Chapter 3

Measure on abstract space

First of all let us quickly recall the key steps we made learning the Lebesgue measure on $(0, 1)$ and $\mathbb{R}^1$.

1. We assigned the measure on the system of all intervals $\mathcal{I}$, setting $\mu(I) := \ell(I)$ for all $I \in \mathcal{I}$. Recall that $\mathcal{I}$ was a semiring.

2. We naturally (and easily) extended the Lebesgue measure to the system of the elementary sets $\mathbf{E}$.
   $\mathbf{E}$ turned out to be a ring. (In case of $(0, 1)$ $\mathbf{E}$ is even algebra; in case of $\mathbb{R}^1$ it is not, since $\mathbb{R}^1$ cannot be represented as a finite union of the bounded intervals).

3. In order to extend the measure to even larger class $\mathcal{L}$, we introduced the outer measure. Outer measure was defined for all(!) subsets of the underlying space $(0, 1)$.
   Either we introduced the inner measure, but its definition was reduced to the outer measure (of the complement).
   Outer measure, in turn, was defined through the measures of the elements from $\mathcal{I}$.

4. We proved that the system $\mathcal{L}$ of the Lebesgue measurable set was a σ-algebra and that the Lebesgue measure was σ-additive.
   As an intermediate step we proved a lot of lemmas and theorems and in partricular the Vall*é*e-Poussin's measurability criterion.
   Also we showed that the outer measure was subadditive and monotonic.

Starting from these observations, one develops the measure theory on an abstract space $\mathbb{X}$.
In principle, two approaches are possible:

1. (axiomatic approach): It is assumed that we already have a σ-algebra $\mathcal{A}$, $\mathcal{A} \subset \mathcal{P}(\mathbb{X})$.
   A pair $(\mathbb{X}, \mathcal{A})$ is called the *measurable space.*
   Further it is assumed that we have a measure[1] μ, defined on $\mathcal{A}$
   A triple $(\mathbb{X}, \mathcal{A}, \mu)$ is called the *measure space.*
   Such approach is common for financial modeling; the measure space[2] is called *probability space* or (sometimes) *the Kolmogorov triple* and usually denoted as $(\Omega, \mathcal{F}, \mathbb{P})$.
   The essence of the axiomatic approach is that the σ-algebra and the measure are assumed to be already given.

2. (constructive approach): We are given an [abstract] underlying space $\mathbb{X}$, a system $\mathcal{S}$ of subsets of $\mathbb{X}$ and either an additive function υ on $\mathcal{S}$ or an outer measure λ^* on $\mathcal{P}(\mathbb{X})$.
   $\mathcal{S}$ is usually a semiring but also can be a ring or an algebra. The objective is to extend the function υ to a measure μ on the σ-algebra $\mathcal{A}$ s.t. $\mathcal{S} \subset \mathcal{A} \subset \mathcal{P}(\mathbb{X})$ and $\mu|_{\mathcal{S}} = \upsilon$

Though in financial modeling we will mostly use axiomatic approach, the constructive one is worth learning too, since it helps to disclose the inherent properties of a measure. I believe that pedagogically it is the best way to let these two appoaches walk hand in hand. So let's rock!
From now on $\mathbb{X}$ denotes the [abstract] space in question.

Definition 3.0.1. Let $\mathcal{S}$ be a semiring, a ring or an algebra. Let $A \in \mathcal{S}$ and let $A_1, A_2, ..., A_n \in \mathcal{S}$ be an arbitrary(!) finite decomposition of A, i.e. $A = \bigsqcup_{k=1}^{n} A_k$.

A function $\upsilon : \mathcal{S} \to [0, \infty]$ is called a *volume* if it holds that
(v1) $\upsilon(\varnothing) = 0$
(v2)

$$\upsilon(A) = \sum_{k=1}^{n} \upsilon(A_k)$$

for all $A \in \mathcal{S}$ and all their finite decompositions.

Remark 3.0.2. If $\upsilon(\cdot)$ is bounded, than (v1) is redundant since from the decomposion $\varnothing = \varnothing \cup \varnothing$ it follows that $\upsilon(\varnothing) = 2\upsilon(\varnothing)$, i.e. $\upsilon(\varnothing) = 0$. Otherwise it may happen that $\upsilon(\cdot) \equiv +\infty$ (though such case is not of practical interest).

---

[1] Soon I give a formal definition of an [abstract] measure.

[2] Additionally, it is endowed with a *filtration* $(F)_t$, i.e. an increasing sequence of σ-algebras. We postpone discussion of filtrations to the chapter on stochastic processes.

Exercise 3.0.3. In words the we mean in the Definition 3.0.1 that A (an arbitrary element of $\mathcal{S}$) can be represented as a finite(!), disjoint(!) union of some other elements of $\mathcal{S}$. There can be many such decompositions but we require the volume of A to be equal for all of them.
1. Compare it to the Lemma 2.2.3.2
2. Prove that there exists *at least one* such decomposition. (Hint: $A = A \cup \varnothing$ will also do).

Definition 3.0.4. Let $\mathcal{A}$ be a σ-algebra on $\mathbb{X}$. A function $\mu : \mathcal{A} \to [0, \infty]$ is called a measure if
(μ1) $\mu(\varnothing) = 0$
(μ2) If $A_1, A_2, ... \in \mathcal{A}$ are pairwise disjoint then

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$$

Remark 3.0.5. The elements of an [arbitrary] σ-algebra are sometimes called *measurable* sets; in the sence that we can measure them if assign a measure according to the Definition 3.0.4

Example 3.0.6. (Some measure spaces)

1. $((0,1), \mathcal{L}, \mu)$, $\mathcal{L}$ is the σ-algebra of Lebesgue-measurable set, μ is the Lebesgue measure. This example is well-known to us from Chapter 2.

2. $((0,1), \mathcal{B}, \mu|_{\mathcal{B}})$, $\mathcal{B}$ is the Borel σ-algebra (which, as so far was stated without proof, is smaller than $\mathcal{L}$ and $\mu|_{\mathcal{B}}$ is the projection of the Lebesgue measure on $\mathcal{B}$. Note that this measure space is incomplete, i.e. if $A \in \mathcal{B}$, $\mu|_{\mathcal{B}}(A) = 0$ and $B \subset A$, it does *not* imply that $B \in \mathcal{B}$.

3. $(\mathbb{X}, \mathcal{P}(\mathbb{X}), \delta_{x_0})$ where $\mathbb{X}$ is an abitrary space, $\mathcal{P}(\mathbb{X})$ is the system of all subsets of $\mathbb{X}$, $x_0 \in \mathbb{X}$ is a point from $\mathbb{X}$ and the *Dirac measure*

$$\delta_{x_0}(A) := \begin{cases} 1, x_0 \in A \\ 0, else \end{cases}$$

   The proof that $(\mathbb{X}, \mathcal{P}(\mathbb{X}), \delta_{x_0})$ is indeed a measure space is left as an *easy* exercise. Can we take an arbitrary σ-algebra on $\mathbb{X}$ instead of $\mathcal{P}(\mathbb{X})$?

4. $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \#)$, where the *counting* measure

$$\#(A) := \{0, 1, 2, ..., +\infty\}$$

   is equal to the number of elements in A.

5. $(\mathbb{X}, \mathcal{L}, \mu)$, $\mathbb{X}$ uncountable, $\mathcal{A} := \{A \subset \mathbb{X} \quad | \quad A \text{ or } A^c \text{ countable } \}$ and

$$\mu(A) := \begin{cases} 0, & A\, is\, countable \\ 1, & else \end{cases}$$

Will the quantity also

$$\widehat{\mu}(A) := \begin{cases} 0, & A\, is\, countable \\ +\infty, & else \end{cases}$$

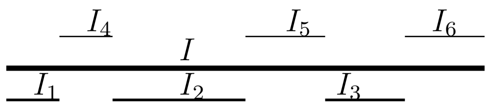
be a measure on $\mathcal{A}$?

Our nearest objective to consider the following problem: If we have a volume $\upsilon(\cdot) : \mathcal{S} \to [0, \infty]$, where $\mathcal{S}$ is a semiring, how can we extend $\upsilon(\cdot)$ to the ring $\mathcal{R}(\mathcal{S})$, i.e. the ring generated by $\mathcal{S}$? This is nothing else but an abstract analog of the case, where we extended μ (equal to ℓ) from $\mathcal{I}$ to $\mathbf{E}$. Further we will discuss, to which requirements should υ satisfy in order to be extended to some σ-algebra $\mathcal{A}$. Note that $\mathcal{A}$ is generally *not* a sigma-algebra, generated by $\mathcal{S}$. Again an analogy with Chapter 2: the Borel σ-algebra was generated by the semiring $\mathcal{I}$ but the Lebesgue σ-algebra of the measurable sets turned out to be larger.

