{\{t\}}$ for $t \geq 0$. □

2.29 Example. Let $G \subset \mathbb{R}^n$ be a bounded domain and $(T(t))_{t\geq 0}$ a positive C_0-semigroup on $C(\overline{G}) = C_b(\overline{G})$. Moreover, assume that $T(t)1 = 1$ for all $t \geq 0$. Then,

$$\mathbb{E}_{s,t}(f) := T(t-s)f \quad \big(f \in C(\overline{G})\big)$$

defines a linear kernel on $C(\overline{G})$, which is continuous from above (see Remark 1.28 a)) for all $0 \leq s < t$. Moreover, for all $0 \leq s < t < u$ and $f \in C(\overline{G})$ we have that

$$\big(\mathbb{E}_{s,t}\mathbb{E}_{t,u}\big)(f) = T(t-s)T(u-t)f = T(u-s)f = \mathbb{E}_{s,u}(f).$$

Hence, there exists a time-homogeneous Markov process $(\Omega, \mathscr{F}, (\mathbb{E}^x)_{x\in S}, (X_t)_{t\geq 0})$ with

$$\mathbb{E}(f(X_t)|\mathscr{F}_s) = \big(T(t-s)f\big)(X_s)$$

for all $0 \leq s < t$ and $f \in \mathscr{L}^\infty(S)$, where $(\mathscr{F}_t)_{t\geq 0}$ is the canonical filtration. An example for such a semigroup is the semigroup generated by the heat equation with homogeneous Neumann boundary conditions on $(0,1)$, i.e.

$$\begin{aligned} u_t(t,x) &= u_{xx}(t,x), && t > 0,\ x \in [0,1],\\ u_x(t,0) &= 0, && t \geq 0,\\ u_x(t,1) &= 0, && t \geq 0,\\ u(0,x) &= u_0(x), && x \in [0,1]. \end{aligned}$$

More precisely, we consider the operator $A\colon D(A)\subset C([0,1])\to C([0,1])$, given by $Au:=u_{xx}$ for $u\in D(A):=\big\{v\in C^2([0,1])\colon v_x(0)=v_x(1)=0\big\}$. For $\mu>0$, the resolvent operator $R(\mu^2,A):=(\mu^2-A)^{-1}$ of A at μ^2 is given by

$$\big(R(\mu^2,A)f\big)(x)=\frac{\cosh(\mu x)}{\sinh\mu}\int_0^1\frac{\cosh(\mu\big(1-y)\big)}{\mu}f(y)\,\mathrm{d}y-\int_0^x\frac{\sinh\big(\mu(x-y)\big)}{\mu}f(y)\,\mathrm{d}y$$

for $f\in C([0,1])$. Then, we have that $\mu^2R(\mu^2,A)1=1$ for all $\mu>0$, and for $f\in C([0,1])$ with $f\geq 0$ we have that

$$\begin{aligned}\big(R(\mu^2,A)f\big)(x)&\geq\frac{1}{\mu\sinh\mu}\int_0^1\Big(\cosh(\mu x)\cosh\big(\mu(1-y)\big)-\sinh\mu\sinh\big(\mu(x-y)\big)\Big)f(y)\,\mathrm{d}y\\&=\frac{1}{\mu\sinh\mu}\int_0^1\cosh\big(\mu(1-x)\big)\cosh(\mu y)f(y)\,\mathrm{d}y\\&\geq 0.\end{aligned}$$

Hence, for all $\mu>0$ we get that $R(\mu^2,A)$ is positive and therefore $R(\lambda,A)\in L(C([0,1]))$ with

$$\|\mu^2R(\lambda,A)\|_{L(C([0,1]))}\leq\mu^2R(\mu^2,A)1=1.$$

As $D(A)$ is dense in $C([0,1])$ by the Stone-Weierstraß Theorem (cf. [70, Theorem 7.32]), we obtain that A generates a positive C_0-semigroup of contractions $(T(t))_{t\geq0}$ on $C([0,1])$ by the Hille-Yosida Theorem (cf. [63, Section 1.3]). Moreover, as $\lambda R(\lambda,A)1=1$ for all $\lambda>0$, we get that

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

2.3 Lévy processes

In the sequel, let G be an abelian group and d be a translation invariant metric on G, i.e.

$$d(x+z, y+z) = d(x, y)$$

for all $x, y, z \in G$. Then, G is a topological group and we further assume that G is Polish. Typical examples for G are $\mathbb{R}^n$, $\mathbb{Z}^n$ and the n-dimensional Torus $\mathbb{T}^n$ for $n \in \mathbb{N}$ or any separable Banach space.

In this section, we consider Lévy processes as a special case of Markov processes. For a detailed discussion of Lévy proceses in the linear case we refer to Applebaum [2] or Sato [72]. As before, we denote by $\mathscr{L}^\infty(G, \mathscr{B})$ the space of all bounded measurable functions $G \to \mathbb{R}$, where $\mathscr{B}$ is the Borel σ-algebra on G, by $C_b(G)$ the space of all bounded continuous functions $G \to \mathbb{R}$ and by $\mathrm{BUC}(G)$ the space of all bounded uniformly continuous functions $G \to \mathbb{R}$.

2.30 Definition. a) Let $\mathcal{E}\colon \mathscr{L}^\infty(G, \mathscr{B}) \to \mathbb{R}$ be an expectation. Then, we say that $\mathcal{E}$ is *tight* if for all $\varepsilon > 0$ there exists a compact set $K \subset G$ with $\mathcal{E}\big(1_{G\setminus K}\big) < \varepsilon$ (cp. Remark 1.47).

b) We say that an operator $\mathscr{S}\colon \mathrm{BUC}(G) \to \mathrm{BUC}(G)$ is a *Markovian convolution* if there exists a expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a random variable $X\colon \Omega \to G$ such that

$$\big(\mathscr{S} f\big)(x) = \mathcal{E}(f(x+X))$$

for all $f \in \mathrm{BUC}(G)$ and $x \in G$. We say that $\mathscr{S}$ is *tight* if $\mathcal{E} \circ X^{-1}$ is tight. Notice that any linear Markovian convolution is tight.

Let $(\Omega, \mathscr{F}, \mathcal{E})$ be a convex expectation space and $Y\colon \Omega \to G$ a random variable. Further, let S be a Polish space and $X\colon \Omega \to S$ a random variable. Then, for all $f \in C_b(S \times G)$ we have that

$$S \to \mathbb{R}, \quad x_0 \mapsto \mathcal{E}(f(x, Y))|_{x=x_0} := \mathcal{E}(f(x_0, Y))$$

is bounded and lower semicontinuous and therefore, $\mathcal{E}\big(\mathcal{E}(f(x,Y))|_{x=X}\big)$ is well-defined. We say that Y is *independent* of X if

$$\mathcal{E}(f(X,Y)) = \mathcal{E}\big(\mathcal{E}(f(x,Y))|_{x=X}\big)$$

for all $f \in C_b(S \times G)$.

2.31 Definition. a) We say that $(\mathscr{S}(t))_{t\geq 0}$ is a *Markovian convolution semigroup* if

(i) $\mathscr{S}(t)$ is a Markovian convolution for all $t \geq 0$,

(ii) $\mathscr{S}(0)f = f$ for all $f \in \mathrm{BUC}(G)$,

(iii) $\mathscr{S}(s+t) = \mathscr{S}(s)\mathscr{S}(t)$ for all $s, t \geq 0$,

(iv) For all $f \in \mathrm{BUC}(G)$ we have that $\lim_{t\searrow 0} \|\mathscr{S}(t)f - f\|_\infty = 0$.

In this case, we say that $(\mathscr{S}(t))_{t\geq 0}$ is *tight* if $\mathscr{S}(t)$ is tight for all $t \geq 0$.

b) Let $(\Omega, \mathscr{F}, \mathcal{E})$ be a convex expectation space. Then, $(X_t)_{t\geq 0}$ is called an $\mathcal{E}$-*Lévy process* if

(i) $X_t \colon \Omega \to G$ is $\mathscr{F}$-$\mathscr{B}$-measurable for all $t \geq 0$,

(ii) We have that $\mathcal{E} \circ X_0^{-1} = \delta_0$, i.e. $\mathcal{E}(f(X_0)) = f(0)$ for all $f \in \mathscr{L}^\infty(G, \mathscr{B})$,

(iii) We have that $\mathcal{E} \circ (X_{s+t} - X_s)^{-1} = \mathcal{E} \circ X_t^{-1}$ for all $s, t \geq 0$,

(iv) For all $s, t \geq 0$, $n \in \mathbb{N}$, $0 \leq t_1 \leq \ldots \leq t_n \leq s$ we have that $X_{s+t} - X_s$ is independent of $(X_{t_1}, \ldots, X_{t_n})$,

(v) $\mathcal{E}(f(X_t)) \to f(0)$ for all $f \in C_b(G)$, i.e. $X_t \to X_0$ in distribution as $t \searrow 0$.

Note that, as any convex Markovian convolution is 1-Lipschitz, condition (iv) in Definition 2.31 a) is equivalent to

(iv′) For all $f \in \mathrm{BUC}(G)$ the mapping $[0, \infty) \to \mathrm{BUC}(G)$, $t \mapsto \mathscr{S}(t)f$ is continuous,

for convex Markovian convolution semigroups.

2.32 Theorem. *a) Let $(\Omega, \mathscr{F}, \mathcal{E})$ be a convex expectation space and $(X_t)_{t \geq 0}$ an $\mathcal{E}$-Lévy process. For $t \geq 0$, $f \in \mathrm{BUC}(G)$ and $x \in G$ define*

$$\big(\mathscr{S}(t)f\big)(x) := \mathcal{E}(f(x + X_t)).$$

Then, $(\mathscr{S}(t))_{t \geq 0}$ is a convex Markovian convolution semigroup.

b) Let $(\mathscr{S}(t))_{t \geq 0}$ be a tight convex Markovian convolution semigroup. Then, there exists a convex expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t \geq 0}$ such that

$$\big(\mathscr{S}(t)f\big)(x) = \mathcal{E}(f(x + X_t))$$

for all $t \geq 0$, $f \in \mathrm{BUC}(G)$ and $x \in G$. Moreover, $\mathcal{E} \circ X_t^{-1}$ is tight for all $t \geq 0$.

Proof. a) By definition, we have that $\mathscr{S}(t)$ is a convex Markovian convolution for all $t \geq 0$. Moreover, $\mathscr{S}(0)f = f$ for all $f \in \mathrm{BUC}(G)$. Therefore $(\mathscr{S}(t))_{t \geq 0}$ satisfies (i) and (ii) in Definition 2.31 a). Let $s, t \geq 0$ and $f \in \mathrm{BUC}(G)$. Then, we have that

$$\mathcal{E}(f(x + (X_{t+s} - X_s)))|_{x=x_0} = \mathcal{E}(f(x_0 + (X_{t+s} - X_s))) = \mathcal{E}(f(x_0 + X_t)) = \big(\mathscr{S}(t)f\big)(x_0)$$

for all $x_0 \in G$. As $X_{t+s} - X_s$ is independent of X_s we thus get that

$$\begin{aligned}\big(\mathscr{S}(t+s)f\big)(x) = \mathcal{E}(f(x + X_{t+s})) &= \mathcal{E}\big(f\big((x + X_s) + (X_{t+s} - X_s)\big)\big) \\ &= \mathcal{E}\big(\big(\mathscr{S}(t)f\big)(x + X_s)\big) = \big(\mathscr{S}(s)\mathscr{S}(t)f\big)(x)\end{aligned}$$

for all $x \in G$, i.e. $\mathscr{S}(t+s) = \mathscr{S}(s)\mathscr{S}(t)$. In order to show that (iv) is satisfied, we first show that for all $C \geq 0$ and $\delta > 0$ we have that

$$\mathcal{E}\big(C 1_{G \setminus B(0,\delta)}(X_t)\big) \to 0$$

as $t \searrow 0$. Let $\varphi \colon G \to \mathbb{R}$ be defined by

$$\varphi(y) := \frac{C}{\delta}\big(d(y, 0) \wedge \delta\big)$$

for $y \in G$. Then, $\varphi \in C_b(G)$ with $\varphi \geq 0$, $\varphi(0) = 0$ and $\varphi(y) = C$ for all $y \in G \setminus B(0,\delta)$. We thus have that

$$0 \leq \mathcal{E}\big(C1_{G\setminus B(0,\delta)}(X_t)\big) = \mathcal{E}\big(\varphi(X_t)1_{G\setminus B(0,\delta)}(X_t)\big) \leq \mathcal{E}(\varphi(X_t)) \to 0$$

as $t \searrow 0$. Now, let $\varepsilon > 0$. As $f \in \mathrm{BUC}(G)$, there exists some $\delta > 0$ such that

$$|f(x+y) - f(x)|_\infty < \frac{\varepsilon}{2}$$

for all $x \in G$ and all $y \in G$ with $d(y,0) < \delta$, thanks to the assumption on the metric d. For $x \in G$ let

$$g_x \colon G \to \mathbb{R}, \quad y \mapsto 1_{B(0,\delta)}(y)\big(f(x+y) - f(x)\big).$$

Then, we have that $g_x \in \mathscr{L}^\infty(G)$ with

$$\|g_x\|_\infty \leq \frac{\varepsilon}{2}$$

for all $x \in G$. Moreover, let $t_0 > 0$ with

$$\mathcal{E}\big(2\|f\|_\infty 1_{G\setminus B(0,\delta)}(X_t)\big) \leq \frac{\varepsilon}{2}$$

for all $0 < t < t_0$. By cash additivity and Lemma 1.17, we get that

$$\begin{aligned}\big|\big(\mathscr{S}(t)f\big)(x) - f(x)\big| &= \big|\mathcal{E}\big(f(x+X_t) - f(x)\big)\big| \leq \mathcal{E}\big(|f(x+X_t) - f(x)|\big) \\ &\leq \mathcal{E}\big(g_x(X_t) + 2\|f\|_\infty 1_{G\setminus B(0,\delta)}(X_t)\big) \\ &\leq \|g_x\|_\infty + \mathcal{E}\big(2\|f\|_\infty 1_{G\setminus B(0,\delta)}(X_t)\big) \leq \varepsilon\end{aligned}$$

for all $0 < t < t_0$, i.e. $\|\mathscr{S}(t)f - f\|_\infty \leq \varepsilon$ for all $0 < t < t_0$.

b) By Lemma 1.48 and Theorem 2.28, there exists a convex expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and a stochastic process $(X_t)_{t\geq 0}$ which satisfies (i), (ii), (iii) and (iv) in Definition 2.31 and

$$\big(\mathscr{S}(t)f\big)(x) = \mathcal{E}(f(x+X_t))$$

for all $f \in C_b(G)$ and $x \in G$. It remains to verify (v) in Definition 2.31. Let $f \in C_b(G)$. Then, by Remark 1.59, there exists a sequence $(\psi_n)_{n\in\mathbb{N}} \subset \mathrm{BUC}(G)$ with $\psi_n \nearrow f$ as $n \to \infty$. Hence, there exists some $n \in \mathbb{N}$ with

$$\big|\big(\mathscr{S}(t)f\big)(0) - \big(\mathscr{S}(t)\psi_n\big)(0)\big| < \frac{\varepsilon}{3}$$

and

$$|f(0) - \psi_n(0)| < \frac{\varepsilon}{3}.$$

Moreover, there exists some $t_0 > 0$ such that

$$\|\mathscr{S}(t)\psi_n - \psi_n\|_\infty < \frac{\varepsilon}{3}$$

for all $0 < t < t_0$. Therefore, we get that

$$|\mathcal{E}(f(X_t)) - f(0)| \leq \big|\big(\mathscr{S}(t)f\big)(0) - \big(\mathscr{S}(t)\psi_n\big)(0)\big| + |f(0) - \psi_n(0)| + \|\mathscr{S}(t)\psi_n - \psi_n\|_\infty < \varepsilon$$

for all $0 < t < t_0$.

