



Universiteit
Leiden
The Netherlands

Gibbs-non-Gibbs transitions and vector-valued integration

Zuijlen, W.B. van

Citation

Zuijlen, W. B. van. (2016, September 7). *Gibbs-non-Gibbs transitions and vector-valued integration*. Retrieved from <https://hdl.handle.net/1887/42823>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/42823>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/42823> holds various files of this Leiden University dissertation.

Author: Zuijlen, W. B. van

Title: Gibbs-non-Gibbs transitions and vector-valued integration

Issue Date: 2016-09-07

Gibbs-non-Gibbs Transitions and Vector-Valued Integration

Willem van Zuijlen

Cover: Wandering, Ricardo Pacheco.

Permission to use the picture has been kindly granted by the artist.

www.ricardopacheco.info

A catalogue record is available from the Leiden University Library.

The research in this thesis was supported by the ERC Advanced Grant VARIS-267356 of Frank den Hollander.

Gibbs-non-Gibbs Transitions and Vector-Valued Integration

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 7 september 2016
klokke 15.00 uur

door

Willem Bernard van Zuijlen
geboren te Zeeland
in 1989

Samenstelling van de promotiecommissie:

1^e Promotor:

Prof. dr. W. Th. F. den Hollander (Universiteit Leiden)

2^e Promotor:

Prof. dr. F. H. J. Redig (TU Delft)

Overige Leden:

Prof. dr. B. Edixhoven (Universiteit Leiden, secretary)

Prof. dr. R. Fernández (Universiteit Utrecht)

Prof. dr. K. Külske (Ruhr-Universität Bochum)

Prof. dr. B. de Pagter (TU Delft)

Contents

1	Introduction	1
§1.1	Introduction to Part I	2
§1.1.1	An informal description of equilibrium statistical mechanics . .	2
§1.1.2	An informal description of Gibbs measures: the Ising model . .	3
§1.1.3	An informal description of phase transitions	6
§1.1.4	A formal description of Gibbs measures	7
§1.1.5	Mean-field sequences and sequential Gibbsianness	11
§1.1.6	Gibbs-non-Gibbs transitions	12
§1.1.7	Main results in Part I	13
§1.2	Introduction to Part II	16
§1.2.1	Informal description of Riemann, Lebesgue, Bochner and Pettis integrability	16
§1.2.2	Main contents of Chapter 5	17
§1.2.3	An example of an integrable function with values in $C(\mathbb{R})$. . .	18
§1.2.4	Main contents of Chapter 6	23
I	Gibbs-non-Gibbs transitions of mean-field systems	25
2	Gibbs-non-Gibbs dynamical transitions for mean-field interacting Brownian motions	27
§2.1	Introduction and main results	28
§2.1.1	Background	28
§2.1.2	Outline	29
§2.1.3	Sequences of finite-volume mean-field Gibbs measures, Potential, Sequentially Gibbs	29
§2.1.4	Brownian motion dynamics	31
§2.1.5	Trajectories of the magnetisation (Intermezzo)	33
§2.1.6	Uniqueness of the minimisers of the rate function	35
§2.1.7	Examples	36
§2.1.8	Discussion	39
§2.2	Proof of Lemma 2.1.3	40
§2.3	Proof of Theorem 2.1.4	41
§2.4	Proof of Lemma 2.1.6	42

§2.4.1	Two preparatory lemmas	42
§2.4.2	Proof of Lemma 2.1.6	44
§2.5	Tools from convex analysis: proof of Theorem 2.1.11	48
Appendix		
§2.A	Key formulas	51
§2.B	Proper weakly continuous regular conditional probabilities	57
§2.C	Miscellaneous: additional material on potentials	58
§2.C.1	Differentiability of the evolved potential	62
3	Level 2 mean-field Gibbs-non-Gibbs transitions	67
§3.1	Introduction and main results	68
§3.1.1	Motivation	68
§3.1.2	Outline	68
§3.2	Some notations and definitions	69
§3.3	Preliminaries	69
§3.3.1	Exchangeability and permutations	69
§3.3.2	Convergence on the compactly supported probabilities	71
§3.3.3	Supports of measures	74
§3.4	Sequentially Gibbs	74
§3.5	Sequences of finite-volume mean-field Gibbs measures	77
§3.5.1	Sequences of finite-volume mean-field Gibbs measures	78
§3.5.2	Conditions on the potential implying sequential Gibbsianness	79
§3.5.3	Further discussions on Hamiltonians	84
§3.6	Transformations of Gibbs measures by probability kernels	86
§3.7	Independent transformations: sufficient condition for sequential Gibbsianness	94
Appendix		
§3.A	Continuity of the limit of functions	98
§3.B	Tilted product regular conditional kernel	99
4	Large deviations of continuous regular conditional probabilities	101
§4.1	Introduction and main results	102
§4.1.1	Literature	102
§4.1.2	Large deviations	103
§4.1.3	Main results	104
§4.1.4	Gibbs-non-Gibbs transitions and future research	106
§4.1.5	Outline	107
§4.2	Notations and conventions	107
§4.3	Regular conditional kernels being product regular conditional kernels	108
§4.4	Weakly continuous kernels	109
§4.5	Some facts about functions with compact sublevel sets	114
§4.6	Large deviations for product regular conditional probabilities	115
§4.7	Large deviations for regular conditional probabilities	122
§4.8	An application to conditional probabilities of empirical distributions	125
§4.9	Examples	130

Appendix	
§4.A An elementary fact about limsup and liminf	134
§4.B Sufficient bounds for large deviation bounds	134
§4.C Proof of Theorem 4.9.2	135
§4.D Miscellaneous: Additional remarks	137
§4.D.1 Regarding the conditions of Theorem 4.6.3	137
§4.D.2 Large deviations over different parameters	138
 II Integrals for functions with values in a partially ordered vector space	 141
5 Integrals for functions with values in a partially ordered vector space	143
§5.1 Introduction	144
§5.2 Some Notation	144
§5.3 The vertical extension	145
§5.4 The lateral extension	150
§5.5 Combining vertical and lateral extensions	158
§5.6 Embedding E in a (slightly) larger space	165
§5.7 Integration for functions with values in $\mathbb{R}$	166
§5.8 Extensions of integrals on simple functions	167
§5.9 Comparison with Bochner- and Pettis integral	173
§5.10 Extensions of Bochner integrable functions	176
§5.11 Discussion	179
Appendix	
§5.A Appendix	180
§5.B Miscellaneous: Additional ideas, remarks etc.	181
§5.B.1 Observations and remarks regarding Sections 5.3 – 5.10	181
§5.B.2 Mediated, splitting and similar notions: their relation to σ -Dedekind completeness	189
§5.B.3 Remarks about convolution	193
§5.B.4 Improper S_{VL} -functions	194
 6 Bochner integrals in ordered vector spaces	 199
§6.1 Introduction	200
§6.2 Notation	200
§6.3 Ordered Banach spaces with closed generating cones	201
§6.4 The Bochner integral on ordered Banach spaces	204
§6.5 An extension of the Bochner integral	208
§6.6 Covers of ordered Banach spaces with closed generating cones	211
§6.7 The integral for an Archimedean ordered vector space	212
§6.8 Comparison with other integrals	214
§6.9 An example: Convolution	216

Appendix	
§6.A Appendix: Absolutely dominating norms	218
§6.B Appendix: Banach cover of Köthe spaces	221
§6.C Miscellaneous: Additional ideas, remarks etc.	222
Bibliography	225
Index	232
Samenvatting	234
Acknowledgements	237
Curriculum Vitae	239



CHAPTER 1

Introduction

This thesis consists of two parts. Part I focusses on Gibbs-non-Gibbs transitions in mean-field systems, Part II focusses on integration for functions with values in a partially ordered vector space.

§1.1 Introduction to Part I

Gibbs measures are key objects in the field of equilibrium statistical mechanics. They describe the macroscopic state of a system of a large number of components that is in equilibrium. It may happen that when the system is transformed, for example, by a stochastic dynamics that runs over a certain time interval, the evolved state is no longer a Gibbs measure (think of a system of particles with a certain initial temperature in equilibrium that is being heated up). Such situations are considered in Part I of the thesis.

In Section 1.1.1 we first give an informal description of equilibrium statistical mechanics. In Section 1.1.2 we describe the idea behind Gibbs measures, considering the example of the Ising model. In Section 1.1.3 we describe the concept of a phase transition. In Section 1.1.4 we give the definition of a Gibbs measure. In Section 1.1.5 we describe mean-field systems and sequential Gibbsianness and relate them to the notion of a Gibbs measure. In Section 1.1.6 we describe the phenomenon of Gibbs-non-Gibbs transitions and pose the main questions related to it. In Section 1.1.7 we list our main contributions in Part I.

Even though the notion of Gibbs measure pervades this introduction, it serves mainly as a motivation for our study on mean-field systems. This introduction is inspired by other PhD-theses that have been written on Gibbs-non-Gibbs transitions, in particular, A. Opoku [71], V. Ermolaev [35], F. Wang [89], and J. Martínez [67].

§1.1.1 An informal description of equilibrium statistical mechanics

Statistical mechanics studies systems consisting of a large number of interacting components with the help of probability theory. Classically, the motion of particles of a gas or a liquid are described by Newton's laws, or equivalently by Hamilton equations. As this is a huge and complex system of differential equations (typically of the order of Avogadro's number, 6×10^{23}), it is impossible to solve it, either rigorously or numerically. However, we are not interested in the precise local microscopic behaviour of all the particles, but rather in their global macroscopic behaviour. If the system of particles is in equilibrium, which means that the macroscopic behaviour does not change in time, then this macroscopic behaviour is well-described by a Gibbs measure. This is a probability measure on the space of all possible configurations of the particles, describing the local laws of the particles in terms of an energy function. The energy of a configuration can be seen as a cost, because a lower energy of a configuration means that it is more likely to occur.

The definition of a Gibbs measure is a natural generalisation of the definition of the canonical ensemble in statistical mechanics, also called the Boltzmann distribution, for systems consisting of infinitely many particles. There are several ways to explain why Gibbs measures represent the macroscopic behaviour in equilibrium, for which we refer to the books by Ruelle [81], Sinai [17], Gallavotti [43], and Bovier [13].

In the next section we informally introduce such energy functions and explain how

Gibbs measures are described in terms of them.

§1.1.2 An informal description of Gibbs measures: the Ising model

Let us first describe the components, called spins, in a finite box, say $\boxplus = \{1, \dots, 6\}^2$. The spins take the value -1 ($\ominus$) or 1 ($\oplus$), so that the set of configurations is $\{-1, 1\}^{\boxplus}$. See Figure 1.1 for an example of a configuration. We consider a parameter $\beta \in [0, \infty)$, called the inverse temperature ($\beta = \frac{1}{T}$, where $T \in (0, \infty]$ denotes the temperature). For a configuration $\sigma = (\sigma(x))_{x \in \boxplus} \in \{-1, 1\}^{\boxplus}$ we write $H(\sigma)$ for its energy, which is given by

$$H(\sigma) = - \sum_{\substack{x, y \in \boxplus \\ |x - y| \leq 1}} \sigma(x)\sigma(y).$$

The function H on $\{-1, 1\}^{\boxplus}$ is called the Hamiltonian. The probability of a configuration σ , is given by

$$\mathbb{P}(\sigma) = \frac{1}{Z} e^{-\beta H(\sigma)},$$

where $Z \in (0, \infty)$ is the normalising constant, chosen such that $\mathbb{P}$ is a probability measure ($Z = \sum_{\sigma \in \{-1, 1\}^{\boxplus}} e^{-\beta H(\sigma)}$).

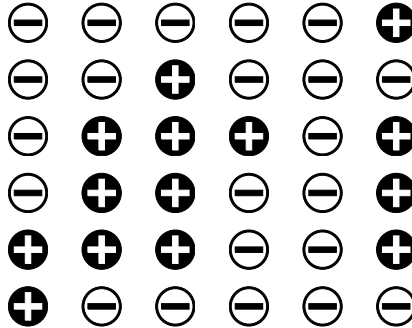


Figure 1.1: An example of a configuration of magnetic spins in a box.

If $\beta = 0$, corresponding to “infinite temperature”, then $\mathbb{P}$ is the uniform distribution on the configuration space $\{-1, 1\}^{\boxplus}$: every σ has probability $\frac{1}{Z}$, where Z is the total number of configurations. However, if $\beta > 0$, then the energy influences the probability $\mathbb{P}$. For a configuration σ , the contribution of the spins at two neighbouring positions x and y in $\boxplus$, $|x - y| = 1$, is -1 if the spins are aligned, $\sigma(x) = \sigma(y)$, and 1 if the spins are not aligned, $\sigma(x) \neq \sigma(y)$. The number of pairs x, y in $\boxplus$ that are neighbours is $4 \times 2 + 4 \times 4 \times 3 + 4 \times 4 \times 4 = 120$. Considering the configurations $\sigma_{\blacksquare}$ and $\sigma_{\blacksquare}$ as in Figure 1.2, we see that $H(\sigma_{\blacksquare}) = -120$ and $H(\sigma_{\blacksquare}) = 120$: each pair $x, y \in \boxplus$ of neighbours contributes 1 to the energy of $\sigma_{\blacksquare}$ and -1 to the energy of $\sigma_{\blacksquare}$. Because

$H(\sigma_{\boxplus})$ is less than $H(\sigma_{\boxminus})$, the configuration $\sigma_{\boxplus}$ has a larger probability than $\sigma_{\boxminus}$. Note that the configuration $\sigma_{\boxplus}$ in Figure 1.3 has the same number of $\oplus$ and $\ominus$, but has a larger probability than $\sigma_{\boxminus}$. This shows the dependence on the spatial geometry of the configuration. We will come back to this in Section 1.1.5 in the context of mean-field systems, where we do not consider spatial geometry.

The β indicates the strength or influence of the energy on the probability. For a large value of β , the typical configurations under $\mathbb{P}$ are those with low energy, i.e., those configurations where the $\oplus$ and $\ominus$ are clustered together. Conversely, for a small value of β , there is not much preference between different configurations, allowing them to be unclustered.

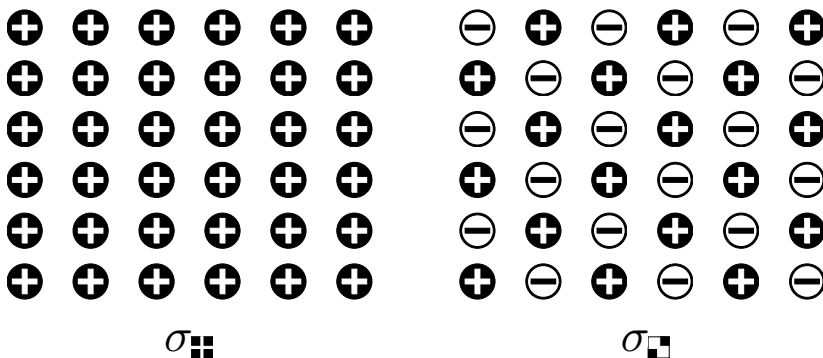


Figure 1.2: A fully aligned configuration and an alternating configuration.

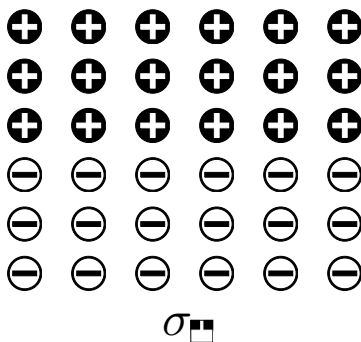


Figure 1.3: A configuration with plus and minus spins sorted.

The systems considered so far consist of a finite number of spins, with values in a finite space and without a boundary condition. In order to extend the notion of probability measure to an infinite number of spins, we will first discuss an example in which we take into account boundary conditions.

Instead of considering the 6×6 box $\boxplus = \{1, \dots, 6\}^2$, we consider the 8×8 box $\{0, 1, \dots, 7\}^2$, where the set of positions outside $\boxplus$ is considered to be the boundary, which we denote by B . So

$$B = \{(x, y) : (x, y) = (0, i), (7, i), (i, 0), (i, 7) \text{ for some } i \in \{0, \dots, 7\}\}.$$

We assume that the spins on the boundary B are fixed and we give the distribution of the configurations in $\boxplus$. Let η denote the boundary condition, i.e., $\eta((0, i))$, $\eta((7, i))$, $\eta((i, 0))$, $\eta((i, 7))$ are elements of $\{-1, 1\}$ for $i \in \{0, \dots, 7\}$ as in Figure 1.4.

$\eta(0,0)$	$\eta(0,1)$	$\eta(0,2)$	$\eta(0,3)$	$\eta(0,4)$	$\eta(0,5)$	$\eta(0,6)$	$\eta(0,7)$
$\eta(1,0)$	+	+	+	+	+	-	$\eta(1,7)$
$\eta(2,0)$	+	+	-	+	+	+	$\eta(2,7)$
$\eta(3,0)$	+	-	-	-	+	-	$\eta(3,7)$
$\eta(4,0)$	+	-	-	+	+	-	$\eta(4,7)$
$\eta(5,0)$	-	-	-	+	+	-	$\eta(5,7)$
$\eta(6,0)$	-	+	+	+	+	+	$\eta(6,7)$
$\eta(7,0)$	$\eta(7,1)$	$\eta(7,2)$	$\eta(7,3)$	$\eta(7,4)$	$\eta(7,5)$	$\eta(7,6)$	$\eta(7,7)$

Figure 1.4: A configuration with boundary condition η .

We choose the following energy function for configurations $\sigma = (\sigma(x))_{x \in \boxplus}$, depending on the boundary condition η :

$$H^\eta(\sigma) = - \sum_{\substack{x, y \in \boxplus \\ |x-y|=1}} \sigma(x)\sigma(y) - \sum_{\substack{x \in \boxplus, z \in B \\ |x-z|=1}} \sigma(x)\eta(z).$$

Then the probability of σ is given by

$$\mathbb{P}(\sigma) = \frac{1}{Z^\eta} e^{-\beta H^\eta(\sigma)}, \quad (1.1)$$

where $Z^\eta = \sum_{\sigma \in \{-1, 1\}^{\boxplus}} e^{-\beta H^\eta(\sigma)}$.

The example of the box $\boxplus = \{1, \dots, 6\}^2$ can of course be extended to general boxes, for example to those of the form $\Lambda_n = \{-n, -n+1, \dots, n-1, n\}^2$. The probability as in (1.1) for a box Λ_n and a boundary configurations η is called the finite-volume Gibbs measure with boundary condition η .

A Gibbs measure is a measure on “a box of infinite size”, $\mathbb{Z}^2$ for example. The finite-volume Gibbs measures with boundary conditions serve as candidate conditional probabilities for a Gibbs measure on the infinite lattice as it forms a consistent family of conditional probabilities. The idea is as follows. For each finite box $\Lambda_n \subset \mathbb{Z}^2$ and each configuration η outside Λ_n , i.e., $\eta \in \{-1, 1\}^{\mathbb{Z}^2 \setminus \Lambda_n}$, we have an energy function $H_{\Lambda_n}^\eta$ on $\{-1, 1\}^{\Lambda_n}$ and a partition sum $Z_{\Lambda_n}^\eta > 0$, similarly as defined above. Let μ be a probability measure on $\{-1, 1\}^{\mathbb{Z}^2}$, which is viewed as the law of the infinite configuration Θ , i.e., $\mathbb{P}_\mu(\Theta \in A) = \mu(A)$ for subsets $A \subset \{-1, 1\}^{\mathbb{Z}^2}$. Then μ is called a Gibbs measure whenever for all n and all $\eta \in \{-1, 1\}^{\mathbb{Z}^2 \setminus \Lambda_n}$ the probability under μ that the configuration Θ inside Λ_n equals σ given that the configuration Θ outside

Λ_n equals η is equal to $\frac{1}{Z_\eta} e^{-\beta H^\eta(\sigma)}$, i.e.,

$$\mathbb{P}_\mu \left(\Theta|_{\Lambda_n} = \sigma \mid \Theta|_{\Lambda_n^c} = \eta \right) = \frac{1}{Z_{\Lambda_n}^\eta} e^{-\beta H_{\Lambda_n}^\eta(\sigma)}. \quad (1.2)$$

The Hamiltonian $H_{\Lambda_n}^\eta$ is the energy function for the boundary condition η (see also Example 1.1.4).

We give a precise definition of a Gibbs measure in Section 1.1.4.

§1.1.3 An informal description of phase transitions

In this section we discuss the occurrence of phase transitions, which are important in statistical mechanics.

The most common phase transitions we experience in daily life concerns water: vaporisation of water to steam occurs at 100 °C, freezing of water to ice at 0 °C. Such a phenomenon, where around a certain value of a parameter there is a drastic change in the behaviour on the macroscopic level, is called a phase transition. Even though experimentally such phase transitions have been found in many systems, there is a rigorous mathematical description for only relatively few models.

Results on phase transitions are available for spins systems on a discrete lattice. Let us consider the example of the Ising model on $\mathbb{Z}^2$ described above in Section 1.1.2. For this model a phase transition occurs in the following sense. There exists a $\beta_c \in (0, \infty)$ such that for all $\beta < \beta_c$ there exist a unique measure μ such that (1.2) holds for all n and choices η , while for $\beta > \beta_c$ there exist two such measures μ_+ and μ_- , with $\mu_+ \neq \mu_-$. Both μ_- and μ_+ are limits as $n \rightarrow \infty$ of finite-volume Gibbs measures on Λ_n with boundary conditions η_n^- and η_n^+ , consisting of only $-1, \ominus$ and $1, \oplus$, respectively.

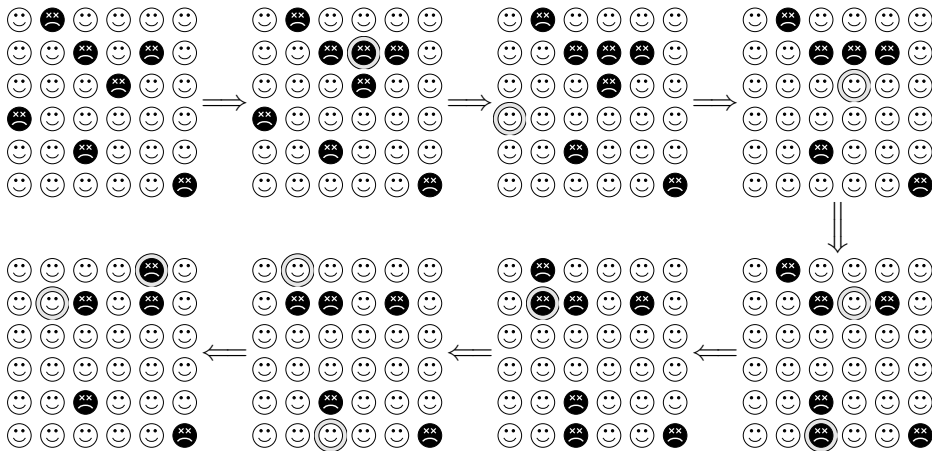


Figure 1.5: Possible evolution of infections and recoveries.

A different type of phase transition occurs in the following interacting particle system, called the contact process. We consider the lattice $\mathbb{Z}^2$ and imagine that every

site represents an individual, being either healthy ☺ or infected ☹. At each time $t \geq 0$ a subset of $\mathbb{Z}^2$ is infected. Over time, infected individuals recover at constant rate 1, while healthy individuals become infected at rate λ times the number of infected neighbours (where rate λ means that the transition from healthy to infected occurs according to an exponential distribution with mean $\frac{1}{\lambda}$). For example:

- (recovery) ☹ $\Rightarrow$ ☺ at rate 1,
- (infection) ☺ ☺ ☹ $\Rightarrow$ ☺ ☹ ☹ at rate 2.

A possible time evolution could look like Figure 1.5. In this model the following phase transition occurs. There exists a critical infection rate λ_c such that for $\lambda < \lambda_c$ the epidemic eventually dies out while for $\lambda > \lambda_c$ the epidemic has a positive probability to persist forever. The intuitive explanation for small λ is that recovery occurs much faster than infection, so that an infection is more likely to disappear than to spread to one the neighbours of the individual. The intuitive explanation for large λ is that the reverse happens. See Liggett [65] for more details on the phase transition of the contact process (and of the Ising model).

Many other phase transitions exists in statistical mechanics, for example, in percolation, in branching processes, transience versus recurrence of random walks, etc.

§1.1.4 A formal description of Gibbs measures

In this section we give the formal definition of a Gibbs measure, and an equivalent description that is used in Section 1.1.5 to indicate the similarities between the notion of Gibbs measure and the notion of sequentially Gibbs for mean-field systems. We refer the reader to the book by Georgii [44] and to van Enter, Fernández and Sokal [32], Fernández [38], for more details.

Let $(E, \mathcal{E})$ be a measurable space and $\mathbb{L}$ a countable set. Let $(\Omega, \mathcal{F})$ be the measurable space given by $\Omega = E^{\mathbb{L}}$ and $\mathcal{F} = \bigotimes_{l \in \mathbb{L}} \mathcal{E}$. For a nonempty $\Lambda \subset \mathbb{L}$ we write $\pi_{\Lambda} : \Omega \rightarrow E^{\Lambda}$ for the restriction map $\pi_{\Lambda}(\omega) = \omega|_{\Lambda}$. We write $\mathfrak{P}_0(\mathbb{L})$ for the set of nonempty finite subsets of $\mathbb{L}$. For probability measures λ on $(E, \mathcal{E})$ we write $\lambda^{\Lambda} = \bigotimes_{l \in \Lambda} \lambda$ for the product measure on $(E^{\Lambda}, \bigotimes_{l \in \Lambda} \mathcal{E})$ for all nonempty $\Lambda \subset \mathbb{L}$. For $\Lambda \subset \mathbb{L}$ and functions $\sigma : A \rightarrow E$, $\omega : B \rightarrow E$ with $\Lambda \subset A$ and $\Lambda^c \subset B$, we use the notation $\sigma \mathbb{1}_{\Lambda} + \omega \mathbb{1}_{\Lambda^c}$ for the function that maps $l \in \Lambda$ to $\sigma(l)$ and $l \in \Lambda^c$ to $\omega(l)$. In the statistical mechanics literature the following shorter notation is common:

$$\sigma_{\Lambda} \omega_{\Lambda^c} = \sigma \mathbb{1}_{\Lambda} + \omega \mathbb{1}_{\Lambda^c}.$$

1.1.1 Definition. An *interaction* on $(\Omega, \mathcal{F})$ is a function $\Phi : \mathfrak{P}_0(\mathbb{L}) \rightarrow M(\mathcal{F})$, or equivalently, a family $\Phi = (\Phi_A)_{A \in \mathfrak{P}_0(\mathbb{L})}$ of measurable functions for which $\Phi(A) = \Phi_A$ is $\mathcal{F}_A$ -measurable for all $A \in \mathfrak{P}_0(\mathbb{L})$. We call the interaction Φ *absolutely summable*

if Φ_A is a bounded function for all $A \in \mathfrak{P}_0(\mathbb{L})$ and if for all $l \in \mathbb{L}$,

$$\sum_{\substack{A \in \mathfrak{P}_0(\mathcal{F}) \\ A \ni l}} \|\Phi_A\|_\infty < \infty.$$

Let $\Lambda \in \mathfrak{P}_0(\mathbb{L})$. For an absolutely summable Φ we define the *Hamiltonian for volume* Λ , $H_\Lambda^\Phi : \Omega \rightarrow \mathbb{R}$, by

$$H_\Lambda^\Phi(\omega) = \sum_{\substack{A \in \mathfrak{P}_0(\mathbb{L}) \\ A \cap \Lambda \neq \emptyset}} \Phi_A(\omega) \quad (\omega \in \Omega). \quad (1.3)$$

Let λ be a probability measure on $(E, \mathcal{E})$. For $\Lambda \in \mathfrak{P}_0(\mathbb{L})$, the *Gibbs distribution in volume* Λ corresponding to Φ and λ , denoted by $\gamma_\Lambda^{\Phi, \lambda} : E^{\Lambda^c} \times \mathcal{F} \rightarrow [0, 1]$, is defined by

$$\gamma_\Lambda^{\Phi, \lambda}(\omega, A) = \frac{1}{Z_\Lambda^\Phi(\omega)} \int_{E^\Lambda} \mathbb{1}_A(\sigma_\Lambda \omega_{\Lambda^c}) e^{-H_\Lambda^\Phi(\sigma_\Lambda \omega_{\Lambda^c})} d\lambda^\Lambda(\sigma),$$

where

$$Z_\Lambda^\Phi(\omega) = \int_{E^\Lambda} e^{-H_\Lambda^\Phi(\sigma_\Lambda \omega_{\Lambda^c})} d\lambda^\Lambda(\sigma).$$

For $\omega \in E^{\Lambda^c}$ the measure $\gamma_\Lambda^{\Phi, \lambda}(\omega, \cdot)$ is called the *Gibbs distribution in volume* Λ with *boundary condition* ω .

A probability measure μ on $(\Omega, \mathcal{F})$ is called a *Gibbs measure for* Φ (and λ) whenever $\gamma_\Lambda^{\Phi, \lambda}$ is a regular conditional probability under μ with respect to π_{Λ^c} for all $\Lambda \in \mathfrak{P}_0(\mathbb{L})$, or equivalently, whenever for all $A \in \mathcal{F}$ the function $\Omega \rightarrow [0, 1], \omega \mapsto \gamma_\Lambda(\omega|_{\Lambda^c}, A)$ is a version of the conditional probability under μ of A with respect to $\mathcal{F}_{\Lambda^c}$, the σ -algebra generated by π_{Λ^c} , i.e., whenever

$$\mu(A \cap \pi_{\Lambda^c}^{-1}(B)) = \int_{E^{\Lambda^c}} \mathbb{1}_B(\zeta) \gamma_\Lambda(\zeta, A) d[\mu \circ \pi_{\Lambda^c}^{-1}](\zeta) \quad \left(B \in \bigotimes_{l \in \Lambda^c} \mathcal{E} \right),$$

or equivalently,

$$\mu(A \cap B) = \int_{E^\mathbb{L}} \mathbb{1}_B(\omega) \gamma_\Lambda(\omega|_{\Lambda^c}, A) d\mu(\omega) \quad (B \in \mathcal{F}_{\Lambda^c}).$$

A probability measure μ on Ω is called a *Gibbs measure* whenever there exists an absolutely summable interaction Φ such that μ is a Gibbs measure for Φ .

1.1.2 Remark. As Φ is absolutely summable, the Hamiltonian H_Λ^Φ is a bounded measurable function, which implies that the function $E^\Lambda \rightarrow \mathbb{R}$ given by

$$\sigma \mapsto e^{-H_\Lambda^\Phi(\sigma_\Lambda \omega_{\Lambda^c})},$$

is integrable for all $\omega \in E^{\Lambda^c}$, so that $Z_\Lambda^\Phi(\omega) \in (0, \infty)$.

1.1.3 Remark. In our definition we assume Φ to be absolutely summable, even though in the literature (like in Georgii [44]) it is only assumed that H_Λ^Φ exists, in the sense that the series in (1.3) converges.

1.1.4 Example (Ising model). Let $E = \{-1, 1\}$, λ be the uniform measure on $\{-1, 1\}$, i.e., $\lambda(\{-1\}) = \lambda(\{1\}) = \frac{1}{2}$, and $\mathbb{L} = \mathbb{Z}^d$ for some $d \in \mathbb{N}$. We define an interaction potential Φ as follows. Let $\Phi_{\{x,y\}}(\omega) = -\beta\omega(x)\omega(y)$ whenever $x, y \in \mathbb{Z}^d$ are neighbours, i.e., $|x - y| = 1$. Let $\Phi_A = 0$ for $A \in \mathfrak{P}_0(\mathbb{L})$ when $\#A \neq 2$ or when $A = \{x, y\}$ and x and y are not neighbours. Then Φ is absolutely summable. This interaction potential Φ corresponds to the Ising model described in Section 1.1.2: for $d = 2$, $\Lambda = \Lambda_n = \{-n, -n+1, \dots, n-1, n\}$ and $\sigma \in E^\Lambda, \eta \in E^{\Lambda^c}$ we have

$$H_\Lambda^\Phi(\sigma_\Lambda \eta_{\Lambda^c}) = H_\Lambda^\eta(\sigma).$$

The notion of a phase transition as explained in Section 1.1.3 is formalised in the following definition.

1.1.5 Definition (Phase transition). An interaction potential Φ is said to exhibit a *phase transition* if there is more than one probability measure that is a Gibbs measure for Φ .

1.1.6 Example (Phase transition for the Ising model). As mentioned in Section 1.1.3 the Ising model undergoes a phase transition in dimension 2. With the formal definitions of the present section we can define the two phases μ_- and μ_+ . Consider the situation as in Example 1.1.4. Let $\eta_n^-, \eta_n^+ \in \{-1, 1\}^{\Lambda_n^c}$ be given by $\eta_n^-(x) = -1$ and $\eta_n^+(x) = 1$ for all $n \in \mathbb{N}$. Then μ_- is the limit $\lim_{n \rightarrow \infty} \gamma_{\Lambda_n}^{\Phi, \lambda}(\eta_n^-, \cdot)$ and μ_+ is the limit $\lim_{n \rightarrow \infty} \gamma_{\Lambda_n}^{\Phi, \lambda}(\eta_n^+, \cdot)$ (see for example Bovier [13, Theorem 4.3.7]).

Using the notion of quasilocality of a specification we formulate in Theorem 1.1.11 an equivalent description for a measure to be Gibbs. We introduce these notions in Definition 1.1.7 and Definition 1.1.10 below.

1.1.7 Definition (Specification). A family $\Gamma = (\gamma_\Lambda)_{\Lambda \in \mathfrak{P}_0(\mathbb{L})}$ of probability kernels $\gamma_\Lambda : \Omega \times \mathcal{F} \rightarrow [0, 1]$ for which

- (i) $\gamma_\Lambda(\cdot, A)$ is $\mathcal{F}_{\Lambda^c}$ -measurable for all $A \in \mathcal{F}$,
- (ii) $\gamma_\Lambda(\omega, B) = \mathbb{1}_B(\omega)$ for all $\omega \in \Omega$ and $B \in \mathcal{F}_{\Lambda^c}$,
- (iii) $\Upsilon \subset \Lambda$ implies $\gamma_\Lambda(\omega, A) = \int \gamma_\Upsilon(\eta, A) d[\gamma_\Lambda(\omega, \cdot)](\eta)$ for all $\omega \in \Omega, A \in \mathcal{F}$,

is called a *specification*.

1.1.8 Example (Gibbsian specification). With $\gamma_\Lambda^{\Phi, \lambda}$ as in Definition 1.1.1, $\Gamma = (\gamma_\Lambda^{\Phi, \lambda})_{\Lambda \in \mathfrak{P}_0(\mathbb{L})}$ is a specification.

1.1.9 Definition. Define $\mathcal{F}_\Lambda$ for a nonempty $\Lambda \subset \mathbb{L}$ to be

$$\mathcal{F}_\Lambda = \left\{ \pi_\Lambda^{-1}(A) : A \in \bigotimes_{l \in \Lambda} \mathcal{E} \right\}.$$

Note that $\mathcal{F}_\Lambda$ is a σ -algebra and that $\Upsilon \subset \Lambda$ implies $\mathcal{F}_\Upsilon \subset \mathcal{F}_\Lambda$.

1.1.10 Definition (Local and quasilocal). A function $f : \Omega \rightarrow \mathbb{R}$ is called *local* if there exists a $\Lambda \in \mathfrak{P}_0(\mathbb{L})$ such that $f(\omega) = f(\eta)$ for $\omega, \eta \in \Omega$ with $\omega|_\Lambda = \eta|_\Lambda$ (i.e., $\omega(l) = \eta(l)$ for all $l \in \Lambda$). Equivalently, f is local if and only if f is $\mathcal{F}_\Lambda$ -measurable for some $\Lambda \in \mathfrak{P}_0(\mathbb{L})$.

A function $f : \Omega \rightarrow \mathbb{R}$ is called *quasilocal* if there exists a sequence $(f_n)_{n \in \mathbb{N}}$ of local functions that converges uniformly to f , i.e., $\lim_{n \rightarrow \infty} \sup_{\omega \in E^\mathbb{L}} |f_n(\omega) - f(\omega)| = 0$. Equivalently, f is quasilocal if and only if

$$\lim_{\Lambda \in \mathfrak{P}_0(\mathbb{L})} \sup_{\substack{\zeta, \eta \in \Omega \\ \zeta|_\Lambda = \eta|_\Lambda}} |f(\zeta) - f(\eta)| = 0. \quad (1.4)$$

A specification $\Gamma = (\gamma_\Lambda)_{\Lambda \in \mathfrak{P}_0(\mathbb{L})}$ is called *quasilocal* if $\omega \mapsto \int_\Omega f \, d\gamma(\omega, \cdot)$ is quasilocal for all bounded quasilocal $f : \Omega \rightarrow \mathbb{R}$ (or equivalently, for all bounded local functions f).

The following theorem is due to Sullivan [87] and Kozlov [58]. See also van Enter, Fernández and Sokal [32, Theorem 2.12] and Fernández [38, Theorem 4.78].

1.1.11 Theorem. *Let λ be a probability measure on $(E, \mathcal{E})$ and let $\Gamma = (\gamma_\Lambda)_{\Lambda \in \mathfrak{P}_0(\mathbb{L})}$ be a specification. Then there exists an absolutely summable interaction Φ such that $\gamma_\Lambda = \gamma_\Lambda^{\Phi, \lambda}$ (as in Definition 1.1.1) for all $\Lambda \in \mathfrak{P}_0(\mathbb{L})$ if and only if*

- (a) Γ is quasilocal.
- (b) For all $\Lambda \in \mathfrak{P}_0(\mathbb{L})$ there exist $\alpha_\Lambda, \beta_\Lambda \in (0, \infty)$ such that

$$\alpha_\Lambda \lambda^\mathbb{L}(A) \leq \gamma_\Lambda(\omega, A) \leq \beta_\Lambda \lambda^\mathbb{L}(A) \quad (\omega \in \Omega, A \in \mathcal{F}_\Lambda). \quad (1.5)$$

1.1.12 Remark. Property (b) of Theorem 1.1.11 is also called *uniformly nonnull* (see van Enter, Fernández and Sokal [32] by Fernández [38]). It implies that $\lambda^\mathbb{L} \ll \gamma_\Lambda(\omega, \cdot)$ and $\gamma_\Lambda(\omega, \cdot) \ll \lambda^\mathbb{L}$, which by the Radon-Nikodym theorem [10, Theorem 3.2.2] means that there exists a $\lambda^\mathbb{L}$ -integrable $h : \Omega \rightarrow (0, \infty)$ such that $\gamma_\Lambda(\omega, \cdot) = h \lambda^\mathbb{L}$.

1.1.13 Remark (The meaning of a quasilocal specification).

Whenever μ is a Gibbs measure with specification $\Gamma = (\gamma_\Lambda)_{\Lambda \in \mathfrak{P}_0(\mathbb{L})}$, the following holds: For all $\varepsilon > 0$, $\Delta \in \mathfrak{P}_0(\mathbb{L})$ and $A \in \mathcal{F}_\Delta$ there exists a $\Lambda \in \mathfrak{P}_0(\mathbb{L})$ with $\Delta \subset \Lambda$ such that, for all $\eta, \omega \in E^{\Delta^c}$ with $\eta|_\Lambda = \omega|_\Lambda$ (see Figure 1.6),

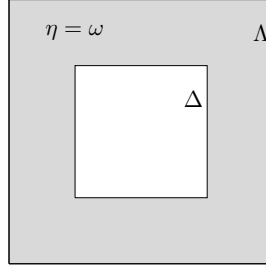
$$|\gamma_\Delta(\eta, A) - \gamma_\Delta(\omega, A)| < \varepsilon.$$

Intuitively, this means that under μ the outcome of the components inside Δ depends only weakly on the outcome of the components that are far from Δ , where far can be understood in terms of distance, as in the example of the Ising model (Example 1.1.4).

1.1.14 Remark (Towards the definition of Gibbs for mean-field systems).

The quasilocality of the specification plays a central role in the motivation of the definition of Gibbsianness for mean-field sequences. Namely, with Γ as in Theorem 1.1.11, we have the following. Let $\omega \in \Omega$ and $\Lambda \in \mathfrak{P}_0(\mathbb{L})$. Then

$$\gamma_\Lambda(\omega_n, A) \rightarrow \gamma_\Lambda(\omega, A) \quad (A \in \mathcal{F}_\Lambda)$$

Figure 1.6: $\eta = \omega$ within Λ .

for all sequences $(\omega_n)_{n \in \mathbb{N}}$ that on each finite box Υ agree with ω for large n , i.e., for all $\Upsilon \in \mathfrak{P}_0(\mathbb{L})$ there exists an N such that $\omega_n|_{\Upsilon} = \omega|_{\Upsilon}$ for all $n \geq N$.

§1.1.5 Mean-field sequences and sequential Gibbsianness

In Remark 1.1.13 we saw that under a Gibbs measure the components interact weakly when they are far from each other. Components in the Ising model that are close to each other, in the sense that they are neighbours, strongly interact with each other, in the sense that if a neighbour of a component takes the value -1 then the component prefers to take the value -1 as well.

A different type of system is a mean-field system. Mean-field stands for “equal interaction between all the components”: among n components, every component interacts with every other component at strength $\frac{1}{n}$. Therefore an increase of the number of components corresponds to a decrease of the interaction strength. Because of this, a mean-field system is not described by a single probability measure, but rather by a sequence of probability measures, indexed by n .

More formally, a mean-field sequence is a sequence $(\mu_n)_{n \in \mathbb{N}}$ of probability measures, where μ_n describes the law of n components, with the property that μ_n is invariant under permutations of the components. A typical form of a mean-field sequence for spins taking values in a subset of $\mathbb{R}$ is a sequence of probability measures $(\mu_n)_{n \in \mathbb{N}}$ given by

$$\mu_n(A) = \frac{1}{Z_n} \int_A e^{-nV \circ m_n(x)} d\lambda^n(x) \quad (A \text{ measurable}), \quad (1.6)$$

where Z_n is the normalising constant, λ a probability measure, $V : \mathbb{R} \rightarrow [0, \infty)$ and $m_n(x) = \frac{1}{n} \sum_{i=1}^n x_i$, called the empirical magnetisation of x . This setting is considered in Chapter 2. (In the literature this is referred to as a level-1 description.) It is also possible to replace “ V ” in (1.6) by “ V_n ”. In a more general setting one considers the quantity $L_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$ instead of m_n , which is called the empirical distribution. This setting is considered in Chapter 3. (In the literature this is referred to as a level-2 description.)

The definition of Gibbsianness for mean-field sequences has been developed by le Ny and Külske [60], although its roots can already be found in Külske [59]. We give

the precise definition in Chapter 2 and Chapter 3. We assume that there exists a continuous regular conditional probability γ_n under μ_n of the first spin x_1 conditional on the other spins $x_2, \dots, x_n$:

$$\gamma_n((x_2, \dots, x_n), dx_1) = \mu_n(dx_1 | x_2, \dots, x_n).$$

The sequence $(\mu_n)_{n \in \mathbb{N}}$ is called sequentially Gibbs whenever for every α there exists a probability measure γ_α such that

$$\gamma_n((x_2, \dots, x_n), \cdot) \rightarrow \gamma_\alpha$$

for all possible choices of $x_2, \dots, x_n$ such that their magnetisation converges to α . The idea behind this is similar to what was described in Remark 1.1.14: the limit of the mean-field kernels is independent of the choice of the sequence $(x_2^n, \dots, x_n^n)$ for which the magnetisation converges to α , i.e., $m_{n-1}(x_2^n, \dots, x_n^n) \rightarrow \alpha$, while the limit of the specification kernels in Remark 1.1.14 is independent of the choice of the sequence $(\omega_n)_{n \in \mathbb{N}}$ such that for all Υ the sequence eventually equals ω on Υ , i.e., $\omega_n|_\Upsilon = \omega|_\Upsilon$ for large n .

§1.1.6 Gibbs-non-Gibbs transitions

Consider a Gibbsian system, either a measure (on a lattice) or mean-field system. When a transformation of the system is made, the main question of interest is the following.

Is the transformed system Gibbsian or not?

One of the first examples of the occurrence of a non-Gibbs measure as the result of a natural transformation applied to the configuration space can be found in Israel [53] and van Enter, Fernández and Sokal [32]. In this example, the law restricted to a subset of the lattice may not be a Gibbs measure, even though the law on the full lattice is. For other transformations, for example majority rules rather than decimation rules, it has also been shown that the transformed measure may cease to be Gibbs.

Considering a transformation by the evolution of certain stochastic dynamics over a certain time interval, we can also consider the following interesting questions.

- What is the first time at which there is a transition from Gibbs to non-Gibbs?
- If a transition from Gibbs to non-Gibbs occurs, then, at a later time, does the system become Gibbs again?
- Is there “short-time Gibbsianness” or can the transition occur at time 0, in the sense that for all $t > 0$ the evolved system is non-Gibbs?

The transformation of Gibbs measures by stochastic dynamics constitute yet another class of examples. In van Enter, Fernández, den Hollander and Redig [30] the Ising model on $\mathbb{Z}^2$ at low temperature is considered. When subjected to independent or

weakly dependent spin-flip dynamics (which can be interpreted as heating of the system), the evolved measure is no longer Gibbs at certain later times. Depending on the magnetic field h , recovery of the Gibbs property occurs ($h \neq 0$) or not ($h = 0$). Later work by the same authors [31] suggested a general approach to Gibbs-non-Gibbs transitions via large deviations. The conjecture is that the evolved measure being Gibbs corresponds to having a unique most likely history for each possible evolved state, where most likely history is interpreted in the sense of large deviations, namely, as a minimiser of a rate function. This paper can be seen as the start of a systematic study of Gibbs-non-Gibbs transitions via large deviations.

For reviews we refer to den Hollander [51], Fernández [38], van Enter, Redig and Verbitskiy [34], le Ny [69] and van Enter [29].

As the geometry of the configuration space causes the study of Gibbs-non-Gibbs transitions for measures to be difficult, we try to develop techniques to study these transitions in mean-field systems. In these systems geometry is not taken into account, which makes them easier to study. In many mean-field systems the Gibbsianness can be explicitly related to a minimising problem in finite dimensions.

§1.1.7 Main results in Part I

Part I consists of two chapters related to Gibbs-non-Gibbs transitions. Chapter 4 considers large deviations for conditional probabilities. The need for such large deviations is motivated by a main result of Chapter 3.

The results in Chapter 2

In Chapter 2 we present our study of mean-field sequences for $\mathbb{R}$ -valued spins. It introduces the notion of sequential Gibbs for $\mathbb{R}$ -valued spins in an analogous way as was done by Külske and le Ny [60]. We show that, whenever the potential V is positive and C^1 , the sequence of finite-volume mean-field Gibbs measures with potential V is sequentially Gibbs. Considering such sequences as the law of our initial system, we evolve the spins according to independent Brownian motions and show that the conjecture in van Enter, Fernández, den Hollander and Redig [31] on the relation between Gibbsianness and large deviations indeed holds: We prove that the evolved system is sequentially Gibbs if and only if the large deviation rate function of the magnetisation of the initial system conditional on the final magnetisation of the evolved system has a unique global minimiser for each final magnetisation. Due to the fact that we consider independent Brownian motions as stochastic dynamics, the rate function has a simple form. In particular, by studying its convexity properties we can derive a criterion in terms on the potential that implies uniqueness of global minimisers. Indeed, an explicit bound on the second difference quotient of the potential holds if and only if the rate function has a unique global minimiser for each final magnetisation. With the help of this criterion a full classification of possible scenarios for the evolution can be given, showing when the evolved sequence is or is not sequential Gibbs. There are four different scenarios: the evolved sequence is Gibbsian for all time; there is short-time Gibbsianness, but there is a transition from

Gibbs to non-Gibbs at some time $t > 0$, where at time t the sequence is Gibbsian; the third scenario differs from the second scenario in that at time t the sequence is not Gibbsian; the fourth scenario is the immediate loss of Gibbsianness. Our example is the first in the literature where a full classification is achieved, and where immediate loss of Gibbsianness occurs.

The results in Chapter 3

In Chapter 3 we present our study of mean-field sequences for spins with values in a general topological space $\mathcal{X}$. Whereas in Chapter 2 the notion of sequentially Gibbs could be described in terms of the magnetisation, when we consider independent diffusions rather than Brownian motions we are forced to work with the empirical distribution, which is done in Chapter 3 (in the literature this is referred to as a level-2 description, while the description via magnetisation is level-1). We show that this more general notion agrees with the notion of sequential Gibbs in terms of the magnetisation. Similarly as in Chapter 2, but now in terms of the empirical distribution, we derive sufficient conditions on the potential for a sequence of finite-volume mean-field Gibbs measures to be sequential Gibbs. After that we consider transformations via general kernels and provide formulas for the conditional probabilities that are needed later on. Whenever the large deviation principle holds under such independent transformations (of the initial empirical distribution conditional on the final empirical distribution), we can formulate a sufficient condition for Gibbsianness of the evolved sequence. Indeed, in line with the proposed conjecture the transformed sequence is sequentially Gibbs if the large deviation rate function of the initial empirical distribution conditional on the final empirical distribution of the evolved system has a unique global minimiser for each possible final empirical distribution. This implication motivates us to study when such large deviation principles hold. This is done in Chapter 4. The other direction seems more cumbersome and might not be as easy to establish in full generality, in the sense that one may need to use intrinsic properties of the transformation to prove that multiple global minimisers imply non-Gibbsianness.

The results in Chapter 4

In Chapter 4 we present our study of large deviations for conditional probabilities. We consider the notion of weakly continuous regular conditional probability kernels and study under which conditions these satisfy a large deviation principle. We apply the resulting theory to conditional probabilities under measures on a product of finite spaces. Let us call one space the initial space and the other space the final space. We derive the large deviation principle for the conditional probability of the empirical distribution on the initial space conditional on the empirical distribution on the final space. This example can be applied to the setting presented in Chapter 3 in the following sense. Considering an initial mean-field sequence on a finite state space and an independent transformation to another finite state space, we obtain a two-layer measure, i.e., a measure on the product space of the initial and the final space. The large deviation principle we obtain in Section 4.8 shows that the conditions of the

main theorem of Section 3.7 are satisfied: in this situation sequential Gibbsianness is implied by uniqueness of the global minimiser of the rate function.

Open problems

We close by listing a few as yet unanswered questions that arise from the research in this thesis. First, coming back to the application in Chapter 4, we would like to extend our large deviation principle to more general spaces than finite spaces. With the help of this extension we would like to apply the theory of Chapter 3 to settings like the evolution of independent diffusions (rather than Brownian motions) or even dependent diffusions (where the dependence is of mean-field type). For such evolutions we would also like to prove the reverse implication, namely, that multiple global minimisers of the rate function imply non-Gibbsianness. More in the spirit of the setting of Chapter 2, we would like to see whether the uniqueness of global minimisers of the rate function corresponds to regularity properties of the potential.

§1.2 Introduction to Part II

The integral is an integral part in many branches of mathematics. One naturally associates it with the problem of computing the area underneath the graph of a function, but it also occurs, for instance, in the theory of dynamical systems, is used to represent functions and operators in Fourier theory, and plays the role of an expectation in the theory of probability.

Various types of integrals have been developed. The best known integrals for real-valued functions are the Riemann and the Lebesgue integral. For functions with values in a Banach space there are the Bochner and the Pettis integral. Many more notions of integrability have been developed for functions with values in the real line, a Banach space or a topological vector space. However, as far as we know, there is no notion of integration for functions with values in a partially ordered vector space (not necessarily a σ -Dedekind complete Riesz space).

In Section 1.2.1 we sketch the ideas behind Riemann, Lebesgue, Bochner and Pettis integrability. In Section 1.2.2 we describe the main contents of Chapter 5. In Section 1.2.3 we give an example of an integrable function in the sense of the theory of Chapter 5. In Section 1.2.4 we describe the main contents of Chapter 6

§1.2.1 Informal description of Riemann, Lebesgue, Bochner and Pettis integrability

The basic idea behind the *Riemann integral* is approximation by simple functions. One way to describe Riemann integrability of a function f on an interval $[a, b]$ is first to define the integral of a simple function and then to call f integrable if the infimum of the integrals of the simple functions majorising f is equal to the supremum of the integrals of the simple functions minorising f .

In a second step - and this is relevant to our present purposes - there is the *improper Riemann integral*. A function f on an interval $[a, \infty)$ is improperly Riemann integrable if f is integrable over every interval $[a, b]$ and

$$\lim_{b \rightarrow \infty} \int_a^b f(x) \, dx$$

exists.

Different is the set-up for the *Lebesgue integral*. Starting from a measure space $(X, \mathcal{A}, \mu)$, typically one first defines the notion of $\mathcal{A}$ -measurability and the integral of an $\mathcal{A}$ -simple function. Then a positive measurable function on X is called integrable if the supremum of the integrals of all minorising simple functions is finite. The notion of integrability extends by linearity.

With the notion of the Lebesgue integral at hand one can define *Bochner integrability* for functions with values in a Banach space. Start with a measure space $(X, \mathcal{A}, \mu)$ and a Banach space E . Introduce the integral of an $\mathcal{A}$ -simple function $X \rightarrow E$. Call a function $X \rightarrow E$ (strongly) measurable if it is the pointwise limit of simple functions. A measurable function $f : X \rightarrow E$ is said to be Bochner integrable

if there exist simple functions $f_1, f_2, \dots$ such that

$$\lim_{n \rightarrow \infty} \int \|f(x) - f_n(x)\| \, dx = 0, \quad (1.7)$$

which is the case if and only if the function $x \mapsto \|f(x)\|$ is (Lebesgue) integrable.

The *Pettis integral* is another integral for functions with values in a Banach space. Pettis integration is different in its set-up from the (improper) Riemann, Lebesgue and Bochner integration, as it is not based on a class of simple functions. Instead, it relies on the dual Banach space. Given a measure space $(X, \mathcal{A}, \mu)$ and a Banach space E , a function $f : X \rightarrow E$ is Pettis integrable with integral $I \in E$ if for every $\alpha \in E'$ the composition $\alpha \circ f : X \rightarrow \mathbb{R}$ is μ -integrable with integral $\alpha(I)$.

§1.2.2 Main contents of Chapter 5

Many standard Banach spaces, like c , ℓ^1 , ℓ^∞ , $L^1(\mu)$ and $C_b(X)$, are partially ordered vector spaces. A vector space E is a partially ordered vector space if for $x, y, a \in E$ and $\lambda \geq 0$, $x \leq y$ implies both $x + a \leq y + a$ and $\lambda x \leq \lambda y$. Many “classical” spaces are partially ordered vector spaces but do not carry natural norms, such as $C(\mathbb{R})$ and, for a measure space $(X, \mathcal{A}, \mu)$, the space of μ -equivalence classes of $\mathcal{A}$ -measurable functions. In Chapter 5 (but also in Chapter 6) we develop a concept of integration for functions with values in a partially ordered vector space. In Section 1.2.3 we give an example of an integrable function with values in $C(\mathbb{R})$.

In Chapter 5, following the ideas of the Riemann, Lebesgue and Bochner integration theories, we start with a class of “integrable” functions Γ and describe two ways to extend Γ : the vertical and the lateral extensions. We consider the following situation: E and F are partially ordered vector spaces, X is a set, Γ is a linear subspace of F^X (the space of all functions $X \rightarrow F$) and $\varphi : \Gamma \rightarrow E$ is a linear order preserving map, the “integral”. The plan is to extend φ to a map $\varphi_* : \Gamma_* \rightarrow E$, where Γ_* is a subspace of F^X that contains Γ . See Figure 1.7 for the general picture one should have in mind and Figure 1.8 for the general picture of an extension. (In the context of the Riemann, Lebesgue, Bochner and Pettis integrals we have $E = F$.)

$$\begin{array}{c} \Gamma \subset F^X \\ \varphi \downarrow \\ E \end{array}$$

Figure 1.7: General situation.

$$\begin{array}{c} \Gamma \subset \Gamma_* \subset F^X \\ \varphi \downarrow \swarrow \varphi_* \\ E \end{array}$$

Figure 1.8: An extension.

The *vertical extension* follows the Riemann approach: We call a function $X \rightarrow F$ (i.e., $f \in F^X$) vertically integrable whenever there exist countable sets $\Lambda, \Upsilon \subset \Gamma$ with $\Lambda \leq f \leq \Upsilon$ and (with $\varphi(\Lambda) = \{\varphi(g) : g \in \Lambda\}$)

$$\sup \varphi(\Lambda) = \inf \varphi(\Upsilon) \quad (\text{in } E). \quad (1.8)$$

In this way we obtain a set Γ_V of functions and we have a natural extension $\varphi_V : \Gamma_V \rightarrow E$, where $\varphi_V(f)$ equals (1.8).

The idea of the *lateral extension* is somewhat similar to the one of the improper Riemann integral. Above, we described the integral of a function f on $[a, \infty)$ as

$$\lim_{b \rightarrow \infty} \int_a^b f(x) \, dx,$$

approximating the domain of integration rather than the values of the function. For positive f , an equivalent description of this integral is

$$\sum_{n=1}^{\infty} \int_{a_n}^{a_{n+1}} f(x) \, dx,$$

where $a = a_1 < a_2 < \dots$ and $a_n \rightarrow \infty$. Here we partition $[a, \infty)$ into subintervals on which f is properly Riemann integrable.

Led by the above, we make the lateral extension as follows. In addition to E, F, X, Γ, φ , we assume that there is given a collection $\mathcal{I}$ of subsets of X . A positive function $f : X \rightarrow f^+$ is called laterally integrable if there exists a partition $(A_n)_{n \in \mathbb{N}}$ of X by elements of $\mathcal{I}$ such that

- $f \mathbb{1}_{A_n} \in \Gamma$ for all n ($\mathbb{1}_A$ denotes the indicator function of A),
- $\sum_n \varphi(f \mathbb{1}_{A_n})$ exists in E , i.e., the set $\{\sum_{n=1}^N \varphi(f \mathbb{1}_{A_n}) : N \in \mathbb{N}\}$ has a supremum in E .

The latter supremum is then denoted by $\varphi_L(f)$. By definition, Γ_L is the linear span of the set of all laterally integrable functions $X \rightarrow F^+$. The functional φ_L extends to an order preserving linear $\varphi_L : \Gamma_L \rightarrow E$.

These extension methods can be combined. Under suitable circumstances one can make Γ_{LV}, Γ_{VL} , etc. (where $\Gamma_{LV} = (\Gamma_L)_V$). Then the question arises whether the inclusions

$$\Gamma_{LV L} \supset \Gamma_{LV} \supset \Gamma_L \supset \Gamma \subset \Gamma_V \subset \Gamma_{VL} \subset \Gamma_{VL V}$$

are proper and whether, for instance, $\Gamma_{LV} = \Gamma_{VL}$. Those and similar problems are treated in Chapter 5. Furthermore, we compare the integrals obtained by combinations of these extensions with the Bochner and the Pettis integral.

§1.2.3 An example of an integrable function with values in $C(\mathbb{R})$

We consider $C(\mathbb{R})$ -valued integrals for functions $\mathbb{R} \rightarrow C(\mathbb{R})$, i.e., $X = \mathbb{R}$, $E = F = C(\mathbb{R})$. For Γ we take the set of all simple functions $\mathbb{R} \rightarrow C(\mathbb{R})$. Here, a function is said to be simple if it is of the form $\sum_{n=1}^N a_n \mathbb{1}_{I_n}$, where $a_1, \dots, a_N \in C(\mathbb{R})$, $I_1, \dots, I_N$ are

bounded intervals and $\mathbb{1}_I$ is the indicator function of I . On Γ we have an “integral” φ given by:

$$\varphi\left(\sum_{n=1}^N a_n \mathbb{1}_{I_n}\right) = \sum_{n=1}^N \lambda(I_n) a_n,$$

where $\lambda(I)$ denotes the length of I .

Define $h \in C(\mathbb{R})$, $g \in C(\mathbb{R})$, and $f : \mathbb{R} \rightarrow C(\mathbb{R})$ by (see Figure 1.9)

$$\begin{aligned} h(x) &= x^2, \\ g(x) &= (1 - |x|) \vee 0, \\ f(x) &\text{ is the function } t \mapsto h(x)g(t - x). \end{aligned}$$

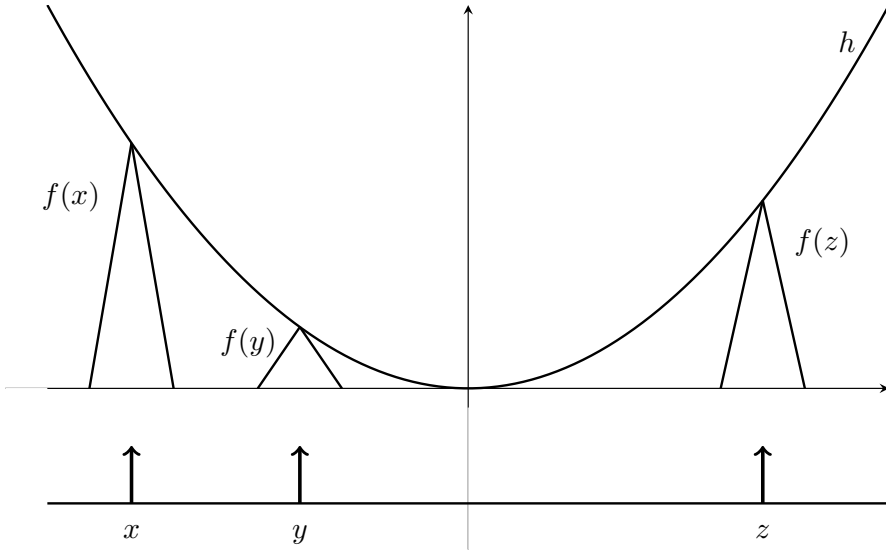


Figure 1.9: Some values of $f(x)$.

Omitting technical details, we will now show that $f \in \Gamma_{LV}$. For the $\mathcal{I}$ that makes the expression “ Γ_L ” meaningful we take the collection of all bounded intervals of $\mathbb{R}$. First, we construct functions $\sigma_1, \sigma_2, \dots$ in Γ_L that majorise f and $\tau_1, \tau_2, \dots$ that minorise f .

For an interval $[a, b)$, the set $\{f(x) : x \in [a, b)\}$ has a supremum and an infimum in $C(\mathbb{R})$. We denote these by

$$\sup_{[a,b)} f, \quad \inf_{[a,b)} f,$$

respectively. Both are continuous functions, vanishing outside $(a - 1, b + 1)$. See for example Figure 1.10, where $[a, b) = [7, 8)$, or Figure 1.11, where $[a, b) = [0, 1)$.

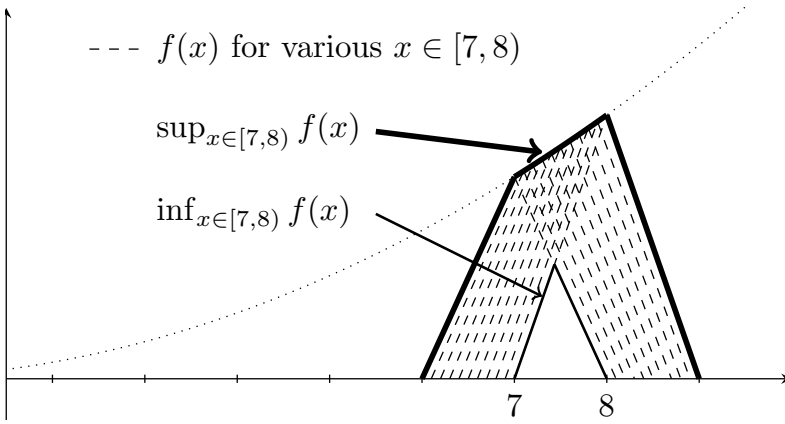


Figure 1.10: The supremum and infimum of f over $[7, 8)$.

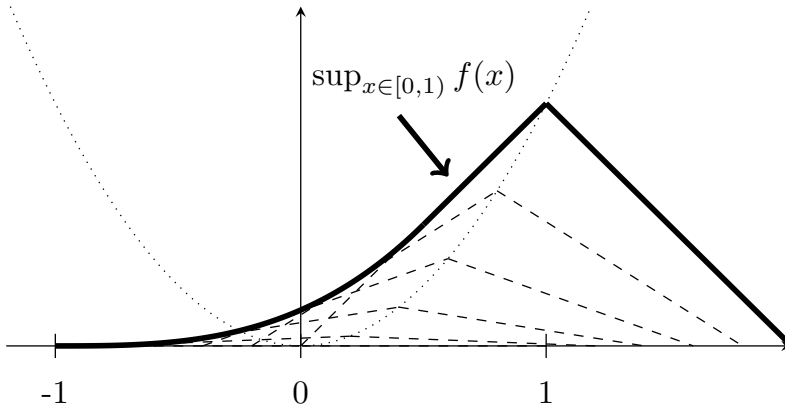


Figure 1.11: The supremum and infimum of f over $[0, 1)$.

Define $\sigma_1, \tau_1 : \mathbb{R} \rightarrow F = C(\mathbb{R})$ as follows. The intervals $[p, p+1)$ for $p \in \mathbb{Z}$ form a partition of $\mathbb{R}$ by sets in $\mathcal{I}$. For $p \in \mathbb{Z}$, put (see Figure 1.12)

$$\sigma_1(x) = \sup_{[p, p+1)} f, \quad \tau_1(x) = \inf_{[p, p+1)} f \quad \text{if } x \in [p, p+1).$$

Then

- $\sigma_1 \geq f \geq \tau_1$,
- For every p , $\sigma_1 \mathbb{1}_{[p,p+1]}$ is a simple function $\mathbb{R} \rightarrow C(\mathbb{R})$, and $\varphi(\sigma_1 \mathbb{1}_{[p,p+1]}) = \sigma_1(p)$.

Following the same pattern, one defines σ_2 and τ_2 in Γ_L using a partition of $\mathbb{R}$ by intervals of length $\frac{1}{2}$, σ_3 and τ_3 with intervals of length $\frac{1}{4}$, etc. (See Figure 1.12.) One obtains $\sigma_1 \geq \sigma_2 \geq \dots \geq f$ and $\tau_1 \leq \tau_2 \leq \dots \leq f$. It is not terribly difficult to show that $\varphi_L(\sigma_n) - \varphi_L(\tau_n) \downarrow 0$ in $C(\mathbb{R})$ (Figure 1.13 gives an impression). This entails

$$\sup_{n \in \mathbb{N}} \varphi_L(\tau_n) = \inf_{n \in \mathbb{N}} \varphi_L(\sigma_n),$$

and $f \in \Gamma_{LV}$. One can show that the function $\varphi_{LV}(f)$ in $C(\mathbb{R})$ is equal to

$$\begin{aligned} t &\mapsto \int_{-\infty}^{\infty} h(x)g(t-x) \, dx \\ &= \int_{-\infty}^{\infty} h(t-x)g(x) \, dx \\ &= \int_{-1}^1 (t-x)^2(1-|x|) \, dx \\ &= \frac{1}{6} + t^2. \end{aligned}$$

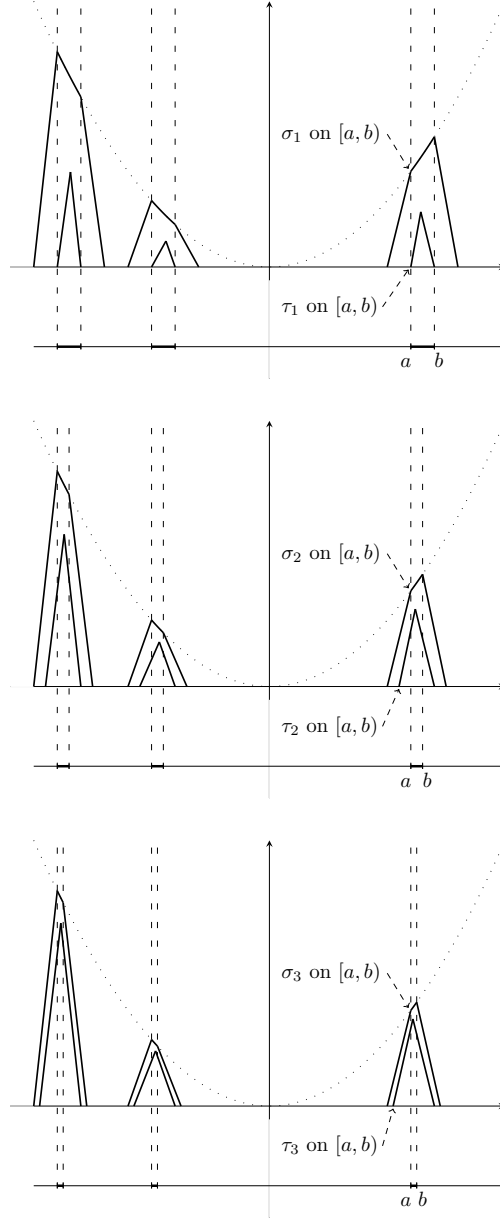


Figure 1.12: σ_n and τ_n for $n = 1, 2, 3$.

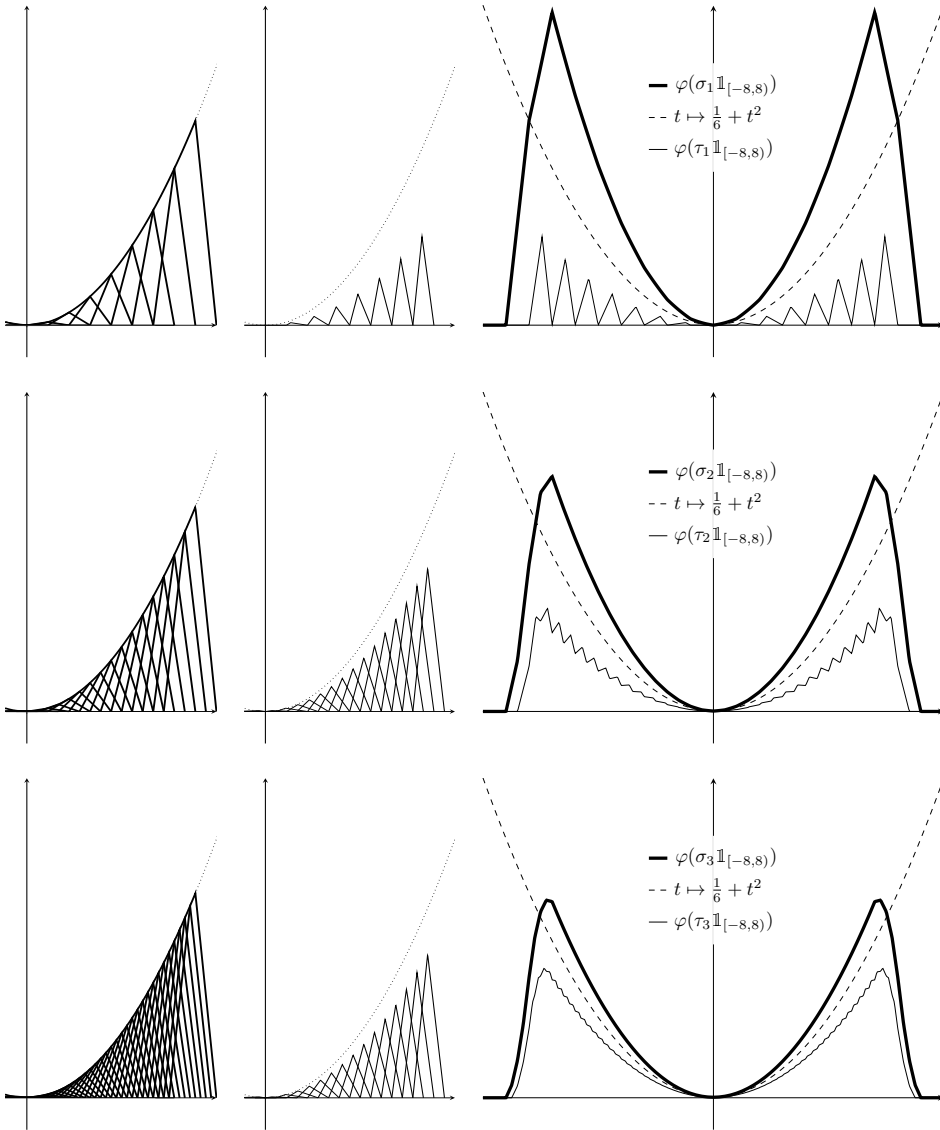


Figure 1.13: $\sigma_n(p)$ and $\tau_n(p)$ for different values of p and $\varphi(\sigma_i \mathbb{1}_{[-8,8]})$ and $\varphi(\tau_i \mathbb{1}_{[-8,8]})$: $n = 1, n = 2$ and $n = 3$.

§1.2.4 Main contents of Chapter 6

In Chapter 6 of Part II we consider Bochner integrals for functions with values in a partially ordered vector space. Whenever a vector space E can be covered by Banach spaces in a nice way, we can define an integral for functions with values in E by saying that a function is integrable whenever its values lie in one of the Banach spaces in the cover and it is Bochner integrable as a function with values in that Banach space. We show that Archimedean ordered vector spaces possess such a Banach cover, consisting of ordered Banach spaces with closed and generating cones. We study the Bochner integral for functions with values in an ordered Banach space with a closed and generating cone, and subsequently we study the integral for functions with values in an Archimedean ordered vector space. Furthermore, we compare this integral with the integrals of Chapter 5.

PART I

GIBBS-NON-GIBBS TRANSITIONS OF MEAN-FIELD SYSTEMS



Gibbs-non-Gibbs dynamical transitions for mean-field interacting Brownian motions

This chapter is based on:

F. den Hollander, R. Redig, and W. van Zuijlen. Gibbs-non-Gibbs dynamical transitions for mean-field interacting Brownian motions. *Stochastic Processes and their Applications*, 125(1):371–400, 2015.

Abstract

We consider a system of real-valued spins interacting with each other through a mean-field Hamiltonian that depends on the empirical magnetisation of the spins. The system is subjected to a stochastic dynamics where the spins perform independent Brownian motions. Using large deviation theory we show that there exists an explicitly computable crossover time $t_c \in [0, \infty]$ from Gibbs to non-Gibbs. We give examples of immediate loss of Gibbsianness ($t_c = 0$), short-time conservation and large-time loss of Gibbsianness ($t_c \in (0, \infty)$), and preservation of Gibbsianness ($t_c = \infty$). Depending on the potential, the system can be Gibbs or non-Gibbs at the crossover time $t = t_c$.

§2.1 Introduction and main results

§2.1.1 Background

Gibbs states are mathematical tools to describe physical interacting particle systems. In the lattice context, a Gibbs measure is a probability measure on the configuration space where the conditional distributions inside a finite subset of the lattice, given that the configuration outside this set is fixed, are described by a Gibbs specification, i.e., by a Boltzmann factor depending on an absolutely summable interaction potential (see Georgii [44, Definition 2.9]). When such systems evolve over time according to a stochastic dynamics, it may happen that the time-evolved state no longer is Gibbs. This phenomenon was originally discovered and described for heating dynamics by van Enter, Fernández, den Hollander and Redig [30]. In this paper, a low-temperature Ising model is subjected to a high-temperature Glauber spin-flip dynamics. The state remains Gibbs for short times, but becomes non-Gibbs after a finite time. If the magnetic field is zero, then Gibbsianness once lost is never recovered. But if the magnetic field is non-zero and small enough, then Gibbsianness is recovered at later times.

By now results of this type are available for a variety of interacting particle systems, both in the lattice setting and in the mean-field setting. Both for heating dynamics and for cooling dynamics estimates are available on transition times, as well as characterisations of the so-called *bad configurations* leading to non-Gibbsianness (i.e., the “points of essential discontinuity of the conditional probabilities”). It has become clear that Gibbs-non-Gibbs transitions are the rule rather than the exception. We refer the reader to the recent overview by van Enter [29].

In many papers non-Gibbsianness is proved by looking at the evolving system at two times, the initial time and the final time, and applying techniques from equilibrium statistical mechanics. This is a *static* approach that does not illuminate the relation between the Gibbs-non-Gibbs phenomenon and the dynamical effects responsible for its occurrence. This unsatisfactory situation was addressed in Enter, Fernández, den Hollander and Redig [31], where possible dynamical mechanisms were proposed and a *program* was put forward to develop a theory of Gibbs-non-Gibbs transitions in terms of *large deviations for trajectories of relevant physical quantities*.

Fernández, den Hollander and Martínez [39], [39], building on earlier work by Külske and Le Ny [60] and Ermolaev and Külske [36], showed that this program can be fully carried out for the Curie-Weiss model of Ising spins subjected to an infinite-temperature spin-flip dynamics, and also for a Kac-type version of the Curie-Weiss model. The present paper extends these works to systems of continuous spins that interact with each other through a *general* mean-field interaction potential and perform independent *Brownian motions*. The fact that we consider Brownian motions allows us to obtain a *complete characterisation* of passages from Gibbs to non-Gibbs. The key notions of interest are *good magnetisations* and *bad magnetisations* in the thermodynamic limit. Gibbsianness corresponds to having only good magnetisations, while non-Gibbsianness corresponds to having at least one bad magnetisation.

§2.1.2 Outline

The definition of Gibbs for mean-field models differs from that for lattice models because the interaction depends on the size of the system and does not have a geometric structure. In Section 2.1.3 we introduce the notions of a sequence of finite-volume mean-field Gibbs measures with a potential, good magnetisations, bad magnetisations and *sequentially Gibbs*, and show that a sequence of finite-volume mean-field Gibbs measures with a continuously differentiable potential is sequentially Gibbs. In Section 2.1.4 we define the Brownian motion dynamics. We show that a magnetisation $\alpha \in \mathbb{R}$ is bad at time t if and only if the large deviation rate function for the magnetisation at time 0 conditional on the magnetisation at time t being α has multiple global minimisers. We further show that the system is sequentially Gibbs at time t if and only if all magnetisations are good at time t . In Section 2.1.5 we show that a magnetisation α is bad at time t if and only if the large deviation rate function for the *trajectory* of the magnetisation conditional on hitting the value α at time t has multiple global minimisers. We further show that different minimising trajectories are different at time 0. In Section 2.1.6 we show that Gibbsianness can be classified in terms of the *second difference quotient of the potential*. With the help of this classification we show that there exists a unique time $t_c \in [0, \infty]$ at which the system changes from Gibbs to non-Gibbs, and give a characterisation of t_c in terms of the potential associated with the starting measures. In Section 2.1.7 we give examples for which $t_c = 0$, $t_c \in (0, \infty)$ and $t_c = \infty$. In Section 2.1.8 we discuss our results and indicate possible future research. Proofs are given in Sections 2.2–2.5. Appendix 2.A collects a few key formulas that are needed along the way. Appendix 2.B contains some background on proper weakly continuous regular conditional probabilities.

§2.1.3 Sequences of finite-volume mean-field Gibbs measures, Potential, Sequentially Gibbs

In this section we give the definition of a sequence of finite-volume mean-field Gibbs measures (Definition 2.1.1), and of good/bad magnetisations and sequentially Gibbs sequences (Definition 2.1.2). We show that a sequentially Gibbs sequence has a weakly continuous specification kernel (Lemma 2.1.3). We show that sequences of finite-volume mean-field Gibbs measures with a continuously differentiable potential are sequentially Gibbs (Theorem 2.1.4).

In what follows, we write $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_{\geq 2} = \mathbb{N} \setminus \{1\}$. For $n \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^n)$ denotes the Lebesgue measurable subsets of $\mathbb{R}^n$, and $\mu_{\mathcal{N}(v,A)}$ denotes the normal distribution on $\mathcal{B}(\mathbb{R}^n)$ with mean vector $v \in \mathbb{R}^n$ and covariance matrix $A \in \mathbb{R}^{n \times n}$. We write I_n for the identity matrix in $\mathbb{R}^{n \times n}$. For $\alpha \in \mathbb{R}$ and $\varepsilon > 0$, $B(\alpha, \varepsilon)$ denotes the open ball of radius ε centered at α .

2.1.1 Definition. For $n \in \mathbb{N}$, the *empirical magnetisation* $m_n: \mathbb{R}^n \rightarrow \mathbb{R}$ is given by

$$m_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i \quad ((x_1, \dots, x_n) \in \mathbb{R}^n). \quad (2.1)$$

For $n \in \mathbb{N}$, let ν_n be a probability measure on $\mathcal{B}(\mathbb{R}^n)$. Let $V: \mathbb{R} \rightarrow [0, \infty)$ be a Borel measurable function. The sequence $(\nu_n)_{n \in \mathbb{N}}$ is called a *sequence of finite-volume mean-field Gibbs measures* with *potential* V and *reference measures* $(\mu_{N(0, I_n)})_{n \in \mathbb{N}}$ when

$$\nu_n(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(x) e^{-n(V \circ m_n)(x)} d\mu_{N(0, I_n)}(x) \quad (A \in \mathcal{B}(\mathbb{R}^n), n \in \mathbb{N}), \quad (2.2)$$

where $Z_n \in (0, \infty)$ is the *normalising constant*.

Note that ν_n in (2.2) does not change when V is replaced by $V + c$ for some $c \in \mathbb{R}$. Therefore our assumption that $V \geq 0$ is equivalent to the assumption that V is bounded from below.

The model described in Definition 2.1.1 is an example of a mean-field model, where the Hamiltonian $(H_n(x) = n(V \circ m_n)(x))$ depends on the magnetisation $(m_n(x))$ only. In general the Hamiltonian of a mean-field model depends on the empirical mean (i.e., on “ $\frac{1}{n} \sum_{i=1}^n \delta_{x_i}$ ”), but we restrict ourselves to the models in Definition 2.1.1.

2.1.2 Definition. For $n \in \mathbb{N}$, let ρ_n be a probability measure on $\mathcal{B}(\mathbb{R}^n)$, and let $\pi_{(2:n)}: \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ be defined by

$$\pi_{(2:n)}(y_1, \dots, y_n) = (y_2, \dots, y_n) \quad ((y_1, \dots, y_n) \in \mathbb{R}^n). \quad (2.3)$$

Suppose that for every $n \in \mathbb{N}_{\geq 2}$ there exists a weakly continuous proper regular conditional probability $\gamma_n: \mathbb{R}^{n-1} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ under ρ_n of the first spin given the other spins, i.e., γ_n is the unique weakly continuous probability kernel for which for all $A \in \mathcal{B}(\mathbb{R})$, $B \in \mathcal{B}(\mathbb{R}^{n-1})$

$$\begin{aligned} \rho_n(A \times B) \\ = \int_{\mathbb{R}^{n-1}} \mathbb{1}_B(y_2, \dots, y_n) \gamma_n((y_2, \dots, y_n), A) d[\rho_n \circ \pi_{(2:n)}^{-1}](y_2, \dots, y_n). \end{aligned} \quad (2.4)$$

See Appendix 2.B for precise definitions and properties of these objects.

- (a) $\alpha \in \mathbb{R}$ is called a *good magnetisation* for the sequence $(\rho_n)_{n \in \mathbb{N}}$ when there exists a probability measure $\gamma_\alpha: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ for which the sequence of measures $(\gamma_n(v_{n-1}, \cdot))_{n \in \mathbb{N}_{\geq 2}}$ weakly converges to γ_α for all sequences $(v_{n-1})_{n \in \mathbb{N}_{\geq 2}}$ with $v_{n-1} \in \mathbb{R}^{n-1}$ for which the empirical magnetisation of v_n converges to α , i.e., $m_{n-1}(v_{n-1}) \rightarrow \alpha$.
- (b) $\alpha \in \mathbb{R}$ is called a *bad magnetisation* when it is not a good magnetisation.
- (c) The sequence $(\rho_n)_{n \in \mathbb{N}}$ is called *sequentially Gibbs* when all $\alpha \in \mathbb{R}$ are good magnetisations.

The notion of Gibbs for a mean-field model was introduced by Klske and Le Ny [60, Definition 2.1] (see also Klske [59]) and is the same as our definition of sequentially Gibbs (even though our definition of good magnetisation is slightly different).

The following lemma shows that, in the thermodynamic limit $n \rightarrow \infty$, the probability measure of the first spin given the magnetisation of the other spins is a transition kernel that depends weakly continuously on the magnetisation of the other spins. This lemma will be proved in Section 2.2.

2.1.3 Lemma. *Let $(\rho_n)_{n \in \mathbb{N}}$ be sequentially Gibbs. With the same notation as in Definition 2.1.2, define $\gamma: \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ by letting $\gamma(\alpha, \cdot) = \gamma_\alpha$. Then $\alpha \mapsto \gamma(\alpha, \cdot)$ is weakly continuous and, consequently, γ is a transition kernel (called the specification kernel).*

Our first main result, whose proof will be given in Section 2.3, shows that a sequence of finite-volume mean-field Gibbs measures with a continuously differentiable potential is sequentially Gibbs.

2.1.4 Theorem. *Let $(\nu_n)_{n \in \mathbb{N}}$ be a sequence of finite-volume mean-field Gibbs measures with continuous potential $V: \mathbb{R} \rightarrow [0, \infty)$.*

(a) *Define $\bar{\gamma}_n: \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ by*

$$\bar{\gamma}_n(\alpha, A) = \frac{\int_{\mathbb{R}} \mathbb{1}_A(x) e^{-nV(\frac{n-1}{n}\alpha + \frac{x}{n})} e^{-x^2/2} dx}{\int_{\mathbb{R}} e^{-nV(\frac{n-1}{n}\alpha + \frac{x}{n})} e^{-x^2/2} dx} \quad (\alpha \in \mathbb{R}, A \in \mathcal{B}(\mathbb{R})). \quad (2.5)$$

Then $\gamma_n: \mathbb{R}^{n-1} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ defined by $\gamma_n(v, A) = \bar{\gamma}_n(m_{n-1}(v), A)$ for $v \in \mathbb{R}^{n-1}$ and $A \in \mathcal{B}(\mathbb{R})$ is the weakly continuous proper conditional probability under ν_n of the first spin given the other spins.

(b) *If V is continuously differentiable on a neighbourhood of $\alpha \in \mathbb{R}$, then $\bar{\gamma}_n(\alpha_n, \cdot)$ converges weakly (even strongly) to $\mu_{\mathcal{N}(-V'(\alpha), 1)}$ for all sequences $(\alpha_n)_{n \in \mathbb{N}}$ that converge to α (in particular, α is a good magnetisation for $(\nu_n)_{n \in \mathbb{N}}$).*

(c) *If V is continuously differentiable, then $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs.*

In Section 2.1.7 we give an example of a non-differentiable potential for which $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures but not sequentially Gibbs (Example 2.1.17, where we write $\mu_{n,0}$ instead of ν_n).

§2.1.4 Brownian motion dynamics

In this section we introduce the Brownian motion dynamics, give the essential tools for identifying good magnetisations (Lemma 2.1.5) and global minimisers of a certain tilted form of the potential (Lemma 2.1.6), and show that a magnetisation is good if and only if the tilted potential has a unique global minimiser (Theorem 2.1.7).

For $n \in \mathbb{N}$, $\mu_{n,0}$ represents the law of the n spins at time $t = 0$. We assume that $(\mu_{n,0})_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures with continuous potential V . Let $\mu_{n,t}$ be the evolved law at time $t \in (0, \infty)$ when the n spins perform independent Brownian motions, i.e.,

$$\mu_{n,t}(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} p_n(t, z, A) e^{-n(V \circ m_n)(z)} d\mu_{\mathcal{N}(0, I_n)}(z) \quad (A \in \mathcal{B}(\mathbb{R}^n)) \quad (2.6)$$

(recall (2.2)), where

$$p_n(t, z, A) = \mu_{\mathcal{N}(z, tI_n)}(A) = (2\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \mathbb{1}_A(y) e^{-\frac{\|y-z\|^2}{2t}} dy \quad (z \in \mathbb{R}^n, A \in \mathcal{B}(\mathbb{R}^n)). \quad (2.7)$$

There exists a weakly continuous proper regular conditional probability $\gamma_{n,t}$ under $\mu_{n,t}$ of the first spin given the other spins for which $\gamma_{n,t}(u, \cdot) = \gamma_{n,t}(v, \cdot)$ for all $u, v \in \mathbb{R}^{n-1}$ with $m_{n-1}(u) = m_{n-1}(v)$ (a proof and an expression for $\gamma_{n,t}$ are given in Appendix 2.A). Therefore we can determine whether or not $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs by looking at the sequence $(\bar{\gamma}_{n,t})_{n \in \mathbb{N}}$ of probability kernels $\mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$, where $\bar{\gamma}_{n,t}(\alpha, \cdot) = \gamma_{n,t}(v, \cdot)$ for all $v \in \mathbb{R}^{n-1}$ and $\alpha \in \mathbb{R}$ with $m_{n-1}(v) = \alpha$ (an expression for $\bar{\gamma}_{n,t}$ is given in Appendix 2.A and also in (2.10)). This is formalized in the following lemma.

2.1.5 Lemma. *Let $t \in (0, \infty)$. Then $\alpha \in \mathbb{R}$ is a good magnetisation for $(\mu_{n,t})_{n \in \mathbb{N}}$ if and only if there exists a probability measure $\gamma_\alpha: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ such that the sequence $(\bar{\gamma}_{n,t}(\alpha_n, \cdot))_{n \in \mathbb{N}}$ converges weakly to γ_α for all sequences $(\alpha_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ that converge to α .*

The function $\eta_{n,t}: \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ defined for $n \in \mathbb{N}$ and $t \in (0, \infty)$ by

$$\eta_{n,t}(\alpha, A) = \frac{\int_{\mathbb{R}} \mathbb{1}_A(s) e^{-n \left[V(s) + \frac{s^2}{2} + \frac{(s-\alpha)^2}{2t} \right]} ds}{\int_{\mathbb{R}} e^{-n \left[V(s) + \frac{s^2}{2} + \frac{(s-\alpha)^2}{2t} \right]} ds} \quad (\alpha \in \mathbb{R}, A \in \mathcal{B}(\mathbb{R})). \quad (2.8)$$

is the weakly continuous proper regular conditional probability of the magnetisation at time 0 given the magnetisation at time t (see Appendix 2.A). By den Hollander [50, Theorem III.17], the sequence $(\eta_{n,t}(\alpha, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate n and rate function

$$r \mapsto V(r) + \frac{r^2}{2} + \frac{(r-\alpha)^2}{2t} - \inf_{s \in \mathbb{R}} \left[V(s) + \frac{s^2}{2} + \frac{(s-\alpha)^2}{2t} \right]. \quad (2.9)$$

(See Dembo and Zeitouni [24] or den Hollander [50] for background on large deviations.) With this notation, $\bar{\gamma}_{n,t}$ can be written as (see Appendix 2.A)

$$\bar{\gamma}_{n,t}(\alpha, B) = \frac{\int_{\mathbb{R}} \mu_{\mathcal{N}(s,t)}(B) g_{n,t}(\alpha, s) d\mu_{\mathcal{N}(0,1)}(s)}{\int_{\mathbb{R}} g_{n,t}(\alpha, s) d\mu_{\mathcal{N}(0,1)}(s)} \quad (\alpha \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})), \quad (2.10)$$

where $g_{n,t}: \mathbb{R}^2 \rightarrow \mathbb{R}$ is given by

$$g_{n,t}(\alpha, s) = \frac{\int_{\mathbb{R}} e^{-n \left[V(r + \frac{1}{n}(s-r)) - V(r) \right]} e^{-V(r)} d[\eta_{n-1,t}(\alpha, \cdot)](r)}{\int_{\mathbb{R}} e^{-V(r)} d[\eta_{n-1,t}(\alpha, \cdot)](r)} \quad (\alpha, s \in \mathbb{R}). \quad (2.11)$$

The following lemma will be proved in Section 2.4.

2.1.6 Lemma. *Let $V \in C^1(\mathbb{R}, [0, \infty))$, $t \in (0, \infty)$ and $\alpha \in \mathbb{R}$.*

- (a) *If (2.9) has a unique global minimiser $q \in \mathbb{R}$, then there exists a $\mu_{\mathcal{N}(0,1)}$ -integrable function $h: \mathbb{R} \rightarrow [0, \infty)$ such that*

$$\begin{aligned} g_{n,t}(\alpha_n, s) &\rightarrow e^{-(s-q)V'(q)} & (s \in \mathbb{R}), \\ g_{n,t}(\alpha_n, s) &\leq h(s) & (n \in \mathbb{N}, s \in \mathbb{R}), \end{aligned} \quad (2.12)$$

for all sequences $(\alpha_n)_{n \in \mathbb{N}}$ that converge to α .

- (b) Let q_1, q_2 be the smallest, respectively, the largest global minimiser of (2.9). Then there exists a $\mu_{\mathcal{N}(0,1)}$ -integrable function $h: \mathbb{R} \rightarrow [0, \infty)$, and sequences $(\alpha_n^1)_{n \in \mathbb{N}}$ and $(\alpha_n^2)_{n \in \mathbb{N}}$ both converging to α , for which (2.12) holds with $q = q_1$, $\alpha_n = \alpha_n^1$ and with $q = q_2$, $\alpha_n = \alpha_n^2$, respectively.

In case (2.9) has multiple minimisers, Lemma 2.1.6(b) implies that there are sequences $(\alpha_n^1)_{n \in \mathbb{N}}$ and $(\alpha_n^2)_{n \in \mathbb{N}}$ that in some sense “select” the smallest and the largest global minimiser of (2.9), respectively. In the proof of Lemma 2.1.6 we will see that this is the case for $\alpha_n^1 = \alpha - \frac{1}{\sqrt{n}}$ and $\alpha_n^2 = \alpha + \frac{1}{\sqrt{n}}$.

Our second main result shows that sequentially Gibbs is equivalent to uniqueness of the global minimiser of (2.9).

2.1.7 Theorem. *Let $V \in C^1(\mathbb{R}, [0, \infty))$. Then for every $t \in (0, \infty)$*

- (a) $\alpha \in \mathbb{R}$ is a good magnetisation for $(\mu_{n,t})_{n \in \mathbb{N}}$ if and only if (2.9) has a unique global minimiser.
- (b) If $\alpha \in \mathbb{R}$ is a good magnetisation for $(\mu_{n,t})_{n \in \mathbb{N}}$, then

$$\gamma_\alpha(B) = \mu_{\mathcal{N}(-V'(q), 1+t)}(B) \quad (B \in \mathcal{B}(\mathbb{R})), \quad (2.13)$$

where γ_α is the (limiting) probability measure as in Definition 2.1.2(a).

- (c) $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if (2.9) has a unique global minimiser for all $\alpha \in \mathbb{R}$.

The claim in Theorem 2.1.7(a) follows from Lemma 2.1.6, (2.10) and Lebesgue’s Dominated Convergence Theorem after we note that if $q_1, q_2 \in \mathbb{R}$ with $q_1 \neq q_2$ are global minimisers of (2.9), then $V'(q_1) - V'(q_2) = (q_2 - q_1)(1 + t^{-1}) \neq 0$. By Lemma 2.1.6(a),

$$\gamma_\alpha(B) = \frac{\int_{\mathbb{R}} \mu_{\mathcal{N}(s,t)}(B) e^{-sV'(q)} d\mu_{\mathcal{N}(0,1)}(s)}{\int_{\mathbb{R}} e^{-sV'(q)} d\mu_{\mathcal{N}(0,1)}(s)} \quad (B \in \mathcal{B}(\mathbb{R})), \quad (2.14)$$

from which it is easy to conclude Theorem 2.1.7(b). Theorem 2.1.7(c) is an immediate consequence of Theorem 2.1.7(a).

§2.1.5 Trajectories of the magnetisation (Intermezzo)

In this section we consider the probability measure on the set of trajectories of the magnetisation between time 0 and time t . We show the equivalence of uniqueness of the minimising magnetisation at time 0 and uniqueness of the minimising trajectory of the magnetisation (Theorem 2.1.8). This characterises good and bad magnetisations in terms of the trajectory of the magnetisation (Corollary 2.1.9).

By considering minimising trajectories instead of minimising initial points of the magnetisation, we obtain a better picture of the effects of the evolution. The name *two-layer model* has been used for a description of the minimisation problem for the magnetisation at time 0 given the magnetisation at time t . As Section 1.6 will confirm, the optimisation problem for the two-layer model is computationally easier.

However, in contrast with obtaining the function (2.9), obtaining the large deviation rate function for the two-layer model for more general dynamics, e.g., independent diffusion processes, might not be so easy and the rate function might not be given by an explicit formula like (2.9). For example, we took advantage of the fact that the transition kernel for the Brownian motion over time t is given explicitly. For more general diffusions this is not the case and we expect it to be necessary to consider the large deviation rate function for the trajectories (with the goal to obtain an implicit formula for the large deviation rate function for the two-layer model in terms of the more explicit large deviation rate function for the trajectories, by means of the contraction principle). We will show that for the case of independent Brownian motions the minimising problem for the two-layer model and the minimising problem for the trajectories are equivalent by showing that the minimising paths for the trajectories are fully determined by their initial point and endpoint.

Let μ_n be the law on $C([0, \infty), \mathbb{R}^n)$ of the paths of the independent Brownian motions performed by the n spins with initial distribution $\mu_{n,0}$. Thus, with $P(x, \cdot)$ denoting the law of the Brownian motion on $C([0, \infty), \mathbb{R})$ starting at $x \in \mathbb{R}$ and $\mathfrak{S}_{C([0, \infty), \mathbb{R}^n)}$ denoting the Skorohod σ -algebra on $C([0, \infty), \mathbb{R}^n)$, we have

$$\mu_n(A) = \int_{\mathbb{R}^n} \left(\bigotimes_{i=1}^n P(x_i, \cdot) \right) (A) \, d\mu_{n,0}(x_1, \dots, x_n) \quad (A \in \mathfrak{S}_{C([0, \infty), \mathbb{R}^n)}). \quad (2.15)$$

Let $t \in (0, \infty)$. Let $Q_{n,t}: \mathbb{R} \times \mathfrak{S}_{C([0,t], \mathbb{R})} \rightarrow [0, 1]$ be the transition kernel where $Q_{n,t}(s, \cdot)$ is the probability measure of a Brownian motion with variance $\frac{1}{n}$ starting at s . We write m_n also for the function $C([0, t], \mathbb{R}^n) \rightarrow C([0, t], \mathbb{R})$ given by

$$m_n(\phi_1, \dots, \phi_n) = \frac{1}{n} \sum_{i=1}^n \phi_i, \quad ((\phi_1, \dots, \phi_n) \in C([0, t], \mathbb{R}^n)). \quad (2.16)$$

Then $Q_{n,t}(s, A) = [\bigotimes_{i=1}^n P(x_i, \cdot)](\pi_{[0,t]}^{-1}(m_n^{-1}(A)))$ for all $A \in \mathfrak{S}_{C([0,t], \mathbb{R})}$ and all $s \in \mathbb{R}$ and $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ with $m_n(x) = s$, where $\pi_{[0,t]}: C([0, \infty), \mathbb{R}^n) \rightarrow C([0, t], \mathbb{R}^n)$ is given by $\pi_{[0,t]}(\phi) = \phi|_{[0,t]}$. We have

$$\mu_n \circ \pi_{[0,t]}^{-1}(m_n^{-1}(A)) = \int_{\mathbb{R}} Q_{n,t}(s, A) \, d[\mu_{n,0} \circ m_n^{-1}](s) \quad (A \in \mathfrak{S}_{C([0,t], \mathbb{R})}). \quad (2.17)$$

Let $\pi_t: C([0, t], \mathbb{R}) \rightarrow \mathbb{R}$ be the projection on the endpoint of the path, i.e., $\pi_t(\phi) = \phi(t)$.

2.1.8 Theorem. *Let $t \in (0, \infty)$.*

- (a) *For every $n \in \mathbb{N}$ there exists a weakly continuous proper regular conditional probability $\rho_n: \mathbb{R} \times \mathfrak{S}_{C([0,t], \mathbb{R})} \rightarrow [0, 1]$ under $\mu_n \circ \pi_{[0,t]}^{-1} \circ m_n^{-1}$ given π_t (given the endpoint of the trajectory).*
- (b) *For all $\alpha \in \mathbb{R}$, $(\rho_n(\alpha, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle (in $C([0, t], \mathbb{R})$ equipped with the uniform topology) with rate n and rate function $C([0, t], \mathbb{R}) \rightarrow$*

$[0, \infty]$ given by

$$\phi \mapsto \begin{cases} V(\phi(0)) + \frac{1}{2}\phi(0)^2 + \frac{1}{2} \int_0^t \dot{\phi}(s)^2 \, ds - C_{t,\alpha}, & \text{if } \phi \in \mathcal{AC}([0, t], \mathbb{R}) \\ & \text{and } \lim_{s \uparrow t} \phi(s) = \alpha, \\ \infty, & \text{otherwise,} \end{cases} \quad (2.18)$$

where $\mathcal{AC}([0, t], \mathbb{R})$ is the set of absolutely continuous functions from $[0, t]$ to $\mathbb{R}$ restricted to $[0, t)$, and $C_{t,\alpha} = \inf_{s_0 \in \mathbb{R}} V(s_0) + \frac{s_0^2}{2} + \frac{(\alpha - s_0)^2}{2t}$.

- (c) For every $\alpha \in \mathbb{R}$, (2.18) has a unique global minimiser if and only if (2.9) has a unique global minimiser.

Proof. The proof of (a) and (b) is given in Appendix 2.A. For (c) we only prove the ‘if’ implication. The function

$$\mathcal{L}^1([0, t], \mathbb{R}) \rightarrow [0, \infty], \quad g \mapsto \int_0^t g^2(s) \, ds \quad (2.19)$$

is strictly convex (on $\mathcal{L}^2([0, t], \mathbb{R})$), since $2ab < a^2 + b^2$ for $a, b \in \mathbb{R}$ with $a \neq b$. Hence, for all $r \in \mathbb{R}$, the path $\psi(s) = r + \frac{r-\alpha}{t}s$ for $s \in [0, t]$ is the unique path that minimises

$$\inf_{\phi \in \mathcal{AC}([0, t], \mathbb{R}), \phi(0)=r, \phi(t)=\alpha} \frac{1}{2} \int_0^t \dot{\phi}^2(s) \, ds. \quad (2.20)$$

In particular (2.20) equals $\frac{(r-\alpha)^2}{2t}$. Hence the infimum of (2.18) over all paths $\phi \in \mathcal{C}([0, t], \mathbb{R})$ with $\phi(0) = r$ is equal to (2.9). $\square$

As a consequence of Theorem 2.1.8, we can refine the result of Theorem 2.1.7.

2.1.9 Corollary. *Let $V \in C^1(\mathbb{R}, [0, \infty))$. Then for every $t \in (0, \infty)$:*

- (a) *For $\alpha \in \mathbb{R}$ the following are equivalent:*
 - (a1) $\alpha \in \mathbb{R}$ is a good magnetisation for $(\mu_{n,t})_{n \in \mathbb{N}}$,
 - (a2) (2.9) has a unique global minimiser,
 - (a3) (2.18) has a unique global minimiser.
- (b) *The following are equivalent:*
 - (b1) $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs,
 - (b2) (2.9) has a unique global minimiser for all $\alpha \in \mathbb{R}$,
 - (b3) (2.18) has a unique global minimiser for all $\alpha \in \mathbb{R}$.

§2.1.6 Uniqueness of the minimisers of the rate function

In this section we give a necessary and sufficient condition in terms of the second difference quotient of V (Definition 2.1.10) to have uniqueness of the global minimisers of (2.9) (Theorem 2.1.11 and Corollary 2.1.12). From this condition it follows that Gibbsianness can never be recovered once it is lost.

2.1.10 Definition. Let $f: \mathbb{R} \rightarrow \mathbb{R}$. The *second difference quotient* of f is the function

$$\begin{aligned} \Phi_2 f: \{(x, y, z) \in \mathbb{R}^3: x < y < z\} &\rightarrow \mathbb{R}, \\ (x, y, z) &\mapsto \frac{1}{z-x} \left(\frac{f(z) - f(y)}{z-y} - \frac{f(y) - f(x)}{y-x} \right). \end{aligned} \quad (2.21)$$

Our third main result, whose proof will be given in Section 2.5, is the following classification of Gibbsianness.