Lemma 3.0.7. *Let $\mathcal{S}$ be a semiring, $A, A_1, A_2, \dots, A_n \in \mathcal{S}$, $A_1, A_2, \dots, A_n \subset A$ and $A_i \cap A_j = \varnothing$ for $i \neq j$. Then there exist $A_{n+1}, A_{n+2}, \dots, A_s$ $(s \geq n)$ such that*

$$A = \bigsqcup_{k=1}^{s} A_k$$

To see the main idea (by the example of our favourite semiring $\mathcal{I}$) look at the picture:



There are $I, I_1, I_2, I_3 \in \mathcal{I}$, $I_1, I_2, I_3 \subset I$ and I_1, I_2, I_3 are pairwise disjoint. The message is that we can find the pairwise disjoint I_4, I_5, I_6 so that $I = \bigsqcup_{k=1}^{6} I_k$.

Proof (by induction). For $n = 1$ the statement follows from (smR3) (Definition 2.2.2.6). Assume that it holds for $n = m$ and consider the disjoint sets $A_1, \cdots A_m, A_{m+1} \in \mathcal{S}$. By assumption it follows

$$A = A_1 \sqcup A_2 \sqcup \ldots \sqcup A_m \sqcup B_1 \sqcup B_2 \sqcup \ldots \sqcup B_p,$$

where all sets $B_q (q = 1, 2, \ldots, p)$ belong to $\mathcal{S}$. Let

$$B_{q1} := A_{m+1} \cap B_q \quad \overset{(smR2)}{\Rightarrow} \quad B_{q1} \in \mathcal{S}$$

(Note: $q1$ is the double index and not q_1). Of course $B_q \supset B_{q1}$ and thus for every $B_q (q = 1, 2, \ldots, p)$ by (smR3) there exists a disjoint decomposition

$$B_q = B_{q1} \sqcup B_{q2} \sqcup \ldots \sqcup B_{qr_q}$$

where r_q depends on q and where $B_{q1}, B_{q2}, \ldots B_{qr_q} \in \mathcal{S}$.
So we have

$$B_q = (A_{m+1} \cap B_q) \underbrace{\sqcup}_{(\dagger)} B_{q2} \sqcup \ldots \sqcup B_{qr_q} = A_{m+1} \underbrace{\sqcup}_{(\star)} \left(\bigsqcup_{j=2}^{r_q} B_{qj} \right)$$

Note that in $(\star)$ we are really allowed to write $\sqcup$ (and not just $\cup$) because if a point $x \in (A_{m+1} \cap B_q)$ then it is obviously in A_{m+1}. On the other hand x can*not* be in $B_{q2} \sqcup \ldots \sqcup B_{qr_q}$ since it would violate a disjoint condition in $(\dagger)$

Hence

$$A = A_1 \sqcup \ldots \sqcup A_m \sqcup A_{m+1} \sqcup \left(\bigsqcup_{q=1}^{p} \left(\bigsqcup_{j=2}^{r_q} B_{qj} \right) \right)$$

So we proved the Lemma for $n = m + 1$ and hence for all n

■

Theorem 3.0.8. *If $\mathcal{S}$ is a semiring then $\mathcal{R}(\mathcal{S})$ (the ring generated by $\mathcal{S}$) coincides with the system of sets $\mathfrak{U}$, whose elements A can be represented as a finite disjoint union*

$$A = \bigsqcup_{k=1}^{n} A_k \quad A_k \in \mathcal{S}$$

Compare: every element from $\mathbf{E}$ can be represented as a finite disjoint union of intervals from $\mathcal{I}$.

Proof. First we show that $\mathfrak{U}$ is a ring. Indeed, if $A, B \in \mathfrak{U}$ then there exist decompositions

$$A = \bigsqcup_{i=1}^{n} A_i, \; B = \bigsqcup_{j=1}^{m} B_j, \quad A_i \in \mathcal{S}, \; B_j \in \mathcal{S}$$

Since $\mathcal{S}$ is a semiring, it holds that

$$C_{ij} := A_i \cap B_j$$

belongs to $\mathcal{S}$. Of course $\{C_{ij}\}_{i=(1..n),j=(1..m)}$ are disjoint, $A_i \supset C_{ij}$ (i is arbitrary but fixed, j runs from 1 to m) and $B_j \supset C_{ij}$ (j is arbitrary but fixed, i runs from 1 to n) By Lemma 3.0.7 there exist decompositions

$$A_i = \bigsqcup_j C_{ij} \sqcup \bigsqcup_{k=1}^{r_i} D_{ik} \qquad B_j = \bigsqcup_i C_{ij} \sqcup \bigsqcup_{l=1}^{s_j} E_{jl}$$

where $D_{ik}, E_{jl} \in \mathcal{S}$. It follows that the following decompositions exist

$$A \cap B = \bigsqcup_{i,j} C_{ij} \qquad A \Delta B = (A \cup B) \setminus (A \cap B) = \bigsqcup_{i,k} D_{ik} \underbrace{\sqcup}_{(\star)} \bigsqcup_{j,l} E_{jl}$$

hence $(A \cap B) \in \mathfrak{U}$ and $(A \Delta B) \in \mathfrak{U}$. In $(\star)$ we write "$\sqcup$" instead of "$\cup$" because $A \cap B = \bigsqcup_{i,j} C_{ij}$ and thus the rests of A and B are disjoint.
Finally it holds

$$A \cup B = (A \Delta B) \Delta (A \cap B) \qquad A \setminus B = A \Delta (A \cap B)$$

thus $\mathfrak{U}$ is a ring.
Second, we readily see that $\mathfrak{U} = \mathcal{R}(\mathcal{S})$, i.e. $\mathfrak{U}$ is the minimal ring containing $\mathcal{S}$. Indeed, assume there exists a ring $\mathfrak{M}$ s.t. $\mathcal{S} \subset \mathfrak{M} \subset \mathfrak{U}$ and there is a set $A \in \mathfrak{U}, A \notin \mathfrak{M}$. By definition of $\mathfrak{U}$ there is a decomposition $A = \bigsqcup_{k=1}^n A_k, \quad A_k \in \mathcal{S}$. But $\mathcal{S} \subset \mathfrak{M}$ hence every $A_k \in \mathfrak{M}$, and since $\mathfrak{M}$ is a ring it follows that $A \in \mathfrak{M}$, a contradiction!

■

Theorem 3.0.9. *For every volume υ on a semiring $\mathcal{S}$ there exist a unique extention m to the ring generated by $\mathcal{S}$, i.e. $\upsilon : \mathcal{S} \to [0, +\infty]$ can be uniquely extended to $m : \mathcal{R}(\mathcal{S}) \to [0, +\infty]$.*

Proof. We have just shown that for every $A \in \mathcal{R}(\mathcal{S})$ there exists a disjoint decomposition

$$A = \bigsqcup_{k=1}^n B_k, \quad B_k \in \mathcal{S} \quad (\star)$$

Define

$$m(A) := \sum_{k=1}^n \upsilon(B_k) \quad (\star\star)$$

One readily sees that m in $(\star\star)$ does not depend on the choice of a decomposition in $(\star)$. Indeed, consider two decompositions

$$A = \bigsqcup_{k=1}^n B_k = \bigsqcup_{j=1}^r C_j, \quad B_i, C_j \in \mathcal{S}$$

Since for any i, j $(B_i \cap C_j) \in \mathcal{S}$ and υ is additive, it follows that

$$\sum_{i=1}^{n} \upsilon(B_i) = \sum_{i=1}^{n}\sum_{j=1}^{r} \upsilon(B_i \cap C_j) = \sum_{j=1}^{r} \upsilon(C_j)$$

We prove now that m is additive. Consider $A = \bigsqcup\limits_{k=1}^{n} B_k$ and $C = \bigsqcup\limits_{l=1}^{p} D_l$, $B_k, D_l \in \mathcal{S}$ and $A \cap C = \varnothing$. Then

$$A \cup C = A \sqcup C = \left(\bigsqcup_{k=1}^{n} B_k\right) \sqcup \left(\bigsqcup_{l=1}^{p} D_l\right)$$

and the additivity of m follows from the additivity of υ.