□

We close this section with two examples of linear Markovian convolution semigroups, which we will pick up later in a nonlinear setup.

2.33 Example. As in Example 2.29, we consider the heat equation in $(0,1)$ but this time with periodic boundary conditions. That is, we consider the boundary value problem

$$\begin{aligned} u_t(t,x) &= u_{xx}(t,x), && t>0,\ x\in[0,1],\\ u(t,0) &= u(t,1), && t\geq 0,\\ u_x(t,0) &= u_x(t,1), && t\geq 0,\\ u(0,x) &= u_0(x), && x\in[0,1]. \end{aligned}$$

Choosing $[0,1)$ as a representation of the 1-dimensional Torus $\mathbb{T}$, this leads to the operator $A\colon D(A)\subset C(\mathbb{T})\to C(\mathbb{T})$, given by $Au:=u_{xx}$ for $u\in D(A):=C^2(\mathbb{T})$. Note that $C(\mathbb{T})=C_b(\mathbb{T})$. Then, for $\mu>0$ the resolvent operator $R(\mu^2,A):=(\mu^2-A)^{-1}$ of A at μ^2 is given by

$$\begin{aligned} \big(R(\mu^2,A)f\big)(x) = \frac{e^\mu}{e^\mu-1}\int_0^1 \frac{e^{\mu(x-y)}}{2\mu}f(y)\,\mathrm{d}y + \frac{1}{e^\mu-1}\int_0^1\frac{e^{-\mu(x-y)}}{2\mu}f(y)\,\mathrm{d}y\\ -\int_0^x \frac{\sinh\big(\mu(x-y)\big)}{\mu}f(y)\,\mathrm{d}y \end{aligned}$$

for $f\in C(\mathbb{T})$. One readily verifies that $R(\mu^2,A)$ is positive with $\mu^2R(\mu^2,A)1=1$ for all $\mu>0$. By the theorem of Stone-Weierstraß, $C^2(\mathbb{T})$ is dense in $C(\mathbb{T})$. Hence, A generates a tight linear Markovian convolution semigroup on $\mathbb{T}$ which can be extended to a Lévy process on $\mathbb{T}$.

2.34 Example. In a similar way as in Example 2.33, considering the boundary value problem

$$\begin{aligned} u_t(t,x) &= u_x(t,x), && t>0,\ x\in[0,1],\\ u(t,0) &= u(t,1), && t\geq 0,\\ u(0,x) &= u_0(x), && x\in[0,1] \end{aligned}$$

leads to the operator $A\colon C^1(\mathbb{T})\subset C(\mathbb{T})\to C(\mathbb{T})$, given by $Au:=u_x$ for $u\in C^1(\mathbb{T})$. Here, for $\mu>0$ the resolvent operator $R(\mu,A)$ is given by

$$\big(R(\mu,A)f\big)(x) = \frac{e^\mu}{e^\mu-1}\int_0^1 e^{\mu(x-y)}f(y)\,\mathrm{d}y - \int_0^x e^{\mu(x-y)}f(y)\,\mathrm{d}y.$$

for $f\in C(\mathbb{T})$. As we saw in Example 2.33, $C^1(\mathbb{T})$ is dense in $C(\mathbb{T})$ and therefore, A generates a tight linear Markovian convolution semigroup on $\mathbb{T}$ which can be extended to a Lévy process on $\mathbb{T}$.

Chapter 3

A semigroup theoretic approach to fully nonlinear PDEs

3.1 Nisio semigroups

In this section, we explicitly construct Levy processes, whose semigroups are a viscosity solutions of fully nonlinear partial differential equations related to stochastic optimal control (see e.g. [38], [60]). Here, we will use a slightly different notion of a viscosity solution, which is made to fit into a semigroup setting. However, in many cases, particularly in the classical setup, this leads to the same class or an even larger class of test functions. We refer to Crandall et al. [21] for the classical definition and a detailed discussion of viscosity solutions. We basically follow an idea by Nisio [59] in order to construct a sublinear Markovian convolution semigroup which results from a given family of linear Markovian convolution semigroups by constant optimization. In [59] Nisio considered strongly continuous semigroups on the space of all bounded measurable functions. However, by a theorem of Lotz (see e.g. [3, Corollary 4.3.19]), all strongly continuous semigroups on the space of all bounded measurable functions already have a bounded generator, which is not suitable for most applications.

Again, let G be an abelian group and d be a translation invariant metric on G, i.e.

$$d(x+z, y+z) = d(x,y)$$

for all $x, y, z \in G$, which makes G a Polish space. We use the notation from Section 2.3. Throughout this section, we make the following two assumptions:

(A1) For each $\lambda \in \Lambda$ let $A_\lambda \colon D(A_\lambda) \subset \mathrm{BUC}(G) \to \mathrm{BUC}(G)$ be the generator of a linear Markovian convolution semigroup $(S_\lambda(t))_{t\geq 0}$ (cf. Definition 2.31 a)).

(A2) The subspace

$$D := \left\{ f \in \bigcap_{\lambda\in\Lambda} D(A_\lambda) \colon \{A_\lambda f \colon \lambda \in \Lambda\} \text{ is bounded and uniformly equicontinuous} \right\}$$

is dense in $\mathrm{BUC}(G)$.

We consider finite partitions $P := \{\pi \subset [0,\infty) \colon 0 \in \pi,\, |\pi| < \infty\}$. For a partition $\pi \in P$, $\pi = \{t_0, t_1, \ldots, t_m\}$ with $0 = t_0 < t_1 < \ldots < t_m$ we set

$$|\pi|_\infty := \max_{j=1,\ldots,m} (t_j - t_{j-1}).$$

Moreover, we define $|\{0\}|_\infty := 0$. The set of partitions with end-point t will be denoted by P_t, i.e. $P_t := \{\pi \in P : \max \pi = t\}$. Note that

$$P = \bigcup_{t \geq 0} P_t.$$

Let $f \in \mathrm{BUC}(G)$. For $h \geq 0$ and $x \in G$ we define

$$\big(\mathcal{E}_h f\big)(x) := \sup_{\lambda \in \Lambda} \big(S_\lambda(h) f\big)(x)$$

and for a partition $\pi \in P$ as above we set

$$\mathcal{E}_\pi f := \mathcal{E}_{t_1 - t_0} \ldots \mathcal{E}_{t_m - t_{m-1}} f$$

with $\mathcal{E}_{\{0\}} f := f$. Note that $\mathcal{E}_h = \mathcal{E}_{\{0,h\}}$ for $h > 0$.

3.1 Lemma. *a) $\mathcal{E}_\pi$ is a sublinear Markovian convolution for all $\pi \in P$.*

b) Let $f \in D$ and let $L_f := \sup_{\lambda \in \Lambda} \|A_\lambda f\|_\infty$. Then,

$$\|\mathcal{E}_{h_1}(f) - \mathcal{E}_{h_2}(f)\|_\infty \leq L_f \cdot |h_1 - h_2| \quad \textit{for all}\, h_1, h_2 \geq 0 \tag{3.1}$$

and

$$\|\mathcal{E}_\pi f - f\|_\infty \leq L_f \cdot \max \pi \quad \textit{for all}\, \pi \in P. \tag{3.2}$$

Proof. a) As $S_\lambda(h)$ is a linear Markovian convolution for all $\lambda \in \Lambda$ and all $h \geq 0$, $\mathcal{E}_h$ is a sublinear Markovian convolution for all $h \geq 0$. By Lemma 2.19 this property is preserved under compositions and therefore, the same holds for $\mathcal{E}_\pi$.

b) First, notice that $L_f < \infty$ as $f \in D$. Let $h_1, h_2 \geq 0$. Then, for all $x \in G$ and all $\lambda_0 \in \Lambda$ we have that

$$\begin{aligned}\big(S_{\lambda_0}(h_1)\big) f(x) - \big(\mathcal{E}_{h_2}(f)\big)(x) &\leq \big(S_{\lambda_0}(h_1)\big) f(x) - \big(S_{\lambda_0}(h_2)\big) f(x) \\ &\leq \sup_{\lambda \in \Lambda} \|S_\lambda(h_1) f - S_\lambda(h_2) f\|_\infty\end{aligned}$$

and therefore, taking the supremum over $\lambda_0 \in \Lambda$, we get that

$$\big(\mathcal{E}_{h_1}(f)\big)(x) - \big(\mathcal{E}_{h_2}(f)\big)(x) \leq \sup_{\lambda \in \Lambda} \|S_\lambda(h_1) f - S_\lambda(h_2) f\|_\infty$$

for all $x \in G$. By symmetry and taking the supremum over all $x \in G$, we thus get that

$$\|\mathcal{E}_{h_1}(f) - \mathcal{E}_{h_2}(f)\|_\infty \leq \sup_{\lambda \in \Lambda} \|S_\lambda(h_1) f - S_\lambda(h_2) f\|_\infty.$$

Let $\lambda \in \Lambda$ and w.l.o.g. let $h_1 < h_2$. As $f \in D(A_\lambda)$, we get that (see e.g. [34, Lemma II.1.3])

$$\begin{aligned}\|S_\lambda(h_1) f - S_\lambda(h_2) f\|_\infty &= \left\| \int_{h_1}^{h_2} S_\lambda(s) A_\lambda f \, \mathrm{d}s \right\|_\infty \leq \int_{h_1}^{h_2} \|S_\lambda(s) A_\lambda f\|_\infty \, \mathrm{d}s \\ &\leq \|A_\lambda f\|_\infty \cdot (h_2 - h_1) \leq L_f \cdot |h_1 - h_2|.\end{aligned}$$

Taking the supremum over all $\lambda \in \Lambda$, we obtain that

$$\|\mathcal{E}_{h_1}(f) - \mathcal{E}_{h_2}(f)\|_\infty \leq L_f \cdot |h_1 - h_2|.$$

Next, we show that

$$\|\mathcal{E}_\pi(f) - f\|_\infty \leq L_f \cdot \max \pi$$

for all $\pi \in P$ by an induction on $\#\pi \in \mathbb{N}$. First, let $\pi \in P$ with $\#\pi = 1$, i.e. $\pi = \{0\}$. Then, we have that

$$\|\mathcal{E}_\pi(f) - f\|_\infty = \|\mathcal{E}_{\{0\}}(f) - f\|_\infty = 0 = L_f \cdot 0 = L_f \cdot \max \pi.$$

Now, let $m \in \mathbb{N}$ and assume that (3.2) holds for all $\pi \in P$ with $\#\pi = m$ and consider $\pi \in P$ with $\#\pi = m+1$ and $t_m := \max \pi$. Then $\pi' := \pi \setminus \{t_m\} \in P$ with $\#\pi' = m$ and $t_{m-1} := \max \pi' \in [0, t_m)$. We thus have that

$$\mathcal{E}_\pi(f) = \mathcal{E}_{\pi'}\mathcal{E}_{t_m - t_{m-1}}(f) \tag{3.3}$$

and therefore, by induction hypothesis, (3.1) and (3.3), we get that

$$\begin{aligned}
\|\mathcal{E}_\pi(f) - f\|_\infty &\leq \|\mathcal{E}_\pi(f) - \mathcal{E}_{\pi'}(f)\|_\infty + \|\mathcal{E}_{\pi'}(f) - f\|_\infty \\
&= \|\mathcal{E}_{\pi'}\mathcal{E}_{t_m - t_{m-1}}(f) - \mathcal{E}_{\pi'}(f)\|_\infty + \|\mathcal{E}_{\pi'}(f) - f\|_\infty \\
&\leq \|\mathcal{E}_{t_m - t_{m-1}}(f) - f\|_\infty + \|\mathcal{E}_{\pi'}(f) - f\|_\infty \\
&\leq L_f \cdot (t_m - t_{m-1}) + L_f \cdot \max \pi' \\
&= L_f \cdot t_m = L \cdot \max \pi.
\end{aligned}$$

□

The following lemma shows that $\mathcal{E}_h$ depends continuously on the step size $h \geq 0$ and that $\mathcal{E}_\pi$ depends continuously on the partition $\pi \in P$.