2.1.11 Theorem. Let $V: \mathbb{R} \rightarrow [0, \infty)$ be lower semicontinuous. Fix $t \in (0, \infty)$. There exists an $\alpha \in \mathbb{R}$ for which (2.9) has multiple global minimisers if and only if $\Phi_2 V \not\geq -\frac{1+t}{2t}$, i.e., if and only if there exist $a, b, c \in \mathbb{R}$ with $a < b < c$ for which $\Phi_2 V(a, b, c) \leq -\frac{1+t}{2t}$. Consequently, there exists a crossover time $t_c \in [0, \infty]$ such that $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for $t \in (0, t_c)$ and not sequentially Gibbs for $t \in (t_c, \infty)$.

At $t = t_c$, $(\mu_{n,t})_{n \in \mathbb{N}}$ may be sequentially Gibbs or not sequentially Gibbs. Both scenarios are possible (see Example 2.1.14). Theorem 2.1.7(c) together with Theorem 2.1.11 yield the following.

2.1.12 Corollary. Let $V \in C^1(\mathbb{R}, [0, \infty))$. Fix $t \in (0, \infty)$. For all $\alpha \in \mathbb{R}$, (2.9) has a unique global minimiser if and only if $\Phi_2 V > -\frac{1+t}{2t}$. Consequently, the following scenarios occur (where $M \in (\frac{1}{2}, \infty)$):

- (a) $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs
 - (a1) for $t \in (0, \infty)$ when $\Phi_2 V \geq -\frac{1}{2}$,
 - (a2) for $t \in (0, (M - \frac{1}{2})^{-1})$ when $\Phi_2 V \geq -M$,
 - (a3) for $t \in (0, (M - \frac{1}{2})^{-1}]$ when $\Phi_2 V > -M$.
- (b) $(\mu_{n,t})_{n \in \mathbb{N}}$ is not sequentially Gibbs
 - (b1) for $t \in ((M - \frac{1}{2})^{-1}, \infty)$ when $\Phi_2 V \not\geq -M$,
 - (b2) for $t \in [(M - \frac{1}{2})^{-1}, \infty)$ when $\Phi_2 V \not\geq -M$,
 - (b3) for $t \in (0, \infty)$ when $\Phi_2 V$ is not bounded from below.

Note that if V is convex, then $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for all $t \in (0, \infty)$. We will see at the end of Section 2.5 that if $V \in C^2(\mathbb{R}, [0, \infty))$, then (a1),(a2) and (b1),(b2) hold with $\Phi_2 V$ replaced by V'' .

§2.1.7 Examples

In this section we give examples of continuously differentiable potentials for each of the scenarios described in Corollary 2.1.12 (Examples 2.1.13–2.1.16).

2.1.13 Example. [Polynomial potentials: $t_c \in (0, \infty]$, sequentially Gibbs at $t = t_c$] Let $m \in \mathbb{N}$, $a_{2m} \in (0, \infty)$, $a_{2m-1}, \dots, a_2, a_1 \in \mathbb{R}$. Let $a_0 \in \mathbb{R}$ be such that

$$V: \mathbb{R} \rightarrow \mathbb{R}, \quad r \mapsto a_{2m} r^{2m} + a_{2m-1} r^{2m-1} + \dots + a_1 r + a_0 \quad (2.22)$$

satisfies $V \geq 0$. Since V'' is a polynomial of even degree, it is bounded from below, say $V'' \geq -M$ for some $M \in (0, \infty)$. Hence, if V is such a polynomial, then the crossover time t_c is strictly positive, i.e., $t_c \in (0, \infty]$. For example, for the potentials $V(r) = 0$, $V(r) = r^2$ and $V(r) = r^4 - \frac{1}{2}r^2 + 1$, $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for all $t \in [0, \infty)$, while for the potentials $V(r) = r^4 - 4r^2 + 3$ and $V(r) = (r^2 - 9)^2$ there exists a $t_c \in (0, \infty)$ for which $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for $t \in [0, t_c)$ and not sequentially Gibbs for $t \in (t_c, \infty)$.

If $m = 1$, then $\Phi_2 V = a_1 > 0$. Hence $t_c = \infty$ by Corollary 2.1.12.

If $m \geq 2$, then V'' is a polynomial of even degree at least 2. Hence, if $\beta = -\frac{1}{2} \inf_{r \in \mathbb{R}} V''(r)$, then the set $\{r \in \mathbb{R} : V''(r) = -2\beta\}$ is finite. By Lemmas 2.5.9–2.5.10, we therefore have that $\Phi_2 V > -\beta$ and $\Phi_2 V \not\geq -M$ for all $M < \beta$. So if $\beta \in (-\infty, \frac{1}{2}]$, then $t_c = \infty$, while if $\beta \in (\frac{1}{2}, \infty)$, then $t_c = (\beta - \frac{1}{2})^{-1}$ and $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for $t = t_c$ by Corollary 2.1.12.

2.1.14 Example. [Other potentials: $t_c \in (0, \infty]$, sequentially Gibbs at $t = t_c$] Consider the potential $V(r) = 2\beta(1 + \cos r)$ for some $\beta \in (0, \infty)$. Then $V'' \geq -2\beta$ and $V'' \not\geq -M$, and hence $\Phi_2 V \geq -\beta$ and $\Phi_2 V \not\geq -M$ for $M < \beta$ (see Lemma 2.5.9). So, for $\beta \in (0, \frac{1}{2}]$ we have $t_c = \infty$, while for $\beta \in (\frac{1}{2}, \infty)$ we have $t_c = (\beta - \frac{1}{2})^{-1}$ by Corollary 2.1.12. Moreover, if $\beta \in (\frac{1}{2}, \infty)$, then by Lemma 2.5.10 it follows that $\Phi_2 V > -\beta$, and hence $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for $t = t_c$.

In the previous two examples the sequence $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs at $t = t_c$. This is not always the case, as we show in Example 2.1.15 below.

2.1.15 Example. [Other potentials: $t_c \in (0, \infty]$, not sequentially Gibbs at $t = t_c$]

Let $g: \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$g(r) = \begin{cases} e^{-\frac{1}{|r|-1}+|r|-1} & r \in (-\infty, -1) \cup (1, \infty), \\ 0 & r \in [-1, 1]. \end{cases} \quad (2.23)$$

Because

$$\frac{d}{dr} e^{-\frac{1}{r}+r} = (1 + r^{-2})e^{-\frac{1}{r}+r} \quad (r \in (0, \infty)), \quad (2.24)$$

by L'Hôpital's rule $\lim_{r \downarrow 0} \frac{d}{dr} e^{-\frac{1}{r}+r} = 0 = \lim_{r \uparrow 0} \frac{d}{dr} 0$. Hence $g \in C^1(\mathbb{R}, [0, \infty))$. Furthermore

$$\frac{d^2}{dr^2} e^{-\frac{1}{r}+r} = r^{-4}(1 - 2r + 2r^2 + r^4)e^{-\frac{1}{r}+r} \geq 0 \quad (r \in (0, \infty)). \quad (2.25)$$

So g is a convex function with $\Phi_2 g \geq 0$ (see Lemma 2.5.4) and $\Phi_2 g|_{[-1,1]} = 0$. Hence $\Phi_2 g \not\geq 0$. Note also that $\lim_{r \rightarrow \infty} r^{-2}e^{-\frac{1}{r}+r} = \infty$ by L'Hôpital's rule. Therefore, for all $\beta \in (0, \infty)$,

$$\lim_{|r| \rightarrow \infty} g(r) - 2\beta r^2 = \infty. \quad (2.26)$$

Let $\beta \in (0, \infty)$ and consider $V \in C^1(\mathbb{R}, [0, \infty))$ given by

$$V(r) = g(r) - \beta r^2 - C_\beta \quad (r \in \mathbb{R}), \quad (2.27)$$

where $C_\beta = \inf_{s \in \mathbb{R}} g(s) - \beta s^2$ (which exists because of (2.26)). By Lemma 2.5.9, $\Phi_2 V \geq -\beta$ and $\Phi_2 V|_{[-1,1]} = -\beta$ and thus also $\Phi_2 V \not\geq -\beta$. So, for $\beta \in (0, \frac{1}{2}]$ we have $t_c = \infty$, while for $\beta \in (\frac{1}{2}, \infty)$ we have $t_c = (\beta - \frac{1}{2})^{-1}$ and $(\mu_{n,t})_{n \in \mathbb{N}}$ is not sequentially Gibbs for $t = t_c$ by Corollary 2.1.12.

2.1.16 Example. [Other potentials: $t_c = 0$]

Consider the potential $V(r) = 1 - \cos(r^2)$. Then

$$\begin{aligned} V'(r) &= 2r \sin(r^2) \quad (r \in \mathbb{R}), \\ V''(r) &= 2 [2r^2 \cos(r^2) - \sin(r^2)] \quad (r \in \mathbb{R}), \\ V''(\pm\sqrt{\pi k}) &= (-1)^k 4\pi k \quad (k \in \mathbb{N}). \end{aligned} \quad (2.28)$$

Hence $V'' \not\geq -M$ for all $M \in (0, \infty)$, and hence $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs for $t = 0$ but not for $t \in (0, \infty)$ (see Corollary 2.1.12).

We end with an example of a sequence of finite-volume mean-field Gibbs measures that is not sequentially Gibbs.

2.1.17 Example. [A sequence of finite-volume mean-field Gibbs measures that is not sequentially Gibbs]

Let $V \in C(\mathbb{R}, [0, \infty))$ be given by $V(r) = |r|$ for $r \in \mathbb{R}$. Then $(\mu_{n,0})_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures, but it is not sequentially Gibbs as we will show. Indeed, for all sequences $(\alpha_n)_{n \in \mathbb{N}}$ and all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} & \int_{\mathbb{R}} \mathbb{1}_A(r) e^{-nV(\frac{n-1}{n}\alpha_n + \frac{r}{n})} e^{-\frac{r^2}{2}} dr \\ &= e^{(n-1)\alpha_n} \int_{-\infty}^{-(n-1)\alpha_n} \mathbb{1}_A(r) e^{-\frac{r^2}{2} + r} dr \\ & \quad + e^{-(n-1)\alpha_n} \int_{-(n-1)\alpha_n}^{\infty} \mathbb{1}_A(r) e^{-\frac{r^2}{2} - r} dr. \end{aligned} \quad (2.29)$$

If $\alpha_n \geq 0$, then by substitution we get (using $\int_{-\infty}^0 e^{-\frac{1}{2}r^2} dr = \sqrt{\frac{\pi}{2}}$):

$$\begin{aligned} e^{(n-1)\alpha_n} \int_{-\infty}^{-(n-1)\alpha_n} e^{-\frac{r^2}{2} + r} dr &\leq \int_{-\infty}^0 e^{-\frac{(r-(n-1)\alpha_n)^2}{2} + r} dr \\ &\leq \sqrt{\frac{\pi}{2}} e^{-\frac{1}{2}((n-1)\alpha_n)^2}. \end{aligned} \quad (2.30)$$

Hence, if $\alpha_n = (n-1)^{-\frac{1}{2}}$ for $n \geq 2$, then $\alpha_n \downarrow 0$, $(n-1)\alpha_n \rightarrow \infty$ and, for $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} 0 &\leq \lim_{n \rightarrow \infty} e^{2(n-1)\alpha_n} \int_{-\infty}^{-(n-1)\alpha_n} \mathbb{1}_A(r) e^{-\frac{r^2}{2} + r} dr \\ &\leq \lim_{n \rightarrow \infty} \sqrt{\frac{\pi}{2}} e^{(n-1)\alpha_n(1 - \frac{1}{2}(n-1)\alpha_n)} = 0, \end{aligned} \quad (2.31)$$

and hence ($\bar{\gamma}_n$ as in (2.5))

$$\begin{aligned} & \lim_{n \rightarrow \infty} \bar{\gamma}_n(\alpha_n, A) \\ &= \lim_{n \rightarrow \infty} \frac{e^{2(n-1)\alpha_n} \int_{-\infty}^{-(n-1)\alpha_n} \mathbb{1}_A(r) e^{-\frac{r^2}{2}+r} dr + \int_{-(n-1)\alpha_n}^{\infty} \mathbb{1}_A(r) e^{-\frac{r^2}{2}-r} dr}{e^{2(n-1)\alpha_n} \int_{-\infty}^{-(n-1)\alpha_n} e^{-\frac{r^2}{2}+r} dr + \int_{-(n-1)\alpha_n}^{\infty} e^{-\frac{r^2}{2}-r} dr} \\ &= \mu_{\mathcal{N}(-1,1)}(A) \quad (A \in \mathcal{B}(\mathbb{R})). \end{aligned} \quad (2.32)$$

Similarly, if $\alpha_n = -(n-1)^{-\frac{1}{2}}$ for $n \geq 2$, then $\alpha_n \uparrow 0$, $(n-1)\alpha_n \rightarrow -\infty$ and $\bar{\gamma}_n(\alpha_n, A) \rightarrow \mu_{\mathcal{N}(1,1)}(A)$ for $A \in \mathcal{B}(\mathbb{R})$. From this we conclude that $(\mu_{n,0})_{n \in \mathbb{N}}$ is not sequentially Gibbs.

§2.1.8 Discussion

1. If V has a power series expansion $V(x) = \sum_{k \in \mathbb{N}} J_k x^k$, $x \in \mathbb{R}$, then

$$-n(V \circ m_n)(x_1, \dots, x_n) = -\sum_{k \in \mathbb{N}} \frac{J_k}{n^{k-1}} \sum_{i_1, \dots, i_k=1}^n \prod_{j=1}^k x_{i_j}, \quad (2.33)$$

i.e., the system with n spins has a mean-field k -spin interaction of strength J_k/n^{k-1} for $k \in \mathbb{N}$. The special case with $J_k \geq 0$ for all $k \in \mathbb{N}$ is called the ferromagnetic model.

2. Redig and Wang [77] analysed our model for a restricted class of potentials. Short-time Gibbsianness (i.e., the time-evolved state is Gibbs up to a strictly positive time) was proved under the condition that the second derivative of the potential exists and is bounded from below. Several scenarios of Gibbs-non-Gibbs transitions were discussed. Our paper considers a very general class of positive potentials and provides the precise connection between bifurcation of minimising trajectories and loss of Gibbsianness.

3. Our paper contains the first example of an initial Gibbs state and a stochastic dynamics for which there is immediate loss of Gibbsianness. For all the models that were considered in the literature so far, short-time Gibbsianness occurs. See e.g. [25], [30], [39], [62], [70].

4. In case the independent Brownian motions are replaced by independent Ornstein-Uhlenbeck processes, we get

$$r \mapsto V(r) + \frac{r^2}{2} + \frac{(r - e^t \alpha)^2}{e^{2t} - 1} - \inf_{s \in \mathbb{R}} V(s) + \frac{s^2}{2} + \frac{(s - e^t \alpha)^2}{e^{2t} - 1} \quad (2.34)$$

instead of (2.9) (cf. [77, Eq. (25)]), and so we obtain completely analogous results (in Corollary 2.1.12 the condition $\Phi_2 V > -\frac{1+t}{2t}$ is replaced by $\Phi_2 V > -(e^{2t} - 1)^{-1}$). In a forthcoming paper we will investigate what happens when the independent Brownian motions are replaced by independent diffusions.

5. For $n \in \mathbb{N}$ and $t > 0$, we can write $\mu_{n,t}$ as (compare with (2.2))

$$\mu_{n,t}(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(x) e^{-n(V_{n,t} \circ m_n)(x)} d\mu_{\mathcal{N}(0, (1+t)I_n)}(x) \quad (A \in \mathcal{B}(\mathbb{R}^n)) \quad (2.35)$$

with

$$V_{n,t}(r) = -\frac{1}{n} \log \left[\int_{\mathbb{R}} e^{-nV(s)} d\mu_{\mathcal{N}(\frac{r}{1+t}, \frac{t}{n(1+t)})}(s) \right] \quad (r \in \mathbb{R}) \quad (2.36)$$

(see (2.103) in Appendix 2.A). The sequence

$$\left(\mu_{\mathcal{N}(\frac{r}{1+t}, \frac{t}{n(1+t)})} \right)_{n \in \mathbb{N}} \quad (2.37)$$

satisfies the large deviation principle with rate n and rate function $s \mapsto \frac{1}{2}(s - \frac{r}{1+t})^2 \left(\frac{1+t}{t} \right)$. Therefore, by Varadhan's Lemma (see den Hollander [50, Theorem III.13]),

$$\lim_{n \rightarrow \infty} V_{n,t}(r) = \inf_{s \in \mathbb{R}} \left[V(s) + \frac{1}{2} \left(s - \frac{r}{1+t} \right)^2 \left(\frac{1+t}{t} \right) \right] \quad (r \in \mathbb{R}). \quad (2.38)$$

Note that, in the context of Definition 2.1.1, we are interested in the behaviour of μ_n for large n only. Therefore, looking back at Definition 2.1.1, we may generalise the notion of a sequence of finite-volume mean-field Gibbs measures with potential V , namely replacing V in (2.2) by a sequence of potentials $(V_n)_{n \in \mathbb{N}}$ that converges to V in an appropriate sense. Then $(\mu_{n,t})_{n \in \mathbb{N}}$ becomes a “generalised” sequence of finite-volume mean-field Gibbs measures with (limiting) potential $V_t(r) = \lim_{n \rightarrow \infty} V_{n,t}(r)$ as given in (2.38). It is then interesting to investigate how the regularity of V_t is related to the sequentially Gibbs property of the sequence $(\mu_{n,t})_{n \in \mathbb{N}}$ (compared to Theorem 2.1.4(c)).

§2.2 Proof of Lemma 2.1.3

2.2.1 Lemma. *Let $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ be metric spaces. Let $f_n: \mathcal{X} \rightarrow \mathcal{Y}$ for $n \in \mathbb{N}$ and suppose that there exists an $f: \mathcal{X} \rightarrow \mathcal{Y}$ such that, for all $x \in \mathcal{X}$ and for all sequences $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{X}$ with $x_n \rightarrow x$ we have $f_n(x_n) \rightarrow f(x)$. Then f is continuous.*

Proof. The proof is elementary. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{X}$ that converges to an element $x \in \mathcal{X}$. We first prove that $f_{k_n}(x_n) \rightarrow f(x)$ for all strictly increasing sequences $(k_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$. To that end, define the sequence $(y_m)_{m \in \mathbb{N}}$ in $\mathcal{X}$ by putting $y_m = x$ for $m \in \mathbb{N} \setminus \{k_n: n \in \mathbb{N}\}$ and $y_{k_n} = x_n$ for $n \in \mathbb{N}$. Then $y_m \rightarrow x$, hence $f_m(y_m) \rightarrow f(x)$, in particular, $f_{k_n}(x_n) = f_{k_n}(y_{k_n}) \rightarrow f(x)$. Since $f_k(x_n) \xrightarrow{k \rightarrow \infty} f(x_n)$ for all $n \in \mathbb{N}$, we can find a strictly increasing sequence $(k_n)_{n \in \mathbb{N}}$ for which $d_{\mathcal{Y}}(f_{k_n}(x_n), f(x_n)) < \frac{1}{n}$ for all $n \in \mathbb{N}$. Hence $d_{\mathcal{Y}}(f(x_n), f(x)) \leq d_{\mathcal{Y}}(f_{k_n}(x_n), f(x_n)) + d_{\mathcal{Y}}(f_{k_n}(x_n), f(x)) \rightarrow 0$. $\square$

Proof of Lemma 2.1.3. The proof of weak continuity of the map $\alpha \mapsto \gamma(\alpha, \cdot)$ is an adaptation of the proof of Lemma 2.2.1. Weak continuity of the map $\alpha \mapsto \gamma(\alpha, \cdot)$

implies continuity of the maps $\alpha \mapsto \int_{\mathbb{R}} f(x) d[\gamma(\alpha, \cdot)](x)$ for $f \in C_b(\mathbb{R})$. For open $A \in \mathcal{B}(\mathbb{R})$ there exists a sequence $(f_n)_{n \in \mathbb{N}}$ in $C_b(\mathbb{R})$ with $f_n \uparrow \mathbb{1}_A$ (point wise). It follows that $\int_{\mathbb{R}} f_n(x) d[\gamma(\alpha, \cdot)](x) \uparrow \gamma(\alpha, A)$ for all $\alpha \in \mathbb{R}$ and open A , and so $\alpha \mapsto \gamma(\alpha, A)$ is measurable for all $A \in \mathcal{B}(\mathbb{R})$ (since the open sets generate the Borel sigma-algebra). $\square$

§2.3 Proof of Theorem 2.1.4

Proof of Theorem 2.1.4. It is not hard to check that γ_n is a regular conditional probability under ρ_n of the first coordinate given the magnetisation of the other coordinates. To see that γ_n is proper and weakly continuous, we refer to Appendix 2.B. Let $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence that converges to α . Let $\delta > 0$ be such that V is continuously differentiable on $B(\alpha, 2\delta)$. Then, by the mean value theorem,

$$\lim_{n \rightarrow \infty} \mathbb{1}_{[-n\delta, n\delta]}(y) \mathbb{1}_A(y) e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} = \mathbb{1}_A(y) e^{-yV'(\alpha)} \quad (y \in \mathbb{R}, A \in \mathcal{B}(\mathbb{R})). \quad (2.39)$$

Let $N \in \mathbb{N}$ be such that $\alpha_n \in B(\alpha, \delta)$ for all $n \geq N$. Then, by the mean value theorem,

$$e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} \leq e^{\sup_{s \in B(\alpha, 2\delta)} |V'(s)| |y|} \quad (y \in [-n\delta, n\delta], n \geq N). \quad (2.40)$$

Since $y \mapsto e^{\sup_{s \in B(\alpha, 2\delta)} |V'(s)| |y|}$ is $\mu_{\mathcal{N}(0,1)}$ -integrable, Lebesgue's Dominated Convergence Theorem implies that for all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{[-n\delta, n\delta]} \mathbb{1}_A(y) e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} e^{-y^2/2} dy \\ &= \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-yV'(\alpha)} e^{-y^2/2} dy. \end{aligned} \quad (2.41)$$

Furthermore (because $n \leq e^n$, $n^2 = n(n-1) + n$ and $V \geq 0$)

$$\begin{aligned} & \int_{[-n\delta, n\delta]^c} e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} e^{-y^2/2} dy \\ &= n \int_{[-\delta, \delta]^c} e^{-n[V(\alpha_n + z) - V(\alpha_n)]} e^{-n^2 z^2/2} dz \\ &\leq n \int_{[-\delta, \delta]^c} e^{nV(\alpha_n)} e^{-n^2 z^2/2} dz \\ &\leq e^{-n[\frac{n-1}{2}\delta^2 - (V(\alpha_n) + 1)]} \int_{[-\delta, \delta]^c} e^{-nz^2/2} dz, \end{aligned} \quad (2.42)$$

where the last term converges to 0 as $n \rightarrow \infty$. So, by (2.41) – (2.42),

$$\int_{\mathbb{R}} \mathbb{1}_A(y) e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} e^{-y^2/2} dy \rightarrow \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-yV'(\alpha)} e^{-y^2/2} dy, \quad (2.43)$$

and hence, by (2.5), $\lim_{n \rightarrow \infty} \bar{\gamma}_n(\alpha_n, A) = \mu_{\mathcal{N}(-V'(\alpha), 1)}(A)$ for all $A \in \mathcal{B}(\mathbb{R})$, i.e., the sequence $(\bar{\gamma}_n(\alpha_n, \cdot))_{n \in \mathbb{N}}$ converges strongly (and hence weakly) to $\mu_{\mathcal{N}(-V'(\alpha), 1)}$. $\square$

§2.4 Proof of Lemma 2.1.6

Section 2.4.1 contains two preparatory lemmas (Lemmas 2.4.1–2.4.2) that provide estimates on $g_{n,t}$ in (2.11). These lemmas will be needed in Section 2.4.2 to give the proof.

§2.4.1 Two preparatory lemmas

Define $I_{t,\alpha}: \mathbb{R} \rightarrow [0, \infty)$ for $t \in (0, \infty)$ and $\alpha \in \mathbb{R}$ by

$$I_{t,\alpha}(r) = V(r) + \left(r - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t} \quad (r \in \mathbb{R}). \quad (2.44)$$

Note that $r \mapsto I_{t,\alpha}(r) - \inf_{s \in \mathbb{R}} I_{t,\alpha}(s)$ is equal to (2.9). Hence (see (2.8))

$$\eta_{n,t}(\alpha, A) = \frac{\int_{\mathbb{R}} \mathbb{1}_A(s) e^{-nI_{t,\alpha}(s)} ds}{\int_{\mathbb{R}} e^{-nI_{t,\alpha}(s)} ds} \quad (\alpha \in \mathbb{R}, A \in \mathcal{B}(\mathbb{R}), n \in \mathbb{N}, t \in (0, \infty)). \quad (2.45)$$

2.4.1 Lemma. *For every $t \in (0, \infty)$ there exists an $L > 0$ such that, for all $n \in \mathbb{N}_{\geq 2}$,*

$$g_{n,t}(\alpha, s) \leq L e^{-\frac{\alpha}{t}s + \frac{1}{4}s^2} G_t(n, \alpha) \quad (\alpha, s \in \mathbb{R}), \quad (2.46)$$

where $G_t: \mathbb{N} \times \mathbb{R} \rightarrow \mathbb{R}$ is given by

$$G_t(n, \alpha) = \frac{\int_{\mathbb{R}} e^{\left(\frac{1+t}{t}\right)^2 z^2} e^{-nV(z)} e^{-(n-1)\left(z - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dz}{\int_{\mathbb{R}} e^{-nV(r)} e^{-(n-1)\left(r - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dr} \quad (n \in \mathbb{N}, \alpha \in \mathbb{R}). \quad (2.47)$$

Consequently, if for a bounded sequence $(\alpha_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ the sequence $(G_t(n, \alpha_n))_{n \in \mathbb{N}}$ is bounded as well, then there exists a $\mu_{\mathcal{N}(0,1)}$ -integrable function $h: \mathbb{R} \rightarrow [0, \infty)$ for which, for all $n \in \mathbb{N}$,

$$g_{n,t}(\alpha_n, s) \leq h(s) \quad (s \in \mathbb{R}). \quad (2.48)$$

Proof. After some elementary computations (see (2.A.4) in Appendix 2.A), we may rewrite (2.11) as

$$g_{n,t}(\alpha, s) = \frac{\frac{n}{n-1} e^{-\frac{\alpha}{t}s} \int_{\mathbb{R}} e^{[-2z^2 + 2(s + \frac{\alpha}{1+t})z - \frac{1}{n-1}(z-s)^2]} e^{\frac{1+t}{2t}} e^{-nV(z)} e^{-(n-1)\left(z - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dz}{\int_{\mathbb{R}} e^{-nV(r)} e^{-(n-1)\left(r - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dr} \quad (2.49)$$

Since $-z^2 + 2\frac{\alpha}{1+t}z = -(z - \frac{\alpha}{1+t})^2 + (\frac{\alpha}{1+t})^2$ and $\frac{1+t}{t}sz \leq \frac{1}{4}s^2 + (\frac{1+t}{t})^2 z^2$, we get

$$g_{n,t}(\alpha, s) \leq 2e^{(\frac{\alpha}{1+t})^2} e^{-\frac{\alpha}{t}s} e^{\frac{1}{4}s^2} \frac{\int_{\mathbb{R}} e^{(\frac{1+t}{t})^2 z^2} e^{-nV(z)} e^{-(n-1)\left(z - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dz}{\int_{\mathbb{R}} e^{-nV(r)} e^{-(n-1)\left(r - \frac{\alpha}{1+t}\right)^2 \frac{1+t}{2t}} dr}, \quad (2.50)$$

which yields (2.46). The claim in (2.48) follows from (2.46) because $s \mapsto L e^{l|s| + \frac{1}{4}s^2}$ is $\mu_{\mathcal{N}(0,1)}$ -integrable for all $l \in \mathbb{R}$. $\square$

2.4.2 Lemma. Let $V \in C^1(\mathbb{R}, [0, \infty))$ and $t \in (0, \infty)$. For all $q, s, \alpha \in \mathbb{R}$, all sequences $(\alpha_n)_{n \in \mathbb{N}}$ with $\alpha_n \rightarrow \alpha$ and all $\varepsilon > 0$, there exist $\delta > 0$, $N \in \mathbb{N}$ and $M > 0$ such that for all $n \geq N$,

$$\begin{aligned} & \left| g_{n,t}(\alpha_n, s) - e^{-(s-q)V'(q)} \right| \vee \left| G_t(n, \alpha_n) - e^{\left(\frac{1+t}{t}\right)^2 q^2} \right| \\ & \leq \varepsilon + \\ & M \frac{\int_{B(q,\delta)^c} e^{2\left(\frac{1+t}{t}\right)^2 (r-\alpha_n)^2} e^{-(n-1)V(r) \wedge V\left(r+\frac{1}{n}(s-r)\right)} e^{-(n-1)\left(r-\frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t}} dr}{\int_{B(q,\delta)} e^{-V(r)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr}, \quad (2.51) \end{aligned}$$

where $G_t(n, \alpha)$ is as in (2.47).

Proof. Let $q, s, \alpha \in \mathbb{R}$, let $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}$ with $\alpha_n \rightarrow \alpha$, and let $\varepsilon > 0$. Let $\gamma > 0$ be such that $|e^y - 1| < \varepsilon e^{(s-q)V'(q)}$ for all $y \in (-\gamma, \gamma)$. Let $\delta \in (0, 1)$ be such that

$$\delta |V'(q)| < \frac{\gamma}{2}, \quad |V'(x) - V'(q)| < \frac{\gamma}{2(|s| + |q| + 1)} \quad (x \in B(q, 2\delta)), \quad (2.52)$$

$$\left| e^{\left(\frac{1+t}{t}\right)^2 r^2} - e^{\left(\frac{1+t}{t}\right)^2 q^2} \right| < \varepsilon \quad (r \in B(q, 2\delta)). \quad (2.53)$$

Let $N \in \mathbb{N}$ be such that $\frac{|s|+|q|+\delta}{N} < \delta$. Then $r + \frac{1}{n}(s-r) \in B(q, 2\delta)$ for all $r \in B(q, \delta)$ and all $n \geq N$. Let $r \in B(q, \delta)$ and $n \geq N$. By the Mean Value Theorem there exists an x in between r and $r + \frac{1}{n}(s-r)$, thus $x \in B(q, 2\delta)$, such that

$$\begin{aligned} \left| e^{-n[V(r+\frac{1}{n}(s-r))-V(r)]} e^{-(s-q)V'(q)} \right| & \leq \left| e^{-(s-r)V'(x)} - e^{-(s-q)V'(q)} \right| \\ & = e^{-(s-q)V'(q)} \left| e^{(s-q)V'(q)-(s-r)V'(x)} - 1 \right|. \quad (2.54) \end{aligned}$$

We show that the right-hand side of (2.54) is less than ε by showing $|(s-q)V'(q) - (s-r)V'(x)| < \gamma$:

$$\begin{aligned} |(s-q)V'(q) - (s-r)V'(x)| & \leq |s-r||V'(q) - V'(x)| + |q-r||V'(q)| \\ & \leq (|s| + |q| + 1)|V'(q) - V'(x)| + \delta |V'(q)| < \gamma, \quad (2.55) \end{aligned}$$

using (2.52). We obtain

$$\left| e^{-n[V(r+\frac{1}{n}(s-r))-V(r)]} - e^{-(s-q)V'(q)} \right| \leq \varepsilon \quad (n \geq N, r \in B(q, \delta)). \quad (2.56)$$

Hence

$$\begin{aligned}
 & \int_{\mathbb{R}} \left| e^{-n[V(r+\frac{1}{n}(s-r))-V(r)]} - e^{-(s-q)V'(q)} \right| \\
 & \quad \times e^{-V(r)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \leq \varepsilon \int_{B(q,\delta)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \quad + e^{-sV'(q)} \int_{B(q,\delta)^c} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \quad + \int_{B(q,\delta)^c} e^{-(n-1)[V(r+\frac{1}{n}(s-r))+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr
 \end{aligned} \tag{2.57}$$

and by (2.53)

$$\begin{aligned}
 & \int_{\mathbb{R}} \left| e^{(\frac{1+t}{t})^2 r^2} - e^{(\frac{1+t}{t})^2 q^2} \right| e^{-V(r)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \leq \varepsilon \int_{B(q,\delta)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \quad + e^{(\frac{1+t}{t})^2 q^2} \int_{B(q,\delta)^c} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr \\
 & \quad + \int_{B(q,\delta)^c} e^{(\frac{1+t}{t})^2 r^2} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr.
 \end{aligned} \tag{2.58}$$

Because $e^{-(n-1)V(r)} \vee e^{-(n-1)V(r+\frac{1}{n}(s-r))} \leq e^{-(n-1)[V(r) \wedge V(r+\frac{1}{n}(s-r))]}$, we obtain

$$\begin{aligned}
 & \left| g_{n,t}(\alpha_n, s) - e^{-sV'(q)} \right| \vee \left| G_t(n, \alpha_n) - e^{(\frac{1+t}{t})^2 q^2} \right| \\
 & \leq \varepsilon + K \frac{\int_{B(q,\delta)^c} e^{(\frac{1+t}{t})^2 r^2} e^{-(n-1)[V(r) \wedge V(r+\frac{1}{n}(s-r))+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr}{\int_{B(q,\delta)} e^{-V(r)} e^{-(n-1)[V(r)+(r-\frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t}]} dr}
 \end{aligned} \tag{2.59}$$

with $K = e^{-sV'(q)} + e^{(\frac{1+t}{t})^2 q^2} + 1$ (see (2.107) in Appendix 2.A). Because $r^2 \leq 2\left(r - \frac{\alpha_n}{1+t}\right)^2 + 2\left(\frac{\alpha_n}{1+t}\right)^2$ and $(\alpha_n)_{n \in \mathbb{N}}$ is bounded, we get (2.51). $\square$

§2.4.2 Proof of Lemma 2.1.6

In the proof we use the identity

$$\begin{aligned}
 & \left(r - \frac{\alpha_n}{1+t} \right)^2 \frac{1+t}{2t} \\
 & = \left(r - \frac{\alpha}{1+t} \right)^2 \frac{1+t}{2t} + \frac{1}{t} \left(r - \frac{\alpha}{1+t} \right) (\alpha - \alpha_n) + \frac{(\alpha - \alpha_n)^2}{2t(1+t)},
 \end{aligned} \tag{2.60}$$

which implies

$$I_{t,\alpha_n}(r) = I_{t,\alpha}(r) + \frac{1}{t} \left(r - \frac{\alpha}{1+t} \right) (\alpha - \alpha_n) + \frac{(\alpha - \alpha_n)^2}{2t(1+t)}. \tag{2.61}$$

Proof of Lemma 2.1.6. Let $s, q \in \mathbb{R}$ be the smallest global minimiser of (2.9), i.e.,

$$q = \inf \left\{ r \in \mathbb{R} : I_{t,\alpha}(r) = \inf_{s \in \mathbb{R}} I_{t,\alpha}(s) \right\}. \quad (2.62)$$

(A similar argument works for the largest global minimiser.) By Lemmas 2.4.1–2.4.2 it suffices to show that, for all $\delta > 0$,

$$\frac{\int_{B(q,\delta)^c} e^{2\left(\frac{1+t}{t}\right)^2 \left(r - \frac{\alpha_n}{1+t}\right)^2} e^{-(n-1)\left[V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t}\right]} dr}{\int_{B(q,\delta)} e^{-V(r)} e^{-(n-1)\left[V(r) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t}\right]} dr} \rightarrow 0. \quad (2.63)$$

For Part (b) we need to consider a particular sequence $(\alpha_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ converging to α , while for Part (a) we need to consider all sequences $(\alpha_n)_{n \in \mathbb{N}}$ converging to α . In both cases, for $\delta > 0$ we provide a sequence $(c_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ for which we check the following three steps, which together yield (2.63):

Step 1: Find $R > 0$, $C_1 > 0$ and $N_1 \in \mathbb{N}$ for which for all $n \geq N_1$

$$\int_{B(0,R)^c} e^{2\left(\frac{1+t}{t}\right)^2 (r - \alpha_n)^2} e^{-(n-1)\left[V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t} - c_n\right]} dr \leq C_1. \quad (2.64)$$

Step 2: Find $C_2 > 0$ and $N_2 \in \mathbb{N}$ for which for all $n \geq N_2$

$$\begin{aligned} \int_{B(q,\delta)^c \cap B(0,R)} e^{2\left(\frac{1+t}{t}\right)^2 (r - \alpha_n)^2} e^{-(n-1)\left[V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t} - c_n\right]} dr \\ \leq C_2. \end{aligned} \quad (2.65)$$

Step 3: Find $N_3 \in \mathbb{N}$ and a sequence $(\Gamma_n)_{n \in \mathbb{N}}$ with $\Gamma_n \rightarrow \infty$ for which for all $n \geq N_3$

$$\int_{B(q,\delta)} e^{-V(r)} e^{-(n-1)\left[V(r) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{(1+t)}{2t} - c_n\right]} dr \geq \Gamma_n. \quad (2.66)$$

Abbreviate

$$c = I_{t,\alpha}(q) = \inf_{r \in \mathbb{R}} I_{t,\alpha}(r) \in [0, \infty). \quad (2.67)$$

• **Step 1** for (a) and (b). For all bounded sequences $(\alpha_n)_{n \in \mathbb{N}}$ (in particular those that converge to α) there exists an $R > 0$ such that

$$\left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1}{2t} > c + 1 \quad (r \in B(0,R)^c, n \in \mathbb{N}). \quad (2.68)$$

Therefore, for all sequences $(c_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$ with $c_n \leq c + 1$ for all $n \in \mathbb{N}$,

$$\begin{aligned} V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} > c_n + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1}{2} \\ (r \in B(0,R)^c, n \in \mathbb{N}). \end{aligned} \quad (2.69)$$

Let $N_1 \in \mathbb{N}$ be such that $N_1 - 1 > 4\left(\frac{1+t}{t}\right)^2 + 1$. Then

$$\begin{aligned} & \int_{B(0,R)^c} e^{2\left(\frac{1+t}{t}\right)^2 \left(r - \frac{\alpha_n}{1+t}\right)^2} e^{-(n-1)\left[V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} - c_n\right]} \, dr \\ & \leq \int_{\mathbb{R}} e^{\left(4\left(\frac{1+t}{t}\right)^2 - (n-1)\right) \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1}{2}} \, dr \\ & \leq \int_{\mathbb{R}} e^{-\left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1}{2}} \, dr = \sqrt{2\pi} \quad (n \geq N_1). \end{aligned} \quad (2.70)$$

• **Step 2 for (a).** Because $\lim_{r \rightarrow \pm\infty} I_{t,\alpha}(r) = \infty$, $I_{t,\alpha}$ is continuous and $I_{t,\alpha}$ attains its global minimum at q , there exists a $\rho \in (0, \frac{1}{5})$ for which

$$I_{t,\alpha}(r) > c + 5\rho \quad (r \in B(q, \delta)^c). \quad (2.71)$$

Here, and in Step 3 for (a) below, we pick $c_n = c + 3\rho$ for $n \in \mathbb{N}$. Note that $c_n \leq c + 1$ for all $n \in \mathbb{N}$. By (2.60) and the continuity of V there exists an $N_2 \in \mathbb{N}$ such that, for all $n \geq N_2$,

$$\begin{aligned} V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} & > I_{t,\alpha}(r) - \rho > c + 4\rho \\ & (r \in B(q, \delta)^c \cap B(0, R)). \end{aligned} \quad (2.72)$$

Moreover, there exists an $\Upsilon > 0$ such that $e^{2\left(\frac{1+t}{t}\right)^2 \left(r - \frac{\alpha_n}{1+t}\right)^2} \leq \Upsilon$ for all $n \in \mathbb{N}$ and all $r \in B(0, R)$. Hence we obtain (2.65) for all $n \geq N_2$ with $C_2 = 2R\Upsilon$ (and $c_n = c + 3\rho$ for $n \in \mathbb{N}$).

• **Step 2 for (b).** Here, and in Step 3 for (b) below, for $n \in \mathbb{N}$, we consider $\alpha_n = \alpha - \frac{1}{\sqrt{n}}$, and

$$\begin{aligned} c_n &= I_{t,\alpha_n}(q) + \frac{\delta}{\sqrt{nt}} \\ &= I_{t,\alpha}(q) + \frac{1}{t} \left(q - \frac{\alpha}{1+t}\right) (\alpha - \alpha_n) + \frac{(\alpha - \alpha_n)^2}{2t(1+t)} + \frac{\delta}{\sqrt{nt}}. \end{aligned} \quad (2.73)$$

Note that $c_n \rightarrow c$, and so there exists an $N_1 \in \mathbb{N}$ for which $N_1 - 1 > 4\left(\frac{1+t}{t}\right)^2 + 1$ and $c_n \leq c + 1$ for $n \geq N_1$ (and thus (2.70) holds). For $r \in B(q, \delta)^c \cap B(0, R)$ we write

$$\begin{aligned} & V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} - c_n \\ &= \left(V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) - V(r)\right) + (I_{t,\alpha_n}(r) - c_n). \end{aligned} \quad (2.74)$$

For the left part (of the right hand side of (2.74)) we have

$$V(r) \wedge V\left(r + \frac{1}{n}(s-r)\right) - V(r) \geq -\frac{1}{n}\Theta \quad (2.75)$$

with $\Theta = (\sup_{u \in B(0, R+|s|)} |V'(u)|)(R + |s|)$. For the right part first note that, by the definition of q and the continuity of $I_{t,\alpha}$, there exists a $\rho > 0$ such that

$$I_{t,\alpha}(r) > I_{t,\alpha}(q) + \rho \quad (r \in (-\infty, q - \delta)). \quad (2.76)$$

Because $(\alpha - \alpha_n) \frac{1}{t} = \frac{1}{\sqrt{nt}}$, by (2.61) we have for the right part, for $r \in B(0, R)$,

$$\begin{aligned} I_{t,\alpha_n}(r) - c_n &= I_{t,\alpha}(r) - I_{t,\alpha}(q) + (r - q) \frac{1}{\sqrt{nt}} - \frac{\delta}{\sqrt{nt}} \\ &\geq \begin{cases} \rho - (R + \delta) \frac{1}{\sqrt{nt}} & r < q - \delta, \\ 0 & r > q + \delta. \end{cases} \end{aligned} \quad (2.77)$$

Let $N_2 \in \mathbb{N}$ be such that $(R + \delta) \frac{1}{\sqrt{nt}} < \rho$ for $n \geq N_2$. Then, for $r \in B(q, \delta)^c \cap B(0, R)$,

$$V(r) \wedge V\left(r + \frac{1}{n}(s - r)\right) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} - c_n \geq -\frac{1}{n}\Theta \quad (n \geq N_2). \quad (2.78)$$

Moreover, there exists a $\Upsilon > 0$ such that $e^{2(\frac{1+t}{t})^2(r - \frac{\alpha_n}{1+t})^2} \leq \Upsilon$ for all $n \in \mathbb{N}$ and all $r \in B(0, R)$. Therefore we obtain (2.65) for all $n \geq N_2$ with $C_2 = 2R\Upsilon e^\Theta$.

• **Step 3 for (a).** For $r \in A = B(q, \delta) \cap \{r \in \mathbb{R} : I_{t,\alpha}(r) < c + \rho\}$ there exists an $N_3 \in \mathbb{N}$ for which $I_{t,\alpha_n}(r) < c + 2\rho$ for all $n \geq N_3$. Hence

$$\begin{aligned} \int_{B(q, \delta)} e^{-V(r)} e^{-(n-1)[V(r) + (r - \frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t} - (c+3\rho)]} dr &\geq e^{(n-1)\rho} \int_A e^{-V(r)} dr \\ &\quad (n \geq N_3). \end{aligned} \quad (2.79)$$

• **Step 3 for (b).** There exists a $K > 0$ such that, for all $n \in \mathbb{N}$ and $r \in B(q, \frac{\delta}{n})$,

$$\begin{aligned} I_{t,\alpha_n}(r) - c_n &= I_{t,\alpha}(r) - I_{t,\alpha}(q) + (r - q) \frac{1}{\sqrt{nt}} - \frac{\delta}{\sqrt{nt}} \\ &< \frac{\delta}{n} \sup_{s \in B(q, \delta)} \left| \frac{d}{ds} I_{t,\alpha}(s) \right| + \frac{\delta}{tn\sqrt{n}} - \frac{\delta}{\sqrt{nt}} \\ &< \frac{1}{n}K - \frac{\delta}{\sqrt{nt}} = \frac{1}{\sqrt{n}} \left(\frac{K}{\sqrt{n}} - \frac{\delta}{t} \right). \end{aligned} \quad (2.80)$$

Let $N_3 \in \mathbb{N}$ be such that $\frac{K}{\sqrt{n}} < \frac{1}{2} \frac{\delta}{t}$ for $n \geq N_3$. Then, for $r \in B(q, \frac{\delta}{n})$,

$$V(r) + \left(r - \frac{\alpha_n}{1+t}\right)^2 \frac{1+t}{2t} - c_n < -\frac{1}{2} \frac{\delta}{t} \frac{1}{\sqrt{n}} \quad (n \geq N_3). \quad (2.81)$$

Let $\kappa > 0$ be such that $e^{-V(r)} > \kappa$ for all $r \in B(q, \delta)$. Then

$$\begin{aligned} \int_{B(q, \frac{\delta}{n})} e^{-V(r)} e^{-(n-1)[V(r) + (r - \frac{\alpha_n}{1+t})^2 \frac{(1+t)}{2t} - c_n]} dr \\ \geq \frac{2\delta}{n} \kappa e^{(\sqrt{n}-1)\frac{1}{2}\frac{\delta}{t}} \quad (n \geq N_3). \end{aligned} \quad (2.82)$$

□

§2.5 Tools from convex analysis: proof of Theorem 2.1.11

In this section we state a definition (Definition 2.5.1) and several lemmas (Lemmas 2.5.2–2.5.8) that are based on convex analysis, and use these to give the proof of Theorem 2.1.11. After that we prove the claim made below Corollary 2.1.12 (Lemma 2.5.9) and make an additional observation (Lemma 2.5.10) that can be used to determine whether $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs at $t = t_c$.

2.5.1 Definition. Let $f: \mathbb{R} \rightarrow \mathbb{R}$. Then $a \in \mathbb{R}$ is called a *supporting point* for f if there exists a linear function $l: \mathbb{R} \rightarrow \mathbb{R}$ with $l(a) = f(a)$ and $l(x) \leq f(x)$, $x \in \mathbb{R}$.

2.5.2 Lemma. Let $f: \mathbb{R} \rightarrow \mathbb{R}$. Then

(a) for $x, y, z \in \mathbb{R}$ with $x < y < z$:

$$\Phi_2 f(x, y, z) = \frac{f(x)}{(x-y)(x-z)} + \frac{f(y)}{(y-x)(y-z)} + \frac{f(z)}{(z-x)(z-y)}. \quad (2.83)$$

(b) for $a, b, c, d \in \mathbb{R}$ with $a < b < c < d$:

$$(d-a)\Phi_2 f(a, b, d) = (b-a)\Phi_2 f(a, b, c) + (d-c)\Phi_2 f(b, c, d), \quad (2.84)$$

$$(d-a)\Phi_2 f(a, c, d) = (c-a)\Phi_2 f(a, b, c) + (d-b)\Phi_2 f(b, c, d). \quad (2.85)$$

(c) for $g: \mathbb{R} \rightarrow \mathbb{R}$, $\theta, \kappa \in \mathbb{R}$:

$$\Phi_2(\theta f + \kappa g) = \theta \Phi_2 f + \kappa \Phi_2 g. \quad (2.86)$$

(d) for $g(x) = x^2$, $\Phi_2 g = 1$ and $\Phi_2 h = 0$ if $h(x) = \alpha x + \beta$ for $\alpha, \beta \in \mathbb{R}$.

Proof. The proof can be done by hand. See also Schikhof [84, Lemma 29.2]. $\square$

2.5.3 Lemma. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ and $y \in \mathbb{R}$. Then the following are equivalent:

(a) y is a supporting point for f ,

$$(b) \quad \frac{f(z) - f(y)}{z - y} \geq \frac{f(y) - f(x)}{y - x} \quad (x, z \in \mathbb{R}, x < y < z),$$

(c) $\Phi_2 f(\cdot, y, \cdot) \geq 0$.

Proof. Straightforward. $\square$

2.5.4 Lemma. A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is convex if and only if $\Phi_2 f \geq 0$. Moreover, f is strictly convex if and only if $\Phi_2 f > 0$.

Proof. See Schikhof and van Rooij [78, Theorem 2.2]. $\square$

2.5.5 Lemma. Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be lower semicontinuous with $\lim_{|x| \rightarrow \infty} f(x) = \infty$. Suppose that f is bounded from below. Then there exists an $a \in \mathbb{R}$ for which $f(a) = \inf_{x \in \mathbb{R}} f(x)$. In particular, a is a supporting point for f .

Proof. Let $c = \inf_{x \in \mathbb{R}} f(x)$. Define $A_n = \{x \in \mathbb{R} : f(x) \leq c + \frac{1}{n}\}$, $n \in \mathbb{N}$. Then A_n is compact and $A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$. Therefore there exists an $a \in \mathbb{R}$ for which $a \in \bigcap_{n \in \mathbb{N}} A_n$. $\square$

2.5.6 Lemma. Let $f: \mathbb{R} \rightarrow [0, \infty)$ be lower semicontinuous with $\lim_{|x| \rightarrow \infty} f(x) = \infty$. Then the following are equivalent:

- (a) There exists an $\alpha \in \mathbb{R}$ for which $x \mapsto f(x) - \alpha x$ has multiple global minimisers.
- (b) There exists a linear $l: \mathbb{R} \rightarrow \mathbb{R}$ for which $\#\{x \in \mathbb{R} : l(x) = f(x)\} \geq 2$ and $l \leq f$.
- (c) There exist $a, b, c \in \mathbb{R}$ with $a < b < c$ and $\Phi_2 f(\cdot, a, \cdot) \geq 0$, $\Phi_2 f(\cdot, c, \cdot) \geq 0$, $\Phi_2 f(\cdot, b, \cdot) \not\geq 0$.
- (d) There exist $a, x, b, y, c \in \mathbb{R}$ with $a \leq x < b < y \leq c$ and $\Phi_2 f(\cdot, a, \cdot) \geq 0$, $\Phi_2 f(\cdot, c, \cdot) \geq 0$, $\Phi_2 f(x, b, y) \leq 0$.

Proof. The equivalence (a) $\iff$ (b) and the implication (d) $\Rightarrow$ (c) are trivial.

(c) $\Rightarrow$ (d). Assume (c). Then there exist $x, y \in \mathbb{R}$ with $x < b < y$ for which $\Phi_2 f(x, b, y) \leq 0$. If $x < a$ and/or $y > c$, then $\Phi_2 f(a, b, y) \leq 0$ and/or $\Phi_2 f(x, b, c) \leq 0$ by Lemma 2.5.2(b). Therefore we may assume that $x \geq a$ and $y \leq c$, i.e., we obtain (d).

(b) $\Rightarrow$ (c). Assume (b). Let $a, c \in \{x \in \mathbb{R} : l(x) = f(x)\}$ with $a < c$. Let $b \in (a, c)$. Then

$$\frac{f(c) - f(a)}{c - a} = \frac{l(c) - l(a)}{c - a} = \frac{l(b) - l(a)}{b - a} \leq \frac{f(b) - f(a)}{b - a}, \quad (2.87)$$

i.e., $\Phi_2 f(a, b, c) \leq 0$.

(d) $\Rightarrow$ (b). Define $w, z \in \mathbb{R}$ by

$$\begin{aligned} w &= \sup\{s \leq b : \Phi_2 f(\cdot, s, \cdot) \geq 0\}, \\ z &= \inf\{s \geq b : \Phi_2 f(\cdot, s, \cdot) \geq 0\}. \end{aligned} \quad (2.88)$$

Because f is lower semicontinuous, we have $\liminf_{s \uparrow w} f(s) \geq f(w)$. Therefore, by Lemma 2.5.2(a), we have, for $q, r \in \mathbb{R}$ with $q < w < r$,

$$\begin{aligned} 0 &\leq \limsup_{s \uparrow w} \Phi_2 f(q, s, r) \\ &= \frac{f(q)}{(q - w)(q - r)} + \frac{f(r)}{(r - w)(r - q)} - \frac{\liminf_{s \uparrow w} f(s)}{(r - w)(w - q)} \leq \Phi_2 f(q, w, r). \end{aligned} \quad (2.89)$$

So $\Phi_2 f(\cdot, w, \cdot) \geq 0$. Similarly $\Phi_2 f(\cdot, z, \cdot) \geq 0$. If $w = b$, then $z = b$, and vice versa.

• Assume that $w = b = z$. Then f is convex and $\Phi_2 f(x, b, y) = 0$. With $l: \mathbb{R} \rightarrow \mathbb{R}$, $s \mapsto f(x) + \frac{f(y) - f(x)}{y - x}(s - x)$ one then has $l \leq f$ and $l(s) = f(s)$ for all $s \in [x, y]$, since

$$\frac{f(b) - f(x)}{b - x} \leq \frac{f(s) - f(b)}{s - b} \leq \frac{f(y) - f(b)}{y - b} = \frac{f(b) - f(x)}{b - x}. \quad (2.90)$$

• Assume that $w < b < z$. Define $l: \mathbb{R} \rightarrow \mathbb{R}$, $s \mapsto f(w) + \frac{f(z) - f(w)}{z - w}(s - w)$. Then $l \leq f$ on $(w, z)^c$. Note that $f - l|_{[w, z]}$ is lower semicontinuous and bounded from below. By

Lemma 2.5.5, it attains its infimum at some $a \in [w, z]$. This a is a supporting point of f , and hence $a = w$ or $a = z$ by Lemma 2.5.3. Thus $l(s) \leq f(s)$ for all $s \in \mathbb{R}$. $\square$

2.5.7 Lemma. *Let $f: \mathbb{R} \rightarrow [0, \infty)$ be lower semicontinuous. Let $r \in \mathbb{R}$ and $\beta > 0$. Then there exist $q, s \in \mathbb{R}$ with $q < r < s$ that are supporting points of $x \mapsto f(x) + \beta x^2$, i.e., $\Phi_2 f(\cdot, q, \cdot) \geq -\beta$, $\Phi_2 f(\cdot, s, \cdot) \geq -\beta$.*

Proof. Since $x \mapsto f(x) + \beta x^2$ is lower semicontinuous and $\lim_{|x| \rightarrow \infty} [f(x) + \beta x^2] = \infty$, by Lemma 2.5.5 there exists an $a \in \mathbb{R}$ for which a is a global minimum and thus a supporting point for $x \mapsto f(x) + \beta x^2$. There exists a (large enough) $\theta > 0$ such that

$$\{x \in \mathbb{R}: f(a) - 1 + \theta(x - r) = \beta x^2\} \quad (2.91)$$

has two elements, say x_1, x_2 with $x_1 < x_2$. By the definition of a , we have $x_1 > r$. By Lemma 2.5.5, there exists an $s \in \mathbb{R}$ that is a global minimum and a supporting point of

$$x \mapsto f(x) + \beta x^2 - (f(a) - 1 + \theta(x - r)). \quad (2.92)$$

Hence s is also a supporting point of $x \mapsto f(x) + \beta x^2$. Because (2.92) is strictly negative on (x_1, x_2) and non-negative on $[x_1, x_2]^c$, we have $s \in [x_1, x_2]$. Therefore $s > r$. There also exists a (small enough) $\theta < 0$ for which (2.91) has two elements. In the same way we can prove that there is an $q < r$ that is also a supporting point of $x \mapsto f(x) + \beta x^2$. The last part of the statement is a consequence of Lemma 2.5.2. $\square$

2.5.8 Lemma. *Let $f: \mathbb{R} \rightarrow [0, \infty)$ be lower semicontinuous and let $\beta \in (0, \infty)$. Then there exists an $\alpha \in \mathbb{R}$ for which $x \mapsto f(x) + \beta x^2 - \alpha x$ has multiple global minimisers if and only if $\Phi_2 f \not\geq -\beta$.*

Proof. This is a consequence of Lemmas 2.5.6–2.5.7. $\square$

Proof of Theorem 2.1.11. The claim in Theorem 2.1.11 follows by applying Lemma 2.5.8 with $\beta = \frac{1+t}{2t}$ to the lower semicontinuous function $r \mapsto V(r) + \frac{1}{2}r^2$. $\square$

The following observation proves the claim made below Corollary 2.1.12.

2.5.9 Lemma. *Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be twice differentiable. Then $f'' \geq 2\beta$ if and only if $\Phi_2 f \geq \beta$ for all $\beta \in \mathbb{R}$.*

Proof. By Lemma 2.5.4, $\Phi_2 g \geq 0$ if and only if g is convex. Since a twice differentiable function g is convex if and only if $g'' \geq 0$, this implies the equivalence $\Phi_2 f \geq 0 \iff f'' \geq 0$. Let $\beta \in \mathbb{R}$ and let $g: \mathbb{R} \rightarrow \mathbb{R}$ be given by $g(r) = f(r) - \beta r^2$. Then, by Lemma 2.5.2, we have $f'' \geq 2\beta \iff g'' \geq 0 \iff \Phi_2 g \geq 0 \iff \Phi_2 f \geq \beta$. $\square$

In contrast to Lemma 2.5.9, we can have $\Phi_2 f > \beta$ but not $f'' > 2\beta$ (take e.g. $\beta = 0$ and $f(x) = x^4$, in which case $\Phi_2 f > 0$ by Lemma 2.5.4 but $f''(0) = 0$). However, according to the next observation the second derivative of f can be used to determine whether $\Phi_2 f > \beta$. This observation can be used to determine whether $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs at $t = t_c$.

2.5.10 Lemma. Let $f: \mathbb{R} \rightarrow \mathbb{R}$. Let $a, b, c \in \mathbb{R}$ with $a < b < c$, and $\beta \in \mathbb{R}$.

- (a) If $\Phi_2 f|_{(a,b)} > \beta$, $\Phi_2 f|_{(a,b]} \geq \beta$, $\Phi_2 f|_{(b,c)} > \beta$, $\Phi_2 f|_{[b,c)} \geq \beta$ and $\Phi_2 f|_{(a,c)}(\cdot, b, \cdot) \geq 0$, then $\Phi_2 f|_{(a,c)} > \beta$.
- (b) If f is upper semicontinuous and $\Phi_2 f|_{(a,b)} \geq \beta$, then $\Phi_2 f|_{[a,b]} \geq \beta$.
- (c) If f is twice differentiable on (a, b) and $f|_{(a,b)}'' > \beta$, then $\Phi_2 f|_{(a,b)} > \beta$.

Proof. Without loss of generality we may assume $b = 0$.

- (a) Let $x, y, z \in (a, c)$. If $x < y < 0 < z$ or $x < 0 < y < z$, then with Lemma 2.5.2(b) we easily get $\Phi_2 f(x, y, z) > \beta$. If $y = 0$, then $x < \frac{x}{2} < 0 < z$, and hence $\Phi_2(x, \frac{x}{2}, 0) > 0$. Again with Lemma 2.5.2(b), we get $\Phi_2(x, 0, z) > 0$.
- (b) If f is upper semicontinuous, then $\limsup_{s \uparrow b} f(s) \leq f(b)$ and $\limsup_{s \downarrow a} f(s) \leq f(a)$. Together with Lemma 2.5.2(a) this proves the second statement.
- (c) If $f|_{(a,b)}'' > 0$, then f is strictly convex, and with Lemma 2.5.4 this implies (c) in case $\beta = 0$. Replacing f by $g(r) = f(r) - \frac{\beta}{2}r^2$, we obtain (c) for $\beta \neq 0$ (see Lemma 2.5.2). $\square$

Appendix

§2.A Key formulas

In this appendix we derive a few formulas that were used in the main body of the paper.

2.A.1. We derive formulas for $\gamma_{n,t}$ and $\bar{\gamma}_{n,t}$ described in Section 1.4.

Inserting (2.7) into (2.6) we get, for $A \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned} \mu_{n,t}(A) &= \frac{1}{Z_n} \int_{\mathbb{R}^n} \left[(2\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \mathbb{1}_A(y) e^{-\frac{\|y-z\|^2}{2t}} dy \right] e^{-n(V \circ m_n)(z)} d\mu_{\mathcal{N}(0, I_n)}(z) \\ &= \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(y) \left[(2\pi)^{-n} t^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{\|y-z\|^2}{2t}} e^{-\frac{\|z\|^2}{2}} e^{-n(V \circ m_n)(z)} dz \right] dy. \end{aligned} \quad (2.93)$$

Since $\frac{\|y-z\|^2}{2t} + \frac{\|z\|^2}{2} = \frac{\|y\|^2}{2(1+t)} + \frac{\|\frac{y}{1+t} - z\|^2(1+t)}{2t}$ for $y, z \in \mathbb{R}^n$, we get, for $A \in \mathcal{B}(\mathbb{R}^n)$,

$$\begin{aligned} \mu_{n,t}(A) &= \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(y) \left[(2\pi)^{-n} t^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{\|y\|^2}{2(1+t)}} e^{-\frac{\|\frac{y}{1+t} - z\|^2(1+t)}{2t}} e^{-n(V \circ m_n)(z)} dz \right] dy. \end{aligned} \quad (2.94)$$

Then it is not hard to check that $\gamma_{n,t}: \mathbb{R}^{n-1} \times \mathcal{B}(\mathbb{R})$ defined for $y_2, \dots, y_n \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$ by

$$\begin{aligned} &\gamma_{n,t}((y_2, \dots, y_n), B) \\ &= \frac{\int_{\mathbb{R}} \mathbb{1}_B(x) e^{-\frac{x^2}{2(1+t)}} \int_{\mathbb{R}^n} e^{-n(V \circ m_n)(z)} d\mu_{\mathcal{N}(\frac{(x, y_2, \dots, y_n)}{1+t}, \frac{t}{1+t} I_n)}(z) dx}{\int_{\mathbb{R}} e^{-\frac{x^2}{2(1+t)}} \int_{\mathbb{R}^n} e^{-n(V \circ m_n)(z)} d\mu_{\mathcal{N}(\frac{(x, y_2, \dots, y_n)}{1+t}, \frac{t}{1+t} I_n)}(z) dx}. \end{aligned} \quad (2.95)$$

is the weakly continuous proper conditional probability under $\mu_{n,t}$ of the first spin given the other spins. Using the identities

$$\mu_{\mathcal{N}}\left(\frac{(x, y_2, \dots, y_n)}{1+t}, \frac{t}{1+t} I_n\right) = \mu_{\mathcal{N}}\left(\frac{x}{1+t}, \frac{t}{1+t}\right) \otimes \mu_{\mathcal{N}}\left(\frac{(y_2, \dots, y_n)}{1+t}, \frac{t}{1+t} I_{n-1}\right), \quad (2.96)$$

$$\mu_{\mathcal{N}}\left(\frac{(y_2, \dots, y_n)}{1+t}, \frac{t}{1+t} I_{n-1}\right) \circ m_{n-1}^{-1} = \mu_{\mathcal{N}}\left(\frac{m_{n-1}(y_2, \dots, y_n)}{1+t}, \frac{t}{(n-1)(1+t)}\right), \quad (2.97)$$

$$m_n(z_1, \dots, z_n) = \frac{z_1}{n} + \frac{n-1}{n} m_{n-1}(z_2, \dots, z_n), \quad (2.98)$$

we obtain the expression

$$\begin{aligned} & \gamma_{n,t}((y_2, \dots, y_n), B) \\ &= \frac{\int_{\mathbb{R}} \mathbb{1}_B(x) \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} d\mu_{\mathcal{N}}\left(\frac{m_{n-1}(y_2, \dots, y_n)}{1+t}, \frac{t}{(n-1)(1+t)}\right)(r) \\ & \quad d\mu_{\mathcal{N}}\left(\frac{x}{1+t}, \frac{t}{1+t}\right)(s) d\mu_{\mathcal{N}(0,1+t)}(x)}{\int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} d\mu_{\mathcal{N}}\left(\frac{m_{n-1}(y_2, \dots, y_n)}{1+t}, \frac{t}{(n-1)(1+t)}\right)(r) \\ & \quad d\mu_{\mathcal{N}}\left(\frac{x}{1+t}, \frac{t}{1+t}\right)(s) d\mu_{\mathcal{N}(0,1+t)}(x)} \end{aligned} \quad (2.99)$$

We see that $\gamma_{n,t}(u, \cdot) = \gamma_{n,t}(v, \cdot)$ for all $u, v \in \mathbb{R}^{n-1}$ with $m_{n-1}(v) = m_{n-1}(u)$. Hence we can define $\bar{\gamma}_{n,t}: \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ by letting $\bar{\gamma}_{n,t}(\alpha, B) = \gamma_{n,t}(v, B)$ for $\alpha \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$, where $v \in \mathbb{R}^{n-1}$ is such that $m_{n-1}(v) = \alpha$, i.e.,

$$\begin{aligned} & \bar{\gamma}_{n,t}(\alpha, B) \\ &= \frac{\int_{\mathbb{R}} \mathbb{1}_B(x) \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} d\mu_{\mathcal{N}}\left(\frac{\alpha}{1+t}, \frac{t}{(n-1)(1+t)}\right)(r) \\ & \quad d\mu_{\mathcal{N}}\left(\frac{x}{1+t}, \frac{t}{1+t}\right)(s) d\mu_{\mathcal{N}(0,1+t)}(x)}{\int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} d\mu_{\mathcal{N}}\left(\frac{\alpha}{1+t}, \frac{t}{(n-1)(1+t)}\right)(r) \\ & \quad d\mu_{\mathcal{N}}\left(\frac{x}{1+t}, \frac{t}{1+t}\right)(s) d\mu_{\mathcal{N}(0,1+t)}(x)} \end{aligned} \quad (2.100)$$

2.A.2. We show that $\eta_{n,t}$ is indeed the weakly continuous proper regular conditional probability of the magnetisation of the n spins at time 0 given the magnetisation at time t .

Let μ_n be the law on $C([0, \infty), \mathbb{R}^n)$ of the paths of the independent Brownian motions performed by the n spins with initial distribution $\mu_{n,0}$, i.e., μ_n is given by (2.15). The joint law of the process at time 0 and time t is given by

$$\mu_{n,(0,t)}(A) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(x, y) d[p_n(t, x, \cdot)](y) d\mu_{n,0}(x) \quad (A \in \mathcal{B}((\mathbb{R}^2)^n)). \quad (2.101)$$

We write m_n also for the function $(\mathbb{R}^2)^n \rightarrow \mathbb{R}^2$ given by

$$m_n((x_1, y_1), \dots, (x_n, y_n)) = \frac{1}{n} \sum_{i=1}^n (x_i, y_i) \quad (x_1, y_1, \dots, x_n, y_n \in \mathbb{R}). \quad (2.102)$$

Let $\bar{\mu}_{n,(0,t)} = \mu_{n,(0,t)} \circ m_n^{-1}$. Since $p_n(t, x, \cdot) \circ m_n^{-1} = \mu_{\mathcal{N}(x, tI_n)} \circ m_n^{-1} = \mu_{\mathcal{N}(m_n(x), \frac{t}{n})}$ and $\mu_{\mathcal{N}(0, I_n)} \circ m_n^{-1} = \mu_{\mathcal{N}(0, \frac{1}{n})}$, we have

$$\begin{aligned} \bar{\mu}_{n,(0,t)}(A) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(s, \alpha) \, d\mu_{\mathcal{N}(s, \frac{t}{n})}(\alpha) e^{-nV(s)} \, d\mu_{\mathcal{N}(0, \frac{1}{n})}(s) \\ &= \frac{1}{\sqrt{2\pi\frac{t}{n}}} \frac{1}{\sqrt{2\pi\frac{1}{n}}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(s, \alpha) e^{-n[V(s) + \frac{s^2}{2} + \frac{(s-\alpha)^2}{2t}]} \, ds \, d\alpha \\ &\quad (A \in \mathcal{B}(\mathbb{R}^2)). \end{aligned} \quad (2.103)$$

From this it follows that $\eta_{n,t}$ given in (2.8) is the weakly continuous proper regular conditional probability under $\bar{\mu}_{n,(0,t)}$ of the first coordinate given the second, i.e., the weakly continuous proper regular conditional probability of the magnetisation of the n spins at time 0 given the magnetisation at time t .

2.A.3. We verify (2.11) and (2.10).

An elementary computation gives that, for $\alpha, s \in \mathbb{R}$, $t \in (0, \infty)$ and $n \in \mathbb{N}$,

$$\begin{aligned} &\int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} \, d\mu_{\mathcal{N}(\frac{\alpha}{1+t}, \frac{t}{(n-1)1+t})}(r) \\ &= \sqrt{\frac{(n-1)(1+t)}{2\pi t}} \int_{\mathbb{R}} e^{-nV(\frac{1}{n}s + \frac{n-1}{n}r)} e^{-(r - \frac{\alpha}{1+t})^2 \frac{(n-1)(1+t)}{2t}} \, dr \\ &= \sqrt{\frac{(n-1)(1+t)}{2\pi t}} e^{-(n-1)\frac{\alpha^2}{1+t}} \\ &\quad \times \int_{\mathbb{R}} e^{-n[V(r + \frac{1}{n}(s-r)) - V(r)]} e^{-V(r)} e^{-(n-1)[V(r) + \frac{r^2}{2} + \frac{(r-\alpha)^2}{2t}]} \, dr. \end{aligned} \quad (2.104)$$

Hence, for $n \in \mathbb{N}$ and $t \in (0, \infty)$, we can write

$$\begin{aligned} \bar{\gamma}_{n,t}(\alpha, B) &= \frac{\int_{\mathbb{R}} \mathbb{1}_B(x) \int_{\mathbb{R}} g_{n,t}(\alpha, s) \, d\mu_{\mathcal{N}(\frac{x}{1+t}, \frac{t}{1+t})}(s) \, d\mu_{\mathcal{N}(0, 1+t)}(x)}{\int_{\mathbb{R}} \int_{\mathbb{R}} g_{n,t}(\alpha, s) \, d\mu_{\mathcal{N}(\frac{x}{1+t}, \frac{t}{1+t})}(s) \, d\mu_{\mathcal{N}(0, 1+t)}(x)} \\ &\quad (\alpha \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})), \end{aligned} \quad (2.105)$$

where $g_{n,t}: \mathbb{R}^2 \rightarrow \mathbb{R}$ is as in (2.11). With Fubini's Theorem we have

$$\begin{aligned} &\int_{\mathbb{R}} \mathbb{1}_B(x) \int_{\mathbb{R}} g_{n,t}(\alpha, s) \, d\mu_{\mathcal{N}(\frac{x}{1+t}, \frac{t}{1+t})}(s) \, d\mu_{\mathcal{N}(0, 1+t)}(x) \\ &= \frac{1}{2\pi\sqrt{t}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_B(x) g_{n,t}(\alpha, s) e^{-(s - \frac{x}{1+t})^2 \frac{1+t}{2t}} e^{-x^2 \frac{1}{2(1+t)}} \, dx \, ds \\ &= \frac{1}{2\pi\sqrt{t}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_B(x) e^{2xs \frac{1}{2t}} e^{-x^2 \frac{1}{2t}} \, dx \right) g_{n,t}(\alpha, s) e^{-s^2 \frac{1+t}{2t}} \, ds \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mu_{\mathcal{N}(s, t)}(B) e^{s^2 \frac{1}{2t}} g_{n,t}(\alpha, s) e^{-s^2 \frac{1+t}{2t}} \, ds \\ &= \int_{\mathbb{R}} \mu_{\mathcal{N}(s, t)}(B) g_{n,t}(\alpha, s) \, d\mu_{\mathcal{N}(0, 1)}(s) \quad (B \in \mathcal{B}(\mathbb{R})). \end{aligned} \quad (2.106)$$

With this we obtain (2.10).

2.A.4. Let $n \in \mathbb{N}_{\geq 2}$ and $t \in (0, \infty)$. Note that, with (2.44) and (2.45), $g_{n,t}$ is given by

$$g_{n,t}(\alpha, s) = \frac{\int_{\mathbb{R}} e^{-n[V(r + \frac{1}{n}(s-r))]} e^{-(n-1)(r - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dr}{\int_{\mathbb{R}} e^{-nV(r)} e^{-(n-1)(r - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dr} \quad (\alpha, s \in \mathbb{R}). \quad (2.107)$$

The numerator equals

$$\begin{aligned} & \int_{\mathbb{R}} e^{-n[V(r + \frac{1}{n}(s-r))]} e^{-(n-1)(r - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dr \\ &= \frac{n}{n-1} \int_{\mathbb{R}} e^{-nV(z)} e^{-(n-1)(\frac{n}{n-1}z - \frac{1}{n-1}s - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dz. \end{aligned} \quad (2.108)$$

Via the identities

$$\begin{aligned} & -(n-1) \left(\frac{n}{n-1}z - \frac{1}{n-1}s - \frac{\alpha}{1+t} \right)^2 \\ &= -(n-1) \left(z + \frac{1}{n-1}(z-s) - \frac{\alpha}{1+t} \right)^2 \\ &= -(n-1) \left(z - \frac{\alpha}{1+t} \right)^2 - 2 \left(z - \frac{\alpha}{1+t} \right) (z-s) - \frac{1}{n-1}(z-s)^2 \\ &= -(n-1) \left(z - \frac{\alpha}{1+t} \right)^2 - 2z^2 + 2 \left(s + \frac{\alpha}{1+t} \right) z - 2 \frac{\alpha}{1+t} s - \frac{1}{n-1}(z-s)^2, \end{aligned} \quad (2.109)$$

we get

$$\begin{aligned} g_{n,t}(\alpha, s) &= \frac{\frac{n}{n-1} e^{-\frac{\alpha}{t}s} \int_{\mathbb{R}} e^{[-2z^2 + 2(s + \frac{\alpha}{1+t})z - \frac{1}{n-1}(z-s)^2] \frac{1+t}{2t}} e^{-nV(z)} e^{-(n-1)(z - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dz}{\int_{\mathbb{R}} e^{-nV(r)} e^{-(n-1)(r - \frac{\alpha}{1+t})^2 \frac{1+t}{2t}} dr} \\ & \quad (\alpha, s \in \mathbb{R}). \end{aligned} \quad (2.110)$$

2.A.5. We give the proof of Theorem 2.1.11, namely we prove (a), i.e., the existence of ρ_n mentioned in Theorem 2.1.8 and prove (b), i.e., that for $\alpha \in \mathbb{R}$ the large deviation principle holds for $(\rho_n(\alpha, \cdot))_{n \in \mathbb{N}}$ with rate n and rate function given in (2.18).

• *Proof of (a), existence of ρ_n .*

Let $\mathfrak{J} = \{(t_0, t_1, \dots, t_k, t) : k \in \mathbb{N}_0, 0 = t_0 < t_1 < \dots < t_k < t\}$ and let $j \in \mathfrak{J}$ be given by $j = (t_0, t_1, \dots, t_k, t)$. Define $\pi_j : C([0, t], \mathbb{R}) \rightarrow \mathbb{R}^{k+2}$ by

$$\pi_j(\phi) = (\phi(t_0), \phi(t_1), \dots, \phi(t_k), \phi(t)) \quad (\phi \in C([0, t], \mathbb{R})). \quad (2.111)$$

Similarly as in item 2.A.2, $\bar{\mu}_{n,j} := \mu_n \circ \pi_{[0,t]}^{-1} \circ \pi_j^{-1} \circ m_n^{-1} = (\mu_n \circ \pi_{[0,t]}^{-1} \circ m_n^{-1}) \circ \pi_j^{-1}$ is given by

$$\begin{aligned} \bar{\mu}_{n,j}(A) &= \int_{\mathbb{R}^{k+2}} \mathbb{1}_A(s_0, s_1, \dots, s_k, s_{k+1}) \sqrt{\frac{n}{2\pi(t-t_k)}} e^{-n \frac{(s_{k+1} - s_k)^2}{2(t-t_k)}} \\ & \quad \times \prod_{i=1}^k \left[\sqrt{\frac{n}{2\pi(t_i - t_{i-1})}} e^{-n \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})}} \right] \\ & \quad \times \frac{1}{Z_n} e^{-n[V(s_0) + \frac{s_0^2}{2}]} ds_{k+1} ds_k \dots ds_1 ds_0 \quad (A \in \mathcal{B}(\mathbb{R}^{k+2})). \end{aligned} \quad (2.112)$$

Then $\rho_{n,t,j}: \mathbb{R} \times \mathcal{B}(\mathbb{R}^{k+1}) \rightarrow [0, 1]$ defined by

$$\begin{aligned} \rho_{n,t,j}(\alpha, A) = & \frac{\int_{\mathbb{R}^{k+1}} \mathbb{1}_A(s_0, s_1, \dots, s_k) e^{-n \frac{(\alpha - s_k)^2}{2(t-t_k)}} \prod_{i=1}^k \left[e^{-n \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})}} \right] \\ & \times e^{-n \left[V(s_0) + \frac{s_0^2}{2} \right]} ds_k \cdots ds_1 ds_0}{\int_{\mathbb{R}^{k+1}} e^{-n \frac{(\alpha - s_k)^2}{2(t-t_k)}} \prod_{i=1}^k \left[e^{-n \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})}} \right] e^{-n \left[V(s_0) + \frac{s_0^2}{2} \right]} ds_k \cdots ds_1 ds_0} \end{aligned} \quad (2.113)$$

is the weakly continuous proper regular conditional probability under $\bar{\mu}_{n,j}$ given the coordinate at time t . By Kolmogorov's Theorem (e.g. Bogachev [10, Theorem 7.7.2]), there exists a measure $\rho_{n,t}(\alpha, \cdot)$ on $C([0, t], \mathbb{R})$ (see e.g. [10, Theorem 7.7.4], it is similar to the fact that the Brownian motion is a process on $C([0, t], \mathbb{R})$, which is stated below [10, Theorem 7.7.4]) for which $\rho_{n,t}(\alpha, \cdot) \circ \pi_j^{-1} = \rho_{n,t,j}(\alpha, \cdot)$ for all $j \in \mathfrak{J}$. Because $\alpha \mapsto \rho_{n,t,j}(\alpha, \cdot)$ is strongly continuous for all $n \in \mathbb{N}$ and $j \in \mathfrak{J}$ (see Appendix 2.B), the map $\alpha \mapsto \rho_{n,t}(\alpha, \cdot)$ is (strongly and hence) weakly continuous, i.e., $\rho_{n,t}$ is the weakly continuous proper regular conditional probability of $\mu_n \circ \pi_{[0,t]}^{-1} \circ m_n^{-1}$ under π_t .

• *Proof of (b), large deviation principle.*

Let $j \in \mathfrak{J}$ be given by $j = (t_0, t_1, \dots, t_k)$ and let $\alpha \in \mathbb{R}$. By den Hollander [50, Theorem III.17], the sequence $(\rho_{n,t,j}(\alpha, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate n and rate function $I_j: \mathbb{R}^{k+1} \rightarrow [0, \infty]$ given by

$$(s_0, s_1, \dots, s_k) \mapsto V(s_0) + \frac{s_0^2}{2} + \left[\sum_{i=1}^k \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})} \right] + \frac{(\alpha - s_k)^2}{2(t - t_k)} - \mathfrak{C}_j, \quad (2.114)$$

where $\mathfrak{C}_j$ is such that (2.114) has infimum 0, i.e.,

$$\mathfrak{C}_j = \inf_{s_0, s_1, \dots, s_k \in \mathbb{R}} \left(V(s_0) + \frac{s_0^2}{2} + \left[\sum_{i=1}^k \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})} \right] + \frac{(\alpha - s_k)^2}{2(t - t_k)} \right). \quad (2.115)$$

We will show that I_j is a good rate function, i.e., I_j has compact level sets. Let $c > 0$. Let $K_0 = \{s_0 \in \mathbb{R} : V(s_0) + \frac{s_0^2}{2} \leq c\}$, $K_i = \{s_i \in \mathbb{R} : \sup_{s_{i-1} \in K_{i-1}} \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})} \leq c\}$ for $i \in \{1, \dots, k\}$ and $K_k^* = \{s_k \in \mathbb{R} : \frac{(\alpha - s_k)^2}{2(t - t_k)} \leq c\}$. All these sets are compact and therefore also the set $\{(s_0, s_1, \dots, s_k) \in \mathbb{R}^{k+1} : s_i \in K_i \text{ for } i \in \{0, \dots, k-1\}, s_k \in K_k^* \cap K_k^*\}$. Since the level sets of I_j are closed, we conclude by this that I_j has compact level sets.

We will show that the constant $\mathfrak{C}_j$, does not depend on j , by showing

$$\mathfrak{C}_j = \mathfrak{C}_{(0,t)} = \inf_{s_0 \in \mathbb{R}} V(s_0) + \frac{s_0^2}{2} + \frac{(\alpha - s_0)^2}{2t} = C_{t,\alpha}. \quad (2.116)$$

First, note that $\frac{(a+b)^2}{c+d} \leq \frac{a^2}{c} + \frac{b^2}{d}$ for $a, b, c, d \in \mathbb{R}$ with $c, d > 0$, since $(da - cb)^2 \geq 0$. By this we conclude that $\left[\sum_{i=1}^k \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})} \right] + \frac{(\alpha - s_k)^2}{2(t - t_k)} \geq \frac{(\alpha - s_0)^2}{2t}$ for all $s_0, s_1, \dots, s_k \in \mathbb{R}$ and thus that $\mathfrak{C}_j \geq \mathfrak{C}_{(0,t)}$. By letting $s_i = \psi(t_i)$ for $i \in \{1, \dots, k\}$, where

$\psi(s) = s_0 + \frac{\alpha - s_0}{t} s$, we get $\left[\sum_{i=1}^k \frac{(s_i - s_{i-1})^2}{2(t_i - t_{i-1})} \right] + \frac{(\alpha - s_k)^2}{2(t - t_k)} = \frac{(\alpha - s_0)^2}{2t}$. Hence we conclude $\mathfrak{C}_j = \mathfrak{C}_{(0,t)}$.

By the Dawson-Gärtner projective limit theorem [24, Theorem 4.6.1] the sequence $(\rho_{n,t}(\alpha, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle on $\mathbb{R}^{[0,t]}$, equipped with the product topology (see the beginning of the proof [24, Theorem 5.1.6] why one can replace the projective limit by this product space) with rate n and rate function $\mathbb{R}^{[0,t]} \rightarrow [0, \infty]$ given by $\phi \mapsto \sup_{j \in \mathbb{J}} I_j(\pi_j(\phi))$, i.e.,

$$\begin{aligned} \phi \mapsto & V(\phi(0)) + \frac{\phi(0)^2}{2} - C_{t,\alpha} + \\ & \sup \left\{ \left[\sum_{i=1}^k \frac{(\phi(t_i) - \phi(t_{i-1}))^2}{2(t_i - t_{i-1})} \right] + \frac{(\alpha - \phi(t_k))^2}{2(t - t_k)} \right. \\ & \left. : k \in \mathbb{N}, 0 < t_1 < \dots < t_k < t \right\}. \end{aligned} \quad (2.117)$$

Note that if $\phi \in \mathcal{AC}([0,t], \mathbb{R})$ and $\phi(s)$ does not converge to α as $s \uparrow t$, then $\sup_{j \in \mathbb{J}} I_j(\phi) = \infty$, since $\sup_{s \in (0,t)} \frac{(\alpha - \phi(s))^2}{2(t-s)} = \infty$. Furthermore, if $\phi \in \mathcal{AC}([0,t], \mathbb{R})$ and $\lim_{s \uparrow t} \phi(s) = \alpha$, then the function $\bar{\phi} : [0,t] \rightarrow \mathbb{R}$ given by $\bar{\phi} = \phi$ on $[0,t)$ and $\bar{\phi}(t) = \alpha$ is an element of $\mathcal{AC}([0,t], \mathbb{R})$ and the supremum on the second line in (2.117) is equal to

$$\sup \left\{ \sum_{i=1}^{k+1} \frac{(\phi(t_i) - \phi(t_{i-1}))^2}{2(t_i - t_{i-1})} : k \in \mathbb{N}, 0 < t_1 < \dots < t_k < t_{k+1} = t \right\}, \quad (2.118)$$

In [24, Proof of Lemma 5.1.6] (with $\Lambda^*(x)$ replaced by x^2) it is shown that this supremum is equal to $\frac{1}{2} \int_0^t \dot{\phi}^2(s) ds$. Furthermore, in [24, Proof of Lemma 5.1.6] (last part) it is also shown that (2.117) equals ∞ when $\phi \notin \mathcal{AC}([0,t], \mathbb{R})$. Hence $(\rho_{n,t}(\alpha, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle on $\mathbb{R}^{[0,t]}$ with rate n and rate function $\mathbb{R}^{[0,t]} \rightarrow [0, \infty]$ given by (2.18). This leaves us to prove that the large deviation principle also holds on $C([0,t], \mathbb{R})$ equipped with the topology of uniform convergence. To prove this, by [24, Theorem 4.1.5(b) and Theorem 4.2.6], it is sufficient to show that $(\rho_{n,t}(\alpha, \cdot))_{n \in \mathbb{N}}$ is exponentially tight in $C([0,t], \mathbb{R})$ equipped with the topology of uniform convergence. The exponential tightness follows in turn by showing that the large deviation rate function in (2.18), which we call J here, has compact level sets in the uniform topology of $C([0,t], \mathbb{R})$, i.e., for $b > 0$ the set

$$K_b = \{ \phi \in \mathcal{AC}([0,t], \mathbb{R}) : \lim_{s \uparrow t} \phi(s) = \alpha, J(\phi) \leq b \} \quad (2.119)$$

is compact. We prove this by using the Arzelà-Ascoli theorem (see e.g. [24, Theorem C.8]). Since K_b is closed it is sufficient to show that K_b is bounded and equicontinuous (actually the Arzelà-Ascoli theorem can not directly be used since $[0,t]$ is not compact, however proving that $\tilde{K}_b = \{ \phi \in \mathcal{AC}([0,t], \mathbb{R}) : \phi(t) = \alpha, \tilde{J}(\phi) \}$ is bounded and equicontinuous in $C([0,t], \mathbb{R})$ suffices, where $\tilde{J}$ is the canonical extension of J to $\mathbb{R}^{[0,t]}$. The proof is similar as showing that K_b is bounded and equicontinuous).

Equicontinuity of K_b . For $\phi \in K_b$ and $u, v \in [0, t]$ with $u < v$ (applying Jensen's inequality)

$$\left(\frac{\phi(u) - \phi(v)}{u - v} \right)^2 \leq \frac{1}{(u - v)} \int_u^v \dot{\phi}(s)^2 \, ds \leq \frac{2b + C_{t,\alpha}}{(u - v)}. \quad (2.120)$$

and since $2m|x| \leq x^2 + m^2$ for all $m > 0$ we have

$$|\phi(u) - \phi(v)| \leq \frac{2b + C_{t,\alpha}}{2m} |u - v| + \frac{m}{2}. \quad (2.121)$$

This implies equicontinuity.

Boundedness of K_b . Let $\psi_1, \dots, \psi_k$ be the global minimisers of the (lower semicontinuous) rate function J . By the proof of (c) we know that ψ_i is the linear function that connects $(0, \psi_i(0))$ and (t, α) , i.e., $\psi_i(s) = \psi_i(0) + \frac{\alpha - \psi_i(0)}{t} s$ for $s \in [0, t]$. Let $m > 0$ be such that $|\psi_i(0)| \leq m$ for all $i \in \{1, \dots, k\}$ and $|\alpha| \leq m$ and such that $V(s_0) + \frac{s_0^2}{2} + \frac{(\alpha - s_0)^2}{2t} \geq C_{t,\alpha} + b$ for $s_0 \in [-m, m]^c$. Suppose that $\phi \in \mathcal{AC}([0, t], \mathbb{R})$ with $\lim_{s \uparrow t} \phi(s) = \alpha$ and $\|\phi\|_\infty \geq m + b + 1$ (where $\|\cdot\|_\infty$ is the supremum norm). Let $u \in (0, t)$ be such that $|\phi(u)| \geq m + b + 1$. Then the optimal path ψ from 0 to t which agrees with ϕ in 0, in u and in t (i.e., $\lim_{s \uparrow t} \psi(s) = \alpha$) is the linear interpolation between the points $(0, \phi(0))$, $(u, \phi(u))$, (t, α) (see the proof of (c)), i.e.,

$$\psi(s) = \begin{cases} \phi(0) + \frac{\phi(u) - \phi(0)}{u} s & s \in [0, u], \\ \phi(u) + \frac{\alpha - \phi(u)}{t - u} (s - u) & s \in [u, t]. \end{cases} \quad (2.122)$$

So then we have

$$\begin{aligned} J(\phi) &\geq J(\psi) \geq I_{(0,u,t)}(\psi) \\ &= V(\phi(0)) + \frac{\phi(0)^2}{2} - C_{t,\alpha} + \frac{(\phi(u) - \phi(0))^2}{2u} + \frac{(\alpha - \phi(u))^2}{2(t - u)} \geq b, \end{aligned} \quad (2.123)$$

since either $\phi(0) \in [-m, m]$ and thus $|\phi(0) - \phi(u)|^2 \geq (b + 1)^2 \geq b$ or $\phi(0) \in [-m, m]^c$ and thus $V(s_0) + \frac{s_0^2}{2} + \frac{(\alpha - s_0)^2}{2t} \geq C_{t,\alpha} + b$. By this we conclude that the set K_b bounded in $\|\cdot\|_\infty$ -norm by $m + b + 1$.

§2.B Proper weakly continuous regular conditional probabilities

2.B.1 Definition. Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces with Borel sigma-algebras $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{Y})$. Equip $\mathcal{X} \times \mathcal{Y}$ with the product topology. Then $\mathcal{B}(\mathcal{X} \times \mathcal{Y}) = \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ (i.e., the smallest sigma-algebra containing all sets $A \times B$ with $A \in \mathcal{B}(\mathcal{X})$ and $B \in \mathcal{B}(\mathcal{Y})$, see [10, Theorem 6.4.2]). Let μ be a probability measure on $\mathcal{B}(\mathcal{X} \times \mathcal{Y})$ and let $\pi: \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ the canonical projection. Then $\gamma: \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ is called a *regular conditional probability* under μ of the first coordinate given the second, when γ is a transition kernel and

$$\mu(A \times B) = \int \mathbb{1}_B(y) \gamma(y, A) \, d[\mu \circ \pi^{-1}](y) \quad (A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})). \quad (2.124)$$

γ is called *proper* when $\gamma(y, \cdot) = 0$ for all $y \in \text{supp}(\mu \circ \pi^{-1})^c$, where

$$\text{supp}(\nu) = \mathcal{Y} \setminus \bigcup \{U \subset \mathcal{Y}: U \text{ is open and } \nu(U) = 0\} \quad (2.125)$$

for measures ν on $\mathcal{B}(\mathcal{Y})$. γ is called *weakly continuous* when the map $\alpha \rightarrow \gamma(\alpha, \cdot)$ is weakly continuous.

2.B.2 Lemma. *With the notation as in Definition 2.B.1, if $\gamma_1, \gamma_2: \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ are two proper regular conditional probabilities under μ of the first coordinate given the second, then $\gamma_1(y, \cdot) = \gamma_2(y, \cdot)$ for $\mu \circ \pi^{-1}$ -a.e. $y \in Y$. Consequently, if there exists a weakly continuous proper regular conditional probability of μ under π , then it is unique.*

Proof. The first statement can be found in Bogachev [10, Section 10.4]. The second statement follows from the fact that if γ_1 and γ_2 are proper regular conditional probabilities, then $\mu(B) = 1$ for $B = \{y \in \text{supp}(\mu): \gamma_1(y, \cdot) = \gamma_2(y, \cdot)\}$, and hence B is dense in $\text{supp}(\mu)$. So if γ_1 and γ_2 are weakly continuous, then $B = \text{supp}(\mu)$, i.e., $\gamma_1 = \gamma_2$. $\square$

We will use the following lemma to conclude that regular conditional probabilities with a continuous bounded density are weakly continuous. This lemma is an easy consequence of Lebesgue's Dominated Convergence Theorem.

2.B.3 Lemma. *Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces with Borel sigma-algebras $\mathcal{B}(\mathcal{X})$ and $\mathcal{B}(\mathcal{Y})$. Let μ be a probability measure on $\mathcal{B}(\mathcal{X})$. Let $f \in C_b(\mathcal{X} \times \mathcal{Y}, \mathbb{R})$. If $\gamma: \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ is given by*

$$\gamma(y, A) = \frac{\int \mathbb{1}_A(x) f(y, x) \, d\mu(x)}{\int f(y, x) \, d\mu(x)} \quad (y \in \mathcal{Y}, A \in \mathcal{B}(\mathcal{X})), \quad (2.126)$$

then γ is weakly continuous (even strongly continuous, i.e., $y \mapsto \gamma(y, A)$ is continuous for all $A \in \mathcal{B}(\mathcal{X})$).

§2.C Miscellaneous: additional material on potentials

In this section we present additional material that has not been published in [52].

In this section $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures with potential $V: \mathbb{R} \rightarrow [0, \infty)$.

2.C.1. For the definition of sequentially Gibbs it is important that there is no ambiguity about the conditional kernel γ_n . The assumption that γ_n is continuous is crucial for the uniqueness. It is therefore reasonable to restrict to continuous V , as by Theorem 2.1.4, a continuous V guarantees continuity of γ_n and thus of $\bar{\gamma}_n$. In this perspective, regarding Theorem 2.C.4, we would like to know whether the following holds for continuous V :

$(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if $V \in C^1(\mathbb{R}, [0, \infty))$.

In Theorem 2.C.4 we show that this statement holds for differentiable potentials. In Corollary 2.C.7 we extend this statement to potentials that possess left and a right derivatives. First, in Theorem 2.C.2, we show that whenever V is differentiable, there exists a sequence $(\alpha_n)_{n \in \mathbb{N}}$ with $\alpha_n \rightarrow \alpha$ such that $\bar{\gamma}_n(\alpha_n, \cdot)$ weakly converges.

2.C.2 Theorem. *Let $\bar{\gamma}_n$ be as in (2.5). If V is differentiable in $\alpha \in \mathbb{R}$, then $\bar{\gamma}_n(\frac{n}{n-1}\alpha, \cdot)$ converges weakly (even strongly) to the measure $\mu_{\mathcal{N}(-V'(\alpha), 1)}$.*

Proof. Let $\delta > 0$ be such that

$$\left| \frac{V(\alpha + h) - V(\alpha)}{h} - V'(\alpha) \right| \leq 1 \quad (h \in (-\delta, \delta)). \quad (2.127)$$

Then we have for all $n \in \mathbb{N}$

$$e^{-n[V(\alpha + \frac{y}{n}) - V(\alpha)]} \leq e^{(|V'(\alpha)| + 1)|y|} \quad (y \in [-n\delta, n\delta]). \quad (2.128)$$

Since $y \mapsto e^{(|V'(\alpha)| + 1)|y|}$ is $\mu_{\mathcal{N}(0, 1)}$ -integrable, Lebesgue's Dominated Convergence Theorem implies that for all $A \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{[-n\delta, n\delta]} \mathbb{1}_A(y) e^{-n[V(\alpha + \frac{y}{n}) - V(\alpha)]} e^{-y^2/2} dy \\ = \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-yV'(\alpha)} e^{-y^2/2} dy. \end{aligned} \quad (2.129)$$

Furthermore (because $n \leq e^n$, $n^2 = n(n-1) + n$ and $V \geq 0$)

$$\begin{aligned} \int_{[-n\delta, n\delta]^c} e^{-n[V(\alpha + \frac{y}{n}) - V(\alpha)]} e^{-y^2/2} dy \\ = n \int_{[-\delta, \delta]^c} e^{-n[V(\alpha + z) - V(\alpha)]} e^{-n^2 z^2/2} dz \\ \leq n \int_{[-\delta, \delta]^c} e^{nV(\alpha)} e^{-n^2 z^2/2} dz \leq e^{-n[\frac{n-1}{2}\delta^2 - (V(\alpha) + 1)]} \int_{[-\delta, \delta]^c} e^{-nz^2/2} dz, \end{aligned} \quad (2.130)$$

where the last term converges to 0 as $n \rightarrow \infty$. So, by (2.129) – (2.130),

$$\int_{\mathbb{R}} \mathbb{1}_A(y) e^{-n[V(\alpha + \frac{y}{n}) - V(\alpha)]} e^{-y^2/2} dy \rightarrow \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-yV'(\alpha)} e^{-y^2/2} dy, \quad (2.131)$$

and hence, by (2.5), $\lim_{n \rightarrow \infty} \bar{\gamma}_n(\frac{n}{n-1}\alpha, A) = \mu_{\mathcal{N}(-V'(\alpha), 1)}(A)$ for all $A \in \mathcal{B}(\mathbb{R})$, i.e., the sequence $(\bar{\gamma}_n(\frac{n}{n-1}\alpha, \cdot))_{n \in \mathbb{N}}$ converges strongly (and hence weakly) to $\mu_{\mathcal{N}(-V'(\alpha), 1)}$. $\square$

2.C.3 Lemma. *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}$ and $x \in \mathbb{R}$. Then $x_n \rightarrow x$ if and only if $(\mu_{\mathcal{N}(x_n, 1)})_{n \in \mathbb{N}}$ weakly converges to $\mu_{\mathcal{N}(x, 1)}$.*

Proof. Whenever $x_n \rightarrow x$, then the weak convergence of the normal distributions is a consequence of [6, Theorem 26.3]. Suppose that $x_n \not\rightarrow x$. The sequence $(x_n)_{n \in \mathbb{N}}$

has a limit point in $[-\infty, \infty]$. By restricting to a subsequence we may assume that $(x_n)_{n \in \mathbb{N}}$ converges to an element $y \in [-\infty, \infty]$, where $y \neq x$. If $y \in \mathbb{R}$, then $(\mu_{\mathcal{N}(x_n, 1)})_{n \in \mathbb{N}}$ weakly converges to $\mu_{\mathcal{N}(y, 1)}$ by [6, Theorem 26.3]. If $y \in \{-\infty, \infty\}$, then $\mu_{\mathcal{N}(x_n, 1)}([-M, M]) \rightarrow 0$ for all $M > 0$. On the other hand $\mu_{\mathcal{N}(x_n, 1)}(\mathbb{R}) = 1$ for all $n \in \mathbb{N}$. This implies that $\mu_{\mathcal{N}(x_n, 1)}$ does not (weakly or strongly) converge to a probability measure on $\mathbb{R}$. $\square$

2.C.4 Theorem. *Suppose that V is differentiable. Then $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if V' is continuous, i.e., $V \in C^1(\mathbb{R}, [0, \infty))$.*

Proof. Suppose that $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs with specification kernel $\gamma : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$. By Lemma 2.1.3 γ is weakly continuous. By Theorem 2.C.2 $\gamma(\alpha, \cdot) = \mu_{\mathcal{N}(-V'(\alpha), 1)}$ for all $\alpha \in \mathbb{R}$. By Lemma 2.C.3 this implies that V' is continuous. The other implication follows by Theorem 2.1.4. $\square$

2.C.5 Remark. Let $\gamma : \mathbb{R} \times \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ be a probability kernel. Consider the following statements for a sequence $(\rho_n)_{n \in \mathbb{N}}$ with probability kernels $(\gamma_n)_{n \in \mathbb{N}}$ as in Definition 2.1.2.

- (a) For all $\alpha \in \mathbb{R}$ there exists a sequence $(\beta_n)_{n \in \mathbb{N}}$ such that $\bar{\gamma}_n(\beta_n, \cdot) \xrightarrow{w} \gamma(\alpha, \cdot)$,
- (b) $\alpha \mapsto \gamma(\alpha, \cdot)$ is weakly continuous.

For the sequence of finite-volume mean-field Gibbs measures $(\nu_n)_{n \in \mathbb{N}}$ with a differentiable potential, Theorem 2.C.4 and Theorem 2.C.2 imply that $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if (a) and (b) hold. It is not clear whether this equivalence holds in general.

2.C.6 Theorem. *Let $\alpha \in \mathbb{R}$. Let $V : \mathbb{R} \rightarrow [0, \infty)$ be measurable. Suppose that the left derivative of V at α , $D_-V(\alpha)$, and its right derivative at α , $D_+V(\alpha)$, exist. There exist sequences $(\alpha_n^+)_{n \in \mathbb{N}}$ and $(\alpha_n^-)_{n \in \mathbb{N}}$ such that $\alpha_n^- < \alpha < \alpha_n^+$ and $\alpha_n^+ \downarrow \alpha$, $\alpha_n^- \uparrow \alpha$ and there exists an $f \in C_b(\mathbb{R})$ such that $\lim_{n \rightarrow \infty} \int f d[\bar{\gamma}_n(\alpha_n^+, \cdot)] \neq \lim_{n \rightarrow \infty} \int f d[\bar{\gamma}_n(\alpha_n^-, \cdot)]$.*

Proof. As both the left and right derivative exist, there exists a $\delta > 0$ and an $M > 0$ such that

$$-M \leq \frac{V(\alpha + h) - V(\alpha)}{h} \leq M \quad (h \in [-2\delta, 2\delta]). \quad (2.132)$$

From this V is also bounded on $[\alpha - 2\delta, \alpha + 2\delta]$. Let $L > 0$ be such that $|V| \leq L$ on $[\alpha - 2\delta, \alpha + 2\delta]$. Let $(\alpha_n)_{n \in \mathbb{N}}$ be a sequence in $[\alpha - \delta, \alpha + \delta]$. Then

$$\mathbb{1}_{[-n\delta, n\delta]}(y) e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} \leq e^{M|y|} \quad (n \in \mathbb{N}, y \in \mathbb{R}). \quad (2.133)$$

For all $y \in \mathbb{R}$

$$\begin{aligned} n [V(\alpha_n + \frac{y}{n}) - V(\alpha_n)] &= (n(\alpha_n - \alpha) + y) \frac{V(\alpha_n + \frac{y}{n}) - V(\alpha)}{\alpha_n + \frac{y}{n} - \alpha} \\ &\quad - n(\alpha_n - \alpha) \frac{V(\alpha_n) - V(\alpha)}{\alpha_n - \alpha}. \end{aligned} \quad (2.134)$$

Hence for all $y \in \mathbb{R}$

$$\begin{aligned} & \lim_{n \rightarrow \infty} -n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)] \\ &= \begin{cases} -yD_+V(\alpha) & \alpha_n = \alpha_n^+, y > -\delta, \\ (\delta - y)D_-V(\alpha) - \delta D_+V(\alpha) & \alpha_n = \alpha_n^+, y < -\delta, \\ (\delta - y)D_+V(\alpha) - \delta D_-V(\alpha) & \alpha_n = \alpha_n^-, y > \delta, \\ -yD_-V(\alpha) & \alpha_n = \alpha_n^-, y < \delta. \end{cases} \end{aligned} \quad (2.135)$$

Furthermore (because $n \leq e^n$, $n^2 = n(n-1) + n$, $V \geq 0$ and $V \leq L$ on $[\alpha - 2\delta, \alpha + 2\delta]$)

$$\begin{aligned} & \int_{[-n\delta, n\delta]^c} e^{-n[V(\alpha_n + \frac{y}{n}) - V(\alpha_n)]} e^{-y^2/2} dy \\ &= n \int_{[-\delta, \delta]^c} e^{-n[V(\alpha_n + z) - V(\alpha_n)]} e^{-n^2 z^2/2} dz \\ &\leq n \int_{[-\delta, \delta]^c} e^{nV(\alpha_n)} e^{-n^2 z^2/2} dz \\ &\leq e^{-n[\frac{n-1}{2} - (L+1)]} \int_{[-\delta, \delta]^c} e^{-nz^2/2} dz. \end{aligned} \quad (2.136)$$

Define $\alpha_n^+ := \alpha + \frac{\delta}{n}$ and $\alpha_n^- := \alpha - \frac{\delta}{n}$ for $n \in \mathbb{N}$. Therefore by the Lebesgue's Dominated Convergence Theorem and (2.133) ($y \mapsto e^{M|y|}$ is $\mu_{\mathcal{N}(0,1)}$ -integrable), for $A \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-n[V(\alpha_n^+ + \frac{y}{n}) - V(\alpha_n^+)]} e^{-y^2/2} dy \\ &= \int_{-\infty}^{-\delta} \mathbb{1}_A(y) e^{(\delta-y)D_-V(\alpha) - \delta D_+V(\alpha)} e^{-y^2/2} dy \\ &\quad + \int_{-\delta}^{\infty} \mathbb{1}_A(y) e^{-yD_+V(\alpha)} e^{-y^2/2} dy \\ &= e^{\delta D_-V(\alpha) - \delta D_+V(\alpha)} \int_{-\infty}^{-\delta} \mathbb{1}_A(y) e^{-yD_-V(\alpha)} e^{-y^2/2} dy \\ &\quad + \int_{-\delta}^{\infty} \mathbb{1}_A(y) e^{-yD_+V(\alpha)} e^{-y^2/2} dy, \end{aligned} \quad (2.137)$$

and similarly

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \mathbb{1}_A(y) e^{-n[V(\alpha_n^- + \frac{y}{n}) - V(\alpha_n^-)]} e^{-y^2/2} dy \\ &= \int_{-\infty}^{\delta} \mathbb{1}_A(y) e^{-yD_-V(\alpha)} e^{-y^2/2} dy \\ &\quad + e^{\delta D_+V(\alpha) - \delta D_-V(\alpha)} \int_{\delta}^{\infty} \mathbb{1}_A(y) e^{-yD_+V(\alpha)} e^{-y^2/2} dy. \end{aligned} \quad (2.138)$$

With

$$Q_- = \int_{-\infty}^{-\delta} e^{-yD_- V(\alpha)} e^{-y^2/2} dy \quad (2.139)$$

$$Q_+ = \int_{\delta}^{\infty} e^{-yD_+ V(\alpha)} e^{-y^2/2} dy \quad (2.140)$$

we obtain for all $A \in \mathcal{B}(\mathbb{R})$ with $A \subset [-\delta, \delta]$

$$\begin{aligned} \gamma_n(\alpha_n^+, A) &\rightarrow \frac{\int_{-\delta}^{\delta} \mathbb{1}_A(y) e^{-yD_+ V(\alpha)} e^{-y^2/2} dy}{\int_{-\delta}^{\delta} e^{-yD_+ V(\alpha)} e^{-y^2/2} dy + e^{\delta D_- V(\alpha) - \delta D_+ V(\alpha)} Q_- + Q_+} \\ &= \frac{\int_{-\delta}^{\delta} \mathbb{1}_A(y) e^{(\delta-y)D_+ V(\alpha)} e^{-y^2/2} dy}{\int_{-\delta}^{\delta} e^{(\delta-y)D_+ V(\alpha)} e^{-y^2/2} dy + e^{\delta D_- V(\alpha)} Q_- + e^{\delta D_+ V(\alpha)} Q_+}, \end{aligned} \quad (2.141)$$

$$\begin{aligned} \gamma_n(\alpha_n^-, A) &\rightarrow \frac{\int_{-\delta}^{\delta} \mathbb{1}_A(y) e^{-yD_- V(\alpha)} e^{-y^2/2} dy}{\int_{-\delta}^{\delta} e^{-yD_- V(\alpha)} e^{-y^2/2} dy + Q_- + e^{\delta D_+ V(\alpha) - \delta D_- V(\alpha)} Q_+} \\ &= \frac{\int_{-\delta}^{\delta} \mathbb{1}_A(y) e^{(\delta-y)D_- V(\alpha)} e^{-y^2/2} dy}{\int_{-\delta}^{\delta} e^{(\delta-y)D_- V(\alpha)} e^{-y^2/2} dy + e^{\delta D_- V(\alpha)} Q_- + e^{\delta D_+ V(\alpha)} Q_+}. \end{aligned} \quad (2.142)$$

As $\lim_{n \rightarrow \infty} \gamma_n(\alpha_n^-, [-\delta, \delta]) \neq \lim_{n \rightarrow \infty} \gamma_n(\alpha_n^+, [-\delta, \delta])$ and $f_k \uparrow \mathbb{1}_{[-\delta, \delta]}$, where $f_k(x) = d(x, [-(1 - \frac{1}{k})\delta, (1 - \frac{1}{k})\delta])$ there exists a k such that

$$\lim_{n \rightarrow \infty} \int f_k d[\gamma_n(\alpha_n^-, \cdot)] \neq \lim_{n \rightarrow \infty} \int f_k d[\gamma_n(\alpha_n^+, \cdot)]. \quad (2.143)$$

□

As a direct consequence of Theorem 2.C.6 and Theorem 2.C.4 we obtain the following.