To show the unicity of m we first of all note an obvious fact that $m(B) = \upsilon(B)$ for all $B \in \mathcal{S}$. If $\bar{m}$ is another extention of υ we still have $\upsilon(B) = m(B) = \bar{m}(B)$. Consider a decompostion $A = \bigsqcup_{k=1}^{n} B_k,\ A \in \mathcal{R}(\mathcal{S})$. It holds

$$\bar{m}(A) = \sum_{k=1}^{n} \bar{m}(B_k) = \sum_{k=1}^{n} m(B_k) = m(A)$$

i.e. $\bar{m}(A) \equiv m(A)$ for all $A \in \mathcal{R}(\mathcal{S})$

■

Actually, we developed an abstract analog of the extention of ℓ from $\mathcal{I}$ to $\mathbf{E}$. Unicity of m relates to Lemma 2.2.3.2 and the additivity is a direct analog of Lemma 2.2.3.3.

There are a couple of nearly obvious but important properties, which follow from the positivity and the additivity of m.

Theorem 3.0.10. *Let m be a volume on a ring $\mathcal{R}_m$, i.e. $m : \mathcal{R}_m \to [0, \infty]$. Let $A, A_1, \dots, A_n \in \mathcal{R}_m$.*

1. *If $\bigsqcup_{k=1}^{n} A_k \subset A$ (disjoint union, i.e. $A_i \cap A_j = \varnothing$ for $i \neq j$) then*

$$\sum_{k=1}^{n} m(A_k) \le m(A)$$

2. *If $\bigcup_{k=1}^{n} A_k \supset A$ then*

$$\sum_{k=1}^{n} m(A_k) \ge m(A)$$

In particular, if $A \subset A'$ and $A, A' \in \mathcal{R}_m$ then $m(A) \le m(A')$

Proof. 1. Since $A, A_1, \dots, A_n$ are pairwise-disjoint it follows from the additivity of m that

$$m(A) = \sum_{k=1}^{n} m(A_k) + m\left(A \setminus \bigsqcup_{k=1}^{n} A_k\right)$$

In turn $m\left(\bigsqcup_{k=1}^{n} A_k\right) \geq 0$ due to the positivity of m, hence the first statement holds true.
2. Further for any $A_1, A_2 \in \mathcal{R}_m$ we have

$$m(A_1 \cup A_2) = m(A_1) + m(A_2) - m(A_1 \cap A_2) \leq m(A_1) + m(A_2)$$

By induction

$$m\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} m(A_k)$$

And again from the additivity of m we obtain

$$m(A) = m\left(\bigcup_{k=1}^{n} A_k\right) - m\left(\bigcup_{k=1}^{n} A_k \setminus A\right) \leq m\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} m(A_k)$$

■

Note that although we proved the theorem for the case of a ring it holds for the semirings too (why?)

If a volume v on a semiring $\mathcal{S}$ is just additive then, generally speaking, we cannot extend it beyond the ring $\mathcal{R}(\mathcal{S})$. But if v is also σ-additive then (as we will see soon) we can proceed to the σ-algebra of measurable sets (which is, as was already mentioned, is even larger than $\mathcal{A}(\mathcal{S})$).
From now on we will say a "*premeasure*" instead of a "sigma-additive volume".

Definition 3.0.11. Let $\mathcal{S}$ be a semiring, a ring or an algebra. Let $A \in \mathcal{S}$ and let $A_1, A_2, \dots$ be an arbitrary countable(!) decomposition of A, i.e. $A = \bigsqcup_{k=1}^{\infty} A_k$ (disjoint union, $A_i \cap A_j = \varnothing$ for $i \neq j$).
A function $m : \mathcal{S} \to [0, \infty]$ is called a *premeasure* if
(m1) $m(\varnothing) = 0$
(m2)

$$m(A) = \sum_{k=1}^{\infty} m(A_k)$$

for all $A \in \mathcal{S}$ and all their countable decompositions.

Compare to the Definition 3.0.1(v2) and you will see that we just switched from finite sums and unions to the countably infinite analogs.
Obviously every measure is a premeasure since a measure is defined on a σ-algebra (Definition 3.0.4) and every σ-algebra conforms to the properties of a semiring, a ring or an algebra. And our nearest objective is, in a certain sense, to show the converse: i.e. that every premeasure can be extended to the measure.

Remark 3.0.12. Just to avoid confusion: as we remember, a ring and an algebra are *not* closed under a countably-infinite union. Moreover, a semiring is not at all closed under the union. It means that if $A_1, A_2, \ldots$ are in $\mathcal{S}$, the set $A = \bigcup_{k=1}^{\infty} A_k$ is *not necessarily* in $\mathcal{S}$. But it *may* be! So if it is and, moreover, $\{A_k\}_{k\in\mathbb{N}}$ are pairwise-disjoint then we require (m2) to hold.
Again, a semiring $\mathcal{I}$ gives us a good example: say, $(0, \frac{1}{3}] \sqcup (\frac{1}{2}, 1) \notin \mathcal{I}$ but

$$\left(0, \frac{1}{2}\right] \bigsqcup \left(\frac{1}{2}, \frac{2}{3}\right] \bigsqcup \left(\frac{2}{3}, \frac{3}{4}\right] \bigsqcup \cdots \bigsqcup \left(\frac{n}{n+1}, \frac{n+1}{n+2}\right] \bigsqcup \cdots = (0,1) \in \mathcal{I}$$

Example 3.0.13. (σ-additive [pre]measures)

1. The Lebesgue premeasure ℓ on $\mathcal{I}$ and the Lebesgue measure μ on $\mathcal{L}$ are σ-additive.

2. Let $\mathbb{X} := \{x_1, x_2, \ldots\}$ be an arbitrary countable set. To every $x_n \in \mathbb{X}$ we assign a *probability* $p_n > 0$ such that $\sum_{n=1}^{\infty} p_n = 1$. You easily check that the measure

$$\mathbb{P}(A) := \sum_{x_n \in A} p_n \qquad A \in \mathcal{P}(\mathbb{X})$$

is a σ-additive measure and $\mathbb{P}(\mathbb{X}) = 1$. This example is important for discrete financial mathematics.

Example 3.0.14. (Longtime promised additive but not σ-additive case)
Let $\mathbb{X} := (0,1) \cap \mathbb{Q}$, i.e. $\mathbb{X}$ is a set of all rational points on $(0,1)$. Define $\mathcal{S} := \mathbb{X} \cap \mathcal{I}$, i.e. $\mathcal{S}$ consists of the intersections of $\mathbb{X}$ with arbitrary intervals (a,b), $(a,b]$, $[a,b)$ and $[a,b]$, where $0 < a \leq b < 1$. One easily sees that $\mathcal{S}$ is a semiring. Indeed, for $A := \mathbb{X} \cap I_1$, $B := \mathbb{X} \cap I_2$ from $\mathcal{S}$ we have

$$A \cap B = [\mathbb{X} \cap I_1] \cap [\mathbb{X} \cap I_2] = \mathbb{X} \cap \underbrace{(I_1 \cap I_2)}_{\in \mathcal{I}}$$

For $A \subset B$ it holds

$$B \setminus A = [\mathbb{X} \cap I_1] \setminus [\mathbb{X} \cap I_2] = \mathbb{X} \cap [I_2 \setminus I_2]$$

But $I_2 \setminus I_1$ can be represented as a disjoint union of intervals from $\mathcal{I}$

For every set $A_I := \mathbb{X} \cap I$ in $\mathcal{S}$ we assign a volume

$$\upsilon(A_I) = \ell(I)$$

υ is additive but not σ-additive since $\upsilon(\mathbb{X}) = 1$ but at the same time $\mathbb{X}$ is a countable(!) union of rational points, each having zero volume.