3.2 Lemma. *a) Let $h \geq 0$ and $(h_n)_{n\in\mathbb{N}} \subset [0,\infty)$ with $h_n \to h$ as $n \to \infty$. Then, for all $f \in \mathrm{BUC}(G)$ we have that*

$$\|\mathcal{E}_h(f) - \mathcal{E}_{h_n}(f)\|_\infty \to 0, \quad n \to \infty.$$

b) Let $m \in \mathbb{N}$ and $\pi = \{t_0, t_1, \ldots, t_m\} \in P$ with $0 = t_0 < \ldots < t_m$. For each $n \in \mathbb{N}$ let $\pi_n = \{t_0^n, t_1^n, \ldots, t_m^n\} \in P$ with $0 = t_0^n < t_1^n < \ldots < t_m^n$ and $t_i^n \to t_i$ as $n \to \infty$ for all $i \in \{1, \ldots, m\}$. Then, for all $f \in \mathrm{BUC}(G)$ we have that

$$\|\mathcal{E}_\pi(f) - \mathcal{E}_{\pi_n}(f)\|_\infty \to 0, \quad n \to \infty.$$

Proof. a) Let $f \in \mathrm{BUC}(G)$ and $\varepsilon > 0$. Then, by Assumption (A2), there exists some $f_0 \in D$ with $\|f - f_0\|_\infty \leq \frac{\varepsilon}{3}$. Let $L_0 := \sup_{\lambda\in\Lambda} \|A_\lambda f_0\|_\infty$. Then, $L_0 < \infty$ as $f_0 \in D$ and we have that

$$\begin{aligned}
\|\mathcal{E}_h(f) - \mathcal{E}_{h_n}(f)\|_\infty &\leq \|\mathcal{E}_h(f) - \mathcal{E}_h(f_0)\|_\infty + \|\mathcal{E}_{h_n}(f) - \mathcal{E}_{h_n}(f_0)\|_\infty + \|\mathcal{E}_h(f_0) - \mathcal{E}_{h_n}(f_0)\|_\infty \\
&\leq \frac{2\varepsilon}{3} + L_0 \cdot |h - h_n| < \varepsilon
\end{aligned}$$

for all $n \in \mathbb{N}$ with $L_0 \cdot |h - h_n| < \frac{\varepsilon}{3}$.

b) First note that the set of all partitions with cardinality $m+1$ can be identified with the set

$$S^m := \big\{(s_1, \ldots, s_m) \in \mathbb{R}^m \colon 0 < s_1 < \ldots < s_m\big\} \subset \mathbb{R}^m.$$

Therefore, the assertion is equivalent to the continuity of the map

$$S^m \to \mathrm{BUC}(G), \quad (s_1, \ldots, s_m) \to \mathcal{E}_{\{0,s_1,\ldots,s_m\}}. \tag{3.4}$$

By part a), the mapping

$$[0,\infty) \to \mathrm{BUC}(G), \quad h \mapsto \mathcal{E}_h(f)$$

is continuous, which implies the continuity of (3.4).

□

Let $f \in \mathrm{BUC}(G)$. In the following, we now consider the limit of $\mathcal{E}_\pi(f)$ when the mesh size of the partition $\pi \in P$ tends to zero. For this, first note that for $h_1, h_2 \geq 0$ and $x \in G$ we have that

$$\begin{aligned}\big(\mathcal{E}_{h_1+h_2}(f)\big)(x) &= \sup_{\lambda\in\Lambda} \big(S_\lambda(h_1+h_2)f\big)(x) = \sup_{\lambda\in\Lambda} \big(S_\lambda(h_1)S_\lambda(h_2)f\big)(x)\\ &\leq \sup_{\lambda\in\Lambda} \big(S_\lambda(h_1)\mathcal{E}_{h_2}(f)\big)(x) = \big(\mathcal{E}_{h_1}\mathcal{E}_{h_2}(f)\big)(x),\end{aligned}$$

which implies the pointwise inequality

$$\mathcal{E}_{\pi_1}(f) \leq \mathcal{E}_{\pi_2}(f) \tag{3.5}$$

for $\pi_1, \pi_2 \in P$ with $\pi_1 \subset \pi_2$. In particular, for $\pi_1, \pi_2 \in P$ and $\pi := \pi_1 \cup \pi_2$ we have that $\pi \in P$ with

$$\big(\mathcal{E}_{\pi_1}(f)\big) \vee \big(\mathcal{E}_{\pi_2}(f)\big) \leq \mathcal{E}_\pi(f). \tag{3.6}$$

Recall that we denote by P_t the set of all finite partitions with end point $t \geq 0$. For $t \geq 0$, $x \in G$ and $f \in \mathrm{BUC}(G)$ we define

$$\big(\mathscr{S}(t)f\big)(x) := \sup_{\pi\in P_t} \big(\mathcal{E}_\pi(f)\big)(x).$$

Note that $\mathscr{S}(0)f = f$ for all $f \in \mathrm{BUC}(G)$. The family $(\mathscr{S}(t))_{t\geq 0}$ is called the *Nisio semigroup* to $(A_\lambda)_{\lambda\in\Lambda}$. In the following, we show that the Nisio semigroup $(\mathscr{S}(t))_{t\geq 0}$ is a sublinear Markovian convolution semigroup.

3.3 Lemma. *a) $\mathscr{S}(t)$ is a sublinear Markovian convolution for all $t \geq 0$.*

b) Let $f \in D$ and $L_f := \sup_{\lambda\in\Lambda} \|A_\lambda f\|_\infty$. Then, for all $t \geq 0$ we have that

$$\|\mathscr{S}(t)f - f\|_\infty \leq L_f \cdot t.$$

c) For all $f \in \mathrm{BUC}(G)$ we have that $\|\mathscr{S}(t)f - f\|_\infty \to 0$ as $t \searrow 0$.

Proof. a) This follows directly from the fact that $\mathcal{E}_\pi$ is a sublinear Markovian convolution for all $\pi \in P$.

b) As $f \in D$, we have that $L_f < \infty$. By Lemma 3.1 b), we get that

$$\|\mathscr{S}(t)f - f\|_\infty \leq \sup_{\pi \in P_t} \|\mathcal{E}_\pi(f) - f\|_\infty \leq L_f \cdot t$$

for all $t > 0$.

c) Let $f \in \mathrm{BUC}(G)$ and $\varepsilon > 0$. As D is dense in $\mathrm{BUC}(G)$, there exists some $f_0 \in D$ with $\|f - f_0\|_\infty \leq \frac{\varepsilon}{3}$. By part b), there exists some $t_0 > 0$ such that

$$\|\mathscr{S}(t)f_0 - f_0\|_\infty \leq \frac{\varepsilon}{3}$$

for all $0 \leq t \leq t_0$. We thus get that

$$\begin{aligned}\|\mathscr{S}(t)f - f\|_\infty &\leq \|\mathscr{S}(t)f - \mathscr{S}(t)f_0\|_\infty + \|\mathscr{S}(t)f_0 - f_0\|_\infty + \|f - f_0\|_\infty \\ &\leq \|\mathscr{S}(t)f_0 - f_0\|_\infty + 2\|f - f_0\|_\infty \leq \varepsilon\end{aligned}$$

for all $0 \leq t \leq t_0$.

□

The following lemma shows that $\mathscr{S}(t)f$ can be obtained by a pointwise monotone approximation with finite partitions letting the mesh size tend to zero.

3.4 Lemma. *Let $t \geq 0$ and $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\pi_n \subset \pi_{n+1}$ for all $n \in \mathbb{N}$ and $|\pi_n|_\infty \searrow 0$ as $n \to \infty$. Then, for all $f \in \mathrm{BUC}(G)$ it holds*

$$\mathcal{E}_{\pi_n}(f) \nearrow \mathscr{S}(t)f, \quad n \to \infty.$$

Proof. For $t = 0$ the statement is trivial. Therefore, assume that $t > 0$ and let

$$\big(\mathcal{E}_\infty(f)\big)(x) := \sup_{n\in\mathbb{N}} \big(\mathcal{E}_{\pi_n}(f)\big)(x)$$

for $f \in \mathrm{BUC}(G)$ and $x \in G$. Then, $\mathcal{E}_\infty$ is a sublinear Markovian convolution. As $\pi_n \subset \pi_{n+1}$ for all $n \in \mathbb{N}$, it follows that

$$\mathcal{E}_{\pi_n}(f) \nearrow \mathcal{E}_\infty(f), \quad n \to \infty$$

for all $f \in \mathrm{BUC}(G)$. Since $(\pi_n)_{n\in\mathbb{N}} \subset P_t$, we obtain that

$$\mathcal{E}_\infty(f) \leq \mathscr{S}(t)f$$

for all $f \in \mathrm{BUC}(G)$. Let $\pi = \{t_0, t_1, \ldots, t_m\} \in P_t$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_m = t$. Since $|\pi_n|_\infty \searrow 0$ as $n \to \infty$, w.l.o.g. we may assume that $\#\pi_n \geq m+1$ for all $n \in \mathbb{N}$. Moreover, let $0 = t_0^n < t_1^n < \ldots < t_m^n = t$ for all $n \in \mathbb{N}$ with $\pi_n' := \{t_0^n, t_1^n, \ldots, t_m^n\} \subset \pi_n$ and $t_i^n \to t_i$ as $n \to \infty$ for all $i \in \{1, \ldots, m\}$. Then, by Lemma 3.2 b), we have that

$$\|\mathcal{E}_\pi(f) - \mathcal{E}_{\pi_n'}(f)\|_\infty \to 0, \quad n \to \infty$$

for all $f \in \mathrm{BUC}(G)$ and therefore,

$$\mathcal{E}_\infty(f) \geq \mathcal{E}_{\pi_n}(f) \geq \mathcal{E}_{\pi_n'}(f) \geq \mathcal{E}_\pi(f) - \|\mathcal{E}_\pi(f) - \mathcal{E}_{\pi_n'}(f)\|_\infty$$

for all $f \in \mathrm{BUC}(G)$. Letting $n \to \infty$ we obtain that $\mathcal{E}_\infty(f) \geq \mathcal{E}_\pi(f)$ for all $f \in \mathrm{BUC}(G)$. Taking the supremum over all $\pi \in P_t$ we thus get that $\mathcal{E}_\infty = \mathscr{S}(t)$. □

3.5 Corollary. *For all $t > 0$ there exists a sequence $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with*

$$\mathcal{E}_{\pi_n}(f) \nearrow \mathscr{S}(t)f$$

as $n \to \infty$ for all $f \in \mathrm{BUC}(G)$.

Proof. For example choose $\pi_n = \left\{\frac{kt}{2^n} : k \in \{0, \ldots, 2^n\}\right\}$ or $\pi_n = \left\{\frac{kt}{n!} : k \in \{0, \ldots, n!\}\right\}$ in Lemma 3.4. □

3.6 Corollary. *For all $t \geq 0$ and $f \in \mathrm{BUC}(G)$ we have that*

$$\mathscr{S}(t)f = \sup_{n\in\mathbb{N}} \mathcal{E}^n_{\frac{1}{n}}(f) = \lim_{n\to\infty} \mathcal{E}^{2^n}_{2^{-n}}(f),$$

where the supremum and the limit are to be understood pointwise.

Next, we show that the family $(\mathscr{S}(t))_{t\geq 0}$ satisfies the semigroup property (cf. Remark 2.26). In the context of stochastic optimal control, the semigroup property is often being referred to as the *dynamic programmig principle.*

3.7 Theorem (Dynamic programming principle)**.** *The family $(\mathscr{S}(t))_{t\geq 0}$ is a Markovian convolution semigroup of sublinear operators. In particular, for all $s, t \geq 0$ we have that*

$$\mathscr{S}(s+t) = \mathscr{S}(s)\mathscr{S}(t). \tag{3.7}$$

Proof. We have already shown all properties of a sublinear Markovian convolution semigroup except for the semigroup property (3.7). Let $f \in \mathrm{BUC}(G)$. If $s = 0$ or $t = 0$ the statement is trivial. Therefore, let $s, t > 0$, $\pi_0 \in P_{s+t}$ and $\pi := \pi_0 \cup \{s\}$. Then, we have that $\pi \in P_{s+t}$ with $\pi_0 \subset \pi$. Hence, by (3.5), we get that

$$\mathcal{E}_{\pi_0}(f) \leq \mathcal{E}_\pi(f).$$

Let $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots t_m = s + t$ with $\pi = \{t_0, \ldots, t_m\}$ and $i \in \{1, \ldots, m\}$ with $t_i = s$. Then, we have that $\pi_1 := \{t_0, \ldots, t_i\} \in P_s$ and $\pi_2 := \{t_i - s, \ldots, t_n - s\} \in P_t$ with

$$\mathcal{E}_{\pi_1} = \mathcal{E}_{t_1 - t_0} \cdots \mathcal{E}_{t_i - t_{i-1}}$$

and

$$\mathcal{E}_{\pi_2} = \mathcal{E}_{t_{i+1} - t_i} \cdots \mathcal{E}_{t_m - t_{m-1}}.$$

We thus get that

$$\begin{aligned}\mathcal{E}_{\pi_0}(f) \leq \mathcal{E}_\pi(f) &= \mathcal{E}_{t_1-t_0} \cdots \mathcal{E}_{t_m - t_{m-1}}(f) = \big(\mathcal{E}_{t_1-t_0} \cdots \mathcal{E}_{t_i - t_{i-1}}\big)\big(\mathcal{E}_{t_{i+1}-t_i} \cdots \mathcal{E}_{t_m-t_{m-1}}(f)\big) \\ &= \mathcal{E}_{\pi_1}\mathcal{E}_{\pi_2}(f) \leq \mathcal{E}_{\pi_1}\big(\mathscr{S}(t)f\big) \leq \mathscr{S}(s)\mathscr{S}(t)f.\end{aligned}$$

Taking the supremum over all $\pi_0 \in P_{s+t}$, we get that $\mathscr{S}(s+t)f \leq \mathscr{S}(s)\mathscr{S}(t)f$.

Now, let $(\pi_n)_{n\in\mathbb{N}} \subset P_t$ with $\mathcal{E}_{\pi_n} f \nearrow \mathscr{S}(t)f$ as $n \to \infty$ (see Corollary 3.5) and fix $\pi_0 \in P_s$. Then, for all $n \in \mathbb{N}$ we have that

$$\pi_n' := \pi_0 \cup \{s + \tau : \tau \in \pi_n\} \in P_{s+t}$$

with $\mathcal{E}_{\pi_n'} = \mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n}$. As $\mathcal{E}_{\pi_0}$ is continuous from below, we get that

$$\mathcal{E}_{\pi_0}\big(\mathscr{S}(t)f\big) = \lim_{n\to\infty} \mathcal{E}_{\pi_0}\mathcal{E}_{\pi_n}(f) = \lim_{n\to\infty} \mathcal{E}_{\pi_n'}(f) \leq \mathscr{S}(s+t)f.$$

Taking the supremum over all $\pi_0 \in P_s$, we get that $\mathscr{S}(s)\mathscr{S}(t)f \leq \mathscr{S}(s+t)f$. □

3.8 Remark. The family $(\mathscr{S}(t))_{t\geq 0}$ in the above theorem is the smallest semigroup which dominates all semigroups $(S_\lambda(t))_{t\geq 0}$ with $\lambda\in\Lambda$.