2.C.7 Corollary. *Suppose that the left and right derivatives of $V : \mathbb{R} \rightarrow [0, \infty)$ exist everywhere in $\mathbb{R}$. Then $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if $V \in C^1(\mathbb{R}, [0, \infty))$.*

§2.C.1 Differentiability of the evolved potential

In item 5 of the Discussion 2.1.8 we showed that the time evolved sequence $(\mu_{n,t})_{n \in \mathbb{N}}$ can be written as

$$\mu_{n,t}(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(x) e^{-n(V_{n,t} \circ m_n)(x)} d\mu_{\mathcal{N}(0, (1+t)I_n)}(x) \quad (A \in \mathcal{B}(\mathbb{R}^n)). \quad (2.144)$$

with $V_{n,t} : \mathbb{R} \rightarrow [0, \infty)$ as in 2.36. We mentioned in (2.38) that the pointwise limit of $V_{n,t}$ exists and equals $V_t : \mathbb{R} \rightarrow [0, \infty)$, with

$$V_t(r) = \inf_{s \in \mathbb{R}} \left[V(s) + \frac{1+t}{2t} \left(s - \frac{r}{1+t} \right)^2 \right] \quad (r \in \mathbb{R}). \quad (2.145)$$

To some extent we show that V_t is effectively the potential of $(\mu_{n,t})_{n \in \mathbb{N}}$ in Corollary 2.C.9. First, recall that (see (2.44))

$$I_{t,r}(s) = V(s) + \frac{1+t}{2t} \left(s - \frac{r}{1+t} \right)^2, \quad (2.146)$$

and that $r \mapsto I_{t,\alpha}(r) - \inf_{s \in \mathbb{R}} I_{t,\alpha}(s)$ is equal to (2.9). Note that

$$V_t(r) = \inf_{s \in \mathbb{R}} I_{t,r}(s) \quad (r \in \mathbb{R}). \quad (2.147)$$

2.C.8 Lemma. *The following are equivalent.*

- (a) $V_t \in C^1(\mathbb{R})$,
- (b) V_t is differentiable,
- (c) $I_{t,\alpha}$ has a unique global minimiser for all $\alpha \in \mathbb{R}$.

Whenever $I_{t,\alpha}$ has a unique global minimiser $q(\alpha)$ for all $\alpha \in \mathbb{R}$, then

$$V'(q(\alpha)) = (1+t)V'_t(\alpha) \quad (\alpha \in \mathbb{R}). \quad (2.148)$$

Proof. The equivalences follow from the theory of Cannarsa and Sinestrari [18] in the following way. In our case we consider $L(s, \xi, \zeta) = \frac{1}{2}\zeta^2$, see (2.20) and the equivalence between (a2) and (a3) of Corollary 2.1.9. It follows that with this L the conditions of [18, Theorem 6.4.3 and Theorem 6.4.9] are satisfied. [18, Theorem 6.4.3] implies that V_t is semiconcave. By [18, Theorem 6.4.9] the minimising paths of the infimum in (2.20) are in one to one correspondence with the reachable gradient (see [18, Definition 3.1.10]). By [18, Theorem 3.3.4 and Theorem 3.3.6] imply that the set of reachable gradients consist of only one element if and only if V_t is differentiable, in which case V_t is continuously differentiable.

Suppose that $I_{t,\alpha}$ has a unique global minimiser $q(\alpha)$ for all $\alpha \in \mathbb{R}$. Then $\Phi_2 V > -\frac{1+t}{2t}$ and thus $\Phi_2 I_{t,\alpha} > 0$ by Lemma 2.5.2, therefore $I_{t,\alpha}$ is strictly convex by Lemma 2.5.4. Therefore the function

$$r \mapsto V'(r) + \frac{1+t}{t}r - \frac{\alpha}{t} \quad (2.149)$$

has a unique zero, which is $q(\alpha)$. So the function (multiply (2.149) by t)

$$r \mapsto tV'(r) + (1+t)r \quad (s \in \mathbb{R}) \quad (2.150)$$

is bijective and continuous and

$$tV'(r) + (1+t)r = \alpha \iff r = q(\alpha) \quad (\alpha \in \mathbb{R}), \quad (2.151)$$

Hence $q^{-1}(r) = tV'(r) + (1+t)r$. Since q^{-1} is bijective and continuous, it is either strictly increasing or strictly decreasing. Hence also q is either strictly increasing or strictly decreasing. By Lebesgue's theorem, q is differentiable almost everywhere. Let $D \subset \mathbb{R}$ be a measurable set with $\lambda(\mathbb{R} \setminus D) = 0$ on which q is differentiable. Then

$$V'_t(\alpha) = V'(q(\alpha)) \cdot q'(\alpha) + \frac{1+t}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) \left(q'(\alpha) - \frac{1}{1+t} \right) \quad (\alpha \in D). \quad (2.152)$$

Since $q(\alpha)$ is the global minimiser of $r \mapsto V(r) + \frac{1+t}{2t}(r - \frac{\alpha}{1+t})^2$, we have

$$V'(q(\alpha)) + \frac{1+t}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) = 0 \quad (\alpha \in D), \quad (2.153)$$

$$-\frac{1+t}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) = V'(q(\alpha)) \quad (\alpha \in D). \quad (2.154)$$

Therefore

$$\begin{aligned} V'_t(\alpha) &= \left[V'(q(\alpha)) + \frac{1+t}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) \right] q'(\alpha) - \frac{1}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) \\ &= -\frac{1}{t} \left(q(\alpha) - \frac{\alpha}{1+t} \right) = \frac{1}{1+t} V'(q(\alpha)) \quad (\alpha \in D). \end{aligned} \quad (2.155)$$

As V_t is continuously differentiable, q is continuous, and D is dense in $\mathbb{R}$,

$$V'_t(\alpha) = \frac{1}{1+t} V'(q(\alpha)). \quad (2.156)$$

2.C.9 Corollary. *Let $\nu_{n,t}$ be the measure on $\mathcal{B}(\mathbb{R}^n)$ defined by*

$$\nu_{n,t}(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(x) e^{-n(V_{n,t} \circ m_n)(x)} d\mu_{\mathcal{N}(0, (1+t)I_n)}(x) \quad (A \in \mathcal{B}(\mathbb{R}^n)). \quad (2.157)$$

Then the following are equivalent

- (a) $(\nu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs,
- (b) $V_t \in C^1(\mathbb{R})$,
- (c) V_t is differentiable,
- (d) $I_{t,\alpha}$ has a unique global minimiser for all $\alpha \in \mathbb{R}$,
- (e) $(\mu_{n,t})_{n \in \mathbb{N}}$ is sequentially Gibbs.

In case $(\nu_{n,t})_{n \in \mathbb{N}}$ and $(\mu_{n,t})_{n \in \mathbb{N}}$ are sequentially Gibbs, their specification kernel is given by

$$\gamma(\alpha, B) = \mu_{\mathcal{N}(-(1+t)V'_t(\alpha), 1+t)}(B) = \mu_{\mathcal{N}(-V'(q(\alpha)), 1+t)}(B) \quad (B \in \mathcal{B}(\mathbb{R})), \quad (2.158)$$

where $q(\alpha)$ is the unique global minimiser of $I_{t,\alpha}$ for all $\alpha \in \mathbb{R}$.

Proof. By [18, Theorem 3.2.1] the left and right derivatives of V exist everywhere in $\mathbb{R}$. Therefore the equivalence between (a) and (b) follows from Corollary 2.C.7 (eventhough Corollary 2.C.7 is based on Theorem 2.1.4, in which the variance of the normal distribution is given by 1 instead of $1+t$, the proof can be adapted to general variances). The equivalence between (d) and (e) follows from Theorem 2.1.7. The equivalences between (b), (c) and (d) are shown in Lemma 2.C.8. (2.158) follows from Lemma 2.C.8. $\square$

2.C.10. In Section 3.5.3 we make an attempt in a more general context to make sense of the fact that a limit of potentials V_n determines the fact whether the sequence with those potentials is sequentially Gibbs.



Level 2 mean-field Gibbs-non-Gibbs transitions

Abstract

We consider mean-field sequences on a general state space $\mathcal{X}$ that are described in terms of empirical distributions. A framework is presented to describe the sequential Gibbsianness of those mean-field sequences. We discuss properties on the potential V , in this case being a function on the space of compactly supported probability measures on $\mathcal{X}$, that imply sequential Gibbsianness of finite-volume mean-field Gibbs measures with potential V . We give formulas for the corresponding conditional probabilities in the case the mean-field sequences are transformed. In case of independent transformations we show that sequential Gibbsianness is implied by uniqueness of global minimisers of the large deviation rate function of the initial empirical distribution given the transformed empirical distribution.

§3.1 Introduction and main results

§3.1.1 Motivation

In Chapter 2 the finite-volume mean-field Gibbs measures (on $\mathbb{R}$) are typically of the form

$$A \mapsto \frac{1}{Z_n} \int_A e^{-nV(m_n(x_1, \dots, x_n))} d\lambda(x_1) \cdots d\lambda(x_n), \quad (3.1)$$

where λ is a probability measure, V a positive measurable function, $m_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i$ and Z_n is the normalising constant. In order not to be only able to consider this m_n , but to be able to replace it, for example, by $(x_1, \dots, x_n) \mapsto \frac{1}{n} \sum_{i=1}^n x_i^2$ or $(x_1, \dots, x_n) \mapsto \frac{1}{n} \sum_{i=1}^n x_i^3$, or any other function that is invariant under exchanging the coordinates, we generalise the definition of sequentially Gibbs to sequences of exchangeable measures. There is a one-to-one correspondence between the equivalence class of a vector $(x_1, \dots, x_n)$ under the equivalence relation described by the exchangeability of the coordinates, and the empirical distribution $\frac{1}{n} \sum_{i=1}^n \delta_{x_i}$. As this empirical distribution captures the exchangeability, we generalise finite-volume mean-field Gibbs measures by allowing them to be of the form

$$A \mapsto \frac{1}{Z_n} \int_A e^{-nV(L_n(x_1, \dots, x_n))} d\lambda(x_1) \cdots d\lambda(x_n), \quad (3.2)$$

where $L_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$, V is now a function on the space of compactly supported probability measures and, instead of $\mathbb{R}$, we consider a general state space (moreover, λ is not any more required to be a probability measure). This framework in terms of the empirical distribution is also called the level-2 setting (by, for example, Ellis [28]). As $m_n(x) = \int s d[L_n(x)](s)$, this generalises the framework of Chapter 2.

In this chapter we provide the mathematical objects and definitions needed to treat this level-2 setting, for general state spaces (not restricted to compact spaces). We follow the large deviation framework of Gibbs-non-Gibbs transitions put forward in van Enter, Fernández, den Hollander and Redig [31].

§3.1.2 Outline

In Section 3.3 we state definitions and basic results that are used in the rest of this chapter. In Section 3.4 we give a definition of sequentially Gibbs that is formulated in terms of the empirical distributions and show that this agrees with previous definitions of Gibbsianness for mean-field sequences (see Külske and Opoku [61] and van Enter, Külske, Opoku and Ruszel [33] for the definition on compact state spaces in terms of the empirical distribution and Ermolaev and Külske [36], Fernández, den Hollander and Martínez [39] and den Hollander, Redig and van Zuijlen [52] for the definition on the state spaces $\{-1, 1\}$ and $\mathbb{R}$ in terms of the magnetisation). In Section 3.5 we discuss sequences of finite-volume mean-field Gibbs measures and show that, whenever their Hamiltonians originate from a potential V , certain differentiability properties of V imply that such sequences are Gibbs. In Section 3.6 we consider transformations of

kernels of sequences of finite-volume mean-field Gibbs measures and derive relations between regular conditional kernels, in particular, when independent transformations or transformations that possess a density function are considered. In Section 3.7 we prove that whenever we have a large deviation principle for the probabilities of the initial empirical distribution conditioned on the final (or transformed) empirical distribution, uniqueness of the global minimiser of this rate function implies sequential Gibbsianness of the transformed sequence of finite-volume mean-field Gibbs sequences.

§3.2 Some notations and definitions

We write $\mathbb{N}$ for the set $\{1, 2, 3, \dots\}$. For a topological space $\mathcal{X}$ we write $\mathcal{B}(\mathcal{X})$ for the Borel- σ -algebra and $\mathcal{P}(\mathcal{X})$ for the set of probability measures $\mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$. For a function $f : \mathcal{X} \rightarrow \mathbb{R}$ and $c \in \mathbb{R}$ we write $[f \leq c] = \{x \in \mathcal{X} : f(x) \leq c\}$. Similarly, we use the notations $[f > c]$, $[f \geq c]$ and $[f < c]$. We write $\mathcal{B}\mathfrak{a}(\mathcal{X})$ for the Baire- σ -algebra on $\mathcal{X}$. For $x \in X$ we write δ_x for the element in $\mathcal{P}(\mathcal{X})$ with $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise. For $\mu \in \mathcal{P}(\mathcal{X})$ we write $\text{supp } \mu = \{x \in \mathcal{X} : \mu(V) > 0 \text{ for all } V \in \mathcal{N}_x\}$ and call this the support of μ . We write $\mathcal{P}_c(\mathcal{X})$ for the set of compactly supported probability measures in $\mathcal{P}(\mathcal{X})$. As short-hand notation for a probability measure $\lambda \in \mathcal{P}(\mathcal{X})$, we write λ^n for the product measure $\bigotimes_{i=1}^n \lambda$ on $\mathcal{B}(\mathcal{X}^n)$. Let $\mathcal{Y}$ be another topological space and $P : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ be a probability kernel. For $\nu \in \mathcal{P}(\mathcal{X})$ we write νP and $\nu \rtimes P$ for the probability measures¹ on $\mathcal{B}(\mathcal{Y})$ and $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$, respectively, given by

$$\nu P(B) = \int_{\mathcal{X}} P(v, B) \, d\nu(v), \quad (3.3)$$

$$\nu \rtimes P(D) = \int_{\mathcal{X}} \int_{\mathcal{Y}} \mathbb{1}_D(v, w) \, d[P(v, \cdot)](w) \, d\nu(v). \quad (3.4)$$

We equip $\mathcal{P}(\mathcal{X})$ with the weak topology and $\mathcal{P}_c(\mathcal{X})$ with the C -weak topology (see Definition 3.3.6), unless indicated otherwise.

§3.3 Preliminaries

In this section we introduce some definitions and a few theorems that serve as a basis for the theory developed in the rest of the chapter.

In this section $\mathcal{X}$ and $\mathcal{Y}$ are topological spaces.

§3.3.1 Exchangeability and permutations

In this section we introduce permutations, exchangeability of a measure on a product space $\mathcal{X}^n$, and the frequently used definition of the empirical distribution of a vector.

¹That $\nu \rtimes P$ is a measure follows from the fact that $w \mapsto \mathbb{1}_D(v, w)$ is $\mathcal{B}(\mathcal{Y})$ -measurable (see Bogachev [10, Theorem 3.4.1]) and that $v \mapsto \int_{\mathcal{Y}} \mathbb{1}_D(v, w) \, d[P(v, \cdot)](w)$ is $\mathcal{B}(\mathcal{X})$ -measurable.

3.3.1 Definition. We write $\text{Sym}(n)$ for the group of permutations on $\{1, \dots, n\}$. For a permutation $\sigma \in \text{Sym}(n)$ we write $\hat{\sigma}$ for the function $\mathcal{X}^n \rightarrow \mathcal{X}^n$ given by $\hat{\sigma}(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$. We define the equivalence relation $\sim_{\text{Sym}}$ on $\mathcal{X}^n$ for $x, y \in \mathcal{X}^n$ by

$$x \sim_{\text{Sym}} y \iff \exists \sigma \in \text{Sym}(n) \left[x = \hat{\sigma}(y) \right]. \quad (3.5)$$

3.3.2 Definition. We define the *empirical distribution function* $L_n : \mathcal{X}^n \rightarrow \mathcal{P}_c(\mathcal{X})$ by

$$L_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} \quad (x = (x_1, \dots, x_n) \in \mathcal{X}^n). \quad (3.6)$$

$L_n(x)$ is called the *empirical distribution* of x . We write $\mathcal{P}_{\text{emp}}^n(\mathcal{X})$ for the set of all empirical distributions, i.e., $\mathcal{P}_{\text{emp}}^n(\mathcal{X}) = L_n(\mathcal{X}^n)$. We use the notation L_n also for the map $L_n : \mathcal{Y}^n \rightarrow \mathcal{P}_c(\mathcal{Y})$ with $L_n(y) = \frac{1}{n} \sum_{i=1}^n \delta_{y_i}$, and similarly $\mathcal{P}_{\text{emp}}^n(\mathcal{Y}) = L_n(\mathcal{Y}^n)$.

The proof of the following fact is left for the reader.

3.3.3 Theorem. For $x, y \in \mathcal{X}^n$, $x \sim_{\text{Sym}} y$ if and only if $L_n(x) = L_n(y)$. Hence $\mathcal{X}^n / \sim_{\text{Sym}} \cong \mathcal{P}_{\text{emp}}^n(\mathcal{X})$.

3.3.4 Definitions. We say that a measure ρ on $\mathcal{X}^n$ is *exchangeable* if $\rho_n = \rho_n \circ \hat{\sigma}$ for all $\sigma \in \text{Sym}(n)$.

3.3.5. Fix $n \in \mathbb{N}$. We make some observations from which one can conclude the following: If ρ is an exchangeable measure on $\mathcal{X}^n$, λ is a measure on $\mathcal{X}$ and $\rho \ll \lambda^n$ and $\lambda^n \ll \rho$, then there exists a $V : \mathcal{P}_{\text{emp}}^n(\mathcal{X}) \rightarrow \mathbb{R}$ such that $V \circ L_n$ is $\mathcal{B}(\mathcal{X}^n)$ -measurable and

$$\rho(A) = \int_A e^{-nV \circ L_n(v)} d\lambda^n(v) \quad (A \in \mathcal{B}(\mathcal{X}^n)). \quad (3.7)$$

- (a) Let $f : \mathcal{X}^n \rightarrow \mathcal{Y}$ be such that $f \circ \hat{\sigma} = f$ for all $\sigma \in \text{Sym}(n)$. Then $f(u) = f(v)$ for all $u, v \in \mathcal{X}^n$ with $L_n(u) = L_n(v)$; thus there is a map $\tilde{f} : \mathcal{P}_{\text{emp}}^n(\mathcal{X}) \rightarrow \mathcal{Y}$ with $f = \tilde{f} \circ L_n$.
- (b) Let ρ be an exchangeable measure on $\mathcal{X}^n$. Let λ be a measure on $\mathcal{X}$. Suppose that $\rho \ll \lambda^n$. Then there is a function $f : \mathcal{P}_{\text{emp}}^n(\mathcal{X}) \rightarrow [0, \infty)$ such that $f \circ L_n$ is measurable and $\rho = (f \circ L_n)\lambda^n$, i.e., $\rho(A) = \int_A f \circ L_n d\lambda^n$ for $A \in \mathcal{B}(\mathcal{X}^n)$.

Indeed, suppose $g : \mathcal{X}^n \rightarrow [0, \infty)$ is measurable and $\rho = g\lambda^n$. Then there exists a set $Z \subset \mathcal{X}^n$ such that $\lambda^n(\mathcal{X}^n \setminus Z) = 0$ and $g \circ \hat{\sigma}(v) = g(v)$ for all $v \in Z$. Let $h = g\mathbb{1}_Z$. By (a) there exists a $\tilde{h} : \mathcal{P}_{\text{emp}}^n(\mathcal{X}) \rightarrow [0, \infty)$ with $\tilde{h} \circ L_n = h$.

- (c) Let ρ and λ be as in (b). Suppose that also $\lambda^n \ll \rho$. Then there exists a function $V : \mathcal{P}_{\text{emp}}^n(\mathcal{X}) \rightarrow \mathbb{R}$ such that $f = e^{-nV \circ L_n}$ satisfies the above.

Indeed, because $\lambda^n \ll \rho$ as well as $\rho \ll \lambda^n$, one may take g and also h as above to be strictly positive.

§3.3.2 Convergence on the compactly supported probabilities

In this section we introduce the notions of weak, vague and C -weak convergence. We discuss the relation of C -weak convergence with elements of $\bigcup_{n \in \mathbb{N}} \mathcal{P}_{emp}^n(\mathcal{X})$, and also introduce the notion of LIM. These definitions are used to define sequential Gibbsianness in Section 3.4.

3.3.6 Definition. A net $(\mu_\alpha)_{\alpha \in \mathbb{A}}$ in $\mathcal{P}(\mathcal{X})$ is said to *converge weakly* to μ ($\mu_\alpha \xrightarrow{w} \mu$) if

$$\int f \, d\mu_\alpha \rightarrow \int f \, d\mu \quad (f \in C_b(\mathcal{X})). \quad (3.8)$$

A net $(\mu_\alpha)_{\alpha \in \mathbb{A}}$ in $\mathcal{P}(\mathcal{X})$ it is said to *converge vaguely* to μ ($\mu_\alpha \xrightarrow{v} \mu$) if

$$\int f \, d\mu_\alpha \rightarrow \int f \, d\mu \quad (f \in C_c(\mathcal{X})). \quad (3.9)$$

A net $(\mu_\alpha)_{\alpha \in \mathbb{A}}$ in $\mathcal{P}_c(\mathcal{X})$ it is said to *converge C -weakly* to μ ($\mu_\alpha \xrightarrow{C-w} \mu$) if

$$\int f \, d\mu_\alpha \rightarrow \int f \, d\mu \quad (f \in C(\mathcal{X})). \quad (3.10)$$

These notions of convergence form a convergence class and therefore have an underlying topology (see Kelley [56, Theorem 2.9]). The topology on $\mathcal{P}(\mathcal{X})$ induced by the weak convergence is called the *weak topology*, the topology on $\mathcal{P}(\mathcal{X})$ induced by the vague convergence is called the *vague topology*, and the topology on $\mathcal{P}_c(\mathcal{X})$ induced by the C -weak convergence is called the *C -weak topology*.

Weak convergence is the same as C -weak convergence in case $\mathcal{X}$ is a compact space.

As is mentioned in Section 3.2, we equip $\mathcal{P}_c(\mathcal{X})$ with the C -weak topology unless indicated otherwise.

Before discussing the different topologies, we discuss in 3.3.8 when these topologies are Hausdorff.

3.3.7 Definition. A Hausdorff space $\mathcal{X}$ is called *perfectly normal* whenever every closed set $C \subset \mathcal{X}$ has the form $C = f^{-1}(\{0\})$ for some $f \in C(\mathcal{X})$.

3.3.8. If the Borel- σ -algebra $\mathcal{B}(\mathcal{X})$ is equal to the Baire- σ -algebra $\mathcal{Ba}(\mathcal{X})$, then every measure is determined by the continuous functions. Indeed, if μ and ν are measures on $\mathcal{X}$ for which $\int f \, d\mu = \int f \, d\nu$ for all $f \in C(\mathcal{X})$, then $\mu = \nu$ by Theorem 4.4.2. This implies that the topologies in Definitions 3.3.6 are Hausdorff topologies whenever $\mathcal{B}(\mathcal{X}) = \mathcal{Ba}(\mathcal{X})$. If $\mathcal{X}$ is perfectly normal, then $\mathcal{B}(\mathcal{X}) = \mathcal{Ba}(\mathcal{X})$ (see [10, Proposition 6.3.4]). All metric spaces are perfectly normal (see for example [10, Proposition 6.3.5]). On the contrary, the Sorgenfrey line, i.e., the space $\mathbb{R}$ with the right half-open interval topology, is perfectly normal but not second countable (see Steen and Seebach [86, Example 51]) and whence not metrisable.

3.3.9 Example. The sequence $(\mu_n)_{n \in \mathbb{N}}$ of compactly supported probability measures on $\mathbb{R}$, with

$$\mu_n = \frac{1}{n}\delta_n + \frac{n-1}{n}\delta_0, \quad (3.11)$$

converges weakly to δ_0 but not C-weakly, as, for $f \in C(\mathbb{R})$ given by $f(x) = x$, $\int f d\mu_n = 1$ while $\int f d\delta_0 = 0$.

3.3.10 Theorem. Let $(\mathcal{X}, d)$ be a σ -compact metric space. Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{P}_c(\mathcal{X})$ and let $\mu \in \mathcal{P}_c(\mathcal{X})$. Then $\mu_n \xrightarrow{C-w} \mu$ if and only if $\mu_n \xrightarrow{w} \mu$ (or $\mu_n \xrightarrow{v} \mu$) and $\bigcup_{n \in \mathbb{N}} \text{supp } \mu_n$ is relatively compact.

Proof. The proof of the “if” part is left for the reader.

We proof the “only if” part. Suppose that $\mathcal{X} = \bigcup_{n \in \mathbb{N}} K_n$ for an increasing sequence of compact sets $K_1 \subset K_2 \subset \dots$. Suppose that $\bigcup_{n \in \mathbb{N}} \text{supp } \mu_n$ is not relatively compact. We show that $(\mu_n)_{n \in \mathbb{N}}$ is not a C-weakly convergent sequence. For all compact K there exists an $m \in \mathbb{N}$ for which $\text{supp } \mu_m \not\subset K$. Then for all $n \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ with $\mu_m(K_n^c) > 0$, and, since $K_k \uparrow \mathcal{X}$, there also exists an $k \in \mathbb{N}$ with $k > m$ for which $\mu_m(K_k \setminus K_n) > 0$. Choose $n_1 < n_2 < n_3 < \dots$ and $m_1 < m_2 < m_3 < \dots$ in $\mathbb{N}$ such that $\mu_{m_i}(K_{n_{i+1}} \setminus K_{n_i}) > 0$ for all $i \in \mathbb{N}$.

For the moment fix $i \in \mathbb{N}$. We show there exists a continuous function g_i with support in $K_{n_{i+1}} \setminus K_{n_i}$ for which $\int g_i d\mu_{n_i} = i$. For $k \in \mathbb{N}$ let $f_k(x) = (kd(x, K_{n_{i+1}}^c \cup K_{n_i})) \wedge 1$ for $x \in \mathcal{X}$, where d denotes the metric on $\mathcal{X}$. Then $\text{supp } f_k \subset K_{n_{i+1}} \setminus K_{n_i}$ for all $k \in \mathbb{N}$. Hence $f_k \uparrow \mathbb{1}_{K_{n_{i+1}} \setminus K_{n_i}}$ and therefore there exists a $k \in \mathbb{N}$ for which $\int f_k d\mu_i > 0$. This proves the existence of such g_i .

Define $g = \sum_{i \in \mathbb{N}} g_i$. Then g is a continuous function for which $\int g d\mu_{n_i} \geq i$ for all $i \in \mathbb{N}$. Hence $\int g d\mu_{n_i}$ does not converge and therefore for all $\mu \in \mathcal{P}_c(\mathcal{X})$ the sequence $(\mu_n)_{n \in \mathbb{N}}$ does not converge C-weakly to μ . $\square$

3.3.11 Remark. By the proof of Theorem 3.3.10 we can also infer that $\mu_n \xrightarrow{C-w} \mu$ if and only if $\bigcup_{n \in \mathbb{N}} \text{supp } \mu_n$ is relatively compact and $\int f d\mu_n \rightarrow \int f d\mu$ for all locally Lipschitz functions f .

3.3.12 Theorem. Let $\mathcal{X}$ be normal. Then $\mathcal{P}_{emp}^n(\mathcal{X})$ is a closed set in $\mathcal{P}_c(\mathcal{X})$. Moreover, if ξ_k and ξ in $\mathcal{P}_{emp}^n(\mathcal{X})$ are such that $\xi_k \xrightarrow{C-w} \xi$, then there exist x_{ki} and z_i in $\mathcal{X}$ for $i \in \{1, \dots, n\}$ such that $\xi_k = L_n((x_{k1}, \dots, x_{kn}))$, $\xi = L_n((z_1, \dots, z_n))$ and $x_{ki} \rightarrow z_i$ for all $i \in \{1, \dots, n\}$. Consequently, $L_n(F)$ is closed in $\mathcal{P}_c(\mathcal{X})$ for every closed set $F \subset \mathcal{X}^n$.

Proof. Let $(v_k)_{k \in \mathbb{N}}$ be a sequence in $\mathcal{X}^n$ such that $L_n(v_k) \xrightarrow{C-w} \xi$ for some $\xi \in \mathcal{P}_c(\mathcal{X})$. We prove $\xi = L_n(v)$ for some $v \in \mathcal{X}^n$. By Theorem 3.3.10 there exists a compact set $K \subset \mathcal{X}$ such that $v_k \in K^n$ for $k \in \mathbb{N}$. Therefore $(v_k)_{k \in \mathbb{N}}$ has a convergent subsequence with a limit $v = (y_1, \dots, y_n) \in \mathcal{X}^n$, so that $\xi = L_n(v)$. As $\mathcal{X}$ is assumed to be normal, the elements of $\{y_i : i \in \{1, \dots, n\}\}$ can be separated by continuous functions. Therefore one can “select the coordinates” of v_k in such way that they converge. $\square$

3.3.13 Definition. Let $\mathcal{X}$ be a topological space and $D_1, D_2, \dots$ be subsets of $\mathcal{X}$. We call $x \in \mathcal{X}$ a $(D_n)_{n \in \mathbb{N}}$ -limit point whenever there exist $x_n \in D_n$ with $x_n \rightarrow x$. We write $\text{LIM}[D_n]$ for the set of $(D_n)_{n \in \mathbb{N}}$ -limit points of $\mathcal{X}$.

3.3.14 Example. Let $\mathcal{X} = [-1, 1]$ and $D_n = \{-1, -1 + \frac{2}{n}, -1 + \frac{4}{n}, \dots, 1 - \frac{2}{n}, 1\}$. Then $\mathcal{X} = \text{LIM } D_n$.

In Theorem 3.3.16 we state that $\text{LIM}[\mathcal{P}_{emp}^n(\mathcal{X})] = \mathcal{P}_c(\mathcal{X})$, whenever $\mathcal{X}$ is a σ -compact metric space. This follows from Lemma 3.3.15. For the proof of Lemma 3.3.15 we use Lemma 4.8.6, which states that $\text{LIM}[\mathcal{P}_{emp}^n(\mathcal{X})] = \mathcal{P}(\mathcal{X})$ if $\mathcal{X}$ is finite.

3.3.15 Lemma. Let $(\mathcal{X}, d)$ be a compact metric space. Then $\mathcal{P}(\mathcal{X}) = \text{LIM}[\mathcal{P}_{emp}^n(\mathcal{X})]$ (equipped with the weak topology). In particular, $\bigcup_{n \in \mathbb{N}} \mathcal{P}_{emp}^n(\mathcal{X})$ is dense in $\mathcal{P}(\mathcal{X})$.

Proof. For $m \in \mathbb{N}$, let $k_m \in \mathbb{N}$ and let $x_1^m, x_2^m, \dots, x_{k_m}^m \in \mathcal{X}$ be such that $\mathcal{X} = \bigcup_{i=1}^{k_m} B(x_i^m, \frac{1}{m})$. Let $A_1^m, \dots, A_{k_m}^m$ be a partition of $\mathcal{X}$ in $\mathcal{B}(\mathcal{X})$ with $A_i^m \subset B(x_i^m, \frac{1}{m})$ (for example, $A_i^m = B(x_i^m, \frac{1}{m}) \setminus [\bigcup_{j=1}^{i-1} B(x_j^m, \frac{1}{m})]$). For $f \in C(\mathcal{X})$, define $S_{f,m} : \mathcal{X} \rightarrow \mathbb{R}$ by

$$S_{f,m} = \sum_{i=1}^{k_m} f(x_i^m) \mathbb{1}_{A_i^m}. \quad (3.12)$$

As $\sup\{|f(x) - f(y)| : x, y \in \mathcal{X}, d(x, y) < \frac{1}{m}\} \rightarrow 0$, we have $\sup\{|S_{f,m}(x) - f(x)| : x \in \mathcal{X}\} \rightarrow 0$. Let $\nu \in \mathcal{P}(\mathcal{X})$. Define $\rho_m := \sum_{i=1}^{k_m} \nu(A_i^m) \delta_{x_i^m}$. Then $\int f d\rho_m = \int S_{f,m} d\nu \rightarrow \int f d\nu$ for all $f \in C(\mathcal{X})$, i.e., $\rho_m \xrightarrow{w} \nu$. Because by Lemma 4.8.6

$$\begin{aligned} \text{LIM} \left[\left\{ \sum_{i=1}^k q_i \delta_{x_i} : q_i \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}, \sum_{i=1}^n q_i = 1 \right\} \right] \\ = \left\{ \sum_{i=1}^n \alpha_i \delta_{x_i} : \alpha_i \in [0, 1], \sum_{i=1}^n \alpha_i = 1 \right\}. \end{aligned} \quad (3.13)$$

Therefore, for each m there exists a sequence $(\xi_n^m)_{n \in \mathbb{N}}$ with $\xi_n^m \in \mathcal{P}_{emp}^n(\mathcal{X})$ such that $\xi_n^m \xrightarrow{w} \rho_m$ as $n \rightarrow \infty$.

Let $\mathfrak{d}$ denote the Prohorov metric on $\mathcal{P}(\mathcal{X})$ (see [7, Appendix III]). Let $n_1 < n_2 < n_3 < \dots$ in $\mathbb{N}$ be such that $\mathfrak{d}(\xi_k^m, \rho_k) < \frac{1}{k}$ for all $m > n_k$ and for each $k \in \mathbb{N}$. Let $\nu_n = \xi_1^n$ for $n \leq n_2$ and $\nu_n = \xi_k^n$ whenever $n_k < n \leq n_{k+1}$ for some $k \geq 2$, i.e., $(\nu_n)_{n \in \mathbb{N}}$ is the sequence

$$\begin{aligned} \xi_1^1, \xi_1^2, \dots, \xi_1^{n_2}, \xi_2^{n_2+1}, \xi_2^{n_2+2}, \dots, \xi_2^{n_3}, \\ \xi_3^{n_3+1}, \xi_3^{n_3+2}, \dots, \xi_3^{n_4}, \xi_4^{n_4+1}, \xi_4^{n_4+2}, \dots, \xi_4^{n_5}, \dots \end{aligned} \quad (3.14)$$

Then $\mathfrak{d}(\nu_n, \nu) \leq \mathfrak{d}(\nu_n, \rho_k) + \mathfrak{d}(\rho_k, \nu) \leq \frac{1}{k} + \mathfrak{d}(\rho_k, \nu)$ for $n_k < n \leq n_{k+1}$, and whence $\nu_n \xrightarrow{w} \nu$. $\square$

3.3.16 Theorem. *Let $(\mathcal{X}, d)$ be a σ -compact metric space. Then*

$$\mathcal{P}_c(\mathcal{X}) = \text{LIM}[\mathcal{P}_{emp}^n(\mathcal{X})] \quad (3.15)$$

(equipped with the C -weak topology). In particular, the set $\bigcup_{n \in \mathbb{N}} \mathcal{P}_{emp}^n(\mathcal{X})$ is dense in $\mathcal{P}_c(\mathcal{X})$.

Proof. This is a consequence of Theorem 3.3.10 and Theorem 3.3.12. $\square$

§3.3.3 Supports of measures

In this section we discuss some relations between absolute continuity of measures and their supports.

3.3.17 Lemma. *Let ν, μ be measures on $\mathcal{B}(\mathcal{X})$ with $\nu \ll \mu$. Then*

- (a) $\text{supp}(\mu) \subset \text{supp}(\nu)$.
- (b) $f\nu \ll f\mu$ and $\text{supp}(f\mu) \subset \text{supp}(f\nu)$ for measurable $f : \mathcal{X} \rightarrow [0, \infty)$.
- (c) $\nu \rtimes P \ll \mu \rtimes P$ and $\text{supp}(\mu \rtimes P) \subset \text{supp}(\nu \rtimes P)$ for kernels $P : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, \infty)$.
(See (3.4) for the notation $\nu \rtimes P$.)
- (d) $\nu \circ \sigma^{-1} \ll \mu \circ \sigma^{-1}$ and $\text{supp}(\mu \circ \sigma^{-1}) \subset \text{supp}(\nu \circ \sigma^{-1})$ for measurable $\sigma : \mathcal{X} \rightarrow \mathcal{Y}$.
- (e) $\text{supp}(\nu \otimes \rho) = \text{supp}(\nu) \times \text{supp}(\rho)$ for measures ρ on $\mathcal{B}(\mathcal{Y})$.

Proof. (a) Follows from the fact that

$$\text{supp } \mu = \bigcap \{C \subset \mathcal{X} : C \text{ is closed and } \mu(\mathcal{X} \setminus C) = 0\}. \quad (3.16)$$

(b) Follows from the Radon-Nikodym theorem.

(c) Follows from (b) as the measure $[\nu \rtimes P](\cdot \times B)$ equals $P(\cdot, B)\nu$. $\square$

§3.4 Sequentially Gibbs

We present the notion of sequentially Gibbs for sequences of exchangeable probability measures (Definition 3.4.1). We give an equivalent formulation for sequentially Gibbs in terms of a kernel that is indexed by the empirical measures. We show that under sequential Gibbsianness the limiting kernel is continuous (Lemma 3.4.5). Then we give an equivalent notion of sequential Gibbsianness (Theorem 3.4.7) whenever there is f -dependence (Definition 3.4.6), and conclude from this that the notion of sequentially Gibbs agrees with Gibbsianness for mean-field models in other studies (Corollary 3.4.9 and 3.4.10).

In this section $\mathcal{X}$ is a perfectly normal space (see Definition 3.3.7). For $n \in \mathbb{N}$, ρ_n is an exchangeable probability measure on $\mathcal{B}(\mathcal{X}^n)$ (see Definition 3.3.4). For all $n \in \mathbb{N}$, we use σ_1 for the projection of $\mathcal{X}^{n+1}$ on the first coordinate, and $\sigma_\perp$ for the projection of $\mathcal{X}^{n+1}$ on all but the first coordinate for all n , i.e.,

$$\sigma_1(x) = x_1, \quad \sigma_\perp(x) = (x_2, \dots, x_{n+1}) \quad \left(x = (x_1, \dots, x_{n+1}) \in \mathcal{X}^{n+1} \right). \quad (3.17)$$

We assume that for every $n \in \mathbb{N}$ there exists a product regular conditional kernel² $\gamma_n: \mathcal{X}^n \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under ρ_{n+1} with respect to $\sigma_\perp$ that is continuous on $\text{supp}(\rho_n \circ \sigma_\perp^{-1})$.

3.4.1 Definition. (Good limiting empirical distributions and sequentially Gibbs). A $\zeta \in \mathcal{P}_c(\mathcal{X})$ is called a *good limiting empirical distribution* for $(\rho_n)_{n \in \mathbb{N}}$ when there exists a probability measure $\gamma_\zeta: \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ for which $\gamma_n(v_n, \cdot) \xrightarrow{w} \gamma_\zeta$ for all sequences $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \text{supp}(\rho_n \circ \sigma_\perp^{-1}) \subset \mathcal{X}^n$ and $L_n(v_n) \xrightarrow{C-w} \zeta$. $(\rho_n)_{n \in \mathbb{N}}$ is called *sequentially Gibbs* when all $\zeta \in \text{LIM}[L_n(\text{supp}(\rho_n \circ \sigma_\perp^{-1}))]$ are good limiting empirical distributions.

3.4.2 Remarks. (1) The assumption that $\mathcal{X}$ is perfectly normal implies that for all n the kernel γ_n is unique in the sense of Theorem 4.4.3. This makes the notion of sequentially Gibbs independent of the “choice” of γ_n .

(2) Definition 3.4.1 is consistent with the definition for compact spaces given by Külske and Opoku [61, Definition 2.2], since for compact $\mathcal{X}$ a sequence in $\mathcal{P}(\mathcal{X})$ converges weakly if and only if it converges C -weakly (Theorem 3.3.10).

(3) The convergence $L_n(v_n) \xrightarrow{C-w} \zeta$ in Definition 3.4.1 may seem arbitrary. In some sense one could as well have taken convergence along a subset of continuous functions. The choice for $C(\mathcal{X})$ instead of, for example, $C_b(\mathcal{X})$ will become clear in Corollary 3.4.9 as, for example, the function $f: \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x$ is in $C(\mathbb{R})$ but not in $C_b(\mathbb{R})$.

3.4.3. Since ρ_n is exchangeable, $\gamma_n(v, \cdot) = \gamma_n(u, \cdot)$ for all $u, v \in \mathcal{X}^n$ with $L_n(u) = L_n(v)$. Then $(v, A) \mapsto \gamma_n(\hat{\sigma}(v), A)$ is also a product regular conditional kernel $\mathcal{X}^n \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under ρ_{n+1} with respect to $\sigma_\perp$ that is continuous on $\text{supp}(\rho_n \circ \sigma_\perp^{-1})$. Hence we have $\gamma_n(\hat{\sigma}(v), \cdot) = \gamma_n(v, \cdot)$ by Theorem 4.4.3. Therefore we can define a kernel $\tilde{\gamma}_n: \mathcal{P}_{emp}^n(\mathcal{X}) \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ by

$$\tilde{\gamma}_n(\xi, A) = \gamma_n(v, A), \quad \text{where } v \in \mathcal{X}^n \text{ is such that } L_n(v) = \xi. \quad (3.18)$$

With this we can reformulate the notion of sequentially Gibbs.

3.4.4 Lemma. $\zeta \in \mathcal{P}_c(\mathcal{X})$ is a good limiting empirical distribution for $(\rho_n)_{n \in \mathbb{N}}$ if and only if there exists a probability measure $\gamma_\zeta: \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ such that $\tilde{\gamma}_n(\zeta_n, \cdot) \xrightarrow{w} \gamma_\zeta$ for all sequences $(\zeta_n)_{n \in \mathbb{N}}$ with $\zeta_n \in \mathcal{P}_{emp}^n(\mathcal{X})$ and $\zeta_n \xrightarrow{C-w} \zeta$.

The following lemma is similar to Lemma 2.1.3, but we use a generalised version of Lemma 2.2.1 for the proof, namely, Theorem 3.A.1.

3.4.5 Lemma. Let $(\rho_n)_{n \in \mathbb{N}}$ be sequentially Gibbs. Define $\gamma: \mathcal{P}_c(\mathcal{X}) \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ by $\gamma(\zeta, A) = \gamma_\zeta(A)$ for $\zeta \in \mathcal{P}(\mathcal{X})$ and $A \in \mathcal{B}(\mathcal{X})$, where γ_ζ is as in Definition 3.4.1. Then γ is a weakly continuous probability kernel, called the specification kernel for $(\rho_n)_{n \in \mathbb{N}}$.

²For the definition of a product regular conditional kernel see Definition 4.3.2.

Proof. We apply Theorem 3.A.1. Let $D_n = \mathcal{P}_{emp}^{2^n}(\mathcal{X})$. Then $D_1 \subset D_2 \subset \dots$. For all $f \in C_b(\mathcal{X})$ and $\zeta \in \mathcal{P}_c(\mathcal{X})$ we have, by Lemma 3.4.4 and Theorem 3.3.16, that $\int f d[\gamma_{2^n}(\zeta_n, \cdot)] \rightarrow \int f d[\gamma(\zeta, \cdot)]$ for all sequences $(\zeta_n)_{n \in \mathbb{N}}$ with $\zeta_n \in D_n$ and $\zeta_n \rightarrow \zeta$. By Theorem 3.A.1 $\zeta \mapsto \int f d[\gamma(\zeta, \cdot)]$ is continuous on $\mathcal{P}_c(\mathcal{X})$. Therefore γ is weakly continuous. $\square$

3.4.6 Definition. Let $f \in C(\mathcal{X}, \mathbb{R}^k)$ for some $k \in \mathbb{N}$. We say that $(\gamma_n)_{n \in \mathbb{N}}$ is *f-dependent* if for all $n \in \mathbb{N}$

$$\gamma_n(u, \cdot) = \gamma_n(v, \cdot) \quad \text{for all } u, v \in \mathcal{X}^n \text{ with } \int f d[L_n(u)] = \int f d[L_n(v)]. \quad (3.19)$$

Whenever $\mathcal{X}$ is a convex subset of $\mathbb{R}$, the sequence $(\gamma_n)_{n \in \mathbb{N}}$ is said to be *magnetisation dependent* if γ_n is *f-dependent* for $f(x) = x$. The *magnetisation* of a vector $v \in \mathcal{X}^n$ is given by $m_n(v) = \int f d[L_n(v)] = \frac{1}{n} \sum_{i=1}^n v_i$.

3.4.7 Theorem. Let $f \in C(\mathcal{X}, \mathbb{R}^k)$ for some $k \in \mathbb{N}$. Then

$$\begin{aligned} & \left\{ \int f d\zeta : \zeta \in \text{LIM}[L_n(\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}))] \right\} \\ & \subset \text{LIM} \left[\left\{ \int f d[L_n(v)] : v \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}) \right\} \right]. \end{aligned} \quad (3.20)$$

Consider the following statements.

- (a) $(\rho_n)_{n \in \mathbb{N}}$ is sequentially Gibbs.
- (b) For all $\alpha \in \text{LIM}[\{\int f d[L_n(v)] : v \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})\}]$ there exists a probability measure $\tilde{\gamma}_\alpha : \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ for which $\gamma_n(v_n, \cdot) \xrightarrow{w} \tilde{\gamma}_\alpha$ for all sequences $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})$ and $\int f d[L_n(v_n)] \rightarrow \alpha$.
- (c) For all $\alpha \in \text{LIM}[\{\int f d[L_n(v)] : v \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})\}]$ there exists a $\zeta \in \text{LIM}[L_n(\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}))]$ such that for all $v_n \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})$ with $\int f d[L_n(v_n)] \rightarrow \alpha$ there exist $w_n \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})$ such that

$$\int f d[L_n(w_n)] = \int f d[L_n(v_n)] \quad (3.21)$$

and $L_n(w_n) \xrightarrow{C-w} \zeta$.

Then (b) $\implies$ (a) and (a) $\mathcal{E}$ (c) $\implies$ (b).

Proof. The inclusion in (3.20) is obvious. Suppose that (b) holds. Let

$$\zeta \in \text{LIM}[L_n(\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}))]. \quad (3.22)$$

Suppose that $(v_n)_{n \in \mathbb{N}}$ is a sequence with $(v_n)_{n \in \mathbb{N}}$ with

$$v_n \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}), \quad L_n(v_n) \xrightarrow{C-w} \zeta. \quad (3.23)$$

Then $\int f d[L_n(v_n)] \rightarrow \int f d\zeta$ and thus $\gamma_n(v_n, \cdot) \xrightarrow{w} \tilde{\gamma}_\alpha$, where $\alpha = \int f d\zeta$.

Suppose that (a) and (c) hold. Let $\alpha \in \text{LIM}[\{ \int f \, d[L_n(v)] : v \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}) \}]$ and let ζ be as in (c). Suppose that $(v_n)_{n \in \mathbb{N}}$ is a sequence with

$$v_n \in \text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}), \quad \int f \, d[L_n(v_n)] \rightarrow \alpha. \quad (3.24)$$

Then for w_n as in (c), by (3.21) we have $\gamma_n(v_n, \cdot) = \gamma_n(w_n, \cdot) \rightarrow \gamma_\zeta$. $\square$

3.4.8. If $\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}) = \mathcal{X}^n$, then $L_n(\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1})) = \mathcal{P}_{emp}^n(\mathcal{X})$, and thus $\text{LIM}[L_n(\text{supp}(\rho_{n+1} \circ \sigma_\perp^{-1}))] = \mathcal{P}_c(\mathcal{X})$ by Theorem 3.3.16.

3.4.9 Corollary. *Let $\mathcal{X} \subset \mathbb{R}$ be a convex set or $\mathcal{X} = \{-1, 1\}$. Suppose that $(\gamma_n)_{n \in \mathbb{N}}$ is magnetisation dependent and that $\text{supp}(\rho_{n+1} \circ \sigma_1^{-1}) = \mathcal{X}$ for all $n \in \mathbb{N}$. Then $(\rho_n)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if for all $\alpha \in \mathcal{X}$ there exists a $\tilde{\gamma}_\alpha : \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ for which $\gamma_n(v_n, \cdot) \xrightarrow{w} \tilde{\gamma}_\alpha$ for all sequences $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \mathcal{X}^n$ and $m_n(v_n) \rightarrow \alpha$.*

Proof. We show that (c) of Theorem 3.4.7 holds.

First assume that $\mathcal{X}$ is convex. Let $\alpha \in \mathcal{X}$. Let $v_n \in \mathcal{X}^n$ be such that $m_n(v_n) \rightarrow \alpha$. For $w_n = (m_n(v_n), \dots, m_n(v_n))$ we have $L_n(w_n) = \delta_{m_n(v_n)} \xrightarrow{C-w} \delta_\alpha$.

Suppose now that $\mathcal{X} = \{-1, 1\}$. Then $m_n(v)$ determines the number of i 's for which $v_i = 1$ by $\#\{i \in \{1, \dots, n\} : v_i = 1\} = \frac{1}{2}(n + nm_n(v))$. Hence

$$L_n(v) = \frac{1}{2}(1 + m_n(v))\delta_{\{1\}} + \frac{1}{2}(1 - m_n(v))\delta_{\{-1\}}. \quad (3.25)$$

Therefore, if $m_n(v_n) \rightarrow \alpha$, then $L_n(v_n) \rightarrow \frac{1}{2}(1 + \alpha)\delta_{\{1\}} + \frac{1}{2}(1 - \alpha)\delta_{\{-1\}}$. $\square$

3.4.10. Corollary 3.4.9 implies that, whenever $\text{supp}(\rho_{n+1} \circ \sigma_1^{-1}) = \mathcal{X}$, the notion of sequentially Gibbs agrees with the notion of sequentially Gibbs in den Hollander, Redig and van Zuijlen [52, Definition 1.2] and with the notion of Gibbs as in Külske and le Ny [60] (used also in Ermolaev and Külske [36], Fernández, den Hollander and Martínez [39]).

§3.5 Sequences of finite-volume mean-field Gibbs measures

In this section we define sequences of finite-volume mean-field Gibbs measures (in Definition 3.5.1) and study under which conditions on the Hamiltonians, in particular, under which conditions on the potential, such sequences are sequentially Gibbs. Whenever the Hamiltonians originate from a potential V , a differentiability property, called continuously $\star$ -differentiability, together with a certain bound, called the $\star$ -dominated bound on V , imply sequential Gibbsianness (see Theorem 3.5.18). Furthermore, we provide conditions under which a potential V satisfies the differentiability and the bound (in Theorem 3.5.11 and Theorem 3.5.12). We relate these properties of the potential to Hamiltonians that have been considered in the Gibbs-non-Gibbs literature (in Example 3.5.15).

In Section 3.5.1 we give the definition of sequences of finite-volume mean-field Gibbs measures. In Section 3.5.2 we introduce the notions $\star$ -differentiability and the $\star$ -dominated bound. We devote Section 3.5.3 to a discussion on the situation where the Hamiltonians do not originate from a potential.

Throughout this section $\mathcal{X}$ is a perfectly normal space and λ is a measure such that $\text{supp}(\lambda) = \mathcal{X}$.

§3.5.1 Sequences of finite-volume mean-field Gibbs measures

3.5.1 Definitions. For $n \in \mathbb{N}$, let ν_n be a probability measure on $\mathcal{B}(\mathcal{X}^n)$. We call a measurable function $H_n : \mathcal{X}^n \rightarrow \mathbb{R}$ a *Hamiltonian* (for volume n with respect to λ) if

$$Z_n := \int_{\mathcal{X}^n} e^{-H_n(v)} d\lambda^n(v) < \infty. \quad (3.26)$$

The sequence $(\nu_n)_{n \in \mathbb{N}}$ is called a *sequence of finite-volume mean-field Gibbs measures* with *Hamiltonians* $(H_n)_{n \in \mathbb{N}}$ (and *reference measure* λ) if

$$\nu_n(A) = \frac{1}{Z_n} \int_{\mathcal{X}^n} \mathbb{1}_A(v) e^{-H_n(v)} d\lambda^n(v) \quad (A \in \mathcal{B}(\mathbb{R}^n), n \in \mathbb{N}). \quad (3.27)$$

Z_n is called the *normalising constant*. A function $V : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty)$ for which $V \circ L_n$ is a Hamiltonian for all $n \in \mathbb{N}$ is called a *potential* (with respect to λ). If V is a potential, then a sequence of Hamiltonians $(H_n)_{n \in \mathbb{N}}$ is said to *originate from the potential* V if $H_n = nV \circ L_n$ for all $n \in \mathbb{N}$. In that case the sequence $(\nu_n)_{n \in \mathbb{N}}$ is called a *sequence of finite-volume mean-field Gibbs measures with potential* V .

3.5.2 Remark. Note that if λ is a probability measure and $H_n : \mathcal{X}^n \rightarrow \mathbb{R}$ is measurable and bounded from below, then H_n is a Hamiltonian with respect to λ .

3.5.3. 3.3.5 indicates that sequences of exchangeable probability measures $(\nu_n)_{n \in \mathbb{N}}$ with $\nu_n \ll \lambda^n$ and $\lambda^n \ll \nu_n$, are sequences of finite-volume mean-field Gibbs measures.

3.5.4 Example (Regarding Remark 3.5.2). Consider the sequence of probability measures $(\nu_n)_{n \in \mathbb{N}}$ on $\mathbb{R}$ as in Definition 2.1.1 given by

$$\nu_n(A) = \frac{1}{Z_n} \int_{\mathbb{R}^n} \mathbb{1}_A(v) e^{-nF \circ m_n(v)} d[\mu_{\mathcal{N}(0,1)}]^n(v), \quad (3.28)$$

where $F : \mathbb{R} \rightarrow [0, \infty)$ is measurable, $m_n(v)$ is the magnetisation of v (see Definition 3.4.6) and $\mu_{\mathcal{N}(0,1)}$ denotes the standard normal distribution. Then $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures with Hamiltonians $(H_n^1)_{n \in \mathbb{N}}$ and reference measure $\mu_{\mathcal{N}(0,1)}$, but also with Hamiltonians $(H_n^2)_{n \in \mathbb{N}}$ and reference measure the Lebesgue measure, where

$$H_n^1(v) = nF(m_n(v)) \quad (v \in \mathbb{R}^n), \quad (3.29)$$

$$H_n^2(v) = nF(m_n(v)) + \sum_{i=1}^n v_i^2/2 \quad (v \in \mathbb{R}^n). \quad (3.30)$$

Note that $(H_n^1)_{n \in \mathbb{N}}$ and $(H_n^1)_{n \in \mathbb{N}}$ originate from potentials V^1 and V^2 , where

$$V^1(\zeta) = F\left(\int x \, d\zeta(x)\right), \quad V^2(\zeta) = V^1(\zeta) + \int x^2 \, d\zeta(x). \quad (3.31)$$

In Example 3.5.15 we will also discuss the Curie-Weiss model.

3.5.5. As $\text{supp}(\lambda) = \mathcal{X}$ and $e^{-H_n} > 0$, the support of ν_n equals that of λ^n (see Lemma 3.3.17), i.e., $\text{supp}(\nu_n) = \mathcal{X}^n$. Therefore $\text{supp}(\nu_{n+1} \circ \sigma_\perp^{-1}) = \mathcal{X}^n$.

3.5.6. Assume that H_n are continuous Hamiltonians for all $n \in \mathbb{N}$. Observe that, by Theorem 4.4.7, $\gamma_n : \mathcal{X}^n \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ given by

$$\gamma_n(v, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-H_{n+1}(x, v)} \, d\lambda(x)}{\int_{\mathcal{X}} e^{-H_{n+1}(x, v)} \, d\lambda(x)}, \quad (3.32)$$

is the³ weakly continuous regular conditional probability under ν_{n+1} with respect to $\sigma_\perp$. γ_n can also be written as

$$\gamma_n(v, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-[H_{n+1}(x, v) - \frac{n+1}{n} H_n(v)]} \, d\lambda(x)}{\int_{\mathcal{X}} e^{-[H_{n+1}(x, v) - \frac{n+1}{n} H_n(v)]} \, d\lambda(x)}. \quad (3.33)$$

Whenever $(H_n)_{n \in \mathbb{N}}$ originates from a potential V , then

$$\gamma_n(v, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-(n+1)[V \circ L_{n+1}(x, v) - V \circ L_n(v)]} \, d\lambda(x)}{\int_{\mathcal{X}} e^{-(n+1)[V \circ L_{n+1}(x, v) - V \circ L_n(v)]} \, d\lambda(x)}.$$

Note that the term in the exponent can be written as

$$\begin{aligned} (n+1)[V \circ L_{n+1}(x, v) - V \circ L_n(v)] \\ = \frac{V(\frac{n}{n+1} L_n(v) + \frac{1}{n+1} [\delta_x - L_n(v)]) - V(L_n(v))}{\frac{1}{n+1}}. \end{aligned} \quad (3.34)$$

By this formula we can already suspect that a sort of differentiability property on V may imply that $\gamma_n(v_n, \cdot)$ converges when $L_n(v_n) \xrightarrow{C-w} \zeta$ as $n \rightarrow \infty$, i.e., that $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs.

§3.5.2 Conditions on the potential implying sequential Gibbsianness

In Theorem 3.5.18 we will see that a sequence of finite-volume mean-field Gibbs measures with potential V is sequentially Gibbs whenever it is $\star$ -differentiable and satisfies the $\star$ -dominated bound.

³Because $\text{supp}(\nu_n) = \mathcal{X}^n$ (see 3.5.5), γ_n is the unique regular conditional probability under ν_{n+1} with respect to $\sigma_\perp$ by Theorem 4.4.3.

3.5.7 Definition. Let $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$. We call V $\star$ -differentiable if for all $\xi, \theta \in \mathcal{P}_c(\mathcal{X})$ the function $[0, 1] \rightarrow \mathbb{R}, \alpha \mapsto V(\xi + \alpha(\theta - \xi))$ is continuous on $[0, 1]$, differentiable on $(0, 1)$ and the right derivative exists at 0. If V is $\star$ -differentiable, then we write $\delta V : \mathcal{P}_c(\mathcal{X}) \times \mathcal{P}_c(\mathcal{X}) \times [0, 1] \rightarrow \mathbb{R}$ for the function given by

$$\begin{aligned}\delta V(\xi, \theta, \alpha) &= \lim_{\substack{h \rightarrow 0 \\ h + \alpha \in [0, 1]}} \frac{V(\xi + (\alpha + h)(\theta - \xi)) - V(\xi + \alpha(\theta - \xi))}{h} & \alpha \in (0, 1), \\ \delta V(\xi, \theta, 0) &= \lim_{\substack{h \downarrow 0 \\ h \in [0, 1]}} \frac{V(\xi + h(\theta - \xi)) - V(\xi)}{h} \\ \delta V(\xi, \theta, 1) &= \delta V(\theta, \xi, 0)\end{aligned}\tag{3.35}$$

If the function $\delta V(\cdot, \theta, \cdot) : \mathcal{P}_c(\mathcal{X}) \times [0, 1] \rightarrow \mathbb{R}$ is continuous for all $\theta \in \mathcal{P}_c(\mathcal{X})$ in the sense that $\xi_n \xrightarrow{C-w} \xi$ and $\alpha_n \rightarrow \alpha$ imply that $\delta V(\xi_n, \theta, \alpha_n) \rightarrow \delta V(\xi, \theta, \alpha)$, then V is called *continuously $\star$ -differentiable*.

3.5.8 Lemma. Let $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ be continuously $\star$ -differentiable. For all $\xi \in \mathcal{P}_c(\mathcal{X})$ and $x \in \mathcal{X}$ there exists an $\alpha \in (0, \frac{1}{n+1})$ such that

$$(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)] = \delta V(\xi, \delta_x, \alpha).\tag{3.36}$$

Consequently, for $\zeta \in \mathcal{P}_c(\mathcal{X})$ and $x \in \mathcal{X}$ and $(\zeta_n)_{n \in \mathbb{N}}$ in $\mathcal{P}_c(\mathcal{X})$ with $\zeta_n \rightarrow \zeta$ in $\mathcal{P}_c(\mathcal{X})$

$$(n+1)[V(\zeta_n + \frac{1}{n+1}[\delta_x - \zeta_n]) - V(\zeta_n)] \rightarrow \delta V(\zeta, \delta_x).\tag{3.37}$$

Proof. For $\xi \in \mathcal{P}_c(\mathcal{X})$ and $x \in \mathcal{X}$, the function $[0, \frac{1}{n+1}] \rightarrow [0, \infty), \alpha \mapsto V(\xi + \alpha[\delta_x - \xi])$, is continuous and differentiable on $(0, \frac{1}{n+1})$. By the mean value theorem there exists an $\alpha \in (0, \frac{1}{n+1})$ for which (3.36). $\square$

3.5.9 Definition. Let $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ be $\star$ -differentiable. We say that V satisfies the $\star$ -dominated bound whenever for all compact $K \subset \mathcal{P}_c(\mathcal{X})$ there exists a λ -integrable function $h : \mathcal{X} \rightarrow [0, \infty)$, an increasing sequence $(\mathcal{X}_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(\mathcal{X})$ with $\mathcal{X}_n \uparrow \mathcal{X}$ (i.e., $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \dots$ and $\bigcup_{n \in \mathbb{N}} \mathcal{X}_n = \mathcal{X}$) with $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ such that for all $n \in \mathbb{N}$

$$e^{-\delta V(\xi, \delta_x, \alpha)} \leq h(x) \quad (\xi \in K, x \in \mathcal{X}_n, \alpha \in [0, \frac{1}{n}]).\tag{3.38}$$

3.5.10. Suppose that $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ is continuously $\star$ -differentiable and satisfies the $\star$ -dominated bound. By Lemma 3.5.8 it follows that, if $K \subset \mathcal{P}_c(\mathcal{X})$ is compact, then there exist an increasing sequence $(\mathcal{X}_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(\mathcal{X})$ with $\mathcal{X}_n \uparrow \mathcal{X}$ and $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ and a λ -integrable function $h : \mathcal{X} \rightarrow [0, \infty)$ such that

$$e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} \leq h(x) \quad (\xi \in K, x \in \mathcal{X}_n, n \in \mathbb{N}).\tag{3.39}$$

In particular, whenever $\zeta \in \mathcal{P}_c(\mathcal{X})$ and $\zeta_n \in \mathcal{P}_{emp}^n(\mathcal{X})$ are such that $\zeta_n \xrightarrow{C-w} \zeta$ there exist such $(\mathcal{X}_n)_{n \in \mathbb{N}}$ and h such that

$$e^{-(n+1)[V(\zeta_n + \frac{1}{n+1}[\delta_x - \zeta_n]) - V(\zeta_n)]} \leq h(x) \quad (x \in \mathcal{X}_n, n \in \mathbb{N}).\tag{3.40}$$

3.5.11 Theorem. Suppose that $\mathcal{X}$ is a compact space. Suppose that $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ is continuously $\star$ -differentiable such that δV is continuous (as a function of three variables). Then V satisfies the $\star$ -dominated bound.

Proof. $\mathcal{P}(\mathcal{X}) = \mathcal{P}_c(\mathcal{X})$ is weakly compact, see Bogachev [10, Theorem 8.9.3] (as in a perfectly normal space all Borel measures are Baire measures, see 3.3.8). Therefore δV , being continuous, is bounded on $\mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$. $\square$

3.5.12 Theorem. Let $k \in \mathbb{N}$, $f \in C(\mathcal{X}, \mathbb{R}^k)$ and $F \in C^1(\mathbb{R}^k)$. Let $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ be given by

$$V(\xi) = F\left(\int f \, d\xi\right). \quad (3.41)$$

Then V is a continuously $\star$ -differentiable potential with

$$\delta V(\xi, \theta, \alpha) = \nabla F\left((1-\alpha) \int f \, d\xi + \alpha \int f \, d\theta\right) \left(\int f \, d\theta - \int f \, d\xi\right). \quad (3.42)$$

In particular,

$$\delta V(\xi, \theta) = \nabla F\left(\int f \, d\xi\right) \left(\int f \, d\theta - \int f \, d\xi\right). \quad (3.43)$$

Suppose that $x \mapsto e^{m|f(x)|_1}$ ($|a|_1 = |a_1| + \dots + |a_k|$ for $a \in \mathbb{R}^k$) is λ -integrable for all $m > 0$. If either $\sup_{v \in \mathbb{R}^k} |\nabla F(v)| < \infty$ or $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ for $\mathcal{X}_n = \{x \in \mathcal{X} : f(x) \in (-n, n)^k\}$, then V satisfies the $\star$ -dominated bound.

Proof. Let $\xi, \theta \in \mathcal{P}_c(\mathcal{X})$ and let $w, z \in \mathbb{R}^k$ be given by $w = \int f \, d\xi$ and $z = \int f \, d\theta$. Then

$$V(\xi + \alpha(\theta - \xi)) = F(w + \alpha(z - w)) \quad (\alpha \in [0, 1]). \quad (3.44)$$

As $F \in C^1(\mathbb{R}^k)$, $\alpha \mapsto V(\xi + \alpha(\theta - \xi))$ is continuous on $[0, 1]$ and differentiable on $(0, 1)$ and the right derivative exists at 0. We have for $\alpha \in [0, 1]$

$$\begin{aligned} \delta V(\xi, \theta, \alpha) &= \lim_{h \rightarrow 0} \frac{F(w + (\alpha + h)(z - w)) - F(w + \alpha(z - w))}{h} \\ &= \nabla F(w + \alpha(z - w))(z - w). \end{aligned} \quad (3.45)$$

This proves (3.42). Since $\xi \mapsto \int f \, d\xi$ is C-weakly continuous and ∇F is continuous by assumption, V is continuously $\star$ -differentiable.

Assume that $x \mapsto e^{m|f(x)|_1}$ is λ -integrable for all $m \in \mathbb{R}$. Let $K \subset \mathcal{P}_c(\mathcal{X})$ be compact. Then the set $\{\int f \, d\xi : \xi \in K\}$ is compact and so $L := \sup_{\xi \in K} |\int f \, d\xi| < \infty$.

• If $M := \sup_{v \in \mathbb{R}^k} |\nabla F(v)| < \infty$, then

$$e^{-\delta V(\xi, \delta_x, \alpha)} \leq e^{M|f(x)|_1 + ML} \quad (x \in \mathcal{X}), \quad (3.46)$$

which proves the $\star$ -dominated bound.

• Suppose that $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$. By definition of $\mathcal{X}_n$ we have $\frac{1}{n} f(x) \in [-1, 1]^k$ for all $x \in \mathcal{X}_n$. Therefore

$$\begin{aligned} & \left\{ (1-\alpha) \int f \, d\xi + \alpha f(x) : \xi \in K, x \in \mathcal{X}_n, \alpha \in [0, \tfrac{1}{n}] \right\} \\ & \subset \left\{ (1-\alpha) \int f \, d\xi + \alpha v : \xi \in K, v \in [-1, 1]^k, \alpha \in [0, 1] \right\}. \end{aligned} \quad (3.47)$$

So with

$$M := \sup \left\{ \left| \nabla F((1-\alpha) \int f \, d\xi + \alpha v) \right| : \xi \in K, v \in [-1, 1]^k, \alpha \in [0, 1] \right\} \quad (3.48)$$

we have

$$e^{-\delta V(\xi, \delta_x, \alpha)} \leq e^{M|f(x)|_1 + ML} \quad (\xi \in K, x \in \mathcal{X}_n, \alpha \in [0, \tfrac{1}{n}]). \quad (3.49)$$

As $M < \infty$, we conclude that V satisfies the $\star$ -dominated bound. $\square$

3.5.13. Note that δV in (3.42) is continuous. Therefore Theorem 3.5.11 is applicable whenever $\mathcal{X}$ is compact.

3.5.14. Note that if $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of finite-volume mean-field Gibbs measures with potential V , and V is as in Theorem 3.5.12, then $(\gamma_n)_{n \in \mathbb{N}}$ as in 3.5.6 is f -dependent.

3.5.15 Examples.

- (a) Consider Example 3.5.4. Both potentials V^1 and V^2 are of the form (3.41), namely with

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = x, \quad g : \mathbb{R} \rightarrow \mathbb{R}^2, \quad g(x) = (x, x^2), \quad (3.50)$$

we have for $G(y, z) = F(y) + z$,

$$V^1(\xi) = F\left(\int f \, d\xi\right), \quad V^2(\xi) = G\left(\int g \, d\xi\right). \quad (3.51)$$

By Theorem 3.5.12 V^1 and V^2 are continuously $\star$ -differentiable whenever F is continuously differentiable. Since $e^{m|x|}$ is $\mu_{\mathcal{N}(0,1)}$ -integrable, V^1 satisfies the $\star$ -dominated bound, as $\frac{1}{n} \log \lambda([-n, n]^c) = -\infty$ (see also (2.42)). On the other hand, V^2 satisfies the $\star$ -dominated bound, but $e^{m(|x|+x^2/2)}$ is not integrable with respect to the Lebesgue measure for $m > 0$.

- (b) In Fernández, den Hollander and Martínez [39] and Ermolaev and Külske [36] Curie-Weiss model is considered. The corresponding Hamiltonians are given by

$$H_n(v) = -\frac{J}{2n} \sum_{i,j=1}^n v_i v_j - h \sum_{i=1}^n v_i \quad (v = (v_1, \dots, v_n) \in \{-1, 1\}^n), \quad (3.52)$$

where $J > 0$ and $h \in \mathbb{R}$. These H_n originate from a potential V of the form (3.41) with $f(x) = x$ and $F(y) = -\frac{J}{2}y^2 - hy$. As $[-1, 1]$ is compact, V satisfies the $\star$ -dominated bound (see Theorem 3.5.11).

3.5.16. In the proof of Lemma 3.5.17 we will use the following fact. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence in $(0, \infty)$. Then $\lim_{n \rightarrow \infty} \frac{1}{n} \log a_n = -\infty$ if and only if $e^{nm} a_n \rightarrow 0$ for all $m > 0$.

3.5.17 Lemma. *Let $f_n, g_n : \mathcal{X} \rightarrow [0, \infty)$ be measurable and let $\mathcal{X}_n$ be a measurable set for all $n \in \mathbb{N}$ and $\mathcal{X}_n \uparrow \mathcal{X}$.*

(a) *If $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ and there is an $m > 0$ such that*

$$f_n(x) \leq e^{nm} \quad (x \in \mathcal{X}_n^c, n \in \mathbb{N}), \quad (3.53)$$

then $\int_{\mathcal{X}_n^c} f_n \, d\lambda \rightarrow 0$.

(b) *If there is an $h \in \mathcal{L}^1(\lambda)$ with*

$$f_n(x), g_n(x) \leq h(x) \quad (x \in \mathcal{X}_n, n \in \mathbb{N}), \quad (3.54)$$

$$|f_n(x) - g_n(x)| \rightarrow 0 \quad (x \in \mathcal{X}), \quad (3.55)$$

then

$$\int_{\mathcal{X}_n} |f_n - g_n| \, d\lambda \rightarrow 0. \quad (3.56)$$

Proof. (a) follows from 3.5.16. (b) follows from Lebesgue's Dominated Convergence Theorem. $\square$

3.5.18 Theorem. *Let $\mathcal{X}$ be a σ -compact locally compact space. Let $V : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$ be a continuous and continuously $\star$ -differentiable potential that satisfies the $\star$ -dominated bound. Then the sequence $(\nu_n)_{n \in \mathbb{N}}$ of finite-volume mean-field Gibbs measures with potential V is sequentially Gibbs with specification kernel given by*

$$\gamma(\zeta, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-\delta V(\zeta, \delta_x)} \, d\lambda(x)}{\int_{\mathcal{X}} e^{-\delta V(\zeta, \delta_x)} \, d\lambda(x)} \quad (\zeta \in \mathcal{P}_c(\mathcal{X}), A \in \mathcal{B}(\mathcal{X})). \quad (3.57)$$

Proof. Let $v_n \in \mathcal{X}^n$ for $n \in \mathbb{N}$ be such that $\zeta_n := L_n(v_n) \xrightarrow{C-w} \zeta$ for some $\zeta \in \mathcal{P}_c(\mathcal{X})$. We prove $\gamma_n(v_n, \cdot) \xrightarrow{w} \gamma(\zeta, \cdot)$ by proving that for all $A \in \mathcal{B}(\mathcal{X})$ (see (3.34))

$$\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-(n+1)[V \circ L_{n+1}(x, v_n) - V \circ L_n(v_n)]} \, d\lambda(x) \rightarrow \int_{\mathcal{X}} \mathbb{1}_A(x) e^{-\delta V(\zeta, \delta_x)} \, d\lambda(x). \quad (3.58)$$

We prove this using Lemma 3.5.17. Define $g : \mathcal{X} \rightarrow [0, \infty)$ by

$$g(x) = e^{-\delta V(\zeta, \delta_x)}. \quad (3.59)$$

Let f_n be given by $f_n(x) = e^{-(n+1)[V \circ L_{n+1}(x, v_n) - V \circ L_n(v_n)]}$. By Lemma 3.5.8 $f_n(x) \rightarrow g(x)$ for all $x \in \mathcal{X}$. By the $\star$ -dominated bound (see 3.5.10) there exist an increasing sequence $(\mathcal{X}_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(\mathcal{X})$ with $\mathcal{X}_n \uparrow \mathcal{X}$ and $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ and a λ -integrable function $h : \mathcal{X} \rightarrow [0, \infty)$ such that $f_n(x) \leq h(x)$ for $x \in \mathcal{X}_n$ and all $n \in \mathbb{N}$. As g is the pointwise limit of the f_n , $g \leq h$ which implies that g is λ -integrable. By continuity of V there is an $m > 0$ such that $f_n(x) \leq e^{(n+1)V \circ L_n(v_n)} \leq e^{(n+1)m}$ for all $x \in \mathcal{X}$. Then by Lemma 3.5.17, $\int_{\mathcal{X}} |f_n - g| \, d\lambda \leq \int_{\mathcal{X}_n^c} f_n \, d\lambda + \int_{\mathcal{X}_n^c} g \, d\lambda + \int_{\mathcal{X}_n} |f_n - g| \, d\lambda \rightarrow 0$, which proves (3.58). $\square$

3.5.19 Theorem. Let $\mathcal{X}$ be a σ -compact locally compact space. Let V_1 and V_2 be continuous and continuously $\star$ -differentiable potentials $\mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty)$ that satisfy the $\star$ -dominated bound. Suppose that $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ are sequences in $[0, \infty)$ for which there exists an $a \in (0, \infty)$ such that

$$\frac{a_n}{n} \rightarrow a, \quad \frac{b_n}{n} \rightarrow 0. \quad (3.60)$$

Then the sequence of finite-volume mean-field Gibbs measures with Hamiltonians given by $H_n = a_n V_1 \circ L_n + b_n V_2 \circ L_n$ is sequentially Gibbs with specification kernel given by

$$\gamma(\zeta, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-a\delta V_1(\zeta, \delta_x)} d\lambda(x)}{\int_{\mathcal{X}} e^{-a\delta V_1(\zeta, \delta_x)} d\lambda(x)} \quad (\zeta \in \mathcal{P}_c(\mathcal{X}), A \in \mathcal{B}(\mathcal{X})). \quad (3.61)$$

Proof. The proof is an adaptation of the proof of Theorem 3.5.18. $\square$

§3.5.3 Further discussions on Hamiltonians

In this section we present our observations made so far on the relation between the Hamiltonians of a sequence of finite-volume mean-field Gibbs measures and the sequential Gibbsianness of this sequence.

3.5.20 Lemma. Suppose that $(\mu_n^1)_{n \in \mathbb{N}}$ and $(\mu_n^2)_{n \in \mathbb{N}}$ are sequences of finite-volume mean-field Gibbs measures with continuous Hamiltonians $(H_n^1)_{n \in \mathbb{N}}$ and $(H_n^2)_{n \in \mathbb{N}}$. If, for all $\zeta \in \mathcal{P}_c(\mathcal{X})$ and all $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \mathcal{X}^n$ and $L_n(v_n) \xrightarrow{C-w} \zeta$,

$$\int_{\mathcal{X}} \left| \frac{e^{-H_{n+1}^1(x, v_n)}}{\int_{\mathcal{X}} e^{-H_{n+1}^1(y, v_n)} d\lambda(y)} - \frac{e^{-H_{n+1}^2(x, v_n)}}{\int_{\mathcal{X}} e^{-H_{n+1}^2(y, v_n)} d\lambda(y)} \right| d\lambda(x) \rightarrow 0, \quad (3.62)$$

then $(\mu_n^1)_{n \in \mathbb{N}}$ is sequentially Gibbs if and only if $(\mu_n^2)_{n \in \mathbb{N}}$ is sequentially Gibbs. If both are sequentially Gibbs, then they have the same specification kernel.

Proof. Let $\gamma_n^1, \gamma_n^2 : \mathcal{X}^n \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ be as in 3.5.6 for μ_n^1 and μ_n^2 , respectively. The convergence in (3.62) implies that if $(v_n)_{n \in \mathbb{N}}$ is such that $v_n \in \mathcal{X}^n$, $\zeta \in \mathcal{P}(\mathcal{X})$, $L_n(v_n) \xrightarrow{C-w} \zeta$ and $\gamma_n^1(v_n, \cdot)$ converges weakly, then $\gamma_n^2(v_n, \cdot)$ converges weakly to the same limit (and vice versa). $\square$

3.5.21 Lemma. Suppose that $f_n, g_n \in \mathcal{L}^1(\lambda)^+$ are such that $\int |f_n - g_n| d\lambda \rightarrow 0$. Moreover, suppose that $F_n := \int f_n d\lambda > 0$ and $G_n := \int g_n d\lambda > 0$ for all $n \in \mathbb{N}$, and that there exists a $\delta > 0$ for which $G_n \geq \delta$ for all $n \in \mathbb{N}$. Then

$$\int \left| \frac{f_n}{F_n} - \frac{g_n}{G_n} \right| d\lambda \rightarrow 0. \quad (3.63)$$

Proof. Let $d_n = |F_n - G_n|$. As $d_n \rightarrow 0$, for convenience we assume that $d_n < \frac{\delta}{3}$ for all $n \in \mathbb{N}$. Then $G_n, F_n, G_n + d_n, G_n - d_n \geq \frac{\delta}{3}$ for all $n \in \mathbb{N}$. We have

$$\left| \frac{f_n(x)}{F_n} - \frac{g_n(x)}{G_n} \right| \leq \left| \frac{f_n(x)}{G_n + d_n} - \frac{g_n(x)}{G_n} \right| + \left| \frac{f_n(x)}{G_n - d_n} - \frac{g_n(x)}{G_n} \right|. \quad (3.64)$$

With e_n being either d_n or $-d_n$, we have

$$\begin{aligned} \left| \frac{f_n(x)}{G_n + e_n} - \frac{g_n(x)}{G_n} \right| &= \left| \frac{f_n(x) - g_n(x)}{G_n + e_n} \right| + \left| \frac{g_n(x)e_n}{G_n(G_n + e_n)} \right| \\ &\leq \frac{3}{\delta} |f_n(x) - g_n(x)| + \frac{3}{\delta} \frac{g_n(x)}{G_n} d_n. \end{aligned} \quad (3.65)$$

Hence

$$\int \left| \frac{f_n(x)}{F_n} - \frac{g_n(x)}{G_n} \right| d\lambda(x) \leq \frac{6}{\delta} \int |f_n - g_n| d\lambda + \frac{6}{\delta} d_n. \quad (3.66)$$

□

3.5.22 Theorem. *Let $\zeta \in \mathcal{P}_c(\mathcal{X})$ and let $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \mathcal{X}^n$ be such that $L_n(v_n) \xrightarrow{C-w} \zeta$. We suppose the following.*

- $\mathcal{X}$ is a σ -compact locally compact space.
- $V : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty)$ is a continuous and continuously $\star$ -differentiable potential.
- $(H_n)_{n \in \mathbb{N}}$ is a sequence of continuous Hamiltonians.
- There is an $m > 0$ such that $H_n(v_n) \leq nm$ for all $n \in \mathbb{N}$.
- There are measurable $\mathcal{X}_n \subset \mathcal{X}$ with $\mathcal{X}_n \uparrow \mathcal{X}$, $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$, and an $h \in \mathcal{L}^1(\lambda)^+$ such that

$$\begin{aligned} \exp(-[H_{n+1}(x, v_n) - H_n(v_n)]) &\leq h(x) \quad (x \in \mathcal{X}_n, n \in \mathbb{N}), \\ \exp(-(n+1)[V \circ L_{n+1}(x, v_n) - V \circ L_n(v_n)]) &\leq h(x) \quad (x \in \mathcal{X}_n, n \in \mathbb{N}), \end{aligned} \quad (3.67)$$

- For all $x \in \mathcal{X}$:

$$\lim_{n \rightarrow \infty} \left| H_{n+1}(x, v_n) - H_n(v_n) - (n+1)[V \circ L_{n+1}(x, v_n) - V \circ L_n(v_n)] \right| = 0. \quad (3.68)$$

Then

$$\int_{\mathcal{X}} \left| \frac{e^{-H_{n+1}(x, v_n)}}{\int_{\mathcal{X}} e^{-H_{n+1}(y, v_n)} d\lambda(y)} - \frac{e^{-(n+1)V \circ L_{n+1}(x, v_n)}}{\int_{\mathcal{X}} e^{-(n+1)V \circ L_{n+1}(y, v_n)} d\lambda(y)} \right| d\lambda(x) \rightarrow 0. \quad (3.69)$$

Consequently, if the above holds for all $\zeta \in \mathcal{P}_c(\mathcal{X})$ and all $(v_n)_{n \in \mathbb{N}}$ with $v_n \in \mathcal{X}^n$ and $L_n(v_n) \xrightarrow{C-w} \zeta$ and if $(\nu_n)_{n \in \mathbb{N}}$ is the sequence of finite-volume mean-field Gibbs measures with Hamiltonians $(H_n)_{n \in \mathbb{N}}$, then $(\nu_n)_{n \in \mathbb{N}}$ is sequentially Gibbs.

Proof. Define $f_n, g_n : \mathcal{X} \times \mathcal{X}^n \rightarrow [0, \infty)$ as

$$f_n(x) = \exp(-[H_{n+1}(x, v_n) - H_n(v_n)]) \quad (x \in \mathcal{X}), \quad (3.70)$$

$$g_n(x) = \exp(-(n+1)[V \circ L_{n+1}(x, v_n) - V \circ L_n(v_n)]) \quad (x \in \mathcal{X}). \quad (3.71)$$

By continuity of V and by the condition on $H_n(v_n)$, there exists an $m > 0$ such that $f_n(x), g_n(x) \leq e^{nm}$ for all $x \in \mathcal{X}$ and $n \in \mathbb{N}$. Note that (3.67) and (3.68) imply

$$f_n(x), g_n(x) \leq h(x) \quad (x \in \mathcal{X}_n, n \in \mathbb{N}), \quad (3.72)$$

$$\lim_{n \rightarrow \infty} |f_n(x) - g_n(x)| = 0 \quad (x \in \mathcal{X}). \quad (3.73)$$

Hence by Lemma 3.5.17 we obtain

$$\int_{\mathcal{X}} |f_n - g_n| \, d\lambda \leq \int_{\mathcal{X}_n^c} f_n \, d\lambda + \int_{\mathcal{X}_n^c} g_n \, d\lambda + \int_{\mathcal{X}_n} |f_n - g_n| \, d\lambda \rightarrow 0. \quad (3.74)$$

Define $F_n = \int f_n \, d\lambda$ and $G_n = \int g_n \, d\lambda$ for $n \in \mathbb{N}$. By Lemma 3.5.21 it is sufficient to show that there exists a $\delta > 0$ such that $G_n \geq \delta$ for all $n \in \mathbb{N}$. But this follows from the fact that $G_n \rightarrow \int e^{-\delta V(\zeta, \delta x)} \, d\lambda(x)$, which in turn follows from the convergence in (3.58) in the proof of Theorem 3.5.18. $\square$

§3.6 Transformations of Gibbs measures by probability kernels

We consider a transformation of a sequence of finite-volume mean-field Gibbs measures by a sequence of probability kernels. We derive relations between regular conditional kernels. In particular, we describe these relations in the case of independent transformations (in Theorem 3.6.3) and in case the transformation possess a density function (in Theorem 3.6.6). We use these relations and formulas in Section 3.7.

Throughout this section we assume the following:

- (i) $\mathcal{X}$ and $\mathcal{Y}$ are perfectly normal topological spaces.
- (ii) $\lambda \in \mathcal{P}(\mathcal{X})$ and $\text{supp}(\lambda) = \mathcal{X}$.
- (iii) $(\nu_n)_{n \in \mathbb{N}}$ is assumed to be a sequence of finite-volume mean-field Gibbs measures with Hamiltonians $(H_n)_{n \in \mathbb{N}}$ as in Definition 3.5.1.
- (iv) H_n is continuous for all $n \in \mathbb{N}$.
- (v) P_n is a probability kernel $\mathcal{X}^n \times \mathcal{B}(\mathcal{Y}^n) \rightarrow [0, 1]$ for all $n \in \mathbb{N}$.
- (vi) Similarly as in Section 3.4, but here with $\mathcal{Y}$ instead of $\mathcal{X}$:

$$\sigma_1(y) = y_1, \quad \sigma_{\perp}(y) = (y_2, \dots, y_{n+1}) \quad \left(y = (y_1, \dots, y_{n+1}) \in \mathcal{Y}^{n+1} \right). \quad (3.75)$$

- (vii) For all $n \in \mathbb{N}$ we write π for the map $\pi : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{Y}^n$ given by $\pi(v, w) = w$.
- (viii) For all $n \in \mathbb{N}$ there exists a product regular conditional kernel $\eta_n^0 : \mathcal{Y}^n \times \mathcal{B}(\mathcal{X}^n) \rightarrow [0, 1]$ under $\lambda^n \times P_n$ with respect to π (of the coordinates in $\mathcal{X}$ given the coordinates in $\mathcal{Y}$) that is weakly continuous on $\text{supp}(\lambda^n P_n)$. (See (3.4) for the notations $\lambda^n \times P_n$ and $\lambda^n P_n$.)
- (ix) For all $n \in \mathbb{N}$ there exists a product regular conditional kernel $\gamma_n^0 : \mathcal{Y}^n \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ under $\lambda^{n+1} P_{n+1}$ with respect to $\sigma_{\perp}$ (of the first coordinate given the other n coordinates) that is weakly continuous on $\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1})$.

We define

(a) $\eta_n : \mathcal{Y}^n \times \mathcal{B}(\mathcal{X}^n) \rightarrow [0, 1]$ by

$$\eta_n(w, A) = \begin{cases} \frac{\int_{\mathcal{X}^n} \mathbb{1}_A(v) e^{-H_n(v)} d[\eta_n^0(w, \cdot)](v)}{\int_{\mathcal{X}^n} e^{-H_n(v)} d[\eta_n^0(w, \cdot)](v)} & w \in \text{supp}(\lambda^n P_n), \\ 0 & w \notin \text{supp}(\lambda^n P_n). \end{cases} \quad (3.76)$$

(b) $\mathcal{Z}_n : \mathcal{Y}^n \rightarrow (0, 1]$ by

$$\mathcal{Z}_n(w) = \begin{cases} \int_{\mathcal{X}^n} e^{-H_n(v)} d[\eta_n^0(w, \cdot)](v) & w \in \text{supp}(\lambda^n P_n), \\ 0 & w \notin \text{supp}(\lambda^n P_n). \end{cases} \quad (3.77)$$

(c) $\gamma_n : \mathcal{Y}^n \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ by

$$\gamma_n(w, D) = \begin{cases} \frac{\int_{\mathcal{Y}} \mathbb{1}_D(y) \mathcal{Z}_{n+1}(y, w) d[\gamma_n^0(w, \cdot)](y)}{\int_{\mathcal{Y}} \mathcal{Z}_{n+1}(y, w) d[\gamma_n^0(w, \cdot)](y)} & w \in \text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}), \\ 0 & w \notin \text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}). \end{cases} \quad (3.78)$$

We address the relations between γ_n , γ_n^0 , $\mathcal{Z}_n$, η_n^0 , η_n , ν_n , λ^n and P_n .

3.6.1. By our assumptions on η_n^0 and γ_n^0 , we have for $A \in \mathcal{B}(\mathcal{X}^n)$, $B \in \mathcal{B}(\mathcal{Y}^n)$ and $D \in \mathcal{B}(\mathcal{Y})$,

$$\lambda^n \rtimes P_n(A \times B) = \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \eta_n^0(w, A) d[\lambda^n P_n](w), \quad (3.79)$$

$$\lambda^{n+1} P_{n+1}(D \times B) = \int_{\mathcal{Y} \times \mathcal{Y}^n} \mathbb{1}_B(w) \gamma_n^0(w, D) d[\lambda^{n+1} P_{n+1}](y, w). \quad (3.80)$$

Moreover, by Theorem 3.3.17 we have

$$\text{supp}(\lambda^n P_n) = \text{supp}(\nu_n P_n), \quad (3.81)$$

$$\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}) = \text{supp}(\nu_{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}). \quad (3.82)$$

3.6.2 Theorem.

(a) η_n is a product regular conditional kernel under $\nu_n \rtimes P_n$ with respect to π (of the coordinates in $\mathcal{X}$ given the coordinates in $\mathcal{Y}$) that is weakly continuous on $\text{supp}(\lambda^n P_n)$.

(b) For $A \in \mathcal{B}(\mathcal{X}^n)$ and $B \in \mathcal{B}(\mathcal{Y}^n)$

$$\nu_n \rtimes P_n(A \times B) = \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \frac{\mathcal{Z}_n(w)}{Z_n} \eta_n(w, A) d[\lambda^n P_n](w). \quad (3.83)$$

(c) $\nu_n P_n$ is absolutely continuous with respect to $\lambda^n P_n$. The Radon-Nikodym derivative is given by

$$\frac{d\nu_n P_n}{d\lambda^n P_n}(w) = \frac{\mathcal{Z}_n(w)}{Z_n} \quad (w \in \mathcal{Y}^n). \quad (3.84)$$

- (d) $Z_n = \int_{\mathcal{Y}^n} \mathcal{Z}_n(w) \, d[\lambda^n P_n](w)$.
- (e) $\mathcal{Z}_n$ is continuous on $\text{supp}(\lambda^n P_n)$.
- (f) γ_n is a product regular conditional kernel under $\nu_{n+1} P_{n+1}$ with respect to $\sigma_\perp$ (of the first coordinate given the other n coordinates) that is weakly continuous on $\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_\perp^{-1}) \cap \text{supp}(\lambda^n P_n)$.

Proof. For (a) see Section 3.B.

(b) For $A \in \mathcal{B}(\mathcal{X}^n)$ and $B \in \mathcal{B}(\mathcal{Y}^n)$ we have

$$\begin{aligned}
 \nu_n \rtimes P_n(A \times B) &= \frac{1}{Z_n} \int_{\mathcal{X}^n \times \mathcal{Y}^n} \mathbb{1}_{A \times B}(v, w) e^{-H_n(v)} \, d[\lambda^n \rtimes P_n](v, w) \\
 &= \frac{1}{Z_n} \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \left[\int_{\mathcal{X}^n} \mathbb{1}_A(v) e^{-H_n(v)} \, d[\eta_n^0(w, \cdot)](v) \right] d[\lambda^n P_n](w) \\
 &= \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \frac{\mathcal{Z}_n(w)}{Z_n} \eta_n(w, A) \, d[\lambda^n P_n](w).
 \end{aligned} \tag{3.85}$$

Hence

$$\nu_n P_n(B) = \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \frac{\mathcal{Z}_n(w)}{Z_n} \, d[\lambda^n P_n](w), \tag{3.86}$$

which implies (c).

(d) is a consequence of (c).

(e) follows from the fact that e^{-H_n} is a bounded continuous function.

(f) For $D \in \mathcal{B}(\mathcal{Y})$ and $B \in \mathcal{B}(\mathcal{Y}^n)$ we have, by (3.80) and (3.86),

$$\begin{aligned}
 &\nu_{n+1} P_{n+1}(D \times B) \\
 &= \int_{\mathcal{Y} \times \mathcal{Y}^n} \mathbb{1}_B(w) \frac{1}{Z_{n+1}} \left[\int_{\mathcal{Y}} \mathbb{1}_D(y) \mathcal{Z}_{n+1}(y, w) \, d[\gamma_n^0(w, \cdot)](y) \right] \\
 &\quad d[\lambda^{n+1} P_{n+1}](z, w) \\
 &= \int_{\mathcal{Y} \times \mathcal{Y}^n} \mathbb{1}_B(w) \gamma_n(w, D) \frac{1}{Z_{n+1}} \left[\int_{\mathcal{Y}} \mathcal{Z}_{n+1}(y, w) \, d[\gamma_n^0(w, \cdot)](y) \right] \\
 &\quad d[\lambda^{n+1} P_{n+1}](z, w) \\
 &= \int_{\mathcal{Y} \times \mathcal{Y}^n} \mathbb{1}_B(w) \gamma_n(w, D) \frac{\mathcal{Z}_{n+1}(z, w)}{Z_{n+1}} \, d[\lambda^{n+1} P_{n+1}](z, w) \\
 &= \int_{\mathcal{Y} \times \mathcal{Y}^n} \mathbb{1}_B(w) \gamma_n(w, D) \, d[\nu_{n+1} P_{n+1}](z, w).
 \end{aligned} \tag{3.87}$$

As γ_n^0 is weakly continuous on $\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_\perp^{-1})$, (f) follows from (3.87) and (e). $\square$

In the following we discuss different scenarios: P_n is the product of kernels (Theorem 3.6.3), in the sense that it is composed by independent transformations (every coordinate undergoes a transformation independent of the others), P_n has a strictly positive density function (Theorem 3.6.6) and P_n is a “mean-field” transformation (Definition 3.6.8 and onwards).

3.6.3 Theorem (Independent transformations). Define $\psi_n : \mathcal{X} \times \mathcal{Y}^n \rightarrow (0, \infty]$ by⁴

$$\psi_n(x, w) := \int_{\mathcal{X}^n} e^{-H_{n+1}(x, v)} d[\eta_n^0(w, \cdot)](v). \quad (3.88)$$

Let $P : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ be a probability kernel. Suppose that $P_n(v, \cdot) = \bigotimes_{i=1}^n P(v_i, \cdot)$ for all $v \in \mathcal{X}^n$. Then

$$\text{supp}(\lambda^n P_n) = \text{supp}(\lambda P)^n = \text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}). \quad (3.89)$$

Moreover, for $w = (w_1, \dots, w_n)$ in $\text{supp}(\lambda P)^n$ and $D \in \mathcal{B}(\mathcal{Y})$

- (a) $\eta_n^0(w, \cdot) = \bigotimes_{i=1}^n \eta_1^0(w_i, \cdot)$.
- (b) $\gamma_n^0(w, \cdot) = \lambda P$ and

$$\gamma_n(w, D) = \frac{\int_{\mathcal{X}} \psi_n(x, w) P(x, D) d\lambda(x)}{\int_{\mathcal{X}} \psi_n(x, w) d\lambda(x)}. \quad (3.90)$$

Proof. It is straightforward to check that $\eta_n^0(w, \cdot) = \bigotimes_{i=1}^n \eta_1^0(w_i, \cdot)$ and $\gamma_n^0(w, \cdot) = \lambda P$. Using (a) and the definition of η_n , we obtain

$$\begin{aligned} \mathcal{Z}_{n+1}(y, w) &= \int_{\mathcal{X}} \int_{\mathcal{X}^n} e^{-H_{n+1}(z, v)} d[\eta_n^0(w, \cdot)](v) d[\eta_1^0(y, \cdot)](z) \\ &= \int_{\mathcal{X}} \psi_n(z, w) d[\eta_1^0(y, \cdot)](z). \end{aligned} \quad (3.91)$$

Hence

$$\begin{aligned} &\int_{\mathcal{Y}} \mathbb{1}_D(y) \mathcal{Z}_{n+1}(y, w) d[\lambda P](y) \\ &= \int_{\mathcal{Y}} \mathbb{1}_D(y) \int_{\mathcal{X}} \psi_n(z, w) d[\eta_1^0(y, \cdot)](z) d[\lambda P](y) \\ &= \int_{\mathcal{X} \times \mathcal{Y}} \mathbb{1}_D(y) \psi_n(x, w) d[\lambda \rtimes P](x, y) = \int_{\mathcal{X}} \psi_n(x, w) P(x, D) d\lambda(x). \end{aligned} \quad (3.92)$$

From this formula we obtain (3.90). $\square$

3.6.4 Remark. The formula for γ_n in terms of ψ_n in Theorem 3.6.3 is comparable to the formula for $\bar{\gamma}_{n,t}$ of Chapter 2 in terms of $g_{n,t}$ as in (2.10) and (2.11).

3.6.5. Let $\kappa \in \mathcal{P}(\mathcal{Y})$ and let $p_n : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow [0, \infty)$ be a measurable function. Suppose that

$$P_n(v, B) = \int_{\mathcal{Y}^n} p_n(v, w) d\kappa^n(w) \quad (v \in \mathcal{X}^n, B \in \mathcal{B}(\mathcal{Y}^n)) \quad (3.93)$$

⁴As $e^{-H_{n+1}}$ is integrable with respect to λ^{n+1} , $(v, w) \mapsto e^{-H_{n+1}(x, v)}$ is integrable with respect to $\lambda^n \rtimes P_n$ and so $\int_{\mathcal{Y}^n} \int_{\mathcal{X}^n} e^{-H_{n+1}(x, v)} d[\eta_n^0(w, \cdot)](v) d[\lambda^n P_n](w) < \infty$ for λ -almost all $x \in \mathcal{X}$.

Then p_n is $\lambda^n \times \kappa^n$ -integrable and

$$\lambda^n \times P_n(A) = \int_{\mathcal{X}^n} \int_{\mathcal{Y}^n} p_n(v, w) \, d\kappa^n(w) \, d\lambda^n(v) \quad (A \in \mathcal{B}(\mathcal{X}^n \times \mathcal{Y}^n)). \quad (3.94)$$

Hence $\lambda^n P_n(B) = \int_{\mathcal{Y}^n} \mathbb{1}_B(w) q_n(w) \, d\kappa^n(w)$ for the density function $q_n : \mathcal{Y}^n \rightarrow [0, \infty)$ given by $q_n(w) = \int_{\mathcal{X}^n} p_n(v, w) \, d\lambda^n(v)$. Assuming that $q_n(w) > 0$ for all $w \in \mathcal{Y}^n$, we obtain (e.g., by Theorem 4.4.7) for $w \in \mathcal{Y}^n$, $A \in \mathcal{B}(\mathcal{X}^n)$ and $D \in \mathcal{B}(\mathcal{Y})$,

$$\eta_n^0(w, A) = \frac{\int_{\mathcal{X}^n} \mathbb{1}_A(v) p_n(v, w) \, d\lambda^n(v)}{\int_{\mathcal{X}^n} p_n(v, w) \, d\lambda^n(v)} = \frac{\int_{\mathcal{X}^n} \mathbb{1}_A(v) p_n(v, w) \, d\lambda^n(v)}{q_n(w)}, \quad (3.95)$$

$$\gamma_n^0(w, D) = \frac{\int_{\mathcal{Y}} \mathbb{1}_D(y) q_{n+1}(y, w) \, d\kappa(y)}{\int_{\mathcal{Y}} q_{n+1}(y, w) \, d\kappa(y)}, \quad (3.96)$$

$$\begin{aligned} \gamma_n(w, D) &= \frac{\int_{\mathcal{Y}} \mathbb{1}_D(z) \mathcal{Z}_{n+1}(z, w) q_{n+1}(z, w) \, d\kappa(z)}{\int_{\mathcal{Y}} \mathcal{Z}_{n+1}(z, w) q_{n+1}(z, w) \, d\kappa(z)} \\ &= \frac{\int_{\mathcal{Y}} \mathbb{1}_D(z) \int_{\mathcal{X}^{n+1}} e^{-H_{n+1}(v)} \, d[\eta_{n+1}^0((z, w), \cdot)](v) q_{n+1}(z, w) \, d\kappa(z)}{\int_{\mathcal{Y}} \int_{\mathcal{X}^{n+1}} e^{-H_{n+1}(v)} \, d[\eta_{n+1}^0((z, w), \cdot)](v) q_{n+1}(z, w) \, d\kappa(z)} \\ &= \frac{\int_{\mathcal{Y}} \mathbb{1}_D(z) \int_{\mathcal{X}^{n+1}} e^{-H_{n+1}(v)} p_{n+1}(v, (z, w)) \, d\lambda^{n+1}(v) \, d\kappa(z)}{\int_{\mathcal{Y}} \int_{\mathcal{X}^{n+1}} e^{-H_{n+1}(v)} p_{n+1}(v, (z, w)) \, d\lambda^{n+1}(v) \, d\kappa(z)} \\ &= \frac{\int_{\mathcal{Y}} \mathbb{1}_D(z) \int_{\mathcal{X}} \int_{\mathcal{X}^n} e^{-H_{n+1}(x, v)} p_{n+1}((x, v), (z, w)) \, d\lambda^n(v) \, d\lambda(x) \, d\kappa(z)}{\int_{\mathcal{Y}} \int_{\mathcal{X}} \int_{\mathcal{X}^n} e^{-H_{n+1}(x, v)} p_{n+1}((x, v), (z, w)) \, d\lambda^n(v) \, d\lambda(x) \, d\kappa(z)}. \end{aligned} \quad (3.97)$$

From the observations in 3.6.5 we obtain the following theorem.