Theorem 3.0.15. *A premeasure m on a semiring $\mathfrak{S}$ can be uniquely extended to the ring $\mathcal{R}(\mathfrak{S})$*

Proof. Let

$$A \in \mathcal{R}(\mathfrak{S}), \qquad B_n \in \mathcal{R}(\mathfrak{S}), \quad n = 1, 2, \dots$$

and

$$A = \bigsqcup_{n=1}^{\infty} B_n \qquad B_s \cap B_r = \varnothing \text{ for } r \neq s$$

Then there exist (Theorem 3.0.8) the sets A_j and B_{ni} from $\mathfrak{S}$ such that

$$A = \bigsqcup_j A_j, \qquad B_n = \bigsqcup_i B_{ni}, \quad n = 1, 2, \dots \qquad i, j \text{ finite}$$

Define $C_{nij} := B_{ni} \cap A_j$. The sets C_{nij} are in $\mathfrak{S}$ and pairwise disjoint. It holds

$$\bigsqcup_j C_{nij} = B_{ni} \cap \bigsqcup_j A_j = B_{ni} \cap A = B_{ni}$$

$$\bigsqcup_{n=1}^{\infty} \bigsqcup_i C_{nij} = \bigsqcup_{n,i} B_{ni} \cap A_j = A \cap A_j = A_j$$

Due to the σ-additivity of m on $\mathfrak{S}$ we obtain

$$m(A_j) = \sum_{n=1}^{\infty} \sum_i m(C_{nij}) \tag{3.1}$$

$$m(B_{ni}) = \sum_j m(C_{nij}) \tag{3.2}$$

And by definition of m on $\mathcal{R}(\mathfrak{S})$

$$m(A) = \sum_j m(A_j) \tag{3.3}$$

$$m(B_n) = \sum_j m(B_{nj}) \tag{3.4}$$

From (3.1) - (3.4) it follows that

$$m(A) = \sum_{n=1}^{\infty} m(B_n)$$

(The sums w.r.t. i and j are finite, the infinite sum w.r.t. n converges.)

■

Lemma 3.0.16. *(Properties of a premeasure on a ring) Let m be a premeasure on a ring $\mathcal{R}$ and $A, A_1, A_2, \ldots \in \mathcal{R}$.*

1. *If $\bigsqcup_{k=1}^{\infty} A_k \subset A$, $A_i \cap A_j = \varnothing$ for $i \neq j$ then*

$$\sum_{k=1}^{\infty} m(A_k) \leq m(A)$$

2. *(σ-subadditivity) If $\bigcup_{k=1}^{\infty} A_k \supset A$ then*

$$\sum_{k=1}^{\infty} m(A_k) \geq m(A)$$

Proof. 1. By Theorem 3.0.10(1) it holds for any n

$$\sum_{k=1}^{n} m(A_k) \leq m(A)$$

So the series is bounded and letting $n \to \infty$ we yield the claim. Note that this property does not depend on the σ-additivity of m (whereas the next one does).
2. Since $\mathcal{R}$ is a ring, the sets

$$B_n = A_n \setminus \bigsqcup_{k=1}^{n-1} A_k \qquad (\star)$$

belong to $\mathcal{R}$. We have (it is useful to recall Figure 2.2.3.3 to see why)

$$A = \bigsqcup_{n=1}^{\infty} B_n, \quad B_i \cap B_j = \varnothing \text{ for } i \neq j \qquad B_n \subset A_n \qquad (\star\star)$$

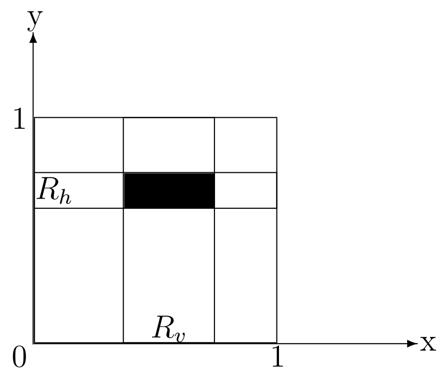
Hence

$$m(A) \leq \sum_{n=1}^{\infty} m(B_n) \leq \sum_{n=1}^{\infty} m(A_n)$$

A careful reader might be at first confused that we consider an union upto $(n-1)$-th element in $(\star)$ but then use a countable union in $(\star\star)$. But as a matter of fact n in $(\star)$ is arbitrary rather than fixed. And by assumption $A_n \in \mathcal{R}$ for any n hence either is B_n.

■

Remark 3.0.17. *(Importance of a semiring)
A curious reader may ask why it is important to use a semiring as a starting point. Why not to try to assign a premeasure m on another system of sets $\mathfrak{S}$ and then to try to extend it - first to the ring (or algebra), generated by $\mathfrak{S}$ and then to some σ-algebra.
The problem is that an extention of m even to the $\mathcal{R}(\mathfrak{S})$ may be not unique.
Here is an example
Let $\mathbb{X} := [0,1] \times [0,1]$ (i.e. $\mathbb{X}$ is a unit square) and let $\mathfrak{S}$ be a system of all vertical and horizontal rectangles R_h and R_v, i.e. such rectangles, whose length or width is equal to 1.



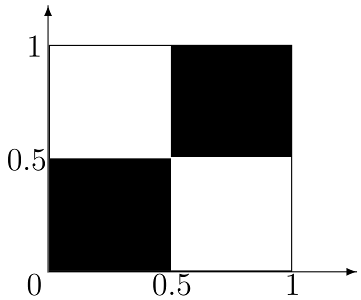
We define a "premeasure" s on $\mathfrak{S}$; for every $R \in \mathfrak{S}$ let $s(R)$ be equal to the area of R. The extention of s to the $\mathcal{R}(\mathfrak{S})$ (the ring generated by $\mathfrak{S}$) is not unique.
First of all let us find out what $\mathcal{R}(\mathfrak{S})$ is. Well, every rectangle in $\mathbb{X}$ (that is

every element of a product-semiring $\mathcal{I}\times\mathcal{I}$) can be obtained as an intersection $R_h \cap R_v$ of two rectangles from $\mathfrak{S}$. Since $\mathcal{R}(\mathfrak{S})$ is closed under intersection, we obtain that $\mathcal{R}(\mathfrak{S}) \supset \mathcal{I}\times\mathcal{I}$. The converse inclusion $\mathfrak{S} \subset \mathcal{R}(\mathcal{I}\times\mathcal{I})$ is obvious since $\mathfrak{S} \subset \mathcal{I}\times\mathcal{I}$ thus the rings generated by $\mathfrak{S}$ and by $\mathcal{I}\times\mathcal{I}$ coincide. By Theorem 3.0.8 any element of $\mathcal{R}(\mathfrak{S})$ can be represented as a finite disjoint union of rectangles.
So an obvious extention of s from $\mathfrak{S}$ to $\mathcal{R}(\mathfrak{S})$ is the following: first the measure s of $R_h \cap R_v$ is defined

$$s(R_h \cap R_v) := s(R_h) \cdot s(R_v)$$

Thus we have a measure on a semiring $\mathcal{I}\times\mathcal{I}$ and this measure is nothing else but area.

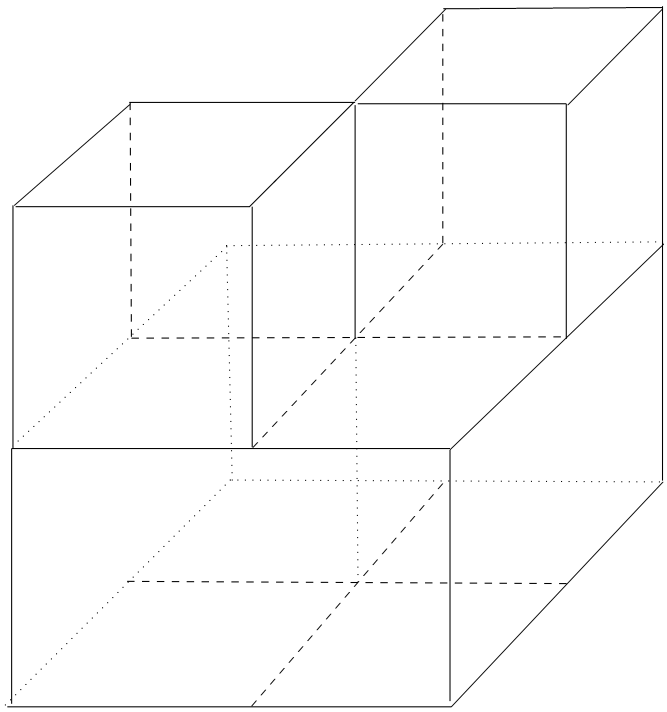
However, we can proceed another way.



Consider an arbitrary rectangle. We multiply the area of its intersection with black quadrants by $\frac{3}{2}$ and by $\frac{1}{2}$ the area of its intersection with white quadrants. You easily check that the measure of every element from $\mathfrak{S}$ is still equal to its area but for arbitrary element from $\mathcal{I}\times\mathcal{I}$ it is not the case. Note that this new measure is *not* translation-invariant (well, we will later prove that in $\mathbb{R}^n$ there is a unique translation-invariant measure, the Lebesgue one).