Using Lemma 2.19, we saw in the proof of Lemma 3.1 that $\mathcal{E}_\pi$ has a dual representation in terms of distributions of stochastic integrals with a finite space-time partition for all $\pi\in P$. As each one of these distributions can be extended to the whole path space by Kolmogorov's extension theorem, there exists a set $\mathcal{Q}\subset \mathrm{ca}^1_+(S^{[0,\infty)},\mathscr{B}^{[0,\infty)})$ such that

$$\big(\mathscr{S}(t)f\big)(0)=\sup_{\mu\in\mathcal{Q}}\mu\big(f\circ \mathrm{pr}_{\{t\}}\big)$$

for all $t\geq 0$ and $f\in \mathrm{BUC}(G)$. The canonical process $\big(\mathrm{pr}_{\{t\}}\big)_{t\geq 0}$ is then an $\mathcal{E}$-Lévy process, where $\mathcal{E}\colon \mathrm{ca}^1_+\big(S^{[0,\infty)},\mathscr{B}^{[0,\infty)}\big)\to\mathbb{R}$ is defined by

$$\mathcal{E}(g)=\sup_{\mu\in\mathcal{Q}}\mu g$$

for all $g\in\mathscr{L}^\infty\big(S^{[0,\infty)},\mathscr{B}^{[0,\infty)}\big)$. We thus obtain the following main theorem.

3.9 Theorem. *There exists a sublinear expectation space $(\Omega,\mathscr{F},\mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\geq 0}$ such that*

$$\big(\mathscr{S}(t)f\big)(x)=\mathcal{E}(f(x+X_t))$$

for all $x\in G$, $t\geq 0$ and $f\in \mathrm{BUC}(G)$.

We now turn our focus on the connection between the Nisio semigroup $(\mathscr{S}(t))_{t\geq 0}$ and fully nonlinear PDEs. Let $f\in D$. As $\{A_\lambda f\colon \lambda\in\Lambda\}\subset \mathrm{BUC}(G)$ is bounded and uniformly equicontinuous, we have that

$$\mathcal{A}f:=\sup_{\lambda\in\Lambda}A_\lambda f\in \mathrm{BUC}(G),$$

where the supremum is to be understood pointwise. In the linear case, i.e. if $|\Lambda|=1$, the vector-valued version of the Fundamental Theorem of Calculus implies that

$$\mathscr{S}(t)f-f=\int_0^t\mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s, \tag{3.8}$$

where the integral that appears on the right hand side is a Bochner integral with values in the Banach space $\mathrm{BUC}(G)$. We refer to [1, Chapter X] or [30, Chapter II] for a discussion of Bochner's integral. We now want to show that the equality in (3.8) can be replaced by the inequality

$$\mathscr{S}(t)f-f\leq\int_0^t\mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s \tag{3.9}$$

in the sublinear case. For this, we first state some technical lemma on Bochner integration, which we will need in order to prove (3.9) in Lemma 3.11 below.

3.10 Lemma. *Let $\mathscr{S}\colon \mathrm{BUC}(G)\to \mathrm{BUC}(G)$ be convex and continuous and $(\Omega,\mathscr{F},\nu)$ be a finite measure space with $\nu\neq 0$. Further, let $g\colon\Omega\to \mathrm{BUC}(G)$ be bounded and $\mathscr{F}$-$\mathscr{B}(\mathrm{BUC}(G))$-measurable with separable range $g(\Omega)$, i.e. g is Bochner integrable. Then, the mapping*

$$\mathscr{S}g\colon\Omega\to \mathrm{BUC}(G),\quad \omega\mapsto\mathscr{S}\big(g(\omega)\big)$$

is bounded and $\mathscr{F}$*-*$\mathscr{B}(\mathrm{BUC}(G))$*-measurable with separable range* $(\mathscr{S}g)(\Omega)$*. Hence,* $\mathscr{S}g$ *is Bochner integrable and we have that*

$$\mathscr{S}\left(\frac{1}{\nu(\Omega)}\int_\Omega g\,\mathrm{d}\nu\right) \le \frac{1}{\nu(\Omega)}\int_\Omega \mathscr{S}g\,\mathrm{d}\nu.$$

Proof. As $\mathscr{S}\colon \mathrm{BUC}(G) \to \mathrm{BUC}(G)$ is continuous, we obtain that $\mathscr{S}g$ is $\mathscr{F}$-$\mathscr{B}(\mathrm{BUC}(G))$-measurable with separable range $(\mathscr{S}g)(\Omega)$. If g is a simple function, $\mathscr{S}g$ is a simple function as well, and the assertion follows from the convexity of the operator $\mathscr{S}$. As $g\colon \Omega \to \mathrm{BUC}(G)$ is $\mathscr{F}$-$\mathscr{B}(\mathrm{BUC}(G))$-measurable with separable range $g(\Omega)$, there exists a sequence of simple functions $(g_n)_{n\in\mathbb{N}}$ with $\|g(\omega) - g_n(\omega)\|_\infty \to 0$ as $n \to \infty$ for all $\omega \in \Omega$. Again, since the map $\mathscr{S}\colon \mathrm{BUC}(G) \to \mathrm{BUC}(G)$ is continuous, we obtain that

$$\|\mathscr{S}g(\omega) - \mathscr{S}g_n(\omega)\|_\infty \to 0$$

as $n \to \infty$. Hence, by definition of Bochner's integral and continuity of $\mathscr{S}\colon \mathrm{BUC}(G) \to \mathrm{BUC}(G)$, we get that

$$\mathscr{S}\left(\frac{1}{\nu(\Omega)}\int_\Omega g\,\mathrm{d}\nu\right) = \lim_{n\to\infty}\mathscr{S}\left(\frac{1}{\nu(\Omega)}\int_\Omega g_n\,\mathrm{d}\nu\right) \le \lim_{n\to\infty}\frac{1}{\nu(\Omega)}\int_\Omega \mathscr{S}g_n\,\mathrm{d}\nu = \frac{1}{\nu(\Omega)}\int_\Omega \mathscr{S}g\,\mathrm{d}\nu.$$

□

3.11 Lemma. *For all* $f \in D$ *and* $t > 0$ *it holds that*

$$\mathscr{S}(t)f - f \le \int_0^t \mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s.$$

Proof. First, note that the mapping $[0,\infty) \to \mathrm{BUC}(G)$, $s \mapsto \mathscr{S}(s)\mathcal{A}f$ is continuous since $\mathcal{A}f \in \mathrm{BUC}(G)$ and therefore, the Bochner integral that appears is well-defined. Moreover, for all $t, h > 0$ we have that

$$\mathcal{E}_h(f) - f = \sup_{\lambda\in\Lambda} S_\lambda(h)f - f = \sup_{\lambda\in\Lambda}\int_0^h S_\lambda(\tau)A_\lambda f\,\mathrm{d}\tau \le \int_0^h \mathscr{S}(\tau)\mathcal{A}f\,\mathrm{d}\tau = \int_t^{t+h}\mathscr{S}(s-t)\mathcal{A}f\,\mathrm{d}s. \tag{3.10}$$

We first show that

$$\mathcal{E}_\pi(f) - f \le \int_0^{\max\pi}\mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s. \tag{3.11}$$

for all $\pi \in P$ by an induction on $m = \#\pi$. If $m = 1$, i.e. if $\pi = \{0\}$, the statement is trivial. Hence, assume that

$$\mathcal{E}_{\pi'}(f) - f \le \int_0^{\max\pi'}\mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s$$

for all $\pi' \in P$ with $\#\pi' = m$ for some $m \in \mathbb{N}$. Let $\pi = \{t_0, t_1, \ldots, t_m\}$ with $0 = t_0 < t_1 < \ldots < t_m$ and $\pi' := \pi \setminus \{t_m\}$. Then, by (3.10) and Lemma 3.10, we obtain that

$$\begin{aligned}\mathcal{E}_\pi(f) - \mathcal{E}_{\pi'}(f) &\le \mathcal{E}_{\pi'}\big(\mathcal{E}_{t_m - t_{m-1}}(f) - f\big) \le \mathcal{E}_{\pi'}\bigg(\int_{t_{m-1}}^{t_m}\mathscr{S}(s-t_{m-1})\mathcal{A}f\,\mathrm{d}s\bigg)\\ &\le \mathscr{S}(t_{m-1})\bigg(\int_{t_{m-1}}^{t_m}\mathscr{S}(s-t_{m-1})\mathcal{A}f\,\mathrm{d}s\bigg) \le \int_{t_{m-1}}^{t_m}\mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s.\end{aligned}$$

Using the induction hypothesis, we thus get that

$$\begin{aligned}\mathcal{E}_\pi(f) - f &= \big(\mathcal{E}_\pi(f) - \mathcal{E}_{\pi'}(f)\big) + \big(\mathcal{E}_{\pi'}(f) - f\big) \le \int_{t_{m-1}}^{t_m} \mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s + \int_0^{t_{m-1}} \mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s\\ &= \int_0^{\max\pi} \mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s.\end{aligned}$$

By (3.11), it follows that

$$\mathcal{E}_\pi(f) - f \le \int_0^t \mathscr{S}(s)\mathcal{A}f\,\mathrm{d}s$$

for all $\pi \in P_t$. Taking the supremum over all $\pi \in P_t$ we obtain the assertion. □

3.12 Lemma. *Let $M \subset \mathrm{BUC}(G)$ be bounded and uniformly equicontinuous. Then,*

$$\sup_{\lambda\in\Lambda}\sup_{g\in M} \|S_\lambda(t)g - g\|_\infty \to 0, \quad t \searrow 0.$$

Proof. Let $\varepsilon > 0$ and $C := \sup_{g\in M}\|g\|_\infty$. Then, there exists some $\delta > 0$ such that

$$\sup_{g\in M} |g(x) - g(y)| \le \varepsilon$$

for all $x, y \in G$ with $d(x,y) \le \delta$. Let $\varphi(y) := \frac{1}{\delta}\big(d(y,0)\wedge\delta\big)$ for all $y \in G$. Then, $\varphi \in \mathrm{BUC}(G)$, $0 \le \varphi \le 1$, $\varphi(0) = 0$ and $\varphi(y) = 1$ for all $y \in G \setminus B(0,\delta)$. Using Lemma 1.17, we thus get that

$$\begin{aligned}\big|\big(S_\lambda(t)g\big)(x) - g(x)\big| &= \big|\big(S_\lambda(t)\big(g(x+\cdot) - g(x)\big)\big)(0)\big| \le \big(S_\lambda(t)|g(x+\cdot) - g(x)|\big)(0)\\ &\le 2C\big(\mathscr{S}(t)\varphi\big)(0) + \varepsilon\end{aligned}$$

for all $\lambda \in \Lambda$, $g \in M$ and $x \in G$. Taking the supremum over all $\lambda \in \Lambda$, $g \in M$ and $x \in G$, we get that

$$\sup_{\lambda\in\Lambda}\sup_{g\in M} \|S_\lambda(t)g - g\|_\infty \le 2C\big(\mathscr{S}(t)\varphi\big)(0) + \varepsilon.$$

Since $\big(\mathscr{S}(t)\varphi\big)(0) \to 0$ as $t \searrow 0$ by Lemma 3.3 c), the assertion follows. □

3.13 Lemma. *For all $f \in D$ it holds that*

$$\lim_{h\searrow 0}\left\|\frac{\mathscr{S}(h)f - f}{h} - \mathcal{A}f\right\|_\infty = 0.$$

Proof. Let $\varepsilon > 0$. By Lemma 3.12 and Lemma 3.3 d), there exists some $h_0 > 0$ such that

$$S_\lambda(h)A_\lambda f - A_\lambda f \ge -\varepsilon$$

for all $\lambda \in \Lambda$ and

$$\mathscr{S}(h)\mathcal{A}f - \mathcal{A}f \le \varepsilon$$

for all $0 < h \le h_0$. Let $0 < h \le h_0$. Then, we get that

$$\mathscr{S}(h)f - f \ge S_\lambda(h)f - f = \int_0^h S_\lambda(s)A_\lambda f\,\mathrm{d}s \ge \big(A_\lambda f - \varepsilon\big)h.$$

Dividing by h and taking the supremum over all $\lambda \in \Lambda$, it follows that

$$\frac{\mathscr{S}(h)f - f}{h} \geq \mathcal{A}f - \varepsilon. \tag{3.12}$$

By Lemma 3.11, we have that

$$\mathscr{S}(h)f - f - h\mathcal{A}f \leq \int_0^h \mathscr{S}(s)\mathcal{A}f \,\mathrm{d}s - h\mathcal{A}f = \int_0^h \mathscr{S}(s)\mathcal{A}f - \mathcal{A}f \,\mathrm{d}s \leq h\varepsilon$$

Again, dividing by $h > 0$ yields that

$$\frac{\mathscr{S}(h)f - f}{h} - \mathcal{A}f \leq \varepsilon.$$

Together with (3.12) this implies that

$$\left\| \frac{\mathscr{S}(h)f - f}{h} - \mathcal{A}f \right\|_\infty \leq \varepsilon.$$

□

3.14 Definition. Let $u\colon [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$. Then, we say that u a *D-viscosity subsolution* of the PDE

$$u_t(t,x) = \big(\mathcal{A}u(t)\big)(x), \quad (t,x) \in (0,\infty) \times G$$

if $u\colon [0,\infty) \to \mathrm{BUC}(G)$ is continuous and for all $t > 0$, $x \in G$ and all differentiable functions $\psi = \psi_{t,x}\colon (0,\infty) \to \mathrm{BUC}(G)$ with $\big(\psi(t)\big)(x) = \big(u(t)\big)(x)$, $\psi(s) \geq u(s)$ and $\psi(s) \in D$ for all $s > 0$ we have that

$$\partial_t \psi(t,x) \leq \big(\mathcal{A}\psi(t)\big)(x).$$

We say that $u(t,x) := \big(u(t)\big)(x)$ is a *D-viscosity supersolution* of the PDE

$$u_t(t,x) = \big(\mathcal{A}u(t)\big)(x), \quad (t,x) \in (0,\infty) \times G$$

if $u\colon [0,\infty) \to \mathrm{BUC}(G)$ is continuous and for all $t > 0$, $x \in G$ and all differentiable functions $\psi = \psi_{t,x}\colon (0,\infty) \to \mathrm{BUC}(G)$ with $\big(\psi(t)\big)(x) = \big(u(t)\big)(x)$, $\psi(s) \leq u(s)$ and $\psi(s) \in D$ for all $s > 0$ we have that

$$\partial_t \psi(t,x) \geq \big(\mathcal{A}\psi(t)\big)(x).$$

We say that u is a *D-viscosity solution* if u is both, a D-viscosity subsolution and a D-viscosity supersolution.