3.6.6 Theorem. *Let $\kappa \in \mathcal{P}(\mathcal{Y})$. Suppose that $p_n : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow (0, \infty)$ is a strictly positive measurable function such that*

$$P_n(v, B) = \int_{\mathcal{Y}^n} \mathbb{1}_B(w) p_n(v, w) \, d\kappa^n(w) \quad (v \in \mathcal{X}^n, B \in \mathcal{B}(\mathcal{Y}^n)). \quad (3.98)$$

Then

$$\gamma_n(w, D) = \frac{\int_{\mathcal{X}} \int_{\mathcal{Y}} \mathbb{1}_D(z) \phi_n(x, z, w) \, d\lambda(x) \, d\kappa(z)}{\int_{\mathcal{Y}} \int_{\mathcal{X}} \phi_n(x, z, w) \, d\lambda(x) \, d\kappa(z)}, \quad (3.99)$$

where ϕ_n is the function $\mathcal{X} \times \mathcal{Y} \times \mathcal{Y}^n \rightarrow (0, \infty)$ given by

$$\phi_n(x, z, w) = \int_{\mathcal{X}^n} e^{-H_{n+1}(x, v)} \frac{p_{n+1}((x, v), (z, w))}{p_n(v, w)} \, d[\eta_n^0(w, \cdot)](v). \quad (3.100)$$

Proof. As p_n is strictly positive, by (3.95),

$$\frac{d\lambda^n}{d\eta_n^0(w, \cdot)}(v) = \frac{q_n(w)}{p_n(v, w)} \quad (v \in \mathcal{X}^n). \quad (3.101)$$

The rest follows from (3.97).

3.6.7. We compare ψ_n with ϕ_n in the situation of Theorem 3.6.6 with independent transformations, as in Theorem 3.6.3: Suppose that $P : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ is a probability kernel such that $P_n(v, \cdot) = \bigotimes_{i=1}^n P(v_i, \cdot)$ for all $v \in \mathcal{X}^n$. Let $\kappa \in \mathcal{P}(\mathcal{Y})$. Suppose that $p : \mathcal{X} \times \mathcal{Y} \rightarrow (0, \infty)$ is a strictly positive measurable function such that

$$P(x, B) = \int_{\mathcal{Y}} \mathbb{1}_B(y) p(x, y) \, d\kappa(y) \quad (x \in \mathcal{X}, B \in \mathcal{B}(\mathcal{Y})). \quad (3.102)$$

Then, with $p_n(v, w) = p(v_1, w_1)p(v_2, w_2) \cdots p(v_n, w_n)$ for $v = (v_1, \dots, v_n) \in \mathcal{X}^n$ and $w = (w_1, \dots, w_n) \in \mathcal{Y}^n$, we obtain (3.98). Moreover, with ϕ_n as in (3.100) and ψ_n as in (3.88), we have for $x \in \mathcal{X}, z \in \mathcal{Y}, w \in \mathcal{Y}^n$ and $D \in \mathcal{B}(\mathcal{Y})$,

$$\phi_n(x, z, w) = p(x, z)\psi_n(x, w), \quad (3.103)$$

$$\psi_n(x, w)P(x, D) = \int_{\mathcal{Y}} \mathbb{1}_D(z)\phi_n(x, z, w) \, d\kappa(z), \quad (3.104)$$

$$\psi_n(x, w) = \int_{\mathcal{Y}} \phi_n(x, z, w) \, d\kappa(z). \quad (3.105)$$

In the mean-field setting it is reasonable to consider mean-field transformations, which heuristically means that the probability to go from v to w (where $v \in \mathcal{X}^n$ and $w \in \mathcal{Y}^n$) is the same as the probability to go from $\hat{\sigma}(v)$ to $\hat{\sigma}(w)$ for every $\sigma \in \text{Sym}(n)$.

3.6.8 Definition. We say that the kernel P_n is a *mean-field transformation* if

$$P_n(v, A) = P_n(\hat{\sigma}(v), \hat{\sigma}(A)) \quad (v \in \mathcal{X}^n, A \in \mathcal{B}(\mathcal{Y}^n), \sigma \in \text{Sym}(n)). \quad (3.106)$$

3.6.9 Lemma. Suppose that P_n is a mean-field transformation. Then, for all $\sigma \in \text{Sym}(n)$,

$$\eta_n^0(\hat{\sigma}(w), \hat{\sigma}(A)) = \eta_n^0(w, A) \quad (w \in \mathcal{Y}^n, A \in \mathcal{B}(\mathcal{X}^n)), \quad (3.107)$$

$$\gamma_n^0(\hat{\sigma}(w), D) = \gamma_n^0(w, D) \quad (w \in \mathcal{Y}^n, D \in \mathcal{B}(\mathcal{Y})). \quad (3.108)$$

If also $H_n \circ \hat{\sigma} = H_n$ for all $\sigma \in \text{Sym}(n)$, then, for all $\sigma \in \text{Sym}(n)$, $\mathcal{Z}_n \circ \hat{\sigma} = \mathcal{Z}_n$ and

$$\eta_n(\hat{\sigma}(w), \hat{\sigma}(A)) = \eta_n(w, A) \quad (w \in \mathcal{Y}^n, A \in \mathcal{B}(\mathcal{X}^n)), \quad (3.109)$$

$$\gamma_n(\hat{\sigma}(w), D) = \gamma_n(w, D) \quad (w \in \mathcal{Y}^n, D \in \mathcal{B}(\mathcal{Y})). \quad (3.110)$$

Proof. Let $\sigma \in \text{Sym}(n)$. By a simple calculation one can prove that $\lambda^n \rtimes P_n(\hat{\sigma}(A) \times \hat{\sigma}(B)) = \lambda^n \rtimes P_n(A \times B)$ for all $A \in \mathcal{B}(\mathcal{X}^n)$ and $B \in \mathcal{B}(\mathcal{Y}^n)$. Then

$$\begin{aligned} & \int_{\mathcal{Y}^n} \mathbb{1}_B(w) \eta_n^0(\hat{\sigma}(w), \hat{\sigma}(A)) \, d[\lambda^n P_n](w) \\ &= \int_{\mathcal{Y}^n} \mathbb{1}_{\hat{\sigma}(B)}(w) \eta_n^0(w, \hat{\sigma}(A)) \, d[\lambda^n P_n](w) \\ &= \lambda^n \rtimes P_n(\hat{\sigma}(A) \times \hat{\sigma}(B)) \\ &= \lambda^n \rtimes P_n(A \times B) \quad (A \in \mathcal{B}(\mathcal{X}^n), B \in \mathcal{B}(\mathcal{Y}^n)). \end{aligned} \quad (3.111)$$

By Theorem 4.4.3 we conclude (3.107). In a similar fashion, we prove (3.108). The statements for η_n , $\mathcal{Z}_n$ and γ_n follow from their definition. $\square$

By Lemma 3.6.9, the kernels η_n^0 , η_n , γ_n^0 depend rather on $L_n(w)$ than on their argument w . For that reason we define in 3.6.10 and in 3.6.13 their corresponding kernel indexed by the probability measures.

3.6.10. Suppose that P_n is a mean-field transformation, as in the setting of Lemma 3.6.9. Then $\eta_n^0(w, L_n^{-1}(A)) = \eta_n^0(\hat{\sigma}(w), L_n^{-1}(A))$ for all $w \in \mathcal{Y}^n$ and $A \in \mathcal{B}(\mathcal{P}_c(\mathcal{X}))$. Therefore we can define a kernel $\tilde{\eta}_n^0 : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{P}_c(\mathcal{X})) \rightarrow [0, 1]$ by

$$\tilde{\eta}_n^0(\xi, A) = \begin{cases} \eta_n^0(w, L_n^{-1}(A)) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}). \end{cases} \quad (3.112)$$

For $A \in \mathcal{B}(\mathcal{P}_c(\mathcal{X}))$ and $B \in \mathcal{B}(\mathcal{P}_c(\mathcal{Y}))$,

$$\begin{aligned} [\lambda^n \rtimes P_n](L_n^{-1}(A) \times L_n^{-1}(B)) &= \int_{\mathcal{Y}^n} \mathbb{1}_{L_n^{-1}(B)}(w) \eta_n^0(w, L_n^{-1}(A)) \, d[\lambda^n P_n](w) \\ &= \int_{\mathcal{P}_c(\mathcal{Y})} \mathbb{1}_B(\xi) \tilde{\eta}_n^0(\xi, A) \, d[\lambda^n P_n \circ L_n^{-1}](\xi). \end{aligned} \quad (3.113)$$

Let $\mathbf{m} : \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \rightarrow \mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{Y})$ be the map that maps a probability measure on $\mathcal{X} \times \mathcal{Y}$ to the pair of its marginals, and let

$$\mu_n = [\lambda^n \rtimes P_n] \circ L_n^{-1} \circ \mathbf{m}^{-1}. \quad (3.114)$$

Note that $\mu_n(A \times B)$ agrees with (3.113). We write π for the projection on the second coordinate of $\mathcal{P}_c(\mathcal{X}) \times \mathcal{P}_c(\mathcal{Y})$. By (3.113) we see that $\tilde{\eta}_n^0$ is the product regular conditional kernel under μ_n with respect to π . By Theorem 3.3.12, $L_n(\text{supp}(\lambda^n P_n))$ is closed and $\tilde{\eta}_n^0$ is weakly continuous on $L_n(\text{supp}(\lambda^n P_n))$.

Moreover, as $\mu_n \circ \pi^{-1} = \lambda^n P_n \circ L_n^{-1}$ by Theorem 4.4.2, we have

$$0 \leq \mu_n \circ \pi^{-1}(\mathcal{P}_c(\mathcal{Y}) \setminus L_n(\text{supp}(\lambda^n P_n))) \leq \lambda^n P_n(\mathcal{X}^n \setminus \text{supp}(\lambda^n P_n)) = 0, \quad (3.115)$$

so that $\text{supp}(\mu_n \circ \pi^{-1}) \subset L_n(\text{supp}(\lambda^n P_n))$. If $C \subset \mathcal{P}_c(\mathcal{X})$ is closed and $\mu_n \circ \pi^{-1}(C) = 0$, i.e., $\lambda^n P_n(L_n^{-1}(C)) = 0$, then $\text{supp}(\lambda^n P_n) \subset L_n^{-1}(C)$, from which we obtain the inclusion $L_n(\text{supp}(\lambda^n P_n)) \subset L_n(L_n^{-1}(C)) \subset C$. With the formula in (3.16) for the support of a measure, we conclude

$$\text{supp}(\mu_n \circ \pi^{-1}) = L_n(\text{supp}(\lambda^n P_n)). \quad (3.116)$$

We summarise the observations of 3.6.10 in the following theorem.

3.6.11 Theorem. Suppose that P_n is a mean-field transformation. Let $\tilde{\eta}_n^0 : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{P}_c(\mathcal{X})) \rightarrow [0, 1]$ be as defined in (3.112). Then $\tilde{\eta}_n^0$ is the product regular conditional kernel under μ_n with respect to π that is weakly continuous on $\text{supp}(\mu_n \circ \pi^{-1}) = L_n(\text{supp}(\lambda^n P_n))$.

3.6.12. Theorem 3.6.11 implies the following. If $\zeta \in \mathcal{P}_c(\mathcal{Y})$ and $I_\zeta : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty]$ is lower semicontinuous, then the sequence $(\eta_n^0(w_n, \cdot) \circ L_n^{-1})_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I_ζ for all sequences $(w_n)_{n \in \mathbb{N}}$ with $w_n \in \text{supp}(\lambda^n P_n)$ and $L_n(w_n) \xrightarrow{C-w} \zeta$ if and only if $(\tilde{\eta}_n^0(\zeta_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I_ζ for all sequences $(\zeta_n)_{n \in \mathbb{N}}$ with $\zeta_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $\zeta_n \xrightarrow{C-w} \zeta$.

3.6.13. Consider again the situation in Lemma 3.6.9, where P_n is a mean-field transformation and $H_n \circ \hat{\sigma} = H_n$ for all $\sigma \in \text{Sym}(n)$. Similarly as in 3.6.10 we can define the kernels

$$\tilde{\eta}_n : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{P}_c(\mathcal{X})) \rightarrow [0, 1], \quad (3.117)$$

$$\tilde{\gamma}_n^0 : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1], \quad (3.118)$$

$$\tilde{\gamma}_n : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1], \quad (3.119)$$

and $\tilde{H}_n : \mathcal{P}_c(\mathcal{X}) \rightarrow \mathbb{R}$, $\tilde{Z}_n : \mathcal{P}_c(\mathcal{Y}) \rightarrow (0, 1]$ by

$$\tilde{\eta}_n(\xi, A) = \begin{cases} \eta_n^0(w, L_n^{-1}(A)) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}), \end{cases} \quad (3.120)$$

$$\tilde{\gamma}_n^0(\xi, D) = \begin{cases} \gamma_n^0(w, D) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}), \end{cases} \quad (3.121)$$

$$\tilde{\gamma}_n(\xi, D) = \begin{cases} \gamma_n(w, D) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}), \end{cases} \quad (3.122)$$

$$\tilde{H}_n(\xi) = \begin{cases} H_n(w) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}), \end{cases} \quad (3.123)$$

$$\tilde{Z}_n(\xi) = \begin{cases} Z_n(w) & \text{if } \xi = L_n(w) \text{ for some } w \in \mathcal{Y}^n, \\ 0 & \text{if } \xi \notin \mathcal{P}_{emp}^n(\mathcal{Y}). \end{cases} \quad (3.124)$$

Similarly to (3.76), (3.77) and (3.78) we have

$$\tilde{\eta}_n(\xi, A) = \begin{cases} \frac{\int_{\mathcal{P}_c(\mathcal{X})} \mathbb{1}_A(\rho) e^{-\tilde{H}_n(\rho)} d[\tilde{\eta}_n^0(\xi, \cdot)](\rho)}{\int_{\mathcal{X}^n} e^{-\tilde{H}_n(\rho)} d[\tilde{\eta}_n^0(\xi, \cdot)](\rho)} & \xi \in L_n(\text{supp}(\lambda^n P_n)), \\ 0 & \xi \notin L_n(\text{supp}(\lambda^n P_n)), \end{cases} \quad (3.125)$$

$$\tilde{Z}_n(\xi) = \begin{cases} \int_{\mathcal{P}_c(\mathcal{X})} e^{-\tilde{H}_n(\rho)} d[\tilde{\eta}_n^0(\xi, \cdot)](\rho) & \xi \in L_n(\text{supp}(\lambda^n P_n)), \\ 0 & \xi \notin L_n(\text{supp}(\lambda^n P_n)), \end{cases} \quad (3.126)$$

$$\tilde{\gamma}_n(\xi, D) = \begin{cases} \frac{\int_{\mathcal{Y}} \mathbb{1}_D(y) \tilde{Z}_{n+1}(\xi + \frac{1}{n+1}[\delta_y - \xi]) d[\tilde{\gamma}_n^0(\xi, \cdot)](y)}{\int_{\mathcal{Y}} \tilde{Z}_{n+1}(\xi + \frac{1}{n+1}[\delta_y - \xi]) d[\tilde{\gamma}_n^0(\xi, \cdot)](y)} & \xi \in L_n(\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1})), \\ 0 & \xi \notin L_n(\text{supp}(\lambda^{n+1} P_{n+1} \circ \sigma_{\perp}^{-1})). \end{cases} \quad (3.127)$$

By Theorem 3.6.11 follows that $\tilde{\eta}_n$ is the product regular conditional kernel under $[\nu_n \times P_n] \circ L_n^{-1} \circ \mathbf{m}^{-1}$ with respect to π that is weakly continuous on $L_n(\text{supp}(\lambda^n P_n))$.

ψ_n , being the function defined in (3.88), can be written as

$$\psi_n(x, w) = \int_{\mathcal{P}_c(\mathcal{X})} e^{-[\tilde{H}_{n+1}(\xi + \frac{1}{n+1}[\delta_x - \xi]) - \tilde{H}_n(\xi)]} d[\tilde{\eta}_n(L_n(w), \cdot)](\xi). \quad (3.128)$$

If, moreover $H_n = nV \circ L_n$, then

$$\psi_n(x, w) = \int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(L_n(w), \cdot)](\xi). \quad (3.129)$$

§3.7 Independent transformations: sufficient condition for sequential Gibbsianness

In this section we consider the situation as in Section 3.6 ((i) – (ix)). Additionally, we assume that H_n originates from some potential V and consider independent transformations: we assume that $P_n(x, \cdot) = \bigotimes_{i=1}^n P(x_i, \cdot)$ for $x = (x_1, \dots, x_n) \in \mathcal{X}^n$, where $P : \mathcal{X} \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ is a probability kernel. We assume that $\text{supp}(\lambda P) = \mathcal{Y}$. In Corollary 3.7.7 we give a sufficient condition for the transformed sequence $(\nu_n P_n)_{n \in \mathbb{N}}$ to be sequential Gibbs, in terms of the rate function of the initial empirical distribution (on $\mathcal{X}$) conditioned on the final (i.e., transformed) empirical distribution (on $\mathcal{Y}$).

First we state two preparatory lemma's, Lemma 3.7.2 and Lemma 3.7.3. Then in Theorem 3.7.6 we prove a convergence result under a large deviation principle, of which Corollary 3.7.7 is a consequence.

3.7.1. By Theorem 3.6.3 we have

$$\text{supp}(\nu_n P_n) = \mathcal{Y}^n = \text{supp}(\nu_{n+1} P_{n+1} \circ \sigma_{\perp}^{-1}). \quad (3.130)$$

Therefore $(\nu_n P_n)_{n \in \mathbb{N}}$ is sequentially Gibbs whenever there exists a specification kernel $\gamma : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ such that for all $\zeta \in \mathcal{P}_c(\mathcal{Y})$ and all $(w_n)_{n \in \mathbb{N}}$ with $w_n \in \mathcal{Y}^n$ and $L_n(w_n) \xrightarrow{C-w} \zeta$ we have the convergence $\gamma_n(w_n, \cdot) \xrightarrow{w} \gamma(\zeta, \cdot)$.

3.7.2 Lemma. Assume that V is continuously $\star$ -differentiable.

- (a) Let $\theta, \xi_0 \in \mathcal{P}_c(\mathcal{X})$ and $\varepsilon > 0$. Then there exists an open neighbourhood $A \subset \mathcal{P}_c(\mathcal{X})$ of ξ_0 and an $N \in \mathbb{N}$ such that

$$|n[V(\xi + \frac{1}{n}(\theta - \xi)) - V(\xi)] - \delta V(\xi_0, \theta)| < \varepsilon \quad (\xi \in A, n \geq N). \quad (3.131)$$

- (b) Let $K \subset \mathcal{P}_c(\mathcal{X})$ be compact and $\theta \in \mathcal{P}_c(\mathcal{X})$. Then there exists an $M > 0$ such that

$$(n+1) \left| V(\xi + \frac{1}{n+1}(\theta - \xi)) - V(\xi) \right| \leq M \quad (\xi \in K, n \in \mathbb{N}). \quad (3.132)$$

Proof. (a) Since δV is continuous, there exists a $\delta > 0$ and an open neighbourhood $A \subset \mathcal{P}_c(\mathcal{X})$ of ξ_0 for which

$$|\delta V(\xi, \theta, \alpha) - \delta V(\xi_0, \theta)| < \varepsilon \quad (\xi \in A, \alpha \in [0, \delta]). \quad (3.133)$$

By Lemma 3.5.8 with $N \in \mathbb{N}$ such that $\frac{1}{N} < \delta$, we obtain (3.131).

(b) follows from Lemma 3.5.8 by letting $M = \sup_{\xi \in K, \alpha \in [0, 1]} |\delta V(\xi, \theta, \alpha)|$. $\square$

3.7.3 Lemma. Let $I : \mathcal{Y} \rightarrow [0, \infty]$ have compact sublevel sets, i.e., $[I \leq c]$ is compact ($[I \leq c] = \{y \in \mathcal{Y} : I(y) \leq c\}$, see Section 3.2). Suppose that there exists a unique y_0 for which $I(y_0) = 0$. Then for every open neighbourhood A of y_0 there exists a $\beta > 0$ for which $I > \beta$ on A^c , i.e., $I(y) > \beta$ for all $y \in A^c$.

§3.7. Independent transformations: sufficient condition for sequential Gibbsianness

Proof. Let A be an open neighbourhood of y_0 . Let $F_n = [I \leq \frac{1}{n}] \cap A^c$ for $n \in \mathbb{N}$. Then $\bigcap_{n \in \mathbb{N}} [I \leq \frac{1}{n}] = \{y_0\}$ implies that $\bigcap_{n \in \mathbb{N}} F_n = \emptyset$. Therefore since F_n is contained in the compact set $[I \leq 1]$ for all $n \in \mathbb{N}$ and $F_1 \supset F_2 \supset \dots$, there exists an $N \in \mathbb{N}$ for which $F_N = \emptyset$, i.e., $I(y) > \frac{1}{N}$ for all $y \in A^c$. $\square$

3.7.4 Lemma. Assume that V is continuous and bounded from below. Let $\zeta \in \mathcal{P}_c(\mathcal{X})$ and $\zeta_n \in \mathcal{P}_{emp}^n(\mathcal{X})$ be such that $\zeta_n \xrightarrow{C-w} \zeta$. Assume that $(\tilde{\eta}_n^0(\zeta_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $J_\zeta : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty]$. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi) = -\inf[V + J_\zeta]. \quad (3.134)$$

Moreover, $(\tilde{\eta}_n(\zeta_n, \cdot))_{n \in \mathbb{N}}$ and the sequence of measures given by

$$A \mapsto \frac{\int_{\mathcal{P}_c(\mathcal{X})} \mathbb{1}_A(\xi) e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)}{\int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)} \quad (3.135)$$

satisfy the large deviation principle with rate function $I_\zeta : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty]$ given by

$$I_\zeta = V + J_\zeta - \inf[V + J_\zeta]. \quad (3.136)$$

Proof. This follows from Varadhan's Lemma (see den Hollander [50, Theorem III.13 and III.17]). $\square$

3.7.5 Remark. Instead of assuming that V is bounded from below, for convenience we will assume V to be positive in Lemma 3.7.6 and Theorem 3.7.7.

3.7.6 Lemma. Assume that V is continuous, positive and continuously $\star$ -differentiable. Let $\zeta \in \mathcal{P}_c(\mathcal{X})$ and $(w_n)_{n \in \mathbb{N}}$ with $w_n \in \mathcal{Y}^n$ be such that $\zeta_n = L_n(w_n) \xrightarrow{C-w} \zeta$. Assume that $(\tilde{\eta}_n^0(\zeta_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with a rate function $J_\zeta : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty]$ that has compact sublevel sets and is continuous on $[J_\zeta < \infty]$. If I_ζ (defined as in Lemma 3.7.4) has a unique global minimiser ξ_0 , then

$$c_n^{-1} \psi_n(x, w_n) \rightarrow e^{-\delta V(\xi_0, \delta_x)} \quad (x \in \mathcal{X}), \quad (3.137)$$

where $c_n \in (0, \infty)$ is given by

$$c_n = \int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi). \quad (3.138)$$

Moreover, if V satisfies the $\star$ -dominated bound, then

$$\gamma_n(w_n, \cdot) \rightarrow \gamma(\zeta, \cdot), \quad (3.139)$$

where $\gamma : \mathcal{P}_c(\mathcal{Y}) \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$ is given by

$$\gamma(\zeta, D) = \frac{\int_{\mathcal{X}} e^{-\delta V(\xi_0, \delta_x)} P(x, D) d\lambda(x)}{\int_{\mathcal{X}} e^{-\delta V(\xi_0, \delta_x)} d\lambda(x)}. \quad (3.140)$$

Proof. Let $x \in \mathcal{X}$. By Lemma 3.7.4, the sequence of probability measures $(\rho_n)_{n \in \mathbb{N}}$ on $\mathcal{P}_c(\mathcal{X})$ given by

$$\rho_n(A) = \frac{\int_{\mathcal{P}_c(\mathcal{X})} \mathbb{1}_A(\xi) e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)}{\int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)} \quad (3.141)$$

satisfies the large deviation principle with rate function I_ζ . Note that (see (3.129))

$$c_n^{-1} \psi_n(x, w_n) = \int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} d\rho_n(\xi). \quad (3.142)$$

Hence for $A, K \in \mathcal{B}(\mathcal{P}_c(\mathcal{X}))$ we have

$$\begin{aligned} & \left| c_n^{-1} \psi_n(x, w_n) - e^{-\delta V(\xi_0, \delta_x)} \right| \\ & \leq \mathfrak{B}_1(n, A) + \mathfrak{B}_2(n, A^c) + \mathfrak{B}_3(n, K \cap A^c) + \mathfrak{B}_4(n, K^c), \end{aligned} \quad (3.143)$$

where

$$\mathfrak{B}_1(n, A) = \int_A \left| e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} - e^{-\delta V(\xi_0, \delta_x)} \right| d\rho_n(\xi), \quad (3.144)$$

$$\mathfrak{B}_2(n, A^c) = e^{-\delta V(\xi_0, \delta_x)} \rho_n(A^c), \quad (3.145)$$

$$\mathfrak{B}_3(n, K \cap A^c) = \int_{K \cap A^c} e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} d\rho_n(\xi), \quad (3.146)$$

$$\mathfrak{B}_4(n, K^c) = \frac{\tilde{\eta}_n^0(\zeta_n, K^c)}{\int_{\mathcal{X}^n} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)}. \quad (3.147)$$

Let $\varepsilon > 0$. We show that there exist $A, K \subset \mathcal{P}_c(\mathcal{X})$ such that the four different summands are less than ε for large n .

(i) By Lemma 3.7.2(a) there exists an open neighbourhood A of ξ_0 and an $N_1 \in \mathbb{N}$ such that

$$\left| e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} - e^{-\delta V(\xi_0, \delta_x)} \right| < \varepsilon \quad (\xi \in A, n \geq N_1). \quad (3.148)$$

This implies that $\mathfrak{B}_1(n, A) < \varepsilon$ for $n \geq N_1$.

(ii) By Lemma 3.7.3 there exists a $\beta > 0$ such that $I_\zeta > \beta$ on A^c . By the large deviation principle $\lim_{n \rightarrow \infty} \rho_n(A^c) = 0$, so that $\mathfrak{B}_2(n, A^c) \rightarrow 0$.

Define

$$\kappa = \inf[V + J_\zeta], \quad (3.149)$$

$$K = [J_\zeta \leq \kappa + 4]. \quad (3.150)$$

(iii) Let $M > 0$ be as in Lemma 3.7.2(b). Then

$$\int_{K \cap A^c} e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} d\rho_n(\xi) \leq M \rho_n(A^c). \quad (3.151)$$

Therefore $\mathfrak{B}_3(n, K \cap A^c) \rightarrow 0$.

§3.7. Independent transformations: sufficient condition for sequential Gibbsianness

(iv) By the large deviation principle for $(\tilde{\eta}_n^0(\zeta_n, \cdot))_{n \in \mathbb{N}}$ and the fact that J_ζ is continuous on $[J_\zeta < \infty]$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \tilde{\eta}_n^0(\zeta_n, K^c) \leq -\inf J_\zeta(\overline{K^c}) = -\inf J_\zeta(K^c). \quad (3.152)$$

By the definition of K , there exists an $N_2 \in \mathbb{N}$ such that

$$\tilde{\eta}_n^0(\zeta_n, K^c) \leq e^{-n(\kappa+3)} \quad (n \geq N_2). \quad (3.153)$$

By (3.134) there exists an $N_3 \in \mathbb{N}$ such that

$$\int_{\mathcal{P}_c(\mathcal{X})} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi) \geq e^{-n(\kappa+1)} \quad (n \geq N_3). \quad (3.154)$$

Hence, for $n \geq \max\{N_2, N_3\}$,

$$\frac{\tilde{\eta}_n^0(\zeta_n, K^c)}{\int_{\mathcal{X}^n} e^{-(n+1)V(\xi)} d[\tilde{\eta}_n^0(\zeta_n, \cdot)](\xi)} \leq \tilde{\eta}_n^0(\zeta_n, K^c) e^{n(\kappa+1)} \leq e^{-2n}, \quad (3.155)$$

and so $\mathfrak{B}_4(n, K^c) \rightarrow 0$.

Suppose that V satisfies the $\star$ -dominated bound. Let $h : \mathcal{X} \rightarrow [0, \infty)$ be λ -integrable and $(\mathcal{X}_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(\mathcal{X})$ with $\mathcal{X}_n \uparrow \mathcal{X}$ and $\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda(\mathcal{X}_n^c) = -\infty$ be such that (3.39) holds for K as in (3.150). Let $L > 0$ be such that $\mathfrak{B}_4(n, K^c) \leq L$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned} c_n^{-1} \psi_n(x, w_n) &\leq \int_K e^{-(n+1)[V(\xi + \frac{1}{n+1}[\delta_x - \xi]) - V(\xi)]} d\rho_n(\xi) + \mathfrak{B}_4(n, K^c) \\ &\leq h(x) + L \quad (x \in \mathcal{X}_n, n \in \mathbb{N}). \end{aligned} \quad (3.156)$$

Moreover, with $M = \sup_{\xi \in K} V(\xi)$ we have for all $x \in \mathcal{X}$

$$c_n^{-1} \psi_n(x, w_n) \leq \int_K e^{(n+1)V(\xi)} d\rho_n(\xi) + L \leq e^{(n+1)M} + L. \quad (3.157)$$

By Lemma 3.5.17 we obtain for all $D \in \mathcal{B}(\mathcal{Y})$

$$\int_{\mathcal{X}} c_n^{-1} \psi_n(x, w_n) P(x, D) d\lambda(x) \rightarrow \int_{\mathcal{X}} e^{-\delta V(\zeta, \delta_x)} P(x, D) d\lambda(x). \quad (3.158)$$

By (3.90) it follows that $\gamma_n(w_n, \cdot) \rightarrow \gamma(\zeta, \cdot)$. $\square$

In the following Corollary we assume additionally that $\mathcal{X}$ is a σ -compact metric space, so that $\mathcal{P}_c(\mathcal{X}) = \text{LIM}[\mathcal{P}_{emp}^n(\mathcal{X})]$ by Theorem 3.3.16. Moreover, by 3.7.1 $\mathcal{P}_c(\mathcal{X}) = \text{LIM}[\text{supp}(\nu_{n+1} P_{n+1} \circ \sigma_\perp^{-1})]$.

3.7.7 Theorem. *Let $\mathcal{X}$ be a σ -compact metric space. Assume that $V : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty)$ is continuous and continuously $\star$ -differentiable and satisfies the $\star$ -dominated bound. Suppose that for all $\zeta \in \mathcal{P}_c(\mathcal{X})$ and $(\zeta_n)_{n \in \mathbb{N}}$ with $\zeta_n \in \mathcal{P}_{emp}^n(\mathcal{X})$ and $\zeta_n \xrightarrow{C-w} \zeta$ the sequence $(\tilde{\eta}_n^0(\zeta_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with a rate function*

$J_\zeta : \mathcal{P}_c(\mathcal{X}) \rightarrow [0, \infty]$ that has compact sublevel sets and is continuous on $[J_\zeta < \infty]$. Suppose, furthermore, that $I_\zeta := V + J_\zeta - \inf[V + J_\zeta]$ has a unique global minimiser ξ_ζ for all $\zeta \in \mathcal{P}_c(\mathcal{X})$. Then the sequence $(\nu_n P_n)_{n \in \mathbb{N}}$ is sequentially Gibbs with specification kernel given by

$$\gamma(\zeta, A) = \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) e^{-\delta V(\xi_\zeta, \delta x)} d\lambda(x)}{\int_{\mathcal{X}} e^{-\delta V(\xi_\zeta, \delta x)} d\lambda(x)} \quad (\zeta \in \mathcal{P}_c(\mathcal{X}), A \in \mathcal{B}(\mathcal{X})). \quad (3.159)$$

Appendix

§3.A Continuity of the limit of functions

3.A.1 Theorem. Let $(\mathcal{X}, d_{\mathcal{X}})$ and $(\mathcal{Y}, d_{\mathcal{Y}})$ be metric spaces. Let $D_1 \subset D_2 \subset \dots$ be an increasing sequence of subsets of $\mathcal{X}$ with $\text{LIM}[D_n] = \mathcal{X}$. Suppose that $f_n : D_n \rightarrow \mathcal{Y}$ and $f : \mathcal{X} \rightarrow \mathcal{Y}$ are such that, for all sequences $(x_n)_{n \in \mathbb{N}}$, with $x_n \in D_n$ and $x_n \rightarrow x$,

$$f_n(x_n) \rightarrow f(x). \quad (3.160)$$

Then f is continuous.

Proof. (1) Suppose that $(k_m)_{m \in \mathbb{N}}$ is a strictly increasing sequence in $\mathbb{N}$ and $x_m \in D_{k_m}$ is such that $x_m \rightarrow x$. Then $f_{k_m}(x_m) \rightarrow f(x)$ (note that $x_m \in D_{k_m}$ as $k_m \geq m$).

Proof of (1): Let $(z_n)_{n \in \mathbb{N}}$ be a sequence with $z_n \in D_n$ and $z_n \rightarrow x$. Let $(y_n)_{n \in \mathbb{N}}$ be the sequence given by

$$y_n = \begin{cases} x_m & n = k_m, \\ z_n & n \neq k_m \text{ for all } m \in \mathbb{N}. \end{cases} \quad (3.161)$$

Then $y_n \rightarrow x$ and so $f_n(y_n) \rightarrow f(x)$. In particular, $f_{k_m}(x_n) = f_{k_m}(y_{k_m}) \rightarrow f(x)$.

(2) If $(x_n)_{n \in \mathbb{N}}$ is a sequence with $x_n \in D_n$, $x \in \mathcal{X}$ and $x_n \rightarrow x$, then $f(x_n) \rightarrow f(x)$.

Proof of (2): Since $f_k(x_n) \xrightarrow{k \rightarrow \infty} f(x_n)$ for all $n \in \mathbb{N}$, we can find a strictly increasing sequence $(k_n)_{n \in \mathbb{N}}$ for which $d_{\mathcal{Y}}(f_{k_n}(x_n), f(x_n)) < \frac{1}{n}$ for all $n \in \mathbb{N}$. Hence $d_{\mathcal{Y}}(f(x_n), f(x)) \leq d_{\mathcal{Y}}(f_{k_n}(x_n), f(x_n)) + d_{\mathcal{Y}}(f_{k_n}(x_n), f(x)) \rightarrow 0$ by (1).

Now assume that $(x_n)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{X}$ converging to an $x \in \mathcal{X}$. For $n \in \mathbb{N}$, let $(y_m^n)_{m \in \mathbb{N}}$ be a sequence with $y_m^n \in D_m$ and $y_m^n \rightarrow x_n$ as $m \rightarrow \infty$. Choose a strictly increasing sequence $(m_n)_{n \in \mathbb{N}}$ in $\mathbb{N}$ for which

$$d_{\mathcal{Y}}(f(y_{m_n}^n), f(x_n)) < \frac{1}{n}, \quad d_{\mathcal{X}}(y_{m_n}^n, x_n) < \frac{1}{n}. \quad (3.162)$$

By considering a sequence $(z_k)_{k \in \mathbb{N}}$ with $z_k \in D_k$ and

$$z_k = y_{m_n}^n \quad (k \in \{m_n, m_n + 1, \dots, m_{n+1} - 1\}), \quad (3.163)$$

we have $f(z_k) \rightarrow f(x)$ as $k \rightarrow \infty$ by (2), and hence $f(y_{m_n}^n) \rightarrow f(x)$ as $n \rightarrow \infty$. It follows that $d_{\mathcal{Y}}(f(x_n), f(x)) \leq d_{\mathcal{Y}}(f(x_n), f(y_{m_n}^n)) + d_{\mathcal{Y}}(f(y_{m_n}^n), f(x)) \rightarrow 0$. $\square$

§3.B Tilted product regular conditional kernel

3.B.1 Lemma. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. Let ζ be a probability measure on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$. Let $f : X \rightarrow [0, \infty)$ and assume that $(x, y) \mapsto f(x)$ is ζ -integrable such that $\int_{X \times Y} f(x) d\zeta(x, y) > 0$. Define $\mu : \mathcal{A} \otimes \mathcal{B} \rightarrow [0, 1]$ to be the probability measure given by

$$\mu(D) = \frac{\int_{X \times Y} \mathbb{1}_D(x, y) f(x) d\zeta(x, y)}{\int_{X \times Y} f(x) d\zeta(x, y)} \quad (D \in \mathcal{A} \otimes \mathcal{B}). \quad (3.164)$$

Define ζ_2, μ_2 to be the probability measures on $\mathcal{B}$ given by

$$\zeta_2(B) = \zeta(X \times B), \quad \mu_2(B) = \mu(X \times B) \quad (B \in \mathcal{B}). \quad (3.165)$$

Suppose that $\theta : Y \times \mathcal{A} \rightarrow [0, 1]$ is a product regular conditional probability kernel under ζ with respect to the projection on the second coordinate. Then $\eta : Y \times \mathcal{A} \rightarrow [0, 1]$ defined by

$$\eta(y, A) = \begin{cases} \frac{\int_X \mathbb{1}_A(x) f(x) d[\theta(y, \cdot)](x)}{\int_X f(x) d[\theta(y, \cdot)](x)} & \text{if } \int_X f(x) d[\theta(y, \cdot)](x) > 0, \\ 0 & \text{if } \int_X f(x) d[\theta(y, \cdot)](x) = 0, \end{cases} \quad (3.166)$$

is a product regular conditional probability kernel under μ with respect to the projection on the second coordinate.

Proof. Write

$$Z = \int_{X \times Y} f(x) d\zeta(x, y). \quad (3.167)$$

Note that $\mu_2 = h\zeta_2$, where $h(y) = \frac{1}{Z} \int_X f(x) d[\theta(y, \cdot)](x)$. For $A \in \mathcal{A}$ and $B \in \mathcal{B}$,

$$\begin{aligned} & \int_{X \times Y} \mathbb{1}_{A \times B}(x, y) d[\eta(y, \cdot)](x) d\mu_2(y) \\ &= \frac{1}{Z} \int_Y \frac{\int_X \mathbb{1}_{A \times B}(x, y) f(x) d[\theta(y, \cdot)](x)}{h(y)} \mathbb{1}_{[h > 0]}(y) h(y) d\zeta_2(y) \\ &= \frac{1}{Z} \int_Y \int_X \mathbb{1}_{A \times B}(x, y) f(x) d[\theta(y, \cdot)](x) d\zeta_2(y) = \mu(A \times B). \end{aligned} \quad (3.168)$$

□



Large deviations of continuous regular conditional probabilities

This chapter is based on a preprint:

W. van Zuijlen. Large deviations of continuous regular conditional probabilities.

Preprint available at <http://arxiv.org/abs/1605.05192>.

Abstract

We study product regular conditional probabilities under measures of two coordinates with respect to the second coordinate that are weakly continuous on the support of the marginal of the second coordinate. Assuming that there exists a sequence of probability measures on the product space that satisfies a large deviation principle, we present necessary and sufficient conditions for the conditional probabilities under these measures to satisfy a large deviation principle. The arguments of these conditional probabilities are assumed to converge. A way to view regular conditional probabilities as a special case of product regular conditional probabilities is presented. This is used to derive conditions for large deviations of regular conditional probabilities. In addition, we derive a Sanov-type theorem for large deviations of the empirical distribution of the first coordinate conditioned on fixing the empirical distribution of the second coordinate.

§4.1 Introduction and main results

In the present paper we study large deviations of probabilities “of the form”

$$\mathbb{P}(X_n \in A \mid Y_n = y_n), \quad (4.1)$$

where $((X_n, Y_n))_{n \in \mathbb{N}}$ is a sequence of couples of random variables that satisfies a large deviation principle and $y_n \rightarrow y$ for some y . As the event $[Y_n = y_n]$ may have probability zero, we make sense of (4.1) in terms of a kernel η_n , so that

$$\eta_n(y_n, A)$$

“represents” (4.1).

Such kernels are called regular conditional probabilities and form an important object in probability theory. The existence of regular conditional probabilities has been studied extensively, for example, by Faden [37] or by Leao, Fragoso and Fuffino [63]. There exist in fact various forms of regular conditional probabilities; namely either with respect to a σ -algebra, with respect to a measurable map, or with respect to the projection on one of the coordinates (in case of a product space).

In order to consider large deviations of conditional probabilities, we have to specify which conditional probability we are considering; the conditional probability may not be unique. However, if a (product) regular conditional probability is weakly continuous on the support of the measure composed with the inverse of the measurable map (or projection), it is unique on that domain. For these (product) regular conditional probabilities it is natural to study their large deviations, whenever the argument of the probability is in the domain on which it is unique. In this paper we study the large deviations in the case when the arguments of these kernels converge, i.e., we study large deviations of $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ for the case that $y_n \rightarrow y$. To the best of our knowledge, current literature does not provide a general condition under which such kernels satisfy a large deviation principle.

§4.1.1 Literature

Some examples in this direction are present. For example in Adams, Dirr, Peletier and Zimmer [1], the large deviation principle is proved for the empirical distribution that is evolved by independent Brownian motions conditioned on their initial empirical distribution to lie in a ball (see [1, Theorem 1]). They proceed by proving that the large deviation principle rate function converges as the radius of the ball converges to zero. For the purpose of this paper, we have to show that the limit of the radius of the ball and the limit belonging to the large deviation principle can be interchanged. Léonard [64] proves the large deviation principle of the empirical distribution that is evolved by independent Brownian motions conditioned on their initial empirical distribution; those initial empirical distributions are assumed to be converging (see [64, Proposition 2.19]). In both papers, the evolved state is conditioned on the initial state, while there is also interest in large deviations of the initial state conditioned on the evolved state. In this paper we prove the large deviation principle in this setting for finite state spaces.

There exist various results on quenched large deviations, i.e., large deviations for regular conditional probabilities in the sense that for almost all realisations of the disorder the conditional probabilities satisfy the large deviation principle with a rate function that does not depend on the disorder. Examples of papers on quenched large deviations are Comets [19] for conditional large deviations of i.i.d. random fields, Greven and den Hollander [45] and Comets, Gantert and Zeitouni [20] for random walks in random environments, Kosygina-Rezakhanlou-Varadhan [57], for a diffusion with a random drift, Rassoul-Agha, Seppäläinen and Yilmaz [76] for polymers in a random potential.

Biggins [5] obtains the large deviation principle for mixtures of probability measures that satisfy the large deviation principle with kernels that satisfy the large deviation principle as their arguments converge. To some extent we complement the article in the opposite direction, in the sense that we assume the large deviation principle of the mixture and derive the large deviation principle of the kernels.

Our main motivation to study the above large deviations lies in the theory of Gibbs-non-Gibbs transitions. There is a correspondence between the large deviation rate function of the conditional probability with respect to the evolved coordinate and the evolved state (measure or sequence) being Gibbs (see van Enter, Fernández, den Hollander and Redig [31]). We refer to Section 4.1.4 for further discussions on Gibbs-non-Gibbs transitions.

§4.1.2 Large deviations

In the literature on large deviations two dominant definitions of large deviation principles are used. One is in terms of a σ -algebra on the topological space, as is done in the book by Dembo and Zeitouni [24] and in the book by Deuschel and Stroock [26], the other is in terms of the topology, i.e., in terms of open and closed sets, as is done in the book by den Hollander [50] and in the book by Rassoul-Agha and Seppäläinen [75]. Whenever one considers the Borel- σ -algebra on the topological space, the two definitions agree.

We define the large deviation lower bound and the large deviation upper bound separately, as in Section 4.1.3 and in Section 4.6 we describe the necessary and sufficient conditions for each of the bounds separately. Moreover, we define them on a set of subsets of the topological space, which is not required to be a σ -algebra. In Remark 4.7.4 we motivate the choice for this definition.

4.1.1 Definition. Let $\mathcal{X}$ be a topological space and $\mathcal{A}$ be a set of subsets of $\mathcal{X}$. Let $I : \mathcal{X} \rightarrow [0, \infty]$ be lower semicontinuous. Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence of probability measures on $\mathcal{A}$. Let $(r_n)_{n \in \mathbb{N}}$ be an increasing sequence in $(0, \infty)$ with $\lim_{n \rightarrow \infty} r_n = \infty$. We say that $(\mu_n)_{n \in \mathbb{N}}$ satisfies a *large deviation lower bound* on $\mathcal{A}$ with rate function I and rates $(r_n)_{n \in \mathbb{N}}$ if

$$\liminf_{n \rightarrow \infty} \frac{1}{r_n} \log \mu_n(A) \geq -\inf I(A^\circ) \quad (A \in \mathcal{A}). \quad (4.2)$$

We say that $(\mu_n)_{n \in \mathbb{N}}$ satisfies a *large deviation upper bound* on $\mathcal{A}$ with rate function

I and rates $(r_n)_{n \in \mathbb{N}}$ if

$$\limsup_{n \rightarrow \infty} \frac{1}{r_n} \log \mu_n(A) \leq -\inf I(\bar{A}) \quad (A \in \mathcal{A}). \quad (4.3)$$

In the rest of the paper we only consider the rates $r_n = n$. However, the theory presented is still valid for general rates $(r_n)_{n \in \mathbb{N}}$. We say that $(\mu_n)_{n \in \mathbb{N}}$ satisfies a *large deviation principle* on $\mathcal{A}$ with rate function I whenever it satisfies both the large deviation lower bound and the large deviation upper bound with rate function I .

We omit “on $\mathcal{A}$ ” whenever $\mathcal{A}$ is the Borel- σ -algebra $\mathcal{B}(\mathcal{X})$ on $\mathcal{X}$. In this case the large deviation lower bound is satisfied if and only if the inequality in (4.2) holds for all open subsets of $\mathcal{X}$ and the large deviation upper bound is satisfied if and only if the inequality in (4.3) holds for all closed subsets of $\mathcal{X}$.

§4.1.3 Main results

See Section 4.3 and Section 4.4 for the definitions of the objects in the statements of the following theorems. In Section 4.6 and Section 4.7 we consider a more general situation. Theorem 4.1.2 is a consequence of Theorem 4.6.9 and Theorem 4.1.3 is a consequence of Theorem 4.7.5.

In this section $\mathcal{X}$ and $\mathcal{Y}$ are a metric spaces.

4.1.2 Theorem. *Let $\pi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ be given by $\pi(x, y) = y$. Suppose that $(\mu_n)_{n \in \mathbb{N}}$ is a sequence of probability measures on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ that satisfies the large deviation principle with rate function $J : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty]$ that has compact sublevel sets. Suppose that for each $n \in \mathbb{N}$ there exists a product regular conditional probability $\eta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under μ_n with respect to π that is weakly continuous on $\text{supp}(\mu_n \circ \pi^{-1})$, which we assume to be nonempty. Let $y \in \mathcal{Y}$ be such that $\inf J(\mathcal{X} \times \{y\}) < \infty$. Define $I : \mathcal{X} \rightarrow [0, \infty]$ by*

$$I(x) = J(x, y) - \inf J(\mathcal{X} \times \{y\}). \quad (4.4)$$

I has compact sublevel sets, and, for each $n \in \mathbb{N}$, η_n is unique on $\text{supp}(\mu_n \circ \pi^{-1})$. Moreover,

$$(A1) \iff (A2) \text{ and } (B1) \iff (B2),$$

where

(A1) *For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ for all n large enough,¹ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation lower bound with rate function I .*

(A2) *For all $x \in \mathcal{X}$ and $r > 0$, with $U = B(x, r)$,*

$$\sup_{\varepsilon > 0} \liminf_{n \rightarrow \infty} \inf_{\substack{z \in \mathcal{Y}, \delta \in (0, \varepsilon) \\ B(z, \delta) \subset B(y, \varepsilon)}} \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} \mid \mathcal{X} \times B(z, \delta)) \geq -\inf I(U). \quad (4.5)$$

¹Meaning that there exists an $N \in \mathbb{N}$ such that $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ for all $n \geq N$.

- (B1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ for all n large enough, the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound with rate function I .
- (B2) For all $x_1, \dots, x_k \in \mathcal{X}$ and $r_1, \dots, r_k > 0$, with $W = \mathcal{X} \setminus [B(x_1, r_1) \cup \dots \cup B(x_k, r_k)]$,

$$\inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \sup_{\substack{z \in \mathcal{Y}, \delta \in (0, \varepsilon) \\ B(z, \delta) \subset B(y, \varepsilon)}} \frac{1}{n} \log \mu_n \left(W^\circ \times \mathcal{Y} \mid \mathcal{X} \times B(z, \delta) \right) \leq -\inf I(W). \quad (4.6)$$

The next theorem is similar to Theorem 4.1.2, but considers the large deviation bounds for regular conditional kernels instead of product regular conditional probabilities.

4.1.3 Theorem. Let $\tau : \mathcal{X} \rightarrow \mathcal{Y}$ be continuous. Suppose that $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of probability measures on $\mathcal{B}(\mathcal{X})$ that satisfies the large deviation principle with rate function $J : \mathcal{X} \rightarrow [0, \infty]$ that has compact sublevel sets. Suppose that for each $n \in \mathbb{N}$ there exists a regular conditional probability $\eta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under ν_n with respect to τ that is weakly continuous on $\text{supp}(\nu_n \circ \tau^{-1})$, which is assumed to be nonempty. Let $y \in \mathcal{Y}$ be such that $\inf J(\tau^{-1}(\{y\})) < \infty$. Define $I : \mathcal{X} \rightarrow [0, \infty]$ by

$$I(x) = \begin{cases} J(x) - \inf J(\tau^{-1}(\{y\})) & \tau(x) = y, \\ \infty & \tau(x) \neq y. \end{cases} \quad (4.7)$$

I has compact sublevel sets, and, for each $n \in \mathbb{N}$, η_n is unique on $\text{supp}(\nu_n \circ \tau^{-1})$. Moreover,

$$(A1) \iff (A2) \text{ and } (B1) \iff (B2),$$

where

- (A1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and $y_n \in \text{supp}(\nu_n \circ \pi^{-1})$ for all n large enough, the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation lower bound with rate function I .
- (A2) For all $x \in \mathcal{X}$ and $r > 0$, with $U = B(x, r)$,

$$\sup_{\varepsilon > 0} \liminf_{n \rightarrow \infty} \inf_{\substack{z \in \mathcal{Y}, \delta \in (0, \varepsilon) \\ B(z, \delta) \subset B(y, \varepsilon)}} \frac{1}{n} \log \nu_n \left(\overline{U} \mid \tau^{-1}(B(z, \delta)) \right) \geq -\inf I(U). \quad (4.8)$$

- (B1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and $y_n \in \text{supp}(\nu_n \circ \pi^{-1})$ for all n large enough, the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound with rate function I .
- (B2) For all $x_1, \dots, x_k \in \mathcal{X}$ and $r_1, \dots, r_k > 0$, with $W = \mathcal{X} \setminus [B(x_1, r_1) \cup \dots \cup B(x_k, r_k)]$,

$$\inf_{\varepsilon > 0} \limsup_{n \rightarrow \infty} \sup_{\substack{z \in \mathcal{Y}, \delta \in (0, \varepsilon) \\ B(z, \delta) \subset B(y, \varepsilon)}} \frac{1}{n} \log \nu_n \left(W^\circ \mid \tau^{-1}(B(z, \delta)) \right) \leq -\inf I(W). \quad (4.9)$$

§4.1.4 Gibbs-non-Gibbs transitions and future research

In this section we discuss the relation between the large deviation results in this paper and Gibbs-non-Gibbs transitions in more detail. In particular, we discuss possible future directions regarding large deviations of conditional kernels.

The following situation for interacting particle systems occurs in the mean-field context (a similar context holds in the context of lattices). The initial system of -so called- spins consists of distributions describing the interaction between spins via a potential V (for each n there is a distribution describing the law of n spins). This initial system is assumed to be Gibbs, which is called sequentially Gibbs in the mean-field context. Allowing the initial state to be transformed, for example, by an evolution of the spins, a question of interest is whether the transformed state is (sequentially) Gibbs. This question has been addressed in the mean-field context by Ermoleav and Külske [36] and by Fernández, den Hollander and Martínez [39] for $\{-1, +1\}$ -valued spins, by den Hollander, Redig and van Zuijlen [52] for $\mathbb{R}$ -valued spins and by Külske and Opoku [61] and van Enter, Külske, Opoku and Ruszel [33] for compactly valued spins. In these papers, independent dynamics of the spins are considered (the evolution of each spin is independent of the evolution of the other spins). Independent dynamics simplify the situation. Namely, the evolved measure on either the product space of the initial and the final space, or -in case of an evolution- the space of trajectories, is a tilted measure of the evolved measure when considering $V = 0$. In this case the measure is a product measure, which means that the spins are independent. As a consequence (this will be clarified in a forthcoming paper) the conditional kernel η_n of the initial state on n spins with respect to the final state (for a fixed potential V) is a tilted version of the conditional kernel η_n^0 of the initial state with respect to the final state of independent spins (i.e., $V = 0$). Because of this tilting, by Varadhan's lemma, $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $V + I_y - \inf(V + I_y)$ if $(\eta_n^0(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I_y . In the forthcoming paper we will prove that the evolved sequence is sequentially Gibbs if $V + I_\zeta$ has a unique global minimiser.

The large deviation principle of $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ has been mentioned in the case of trajectories in [36, Corollary 2.4] and -as a corollary of that theorem- for the case of the product space of the initial and the final space in [39, Corollary 1.3]. However, no proof was given. Theorem 4.8.2 provides a rigorous proof of the large deviation principle statement in [39, Corollary 1.3]. In this paper we do not provide a rigorous proof of [36, Corollary 2.4]. But Theorem 4.1.3 may be used, as the conditioning on the final state is a regular conditional kernel with respect to the map $\tau : C([0, T], \mathcal{X}) \rightarrow \mathcal{X}$, $\tau(f) = f(T)$.

In order to deal with empirical distributions (and not with magnetisations as is done in [52]), in future research we strive to 'extend' the statement of Theorem 4.8.2 to infinite and possibly non-compact state spaces. In the case of non-compact spaces it may be that topologies on the space of probability measures are considered that are not metrisable.

§4.1.5 Outline

We list some notations, definitions and assumptions in Section 4.2. In Section 4.3 we give and compare the notions of regular conditional kernels, we show that a regular conditional kernel under a measure ν is in fact a product regular conditional kernel under a measure that is related to ν . In Section 4.4 we introduce and study weakly continuous regular conditional kernels. In Section 4.5 we present some facts about lower semicontinuous functions with compact sublevel sets. Relying on the results of Sections 4.4 and 4.5, in Section 4.6 we present results on large deviation bounds for product regular conditional probabilities, in particular, necessary and sufficient conditions for these bounds to hold. In Section 4.7 we discuss how to obtain large deviation bounds for regular conditional probabilities from the results in Section 4.6. In Section 4.8 we apply the theory to obtain the large deviation principle for the empirical density of the first coordinate given the empirical density of the second coordinate, for independent and identically distributed pairs of random variables. In Section 4.9 we give some examples. We also include an example for which the conditions are not satisfied. For this example we compare the quenched large deviations with large deviations of the weakly continuous regular conditional probabilities and comment on the difference with an example by La Cour and Schieve [22]. In appendices 4.A and 4.B we state some general results considering large deviations bounds that are used in the different sections. In appendix 4.C we provide the proof of a theorem on which the examples of Section 4.9 rely.

§4.2 Notations and conventions

$\mathbb{N} = \{1, 2, 3, \dots\}$. For a topological space $\mathcal{X}$ we write $\mathcal{B}(\mathcal{X})$ for the Borel- σ -algebra and $\mathcal{P}(\mathcal{X})$ and $\mathcal{M}(\mathcal{X})$ for the spaces of probability and signed measures on $\mathcal{B}(\mathcal{X})$, respectively. For $A \subset \mathcal{X}$ we write A° for the interior of A and $\bar{A}$ for the closure of A . For $x \in \mathcal{X}$ we write δ_x for the element in $\mathcal{P}(\mathcal{X})$ with $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise. For $x \in \mathcal{X}$ we write $\mathcal{N}_x$ for the set of $\mathcal{B}(\mathcal{X})$ -measurable neighbourhoods of x . For a $\mu \in \mathcal{M}(\mathcal{X})$ we write $\text{supp } \mu = \{x \in \mathcal{X} : |\mu|(V) > 0 \text{ for all } V \in \mathcal{N}_x\}$ and call this the support of μ . For a function f from a set $\mathcal{X}$ into $\mathbb{R}$ and $c \in \mathbb{R}$ we write $[f \geq c] = \{x \in \mathcal{X} : f(x) \geq c\}$. Similarly, we use the notations $[f > c]$, $[f \leq c]$ and $[f < c]$. Whenever $(x_\iota)_{\iota \in \mathbb{I}}$ is a net, where $\mathbb{I}$ is a directed set by (a direction) $\preceq$, we write $\liminf_{\iota \in \mathbb{I}} x_\iota = \sup_{\iota_0 \in \mathbb{I}} \inf_{\iota \succeq \iota_0, \iota \in \mathbb{I}} x_\iota$ (similarly $\limsup$). In particular, if $\mathcal{V} \subset \mathcal{N}_x$ and $\bigcap \mathcal{V} = \{x\}$ and $f : \mathcal{V} \rightarrow \mathbb{R}$ we write $\liminf_{V \in \mathcal{V}} f(V) = \sup_{V_0 \in \mathcal{V}} \inf_{V \subset V_0, V \in \mathcal{V}} f(V)$ (i.e., we consider $(f(V))_{V \in \mathcal{V}}$ as a net where $\mathcal{V}$ is directed by $\supset$ (as $\preceq$)).

Whenever we write $\mu(A|B)$ we implicitly assume that it is well defined (as $\mu(A \cap B)/\mu(B)$), i.e., that $\mu(B) \neq 0$.

We use the conventions $\log 0 = -\infty$ and $\inf I(\emptyset) = \infty$ whenever I is a function with values in $[0, \infty]$.

All measures in this paper are signed measures, unless mentioned otherwise.

§4.3 Regular conditional kernels being product regular conditional kernels

In this section we introduce the notion of a (product) regular conditional kernel. For an extensive study on regular conditional kernels see Bogachev [10, Section 10.4]. The notion of a product regular conditional kernel does not appear in [10], but it does in Faden [37] and in Leao, Frago and Fuffino [63]. Besides giving definitions we make a few observations, of which Theorem 4.3.6 is used later on to derive statements of regular conditional kernels from statements of product regular conditional kernels.

In this section $(X, \mathcal{A})$, $(Y, \mathcal{B})$ are measurable spaces, ν is a measure on $\mathcal{A}$ and μ is a measure on $\mathcal{A} \otimes \mathcal{B}$, $\tau : X \rightarrow Y$ is measurable and $\pi : X \times Y \rightarrow Y$ is given by $\pi(x, y) = y$.

4.3.1 Definition. A function $\eta : Y \times \mathcal{A} \rightarrow \mathbb{R}$ is called a $(\mathcal{B})$ -kernel if $\eta(\cdot, A)$ is $(\mathcal{B})$ -measurable for all $A \in \mathcal{A}$ and $\eta(y, \cdot)$ is a measure for all $y \in Y$. A kernel η is called a *probability kernel* if $\eta(y, \cdot)$ is a probability measure for all $y \in Y$.

4.3.2 Definition. Let $\eta : Y \times \mathcal{A} \rightarrow \mathbb{R}$ be a (probability) kernel.

- (a) η is called a *regular conditional kernel* (*regular conditional probability*) under ν with respect to τ if

$$\nu(F \cap \tau^{-1}(B)) = \int_Y \mathbb{1}_B(y) \eta(y, F) \, d[|\nu| \circ \tau^{-1}](y) \quad (F \in \mathcal{F}, B \in \mathcal{B}). \quad (4.10)$$

- (b) η is called a *product regular conditional kernel* (*product regular conditional probability*) under μ with respect to π if

$$\mu(A \times B) = \int_Y \mathbb{1}_B(y) \eta(y, A) \, d[|\mu| \circ \pi^{-1}](y) \quad (A \in \mathcal{A}, B \in \mathcal{B}). \quad (4.11)$$

4.3.3. Suppose that $\mathcal{E}$ is a sub- σ -algebra of $\mathcal{F}$. Let $(Y, \mathcal{B}) = (X, \mathcal{E})$ and $\text{Id} : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ be the identity map. In agreement of [10, Definition 10.4.1] a kernel $\eta : Y \times \mathcal{A} \rightarrow \mathbb{R}$ is a *regular conditional kernel* under μ with respect to $\mathcal{E}$ if and only if η is a regular conditional kernel under μ with respect to Id .

4.3.4. Consider the two kernels $\eta : Y \times \mathcal{A} \rightarrow \mathbb{R}$ and $\xi : Y \times (\mathcal{A} \otimes \mathcal{B}) \rightarrow \mathbb{R}$, corresponding to each other by the formulas $\xi(y, F) = \int_{\mathcal{X}} \mathbb{1}_F(x, y) \, d[\eta(y, \cdot)](x)$ and $\eta(y, A) = \xi(y, A \times Y)$. Then ξ is a regular conditional kernel under μ given π if and only if η is a product regular conditional kernel under μ given π .

In general, $X \times Y$ may be equipped with a σ -algebra $\mathcal{F}$ different from $\mathcal{A} \otimes \mathcal{B}$. In this situation, where μ is a measure on $\mathcal{F}$ and π is $\mathcal{F}$ -measurable the above correspondence cannot be used in general to reduce statements about product regular conditional kernels to statements about regular conditional kernels. See also example 4.4.5.

On the other hand, regular conditional probabilities can be seen as special cases of product regular conditional probabilities, see Theorem 4.3.6. In the present paper we use this to derive Theorem 4.1.3 from Theorem 4.1.2 but also Theorem 4.7.5 from Theorem 4.6.9.

4.3.5 Remark. If $\mathcal{A}$ is generated by a countable set, two regular conditional probabilities under a measure with respect to a σ -algebra (see 4.3.3) are almost everywhere equal (see Bogachev [10, Theorem 10.4.3]). Similarly one could state an analogous statement for regular conditional kernels with respect to measurable maps and for product regular conditional kernels. In Theorem 4.4.3 we prove that (product) regular conditional kernels are unique on the domain on which they are weakly continuous, in case the underlying topological space is perfectly normal. For such space the Borel- σ -algebra may not be generated by a countable set.²

4.3.6 Theorem. (a) *There exists a measure $\tilde{\mu}$ on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ for which $\tilde{\mu}(A \times B) = \nu(A \cap \tau^{-1}(B))$.*

(b) *$\eta : Y \times \mathcal{A} \rightarrow \mathbb{R}$ is a regular conditional kernel under ν with respect to τ if and only if η is a product conditional kernel under $\tilde{\mu}$ with respect to π .*

Proof. (a) We may assume ν to be positive, since $\nu = \nu^+ - \nu^-$. Let $\mathcal{E}$ be the set that consists of $\bigcup_{i=1}^n A_i \times B_i$, where $n \in \mathbb{N}$ and $A_i \in \mathcal{A}, B_i \in \mathcal{B}$ are such that $A_1 \times B_1, \dots, A_n \times B_n$ are disjoint. Define $\nu^* : \mathcal{E} \rightarrow [0, \infty)$ by $\nu^*(\bigcup_{i=1}^n A_i \times B_i) = \nu(\bigcup_{i=1}^n A_i \cap \tau^{-1}(B_i))$ for $A_1, \dots, A_n \in \mathcal{A}$ and $B_1, \dots, B_n \in \mathcal{B}$ as above. Checking that $\mathcal{E}$ is a ring of sets and that ν^* is σ -additive is left for the reader. The existence and unicity of the extension $\tilde{\mu}$ follows from the Carathéodory Theorem (see Halmos [46, Section 13, Theorem A]).

(b) Follows from by definition of $\tilde{\mu}$ (note that $\nu \circ \tau^{-1} = \tilde{\mu} \circ \pi^{-1}$). $\square$

§4.4 Weakly continuous kernels

In this section we introduce the notion of weak continuity for kernels on topological spaces. In Theorem 4.4.3 we show uniqueness of (product) regular conditional kernels that are weakly continuous. In Theorem 4.4.6 and Theorem 4.4.7 we describe conditions that imply the existence of weakly continuous regular conditional probabilities. Similarly as is done in the Portmanteau Theorem when one considers metric spaces, weak convergence implies lower bounds for open sets and upper bounds for closed sets, as is shown in Theorem 4.4.9. As described in Lemma 4.4.10 and Lemma 4.4.11 these $\liminf$ and $\limsup$ bounds imply bounds for (product) regular conditional probabilities on which the results of Sections 4.6 and 4.7 are based.

In this section $\mathcal{X}$ and $\mathcal{Y}$ are topological spaces, ν is a measure on $\mathcal{B}(\mathcal{X})$ and μ is a measure on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$, $\tau : \mathcal{X} \rightarrow \mathcal{Y}$ is measurable and $\pi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ is given by $\pi(x, y) = y$.

4.4.1 Definition. We equip the space of measures, $\mathcal{M}(\mathcal{X})$, with the weak topology (generated by $C_b(\mathcal{X})$, and denoted by $\sigma(\mathcal{M}(\mathcal{X}), C_b(\mathcal{X}))$ as in the book of Schaefer [83, Chapter II, Section 5]). In this topology, a net $(\mu_\iota)_{\iota \in I}$ in $\mathcal{M}(\mathcal{X})$ converges to a μ in $\mathcal{M}(\mathcal{X})$ if $\int_{\mathcal{X}} f d\mu_\iota \rightarrow \int_{\mathcal{X}} f d\mu$ for all $f \in C_b(\mathcal{X})$.

Let $D \subset \mathcal{Y}$. A kernel $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}$ is called *weakly continuous* on D if the

²The Sorgenfrey line, the space $\mathbb{R}$ with the right half-open interval topology, is perfectly normal but not second countable (see Steen and Seebach [86, Example 51]).

map $D \rightarrow \mathcal{M}(\mathcal{X})$ given by $y \mapsto \eta(y, \cdot)$ is continuous in the weak topology. η is called *weakly continuous* if η is weakly continuous on $\mathcal{Y}$.

4.4.2 Theorem. *Let $\mathcal{X}$ be a perfectly normal³ space and $\mu \in \mathcal{M}(\mathcal{X})$. Then*

$$\text{supp } \mu = \left\{ x \in \mathcal{X} : \int_{\mathcal{X}} f \, d|\mu| > 0 \text{ for all } f \in C(\mathcal{X}, [0, 1]) \text{ with } f(x) > 0 \right\}. \quad (4.12)$$

Moreover, $|\mu|(\mathcal{X} \setminus \text{supp}(\mu)) = 0$.⁴ As a consequence, $\mu = 0$ if and only if $\int_{\mathcal{X}} f \, d|\mu| = 0$ for all $f \in C_b(\mathcal{X})$.

Proof. We may assume μ is positive. Let $x \in \text{supp } \mu$. Then $\mu(V) > 0$ for all $V \in \mathcal{N}_x$. Let $f \in C(\mathcal{X}, [0, 1])$ be such that $f(x) > 0$. Then $V = f^{-1}(0, \infty)$ has strictly positive measure. Since $\mu(V) = \lim_{n \rightarrow \infty} \int_{\mathcal{X}} \min\{nf, 1\} \, d\mu$, there exists an n such that $\int_{\mathcal{X}} \min\{nf, 1\} \, d\mu > 0$. Consequently, as $f \geq \frac{1}{n} \min\{nf, 1\}$, we have $\int_{\mathcal{X}} f \, d\mu > 0$.

Let $x \in \mathcal{X}$ be such that $\int_{\mathcal{X}} f \, d\mu > 0$ for all $f \in C(\mathcal{X}, [0, 1])$ with $f(x) > 0$. Let $V \in \mathcal{N}_x$. As $V = f^{-1}(0, \infty)$ for some $f \in C(\mathcal{X}, [0, 1])$, we have $\mu(V) \geq \int_{\mathcal{X}} f \, d\mu > 0$. $\square$

4.4.3 Theorem. *Suppose that $\mathcal{X}$ is a perfectly normal space.*

- (a) *Let η and ζ be regular conditional kernels under ν with respect to τ that are weakly continuous on $\text{supp}(|\nu| \circ \tau^{-1})$. Then $\eta(y, \cdot) = \zeta(y, \cdot)$ for all $y \in \text{supp}(|\nu| \circ \tau^{-1})$. If ν is a probability measure, then $\eta(y, \cdot)$ is a probability measure for all $y \in \text{supp}(|\nu| \circ \tau^{-1})$.*
- (b) *Let η and ζ be product regular conditional kernels under μ with respect to π that are weakly continuous on $\text{supp}(|\mu| \circ \pi^{-1})$. Then $\eta(y, \cdot) = \zeta(y, \cdot)$ for all $y \in \text{supp}(|\mu| \circ \pi^{-1})$. If μ is a probability measure, then $\eta(y, \cdot)$ is a probability measure for all $y \in \text{supp}(|\mu| \circ \pi^{-1})$.*

Proof. We prove (a), the proof of (b) is similar (replace “ $|\nu| \circ \tau^{-1}$ ” by “ $|\mu| \circ \pi^{-1}$ ”). To prove $\eta = \zeta$ on $D = \text{supp}(|\nu| \circ \tau^{-1})$, by Theorem 4.4.2, it is sufficient to prove $\int_{\mathcal{X}} f \, d\eta(y, \cdot) = \int_{\mathcal{X}} f \, d\zeta(y, \cdot)$ for all $y \in D$ and all $f \in C_b(\mathcal{X})$. Let $f \in C_b(\mathcal{X})$. Because f is the uniform limit of simple functions, one has for all $B \in \mathcal{B}(\mathcal{Y})$

$$\int_{\mathcal{Y}} \mathbb{1}_B(y) \left[\int_{\mathcal{X}} f \, d\eta(y, \cdot) \right] d[|\nu| \circ \tau^{-1}](y) = \int_{\mathcal{Y}} \mathbb{1}_B(y) \left[\int_{\mathcal{X}} f \, d\zeta(y, \cdot) \right] d[|\nu| \circ \tau^{-1}](y). \quad (4.13)$$

Therefore there exists a set $Z \in \mathcal{B}(\mathcal{Y})$ with $|\nu| \circ \tau^{-1}(\mathcal{Y} \setminus Z) = 0$ such that

$$\int_{\mathcal{X}} f \, d\eta(z, \cdot) = \int_{\mathcal{X}} f \, d\zeta(z, \cdot) \quad (z \in Z). \quad (4.14)$$

Since both $y \mapsto \int_{\mathcal{X}} f \, d\eta(y, \cdot)$ and $y \mapsto \int_{\mathcal{X}} f \, d\zeta(y, \cdot)$ are weakly continuous on D , and Z is dense in D by Theorem 4.4.2, we have $\int_{\mathcal{X}} f \, d\eta(y, \cdot) = \int_{\mathcal{X}} f \, d\zeta(y, \cdot)$ for all $y \in D$. The second statement is proved by taking $f = \mathbb{1}_{\mathcal{X}}$. $\square$

³Perfectly normal means that every open set in $\mathcal{X}$ is equal to $f^{-1}((0, \infty))$ for some $f \in C(\mathcal{X})$. All metric spaces are perfectly normal; Bogachev [10, Proposition 6.3.5].

⁴This is not true in general. For an example see Bogachev [10, Example 7.1.3].

4.4.4. When η is a regular conditional kernel under ν with respect to τ , the value of the function $\eta(\cdot, A)$ on the complement of $\text{supp}(|\nu| \circ \tau^{-1})$ is not determined, in the sense that, if $\tilde{\eta}$ is a kernel with $\tilde{\eta}(y, \cdot) = \eta(y, \cdot)$ for all $y \in \text{supp}(|\nu| \circ \tau^{-1})$, then $\tilde{\eta}$ is also a regular conditional kernel under ν with respect to τ .

For example $\tilde{\eta}$ given by $\tilde{\eta}(y, \cdot) = \eta(y, \cdot)$ for $y \in \text{supp}(|\nu| \circ \tau^{-1})$ and $\tilde{\eta}(y, \cdot) = \delta_x$ for $y \in \text{supp}(|\nu| \circ \tau^{-1})^c$ for some chosen $x \in \mathcal{X}$, is such regular conditional kernel.

Whence if ν is a probability measure and there exists a regular conditional kernel under ν with respect to τ that is weakly continuous on $\text{supp}(|\nu| \circ \tau^{-1})$, then we may as well assume this kernel to be a probability kernel. A similar statement is true for product regular conditional kernels.

4.4.5. By Theorem 4.3.6 statement (a) of Theorem 4.4.3 is a consequence of statement (b). In an attempt to reduce statement (b) to statement (a) the following problem occurs to the correspondence between regular conditional kernels and product regular conditional kernels that is mentioned in 4.3.4.

The Borel- σ -algebra of $\mathcal{X} \times \mathcal{Y}$, i.e., $\mathcal{B}(\mathcal{X} \times \mathcal{Y})$ may be strictly larger as $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ (see, e.g., Bogachev [10, Lemma 6.4.1 and Example 6.4.3]). If this is the case, i.e., $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y}) \subsetneq \mathcal{B}(\mathcal{X} \times \mathcal{Y})$, and $\mathcal{B}(\mathcal{X} \times \mathcal{Y})$ equals the Baire- σ -algebra on $\mathcal{X} \times \mathcal{Y}$, i.e., the smallest σ -algebra that makes all continuous function $\mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ measurable; then there exists a continuous function $f \in C(\mathcal{X} \times \mathcal{Y})$ that is not $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ -measurable. Composing the function f with arctan, we obtain a $g \in C_b(\mathcal{X} \times \mathcal{Y})$ that is not measurable with respect to $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$. So if $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}$ is a product regular conditional kernel under μ with respect to π , and $\xi : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y}) \rightarrow \mathbb{R}$ is as in Example 4.3.4 then g is not integrable with respect to $\xi(y, \cdot)$ for any $y \in \mathcal{Y}$.

$\mathcal{B}(\mathcal{X} \times \mathcal{Y})$ equals the Baire- σ -algebra if $\mathcal{X} \times \mathcal{Y}$ is a metric space (Bogachev [10, Proposition 6.3.4]). Therefore $\mathcal{X} = \mathcal{Y} = \mathbb{R}^{\mathbb{R}}$ equipped with the discrete topology form an example for which the above is the case.

We state two theorems (Theorem 4.4.6 and Theorem 4.4.7) showing the existence of product regular conditional probabilities that are weakly continuous on $\text{supp}(|\mu| \circ \pi^{-1})$.

4.4.6 Theorem. *Suppose that $\mathcal{Y}$ is countable and equipped with the discrete topology. Then $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}$ defined by*

$$\eta(y, A) = \begin{cases} \mu(A \times \mathcal{Y} | \mathcal{X} \times \{y\}) & \mu(\mathcal{X} \times \{y\}) \neq 0, \\ 0 & \mu(\mathcal{X} \times \{y\}) = 0, \end{cases} \quad (4.15)$$

is a product regular conditional kernel under μ with respect to π that is weakly continuous on $\text{supp}(|\mu| \circ \pi^{-1})$.

Proof. Follows from the fact that $\mu(A \times B) = \sum_{y \in B} \mu(A \times \{y\})$ for $A \in \mathcal{B}(\mathcal{X})$, $B \in \mathcal{B}(\mathcal{Y})$. $\square$

The following theorem is an easy consequence of Lebesgue's Dominated Convergence Theorem.

4.4.7 Theorem. Let λ be a probability measure on $\mathcal{B}(\mathcal{X})$. Let $D \subset \mathcal{Y}$. Let $f : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty)$ be a bounded $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ -measurable function such that $y \mapsto f(x, y)$ is continuous on D and equal to zero on $\mathcal{Y} \setminus D$ for λ -almost all $x \in \mathcal{X}$. Suppose that $\int_{\mathcal{X}} f(x, y) d\lambda(x) > 0$ for all $y \in D$. If $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ is given by

$$\eta(y, A) = \begin{cases} \frac{\int_{\mathcal{X}} \mathbb{1}_A(x) f(x, y) d\lambda(x)}{\int_{\mathcal{X}} f(x, y) d\lambda(x)} & y \in D, \\ 0 & y \notin D. \end{cases} \quad (4.16)$$

then η is weakly continuous on D (even strongly continuous, i.e., $y \mapsto \eta(y, A)$ is continuous for all $A \in \mathcal{B}(\mathcal{X})$). Let κ be a probability measure on $\mathcal{B}(\mathcal{Y})$ and assume $D = \text{supp } \kappa$. Then η is a product regular conditional kernel under

$$\mu : \mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1], \quad \mu(A) = \frac{\int_{\mathcal{X} \times \mathcal{Y}} \mathbb{1}_A f d[\lambda \otimes \kappa]}{\int_{\mathcal{X} \times \mathcal{Y}} f d[\lambda \otimes \kappa]} \quad (4.17)$$

with respect to π , that is weakly continuous on $D = \text{supp}(|\mu| \circ \pi^{-1})$.

4.4.8. In case $\mathcal{Y}$ is first countable, the notion of open and closed sets and continuity of functions $\mathcal{Y} \rightarrow \mathbb{R}$ is characterised by the convergence of sequences. Therefore the following are equivalent for a kernel $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}$

- (a) η is weakly continuous in y .
- (b) For all $(y_n)_{n \in \mathbb{N}}$ in $\mathcal{Y}$ with $y_n \rightarrow y$ one has $\eta(y_n, \cdot) \xrightarrow{w} \eta(y, \cdot)$.

In Section 4.6 the condition (b) of Theorem 4.4.9 is one of the key assumptions. If $\mathcal{X}$ is a metric space, this property follows from the weak continuity as in the Portmanteau Theorem. We state this in Theorem 4.4.9.

4.4.9 Theorem. Let $\eta : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow \mathbb{R}$ be a probability kernel. Let $D \subset \mathcal{Y}$, $y \in D$ and $\mathcal{V} \subset \mathcal{N}_y$ be such that $\bigcap \mathcal{V} = \{y\}$. Consider the following conditions.

- (a) $D \rightarrow \mathcal{M}(\mathcal{X})$, $y \mapsto \eta(y, \cdot)$ is weakly continuous in y .
 - (b) $\liminf_{\iota \in \mathbb{I}} \eta(y_\iota, G) \geq \eta(y, G)$ for all open $G \subset \mathcal{X}$ and $(y_\iota)_{\iota \in \mathbb{I}}$ in D with $y_\iota \rightarrow y$.
 - (c) $\limsup_{\iota \in \mathbb{I}} \eta(y_\iota, F) \leq \eta(y, F)$ for all closed $F \subset \mathcal{X}$ and $(y_\iota)_{\iota \in \mathbb{I}}$ in D with $y_\iota \rightarrow y$.
 - (d) $\sup_{V \in \mathcal{V}} \inf_{v \in V \cap D} \eta(v, G) \geq \eta(y, G)$ for all open sets $G \subset \mathcal{X}$.
 - (e) $\inf_{V \in \mathcal{V}} \sup_{v \in V \cap D} \eta(v, F) \leq \eta(y, F)$ for all closed sets $F \subset \mathcal{X}$.
- (b), (c), (d), (e) are equivalent. If $\mathcal{X}$ is metrisable, then (a) implies (b). If $\mathcal{X}$ is metrisable and $\mathcal{Y}$ is first countable, then (a) is equivalent to (b) and hence to (c), (d) and (e).

Proof. We leave it to the reader to check the equivalences between (b), (c), (d), (e). If $\mathcal{X}$ is a metric space, one can follow the lines of the Portmanteau Theorem in the book of Billingsley [7, Theorem 2.1] for the implication (a) implies (b), the fact that the measures in the proof are indexed by the natural numbers instead of a general directed set $\mathbb{I}$ does not affect the argument. The proof of (b) $\implies$ (a) in the book of Billingsley relies on the Lebesgue Dominated Convergence theorem. But when $\mathcal{Y}$ is first countable, one can restrict to sequences (see 4.4.8) and obtain the implication (b) $\implies$ (a) as is done in the book of Billingsley. $\square$

4.4.10 Lemma. Assume that μ is a probability measure. Let η be a product regular conditional probability under μ with respect to π . Write $D = \text{supp}(\mu \circ \pi^{-1})$ and let $y \in D$. Then for every $U \in \mathcal{N}_y$ one has $\mu(\mathcal{X} \times U) > 0$ and

$$\inf_{v \in U \cap D} \eta(v, A) \leq \mu(A \times \mathcal{Y} | \mathcal{X} \times U) \leq \sup_{v \in U \cap D} \eta(v, A) \quad (A \in \mathcal{B}(\mathcal{X})). \quad (4.18)$$

Moreover, if $\mathcal{V} \subset \mathcal{N}_y$ is such that $\bigcap \mathcal{V} = \{y\}$ and η satisfies (b) of Theorem 4.4.9, then

$$\liminf_{V \in \mathcal{V}} \mu(G \times \mathcal{Y} | \mathcal{X} \times V) \geq \eta(y, G) \text{ for all open } G \subset \mathcal{X}, \quad (4.19)$$

$$\limsup_{V \in \mathcal{V}} \mu(F \times \mathcal{Y} | \mathcal{X} \times V) \leq \eta(y, F) \text{ for all closed } F \subset \mathcal{X}. \quad (4.20)$$

Proof. Let $U \in \mathcal{N}_y$. Since $y \in D = \text{supp}(\mu \circ \pi^{-1})$ one has $\mu(\mathcal{X} \times U) > 0$. (4.18) follows from the fact that for all $A \in \mathcal{B}(\mathcal{X})$

$$\begin{aligned} \frac{\mu(A \times U)}{\mu(\mathcal{X} \times U)} &= \frac{\int_{\mathcal{Y}} \mathbb{1}_U(y) \eta(y, A) \, d[\mu \circ \pi^{-1}](y)}{\int_{\mathcal{Y}} \mathbb{1}_U(y) \, d[\mu \circ \pi^{-1}](y)} \\ &= \frac{\int_{\mathcal{Y}} \mathbb{1}_{U \cap D}(y) \eta(y, A) \, d[\mu \circ \pi^{-1}](y)}{\int_{\mathcal{Y}} \mathbb{1}_{U \cap D}(y) \, d[\mu \circ \pi^{-1}](y)}. \end{aligned} \quad (4.21)$$

For an open $G \subset \mathcal{X}$ we have for $\mathcal{V}$ as above

$$\liminf_{V \in \mathcal{V}} \mu(G \times \mathcal{Y} | \mathcal{X} \times V) \geq \liminf_{V \in \mathcal{V}} \inf_{v \in V \cap D} \eta(v, G) = \sup_{V \in \mathcal{V}} \inf_{v \in V \cap D} \eta(v, G). \quad (4.22)$$

Thus (4.19) follows when assuming (b) of Theorem 4.4.9. Similarly, one obtains (4.20). $\square$

For a regular conditional probability we have a similar statement, see Lemma 4.4.11. The proof can be done following the lines of the proof of Lemma 4.4.10 or as a consequence of Lemma 4.4.10 using Theorem 4.3.6.

4.4.11 Lemma. Assume that ν is a probability measure. Let η be a regular conditional probability under ν with respect to τ . Write $D = \text{supp}(\nu \circ \tau^{-1})$ and let $y \in D$. Then for every $U \in \mathcal{N}_y$ one has $\nu(\tau^{-1}(U)) > 0$ and

$$\inf_{v \in U \cap D} \eta(v, A) \leq \nu(A | \tau^{-1}(U)) \leq \sup_{v \in U \cap D} \eta(v, A) \quad (A \in \mathcal{B}(\mathcal{X})). \quad (4.23)$$

Moreover, if $\mathcal{V} \subset \mathcal{N}_y$ is such that $\bigcap \mathcal{V} = \{y\}$ and η satisfies (b) of Theorem 4.4.9, then

$$\liminf_{V \in \mathcal{V}} \nu(G | \tau^{-1}(V)) \geq \eta(y, G) \text{ for all open } G \subset \mathcal{X}, \quad (4.24)$$

$$\limsup_{V \in \mathcal{V}} \nu(F | \tau^{-1}(V)) \leq \eta(y, F) \text{ for all closed } F \subset \mathcal{X}. \quad (4.25)$$

§4.5 Some facts about functions with compact sublevel sets

In this section we present some facts for functions with compact sublevel sets which are used in Sections 4.6, 4.7 and 4.8.

In this section $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$ are topological spaces.

4.5.1 Definition. Let $J : \mathcal{X} \rightarrow [0, \infty]$. We call the set $[J \leq \alpha]$ (see Section 4.2) a sublevel set of J for $\alpha \in [0, \infty)$. J is said to be *lower semicontinuous* if all sublevels of J are closed. J is said to have *compact sublevel sets* if all sublevels of J are compact.

4.5.2. Let $J : \mathcal{X} \rightarrow [0, \infty]$ be lower semicontinuous. Then

$$J(x) = \sup_{G \in \mathcal{N}_x} \inf J(G). \quad (4.26)$$

Indeed, for all $\alpha < J(x)$ the set $[J > \alpha]$ is open and contains x .

Hence, a function $J : \mathcal{X} \rightarrow [0, \infty]$ is lower semicontinuous if and only if

$$\liminf_{\iota \in \mathbb{I}} J(x_\iota) \geq J(x) \quad (4.27)$$

for all $x \in \mathcal{X}$ and all nets $(x_\iota)_{\iota \in \mathbb{I}}$ in $\mathcal{X}$ that converge to x .

4.5.3 Lemma. Let $\tau : \mathcal{Z} \rightarrow \mathcal{Y}$ be continuous. Let $J : \mathcal{Z} \rightarrow [0, \infty]$ have compact sublevel sets. Let $y \in \mathcal{Y}$ and $\mathcal{V} \subset \mathcal{N}_y$, $\bigcap \mathcal{V} = \{y\}$. Let $F \subset \mathcal{Z}$ be closed. Then

$$\liminf_{V \in \mathcal{V}} \inf J(F \cap \tau^{-1}(\overline{V})) = \inf J(F \cap \tau^{-1}(\{y\})). \quad (4.28)$$

Consequently, if $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, then, for all closed $F \subset \mathcal{X}$ with $\inf J(F \times \{y\}) < \infty$,

$$\liminf_{V \in \mathcal{V}} \inf J(F \times \overline{V}) = \inf J(F \times \{y\}). \quad (4.29)$$

Proof. The $\leq$ inequality in (4.28) is immediate. Because $\liminf_{V \in \mathcal{V}} \inf J(F \cap \tau^{-1}(\overline{V})) \geq \liminf_{V \in \mathcal{N}_y} \inf J(F \cap \tau^{-1}(\overline{V}))$, it is sufficient to prove

$$\alpha := \liminf_{V \in \mathcal{N}_y} \inf J(F \cap \tau^{-1}(\overline{V})) \geq \inf J(F \cap \tau^{-1}(\{y\})). \quad (4.30)$$

Note that $\alpha = \sup_{V \in \mathcal{N}_y} \inf J(F \cap \tau^{-1}(\overline{V}))$. If $\alpha = \infty$, there is nothing to prove. Suppose that $\alpha < \infty$. Whence $F \cap \tau^{-1}(\overline{V}) \cap [J \leq \alpha + \varepsilon] \neq \emptyset$ for all $V \in \mathcal{N}_y$ and all $\varepsilon > 0$. Since $[J \leq \alpha + \varepsilon]$ is compact, this implies that $\bigcap_{V \in \mathcal{N}_y} F \cap \tau^{-1}(\overline{V}) \cap [J \leq \alpha + \varepsilon] \neq \emptyset$, i.e., $\inf J(F \cap \tau^{-1}(\{y\})) \leq \alpha + \varepsilon$ for all $\varepsilon > 0$. $\square$

4.5.4. The assumption that τ be continuous is not redundant; e.g., consider $\mathcal{Y} = \mathcal{Z} = [0, 1]$ and $J = \mathbb{1}_{(\frac{1}{2}, 1]}$ and τ given by $\tau(0) = 0$, $\tau(1) = 1$ and $\tau(x) = 1 - x$ for $x \in (0, 1)$, $F = [0, 1]$ and $y = 1$. Then, for all neighbourhoods V of y , $\tau^{-1}(V)$ contains the interval $(0, \varepsilon)$ for some $\varepsilon > 0$, whence $\inf J(F \cap \tau^{-1}(V)) = 0$ but $\inf J(F \cap \tau^{-1}(\{y\})) = J(1) = 1$.

4.5.5 Lemma. Let $\mathcal{X}$ be normal and let $\mathcal{G}$ be a basis for the topology of $\mathcal{X}$. Let $J : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty]$ have compact sublevel sets.

- (a) For all open $G \subset \mathcal{X}$ and $\varepsilon > 0$ there exists a $U \in \mathcal{G}$ with $U \subset \overline{U} \subset G$ such that

$$\inf J(G \times \{y\}) + \varepsilon \geq \inf J(U \times \{y\}). \quad (4.31)$$

- (b) For all closed $F \subset \mathcal{X}$ and $\alpha < \inf J(F \times \{y\})$, there exists $U_1, \dots, U_k \in \mathcal{G}$ such that with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$ one has $F \subset W^\circ \subset W$ and

$$\alpha < \inf J(W \times \{y\}) \leq \inf J(W^\circ \times \{y\}) \leq \inf J(F \times \{y\}). \quad (4.32)$$

Proof. (a) Let $\varepsilon > 0$. Let $x \in G$ be such that $J(x, y) \leq \inf J(G \times \{y\}) + \varepsilon$. Since $\mathcal{X}$ is a normal topological space, there exists an open set U with $x \in U \subset \overline{U} \subset G$. Because $\mathcal{G}$ is a basis, U may be chosen in $\mathcal{G}$. Then $\inf J(G \times \{y\}) + \varepsilon \geq J(x, y) \geq \inf J(U \times \{y\})$.

(b) Let $\beta > \alpha$ be such that $\beta < \inf J(F \times \{y\})$. The set $K := \{x \in \mathcal{X} : J(x, y) \leq \beta\}$ is a compact set that is disjoint from F . Whence there exists disjoint open $U, V \subset \mathcal{X}$ with $K \subset U$ and $F \subset V$. Since $\mathcal{G}$ is a basis and K is compact, there exists $U_1, \dots, U_k$ in $\mathcal{G}$ with $K \subset U_1 \cup \dots \cup U_k \subset U$. Then $\overline{U_1 \cup \dots \cup U_k} \cap V = \emptyset$. Whence with $W := \mathcal{X} \setminus \overline{U_1 \cup \dots \cup U_k}$ one has $F \subset W^\circ$ and $W \subset \mathcal{X} \setminus K$, which implies $\inf J(W \times \{y\}) \geq \beta > \alpha$. $\square$

§4.6 Large deviations for product regular conditional probabilities

In this section we consider the following situation.

- (i) $\mathcal{X}$ and $\mathcal{Y}$ are topological spaces, where $\mathcal{X}$ is normal.
- (ii) $\mathcal{G}$ is a basis for the topology of $\mathcal{X}$ and $\mathcal{H}$ is a basis for the topology of $\mathcal{Y}$.
- (iii) $\pi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ is given by $\pi(x, y) = y$.
- (iv) $(\mu_n)_{n \in \mathbb{N}}$ is a sequence of probability measures on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ satisfying the large deviation principle on $\{A \times B : A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})\}$ with a rate function $J : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty]$ that has compact sublevel sets.
- (v) For each $n \in \mathbb{N}$ we assume the following: $\text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$,⁵ there exists a product regular conditional probability $\eta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under μ_n with respect to π , which satisfies the following continuity condition (see Theorem 4.4.9):

$$\begin{aligned} \liminf_{\iota \in \mathbb{I}} \eta_n(y_\iota, G) &\geq \eta_n(y, G) \text{ for all open } G \subset \mathcal{X} \\ \text{and } (y_\iota)_{\iota \in \mathbb{I}} &\text{ in } \text{supp}(\mu_n \circ \pi^{-1}) \text{ with } y_\iota \rightarrow y. \end{aligned} \quad (4.33)$$

- (vi) Let $y \in \mathcal{Y}$. We assume that $\inf J(\mathcal{X} \times \{y\}) < \infty$ and that there exist $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ with $y_n \rightarrow y$. We define $I : \mathcal{X} \rightarrow [0, \infty]$ by

$$I(x) = J(x, y) - \inf J(\mathcal{X} \times \{y\}). \quad (4.34)$$

⁵As we are considering large deviation bound for $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ we want such y_n to exist. Instead of this condition one could of course deal with the situation where $\text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$ for all $n \geq N$ for some large N and consider sequences $(y_n)_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ for $n \geq N$.

In this section we derive necessary and sufficient conditions for the large deviation bounds with rate function I for sequences of the form $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$. We prove this for general topological spaces instead of metric spaces as it does not cost more effort.

In Theorem 4.6.3 we consider a fixed sequence $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and describe equivalent conditions for the lower and upper large deviation bound to hold.

We are interested in the question whether for all sequences $(y_n)_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the lower and upper large deviation bound with rate function I . In Theorem 4.6.9 we give equivalent⁶ and sufficient conditions for these bounds in a way that does not depend on sequences $(y_n)_{n \in \mathbb{N}}$ and the sets $(\mathcal{V}_n)_{n \in \mathbb{N}}$ as in Theorem 4.6.3.

Finally in 4.6.12 we comment on deriving Theorem 4.1.2 from Theorem 4.6.9.

But first we consider specific situations, providing a simple proof of the large deviation bounds with rate function I for sequences of the form $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$. Namely, we consider the case that $\mathcal{Y}$ is a discrete space (Theorem 4.6.1) and the case where μ_n is a product measure for all $n \in \mathbb{N}$ (Theorem 4.6.2).

4.6.1 Theorem. *Suppose that $\mathcal{Y}$ is countable and equipped with the discrete topology. Let $y \in \mathcal{Y}$ be such that $\inf J(\mathcal{X} \times \{y\}) < \infty$. For all $(y_n)_{n \in \mathbb{N}}$ in $\mathcal{Y}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I .*

Proof. This basically follows from the following inequalities which follow from the large deviation principle and from Theorem 4.4.6.

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G \times \{y\}) \geq -\inf J(G \times \{y\}) \text{ for all open } G \subset \mathcal{X}, \quad (4.35)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F \times \{y\}) \leq -\inf J(F \times \{y\}) \text{ for all closed } F \subset \mathcal{X}. \quad (4.36)$$

□

4.6.2 Theorem (Independent coordinates). *Suppose that $\mathcal{X}$ and $\mathcal{Y}$ are second countable and $\mathcal{Y}$ is regular. Suppose that $\mu_n = \mu_n^1 \otimes \mu_n^2$ for some μ_n^1 on $\mathcal{B}(\mathcal{X})$ and μ_n^2 on $\mathcal{B}(\mathcal{Y})$ for all $n \in \mathbb{N}$. Then $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I for all sequences $(y_n)_{n \in \mathbb{N}}$ in $\mathcal{Y}$. In particular, $\eta_n(y_n, \cdot) = \mu_n^1$ and $I(x) = \inf J(\{x\} \times \mathcal{Y})$.*

Proof. It is straightforward to see that $\eta_n(y, \cdot) = \mu_n^1$ for all $y \in \mathcal{Y}$. $(\mu_n^1)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $J_1(x) := \inf J(\{x\} \times \mathcal{Y})$. Indeed, for an open set $G \subset \mathcal{X}$ and a closed set $F \subset \mathcal{X}$ we have

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(G) = \liminf_{n \rightarrow \infty} \frac{1}{n} \log (\mu_n^1 \otimes \mu_n^2)(G \times \mathcal{Y}) \geq -\inf J(G \times \mathcal{Y}), \quad (4.37)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(F) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log (\mu_n^1 \otimes \mu_n^2)(F \times \mathcal{Y}) \leq -\inf J(F \times \mathcal{Y}). \quad (4.38)$$

Similarly, $(\mu_n^2)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $J_2(z) := \inf J(\mathcal{X} \times \{z\})$. J_1 and J_2 are lower semicontinuous, which can be concluded by 4.5.2

⁶Under the condition that $\mathcal{Y}$ is first countable.

and Lemma 4.5.3, as for example, $\lim_{l \in \mathbb{I}} z_l = z$ implies $\liminf_{l \in \mathbb{I}} \inf J(\mathcal{X} \times \{z_l\}) \geq \liminf_{V \in \mathcal{N}_z} \inf J(\mathcal{X} \times V)$. Using Theorem 4.B.4 it is not difficult to prove that $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $(x, z) \mapsto J_1(x) + J_2(z)$, so that (see Rassoul-Agha and Seppäläinen [75, Theorem 2.18]) $J(x, z) = J_1(x) + J_2(z)$, and thus $I(x) = J(x, y) - \inf J(\mathcal{X} \times \{y\}) = J_1(x) = \inf J(\{x\} \times \mathcal{Y})$ for all $x \in \mathcal{X}, z \in \mathcal{Y}$. $\square$

4.6.3 Theorem. *Let $(y_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{Y}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ that converges to y . For $n \in \mathbb{N}$ let $\mathcal{V}_n \subset \mathcal{N}_{y_n}$ be such that $\bigcap \mathcal{V}_n = \{y_n\}$. Then (a2) $\iff$ (a3) $\iff$ (a1) and (b2) $\iff$ (b3) $\iff$ (b1)*

(a1) *For all open $G \subset \mathcal{X}$*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, G) \geq -\inf I(G). \quad (4.39)$$

(a2) *For all $U \in \mathcal{G}^7$*

$$\liminf_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U). \quad (4.40)$$

(a3) *For all open $U \subset \mathcal{X}$ one has*

$$\liminf_{n \rightarrow \infty} \liminf_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U). \quad (4.41)$$

(b1) *For all closed $F \subset \mathcal{X}$*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, F) \leq -\inf I(F). \quad (4.42)$$

(b2) *For all $U_1, \dots, U_k \in \mathcal{G}$ one has for $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$*

$$\limsup_{n \rightarrow \infty} \liminf_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \leq -\inf I(W). \quad (4.43)$$

(b3) *For all closed $W \subset \mathcal{X}$*

$$\limsup_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V) \leq -\inf I(W). \quad (4.44)$$

Proof. The implications (a3) $\implies$ (a2) and (b3) $\implies$ (b2) are immediate.

(a1) $\implies$ (a3) Let $U \subset \mathcal{X}$ be an open set. By Lemma 4.4.10, (4.19),

$$\liminf_{n \rightarrow \infty} \liminf_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V) \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, U). \quad (4.45)$$

(b1) $\implies$ (b3) Let $W \subset \mathcal{X}$ be a closed set. By Lemma 4.4.10, (4.20),

$$\limsup_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, W). \quad (4.46)$$

⁷Note that $\mu_n(\mathcal{X} \times V) > 0$ for all $n \in \mathbb{N}$ and $V \in \mathcal{N}_{y_n}$, as $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$.

(a2) $\implies$ (a1). Let $G \subset \mathcal{X}$ be open. Let $\varepsilon > 0$ and U be as in Lemma 4.5.5(a). Then we obtain using Lemma 4.4.10

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, G) &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, \overline{U}) \\ &\geq \liminf_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \\ &\geq -\inf I(U) = -\inf J(U \times \{y\}) + \inf J(\mathcal{X} \times \{y\}) \\ &\geq -\inf J(G \times \{y\}) + \inf J(\mathcal{X} \times \{y\}) - \varepsilon. \end{aligned} \quad (4.47)$$

As this holds for all $\varepsilon > 0$, we conclude (4.39).

(b2) $\implies$ (b1). Let $\alpha < \inf J(F \times \{y\})$ and $U_1, \dots, U_k$ and W be as in Lemma 4.5.5(b). Then we obtain using Lemma 4.4.10

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, F) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, W^\circ) \\ &\leq \limsup_{n \rightarrow \infty} \liminf_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \\ &\leq -\inf I(W) \leq -\alpha + \inf J(\mathcal{X} \times \{y\}). \end{aligned} \quad (4.48)$$

As this holds for all $\alpha < \inf J(F \times \{y\})$, we conclude (4.42). $\square$

4.6.4 (Fixed y). Note that if $y_n = y$ for all $n \in \mathbb{N}$, one can take $\mathcal{V}_n = \mathcal{V}$ for a $\mathcal{V} \subset \mathcal{N}_y$ with $\bigcap \mathcal{V} = \{y\}$. Then Theorem 4.6.3 implies that $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I if and only if (a2) and (b2) hold (with $\mathcal{V}_n = \mathcal{V}$).

4.6.5. Let $(y_n)_{n \in \mathbb{N}}$ in $\mathcal{Y}$ be such that $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $y_n \rightarrow y$. From Theorem 4.6.3 we derive that (a2) holds for some $\mathcal{V}_n \subset \mathcal{N}_{y_n}$ with $\bigcap \mathcal{V}_n = \{y_n\}$ if and only if (a2) holds for all such $\mathcal{V}_n$. Similarly, (b2) holds for some $\mathcal{V}_n \subset \mathcal{N}_{y_n}$ with $\bigcap \mathcal{V}_n = \{y_n\}$ if and only if (b2) holds for all such $\mathcal{V}_n \subset \mathcal{N}_{y_n}$.

In Lemma 4.6.7, we give a consequence of the large deviation principle of $(\mu_n)_{n \in \mathbb{N}}$. In Theorem 4.6.9 and Theorem 4.6.10 we use this to formulate sufficient conditions for upper or lower large deviation bound on sequences $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ with $y_n \rightarrow y$ and sequences $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$.

We assumed $\mathcal{X}$ to be normal in this section. For Lemma 4.6.7 this assumption can be dropped.

4.6.6. For all neighbourhoods V of y one has by the large deviation principle

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \times V) \geq -\inf J(\mathcal{X} \times V^\circ) \geq -\inf J(\mathcal{X} \times \{y\}) > -\infty. \quad (4.49)$$

In particular, there exists an $N \in \mathbb{N}$ such that $\mu_n(\mathcal{X} \times V) > 0$ for all $n \geq N$. Therefore $\mu_n(G \times \mathcal{Y} | \mathcal{X} \times V)$ is well-defined for large n .