If you are still confused with this non-translationinvariant "area", the 3rd dimention will help you to get a clue. Look at the picture below. You may conceive this new measure as a *volume* of a parallelepiped with the base lying within our unit square. If you start shifting a rectangle from $\mathfrak{S}$, its volume will remain constant since (due to symmetry of our spatial figure) one half of its base will always be under the lower "roof" and another half under the higher one. But for an arbitrary rectangle from $\mathcal{I}\times\mathcal{I}$ it is of course not the case.

And the last note to this highly edifying example: we have connected here the area and the volume. But they are the measures in, respectively, $\mathbb{R}^2$ and $\mathbb{R}^3$. This is an "overture" to very important topics: Fubini and Radon-Nikodym Theorems. These theorems, however, can be learnt only after getting familiar with measurable functions.



So far we have learnt howto extend a volume or a premeasure from a semiring $\mathfrak{S}$ to the ring $\mathcal{R}(\mathfrak{S})$ generated by this semiring. If we deal with an additive but not σ-additive volume then, generally speaking, we cannot proceed further. In case of a premeasure (which is by definition σ-additive) we can go substantively further and extend the premeasure to, in a sence, the maximal system of sets (what do you expect this system to be?).

From now on we *could have* distinguished two cases as a starting point: a semiring with and without *unity*. A unity is a set $U \in \mathfrak{S}$ such that $U \cap B = B$ for any $B \in \mathfrak{S}$. In other words, U is a maximal element and contains all other elements of $\mathfrak{S}$. In the previous chapter we have comprehensively considered the semiring $\mathcal{I}$ whose unity was an interval $(0, 1)$. On the other hand, if we take not $(0, 1)$ but $\mathbb{R}^1$ as an underlying space, we can consider a semiring $\mathcal{I}_{\mathbb{R}^1}$, which consists of all bounded(!) intervals (a, b), $(a, b]$, $[a, b)$ and $[a, b]$ (prove that $\mathcal{I}_{\mathbb{R}^1}$ is indeed a semiring!). Obviously, $\mathcal{I}_{\mathbb{R}^1}$ has no unity since whatever

a and b we take, there is an even larger interval with endpoints, say, $a-1$ and $b+1$. A unity in this case would be $\mathbb{R}^1$ itself, but it is unbounded and thus does not belong[3] to $\mathcal{I}_{\mathbb{R}^1}$.

We, however, will *not* distinguish between the cases with/without unity. Instead, we proceed according to a powerful *Carathéodory construction.* At the beginning it may seem to be too abstract but later you will see its connection to the Valleé-Poussin's measurability criterion.
Henceforward we will actively deal with an outer measure. Contrary, we will *not* need an inner measure (since the latter makes sense only on a semiring with unity, recall Section 2.4 from the previous chapter).

Definition 3.0.18. Let $\mathbb{X}$ be an abstract space.
A function $\lambda^* : \mathcal{P}(\mathbb{X}) \to [0, \infty]$ is called an *outer measure* if
(om1) $\lambda^*(\varnothing) = 0$
(om2) $E, E_1, E_2, ... \subset \mathbb{X}$ and $E \subset \bigcup\limits_{k=1}^{\infty} E_k$ then $\lambda^*(E) \le \sum\limits_{k=1}^{\infty} \lambda^*(E_k)$

Remark 3.0.19.

1. λ^* is defined on $\mathcal{P}(\mathbb{X})$ i.e. on *all* subsets of $\mathbb{X}$.
2. The union $\bigcup_{k=1}^{\infty} E_k$ in (om2) is at most countable and not necessarily disjoint.
3. The property (om2) is called the subadditivity, so an outer measure is subadditive. But it does *not* have to be σ-additive and even additive.

Exercise 3.0.20.

1. In the previous chapter (Definition 2.2.3.1) we introduced an outer measure on $(0, 1)$. Prove that it is consistent with Definition 3.0.18.
2. Prove the monotonicity of the outer measure, i.e. the fact that if $A \subset B \subset \mathbb{X}$ then $\lambda^*(A) \le \lambda^*(B)$. (Hint: The proof is very straightforward but if you nevertheless expirience difficulty your should re-read the proof of Lemma 2.2.3.15)

In the Definition 2.2.3.1 we obtained the outer measure λ^* by means of a [pre]measure ℓ defined on the semiring $\mathcal{I}$. Now we consider an abstract analog of this approach.

[3] though as you (should) remember, $\mathbb{R}^1$ can be obtained as a countable union of the intervals from $\mathcal{I}_{\mathbb{R}^1}$; $\mathbb{R}^1 = \bigcup_{k=1}^{\infty} (-k, k)$. Whatever k we take, $(-k, k)$ is still bounded.

Theorem 3.0.21. *Let m be a premeasure on a semiring $\mathcal{S}$ and let $\lambda^* : \mathcal{P}(\mathbb{X}) \to [0, \infty]$ be defined as follows:*

$$\lambda^*(E) := inf \left\{ \sum_{k=1}^{\infty} m(S_k) \right\}$$

if there exists an at most countable covering of E from $\mathcal{S}$, i.e. $S_1, S_2, \ldots \in \mathcal{S}$ such that $E \subset \bigcup_{k=1}^{\infty} S_k$.
Otherwise we assign

$$\lambda^*(E) := +\infty$$

Then λ^* is an outer measure on $\mathbb{X}$ and $\lambda^*(S) = m(S)$ for all $S \in \mathcal{S}$.

Proof. First of all we are allowed to assume (Theorem 3.0.15) that m is extended to $\mathcal{R}(\mathcal{S})$.
(om1) holds since by definition $m(\varnothing) = 0$ and $\varnothing \in \mathcal{S}$. (om2) must be checked only for the case $\sum_{k=1}^{\infty} \lambda^*(E_k) < +\infty$. For this case we fix $\varepsilon > 0$ and for every E_k find a covering S_{kn} $(n = 1, 2, \ldots)$ from $\mathcal{S}$ such that

$$\sum_{n=1}^{\infty} m(S_{kn}) < \lambda^*(E_k) + \frac{\varepsilon}{2^k}$$

Such covering does exist since λ^* is defined as an *infimum* of all coverings from $\mathcal{S}$. The sets S_{kn} $(k, n = 1, 2, \ldots)$ constitute a countable covering of a set E, it holds

$$E \subset \bigcup_{k=1}^{\infty} E_k \subset \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} S_{kn}$$

Thus

$$\lambda^*(E) \leq \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} m(S_{kn})$$

and on the other hand

$$\sum_{k=1}^{\infty} \sum_{n=1}^{\infty} m(S_{kn}) < \sum_{k=1}^{\infty} \left(\lambda^*(E_k) + \frac{\varepsilon}{2^k} \right) \leq \sum_{k=1}^{\infty} \lambda^*(E_k) + \varepsilon$$

Since ε is arbitrary we yield that

$$\lambda^*(E) \leq \sum_{k=1}^{\infty} \lambda^*(E_k)$$

so λ^* is an outer measure.

In order to prove equality

$$\lambda^*(S) = m(S) \qquad \forall S \in \mathcal{S}$$

we at first note that $\lambda^*(S) \leq m(S)$ since S is covered by itself (and the λ^* is defined through *infimum* of coverings). The converse inequality follows from the Lemma 3.0.16: $m(S) \leq \sum_{k=1}^{\infty} m(S_k)$ for any covering of S thus $m(S) \leq \lambda^*(S)$

■

We say that the outer measure λ^* is *generated by the measure* m .

Now let λ^* be an arbitrary outer measure on $\mathbb{X}$. Take two sets $A, E \subset \mathbb{X}$. We say that A splits E *nicely* if

$$\lambda^*(E) = \lambda^*(E \cap A) + \lambda^*(E \cap A^c) \tag{3.5}$$

Usually it is difficult to check the equality in (3.5) directly, so we have to check "$\leq$" and "$\geq$" conditions. In is clear that by the subadditivity of an outer measure "$\leq$" always holds true. That's why in order to state that A splits E nicely we have to check only that

$$\lambda^*(E) \geq \lambda^*(E \cap A) + \lambda^*(E \cap A^c) \qquad (\dagger)$$

whereas this may be a non-trivial task only if $\lambda^*(E) < +\infty$. Since $A = (A^c)^c$ it follows that if A splits E nicely then either does A^c.