3.15 Remark. Consider the case $G = \mathbb{R}^d$. Let $\psi\colon (0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ be differentiable w.r.t. the variable t and assume that $\partial_t\psi\colon (0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ is uniformly in x continuous w.r.t. t, i.e. for all $t > 0$ we have that

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in \mathbb{R}^d\, \forall s > 0, |t-s| < \delta\colon |\partial_t\psi(t,x) - \partial_t\psi(s,x)| < \varepsilon.$$

Then,

$$\sup_{x \in \mathbb{R}^d} \left| \frac{\psi(t+h,x) - \psi(t,x)}{h} - \partial_t\psi(t,x) \right| \to 0$$

as $h \to 0$ for all $t > 0$. If we assume that, in addition, $\psi(s,\cdot) \in \mathrm{BUC}^2(\mathbb{R}^d)$ for all $s > 0$, then $\psi\colon (0,\infty) \to \mathrm{BUC}(\mathbb{R}^d)$ is differentiable with $\psi(s) \in \mathrm{BUC}^2(\mathbb{R}^d)$ for all $s > 0$ using the identification $\big(\psi(s)\big)(x) := \psi(s,x)$. In particular, every function $\psi \in C_b^{2,3}([0,\infty)\times\mathbb{R}^d)$ satisfies these conditions. The space $C_b^{2,3}([0,\infty)\times\mathbb{R}^d)$ is often considered as a class of test functions in the framework of viscosity solutions, see e.g. [50].

We are now able to state the second main theorem of this section, which asserts that the Nisio semigroup $(\mathscr{S}(t))_{t\geq 0}$ defines a D-viscosity solution of the fully nonlinear PDE

$$u_t(t,x) = \sup_{\lambda\in\Lambda} \big(A_\lambda u(t)\big)(x), \quad (t,x) \in (0,\infty)\times G.$$

3.16 Theorem. *Let $u_0 \in \mathrm{BUC}(G)$ and define $u(t) := \mathscr{S}(t)u_0$ for all $t \geq 0$. Using the notation $u(t,x) := \big(u(t)\big)(x)$ for $t \geq 0$ and $x \in G$, it holds that u is a D-viscosity solution of the fully nonlinear PDE*

$$u_t(t,x) = \sup_{\lambda\in\Lambda} \big(A_\lambda u(t)\big)(x), \quad (t,x) \in (0,\infty)\times G$$

with $u(0,x) = u_0(x)$ for all $x \in G$.

Proof. Let $t > 0$, $x \in G$ and $\psi\colon (0,\infty) \to \mathrm{BUC}(G)$ be differentiable with $\big(\psi(t)\big)(x) = \big(u(t)\big)(x)$, $\psi(s) \geq u(s)$ and $\psi(s) \in D$ for all $s > 0$. Then, by Proposition 3.7, for all $0 < h < t$ we get that

$$\begin{aligned} 0 &= \frac{\mathscr{S}(h)\mathscr{S}(t-h)u_0 - \mathscr{S}(t)u_0}{h} = \frac{\mathscr{S}(h)u(t-h) - u(t)}{h} \\ &\leq \frac{\mathscr{S}(h)\psi(t-h) - u(t)}{h} \leq \frac{\mathscr{S}(h)\big(\psi(t-h)-\psi(t)\big) + \mathscr{S}(h)\psi(t) - u(t)}{h} \\ &= \mathscr{S}(h)\bigg(\frac{\psi(t-h)-\psi(t)}{h}\bigg) + \frac{\mathscr{S}(h)\psi(t)-\psi(t)}{h} + \frac{\psi(t)-u(t)}{h}. \end{aligned}$$

Let $\varepsilon > 0$. Then, by Lemma 3.13, there exists some $0 < h_0 < t$ such that

$$\frac{\mathscr{S}(h)\psi(t)-\psi(t)}{h} \leq \mathcal{A}\psi(t) + \frac{\varepsilon}{3},$$

$$\frac{\psi(t-h)-\psi(t)}{h} \leq -\psi_t(t) + \frac{\varepsilon}{3}$$

and

$$\mathscr{S}(h)\big(-\psi_t(t)\big) \leq -\psi_t(t) + \frac{\varepsilon}{3}$$

for all $0 < h < h_0$. We thus get that

$$0 \leq \mathscr{S}(h)\big(-\psi_t(t)\big) + \mathcal{A}\psi(t) + \frac{2\varepsilon}{3} + \frac{\psi(t)-u(t)}{h} \leq -\psi_t(t) + \mathcal{A}\psi(t) + \varepsilon + \frac{\psi(t)-u(t)}{h}$$

for all $0 < h < h_0$. As $\big(\psi(t)\big)(x) = u(t,x) = \big(u(t)\big)(x)$, we obtain that

$$0 \leq -\big(\psi_t(t)\big)(x) + \big(\mathcal{A}\psi(t)\big)(x) + \varepsilon.$$

Letting $\varepsilon \searrow 0$, we obtain that u is a D-viscosity subsolution.

Now, let $\psi\colon (0,\infty) \to \mathrm{BUC}(G)$ be differentiable with $\big(\psi(t)\big)(x) = \big(u(t)\big)(x)$, $\psi(s) \in D$ and $\psi(s) \leq u(s)$ for all $s > 0$. Then, by Proposition 3.7, for all $0 < h < t$ we get that

$$\begin{aligned}
0 &= \frac{\mathscr{S}(t)u_0 - \mathscr{S}(h)\mathscr{S}(t-h)u_0}{h} = \frac{u(t) - \mathscr{S}(h)u(t-h)}{h} \leq \frac{u(t) - \mathscr{S}(h)\psi(t-h)}{h}\\
&= \frac{u(t)-\psi(t)}{h} + \frac{\psi(t) - \mathscr{S}(h)\psi(t)}{h} + \frac{\mathscr{S}(h)\psi(t) - \mathscr{S}(h)\psi(t-h)}{h}\\
&\leq \frac{u(t)-\psi(t)}{h} + \frac{\psi(t) - \mathscr{S}(h)\psi(t)}{h} + \mathscr{S}(h)\bigg(\frac{\psi(t)-\psi(t-h)}{h}\bigg).
\end{aligned}$$

Let $\varepsilon > 0$. Then, by Lemma 3.13, there exists some $0 < h_0 < t$ such that

$$\frac{\psi(t) - \mathscr{S}(h)\psi(t)}{h} \leq -\mathcal{A}\psi(t) + \frac{\varepsilon}{3},$$

$$\frac{\psi(t) - \psi(t-h)}{h} \leq \psi_t(t) + \frac{\varepsilon}{3}$$

and

$$\mathscr{S}(h)\big(\psi_t(t)\big) \leq \psi_t(t) + \frac{\varepsilon}{3}$$

for all $0 < h < h_0$. We thus get that

$$0 \leq \frac{u(t)-\psi(t)}{h} - \mathcal{A}\psi(t) + \mathscr{S}(h)\big(\psi_t(t)\big) + \frac{2\varepsilon}{3} \leq \frac{u(t)-\psi(t)}{h} - \mathcal{A}\psi(t) + \psi_t(t) + \varepsilon$$

for all $0 < h < h_0$. As $\big(\psi(t)\big)(x) = \big(u(t)\big)(x)$, we obtain that

$$0 \leq -\big(\mathcal{A}\psi(t)\big)(x) + \big(\psi_t(t)\big)(x) + \varepsilon.$$

Letting $\varepsilon \searrow 0$, we obtain that u is a D-viscosity supersolution. □

3.17 Corollary. *Let $u_0 \in D$ with $u(t) := \mathscr{S}(t)u_0 \in D$ for all $t \geq 0$. Then u is a classical solution of the fully nonlinear PDE*

$$\begin{aligned}
u_t(t) &= \mathcal{A}u(t), \quad t \geq 0,\\
u(0) &= u_0.
\end{aligned}$$

Moreover, we have that $u \in C^1([0,\infty);\mathrm{BUC}(G))$ and $\mathcal{A}u \in C([0,\infty);\mathrm{BUC}(G))$.

We close this section with a couple of examples that illustrate where and how the previous results can be applied.

3.18 Example (Compound Poisson processes)**.** Let $\lambda \geq 0$, $\mu \in \mathrm{ca}_+^1(G,\mathscr{B})$ and

$$\big(A_{\lambda,\mu}f\big)(x) := \lambda \int_G f(x+y) - f(x)\,\mathrm{d}\mu(y)$$

for $f \in \mathrm{BUC}(G)$ and $x \in G$. Then, $A_{\lambda,\mu}\colon \mathrm{BUC}(G) \to \mathrm{BUC}(G)$ is a bounded linear operator which satisfies the *strong maximum principle*, i.e. for all functions $f \in \mathrm{BUC}(G)$ and $x_0 \in G$

with $f(x_0) = \max_{x\in G} f(x)$ we have that $\big(A_{\mu,\lambda} f\big)(x_0) \leq 0$. Let $t \geq 0$ be fixed and $S_{\lambda,\mu}$ be the exponential of the bounded linear operator $A_{\lambda,\mu}$, i.e.

$$S_{\lambda,\mu}(t) := e^{tA_{\lambda,\mu}} \in L(\mathrm{BUC}(G)).$$

Then, by the Theorem of Kolmogorov, there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, an i.i.d. sequence $(Y_i)_{i\in\mathbb{N}}$ with $Y_i\colon \Omega \to G$ and $Y_i \sim \mu$ and a Poisson process $(N_t)_{t\geq 0}$ which is independent of $(Y_i)_{i\in\mathbb{N}}$. Recall that a Poisson process is a Lévy process with $\mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$ for all $t \geq 0$ and $n \in \mathbb{N}$. We then define

$$J_t := \sum_{i=1}^{N_t} Y_i.$$

The resulting process $(J_t)_{t\geq 0}$ is then called a *compound Poisson process*. Moreover, for all $t > 0$, $f \in \mathrm{BUC}(G)$ and $x \in G$ we have that

$$\big(S_{\lambda,\mu}(t)f\big)(x) = \mathbb{E}(f(x + J_t)) = \int_G f(x+y)\,\mathrm{d}\big(\mathbb{P}\circ J_t^{-1}\big)(y).$$

In fact, we have that

$$\big((A_{\lambda,\mu} + \lambda)f\big)(x) = \lambda \int_G f(x+y)\,\mathrm{d}\mu(y).$$

This implies that

$$\begin{aligned}
\mathbb{E}(f(x+J_t)) &= \sum_{n=0}^{\infty} \mathbb{E}(f(x + Y_1 + \ldots + Y_n))e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} t^n\lambda^n \int_G \cdots \int_G f(x + y_1 + \ldots + y_n)\,\mathrm{d}\mu(y_1)\cdots\mathrm{d}\mu(y_n) \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} \frac{\big(t^n(A_{\lambda,\mu}+\lambda)^n f\big)(x)}{n!} = e^{-\lambda t}\big(e^{t(A_{\lambda,\mu}+\lambda)}f\big)(x) \\
&= \big(e^{tA_{\lambda,\mu}}f\big)(x) = \big(S_{\lambda,\mu}(t)f\big)(x)
\end{aligned}$$

for all $f \in \mathrm{BUC}(G)$ and all $x \in G$. Note that

$$S_{0,\mu}(t)f = f$$

for all $f \in \mathrm{BUC}(G)$. Now, assume that $\Lambda \subset [0,\infty)$ is bounded and $\mathcal{Q} \subset \mathrm{ca}_+^1(G, \mathscr{B})$. Then, we have that

$$\big\{f \in \mathrm{BUC}(G)\colon \{A_{\lambda,\mu}f\colon (\lambda,\mu) \in \Lambda\times\mathcal{Q}\} \text{ is bounded and uniformly equicontinuous}\big\} = \mathrm{BUC}(G)$$

and therefore, the assumptions (A1) and (A2) are satisfied with $D = \mathrm{BUC}(G)$. Hence, there exists a nonlinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\geq 0}$ such that for all $u_0 \in \mathrm{BUC}(G)$ the function

$$u(t,x) := \big(u(t)\big)(x) := \mathcal{E}(u_0(x + X_t)) \quad (t \geq 0, x \in G)$$

is the unique classical solution of the fully nonlinear integrodifferential equation

$$\begin{aligned} u_t(t,x) &= \sup_{(\lambda,\mu)\in\Lambda\times\mathcal{Q}} \big(A_{\lambda,\mu}u(t)\big)(x), \quad (t,x)\in[0,\infty)\times G,\\ u(0,x) &= u_0(x), \quad x\in G. \end{aligned}$$

Moreover, we have that $u\in C^1([0,\infty);\mathrm{BUC}(G))$ and $\mathcal{A}u\in C([0,\infty);\mathrm{BUC}(G))$, where we have set $(\mathcal{A}v)(x) := \sup_{(\lambda,\mu)\in\Lambda\times\mathcal{Q}}(A_{\lambda,\mu}v)(x)$ for $x\in G$ and $v\in\mathrm{BUC}(G)$.