4.6.7 Lemma.

(a) For open $G \subset \mathcal{X}$

$$\liminf_{V \in \mathcal{N}_y} \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}: \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(G \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(G). \quad (4.50)$$

(b) For closed $F \subset \mathcal{X}$

$$\limsup_{V \in \mathcal{N}_y} \limsup_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}; \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(F \times \mathcal{Y} | \mathcal{X} \times V) \leq -\inf I(F). \quad (4.51)$$

Proof. (a). Let $\varepsilon > 0$. By Lemma 4.5.3, there exists a $V_0 \in \mathcal{N}_y$ such that for all $V \in \mathcal{N}_y$ with $V \subset V_0$

$$\inf J(\mathcal{X} \times \{y\}) \geq \inf J(\mathcal{X} \times \bar{V}) \geq \inf J(\mathcal{X} \times \bar{V}_0) \geq \inf J(\mathcal{X} \times \{y\}) - \varepsilon. \quad (4.52)$$

Let $V \in \mathcal{N}_y$ be such that $V \subset V_0$. As $\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \times V) > -\infty$ (see 4.6.6) we can “split the lim inf in two” and we get by the large deviation principle and by (4.52)

$$\begin{aligned} \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}; \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(G \times \mathcal{Y} | \mathcal{X} \times V) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G \times V) - \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \times V) \\ &\geq -\inf J(G \times \{y\}) + \inf J(\mathcal{X} \times \bar{V}) \geq -\inf I(G) - \varepsilon. \end{aligned} \quad (4.53)$$

(b). Let $\alpha < \inf J(F \times \{y\})$. There exists a neighbourhood V_0 of y such that for all neighbourhoods V of y with $V \subset V_0$

$$\inf J(F \times \{y\}) \geq \inf J(F \times \bar{V}) \geq \inf J(F \times \bar{V}_0) \geq \alpha. \quad (4.54)$$

Let $V \in \mathcal{N}_y$ be such that $y \in V \subset V_0$. Similarly as above we get

$$\limsup_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}; \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(F \times \mathcal{Y} | \mathcal{X} \times V) \leq -\alpha + \inf J(\mathcal{X} \times \{y\}). \quad (4.55)$$

□

4.6.8 Theorem. I has compact sublevel sets.

Proof. $[I \leq c] = \pi([J \leq c + \inf J(\mathcal{X} \times \{y\})])$. □

4.6.9 Theorem. We have

$$(A5) \implies (A4) \iff (A3) \implies (A2) \implies (A1),$$

and, if $\mathcal{Y}$ is first countable, then

$$(A1) \iff (A2),$$

where

(A1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation lower bound with rate function I .

(A2) For all $U \in \mathcal{G}$

$$\sup_{V_0 \in \mathcal{N}_y} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U). \quad (4.56)$$

(A3) For all $U \in \mathcal{G}$

$$\begin{aligned} & \sup_{V_0 \in \mathcal{N}_y} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \\ & \geq \liminf_{V \in \mathcal{N}_y} \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}; \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V). \end{aligned} \quad (4.57)$$

(A4) For all $U \in \mathcal{G}$ we have $\forall Z_0 \in \mathcal{N}_y \forall \varepsilon > 0 \exists V_0 \in \mathcal{N}_y \exists Z \in \mathcal{N}_y, Z \subset Z_0 \forall M \exists m \geq M \exists N \forall n \geq N \forall V \in \mathcal{H}, V \subset V_0, V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$:

$$\frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq \frac{1}{m} \log \mu_m(U \times \mathcal{Y} | \mathcal{X} \times Z) - \varepsilon. \quad (4.58)$$

(A5) For all $U \in \mathcal{G}$ we have $\forall \varepsilon > 0 \forall V_0 \in \mathcal{N}_y \exists N \in \mathbb{N} \forall n \geq N \forall V \in \mathcal{H}, V \subset V_0, V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$:

$$\mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq e^{-n\varepsilon} \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_0). \quad (4.59)$$

Moreover,

$$(B5) \implies (B4) \iff (B3) \implies (B2) \implies (B1),$$

and, if $\mathcal{Y}$ is first countable, then

$$(B1) \iff (B2),$$

where

(B1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound with rate function I .

(B2) For all $U_1, \dots, U_k \in \mathcal{G}$ one has for $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$

$$\inf_{V_0 \in \mathcal{N}_y} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \leq -\inf I(W). \quad (4.60)$$

(B3) For all $U_1, \dots, U_k \in \mathcal{G}$ with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$

$$\begin{aligned} & \inf_{V_0 \in \mathcal{N}_y} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \\ & \leq \limsup_{V \in \mathcal{N}_y} \limsup_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}; \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V). \end{aligned} \quad (4.61)$$

(B4) For all $U_1, \dots, U_k \in \mathcal{G}$ with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$ we have $\forall Z_0 \in \mathcal{N}_y \forall \varepsilon > 0 \exists V_0 \in \mathcal{N}_y \exists Z \in \mathcal{N}_y, Z \subset Z_0 \forall M \exists m \geq M \exists N \forall n \geq N \forall V \in \mathcal{H}, V \subset V_0, V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$:

$$\frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \leq \frac{1}{m} \log \mu_m(U \times \mathcal{Y} | \mathcal{X} \times Z) + \varepsilon. \quad (4.62)$$

(B5) For all $U_1, \dots, U_k \in \mathcal{G}$ with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$ we have $\forall \varepsilon > 0 \forall V_0 \in \mathcal{N}_y \exists N \in \mathbb{N} \forall n \geq N \forall V \in \mathcal{H}, V \subset V_0, V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$:

$$\mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \leq e^{n\varepsilon} \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V_0) \quad (4.63)$$

Proof. The proofs of $(B5) \implies (B4) \iff (B3) \implies (B2) \implies (B1)$ and of $(B1) \implies (B2)$ are similar to the proofs of the following implications.

(A4) $\iff$ (A3) follows by definition of $\sup$, $\inf$, $\limsup$ and $\liminf$.

(A5) $\implies$ (A3) Let $U \in \mathcal{G}$. Assuming (A5) we obtain $\forall \varepsilon > 0 \forall V_0 \in \mathcal{N}_y \exists N \in \mathbb{N} \forall n \geq N$ one has $\mu_n(\mathcal{X} \times V_0) > 0$ and

$$\inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_0) - \varepsilon. \quad (4.64)$$

So $\forall \varepsilon > 0 \forall V_0 \in \mathcal{N}_y$

$$\begin{aligned} \sup_{Z \in \mathcal{N}_y} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset Z \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \\ \geq \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}: \mu_n(\mathcal{X} \times V_0) > 0}} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_0) - \varepsilon. \end{aligned} \quad (4.65)$$

(A3) $\implies$ (A2) Follows by Lemma 4.6.7.

(A2) $\implies$ (A1). Suppose that (A2) holds. Let $U \in \mathcal{G}$ with $\inf J(U \times \{y\}) < \infty$ and let $\varepsilon > 0$. Let $V_0 \in \mathcal{N}_y$ and $N \in \mathbb{N}$ be such that $\frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U) - \varepsilon$ for all $n \geq N$ and all $V \in \mathcal{H}$ with $V \subset V_0$ and $V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$. Let $(y_n)_{n \in \mathbb{N}}$ be such that $y_n \in \text{supp}(\mu_n \circ \pi^{-1})$ and $y_n \rightarrow y$. Let $N_0 \geq N$ be such that $y_n \in V_0$ for all $n \geq N_0$. Then for all $n \geq N_0$ and $V \in \mathcal{N}_{y_n} \cap \mathcal{H}$ with $V \subset V_0$ we have $\frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U) - \varepsilon$. This implies (a2) of Theorem 4.6.3 (with $\mathcal{V}_n = \mathcal{N}_{y_n} \cap \mathcal{H}$).

(A1) $\implies$ (A2) (assuming $\mathcal{Y}$ is first countable). Suppose that (A2) does not hold. Let $(V_m)_{m \in \mathbb{N}}$ be a decreasing sequence in $\mathcal{H}$ with $\bigcap_{m \in \mathbb{N}} V_m = \{y\}$. Then there exists a $U \in \mathcal{G}$ with $\inf J(U \times \{y\}) < \infty$ and an $\alpha > \inf I(U)$ such that for all $M \in \mathbb{N}$ and $N \in \mathbb{N}$ there exists an $n \geq N$ and a $V \in \mathcal{H}$ with $V \subset V_M$ and $V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset$ such that

$$\frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V) \leq -\alpha. \quad (4.66)$$

Let $\beta < \alpha$ be such that $\beta > \inf I(U)$. By Lemma 4.4.10 we have

$$\inf_{z \in V \cap \text{supp}(\mu_n \circ \pi^{-1})} \frac{1}{n} \log \eta_n(z, U) \leq \frac{1}{n} \log \mu_n(\bar{U} \times \mathcal{Y} | \mathcal{X} \times V). \quad (4.67)$$

For each $m \in \mathbb{N}$ there exists an n_m and a $y_{n_m} \in V_m \cap \text{supp}(\mu_{n_m} \circ \pi^{-1})$ such that

$$\frac{1}{n_m} \log \eta_{n_m}(y_{n_m}, U) \leq -\beta. \quad (4.68)$$

We may choose $n_1 < n_2 < n_3 < \dots$. With $y_k = y$ for $k \notin \{n_m : m \in \mathbb{N}\}$ we have $y_n \rightarrow y$ and

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \eta_n(y_n, U) \leq \liminf_{m \rightarrow \infty} \frac{1}{n_m} \log \eta_{n_m}(y_{n_m}, U) \leq -\beta. \quad (4.69)$$

Therefore (a1) of Theorem 4.6.3 does not hold, which implies that (A1) does not hold. $\square$

We can also use Lemma 4.6.7 and Theorem 4.6.3 (see also 4.6.4) to obtain sufficient conditions for the lower or upper large deviation bounds for $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$.

4.6.10 Theorem. *Let $\mathcal{V} \subset \mathcal{N}_y$ be such that $\bigcap \mathcal{V} = \{y\}$.*

(a) *Suppose that for all $U \in \mathcal{G}$ with $\inf J(U \times \{y\}) < \infty$*

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}} \frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \\ & \geq \liminf_{V \in \mathcal{N}_y} \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}: \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V). \end{aligned} \quad (4.70)$$

Then $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation lower bound with rate function I .

(b) *Suppose that for all $U_1, \dots, U_k \in \mathcal{G}$ with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$*

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \liminf_{V \in \mathcal{V}} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \\ & \leq \limsup_{V \in \mathcal{N}_y} \limsup_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}: \mu_n(\mathcal{X} \times V) > 0}} \frac{1}{n} \log \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V). \end{aligned} \quad (4.71)$$

Then $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound with rate function I .

4.6.11. (4.70) and (4.71) hold for example when $\forall \varepsilon > 0 \forall V_0 \in \mathcal{V} \exists N \in \mathbb{N} \forall n \geq N \forall V \in \mathcal{V}, V \subset V_0$:

$$\mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq e^{-n\varepsilon} \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V_0), \quad (4.72)$$

$$\mu_n(W^\circ \times \mathcal{Y} | \mathcal{X} \times V) \leq e^{n\varepsilon} \mu_n(W \times \mathcal{Y} | \mathcal{X} \times V_0), \quad (4.73)$$

respectively.

4.6.12. Theorem 4.1.2 is a consequence of Theorem 4.4.9, Theorem 4.6.8 and Theorem 4.6.9 with $\mathcal{G} = \{B(x, r) : x \in \mathcal{X}, r > 0\}$ and $\mathcal{H} = \{B(y, \delta) : y \in \mathcal{Y}, \delta > 0\}$.

§4.7 Large deviations for regular conditional probabilities

In this section $\mathcal{X}$ and $\mathcal{Y}$ are topological spaces, $(\nu_n)_{n \in \mathbb{N}}$ is a sequence of probability measures on $\mathcal{B}(\mathcal{X})$ that satisfies the large deviation principle with rate function $K : \mathcal{X} \rightarrow [0, \infty]$ and $\tau : \mathcal{X} \rightarrow \mathcal{Y}$ is continuous. For more assumptions, see 4.7.2.

We derive the analogous statements as in Section 4.6 but for regular conditional kernels instead of product regular conditional kernels (4.7.3 and Theorem 4.7.5). First we show that with μ_n the probability measure corresponding on the product space corresponding to ν_n as in Theorem 4.3.6, the sequence $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with a rate function described in terms of K (Theorem 4.7.1).

If $(\eta_n)_{n \in \mathbb{N}}$ are regular conditional probabilities under $(\nu_n)_{n \in \mathbb{N}}$ given τ , then one could also follow the proofs in Section 4.6 for the product regular conditional probabilities to obtain similar results for large deviations for sequences that are of the

form $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$. Instead, we make the approach via Theorem 4.3.6 to translate the results to the setting of regular conditional probabilities.

4.7.1 Theorem. *For all $n \in \mathbb{N}$ let μ_n be the probability measure on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ for which $\mu_n(A \times B) = \nu_n(A \cap \tau^{-1}(B))$ for $A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})$ (as in Theorem 4.3.6). Then $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle on $\{A \times B : A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})\}$ with rate function $J : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty]$ given by*

$$J(x, y) = \begin{cases} K(x) & \tau(x) = y, \\ \infty & \tau(x) \neq y. \end{cases} \quad (4.74)$$

If K has compact sublevel sets, then so does J .

Proof. By definition of J we have

$$\inf K(A \cap \tau^{-1}(B)) = \inf J(A \times B) \quad (A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})). \quad (4.75)$$

Let $A \in \mathcal{B}(\mathcal{X})$ and $B \in \mathcal{B}(\mathcal{Y})$. Then

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) &= \liminf_{n \rightarrow \infty} \frac{1}{n} \log \nu_n(A \cap \tau^{-1}(B)) \\ &\geq -\inf K((A \cap \tau^{-1}(B))^\circ). \end{aligned} \quad (4.76)$$

We have $(A \cap \tau^{-1}(B))^\circ = A^\circ \cap \tau^{-1}(B)^\circ$ and $\tau^{-1}(B)^\circ \supset \tau^{-1}(B^\circ)$, whence

$$\begin{aligned} \inf K((A \cap \tau^{-1}(B))^\circ) &\leq \inf K(A^\circ \cap \tau^{-1}(B^\circ)) \\ &= \inf J(A^\circ \times B^\circ) = \inf J((A \times B)^\circ). \end{aligned} \quad (4.77)$$

Similarly

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log \nu_n(A \cap \tau^{-1}(B)) \\ &\leq -\inf K(\overline{A \cap \tau^{-1}(B)}). \end{aligned} \quad (4.78)$$

We have $\overline{A \cap \tau^{-1}(B)} \subset \overline{A} \cap \overline{\tau^{-1}(B)}$ and $\overline{\tau^{-1}(B)} \subset \tau^{-1}(\overline{B})$, whence

$$\inf K(\overline{A \cap \tau^{-1}(B)}) \geq \inf K(\overline{A} \cap \tau^{-1}(\overline{B})) = \inf J(\overline{A} \times \overline{B}). \quad (4.79)$$

Suppose that K has compact sublevel sets. Let $c \geq 0$. Then $[J \leq c]$ is contained in the compact set $[K \leq c] \times \tau([K \leq c])$. By Theorem 4.6.8 I has compact sublevel sets. $\square$

4.7.2. In the rest of this section $\mathcal{X}$ is normal, $\mathcal{G}, \mathcal{H}, \pi$ are as in (ii) and (iii) of Section 4.6. Furthermore similarly to (v) and (vi) of Section 4.6 we assume the following.

(v)* For each $n \in \mathbb{N}$ we assume the following: $\text{supp}(\nu_n \circ \tau^{-1}) \neq \emptyset$, there exists a regular conditional probability $\eta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ under ν_n with respect to τ , satisfying the continuity condition (4.33).

(vi)* Let $y \in \mathcal{Y}$. We assume that $\inf K(\tau^{-1}(\{y\})) < \infty$ and that there exist $y_n \in \text{supp}(\nu_n \circ \tau^{-1})$ with $y_n \rightarrow y$. Let $I : \mathcal{X} \rightarrow [0, \infty]$ be given by

$$\begin{aligned} I(x) &= J(x, y) - \inf J(\mathcal{X} \times \{y\}) \\ &= \begin{cases} K(x) - \inf K(\tau^{-1}(\{y\})) & \tau(x) = y, \\ \infty & \tau(x) \neq y. \end{cases} \end{aligned} \quad (4.80)$$

4.7.3. As by Theorem 4.3.6 η_n is the product regular conditional kernel under μ_n with respect to π , by Theorem 4.7.1 $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle on $\{A \times B : A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})\}$ with rate function J , and $\inf J(\mathcal{X} \times \{y\}) = \inf K(\tau^{-1}(\{y\})) < \infty$ and μ_n and η_n are as in Section 4.6 (in the sense that (iv), (v), (vi) hold). Therefore we can translate the results of Section 4.6, but also the results of Section 4.4 and Section 4.5, using for example (4.75), $\nu_n \circ \tau^{-1} = \mu_n \circ \pi^{-1}$ and that for $V \in \mathcal{B}(\mathcal{Y})$ with $\nu_n(\tau^{-1}(V)) > 0$ and for $A \in \mathcal{B}(\mathcal{X})$

$$\nu_n(A|\tau^{-1}(V)) = \mu_n(A \times \mathcal{X}|\mathcal{X} \times V). \quad (4.81)$$

In this sense also Theorem 4.1.3 follows from Theorem 4.1.2. We present some of the equivalent statements of Theorem 4.6.9 in Theorem 4.7.5.

4.7.4 Remark. Because of the relation between μ_n and ν_n and between K and J , in Theorem 4.7.1 we were able to prove the large deviation principle on $\{A \times B : A \in \mathcal{B}(\mathcal{X}), B \in \mathcal{B}(\mathcal{Y})\}$. Whether it can be extended to the large deviation principle on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ is a priori not clear. However, for the purpose of using the results of Section 4.6 this is not required (as only (iv) of Section 4.6 is required). This is the main reason to define the large deviation bounds as in Definition 4.1.1.

4.7.5 Theorem. (A3) $\implies$ (A2) $\implies$ (A1). If $\mathcal{Y}$ is first countable, then (A1) $\iff$ (A2).

(A1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\nu_n \circ \tau^{-1})$ and $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation lower bound with rate function I .

(A2) For all $U \in \mathcal{G}$

$$\sup_{V_0 \in \mathcal{N}_y} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\nu_n \circ \tau^{-1}) \neq \emptyset}} \frac{1}{n} \log \nu_n(\overline{U}|\tau^{-1}(V)) \geq -\inf I(U). \quad (4.82)$$

(A3) For all $U \in \mathcal{G}$

$$\begin{aligned} & \sup_{V_0 \in \mathcal{N}_y} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\nu_n \circ \tau^{-1}) \neq \emptyset}} \frac{1}{n} \log \nu_n(\overline{U}|\tau^{-1}(V)) \\ & \geq \liminf_{V \in \mathcal{N}_y} \liminf_{\substack{n \rightarrow \infty \\ n \in \mathbb{N} : \nu_n(\tau^{-1}(V)) > 0}} \frac{1}{n} \log \nu_n(U|\tau^{-1}(V)). \end{aligned} \quad (4.83)$$

(B3) $\implies$ (B2) $\implies$ (B1). If $\mathcal{Y}$ is first countable then (B1) $\iff$ (B2).

(B1) For all $(y_n)_{n \in \mathbb{N}}$ with $y_n \in \text{supp}(\nu_n \circ \tau^{-1})$ and $y_n \rightarrow y$ the sequence $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound with rate function I .

(B2) For all $U_1, \dots, U_k \in \mathcal{G}$ one has for $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$

$$\inf_{V_0 \in \mathcal{N}_y} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\nu_n \circ \tau^{-1}) \neq \emptyset}} \frac{1}{n} \log \nu_n(W^\circ | \tau^{-1}(V)) \leq -\inf I(W). \quad (4.84)$$

(B3) For all $U_1, \dots, U_k \in \mathcal{G}$ with $W = \mathcal{X} \setminus (U_1 \cup \dots \cup U_k)$

$$\begin{aligned} & \inf_{V_0 \in \mathcal{N}_y} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\nu_n \circ \tau^{-1}) \neq \emptyset}} \frac{1}{n} \log \nu_n(W^\circ | \tau^{-1}(V)) \\ & \leq \limsup_{V \in \mathcal{N}_y} \limsup_{\substack{n \rightarrow \infty \\ n \in \mathbb{N}: \nu_n(\tau^{-1}(V)) > 0}} \frac{1}{n} \log \nu_n(W | \tau^{-1}(V)). \end{aligned} \quad (4.85)$$

§4.8 An application to conditional probabilities of empirical distributions on finite sets

In terms of random variables, Sanov's Theorem gives us the large deviation principle of empirical densities $\frac{1}{n} \sum_{i=1}^n \delta_{X_i}$, where $X_1, X_2, \dots$ are independent and identically distributed random variables. We consider large deviations of $\frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ conditioning on $\frac{1}{n} \sum_{i=1}^n \delta_{Y_i} = \psi_n$, where $(X_1, Y_1), (X_2, Y_2), \dots$ are independent and identically distributed couples of random variables, both random variables attaining their values in a finite set. This large deviation principle is formalised in Theorem 4.8.2.

In this section we consider the following.

- Let $\mathcal{R}$ and $\mathcal{S}$ be finite sets equipped with the discrete topology (discrete metric). Let $\mathcal{P}(\mathcal{R}), \mathcal{P}(\mathcal{S})$ and $\mathcal{P}(\mathcal{R} \times \mathcal{S})$ be equipped by the weak topology and let $\mathfrak{d}$ denote the Prohorov metric (see Billingsley [7, Appendix III]) on each of the spaces.
- Let $\lambda \in \mathcal{P}(\mathcal{R} \times \mathcal{S})$. We assume $\lambda(\mathcal{R} \times \{s\}) > 0$ for all $s \in \mathcal{S}$.
- For $n \in \mathbb{N}$ let $L_n : \mathcal{R}^n \rightarrow \mathcal{P}(\mathcal{R})$ be given by $L_n(r) = \frac{1}{n} \sum_{i=1}^n \delta_{r_i}$ for $r = (r_1, \dots, r_n) \in \mathcal{R}^n$.
- Write $\mathcal{P}_{emp}^n(\mathcal{R}) = L_n(\mathcal{R}^n) = \{\frac{1}{n} \sum_{i=1}^n \delta_{r_i} : r_1, \dots, r_n \in \mathcal{R}\}$, similarly $\mathcal{P}_{emp}^n(\mathcal{S}) = L_n(\mathcal{S}^n)$ and $\mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S}) = L_n((\mathcal{R} \times \mathcal{S})^n)$.
- Let $\mathfrak{m} : \mathcal{P}(\mathcal{R} \times \mathcal{S}) \rightarrow \mathcal{P}(\mathcal{R}) \times \mathcal{P}(\mathcal{S})$ be the map that maps a measure in $\mathcal{P}(\mathcal{R} \times \mathcal{S})$ onto the pair of its marginals, i.e., $\mathfrak{m}$ is given by

$$\mathfrak{m}(\xi) = (\xi(\cdot \times \mathcal{S}), \xi(\mathcal{R} \times \cdot)). \quad (4.86)$$

- Let $\pi : \mathcal{P}(\mathcal{R}) \times \mathcal{P}(\mathcal{S}) \rightarrow \mathcal{P}(\mathcal{S})$ be the map given by $\pi(\xi, \zeta) = \zeta$.
- Let μ_n be the probability measure on $\mathcal{B}(\mathcal{P}(\mathcal{R})) \otimes \mathcal{B}(\mathcal{P}(\mathcal{S}))$ defined by $\mu_n = (\bigotimes_{i=1}^n \lambda) \circ L_n^{-1} \circ \mathfrak{m}^{-1}$, so that for $A \in \mathcal{B}(\mathcal{P}(\mathcal{R}))$ and $B \in \mathcal{B}(\mathcal{P}(\mathcal{S}))$

$$\mu_n(A \times B) = \left(\bigotimes_{i=1}^n \lambda \right) (L_n^{-1}(A) \times L_n^{-1}(B)). \quad (4.87)$$

- Define $\theta : \mathcal{S} \times \mathcal{B}(\mathcal{R}) \rightarrow [0, 1]$ by $\theta(s, A) = \lambda(A \times \mathcal{S} | \mathcal{R} \times \{s\})$.

- Define $\eta_n : \mathcal{P}(\mathcal{S}) \times \mathcal{B}(\mathcal{P}(\mathcal{R})) \rightarrow [0, 1]$ by

$$\eta_n(\xi, A) = \begin{cases} [\bigotimes_{i=1}^n \theta(s_i, \cdot)] \circ L_n^{-1}(A) & \xi \in \mathcal{P}_{emp}^n(\mathcal{S}), \xi = L_n(s_1, \dots, s_n) \\ & \text{for } s_1, \dots, s_n \in \mathcal{S}, \\ 0 & \xi \notin \mathcal{P}_{emp}^n(\mathcal{S}). \end{cases} \quad (4.88)$$

- Let $J : \mathcal{P}(\mathcal{R}) \times \mathcal{P}(\mathcal{S}) \rightarrow [0, \infty]$ be given by

$$J(\rho, \sigma) = \inf_{\xi \in \mathfrak{m}^{-1}(\{(\rho, \sigma)\})} H(\xi|\lambda). \quad (4.89)$$

where $H(\xi|\lambda)$ is the relative entropy of ξ with respect to λ ([24, Definition 2.1.5]).

- Let $\psi \in \mathcal{P}(\mathcal{S})$ be such that

$$\inf_{\xi \in \mathfrak{m}^{-1}(\mathcal{P}(\mathcal{R}) \times \{\psi\})} H(\xi|\lambda) < \infty. \quad (4.90)$$

4.8.1. We present some fact which follow from the assumptions with little effort; to some facts we give some explanation or references.

- (a) $\mathcal{P}_{emp}^n(\mathcal{S})$ is closed in $\mathcal{P}(\mathcal{S})$. Moreover, if ξ_k and ξ in $\mathcal{P}_{emp}^n(\mathcal{S})$ are such that $\xi_k \rightarrow \xi$, then there exist s_{ki} and q_i in $\mathcal{S}$ for $i \in \{1, \dots, n\}$ such that $\xi_k = L_n((s_{k1}, \dots, s_{kn}))$, $\xi = L_n((q_1, \dots, q_n))$ and $s_{ki} \rightarrow q_i$ for all $i \in \{1, \dots, n\}$.
- (b) $\text{supp}(\mu_n \circ \pi^{-1}) = \mathcal{P}_{emp}^n(\mathcal{S})$.
- (c) η_n is a product regular conditional kernel under μ_n with respect to π that is weakly continuous on $\mathcal{P}_{emp}^n(\mathcal{S})$.
- (d) $(\bigotimes \lambda^n \circ L_n^{-1})_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $H(\cdot|\lambda)$.
- (e) $\mathfrak{m}$ is continuous.
- (f) $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function J .

(a) follows from the fact that $\mathcal{S}$ is a finite space. (b) follows from (a), from the fact that the complement of $\mathcal{P}_{emp}^n(\mathcal{S})$ has $\mu_n \circ \pi^{-1}$ -measure zero and because $\mu_n \circ \pi^{-1}(\{L_n(s)\}) > 0$ for all $s \in \mathcal{S}^n$, which is due to the assumptions on λ . (c) follows by a straightforward calculation, the continuity follows from (a). For (d) see Sanov's Theorem (Dembo and Zeitouni [24, Theorem 6.2.10]). (e) follows from the fact that if $\xi_n \rightarrow \xi$ in $\mathcal{P}(\mathcal{R} \times \mathcal{S})$, then the $\mathcal{R}$ - and $\mathcal{S}$ -marginals of ξ_n converge to the $\mathcal{R}$ - and $\mathcal{S}$ -marginals of ξ , respectively. Then (f) follows from (e) and (d) by the contraction principle [24, Theorem 4.2.1].

In the rest of this section we prove the following theorem.

4.8.2 Theorem. *For all $(\psi_n)_{n \in \mathbb{N}}$ with $\psi_n \in \mathcal{P}_{emp}^n(\mathcal{S})$ and $\psi_n \rightarrow \psi$ the sequence $(\eta_n(\psi_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $I : \mathcal{P}(\mathcal{R}) \rightarrow [0, \infty]$, given by*

$$I(\phi) = \inf_{\xi \in \mathfrak{m}^{-1}(\{(\phi, \psi)\})} H(\xi|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(\mathcal{P}(\mathcal{R}) \times \{\psi\})} H(\xi|\lambda). \quad (4.91)$$

I is continuous on $[I < \infty]$.

As $\mathcal{P}(\mathcal{S})$ is first countable, it is sufficient to show that (A2) and (B2) of Theorem 4.6.9 hold. In 4.8.4 we use the bounds of Lemma 4.8.3 to derive other bounds which imply (A2) and (B2). The continuity of I follows by continuity of the map $\nu \mapsto H(\nu|\lambda)$ (Lemma 4.8.5).

4.8.3 Lemma. [24, Lemma 2.1.9] For $\nu \in \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ one has, with $M = \#\mathcal{R}\#\mathcal{S}$,

$$(n+1)^{-M} e^{-nH(\nu|\lambda)} \leq \left[\bigotimes_{i=1}^n \lambda \right] (L_n^{-1}(\{\nu\})) \leq e^{-nH(\nu|\lambda)}. \quad (4.92)$$

4.8.4. From Lemma 4.8.3 we obtain the following bounds for $A \in \mathcal{B}(\mathcal{P}(\mathcal{R}))$ and $B \in \mathcal{B}(\mathcal{P}(\mathcal{S}))$.

$$\begin{aligned} \mu_n(A \times B) &\leq \#L_n^{-1}(A) \#L_n^{-1}(B) e^{-n \inf_{\nu \in \mathfrak{m}^{-1}(A \times B) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda)} \\ &\leq (n+1)^M e^{-n \inf_{\nu \in \mathfrak{m}^{-1}(A \times B)} H(\nu|\lambda)}, \end{aligned} \quad (4.93)$$

$$\mu_n(A \times B) \geq (n+1)^{-M} e^{-n \inf_{\nu \in \mathfrak{m}^{-1}(A \times B) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda)}. \quad (4.94)$$

Whence

$$\begin{aligned} &\frac{1}{n} \log \left[(n+1)^{-2M} \right] - \left[\inf_{\nu \in \mathfrak{m}^{-1}(A \times B) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(\mathcal{R} \times B)} H(\xi|\lambda) \right] \\ &\leq \frac{1}{n} \log \mu_n(A \times \mathcal{S} | \mathcal{R} \times B) \\ &\leq \frac{1}{n} \log \left[(n+1)^{2M} \right] - \left[\inf_{\nu \in \mathfrak{m}^{-1}(A \times B)} H(\nu|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(\mathcal{R} \times B) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\xi|\lambda) \right]. \end{aligned} \quad (4.95)$$

In order to derive (A2) and (B2) of Theorem 4.6.9 we make the following observation. By (4.95) we have for an open U and a closed W that if for both $A = U$ and $C = \mathcal{R}$ as well as $A = \mathcal{R}$ and $C = W$ we have

$$\begin{aligned} &\inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \mathcal{P}_{emp}^n(\mathcal{S}) \neq \emptyset}} \left[\inf_{\nu \in \mathfrak{m}^{-1}(\overline{A} \times V) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(C^\circ \times V)} H(\xi|\lambda) \right] \\ &\leq \inf_{\nu \in \mathfrak{m}^{-1}(A \times \{\psi\})} H(\nu|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(C \times \{\psi\})} H(\xi|\lambda), \end{aligned} \quad (4.96)$$

then

$$\sup_{V_0 \in \mathcal{N}_\psi} \liminf_{n \rightarrow \infty} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{S} | \mathcal{R} \times V) \geq -\inf I(U), \quad (4.97)$$

$$\inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \frac{1}{n} \log \mu_n(W^\circ \times \mathcal{S} | \mathcal{R} \times V) \leq -\inf I(W), \quad (4.98)$$

As

$$\begin{aligned}
 & \inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \mathcal{P}_{emp}^n(\mathcal{S}) \neq \emptyset}} \left[\inf_{\nu \in \mathfrak{m}^{-1}(A \times V) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) - \inf_{\xi \in \mathfrak{m}^{-1}(C \times V)} H(\xi|\lambda) \right] \\
 & \leq \inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \mathcal{P}_{emp}^n(\mathcal{S}) \neq \emptyset}} \inf_{\nu \in \mathfrak{m}^{-1}(A \times V) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) \\
 & \quad - \sup_{V_0 \in \mathcal{N}_\psi} \inf_{\substack{V \in \mathcal{H}, V \subset V_0 \\ V \cap \text{supp}(\mu_n \circ \pi^{-1}) \neq \emptyset}} \inf_{\xi \in \mathfrak{m}^{-1}(C \times V)} H(\xi|\lambda) \\
 & \leq \inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\zeta \in \mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0} \inf_{\nu \in \mathfrak{m}^{-1}(A \times \{\zeta\}) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) \\
 & \quad - \sup_{V \in \mathcal{N}_\psi} \inf_{\xi \in \mathfrak{m}^{-1}(C \times V)} H(\xi|\lambda), \tag{4.99}
 \end{aligned}$$

(4.96) holds (for both $A = U$ and $C = \mathcal{R}$ as well as for $A = \mathcal{R}$ and $C = W$, where U is open and W is closed) if for all open U and all closed W

$$\begin{aligned}
 & \inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\zeta \in \mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0} \inf_{\nu \in \mathfrak{m}^{-1}(U \times \{\zeta\}) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) \\
 & \leq \inf_{\nu \in \mathfrak{m}^{-1}(U \times \{\psi\})} H(\nu|\lambda), \tag{4.100}
 \end{aligned}$$

$$\sup_{V \in \mathcal{N}_\psi} \inf_{\xi \in \mathfrak{m}^{-1}(W \times V)} H(\xi|\lambda) \geq \inf_{\xi \in \mathfrak{m}^{-1}(W \times \{\psi\})} H(\xi|\lambda). \tag{4.101}$$

(4.101) is a consequence of Lemma 4.5.3, as $\mathfrak{m}^{-1}(W \times V) = \mathfrak{m}^{-1}(W \times \mathcal{P}(\mathcal{S})) \cap \mathfrak{m}^{-1}(\mathcal{P}(\mathcal{R}) \times V)$, the set $F = \mathfrak{m}^{-1}(W \times \mathcal{P}(\mathcal{S}))$ is closed for closed W , $\mathfrak{m}^{-1}(\mathcal{P}(\mathcal{R}) \times V) = (\pi \circ \mathfrak{m})^{-1}(V)$ and $\pi \circ \mathfrak{m}$ is continuous. The proof of inequality (4.100) requires a little more attention. First we present some facts which are used to prove this inequality in Lemma 4.8.8.

4.8.5 Lemma. [24, Remark below Definition 2.1.5] *The map $\nu \mapsto H(\nu|\lambda)$ is continuous on $[H(\cdot|\lambda) < \infty]$. In particular, for all $\varepsilon > 0$ and $\xi \in \mathcal{P}(\mathcal{R} \times \mathcal{S})$ there exists a $\Theta \in \mathcal{N}_\xi$ such that*

$$H(\nu|\lambda) - \varepsilon \leq H(\xi|\lambda) \quad (\nu \in \Theta \cap [H(\cdot|\lambda) < \infty]). \tag{4.102}$$

Consequently, I as in (4.91) is continuous on $[J < \infty]$.

4.8.6 Lemma. (a) *Let $k, l \in \mathbb{N}$ and $\zeta \in \mathcal{P}_{emp}^k(\mathcal{S})$. For all $m \geq kl$ there exists a $\nu \in \mathcal{P}_{emp}^m(\mathcal{S})$ such that $\mathfrak{d}(\nu, \zeta) < \frac{1}{l}$.*

(b) *For all open $\Theta \subset \mathcal{P}(\mathcal{S})$ there exists an $N \in \mathbb{N}$ such that $\mathcal{P}_{emp}^n(\mathcal{S}) \cap \Theta \neq \emptyset$ for all $n \geq N$.*

Proof. (a) Let $i \in \{1, \dots, k\}$. Let $\xi \in \mathcal{P}_{emp}^i(\mathcal{S})$. Then the measure $\frac{lk}{lk+i}\zeta + \frac{i}{lk+i}\xi$ is an element of $\mathcal{P}_{emp}^{lk+i}(\mathcal{S})$. For every $A \subset \mathcal{S}$

$$\left| \left[\frac{lk}{lk+i}\zeta + \frac{i}{lk+i}\xi \right](A) - \zeta(A) \right| \leq 2 \frac{i}{lk+i} \leq 2 \frac{k}{lk} = \frac{2}{l}. \tag{4.103}$$

By definition of the Prohorov metric, this implies $\mathfrak{d}([\frac{lk}{lk+i}\zeta + \frac{i}{lk+i}\xi], \zeta) \leq \frac{2}{l}$.

(b) Let $\xi \in \mathcal{P}(\mathcal{S})$ and $\delta > 0$ be such that $B(\xi, \delta) \subset \Theta$. For each $\xi \in \mathcal{P}(\mathcal{S})$ there is a $k \in \mathbb{N}$ and a $\zeta \in \mathcal{P}_{emp}^k(\mathcal{S})$ such that $\mathfrak{d}(\zeta, \xi) < \frac{\delta}{2}$. Because of this (b) follows from (a) by letting l be such that $\frac{1}{l} < \frac{\delta}{2}$ and $N = lk$. $\square$

4.8.7 Lemma. *Let $\xi \in \mathcal{P}(\mathcal{R} \times \mathcal{S})$, $\pi \circ \mathbf{m}(\xi) = \psi$ and $\xi \ll \lambda$. For all $\delta > 0$ there exists a $\kappa > 0$ and an $N \in \mathbb{N}$ such that for all $n \geq N$ and all $\zeta \in \mathcal{P}_{emp}^n(\mathcal{S})$ with $\mathfrak{d}(\zeta, \psi) < \kappa$ there is a $\nu \in \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ with*

$$\pi \circ \mathbf{m}(\nu) = \zeta, \quad \nu \ll \lambda, \quad \mathfrak{d}(\nu, \xi) < \delta, \quad \mathfrak{d}(\nu(\cdot \times \mathcal{S}), \xi(\cdot \times \mathcal{S})) < \delta. \quad (4.104)$$

Proof. In this proof, for a measure $\xi \in \mathcal{P}(\mathcal{R} \times \mathcal{S})$, we write $\xi_{rs} = \xi(\{(r, s)\})$, so that $\xi = \sum_{rs} \xi_{rs} \delta_{(r,s)}$ where we use the short-hand notation “ $\sum_{rs}$ ” instead of “ $\sum_{r \in \mathcal{R}, s \in \mathcal{S}}$ ”. Let $M = \#\mathcal{R}\#\mathcal{S}$. Note that

$$\mathfrak{d}(\xi, \nu) \leq M \max_{r \in \mathcal{R}, s \in \mathcal{S}} |\xi_{rs} - \nu_{rs}| \quad (\xi, \nu \in \mathcal{P}(\mathcal{R} \times \mathcal{S})). \quad (4.105)$$

Let $\kappa > 0$ and $n \in \mathbb{N}$. We first give an estimation by which it is clear which κ and N one should choose. By the assumptions on λ for every $s \in \mathcal{S}$ there exists a $r_s \in \mathcal{R}$ with $\lambda_{r_s s} > 0$.

First we show that there exists a $\xi^* \in \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ with $\xi^* \ll \xi$ and $|\xi_{rs}^* - \xi_{rs}| \leq \frac{2}{n}$ for all $r \in \mathcal{R}$ and $s \in \mathcal{S}$. For each pair $(r, s) \in \mathcal{R} \times \mathcal{S}$ with $\xi_{rs} > 0$ we can choose a $\xi'_{rs} \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ such that $|\xi_{rs} - \xi'_{rs}| < \frac{1}{n}$. By letting $\xi_{rs}^* = 0$ when $\xi_{rs} = 0$ and add or subtract $\frac{1}{n}$ to some of the ξ'_{rs} we obtain a collection of $\xi_{rs}^* \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ with $\sum_{rs} \xi_{rs}^* = 1$ and $|\xi_{rs}^* - \xi_{rs}| \leq \frac{2}{n}$ and $\xi_{rs}^* = 0$ whenever $\xi_{rs} = 0$ for all $r \in \mathcal{R}$ and $s \in \mathcal{S}$.

Let $\xi \in \mathcal{P}(\mathcal{R} \times \mathcal{S})$. Suppose that $\zeta \in \mathcal{P}_{emp}^n(\mathcal{S})$ is such that $|\zeta_s - \sum_r \xi_{rs}| < \kappa$. Then $|\zeta_s - \sum_r \xi_{rs}^*| < \kappa + \frac{2}{n}M$. We construct a $\nu \in \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ by defining the ν_{rs} by each s separately. Let $s \in \mathcal{S}$. If $\zeta_s - \sum_r \xi_{rs}^* < 0$, then we choose $\nu_{rs} \leq \xi_{rs}^*$ with $\nu_{rs} \in \{0, \frac{1}{n}, \dots, 1\}$ in such way that $\sum_r \nu_{rs} = \zeta_s$ (note that $|\nu_{rs} - \xi_{rs}^*| \leq |\zeta_s - \sum_r \xi_{rs}^*|$). While, if $\zeta_s - \sum_r \xi_{rs}^* \geq 0$, then we let $\nu_{rs} = \xi_{rs}^*$ for all $r \neq r_s$ and we let $\nu_{r_s s} = \xi_{r_s s}^* + \zeta_s - \sum_r \xi_{rs}^*$ (so that $\sum_r \nu_{rs} = \zeta_s$). As $\xi^* \ll \xi$ and $\xi \ll \lambda$, by the construction of ν we have $\nu \ll \lambda$. Moreover, we have $\pi \circ \mathbf{m}(\nu) = \zeta$ and

$$\begin{aligned} \max_{r \in \mathcal{R}, s \in \mathcal{S}} |\nu_{rs} - \xi_{rs}| &\leq \max_{s \in \mathcal{S}} \left| \zeta_s - \sum_r \xi_{rs}^* \right| + \max_{r \in \mathcal{R}, s \in \mathcal{S}} |\xi_{rs}^* - \xi_{rs}| \\ &\leq \kappa + \frac{2}{n}M + \frac{2}{n}. \end{aligned} \quad (4.106)$$

Which implies by (4.105)

$$\mathfrak{d}(\nu, \xi) \leq M\kappa + \frac{2}{n}(M^2 + M). \quad (4.107)$$

Moreover, as $|\sum_s \nu_{rs} - \sum_s \xi_{rs}| \leq M \max_{s \in \mathcal{S}} |\nu_{rs} - \xi_{rs}|$,

$$\mathfrak{d}(\nu(\cdot \times \mathcal{S}), \xi(\cdot \times \mathcal{S})) \leq M \max_{r \in \mathcal{R}} \left| \sum_s \nu_{rs} - \sum_s \xi_{rs} \right| \leq M^2\kappa + \frac{2}{n}(M^3 + M^2). \quad (4.108)$$

By choosing $\kappa > 0$ and $N \in \mathbb{N}$ such that $M^2\kappa + \frac{2}{n}(M^3 + M^2) < \delta$ the proof is complete. $\square$

4.8.8 Lemma. For all open $U \subset \mathcal{R}$

$$\begin{aligned} 0 &\leq \inf_{V_0 \in \mathcal{N}_\psi} \limsup_{n \rightarrow \infty} \sup_{\zeta \in \mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0} \inf_{\nu \in \mathbf{m}^{-1}(U \times \{\zeta\}) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})} H(\nu|\lambda) \\ &\leq \inf_{\nu \in \mathbf{m}^{-1}(U \times \{\psi\})} H(\nu|\lambda). \end{aligned} \quad (4.109)$$

Proof. We assume $\inf_{\nu \in \mathbf{m}^{-1}(U \times \{\psi\})} H(\nu|\lambda) < \infty$. Let $\xi \in \mathbf{m}^{-1}(U \times \{\psi\})$ be such that $H(\xi|\lambda) < \infty$. Let $\varepsilon > 0$. We show there exists a $V_0 \in \mathcal{N}_\psi$ and an $N \in \mathbb{N}$ such that for all $n \geq N$ the set $\mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0$ is not empty and for all $\zeta \in \mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0$ there exists a $\nu \in \mathbf{m}^{-1}(U \times \{\zeta\}) \cap \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ with

$$H(\nu|\lambda) - \varepsilon \leq H(\xi|\lambda). \quad (4.110)$$

Let δ be such that (see Lemma 4.8.5)

$$B(\xi(\cdot \times \mathcal{S}), \delta) \subset U, \quad (4.111)$$

$$H(\nu|\lambda) - \varepsilon \leq H(\xi|\lambda) \quad (\nu \in B(\xi, \delta) \cap [H(\cdot|\lambda) < \infty]). \quad (4.112)$$

Then let $\kappa > 0$ and $N \in \mathbb{N}$ be as in Lemma 4.8.7. Let $V_0 = B(\psi, \kappa)$. By Lemma 4.8.6 we may assume that N is large enough such that $\mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0 \neq \emptyset$. Let $n \geq N$ and $\zeta \in \mathcal{P}_{emp}^n(\mathcal{S}) \cap V_0$. By Lemma 4.8.7 there exists a $\nu \in \mathcal{P}_{emp}^n(\mathcal{R} \times \mathcal{S})$ with $\pi \circ \mathbf{m}(\nu) = \zeta$, $\nu \ll \lambda$ and $\nu(\cdot \times \mathcal{S}) \in B(\xi(\cdot \times \mathcal{S}), \delta)$, $\nu \in B(\xi, \delta)$, i.e., by (4.111), $\nu \in \mathbf{m}^{-1}(U \times \{\zeta\})$. $\nu \ll \lambda$ implies $\nu \in [H(\cdot|\lambda) < \infty]$, thus with (4.112) we obtain (4.110). $\square$

§4.9 Examples

In Section 4.8 we showed that the regular conditional kernel η_n as in (4.88) satisfies (A1) and (B1) of Theorem 4.6.9 by showing that (A2) and (B2) of that theorem hold. This is not always the most optimal approach; in Example 4.9.1 we show that for a specific example of Gaussian measures the expression of η_n allows us to derive (A1) and (B1) directly.

Furthermore, relying on Theorem 4.9.2, in Example 4.9.4, we give an example of a $(\eta_n)_{n \in \mathbb{N}}$ for which (A1) of Theorem 4.6.9 does not hold. In Remark 4.9.5 we mention that for the one choice of measures in Example 4.9.4 a quenched large deviation principle is satisfied, while for the other choice of measures there is no quenched large deviation principle. In Example 4.9.6 we show that for a choice of measures as in Example 4.9.4 the conditional regular kernel in a specific chosen point does not satisfy any large deviation principle. In Remark 4.9.7 we discuss exponential tightness of the regular conditional kernel. In Remark 4.9.8 we discuss the differences between the present paper and the paper of La Cour and Schieve [22].

4.9.1 Example. Let $r \neq 0$, $Z_n := \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-\frac{n}{2}(x^2 - 2rxy + y^2)} dx dy$ and consider the sequence of probability measures $(\mu_n)_{n \in \mathbb{N}}$ on $\mathcal{B}(\mathbb{R} \times \mathbb{R})$ determined by

$$\mu_n(A \times B) = \frac{1}{Z_n} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{A \times B}(x, y) e^{-\frac{n}{2}(x^2 - 2rxy + y^2)} dx dy \quad (A, B \in \mathcal{B}(\mathbb{R})). \quad (4.113)$$

The sequence satisfies the large deviation principle with rate function $J : \mathbb{R}^2 \rightarrow [0, \infty]$ given by $J(x, y) = \frac{1}{2}(x^2 - 2rxy + y^2)$. By Theorem 4.4.7 η_n given by

$$\eta_n(y, A) = \frac{\int_A e^{-\frac{n}{2}(x^2 - 2rxy)} dx}{\int_{\mathbb{R}} e^{-\frac{n}{2}(x^2 - 2rxy)} dx} = \frac{\int_A e^{-\frac{n}{2}(x - ry)^2} dx}{\int_{\mathbb{R}} e^{-\frac{n}{2}(x - ry)^2} dx}, \quad (4.114)$$

is the weakly continuous product regular conditional probability under μ_n with respect to the projection on the $\mathcal{Y}$ -coordinate. If $y_n \rightarrow y$, one can show that for $\lambda \in \mathbb{R}$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{R}} e^{n\lambda x} d[\eta_n(y_n, \cdot)](x) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \int_{\mathbb{R}} e^{n\lambda x} d[\eta_n(y, \cdot)](x) \\ &= \lambda ry + \frac{1}{2}\lambda^2. \end{aligned} \quad (4.115)$$

Then by the Gärtner-Ellis Theorem (see for example Dembo and Zeitouni [24, Theorem 2.3.6]) we conclude that $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with the same rate function as the one of the large deviation principle of $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$, which is $x \mapsto (x - ry)^2$. Note that this equals $J(x, y) - \inf J(\mathbb{R} \times \{y\})$ because of the equality $x^2 - 2rxy + y^2 = (x - ry)^2 + (1 - r^2)y^2$.

The proof of the following theorem can be found in Appendix 4.C.

4.9.2 Theorem. *Let $\mathcal{X}$ and $\mathcal{Y}$ be separable metric spaces. Let $(\mu_n^1)_{n \in \mathbb{N}}$ and $(\mu_n^2)_{n \in \mathbb{N}}$ be sequences of probability measures on $\mathcal{B}(\mathcal{X})$. Let $(\nu_n)_{n \in \mathbb{N}}$ be a sequence of probability measures on $\mathcal{B}(\mathcal{Y})$ that satisfies the large deviation principle with a rate function $L : \mathcal{Y} \rightarrow [0, \infty]$. Suppose that $y \in \mathcal{Y}$ and $W_n \in \mathcal{N}_y$ are such that $\bigcap_{n \in \mathbb{N}} W_n = \{y\}$ and $\alpha_n : \mathcal{Y} \rightarrow [0, 1]$ is a continuous function with $\alpha_n(y) = 0$ and $\alpha_n = 1$ on $\mathcal{Y} \setminus W_n$ such that*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(\int_{W_n} \alpha_n d\nu_n \right) = 0, \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(\int_{W_n} (1 - \alpha_n) d\nu_n \right) = 0. \quad (4.116)$$

Assume $(\mu_n^1)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I . Assume furthermore that for all open $A \subset \mathcal{X}$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(A) \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^2(A), \quad (4.117)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(\mathcal{X} \setminus A) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^2(\mathcal{X} \setminus A). \quad (4.118)$$

Let μ_n be the probability measure on $\mathcal{B}(\mathcal{X}) \otimes \mathcal{B}(\mathcal{Y})$ for which for $A \in \mathcal{B}(\mathcal{X})$, $B \in \mathcal{B}(\mathcal{Y})$

$$\mu_n(A \times B) = \mu_n^1(A) \int_{\mathcal{Y}} \mathbb{1}_B \alpha_n d\nu_n + \mu_n^2(A) \int_{\mathcal{Y}} \mathbb{1}_B (1 - \alpha_n) d\nu_n. \quad (4.119)$$

Then $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $J : \mathcal{X} \times \mathcal{Y} \rightarrow [0, \infty]$ given by $J(x, y) = I(x) + L(y)$. $\eta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ defined by

$$\eta_n(y, A) = \alpha_n(y) \mu_n^1(A) + (1 - \alpha_n(y)) \mu_n^2(A) \quad (4.120)$$

is the weakly continuous product regular conditional probability under μ_n with respect to $\pi : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ given by $\pi(x, y) = y$.

Note that $I(x) = J(x, y) - \inf J(\mathcal{X} \times \{y\})$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$.

4.9.3 Examples. We give examples of $\mathcal{Y}, W_n, \alpha_n, \nu_n$ and L such that (4.116) of Theorem 4.9.2 is satisfied and $(\nu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function L .

- (a) Let $\mathcal{Y} = [0, \infty)$, $\alpha_n(y) = \min\{ny, 1\}$ for $y \in \mathcal{Y}$ and for $B \in \mathcal{B}([0, \infty))$ let $\nu_n(B) = \int_0^\infty \mathbb{1}_B(y) n e^{-ny} dy$. Then $\int_0^\infty \alpha_n d\nu_n = 1 - 2e^{-1}$ and $\int_0^\infty (1 - \alpha_n) d\nu_n = e^{-1}$. Therefore with this ν_n , α_n and $W_n = [-\frac{1}{n}, \frac{1}{n}]$ (4.116) is satisfied. Moreover $(\nu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $L : \mathcal{Y} \rightarrow [0, \infty]$, $L(y) = y$ (this follows from example by the Gärtner-Ellis Theorem [24, Theorem 2.3.6]).
- (b) Let $\mathcal{Y} = \mathbb{R}$ and $\nu_n = \mu_{\mathcal{N}(0, \frac{1}{n})}$ (the Gaussian measure corresponding to a $\mathcal{N}(0, \frac{1}{n})$ distributed random variable). Then there exists a decreasing sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ with $\varepsilon_n \downarrow 0$, such that with $W_n = [-\varepsilon_n, \varepsilon_n]$ there exist functions α_n as in Theorem 4.9.2 such that (4.116) is satisfied (see the postscript). With $\nu_n^0 = \frac{1}{2}\delta_0 + \frac{1}{2}\nu_n$ instead of ν_n , (4.116) is also satisfied. Moreover, $(\nu_n)_{n \in \mathbb{N}}$ and $(\nu_n^0)_{n \in \mathbb{N}}$ (use Lemma 4.A.1) satisfy the large deviation principle with rate function $L : \mathcal{Y} \rightarrow [0, \infty]$, $L(y) = \frac{1}{2}y^2$.
- Postscript. Let $\beta = \nu_1([-1, 1])$. Let $\kappa_n = \frac{1}{\sqrt{n}}$. Then $\nu_n[-\kappa_n, \kappa_n] = \beta$ for all $n \in \mathbb{N}$. Let $\phi_\varepsilon : \mathbb{R} \rightarrow [0, 1]$ be defined by $\phi_\varepsilon(z) = \min\{\varepsilon^{-1}|z|, 1\}$. Then $\lim_{\varepsilon \downarrow 0} \int_{[-\kappa_n, \kappa_n]} \phi_\varepsilon d\nu_1 = \beta$, $\lim_{\varepsilon \downarrow 0} \int_{[-\kappa_n, \kappa_n]} 1 - \phi_\varepsilon d\nu_1 = 0$ and

$$\int_{[-\kappa_n, \kappa_n]} \phi_{\kappa_n} d\nu_n < \int_{[-\kappa_n, \kappa_n]} 1 - \phi_{\kappa_n} d\nu_n. \quad (4.121)$$

Therefore, for all $n \in \mathbb{N}$, there exists an $\varepsilon_n \in (0, \kappa_n)$ such that

$$\int_{[-\kappa_n, \kappa_n]} \phi_{\varepsilon_n} d\nu_n = \frac{1}{2}\beta = \int_{[-\kappa_n, \kappa_n]} 1 - \phi_{\varepsilon_n} d\nu_n, \quad (4.122)$$

With $\alpha_n = \phi_{\varepsilon_n}$, (4.116) as in Theorem 4.9.2 is satisfied.

4.9.4 Example. With $\mathcal{X} = \mathbb{R}$, $\mu_n^1 = \mu_{\mathcal{N}(0, \frac{1}{n})}$, $\mu_n^2 = \delta_{\frac{1}{n}}$ and $I(x) = \frac{1}{2}x^2$ for $x \in \mathbb{R}$ and $\mathcal{Y}, \nu_n$ (or ν_n^0), α_n , W_n and L as in Examples 4.9.3 (a) or (b) the conditions of Theorem 4.9.2 are satisfied (note that $(\delta_{\frac{1}{n}})_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function $H : \mathbb{R} \rightarrow [0, \infty]$ given by $H(0) = 0$ and $H(x) = \infty$ for $x \neq 0$).

Then $\eta_n(0, \cdot) = \delta_{\frac{1}{n}}$ and $\eta_n(\varepsilon_n, \cdot) = \mu_{\mathcal{N}(0, \frac{1}{n})}$ for all $n \in \mathbb{N}$. Whence $(\eta_n(0, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function H and $(\eta_n(\varepsilon_n, \cdot))_{n \in \mathbb{N}}$ (and also $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$ for $y > 0$) satisfies the large deviation principle with rate function I . Because $I \geq H$, the sequence $(\eta_n(0, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation upper bound not only with I but also with H instead of I . Therefore (b1) of Theorem 4.6.3 holds in case $y_n = 0$ for all n . Since $(\eta_n(0, \cdot))_{n \in \mathbb{N}}$ does not satisfy the large deviation principle with rate function I , (a1) of Theorem 4.6.3 does not hold. Therefore for any decreasing sequence $(V_m)_{m \in \mathbb{N}}$ in $\mathcal{N}_0$ with $\bigcap_{m \in \mathbb{N}} V_m = \{0\}$ there exists an open set U with $\inf I(U) < \infty$ with

$$\liminf_{n \rightarrow \infty} \limsup_{m \rightarrow \infty} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_m) < -\inf I(U). \quad (4.123)$$

We illustrate this for $\mathcal{Y}, \alpha_n, W_n, \nu_n$ and L as in Examples 4.9.3(a): For $V_m = [0, \frac{1}{m})$, $U = (1, \infty)$ we get for $m \geq n$

$$\mu_n(U \times V_m) = \mu_{\mathcal{N}(0, \frac{1}{n})}(U) \int_0^{\frac{1}{m}} ny \cdot ne^{-ny} dy, \quad (4.124)$$

$$\mu_n(\mathcal{X} \times V_m) = \int_0^{\frac{1}{m}} ne^{-ny} dy. \quad (4.125)$$

Since $\int_0^{\frac{1}{m}} ny \cdot ne^{-ny} dy \leq \frac{n}{m} \int_0^{\frac{1}{m}} ne^{-ny} dy$ we get

$$\mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_m) \leq \frac{n}{m} \mu_{\mathcal{N}(0, \frac{1}{n})}(U) \quad (4.126)$$

which converges to zero as $m \rightarrow \infty$, which implies

$$\limsup_{m \rightarrow \infty} \frac{1}{n} \log \mu_n(U \times \mathcal{Y} | \mathcal{X} \times V_m) = -\infty < -\frac{1}{2} = -\inf I(U). \quad (4.127)$$

4.9.5 Remark (Quenched large deviations). Consider the situation as in Example 4.9.4. For all $n \in \mathbb{N}$ we have the following. If $\zeta_n : \mathcal{Y} \times \mathcal{B}(\mathcal{X}) \rightarrow [0, 1]$ is a product regular conditional probability under μ_n with respect to π , then $\zeta_n(y, \cdot) = \eta_n(y, \cdot)$ for $[\mu_n \circ \pi^{-1}]$ -almost all y (see Remark 4.3.5).

Whence, with ν_n as in Examples 4.9.3 (a) or (b), we have a quenched large deviation principle of the conditional probability with respect to the second coordinate with rate function I ; for every product regular conditional probability ζ_n under μ_n with respect to π there exists a $Z \subset \mathcal{Y}$ with $\mu_n \circ \pi^{-1}(Z) = \nu_n(Z) = 1$ such that $(\zeta_n(y, \cdot))_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I for all $y \in Z$.

However, with ν_n^0 as in Examples 4.9.3(b) instead of ν_n for such ζ one has $\zeta_n(0, \cdot) = \eta_n(0, \cdot)$ as $\nu_n^0(\{0\}) > 0$. Thus in this case we do not have such a quenched large deviation principle.

4.9.6 Example. With $\mathcal{X} = \mathbb{N}$, $\mu_n^1 = \sum_{k \in \mathbb{N}} 2^{-k} \delta_k$, $\mu_n^2 = \delta_n$ and $I(x) = 0$ for $x \in \mathbb{N}$ as in Example 4.9.4, and $\mathcal{Y}, W_n, \alpha_n, \nu_n$ and L as in Examples 4.9.3(a) or (b), the conditions of Theorem 4.9.2 are satisfied. In this case $(\eta_n(0, \cdot))_{n \in \mathbb{N}}$ does not satisfy a large deviation principle.

4.9.7 Remark. (Exponential tightness of the regular conditional kernel).

Considering the situation as in Theorem 4.9.2, we would like to mention that if $(\mu_n^1)_{n \in \mathbb{N}}$ is exponentially tight, then so is $(\mu_n)_{n \in \mathbb{N}}$ since $\mu_n(K_1^c \times K_2^c) = \mu_n^1(K_1^c) \nu_n(K_2^c)$ for large n and (compact) $K_1 \subset \mathcal{X}, K_2 \subset \mathcal{Y}$. Similarly $(\eta_n(y, \cdot))_{n \in \mathbb{N}}$ is exponentially tight for all $y > 0$ since $\eta_n(y, K^c) = \mu_n^1(K^c)$ for large n and compact $K \subset \mathcal{X}$. However, as is the case in Example 4.9.6, $(\eta_n(y_n, \cdot))_{n \in \mathbb{N}}$ need not be exponentially tight for all converging sequences $(y_n)_{n \in \mathbb{N}}$ (e.g., if $(\mu_n^2)_{n \in \mathbb{N}}$ is not exponentially tight, then $(\eta_n(0, \cdot))_{n \in \mathbb{N}}$ is neither).

4.9.8 Remark. Example 4.9.4 with ν_n (or ν_n^0) and α_n as in Examples 4.9.3(b) fits the assumptions made in Section 4 of La Cour and Schieve [22].⁸ In that paper it

⁸The logarithmic moment generating function (see Dembo and Zeitouni [24, Assumption 2.3.2]) is given by $(x, y) \mapsto \frac{1}{2}x^2 + \frac{1}{2}y^2$, whence the Hessian of it equals the identity matrix and is therefore invertible. In [22] is mentioned that one can not proceed the conditioning on all elements, but only those that equal the derivative of $y \mapsto \frac{1}{2}y^2$ at a certain point are considered, of which 0 is an example.

is claimed that the law of the first coordinate conditioned on the second coordinate satisfies the large deviation principle with the rate function I . Their notion of conditioning on y is “condition on an arbitrarily small neighbourhood around y ”. This approach needs to be justified. Our results are different, as by Example 4.9.4 the conditioned kernel in 0, $\eta_n(0, \cdot)$ does not satisfy the large deviation principle with the rate function I (even in the sense of quenched large deviations as discussed in Remark 4.9.5).

Appendix

§4.A An elementary fact about limsup and liminf

4.A.1 Lemma. *Let $k \in \mathbb{N}$ and $a_n^i \in [0, \infty)$ for all $n \in \mathbb{N}$ and $i \in \{1, \dots, k\}$. If there exists an $N \in \mathbb{N}$ such that $\max_{i \in \{1, \dots, k\}} a_n^i > 0$ for all $n \geq N$, then*⁹

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \log \left(\sum_{i=1}^k a_n^i \right) - \max_{i \in \{1, \dots, k\}} \frac{1}{n} \log a_n^i \right) = 0, \quad (4.128)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{i=1}^k a_n^i \right) = \max_{i \in \{1, \dots, k\}} \limsup_{n \rightarrow \infty} \frac{1}{n} \log a_n^i, \quad (4.129)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(\sum_{i=1}^k a_n^i \right) = \max_{i \in \{1, \dots, k\}} \liminf_{n \rightarrow \infty} \frac{1}{n} \log a_n^i. \quad (4.130)$$

Proof. (4.128), (4.129) and (4.130) follow from the inequality

$$\begin{aligned} \max_{i \in \{1, \dots, k\}} \frac{1}{n} \log a_n^i &\leq \frac{1}{n} \log \left(\sum_{i=1}^k a_n^i \right) \leq \frac{1}{n} \log \left(k \max_{i \in \{1, \dots, k\}} a_n^i \right) \\ &\leq \frac{1}{n} \log k + \max_{i \in \{1, \dots, k\}} \frac{1}{n} \log(a_n^i). \end{aligned} \quad (4.131)$$

□

§4.B Sufficient bounds for large deviation bounds

Let $\mathcal{X}$ be a topological space. Let $I : \mathcal{X} \rightarrow [0, \infty]$ have compact sublevel sets. Let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence of probability measures on $\mathcal{B}(\mathcal{X})$.

4.B.1 Lemma. *Let $(F_m)_{m \in \mathbb{N}}$ be a decreasing sequence of closed sets with $F = \bigcap_{m \in \mathbb{N}} F_m$. Then*

$$\sup_{m \in \mathbb{N}} \inf I(F_m) = \inf I(F). \quad (4.132)$$

⁹Equation (4.129) can also be found in Dembo and Zeitouni [24, Theorem 1.2.15]

Proof. Let $c := \sup_{m \in \mathbb{N}} \inf I(F_m)$. Note that $c \leq \inf I(F)$. If $c = \infty$ there is nothing to prove. Assume that $c < \infty$. Let K be the compact set $[I \leq c]$. Then $F_m \cap K \neq \emptyset$ for all $m \in \mathbb{N}$, whence $F \cap K \neq \emptyset$ and thus $\inf I(F) \leq c$. $\square$

4.B.2. For Lemma 4.B.1 the condition that I has compact sublevel sets is not redundant. For example: Let $I : \mathbb{N} \cup \{0\} \rightarrow [0, \infty]$ be given by $I(0) = 1$ and $I(x) = 0$ for $x \in \mathbb{N}$. Then for $F_m = \{0\} \cup \{m, m+1, \dots\}$ and $F = \{0\}$ one has $\sup_{m \in \mathbb{N}} \inf I(F_m) = 0$ and $\inf I(F) = 1$.

4.B.3 Lemma.

(a) $\mathcal{U}$ be a set of open subsets of $\mathcal{X}$. Suppose that for all $G \in \mathcal{U}$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq -\inf I(G). \quad (4.133)$$

Then $G = \bigcup \mathcal{U}$ satisfies (4.133) as well.¹⁰

(b) Let $F_1, F_2, \dots$ be closed. Suppose that for all $F \in \{F_m : m \in \mathbb{N}\}$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq -\inf I(F). \quad (4.134)$$

Then $F = \bigcap_{m \in \mathbb{N}} F_m$ satisfies (4.134) as well.

Proof.

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n\left(\bigcup \mathcal{U}\right) \geq \sup_{G \in \mathcal{U}} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(G) \geq \sup_{G \in \mathcal{U}} (-\inf I(G)), \quad (4.135)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n\left(\bigcap_{m \in \mathbb{N}} F_m\right) \leq \inf_{m \in \mathbb{N}} \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F_m) \leq \inf_{m \in \mathbb{N}} (-\inf I(F_m)). \quad (4.136)$$

Now apply Lemma 4.B.1. $\square$

As a consequence of Lemma 4.B.3 we obtain the following.

4.B.4 Theorem. Suppose that $\mathcal{G}$ is a basis for the topology on $\mathcal{X}$, such that (4.133) holds for all $G \in \mathcal{G}$ and (4.134) holds for $F = \mathcal{X} \setminus G$. Suppose that every open G can be written as countable union of elements in $\mathcal{G}$. Then $(\mu_n)_{n \in \mathbb{N}}$ satisfies the large deviation principle with rate function I .

§4.C Proof of Theorem 4.9.2

Proof of Theorem 4.9.2. As $\mathcal{X}$ and $\mathcal{Y}$ are separable metric spaces, every open subset of $\mathcal{X} \times \mathcal{Y}$ is a countable union of elements of the form $A \times B$ where $A \subset \mathcal{X}$ is open and $B \in \mathcal{H}$, where (with d_Y the metric on $\mathcal{Y}$)

$$\mathcal{H} = \{B(y, \delta) : \delta > 0\} \cup \{B(z, \delta) : z \neq y, 0 < \delta < d_Y(y, z)\}. \quad (4.137)$$

¹⁰This can also be found in O'Brien [14, Proposition 2.1].

We use Theorem 4.B.4 to prove the large deviation bounds. Note first that $(\mathcal{X} \times \mathcal{Y}) \setminus (A \times B) = (\mathcal{X} \times (\mathcal{Y} \setminus B)) \cup ((\mathcal{X} \setminus A) \times \mathcal{Y})$, that $\min\{\inf I(\mathcal{X} \setminus A), \inf L(\mathcal{Y} \setminus B)\} = \inf_{(x,y) \in (\mathcal{X} \times \mathcal{Y}) \setminus (A \times B)} I(x) + L(y)$ and that by (4.129)

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((\mathcal{X} \times \mathcal{Y}) \setminus (A \times B)) \\ & \leq \max \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \times (\mathcal{Y} \setminus B)), \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((\mathcal{X} \setminus A) \times \mathcal{Y}) \right\}. \end{aligned}$$

Using this and Theorem 4.B.4 it is sufficient to show that for all open sets $A \subset \mathcal{X}$ and $B \subset \mathcal{Y}$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(\mathcal{X} \times (\mathcal{Y} \setminus B)) \leq -\inf L(\mathcal{Y} \setminus B), \quad (4.138)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((\mathcal{X} \setminus A) \times \mathcal{Y}) \leq -\inf I(\mathcal{X} \setminus A) \quad (4.139)$$

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) \geq -\inf I(A) - \inf L(B). \quad (4.140)$$

Let $A \subset \mathcal{X}$ be open and $B \in \mathcal{H}$.

- (4.138) follows from the fact that $\mu_n(\mathcal{X} \times (\mathcal{Y} \setminus B)) = \nu_n(\mathcal{Y} \setminus B)$.
- (4.139) follows from the fact that by (4.116), (4.118) and (4.129) we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n((\mathcal{X} \setminus A) \times \mathcal{Y}) \\ & = \max \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(\mathcal{X} \setminus A), \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^2(\mathcal{X} \setminus A) \right\} \\ & = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(\mathcal{X} \setminus A) \leq -\inf I(\mathcal{X} \setminus A), \end{aligned} \quad (4.141)$$

- (4.140) follows by separating two cases (as either $y \in B$ or $y \notin \overline{B}$):

If $y \notin \overline{B}$, then $W_n \cap B = \emptyset$ and so $\mu_n(A \times B) = \mu_n^1(A) \nu_n(B)$ for large n , whence

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) = \liminf_{n \rightarrow \infty} \left(\frac{1}{n} \log \mu_n^1(A) + \frac{1}{n} \log \nu_n(B) \right). \quad (4.142)$$

Suppose that $y \in B$, i.e., $W_n \subset B$ for large n . By (4.130) we obtain

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) \\ & = \max \left\{ \liminf_{n \rightarrow \infty} \left(\frac{1}{n} \log \mu_n^1(A) + \frac{1}{n} \log \left(\int_{W_n} \alpha_n d\nu_n \right) \right), \right. \\ & \quad \liminf_{n \rightarrow \infty} \left(\frac{1}{n} \log \mu_n^2(A) + \frac{1}{n} \log \left(\int_{W_n} (1 - \alpha_n) d\nu_n \right) \right), \\ & \quad \left. \liminf_{n \rightarrow \infty} \left(\frac{1}{n} \log \mu_n^1(A) + \frac{1}{n} \log \left(\int_{W_n^c} \mathbb{1}_B d\nu_n \right) \right) \right\} \end{aligned} \quad (4.143)$$

Using that $\liminf_{n \rightarrow \infty} \frac{1}{n} \log \left(\int_{W_n^c} \mathbb{1}_B d\nu_n \right) \leq 0$ together with (4.116) and (4.117), we obtain

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(A \times B) & = \max \left\{ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^1(A), \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n^2(A) \right\} \\ & \geq -\inf I(A). \end{aligned} \quad (4.144)$$

Because $\inf L(B) \geq 0$, we conclude (4.140).

We leave it to the reader to check that η_n is the weakly continuous product regular conditional probability under μ_n with respect to π . $\square$

§4.D Miscellaneous: Additional remarks

In Section 4.D.1 we comment on the conditions of Theorem 4.6.3 and show that one cannot expect a straightforward reduction of these conditions to be feasible. In Section 4.D.2 we comment on the difference between large deviations of sequences and large deviations of nets of probability measures.

§4.D.1 Regarding the conditions of Theorem 4.6.3

4.D.1. We make an observation regarding condition (a2) of Theorem 4.6.3, which reads as follows: For all $U \in \mathcal{G}$

$$\liminf_{n \rightarrow \infty} \limsup_{V \in \mathcal{V}_n} \frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U). \quad (4.145)$$

This holds if and only if $\forall U \in \mathcal{G}, \forall \varepsilon > 0 \exists N \forall n \geq N \forall V_0 \in \mathcal{V}_n \exists V \in \mathcal{V}_n, V \subset V_0$:

$$\frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) \geq -\inf I(U) - \varepsilon. \quad (4.146)$$

Note that

$$\frac{1}{n} \log \mu_n(\overline{U} \times \mathcal{Y} | \mathcal{X} \times V) = \frac{1}{n} \log \mu_n(\overline{U} \times V) - \frac{1}{n} \log \mu_n(\mathcal{X} \times V), \quad (4.147)$$

$$\inf I(U) = \inf J(U \times \{y\}) - \inf J(\mathcal{X} \times \{y\}). \quad (4.148)$$

By the large deviation principle of $(\mu_n)_{n \in \mathbb{N}}$ and Lemma 4.5.3 we have $\forall U \in \mathcal{G}, \forall \delta > 0, \exists N \forall n \geq N \forall V_0 \in \mathcal{V}_n \exists V \in \mathcal{V}_n, V \subset V_0$:

$$\frac{1}{n} \log \mu_n(\mathcal{X} \times V) \leq \frac{1}{n} \log \mu_n(\mathcal{X} \times \overline{V}) \leq -\inf J(\mathcal{X} \times \{y\}) + \delta. \quad (4.149)$$

With this the following would be sufficient to imply condition (a2) of Theorem 4.6.3: $\forall U \in \mathcal{G}, \forall \varepsilon > 0, \exists N \forall n \geq N \forall V_0 \in \mathcal{V}_n \exists V \in \mathcal{V}_n, V \subset V_0$:

$$\frac{1}{n} \log \mu_n(\overline{U} \times V) \geq -\inf J(U \times \{y\}) - \varepsilon. \quad (4.150)$$

We mention that we cannot expect this condition to hold in general. Indeed, if $\mathcal{Y}$ is first countable, then we may assume that $\mathcal{V}_n$ is a countable set consisting of decreasing sets V_m^n with $\bigcap_{m \in \mathbb{N}} V_m^n = \{y_n\}$, and therefore the above condition regarding (4.150) implies that for all $U \in \mathcal{G}$ with $\inf J(U \times \{y\}) < \infty$ and for n large enough $\mu_n(U \times \{y_n\}) = \lim_{m \rightarrow \infty} \mu_n(U \times V_m^n) > 0$.

The approach mentioned here has been used in Section 4.8 (see 4.8.4).

§4.D.2 Large deviations over different parameters

In large deviation theory there exist basically two formulations of the large deviation principle. One is in terms of sequences, as in the books by den Hollander [50] and by Rassoul-Agha and Seppäläinen [75], the other is in terms of nets indexed by $(0, \infty)$, as in the books by Dembo and Zeitouni [24] and by Deuschel and Stroock [26]. In Theorem 4.D.2 we address the relation between these formulations.

Suppose that $D \subset (0, \infty)$ is such that $0 \in \overline{D}$. Let $\mathcal{X}$ be a topological space. Let $(\mu_\varepsilon)_{\varepsilon \in D}$ be probability measures on $\mathcal{B}(\mathcal{X})$.

4.D.2 Theorem. For $A \in \mathcal{B}(\mathcal{X})$

$$\liminf_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(A) \geq -\inf I(A^\circ) \quad (4.151)$$

if and only if for all sequences $(\varepsilon_n)_{n \in \mathbb{N}}$ in D with $\varepsilon_n \rightarrow 0$

$$\liminf_{n \rightarrow \infty} \varepsilon_n \log \mu_{\varepsilon_n}(A) \geq -\inf I(A^\circ). \quad (4.152)$$

Similarly

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(A) \leq -\inf I(\overline{A}) \quad (4.153)$$

if and only if for all sequences $(\varepsilon_n)_{n \in \mathbb{N}}$ in D with $\varepsilon_n \rightarrow 0$

$$\limsup_{n \rightarrow \infty} \varepsilon_n \log \mu_{\varepsilon_n}(A) \leq -\inf I(\overline{A}). \quad (4.154)$$

Proof. Elementary. □

PART II

INTEGRALS FOR FUNCTIONS WITH VALUES IN A PARTIALLY ORDERED VECTOR SPACE



Integrals for functions with values in a partially ordered vector space

This chapter is based on:

A. van Rooij and W. van Zuijlen. Integrals for functions with values in partially ordered vector spaces. To appear in Positivity, preprint available at <http://arxiv.org/abs/1505.05997>

Abstract

We consider integration of functions with values in a partially ordered vector space, and two notions of extension of the space of integrable functions. Applying both extensions to the space of real valued simple functions on a measure space leads to the classical space of integrable functions.

§5.1 Introduction

For functions with values in a Banach space there exist several notions of integration. The best known are the Bochner and Pettis integrals (see [9] and [74]). These have been thoroughly studied, yielding a substantial theory (see Chapter III in the book by E. Hille and R.S. Phillips, [48]).

As far as we know, there is no notion of integration for functions with values in a partially ordered vector space; not necessarily a σ -Dedekind complete Riesz space. In this paper we present such a notion. The basic idea is the following. (Here, E is a partially ordered vector space in which our integrals take their values.)

In the style of Daniell [23] and Bourbaki [12, Chapter 3,4], we do not start from a measure space but from a set X , a collection Γ of functions $X \rightarrow E$, and a functional $\varphi : \Gamma \rightarrow E$, our “elementary integral”. We describe two procedures for extending φ to a larger class of functions $X \rightarrow E$. The first (see Section 5.3), the “vertical extension”, is analogous to the usual construction of the Riemann integral, proceeding from the space of simple functions. The second (see Section 5.4), the “lateral extension”, is related to the improper Riemann integral.

In Section 5.5 we investigate what happens if one repeatedly applies those extension procedures, without considering the space E to be σ -Dedekind complete or even Archimedean. However, under some mild conditions on E one can embed E into a σ -Dedekind complete space. In Section 5.6 we discuss the extensions procedures in the larger space. Section 5.7 and Section 5.8 treat the situation in which Γ consists of the simple E -valued functions on a measure space. (In Section 5.7 we have $E = \mathbb{R}$.) In Section 5.9 we consider connections of our extensions with the Bochner and the Pettis integrals for the case where E is a Banach lattice. In Section 5.10 we apply our extensions to the Bochner integral. For more alternative approaches we refer to the discussion in Section 5.11.

§5.2 Some Notation

$\mathbb{N}$ is $\{1, 2, 3, \dots\}$.

Let X be a set. We write $\mathcal{P}(X)$ for the set of subsets of X . For a subset A of X :

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

As a shorthand notation we write $\mathbb{1} = \mathbb{1}_X$.

Let E be a vector space. We write $x = (x_1, x_2, \dots)$ for functions $x : \mathbb{N} \rightarrow E$ (i.e., elements of $E^{\mathbb{N}}$) and we define

$$c_{00}[E] = \{x \in E^{\mathbb{N}} : \exists N \ \forall n \geq N [x_n = 0]\}, \quad c_{00} = c_{00}[\mathbb{R}]$$

We write c_0 for the set of sequences in $\mathbb{R}$ that converge to 0, c for the set of convergent sequences in $\mathbb{R}$, $\ell^\infty(X)$ for the set of bounded functions $X \rightarrow \mathbb{R}$, ℓ^∞ for $\ell^\infty(\mathbb{N})$, and ℓ^1 for the set of absolutely summable sequences in $\mathbb{R}$. We write e_n for the element

$1_{\{n\}}$ of $\mathbb{R}^{\mathbb{N}}$.

For a complete σ -finite measure space $(X, \mathcal{A}, \mu)$ we write $\mathcal{L}^1(\mu)$ for the space of integrable functions, $L^1(\mu) = \mathcal{L}^1(\mu)/\mathcal{N}$ where $\mathcal{N}$ denotes the space of functions that are zero μ -a.e. Moreover we write $L^\infty(\mu)$ for the space of equivalence classes of measurable functions that are bounded almost everywhere.

For a subset Γ of a partially ordered vector space Ω , we write $\Gamma^+ = \{f \in \Gamma : f \geq 0\}$. If $\Lambda, \Upsilon \subset \Omega$ and $f \leq g$ for all $f \in \Lambda$ and $g \in \Upsilon$ we write $\Lambda \leq \Upsilon$; if $\Lambda = \{f\}$ we write $f \leq \Upsilon$ instead of $\{f\} \leq \Upsilon$ etc. For a sequence $(h_n)_{n \in \mathbb{N}}$ in a partially ordered vector space we write $h_n \downarrow 0$ if $h_1 \geq h_2 \geq h_3 \geq \dots$ and $\inf_{n \in \mathbb{N}} h_n = 0$.

§5.3 The vertical extension

Throughout this section, E and Ω are partially ordered vector spaces, $\Gamma \subset \Omega$ is a linear subspace and $\varphi : \Gamma \rightarrow E$ is order preserving and linear. Additional assumptions are given in 5.3.14.

5.3.1 Definition. Define

$$\Gamma_v = \left\{ f \in \Omega : \sup_{\sigma \in \Gamma : \sigma \leq f} \varphi(\sigma) = \inf_{\tau \in \Gamma : \tau \geq f} \varphi(\tau) \right\}, \quad (5.1)$$

and $\varphi_v : \Gamma_v \rightarrow E$ by

$$\varphi_v(f) = \sup_{\sigma \in \Gamma : \sigma \leq f} \varphi(\sigma) \quad (f \in \Gamma_v). \quad (5.2)$$

Note: If $f \in \Omega$ and there exist subsets $\Lambda, \Upsilon \subset \Gamma$ with $\Lambda \leq f \leq \Upsilon$ such that $\sup \varphi(\Lambda) = \inf \varphi(\Upsilon)$, then $f \in \Gamma_v$ and $\varphi_v(f) = \inf \varphi(\Upsilon)$.

5.3.2. The following observations are elementary.

- (a) $\Gamma \subset \Gamma_v$ and $\varphi_v(\tau) = \varphi(\tau)$ for all $\tau \in \Gamma$.
- (b) Γ_v is a partially ordered vector space and φ_v is a linear order preserving map¹.
- (c) $(\Gamma_v)_v = \Gamma_v$ and $(\varphi_v)_v = \varphi_v$.
- (d) If Π is a subset of Γ , then $\Pi_v \subset \Gamma_v$.

Of more importance to us than Γ_v and φ_v is the following variation in which we consider only countable subsets of Γ .

5.3.3 Definition. Let Γ_V be the set consisting of those f in Ω for which there exist countable sets $\Lambda, \Upsilon \subset \Gamma$ with $\Lambda \leq f \leq \Upsilon$ such that

$$\sup \varphi(\Lambda) = \inf \varphi(\Upsilon). \quad (5.3)$$

From the remark following Definition 5.3.1 it follows that Γ_V is a subset of Γ_v and that (for f and Λ as above) $\varphi_v(f)$ is equal to $\sup \varphi(\Lambda)$. We will write $\varphi_V = \varphi_v|_{\Gamma_V}$. We call Γ_V the *vertical extension*² under φ of Γ and φ_V the *vertical extension* of φ .

¹This follows from the following fact: Let $A, B \subset E$. If A and B have suprema (infima) in E , then so does $A + B$ and $\sup(A + B) = \sup A + \sup B$ ($\inf(A + B) = \inf A + \inf B$).

²One could also define the vertical extension in case $E, \Omega, \Gamma \subset \Omega$ are partially ordered sets (not necessarily vector spaces) and $\varphi : \Gamma \rightarrow E$ is an order preserving map.

In what follows we will only consider φ_V and not φ_v . However, most of the theory presented can be developed similarly for φ_v . (For comments see 5.11.2.)

5.3.4 Example. Γ_V is the set of Riemann integrable functions on $[0, 1]$ and φ_V is the Riemann integral in case $E = \mathbb{R}$, $\Omega = \mathbb{R}^{[0,1]}$ and Γ is the linear span of $\{\mathbb{1}_I : I \text{ is an interval in } [0, 1]\}$ and φ is the Riemann integral on Γ .

5.3.5. In analogy with 5.3.2 we have the following.

- (a) $\Gamma \subset \Gamma_V$ and $\varphi_V(\tau) = \varphi(\tau)$ for all $\tau \in \Gamma$.
- (b) Γ_V is a partially ordered vector space and φ_V is a linear order preserving map.
- (c) $(\Gamma_V)_V = \Gamma_V$ and $(\varphi_V)_V = \varphi_V$.
- (d) If $\Pi \subset \Gamma$, then $\Pi_V \subset \Gamma_V$.

5.3.6 Definition. Let D be a linear subspace of E . D is called *mediated* in E if the following is true:

If A and B are countable subsets of D such that $\inf A - B = 0$ in E , then A has an infimum (and consequently B has a supremum and $\inf A = \sup B$). (5.4)

D is mediated in E if and only if the following requirement (equivalent with order completeness in the sense of [72], for $D = E$) is satisfied

If A and B are countable subsets of D such that $\inf A - B = 0$ in E , then there exists an $h \in E$ with $B \leq h \leq A$. (5.5)

We say that E is *mediated* if E is mediated in itself.

Note: if D is mediated in E , then so is every linear subspace of D . Every σ -Dedekind complete E is mediated, but so is $\mathbb{R}^2$, ordered lexicographically. Also, c_{00} and c_0 are mediated in c , but c is not mediated.

With this the following lemma is a tautology.

5.3.7 Lemma. Suppose $\varphi(\Gamma)$ is mediated in E . Let $f \in \Omega$. Then $f \in \Gamma_V$ if and only if there exist countable sets $\Lambda, \Upsilon \subset \Gamma$ with $\Lambda \leq f \leq \Upsilon$ such that

$$\inf_{\tau \in \Upsilon, \sigma \in \Lambda} \varphi(\tau - \sigma) = 0. \quad (5.6)$$

The next example shows that Γ_V is not necessarily a Riesz space even if E and Γ are. However, see Corollary 5.3.10.

5.3.8 Example. Consider $E = c$, $\Gamma = c \times c$, $\Omega = \ell^\infty \times \ell^\infty$. Let $\varphi : \Gamma \rightarrow c$ be given by $\varphi(f, g) = f + g$. For all $f \in \ell^\infty$ there are $h_1, h_2, \dots \in c$ with $h_n \downarrow f$. It follows that, $\Gamma_V = \{(f, g) \in \ell^\infty \times \ell^\infty : f + g \in c\}$. Note that Γ_V is not a Riesz space since for every $f \in \ell^\infty$ with $f \geq 0$ and $f \notin c$ we have $(f, -f) \in \Gamma_V$ but $(f, -f)^+ = (f, 0) \notin \Gamma_V$.

5.3.9 Lemma. Suppose $\varphi(\Gamma)$ is mediated in E . Let $\Theta : \Omega \rightarrow \Omega$ be an order preserving map with the properties:

- if $\sigma, \tau \in \Gamma$ and $\sigma \leq \tau$, then $0 \leq \Theta(\tau) - \Theta(\sigma) \leq \tau - \sigma$;

- $\Theta(\Gamma) \subset \Gamma_V$.

Then $\Theta(\Gamma_V) \subset \Gamma_V$.

Proof. Let $f \in \Gamma_V$ and let $\Lambda, \Upsilon \subset \Gamma$ be countable sets with $\Lambda \leq f \leq \Upsilon$ satisfying (5.6). Then $\Theta(\Lambda) \leq \Theta(f) \leq \Theta(\Upsilon)$ and

$$\inf_{\tau \in \Theta(\Upsilon), \sigma \in \Theta(\Lambda)} \varphi(\tau - \sigma) = \inf_{\tau \in \Upsilon, \sigma \in \Lambda} \varphi(\Theta(\tau) - \Theta(\sigma)) \leq \inf_{\tau \in \Upsilon, \sigma \in \Lambda} \varphi(\tau - \sigma) = 0. \quad (5.7)$$

□

5.3.10 Corollary. Suppose that $\varphi(\Gamma)$ is mediated in E . Suppose Ω is a Riesz space and Γ is a Riesz subspace of Ω . Then so is Γ_V .

Proof. Apply Theorem 5.3.9 with $\Theta(\omega) = \omega^+$. □

5.3.11. If Γ is a directed set, i.e., $\Gamma = \Gamma^+ - \Gamma^+$, then so is Γ_V . Indeed, if $f \in \Gamma_V$, then there exist $\sigma, \tau \in \Gamma^+$ such that $f \geq \tau - \sigma$ and thus $f = (f + \sigma) - \sigma \in \Gamma_V^+ - \Gamma_V^+$.

5.3.12. In the last part of this section we will consider a situation in which Ω has some extra structure. But first we briefly consider the case where E is a Banach lattice with σ -order continuous norm. As it turns out, such an E is mediated (see Theorem 5.4.24), but is not necessarily σ -Dedekind complete (consider the Banach lattice $C(X)$ where X is the one-point compactification of an uncountable discrete space). For such E we describe Γ_V in terms of the norm.

5.3.13 Theorem. Let E be a Banach lattice with a σ -order continuous norm. Let Ω be a Riesz space and Γ be a Riesz subspace of Ω . For $f \in \Omega$ we have: $f \in \Gamma_V$ if and only if for every $\varepsilon > 0$ there exist $\sigma, \tau \in \Gamma$ with $\sigma \leq f \leq \tau$ and $\|\varphi(\tau) - \varphi(\sigma)\| < \varepsilon$.

Proof. First, assume $f \in \Gamma_V$. As Γ is a Riesz subspace of Ω there exist sequences $(\sigma_n)_{n \in \mathbb{N}}$ and $(\tau_n)_{n \in \mathbb{N}}$ in Γ such that $\sigma_n \uparrow, \tau_n \downarrow$,

$$\sigma_n \leq f \leq \tau_n \quad (n \in \mathbb{N}), \quad \sup_{n \in \mathbb{N}} \varphi(\sigma_n) = \inf_{n \in \mathbb{N}} \varphi(\tau_n). \quad (5.8)$$

Then $\varphi(\tau_n - \sigma_n) \downarrow 0$ in E , so $\|\varphi(\tau_n) - \varphi(\sigma_n)\| \downarrow 0$ and we are done.

The converse: For each $n \in \mathbb{N}$, choose $\sigma_n, \tau_n \in \Gamma$ for which

$$\sigma_n \leq f \leq \tau_n, \quad \|\varphi(\tau_n) - \varphi(\sigma_n)\| \leq n^{-1}. \quad (5.9)$$

Setting $\sigma'_n = \sigma_1 \vee \cdots \vee \sigma_n$ and $\tau'_n = \tau_1 \wedge \cdots \wedge \tau_n$ we have, for each $n \in \mathbb{N}$

$$\sigma'_n, \tau'_n \in \Gamma, \quad \sigma'_n \leq f \leq \tau'_n. \quad (5.10)$$

If $n \geq N$, then $0 \leq \sigma'_n - \sigma'_N \leq f - \sigma_N \leq \tau_N - \sigma_N$, whence $\|\varphi(\sigma'_n) - \varphi(\sigma'_N)\| \leq \|\varphi(\tau_N) - \varphi(\sigma_N)\| \leq N^{-1}$. Thus, the sequence $(\varphi(\sigma'_n))_{n \in \mathbb{N}}$ converges in the sense of the norm. So does $(\varphi(\tau'_n))_{n \in \mathbb{N}}$. Their limits are the same element a of E , and, since $\sigma'_n \uparrow, \tau'_m \downarrow$, we see that $a = \sup_{n \in \mathbb{N}} \varphi(\sigma'_n) = \inf_{m \in \mathbb{N}} \varphi(\tau'_m)$. □

5.3.14. In the rest of this section Ω is the collection F^X of all maps of a set X into a partially ordered vector space F .

5.3.15. A function $g : X \rightarrow \mathbb{R}$ determines a multiplication operator $f \mapsto gf$ in Ω . We investigate the collection of all functions g for which

$$f \in \Gamma_V \implies gf \in \Gamma_V, \quad (5.11)$$

and, for given f , the behaviour of the map $g \mapsto \varphi_V(gf)$.

5.3.16. For an algebra of subsets of X , $\mathcal{A} \subset \mathcal{P}(X)$ we write $[\mathcal{A}]$ for the Riesz space of all $\mathcal{A}$ -step functions, i.e., functions of the form $\sum_{i=1}^n \lambda_i \mathbb{1}_{A_i}$ for $n \in \mathbb{N}$, $\lambda_i \in \mathbb{R}$, $A_i \in \mathcal{A}$ for $i \in \{1, \dots, n\}$. Define the collection of functions $[\mathcal{A}]^o$ by

$$[\mathcal{A}]^o = \{f \in \mathbb{R}^X : \text{there are } (s_n)_{n \in \mathbb{N}} \text{ in } [\mathcal{A}] \text{ and } (j_n)_{n \in \mathbb{N}} \text{ in } [\mathcal{A}]^+ \text{ for which } |f - s_n| \leq j_n \text{ and } j_n \downarrow 0 \text{ pointwise}\}. \quad (5.12)$$

(This $[\mathcal{A}]^o$ is the vertical extension of $[\mathcal{A}]$ obtained by, in Definition 5.3.3, choosing $E = \mathbb{R}^X$, $\Omega = \mathbb{R}^X$, $\Gamma = [\mathcal{A}]$, $\varphi(f) = f$ ($f \in \Gamma$.) Note that $[\mathcal{A}]$ and $[\mathcal{A}]^o$ are Riesz spaces, and uniform limits of elements of $[\mathcal{A}]$ are in $[\mathcal{A}]^o$. (Actually, $[\mathcal{A}]^o$ is uniformly complete.) Furthermore, $[\mathcal{A}]^o$ contains every bounded function f with $\{x \in X : f(x) \leq s\} \in \mathcal{A}$ for all $s \in \mathbb{R}$. In case $\mathcal{A}$ is a σ -algebra, $[\mathcal{A}]^o$ is precisely the collection of all bounded $\mathcal{A}$ -measurable functions.

5.3.17 Lemma. Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra of subsets of a set X . Suppose that $(g_n)_{n \in \mathbb{N}}$ is a sequence in $[\mathcal{A}]^o$ for which $g_n \downarrow 0$ pointwise. Then there exists a sequence $(j_n)_{n \in \mathbb{N}}$ in $[\mathcal{A}]$ with $j_n \geq g_n$ and $j_n \downarrow 0$ pointwise.

Proof. For all $n \in \mathbb{N}$ there exists a sequence $(s_{nk})_{k \in \mathbb{N}}$ in $[\mathcal{A}]$ with $s_{nk} \geq g_n$ for all $k \in \mathbb{N}$ and $s_{nk} \downarrow_k g_n$ pointwise. Since $(g_n)_{n \in \mathbb{N}}$ is a decreasing sequence, we have $s_{mk} \geq g_n$ for all $m \leq n$ and all $k \in \mathbb{N}$. Hence $j_n := \inf_{m, k \leq n} s_{mk}$ is an element in $[\mathcal{A}]$ with $j_n \geq g_n$. Clearly $j_n \downarrow$ and $\inf_{n \in \mathbb{N}} j_n = \inf_{n \in \mathbb{N}} \inf_{m, k \leq n} s_{mk} = \inf_{n \in \mathbb{N}} \inf_{k \in \mathbb{N}} s_{nk} = \inf_{n \in \mathbb{N}} g_n = 0$. $\square$

The following lemma is a consequence of Lemma 5.3.9.

5.3.18 Lemma. Define the algebra

$$\mathcal{A} = \{A \subset X : f \mathbb{1}_A \in \Gamma \text{ for } f \in \Gamma\}. \quad (5.13)$$

If $\varphi(\Gamma)$ is mediated in E , then

$$f \mathbb{1}_A \in \Gamma_V \quad (f \in \Gamma_V, A \in \mathcal{A}). \quad (5.14)$$

5.3.19 Definition. E is called *Archimedean*³ (see Peressini [73]) if for all $a, b \in E$ the following holds: if $na \leq b$ for all $n \in \mathbb{N}$, then $a \leq 0$.

³In some places, e.g., Birkhoff [8], the term ‘integrally closed’ is used.

5.3.20 Definition. A sequence $(a_n)_{n \in \mathbb{N}}$ in E is called *order convergent* to an element $a \in E$ if there exists a sequence $(h_n)_{n \in \mathbb{N}}$ in E^+ with $h_n \downarrow 0$ and $-h_n \leq a - a_n \leq h_n$.

Notation: $a_n \xrightarrow{o} a$.

5.3.21 Theorem. Let $\mathcal{A}$ be as in (5.13). Suppose that E is Archimedean, Γ is directed and $\varphi(\Gamma)$ is mediated in E . Furthermore assume φ has the following continuity property.

If $A_1, A_2, \dots$ in $\mathcal{A}$ are such that $A_1 \supset A_2 \supset \dots$ and $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$,
then $\varphi(f \mathbb{1}_{A_n}) \downarrow 0$ for all $f \in \Gamma^+$. (5.15)

- (a) $gf \in \Gamma_V$ for all $g \in [\mathcal{A}]^o$ and all $f \in \Gamma_V$.
- (b) Let $g \in [\mathcal{A}]^o$ and let $(g_n)_{n \in \mathbb{N}}$ be a sequence in $[\mathcal{A}]^o$ for which there is a sequence $(j_n)_{n \in \mathbb{N}}$ in $[\mathcal{A}]^{o+}$ with $-j_n \leq g_n - g \leq j_n$ and $j_n \downarrow 0$ pointwise. Then

$$\varphi_V(g_n f) \xrightarrow{o} \varphi_V(gf) \quad (f \in \Gamma_V). \quad (5.16)$$

(Order convergence in the sense of E .)

Proof. We first prove the following:

(★) Let $f \in \Gamma_V^+$. Let $(g_n)_{n \in \mathbb{N}}$ be a sequence in $[\mathcal{A}]^o$ for which $g_n f \in \Gamma_V$ for all $n \in \mathbb{N}$ and $g_n \downarrow 0$ pointwise. Then

$$\varphi_V(g_n f) \downarrow 0. \quad (5.17)$$

Let $\sigma \in \Gamma^+$, $\sigma \geq f$. It follows from Lemma 5.3.17 that we may assume $g_n \in [\mathcal{A}]$ for all $n \in \mathbb{N}$. For all $n \in \mathbb{N}$ we have $0 \leq \varphi_V(g_n f) \leq \varphi_V(g_n \sigma)$, so we are done if $\varphi_V(g_n \sigma) \downarrow 0$. Let $h \in E$, $h \leq \varphi_V(g_n \sigma)$ for all $n \in \mathbb{N}$; we prove $h \leq 0$.

Take $\varepsilon > 0$. For each $n \in \mathbb{N}$, set $A_n = \{x \in X : g_n(x) \geq \varepsilon\}$. Then $A_n \in \mathcal{A}$ for $n \in \mathbb{N}$ and $A_1 \supset A_2 \supset \dots$ and $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$. Putting $M = \|g_1\|_\infty$ we see that

$$g_n \leq \varepsilon \mathbb{1}_X + M \mathbb{1}_{A_n} \quad (n \in \mathbb{N}), \quad (5.18)$$

whence

$$h \leq \varphi_V(g_n \sigma) \leq \varepsilon \varphi(\sigma) + M \varphi(\mathbb{1}_{A_n} \sigma) \quad (n \in \mathbb{N}). \quad (5.19)$$

By the continuity property of φ , $h \leq \varepsilon \varphi(\sigma)$. As this is true for each $\varepsilon > 0$ and E is Archimedean, we obtain $h \leq 0$.

(a) Since Γ_V is directed (see 5.3.11) it is sufficient to consider $f \in \Gamma_V^+$. Let $g \in [\mathcal{A}]^o$. There are sequences of step functions $(h_n)_{n \in \mathbb{N}}$ and $(j_n)_{n \in \mathbb{N}}$ for which $h_n \uparrow g$, $j_n \downarrow 0$ and thus $j_n - h_n \downarrow 0$. By Lemma 5.3.18 $h_n f, j_n f \in \Gamma_V$ for all $n \in \mathbb{N}$. Then $h_n f \leq gf \leq j_n f$ for $n \in \mathbb{N}$ and $\inf_{n \in \mathbb{N}} \varphi_V((j_n - h_n)f) = 0$ by (★). By Lemma 5.3.7 and 5.3.5(c) we obtain that $gf \in \Gamma_V$.

(b) It is sufficient to consider $f \in \Gamma_V^+$. By (a) we may also assume $g = 0$. But then (b) follows from (★). $\square$

5.3.22 Remark. Consider the situation in Theorem 5.3.21. Suppose $\mathcal{B} \subset \mathcal{A}$ is a σ -algebra. Then all bounded $\mathcal{B}$ -measurable functions lie in $[\mathcal{A}]^o$. If $(g_n)_{n \in \mathbb{N}}$ is a bounded sequence of bounded $\mathcal{B}$ -measurable functions that converges pointwise to a function g , then the condition of Theorem 5.3.21(b) is satisfied.

5.3.23 Remark. In the next section we will consider a situation similar to the one of Theorem 5.3.21, in which $\mathcal{A}$ is replaced by a subset $\mathcal{I}$ that is closed under taking finite intersections. We will also adapt the continuity property on φ (see 5.4.3).

§5.4 The lateral extension

The construction described in Definition 5.3.3 is reminiscent of the Riemann integral and, indeed, the Riemann integral is a special case (see Example 5.3.4).

In the present section we consider a type of extension, analogous to the *improper* Riemann integral. One usually defines the improper integral of a function f on $[0, \infty)$ to be

$$\lim_{s \rightarrow \infty} \int_0^s f(x) \, dx, \quad (5.20)$$

approximating the domain, not the values of f .

For our purposes a more convenient description of the same integral would be

$$\sum_{n=1}^{\infty} \int_{a_n}^{a_{n+1}} f(x) \, dx, \quad (5.21)$$

where $0 = a_1 < a_2 < \dots$ and $a_n \rightarrow \infty$. Here the domain is split up into manageable pieces.

Splitting up the domain is the basic idea we develop in this section. (This may explain our use of the terms “vertical” and “lateral”.)

Throughout this section, X is a set, E and F are partially ordered vector spaces, Γ is a directed⁴ linear subspace of F^X , and φ is a linear order preserving map $\Gamma \rightarrow E$. (With $\Omega = F^X$, all considerations of Section 5.3 are applicable.)

$$\begin{array}{c} \Gamma \subset F^X \\ \varphi \downarrow \\ E \end{array}$$

Furthermore, $\mathcal{I}$ is a collection of subsets of X , closed under taking finite intersections. See Definition 5.4.1 and Definition 5.4.2 for two more assumptions.

⁴For the construction of the lateral extension, one does not need to assume that Γ is directed. However, as one can see later on in the construction, the only part of Γ that matters for the extension is $\Gamma^+ - \Gamma^+$.

As a shorthand notation, if $(a_n)_{n \in \mathbb{N}}$ is a sequence in E^+ and $\{\sum_{n=1}^N a_n : N \in \mathbb{N}\}$ has a supremum, we denote this supremum by

$$\sum_n a_n. \quad (5.22)$$

5.4.1 Definition. A disjoint sequence $(A_n)_{n \in \mathbb{N}}$ of elements in $\mathcal{I}$ whose union is X is called a *partition*. If $(A_n)_{n \in \mathbb{N}}$ and $(B_n)_{n \in \mathbb{N}}$ are partitions and for all $n \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ for which $B_n \subset A_m$, then $(B_n)_{n \in \mathbb{N}}$ is called a *refinement* of $(A_n)_{n \in \mathbb{N}}$. Note that if $(A_n)_{n \in \mathbb{N}}$ and $(B_n)_{n \in \mathbb{N}}$ are partitions then there exists a refinement of both $(A_n)_{n \in \mathbb{N}}$ and $(B_n)_{n \in \mathbb{N}}$ (e.g., a partition that consists of all sets of the form $A_n \cap B_m$ with $n, m \in \mathbb{N}$).

We assume that there exists at least one partition.

5.4.2 Definition. We call a linear subspace Δ of F^X *stable* (under $\mathcal{I}$) if

$$f \mathbb{1}_A \in \Delta \quad (f \in \Delta, A \in \mathcal{I}). \quad (5.23)$$

If Δ is a stable space, then a linear and order preserving map $\omega : \Delta \rightarrow E$ is said to be *laterally extendable* if for all partitions $(A_n)_{n \in \mathbb{N}}$

$$\omega(f) = \sum_n \omega(f \mathbb{1}_{A_n}) \quad (\text{see (5.22)}) \quad (f \in \Delta^+). \quad (5.24)$$

We assume Γ is stable and φ is laterally extendable.

5.4.3. In the situation of Theorem 5.3.21 we can choose $\mathcal{I} = \mathcal{A}$; then (5.15) is precisely the lateral extendability of φ .

5.4.4 Example. For any partially ordered vector space F and a linear subspace $E \subset F$, the following choices lead to a system fulfilling all of our assumptions: $X = \mathbb{N}$, $\mathcal{I} = \mathcal{P}(\mathbb{N})$, $\Gamma = c_{00}[E]$ (see Section 5.2), $\varphi(f) = \sum_{n \in \mathbb{N}} f(n)$ for $f \in \Gamma$.