Definition 3.0.22. (Measurability by Carathéodory) The set $A \subset \mathbb{X}$ is called λ^*-measurable (or simply *measurable*) if A splits nicely every(!) set $E \subset \mathbb{X}$. We denote by μ the projection of λ^* from $\mathcal{P}(\mathbb{X})$ to the system $\mathcal{A}$ of measurable sets, i.e.

$$\mu(A) := \lambda^*(A) \text{ for any } \textit{measurable}\, A \subset \mathbb{X}$$

Before we proceed further let us halt for a while. At first glance the Carathéodory's approach may seem to be ugly, indeed it is very abstract and moreover even in the simplest case $\mathbb{X} := \mathbb{R}^1$ we have to check (†) for whole $\mathcal{P}(\mathbb{R}^1)$, whose cardinality is higher than continuum!
Well, yes it *is* abstract. And I am not able to give an example, which makes it "concreter" (maybe you will be). As to the anxiety w.r.t. the technical issues, so I say they are groundless! Indeed, recall Chapter 1 and our long way to the fact that the system of measurable sets is a σ-algebra. By means of the Carathéodory's approach we get it directly from the definitions of the outer measure, measure and measurability!

Theorem 3.0.23. *The system $\mathcal{A}$ of all λ^*-measurable subsets of $\mathbb{X}$ is a σ-algebra and μ is a measure on $\mathcal{A}$*

Proof. As a first step we show that $\mathcal{A}$ is an algebra and μ is additive. Since for any $E \subset \mathbb{X}$ we have

$$\lambda^*(E) = \lambda^*(E \cap \mathbb{X}) + \lambda^*(E \cap \mathbb{X}^c) = \lambda^*(E \cap \mathbb{X}) + \lambda^*(E \cap \varnothing)$$

both $\mathbb{X}$ and $\varnothing$ belong to $\mathcal{A}$. Moreover, if $A \in \mathcal{A}$ then either $A^c \in \mathcal{A}$. Now let $A_1, A_2 \in \mathcal{A}$ and $B := A_1 \cap A_2$. Let us show that B splits any $E \subset \mathbb{X}$ nicely. It holds

$$\lambda^*(E) = \lambda^*(E\cap A_1)+\lambda^*(E\cap A_1^c) = \lambda^*(E\cap A_1\cap A_2)+\lambda^*(E\cap A_1\cap A_2^c)+\lambda^*(E\cap A_1^c)$$

On the other hand

$$\lambda^*(E \cap B) + \lambda^*(E \cap B^c) = \lambda^*(E \cap B) + \lambda^*(E \cap B^c \cap A_1) + \lambda^*(E \cap B^c \cap A_1^c)$$

But $B^c \supset A^c$, $B^c \cap A_1 = A_1 \cap A_2^c$ thus the right-hand parts in both equalities are the same and

$$\lambda^*(E) = \lambda^*(E \cap B) + \lambda^*(E \cap B^c)$$

Thus (compare to Remark 2.2.2.2) $\mathcal{A}$ is an algebra.

Let $A_1, A_2 \in \mathcal{A}$, $A_1 \cap A_2 = \varnothing$ and $A := A_1 \sqcup A_2$. Then for any $E \subset \mathbb{X}$ we have

$$\lambda^*(E \cap A) = \lambda^*(E \cap A \cap A_1) + \lambda^*(E \cap A \cap A_1^c)$$

But $A \cap A_1 = A_1$, $A \cap A_1^c = A_2$ hence

$$\lambda^*(E \cap A) = \lambda^*(E \cap A_1) + \lambda^*(E \cap A_2) \tag{3.6}$$

In the special case when $E = A$ we use the fact that $A, A_1, A_2 \in \mathcal{A}$ and obtain

$$\mu(A) = \lambda^*(A) = \lambda^*(A_1) + \lambda^*(A_2) = \mu(A_1) + \mu(A_2)$$

thus μ is additive.

Finally we show that $\mathcal{A}$ is a σ-algebra and μ is σ-additive. Consider at first a disjoint sequence $A_k \in \mathcal{A}$ ($k \in \mathbb{N}$) and let $A := \bigsqcup_{k=1}^{\infty} A_k$. For any $p \in \mathbb{N}$ we assign $B_p := \bigsqcup_{k=1}^{p} A_k$ (p is arbitrary but finite). Since $\mathcal{A}$ is an algebra, we obtain that $B_p \in \mathcal{A}$ which means that for any $E \subset \mathbb{X}$

$$\lambda^*(E) = \lambda^*(E \cap B_p) + \lambda^*(E \cap B_p^c)$$

Applying (3.6) to the first summand at the right-hand side yields

$$\lambda^*(E) = \sum_{k=1}^{p} \lambda^*(E \cap A_k) + \lambda^*(E \cap B_p^c)$$

Since $B_p^c \supset A^c$, it holds by the monotonicity of the outer measure that

$$\lambda^*(E) \geq \sum_{k=1}^{p} \lambda^*(E \cap A_k) + \lambda^*(E \cap A^c)$$

The left-hand side does not depend on p and the inequality holds for any p. Hence we are allowed to let $p \to \infty$ and using the subadditivity of an outer measure (for $\bigsqcup_{k=1}^{\infty} (E \cap A_k) \supset (E \cap A)$) we come to (recall (†))

$$\lambda^*(E) \geq \sum_{k=1}^{\infty} \lambda^*(E \cap A_k) + \lambda^*(E \cap A^c) \geq \lambda^*(E \cap A) + \lambda^*(E \cap A^c)$$

The converse inequality holds again by the subadditivity of λ^* thus A splits E nicely, that is, A is measurable or in other words $A \in \mathcal{A}$. For $E = A$ we obtain

$$\mu(A) = \lambda^*(A) = \sum_{k=1}^{\infty} \lambda^*(A_k) = \sum_{k=1}^{\infty} \mu(A_k)$$

To see that $\mathcal{A}$ is a σ-algebra, consider an arbitrary (not necessarily disjoint) sequence $A_k \in \mathcal{A}$ $(k \in \mathbb{N})$ and let $A := \bigcup_{k=1}^{\infty} A_k$. Recalling the Figure 2.2.3.3 you easily construct a sequence of *disjoint* sets $\widehat{A_k}$ from $\mathcal{A}$ such that $A = \bigsqcup_{k=1}^{\infty} \widehat{A_k}$ hence, as we just proved, $A \in \mathcal{A}$.

■

Theorem 3.0.23 demonstrates an elegance of Carathéodory's approach. In chapter 2 we did a lot of auxiliary stuff to come to the σ-algebra of the measurable sets, whereas now we obtain it directly from the definition. But of course one can appreciate this beauty only in such comparison.

From now on we say that the measure μ on the σ-algebra $\mathcal{A}$ is *generated* by the outer measure λ^*.

Theorem 3.0.24. *1. If $\lambda^*(A) = 0$ then $A \in \mathcal{A}$*

Indeed, in this case by the monotonicity of the outer measure it holds that $\lambda^*(E) \geq \lambda^*(E \cap A) + \lambda^*(E \cap A^c)$ since $(E \cap A^c) \subset E$ and $\lambda^*(E \cap A) \leq \lambda^*(A) = 0$

2. *(Completeness of the measure, generated by an outer measure).*
   If $A \in \mathcal{A}$, $\mu(A) = 0$ and $B \subset A$ then $B \in \mathcal{A}$.
   Since A is measurable, we have $\lambda^*(A) = \mu(A)$ and by the monotonicity of an outer measure we obtain that $\lambda^*(B) = 0$. It remains to apply the previos statement.

3. *If $A_1 \subset A \subset A_2$, $A_1, A_2 \in \mathcal{A}$ and $\mu(A_1) = \mu(A_2) < \infty$ then $A \in \mathcal{A}$ and $\mu(A_1) = \mu(A) = \mu(A_2)$*
   We note that $(A \setminus A_1) \subset (A_2 \setminus A_1)$ but $(A_2 \setminus A_1) \in \mathcal{A}$ and $\mu(A_2 \setminus A_1) = 0$ hence $(A \setminus A_1) \in \mathcal{A}$ and $A = A_1 \cup (A \setminus A_1)$ is measurable too.

■

Remark 3.0.25. Once again it is worth reminding that incomplete measures *do* exist too. Theorem 3.0.24(2) holds when we obtain a measure μ by means of the outer measure λ^*. But if we assign a measure in some other way, it may be incomplete (later we will see an example).