3.19 Example (Lévy processes in $\mathbb{R}^d$)**.** Let $d\in\mathbb{N}$ and Λ be a set of Lévy triplets (b,Σ,μ), i.e. $b\in\mathbb{R}^d$, $\Sigma\in\mathbb{R}^{d\times d}$ is a symmetric positive semidefinite matrix and μ is a Lévy measure. For any Lévy triplet (b,Σ,μ) consider

$$\big(A_{b,\Sigma,\mu}f\big)(x) := b\cdot\nabla f(x) + \frac{1}{2}\mathrm{tr}\big(\Sigma\nabla^2 f(x)\big) + \int_{\mathbb{R}^d} f(x+y)-f(x)-\nabla f(x)\cdot h(y)\,\mathrm{d}\mu(y)$$

for $x\in\mathbb{R}^d$ and $f\in D(A_{b,\Sigma,\mu}) := \big\{g\in\mathrm{BUC}(\mathbb{R}^d)\,\big|\,A_{b,\Sigma,\mu}g\in\mathrm{BUC}(\mathbb{R}^d)\big\}$. Here, $h\colon\mathbb{R}^d\to\mathbb{R}^d$ is given by

$$h(y) := \begin{cases} y, & |y|\le 1,\\ 0, & |y|>1 \end{cases}$$

for all $y\in\mathbb{R}^d\setminus\{0\}$. Then, by the Lévy-Khintchine formula (see e.g. [72, Section 2.8] or [2, Section 1.2.4]), for each Lévy triplet (b,Σ,μ) the operator $A_{b,\Sigma,\mu}$ generates a Markovian convolution semigroup $(S_{b,\Sigma,\mu}(t))_{t\ge0}$ of linear operators on $\mathrm{BUC}(\mathbb{R}^d)$ and for any Markovian convolution semigroup $(S(t))_{t\ge0}$ of linear operators on $\mathrm{BUC}(\mathbb{R}^d)$ there exists a Lévy triplet (b,Σ,μ) with $(S(t))_{t\ge0}=(S_{b,\Sigma,\mu}(t))_{t\ge0}$. In particular, (A1) is satisfied. Assume that

$$C := \sup_{(b,\Sigma,\mu)\in\Lambda}\left(1+|b|+|\Sigma|+\int_{\mathbb{R}^d}1\wedge|y|^2\,\mathrm{d}\mu(y)\right)<\infty. \tag{3.13}$$

We will verify that under this condition the assumption (A2) is satisfied. Let

$$D := \left\{f\in\bigcap_{(b,\Sigma,\mu)\in\Lambda}D(A_{b,\Sigma,\mu})\colon \{A_{b,\Sigma,\mu}f\,|\,(b,\Sigma,\mu)\in\Lambda\}\text{ is bounded and unif. equicontinuous}\right\}.$$

As $\mathrm{BUC}^2(\mathbb{R}^d)$ is dense in $\mathrm{BUC}(\mathbb{R}^d)$, it suffices to show that $\mathrm{BUC}^2(\mathbb{R}^d)\subset D$. Let $f\in\mathrm{BUC}^2(\mathbb{R}^d)$. Then, we have that $f\in D(A_{b,\Sigma,\mu})$ for any Lévy triplet $(b,\Sigma,\mu)\in\Lambda$. Moreover, by Taylor's theorem, we have that

$$\begin{aligned} \big|\big(A_{b,\Sigma,\mu}f\big)(x)\big| &\le |b|\cdot|\nabla f(x)| + |\Sigma|\cdot|\nabla^2 f(x)| + \int_{\mathbb{R}^d}|f(x+y)-f(x)-\nabla f(x)\cdot h(y)|\,\mathrm{d}y\\ &\le |b|\|\nabla f\|_\infty + |\Sigma|\|\nabla^2 f\|_\infty + \max\big\{2\|f\|_\infty,\|\nabla^2 f\|_\infty\big\}\int_{\mathbb{R}^d}1\wedge|y|^2\,\mathrm{d}\mu(y)\\ &\le 2C\max\big\{\|f\|_\infty,\|\nabla f\|_\infty,\|\nabla^2 f\|_\infty\big\} \end{aligned}$$

for all $x\in\mathbb{R}^d$ and all $(b,\sigma,\mu)\in\Lambda$, i.e.

$$\sup_{(b,\Sigma,\mu)\in\Lambda}\|A_{b,\Sigma,\mu}f\|_\infty \le 2C\max\big\{\|f\|_\infty,\|\nabla f\|_\infty,\|\nabla^2 f\|_\infty\big\}. \tag{3.14}$$

Let $\varepsilon > 0$. As $f \in \mathrm{BUC}^2(\mathbb{R}^d)$, there exists some $\delta > 0$ such that for all $x, z \in \mathbb{R}^d$ with $|x-z| \le \delta$ we have that

$$2C \cdot \left(\sup_{u \in \mathbb{R}^d} |f(x+u) - f(x)| \right) \le \varepsilon,$$

$$2C \cdot \left(\sup_{u \in \mathbb{R}^d} |\nabla f(x+u) - \nabla f(z+u)| \right) \le \varepsilon$$

and

$$2C \cdot \left(\sup_{u \in \mathbb{R}^d} |\nabla^2 f(x+u) - \nabla^2 f(z+u)| \right) \le \varepsilon.$$

Let $x, z \in \mathbb{R}^d$ be fixed with $|x-z| \le \delta$ and let $g \colon \mathbb{R}^d \to \mathbb{R}$ be defined by

$$g(u) := f(x+u) - f(z+u)$$

for all $u \in \mathbb{R}^d$. Then, we have that $g \in \mathrm{BUC}^2(\mathbb{R}^d)$ with

$$2C \max\left\{ \|g\|_\infty, \|\nabla g\|_\infty, \|\nabla^2 g\|_\infty \right\} \le \varepsilon.$$

By (3.14) we thus get that

$$\left| \big(A_{b,\Sigma,\mu} f\big)(x) - \big(A_{b,\Sigma,\mu} f\big)(z) \right| = \left| \big(A_{b,\Sigma,\mu} g\big)(0) \right| \le 2C \max\left\{ \|g\|_\infty, \|\nabla g\|_\infty, \|\nabla^2 g\|_\infty \right\} \le \varepsilon$$

for all $(b, \Sigma, \mu) \in \Lambda$. This shows that (A2) is satisfied. Therefore, there exists a nonlinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t \ge 0}$ such that for all $u_0 \in \mathrm{BUC}(\mathbb{R}^d)$ the function

$$u(t,x) := \big(u(t)\big)(x) := \mathcal{E}(u_0(x + X_t)) \quad (t \ge 0, x \in \mathbb{R}^d)$$

is a $\mathrm{BUC}^2(\mathbb{R}^d)$-viscosity solution of the fully nonlinear PDE

$$\begin{aligned} u_t(t,x) &= \sup_{(b,\Sigma,\mu) \in \Lambda} \big(A_{b,\Sigma,\mu} u(t)\big)(x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d, \\ u(0,x) &= u_0(x), \quad x \in \mathbb{R}^d. \end{aligned}$$

Note that condition (3.13) does not exclude any Lévy measure. In particular, if Λ is finite, (3.13) is always fulfilled. Hence, we may make a completely degenerate choice of operators, e.g. $A_1 := \mathrm{div} := \sum_{i=1}^d \partial_{x_i}$, $A_2 := \Delta := \sum_{i=1}^d \partial_{x_i}^2$ and A_3 given by

$$A_3 f := (2\pi)^{-d/2} \int_{\mathbb{R}^d} \big(f(x+y) - f(x)\big) e^{-|y|^2/2} \, \mathrm{d}y$$

for $f \in \mathrm{BUC}(\mathbb{R}^d)$. If we additionally assume that for all $\varepsilon > 0$ there exists some $M > 0$ such that

$$\sup_{(b,\Sigma,\mu) \in \Lambda} \mu\big(\mathbb{R}^d \setminus B(0,M)\big) < \varepsilon \tag{3.15}$$

and

$$\lim_{\delta \to 0} \sup_{(b,\Sigma,\mu) \in \Lambda} \int_{B(0,\delta)} |z|^2 \, \mathrm{d}\mu(z) = 0 \tag{3.16}$$

we obtain the uniqueness of the viscosity solution of the PDE

$$\begin{aligned} u_t(t,x) &= \sup_{(b,\Sigma,\mu)\in\Lambda} \big(A_{b,\Sigma,\mu}u(t)\big)(x), \quad (t,x)\in(0,\infty)\times\mathbb{R}^d,\\ u(0,x) &= u_0(x), \quad x\in\mathbb{R}^d \end{aligned}$$

by Remark 3.15 and [50, Corollary 53]. A similar result is obtained in Neufeld-Nutz [58]. Note that the conditions (1.3) and (1.4) in [58] imply (3.13), (3.15) and (3.16). However, while conditions (1.3) and (1.4) in [58] exclude all Lévy triplets with non-integrable jumps, (3.13), (3.15) and (3.16) do not exclude any Lévy triplet at all. In particular, for finite Λ the conditions (3.13), (3.15) and (3.16) are still always satisfied. Moreover, we would like to mention that contidition (3.13) is the minimal condition in order that the differential operator in the above PDE is well-defined.

3.20 Example. Let $A\colon D(A)\subset \mathrm{BUC}(G)\to\mathrm{BUC}(G)$ be the generator of a Markovian convolution semigroup $(S(t))_{t\geq 0}$ and $\Lambda\subset[0,\infty)$ be bounded. For all $\lambda\in\Lambda$ let $A_\lambda:=\lambda A$ and $S_\lambda(t):=S(\lambda t)$ for $t\geq 0$. Then, the assumptions (A1) and (A2) are satisfied with $D=D(A)$. Hence, there exists a nonlinear expectation space $(\Omega,\mathscr{F},\mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\geq 0}$ such that for all $u_0\in\mathrm{BUC}(G)$ the function

$$u(t,x):=\big(u(t)\big)(x):=\mathcal{E}(u_0(x+X_t)) \quad (t\geq 0, x\in G)$$

is a $D(A)$-viscosity solution of the fully nonlinear PDE

$$\begin{aligned} u_t(t,x) &= \sup_{\lambda\in\Lambda}\big(\lambda A u(t)\big)(x), \quad (t,x)\in(0,\infty)\times G,\\ u(0,x) &= u_0(x), \quad x\in G. \end{aligned}$$

For example, A could be the generator which is related to a cylindrical Wiener process on a separable Hilbert space (see e.g. [22] or [57]). The obtained Lévy process can be seen as a cylindrical G-Wiener process.

3.21 Example (Approximating the Brownian Motion by compound Poisson processes)**.** For all $h>0$ let $\mu_h:=\frac{1}{h^2}\delta_h$ and consider $\Lambda:=\big\{(0,0,\mu_h)\,|\,h>0\big\}$ in Example 3.19 with $d=1$. Then, we have that

$$\sup_{h>0}\int_{\mathbb{R}} 1\wedge|y|^2\,\mathrm{d}\mu_h(y)=\sup_{h>0}\frac{1}{h^2}\int_{\mathbb{R}}|y|^2\,\mathrm{d}\delta_h(y)=1.$$

Note that

$$\left\|A_{(0,0,\mu_h)}f-\frac12 f''\right\|_\infty=\sup_{x\in\mathbb{R}}\left|\frac{f(x+h)-f(x)-f'(x)h}{h^2}-\frac12 f''(x)\right|\to 0$$

as $h\searrow 0$ for all $f\in\mathrm{BUC}^2(\mathbb{R})$. By Example 3.19, there exists a nonlinear expectation space $(\Omega,\mathscr{F},\mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\geq 0}$ such that for all $u_0\in\mathrm{BUC}(\mathbb{R})$ the function

$$u(t,x):=\big(u(t)\big)(x):=\mathcal{E}(u_0(x+X_t)) \quad (t\geq 0, x\in\mathbb{R})$$

is a $\mathrm{BUC}^2(\mathbb{R})$-viscosity solution of the fully nonlinear PDE

$$\begin{aligned} u_t(t,x) &= \sup_{0<h\leq 1}\big(A_{0,0,\mu_h}u(t)\big)(x), \quad (t,x)\in(0,\infty)\times\mathbb{R},\\ u(0,x) &= u_0(x), \quad x\in\mathbb{R}. \end{aligned}$$

Note that the assumptions in Example 3.18 are not fulfilled and therefore u is not a classical solution.

3.22 Example (Cauchy distributed jumps)**.** For $\alpha > 0$ let μ_α be given by

$$\mu_\alpha\big((-\infty, b)\big) := \frac{\alpha}{\pi}\int_{-\infty}^{b} \frac{1}{y^2+\alpha^2}\,\mathrm{d}y = \frac{1}{2} + \arctan\left(\frac{b}{\alpha}\right)$$

for $b \in \mathbb{R}$. Let $A \subset (0,\infty)$. Then, we have that

$$\sup_{\alpha\in A}\int_{\mathbb{R}} 1 \wedge |y|^2\,\mathrm{d}\mu_\alpha(y) \le \sup_{\alpha\in A}\mu_\alpha(\mathbb{R}) = 1.$$

By Example 3.18 with $G = \mathbb{R}$ (using the notation from Example 3.19), there exists a nonlinear expectation space $(\Omega, \mathscr{F}, \mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\ge0}$ such that for all $u_0 \in \mathrm{BUC}(\mathbb{R})$ the function

$$u(t,x) := \big(u(t)\big)(x) := \mathcal{E}(u_0(x + X_t)) \quad (t \ge 0, x \in \mathbb{R})$$

is the unique classical solution of the fully nonlinear PDE

$$\begin{aligned} u_t(t,x) &= \sup_{\alpha\in A}\big(A_{0,0,\mu_\alpha}u(t)\big)(x), \quad (t,x) \in (0,\infty)\times\mathbb{R},\\ u(0,x) &= u_0(x), \quad x \in \mathbb{R} \end{aligned}$$

with $u \in C^1([0,\infty);\mathrm{BUC}(\mathbb{R}))$ and $\mathcal{A}u \in C([0,\infty);\mathrm{BUC}(\mathbb{R}))$.