5.4.5 Definition. Let Δ be a stable subspace of F^X and let $\omega : \Delta \rightarrow E$ be a laterally extendable linear order preserving map. Let $(A_n)_{n \in \mathbb{N}}$ be a partition, and $f : X \rightarrow F$. We call $(A_n)_{n \in \mathbb{N}}$ a *partition for f* (occasionally *Δ -partition for f*) if

$$f \mathbb{1}_{A_n} \in \Delta \quad (n \in \mathbb{N}). \quad (5.25)$$

A function $f : X \rightarrow F$ is said to be *partially in Δ* if there exists a partition for f . For $f : X \rightarrow F^+$, $(A_n)_{n \in \mathbb{N}}$ is called a *ω -partition for f* if it is a partition for f and if

$$\sum_n \omega(f \mathbb{1}_{A_n}) \text{ exists.} \quad (5.26)$$

A function $f : X \rightarrow F^+$ that is partially in Δ is called *laterally ω -integrable* if there exists a ω -partition for f .

5.4.6 Example. Consider the situation of Example 5.4.4. A function $x : \mathbb{N} \rightarrow F$ is partially in Γ if and only if $x_n \in E$ for every $n \in \mathbb{N}$. If $x \geq 0$, then x is laterally integrable if $x_n \in E$ for every $n \in \mathbb{N}$ and $\sum_n x_n$ exists in E .

5.4.7. Naturally, we wish to use (5.26) to define an integral for f . For that we have to show the supremum to be independent of the choice of the partition $(A_n)_{n \in \mathbb{N}}$.

5.4.8 Lemma.

- (a) Let $f : X \rightarrow F$ and let $(A_n)_{n \in \mathbb{N}}$ be a partition for f . If $(B_n)_{n \in \mathbb{N}}$ is a partition that is a refinement of $(A_n)_{n \in \mathbb{N}}$, then $(B_n)_{n \in \mathbb{N}}$ is a partition for f .
- (b) Let $f : X \rightarrow F^+$ and let $(A_n)_{n \in \mathbb{N}}$ and $(B_m)_{m \in \mathbb{N}}$ be partitions for f . Then the sets

$$\left\{ \sum_{n=1}^N \varphi(f \mathbb{1}_{A_n}) : N \in \mathbb{N} \right\} \text{ and } \left\{ \sum_{m=1}^M \varphi(f \mathbb{1}_{B_m}) : M \in \mathbb{N} \right\} \quad (5.27)$$

have the same upper bounds in E .

Proof. We leave the proof of (a) to the reader. Let u be an upper bound for the set $\{\sum_{n=1}^N \varphi(f \mathbb{1}_{A_n}) : N \in \mathbb{N}\}$; it suffices to prove that u is an upper bound for $\{\sum_{m=1}^M \varphi(f \mathbb{1}_{B_m}) : M \in \mathbb{N}\}$. Take $M \in \mathbb{N}$; we are done if $u \geq \sum_{m=1}^M \varphi(f \mathbb{1}_{B_m})$, i.e., if $u \geq \varphi(f \mathbb{1}_B)$ where $B = B_1 \cup \dots \cup B_M$. But $f \mathbb{1}_B \in \Gamma$ so $\varphi(f \mathbb{1}_B) = \sum_n \varphi(f \mathbb{1}_B \mathbb{1}_{A_n}) = \sup_{N \in \mathbb{N}} \sum_{n=1}^N \varphi(f \mathbb{1}_B \mathbb{1}_{A_n})$, whereas, for each $N \in \mathbb{N}$

$$\sum_{n=1}^N \varphi(f \mathbb{1}_B \mathbb{1}_{A_n}) \leq \sum_{n=1}^N \varphi(f \mathbb{1}_{A_n}) \leq u. \quad (5.28)$$

□

5.4.9 Theorem. Let $f : X \rightarrow F^+$ be laterally φ -integrable. Then every partition for f is a φ -partition for f . There exists an $a \in E^+$ such that for every partition $(A_n)_{n \in \mathbb{N}}$ for f ,

$$a = \sum_n \varphi(f \mathbb{1}_{A_n}). \quad (5.29)$$

If $f \in \Gamma^+$, then $a = \varphi(f)$.

Proof. This is a consequence of Lemma 5.4.8(b). □

5.4.10 Definition. For a laterally φ -integrable $f : X \rightarrow F^+$ we call the element $a \in E^+$ for which (5.29) holds its φ_L -integral and denote it by $\varphi_L(f)$. For the moment, denote by $(\Gamma^+)_L$ the set of all laterally φ -integrable functions $f : X \rightarrow F^+$. We proceed to extend φ_L to a linear function defined on the linear hull of $(\Gamma^+)_L$, see Definition 5.4.14.

5.4.11. The assumptions that Γ is stable and φ is laterally extendable are crucial for the fact that the φ_L -integral of a laterally φ -integrable function is independent of the choice of a φ -partition (see Lemma 5.4.8(b)).

5.4.12. We will use the following rules for a partially ordered vector space E :

$$a_n \uparrow a, b_n \uparrow b \implies a_n + b_n \uparrow a + b \quad (a_n, b_n, a, b \in E), \quad (5.30)$$

$$a_n \uparrow, b_n \uparrow b, a_n + b_n \uparrow a + b \implies a_n \uparrow a \quad (a_n, b_n, a, b \in E). \quad (5.31)$$

5.4.13. (Extending φ_L) Define $\Gamma_L = \{f_1 - f_2 : f_1, f_2 \in (\Gamma^+)_L\}$.

Step 1. Let $f, g \in (\Gamma^+)_L$. There exists an $(A_n)_{n \in \mathbb{N}}$ that is a φ -partition for f and for g . By defining $a_N = \sum_{n=1}^N \varphi(f \mathbb{1}_{A_n})$ and $b_N = \sum_{n=1}^N \varphi(g \mathbb{1}_{A_n})$ for $N \in \mathbb{N}$, by (5.30) we obtain $f + g \in (\Gamma^+)_L$ with $\varphi_L(f + g) = \varphi_L(f) + \varphi_L(g)$.

Consequently, Γ_L is a vector space, containing $(\Gamma^+)_L$.

Step 2. If $g_1, g_2, h_1, h_2 \in (\Gamma^+)_L$ and $g_1 - g_2 = h_1 - h_2$, then $g_1 + h_2 = g_2 + h_1$ so that, by the above, $\varphi_L(g_1) - \varphi_L(h_1) = \varphi_L(g_2) - \varphi_L(h_2)$.

Hence, φ_L extends to a linear function $\Gamma_L \rightarrow E$ (also denoted by φ_L).

Step 3. Let $f, g \in (\Gamma^+)_L$ and $f \leq g$. By defining a_N and b_N as in step 1 and $c_N = b_N - a_N$, by (5.31) we infer that $g - f \in (\Gamma^+)_L$.

Thus, if $f \in \Gamma_L$ and $f \geq 0$, then $f \in (\Gamma^+)_L$. Briefly: $(\Gamma^+)_L$ is Γ_L^+ , the positive part of Γ_L .

5.4.14 Definition. A function $f : X \rightarrow F$ is called *laterally φ -integrable* if $f \in \Gamma_L$ (see 5.4.13), i.e., if there exist $f_1, f_2 \in (\Gamma^+)_L$ for which $f = f_1 - f_2$. The φ_L -integral of such a function is defined by $\varphi_L(f) = \varphi_L(f_1) - \varphi_L(f_2)$.

φ_L is a function $\Gamma_L \rightarrow E$ and is called the *lateral extension* of φ . The set of laterally φ -integrable functions, Γ_L , is called the *lateral extension* of Γ under φ .

Note that, thanks to Step 3 of 5.4.13, this definition of “laterally φ -integrable” does not conflict with the one given in Definition 5.4.10.

5.4.15. Like for the vertical extension, we have the following elementary observations:

- (a) $\Gamma \subset \Gamma_L$ ⁵ and $\varphi_L(\tau) = \varphi(\tau)$ for all $\tau \in \Gamma$.
 - (b) Γ_L is a directed partially ordered vector space and φ_L is a linear order preserving function on Γ_L .
 - (c) If Π is a directed linear subspace of F^X and $\Pi \subset \Gamma$, then $\Pi_L \subset \Gamma_L$.
- $((\Gamma_L)_L)$ is not so easy. See Theorem 5.4.18 and Example 5.4.19.)

In case E is a Banach lattice with σ -order continuous norm, for Γ_L^+ we have an analogue of Theorem 5.3.13.

5.4.16 Lemma. Suppose E is a Banach lattice with σ -order continuous norm. Let $f : X \rightarrow F^+$. Then f lies in Γ_L^+ if and only if there exists a Γ -partition $(A_n)_{n \in \mathbb{N}}$ for f such that the sequence $(\varphi(f \mathbb{1}_{A_n}))_{n \in \mathbb{N}}$ has a sum in the sense of the norm, in which case $\varphi_L(f)$ is this sum.

Proof. The “only if” part follows by definition of Γ_L and the σ -order continuity of the norm. For the “if” part; this follows from the fact that if $a_n \uparrow$ and $\|a_n - a\| \rightarrow 0$ for $a, a_1, a_2, \dots \in E$, then $a_n \uparrow a$. $\square$

We will now investigate conditions under which φ_L and φ_V themselves are laterally extendable. (For that, their domains have to be able to play the role of Γ , so they have to be stable.) First a useful lemma:

⁵Note that for this inclusion it is necessary that Γ be directed.

5.4.17 Lemma. *Let $f \in \Gamma_L$. Then there exists a partition $(A_n)_{n \in \mathbb{N}}$ for f such that every refinement $(B_m)_{m \in \mathbb{N}}$ of it (is a partition for f and) has this property:*

$$h \in E, h \geq \sum_{m=1}^M \varphi(f \mathbb{1}_{B_m}) \text{ for all } M \in \mathbb{N} \implies h \geq \varphi_L(f). \quad (5.32)$$

Proof. Write $f = f_1 - f_2$ with $f_1, f_2 \in \Gamma_L^+$. Let $(A_n)_{n \in \mathbb{N}}$ be a partition for f_1 and f_2 , and let $(B_m)_{m \in \mathbb{N}}$ be a refinement of $(A_n)_{n \in \mathbb{N}}$. Note that $(B_m)_{m \in \mathbb{N}}$ is a partition for f_1 and f_2 . Let h be an upper bound for $\{\sum_{m=1}^M \varphi(f \mathbb{1}_{B_m}) : M \in \mathbb{N}\}$ in E . For all $M \in \mathbb{N}$,

$$h + \sum_{m=1}^M \varphi(f_2 \mathbb{1}_{B_m}) \geq \sum_{m=1}^M \varphi(f \mathbb{1}_{B_m}) + \sum_{m=1}^M \varphi(f_2 \mathbb{1}_{B_m}) = \sum_{m=1}^M \varphi(f_1 \mathbb{1}_{B_m}). \quad (5.33)$$

Taking the supremum over M yields $h + \varphi_L(f_2) \geq \varphi_L(f_1)$, i.e., $h \geq \varphi_L(f)$. $\square$

5.4.18 Theorem.

(a) *Suppose Γ_L is stable. Then φ_L is laterally extendable, i.e.,*

$$\varphi_L(f) = \sum_n \varphi_L(f \mathbb{1}_{A_n}) \quad (5.34)$$

for every $f \in \Gamma_L^+$ and every φ_L -partition $(A_n)_{n \in \mathbb{N}}$ for f . Therefore $(\Gamma_L)_L = \Gamma_L$ and $(\varphi_L)_L = \varphi_L$.

(b) *Suppose Γ_V is stable. Then φ_V is laterally extendable. (For $(\Gamma_V)_L$ see Section 5.5.)*

Proof. (a) Let $f \in \Gamma_L^+$ and let $(B_n)_{n \in \mathbb{N}}$ be a φ_L -partition for f . Let $(A_n)_{n \in \mathbb{N}}$ be the partition for f as in Lemma 5.4.17. Then form a common refinement of $(B_n)_{n \in \mathbb{N}}$ and $(A_n)_{n \in \mathbb{N}}$ and apply Lemma 5.4.17.

(b) Let $f \in \Gamma_V^+$ and let $(A_n)_{n \in \mathbb{N}}$ be a partition. Let $h \in E, h \geq \sum_{n=1}^N \varphi_V(f \mathbb{1}_{A_n})$ for every $N \in \mathbb{N}$. We wish to prove $h \geq \varphi_V(f)$, which will be the case if $h \geq \varphi(\sigma)$ for every $\sigma \in \Gamma$ with $\sigma \leq f$. For that apply Lemma 5.4.17 to σ . $\square$

The following shows that Γ_L may not be stable, in which case there is no $(\Gamma_L)_L$. (However, see Theorem 5.4.25(a).)

5.4.19 Example. Consider the situation in Example 5.4.4 and assume there is an $a : \mathbb{N} \rightarrow E^+$ such that $\sum_n a_n$ exists in F and $\sum_n a_{2n}$ does not (e.g. $E = F = c$ and $a_n = e_n = \mathbb{1}_{\{n\}}$). By Example 5.4.6 a lies in Γ_L but $b = (0, a_2, 0, a_4, \dots)$ does not; but $b = a \mathbb{1}_{\{2, 4, 6, \dots\}}$ and $\{2, 4, 6, \dots\} \in \mathcal{I}$. (Actually, the existence of such an $a : \mathbb{N} \rightarrow E^+$ is equivalent to E not being “splitting” in F ; see Definition 5.4.21 and (5.36).)

5.4.20 Remark. Γ_V may not be stable either. With $E = c, F = \ell^\infty, X = \{1, 2\}$, $\Gamma = c \times c$ and $\varphi(f, g) = f + g$ (as in Example 5.3.8), the space Γ_V is not stable for $\mathcal{I} = \mathcal{P}(X)$.

5.4.21 Definition. Let D be a linear subspace of E . D is called *splitting* in E if the following is true:

$$\begin{aligned} &\text{If } (a_n)_{n \in \mathbb{N}} \text{ and } (b_n)_{n \in \mathbb{N}} \text{ are sequences in } D \text{ with } 0 \leq a_n \leq b_n \text{ for } n \in \mathbb{N} \\ &\text{and } \sum_n b_n \text{ exists in } E, \text{ then so does } \sum_n a_n. \end{aligned} \quad (5.35)$$

It is not difficult to see that D is splitting in E if and only if

$$\begin{aligned} &\text{If } (a_n)_{n \in \mathbb{N}} \text{ is a sequence in } D^+ \text{ and } \sum_n a_n \text{ exists in } E, \\ &\text{then so does } \sum_n \mathbb{1}_A(n) a_n \text{ for all } A \subset \mathbb{N}. \end{aligned} \quad (5.36)$$

If D is splitting in E , then so is every linear subspace of D . If E is σ -Dedekind complete, then E is also splitting. More generally, D is splitting in E if every bounded increasing sequence in D has a supremum in E . Also, $\mathbb{R}^2$ with the lexicographical ordering is splitting.

In Theorem 5.4.25 we will see what is the use of this concept. First, we have a look at the connection between “splitting” and “mediated”.

5.4.22 Lemma. Suppose D is a linear subspace of E . Consider the condition:

$$\begin{aligned} &\text{For all sequences } (a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \text{ in } D: \\ &a_n \downarrow, b_n \uparrow, \inf_{n \in \mathbb{N}} a_n - b_n = 0 \implies \inf_{n \in \mathbb{N}} a_n = \sup_{n \in \mathbb{N}} b_n. \end{aligned} \quad (5.37)$$

(The infima and suprema in (5.37) are to be taken in E .) If D is either splitting or mediated in E , then (5.37) holds. Conversely, (5.37) implies that D is splitting if $D = E$, whereas (5.37) implies that D is mediated in E if E is a Riesz space and D is a Riesz subspace of E .

Proof. It will be clear that mediatedness implies (5.37) and vice versa if E is a Riesz space and D a Riesz subspace of E .

If D is splitting in E and $a_n \downarrow, b_n \uparrow$ and $\inf a_n - b_n = 0$, then $\sum_n b_{n+1} - b_n + a_n - a_{n+1} = a_1 - b_1$. Hence (5.37) holds.

Suppose $D = E$ and (5.37) holds. Let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be sequences in D with $0 \leq a_n \leq b_n$ for $n \in \mathbb{N}$ such that $\sum_n b_n$ exists. Let $z = \sum_n b_n$, $A_n = \sum_{i=1}^n a_i$, $C_n = \sum_{i=1}^n b_i - a_i$ for $n \in \mathbb{N}$. Then $A_n \uparrow, C_n \uparrow$ and $z - C_n - A_n \downarrow 0$ (note that $z - C_n \in D$). Hence $\sup_{n \in \mathbb{N}} A_n = \sum_n a_n$ exists. $\square$

5.4.23. (a) If E is a Riesz space, then every splitting Riesz subspace is mediated in E .

(b) If E is mediated, then it is splitting. The converse is also true if E is a Riesz space.

(c) c_{00} is mediated in c , not splitting in c (with $D = E = c$ also (5.37) is not satisfied).

(d) If D is the space of all polynomial functions on $[0, 1]$ with degree at most 2 and $E = C[0, 1]$, then D is splitting in E , but not mediated in E . (Actually, D is splitting, but not mediated.)

D is splitting (and satisfies (5.37) with $E = D$): If $u_n \in E^+$, $u_n \uparrow$ and $u_n \leq \mathbb{1}$, then $|u_n(x) - u_n(y)| \leq 4|x - y|$ as can be concluded from the postscript in Example 5.5.15. Therefore the pointwise supremum is continuous. It is even in D since $u_n(x) = a_n x^2 + b_n x + c_n$, where a_n, b_n, c_n are linear combinations of $u_n(0), u_n(\frac{1}{2}), u_n(1)$ (see also the postscript in Example 5.5.15).

D is not mediated: For example one can find countable $A, B \subset E$ for which $\mathbb{1}_{[\frac{1}{2}, 1]}$ is pointwise the infimum of A and $\mathbb{1}_{(\frac{1}{2}, 1]}$ is pointwise the supremum of B , then $\inf A - B = 0$, but there is no $h \in E$ with $B \leq h \leq A$.

5.4.24 Theorem. *Let E be a Banach lattice with σ -order continuous norm. Then E is both mediated and splitting.*

Proof. Suppose $a_n, b_n \in E$ with $0 \leq a_n \leq b_n$ for $n \in \mathbb{N}$. Suppose that $\{\sum_{n=1}^N b_n : N \in \mathbb{N}\}$ has a supremum s in E . We prove that $\{\sum_{n=1}^N a_n : N \in \mathbb{N}\}$ has a supremum in E . Since the norm is σ -order continuous, we have $\|s - \sum_{n=1}^N b_n\| \rightarrow 0$. In particular we get that for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n, m \geq N$ with $m > n$ we have $\|\sum_{i=n}^m b_i\| < \varepsilon$ and thus $\|\sum_{i=n}^m a_i\| < \varepsilon$. From this we infer that $(\sum_{n=1}^N a_n)_{N \in \mathbb{N}}$ converges in norm. Therefore it has a supremum in E . Thus E is splitting. By Lemma 5.4.22 E is mediated. $\square$

5.4.25 Theorem.

- (a) $\varphi(\Gamma)$ splitting in $E \implies \Gamma_L$ is stable and φ_L is laterally extendable.
- (b) $\varphi(\Gamma)$ mediated in $E \implies \Gamma_V$ is stable and φ_V is laterally extendable.
- (c) $\varphi(\Gamma)$ splitting in E and $\varphi_L(\Gamma_L)$ mediated in $E \implies (\Gamma_L)_V$ is stable and $(\varphi_L)_V$ is laterally extendable.

Proof. (a) Let $f \in \Gamma_L$, $B \in \mathcal{I}$; we prove $f \mathbb{1}_B \in \Gamma_L$. (This is sufficient by Theorem 5.4.18(a).) Without loss of generality, assume $f \geq 0$. Choose a φ -partition $(A_n)_{n \in \mathbb{N}}$ for f . Now apply (5.35) to

$$a_n := \varphi(f \mathbb{1}_{A_n \cap B}), \quad b_n := \varphi(f \mathbb{1}_{A_n}) \quad (n \in \mathbb{N}). \quad (5.38)$$

(b) follows from Lemma 5.3.18 and Theorem 5.4.18(b).

(c) By (a) Γ_L is stable and φ_L is laterally extendable. Hence we can apply (b) to Γ_L and φ_L (instead of Γ and φ) and obtain (c). $\square$

5.4.26. To some extent, the assumption of Theorem 5.4.25(a) is minimal. Indeed, in the situation of Example 5.4.4, we see that Γ_L is stable if and only if E (which is $\varphi(\Gamma)$) is splitting in F (see (5.36)).

In Theorem 5.4.25(c) we assumed that $\varphi_L(\Gamma_L)$ (and thus also $\varphi(\Gamma)$) was mediated in E . It may happen that $\varphi(\Gamma)$ is mediated in E , but $\varphi_L(\Gamma_L)$ is not, as Example 5.4.27 illustrates. However, splitting is preserved under the lateral extension and mediation is preserved under the vertical extension, see Theorem 5.4.28.

5.4.27 Example. Let $X = \mathbb{N}$, $\mathcal{I} = \mathcal{P}(\mathbb{N})$, $E = F = c$. Let $\Gamma = c_{00}[c_{00}]$ (see Section 5.2) and $\varphi : \Gamma \rightarrow E$ be given by $\varphi(f) = \sum_{n \in \mathbb{N}} f(n)$. Then $\varphi(\Gamma) = c_{00}$, which is

mediated in c . A function $f : \mathbb{N} \rightarrow c$ is partially in Γ if and only if $f(\mathbb{N}) \subset c_{00}$. For $x \in c^+$ the function given by $f(n) = x(n)\mathbb{1}_{\{n\}}$ for $n \in \mathbb{N}$ lies in Γ_L , and $\varphi_L(f) = x$. It follows that $\varphi_L(\Gamma_L)$ is c , which is not mediated in c .

5.4.28 Theorem.

- (a) If $\varphi(\Gamma)$ is splitting in E , then so is $\varphi_L(\Gamma_L)$.
- (b) If $\varphi(\Gamma)$ is mediated in E , then so is $\varphi_V(\Gamma_V)$.

Proof. (a) Suppose $a_n \in \varphi_L(\Gamma_L)^+$ for $n \in \mathbb{N}$ and $\sum_n a_n$ exists. Let $A \subset \mathbb{N}$. For all $n \in \mathbb{N}$ there exist $b_{n1}, b_{n2}, \dots \in \varphi(\Gamma)^+$ with $a_n = \sum_m b_{nm}$. Hence $\sum_n a_n = \sum_{n,m} b_{nm}$ and so $\sum_{n,m} \mathbb{1}_{A \times \mathbb{N}}(n, m)b_{nm} = \sum_n \mathbb{1}_A(n)a_n$ exists in E .

(b) Suppose $A, B \subset \varphi_V(\Gamma_V)$ are countable sets with $\inf A - B = 0$. For all $a \in A$ and $b \in B$ there exist countable sets $\Upsilon_a, \Lambda_b \subset \Gamma$ with $a = \inf \varphi(\Upsilon_a), b = \sup \varphi(\Lambda_b)$. Then $\inf \varphi(\bigcup_{a \in A} \Upsilon_a - \bigcup_{b \in B} \Lambda_b) = 0$ and thus $\inf A = \inf \varphi(\bigcup_{a \in A} \Upsilon_a) = \sup \varphi(\bigcup_{b \in B} \Lambda_b) = \sup B$. $\square$

5.4.29. For a Riesz space F we will now investigate under which conditions the space Γ_L is a Riesz subspace of F^X . The next example shows that even if E is a Riesz space and Γ is a Riesz subspace of F^X , Γ_L may not be one. However, see Theorem 5.4.32.

5.4.30 Example. Let a, b be as in Example 5.4.19; this time put $d = (0, a_1 + a_2, 0, a_3 + a_4, \dots)$. Then $a, d \in \Gamma_L$ but $a \wedge d = b \notin \Gamma_L$.

Hence, in Example 5.4.4, if F is a Riesz space and E is not splitting in F , then Γ_L is not a Riesz subspace of F^X . As we will see in Theorem 5.4.32, considering the situation of Example 5.4.4: Γ_L is a Riesz subspace of F^X if and only if E is splitting in F .

5.4.31 Lemma. Let $f : X \rightarrow F$ be partially in Γ .

- (a) If f is in Γ_{LV} , then $f \in \Gamma_L$.
- (b) Suppose $\varphi(\Gamma)$ is splitting in E . If $g \leq f \leq h$ for certain $g, h \in \Gamma_L$, then $f \in \Gamma_L$.

Proof. (a) By the definition of Γ_{LV} there exists a $\rho \in \Gamma_L$ with $\rho \leq f$. Then $f - \rho$ is partially in Γ , $f - \rho \in \Gamma_{LV}$, and we are done if $f - \rho \in \Gamma_L$. Hence we may assume $f \geq 0$.

Let $(A_n)_{n \in \mathbb{N}}$ be a partition for f ; we prove $\sum_n \varphi(f\mathbb{1}_{A_n}) = \varphi_{LV}(f)$. It will be clear that $\sum_{n=1}^N \varphi(f\mathbb{1}_{A_n}) \leq \varphi_{LV}(f)$ for $N \in \mathbb{N}$. For the reverse inequality let $h \in E$ be an upper bound for $\{\sum_{n=1}^N \varphi(f\mathbb{1}_{A_n}) : N \in \mathbb{N}\}$. It suffices to show that h must be an upper bound for $\{\varphi_L(\sigma) : \sigma \in \Gamma_L, \sigma \leq f\}$.

Take a $\sigma \in \Gamma_L$ with $\sigma \leq f$. If $(B_n)_{n \in \mathbb{N}}$ is any refinement of $(A_n)_{n \in \mathbb{N}}$ that is a φ -partition for σ , then for all $M \in \mathbb{N}$ there exists an $N \in \mathbb{N}$ with $B_1 \cup \dots \cup B_M \subset A_1 \cup \dots \cup A_N$, so that $h \geq \sum_{n=1}^N \varphi(f\mathbb{1}_{A_n}) \geq \sum_{m=1}^M \varphi(f\mathbb{1}_{B_m}) \geq \sum_{m=1}^M \varphi(\sigma\mathbb{1}_{B_m})$. It follows from Lemma 5.4.17, applied to σ , that the partition $(B_m)_{m \in \mathbb{N}}$ can be chosen so that this implies $h \geq \varphi_L(\sigma)$.

(b) As $h - g \in \Gamma_L$ and $0 \leq f - g \leq h - g$, we may (and do) assume $g = 0$. Let $(A_n)_{n \in \mathbb{N}}$ be a partition for f that is also a φ -partition for h . Now just apply (5.35) to

$$a_n := \varphi(f\mathbb{1}_{A_n}), \quad b_n := \varphi(h\mathbb{1}_{A_n}) \quad (n \in \mathbb{N}). \quad (5.39)$$

$\square$

As a consequence of Lemma 5.4.31:

5.4.32 Theorem. *Let F be a Riesz space and Γ be a Riesz subspace of F^X . The functions $X \rightarrow F$ that are partially in Γ form a Riesz space, Ξ . If $\varphi(\Gamma)$ is splitting in E , then Γ_L is a Riesz ideal in Ξ , in particular, Γ_L is a Riesz space.*

In the classical integration theory and the Bochner integration theory one starts with considering a measure space $(X, \mathcal{A}, \mu)$ and simple functions on X with values in $\mathbb{R}$ or in a Banach space. One defines an integral on these simple functions using the measure and extends this integral to a larger class of integrable functions. In 5.4.33 we will follow a similar procedure, replacing $\mathbb{R}$ or the Banach space with E and applying the lateral extension. In Section 5.8 we will treat such extensions in more detail.

5.4.33. Suppose $(X, \mathcal{A}, \mu)$ is a σ -finite complete measure space and suppose E is directed. Let $F = E$. For $\mathcal{I}$ we choose $\{A \in \mathcal{A} : \mu(A) < \infty\}$. The σ -finiteness of μ guarantees the existence of a partition (and vice versa).

We say that a function $f : X \rightarrow E$ is *simple* if there exist $N \in \mathbb{N}$, $a_1, \dots, a_N \in E$, $A_1, \dots, A_N \in \mathcal{I}$ for which

$$f = \sum_{n=1}^N a_n \mathbb{1}_{A_n}. \quad (5.40)$$

The simple functions form a stable directed linear subspace S of E^X , which is a Riesz subspace of E^X in case E is a Riesz space.

For a given f in S one can choose a representation (5.40) in which the sets $A_1, \dots, A_N$ are pairwise disjoint; thanks to the σ -finiteness of μ one can choose them in such a way that they occur in a partition $(A_n)_{n \in \mathbb{N}}$.

This S is going to be our Γ . We define $\varphi : S \rightarrow E$ by

$$\varphi(f) = \sum_{n=1}^N \mu(A_n) a_n, \quad (5.41)$$

where f, N, A_n, a_n are as in (5.40). The σ -additivity of μ is (necessary and) sufficient to show that S is laterally extendable.

A function $f : X \rightarrow E$ is partially in S if and only if there exist a partition $(A_n)_{n \in \mathbb{N}}$ and a sequence $(a_n)_{n \in \mathbb{N}}$ in E for which

$$f = \sum_{n \in \mathbb{N}} a_n \mathbb{1}_{A_n}. \quad (5.42)$$

An f as in (5.42) with $f \geq 0$ that is partially in S is an element of S_L if and only if $\sum_n \mu(A_n) a_n$ exists in E . (See Theorem 5.4.9.)

§5.5 Combining vertical and lateral extensions

In this section $E, F, X, \mathcal{I}, \Gamma, \varphi$ are as in Section 5.4.

As we have seen, the lateral extension differs from the vertical extension in the sense that the vertical extensions of Γ and φ can always be made, but for lateral extension we had to assume the space Γ to be stable and φ to be laterally extendable (see 5.4.11). In this section we investigate when one can make a lateral extension of another (say vertical) extension. Furthermore we will compare different extensions and combinations of extensions.

Instead of $(\Gamma_L)_V$ and $((\Gamma_L)_V)_L$ we write Γ_{LV} and $\Gamma_{LV L}$; similarly $\varphi_{LV} = (\varphi_L)_V$ etc.

5.5.1. By Theorem 5.4.18 the following holds for a stable directed linear subspace Δ of F^X and a laterally extendable order preserving linear map $\omega : \Delta \rightarrow E$: If Δ_L is stable, then ω_L is laterally extendable (and so Δ_{LL} exists). If Δ_V is stable, then ω_V is laterally extendable (and so Δ_{VL} exists). We will use these facts without explicit mention.

5.5.2. The following statements follow from the definitions and theorems we have:

- (a) $\Gamma_V \subset \Gamma_{LV}$ and $\varphi_{LV} = \varphi_V$ on Γ_V .
- (b) $\Gamma_L \subset \Gamma_{LV}$ and $\varphi_{LV} = \varphi_L$ on Γ_L .
- (c) $\varphi_V = \varphi_L$ on $\Gamma_L \cap \Gamma_V$.

For (d), (e) and (f) let Γ_V be stable.

- (d) $\Gamma_{LV} \subset \Gamma_{VL V}$ and $\varphi_{VL V} = \varphi_{LV}$ on Γ_{LV} .
- (e) $\Gamma_{VL} \subset \Gamma_{VL V}$ and $\varphi_{VL V} = \varphi_{VL}$ on Γ_{VL} .
- (f) $\varphi_{LV} = \varphi_{VL}$ on $\Gamma_{LV} \cap \Gamma_{VL}$.

Observe that as a consequence of (a) and (b): If $f \in \Gamma_L$ and $g \in \Gamma_V$ and $f \leq g$ (or $f \geq g$), then $\varphi_L(f) \leq \varphi_V(g)$ (or $\varphi_L(f) \geq \varphi_V(g)$). Moreover, as a consequence of (c) and (d); if Γ_V is stable: If $f \in \Gamma_{LV}$ and $g \in \Gamma_{VL V}$ and $f \leq g$ (or $f \geq g$), then $\varphi_{LV}(f) \leq \varphi_{VL V}(g)$ (or $\varphi_{LV}(f) \geq \varphi_{VL V}(g)$).

5.5.3. Note that if Γ is stable and φ is laterally extendable, then we can extend Γ to Γ_V, Γ_L and Γ_{LV} . If, moreover, Γ_V is stable, then we can also extend Γ to Γ_{VL} and $\Gamma_{VL V}$. However, “more stability” will not give us larger extensions than $\Gamma_{VL V}$. Indeed, if Γ_{LV} is stable then $\Gamma_{LV} \subset \Gamma_{LV L} = \Gamma_{VL}$ (see Theorem 5.5.8). If moreover $\Gamma_{VL V}$ is stable, then even $\Gamma_{VL V L} = \Gamma_{VL V} = \Gamma_{VL}$.

5.5.4 Lemma.

- (a) If $f \in \Gamma_{LV}^+$, then there exists a countable $\Lambda \subset \Gamma$ with $\Lambda \leq f$ and $\varphi_{LV}(f) = \sup \varphi(\Lambda)$.
- (b) If Γ_V is stable and $f \in \Gamma_{VL}^+$, then there exists a countable $\Lambda \subset \Gamma$ with $\Lambda \leq f$ and $\varphi_{VL}(f) = \sup \varphi(\Lambda)$.

Proof. (a) There exist $\sigma_1, \sigma_2, \dots$ in Γ_L with $\sigma_n \leq f$ for all $n \in \mathbb{N}$ and $\sup_{n \in \mathbb{N}} \varphi_L(\sigma_n) = \varphi_{LV}(f)$. Hence, we are done if for every σ in Γ_L with $\sigma \leq f$ there is a countable set $\Lambda_\sigma \subset \{\rho \in \Gamma : \rho \leq f\}$ such that every upper bound for $\varphi(\Lambda_\sigma)$ majorizes $\varphi_L(\sigma)$. But that is not hard to prove. For such a σ , by Lemma 5.4.17 there exists a partition

$(B_m)_{m \in \mathbb{N}}$ for which (5.32) holds. Now let Λ_σ be $\{\sum_{m=1}^M \sigma \mathbb{1}_{B_m} : M \in \mathbb{N}\}$.
 (b) Suppose Γ_V is stable. Let $(A_n)_{n \in \mathbb{N}}$ be a φ_V -partition for f . Then the set $\Lambda_f = \{\sum_{n=1}^N f \mathbb{1}_{A_n} : N \in \mathbb{N}\}$ is a countable subset of Γ_V and $\sup \varphi_V(\Lambda_f) = \varphi_{VL}(f)$. Moreover, for every $N \in \mathbb{N}$ there is a countable set $\Lambda_N \subset \{\sigma \in \Gamma : \sigma \leq \sum_{n=1}^N f \mathbb{1}_{A_n}\}$ for which $\sup \varphi(\Lambda_N) = \varphi_V(\sum_{n=1}^N f \mathbb{1}_{A_n})$. Take $\Lambda = \bigcup_{N \in \mathbb{N}} \Lambda_N$. $\square$

5.5.5 Theorem. For (b),(c),(d) and (e) let Γ_V be stable and f be partially in Γ_V .

(a) If $f \in \Gamma_{LV}$, then

$$f \in \Gamma_V \iff \text{there exist } \pi, \rho \in \Gamma \text{ with } \pi \leq f \leq \rho^6.$$

(b) If $f \in \Gamma_{VL}$, then

$$f \in \Gamma_V \iff \text{there exist } \pi, \rho \in \Gamma \text{ with } \pi \leq f \leq \rho^6.$$

(c) $f \in \Gamma_{LV} \iff f \in \Gamma_{VL}$ and there exist $\pi, \rho \in \Gamma_L$ with $\pi \leq f \leq \rho$.

(d) If $\varphi_V(\Gamma_V)$ is splitting in E , then

$$f \in \Gamma_{VL} \iff \text{there exist } \pi, \rho \in \Gamma_{VL} \text{ with } \pi \leq f \leq \rho.$$

(e) If $\varphi_V(\Gamma_V)$ is splitting in E , then

$$f \in \Gamma_{VL} \cap \Gamma_{LV} \iff \text{there exist } \pi, \rho \in \Gamma_L \text{ with } \pi \leq f \leq \rho.$$

Proof. The proofs of (a) and (b) are similar to the proof of (c) and therefore omitted.

(c) $\Leftarrow$: By Lemma 5.5.4 (b) there exist countable sets $\Lambda, \Upsilon \subset \Gamma$ with $\Lambda \leq f - \pi$ and $\Upsilon \leq \rho - f$ for which $\sup \varphi(\Lambda) = \varphi_{VL}(f - \pi)$ and $\sup \varphi(\Upsilon) = \varphi_{VL}(\rho - f)$. Then $\Lambda + \pi$ and $\rho - \Upsilon$ are countable subsets of Γ_L with $\Lambda + \pi \leq f \leq \rho - \Upsilon$ and $\sup \varphi_L(\Lambda + \pi) = \varphi_{VL}(f) = \inf \varphi_L(\rho - \Upsilon)$. Hence $f \in \Gamma_{LV}$.

$\Rightarrow$: Let $f \in \Gamma_{LV}$ and be partially in Γ_V . There exists a $\pi \in \Gamma_L$ for which $f - \pi \in \Gamma_{LV}^+$, hence we may assume $f \geq 0$. Let $(A_n)_{n \in \mathbb{N}}$ be a Γ_V -partition for f , i.e., $f \mathbb{1}_{A_n} \in \Gamma_V$ and thus $\varphi_{LV}(f \mathbb{1}_{A_n}) = \varphi_V(f \mathbb{1}_{A_n})$ for all $n \in \mathbb{N}$ (see 5.5.2(a)). Then $\varphi_{LV}(f) \geq \sum_{n=1}^N \varphi_V(f \mathbb{1}_{A_n})$ for all $N \in \mathbb{N}$. Let $h \in E$ be such that $h \geq \sum_{n=1}^N \varphi_V(f \mathbb{1}_{A_n})$ for all $N \in \mathbb{N}$. From Lemma 5.4.17 we infer that $h \geq \varphi_L(\sigma)$ for every $\sigma \in \Gamma_L$ with $\sigma \leq f$. We conclude that $\sum_n \varphi_V(f \mathbb{1}_{A_n}) = \varphi_{LV}(f)$, i.e., $f \in \Gamma_{VL}$.

(d) $\Leftarrow$: We may assume $\pi = 0$. Let $(A_n)_{n \in \mathbb{N}}$ be a φ_V -partition for ρ with $f \mathbb{1}_{A_n} \in \Gamma_V$ for all $n \in \mathbb{N}$. Then $0 \leq \varphi_V(f \mathbb{1}_{A_n}) \leq \varphi_V(\rho \mathbb{1}_{A_n})$ for all $n \in \mathbb{N}$ and $\sum_n \varphi_V(\rho \mathbb{1}_{A_n})$ exists in E . Hence, so does $\sum_n \varphi_V(f \mathbb{1}_{A_n})$, i.e., $f \in \Gamma_{VL}$.

(e) is a consequence of (c) and (d). $\square$

In the following example all functions in Γ_{LV} are partially in Γ_V .

5.5.6 Example. Consider $X = \mathbb{N}, \mathcal{I} = \mathcal{P}(\mathbb{N}), E = F$; let D be a linear subspace of E and let D_V be the vertical extension of D with respect to the inclusion map $D \rightarrow E$. Let $\Gamma = c_{00}[D]$ and $\varphi : \Gamma \rightarrow E$ be $\varphi(f) = \sum_{n \in \mathbb{N}} f(n)$. Then $\Gamma_V = c_{00}[D_V]$.

⁶By the definition of ideal in [11] or [40] (note that Γ_V is directed) this means that Γ_V is the smallest ideal in Γ_{LV} (and for (b); in Γ_{VL}) that contains Γ .

Let $f \in \Gamma_{LV}$. We will show that $f(k) \in D_V$ and thus that f is partially in Γ_V . Let $\sigma_n, \tau_n \in \Gamma_L$ be such that $\sigma_n \leq f \leq \tau_n$ and $\inf_{n \in \mathbb{N}} \varphi(\tau_n) = \sup_{n \in \mathbb{N}} \varphi(\sigma_n)$. Then $\inf_{n \in \mathbb{N}} (\tau_n(k) - \sigma_n(k)) \leq \inf_{n \in \mathbb{N}} \varphi(\tau_n - \sigma_n) = 0$. Since $\sigma_n(k), \tau_n(k) \in D$ for all $n \in \mathbb{N}$, we have $f(k) \in D_V$.

Thus every $f \in \Gamma_{LV}$ is partially in Γ_V . Since Γ_V is stable, by Theorem 5.5.5(c) we conclude that $\Gamma_{LV} \subset \Gamma_{VL}$.

5.5.7 Lemma. *Suppose that Γ_{LV} is stable. Then every $f \in \Gamma_{LV}$ is partially in Γ_V .*

Proof. Let $f \in \Gamma_{LV}$ and let $\pi, \rho \in \Gamma_L$ be such that $\pi \leq f \leq \rho$. Let $(A_n)_{n \in \mathbb{N}}$ be a φ -partition for both π and ρ . Then $f \mathbb{1}_{A_n} \in \Gamma_{LV}$ and $\pi \mathbb{1}_{A_n} \leq f \mathbb{1}_{A_n} \leq \rho \mathbb{1}_{A_n}$ for all $n \in \mathbb{N}$. By Theorem 5.5.5(a) we conclude that $f \mathbb{1}_{A_n} \in \Gamma_V$. $\square$

5.5.8 Theorem. *Suppose that Γ_V and Γ_{LV} are stable. Then $\Gamma_{LV} \subset \Gamma_{VL} = \Gamma_{LV L}$. Write $\bar{\Gamma} = \Gamma_{VL}$ and $\bar{\varphi} = \varphi_{VL}$. If $\bar{\Gamma}$ is stable, then $\bar{\Gamma}_L = \bar{\Gamma}$ and $\bar{\varphi}_L = \bar{\varphi}$. If $\bar{\Gamma}_V$ is stable, then $\bar{\Gamma}_V = \bar{\Gamma}$ and $\bar{\varphi}_V = \bar{\varphi}$. In particular, if $\varphi_L(\Gamma_L)$ is mediated in E and $\varphi_V(\Gamma_V)$ is splitting in E , then Γ_V, Γ_{LV} and Γ_{VL} are stable (see Theorem 5.4.25) and thus $\Gamma_{LV} \subset \bar{\Gamma}$, $\bar{\Gamma} = \bar{\Gamma}_V = \bar{\Gamma}_L$, $\bar{\varphi} = \bar{\varphi}_V = \bar{\varphi}_L$, so $\bar{\Gamma} = \bar{\Gamma}$ (and $\bar{\varphi} = \bar{\varphi}$).*

Proof. The inclusion $\Gamma_{LV} \subset \Gamma_{VL}$ follows by Theorem 5.5.5(c) and Lemma 5.5.7. We prove $\Gamma_{LV L} \subset \Gamma_{VL}$. For $f \in \Gamma_{LV L}^+$ there is a φ_{LV} -partition for f and since $\Gamma_{LV} \subset \Gamma_{VL}$ this is also a φ_{VL} -partition for f , hence there exists a φ_V -partition for f , i.e., $f \in \Gamma_{VL}$.

Suppose $\bar{\Gamma}$ is stable. Then $\bar{\Gamma}_L = (\Gamma_{VL})_L = \Gamma_{VL} = \bar{\Gamma}$ and $\bar{\varphi}_L = \bar{\varphi}$ by Theorem 5.4.18(a).

Suppose $\bar{\Gamma}_V$ to be stable. As Γ_V is stable we can apply the first part of the theorem to Γ_V instead of Γ . Indeed, $(\Gamma_V)_V$ and $(\Gamma_V)_{LV}$ are stable, since $(\Gamma_V)_V = \Gamma_V$ and $(\Gamma_V)_{LV} = \bar{\Gamma}_V$. Hence, $(\Gamma_V)_{LV} \subset (\Gamma_V)_{VL} = \Gamma_{VL}$, i.e., $\bar{\Gamma}_V \subset \bar{\Gamma}$ (and $\bar{\varphi}_V = \bar{\varphi}$).

Suppose $\varphi_L(\Gamma_L)$ is mediated in E and $\varphi_V(\Gamma_V)$ is splitting in E . Then Γ_L, Γ_V and Γ_{LV} are stable by Theorem 5.4.25(a),(b) and (c). Consequently, again by Theorem 5.4.25(b) Γ_{VL} is stable. $\square$

5.5.9 Corollary. *Suppose E is mediated (and thus splitting), $\bar{\Gamma} = \Gamma_{VL}$. Then $\bar{\Gamma} = \bar{\Gamma}_V = \bar{\Gamma}_L$, so $\bar{\Gamma} = \bar{\Gamma}$ (and $\bar{\varphi} = \bar{\varphi}$).*

At the end of Section 5.5 we will show that sometimes $\Gamma_{VL} \subsetneq \Gamma_{LV}$ (Example 5.5.14) and sometimes $\Gamma_{LV} \subsetneq \Gamma_{VL}$ (Example 5.5.15). Note that this implies that $\Gamma_{VL V}$ can be strictly larger than either Γ_{VL} or Γ_{LV} .

Theorem 5.5.8 raises the question whether stability of Γ_V entails $\Gamma_{VL} \subset \Gamma_{LV}$. In general the answer is negative; see Example 5.5.15. In Theorem 5.5.10 we give conditions sufficient for the inclusion.

5.5.10 Theorem. *Suppose Γ_V is stable. Consider these two statements.*

- (a) *For every $f \in \Gamma_{VL}^+$ there is a ρ in Γ_L^+ with $f \leq \rho$.*

(b) E satisfies:

If $Y_1, Y_2, \dots \subset E$ are nonempty countable with $\inf Y_n = 0$ for all $n \in \mathbb{N}$,
then there exist $y_1 \in Y_1, y_2 \in Y_2, \dots$ such that $\sum_n y_n$ exists in E . (5.43)

If (a) is satisfied, then $\Gamma_{VL} \subset \Gamma_{LV}$. (b) implies (a).

Proof. If (a) is satisfied, then by Theorem 5.5.5(c) follows that $\Gamma_{VL} \subset \Gamma_{LV}$.
Suppose (b). Let $f \in \Gamma_{VL}^+$. Let $(A_n)_{n \in \mathbb{N}}$ be a φ_V -partition for f . For $n \in \mathbb{N}$, let $\Upsilon_n \subset \Gamma$ be a countable set with $f \mathbb{1}_{A_n} \leq \Upsilon_n$ and

$$\varphi_V(f \mathbb{1}_{A_n}) = \inf \varphi(\Upsilon_n). \quad (5.44)$$

We may assume $\sigma \mathbb{1}_{A_n} = \sigma$ for all $\sigma \in \Upsilon_n$. Choose $\sigma_n \in \Upsilon_n$ for $n \in \mathbb{N}$ such that $\sum_n (\varphi(\sigma_n) - \varphi_V(f \mathbb{1}_{A_n}))$ and thus $\sum_n \varphi(\sigma_n)$ exist in E . Then $\rho := \sum_{n \in \mathbb{N}} \sigma_n$ is in Γ_L^+ with $f \leq \rho$. $\square$

5.5.11. We will discuss examples of spaces E for which (5.43) holds.

(I) If E is a Banach lattice with σ -order continuous norm, then E satisfies (5.43) (one can find $y_n \in Y_n$ with $\|y_n\| \leq 2^{-n}$).

(II) Let $(X, \mathcal{A}, \mu)$ be a complete σ -finite measure space and assume there exists a $g \in L^1(\mu)$ with $g > 0$ μ -a.e.. Then the space E of equivalence classes of measurable functions $X \rightarrow \mathbb{R}$ satisfies (5.43): It is sufficient to prove that if $Z_1, Z_2, \dots \subset E$ are nonempty countable with $\inf Z_n = 0$ for all $n \in \mathbb{N}$, then there exists $z_1 \in Z_1, z_2 \in Z_2, \dots$ and a $z \in E$ such that $z_n \leq z$ for all $n \in \mathbb{N}$ (for Z_n take $2^n Y_n$). One can prove that such a z exists by mapping the equivalence classes of measurable functions into $L^1(\mu)$ by the order isomorphism $f \mapsto (\arctan \circ f)g$.

(III) $\mathbb{R}^{\mathbb{N}}$ is a special case of (II), therefore satisfies (5.43).

5.5.12 Theorem. Let E be mediated and splitting and satisfy (5.43) (e.g. E be a Banach lattice with σ -order continuous norm (Theorem 5.4.24), or E is the space mentioned in 5.5.11(II)). Then Γ_V is stable and $\Gamma_{VL} = \Gamma_{LV}$, $\varphi_{VL} = \varphi_{LV}$.

Proof. This is a consequence of Theorem 5.5.8 and Theorem 5.5.10. $\square$

For a Riesz space F and a Riesz subspace Γ of F^X we will now investigate under which conditions on $\varphi(\Gamma)$, $\varphi_L(\Gamma_L)$ and $\varphi_V(\Gamma_V)$ the spaces Γ_{LV} and Γ_{VL} are Riesz subspaces of F^X .

5.5.13 Theorem. Suppose F is a Riesz space and Γ is a Riesz subspace of F^X . If $\varphi(\Gamma)$ is splitting in E and $\varphi_L(\Gamma_L)$ is mediated in E , then Γ_{LV} is a Riesz subspace of F^X . If $\varphi(\Gamma)$ is mediated in E and $\varphi_V(\Gamma_V)$ is splitting in E , then Γ_{VL} is a Riesz subspace of F^X .

In particular, if E is mediated (and thus splitting), then both Γ_{LV} and Γ_{VL} are Riesz subspaces of F^X .

Proof. Note first that if $\varphi(\Gamma)$ is mediated in E , then Γ_V is stable by Theorem 5.4.25(b). For a proof, combine Theorem 5.4.32 and Corollary 5.3.10. $\square$

The next example illustrates that Γ_{LV} is not always included in Γ_{VL} (given that Γ_V is stable) even if E and F are Riesz spaces and $\Gamma, \Gamma_{LV}, \Gamma_{VL}$ Riesz subspaces of F^X .

5.5.14 Example. $[\Gamma_{VL} \subsetneq \Gamma_{LV} = \Gamma_{VLV}]$

For an element $b = (\beta_1, \beta_2, \dots)$ of $\mathbb{R}^{\mathbb{N}}$ we write $b = \sum_{n \in \mathbb{N}} \beta_n e_n$. Consider $X = \{0, 1, 2, \dots\}$ and $\mathcal{I} = \mathcal{P}(X)$. Let $E = c$, $F = \mathbb{R}^{\mathbb{N}}$, $\Omega = F^X$. We view the elements of Ω as sequences $(a, b_1, b_2, \dots)$ with $a, b_1, b_2, \dots \in \mathbb{R}^{\mathbb{N}}$.

Define sets $\Gamma \subset \Theta \subset \Omega$ and a map $\Phi : \Theta \rightarrow \mathbb{R}^{\mathbb{N}}$ by

$$\Theta = \{(a, \beta_1 e_1, \beta_2 e_2, \dots) : a \in c, \beta_1, \beta_2, \dots \in \mathbb{R}\}, \quad (5.45)$$

$$\Phi(a, \beta_1 e_1, \beta_2 e_2, \dots) = a + \sum_{n \in \mathbb{N}} \beta_n e_n \quad (a \in c, \beta_1, \beta_2, \dots \in \mathbb{R}), \quad (5.46)$$

$$\Gamma = \{(a, \beta_1 e_1, \beta_2 e_2, \dots) : a \in c, (\beta_1, \beta_2, \dots) \in c_{00}\}. \quad (5.47)$$

Then $\Phi(\Gamma) = c = E$; let $\varphi = \Phi|_{\Gamma}$. From the definition it is easy to see that Γ is stable and φ is laterally extendable. We leave it to the reader to verify that $\Gamma_V = \Gamma$,

$$\Gamma_L = \{(a, \beta_1 e_1, \beta_2 e_2, \dots) : a \in c, (\beta_1, \beta_2, \dots) \in c\} \quad (5.48)$$

and $\varphi_L = \Phi$ on Γ_L .

It follows that Γ_V is stable and $\Gamma_{VL} = \Gamma_L \subset \Gamma_{LV} = \Gamma_{VLV}$. We prove $\Gamma_{VL} \neq \Gamma_{LV}$. To this end, define $h \in \Omega$ by

$$\begin{cases} h(n) = (-1)^n e_n & (n = 1, 2, \dots), \\ h(0) = -\sum_{n \in \mathbb{N}} h(n) = -\sum_{n \in \mathbb{N}} (-1)^n e_n. \end{cases} \quad (5.49)$$

As $h(0) \notin c$ we have $h \mathbb{1}_{\{0\}} \notin \Gamma$; in particular, h is not partially in Γ , so $h \notin \Gamma_L = \Gamma_{VL}$. It remains to prove $h \in \Gamma_{LV}$.

For $k \in \mathbb{N}$, define $\tau_k, \sigma_k : X \rightarrow \mathbb{R}^{\mathbb{N}}$:

$$\begin{cases} \tau_k(0) = -\sum_{n=1}^k (-1)^n e_n + \sum_{n=k+1}^{\infty} e_n, \\ \tau_k(n) = h(n) = (-1)^n e_n & (n = 1, \dots, k), \\ \tau_k(n) = e_n & (n = k+1, k+2, \dots), \end{cases} \quad (5.50)$$

$$\begin{cases} \sigma_k(0) = -\sum_{n=1}^k (-1)^n e_n - \sum_{n=k+1}^{\infty} e_n, \\ \sigma_k(n) = h(n) = (-1)^n e_n & (n = 1, \dots, k), \\ \sigma_k(n) = -e_n & (n = k+1, k+2, \dots). \end{cases} \quad (5.51)$$

Then $\tau_k, \sigma_k \in \Gamma_L$, $\tau_k \geq h \geq \sigma_k$, $\varphi_L(\tau_k) = \Phi(\tau_k) = 2 \sum_{n > k} e_n$, $\varphi_L(\sigma_k) = -2 \sum_{n > k} e_n$, so $\inf_{k \in \mathbb{N}} \varphi_L(\tau_k) = \sup_{k \in \mathbb{N}} \varphi_L(\sigma_k) = 0$, and $h \in \Gamma_{LV}$.

The next example illustrates that Γ_{VL} is not always included in Γ_{LV} ; it provides an example of an $f \in \Gamma_{VL}^+$ for which there exist no $\rho \in \Gamma_L^+$ with $f \leq \rho$ (see Theorem 5.5.5(c)).

5.5.15 Example. $[\Gamma_{LV} \subsetneq \Gamma_{VL}]$

Let $E = C[0, 1]$ and let $D \subset C[0, 1]$ be the set of polynomials of degree ≤ 2 . The set

D is order dense⁷ in $C[0, 1]$ (see [42, Example 4.4]). Hence, for all $f \in E$ there exist $(g_n)_{n \in \mathbb{N}}, (h_n)_{n \in \mathbb{N}}$ in D with $f = \inf_{n \in \mathbb{N}} g_n = \sup_{n \in \mathbb{N}} h_n$. Therefore E is the vertical extension of D with respect to the inclusion map $D \rightarrow E$.

Take $X = \mathbb{N}$, $\mathcal{I} = \mathcal{P}(\mathbb{N})$, $F = E = C[0, 1]$, $\Gamma = c_{00}[D] \subset F^{\mathbb{N}} = E^{\mathbb{N}}$ and let $\varphi : \Gamma \rightarrow E$ be given by $\varphi(f) = \sum_{n \in \mathbb{N}} f(n)$. Since this situation is the same as in Example 5.5.6 with $D_V = E$, we have $\Gamma_V = c_{00}[E]$ and $\Gamma_{LV} \subset \Gamma_{VL}$.

Furthermore (see 5.4.6)

$$\Gamma_L^+ = \{f \in (D^+)^{\mathbb{N}} : \sum_n f(n) \text{ exists in } E\}, \quad (5.52)$$

$$\Gamma_{VL}^+ = \{f \in (E^+)^{\mathbb{N}} : \sum_n f(n) \text{ exists in } E\}. \quad (5.53)$$

We construct an $f \in \Gamma_{VL}^+$ that is not in Γ_{LV} . For $n \in \mathbb{N}$ let f_n be the ‘tent’ function defined by

$$\begin{aligned} f_n(0) &= 0; & f_n\left(\frac{1}{n}\right) &= 1; & f_n\left(\frac{1}{i}\right) &= 0 \quad \text{if } i \in \mathbb{N}, i \neq n; \\ f_n &\text{ is affine on the interval } \left[\frac{1}{1+i}, \frac{1}{i}\right] \text{ for all } i \in \mathbb{N}. \end{aligned} \quad (5.54)$$

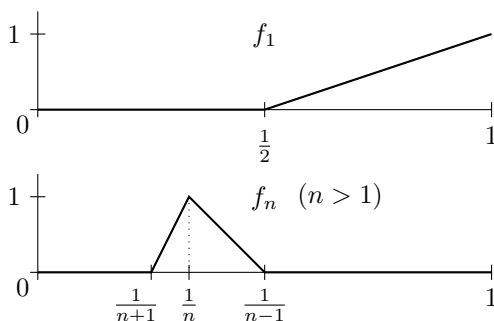


Figure 5.1: Graph of f_n .

Then $\sum_{n=1}^{\infty} f_n = \mathbb{1}_{(0,1]}$ pointwise, so $\sum_n f_n = \mathbb{1}$ in $C[0, 1]$. Hence $f = (f_1, f_2, f_3, \dots) \in \Gamma_{VL}^+$.

We will prove that $f \notin \Gamma_{LV}$; by showing there exists no $\rho \in \Gamma_L$ for which $f \leq \rho$.

Suppose $\rho \in \Gamma_L$ and $f \leq \rho$. Then $\rho = (\rho_1, \rho_2, \dots)$ where $\rho_1, \rho_2, \dots$ are elements of D^+ and $j = \sum_n \rho_n$ exists in $E = C[0, 1]$. Let M be the largest value of j . Every ρ_n is a quadratic function that maps $[0, 1]$ into $[0, M]$. Consequently (see the postscript)

$$|\rho_n(x) - \rho_n(y)| \leq 4M|x - y| \quad (x, y \in [0, 1], n \in \mathbb{N}). \quad (5.55)$$

In particular, $\rho_n(0) \geq \rho_n\left(\frac{1}{n}\right) - 4M\frac{1}{n} \geq f_n\left(\frac{1}{n}\right) - 4M\frac{1}{n} = 1 - 4M\frac{1}{n} \geq \frac{1}{2}$ for $n \geq 8M$. As $j(0) \geq \sum_{n \geq N} \rho_n(0)$ for all $N \in \mathbb{N}$, this is a contradiction.

Postscript. Let $h : x \mapsto ax^2 + bx + c$ be a quadratic function on $[0, 1]$ and $0 \leq h(x) \leq M$

⁷A subspace D of a partially ordered vector space E is called *order dense* in E if $x = \sup\{d \in D : d \leq x\}$ (and thus $x = \inf\{d \in D : d \geq x\}$) for all $x \in E$.

for all x ; we prove $|h'(x)| \leq 4M$ for all $x \in [0, 1]$. Since the derivative is either decreasing or increasing, we have $|h'(x)| \leq \max\{|h'(0)|, |h'(1)|\}$. Now $h'(0) = b = 4h(\frac{1}{2}) - h(1) - 3h(0)$ and $h'(1) = 2a + b = 3h(1) + h(0) - 4h(\frac{1}{2})$. Since $|h(x) - h(y)| \leq M$ for all $x, y \in [0, 1]$, we get the bounds $|h'(0)| \leq 4M$ and $|h'(1)| \leq 4M$ as desired.

5.5.16. Observe that Γ_{VL} in Example 5.5.15 is not stable since $(f_1, 0, f_3, 0, \dots) \notin \Gamma_{VL}$.

§5.6 Embedding E in a (slightly) larger space

In this section $E, F, X, \mathcal{I}, \Gamma, \varphi$ are as in Section 5.4.

Suppose $E^\bullet$ is another partially ordered vector space and $E \subset E^\bullet$. Consider $\varphi^\bullet : \Gamma \rightarrow E^\bullet$, where $\varphi^\bullet(f) = \varphi(f)$ for $f \in \Gamma$.

Write $\Gamma_V^\bullet$ for the vertical extension of Γ with respect to $\varphi^\bullet$. If $\varphi^\bullet$ is laterally extendable, write $\Gamma_L^\bullet$ for the lateral extension of Γ with respect to $\varphi^\bullet$, $\Gamma_{LV}^\bullet$ for the vertical extension of $\Gamma_L^\bullet$ with respect to $\varphi_L^\bullet$. Similarly, if $\Gamma_V^\bullet$ is stable, we introduce the notations $\Gamma_{VL}^\bullet$ and $\Gamma_{VLV}^\bullet$.

It is not generally the case that $\Gamma_V \subset \Gamma_V^\bullet$ or $\Gamma_L \subset \Gamma_L^\bullet$, but a natural restriction on $E^\bullet$ helps; see Theorem 5.6.2.

For $E^\bullet$ we can choose to be a Dedekind complete Riesz space in which countable suprema of E are preserved, in case E is Archimedean and directed (see 5.6.3). In this situation, in some sense, $\Gamma_{VL}^\bullet$ is the largest extension one can obtain.

5.6.1 Definition. Let D be a subspace of a partially ordered vector space P . Then we say that *countable suprema in D are preserved in P* if the following implication holds for all $a \in D$ and all countable $A \subset D$

$$A \text{ has supremum } a \text{ in } D \implies A \text{ has supremum } a \text{ in } P. \quad (5.56)$$

Note that the reverse implication holds always.

The following theorem is a natural consequence.

5.6.2 Theorem. *Suppose that countable suprema in E are preserved in $E^\bullet$. Then $\varphi^\bullet$ is laterally extendable and*

$$f \in \Gamma_V \iff f \in \Gamma_V^\bullet \text{ and } \varphi_V^\bullet(f) \in E, \quad (5.57)$$

$$f \in \Gamma_L \iff f \in \Gamma_L^\bullet \text{ and } \varphi_L^\bullet(f) \in E, \quad (5.58)$$

$$\varphi_V^\bullet(f) = \varphi_V(f) \text{ for } f \in \Gamma_V, \quad \varphi_L^\bullet(f) = \varphi_L(f) \text{ for } f \in \Gamma_L, \quad (5.59)$$

$$\Gamma_{LV} \subset \Gamma_{LV}^\bullet, \quad \varphi_{LV}^\bullet(f) = \varphi_{LV}(f) \text{ for } f \in \Gamma_{LV}. \quad (5.60)$$

Suppose Γ_V and $\Gamma_V^\bullet$ are stable. Then

$$\Gamma_{VL} \subset \Gamma_{VL}^\bullet, \quad \varphi_{VL}^\bullet(f) = \varphi_{VL}(f) \text{ for } f \in \Gamma_{VL}, \quad (5.61)$$

$$\Gamma_{VLV} \subset \Gamma_{VLV}^\bullet, \quad \varphi_{VLV}^\bullet(f) = \varphi_{VLV}(f) \text{ for } f \in \Gamma_{VLV}. \quad (5.62)$$

5.6.3. Under the assumptions made in Section 5.4 Γ is directed, thus so are Γ_L , Γ_V (see 5.3.11) and Γ_{LV} (etc.). Hence $\varphi_V(\Gamma_V)$, $\varphi_L(\Gamma_L)$, $\varphi_{LV}(\Gamma_{LV})$ (etc.) are all subsets of $E^+ - E^+$. For this reason we may assume that E itself is directed.

Then under the (rather general) assumption that E is also Archimedean (see Definition 5.3.19), E can be embedded in a Dedekind complete Riesz space such that suprema and infima in E are preserved, as we state in Theorem 5.6.4.

Consequently, choosing such a Dedekind complete Riesz space for $E^\bullet$ one has the following: $\Gamma_V^\bullet$, $\Gamma_{LV}^\bullet$, $\Gamma_{VL}^\bullet$, $\Gamma_{V LV}^\bullet$ are stable and $\Gamma_{LV}^\bullet \subset \Gamma_{VL}^\bullet =: \bar{\Gamma}^\bullet$, $\bar{\Gamma}_L^\bullet = \bar{\Gamma}_V^\bullet = \bar{\Gamma}^\bullet$ and $\bar{\varphi}_L^\bullet = \bar{\varphi}_V^\bullet = \bar{\varphi}^\bullet$, where $\bar{\varphi}^\bullet := \varphi_{VL}^\bullet$ (see 5.5.8). Moreover, one has (5.60) and if Γ_V is stable; (5.61) and (5.62). For this reason one may consider $\bar{\Gamma}^\bullet$ and $\bar{\varphi}^\bullet$ instead of Γ_{LV} and φ_{LV} , instead of $\Gamma_{LV}^\bullet$ and $\varphi_{LV}^\bullet$ or instead of $\Gamma_{V LV}$ and $\varphi_{V LV}$, indeed $\bar{\Gamma}^\bullet$ contains all of the other extensions and $\bar{\varphi}^\bullet$ agrees with all integrals.

5.6.4 Theorem. [73, Chapter 4, Theorem 1.19]

Let E be an Archimedean directed partially ordered vector space. Then E can be embedded in a Dedekind complete Riesz space $\hat{E}$:

There exists an injective linear $\gamma : E \rightarrow \hat{E}$ for which

- (a) $a \geq 0 \iff \gamma(a) \geq 0$,
- (b) $\gamma(E)$ is order dense in $\hat{E}$ (for the definition of order dense see the seventh footnote).

Consequently, suprema in $\gamma(E)$ are preserved in $\hat{E}$.

§5.7 Integration for functions with values in $\mathbb{R}$

In this section $(X, \mathcal{A}, \mu)$ is a complete σ -finite measure space and $E = F = \mathbb{R}$.

We write S for the vector space of simple functions from X to $\mathbb{R}$ (see 5.4.33). Since $\mathbb{R}$ is a Banach lattice with σ -order continuous norm, S_V is stable and $S_{LV} = S_{VL}$, $\varphi_{LV} = \varphi_{VL}$ (by Theorem 5.5.12). We write $\bar{S} = S_{VL}$ and $\bar{\varphi} = \varphi_{VL}$.

5.7.1 Theorem. $\bar{S} = \mathcal{L}^1(\mu)$ and $\bar{\varphi}(f) = \int f \, d\mu$ for all $f \in \bar{S}$.

Proof. We prove that $S_{VL}^+ \subset \mathcal{L}^1(\mu)^+ \subset S_{LV}^+$ and that $\varphi_{LV}(f) = \int f \, d\mu$ for all $f \in \mathcal{L}^+(\mu)$.

S_V consists of the bounded integrable functions f for which $\{x \in X : f(x) \neq 0\}$ has finite measure. By monotone convergence, we have $f \in \mathcal{L}^1(\mu)$ for every $f \in S_{VL}^+$.

Conversely, let $f \in \mathcal{L}^1(\mu)^+$; we prove $f \in S_{LV}^+$ and $\varphi_{LV}(f) = \int f \, d\mu$. Let $t \in (1, \infty)$. For $n \in \mathbb{Z}$, put $A_n = \{x \in X : t^n \leq f(x) < t^{n+1}\}$. Then $(A_n)_{n \in \mathbb{Z}}$ forms a partition. Define $g := \sum_{n \in \mathbb{Z}} t^n \mathbb{1}_{A_n}$ and $h := tg$; then $g \leq f \leq h$. Since

$$\sum_{n \in \mathbb{Z}} t^n \mu(A_n) \leq \sum_{n \in \mathbb{Z}} \int f \mathbb{1}_{A_n} \, d\mu = \int f \, d\mu, \quad (5.63)$$

we have $g \in S_L$ and $\varphi_L(g) \leq \int f \, d\mu$. Also, $h = tg \in S_L$, and $\varphi_L(h) - \varphi_L(g) = (t-1)\varphi_L(g) \leq (t-1) \int f \, d\mu$. By this and Lemma 5.3.7 it follows that $f \in S_{LV}$ and $\varphi_{LV}(f) = \int f \, d\mu$. $\square$

§5.8 Extensions of integrals on simple functions

In this section E is a directed partially ordered vector space, $(X, \mathcal{A}, \mu)$ is a complete σ -finite measure space and $\mathcal{I}, S, \varphi$ are as in 5.4.33 ($F = E$).

In 5.8.1–5.8.8 for f in S_{LV} or S_{VL} we discuss the relation between f being almost everywhere equal to zero and f having integral zero (i.e., either $\varphi_{LV}(f) = 0$ or $\varphi_{VL}(f) = 0$).

In 5.8.9 we show that under some conditions a function in S_V multiplied with an integrable function with values in $\mathbb{R}$ is a function in S_{LV} .

In 5.8.11–5.8.13 we investigate the relation between the “ LV ”-extension on simple functions with respect to μ and ν , where $\nu = h\mu$ for some measurable $h : X \rightarrow [0, \infty)$.

In 5.8.14 we discuss the relation between the “ LV ”-extension simple functions with values in E or in another partially ordered vector space F , when one makes the composition of a function in the extension with a σ -order continuous linear map $E \rightarrow F$.

In 5.8.15–5.8.17 we will prove that under certain conditions on X the function $x \mapsto F(x, \cdot)$ is in S_V for all $F \in C(X \times T)$ and we relate that to convolution of certain finite measures with continuous functions on a topological group.

5.8.1 Theorem. *Let $f : X \rightarrow E$ and $f = 0$ a.e.. If $f \in S_{LV}$, then $\varphi_{LV}(f) = 0$. If S_V is stable and $f \in S_{VLV}$, then $\varphi_{VLV}(f) = 0$.*

Proof. Let $B = \{x \in X : f(x) \neq 0\}$. Then $B \in \mathcal{A}$ and $\mu(B) = 0$.

(I) Assume $f \in S_V$. Choose $\sigma, \tau \in S$ with $\sigma \leq f \leq \tau$. Then $\sigma \mathbb{1}_B, \tau \mathbb{1}_B \in S$, $\sigma \mathbb{1}_B \leq f \leq \tau \mathbb{1}_B$, and $\varphi(\sigma \mathbb{1}_B) = \varphi(\tau \mathbb{1}_B) = 0$. Hence $\varphi_V(f) = 0$.

(II) Suppose $\sigma \in S_L^+$ and $(A_n)_{n \in \mathbb{N}}$ is a φ -partition for σ . Then $\sigma \mathbb{1}_{A_n \cap B} \in S^+$ for all $n \in \mathbb{N}$ and $\sum_n \varphi(\sigma \mathbb{1}_{A_n \cap B}) = 0$, i.e., $\sigma \mathbb{1}_B \in S_L^+$ with $\varphi_L(\sigma \mathbb{1}_B) = 0$. In particular, if $f \in S_L$ then $\varphi_L(f) = 0$.

(III) Assume $f \in S_{LV}$. With (II) one can repeat the argument of (I) with S replaced by S_L and conclude $\varphi_{LV}(f) = 0$.

(IV) Suppose S_V is stable and $f \in S_{VLV}$. One can repeat the argument in (III) with S replaced by S_V and conclude $\varphi_{VLV}(f) = 0$. $\square$

5.8.2 Definition. A subset $D \subset E$ is called *order bounded* if there are $a, b \in E$ for which $a \leq D \leq b$.

5.8.3 Theorem. *Let $f \in S_{LV}$ or (assuming S_V is stable) $f \in S_{VLV}$. Then there exists a partition $(A_n)_{n \in \mathbb{N}}$ such that each set $f(A_n)$ is order bounded.*

Proof. There exists a partition $(A_n)_{n \in \mathbb{N}}$ such that for all $n \in \mathbb{N}$ there exist $h_n, g_n \in S$ for which $h_n \leq f \mathbb{1}_{A_n} \leq g_n$. Choose $a_n, b_n \in E$ for which $a_n \leq h_n(x)$ and $g_n(x) \leq b_n$ for all $x \in X$. Then $a_n \leq f(x) \leq b_n$ for $n \in \mathbb{N}$, $x \in A_n$. $\square$

5.8.4 Theorem. *Let $f : X \rightarrow E$ and $f = 0$ a.e.. Suppose there exists a partition $(A_n)_{n \in \mathbb{N}}$ such that for every $n \in \mathbb{N}$ the subset $f(A_n)$ of E is order bounded. Then $f \in S_{LV}$ and if S_V is stable then also $f \in \Gamma_{VL}$.*

Proof. Choose $a_1, a_2, \dots$ and $b_1, b_2, \dots$ in E such that

$$a_n \leq f(x) \leq b_n \quad (n \in \mathbb{N}, x \in A_n). \quad (5.64)$$

Let $B = \{x \in X : f(x) \neq 0\}$. Then $B \in \mathcal{A}$ and $\mu(B) = 0$. Hence $g := \sum_{n \in \mathbb{N}} a_n \mathbb{1}_{A_n \cap B}$ and $h := \sum_{n \in \mathbb{N}} b_n \mathbb{1}_{A_n \cap B}$ are elements of S_L with $\varphi(g) = 0$ and $\varphi_L(h) = 0$. As $g \leq f \leq h$, we get $f \in S_{LV}$ and if S_V is stable also $f \in S_{VL}$. $\square$

For a real valued function $f : X \rightarrow \mathbb{R}$ with $f \geq 0$ and $\int f \, d\mu = 0$ we have $f = 0$ a.e.. We will give an example of a $f \in S_V^+$ with $\varphi_V(f) = 0$ but which is nowhere zero (Example 5.8.8). On the positive side, in Theorem 5.8.7 we show that $f = 0$ a.e. if $f \in S_{LV}^+$ and $\varphi_{LV}(f) = 0$ provided that E satisfies a certain separability condition.

5.8.5 Definition. We call a subset D of $E^+ \setminus \{0\}$ *pervasive*⁸ in E if for all $a \in E$ with $a > 0$ there exists a $d \in D$ such that $0 < d \leq a$. We say that E *possesses a pervasive subset* if there exists a pervasive $D \subset E^+ \setminus \{0\}$.

5.8.6 Example. The Riesz spaces $\mathbb{R}^{\mathbb{N}}, \ell^\infty, c, c_0, \ell^1$ and c_{00} possess countable pervasive subsets. Indeed, in each of them the set $\{\lambda e_n : \lambda \in \mathbb{Q}^+, \lambda > 0, n \in \mathbb{N}\}$ is pervasive. If $\mathcal{X}$ is a completely regular topological space, then $C(\mathcal{X})$ has a countable pervasive subset if and only if $\mathcal{X}$ has a countable base. (If $D \subset E^+ \setminus \{0\}$ is countable and pervasive, then $\mathfrak{U} = \{f^{-1}(0, \infty) : f \in D\}$ is a countable base; vice versa if $\mathfrak{U}$ is a countable base then with choosing an f_U in $C(X)^+$ for each $U \in \mathfrak{U}$ with $f_U = 0$ on U^c and $f_U(x) = 1$ for some $x \in U$, the set $D = \{\varepsilon f_U : \varepsilon \in \mathbb{Q}, \varepsilon > 0, U \in \mathfrak{U}\}$ is pervasive.)

$L^1(\lambda)$ and $L^\infty(\lambda)$ do not possess countable pervasive subsets, considering the Lebesgue measure space $(\mathbb{R}, \mathcal{M}, \lambda)$. (Suppose one of them does. Then one can prove the existence of non-negligible measurable sets $A_1, A_2, \dots \in \mathcal{M}$ such that every non-negligible measurable set contains an A_n , whereas $\lambda(A_n) < 2^{-n}$ for all $n \in \mathbb{N}$. Putting $C = \mathbb{R} \setminus \bigcup_{n \in \mathbb{N}} A_n$ we have a non-negligible measurable set that contains no A_n : a contradiction.)

5.8.7 Theorem. Let E possess a countable pervasive subset D . Let $f \in S_{LV}$. Let $\Lambda, \Upsilon \subset S_L$ be countable sets such that $\Lambda \leq f \leq \Upsilon$ and $\sup \varphi_L(\Lambda) = \inf \varphi_L(\Upsilon)$. Then for almost all $x \in X$

$$\sup_{g \in \Lambda} g(x) = f(x) = \inf_{h \in \Upsilon} h(x). \quad (5.65)$$

Consequently, if $f \in S_{LV}^+$ and $\varphi_{LV}(f) = 0$, then $f = 0$ a.e.. (However, see Example 5.8.8.)

Proof. (I) First, as a special case (namely $f = 0$), let $(\tau_n)_{n \in \mathbb{N}}$ be a sequence in S_L with $\tau_n \geq 0$ for all $n \in \mathbb{N}$ and $\inf_{n \in \mathbb{N}} \varphi_L(\tau_n) = 0$. We prove that $\inf_{n \in \mathbb{N}} \tau_n(x) = 0$ for almost all $x \in X$, by proving that $\mu(A) = 0$, where A is the complement of the set $\{x \in X : \inf_{n \in \mathbb{N}} \tau_n(x) = 0\}$. Indeed, for this A we have

$$A = \bigcup_{d \in D} A_d, \quad \text{with} \quad A_d = \bigcap_{n \in \mathbb{N}} \{x \in X : d \leq \tau_n(x)\}. \quad (5.66)$$

⁸Our use of the term is similar to the one of O. van Gaans and A. Kalauch in [41, Definition 2.3].

Note that for all $n \in \mathbb{N}$ and $d \in D$ the set $\{x \in X : d \leq \tau_n(x)\}$ is measurable. Furthermore, for all $d \in D$ we have:

$$d\mu(A_d) = \varphi(d\mathbb{1}_{A_d}) \leq \varphi_L(\tau_n) \quad (n \in \mathbb{N}). \quad (5.67)$$

Hence $\mu(A_d) = 0$ for all $d \in D$ and thus $\mu(A) = 0$.

(II) Suppose that $\Lambda, \Upsilon \subset \Gamma_L$ are countable sets such that $\Lambda \leq f \leq \Upsilon$, $\sup \varphi_L(\Lambda) = \inf \varphi_L(\Upsilon)$. Then $\inf \varphi_L(\Upsilon - \Lambda) = 0$, so by (I) $\inf_{g \in \Upsilon, h \in \Lambda} (g(x) - h(x)) = 0$ for almost all $x \in X$. $\square$

5.8.8 Example. We give an example of a $f \in S_V^+$ with $\varphi_V(f) = 0$, where $f \neq 0$ everywhere. Let $([0, 1], \mathcal{M}, \lambda)$ be the Lebesgue measure space with underlying set $[0, 1]$. Let $E = \ell^\infty([0, 1])$ (see Section 5.2). Let $f : \mathbb{R} \rightarrow E^+$ be defined by $f(t) = \mathbb{1}_{\{t\}}$ for $t \in [0, 1]$. Note that f is not partially in S . We will show $f \in S_V$. For $n \in \mathbb{N}$ make $\tau_n \in S$:

$$\tau_n(t) = \mathbb{1}_{[\frac{i-1}{n}, \frac{i}{n})} \quad \text{if } i \in \{1, \dots, n\}, t \in [\frac{i-1}{n}, \frac{i}{n}). \quad (5.68)$$

Then $\varphi(\tau_n) = \frac{1}{n} \mathbb{1}_{[0, 1]}$ and $0 \leq f \leq \tau_n$ for $n \in \mathbb{N}$, so $f \in S_V$ and $\varphi_V(f) = 0$. But $f(t) \neq 0$ for all t .

5.8.9 Theorem. Let E be Archimedean and mediated. Let $f : X \rightarrow E$ and $g : X \rightarrow \mathbb{R}$. We write gf for the function $x \mapsto g(x)f(x)$. Then

- (a) $f \in S_V$ and g is bounded and measurable $\implies gf \in S_V$.
- (b) f is partially in S_V and g is measurable $\implies gf$ is partially in S_V .
- (c) $f \in S_V$ and $g \in \mathcal{L}^1(\mu) \implies gf \in S_{LV}$.
- (d) $f \in S_{VL}$ and g is bounded and measurable $\implies gf \in S_{VL}$.
- (e) $f \in S_{VL}$, $f(X)$ is order bounded and $g \in \mathcal{L}^1(\mu) \implies gf \in S_{VL}$.

Proof. E is splitting (see 5.4.23(b)).

- (a) is a consequence of Theorem 5.3.21(a) (see also Remark 5.3.22).
- (b) Let $(A_n)_{n \in \mathbb{N}}$ be a partition such that $f\mathbb{1}_{A_n} \in S_V$ and $g\mathbb{1}_{A_n}$ is bounded for all $n \in \mathbb{N}$. By (a) every $gf\mathbb{1}_{A_n}$ lies in S_V . Then gf is partially in S_V .
- (c) Assume $f \geq 0$ and $g \geq 0$. Choose (see the proof of Theorem 5.7.1) a partition $(A_n)_{n \in \mathbb{N}}$ and numbers $\lambda_1, \lambda_2, \dots$ in $[0, \infty)$ with

$$\tau := \sum_{n \in \mathbb{N}} \lambda_n \mathbb{1}_{A_n} \geq g, \quad \sum_{n \in \mathbb{N}} \lambda_n \mu(A_n) < \infty. \quad (5.69)$$

Then $\tau s \in S_L$ for all $s \in S$. Choose $s \in S$ with $s \geq f$. Then $0 \leq gf \leq \tau s$. From Theorem 5.5.5(e) and (b) it follows that $gf \in S_{LV}$.

- (d) Assume $f \geq 0$ and $0 \leq g \leq \mathbb{1}$. Using (b), choose a partition $(A_n)_{n \in \mathbb{N}}$ with $f\mathbb{1}_{A_n} \in S_V$ and $gf\mathbb{1}_{A_n} \in S_V$ for all $n \in \mathbb{N}$. Then

$$0 \leq \varphi_V(gf\mathbb{1}_{A_n}) \leq \varphi_V(f\mathbb{1}_{A_n}) \quad (n \in \mathbb{N}). \quad (5.70)$$

Since $\sum_n \varphi_V(f\mathbb{1}_{A_n})$ exists and E is splitting, $\sum_n \varphi_V(gf\mathbb{1}_{A_n})$ exists.

- (e) Assume $f \geq 0$ and $g \geq 0$. Choose $a \in E^+$ with $f(x) \leq a$ for all $x \in X$. Choose a

partition $(A_n)_{n \in \mathbb{N}}$ and $\lambda_1, \lambda_2, \dots \in [0, \infty)$ with

$$gf \mathbb{1}_{A_n} \in S_V \quad (n \in \mathbb{N}), \quad (5.71)$$

$$g \leq \sum_{n \in \mathbb{N}} \lambda \mathbb{1}_{A_n}, \quad \sum_{n \in \mathbb{N}} \lambda_n \mu(A_n) < \infty \quad (\text{see the proof of Theorem 5.7.1}). \quad (5.72)$$

Then

$$gf \mathbb{1}_{A_n} \leq \lambda_n a \mathbb{1}_{A_n} \quad (n \in \mathbb{N}), \quad (5.73)$$

$$\varphi_V(\lambda_n a \mathbb{1}_{A_n}) = \varphi(\lambda_n a \mathbb{1}_{A_n}) = \lambda_n \mu(A_n) a \quad (n \in \mathbb{N}), \quad (5.74)$$

so $\sum_n \varphi_V(\lambda_n a \mathbb{1}_{A_n})$ exists and so does $\sum_n \varphi_V(gf \mathbb{1}_{A_n})$. $\square$

5.8.10. In Lemma 5.8.11, Theorem 5.8.12 and Theorem 5.8.13 we investigate the relation between the extensions S_{LV} generated by two different measures, namely μ and $h\mu$ for a measurable function $h : X \rightarrow [0, \infty)$.

Note that for such a function h and all $s \in (1, \infty)$ there exists a $j : X \rightarrow [0, \infty)$ that is partially in the space of simple functions $X \rightarrow [0, \infty)$, i.e., $j = \sum_{n \in \mathbb{N}} \alpha_n \mathbb{1}_{A_n}$ for a partition $(A_n)_{n \in \mathbb{N}}$ and $(\alpha_n)_{n \in \mathbb{N}}$ in $[0, \infty)$ (or in the language of 5.3.16 j is partially in $[\mathcal{A}]$) for which $j \leq h \leq sj$. In the following (5.8.11, 5.8.12 and 5.8.13) we will write $\mathcal{I}^\mu$, S^μ and φ^μ instead of $\mathcal{I}$, S and φ and, similarly for another measure ν on $(X, \mathcal{A})$, we write $\mathcal{I}^\nu$, S^ν and φ^ν according to 5.4.33 with ν instead of μ .

5.8.11 Lemma. Suppose E is splitting. Let $h : X \rightarrow [0, \infty)$ be measurable, $\nu := h\mu$. Let $s \in (1, \infty)$ and let $j : X \rightarrow [0, \infty)$ be partially in $[\mathcal{A}]$ and such that $j \leq h \leq sj$. Let $f \in S_L^{\nu+}$. Then $jf \in S_L^\mu$ and $\varphi_L^\mu(jf) \leq \varphi_L^\nu(f) \leq s\varphi_L^\mu(jf)$.

Proof. Assume $(A_n)_{n \in \mathbb{N}}$ is a partition for j and a φ^μ -partition for f (so $(A_n)_{n \in \mathbb{N}}$ is in $\mathcal{I}^\nu \cap \mathcal{I}^\mu$, i.e., $\mu(A_n), \nu(A_n) < \infty$ for all $n \in \mathbb{N}$). Choose $(\alpha_n)_{n \in \mathbb{N}}$ in $[0, \infty)$ and $(b_n)_{n \in \mathbb{N}}$ in E^+ such that

$$j = \sum_{n \in \mathbb{N}} \alpha_n \mathbb{1}_{A_n}, \quad f = \sum_{n \in \mathbb{N}} b_n \mathbb{1}_{A_n}. \quad (5.75)$$

Then $jf = \sum_{n \in \mathbb{N}} \alpha_n b_n \mathbb{1}_{A_n}$ and thus is in S_L^μ if $\sum_n \mu(A_n) \alpha_n b_n$ exists in E . For each $n \in \mathbb{N}$

$$0 \leq \mu(A_n) \alpha_n = \int j \mathbb{1}_{A_n} d\mu \leq \int h \mathbb{1}_{A_n} d\mu = \nu(A_n), \quad (5.76)$$

whence $0 \leq \mu(A_n) \alpha_n b_n \leq \nu(A_n) b_n$. Because $f \in S_L^{\nu+}$, $\sum_n \nu(A_n) b_n$ exists in E . Since E is splitting also $\sum_n \mu(A_n) \alpha_n b_n$ exists in E , i.e., $jf \in S_L^\mu$.

Furthermore, $\varphi_L^\mu(jf) = \sum_n \mu(A_n) \alpha_n b_n \leq \sum_n \nu(A_n) b_n = \varphi_L^\nu(f)$. On the other hand, we get $\mu(A_n) \alpha_n = \int j \mathbb{1}_{A_n} d\mu \geq \frac{1}{s} \int h \mathbb{1}_{A_n} d\mu = \frac{1}{s} \nu(A_n)$ for each $n \in \mathbb{N}$: it follows that $\varphi_L^\mu(jf) \geq \frac{1}{s} \varphi_L^\nu(f)$. $\square$

5.8.12 Theorem. Let E be Archimedean and splitting. Let $h : X \rightarrow [0, \infty)$ be measurable, $\nu := h\mu$.

- (a) $f \in S_{LV}^\nu \implies hf \in S_{LV}^\mu, \varphi_{LV}^\mu(hf) = \varphi_{LV}^\nu(f),$
 (b) $f \in S_{VL}^\nu \implies hf \in S_{VL}^\mu, \varphi_{VL}^\mu(hf) = \varphi_{VL}^\nu(f).$

Proof. Since both S_{LV}^ν and S_{VL}^ν are directed, we assume $f \geq 0$.

(a) Let $f \in S_{LV}^{\nu+}$. For $n \in \mathbb{N}$ let j_n be partially in $[\mathcal{A}]$ and such that $j_n \leq h \leq (1 + \frac{1}{n})j_n$. Let $\Lambda, \Upsilon \subset S_L^\nu$ be countable sets with $\Lambda \leq f \leq \Upsilon$ be such that $\sup \varphi_L^\nu(\Lambda) = \varphi_{LV}^\nu(f) = \inf \varphi_L^\nu(\Upsilon)$. Then for all $\sigma \in \Lambda$ (note that $\sigma \in S_L^{\nu+} - S_L^{\nu+}$), $\tau \in \Upsilon$ and $n \in \mathbb{N}$ we have $j_n \sigma \leq hf \leq (1 + \frac{1}{n})j_n \tau$ and by Lemma 5.8.11 $j_n \sigma$ and $(1 + \frac{1}{n})j_n \tau$ are in S_L^μ . Therefore we are done if both $\inf_{n \in \mathbb{N}, \sigma \in \Lambda, \tau \in \Upsilon} \varphi_L^\mu((1 + \frac{1}{n})j_n \tau - j_n \sigma) = 0$ and $\varphi_L^\mu(j_n \sigma) \leq \varphi_{LV}^\nu(f) \leq \varphi_L^\mu((1 + \frac{1}{n})j_n \tau)$ for all $n \in \mathbb{N}$ and all $\sigma \in \Lambda, \tau \in \Upsilon$. By Lemma 5.8.11 applied repeatedly we have

$$\begin{aligned} 0 &\leq \varphi_L^\mu((1 + \frac{1}{n})j_n \tau - j_n \sigma) = \varphi_L^\mu(j_n \tau - j_n \sigma) + \frac{1}{n} \varphi_L^\mu(j_n \tau) \\ &\leq \varphi_L^\nu(\tau - \sigma) + \frac{1}{n} \varphi_L^\nu(\tau), \end{aligned} \quad (5.77)$$

which has infimum 0 since E is Archimedean and $\inf_{\tau \in \Upsilon, \sigma \in \Lambda} \varphi_L^\nu(\tau - \sigma) = 0$. On the other hand, by Lemma 5.8.11,

$$\varphi_L^\mu(j_n \sigma) \leq \varphi_L^\nu(\sigma) \leq \varphi_{LV}^\nu(f) \leq \varphi_L^\nu(\tau) \leq (1 + \frac{1}{n}) \varphi_L^\mu(j_n \tau) \quad (n \in \mathbb{N}, \sigma \in \Lambda, \tau \in \Upsilon). \quad (5.78)$$

(b) Let $f \in S_{VL}^{\nu+}$. Choose a partition $(A_n)_{n \in \mathbb{N}}$ with $f \mathbb{1}_{A_n} \in S_V^\nu$ for $n \in \mathbb{N}$. By (a), $hf \mathbb{1}_{A_n} \in S_{LV}^\mu$ for $n \in \mathbb{N}$; by Lemma 5.5.7 $hf \mathbb{1}_{A_n}$ is partially in S_V^μ . Therefore we can choose a partition $(B_n)_{n \in \mathbb{N}}$ with

$$f \mathbb{1}_{B_n} \in S_V^\nu, \quad hf \mathbb{1}_{B_n} \in S_V^\mu \quad (n \in \mathbb{N}). \quad (5.79)$$

By (a), $\varphi_V^\nu(f \mathbb{1}_{B_n}) = \varphi_V^\mu(hf \mathbb{1}_{B_n})$ for all $n \in \mathbb{N}$. But $f \in S_{VL}^{\nu+}$, so

$$\varphi_{VL}^\nu(f) = \sum_n \varphi_V^\nu(f \mathbb{1}_{B_n}) = \sum_n \varphi_V^\mu(hf \mathbb{1}_{B_n}). \quad (5.80)$$

Then $hf \in S_{VL}^\mu$ and $\varphi_{VL}^\mu(hf) = \varphi_{VL}^\nu(f)$. □

5.8.13 Theorem. *Let E be Archimedean and splitting. Let $h : X \rightarrow [0, \infty)$ be measurable, $\nu := h\mu$, $A = \{x \in X : h(x) > 0\}$. Let $f : X \rightarrow E$ be such that $hf \in S_{LV}^\mu$. Then $f \mathbb{1}_A \in S_{LV}^\nu$.*

Proof. Define $h^* : X \rightarrow [0, \infty)$ by

$$h^*(x) = \begin{cases} \frac{1}{h(x)} & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \quad (5.81)$$

Then h^* is measurable and $hh^* = \mathbb{1}_A$ and $\mathbb{1}_A = \mathbb{1} \nu$ -a.e..

hf is in S_L^μ and thus in $S_L^{\mathbb{1} A \mu}$, and since $\mathbb{1}_A \mu = h^* \nu$, also $hf \in S_L^{h^* \nu}$. By Theorem 5.8.12, applied to $h^*, h^* \nu, \nu, hf$ instead of h, ν, μ, f , the function $h^* hf$ is an element of S_{LV}^ν . But $h^* hf = \mathbb{1}_A f$. □

In Theorem 5.8.14 we show that extensions of simple functions with values in E composed with a σ -order continuous linear map $E \rightarrow F$ are extensions of simple functions with values in F (where E and F are Riesz spaces).

5.8.14 Theorem. *Let E and F be Riesz spaces. Let S^E and φ^E be as in 5.4.33, and let S^F and φ^F be defined analogously. Let $\mathcal{L}_c(E, F)$ denote the set of σ -order continuous linear functions $E \rightarrow F$ and $E_c^\sim = \mathcal{L}_c(E, \mathbb{R})$ (definition and notation as in Zaanen [91, Chapter 12, Section 84]). Let $f \in S_{LV}^E$. Then $\alpha \circ f \in S_{LV}^F$ for all $\alpha \in \mathcal{L}_c(E, F)$ and*

$$\alpha(\varphi_{LV}^E(f)) = \varphi_{LV}^F(\alpha \circ f). \quad (5.82)$$

In particular, $\alpha \circ f$ is integrable for all $\alpha \in E_c^\sim$, and $\alpha(\varphi_{LV}^E(f)) = \int \alpha \circ f \, d\mu$.

Proof. Suppose $\alpha \in \mathcal{L}_c(E, F)^+$. Let $\tau \in S_L^{E+}$. Suppose $\tau = \sum_{n \in \mathbb{N}} a_n \mathbb{1}_{A_n}$ for some partition $(A_n)_{n \in \mathbb{N}}$ and a sequence $(a_n)_{n \in \mathbb{N}}$ in E^+ . Then $\alpha(\varphi_L^E(\tau)) = \alpha(\sum_n \mu(A_n) a_n) = \sum_n \mu(A_n) \alpha(a_n)$. Thus $\alpha \circ \tau$ is in S_{LV}^F with $\alpha(\varphi_L^E(\tau)) = \varphi_L^F(\alpha \circ \tau)$. Let $(\sigma_n)_{n \in \mathbb{N}}, (\tau_n)_{n \in \mathbb{N}}$ be sequences in S_L^E with $\sigma_n \leq f \leq \tau_n$, $\sigma_n \uparrow$, $\tau_n \downarrow$ and $\varphi_{LV}^E(f) = \sup_{n \in \mathbb{N}} \varphi_L^E(\sigma_n) = \inf_{n \in \mathbb{N}} \varphi_L^E(\tau_n)$. Then we have $\alpha(\varphi_{LV}^E(f)) = \sup_{n \in \mathbb{N}} \alpha(\varphi_L^E(\sigma_n)) = \sup_{n \in \mathbb{N}} \varphi_L^F(\alpha \circ \sigma_n)$ and $\alpha(\varphi_{LV}^E(f)) = \inf_{n \in \mathbb{N}} \alpha(\varphi_L^E(\tau_n)) = \inf_{n \in \mathbb{N}} \varphi_L^F(\alpha \circ \tau_n)$. Since $\alpha \circ \sigma_n \leq \alpha \circ f \leq \alpha \circ \tau_n$ for all $n \in \mathbb{N}$, we conclude that $\alpha \circ f \in (S^F)_{LV}$ (see Theorem 5.7.1) with $\alpha(\varphi_{LV}^E(f)) = \varphi_{LV}^F(f)$. $\square$

Theorem 5.8.14 will be used in Section 5.9 to compare the integrals φ_{LV} and φ_{VL} with the Pettis integral.

Before proving Theorem 5.8.16 we state (in Theorem 5.8.15) that there is an equivalent formulation for a function F to be in $C(X \times T)$ whenever X, T are topological spaces and X is compact.

5.8.15 Theorem. [85, Theorem 7.7.5] *Let X be a compact and let T be a topological space. Let $F : X \times T \rightarrow \mathbb{R}$ be such that $F(\cdot, t) \in C(X)$ for all $t \in T$. Then $F \in C(X \times T)$ if and only if $t \mapsto F(\cdot, t)$ is continuous, where $C(X)$ is equipped with the supremum norm. Consequently, if $A \subset X$ is a compact set, then $t \mapsto \sup F(A, t)$ and $t \mapsto \inf F(A, t)$ are continuous.*

5.8.16 Theorem. *Let (X, d, μ) be a compact metric probability space. Let T be a topological space and $F \in C(X \times T)$. The function $H : X \rightarrow C(T)$ given by $H(x) = F(x, \cdot)$ is an element of S_V . Furthermore, for $t \in T$, $x \mapsto F(x, t)$ is integrable and*

$$[\varphi_V(H)](t) = \int F(x, t) \, d\mu(x) \quad (t \in T). \quad (5.83)$$

Proof. For $k \in \mathbb{N}$ let $A_{k1}, \dots, A_{kn_k}$ be a partition of X with $\text{diam } A_{ki} \leq k^{-1}$. Define

$$\Delta_k(t) = \sup_{x, y \in X, d(x, y) < k^{-1}} |F(x, t) - F(y, t)| \quad (t \in T). \quad (5.84)$$

Since $x \mapsto F(x, t)$ is uniformly continuous for all $t \in T$, $\Delta_k(t) \downarrow 0$ for all $t \in T$. By Theorem 5.8.15 $t \mapsto \sup F(A_{ki}, t)$ and $t \mapsto \inf F(A_{ki}, t)$ are continuous for all $k \in \mathbb{N}$ and $i \in \{1, \dots, n_k\}$. For $k \in \mathbb{N}$ let $h_k, l_k : X \rightarrow C(T)$ be given by

$$\begin{aligned} h_k(x) &= t \mapsto \sup F(A_{ki}, t) & (x \in A_{ki}), \\ l_k(x) &= t \mapsto \inf F(A_{ki}, t) & (x \in A_{ki}). \end{aligned} \quad (5.85)$$

Then $h_k, l_k \in S$ and $(h_k(x))(t) \geq F(x, t) \geq (l_k(x))(t)$ for all $x \in X$, $t \in T$. For $x \in A_{ki} \cap A_{mj}$ and $t \in T$

$$\begin{aligned} (h_k(x) - l_m(x))(t) &= \sup F(A_{ki}, t) - \inf F(A_{mj}, t) \\ &\leq \sup\{F(u, t) - F(v, t) : u, v \in A_{ki} \cup A_{mj}\} \leq \Delta_{k \wedge m}(t). \end{aligned} \quad (5.86)$$

Let $a_k = \varphi(h_k)$ and $b_k = \varphi(l_k)$ for $k \in \mathbb{N}$. Then $0 \leq a_k(t) - b_m(t) \leq \Delta_{k \wedge m}(t)$ for all $k, m \in \mathbb{N}$ and $\inf_{k, m \in \mathbb{N}} a_k(t) - b_m(t) \leq \inf_{k \in \mathbb{N}} \Delta_k(t) = 0$. Since $a_k, b_k \in C(T)$ and $\sup_{n \in \mathbb{N}} b_n(t) = \inf_{n \in \mathbb{N}} a_n(t)$ for all $t \in T$, the function $t \mapsto \inf_{n \in \mathbb{N}} a_n(t)$ is continuous, i.e., $x \mapsto F(x, \cdot)$ is an element of S_V . Furthermore, we conclude that the function $x \mapsto F(x, t)$ is integrable (by Theorem 5.7.1) and conclude (5.83). $\square$

5.8.17 Example. Consider a metrisable locally compact group G . Let $X \subset G$ be a compact set and μ be a finite (positive) measure on $\mathcal{B}(X)$, the Borel- σ -algebra of X . Let $g \in C(G)$. Define the convolution of g and μ to be the function $g * \mu : G \rightarrow \mathbb{R}$ given by $g * \mu(t) = \int g(tx^{-1}) d\mu(x)$ for $t \in G$. For $x \in X$, let $L_x g \in C(G)$ be the function $t \mapsto g(tx^{-1})$. Then by Theorem 5.8.16, the function $f : X \rightarrow C(G)$ given by $f(x) = L_x g$ is in S_V and $g * \mu = \varphi_V(f) \in C(G)$.

§5.9 Comparison with Bochner- and Pettis integral

We consider the situation of Section 5.8, with an E that has the structure of a Banach lattice. We write $\|\cdot\|$ for the norm on E and E' for the norm dual of E . Then, next to our φ_{LV} (and other extensions) there are the Bochner and the Pettis integrals. (We refer the reader to Hille and Phillips [48, Section 3.7] for background on both integrals.) We denote the set of Bochner (Pettis) integrable functions from the measure space $(X, \mathcal{A}, \mu)$ into the Banach lattice E by $\mathfrak{B}$ ($\mathfrak{P}$) and the Bochner (Pettis) integral of an integrable function f by $\mathfrak{b}(f)$ ($\mathfrak{p}(f)$).

5.9.1. By definition of the Bochner integral, where one also starts with defining the integral on simple functions: $S \subset \mathfrak{B}$ and $\varphi = \mathfrak{b}$ on S . Since $\mathfrak{B} \subset \mathfrak{P}$ and $\mathfrak{b} = \mathfrak{p}$ on $\mathfrak{B}$ we also have $S \subset \mathfrak{P}$ with $\varphi = \mathfrak{p}$ on S .

5.9.2. The following is used in this section. The norm dual of E is equal to the order dual, i.e., $E' = E^{\sim}$. Moreover, for $x, y \in E$ (see de Jonge and van Rooij [55, Theorem 10.2])

$$x \leq y \quad \Longleftrightarrow \quad \alpha(x) \leq \alpha(y) \text{ for all } \alpha \in E^{\sim+}. \quad (5.87)$$

This implies that for a sequence $(y_n)_{n \in \mathbb{N}}$ and x, y in E :

$$\inf_{n \in \mathbb{N}} \alpha(y_n) = 0 \text{ for all } \alpha \in E^{\sim+} \quad \implies \quad \inf_{n \in \mathbb{N}} y_n = 0. \quad (5.88)$$

5.9.3 Theorem. *Let $f \in \mathfrak{P}^+$ and f be partially in S . Then $f \in S_L^+$ and $\mathfrak{p}(f) = \varphi_L(f)$.*

Proof. Let $(A_n)_{n \in \mathbb{N}}$ be a partition for which $f_n := f \mathbb{1}_{A_n} \in S$. Then for every $\alpha \in E^{\sim+}$

$$\alpha(\mathfrak{p}(f)) = \int \alpha \circ f \, d\mu = \sum_{n \in \mathbb{N}} \int \alpha \circ f_n \, d\mu = \sum_{n \in \mathbb{N}} \alpha(\varphi(f_n)). \quad (5.89)$$

Hence $\inf_{N \in \mathbb{N}} \alpha(\mathfrak{p}(f) - \sum_{n=1}^N \varphi(f_n)) = 0$ and thus $\mathfrak{p}(f) = \sum_n \varphi(f_n)$ (see (5.88)). $\square$

5.9.4 Theorem. *Let $f \in \mathfrak{P}$. Then the following holds.*

(a) *If $g \in S_{LV}$ and $f \leq g$, then $\mathfrak{p}(f) \leq \varphi_{LV}(g)$.*

(b) *If S_V is stable, $g \in S_{VLV}$ and $f \leq g$, then $\mathfrak{p}(f) \leq \varphi_{VLV}(g)$.*

Consequently, $\mathfrak{p} = \varphi_{LV}$ on $\mathfrak{P} \cap S_{LV}$, and $\mathfrak{p} = \varphi_{VLV}$ on $\mathfrak{P} \cap S_{VLV}$ if S_V is stable.

The statements in (a) and (b) remain valid by replacing all “ $\leq$ ” by “ $\geq$ ”.

Proof. It will be clear that if $g \in S$ and $f \leq g$, then $g \in \mathfrak{P}$ and hence $\mathfrak{p}(f) \leq \mathfrak{p}(g) = \varphi(g)$.

If $g \in S_V$ and $f \leq g$, then there exists an $\Upsilon \subset S$ with $g \leq \Upsilon$ and $\varphi_V(g) = \inf \varphi(\Upsilon) = \inf \mathfrak{p}(\Upsilon) \geq \mathfrak{p}(f)$.