Theorem 3.0.26. *(λ^*-measurability criterion) Let $D \subset \mathbb{X}$. If for any $\varepsilon > 0$ there exist $A, B \in \mathcal{A}$ such that $A \subset D \subset B$ and $\mu(B \setminus A) < \varepsilon$ then $D \in \mathcal{A}$*

Proof. For an arbitrary $n \in \mathbb{N}$ we can find $A_n, B_n \in \mathcal{A}$ such that

$$A_n \subset D \subset B_n \quad and \quad \mu(B_n \setminus A_n) < \frac{1}{n}$$

Let

$$A := \bigcup_{n=1}^{\infty} A_n \qquad B := \bigcap_{n=1}^{\infty} B_n$$

Then $A, B \in \mathcal{A}$, $A \subset D \subset B$ and $(B \setminus A) \subset (B_n \setminus A_n)$ for any n. Hence $\mu(B \setminus A) = 0$ and due to Theorem 3.0.24(2) $(D \setminus A) \in \mathcal{A}$ and we obtain $D = ((D \setminus A) \cup A) \in \mathcal{A}$

■

Remark 3.0.27. In Theorem 3.0.26 we tightly constrict D between two measurable sets A and B. Though we call it "a criterion", it has just a technical importance. We can easily show that the converse statement holds too, i.e. that for any measurable set $D \subset \mathbb{X}$ there exist $A, B \subset \mathbb{X}$ such that $A \subset D \subset B$ and $\mu(B \setminus A) < \varepsilon$ for all $\varepsilon > 0$. This is trivial ☺ since we can always set $A := B := D$.

On the other hand we considered in Chapter 2 the Vallée-Poussin's measurability criterion, where we tightly covered D with an elementary set E

(i.e. with an element of a ring $\mathbf{E}$). In the [more general] Carathéodory construction neither the Vallée-Poussin's criterion plays such a fundamental role since *theoretically* we have nothing to do with (semi)rings, all what we need is an outer measure. But where we take this outer measure from?! So *in practice* we usually need a semiring with a premeasure on it in order to construct an outer measure on the $\mathbb{X}$, recall Theorem 3.0.21. Hence right now we will consider a transition "premeasure on a semiring → outer measure → σ-algebra of measurable sets" and afterthat clearly state the interconnection between Vallée-Poussin's criterion and Carathéodory construction.

Theorem 3.0.28. *Let $\mathcal{S}$ be a semiring of the sets from the space $\mathbb{X}$, m be a premeasure on $\mathcal{S}$, λ^* be an outer measure generated by m and μ be a projection of λ^* on the σ-algebra $\mathcal{A}$ of λ^*-measurable sets. Then $\mathcal{S} \subset \mathcal{A}$ and $\mu(S) = m(S)$ for all $S \in \mathcal{S}$.*

Proof. The statement $\mu(S) = m(S)$ for all $S \in \mathcal{S}$ follows from the Theorem 3.0.21 and Definition 3.0.22.
To prove the inclusion $\mathcal{S} \subset \mathcal{A}$ we proceed as follows: let $S \in \mathcal{S}$ and $E \subset \mathbb{X}$. We need to show that

$$\lambda^*(E) \geq \lambda^*(E \cap S) + \lambda^*(E \cap S^c)$$

and it suffices (why?) to consider the case $\lambda^*(E) < \infty$. According to the construction of λ^* we can for any $\varepsilon > 0$ find $S_n \in \mathcal{S}$ $(n = 1, 2, \ldots)$ so that

$$E \subset \bigcup_n S_n \qquad \sum_n m(S_n) < \lambda^*(E) + \varepsilon$$

Obiously

$$(E \cap S) \subset \left(S \cap \bigcup_n S_n \right) = \bigcup_n (S_n \cap S) \qquad (E \cap S^c) \subset \left(S^c \cap \bigcup_n S_n \right) = \bigcup_n (S_n \cap S^c)$$

Since a semiring is $\cap$-stable, $(S_n \cap S) \in \mathcal{S}$ thus for any n

$$\lambda^*(E \cap S) \leq \sum_n m(S_n \cap S) \tag{3.7}$$

The sets $S_n \cap S^c$ can be represented as

$$(S_n \cap S^c) = S_n \setminus (S_n \cap S)$$

and of cource $S_n \supset (S_n \cap S^c)$. Hence for any n there exists a finite sequence of disjoint sets $\{C_{nk}\}_k \in \mathcal{S}$ such that $(S_n \cap S^c) = \bigsqcup_k C_{nk}$. By the additivity of m it holds

$$\sum_k m(C_{nk}) = m(S_n) - m(S_n \cap S)$$

Further we have

$$(E \cap S^c) \subset \bigcup_n \bigsqcup_k C_{nk}$$

thus

$$\lambda^*(E \cap S^c) \leq \sum_n \sum_k m(C_{nk}) = \sum_n m(S_n) - \sum_n m(S_n \cap S) \tag{3.8}$$

Summing up (3.7) and (3.8) we obtain

$$\lambda^*(E \cap S) + \lambda^*(E \cap S^c) \leq \sum_n m(S_n) < \lambda^*(E) + \varepsilon$$

and the claim holds because ε is arbitrary.

■

The usefulness of the Theorem 3.0.28 cannot be overstated. For example, in case of $\mathbb{X} := (0,1)$ it remains to show that $\mathcal{I}$ is a semiring and the length ℓ is a premeasure on $\mathcal{I}$. Then we automatically come to the σ-algebra $\mathcal{L}$ of Lebesgue-measurable sets on $(0,1)$.
Another example, which we are going consider right now, is the space $\mathbb{R}^n$. All we have to do is to find a semiring with a premeasure; have you already thought about n-dimensional parallelepipeds and their volume?!
But the greatest importance of the Theorem 3.0.28 is that it gives us a deeper understanding of stochastic processes and probability spaces a.k.a. Kolmogorov triples $(\Omega, \mathcal{F}, \mathbb{P})$. For now we will just cast a glance on how we come from a stochastic process to "a semiring of its finite-dimensional distribution" (aha, $\mathbb{R}^n$!) and then to the σ-algebra of *events* $\mathcal{F}$. Of course this will be a superficial overview because we have, so far, given a rigorous definition neither to a stochastic process nor even to a probability space. But I do believe such informal introduction is quite suitable here. And even necessary (to compensate your motivation after a long series of abstract stuff).

But before we start with [Lebesgue] measure in $\mathbb{R}^n$ and then with stochastic processes, we need to make an important remark. Example 2.2.2.5 in Chapter 2 gives us *two* examples of a semiring: a system $\mathcal{H}$ of half-open intervals $[a,b)$ and $\mathcal{I}$. So the question is, which semiring should we start with?! It turns out, we can start with what we like - either with $\mathcal{H}$ or with $\mathcal{I}$ - the result will be the same. Later we will take time for a formal proof of this statement but for now just recall Theorem 3.0.15: since $\mathcal{R}(\mathcal{H}) = \mathcal{R}(\mathcal{I})$ (i.e. the rings generated by $\mathcal{H}$ and $\mathcal{I}$ coincide) and we consider the same premeasure on both $\mathcal{H}$ and $\mathcal{I}$, we will yield the same extention of this premeasure to the ring $\mathcal{R} := \mathcal{R}(\mathcal{H}) = \mathcal{R}(\mathcal{I})$. So one can believe that for the σ-algebra $\mathcal{A}(\mathcal{R})$ everything will be as good as for $\mathcal{R}$.

Definition 3.0.29. Assume there are given n pairs $a_i, b_i \in \mathbb{R}$ and $a_i \le b_i$. A parallelepiped in n-dimensional [Euclidean] space is a set consisting of points (or vectors) $x = (x_1, \dots, x_n)$ such that $a_1 <\leqslant x_1 <\leqslant b_1, \dots, a_n <\leqslant x_n <\leqslant b_n$ ($<\leqslant$ means "less" *or* "less or equal").

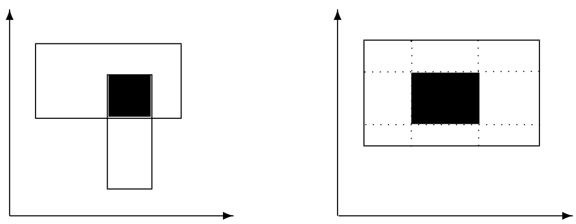
Definition 3.0.30. An n-dimensional parallelepiped such that $a_i \le x_i < b_i$ for all $i = 1, \dots, n$ if called a *right-open (left-closed) cell.*

Remark 3.0.31. Any n-dimensional parallelepiped is nothing else but a *Cartesian product* of some $I_1, \dots, I_n \in \mathcal{I}_{\mathbb{R}}$; as usually $\mathcal{I}_{\mathbb{R}}$ means a semiring of all intervals in $\mathbb{R}^1$.
Respectively, any right-open cell is a Cartesian product of some $H_1, \dots, H_n \in \mathcal{H}_{\mathbb{R}}$, where $\mathcal{H}_{\mathbb{R}}$ stands for a semiring of intervals $[a, b) \subset \mathbb{R}^1$.