3.23 Example (Lévy processes on the d-dimensional Torus $\mathbb{T}^d$)**.** Let $d \in \mathbb{N}$ and $\mathbb{T}^d$ denote the d-dimensional Torus, represented by $[-\pi,\pi)^d$. We say that (b,Σ,μ,ν) is a *Lévy quadruple* if $b \in \mathbb{R}^d$, $\Sigma \in \mathbb{R}^{d\times d}$ is a symmetric positive semidefinite matrix, μ is a positive finite measure on $\mathbb{T}^d$ and ν is a positive measure on $\mathbb{T}^d$ with $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{T}^d} |y|^2\,\mathrm{d}\nu(y) < \infty.$$

For each Lévy quadruple (b,Σ,μ,ν) we then consider

$$\begin{aligned}\big(A_{b,\Sigma,\mu,\nu}f\big)(x) := b\cdot\nabla f(x) + \frac{1}{2}\mathrm{tr}\big(\Sigma\nabla^2 f(x)\big) + \int_{\mathbb{T}^d} f(x+y) - f(x)\,\mathrm{d}\mu(y)\\ + \int_{\mathbb{T}^d} f(x+y) - f(x) - \nabla f(x)\cdot y\,\mathrm{d}\nu(y)\end{aligned}$$

for $x \in \mathbb{R}^d$ and $f \in D(A_{b,\Sigma,\mu,\nu}) := \big\{g \in C(\mathbb{T}^d)\,\big|\, A_{b,\Sigma,\mu,\nu}g \in C(\mathbb{T}^d)\big\}$. Then, for any Lévy quadruple (b,Σ,μ,ν) the operator $A_{b,\Sigma,\mu,\nu}$ generates a Markovian convolution semigroup $(S_{b,\Sigma,\mu,\nu}(t))_{t\ge0}$ of linear operators on $\mathbb{T}^d$. We also refer to [51] for a Lévy-Khintchine formula on compact Lie groups, called Hunt's formula, as well as Example 2.33 and 2.34 for an explicit computation of the resolvent in one dimension for the levy quadruples $(0,2,0,0)$ and $(1,0,0,0)$, respectively.

We start by proving that for any Lévy quadruple (b,Σ,μ,ν) the operator $A_{b,\Sigma,\mu,\nu}$ generates a Markovian convolution semigroup $(S_{b,\Sigma,\mu,\nu}(t))_{t\geq 0}$ of linear operators on $\mathbb{T}^d$. In order to do so, let (b,Σ,μ,ν) be a Lévy quadruple. Then, we define

$$\eta f := \int_{(\pi,\pi]^d} f(y)\,\mathrm{d}\nu(y) + \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \lambda_k \int_{(-\pi,\pi]^d} f(x+2\pi k)\,\mathrm{d}\mu(y)$$

for $f \in \mathscr{L}^\infty(\mathbb{R}^d)$, where $\lambda_k \geq 0$ for all $k \in \mathbb{Z}^d \setminus \{0\}$ and $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \lambda_k = 1$. Then, η is a Lévy measure on $\mathbb{R}^d$ (cp. Example 3.19), and therefore, (b,Σ,η) is a Lévy triplet. Hence, there exists a Markovian convolution semigroup $(S_{b,\Sigma,\eta}(t))_{t\geq 0}$ of linear operators on $\mathbb{R}^d$ with generator $A_{b,\Sigma,\eta}$. As the space $C(\mathbb{T}^d)$ of all 2π-periodic continuous functions is a closed subspace of $\mathrm{BUC}(\mathbb{R}^d)$, which is invariant under $S_{b,\Sigma,\eta}(t)$ for all $t \geq 0$, we obtain that

$$S(t) := \big(S_{b,\Sigma,\eta}(t)\big)|_{C(\mathbb{T}^d)} \quad (t \geq 0)$$

defines a Markovian convolution semigroup of linear operators on $\mathbb{T}^d$. Let A denote the generator of the semigroup $(S(t))_{t\geq 0}$. As $C(\mathbb{T}^d)$ is a closed subspace of $\mathrm{BUC}(\mathbb{R}^d)$ and $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} \lambda_k = 1$, we get that

$$Af = A_{b,\Sigma,\eta} f = A_{b,\Sigma,\mu,\nu} f$$

for all $f \in D(A)$. In particular, $A_{b,\Sigma,\mu,\nu}$ is the generator of $(S(t))_{t\geq 0}$.

Now, let Λ be a set of Lévy quadruples with

$$\sup_{(b,\Sigma,\mu,\nu)\in\Lambda} \left(1 + |b| + |\Sigma| + \mu(\mathbb{T}^d) + \int_{\mathbb{T}^d} |y|^2\,\mathrm{d}\nu(y)\right) < \infty. \tag{3.17}$$

Let $f \in C^2(\mathbb{T}^d) = \mathrm{BUC}^2(\mathbb{T}^d)$. Then, in a similar way as in Example 3.19, one can show that $\{A_{b,\Sigma,\mu,\nu} f \colon (b,\Sigma,\mu,\nu) \in \Lambda\}$ is bounded and equicontinuous. Therefore, the assumptions (A1) and (A2) are satisfied. Hence, there exists a nonlinear expectation space $(\Omega,\mathscr{F},\mathcal{E})$ and an $\mathcal{E}$-Lévy process $(X_t)_{t\geq 0}$ such that for all $u_0 \in C(\mathbb{T}^d)$ the function

$$u(t,x) := \big(u(t)\big)(x) := \mathcal{E}(u_0(x+X_t)) \quad (t \geq 0, x \in \mathbb{T}^d)$$

is a $C^2(\mathbb{T}^d)$-viscosity solution of the fully nonlinear PDE

$$\begin{aligned} u_t(t,x) &= \sup_{(b,\Sigma,\mu,\nu)\in\Lambda} \big(A_{b,\Sigma,\mu,\nu} u(t)\big)(x), \quad (t,x) \in (0,\infty)\times\mathbb{T}^d, \\ u(0,x) &= u_0(x), \quad x \in \mathbb{T}^d. \end{aligned}$$

We would like to mention at this point that, using Hunt's formula, it is also thinkable to consider fully nonlinear PDEs of this type on compact Lie groups under a condition which is similar to (3.17).

3.2 Continuous-time Markov chains

In this section, we consider time-homogeneous continuous-time Markov chains with a finite state space G with $d := |G| \in \mathbb{N}$ endowed with the discrete topology 2^G. Then, G is a Polish space and $\mathscr{L}^\infty(G, 2^G) = \mathrm{BUC}(G)$ can be identified by $\mathbb{R}^d$ since we may w.l.o.g. assume that $G = \{1, \dots, d\}$. We therefore equip $\mathbb{R}^d$ with the supremum norm $\|\cdot\|_\infty$ and the operator norm $\|Q\|$ for a matrix $Q \in \mathbb{R}^{d\times d}$ will always be w.r.t. this norm. In the linear case, every time-homogeneous continuous-time Markov chain can be related to a so called *Q-matrix* and vice versa. We briefly illustrate this relation and refer to Norris [61] for a detailed discussion of Markov chains.

A matrix $Q = (q_{ij})_{1\le i,j\le d} \in \mathbb{R}^{d\times d}$ is called a *Q-matrix* if it satisfies the following conditions:

(i) $q_{ii} \le 0$ for all $i \in \{1, \dots, d\}$,

(ii) $-q_{ij} \le 0$ for all $i, j \in \{1, \dots, d\}$ with $i \ne j$,

(iii) $Q1 = 0$, i.e. $\sum_{j=1}^d q_{ij} = 0$ for all $i \in \{1, \dots, d\}$.

Let $Q \in \mathbb{R}^{d\times d}$ and $S_Q(t) := e^{tQ}$ be the matrix exponential of tQ for all $t \ge 0$. Then, it is well-known that $(S_Q(t))_{t\ge 0}$ defines a continuous semigroup of linear operators on $\mathbb{R}^d = \mathscr{L}^\infty(G, 2^G)$ with generator Q. Moreover, one can show that Q is a Q-matrix if and only of $S_Q(t)$ is a kernel for all $t \ge 0$ (see e.g. [61, Theorem 2.1.2]). Note that $S_Q(t)$ is a kernel if and only if $S_Q(t) \in \mathbb{R}^{d\times d}$ is a *stochastic matrix*, i.e.

(i) $e_i^T S_Q(t) e_j \ge 0$ for all $i, j \in \{1, \dots, d\}$,

(ii) $S_Q(t)1 = 1$.

Therefore, by Example 2.29, every Q-matrix Q can be uniquely related to a time-homogeneous Markov process $(\Omega, \mathscr{F}, (\mathbb{E}^x)_{x\in G}, (X_t)_{t\ge 0})$ with state space G and vice verca. In this case, restricting the time variable to $\mathbb{N}_0$ yields a discrete-time Markov chain with *transition matrix* $P := S_Q(1)$ (see also Example 2.20).

In the sequel, we show that a similar relation holds in the sublinear case using duality theory and the construction from Section 3.1. We start with the definition of a *Q-operator*. Throughout, we denote by e_i the i-th unit vector in $\mathbb{R}^d$ and by α the vector which is constantly equal to $\alpha \in \mathbb{R}$, i.e. $\alpha = \sum_{i=1}^d \alpha e_i$.

3.24 Definition. A mapping $\mathcal{Q}\colon \mathbb{R}^d \to \mathbb{R}^d$ is called a *Q-operator* if the following conditions are satisfied:

(i) $e_i^T \mathcal{Q}(e_i) \le 0$ for all $i \in \{1, \dots, d\}$

(ii) $e_i^T \mathcal{Q}(-e_j) \le 0$ for all $i, j \in \{1, \dots, d\}$ with $i \ne j$,

(iii) $\mathcal{Q}(\alpha) = 0$ for all $\alpha \in \mathbb{R}$.

We say that a set $\mathscr{P} \subset \mathbb{R}^{d\times d}$ is *bounded* if

$$\sup_{1\le i,j\le d} \sup_{Q\in\mathscr{P}} |q_{ij}| < \infty.$$

For sublinear Q-operators we then have the following characterization.

3.25 Proposition. *Let $\mathcal{Q}\colon \mathbb{R}^d \to \mathbb{R}^d$ be a sublinear mapping. Then, the following statements are equivalent:*

a) $\mathcal{Q}$ is a Q-operator,

b) There exists a bounded set $\mathscr{P} \subset \mathbb{R}^{d\times d}$ of Q-matrices such that

$$\mathcal{Q}(f) = \max_{Q\in\mathscr{P}} Qf \quad (f \in \mathbb{R}^d),$$

where the maximum is understood componentwise.

Proof. First assume that $\mathcal{Q}$ is a Q-operator. Then, by duality theory in $\mathbb{R}^d$, there exists a set $\mathscr{P} \subset \mathbb{R}^{d\times d}$ with

$$\mathcal{Q}f = \max_{Q\in\mathscr{P}} Qf \quad (f \in \mathbb{R}^d).$$

Notice that $\mathcal{Q}$ is sublinear in every component. It remains to show that $\mathscr{P}$ is bounded and only consists of Q-matrices. Let $Q = (q_{ij})_{1\le i,j\le d} \in \mathscr{P}$ and $i,j \in \{1,\ldots,d\}$ with $i \neq j$. Then, it holds that

$$q_{ii} \le e_i\mathcal{Q}(e_i) \le 0 \quad \text{and} \quad -e_i\mathcal{Q}(-e_i) \le q_{ii},$$

which shows that $-e_i\mathcal{Q}(-e_i) \le q_{ii} \le q_{ii} \le 0$. Moreover, we have that

$$-q_{ij} \le e_i^T\mathcal{Q}(-e_j) \le 0 \quad \text{and} \quad q_{ij} \le e_i\mathcal{Q}(e_j),$$

which implies that $-e_i\mathcal{Q}(e_j) \le -q_{ij} \le 0$. As $\mathcal{Q}(\alpha) = 0$ for all $\alpha \in \mathbb{R}$, it follows that $Q1 = 0$. This shows that Q is a Q-matrix and that the set $\mathscr{P}$ is bounded.

Now, assume that

$$\mathcal{Q}(f) = \sup_{Q\in\mathscr{P}} Qf \quad (f \in \mathbb{R}^d),$$

where $\mathscr{P} \subset \mathbb{R}^{d\times d}$ is a bounded set of Q-matrices. Then,

$$\mathcal{Q}(\alpha) = \sup_{Q\in\mathscr{P}} Q\alpha = 0$$

for all $\alpha \in \mathbb{R}$. Moreover,

$$e_i\mathcal{Q}(e_i) = \sup_{Q\in\mathscr{P}} q_{ii} \le 0$$

for all $i \in \{1,\ldots,d\}$ and

$$e_i\mathcal{Q}(-e_j) = \sup_{Q\in\mathscr{P}} -q_{ij} \le 0$$

for all $i,j \in \{1,\ldots,d\}$ with $i \neq j$. □

3.26 Remark. The set $\mathscr{P}$ in Proposition 3.25 can be chosen to be convex and compact by considering the closed convex hull of $\mathscr{P}$. Notice that $\mathbb{R}^{d\times d}$ is finite-dimensional and therefore, compactness is equivalent to closedness and boundedness.

3.27 Corollary. *Let $\mathcal{Q}\colon \mathbb{R}^d \to \mathbb{R}^d$ be a sublinear Q-operator. Then, $\mathcal{Q}$ is Lipschitz continuous.*

Proof. Let $\mathscr{P} \subset \mathbb{R}^{d\times d}$ be a bounded set of Q-matrices with

$$\mathcal{Q}(f) = \max_{Q\in\mathscr{P}} Qf$$

for all $f \in \mathbb{R}^d$. As $\mathscr{P} \subset \mathbb{R}^{d\times d}$ is bounded and all norms on $\mathbb{R}^{d\times d}$ are equivalent, it follows that

$$L := \sup_{Q\in\mathscr{P}} \|Q\| < \infty.$$

Let $f, g \in \mathbb{R}^d$ and $i \in \{1, \ldots, d\}$. Then, there exists a Q-matrix $Q \in \mathscr{P}$ with $e_i^T Qf = e_i^T \mathcal{Q}(f)$. It follows that

$$e_i^T\mathcal{Q}(f) - e_i^T\mathcal{Q}(g) = e_i^T Qf - e_i^T\mathcal{Q}(g) \leq e_i^T Q(f-g) \leq \|Q\| \cdot \|f-g\|_\infty \leq L \cdot \|f-g\|_\infty.$$

By a symmetry argument, we obtain that

$$|e_i^T\mathcal{Q}(f) - e_i^T\mathcal{Q}(g)| \leq L \cdot \|f-g\|_\infty,$$

which implies that $\|\mathcal{Q}(f) - \mathcal{Q}(g)\|_\infty \leq L \cdot \|f-g\|_\infty$. □

3.28 Remark. Let $\mathcal{Q}\colon \mathbb{R}^d \to \mathbb{R}^d$ be a sublinear Q-operator. As $\mathcal{Q}$ is Lipschitz continuous, the Picard-Lindelöf Theorem asserts that the nonlinear ordinary differential equation (ODE)

$$u'(t) = \mathcal{Q}(u(t)), \quad t \geq 0$$

with $u(0) = u_0 \in \mathbb{R}^d$ has a unique solution $u \in C^1([0,\infty);\mathbb{R}^d)$.