Let $g \in S_L$ and assume $f \leq g$. Let $g_1, g_2 \in S_L^+$ be such that $g = g_1 - g_2$. Let $(B_i)_{i \in \mathbb{N}}$ be a φ -partition for both g_1 and g_2 . Write $A_n = \bigcup_{i=1}^n B_i$ for $n \in \mathbb{N}$. Let $\alpha \in E^{\sim+}$. $\alpha \circ (f \mathbb{1}_A) = (\alpha \circ f) \mathbb{1}_A$ for every $A \in \mathcal{A}$, so that $\alpha \circ (f \mathbb{1}_A)$ is integrable. Thus, for $n \in \mathbb{N}$ we have

$$\begin{aligned} \int (\alpha \circ f) \mathbb{1}_{A_n} \, d\mu &= \int \alpha \circ (f \mathbb{1}_{A_n}) \, d\mu \leq \int \alpha \circ (g \mathbb{1}_{A_n}) \, d\mu \\ &= \int \alpha \circ g_1 \mathbb{1}_{A_n} \, d\mu - \int \alpha \circ g_2 \mathbb{1}_{A_n} \, d\mu \\ &= \alpha(\varphi(g_1 \mathbb{1}_{A_n})) - \alpha(\varphi(g_2 \mathbb{1}_{A_n})) \\ &\leq \alpha(\varphi(g_1 \mathbb{1}_{A_m})) - \alpha(\varphi(g_2 \mathbb{1}_{A_k})) \quad (k, m \in \mathbb{N}, k < n < m). \end{aligned} \quad (5.90)$$

Which implies that $\int (\alpha \circ f) \mathbb{1}_{A_n} \, d\mu + \alpha(\varphi(g_2 \mathbb{1}_{A_k})) \leq \alpha(\varphi_L(g_1))$ as soon as $k < n$. By letting n tend to ∞ (as $\int (\alpha \circ f) \mathbb{1}_{A_n} \, d\mu \rightarrow \int \alpha \circ f \, d\mu = \alpha(\mathfrak{p}(f))$), for each $k \in \mathbb{N}$ we obtain

$$\alpha(\mathfrak{p}(f)) \leq \alpha(\varphi_L(g_1) - \varphi(g_2 \mathbb{1}_{A_k})). \quad (5.91)$$

This holds for all $\alpha \in E^{\sim+}$, so

$$\mathfrak{p}(f) \leq \varphi_L(g_1) - \varphi(g_2 \mathbb{1}_{A_k}). \quad (5.92)$$

This, in term is true for every k , so $\mathfrak{p}(f) \leq \varphi_L(g)$.

We leave it to check that the preceding lines can be repeated with S_V , S_L or S_{VL} instead of S . $\square$

5.9.5 Theorem. Suppose $\|\cdot\|$ is σ -order continuous. Write $\bar{S} = S_{LV} = S_{VL}$ and $\bar{\varphi} = \varphi_{LV} = \varphi_{VL}$ (see Theorem 5.5.12).

- (a) Then $\bar{S} \subset \mathfrak{P}$. Consequently, if f is essentially separably valued and in $\bar{S}$, then $f \in \mathfrak{B}$. In particular, $S_L \subset \mathfrak{B}$.
- (b) Suppose there exists an $\alpha \in E_c^+$ with the property that if $b \in E$ and $b > 0$, then $\alpha(b) > 0$. Then $\mathfrak{B}_V \subset \mathfrak{B}$. Consequently, $\bar{S} \subset \mathfrak{B}$.

Proof. (a) Because $\|\cdot\|$ is σ -order continuous, $E' = E_c^\sim$. Therefore Theorem 5.8.14 implies that $\bar{S} \subset \mathfrak{P}$.

Note that $S_L \subset \mathfrak{B}$. Since $\mathfrak{B}$ is a Riesz ideal in the space of strongly measurable functions $X \rightarrow E$, an $f \in \bar{S}$ is an element of $\mathfrak{B}$ if it is essentially separably valued, since there are elements $\sigma, \tau \in S_L$ with $\sigma \leq f \leq \tau$ and f is weakly measurable since $f \in \mathfrak{P}$.

(b) Suppose $f \in \mathfrak{B}_V$ and $\sigma_n, \tau_n \in \mathfrak{B}$ are such that $\sigma_n \leq f \leq \tau_n$ for $n \in \mathbb{N}$, $\sigma_n \uparrow, \tau_n \downarrow$ and $\sup_{n \in \mathbb{N}} \mathfrak{b}(\sigma_n) = \mathfrak{b}_V(f) = \inf_{n \in \mathbb{N}} \mathfrak{b}(\tau_n)$. Then $\inf_{n \in \mathbb{N}} \int \alpha \circ (\tau_n - \sigma_n) d\mu = \alpha(\inf_{n \in \mathbb{N}} \mathfrak{b}(\tau_n - \sigma_n)) = 0$ and therefore $\alpha(\inf_{n \in \mathbb{N}} (\tau_n - \sigma_n)) = \inf_{n \in \mathbb{N}} \alpha \circ (\tau_n - \sigma_n)$ is integrable with integral equal to zero. Therefore $\inf_{n \in \mathbb{N}} (\tau_n - \sigma_n) = 0$ a.e., hence $\tau_n \rightarrow f$ a.e.. Therefore f is strongly measurable and thus $f \in \mathfrak{B}$ by (a). By (a) $S_L \subset \mathfrak{B}$, hence $\bar{S} = S_{LV} \subset \mathfrak{B}$. $\square$

5.9.6. For the next theorem we write $S_{\mathbb{R}}$ for the space of simple functions $X \rightarrow \mathbb{R}$. Note that if $u \in E^+$ and π is an element of $(S_{\mathbb{R}})_L^+$, thus of $\mathcal{L}^1(\mu)$, then $u\pi \in S_L$ and $\varphi_L(u\pi) = u \int \pi d\mu$.

5.9.7 Theorem. Let Y be a compact Hausdorff space and let $E = C(Y)$ be equipped with the supremum norm, $\|\cdot\|_\infty$. Then $\mathfrak{B} \subset S_{LV}$ and $\varphi_{LV} = \mathfrak{b}$ on $\mathfrak{B}$.

Proof. Let $f \in \mathfrak{B}$ and $(s_n)_{n \in \mathbb{N}}$ be a sequence of simple functions $X \rightarrow E$ such that $\int \|f - s_n\|_\infty d\mu \leq \frac{1}{n2^n}$ for all n . Then there exists an integrable function $g : X \rightarrow [0, \infty)$ with $g \geq \sum_{n \in \mathbb{N}} n \|f - s_n\|_\infty \mu$ -a.e., and thus $g \geq n \|f - s_n\|_\infty \mu$ -a.e.. By Theorem 5.7.1 there exists a $\pi \in (S_{\mathbb{R}})_L^+$ (see 5.9.6) with $\pi \geq g$. Then, let $Z \in \mathcal{A}$ with $\mu(A) = 0$ be such that $\|f(x) - s_n(x)\|_\infty \leq \frac{1}{n} \pi(x)$ for all $x \in X \setminus Z$. We may, by replacing s_n by $s_n \mathbb{1}_{X \setminus Z}$, assume that $s_n = 0$ on Z . With $Z_k = \{x \in Z : k-1 < \|f(x)\|_\infty \leq k\}$ and ρ the element in S_L^+ with $\rho(x) = 0$ for $x \notin Z$ and $\rho(x) = k \mathbb{1}_Y$ for $x \in Z_k$, we have

$$s_n - \frac{1}{n} \mathbb{1}_Y \pi - \rho \leq f \leq s_n + \frac{1}{n} \mathbb{1}_Y \pi + \rho \quad (n \in \mathbb{N}). \quad (5.93)$$

Because $\varphi_L(\rho) = 0$, $\varphi_L(\frac{1}{n} \mathbb{1}_Y \pi) \rightarrow 0$, $\varphi_L(s_n) = \mathfrak{b}(s_n)$ and $\mathfrak{b}(s_n) \rightarrow \mathfrak{b}(f)$, both $\varphi_L(s_n - \frac{1}{n} \mathbb{1}_Y \pi - \rho)$ and $\varphi_L(s_n + \frac{1}{n} \mathbb{1}_Y \pi + \rho)$ converge to $\mathfrak{b}(f)$. Whence

$$\mathfrak{b}(f) = \sup_{n \in \mathbb{N}} \varphi_L(s_n - \frac{1}{n} \mathbb{1}_Y \pi - \rho) = \inf_{n \in \mathbb{N}} \varphi_L(s_n + \frac{1}{n} \mathbb{1}_Y \pi + \rho). \quad (5.94)$$

Thus $f \in S_{LV}$ and $\varphi_{LV}(f) = \mathfrak{b}(f)$. $\square$

By the Yosida Representation Theorem the following is an immediate consequence.

5.9.8 Corollary. Let E be a Archimedean Riesz space with strong unit u and assume E is uniformly complete, i.e., E is a Banach lattice under the norm $\|\cdot\|_u$ given by $\|x\|_u = \inf\{\lambda \in [0, \infty) : |x| \leq \lambda u\}$. Then $\mathfrak{B} \subset S_{LV}$ and $\varphi_{LV} = \mathfrak{b}$ on $\mathfrak{B}$.

5.9.9 Examples. (I) Take $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, and let μ be the counting measure. We have $S = c_{00}[E]$; $S_V = c_{00}[E]$; all functions $\mathbb{N} \rightarrow E$ are partially in S ; $\bar{S} := S_{LV} = S_{VL} = S_L$ (see Theorem 5.5.5(c)) and $\bar{S}^+$ consists precisely of the functions $f : \mathbb{N} \rightarrow E^+$ for which $\sum_n f(n)$ exists in the sense of the ordering. On the other hand, $f : \mathbb{N} \rightarrow E$ is Bochner integrable if and only if $\sum_{n=1}^{\infty} \|f(n)\| < \infty$.

- If $\|\cdot\|$ is a σ -order continuous norm, then $\mathfrak{B} \subset \bar{S}$.

- Moreover $\|\cdot\|$ is equivalent to an abstract L-norm if and only if $\mathfrak{B} = \bar{S}$ (since, if $\mathfrak{B} = \bar{S}$, the following holds: if $x_1, x_2, \dots \in E^+$ and $\sum_n x_n$ exists, then $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$, see Theorem 5.A.1).

- For $E = c_0$ there exists an $f \in \mathfrak{B}$ that is not in $\bar{S}$. For example $f : \mathbb{N} \rightarrow c_0$ given by

$$f = (e_1, -e_1, e_2, -e_2, e_3, -e_3, \dots) \quad (5.95)$$

is Pettis integrable since $c'_0 \cong \ell^1$ has basis $\{\delta_n : n \in \mathbb{N}\}$ where $\delta_n(x) = x(n)$ and $\sum_{m \in \mathbb{N}} \delta_n(f(m)) = 0$ for all $m \in \mathbb{N}$. c_0 is σ -Dedekind complete and thus by Theorem 5.4.32 the set $\bar{S}$ is a Riesz space. However, $|f|$ is not in $\bar{S}$ and therefore neither f is.

- For $E = c$ there exists an $f \in \bar{S}$ that is not in $\mathfrak{B}$ and not in $\mathfrak{P}$: Consider for example $f : n \mapsto e_n$. It is an element of $\bar{S}$ but not of $\mathfrak{B}$. It is not even Pettis integrable. (Suppose it is, and its integral is a . Then for all $u \in c'$ we have $u(a) = \int u \circ f \, d\mu = \sum_{n=1}^{\infty} u(f(n)) = \sum_{n=1}^{\infty} u(e_n)$. Letting u be the coordinate functions, we see that $a(n) = 1$ for all $n \in \mathbb{N}$; letting u be $x \mapsto \lim_{n \rightarrow \infty} x(n)$ we have a contradiction.)

(II) $\mathfrak{B} \not\subset S_{VLV}$. Let $(\mathbb{R}, \mathcal{M}, \lambda)$ be the Lebesgue measure space. Let E be the σ -Dedekind complete Riesz space $L^1(\lambda)$. Let $g \in L^1(\lambda)$ be the equivalence class of the function that equals $t^{-\frac{1}{2}}$ for $0 < t \leq 1$ and equals 0 for other t . Let $L_x g(t) = g(t - x)$ for $x \in \mathbb{R}$. Then the function $f : \mathbb{R} \rightarrow L^1(\lambda)$ for which $f(x) = \mathbb{1}_{[0,1]}(x) L_x g$ is Bochner integrable (f is continuous in the $\|\cdot\|_1$ norm (because $\|L_\varepsilon g - g\|_1 = 2\sqrt{\varepsilon}$ for $\varepsilon > 0$) and $\int \|f(x)\|_1 \, d\lambda(x) = \int \mathbb{1}_{[0,1]}(x) \int |g(t - x)| \, d\lambda(t) \, d\lambda(x) = \|g\|_1 < \infty$) but no element of S_{VLV} (by Theorem 5.8.3).

§5.10 Extensions of Bochner integrable functions

Consider the situation of Section 5.9.

As we have seen in Examples 5.9.9, e.g., (5.95), the set of Pettis integrable functions need not be stable. We show that $\mathfrak{B}$ is stable and $\mathfrak{b}$ is laterally extendable. Furthermore we give an example of an $f \in \mathfrak{B}_{LV}$ that is neither in S_{VLV} , nor in $\mathfrak{B}_L$ or $\mathfrak{B}_V$.

5.10.1 Theorem. $\mathfrak{B}$ is stable and $\mathfrak{b}$ is laterally extendable.

Proof. Note that $f \mathbb{1}_B \in \mathfrak{B}$ for all $f \in \mathfrak{B}$ and $B \in \mathcal{A}$ (since $f \mathbb{1}_B$ is strongly measurable and $\|f \mathbb{1}_B\|$ is integrable), i.e., $\mathfrak{B}$ is stable. Let $(A_n)_{n \in \mathbb{N}}$ be a partition in $\mathcal{A}$ of X . Let $f : X \rightarrow E^+$ be a Bochner integrable function. Then $\int \|f\| \, d\mu < \infty$ and with $B_n = A_1 \cup \dots \cup A_n$ and Lebesgue's Dominated Convergence Theorem we obtain

$$\left\| \mathfrak{b}\left(f - \sum_{n=1}^N f \mathbb{1}_{A_n}\right) \right\| \leq \int \|f(x) - \mathbb{1}_{B_N}(x) f(x)\| \, d\mu(x) \rightarrow 0. \quad (5.96)$$

Thus

$$\mathfrak{b}(f) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \mathfrak{b}(f \mathbb{1}_{A_n}) = \sum_n \mathfrak{b}(f \mathbb{1}_{A_n}). \quad (5.97)$$

We conclude that $\mathfrak{b}$ is laterally extendable. $\square$

In the following situation we have $\mathfrak{B}_L = \mathfrak{B} = \mathfrak{B}_V$.

5.10.2 Lemma. *Let E be a Banach lattice with an abstract L -norm (i.e., $\|a+b\| = \|a\| + \|b\|$ for $a, b \in E^+$).*

(a) *Then*

$$\|\mathfrak{b}(f)\| = \int \|f\| \, d\mu \quad (f \in \mathfrak{B}^+). \quad (5.98)$$

(b) $\mathfrak{B}_L = \mathfrak{B}$.

(c) *There exists an $\alpha \in E_c^{\sim+}$ as in Theorem 5.9.5(b). Consequently, $\mathfrak{B}_V = \mathfrak{B}$.*

Proof. (a) It is clear that $\|\mathfrak{b}(f)\| = \int \|f\| \, d\mu$ for $f \in S^+$, hence by limits for all $f \in \mathfrak{B}^+$.

(b) Suppose $f \in \mathfrak{B}_L^+$. Let $(A_n)_{n \in \mathbb{N}}$ be a $\mathfrak{b}$ -partition for f , write $f_n = f \mathbb{1}_{A_n}$. Then $\|\sum_{n=1}^N f_n - f\| \rightarrow 0$, hence f is strongly measurable. Moreover, since $\|\cdot\|$ is σ -order continuous $\|\sum_{n=1}^N \mathfrak{b}(f_n) - \mathfrak{b}_L(f)\| \rightarrow 0$, hence $\sum_{n=1}^N \|\mathfrak{b}(f_n)\| \rightarrow \mathfrak{b}_L(f)$. Using (a) we obtain $\int \|f\| \, d\mu = \sum_{n \in \mathbb{N}} \int \|f_n\| \, d\mu = \sum_{n \in \mathbb{N}} \|\mathfrak{b}(f_n)\| < \infty$, i.e., $f \in \mathfrak{B}$.

(c) Extend $\alpha : E^+ \rightarrow \mathbb{R}$ given by $\alpha(b) = \|b\|$ to a linear map on E . $\square$

5.10.3. Consider the situation of Example 5.8.8. Since $S \subset \mathfrak{B}$ and $\varphi(h) = \mathfrak{b}(h)$ for $h \in S$: $f \in \mathfrak{B}_V$. The function f is not essentially separably-valued (i.e., $f(X \setminus A)$ is not separable for all null sets $A \in \mathcal{A}$), hence f (and thus g) is not strongly measurable (see [48, Theorem 3.5.2]). Hence f is not Bochner integrable, i.e., $f \in \mathfrak{B}_V$ but $f \notin \mathfrak{B}$.

In a similar way as has been shown in Example 5.8.8, one can show that $g : \mathbb{R} \rightarrow E^+$ defined by $g(t) = \mathbb{1}_{\{t\}}$ for $t \in \mathbb{R}$ is in S_{LV} . Then $g \in \mathfrak{B}_{LV}$ but $g \notin \mathfrak{B}_V$.

5.10.4. All $f \in \mathfrak{B}_L$ are strongly measurable. Therefore for $f \in \mathfrak{B}_L$ we have $f \notin \mathfrak{B}$ if and only if $\int \|f\| \, d\mu = \infty$.

The following example illustrates that by extending the Bochner integrable functions one can obtain more than by extending the simple functions.

5.10.5 Example. $[\psi \in \mathfrak{B}_V, \psi \notin \mathfrak{B}]$

Let $X = [2, 3]$, let $\mathcal{A}$ be the set of Lebesgue measurable subsets of X and μ be the Lebesgue measure on X . Let M denote the set of equivalence classes of measurable functions $\mathbb{R} \rightarrow \mathbb{R}$. Let

$$E = \left\{ f \in M : \sup_{x \in \mathbb{R}} \int_x^{x+1} |f| < \infty \right\}, \quad \|\cdot\| : E \rightarrow [0, \infty), \quad \|f\| = \sup_{x \in \mathbb{R}} \int_x^{x+1} |f|. \quad (5.99)$$

Then E equipped with the norm $\|\cdot\|$ is a Banach lattice. E is an ideal in M and therefore σ -Dedekind complete (hence S_V is stable; 5.4.25). The norm $\|\cdot\|$ is not σ -order continuous.

For $a \in \mathbb{R}$, $c > 0$ define $S_{a,c} : X \rightarrow E^+$ by $S_{a,c}(x) = \mathbb{1}_{(a+cx, \infty)}$. If $x, y \in X$ with $y > x$ then $\|S_{a,c}(x) - S_{a,c}(y)\| \leq \|\mathbb{1}_{(a+cx, a+cy]}\| \leq c|x - y|$, so $S_{a,c}$ is continuous and therefore strongly measurable. Furthermore $\|S_{a,c}(x)\| = 1$ for all $x \in X$, i.e., $x \mapsto \|S_{a,c}(x)\|$ is integrable. Thus $S_{a,c}$ is Bochner integrable. For $d, e \in \mathbb{R}$ with $e > d$ the map $E \rightarrow \mathbb{R}$, $f \mapsto \int_d^e f$ is a continuous linear functional. Therefore

$$\int_d^e \mathbf{b}(S_{a,c}) = \int_X \int_d^e (S_{a,c}(x))(t) \, dt \, dx = \int_d^e \int_X (S_{a,c}(x))(t) \, dx \, dt. \quad (5.100)$$

Since this holds for all $d, e \in \mathbb{R}$ with $e > d$, for $t \in \mathbb{R}$ we have

$$(\mathbf{b}(S_{a,c}))(t) = \int_X (S_{a,c}(x))(t) \, dx = \int_2^3 \mathbb{1}_{(a+cx, \infty)}(t) \, dx = \left(\frac{t-a}{c} \wedge 3 - 2\right) \vee 0. \quad (5.101)$$

For $k \in \mathbb{N}$ define $r_k, R_k : X \rightarrow E$ by

$$R_k := S_{0,k}, \quad r_k := S_{0,k} - S_{1,k}. \quad (5.102)$$

For $x \in X$ and $k \in \mathbb{N}$, $r_k(x) = \mathbb{1}_{(kx, kx+1]}$ and $kx + 1 < (k+1)x$. Define

$$\psi(x) := \mathbb{1}_{\bigcup_{k \in \mathbb{N}} (kx, kx+1]} = \sum_{k \in \mathbb{N}} r_k(x), \quad \sigma_n := \sum_{k=1}^n r_k, \quad \tau_n := \sum_{k=1}^n r_k + R_{n+1}. \quad (5.103)$$

Note that $\sigma_n \leq \psi \leq \tau_n$ and $\sigma_n, \tau_n \in \mathfrak{B}$ all for $n \in \mathbb{N}$. Since E is σ -Dedekind complete and therefore mediated, from the fact that

$$\inf_{n \in \mathbb{N}} \mathbf{b}(\tau_n - \sigma_n) = \inf_{n \in \mathbb{N}} \mathbf{b}(R_{n+1}) = 0, \quad (5.104)$$

it follows that $\psi \in \mathfrak{B}_V$. However, $\psi \notin \mathfrak{B}$ since ψ is not essentially separably valued: Let $x, y \in X$, $x < y$. We prove $\|\psi(x) - \psi(y)\| \geq 1$. For $k \in \mathbb{N}$:

$$\begin{aligned} k-1 \leq \frac{1}{y-x} < k &\implies \begin{cases} 1 + (k-1)y \leq kx, \\ 1 + kx < ky, \end{cases} \\ &\implies (kx, kx+1] \cap \bigcup_{i \in \mathbb{N}} (iy, iy+1] = \emptyset. \end{aligned} \quad (5.105)$$

Hence $\|\psi(x) - \psi(y)\| \geq 1$ for all $x, y \in X$ with $x \neq y$.

So ψ is an element of $\mathfrak{B}_V$ but not of $\mathfrak{B}$ (and neither of $\mathfrak{B}_L$).

5.10.6 Example. $[f \in \mathfrak{B}_{LV}, f \notin \mathfrak{B}_L, f \notin \mathfrak{B}_V, f \notin S_{VLV}]$

Let $(X, \mathcal{A}, \mu)$ be the Lebesgue measure space $(\mathbb{R}, \mathcal{M}, \lambda)$. Let E and ψ be as in Example 5.10.5. Define $u : \mathbb{R} \rightarrow E$ by

$$u(x) = \begin{cases} \psi(x) & x \in [2, 3], \\ 0 & \text{otherwise.} \end{cases} \quad (5.106)$$

Then u is an element of $\mathfrak{B}_V$ and not of $\mathfrak{B}_L$. As we have seen in Examples 5.9.9(II) there exists a g in $L^1(\lambda)$ and thus in E such that $v : x \mapsto \mathbb{1}_{[0,1]}(x)L_x g$ is an element of $\mathfrak{B}$ that is not an element of S_{VLV} . Furthermore $w : \mathbb{R} \rightarrow E$ given by $w(x) = \mathbb{1}_{(n,n+1]}$ for $x \in (n, n+1]$ is an element of $\mathfrak{B}_L$ and not of $\mathfrak{B}_V$. Therefore $f = u + v + w$ is an element of $\mathfrak{B}_{LV}$ (and thus of $\mathfrak{B}_{VL}$; see Theorem 5.5.8) but is neither an element of S_{VLV} nor of $\mathfrak{B}_V$ or $\mathfrak{B}_L$.

§5.11 Discussion

Of course, to some extent our approach is arbitrary. We mention some alternatives, with comments.

5.11.1. The reader may have wondered why in our definition of the lateral extension the sets A_n are required not only to be disjoint but also to cover X (i.e., to form a partition). Without the covering of X the definition remains perfectly meaningful, but the sum of two positive laterally integrable functions need not be laterally integrable, even in quite natural situations. (E.g., take $E = F = \mathbb{R}$ and $X = [0, 1]$; let $\mathcal{I}$ be the ring generated by the open intervals, Γ the space of all Riemann integrable functions on $[0, 1]$, and φ the Riemann integral. If f is the indicator of the Cantor set, then $\mathbb{1} - f$ is laterally integrable but $2\mathbb{1} - f$ is not.)

5.11.2. For the vertical extension we have, somewhat artificially, introduced a countability restriction leading us from φ_v to φ_V ; see Definition 5.3.3. In some sense, φ_v would have served as well as φ_V . In order to get a non-void theory, however, we would need a much stronger (but analogous) condition than “mediatedness”, restricting our world drastically.

5.11.3. A different approach to both the vertical and the lateral extension, closer to Daniell and Bourbaki, could run as follows. Starting from the situation of 5.3.14, call a function $X \rightarrow F^+$ “integrable” if there exist $f_1, f_2, \dots \in \Gamma^+$ such that

$$\begin{cases} f_n \uparrow f \text{ in } F^X, \\ \sup_{n \in \mathbb{N}} \varphi(f_n) \text{ exists in } E, \end{cases} \quad (5.107)$$

then define the “integral” $\overline{\varphi}(f)$ of f by

$$\overline{\varphi}(f) := \sup_{n \in \mathbb{N}} \varphi(f_n). \quad (5.108)$$

This definition is meaningful only if, in the above situation

$$g \in \Gamma^+, g \leq f, \implies \varphi(g) \leq \sup_{n \in \mathbb{N}} \varphi(f_n) \quad (5.109)$$

which in a natural way leads to the requirement that Γ be a lattice and that φ be continuous in the following sense:

$$h_1, h_2, \dots \in \Gamma^+, h_n \downarrow 0 \implies \varphi(h_n) \downarrow 0. \quad (5.110)$$

These conditions lead to a sensible theory, but again we consider them as too restrictive. (See Example II.2.4 in the thesis of G. Jeurink [54] for an example of a Γ that consists of simple functions on a measure space with values in a $C(X)$ for which (5.110) does not hold for the standard integral on simple functions (see 5.4.33).)

Acknowledgements

The authors are grateful to O. van Gaans for valuable discussions. W.B. van Zuijlen is supported by ERC Advanced Grant VARIS-267356.

Appendix

§5.A Appendix

5.A.1 Theorem. *Let E be a Banach lattice with the property*

$$\text{If } x_1, x_2, \dots \in E^+ \text{ and } \sum_n x_n \text{ exists, then } \sum_{n \in \mathbb{N}} \|x_n\| < \infty. \quad (5.111)$$

Then the norm $\|\cdot\|$ is equivalent to an L -norm.

The proof uses the following lemma.

5.A.2 Lemma. *Let E be a Banach lattice that satisfies (5.111). Then there exists a $C > 0$ such that*

$$x_1, x_2, \dots \in E^+, \sum_n x_n \text{ exists} \implies \sum_{n \in \mathbb{N}} \|x_n\| \leq C \left\| \sum_n x_n \right\|. \quad (5.112)$$

Proof. Suppose not. For $i \in \mathbb{N}$ let $x_{i1}, x_{i2}, \dots \in E^+$, $\sum_n x_{in} = b_i$ and $\sum_{n \in \mathbb{N}} \|x_{in}\| > 2^i \|b_i\|$ and $\|b_i\| = 2^{-i}$. Then $\sum_{i \in \mathbb{N}} \|b_i\| < \infty$, so $\sum_i b_i$ exists. As $\sum_i b_i = \sum_i \sum_n x_{in}$, by (5.111) we get $\infty > \sum_{i \in \mathbb{N}} \sum_{n \in \mathbb{N}} \|x_{in}\| > \sum_{i \in \mathbb{N}} 2^i \|b_i\| = \infty$. $\square$

Proof of Theorem 5.A.1. By Lemma 5.A.2 we can define $p : E \rightarrow [0, \infty)$,

$$p(x) = \sup \left\{ \sum_{n \in \mathbb{N}} \|x_n\| : x_1, x_2, \dots \in E^+, \sum_n x_n \leq |x| \right\}, \quad (5.113)$$

obtaining $p(x) = p(|x|)$, $p(tx) = |t|p(x)$, $\|x\| \leq p(x) \leq C\|x\|$ for all $x \in E$, $t \in \mathbb{R}$ (with C as in Lemma 5.A.2) and $p(x) \leq p(y)$ for $x, y \in E^+$ with $x \leq y$.

Let $x, y \in E^+$; we prove $p(x + y) = p(x) + p(y)$.

- For $\varepsilon > 0$ choose $x_1, x_2, \dots, y_1, y_2, \dots \in E^+$, $\sum_n x_n \leq x$, $\sum_n y_n \leq y$, $\sum_{n \in \mathbb{N}} \|x_n\| \geq p(x) - \varepsilon$, $\sum_{n \in \mathbb{N}} \|y_n\| \geq p(y) - \varepsilon$. Considering the sequence $x_1, y_1, x_2, y_2, \dots$ we find $\sum_{n \in \mathbb{N}} (\|x_n\| + \|y_n\|) \leq p(x + y)$. Hence $p(x + y) \geq p(x) + p(y)$.

- On the other hand: Let $z_1, z_2, \dots \in E^+$, $\sum_n z_n \leq x + y$; we prove $\sum_{n \in \mathbb{N}} \|z_n\| \leq p(x) + p(y)$. Define u_n, v_n by

$$u_1 + \dots + u_n = (z_1 + \dots + z_n) \wedge x, \quad v_n = z_n - u_n \quad (n \in \mathbb{N}). \quad (5.114)$$

Then $(z_1 + \cdots + z_n) \wedge x - z_n = (z_1 + \cdots + z_n - z_n) \wedge (x - z_n) \leq (z_1 + \cdots + z_{n-1}) \wedge x$, implying $u_n - z_n \leq 0$; and $(z_1 + \cdots + z_n) \wedge x \geq (z_1 + \cdots + z_{n-1}) \wedge x$, implying $u_n \geq 0$. Thus

$$u_n \geq 0, v_n \geq 0 \quad (n \in \mathbb{N}), \quad (5.115)$$

$\sum_{n \in \mathbb{N}} \|u_n\| \leq \sum_{n \in \mathbb{N}} \|z_n\| < \infty$, so $\sum_n u_n$ exists; $\sum_n u_n \leq x$, and $\sum_{n \in \mathbb{N}} \|u_n\| \leq p(x)$. $\sum_{n \in \mathbb{N}} \|v_n\| \leq \sum_{n \in \mathbb{N}} \|z_n\| < \infty$, so $\sum_n v_n$ exists. For every $n \in \mathbb{N}$, $z_1 + \cdots + z_n \leq (z_1 + \cdots + z_n + y) \wedge (x + y) = (z_1 + \cdots + z_n) \wedge x + y = u_1 + \cdots + u_n + y$, so $v_1 + \cdots + v_n \leq y$; then $\sum_n v_n \leq y$ and $\sum_{n \in \mathbb{N}} \|v_n\| \leq p(y)$. Thus $\sum_{n \in \mathbb{N}} \|z_n\| \leq \sum_{n \in \mathbb{N}} \|u_n\| + \sum_{n \in \mathbb{N}} \|v_n\| \leq p(x) + p(y)$. $\square$

§5.B Miscellaneous: Additional ideas, remarks etc.

In this section we present additional material that has not been published in [80]. First, in Section 5.B.1, we make observations regarding Sections 5.3 – 5.10. In Section 5.B.2 we compare ‘mediated’ and ‘splitting’ to what we call ‘dividing’ and relate them to σ -Dedekind completeness, and we discuss a similar notion without a countability restriction and relate that to Dedekind completeness. In Section 5.B.3 we make some observations regarding convolution. In Section 5.B.4 we discuss improperly integrable functions.

§5.B.1 Observations and remarks regarding Sections 5.3 – 5.10

Observations regarding Section 5.3

In connection with the question whether the vertical extension of the space Γ is a Riesz space (see Example 5.3.8):

5.B.1. $\varphi(\Gamma)$ does not need to be a Riesz space of E , even if E, Ω are Riesz spaces and Γ is a Riesz subspace of Ω :

Consider $\Omega = \Gamma = \mathbb{R}^2$, $E = \mathbb{R}^{\mathbb{R}}$ and let $\varphi : \Gamma \rightarrow \mathbb{R}^{\mathbb{R}}$ be the map for which $\varphi(a, b)$ is the element of $\mathbb{R}^{\mathbb{R}}$ given by $x \mapsto ax^2 + b$ for $a, b \in \mathbb{R}$.

In Example 5.3.8 we gave E, Ω, Γ and φ for which Γ_V is not a Riesz space. Here, $\Omega = F^X$ for some set X and a partially ordered vector space F with $F \neq E$. In Example 5.B.2 we will also consider $E, F, \Omega, \Gamma, \varphi$ with $\Omega = F^X$ for some set X , but with $E = F$, and show that Γ_V is not a Riesz space even though E is a Riesz space and Γ is a Riesz subspace of E^X .

5.B.2 Example. Consider $E = c$, $\Omega = E^{\mathbb{N}}$ and let $\Gamma \subset \Omega$ be the following partially ordered vector space

$$\{(a_1, a_2, \dots, a_n, \lambda \mathbb{1}, \lambda \mathbb{1}, \dots) : n \in \mathbb{N}, a_1, \dots, a_n \in c, \lambda \in \mathbb{R}\}. \quad (5.116)$$

Let $\psi : \Gamma \rightarrow c$ (ψ takes the role of φ , the notation φ will be used in Example 5.B.3) be the map for which the k -th component of $\psi(f)$ equals

$$[\psi(f)](k) = \sum_{i=0}^{\infty} 2^{-i} f_{k+i}(k) \quad (k \in \mathbb{N}), \quad (5.117)$$

for $f = (f_1, f_2, f_3, \dots) \in \Gamma$. Let

$$g = (e_1, e_2, e_3, e_4, \dots), \quad h = (0, e_1 + 2e_2, 0, e_3 + 2e_4, \dots). \quad (5.118)$$

Upon defining

$$\tau_n = (f_1, f_2, \dots, f_n, 2\mathbb{1}, 2\mathbb{1}, \dots), \quad \sigma_n = (f_1, f_2, \dots, f_n, 0, 0, \dots) \quad (n \in \mathbb{N}), \quad (5.119)$$

one can calculate

$$\psi(\sigma_n)(k) = \begin{cases} f_k(k) + 2^{-1}f_{k+1}(k) + \dots + 2^{k-n}f_n(k) & k \leq n, \\ 0 & k > n, \end{cases} \quad (5.120)$$

$$\psi(\tau_n - \sigma_n)(k) = \begin{cases} 4 & k > n, \\ 2^{1+k-n} & k \leq n, \end{cases} \quad (5.121)$$

so that for $f = g$ or $f = h$, $\inf_{n \in \mathbb{N}} \psi(\tau_n) = \mathbb{1} = \sup_{n \in \mathbb{N}} \psi(\sigma_n)$. Hence both g, h lie in Γ_V . Let $f \in E^{\mathbb{N}}$ be defined by

$$f := g \vee h - g = (0, 2e_2, 0, 2e_4, 0, 2e_6, \dots). \quad (5.122)$$

Then for τ_n and σ_n as in (5.119), the sets $\{\psi(\tau_n) : n \in \mathbb{N}\}$ and $\{\psi(\sigma_n) : n \in \mathbb{N}\}$ do not have an infimum and a supremum in c , respectively. This implies (we leave it to the reader to check this) that f is not an element of Γ_V . Hence Γ_V is not a Riesz space.

As has been shown in Example 5.3.8 and has been mentioned in Remark 5.4.20, Γ_V need not be stable if Γ is stable and φ is laterally extendable, and Γ_V need not be a Riesz subspace of Ω even if Ω and E are Riesz spaces and Γ is a Riesz subspace of Ω . The Γ in Example 5.B.2 is not a set of simple functions. The following example illustrates that S_V need not be stable and need not be a Riesz subspace of E^X even if E is a Riesz space. Therefore, the stability assumption of S_V is not redundant. (Note that here we consider the space of simple functions S , instead of Γ as in Remark 5.4.20.)

5.B.3 Example. Consider the complete measure space $((0, 1], \mathcal{M}, \lambda)$, where $\mathcal{M}$ is the σ -algebra of Lebesgue measurable sets and λ the Lebesgue measure. Consider the Riesz space $E = c$. Let S and φ be as in 5.4.33 ($F = E$). We will show with the aid of Example 5.B.2 that S_V is not stable. Let $A_n = (2^{-n}, 2^{1-n}]$ for $n \in \mathbb{N}$, note that $\lambda(A_n) = 2^{-n}$. Let Γ and ψ be as in Example 5.B.2. Consider the map $\Theta : \Gamma \rightarrow \mathbb{R}^{[0,1]}$ given for $f = (f_1, f_2, f_3, \dots) \in \Gamma$ by

$$\Theta(f)(x) = 2^n f_n \quad (x \in A_n). \quad (5.123)$$

It is not difficult to check that $\Theta(\Gamma) \subset S$ and that $\psi(f) = \varphi(\Theta(f))$ for $f \in \Gamma$. By considering g and h as in (5.118), one can show in analogy with Example 5.B.2 that $\Theta(g)$ and $\Theta(h)$ are in S_{LV} but $\Theta(g) \vee \Theta(h) - \Theta(g)$ is not. Moreover, as $\Theta(g)$ is an element of S_V , the latter is not stable.

5.B.4 Remark. In 5.3.12 there is given a Banach lattice that has a σ -order continuous norm and is not σ -Dedekind complete. Another example is ℓ^∞/c_0 . We prove this in the following lemma.

5.B.5 Lemma.

- (a) *The quotient norm on ℓ^∞/c_0 is σ -order continuous.*
- (b) *ℓ^∞/c_0 is not σ -Dedekind complete.*

Proof. (a) Let $Q : \ell^\infty \rightarrow \ell^\infty/c_0$ be the quotient map. First we prove:

Let $F \subset \mathbb{N}$ be finite, $w \in (\ell^\infty)^+$, $\|Qw\| > 1$. Then there exists a $p \in \mathbb{N} \setminus F$ with $w(p) > 1$.

Indeed, $Q(w\mathbb{1}_F) = 0$ so $1 < \|Qw\| = \|Q(w\mathbb{1}_{\mathbb{N} \setminus F})\| \leq \|w\mathbb{1}_{\mathbb{N} \setminus F}\|_\infty = \sup_{p \in \mathbb{N} \setminus F} w(p)$. This proves the claim.

Now let $u_1 \geq u_2 \geq \dots$ be in $(\ell^\infty/c_0)^+$, $\|u_n\| > 1$ for all $n \in \mathbb{N}$, we prove there is an $a \in \ell^\infty/c_0$ with $0 < a \leq u_n$ for all $n \in \mathbb{N}$.

Choose $v_1, v_2, \dots \in \ell^\infty$, $Qv_n = u_n$. Define $w_n := |v_1| \wedge \dots \wedge |v_n|$. Then $w_1 \geq w_2 \geq \dots$ in $(\ell^\infty)^+$, $Qw_n = u_n$. With the proved statement above there exist $p_1 < p_2 < \dots$ in $\mathbb{N}$ with $w_n(p_n) > 1$ for all $n \in \mathbb{N}$. Set $b = \mathbb{1}_{\{p_1, p_2, \dots\}}$, $a = Qb$. Then $a > 0$ in ℓ^∞/c_0 . For every $n \in \mathbb{N}$: If $k \geq n$ then $w_n(p_k) \geq w_k(p_k) > 1$, so $w_n \geq \mathbb{1}_{\{p_n, p_{n+1}, \dots\}}$, $Qw_n \geq Qb$ and $u_n \geq a$ for all $n \in \mathbb{N}$.

(b) Let $A_1, A_2, \dots$ be disjoint infinite subsets of $\mathbb{N}$. For $n \in \mathbb{N}$ put $a_n = \mathbb{1}_{A_n}$. We prove that $\{Qa_n : n \in \mathbb{N}\}$ has no supremum in ℓ^∞/c_0 .

To this end, let $a \in (\ell^\infty)^+$ be such that $Qa \geq Qa_n$ for all n ; we make a $b \in \ell^\infty$ with $Qb < Qa$, $Qb \geq Qa_n$ for all n .

For every n we have $0 = (Qa_n - Qa)^+ = Q((a_n - a)^+)$, so $(a_n - a)^+ \in c_0$. Then there exist only finitely many $k \in \mathbb{N}$ with $a_n - a > \frac{1}{2}$. We can choose $p_n \in A_n$ with $(a_n - a)(p_n) \leq \frac{1}{2}$, i.e.,

$$a_n(p_n) = 1, \quad a(p_n) \geq \frac{1}{2}. \quad (5.124)$$

Form $P = \{p_n : n \in \mathbb{N}\}$ and define $b \in (\ell^\infty)^+$ by

$$b(k) = \begin{cases} 0 & k \in P, \\ a(k) & k \notin P. \end{cases} \quad (5.125)$$

Then $a - b \geq \frac{1}{2}\mathbb{1}_P$. As P is an infinite set it follows that $Qa - Qb \geq \frac{1}{2}Q\mathbb{1}_P > 0$ and $Qb < Qa$.

Take $n \in \mathbb{N}$. Then $a - b = a\mathbb{1}_P$, so that $(a - b) \wedge a_n$ is a scalar multiple of $\mathbb{1}_{\{p_n\}}$ and lies in c_0 . Hence, $0 = (Qa - Qb) \wedge Qa_n \geq (Qa_n - Qb) \wedge Qa_n = Qa_n - Qb$ and $Qb \geq a_n$. $\square$

The following theorem is an addendum to Remark 5.B.4 and Lemma 5.B.5.

5.B.6 Theorem. *Let X be a compact Hausdorff space. Then the following are equivalent:*

- (a) *The norm on $C(X)$ is σ -order continuous.*
- (b) *Every nonempty G_δ set has a nonempty interior.*
- (c) *$f \in C(X)$, $\{x \in X : f(x) = 0\} \neq \emptyset$ implies that the interior of $\{x \in X : f(x) = 0\}$ is not empty.*

Proof. (a) $\Rightarrow$ (b). Let Y be a nonempty G_δ set. Let $y \in Y$ and U_n be open sets with $Y = \bigcap_{n \in \mathbb{N}} U_n$. Choose $f_n \in C(X)^+$ with $f_n(y) = 1$ and $f_n = 0$ on $X \setminus U_n$. Let $g_n = f_1 \wedge \cdots \wedge f_n$. Then $g_n \in C(X)^+$, $g_n \downarrow$ and $\|g_n\|_\infty \geq g_n(y) = 1$. Therefore there exists a $g \in C(X)$ with $0 < g \leq g_n \leq f_n$ for all $n \in \mathbb{N}$. Then

$$\emptyset \neq \{x \in X : g(x) \neq 0\} \subset \bigcap_{n \in \mathbb{N}} \{x \in X : f_n(x) \neq 0\} \subset \bigcap_{n \in \mathbb{N}} U_n = Y. \quad (5.126)$$

(b) $\Rightarrow$ (c). As $\{x \in X : f(x) \neq 0\} = \bigcap_{n \in \mathbb{N}} \{x \in X : |f(x)| > \frac{1}{n}\}$, it is a G_δ set.

(c) $\Rightarrow$ (a). Let $f_n \downarrow$ in $C(X)^+$ and $\|f_n\|_\infty \geq 1$ for all $n \in \mathbb{N}$. Let $g_n = f_n \wedge \mathbb{1}$ for all $n \in \mathbb{N}$. Then $g_n \in C(X)$, $0 \leq g_n \leq \mathbb{1}$, $g_n \downarrow$, $\|g_n\|_\infty = 1$. By compactness there is a x_0 in $\bigcap_{n \in \mathbb{N}} \{x \in X : g_n(x) = 1\}$. Make $g := \mathbb{1} - \sum_{n \in \mathbb{N}} 2^{-n} g_n$; then $g(x_0) = 0$. There is a nonempty open set $U \subset \{x \in X : g(x) = 0\} \subset \{x \in X : g_n(x) = 1\} \subset \{x \in X : f_n(x) \geq 1\}$ for all $n \in \mathbb{N}$. Then $f_n \not\downarrow 0$. $\square$

5.B.7 Remark. With the $\mathcal{A}$ in Lemma 5.3.18 replaced by $\{A \subset X : f \mathbb{1}_A \in \Gamma_V \text{ for } f \in \Gamma\}$ the statement of Lemma 5.3.18 is still valid. However, for this set the continuity property in (5.15) would have to be reformulated as a continuity property of φ_V to obtain similar results as in Theorem 5.3.21.

5.B.8 Remark. As we stated in Remark 5.3.22, if $\mathcal{B} \subset \mathcal{A}$ is a σ -algebra and $(g_n)_{n \in \mathbb{N}}$ is a bounded sequence of bounded $\mathcal{B}$ -measurable functions that converges pointwise to g , then the condition of Theorem 5.3.21(b) is satisfied. For general sequences $(g_n)_{n \in \mathbb{N}}$ in $[\mathcal{A}]^o$ that converge pointwise to a $g \in [\mathcal{A}]^o$, this need not be the case:

- For $X = [0, 1]$ and $\mathcal{B}$ the Borel σ -algebra on X :

$$g_n(x) = \begin{cases} n & x \in [0, \frac{1}{n}], \\ 0 & \text{otherwise} \end{cases} \quad (5.127)$$

is an element in $[\mathcal{B}]^o$ for all $n \in \mathbb{N}$ and $g_n \rightarrow 0$ pointwise. However, there do not exist $(j_n)_{n \in \mathbb{N}}$ in $[\mathcal{B}]^{o+}$ satisfying the condition of Theorem 5.3.21(b).

- We present another example, now with a sequence $(g_n)_{n \in \mathbb{N}}$ that is bounded. Let X be the Cantor set, $\mathcal{A}$ be the algebra consisting of the clopen subsets of X . Choose $\{q_1, q_2, \dots\}$ dense in X , $q_n \neq q_m$ ($n \neq m$); then $\mathbb{1}_{\{q_n\}} \in [\mathcal{A}]^o$, $\mathbb{1}_{\{q_n\}} \rightarrow 0$.

Suppose $j_1, j_2, \dots \in [\mathcal{A}]^o$, $j_n \geq \mathbb{1}_{\{q_n\}}$, $j_n \downarrow 0$ pointwise. Then $j_m \geq \mathbb{1}_{\{q_n\}}$ if $m \leq n$, so $j_m \geq 1$ on $\{q_m, q_{m+1}, \dots\}$: $j_m \geq 1$ at every continuity point of j_m .

But j_m lies in the first Baire class, hence it is continuous at every point of some dense set ([78, Theorem 11.4]).

On the other hand, order convergence in $[\mathcal{A}]^o$ does not need to imply pointwise convergence, as the following example illustrates.

- $\mathcal{A} = \{A \subset [0, 1] : \text{there exists an } \varepsilon > 0 \text{ s.t. } [0, \varepsilon) \subset A \text{ or } [0, \varepsilon) \subset A^c\}$, $g_n = \mathbb{1}_{[0, \frac{1}{n}]}$, then $g_n \downarrow 0$ in $[\mathcal{A}]^o$ but $\mathbb{1}_{[0, \frac{1}{n}]}(0) = 1$ for all $n \in \mathbb{N}$.

5.B.9 Remark. As has been mentioned in 5.3.16 $[\mathcal{A}]^o$ is uniformly complete. We will give a proof of this.

5.B.10 Theorem. *For an algebra of subsets of X , $\mathcal{A}$, the space $[\mathcal{A}]^o$ is uniformly complete.*

Proof. Let $f_p \in [\mathcal{A}]^o$ for $p \in \mathbb{N}$, $|f_p| \leq 2^{-p} \mathbb{1}$, $f = \sum_{p \in \mathbb{N}} f_p$. We claim $f \in [\mathcal{A}]^o$. For each p , choose $s_{pn} \in [\mathcal{A}]$, $j_{pn} \in [\mathcal{A}]^+$, $|f_p - s_{pn}| \leq j_{pn}$ and $j_{p1} \geq j_{p2} \geq \dots \rightarrow 0$. We may choose them such that $|s_{pn}| \leq 2^{-p} \mathbb{1}$ and $0 \leq j_{pn} \leq 2^{-p+1} \mathbb{1}$. Put $s_N = \sum_{p \leq N} s_{pN}$ and $j_N = 2^{-N+1} \mathbb{1} + \sum_{p \leq N} j_{pN}$ for $N \in \mathbb{N}$, then $s_N \in [\mathcal{A}]$ and $j_N \in [\mathcal{A}]^+$. Then

- $|f - s_N| \leq |\sum_{p > N} f_p| + |\sum_{p \leq N} f - s_{pN}| \leq \sum_{p > N} 2^{-p} \mathbb{1} + \sum_{p \leq N} j_{pN} \leq j_N$.
- $j_N - j_{N+1} = 2^{-N} \mathbb{1} + \sum_{p \leq N} (j_{pN} - j_{p, N+1}) - j_{N+1, N+1} \geq 0$.
- Let $x \in X$, $\varepsilon > 0$. We prove $j_N(x) \leq \varepsilon$ for large N . Choose $K \in \mathbb{N}$ with $2^{-K+2} < \frac{\varepsilon}{3}$. Let N_0 be such that $2^{-N+1} < \frac{\varepsilon}{3}$ for $N \geq N_0$ and $\sum_{p \leq K} j_{pN}(x) < \frac{\varepsilon}{3}$; so that for $N \geq N_0$

$$\begin{aligned} j_N(x) &\leq 2^{-N+1} + \sum_{p \leq N} j_{pN}(x) \\ &< \frac{\varepsilon}{3} + \sum_{p \leq K} j_{pN}(x) + \sum_{p > K} j_{pN}(x) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \sum_{p > K} 2^{-p+1} < \varepsilon. \end{aligned} \tag{5.128}$$

Observations regarding Section 5.5

5.B.11 Remark. A minor remark on Theorem 5.5.8.

The necessity of the stability assumptions of Theorem 5.5.8 can be further investigated: There is not yet a positive answer to the following questions;

- Is the stability assumption that Γ_{LV} necessary to imply $\Gamma_{LV} \subset \Gamma_{VL}$?
- Is the stability assumption that $\bar{\Gamma}_V$ necessary to imply $\bar{\Gamma}_V = \bar{\Gamma}_V$?

The Examples 5.5.15 and 5.5.14 do not provide additional insight to this question.

Observations regarding Section 5.6

5.B.12 Examples. (I) Let $\Omega = \mathbb{R}^{[0,1]}$, $\Gamma = E = C[0, 1]$ and φ the identity $C[0, 1] \rightarrow C[0, 1]$. Let $\mathbb{D} \subset [0, 1]$ be the Cantor set. Then $\mathbb{1}_{\mathbb{D}} \in \Gamma_V$. With $E^\bullet = \mathbb{R}^{[0,1]}$: $\mathbb{1}_{\mathbb{D}} \notin \Gamma_V^\bullet$.

(II) Consider $E^\bullet = c \times \mathbb{R}$, $E = \{(x, \lambda) \in E^\bullet : \lambda = \lim_{n \rightarrow \infty} x(n)\}$. Let $X = \mathbb{N}$, $\mathcal{I} = \mathcal{P}(\mathbb{N})$, $\Gamma = c_{00}[E]$ and $\varphi = \sum_n$. Let $f : \mathbb{N} \rightarrow E$ be given by $f(n) = (e_n, 0)$. Then the set $\{(\sum_{n=1}^N e_n, 0) : N \in \mathbb{N}\}$ has supremum $(\mathbb{1}, 1)$ in E but supremum $(\mathbb{1}, 0)$ in $E^\bullet$. Hence, $f \in \Gamma_L$ and $f \in \Gamma_L^\bullet$ and $\varphi_L(f) = (\mathbb{1}, 1) \neq (\mathbb{1}, 0) = \varphi_L^\bullet(f)$.

(III) Let $E = C[0, 1]$ and $E^\bullet$ be the space of all bounded functions on $[0, 1]$ that are continuous at 0. We will show existence of an increasing sequence $(a_n)_{n \in \mathbb{N}}$ in E that has a supremum in E but not in $E^\bullet$: then with $X, \mathcal{I}, \Gamma, \varphi$ as in (I) the function $f = \sum_{n \in \mathbb{N}} (a_{n+1} - a_n) \mathbb{1}_{\{n\}}$ is in Γ_L but not in $\Gamma_L^\bullet$. Put $S = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ and let $(q_n)_{n \in \mathbb{N}}$ be an enumeration of $(\mathbb{Q} \cap (0, 1]) \setminus S$. For each $n \in \mathbb{N}$, choose u_n in $C[0, 1]^+$ such that $u_n \leq \mathbb{1}$, $u_n(q_n) = 1$, $u_n(s) = 0$ ($s \in S$), and let $a_n = u_1 \vee \dots \vee u_n$. The sequence a_n has supremum $\mathbb{1}$ in $C[0, 1]$. If it has a supremum, b , in $E^\bullet$, then $b(t) = \sup_{n \in \mathbb{N}} a_n(t) = \sup_{n \in \mathbb{N}} u_n(t)$ for $t \in (0, 1]$; so $b(q) = 1$ for $q \in (\mathbb{Q} \cap (0, 1]) \setminus S$, but $b(s) = 0$ for $s \in S$; then b cannot be continuous at 0.

Observations regarding Section 5.8

5.B.13. In the situation of Section 5.8: if $f \in S_L^+$ then $f(X) \subset E_a$, where E_a is the principal ideal of E generated by the integral $a = \varphi_L(f)$. Consequently, every $f \in S_{LV}^+$ has a range that is contained in a principal ideal of E (since f is dominated by a function in S_L).

In Section 5.8 we have considered the situation where $(X, \mathcal{A}, \mu)$ is a complete σ -finite measure space and $\mathcal{I} = \{A \in \mathcal{A} : \mu(A) < \infty\}$ and S and φ as in 5.4.33. However one could also consider $\mathcal{I}$ to be a subset of $\{A \in \mathcal{A} : \mu(A) < \infty\}$ and Γ to be a subset of S .

5.B.14 Example. [Extending the Riemann integral]

Let $(X, \mathcal{A}, \mu)$ be the Lebesgue measure space $(\mathbb{R}, \mathcal{M}, \lambda)$. Let $\mathcal{I}$, S and φ be as in 5.4.33.

Let

$$\mathcal{I}^* = \{\emptyset, \{a\}, [a, b], (a, b], [a, b), (a, b) : a, b \in \mathbb{R}, a < b\} \quad (5.129)$$

be the set of bounded intervals in $\mathbb{R}$. Note that $\mathcal{I}^*$ is closed under taking finite intersections and is a subset of $\mathcal{I} = \{A \in \mathcal{A} : \mu(A) < \infty\}$. In analogy to 5.4.33 let $S^* \subset S$ be the set of simple functions as in (5.40) with $A_1, \dots, A_n \in \mathcal{I}^*$. Then S_V^* is the set of Riemann integrable functions on $\mathbb{R}$ (the union of the sets of Riemann integrable functions on $[-n, n]$ for $n \in \mathbb{N}$). By Theorem 5.5.12 $S_{VL}^* = S_{LV}^*$. Let Θ be the smallest vector space that contains all functions f for which f^+ and f^- are improperly Riemann integrable. Because every improperly Riemann integrable function $f \geq 0$ is φ -partially in S_V^* , and $S_V^* \subset \Theta$ we have $\Theta_L = S_{VL}^* = S_{LV}^*$ by Theorem 5.4.25 and Theorem 5.5.12.

Observations regarding Section 5.9

5.B.15 Remark. A minor remark on Theorem 5.9.5.

The theorem states that the inclusion $\overline{S} \subset \mathfrak{P}$ holds in case E is a Banach lattice with a σ -continuous norm. Under an additional assumption one has $\overline{S} \subset \mathfrak{B}$. However, it is not clear whether one really needs to make an additional assumption, in the sense that we do not know whether there exists an example of an Banach lattice with

σ -order continuous norm E and an f which is in $\overline{S}$ but not in $\mathfrak{B}$. (Note that if such f exists, it cannot be essentially separably valued.)

Observations regarding Section 5.10

The following is another example of a function $f \in \mathfrak{B}_V$ with $f \notin \mathfrak{B}$. However, contrary to Example 5.10.6, the function f is an element of S_{LV} by Corollary 5.9.8.

5.B.16 Example. Let X be $\mathbb{R}$, $\mathcal{A}$ the Borel- σ -algebra of $\mathbb{R}$, and μ a finite measure on $\mathcal{A}$. For E we take the Banach lattice $L^\infty(\mathbb{R})$ (the space of equivalence classes of functions that are bounded almost everywhere).

If $u \in L^\infty(\mathbb{R})$, then for $x \in \mathbb{R}$ we let $L_x u$ be the function $t \mapsto u(t - x)$ ($t \in \mathbb{R}$), and we define $Lu : \mathbb{R} \rightarrow L^\infty(\mathbb{R})$ by

$$(Lu)(x) := L_x u \quad (x \in \mathbb{R}). \quad (5.130)$$

If $u \in C_0(\mathbb{R})$ (the space of continuous functions that converge to zero at $\pm\infty$), then Lu is continuous and therefore Bochner integrable with respect to μ . Its integral $\int L_x u \, d\mu(x)$ is the convolution product $\mu * u$.

Now let w be $\mathbb{1}_{[0,1]}$. The function $Lw : \mathbb{R} \rightarrow L^\infty(\mathbb{R})$ is not Bochner integrable, as it is not essentially separably valued because $\|(Lw)(s) - (Lw)(t)\|_\infty = \|L_s w - L_t w\|_\infty = 1$ as soon as $s, t \in \mathbb{R}$, $s \neq t$. We show that Lw lies in $\mathfrak{B}_V$. Note that $L^\infty(\mathbb{R})$ is mediated, so that we can use Lemma 5.3.7.

For $n \in \mathbb{N}$ choose $u_n, v_n \in C_0(\mathbb{R})$ with

$$\mathbb{1}_{[\frac{1}{n}, 1 - \frac{1}{n}]} \leq u_n \leq \mathbb{1}_{[0,1]} \leq v_n \leq \mathbb{1}_{[-\frac{1}{n}, 1 + \frac{1}{n}]} \quad (5.131)$$

Lu_n and Lv_n lie in $\mathfrak{B}$, and $Lu_n \leq Lw \leq Lv_n$. With the inequalities

$$0 \leq v_n - u_n \leq \mathbb{1}_{[-\frac{1}{n}, \frac{1}{n}]} + \mathbb{1}_{[1 - \frac{1}{n}, 1 + \frac{1}{n}]} \quad (5.132)$$

one obtains, for each $t \in \mathbb{R}$:

$$0 \leq (\mathfrak{b}(Lv_n) - \mathfrak{b}(Lu_n))(t) \leq \mu([t - 1 - \frac{1}{n}, t - 1 + \frac{1}{n}]) + \mu([t - \frac{1}{n}, t + \frac{1}{n}]), \quad (5.133)$$

so that

$$0 \leq \inf_{n \in \mathbb{N}} (\mathfrak{b}(Lv_n) - \mathfrak{b}(Lu_n))(t) \leq \mu(\{t - 1\}) + \mu(\{t\}). \quad (5.134)$$

Then

$$\inf_{n \in \mathbb{N}} (\mathfrak{b}(Lv_n) - \mathfrak{b}(Lu_n))(t) = 0 \quad (5.135)$$

for all but countably many values of t , and

$$\inf_{n \in \mathbb{N}} (\mathfrak{b}(Lv_n) - \mathfrak{b}(Lu_n)) = 0 \text{ in } L^\infty(\mathbb{R}). \quad (5.136)$$

Thus, $Lw \in \mathfrak{B}_V$. (Of course, $\mathfrak{b}_V(Lw)$ is just $\mu * w$.)

Observations regarding Section 5.11

We mentioned in 5.11.3 that Example II.2.4 in the thesis of G. Jeurnink [54] is an example of a Γ that consists of simple functions on a measure space with values in a $C(X)$ for which (5.110) does not hold for the standard integral on simple functions (see 5.4.33).

Theorem 5.B.17 is a rewritten version of Example II.2.4 in the thesis of G. Jeurnink [54]. In 5.B.18 we discuss the property that $C[0, 1]$ has by Theorem 5.B.17, and give a few examples of spaces that do or do not have that property.

5.B.17 Theorem. *Let μ be the Haar measure on $X = \{0, 1\}^{\mathbb{N}}$. Let E be the Riesz space $C(X)$ and φ be the integral as in 5.4.33. Then there exists a decreasing sequence $(s_n)_{n \in \mathbb{N}}$ of simple functions $X \rightarrow C(X)$ with*

$$\begin{cases} s_n(a) \downarrow 0 & \text{for all } a \in \{0, 1\}^{\mathbb{N}}, \\ \varphi(s_n) \geq \frac{1}{2} \mathbb{1} & \text{in } C(X) \text{ for all } n \in \mathbb{N}. \end{cases} \quad (5.137)$$

Proof. The following notations will be used in this example:

For $a_1, \dots, a_n \in \{0, 1\}$:

$$\langle a_1, \dots, a_n \rangle = \{x \in X : x_1 = a_1, \dots, x_n = a_n\}, \quad (5.138)$$

$$\langle a_1, \dots, a_n \rangle^* = \{x \in X : x_{n+1} = a_1, \dots, x_{2n} = a_n\}. \quad (5.139)$$

A set of the form $\langle a_1, \dots, a_n \rangle$ is called a cylinder. It has measure $\mu(\langle a_1, \dots, a_n \rangle) = 2^{-n}$. Notice that the set of cylinders forms a basis for the topology of X .

Observe also that for every $n \in \mathbb{N}$ the sets $\langle a_1, \dots, a_n \rangle^*$, where $a_1, \dots, a_n \in \{0, 1\}$, form a disjoint cover of X .

★ For $n \in \mathbb{N}$ define $t_n : X \rightarrow E$, by

$$(t_n(x))(y) = \begin{cases} 1 & \text{if } (y_{n+1}, \dots, y_{2n}) = (x_1, \dots, x_n), \\ 0 & \text{otherwise.} \end{cases}$$

Let $a_1, \dots, a_n \in \{0, 1\}$. Then $t_n(x) = \mathbb{1}_{\langle a_1, \dots, a_n \rangle^*}$ for all $x \in \langle a_1, \dots, a_n \rangle$. Thus t_n is a simple function with

$$\begin{aligned} \varphi(t_n) &= \sum_{a_1, \dots, a_n \in \{0, 1\}} \mu(\langle a_1, \dots, a_n \rangle) \mathbb{1}_{\langle a_1, \dots, a_n \rangle^*} \\ &= 2^{-n} \sum_{a_1, \dots, a_n \in \{0, 1\}} \mathbb{1}_{\langle a_1, \dots, a_n \rangle^*} = 2^{-n} \mathbb{1}_X. \end{aligned}$$

★ For $n \geq 2$ let s_n be the simple function given by $s_n = \mathbb{1} - (t_2 \vee \dots \vee t_n)$, where $\mathbb{1}$ denotes the function $X \rightarrow E$, $x \mapsto \mathbb{1}_X$. We have $s_n \downarrow$ and $s_n \geq 0$ for all $n \in \mathbb{N}$. Then $\varphi(s_n) \geq \varphi(\mathbb{1} - (t_2 + \dots + t_n)) = \mathbb{1}_X - (2^{-2} \mathbb{1}_X + \dots + 2^{-n} \mathbb{1}_X) \geq \frac{1}{2} \mathbb{1}_X$.

★ Finally we show $s_n \downarrow 0$, i.e., $s_n(x) \downarrow 0$ (in E) for all $x \in X$. For this it is sufficient to prove that $\{y \in X : (s_n(x))(y) \downarrow 0\}$ is dense in X for all $x \in X$. Let $x \in X$. We will even show that the set $S = \{y \in X : \exists n [(s_n(x))(y) = 0]\}$ is dense in X . Let $a_1, \dots, a_n \in \{0, 1\}$ for some $n \in \mathbb{N}$. Then $y = (a_1, \dots, a_n, x_1, \dots, x_n, 0, 0, \dots)$ is an element of $\langle a_1, \dots, a_n \rangle$ and also an element of S , because $(t_n(x))(y) = 1$. $\square$

5.B.18. Let us call a partially ordered vector space E *uncooperative* if there exist a sequence $(f_n)_{n \in \mathbb{N}}$ of simple functions $\{0, 1\}^{\mathbb{N}} \rightarrow E$ and an $e \in E^+$, $e \neq 0$, such that

$$\begin{cases} f_n(a) \downarrow 0 & \text{for all } a \in \{0, 1\}^{\mathbb{N}}, \\ \varphi(f_n) \geq e & \text{in } E. \end{cases} \quad (5.140)$$

Then $C(\{0, 1\}^{\mathbb{N}})$ is uncooperative. In general:

A partially ordered vector space E is uncooperative if there exists a nonzero σ -order continuous positive linear map $\Phi : C(\{0, 1\}^{\mathbb{N}}) \rightarrow E$.

(Note that $\Phi(\varphi(s)) = \int \Phi \circ s \, d\mu$ for every simple function $s : \{0, 1\}^{\mathbb{N}} \rightarrow C(\{0, 1\}^{\mathbb{N}})$ and every linear $\Phi : C(\{0, 1\}^{\mathbb{N}}) \rightarrow E$.)

For an E being either $C(X)$ or $C_b(X)$, where X is a completely regular topological space for which there exists a continuous open map $X \rightarrow \{0, 1\}^{\mathbb{N}}$ this is the case. Examples of such X are $(0, 1)$, $[0, 1]$, $\mathbb{R}$, the Cantor set, $[0, 1]^2$ and $\mathbb{R}^{\mathbb{N}}$.

An example of a Dedekind complete space that is uncooperative is the Dedekind completion of $C[0, 1]$ (see Theorem 5.6.4).

On the other hand, if E is a Riesz space for which $E_c^\sim$ separates the points of E , then E is not uncooperative. (E.g., $L^1(\mathbb{R})$, ℓ^∞ , c , $F^\sim$ for any Riesz space F .) From the fact that $L^\infty(\mathbb{R})$ is such an E it follows that the space of all equivalence classes of Lebesgue measurable functions on $\mathbb{R}$ is not uncooperative.

§5.B.2 Mediated, splitting and similar notions: their relation to σ -Dedekind completeness

In Definition 5.B.19 we present the notion dividing, which is stronger than mediated and splitting. We compare the relations between them and the relation with σ -Dedekind completeness (5.B.20 – 5.B.29). In 5.B.30 – 5.B.34 we discuss similar properties without the ‘countability restriction’.

In this section we assume D to be a linear subspace of E .

5.B.19 Definition. D is said to be *dividing* in E if the following is true:

$$\begin{aligned} &\text{If } A \text{ and } B \text{ are countable subsets of } D \text{ such that} \\ &A + B \text{ has a supremum in } E, \text{ then so do } A \text{ and } B. \end{aligned} \quad (5.141)$$

We say E is *dividing* if E is dividing in itself.

We investigate the relations between dividing, splitting and mediated. First, in Theorem 5.B.20, Theorem 5.B.21 and Theorem 5.B.22 we give equivalent formulations of mediated, splitting and dividing.

5.B.20 Theorem. *The following are equivalent*

- (a) D is splitting in E .
- (b) For all sequences $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ in D with $a_n \geq 0$, $b_n \geq 0$ for which $\{\sum_{n=1}^N a_n + b_n : N \in \mathbb{N}\}$ has a supremum in D , $\{\sum_{n=1}^N a_n : N \in \mathbb{N}\}$ has a supremum in E .

- (c) For all sequences $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ in D with $a_n \uparrow, b_n \uparrow$ for which $\{a_n + b_n : n \in \mathbb{N}\}$ has a supremum in E , the set $\{a_n : n \in \mathbb{N}\}$ has a supremum in E .

5.B.21 Theorem. *The following are equivalent*

- (a) D is mediated in E .
 (b) If A and B are countable subsets of D with $\sup(A+B) \in D$ (supremum in E), then A and B have a supremum in E .

Proof. Suppose (a). Suppose A and B are countable subsets of D with $\sup(A+B) \in D$. Let $z = \sup(A+B)$. Then $\inf((z-B)-A) = 0$, and so A has an infimum in E . (a) follows from (b) because $\inf(A-B) = 0$ if and only if $\sup(B-A) = 0$. $\square$

5.B.22 Theorem. *The following are equivalent*

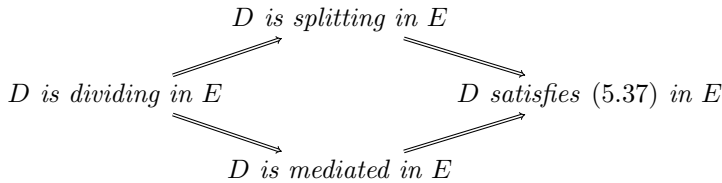
- (a) D is dividing in E .
 (b) For all sequences $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ in D for which $\{\sum_{n=1}^N \sum_{m=1}^M a_n + b_m : N, M \in \mathbb{N}\}$ has a supremum in E , the set $\{\sum_{n=1}^N a_n : N \in \mathbb{N}\}$ has a supremum in E .

One could use the equivalences described above together with Lemma 5.4.22 to prove the following implications.

5.B.23 Corollary. *Let D be a linear subspace of E .*

- (a) *If D is dividing in E , then D is both mediated and splitting in E .*
 (b) *Suppose E is a Riesz space and D is a Riesz subspace of E . If D is splitting in E , then D satisfies (5.37) in E and hence D is mediated in E .*

In a diagram



If E is a Riesz space and D a Riesz subspace of E , then we have the following diagram

$$D \text{ is splitting in } E \implies D \text{ satisfies (5.37) in } E \iff D \text{ is mediated in } E$$

Suppose E is a Riesz space. Then the following are equivalent:

- (c) E is dividing.
 (d) E is splitting.
 (e) E is mediated.
 (f) E satisfies (5.37) with $D = E$.

We already proved that a Banach lattice with a σ -order continuous norm is mediated and splitting (Theorem 5.4.24). It is even dividing, see the next theorem, which is a consequence of Theorem 5.4.24 and Corollary 5.B.23.

5.B.24 Theorem. *A Banach lattice E with σ -order continuous norm $\|\cdot\|$ is dividing.*

5.B.25 Remark. 5.4.23 (c) and (d) provide examples of D that are mediated or splitting in E but not dividing in E . However, this does not provide an example of a D that is both mediated and splitting in E but not dividing in E .

5.B.26 Definition. E is said to be *filling* if for all countable sets $A \subset E$ and $z \in E$ for which $A \leq z$ there exists a countable set B such that $\sup(A + B) = z$.

5.B.27 Theorem. *E is σ -Dedekind complete if and only if E is dividing and filling.*

Proof. “only if”. Suppose E is σ -Dedekind complete. It will be clear that E is dividing. Let $A \subset E$ be countable and $z \in E$ be such that $A \leq z$. Then $\sup A$ exists. Define $b := z - \sup A$. Then $\sup(A + \{b\}) = \sup A + b = z$.

“if”. Suppose E is dividing and filling. Suppose $A \subset E$ and $z \in E$ is such that $A \leq z$. Let $B \subset E$ be such that $\sup(A + B) = z$. Then because E is dividing, A has a supremum in E . $\square$

5.B.28 Example. [A filling space that is not dividing]

Since c is not σ -Dedekind complete, by Theorem 5.B.27 we only have to mention that c is filling. If $(a_n)_{n \in \mathbb{N}}$ is a sequence in c and $z \in E$ is such that $a_n \uparrow$ and $a_n \leq z$ ($n \in \mathbb{N}$) then $b_n := (\beta_1, \beta_2, \dots, \beta_n, 0, 0, \dots)$ with $\beta_i := z(i) - \sup_{n \in \mathbb{N}} a_n(i)$ is such that $a_n + b_n \uparrow z$.

5.B.29 Example. [A non filling Archimedean dividing space] Let $I \subset \mathbb{R}$ be an interval and let $\mathfrak{B}^1$ be the set of functions $I \rightarrow \mathbb{R}$ that are of the first Baire class (i.e., those functions that are the pointwise limit of continuous functions). Then the following holds (see van Rooij and Schikhof [78, Corollary 11,13]):

If $f : I \rightarrow \mathbb{R}$ and there are $(g_n)_{n \in \mathbb{N}}$ and $(h_n)_{n \in \mathbb{N}}$ in $\mathfrak{B}^1$ for which $g_n \downarrow$, $h_n \uparrow$ and

$$\sup_{n \in \mathbb{N}} h_n(t) = f(t) = \inf_{n \in \mathbb{N}} g_n(t) \quad (t \in I), \quad (5.142)$$

then $f \in \mathfrak{B}^1$. This implies that $\mathfrak{B}^1$ is dividing (note that the infimum and the pointwise infimum of first Baire class functions agree, since $\mathbb{1}_{\{t\}} \in \mathfrak{B}^1$ for $t \in I$). $\mathfrak{B}^1$ is also Archimedean, but not σ -Dedekind complete (consider the bounded set $\{\mathbb{1}_{\{q\}} : q \in \mathbb{Q} \cap I\}$) and thus not filling.

We will make a few remarks about the relation between Dedekind completeness and a similar notion to being dividing (one without the “countability restriction”). In Theorem 5.B.34 we discuss a consequence of a similar notion to being mediated (also one without the “countability restriction”).

5.B.30. Consider the following condition

$$\begin{aligned} &\text{If } A \text{ and } B \text{ are subsets of } D \text{ such that} \\ &A + B \text{ has a supremum in } E, \text{ then so do } A \text{ and } B. \end{aligned} \quad (5.143)$$

The analogous notion of filling (Definition 5.B.26) without the “countability restriction” is familiar, as the following theorem shows.

5.B.31 Theorem. *The following are equivalent.*

- (a) *E is Archimedean.*
- (b) *For all nonempty sets $A \subset E$ and $z \in E$ for which $A \leq z$ there exists a set B such that $\sup(A + B) = z$.*
- (c) *For all nonempty countable sets $A \subset E$ and $z \in E$ for which $A \leq z$ there exists a set B such that $\sup(A + B) = z$.*

Consequently, if E is filling then E is Archimedean.

Proof. First note that (b) (and also (c), similarly) is equivalent to the following: For all nonempty sets $A \subset E^+$ there exists a set $B \subset E$ such that $\inf A + B = 0$. Suppose (a), E is Archimedean. Let $A \subset E^+$ and define $B = \{b \in E : b \geq -A\}$. Then $B + A \geq 0$. Let $h \leq B + A$, i.e., $h - b \leq A$ and thus $b - h \in B$ for all $b \in B$. Since $0 \in B$, $-h \in B$ and similarly $-nh \in B$ for all $n \in \mathbb{N}$. Hence $nh \leq a$ for all $a \in A$ by definition of B . Because E is Archimedean, $h \leq 0$ and thus $\inf A + B = 0$. This proves (a) $\Rightarrow$ (b). The implication (b) $\Rightarrow$ (c) is immediate.

Suppose (c) holds. Let $a \in E^+$ and let $A = \{a, \frac{1}{2}a, \frac{1}{3}a, \dots\}$. Let $B \subset E$ be such that $\inf A + B = 0$. Let $c \leq A$, then $c \leq 2A + B$, since $A + B \geq 0$. $2A + B$ and $A + B$ have the same lower bounds, hence $c \leq 0$ and thus $\inf A = 0$. We conclude (c) $\Rightarrow$ (a). $\square$

Example 5.B.29 illustrates that E can be Archimedean but not filling.

5.B.32 Theorem. *The space E is Dedekind complete if and only if E is Archimedean and satisfies (5.143) (with $D = E$).*

5.B.33. We already mentioned in Theorem 5.6.4 that an Archimedean directed partially ordered vector space can be embedded in a Dedekind complete space. As a dividing space or a space E that satisfies (5.143) need not be Archimedean (e.g., $\mathbb{R}^2$ with the lexicographical ordering), one could wonder whether there exists an embedding of a (not necessarily Archimedean) partially ordered vector space in a dividing space.

5.B.34 Theorem. *Suppose that the following holds:*

If A and B are subsets of D such that $\inf A - B = 0$ in E , then

$$A \text{ has an infimum (and consequently } B \text{ has a supremum and } \inf A = \sup B).$$
(5.144)

Then a net $(x_\alpha)_{\alpha \in \mathbb{A}}$ in D order converges to an element $x \in E$ if and only if there exists a net $(h_\alpha)_{\alpha \in \mathbb{A}}$ with $h_\alpha \downarrow 0$ such that

$$\sup_{\alpha \in A} x_\alpha - h_\alpha = x = \inf_{\alpha \in A} x_\alpha + h_\alpha. \quad (5.145)$$

Proof. Let $x_\alpha \rightarrow x$. Let $(h_\alpha)_{\alpha \in \mathbb{A}}$ be such that $-h_\alpha \leq x - x_\alpha \leq h_\alpha$ and $h_\alpha \downarrow 0$. Then $x_\alpha - x_\beta = x - x_\beta + x_\alpha - x \leq h_\alpha + h_\beta$. Therefore

$$\inf_{\alpha, \beta} (x_\alpha + h_\alpha) - (x_\beta - h_\beta) = \inf_{\alpha, \beta} (x_\alpha - x_\beta) + h_\alpha + h_\beta \leq \inf_{\alpha, \beta} 2(h_\alpha + h_\beta) = 0. \quad (5.146)$$

By assumption, $\inf x_\alpha + h_\alpha$ and $\sup_\beta x_\beta - h_\beta$ exist and are equal to x . $\square$

§5.B.3 Remarks about convolution

In this section we consider measurable functions $\mathbb{R} \rightarrow \mathbb{R}$. $L_x f$ is $t \mapsto f(t - x)$ and $[g \geq c] = \{x : g(x) \geq c\}$. By Corollary 5.B.38 we conclude in 5.B.39 that the only possible measurable functions f for which $x \mapsto \mathbb{1}_{[0,1]}(x)L_x f$ is an element of S_{LV} or S_{VLV} ($\mathbb{R}$ is considered to be equipped with the Lebesgue measure), are those that are essentially bounded on every bounded subset of $\mathbb{R}$.

5.B.35 Lemma. *Let f and g be measurable $\mathbb{R} \rightarrow [0, \infty)$, f not essentially bounded. Let $A \subset \mathbb{R}$ be a measurable set with positive measure and assume*

$$L_x f \leq g \quad \text{a.e.} \quad (x \in A). \quad (5.147)$$

Then $\lambda[g \geq n] = \infty$ for all $n \in \mathbb{N}$.

Proof. Assume $\lambda(A) < \infty$.

(I) Let $c \in [0, \infty)$. Take $S \subset [f \geq c]$, $0 < \lambda(S) < \infty$, and $T := [g \geq c]$. The function $(x, y) \mapsto \mathbb{1}_A(x)\mathbb{1}_S(y)\mathbb{1}_{T^c}(x+y)$ is integrable over $\mathbb{R}^2$. Fubini leads to

$$\int \mathbb{1}_A(x)\lambda(S \setminus (T - x)) \, dx = \int \mathbb{1}_S(y)\lambda(A \setminus (T - y)) \, dy. \quad (5.148)$$

For $x \in A$ we have $g \geq L_x f \geq cL_x \mathbb{1}_S = c\mathbb{1}_{S+x}$ a.e., so $S + x \subset T$ a.e., and $\lambda(S \setminus (T - x)) = 0$. Then the left hand member of (5.148) is 0. As S has positive measure, there is a $y \in S$ with $\lambda(A \setminus (T - y)) = 0$. Then $\lambda(A) \leq \lambda(T - y) = \lambda(T)$.

(II) Thus, $\lambda(A) \leq \lambda[g \geq c]$ for all $c \in [0, \infty)$. Now $[g \geq n] \downarrow \emptyset$ but not $\lambda[g \geq n] \downarrow 0$. Hence $\lambda[g \geq n] = \infty$ for all $n \in \mathbb{N}$. $\square$

5.B.36 Corollary. *Let f and g be measurable $\mathbb{R} \rightarrow [0, \infty)$. Suppose*

$$L_x f \leq g \quad \text{a.e.} \quad (x \in A) \quad (5.149)$$

for some measurable set A of positive measure. Then f is essentially bounded on every bounded subset of $\mathbb{R}$.

Also, if B is a measurable set such that f is unbounded on B , then $\lambda(A + B) = \infty$.

Proof. We may assume that A is bounded. Suppose B is a bounded set such that $f\mathbb{1}_B$ is not essentially bounded. As

$$L_x(f\mathbb{1}_B) \leq g\mathbb{1}_{A+B} \quad (x \in A) \quad (5.150)$$

we have $\lambda[g\mathbb{1}_{A+B} \geq 1] = \infty$. Contradiction. $\square$

5.B.37. As a consequence of the previous corollary: If $L_x f \leq g$ a.e. for all $x \in A$ and f is not essentially bounded on bounded subsets of $\mathbb{R}$, then $\lambda(A) = 0$.

Let $M(\mathbb{R})$ be the space of (classes of) measurable functions on $\mathbb{R}$.

5.B.38 Corollary. *Let $f \in M(\mathbb{R})$, let $A \subset \mathbb{R}$ be measurable, $\lambda(A) > 0$. Suppose there exist $w_1 \leq w_2 \leq \dots$ in $M(\mathbb{R})^+$ with*

$$L_x f \in \bigcup_n [-w_n, w_n] \quad (x \in A). \quad (5.151)$$

Then f is essentially bounded on every bounded subset of $\mathbb{R}$.

Proof. For $n \in \mathbb{N}$, put $A_n := \{x \in A : L_x f \in [-w_n, w_n]\}$. Each A_n is measurable and some A_n has positive measure. Now apply Corollary 5.B.36. $\square$

5.B.39. Suppose that $E \subset M(\mathbb{R})$ is a partially ordered vector space. Theorem 5.8.3 shows that for every element of S_{LV} or S_{VLV} there exists a partition $(A_n)_{n \in \mathbb{N}}$ such that $f(A_n)$ is order bounded for each n . Therefore, if $x \mapsto \mathbb{1}_{[0,1]}(x)L_x f$ is an element of S_{LV} or S_{VLV} , Corollary 5.B.38 then implies that f is essentially bounded on every bounded subset of $\mathbb{R}$. Note that the φ_{LV} or φ_{VLV} integral of $x \mapsto \mathbb{1}_{[0,1]}(x)L_x f$ equals the convolution $\mathbb{1}_{[0,1]} * f$.

The following theorem is a special case of Theorem 6.9.1. It implies that for an $f \in C(\mathbb{R})$ the function $x \mapsto \mathbb{1}_{[0,1]}(x)L_x f$ is Bochner integrable. We refer for the details to Theorem 6.9.1.

5.B.40 Theorem. *For every $f \in C(\mathbb{R})$ there exists a $w \in C(\mathbb{R})^+$ such that*

$$\begin{aligned} &\text{Every } L_x f \quad (x \in \mathbb{R}) \text{ lies in the principal ideal } C(\mathbb{R})_w; \\ &x \mapsto L_x f \text{ is continuous relative to } \|\cdot\|_w. \end{aligned} \quad (5.152)$$

5.B.41. Thus, contrary to Corollary 5.B.38: if f is bounded on bounded sets, there is a continuous function larger than f , and by Theorem 5.B.40 $\{L_x f : x \in \mathbb{R}\}$ is contained in a principal ideal of the space of all locally essentially bounded functions.

§5.B.4 Improper S_{VL} -functions

In this section $(X, \mathcal{A}, \mu)$ and $S, \mathcal{I}, \varphi$ be as in Section 5.8 and E is a mediated Riesz space. We assume S_V to be stable.

Note that E is also splitting (see 5.4.23(b)). Therefore $S_{LV} \subset S_{VL}$ by Theorem 5.5.8. In this section we will only consider the (larger) space S_{VL} together with φ_{VL} . However, one could also consider S_{LV} and φ_{LV} in the statements of this section everywhere where S_{VL} and φ_{VL} are considered.

5.B.42 Lemma. *Let $f \in S_{VL}^+$. Then $f \wedge u \mathbb{1}$ is in S_{VL} for all $u \in E^+$ and*

$$\varphi_{VL}(f) = \sup\{\varphi_{VL}(f \wedge u \mathbb{1}) \, d\mu : u \in E^+\}. \quad (5.153)$$

Proof. Take $u \in E^+$. By Lemma 5.3.9 applied to $\Theta(g) = g \wedge u \mathbb{1}$ we have that $g \wedge u \mathbb{1}$ is an element of S_V for $g \in S_V$. Due to the fact that E is splitting (see 5.4.23(b)), $g \wedge u \mathbb{1}$ is an element of S_{VL} for all $g \in S_{VL}^+$.

Let $f \in S_{VL}^+$ and let $(A_n)_{n \in \mathbb{N}}$ a φ -partition for f . Let $B_N = \bigcup_{n \geq N} A_n$. Then

$$\varphi_{LV}(f \mathbb{1}_{B_N}) \leq \sum_{n \geq N} \varphi_V(f \mathbb{1}_{A_n}) \downarrow 0. \quad (5.154)$$

For $n \in \mathbb{N}$ let $a_n \in E^+$ be such that $f \mathbb{1}_{A_n} \leq a_n \mathbb{1}_{A_n}$. If $N \in \mathbb{N}$ and $u = a_1 \vee \cdots \vee a_N$, then $f \leq u \mathbb{1}$ on $X \setminus B_N$, so

$$\varphi_{LV}(f) - \varphi_{LV}(f \wedge u \mathbb{1}) \leq \varphi_{LV}(f \mathbb{1}_{B_N}). \quad (5.155)$$

Hence $\inf\{\varphi_{LV}(f) - \varphi_{LV}(f \wedge u \mathbb{1}) : u \in E^+\} \leq \inf\{\varphi_{LV}(f \mathbb{1}_{B_N}) : N \in \mathbb{N}\} = 0$. $\square$

5.B.43 Definition. We call a function $f : X \rightarrow E^+$ *improperly in S_{VL}* with integral I if

- $f \wedge u \mathbb{1}$ is in S_{VL} for all $u \in E^+$,
- $I = \sup\{\varphi_{VL}(f \wedge u) : u \in E^+\}$.

5.B.44 Example. Consider the function $f : \mathbb{R} \rightarrow L^1(\lambda)$ in Examples 5.9.9(II). f is Bochner integrable but not an element of S_{VLV} (see also 5.B.39). For $k \in \mathbb{N}$ let $f_k : \mathbb{R} \rightarrow L^1(\lambda)$ be given by

$$f_k(x) = f(x) \wedge k \mathbb{1}. \quad (5.156)$$

Then f_k as a map $\mathbb{R} \rightarrow L^1(\lambda)$ is an element of S_V and $\sup_{k \in \mathbb{N}} \varphi_V(f_k) = \mathfrak{b}(f)$:

Let $k \in \mathbb{N}$ and note that $f_k(x) = \mathbb{1}_{[0,1]}(x) L_x(g \wedge k \mathbb{1})$. For $n \in \mathbb{N}$ let $I_{n,1}, \dots, I_{n,2^n}$ be the intervals

$$I_{n,i} = [(i-1)2^{-n}, i2^{-n}) \quad (i \in \{1, \dots, 2^n - 1\}), \quad (5.157)$$

$$I_{n,2^n} = [1 - 2^{-n}, 1]. \quad (5.158)$$

We follow a similar approach as in the proof of Theorem 5.8.16. For $n \in \mathbb{N}$ let $h_n, l_n : X \rightarrow L^1(\lambda)$ be given by

$$h_n(x) = k \mathbb{1}_{I_{n,i}} + L_{i2^{-n}}(g \wedge k \mathbb{1}) \quad (x \in I_{n,i}), \quad (5.159)$$

$$l_n(x) = L_{(i-1)2^{-n}}(g \wedge k \mathbb{1}) \quad (x \in I_{n,i}). \quad (5.160)$$

Then $h_n, l_n \in S$ and $h_n \geq f_k \geq l_n$ for all n . We show that $f \in S_V$ by showing $\inf_{n \in \mathbb{N}} \varphi(h_n - l_n) = 0$ (as $L^1(\lambda)$ is σ -Dedekind complete, this suffices by Lemma 5.3.7).

$$\begin{aligned} \varphi(h_n - l_n) &= 2^{-n} \left[\sum_{i=1}^{2^n} k \mathbb{1}_{I_{n,i}} + L_{i2^{-n}}(g \wedge k \mathbb{1}) - L_{(i-1)2^{-n}}(g \wedge k \mathbb{1}) \right] \\ &= 2^{-n} \left[k \mathbb{1}_{[0,1]} + L_1(g \wedge k \mathbb{1}) - (g \wedge k \mathbb{1}) \right] \leq 2^{1-n} k \mathbb{1}_{[0,1]}. \end{aligned} \quad (5.161)$$

Therefore, $f_k \in S_V$. As f_k is also Bochner integrable (see 6.4.4(a); f_k is continuous by [47, Theorem 20.4] and so $f_k(\mathbb{R})$ is separable), by Theorem 5.9.4 we have $\mathfrak{b}(f_k) = \varphi_V(f_k)$.

As $\lim_{k \rightarrow \infty} \|f(x) - f_k(x)\|_{L^1(\lambda)} = 0$ and $\|f_k(x)\|_{L^1(\lambda)} \leq \|f(x)\|_{L^1(\lambda)}$ for all $x \in \mathbb{R}$, we have $\sup_{k \in \mathbb{N}} \varphi_V(f_k) = \lim_{k \rightarrow \infty} \mathfrak{b}(f_k) = \mathfrak{b}(f)$ by Lebesgue's Dominated Convergence Theorem (see also Zaanen [90, Ch. 6 Section 31 Theorem 4]).

5.B.45 Theorem. *Let $f : X \rightarrow E^+$ be improperly in S_{VL} . Suppose $f \leq h$ for some $h \in S_{VL}^+$. Then $f \in S_{VL}^+$.*

Proof. There exists a partition $(A_n)_{n \in \mathbb{N}}$ on which h and hence f is bounded, i.e., for all $n \in \mathbb{N}$ there exists an $a_n \in E^+$ with $0 \leq f \mathbb{1}_{A_n} \leq h \mathbb{1}_{A_n} \leq a_n$. Since $f \mathbb{1}_{A_n} = (f \mathbb{1}_{A_n}) \wedge (a_n \mathbb{1}) = (f \wedge a_n \mathbb{1}) \mathbb{1}_{A_n}$ this implies $f \mathbb{1}_{A_n}$ is an element of Γ_{VL} (actually of Γ_V , see Theorem 5.5.5). Because E is splitting, this implies $f \in S_{VL}^+$. $\square$

5.B.46. Definition 5.B.43 would be a good ingredient for another extension. However, if f and g are $X \rightarrow E^+$ functions that are improperly in S_{VL} , then it is not clear whether $f + g$ is improperly in S_{VL} .



CHAPTER 6

Bochner integrals in ordered vector spaces

This chapter is based on a preprint:

A. van Rooij and W. van Zuijlen. Bochner integrals in ordered vector spaces. Preprint available at <http://arxiv.org/abs/1604.06341>

Abstract

We present a natural way to cover an Archimedean directed ordered vector space E by Banach spaces and extend the notion of Bochner integrability to functions with values in E . The resulting set of integrable functions is an Archimedean directed ordered vector space and the integral is an order preserving map.

§6.1 Introduction

We extend the notion of Bochner integrability to functions with values in a vector space E that may not itself be a Banach space but is the union of a collection $\mathbb{B}$ of Banach spaces.

The idea is the following. We call a function f , defined on a measure space X and with values in E , “integrable” if for some D in $\mathbb{B}$ all values of f lie in D and f is Bochner integrable as a function $X \rightarrow D$. Of course, one wants a certain consistency: the “integral” of such an f should be independent of the choice of D .

In [88], Thomas obtains this consistency by assuming a Hausdorff locally convex topology on E , entailing many continuous linear functions $E \rightarrow \mathbb{R}$. Their restrictions to the Banach spaces that constitute $\mathbb{B}$ enable one to apply Pettis integration, which leads to the desired uniqueness.

Our approach is different, following a direct-limit-like construction. We assume E to be an ordered vector space with some simple regularity properties (Archimedean, directed) and show that E is the union of a certain increasing system $\mathbb{B}$ of Banach spaces with closed, generating positive cones (under the ordering of E). Uniqueness of the integral follows from properties of such ordered Banach spaces. Moreover, the integrable functions form a vector space and the integral is linear and order preserving.

In Section 6.3 we study ordered Banach spaces with closed generating cones. We give certain properties which can be used to give an alternative proof of a classical theorem which states that every order preserving linear map is continuous, and generalise it to order bounded linear maps. In Section 6.4 we study Bochner integrable functions with values in an ordered Banach space with closed generating cone. In Section 6.5 we present the definition of a Banach cover and the definition of the extension of the Bochner integral to functions with values in a vector space that admits a Banach cover. In Section 6.6 we show that an Archimedean ordered vector space possesses a Banach cover consisting of ordered Banach spaces with closed generating cones. In Section 6.7 we study integrable functions with values in Archimedean ordered vector spaces. In Section 6.8 we compare the integral with integrals considered in [80]. In Section 6.9 we present an application to view the convolution as an integral.

§6.2 Notation

$\mathbb{N}$ is $\{1, 2, \dots\}$. We write “for all n ” instead of “for all $n \in \mathbb{N}$ ”. To avoid confusion:

- An “order” is a “partial order”.
- We call an ordered vector space *Archimedean* (see Peressini [73]) if for all $a, b \in E$ the following holds: if $na \leq b$ for all $n \in \mathbb{N}$, then $a \leq 0$. (In some places, e.g., Birkhoff [8], such spaces are said to be ‘integrally closed’.)

As is common in literature, our notations do not distinguish between a function on a measure space and the class of that function.

§6.3 Ordered Banach spaces with closed generating cones

In this section we describe properties of ordered Banach spaces with closed generating cones. Using these properties we prove in Theorem 6.3.11 that an order bounded map between ordered Banach spaces with closed generating cones is continuous.

6.3.1 Definition. A *normed ordered vector space* is a normed vector space with an order that makes it an ordered vector space. An *ordered Banach space* is a Banach space that is a normed ordered vector space.

A priori there is no connection between the ordering and the norm of a normed ordered vector space. One reasonable and useful connection is the assumption that the (positive) cone be closed.

6.3.2 Theorem. Let E be a normed ordered vector space. E^+ is closed if and only if

$$x \leq y \iff \alpha(x) \leq \alpha(y) \text{ for all } \alpha \in (E')^+. \quad (6.1)$$

Consequently, whenever E^+ is closed then $(E')^+$ separates the points of E and E is Archimedean.

Proof. Since E^+ is convex, E^+ is closed if and only if it is weakly closed (i.e., $\sigma(E, E')$ -closed); see [21, Theorem V.1.4]. The rest follows by [2, Theorem 2.13(3&4)]. $\square$

The following theorem is due to Andô [3]. See also [2, Corollary 2.12].

6.3.3 Theorem. Let D be an ordered Banach space with a closed generating¹ cone D^+ . There exists a $C > 0$ such that

$$C\|x\| \geq \inf\{\|a\| : a \in D^+, -a \leq x \leq a\} \quad (x \in E). \quad (6.2)$$

6.3.4 Definition. Let D be a directed² ordered Banach space. If $C > 0$ is such that (6.2) holds, then we say that the norm $\|\cdot\|$ is *C-absolutely dominating*.³ We say that a norm $\|\cdot\|$ is *absolutely dominating* if it is C -absolutely dominating for some $C > 0$.

On a Banach lattice the norm is 1-absolutely dominating. Actually for Banach lattices there is equality in (6.2).

6.3.5. We refer the reader to Appendix 6.A for the following facts: If $\|\cdot\|$ is C -absolutely dominating on a directed ordered Banach space D , then $C \geq 1$. Whenever there exists a absolutely dominating norm, then for all $\varepsilon > 0$ there exists an equivalent $(1 + \varepsilon)$ -absolutely dominating norm. All norms on a directed ordered vector space D that make D complete and D^+ closed are equivalent (see 6.6.2).

¹ D^+ is generating if $D = D^+ - D^+$.

² D is directed if $D = D^+ - D^+$, i.e., if D^+ is generating.

³Batty and Robinson [4] call the cone D^+ *approximately C-absolutely dominating*, and, Messerschmidt [68] calls D *approximately C-absolutely conormal*, if the norm on D is C -absolutely dominating.

6.3.6. Let D be a directed ordered Banach space. Then $\|\cdot\|$ is absolutely dominating if and only if the (convex) set

$$\bigcup_{a \in D^+, \|a\| \leq 1} [-a, a] \quad (6.3)$$

is a neighbourhood of 0.

6.3.7 Definition. (See [66, §16]) Let E be an ordered vector space. We say that a sequence $(x_n)_{n \in \mathbb{N}}$ in E *converges uniformly* to an element $x \in E$ (notation: $x_n \xrightarrow{u} x$) whenever there exist $a \in E^+$, $\varepsilon_n \in (0, \infty)$ with $\varepsilon_n \rightarrow 0$ and

$$-\varepsilon_n a \leq x_n - x \leq \varepsilon_n a \quad (n \in \mathbb{N}). \quad (6.4)$$

Note that one may replace “ $\varepsilon_n \rightarrow 0$ ” by “ $\varepsilon_n \downarrow 0$ ”. If E is Archimedean then the x as above is unique. We will only consider such convergence in Archimedean spaces. We say that a sequence $(x_n)_{n \in \mathbb{N}}$ in E is a *uniformly Cauchy sequence* if there exists an $a \in E^+$ such that for all $\varepsilon > 0$ there exists an N such that $-\varepsilon a \leq x_n - x_m \leq \varepsilon a$ for all $n, m \geq N$. E is called *uniformly complete* whenever it is Archimedean and all uniformly Cauchy sequences converge uniformly.

6.3.8 Lemma. Let D_1, D_2 be ordered vector spaces and $T : D_1 \rightarrow D_2$. If T is linear and order bounded, then T preserves uniform convergence.

Proof. Suppose $x_n \in D_1$, $\varepsilon_n \in (0, 1)$ and $a \in D_1^+$ are such that $-\varepsilon_n a \leq x_n \leq \varepsilon_n a$ and $\varepsilon_n \downarrow 0$. Let $b \in D_2^+$ be such that $T[-a, a] \subset [-b, b]$. Then $\varepsilon_n^{-1} T x_n \in [-b, b]$, i.e., $-\varepsilon_n b \leq T x_n \leq \varepsilon_n b$ for all n . $\square$

6.3.9 Theorem. Let D be an ordered Banach space. Consider the following conditions.

- (i) $\|\cdot\|$ is absolutely dominating.
- (ii) If $a_1, a_2, \dots \in D^+$, $\sum_{n \in \mathbb{N}} \|a_n\| < \infty$, then there exist $a \in D^+$, $\varepsilon_n \in (0, \infty)$ with $\varepsilon_n \rightarrow 0$ such that $a_n \leq \varepsilon_n a$ for all n .⁴
- (iii) If $x_1, x_2, \dots \in D$, $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$, then $x_n \xrightarrow{u} 0$.

If D^+ is closed, then D satisfies (ii).

Suppose D is Archimedean and directed. Then the following are equivalent.

- (a) D^+ is closed.
- (b) D satisfies (i) and (ii).
- (c) D satisfies (iii).

Proof. Suppose D^+ is closed. Let $a_1, a_2, \dots \in D^+$, $\sum_{n \in \mathbb{N}} \|a_n\| < \infty$. Choose $\varepsilon_n \in (0, \infty)$, $\varepsilon_n \rightarrow 0$ such that $\sum_{n \in \mathbb{N}} \varepsilon_n^{-1} \|a_n\| < \infty$. By norm completeness, $a := \sum_{n \in \mathbb{N}} \varepsilon_n^{-1} a_n$ exists. Because D^+ is closed $a \geq \varepsilon_n^{-1} a_n$ for all n .

⁴For normed Riesz spaces property (ii) is equivalent to what is called the weak Riesz-Fischer condition (see Zaanen [91, Ch. 14, §101]).

(a) $\implies$ (b). (ii) is implied by the above argument. By Theorem 6.3.3 we have (i).
 (b) $\implies$ (c). Let $x_1, x_2, \dots \in D$, $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$. Using (i) let $a_1, a_2, \dots \in D^+$ with $\sum_{n \in \mathbb{N}} \|a_n\| < \infty$ be such that $-a_n \leq x_n \leq a_n$. By (ii) it then follows that $x_n \xrightarrow{u} 0$.
 (c) $\implies$ (a). Take b in the closure of D^+ . For $n \in \mathbb{N}$, choose $x_n \in D^+$, $\sum_{n \in \mathbb{N}} \|x_n - b\| < \infty$. By (iii) there exist $a \in D^+$, $\varepsilon_n \in (0, \infty)$ with $\varepsilon_n \rightarrow 0$ and $-\varepsilon_n a \leq x_n - b \leq \varepsilon_n a$, so that $b \geq x_n - \varepsilon_n a \geq -\varepsilon_n a$. Then $b \geq 0$ because D is Archimedean. $\square$

6.3.10 Lemma. *Let D be an ordered Banach space for which D^+ is closed. Then*

$$x_1, x_2, \dots \in D, a \in D, b \in D, \|x_n - a\| \rightarrow 0, x_n \xrightarrow{u} b \implies a = b. \quad (6.5)$$

Proof. Assume $b = 0$. There exist $c \in D^+$, $\varepsilon_n \in (0, \infty)$ with $\varepsilon_n \downarrow 0$ and $-\varepsilon_n c \leq x_n \leq \varepsilon_n c$. Then $-\varepsilon_n c \leq x_n \leq \varepsilon_n c$ whenever $n \geq N$. Since D^+ is closed, $-\varepsilon_n c \leq a \leq \varepsilon_n c$ for all N , and so, as D is Archimedean by Theorem 6.3.2, $a = 0$. $\square$

With this we can easily prove the following theorem. In a recently published book by Aliprantis and Tourky [2]⁵ but also in older Russian papers by Wulich [15] one can find the proof that an order preserving linear map is continuous (see [2, Theorem 2.32] or combine [15, Theorem III.2.2] (which states the result for $D_2 = \mathbb{R}$) and [16, Theorem VI.2.1]). Theorem 6.3.11 is more general in the sense that it states that linear order bounded maps are continuous.

6.3.11 Theorem. *Let D_1, D_2 be ordered Banach spaces. Suppose D_1^+ is closed and generating and D_2^+ is closed. Let $T : D_1 \rightarrow D_2$ be linear and order bounded. Then T is continuous. Consequently, if T is an order isomorphism then it is a homeomorphism.*

Proof. Let $x_1, x_2, \dots \in D_1$, $x_n \rightarrow 0$ and suppose $Tx_n \rightarrow c$ for some $c \in D_2$. If from this we can prove $c = 0$, then by the Closed Graph Theorem T will be continuous. We may assume $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$. Then $x_n \xrightarrow{u} 0$ in D_1 (Theorem 6.3.9), so $Tx_n \xrightarrow{u} 0$ in D_2 (Lemma 6.3.8). Hence, $c = 0$ according to Lemma 6.3.10. $\square$

We present a consequence as has been done for order preserving linear maps in [16, Theorem VI.2.2].

6.3.12 Corollary. *Let D_1 be an ordered Banach space with closed generating cone. Let D_2 be an ordered normed vector space with a normal cone. Let $T : D_1 \rightarrow D_2$ be linear and order bounded. Then T is continuous.*

Proof. As D_2^+ is normal, we have $D_2^\sim = D_2'$ (see [2, Corollary 2.27]). Let Δ be the unit ball in D_1 . Then for all $\phi \in D_2'$, the map $\phi \circ T$ is an order bounded functional, whence continuous by Theorem 6.3.11. Thus $\phi \circ T(\Delta)$ is bounded for all $\phi \in D_2'$. By an application of the Principle of Uniform Boundedness (see [21, Corollary III.4.3]), $T(\Delta)$ is norm bounded. $\square$

⁵For complete metrisable ordered vector spaces.

§6.4 The Bochner integral on ordered Banach spaces

In this section $(X, \mathcal{A}, \mu)$ **is a complete σ -finite measure space, with $\mu \neq 0$. For more assumptions see 6.4.3 and 6.4.7.**

We define the integral of simple functions in Definition 6.4.1 and recall the definition and facts on Bochner integrability in Definition 6.4.2 and 6.4.4. After that, we consider an ordered Banach space D . In Theorems 6.4.5 and 6.4.6 we describe the order structure of the space $\mathfrak{B}_D$ of Bochner integrable functions. In 6.4.8 we summarise the results of 6.4.9 – 6.4.16, in which we compare closedness and generatingness of the positive cones of D and $\mathfrak{B}_D$.