Throughout the rest of this section, let $\mathcal{Q}$ be a fixed sublinear Q-operator and $\mathscr{P} \subset \mathbb{R}^{d\times d}$ be a bounded set of Q-matrices with

$$\mathcal{Q}(f) = \max_{Q\in\mathscr{P}} Qf$$

for all $f \in \mathbb{R}^d$ (see Proposition 3.25). For all $Q \in \mathscr{P}$ let $(S_Q(t))_{t\geq 0}$ be the continuous semigroup of linear operators on $\mathbb{R}^d$ with generator Q. For all $h \geq 0$ we define

$$\mathcal{E}_h(f) := \sup_{Q\in\mathscr{P}} S_Q(h)f,$$

where the supremum is componentwise. Then, $\mathcal{E}_h$ defines a sublinear kernel from $\mathbb{R}^d$ to $\mathbb{R}^d$ for all $h \geq 0$. For a partition $\pi = \{t_0, t_1, \ldots, t_m\} \in P$ with $m \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_m$ (cf. Section 3.1), we set

$$\mathcal{E}_\pi := \mathcal{E}_{t_1-t_0} \ldots \mathcal{E}_{t_m-t_{m-1}}$$

and $\mathcal{E}_{\{0\}} := \mathcal{E}_0$. We finally define the Nisio semigroup

$$\mathscr{S}(t)f := \sup_{\pi\in P_t} \mathcal{E}_\pi(f)$$

for all $f \in \mathbb{R}^d$ and $t \geq 0$, where the supremum again is to be understood componentwise. Then, with almost literally the same proofs as in Section 3.1, replacing Λ by $\mathscr{P}$, we obtain the following theorem.

3.29 Theorem. *The family $(\mathscr{S}(t))_{t\geq 0}$ defines a continuous semigroup of kernels from $\mathbb{R}^d$ to $\mathbb{R}^d$.*

Using Theorem 2.28, Theorem 3.29 implies that there exists a sublinear time-homogeneous Markov process $(\Omega, \mathscr{F}, (\mathcal{E}^x)_{x\in G}, (X_t)_{t\geq 0})$ such that

$$\big(\mathscr{S}(t)f\big)(x) = \mathcal{E}^x(f(X_t))$$

for all $x \in G$, $t \geq 0$ and $f \in \mathbb{R}^d$. Restricting the time parameter of this process to $\mathbb{N}_0$, leads to a discrete-time Markov chain with transition operator $\mathscr{S}(1)$ (cf. Example 2.20).

The following theorem states that the family $(\mathscr{S}(h))_{h\geq 0}$, we constructed so far, is differentiable at zero. The proof is very similar to the proof of Theorem 3.13.

3.30 Theorem. *For all $f \in \mathbb{R}^d$ it holds that*

$$\left\| \frac{\mathscr{S}(h)f - f}{h} - \mathscr{Q}f \right\|_\infty \to 0, \quad h \searrow 0.$$

Proof. Let $\varepsilon > 0$. By Lemma 3.12 and Lemma 3.3 d), there exists some $h_0 > 0$ such that

$$\|S_Q(h)Qf - Qf\|_\infty \leq \|S_Q(h) - I_d\| \cdot \|Qf\|_\infty \leq \big(e^{\|Q\|h} - 1\big)\|Q\| \cdot \|f\|_\infty \leq \varepsilon$$

and

$$\mathscr{S}(h)\mathscr{Q}f - \mathscr{Q}f \leq \varepsilon$$

for all $0 < h \leq h_0$. Let $0 < h \leq h_0$. Then, we get that

$$\mathscr{S}(h)f - f \geq S_Q(h)f - f = \int_0^h S_Q(s)Qf \,\mathrm{d}s \geq \big(Qf - \varepsilon\big)h.$$

Dividing by h and taking the supremum over all $Q \in \mathscr{P}$, it follows that

$$\frac{\mathscr{S}(h)f - f}{h} \geq \mathscr{Q}f - \varepsilon. \tag{3.18}$$

By Lemma 3.11, we have that

$$\mathscr{S}(h)f - f - h\mathscr{Q}f \leq \int_0^h \mathscr{S}(s)\mathscr{Q}f \,\mathrm{d}s - h\mathscr{Q}f = \int_0^h \mathscr{S}(s)\mathscr{Q}f - \mathscr{Q}f \,\mathrm{d}s \leq h\varepsilon$$

Again, dividing by $h > 0$ yields that

$$\frac{\mathscr{S}(h)f - f}{h} - \mathscr{Q}f \leq \varepsilon.$$

Together with (3.18) this implies that

$$\left\| \frac{\mathscr{S}(h)f - f}{h} - \mathscr{Q}f \right\|_\infty \leq \varepsilon.$$

□

We therefore obtain the second main result of this section which asserts that the semigroup $(\mathscr{S}(t))_{t\geq 0}$ is the unique classical solution to the ODE considered in Remark 3.28. This is a direct consequence of the uniqueness obtained from the Picard-Lindelöf Theorem.

3.31 Corollary. *Let $u_0 \in \mathbb{R}^d$ and $u(t) := \mathscr{S}(t)u_0$ for $t \geq 0$. Then, $u \in C^1([0,\infty);\mathbb{R}^d)$ is the unique classical solution of the ODE*

$$u'(t) = \mathscr{Q}u(t), \quad t \geq 0$$

with $u(0) = u_0 \in \mathbb{R}^d$.

Bibliography

[1] H. Amann and J. Escher. *Analysis. III*. Birkhäuser Verlag, Basel, 2009. Translated from the 2001 German original by Silvio Levy and Matthew Cargo.

[2] D. Applebaum. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2009.

[3] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2011.

[4] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[5] D. Bartl. Pointwise dual representation of dynamic convex expectations. *Preprint*.

[6] H. Bauer. *Probability theory*, volume 23 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1996. Translated from the fourth (1991) German edition by Robert B. Burckel and revised by the author.

[7] H. Bauer. *Measure and integration theory*, volume 26 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2001. Translated from the German by Robert B. Burckel.

[8] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control*, volume 139 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978. The discrete time case.

[9] D. Bierlein. Über die Fortsetzung von Wahrscheinlichkeitsfeldern. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 1:28–46, 1962/1963.

[10] P. Billingsley. *Weak convergence of measures: Applications in probability*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1971. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 5.

[11] V. I. Bogachev. *Measure theory. Vol. I*. Springer-Verlag, Berlin, 2007.

[12] V. I. Bogachev. *Measure theory. Vol. II*. Springer-Verlag, Berlin, 2007.

[13] N. Bourbaki. *Theory of sets*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2004. Reprint of the 1968 English translation [Hermann, Paris; MR0237342].

[14] X. Cabré and L. A. Caffarelli. Regularity for viscosity solutions of fully nonlinear equations $F(D^2u) = 0$. *Topol. Methods Nonlinear Anal.*, 6(1):31–48, 1995.

[15] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and A. Rustichini. Niveloids and their extensions: risk measures on small domains. *J. Math. Anal. Appl.*, 413(1):343–360, 2014.

[16] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electron. J. Probab.*, 11:no. 3, 57–106, 2006.

[17] P. Cheridito, M. Kupper, and L. Tangpi. Representation of increasing convex functionals with countably additive measures. *Preprint*, 2015.

[18] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs. *Comm. Pure Appl. Math.*, 60(7):1081–1110, 2007.

[19] G. Choquet. Forme abstraite du théorème de capacitabilité. *Ann. Inst. Fourier. Grenoble*, 9:83–89, 1959.

[20] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probab. Theory Related Fields*, 123(1):1–27, 2002.

[21] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[22] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[23] F. Delbaen. *Coherent risk measures*. Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa, 2000.

[24] F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, Berlin, 2002.

[25] F. Delbaen. The structure of m-stable sets and in particular of the set of risk neutral measures. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*, volume 1874 of *Lecture Notes in Math.*, pages 215–258. Springer, Berlin, 2006.

[26] F. Delbaen, S. Peng, and E. Rosazza Gianin. Representation of the penalty term of dynamic concave utilities. *Finance Stoch.*, 14(3):449–472, 2010.

[27] C. Dellacherie and P.-A. Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York; North-Holland Publishing Co., Amsterdam-New York, 1978.

[28] R. Denk, M. Kupper, and M. Nendel. Kolmogorov type and general extension results for nonlinear expectations. *Preprint*, 2016.

[29] R. Denk, M. Kupper, and M. Nendel. A semigroup theoretic approach to fully nonlinear PDEs. *In Preparation*, 2017.

[30] J. Diestel and J. J. Uhl, Jr. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[31] R. Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 2005.

[32] R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[33] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[34] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[35] K.-J. Engel and R. Nagel. *A short course on operator semigroups*. Universitext. Springer, New York, 2006.

[36] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[37] K. Fan. Minimax theorems. *Proc. Nat. Acad. Sci. U. S. A.*, 39:42–47, 1953.

[38] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[39] H. Föllmer and C. Klüppelberg. Spatial risk measures: local specification and boundary risk. In *Stochastic analysis and applications 2014*, volume 100 of *Springer Proc. Math. Stat.*, pages 307–326. Springer, Cham, 2014.

[40] H. Föllmer and I. Penner. Convex risk measures and the dynamics of their penalty functions. *Statist. Decisions*, 24(1):61–96, 2006.

[41] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance Stoch.*, 6(4):429–447, 2002.

[42] H. Föllmer and A. Schied. Robust preferences and convex measures of risk. In *Advances in finance and stochastics*, pages 39–56. Springer, Berlin, 2002.

[43] H. Föllmer and A. Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011. An introduction in discrete time.

[44] M. Frittelli and E. Rosazza Gianin. Dynamic convex risk measures. *Risk measures for the 21st century*, pages 227–248, 2004.

[45] M. Frittelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486, 2002.

[46] D. J. Hartfiel. *Markov set-chains*, volume 1695 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1998.

[47] F. Hausdorff. Summationsmethoden und Momentfolgen. I. *Math. Z.*, 9(1-2):74–109, 1921.

[48] F. Hausdorff. Summationsmethoden und Momentfolgen. II. *Math. Z.*, 9(3-4):280–299, 1921.

[49] H. Heuser. *Funktionalanalysis*. Mathematische Leitfäden. [Mathematical Textbooks]. B. G. Teubner, Stuttgart, third edition, 1992. Theorie und Anwendung. [Theory and application].

[50] M. Hu and S. Peng. G-lévy processes under sublinear expectations. 2009.

[51] G. A. Hunt. Semi-groups of measures on Lie groups. *Trans. Amer. Math. Soc.*, 81:264–293, 1956.

[52] H. Ishii and P.-L. Lions. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Differential Equations*, 83(1):26–78, 1990.

[53] N. Jacob. *Pseudo differential operators and Markov processes. Vol. I*. Imperial College Press, London, 2001. Fourier analysis and semigroups.

[54] N. Jacob. *Pseudo differential operators and Markov processes. Vol. III*. Imperial College Press, London, 2005. Markov processes and applications.

[55] K. Jänich. *Topology*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1984. With a chapter by Theodor Bröcker, Translated from the German by Silvio Levy.

[56] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.

[57] W. Liu and M. Röckner. *Stochastic partial differential equations: an introduction*. Universitext. Springer, Cham, 2015.

[58] A. Neufeld and M. Nutz. Nonlinear Lévy processes and their characteristics. *Trans. Amer. Math. Soc.*, 369(1):69–95, 2017.

[59] M. Nisio. On a non-linear semi-group attached to stochastic optimal control. *Publ. Res. Inst. Math. Sci.*, 12(2):513–537, 1976/77.

[60] M. Nisio. *Stochastic control theory*, volume 72 of *Probability Theory and Stochastic Modelling*. Springer, Tokyo, second edition, 2015. Dynamic programming principle.

[61] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.

[62] E. Pardoux. *Markov processes and applications*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester; Dunod, Paris, french edition, 2008. Algorithms, networks, genome and finance.

[63] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[64] S. Peng. Filtration consistent nonlinear expectations and evaluations of contingent claims. *Acta Math. Appl. Sin. Engl. Ser.*, 20(2):191–214, 2004.

[65] S. Peng. Nonlinear expectations and nonlinear Markov chains. *Chinese Ann. Math. Ser. B*, 26(2):159–184, 2005.

[66] S. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 541–567. Springer, Berlin, 2007.

[67] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Process. Appl.*, 118(12):2223–2253, 2008.

[68] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.

[69] W. Rudin. *Functional analysis*. McGraw-Hill Book Co., New York-Düsseldorf-Johannesburg, 1973. McGraw-Hill Series in Higher Mathematics.

[70] W. Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976. International Series in Pure and Applied Mathematics.

[71] W. Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[72] K. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[73] W. Schachermayer. Optimal investment in incomplete financial markets. In *Mathematical finance—Bachelier Congress, 2000 (Paris)*, Springer Finance, pages 427–462. Springer, Berlin, 2002.

[74] H. H. Schaefer. *Topological vector spaces*. Springer-Verlag, New York-Berlin, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.

[75] A. Schied. Risk measures and robust optimization problems. *Stoch. Models*, 22(4):753–831, 2006.

[76] H. M. Soner, N. Touzi, and J. Zhang. Martingale representation theorem for the G-expectation. *Stochastic Process. Appl.*, 121(2):265–287, 2011.

[77] H. M. Soner, N. Touzi, and J. Zhang. Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.*, 16:no. 67, 1844–1879, 2011.

[78] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[79] F. Trèves. *Topological vector spaces, distributions and kernels*. Dover Publications, Inc., Mineola, NY, 2006. Unabridged republication of the 1967 original.

[80] B. Z. Vulikh. *Introduction to the theory of partially ordered spaces*. Translated from the Russian by Leo F. Boron, with the editorial collaboration of Adriaan C. Zaanen and Kiyoshi Iséki. Wolters-Noordhoff Scientific Publications, Ltd., Groningen, 1967.

[81] D. Werner. *Funktionalanalysis*. Springer-Verlag, Berlin, extended edition, 2000.

[82] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[83] J. Yong and X. Y. Zhou. *Stochastic controls*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.