6.4.1 Definition. Let E be a vector space. We say that a function $f : X \rightarrow E$ is *simple* if there exist $N \in \mathbb{N}$, $a_1, \dots, a_N \in E$, $A_1, \dots, A_N \in \mathcal{A}$ with $\mu(A_1), \dots, \mu(A_N) < \infty$ for which

$$f = \sum_{n=1}^N a_n \mathbb{1}_{A_n}. \quad (6.6)$$

The simple functions form a linear subspace S of E^X , which is a Riesz subspace of E^X in case E is a Riesz space. We define $\varphi : S \rightarrow E$ by

$$\varphi(f) = \sum_{n=1}^N \mu(A_n) a_n, \quad (6.7)$$

where f, N, A_n, a_n are as in (6.6). $\varphi(f)$ is called the *integral* of f . We write $S_{\mathbb{R}}$ for the linear space of simple functions $X \rightarrow \mathbb{R}$.

Definition 6.4.2 and the facts in 6.4.4 can be found in Chapter III in the book by E. Hille and R.S. Phillips, [48].

6.4.2 Definition. Let $(D, \|\cdot\|)$ be a Banach space. A function $f : X \rightarrow D$ is called *Bochner integrable* whenever there exists a sequence of simple functions $(s_n)_{n \in \mathbb{N}}$ such that

$$\int \|f(x) - s_n(x)\| \, d\mu(x) \rightarrow 0. \quad (6.8)$$

Then the sequence $(\varphi(s_n))_{n \in \mathbb{N}}$ converges. Its limit is independent of the choice of the sequence $(s_n)_{n \in \mathbb{N}}$ and is called the *Bochner integral* of f .

6.4.3. For the rest of this section, D is a Banach space (with norm $\|\cdot\|$), and, we write $\mathfrak{B}$ (or $\mathfrak{B}_D$) for the Banach space of classes of Bochner integrable functions $X \rightarrow D$, with norm $\|\cdot\|_{\mathfrak{B}}$ (see 6.4.4(b)). We write $\mathfrak{b}$ (or $\mathfrak{b}_D$) for the Bochner integral on $\mathfrak{B}$.

6.4.4. Some facts on the Bochner integrable functions:

- (a) [48, Theorem 3.7.4] & [82, Proposition 2.15] If $f : X \rightarrow D$ is Borel measurable, $f(X)$ is separable and $\int \|f\| \, d\mu < \infty$, then f is Bochner integrable.

- (b) [48, Theorem 3.7.8] $\mathfrak{B}$ is a Banach space under the norm $\|\cdot\|_{\mathfrak{B}} : \mathfrak{B} \rightarrow [0, \infty)$ given by

$$\|f\|_{\mathfrak{B}} = \int \|f\| \, d\mu. \quad (6.9)$$

- (c) [48, Theorem 3.7.12] Let E be a Banach space and $T : D \rightarrow E$ be linear and continuous. If $f \in \mathfrak{B}_D$, then $T \circ f \in \mathfrak{B}_E$ and

$$T(\mathfrak{b}_D(f)) = \mathfrak{b}_E(T \circ f). \quad (6.10)$$

- (d) [48, Theorem 3.7.9] Let $f_1, f_2, \dots$ be in $\mathfrak{B}$, $f : X \rightarrow D$ and $h \in \mathcal{L}^1(\mu)^+$. Suppose that $f_n(x) \rightarrow f(x)$ and $\|f_n(x)\| \leq h(x)$ for μ -almost all $x \in X$. Then $f \in \mathfrak{B}$ and

$$\mathfrak{b}(f_n) \rightarrow \mathfrak{b}(f). \quad (6.11)$$

- (e) If D is an ordered Banach space then $\mathfrak{B}$ is an ordered Banach space under the ordering given by

$$f \leq g \quad (\text{in } \mathfrak{B}) \iff f \leq g \quad \mu - \text{a.e.} \quad (6.12)$$

6.4.5 Theorem. *Let D be a Banach lattice. Then $\mathfrak{B}_D$ is a Banach lattice and $\mathfrak{b}$ is linear and order preserving.*

Proof. The Bochner integrable functions form a Riesz space because of the inequality $\| |x| - |y| \| \leq \|x - y\|$. $\square$

6.4.6 Theorem. *Let D be an ordered Banach space for which D^+ is closed. Then $\mathfrak{b}$ is order preserving.*

Proof. Let $f \in \mathfrak{B}$ and $f \geq 0$. Then

$$\alpha(\mathfrak{b}(f)) = \int \alpha \circ f \, d\mu \geq 0 \quad (\alpha \in (D')^+). \quad (6.13)$$

Whence $\mathfrak{b}(f) \geq 0$, by Theorem 6.3.2. $\square$

6.4.7. For the rest of this section, D is an ordered Banach space.

6.4.8. The following is a list of results presented in 6.4.9 – 6.4.16.

- (a) $\mathfrak{B}^+$ is closed if and only if D^+ is (6.4.9).
- (b) If $\mathfrak{B}$ is directed, then so is D (straightforward, see also Theorem 6.4.12).
- (c) Let $C > 0$ and D^+ be closed and generating. If $\|\cdot\|$ is C -absolutely dominating, then so is $\|\cdot\|_{\mathfrak{B}}$ (Lemma 6.4.11).
- (d) Let $C > 0$. If $\mathfrak{B}$ is directed and $\|\cdot\|_{\mathfrak{B}}$ is C -absolutely dominating, then so is $\|\cdot\|$ (Theorem 6.4.12).

- (e) $\mathfrak{B}^+$ is closed and generating if and only if D^+ is closed and generating (Theorem 6.4.16).
- (f) If there exist disjoint $A_1, A_2, \dots$ in $\mathcal{A}$ with $0 < \mu(A) < \infty$ for all n : If $\mathfrak{B}^+$ is generating, then so is D^+ and $\|\cdot\|$ is absolutely dominating (Corollary 6.4.14).
- (g) If no such $A_1, A_2, \dots$ exist: $\mathfrak{B}^+$ is generating if and only if D^+ is generating (6.4.15).

6.4.9. Whenever $f \in \mathfrak{B}$, $f_n \in \mathfrak{B}^+$ with $\int \|f - f_n\| d\mu \rightarrow 0$, then there exist Bochner integrable $g_n \geq 0$ with $\sum_{n \in \mathbb{N}} \int \|f - g_n\| d\mu < \infty$; this implies $g_n \rightarrow f$ μ -almost everywhere. So whenever D^+ is closed this implies $f \geq 0$ μ -almost everywhere.

We infer that $\mathfrak{B}^+$ is closed whenever D^+ is.

On the other hand, if $\mathfrak{B}^+$ is closed then so is D^+ . Indeed, let $A \in \mathcal{A}$, $0 < \mu(A) < \infty$. If $a_n \in D^+$ and $a_n \rightarrow a$, then $\|a \mathbb{1}_A - a_n \mathbb{1}_A\|_{\mathfrak{B}} \rightarrow 0$. Therefore $a \in D^+$.

6.4.10 Lemma. Suppose D^+ is generating and $C > 0$ is such that $\|\cdot\|$ is C -absolutely dominating. Let $f : X \rightarrow D$ be simple, let $\varepsilon > 0$. Then there exists a simple $g : X \rightarrow D^+$ with $-g \leq f \leq g$ and $\int \|g\| d\mu \leq C \int \|f\| d\mu + \varepsilon$.

Proof. Write $f = \sum_{n=1}^N x_n \mathbb{1}_{A_n}$ with disjoint sets $A_1, \dots, A_N \in \mathcal{A}$ of finite measure and $x_1, \dots, x_N \in D$. Let $\kappa = \mu(\bigcup_{n=1}^N A_n)$ and assume $\kappa > 0$. For each n , choose $a_n \in D^+$, $-a_n \leq x_n \leq a_n$, $\|a_n\| \leq C\|x_n\| + \frac{\varepsilon}{\kappa}$. Put $g = \sum_{n=1}^N a_n \mathbb{1}_{A_n}$. Then $-g \leq f \leq g$ and $\int \|g\| d\mu = \sum_{n=1}^N \|a_n\| \mu(A_n) \leq \sum_{n=1}^N (C\|x_n\| + \frac{\varepsilon}{\kappa}) \mu(A_n) = C \int \|f\| d\mu + \varepsilon$. $\square$

6.4.11 Lemma. Suppose D^+ is closed and generating and $C > 0$. Then $\mathfrak{B}^+$ is closed and generating. If $\|\cdot\|$ is C -absolutely dominating, then so is $\|\cdot\|_{\mathfrak{B}}$.

Proof. $\mathfrak{B}^+$ is closed by 6.4.9. Assume that $\|\cdot\|$ is C -absolutely dominating. Let $f : X \rightarrow D$ be Bochner integrable and let $\varepsilon > 0$. We prove there exists a Bochner integrable $g : X \rightarrow D^+$ with $-g \leq f \leq g$ μ -a.e. and $\int \|g\| d\mu \leq C(\int \|f\| d\mu + \varepsilon)$. Choose simple functions $s_1, s_2, \dots$ with $\int \|f - s_n\| d\mu < \varepsilon 2^{-n-1}$. Then $s_n \rightarrow f$ pointwise outside a μ -null set Y .

Define $f_1 = s_1$, $f_n = s_n - s_{n-1}$ for $n \in \{2, 3, \dots\}$. Observe

$$\begin{aligned} \int \|f_1\| d\mu &< \int \|f\| d\mu + \varepsilon 2^{-2}, \\ \int \|f_n\| d\mu &< \varepsilon 2^{-n-1} + \varepsilon 2^{-n-2} < \varepsilon 2^{-n} \quad (n \in \{2, 3, \dots\}). \end{aligned} \quad (6.14)$$

For each n , choose a simple $g_n : X \rightarrow D^+$ with $-g_n \leq f_n \leq g_n$ such that

$$\begin{aligned} \int \|g_1\| d\mu &< C \left(\int \|f\| d\mu + \varepsilon 2^{-2} \right), \\ \int \|g_n\| d\mu &< C \varepsilon 2^{-n} \quad (n \in \{2, 3, \dots\}). \end{aligned} \quad (6.15)$$

As $\sum_{n \in \mathbb{N}} \int \|g_n\| d\mu < \infty$ there is a μ -null set $Z \subset X$ for which

$$\sum_{n \in \mathbb{N}} \|g_n(x)\| < \infty \quad (x \in X \setminus Z). \quad (6.16)$$

Put $X_0 := X \setminus (Y \cup Z)$. Define $g : X \rightarrow D^+$ by

$$g(x) = \begin{cases} \sum_{n \in \mathbb{N}} g_n(x) & x \in X_0, \\ 0 & x \notin X_0. \end{cases} \quad (6.17)$$

Then g is Bochner integrable and $\int \|g\| \, d\mu \leq \sum_{n \in \mathbb{N}} \int \|g_n\| \, d\mu \leq C(\int \|f\| \, d\mu + \varepsilon)$.

Moreover, for $x \in X_0$ we have $-g(x) \leq \sum_{n=1}^N f_n(x) = s_N(x) \leq g(x)$ for all N , whereas $s_N(x) \rightarrow f(x)$ since $x \notin Y$. From the closedness of D^+ it follows that $-g \leq f \leq g$ on X_0 . $\square$

In the following theorems (6.4.12, 6.4.13, 6.4.14 and 6.4.15) we derive properties of D from properties of $\mathfrak{B}$. In Theorem 6.4.16 we show that D has a closed and generating cone if and only if $\mathfrak{B}$ does.

6.4.12 Theorem. *Assume $\mathfrak{B}$ is directed. Let $C > 0$ and suppose $\|\cdot\|_{\mathfrak{B}}$ is C -absolutely dominating. Then D is directed and $\|\cdot\|$ is C -absolutely dominating.*

Proof. Let $x \in D$ and $x \neq 0$. Let $A \in \mathcal{A}$ with $0 < \mu(A) < \infty$. Then $x\mathbb{1}_A \in \mathfrak{B}$, $\|x\mathbb{1}_A\|_{\mathfrak{B}} = \mu(A)\|x\|$. Let $C' > C$. There is a $g \in \mathfrak{B}$ with $-g \leq x\mathbb{1}_A \leq g$, $\int \|g(t)\| \, d\mu(t) \leq C'\|x\mathbb{1}_A\|_{\mathfrak{B}}$. Then $\int_A \|g(t)\| \, d\mu(t) \leq C'\mu(A)\|x\|$, so $\mu\{t \in A : \|g(t)\| > C'\|x\|\} < \mu(A)$. In particular, there is a $t \in A$ with $\|g(t)\| \leq C'\|x\|$ and $-g(t) \leq x \leq g(t)$. $\square$

6.4.13 Theorem. *Let D be an ordered Banach space such that the Bochner integrable functions $\mathbb{N} \rightarrow D$ form a directed space. Then D is directed and $\|\cdot\|$ is absolutely dominating.*

Proof. D is directed. In case $\|\cdot\|$ is not absolutely dominating, there exist $x_1, x_2, \dots \in D$ such that for every n

$$2^n \|x_n\| \leq \inf\{\|a\| : a \in D^+, -a \leq x_n \leq a\}$$

and $\|x_n\| = 2^{-n}$. Then $n \mapsto x_n$ is Bochner integrable, so, by our assumption, there exist $a_n \in D^+$ with $-a_n \leq x_n \leq a_n$ for all n and $\sum_{n \in \mathbb{N}} \|a_n\| < \infty$ which is false. $\square$

6.4.14 Corollary. *Suppose there exist disjoint $A_1, A_2, \dots$ in $\mathcal{A}$ with $0 < \mu(A_n) < \infty$ for all n . Suppose $\mathfrak{B}$ is directed. Then D is directed and $\|\cdot\|$ is absolutely dominating.*

Proof. This follows from Theorem 6.4.13 since $f \mapsto \sum_{n \in \mathbb{N}} f(n)\mathbb{1}_{A_n}$ forms an isometric order preserving isomorphism from the Bochner integrable functions $\mathbb{N} \rightarrow D$ into $\mathfrak{B}$. $\square$

6.4.15. Whenever there do not exist $A_1, A_2, \dots$ as in Corollary 6.4.14, then $X = A_1 \cup \dots \cup A_N$, where $A_1, \dots, A_N$ are disjoint atoms. Let $\alpha_n = \mu(A_n) \in (0, \infty)$. Define a norm $\|\cdot\|_N$ on D^N by $\|x\|_N = \sum_{n=1}^N \|x_n\|$. Then $T : \mathfrak{B} \rightarrow D^N$ defined by $T(\sum_{n=1}^N x_n \mathbb{1}_{A_n}) = (\alpha_1 x_1, \dots, \alpha_N x_N)$, is an isometric order preserving isomorphism. Therefore,

$$D^+ \text{ is generating} \iff (D^N)^+ \text{ is generating} \iff \mathfrak{B}^+ \text{ is generating.} \quad (6.18)$$

6.4.16 Theorem. D^+ is closed and generating if and only if $\mathfrak{B}^+$ is. Moreover, if D^+ is closed and generating and $C > 0$, then $\|\cdot\|$ is C -absolutely dominating if and only if $\|\cdot\|_{\mathfrak{B}}$ is.

Proof. This follows from Lemma 6.4.11 and Theorem 6.4.12. $\square$

§6.5 An extension of the Bochner integral

We present the definition of a Banach cover (Definition 6.5.2), some examples, and use this notion to extend the Bochner integral to functions with values in such a vector space (6.5.6).

The next result follows by definition of the Bochner integral or by 6.4.4(c).

6.5.1 Theorem. Let D_1 and D be Banach spaces and suppose $D_1 \subset D$ and the inclusion map is continuous. Suppose $f : X \rightarrow D$ has values in D_1 and is Bochner integrable as map $X \rightarrow D_1$ with integral I . Then f is Bochner integrable as a map $X \rightarrow D$ with integral I .

6.5.2 Definition. Let E be a vector space. Suppose $\mathbb{B}$ is a collection of Banach spaces whose underlying vector spaces are linear subspaces of E . $\mathbb{B}$ is called a *Banach cover* of E if $\bigcup \mathbb{B} = E$ and for all D_1 and D_2 in $\mathbb{B}$ there exists a $D \in \mathbb{B}$ with $D_1, D_2 \subset D$ such that both inclusion maps $D_1 \rightarrow D$ and $D_2 \rightarrow D$ are continuous.

For a $D \in \mathbb{B}$, we write $\|\cdot\|_D$ for its norm if not indicated otherwise.

If E is an ordered vector space, then an *ordered Banach cover* is a Banach cover whose elements are seen as ordered subspaces of E .

6.5.3. A bit of pedantry: strictly speaking, a Banach space is a couple $(D, \|\cdot\|)$ consisting of a vector space D and a norm $\|\cdot\|$. One usually talks about the “Banach space D ”, the norm being understood. Mostly, we adopt that convention but not always. In the context of the above definitions one has to be careful. A Banach cover may contain several Banach spaces with the same underlying vector space, so that a formula like “ $D_1, D_2 \subset D$ ” really is ambiguous. What we mean is only an inclusion relation between the vector spaces and no connection between the norms is assumed a priori.

However, suppose $(D_1, \|\cdot\|_1)$ and $(D_2, \|\cdot\|_2)$ are elements of a Banach cover $\mathbb{B}$ and $D_1 = D_2$. There is a Banach space $(D, \|\cdot\|_D)$ in $\mathbb{B}$ with $D_1, D_2 \subset D$ and with continuous inclusion maps. If a sequence $(x_n)_{n \in \mathbb{N}}$ in $D_1 (= D_2)$ is $\|\cdot\|_1$ -convergent to a and $\|\cdot\|_2$ -convergent to b , then it is $\|\cdot\|_D$ -convergent to a and b , so $a = b$. Hence, by the Closed Graph Theorem the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.

Similarly, if $(D_1, \|\cdot\|_1)$ and $(D_2, \|\cdot\|_2)$ are elements of a Banach cover and $D_1 \subset D_2$, then the inclusion map $D_1 \rightarrow D_2$ automatically is $\|\cdot\|_1$ - $\|\cdot\|_2$ -continuous.

6.5.4 Example. Let E be a uniformly complete Riesz space. The set of principal ideals $\mathbb{B} = \{(E_u, \|\cdot\|_u) : u \in E^+, u \neq 0\}$ is an ordered Banach cover of E : for $u, v \in E^+$ with $u, v > 0$ one has $E_u, E_v \subset E_{u+v}$ and $\|\cdot\|_{u+v} \leq \|\cdot\|_u$ on E_u .

6.5.5 Example. Let $(Y, \mathcal{B}, \nu)$ be a complete σ -finite measure space. Let M be the space of classes of measurable functions $Y \rightarrow \mathbb{R}$. A function $\rho : M \rightarrow [0, \infty]$ is called an *function norm* if (i) $\rho(f) = 0 \iff f = 0$ a.e., (ii) $\rho(\alpha f) = |\alpha|\rho(f)$ (where $0 \cdot \infty = 0$), (iii) $\rho(|f|) = \rho(f)$, (iv) $\rho(f + g) \leq \rho(f) + \rho(g)$, (v) $0 \leq f \leq g$ a.e. implies $\rho(f) \leq \rho(g)$ for $f, g \in M$ and $\alpha \in \mathbb{R}$. For such function norm ρ the set

$$L_\rho = \{f \in M : \rho(f) < \infty\} \quad (6.19)$$

is a normed Riesz space called a *Köthe space* (see [55, Ch. III §18] or [66, Ch. 1 §9]).

A Köthe space L_ρ is complete, i.e., a Banach lattice if and only if for all $u_n \in L_\rho^+$ with $\sum_{n \in \mathbb{N}} \rho(u_n) < \infty$, $\sum_{n \in \mathbb{N}} u_n$ is an element of L_ρ ([55, Theorem 19.3]). Examples of complete Köthe spaces are Orlicz spaces ([55, §20]). In particular, the Banach spaces $L^p(\nu)$ for $p \geq 1$ are Köthe spaces. We introduce other examples:

For $w \in M^+$ with $w > 0$ a.e. define $\rho_w : M \rightarrow [0, \infty]$ by

$$\rho_w(f) := \int |f|w \, d\nu. \quad (6.20)$$

Then ρ_w is a function norm. Both the set $\{L_{\rho_w} : w \in M^+, w > 0 \text{ a.e.}\}$ and the set of all complete Köthe spaces are Banach covers of M (see Appendix 6.B).

6.5.6. Let E be a vector space with a Banach cover $\mathbb{B}$. Let $(X, \mathcal{A}, \mu)$ be a complete σ -finite measure space, $\mu \neq 0$.

- (1) For $D \subset \mathbb{B}$ denote by B_D the vector space of all Bochner integrable functions $X \rightarrow D$, and, by b_D the Bochner integral $B_D \rightarrow D$.
- (2) Let $D_1, D_2 \in \mathbb{B}$, $f_1 \in B_{D_1}$, $f_2 \in B_{D_2}$, $f_1 = f_2$ μ -a.e. Then $b_{D_1}(f_1) = b_{D_2}(f_2)$. *Proof.* Choose $D \in \mathbb{B}$ as in Definition 6.5.2. Then $f_1, f_2 \in B_D$ and $b_{D_1}(f_1) = b_D(f_1) = b_D(f_2) = b_{D_2}(f_2)$.
- (3) If $D_1, D_2 \in \mathbb{B}$ have the same underlying vector space, then $B_{D_1} = B_{D_2}$ since the identity map $D_1 \rightarrow D_2$ is a homeomorphism (see 6.5.3).
- (4) We call a function $f : X \rightarrow E$ $\mathbb{B}$ -integrable if there is a $D \in \mathbb{B}$ such that f is μ -a.e. equal to some element of B_D .
- (5) By $\mathfrak{U}$ we indicate the vector space of all μ -equivalence classes of $\mathbb{B}$ -integrable functions.

For $D \in \mathbb{B}$ we have a natural map $T_D : B_D \rightarrow \mathfrak{U}$, assigning to every element of B_D its μ -equivalence class. The space $T_D(B_D)$ is a Banach space, which we indicate by $\mathfrak{B}_D$.⁶ We write $\mathfrak{b}_D$ for the map $\mathfrak{B}_D \rightarrow D$ determined by

$$\mathfrak{b}_D(T_D(f)) = b_D(f) \quad (f \in B_D). \quad (6.21)$$

$\mathfrak{U}$ is the union of the sets $\mathfrak{B}_D$. By (2) there is a unique $\mathfrak{u} : \mathfrak{U} \rightarrow E$ determined by

$$\mathfrak{u}(f) = \mathfrak{b}_D(f) \quad (D \in \mathbb{B}, f \in \mathfrak{B}_D). \quad (6.22)$$

The above leads to the following theorem.

⁶Even though we use the same notation as in Section 6.4, see 6.4.3, the meaning of $\mathfrak{B}_D$ is slightly different.

6.5.7 Theorem. $\mathfrak{U}$ is a vector space, $\mathfrak{u}$ is linear and $\{\mathfrak{B}_D : D \in \mathbb{B}\}$ is a Banach cover of $\mathfrak{U}$.

6.5.8.

- (1) For any vector space the finite dimensional linear subspaces form a Banach cover.
- (2) If E is a Banach space, then $\{E\}$ is a Banach cover.
- (3) If E is a Banach space and $\mathbb{B}$ is a Banach cover with $E \in \mathbb{B}$, then $\mathfrak{U}$ is just the space of (classes of) Bochner integrable functions $X \rightarrow E$ and $\mathfrak{u}$ is the Bochner integral.
- (4) A special case of (3): Let E be a uniformly complete Riesz space with a unit e and let $\mathbb{B}$ be the Banach cover of principal ideals as in Example 6.5.4. Then $(E, \|\cdot\|_e) \in \mathbb{B}$.

6.5.9. Let E be a vector space and $\mathbb{B}_1$ and $\mathbb{B}_2$ be Banach covers of E . Suppose that for all $D_1 \in \mathbb{B}_1$ there exists a $D_2 \in \mathbb{B}_2$ with $D_1 \subset D_2$ such that the inclusion map is continuous. Write $\mathfrak{U}_i, \mathfrak{u}_i$ for the set of $\mathbb{B}_i$ -integrable functions and the $\mathbb{B}_i$ -integral, for $i \in \{1, 2\}$. Then $\mathfrak{U}_1 \subset \mathfrak{U}_2$ and $\mathfrak{u}_1 = \mathfrak{u}_2$ on $\mathfrak{U}_1$.

6.5.10 Example (Different covers and different integrals). Let $(E, \|\cdot\|)$ be an infinite dimensional Banach space. Let $T : E \rightarrow E$ a linear bijection that is not continuous; say, there exist $x_n \in E$ with $\|x_n\| = 2^{-n}$, $Tx_n = x_n$, $T(\sum_{n \in \mathbb{N}} x_n) \neq \sum_{n \in \mathbb{N}} x_n$. Define $\|x\|_T = \|Tx\|$ for $x \in E$. Then $(E, \|\cdot\|_T)$ is a Banach space. The map $f : \mathbb{N} \rightarrow E$ given by $f(n) = x_n$ is Bochner integrable in $(E, \|\cdot\|)$ and in $(E, \|\cdot\|_T)$, but the integrals do not agree. Whence with $\mathbb{B}_1 = \{(E, \|\cdot\|)\}$ and $\mathbb{B}_2 = \{(E, \|\cdot\|_T)\}$ we have $f \in \mathfrak{U}_1 \cap \mathfrak{U}_2$ but $\mathfrak{u}_1(f) \neq \mathfrak{u}_2(f)$.

As an immediate consequence of Theorem 6.4.5 we obtain the following theorem.

6.5.11 Theorem. Suppose E is a Riesz space and $\mathbb{B}$ is a Banach cover of E that consists of Banach lattices that are Riesz subspaces of E . Then $\mathfrak{U}$ is a Riesz space and $\mathfrak{u}$ is order preserving.

Proof. This is a consequence of Theorem 6.4.5. □

6.5.12. Whenever E is an ordered vector space and $\mathbb{B}$ an ordered Banach cover of E , then $\mathfrak{U}$ is an ordered vector space. In order for $\mathfrak{u}$ to be order preserving, one needs a condition on $\mathbb{B}$. This and other matters will be treated in §6.7. A sufficient condition turns out to be closedness of D^+ for every $D \in \mathbb{B}$ (see Theorem 6.4.6 and Theorem 6.7.1). First we will see in §6.6 that all Archimedean directed ordered vector spaces admit such ordered Banach covers. (The Archimedean property is necessary as follows easily from Theorem 6.3.2).

6.5.13. Whenever E is a vector space and $\mathbb{B}$ is a Banach cover of E , then the set

$$\{A \subset E : \text{there exists a } D \in \mathbb{B} \text{ such that } A \text{ is bounded in } D\} \quad (6.23)$$

forms a bornology on E (we refer to the book of Hogbe-Nlend [49] for the theory of bornologies).

§6.6 Covers of ordered Banach spaces with closed generating cones

In this section E is an Archimedean directed ordered vector space. $\mathbb{B}$ is the collection of all ordered Banach spaces that ordered linear subspaces of E whose cones are closed and generating.

We intend to prove that $\mathbb{B}$ is a Banach cover of E (Theorem 6.6.5).

6.6.1 Lemma. *Let $(D_1, \|\cdot\|_1)$, $(D_2, \|\cdot\|_2)$ be in $\mathbb{B}$. Let $z_1, z_2, \dots \in D_1 \cap D_2$, $a \in D_1$, $b \in D_2$, $\|z_n - a\|_1 \rightarrow 0$, $\|z_n - b\|_2 \rightarrow 0$. Then $a = b$.*

Proof. We may assume $\sum_{n \in \mathbb{N}} \|z_n - a\|_1 < \infty$ and $\sum_{n \in \mathbb{N}} \|z_n - b\|_2 < \infty$. Then $z_n \xrightarrow{u} a$ in D_1 and $z_n \xrightarrow{u} b$ in D_2 by Theorem 6.3.9. Then $z_n \xrightarrow{u} a$ and $z_n \xrightarrow{u} b$ in E . Because E is Archimedean, $a = b$. $\square$

6.6.2. If D is an ordered Banach space with closed generating cone D^+ , under each of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$, then these norms are equivalent. Indeed, the identity map $(D, \|\cdot\|_1) \rightarrow (D, \|\cdot\|_2)$ has a closed graph by Lemma 6.6.1.

6.6.3 Theorem. *Let $(D_1, \|\cdot\|_1)$, $(D_2, \|\cdot\|_2)$ be in $\mathbb{B}$. $D_1 + D_2$ is an ordered Banach space with closed generating cone under the norm $\|\cdot\| : D_1 + D_2 \rightarrow [0, \infty)$ defined by*

$$\|z\| := \inf\{\|x\|_1 + \|y\|_2 : x \in D_1, y \in D_2, z = x + y\}. \quad (6.24)$$

Moreover, if $C > 0$ and $\|\cdot\|_1$ and $\|\cdot\|_2$ are C -absolutely dominating, then so is $\|\cdot\|$.

Proof. $D_1 \times D_2$ is a Banach space under the norm $(x, y) \mapsto \|x\|_1 + \|y\|_2$. From Lemma 6.6.1 it follows that $\Delta := \{(a, b) \in D_1 \times D_2 : a = -b\}$ is closed in $D_1 \times D_2$. Then $D_1 \times D_2 / \Delta$ is a Banach space under the quotient norm. This means that $D_1 + D_2$ is a Banach space under $\|\cdot\|$. (In particular, $\|\cdot\|$ is a norm.)

Since $D_1^+ + D_2^+ \subset (D_1 + D_2)^+$, the latter is generating.

We prove that $(D_1 + D_2)^+$ is closed. Let $u_1, u_2, \dots \in D_1 + D_2$, $\sum_{n \in \mathbb{N}} \|u_n\| < \infty$; we prove $u_n \xrightarrow{u} 0$. Choose $x_n \in D_1$, $y_n \in D_2$ with $u_n = x_n + y_n$ and $\sum_{n \in \mathbb{N}} \|x_n\|_1 < \infty$, $\sum_{n \in \mathbb{N}} \|y_n\|_2 < \infty$. Then, see Theorem 6.3.9, $x_n \xrightarrow{u} 0$ in D_1 and $y_n \xrightarrow{u} 0$ in D_2 . It follows that $u_n \xrightarrow{u} 0$ in $D_1 + D_2$. By Theorem 6.3.9 it follows that $(D_1 + D_2)^+$ is closed.

Suppose $C > 0$ is such that $\|\cdot\|_1$ and $\|\cdot\|_2$ are C -absolutely dominating. Let $z \in D_1 + D_2$, $\varepsilon > 0$. Choose $x \in D_1$, $y \in D_2$ with $z = x + y$, $\|x\|_1 + \|y\|_2 \leq \|z\| + \frac{\varepsilon}{3}$. Choose $a \in D_1^+$ with $-a \leq x \leq a$, $\|a\|_1 < C\|x\|_1 + \frac{\varepsilon}{3}$ and $b \in D_2^+$ with $-b \leq y \leq b$, $\|b\|_2 < C\|y\|_2 + \frac{\varepsilon}{3}$. Set $c = a + b$. Then $c \in (D_1 + D_2)^+$, $-c \leq z \leq c$ and $\|c\| \leq \|a\|_1 + \|b\|_2 < C\|x\|_1 + C\|y\|_2 + 2\frac{\varepsilon}{3} < C\|z\| + \varepsilon$. $\square$

6.6.4. Let $x \in E$. (We make a $D \in \mathbb{B}$ with $x \in D$.) Choose $a \in E$ such that $-a \leq x \leq a$. Let $D = \mathbb{R}(a - x) + \mathbb{R}(a + x) = \mathbb{R}a + \mathbb{R}x$. D is a directed ordered vector space. Define $\|\cdot\| : D \rightarrow [0, \infty)$ by

$$\|y\| := \inf\{s \geq 0 : -sa \leq y \leq sa\}. \quad (6.25)$$

Then $\|\cdot\|$ is a norm on D , $-\|y\|a \leq y \leq \|y\|a$ for all $y \in D$ and $\|a\| = 1$. Thus $(D, \|\cdot\|)$ is a directed ordered Banach space. Moreover D^+ is closed: Let $y \in D$, $y_1, y_2, \dots \in D^+$, $\|y - y_n\| < \frac{1}{n}$. Then $y \geq y_n - \frac{1}{n}a \geq -\frac{1}{n}a$, so $y \geq 0$.

$\|\cdot\|$ is 1-absolutely dominating: $\|a\| = 1$, so $\inf\{\|c\| : c \in D^+, -c \leq y \leq c\} \leq \inf\{s \geq 0 : -sa \leq y \leq sa\} = \|y\|$.

Even $\|y\| = \inf\{\|c\| : c \in D^+, -c \leq y \leq c\}$: For $c \in D^+$ with $-c \leq y \leq c$ and $s \geq 0$ such that $c \leq sa$ we have $-sa \leq -c \leq y \leq c \leq sa$ and so $\|y\| \leq \|c\|$.

6.6.5 Theorem. $\mathbb{B}$ is a Banach cover of E . Moreover,

$$\{D \in \mathbb{B} : \|\cdot\|_D \text{ is 1-absolutely dominating}\} \quad (6.26)$$

is a Banach cover of E .

Proof. By 6.6.4 each element of E is contained in an ordered Banach space with closed generating cone (with a 1-absolutely dominating norm). By Theorem 6.6.3 and by definition of the norm, $\mathbb{B}$ forms a Banach cover of E . $\square$

6.6.6. It is reasonable to ask if an analogue of Theorem 6.6.5 holds in the world of Riesz spaces: does every Archimedean Riesz space have a Banach cover consisting of Riesz spaces? The answer is negative.

Let E be the Riesz space of all functions f on $\mathbb{N}$ for which there exist $N \in \mathbb{N}$ and $r, s \in \mathbb{R}$ such that $f(n) = sn + r$ for $n \geq N$. Suppose E has a Banach cover $\mathbb{B}$ consisting of Riesz subspaces of E . There is a $D \in \mathbb{B}$ that contains the constant function $\mathbb{1}$ and the identity map $i : \mathbb{N} \rightarrow \mathbb{N}$. For every $n \in \mathbb{N}$,

$$\mathbb{1}_{\{1, \dots, n\}} = \mathbb{1} \vee (n+1)\mathbb{1} - i \vee n\mathbb{1} \in D. \quad (6.27)$$

It follows that $D = E$, so E is a Banach space under some norm.

But E is the union of an increasing sequence $D_1 \subset D_2 \subset \dots$ of finite dimensional -hence, closed- linear subspaces:

$$D_n = \mathbb{R}\mathbb{1} + \mathbb{R}i + \mathbb{R}\mathbb{1}_{\{1\}} + \dots + \mathbb{R}\mathbb{1}_{\{n\}}. \quad (6.28)$$

By Baire's Category Theorem, some D_n has nonempty interior in E . Then $E = D_n$ and we have a contradiction.

6.6.7. In Theorem 6.6.5 we single out one particular Banach cover $\mathbb{B}$. If we consider only Banach covers consisting of directed spaces with closed cones, this $\mathbb{B}$ is the largest and gives us the largest collection of integrable functions. Without directedness there may not be a largest Banach cover. For instance, consider Example 6.5.10. Impose on E the trivial ordering ($x \leq y$ if and only if $x = y$). Then $E^+ = \{0\}$, and both $\mathbb{B}_1$ and $\mathbb{B}_2$ consist of Banach spaces with closed (but not generating) cones.

§6.7 The integral for an Archimedean ordered vector space

As a consequence of Theorem 6.4.6 we obtain the following extension of Theorem 6.5.11.

6.7.1 Theorem. *Let E be an ordered vector space with an ordered Banach cover $\mathbb{B}$ so that D^+ is closed for all $D \in \mathbb{B}$. Then $\mathbf{u}$ is order preserving. Moreover, E and $\mathfrak{A}$ are Archimedean.*

6.7.2 Lemma. *Let D be an ordered Banach space with a closed generating cone D^+ . Let T be a linear order preserving map of D into an Archimedean ordered vector space H . Then $\ker T$ is closed and $T(D)$ equipped with the norm $\|\cdot\|_q$ given by*

$$\|z\|_q = \inf\{\|x\| : x \in D, Tx = z\}, \quad (6.29)$$

has a closed generating cone $T(D)^+$.

Proof. (I) Let $x_1, x_2, \dots \in \ker T$, $x \in D$, $x_n \rightarrow x$; we prove $x \in \ker T$. We assume $\sum_{n \in \mathbb{N}} \|x - x_n\| < \infty$. By Theorem 6.3.9 $x_n \xrightarrow{u} x$. By Lemma 6.3.8 $Tx_n \xrightarrow{u} Tx$, so $Tx = 0$.

(II) $D/\ker T$ is a Banach space under the quotient norm $\|\cdot\|_Q$. The formula $x + \ker T \mapsto Tx$ describes a linear bijection $D/\ker T \rightarrow T(D)$ and

$$\|x + \ker T\|_Q = \|Tx\|_q \quad (x \in D). \quad (6.30)$$

It follows that $\|\cdot\|_q$ is indeed a norm, turning $T(D)$ into a Banach space.

$T(D^+) \subset T(D)^+$, whence $T(D)$ is directed.

We prove that $T(D)$ satisfies (iii) of Theorem 6.3.9: Let $z_1, z_2, \dots \in T(D)$, $\sum_{n \in \mathbb{N}} \|z_n\|_q < \infty$. Choose $x_n \in D$ such that $Tx_n = z_n$, $\sum_{n \in \mathbb{N}} \|x_n\| < \infty$. Using (iii) for D , $x_n \xrightarrow{u} 0$. Then $z_n = Tx_n \xrightarrow{u} 0$ by Lemma 6.3.8. $\square$

6.7.3. In the proof of Lemma 6.7.2 we mentioned the inclusion $T(D^+) \subset T(D)^+$. This inclusion can be strict. Take $D = H = \mathbb{R}^2$, $T(x, y) = (x, x + y)$. Then $T(D^+) \neq T(D)^+$.

From Theorem 6.3.11, 6.4.4(c) and Lemma 6.7.2 we get:

6.7.4 Theorem. *Let E_1, E_2 be ordered vector spaces, E_i endowed with the Banach cover $\mathbb{B}_i$ consisting of the ordered Banach spaces with closed generating cones. Let $T : E_1 \rightarrow E_2$ be linear and order preserving. If $f : X \rightarrow E_1$ is $\mathbb{B}_1$ -integrable, then $T \circ f : X \rightarrow E_2$ is $\mathbb{B}_2$ -integrable, and $\mathbf{u}_2(T \circ f) = T(\mathbf{u}_1(f))$.*

6.7.5. In view of Theorem 6.3.11 the reader may wonder why in Theorem 6.7.4 T is required to be order preserving and not just order bounded, the more so because of the following considerations. Let D and H be as in Lemma 6.7.2 and T be a linear order bounded map of D into H . As the implication (c) $\implies$ (a) of Theorem 6.3.9 is valid for Archimedean (but not necessarily directed) D , following the lines of the proof of Lemma 6.7.2 $\ker T$ is closed and $T(D)$ equipped with the norm as in (6.29) has a closed cone $T(D)^+$. However, we also need $T(D)$ to be directed and order boundedness of T is no guarantee for that.

An alternative approach might be to drop the directedness condition on the spaces that constitute $\mathbb{B}$. However, the ordered Banach spaces with closed cones may not form a Banach cover.

For an example, let E be ℓ^∞ and let $\mathbb{B}$ be the collection of all ordered Banach spaces that are subspaces of ℓ^∞ and have closed cones. We make $D_1, D_2 \in \mathbb{B}$. For D_1 we take ℓ^∞ with the usual norm $\|\cdot\|_\infty$. Choose a linear bijection $T: \ell^\infty \rightarrow \ell^\infty$ that is not continuous. For $a \in \ell^\infty$ put $a' = (a_1, -a_1, a_2, -a_2, \dots)$. For D_2 we take the vector space $\{a' : a \in \ell^\infty\}$ with the norm $\|\cdot\|_T$ given by $\|a'\|_T = \|Ta\|_\infty$. Then D_2 is a Banach space and D_2^+ , being $\{0\}$, is closed in D_2 . Suppose $\mathbb{B}$ is a Banach cover. Let D be as in Definition 6.5.2. By the Open Mapping Theorem the identity map $D_1 \rightarrow D$ is a homeomorphism. By the continuity of the inclusion map $D_2 \rightarrow D$ there exists a number c such that $\|a'\|_T \leq c\|a'\|_\infty$ for all $a \in \ell^\infty$. Then $\|Ta\|_\infty \leq c\|a'\|_\infty \leq c\|a\|_\infty$ for $a \in \ell^\infty$, so T is continuous. Contradiction.

6.7.6 Theorem. *Let E be an ordered vector space such that $E^\sim$ separates the points of E .⁷ Assume $\mathbb{B}$ is a Banach cover consisting of ordered Banach spaces with closed generating cones. Let $f \in \mathfrak{U}$. Then $\alpha \circ f \in \mathcal{L}^1(\mu)$ for all $\alpha \in E^\sim$. Moreover, $I \in E$ is such that*

$$\alpha(I) = \int \alpha \circ f \, d\mu \quad \text{for all } \alpha \in E^\sim, \quad (6.31)$$

if and only if $I = u(f)$.

Proof. Let $f \in \mathfrak{U}$ and let $I = u(f)$. By Theorem 6.7.4 $\alpha \circ f \in \mathcal{L}^1(\mu)$ for all $\alpha \in E^\sim$ and (6.31) holds. $I = u(f)$ is the only element of E for which (6.31) holds because $E^\sim$ separates the points of E . $\square$

6.7.7 Remark. Functions with values in a Banach space that are Bochner integrable are also Pettis integrable. To some extent the statement of Theorem 6.7.6 is similar. Indeed, the definition of Pettis integrability could be generalised for vector spaces V which are equipped with a set S of linear maps $V \rightarrow \mathbb{R}$ that separates the points of V , in the sense that one calls a function $f: X \rightarrow V$ Pettis integrable if $\alpha \circ f \in \mathcal{L}^1(\mu)$ for all $\alpha \in S$ and there exists a $I \in V$ such that $\alpha(I) = \int \alpha \circ f \, d\mu$ for all $\alpha \in S$. Then Theorem 6.7.6 implies that every $f \in \mathfrak{U}$ is Pettis integrable when considering $V = E$ and $S = E^\sim$. Observe, however, that even for a Riesz space E , $E^\sim$ may be trivial (see, e.g., [55, 5.A]).

§6.8 Comparison with other integrals

In this section $(X, \mathcal{A}, \mu)$ is a complete σ -finite measure space and E is a directed ordered vector space with an ordered Banach cover $\mathbb{B}$ so that D^+ is closed for each $D \in \mathbb{B}$.

In 6.5.6 we have introduced an integral u on a space $\mathfrak{U}$ of $\mathbb{B}$ -integrable functions $X \rightarrow E$.⁸ In [80], starting from a natural integral φ on the space S of all simple functions $X \rightarrow E$ we have made integrals $\varphi_V, \varphi_L, \varphi_{LV}, \dots$ on spaces $S_V, S_L, S_{LV}, \dots$

⁷We write $E^\sim$ for the space of order bounded linear maps $E \rightarrow \mathbb{R}$.

⁸In this section we close an eye for the difference between a function and its equivalence class. There will be no danger of confusion.

There is an elementary connection: S is part of $\mathfrak{U}$ and $\mathfrak{u}$ coincides with φ on S . (Indeed, let $f \in S$. Being a finite set, $f(X)$ is contained in D for some $D \in \mathbb{B}$. Then f is Bochner integrable as a map $X \rightarrow D$.)

In general, S_V and S_L are not subsets of $\mathfrak{U}$, but we can prove that $\mathfrak{u}$ coincides with φ_V on $S_V \cap \mathfrak{U}$ and with φ_L on $S_L \cap \mathfrak{U}$. Better than that: $\mathfrak{u}$ is “compatible” with φ_V in the sense that $\mathfrak{u}$ and φ_V have a common order preserving linear extension $S_V + \mathfrak{U} \rightarrow E$. Similarly, $\mathfrak{u}$ is “compatible with $\varphi_L, \varphi_{LV}, \dots$ ”.

6.8.1 Lemma. (a) Let $f \in \mathfrak{U}, g \in S_V, f \leq g$. Then $\mathfrak{u}(f) \leq \varphi_V(g)$.

(b) Let $f \in \mathfrak{U}, g \in S_L, f \leq g$. Then $\mathfrak{u}(f) \leq \varphi_L(g)$.

Proof. (a) By the definition of φ_V and by the text preceding this lemma we have $\varphi_V(g) = \inf\{\varphi(h) : h \in S, h \geq g\} = \inf\{\mathfrak{u}(h) : h \in S, h \geq g\}$. As $g \geq f$ and $\mathfrak{u}$ is order preserving (Theorem 6.7.1), it follows that $\varphi_V(g) \geq \mathfrak{u}(f)$.

(b) Let $g \in S_L$ and assume $f \leq g$. Let $g_1, g_2 \in S_L^+$ be such that $g = g_1 - g_2$. Let $(B_i)_{i \in \mathbb{N}}$ be a φ -partition for both g_1 and g_2 . Write $A_n = \bigcup_{i=1}^n B_i$ for $n \in \mathbb{N}$. Then $f \mathbb{1}_{A_n} \leq g \mathbb{1}_{A_n}$, thus by (a) (and Theorem 6.7.1)

$$\begin{aligned} \mathfrak{u}(f \mathbb{1}_{A_n}) &\leq \mathfrak{u}(g \mathbb{1}_{A_n}) = \varphi(g \mathbb{1}_{A_n}) = \varphi(g_1 \mathbb{1}_{A_n}) - \varphi(g_2 \mathbb{1}_{A_n}) \\ &\leq \varphi_L(g_1) - \varphi(g_2 \mathbb{1}_{A_k}) \quad (k \in \mathbb{N}, k < n). \end{aligned} \quad (6.32)$$

Which implies $\mathfrak{u}(f \mathbb{1}_{A_n}) + \varphi(g_2 \mathbb{1}_{A_k}) \leq \varphi_L(g_1)$ for all $k < n$. Then letting n tend to ∞ (apply 6.4.4(d): $f(x) \mathbb{1}_{A_n}(x) \rightarrow f(x)$ for all $x \in X$) we obtain

$$\mathfrak{u}(f) \leq \varphi_L(g_1) - \varphi(g_2 \mathbb{1}_{A_k}) \quad (k \in \mathbb{N}), \quad (6.33)$$

from which we conclude $\mathfrak{u}(f) \leq \varphi_L(g)$. $\square$

6.8.2 Theorem.

(a) If $g \in S_{LV}$ and $f \leq g$, then $\mathfrak{u}(f) \leq \varphi_{LV}(g)$.

(b) If S_V is stable, $g \in S_{VLV}$ and $f \leq g$, then $\mathfrak{u}(f) \leq \varphi_{VLV}(g)$.

Proof. Follow the lines of the proof of the lemma with S_V, S_L or S_{VL} instead of S . $\square$

6.8.3 (Comments on Theorem 6.8.2).

(1) The theorem supersedes the lemma because $S_V + S_L \subset S_{LV}$.

(2) As a consequence, $\mathfrak{u} = \varphi_{LV}$ on $\mathfrak{U} \cap S_{LV}$, and $\mathfrak{u} = \varphi_{VLV}$ on $\mathfrak{U} \cap S_{VLV}$ if S_V is stable.

(3) Recall that stability of S_V is necessary for the existence of S_{VLV} .

6.8.4 Theorem. Let E be a uniformly complete Riesz space and $\mathbb{B}$ be the Banach cover of principal ideals (see Example 6.5.4). $\mathfrak{U}$ is a linear subspace of S_{LV} and $\mathfrak{u} = \varphi_{LV}$ on $\mathfrak{U}$.

Proof. Let $u \in E^+$ and let $f : X \rightarrow E_u$ be Bochner integrable. We prove $f \in S_{LV}$.

For simplicity of notation, put $D = E_u$. Let S^D be the space of simple functions $X \rightarrow D$. By [80, Corollary 9.8] we have $f \in (S^D)_{LV}$ and $\mathbf{u}(f) = \varphi_{LV}(f)$. Since D is a Riesz ideal in E , the identity map $D \rightarrow E$ is order continuous. Then [80, Theorem 8.14] implies $f \in S_{LV}$. $\square$

Contrary to Theorem 6.8.4, in [80, Example 9.9(II)] $\mathfrak{U}$ is not a linear subspace of S_{LV} . Then next example shows, in the context of Theorem 6.8.4, that the inclusion may be strict.

6.8.5 Example ($\mathfrak{U} \subsetneq S_{LV}$). For $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ and μ the counting measure and $E = c$. As is mentioned in [80, Examples 9.9(I)], the function $n \mapsto \mathbb{1}_{\{n\}}$ is an element of S_{LV} but not Bochner integrable. With $\mathbb{B}$ the Banach cover of principal ideals, the function $n \mapsto \mathbb{1}_{\{n\}}$ is not $\mathbb{B}$ -integrable (see 6.5.8(4)).

§6.9 An example: Convolution

To illustrate the $\mathbb{B}$ -integral as an extension of the Bochner integral we consider the following situation. (This introduction requires some knowledge of harmonic analysis on locally compact groups, the balance of this section does not.)

Let G be a locally compact group. For $f : G \rightarrow \mathbb{R}$ and $x \in G$ we let $L_x f : G \rightarrow \mathbb{R}$ be the function $y \mapsto f(x^{-1}y)$.

For a finite measure μ on G and f in $L^1(G)$ one defines their convolution product to be the element $\mu * f$ of $L^1(G)$ given for almost every $y \in G$ by

$$(\mu * f)(y) = \int f(x^{-1}y) \, d\mu(x) = \int (L_x f)(y) \, d\mu(x). \quad (6.34)$$

The map $x \mapsto L_x f$ of G into $L^1(G)$ is continuous and bounded, hence Bochner integrable with respect to μ . It is not very difficult to prove that

$$\mu * f = \int L_x f \, d\mu(x). \quad (6.35)$$

Similar statements are true for other spaces of functions instead of $L^1(G)$, such as $L^p(G)$, with $1 < p < \infty$, and $C_0(G)$, the space of continuous functions that vanish at infinity.

But consider the space $C(G)$ of all continuous functions on G . The integrals $\int f(x^{-1}y) \, d\mu(x)$ will not exist for all $f \in C(G)$, $y \in G$ and all finite measures μ , but they do if μ has compact support. Thus, one can reasonably define $\mu * f$ for $f \in C(G)$ and compactly supported μ . However, there is no natural norm on $C(G)$ (except, of course, if G is compact), so we cannot speak of $\int L_x f \, d\mu(x)$ as a Bochner integral. We will see that, at least for σ -compact G , it is a $\mathbb{B}$ -integral where $\mathbb{B}$ is the Banach cover of $C(G)$ that consists of the principal (Riesz) ideals.

6.9.1 Theorem. *Let G be a σ -compact locally compact group. For every $f \in C(G)$ there exists a $w \in C(G)^+$ such that*

$$\begin{aligned} & \text{Every } L_x f \quad (x \in G) \text{ lies in the principal ideal } C(G)_w; \\ & x \mapsto L_x f \text{ is continuous relative to } \|\cdot\|_w. \end{aligned} \quad (6.36)$$

Proof. Choose compact $K_1 \subset K_2 \subset \dots$ such that K_1 is a neighbourhood of e ; $K_n = K_n^{-1}$; $K_n K_n \subset K_{n+1}$, $\bigcup_{n \in \mathbb{N}} K_n = G$.

For $x \in G$, define $[x]$ to be the smallest n with $x \in K_n$. Then $[x] = [x^{-1}]$, $[xy] \leq 1 + [x] \vee [y]$ for all $x, y \in G$ by definition.

Let $f \in C(G)$. Define $u, v : G \rightarrow [1, \infty)$ as follows:

$$u(x) := 1 + \sup |f(K_{n+1})|, \quad v(x) = [x]u(x) \quad \text{if } [x] = n. \quad (6.37)$$

(1) If $x, y \in G$, $[x] \leq [y] = n$, then $x, y \in K_n$, so

$$|f(xy)| \leq \sup |f(K_{n+1})| \leq u(y). \quad (6.38)$$

(2) Hence, for all $x, y \in G$: $|f(x^{-1}y)| \leq u(x) \vee u(y) \leq u(x)v(y)$, i.e., $|L_x f| \leq u(x)v$.

Let $a \in G$, $\varepsilon > 0$. In (5), using (3) and (4), we show the existence of a neighbourhood U of e with

$$x \in aU \implies |L_x f - L_a f| \leq \varepsilon v. \quad (6.39)$$

Choose a p with $a \in K_p$.

(3) K_1 contains an open set V containing e . We make a $q \in \mathbb{N}$ with

$$x \in aV \implies |L_x f - L_a f| \leq \varepsilon v \text{ on } G \setminus K_q. \quad (6.40)$$

Let $x \in aV$. Then $x, a \in K_{p+1}$, so $[x], [a] \leq p+1$. By (1):

$$|(L_x f - L_a f)(y)| \leq |f(x^{-1}y)| + |f(a^{-1}y)| \leq 2u(y) \text{ if } [y] \geq p+1. \quad (6.41)$$

Moreover, $\varepsilon v(y) = \varepsilon[y]u(y) \geq 2u(y)$ if $[y] \geq \frac{2}{\varepsilon}$. Take $q \in \mathbb{N}$ with $q \geq p+1$ and $q \geq \frac{2}{\varepsilon}$. For $y \in G \setminus K_q$ we have $[y] > q$, so $|(L_x f - L_a f)(y)| \leq \varepsilon v(y)$.

(4) We show there exists an open set W containing e with

$$x \in aW \implies |L_x f - L_a f| \leq \varepsilon v \text{ on } K_q. \quad (6.42)$$

The function $(x, y) \mapsto |(L_x f - L_a f)(y)|$ on $G \times G$ is continuous and K_q is compact. Hence by [80, Theorem 8.15] the function $G \mapsto [0, \infty)$

$$x \mapsto \sup_{y \in K_q} |(L_x f - L_a f)(y)| \quad (6.43)$$

is continuous. Its value at a is 0, so there exists an open set W containing e with

$$\sup_{y \in K_q} |(L_x f - L_a f)(y)| \leq \varepsilon \quad (x \in aW). \quad (6.44)$$

As $v(y) \geq 1$ for all y we obtain (6.42).

(5) With $U = V \cap W$ we have

$$x \in aU \implies |L_x f - L_a f| \leq \varepsilon v \text{ on } G. \quad (6.45)$$

Therefore, to prove the theorem it is sufficient to show there exists a continuous function $w \geq v$:

(6) Set $\alpha_n = n(1 + \sup |f(K_{n+1})|)$ for all n ; then

$$[x] = n \implies v(x) = \alpha_n. \quad (6.46)$$

Put $K_0 = \emptyset$. For all $n \in \mathbb{N}$ we have $VK_{n-1} \subset K_1 K_{n-1} \subset K_n$, so K_{n-1} is a subset of K_n° , the interior of K_n . By Urysohn [27, Theorem VII.4.1 and Theorem XI.1.2] for all n there is a continuous $g_n : G \rightarrow [0, 1]$ with

$$g_n = 0 \text{ on } K_{n-1}, \quad g_n = 1 \text{ on } G \setminus K_n^\circ \supset G \setminus K_n. \quad (6.47)$$

Let $b \in G$. There is a l with $b \in K_l$. bV is an open set containing b . As $bV \subset K_{l+1}$ and $g_n = 0$ on K_{n-1} we have: $g_n = 0$ on bV as soon as $n \geq l + 2$.

Hence, $w := \sum_{n \in \mathbb{N}} \alpha_{n+1} g_n$ is a continuous function $G \rightarrow [0, \infty)$. For every $x \in G$ there is an n with $[x] = n$; then $x \notin K_{n-1}$, $g_{n-1}(x) = 1$ and $w(x) \geq \alpha_n = v(x)$. $\square$

6.9.2 Theorem. *Let G be a σ -compact locally compact group. Let μ be a finite measure on the Borel σ -algebra of G with a compact support. Let $\mathbb{B}$ be the Banach cover of $C(G)$ consisting of the principal ideals as in Example 6.5.4. Then for every $f \in C(G)$ the function $x \mapsto L_x f$ is $\mathbb{B}$ -integrable and its integral is the “convolution product” $\mu * f$:*

$$(\mu * f)(y) = \int f(x^{-1}y) \, d\mu(x) \quad (y \in G). \quad (6.48)$$

Proof. By Theorem 6.9.1 there exists a $w \in C(G)^+$ such that (6.36) holds. This implies that the map $x \mapsto L_x f$ is Borel measurable and $\{L_x f : x \in G\}$ is separable in $C(G)_w$. As $x \mapsto \|L_x f\|_w$ is continuous and thus bounded on the support of μ , the map $x \mapsto L_x f$ is $\mathbb{B}$ -integrable (see 6.4.4(a)). That the integral is equal to $\mu * f$ follows by 6.4.4(c). $\square$

6.9.3 Remark. Theorem 6.9.2 compares to [80, Example 8.17] in the sense that in both situations the convolution is equal to an integral of the translation. Though the situation is slightly different in the sense that in Theorem 6.9.2 we consider σ -compact locally compact groups, while in [80, Example 8.17] we considered metric locally compact groups (the fact that in [80, Example 8.17] $L_x f(t) = f(tx^{-1})$ is reminiscent for its statement).

Appendix

§6.A Appendix: Absolutely dominating norms

In this section, $C > 0$ and D is an ordered Banach space with closed generating cone and with a C -absolutely dominating norm $\|\cdot\|$.

As is mentioned in 6.3.5, we show that $C \geq 1$ (6.A.1) and that for every $\varepsilon > 0$ there exists an equivalent $(1 + \varepsilon)$ -absolutely norm (Theorem 6.A.3). Furthermore, we discuss (in 6.A.4 – 6.A.10) whenever there exists an equivalent norm $\|\cdot\|_1$ for which

$$\|x\|_1 = \inf\{\|a\|_1 : a \in D^+, -a \leq x \leq a\} \quad (x \in D). \quad (6.49)$$

This is done by means of the norm $\mathcal{N}$ introduced in 6.A.2. Example 6.A.11 illustrates that the existence of such equivalent norm may fail.

6.A.1 (C has to be ≥ 1). Suppose that $C < 1$. Choose $C' > 0$ such that $C < C' < 1$. For all $a \in D^+$ with $a \neq 0$ there exists a $b \in D^+$ with $a \leq b$ and $\|b\| \leq C'\|a\|$. Let $a \in D^+$ with $\|a\| = 1$. Iteratively one obtains a sequence $a \leq a_1 \leq a_2 \leq \dots$ with $\|a_1\| \leq C'$ and $\|a_{n+1}\| \leq C'\|a_n\|$ for all n . Then $\sum_{n \in \mathbb{N}} \|a_n\| < \infty$ and thus $a_n \xrightarrow{u} 0$ by Theorem 6.3.9, which contradicts $0 < a \leq a_n$.

6.A.2. Define $\mathcal{N} : D \rightarrow [0, \infty)$ by

$$\mathcal{N}(x) = \inf\{\|a\| : a \in D^+, -a \leq x \leq a\}. \quad (6.50)$$

$\mathcal{N}$ is a seminorm, and actually a norm because (see Theorem 6.3.9)

$$\mathcal{N}(x) = 0 \iff \text{there is an } a^* \in D^+ \text{ with } -\frac{1}{n}a^* \leq x \leq \frac{1}{n}a^* \quad (n \in \mathbb{N}). \quad (6.51)$$

Because $\|\cdot\|$ is C -absolutely dominating one has $\mathcal{N} \leq C\|\cdot\|$.

6.A.3 Theorem. For all $\varepsilon > 0$ there exists an equivalent norm ρ on D , for which

$$(1 + \varepsilon)^2 \rho(x) \geq \inf\{\rho(a) : a \in D^+, -a \leq x \leq a\} \quad (x \in D). \quad (6.52)$$

Proof. Define $\rho := \varepsilon \mathcal{N} + \|\cdot\|$. ρ is a norm which is equivalent to $\|\cdot\|$, since $\mathcal{N} \leq C\|\cdot\|$.

Let $x \in D$, $x \neq 0$. Choose $a \in D^+$ with $-a \leq x \leq a$ such that $\|a\| \leq (1 + \varepsilon)\mathcal{N}(x)$. Note that by definition of $\mathcal{N}$ we have $\mathcal{N}(a) \leq \|a\|$. Whence

$$\frac{\rho(a)}{\rho(x)} = \frac{\mathcal{N}(a) + \varepsilon\|a\|}{\mathcal{N}(x) + \varepsilon\|x\|} \leq \frac{\mathcal{N}(a) + \varepsilon\|a\|}{\mathcal{N}(x)} \leq \frac{\|a\| + \varepsilon\|a\|}{\mathcal{N}(x)} \leq (1 + \varepsilon) \frac{\|a\|}{\mathcal{N}(x)} \leq (1 + \varepsilon)^2. \quad (6.53)$$

□

6.A.4. Suppose $\|\cdot\|_1$ is a norm equivalent to $\|\cdot\|$ for which there exists a $C > 0$ such that

$$C\|x\|_1 = \inf\{\|a\|_1 : a \in D^+, -a \leq x \leq a\} \quad (x \in D). \quad (6.54)$$

Then it is straightforward to show that $\mathcal{N}$ is equivalent to $\|\cdot\|$.

6.A.5. $\mathcal{N}$ is equivalent to $\|\cdot\|$ if and only if there exists a $c > 0$ such that $\mathcal{N} \geq c\|\cdot\|$. The latter is true if and only if

$$-\frac{1}{n}a^* \leq x_n \leq \frac{1}{n}a^* \quad (n \in \mathbb{N}) \implies \lim_{n \rightarrow \infty} \|x_n\| = 0. \quad (6.55)$$

6.A.6 Theorem.

$$\mathcal{N}(x) = \inf\{\mathcal{N}(a) : a \in D^+, -a \leq x \leq a\}. \quad (6.56)$$

Proof. ($\leq$) Take $a \in D^+$, $-a \leq x \leq a$; we prove $\mathcal{N}(x) \leq \mathcal{N}(a)$. For all $b \geq a$, $-b \leq x \leq b$, whence $\mathcal{N}(x) \leq \|b\|$. Thus $\mathcal{N}(x) \leq \inf\{\|b\| : a \leq b\} = \mathcal{N}(a)$, the latter by definition of $\mathcal{N}(a)$.

($\geq$) Take $\varepsilon > 0$. Choose $a \in D^+$, $-a \leq x \leq a$, $\|a\| \leq \mathcal{N}(x) + \varepsilon$. Then $\mathcal{N}(a) \leq \|a\| \leq \mathcal{N}(x) + \varepsilon$. $\square$

6.A.7 Theorem. [2, Theorem 2.38] For an ordered normed vector space E the following are equivalent.

- (a) The cone E^+ is normal.
- (b) The normed space E admits an equivalent monotone norm.
- (c) There is a $c > 0$ such that $0 \leq x \leq y$ implies $\|x\| \leq c\|y\|$.

6.A.8. Suppose D is an ordered Banach space with closed generating cone and suppose there exists a $c > 0$ such that

$$\|x\| \leq c \inf\{\|a\| : a \in D^+, -a \leq x \leq a\} = c\mathcal{N}(x). \quad (6.57)$$

Then for $x \in D^+$, $a \in D^+$ with $x \leq a$ one has $\|x\| \leq c\|a\|$.

6.A.9. Suppose D is an ordered Banach space with closed generating cone and $\|\cdot\|$ is monotone. Let $x \in D$ and $a \in D^+$ be such that

$$-a \leq x \leq a. \quad (6.58)$$

Then $0 \leq x + a \leq 2a$ and whence $\|x\| \leq \|x + a\| + \|a\| \leq 3\|a\|$. Thus

$$\|x\| \leq 3 \inf\{\|a\| : a \in D^+, -a \leq x \leq a\} = 3\mathcal{N}(x). \quad (6.59)$$

We conclude:

6.A.10 Theorem. Let D be an ordered Banach space with closed generating cone. The following are equivalent

- (a) There exist a norm $\|\cdot\|_1$ that is equivalent to $\|\cdot\|$ for which

$$\|x\|_1 = \inf\{\|a\|_1 : a \in D^+, -a \leq x \leq a\} \quad (x \in D). \quad (6.60)$$

- (b) There exists a $c > 0$ such that $\mathcal{N} \geq c\|\cdot\|$.
- (c) There exists a monotone norm that is equivalent to $\|\cdot\|$.
- (d) E^+ is normal.

In the following we give an example of an ordered Banach space with closed generating cone D for which none of (a) – (d) of Theorem 6.A.10 holds.

6.A.11 Example. Let D be ℓ^1 with its natural norm. Define $T : \ell^1 \rightarrow \mathbb{R}^{\mathbb{N}}$ by

$$Tx = (x_1, x_1 + x_2, x_1 + x_2 + x_3, \dots). \quad (6.61)$$

As T is linear, D is an ordered vector space under the relation $\preceq$

$$x \preceq y \iff Tx \leq Ty. \quad (6.62)$$

The positive cone of ℓ^1 is included in D^+ , whence D is directed. Moreover, D^+ is closed and so $\|\cdot\|$ is absolutely dominating. With $x_n = (1, -1, 1, -1, \dots, \pm 1, 0, 0, \dots)$ and $a = (1, 0, 0, \dots)$ we have $-a \leq x_n \leq a$ and $\|x_n\| = n$, $\|a\| = 1$.

§6.B Appendix: Banach cover of Köthe spaces

In this section $(Y, \mathcal{B}, \nu)$ is a complete σ -finite measure space and M is the space of classes of measurable functions $Y \rightarrow \mathbb{R}$ as in Example 6.5.5.

6.B.1 Lemma. $\{L_{\rho_w} : w \in M^+, w > 0 \text{ a.e.}\}$ is a Banach cover of M .

Proof. L_{ρ_w} is complete since $f \mapsto fw$ is an isometric bijection $M_w \rightarrow L^1(\nu)$.

Let $f \in M$. We show there exists a $w \in M^+$, $w > 0$ a.e., with $f \in L_{\rho_w}$. By the σ -finiteness of ν there is a $u \in L^1(\nu)$, $u > 0$ a.e. Put $w = (|f| + 1)^{-1}u$. Then $w \in M^+$, $w > 0$ a.e., and $f \in L_{\rho_w}$ because $\int |f|w \, d\nu \leq \int u \, d\nu < \infty$.

If $w, v \in M^+$, $w > 0, v > 0$ a.e., then $w \wedge v > 0$ a.e., $L_{\rho_{w \wedge v}}$ is a subset of L_{ρ_w} and L_{ρ_v} and $\rho_{w \wedge v} \leq \rho_w$ on L_{ρ_w} and $\rho_{w \wedge v} \leq \rho_v$ on L_{ρ_v} . $\square$

6.B.2 Theorem. The complete Köthe spaces form a Banach cover of M .

Proof. Let ρ be a function norm and L_ρ be complete. If $f_n \in M^+$ for $n \in \mathbb{N}$ and $\sum_{n \in \mathbb{N}} \rho(f_n) < \infty$, then $\sum_{n \in \mathbb{N}} |f_n| \in L_\rho$, so $\sum_{n \in \mathbb{N}} |f_n| < \infty$ a.e. and $f_n \rightarrow 0$ a.e. We will use this fact below.

By Lemma 6.B.1 it suffices to prove the following. Let ρ_1, ρ_2 be function norms, L_{ρ_1} and L_{ρ_2} complete. We make a function norm ρ such that L_ρ is complete and $\rho \leq \rho_1, \rho \leq \rho_2$. (Then $L_{\rho_1}, L_{\rho_2} \subset L_\rho$ and we are done.)

Define $\rho : M \rightarrow [0, \infty]$ by

$$\rho(f) = \inf\{\rho_1(g) + \rho_2(h) : g, h \in M^+, g + h \geq |f|\}. \quad (6.63)$$

If $\rho(f) = 0$, choose g_n, h_n with $g_n + h_n \geq |f|$, $\rho_1(g_n) + \rho_2(h_n) \leq 2^{-n}$. Then (by the above), $g_n \rightarrow 0$ a.e., $h_n \rightarrow 0$ a.e. Hence, $f = 0$ a.e.

It follows easily that ρ is a function norm. Obviously, $\rho \leq \rho_1, \rho \leq \rho_2$. For the completeness of L_ρ : Let $u_1, u_2, \dots \in L_\rho^+$, $\sum_{n \in \mathbb{N}} \rho(u_n) < \infty$. Choose $g_n, h_n \in M^+$, $g_n + h_n \geq u_n$, $\rho_1(g_n) + \rho_2(h_n) < \rho(u_n) + 2^{-n}$. Then $\sum_{n \in \mathbb{N}} \rho_1(g_n) < \infty$, so $\sum_{n \in \mathbb{N}} g_n \in L_{\rho_1}$, $\rho_1(\sum_{n \in \mathbb{N}} g_n) < \infty$. Similarly $\rho_2(\sum_{n \in \mathbb{N}} h_n) < \infty$. Then $\sum_{n \in \mathbb{N}} u_n \in L_\rho$. $\square$

§6.C Miscellaneous: Additional ideas, remarks etc.

6.C.1. Let D be a directed ordered Banach space. By Theorem 6.3.2:

$$D^+ \text{ is closed} \implies (D')^+ \text{ and } D^{\sim+} \text{ separate the points of } D. \quad (6.64)$$

By Theorem 6.3.3:

$$D^+ \text{ is closed} \implies \|\cdot\| \text{ is absolutely dominating.} \quad (6.65)$$

Moreover, the following implication holds

$$D^{\sim+} \text{ separates the point of } D \implies D \text{ is Archimedean.} \quad (6.66)$$

By Theorem 6.3.11:

$$D^+ \text{ is closed} \implies D^{\sim} \subset D'. \quad (6.67)$$

Example 6.C.2 illustrates that the reverse implications of (6.64) and (6.65) may fail. Example 6.C.3 illustrates that the reverse implications of (6.66) and (6.67) may fail as well.

6.C.2 Example. We give an example of an ordered Banach space D with

- (a) D^+ is generating,
- (b) D^+ is not closed,
- (c) $\|\cdot\|$ is absolutely dominating,
- (d) $(D')^+$ and $D^{\sim+}$ separate the points of D .

Take $D = \mathbb{R}^2$ with the Euclidean norm but with the following unorthodox ordering:

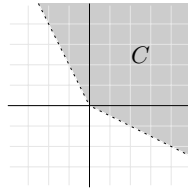


Figure 6.C.1: The cone C .

For $x = (x_1, x_2) \in \mathbb{R}^2$ put $f(x) = 2x_1 + x_2$ and $g(x) = x_1 + 2x_2$. Define the cone C as follows (see Figure 6.C.1): $x \in C$ if and only if both $f(x) > 0$ and $g(x) > 0$, or $x = 0$. Then under the order, $\leq_C$, determined by C (i.e. $x \leq_C y \iff y - x \in C$), the norm is absolutely dominating because $(\mathbb{R}^2)^+ \subset C$. But $C (= D^+)$, is not closed.

Both f, g are in $D^{\sim+}$. As $x > 0$ if and only if $f(x) > 0$ and $g(x) > 0$, $D^{\sim+}$ separates the points of D .

6.C.3 Example. We give an example of an ordered Banach space D with

- (a) D is Archimedean,

- (b) $D^\sim = \{0\}$,
- (c) D^+ is not closed.

Let M be the space of equivalence classes of Borel functions on $[0, 1]$. Let $\mathfrak{x}$ be the identity map $[0, 1] \rightarrow \mathbb{R}$, in M or in $C[0, 1]$. Moreover, let $\#X$ denote the cardinality of X .

We show that there exists a linear bijection $T : C[0, 1] \rightarrow M$ with $T(\mathfrak{x}^n) = n \mathfrak{x}^n$ for $n \in \mathbb{N}$. If T is such linear bijection, then with D being $C[0, 1]$ equipped with the norm $\|\cdot\|_\infty$ but ordered such that T is an order isomorphism, we have $D^\sim = \{0\}$ because $M^\sim = \{0\}$ (see de Jonge and van Rooij [55, 5.A]), and D^+ is not $\|\cdot\|_\infty$ -closed as T is not continuous (see Theorem 6.3.11).

We can make T by the following steps. A Hamel basis for a vector space E is a maximal linearly independent subset of E .

- (a) Let X be a Hamel base of E and let $c_{00}(X)$ be the space of all functions $X \rightarrow \mathbb{R}$ with finite support. Then the formula $x \mapsto \mathbb{1}_{\{x\}}$ determines a linear bijection $E \rightarrow c_{00}(X)$.
- (b) Hence: Let E, F be vector spaces with Hamel bases X, Y . Then every bijection $X \rightarrow Y$ extends to a linear bijection $E \rightarrow F$.
- (c) If Z is a separable metric space then $\#Z \leq \#\mathbb{R}$. *Proof.* Choose a dense set $\{a_1, a_2, \dots\}$. For $z \in Z$ choose $n(1), n(2), \dots$ with $a_{n(i)} \rightarrow z$, and put $S_z = \{n(1), n(2), \dots\}$. Then $z \mapsto S_z$ is an injection $Z \rightarrow \mathcal{P}(\mathbb{N})$, so $\#Z \leq \#\mathcal{P}(\mathbb{N}) = \#\mathbb{R}$.
- (d) The functions $\mathfrak{x}^s$, $s \in [0, \infty)$, are linearly independent in $C[0, 1]$, so there is a Hamel basis X in $C[0, 1]$ with $\mathfrak{x}^s \in X$ for all $s \in [0, \infty)$. Then by (c)

$$\#\mathbb{R} \leq \#X \leq \#C[0, 1] \leq \#\mathbb{R}. \quad (6.68)$$

Similarly, there is a Hamel basis Y in M with $\mathfrak{x}^s \in Y$ for $s \in [0, \infty)$, whence (using the arctan for the inclusion $M \rightarrow L^\infty[0, 1]$ and (c) for the last inequality)

$$\#\mathbb{R} \leq \#Y \leq \#M \leq \#L^\infty[0, 1] \leq \#L^2[0, 1] \leq \#\mathbb{R}. \quad (6.69)$$

- (e) Putting $X' := X \setminus \{\mathbb{1}, \mathfrak{x}, \mathfrak{x}^2, \dots\}$ and $Y' := Y \setminus \{\mathbb{1}, \mathfrak{x}, \mathfrak{x}^2, \dots\}$ we see that $\#X' = \#\mathbb{R} = \#Y'$. Then there is a bijection $X' \rightarrow Y'$, extending to a bijection $X \rightarrow Y$ with

$$\mathfrak{x}^n \rightarrow n \mathfrak{x}^n \quad (n \in \mathbb{N}), \quad (6.70)$$

and to the desired linear bijection $C[0, 1] \rightarrow M$ (see (b)).

6.C.4. (ii) of Theorem 6.3.9 can not be extended to

$$\|x_n\| \rightarrow 0 \implies x_n \xrightarrow{u} 0 \quad (6.71)$$

or even

$$\|x_n\| \rightarrow 0 \implies \{x_n : n \in \mathbb{N}\} \text{ is order bounded.} \quad (6.72)$$

See $D = L^1(\mathbb{R})$.

6.C.5 Lemma. *If $c \in E$ and if D is a linear subspace of E equipped with a norm that turns it into an ordered Banach space with closed generating cone, then $\{x \in D : x \leq c\}$ is closed in D .*

Proof. Let $(x_n)_{n \in \mathbb{N}}$ and x in D be such that $x_n \rightarrow x$ in D and $x_n \leq c$ for all n . Let $a \in E^+$ be such that $-a \leq c \leq a$ and let $D_* = \mathbb{R}a + \mathbb{R}c$ be equipped with the norm as in 6.6.4. Then $x_n \rightarrow x$ in $D + D_*$ and so $c - x_n \rightarrow c - x$ in $D + D_*$. Whence $c - x \in (D + D_*)^+$. $\square$

Bibliography

- [1] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a Large-Deviations Principle to the Wasserstein Gradient Flow: A New Micro-Macro Passage. *Communications in Mathematical Physics*, 307(3):791–815, 2011.
- [2] C. D. Aliprantis and R. Tourky. *Cones and Duality*, volume 84 of *Graduate Studies in Mathematics*. American Mathematical Society, 2007.
- [3] T. Andô. On fundamental properties of a Banach space with a cone. *Pacific Journal of Mathematics*, 12:1163–1169, 1962.
- [4] C. J. K. Batty and D. W. Robinson. Positive One-Parameter Semigroups on Ordered Banach Spaces. *Acta Applicandae Mathematicae. An International Journal on Applying Mathematics and Mathematical Applications*, 2(3-4):221–296, 1984.
- [5] J. D. Biggins. Large deviations for mixtures. *Electronic Communications in Probability*, 9:60–71 (electronic), 2004.
- [6] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, second edition, 1986.
- [7] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [8] G. Birkhoff. *Lattice Theory*. Third edition. American Mathematical Society Colloquium Publications, Vol. XXV. American Mathematical Society, Providence, R.I., 1967.
- [9] S. Bochner. Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind. *Fundamenta Mathematicae*, 20:262–276, 1933.
- [10] V. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.
- [11] F. Bonsall. Sublinear functionals and ideals in partially ordered vector spaces. *Proceedings of the London Mathematical Society. Third Series*, 4(3):402–418, 1954.
- [12] N. Bourbaki. *Livre VI Intégration*. Hermann, 1952.
- [13] A. Bovier. *Statistical Mechanics of Disordered Systems*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2006. A mathematical perspective.

- [14] G. L. O'Brien. Sequences of capacities, with connections to large-deviation theory. *Journal of Theoretical Probability*, 9(1):19–35, 1996.
- [15] т. Вулих. Введение в теорию конусов в нормированных пространствах. *Изд-во Калинин. ун-та*, pages 1–84, 1977. B.Z. Wulich. Einführung in die Theorie der Kegel in normierten Räumen. (Russisch). Izd. Staatl. Universität Kalinin. 1977 (handschriftl. deutsche Übersetzung: M.R.Weber).
- [16] т. Вулих. Специальные вопросы геометрии конусов в нормированных пространствах. *Изд-во Калинин. ун-та*, pages 1–84, 1978. B.Z. Wulich. Spezielle Probleme der Geometrie von Kegeln in normierten Räumen. (Russisch). Izd. Staatl. Universität Kalinin. 1978 (handschriftl. deutsche Übersetzung: M.R.Weber).
- [17] L. A. Bunimovich, I. P. Cornfeld, R. L. Dobrushin, M. V. Jakobson, N. B. Maslova, Y. B. Pesin, Y. G. Sinai, Y. M. Sukhov, and A. M. Vershik. *Dynamical systems. II*, volume 2 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1989. Ergodic theory with applications to dynamical systems and statistical mechanics, Edited and with a preface by Sinai, Translated from the Russian.
- [18] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Inc., Boston, MA, 2004.
- [19] F. Comets. Large Deviation Estimates for a Conditional Probability Distribution. Applications to Random Interaction Gibbs Measures. *Probability Theory and Related Fields*, 80(3):407–432, 1989.
- [20] F. Comets, N. Gantert, and O. Zeitouni. Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probability Theory and Related Fields*, 118(1):65–114, 2000.
- [21] J. B. Conway. *A Course in Functional Analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.
- [22] B. R. La Cour and W. C. Schieve. A General Conditional Large Deviation Principle. *Journal of Statistical Physics*, 161(1):123–130, 2015.
- [23] P. J. Daniell. A general form of integral. *Annals of Mathematics. Second Series*, 19(4):279–294, 1918.
- [24] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [25] D. Dereudre and S. Roelly. Propagation of Gibbsianness for infinite-dimensional gradient Brownian diffusions. *Journal of Statistical Physics*, 121(3-4):511–551, 2005.

-
- [26] J.-D. Deuschel and D. W. Stroock. *Large Deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1989.
- [27] J. Dugundji. *Topology*. Allyn and Bacon, Inc., Boston, Mass., 1966.
- [28] R. S. Ellis. *Entropy, Large Deviations and Statistical Mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original.
- [29] A. C. D. van Enter. On the prevalence of non-Gibbsian states in mathematical physics. *IAMP News Bulletin*, April 2012:15–24, 2012.
- [30] A. C. D. van Enter, R. Fernández, F. den Hollander, and F. Redig. Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures. *Communications in Mathematical Physics*, 226(1):101–130, 2002.
- [31] A. C. D. van Enter, R. Fernández, F. den Hollander, and F. Redig. A large-deviation view on dynamical Gibbs-non-Gibbs transitions. *Moscow Mathematical Journal*, 10(4):687–711, 2010.
- [32] A. C. D. van Enter, R. Fernández, and A. D. Sokal. Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory. *Journal of Statistical Physics*, 72(5-6):879–1167, 1993.
- [33] A. C. D. van Enter, C. Külske, A. A. Opoku, and W. M. Ruszel. Gibbs-non-Gibbs properties for n-vector lattice and mean-field models. *Brazilian Journal of Probability and Statistics*, 24(2):226–255, 2010.
- [34] A. C. D. van Enter, F. Redig, and E. Verbitskiy. Gibbsian and non-Gibbsian states at Eurandom. *Statistica Neerlandica. Journal of the Netherlands Society for Statistics and Operations Research*, 62(3):331–344, 2008.
- [35] V. Ermolaev. *Gibbsian Properties: Examples of Failures and an Application*. PhD thesis, Rijksuniversiteit Groningen, 2011.
- [36] V. Ermolaev and C. Külske. Low-temperature dynamics of the Curie-Weiss model: Periodic orbits, multiple histories, and loss of Gibbsianness. *Journal of Statistical Physics*, 141(5):727–756, 2010.
- [37] A. Faden. The Existence of Regular Conditional Probabilities: Necessary and Sufficient Conditions. *The Annals of Probability*, 13(1):288–298, 1985.
- [38] R. Fernández. Gibbsianness and non-Gibbsianness in lattice random fields. In *Mathematical statistical physics*, pages 731–799. Elsevier B. V., Amsterdam, 2006.
- [39] R. Fernández, F. den Hollander, and J. Martínez. Variational description of Gibbs-non-Gibbs dynamical transitions for the Curie-Weiss model. *Communications in Mathematical Physics*, 319(3):703–730, 2013.

- [40] L. Fuchs. Riesz Vector Spaces and Riesz Algebras. *Séminaire Dubreil. Algèbre et théorie des nombres*, 19(23-24):1–9, 1965-1966.
- [41] O. van Gaans and A. Kalauch. Bands in Pervasive Pre-Riesz Spaces. *Operators and Matrices*, 2(2):177–191, 2008.
- [42] O. van Gaans and A. Kalauch. Ideals and bands in pre-Riesz spaces. *Positivity. An International Mathematics Journal Devoted to Theory and Applications of Positivity*, 12(4):591–611, 2008.
- [43] G. Gallavotti. *Statistical Mechanics*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1999. A short treatise.
- [44] H.-O. Georgii. *Gibbs Measures and Phase Transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1988.
- [45] A. Greven and F. den Hollander. Large deviations for a random walk in random environment. *The Annals of Probability*, 22(3):1381–1428, 1994.
- [46] P. R. Halmos. *Measure Theory*. Springer, 1974.
- [47] E. Hewitt and K. A. Ross. *Abstract harmonic analysis. Vol. I*, volume 115 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, second edition, 1979. Structure of topological groups, integration theory, group representations.
- [48] E. Hille and R. S. Phillips. *Functional Analysis and Semi-Groups*, volume 31 of *Colloquium publications*. American Mathematical Society, Providence, R. I., 1957.
- [49] H. Hogbe-Nlend. *Bornologies and Functional Analysis*, volume 62 of *Notas de matemática*. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. Introductory course on the theory of duality topology-bornology and its use in functional analysis.
- [50] F. den Hollander. *Large Deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.
- [51] F. den Hollander. Gibbs under stochastic dynamics? *Markov Processes and Related Fields*, 10(3):507–516, 2004.
- [52] F. den Hollander, R. Redig, and W. van Zuijlen. Gibbs-non-Gibbs dynamical transitions for mean-field interacting Brownian motions. *Stochastic Processes and their Applications*, 125(1):371–400, 2015.
- [53] R. Israel. Banach algebras and Kadanoff transformations. In J. Fritz, J. Lebowitz, and D. Száz, editors, *Random Fields*, volume 2, pages 593–608. Esztergom, North-Holland, Amsterdam, 1981.
- [54] G. Jeurnink. *Integration of functions with values in a Banach lattice*. PhD thesis, University of Nijmegen, 1982.

-
- [55] E. de Jonge and A. C. M. van Rooij. *Introduction to Riesz spaces*. Mathematisch Centrum, Amsterdam, 1977.
 - [56] J. L. Kelley. *General topology*. D. Van Nostrand Company, Inc., Toronto-New York-London, 1955.
 - [57] E. Kosygina, F. Rezakhanlou, and S. R. S. Varadhan. Stochastic homogenization of Hamilton-Jacobi-Bellmann equations. *Communications on Pure and Applied Mathematics*, 59(10):1489–1521, 2006.
 - [58] O. K. Kozlov. A Gibbs description of a system of random variables. *Akademiya Nauk SSSR. Institut Problem Peredachi Informatsii Akademii Nauk SSSR. Problemy Peredachi Informatsii*, 10(3):94–103, 1974. Russian.
 - [59] C. Külske. Analogues of non-Gibbsianness in joint measures of disordered mean field models. *Journal of Statistical Physics*, 112(5-6):1079–1108, 2003.
 - [60] C. Külske and A. Le Ny. Spin-flip dynamics of the Curie-Weiss model: loss of Gibbsianness with possibly broken symmetry. *Communications in Mathematical Physics*, 271(2):431–454, 2007.
 - [61] C. Külske and A. A. Opoku. Continuous spin mean-field models: Limiting kernels and Gibbs properties of local transforms. *Journal of Mathematical Physics*, 49(12):125215, 31, 2008.
 - [62] C. Külske and F. Redig. Loss without recovery of Gibbsianness during diffusion of continuous spins. *Probability Theory and Related Fields*, 135(3):428–456, 2006.
 - [63] D. Leao, Jr., M. Fragoso, and P. Ruffino. Regular Conditional Probability, Disintegration of Probability and Radon Spaces. *Proyecciones. Journal of Mathematics*, 23(1):15–29, 2004.
 - [64] C. Léonard. A large deviation approach to optimal transport. [arXiv.org/abs/0710.1461v1](https://arxiv.org/abs/0710.1461v1), 2007.
 - [65] T. M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1985.
 - [66] W. A. J. Luxemburg and A. C. Zaanen. *Riesz spaces. Vol. I*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., New York, 1971. North-Holland Mathematical Library.
 - [67] J. Martínez. *Dynamical Gibbs-non-Gibbs transitions and Brownian percolation*. PhD thesis, Leiden University, 2014.
 - [68] M. Messerschmidt. Normality of spaces of operators and quasi-lattices. *Positivity. An International Mathematics Journal Devoted to Theory and Applications of Positivity*, 19(4):695–724, 2015.

- [69] A. Le Ny. Introduction to (generalized) Gibbs measures. *Ensaio Matemáticos [Mathematical Surveys]*, 15:1–126, 2008.
- [70] A. Le Ny and F. Redig. Short time conservation of Gibbsianness under local stochastic evolutions. *Journal of Statistical Physics*, 109(5-6):1073–1090, 2002.
- [71] A. Opoku. *On Gibbs Properties of Transforms of Lattice and Mean-Field Systems*. PhD thesis, Rijksuniversiteit Groningen, 2009.
- [72] B. de Pagter. On z -ideals and d -ideals in Riesz spaces III. *Koninklijke Nederlandse Akademie van Wetenschappen. Indagationes Mathematicae*, 43(4):409–422, 1981.
- [73] A. L. Peressini. *Ordered Topological Vector Spaces*. Harper & Row, Publishers, New York-London, 1967.
- [74] B. J. Pettis. On integration in vector spaces. *Transactions of the American Mathematical Society*, 44(2):277–304, 1938.
- [75] F. Rassoul-Agha and T. Seppäläinen. *A Course on Large Deviations with an Introduction to Gibbs Measures*, volume 162 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.
- [76] F. Rassoul-Agha, T. Seppäläinen, and A. Yilmaz. Quenched free energy and large deviations for random walk in random potential. *Communications on Pure and Applied Mathematics*, 66(2):202–244, 2013.
- [77] F. Redig and F. Wang. Gibbs-non-Gibbs transitions via large deviations: computable examples. *Journal of Statistical Physics*, 147(6):1094–1112, 2012.
- [78] A. C. M. van Rooij and W. H. Schikhof. *A Second Course on Real Functions*. Cambridge University Press, Cambridge-New York, 1982.
- [79] A. van Rooij and W. van Zuijlen. Bochner integrals in ordered vector spaces. Preprint available at <http://arxiv.org/abs/1604.06341>.
- [80] A. van Rooij and W. van Zuijlen. Integrals for functions with values in partially ordered vector spaces. To appear in *Positivity*, preprint available at <http://arxiv.org/abs/1505.05997>.
- [81] D. Ruelle. *Statistical Mechanics: Rigorous Results*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.
- [82] R. A. Ryan. *Introduction to tensor products of Banach spaces*. Springer Monographs in Mathematics. Springer-Verlag London, Ltd., London, 2002.
- [83] H. H. Schaefer. *Topological Vector Spaces*. Springer-Verlag, New York-Berlin, 1971. Third printing corrected, Graduate Texts in Mathematics, Vol. 3.
- [84] W. H. Schikhof. *Ultrametric Calculus*, volume 4 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1984. An introduction to p -adic analysis.

-
- [85] Z. Semadeni. *Banach Spaces of Continuous Functions: Vol. I*. PWN—Polish Scientific Publishers, Warsaw, 1971. Monografie Matematyczne, Tom 55.
 - [86] L. A. Steen and J. A. Seebach, Jr. *Counterexamples in Topology*. Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1970.
 - [87] W. G. Sullivan. Potentials for almost Markovian random fields. *Communications in Mathematical Physics*, 33:61–74, 1973.
 - [88] G. E. F. Thomas. Integration of Functions With Values in Locally Convex Suslin Spaces. *Transactions of the American Mathematical Society*, 212:61–81, 1975.
 - [89] F. Wang. *Dynamical Gibbs-non-Gibbs transitions: a study via coupling and large deviations*. PhD thesis, Leiden University, 2012.
 - [90] A. C. Zaanen. *Integration*. North-Holland Publishing Co., Amsterdam; Interscience Publishers John Wiley & Sons, Inc., New York, 1967. Completely revised edition of An introduction to the theory of integration.
 - [91] A. C. Zaanen. *Riesz Spaces II*, volume 30 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1983.
 - [92] W. van Zuijlen. Large deviations of continuous regular conditional probabilities. Preprint available at <http://arxiv.org/abs/1605.05192>.

Index

- C -weak convergence, 71
- C -weak topology, 71
- $\mathbb{B}$ -integrable, 209
- ω -partition for, 151
- $\star$ -differentiable, 80
- $\star$ -dominated bound, 80
- φ_L -integral, 152
- f -dependent, 76

- absolutely dominating, 201
- absolutely summable interaction, 7
- approximately C -absolutely conormal, 201
- approximately C -absolutely dominating, 201
- Archimedean, 148, 200

- bad magnetisation, 30
- Banach cover, 208
- Bochner integrability, 16
- Bochner integrable, 204
- Bochner integral, 204

- C -absolutely dominating, 201
- compact sublevel sets, 114
- continuously $\star$ -differentiable, 80
- countable suprema are preserved, 165

- dividing, 189

- empirical distribution, 70
- empirical distribution function, 70
- empirical magnetisation, 29
- exchangeable, 70

- filling, 191
- function norm, 209

- Gibbs distribution, 8
- Gibbs measure, 8

- good limiting empirical distribution, 75
- good magnetisation, 30

- Hamiltonian, 8, 78

- improper Riemann integral, 16
- improperly in, 195
- integral, 204
- interaction, 7

- Köthe space, 209
- kernel, 108

- large deviation lower bound, 103
- large deviation principle, 104
- large deviation upper bound, 103
- lateral extension, 18, 153
- laterally ω -integrable, 151
- laterally φ -integrable, 153
- laterally extendable, 151
- Lebesgue integral, 16
- limit point of sequence of sets, 73
- local function, 10
- lower semicontinuous, 114

- magnetisation, 76
- magnetisation dependent, 76
- mean-field transformation, 91
- mediated, 146

- normalising constant, 30, 78
- normed ordered vector space, 201

- order bounded, 167
- order convergent, 149
- order dense, 164
- ordered Banach cover, 208
- ordered Banach space, 201
- originate from the potential, 78

-
- partially in, 151
 - partition, 151
 - partition for a function, 151
 - perfectly normal, 71
 - pervasive, 168
 - Pettis integral, 17
 - phase transition, 9
 - possesses a pervasive subset, 168
 - potential, 30, 78
 - probability kernel, 108
 - product regular conditional kernel, 108
 - product regular conditional probability, 108
 - proper, 58
 - quasilocal function, 10
 - quasilocal specification, 10
 - reference measure, 78
 - reference measures, 30
 - refinement of a partition, 151
 - regular conditional kernel, 108
 - regular conditional probability, 57, 108
 - Riemann integral, 16
 - second difference quotient, 36
 - sequence of finite-volume mean-field Gibbs measures, 30, 78
 - sequence of finite-volume mean-field Gibbs measures with potential, 78
 - sequentially Gibbs, 30, 75
 - simple function, 158, 204
 - specification, 9
 - specification kernel, 31, 75
 - splitting, 155
 - stable, 151
 - supporting point, 48
 - two-layer model, 33
 - uncooperative, 189
 - uniformly Cauchy sequence, 202
 - uniformly complete, 202
 - uniformly convergence, 202
 - uniformly nonnull, 10
 - vague convergence, 71
 - vague topology, 71
 - vertical extension, 17, 145
 - weak convergence, 71
 - weak topology, 71
 - weakly continuous, 58, 109

Samenvatting

In dit proefschrift komen twee onderwerpen aan bod. In Deel I beschouwen we Gibbs-niet-Gibbs overgangen, in het bijzonder voor gemiddeld-veld-interacties. In Deel II beschouwen we vectorintegratie, in het bijzonder integratie van functies met waarden in een partieel geordende vectorruimte.

Deel I: Gibbs-niet-Gibbs overgangen

Een Gibbsmaat is een wiskundig object dat gebruikt wordt om de evenwichtstoestand van een systeem met een groot aantal interagerende componenten te beschrijven. De componenten kunnen uit magnetische spins bestaan, waaraan men een waarde toekent, bijvoorbeeld $+1$ of -1 . Over het algemeen liggen de componenten op een rooster, bijvoorbeeld op het kwadratisch rooster $\mathbb{Z}^2$. De Gibbsmaat beschrijft de kans van een configuratie (bijvoorbeeld een element van $\{-1, +1\}^{\mathbb{Z}^2}$) aan de hand van een energiefunctie, die aan iedere configuratie een energie toekent. Deze beschrijving is gebaseerd op de zogenoemde Maxwell-Boltzmann-Gibbs formule. Er is nog een andere manier om Gibbsmaten te karakteriseren; het zijn namelijk de maten die aan een zekere “regulariteitsconditie” voor conditionele kansen voldoen.

Indien men het systeem dat wordt beschreven door een Gibbsmaat verandert, wat neerkomt op het feit dat het nieuwe systeem beschreven wordt door een transformatie van de Gibbsmaat, dan kan het zo zijn dat de getransformeerde maat zelf geen Gibbsmaat meer is, ofwel, dat deze maat niet aan de “regulariteitsconditie” voldoet. Een van de belangrijke vragen is derhalve wanneer de getransformeerde maat een Gibbsmaat is, en dus wel aan de “regulariteitsconditie” voldoet, en wanneer niet.

Een centraal voorbeeld van zo’n transformatie is de evolutie van het systeem volgens een stochastische dynamica over een zeker tijdsinterval. In dit geval vragen we ons dus af op welke tijdstippen de tijdsgeëvolueerde maat Gibbs is en op welke tijdstippen niet. Deze en gerelateerde vragen worden behandeld in Deel I van het proefschrift. We beschouwen in dit proefschrift alleen de gemiddeld-veld-modellen. In deze modellen wordt de ruimtelijke structuur achterwege gelaten: iedere component is in gelijke mate afhankelijk van iedere andere component.

In Hoofdstuk 2 beschouwen we gemiddeld-veld-modellen van componenten met waarden in de reële rechte $\mathbb{R}$. We geven een conditie op de potentiaal (continue differentieerbaarheid) waaronder het initiële systeem Gibbs is. De evolutie die we beschouwen is die van onafhankelijke Brownse bewegingen. We laten zien dat het tijdsgeëvolueerde systeem Gibbs is dan en slechts dan als de entropiefunctie van de onderliggende conditionele kans van de initiële configuratie gegeven de geëvolueerde configuratie een unieke minimiser heeft. Dit bewijst het algemene vermoeden over dit verband voor de gemiddeld-veld-modellen. Deze entropiefunctie valt dusdanig

te ontleiden dat men een noodzakelijke en voldoende voorwaarde op de potentiaal kan geven waaronder het tijdsgeëvolueerde systeem Gibbs is. Dit legt op zijn beurt expliciet vast welke mogelijke scenario's er zijn voor de tijdslijn van Gibbs of niet-Gibbs. Zo is het voor dit model niet mogelijk om weer Gibbs te worden nadat het niet-Gibbs geworden is. Bij ieder mogelijk scenario geven we een voorbeeld.

In Hoofdstuk 3 beschouwen we een algemener niveau voor gemiddeld-veld-modellen, ook wel het empirische niveau of niveau 2 genoemd. Om algemenere stochastische dynamica's en algemenere potentialen te gebruiken is het noodzakelijk om dit niveau te beschouwen. Dit brengt wel de noodzaak met zich mee om een andere taal en andere wiskundige objecten te introduceren, vooral ook om die passend te maken bij de reeds bekende artikelen over het empirische niveau die compacte toestandruimten beschouwen. Analooq aan Hoofdstuk 1 geven we een vergelijkbare conditie op de potentiaal waaronder het initiële systeem Gibbs is. We leiden formules af voor de conditionele kansen onder transformaties, die op hun beurt gebruikt worden om een algemenere implicatie van uniciteit van de globale minimiser van de entropiefunctie van deze conditionele kansen tot het Gibbs zijn van het getransformeerde systeem te bewijzen. Zo'n entropiefunctie hoort bij het principe van grote afwijkingen. Het is niet vanzelfsprekend dat de conditionele kansen aan zo'n principe voldoen.

In Hoofdstuk 4 gaan we in op de vraag of conditionele kansen aan het principe van grote afwijkingen voldoen. We leiden een noodzakelijke en voldoende voorwaarde daarvoor af. Met behulp van een voorbeeld illustreren we de noodzaak van deze voorwaarden. Daarnaast bewijzen we dat conditionele kansen van zogenaamd "Sanov" type aan het principe van grote afwijkingen voldoen indien de toestandruimte eindig is.

Deel II: Vectorintegratie

Het begrip integraal is welbekend onder zowel wiskundigen als niet-wiskundigen. Ook al wordt het door de niet-wiskundigen vaak gezien als een manier om de oppervlakte onder een grafiek uit te rekenen, heeft het voor de wiskundigen ook andere toepassingen. Indien men functies met waarden in een vectorruimte beschouwt, dan spreekt men wel over vectorintegratie. Men beschouwt vectorruimten met extra structuur en gebruikt deze structuur om tot een degelijke integraal te komen. Zo maken de Bochner- en de Pettis-integraal gebruik van eigenschappen van de Banachruimte, in het bijzonder van de norm, om tot een integraal te komen voor functies met waarden in een Banachruimte. In Deel II beschouwen we geordende vectorruimten (niet noodzakelijk met een norm) en gebruiken de ordening om tot een integraal te komen voor functies met waarden in zo'n vectorruimte.

In Hoofdstuk 5 introduceren we twee uitbreidingen van integralen. We noemen deze de verticale en de horizontale uitbreiding. De manier waarop de verticale uitbreiding is opgebouwd is vergelijkbaar met de opzet van de Riemann-integraal, die van de laterale uitbreiding is vergelijkbaar met de opzet van de oneigenlijke Riemann-integraal. We laten zien onder welke omstandigheden deze verschillende uitbreidingen gecombineerd kunnen worden en hoe deze verschillende combinaties in relatie met elkaar staan. We besteden extra aandacht aan uitbreidingen van integralen van

simpele functies op een maatruimte. In het geval dat de ruimte niet alleen geordend is maar ook een Banachruimte is, vergelijken we de uitbreidingen met de Bochner- en de Pettis-integraal.

In Hoofdstuk 6 beschouwen we Bochner-integralen voor functies met waarden in een geordende vectorruimte. Indien een vectorruimte op een mooie manier overdekt kan worden door Banachruimten, kunnen we een integraal definiëren voor functies met waarden in deze vectorruimte. We laten zien dat Archimedische geordende vectorruimten zulke overdekkingen hebben bestaande uit Banachruimten met gesloten en genererende kegels. We bestuderen eerst Bochner-integreerbare functies met waarden in een Banachruimte die een gesloten en genererende kegel heeft. Daarna bestuderen we de integraal voor functies met waarden in een Archimedische geordende vectorruimte. Ten slotte vergelijken we deze integraal met de integralen uit Hoofdstuk 5.

Acknowledgements

This thesis is the product of four years of PhD studies. It portrays the research on which I worked during these years, and represents only a small fraction of the process that took me from a Master of Science to a Doctor of Philosophy, a PhD. Many people contributed to this thesis, and this whole process.

I am grateful to the members of the reading committee for their time spent in studying the manuscript in order to come to their judgement.

I thank my supervisor Frank den Hollander, who met with me on a regular basis to discuss the progress of my research. He guided me by asking the right questions and gave shape to the different projects I worked on. It is always a pleasure to discuss work with Frank, as he gives stimulating feedback and shows appreciation of my own input and ideas. I admire the way he organises his research and all the activities he is involved in.

I also thank my supervisor Frank Redig. With him I had less frequent but longer discussions, during which we could explore the details and the difficulties of proofs, where his enthusiasm and bursts of ideas were very helpful and encouraging. Frank always had alternative approaches and attempts at hand, also when I had no clear idea of how to make progress on a problem. His easy and open way of supervision, accompanied by the occasional Belgian humour, is very pleasant.

The different approaches and guidances of the two Franks formed a good combination of styles in supervision for the project that has led to Part I.

The supervisor of my Master's thesis, Arnoud van Rooij, became my collaborator on the project that has led to Part II. By means of telephone conversations and visits to Nijmegen and Leiden, we extended the ideas from my Master's thesis to a general approach. Besides this project, we share many interests in mathematics. Arnoud was always open to talk about other matters and give feedback, which I enjoyed a lot. His way of approaching mathematics -with much patience- is something I consider exemplary.

Special thanks to the secretaries of Nijmegen for facilitating the mathematical communication between Arnoud and me.

Aside from my direct supervisors and collaborator, I also benefited greatly from all the mathematicians I met and talked with during my PhD.

I am grateful for having been able to travel to many places, to attend conferences and summer schools, and to visit people for a week.

I especially enjoyed the environment in Leiden I entered four years ago, with people of many distinct nationalities and different cultures. I benefited from the open atmosphere, where I could walk into offices unannounced to find people who were willing to listen to me and to help me find answers to questions. This includes the offices of the support staff as well.

Besides doing math, I needed to have some distraction. I thank all those who contributed to that: those with whom I did sports, danced, had dinner, had some drinks, played games, laughed; simply those with whom I spent time in any nice way. In particular, I thank my close friends, whose doors are always open and who are always supportive.

Last but not least, I thank my family, who always provide a warm home for me, all the support and love one could wish for.

Curriculum Vitae

Willem Bernard van Zuijlen was born in Zeeland in 1989. After graduating in 2007 from Udens College in Uden, he studied mathematics at the Radboud University in Nijmegen. There he obtained a Bachelor of Science in Mathematics in 2010 (*cum laude*) and a Master of Science in Mathematics in 2012 (*summa cum laude*). He wrote his Master's thesis, entitled *Integration of functions with values in a Riesz space*, under the supervision of Prof. dr. A.C.M. van Rooij. In 2012 he moved to the University of Leiden and started his PhD research under the supervision of Prof. dr. Frank den Hollander and Prof. dr. Frank Redig, funded by the ERC Advanced Grant of Prof. dr. Frank den Hollander. He accepts a postdoc position at the Humboldt University of Berlin, starting from October 2016.