Definition 3.0.32. A *volume* v of an n-dimensional parallelepiped P is defined as

$$v(P) := \prod_{i=1}^{n} (b_i - a_i) = \ell(I_1) \times \ldots \times \ell(I_n)$$

Our nearest objective is to prove that the systems of all parallelepipeds (resp. cells) in $\mathbb{R}^n$ are semirings and that v is a premeasure on them.
One way, which is rather popular in literature is to do it directly, i.e. first we show that an intersection of two parallelepipeds (cells) is again a paralleliped (a cell) and if $P \subset \hat{P}$ then there exist *disjoint* parallelepipeds (cells) $P_1, \dots, P_k$ such that $\hat{P} \setminus P = \bigsqcup_{j=1}^{k} P_j$



Further one has to show the additivity and then the σ-additivity of v, whereas even limiting himself to this special case, one would expirience annoying technical difficulties (if you doubt, try to prove that if $P = \bigsqcup_{j=1}^{k} P_j$

then $\upsilon(P) = \sum_{j=1}^{k} \upsilon(P_j)$).

So we proceed another way around, namely, we are going to prove that a Cartesian product of semirings is a semiring too and then introduce the product-premeasures (so υ turns out to be a product-premeasure $\underbrace{\ell \times \ell \times \ldots \times \ell}_{n\,times}$). Such approach is not more complicated as a direct proof for the semirings in $\mathbb{R}^n$. Indeed, they share a lot of common features (just as any abstract semiring has much common with $\mathcal{I}$). So our approach is advantageous because it proves a more general statement.

Theorem 3.0.33. *If $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n$ are semirings then their Cartesian product $\mathcal{S} := \mathcal{S}_1 \times \mathcal{S}_2 \ldots \times \mathcal{S}_n$ is a semiring too.*
$\mathcal{S}$ is called a product semiring.

Proof. It is enough to prove the case $n = 2$ and then apply induction.
I. Let $A, B \in \mathcal{S}_1 \times \mathcal{S}_2$, which means that

$$A := A_1 \times A_2, \qquad A_1 \in \mathcal{S}_1 \qquad A_2 \in \mathcal{S}_2$$

$$B := B_1 \times B_2, \qquad B_1 \in \mathcal{S}_1 \qquad B_2 \in \mathcal{S}_2$$

Just to make it clear

$$A = \{(a_1, a_2) \mid a_1 \in A_1, a_2 \in A_2\} \qquad B = \{(b_1, b_2) \mid b_1 \in B_1, b_2 \in B_2\}$$

Then

$$A \cap B = \{(c_1, c_2) \mid c_1 \in (A_1 \cap B_1), c_2 \in (A_2 \cap B_2)\} = \underbrace{(A_1 \cap B_1)}_{\mathcal{S}_1} \times \underbrace{(A_2 \cap B_2)}_{\mathcal{S}_2}$$

i.e.

$$(A \cap B) \in \mathcal{S}$$

II. Assume additionally that $B_1 \subset A_1$ and $B_2 \subset A_2$. Since $\mathcal{S}_1$ and $\mathcal{S}_2$ are semirings, there are the decompositions

$$A_1 = B_1 \sqcup B_1^{(1)} \sqcup \ldots \sqcup B_1^{(k)} \qquad A_2 = B_2 \sqcup B_2^{(1)} \sqcup \ldots \sqcup B_2^{(l)}$$

But then $A = A_1 \times A_2$
$= (B_1 \times B_2) \sqcup (B_1 \times B_2^{(1)}) \sqcup \ldots \sqcup (B_1 \times B_2^{(l)})$
$\bigsqcup (B_1^{(1)} \times B_2) \sqcup (B_1^{(1)} \times B_2^{(1)}) \sqcup \ldots \sqcup (B_1^{(1)} \times B_2^{(l)})$
$\ldots\ldots\ldots$
$\bigsqcup (B_1^{(k)} \times B_2) \sqcup (B_1^{(k)} \times B_2^{(1)}) \sqcup \ldots \sqcup (B_1^{(k)} \times B_2^{(l)})$

In words this clumsy formula means that we start B_1 and consider its Cartesian products with $B_2, B_2^{(1)}, \ldots, B_2^{(l)}$.
Then We continue with $B_1^{(1)}$ and again make a product with $B_2, B_2^{(1)}, \ldots, B_2^{(l)}$.
We finish with a product $B_1^{(k)}$ and $B_2, B_2^{(1)}, \ldots, B_2^{(l)}$ and then take unions of all Cartesian products. What we see is that the first member of this decomposition $(B_1 \times B_2) =: B \in \mathcal{S}$ and all following members belong to $\mathcal{S}$ too.
Hence $A \setminus B$ can be represented as a disjoint union of some sets from $\mathcal{S}$.

■

Remark 3.0.34. Recall Remark 2.2.2.4 where we state that a [not necessarily finite or even countable] intersection of σ-algebras (algebras, rings) is a σ-algebra (an algebra, a ring) but for semirings it does not hold true.
Theorem 3.0.33 shows us, in a sence, the "reversed situation". A finite(!) Cartesian product of semirings (sic! product, not a union) will be a semiring. But for the rings, algebras and σ-algebras it is generally *not* the case.

Lemma 3.0.35. *Let $\mathcal{S}$ be a semiring and consider an arbitrary (in particular, not necessarily disjoint) sequence $A_1, A_2, \ldots, A_n \in \mathcal{S}$.*
There exists a disjoint *sequence $B_1, B_2, \ldots, B_t \in \mathcal{S}$ such that every A_k can be represented as a union*

$$A_k = \bigsqcup_{s \in M_k} B_s$$

where M_k is a subsequence of $\{1, 2, \ldots, t\}$.

Proof. For the case $n = 1$ the statement is trivial, we just set $t := 1$ and $B_1 := A_1$. Assume it holds for $n = m$ and consider a sequence $A_1, \ldots, A_m, A_{m+1} \in \mathcal{S}$. Let $B_1, B_2, \ldots, B_t \in \mathcal{S}$ be the sets satisfying to the assumption w.r.t. the sets $A_1, \ldots, A_m$. Define

$$B_{s_1} := A_{m+1} \cap B_s$$

(s runs from 1 to t and some B_{s1} may be empty). Due to Lemma 3.0.7 we have a decomposition

$$A_{m+1} = \left(\bigsqcup_{s=1}^{t} B_{s1} \right) \sqcup \left(\bigsqcup_{p=1}^{q} \widetilde{B_p} \right) \qquad \widetilde{B_p} \in \mathcal{S}$$

and since $B_{s1} \subset B_s$ it holds by (smR2)

$$B_s = B_{s1} \sqcup B_{s2} \sqcup \ldots \sqcup B_{s\,f(s)} \qquad B_{sj} \in \mathcal{S} \quad (j = 1, ..f(s))$$

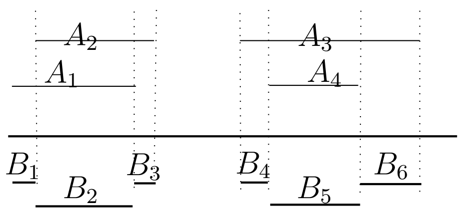
But then

$$A_k = \bigsqcup_{s\in M_k} \bigsqcup_{j=1}^{f(s)} B_{sj} \qquad k = (1,\ldots,m)$$

and moreover B_{sj} and $\widetilde{B}_p$ are pairwise disjoint (B_{s1}s and $\widetilde{B}_p$s constitute a disjoint decomposition of A_{m+1} whereas $B_{s2},\ldots B_{sj}$ are disjoint with each other and with A_{m+1} too).
Thus the sets B_{sj} and $\widetilde{B}_p$ satisfy the Lemma w.r.t. $A_1,\ldots,A_m,A_{m+1}$

■

Again there is a nice geometric interpretation (if you think of $\mathcal{I}$). You see that any A_k can be represented as a disjoint union of some Bs.



Theorem 3.0.36. *Let $\mathcal{S}_1,\mathcal{S}_2,\ldots,\mathcal{S}_n$ be semirings and $\upsilon_1,\upsilon_2,\ldots,\upsilon_n$ be volumes on them (in the sense of Definition 3.0.1). Consider a product semiring*

$$S := \mathcal{S}_1 \times \mathcal{S}_2 \times \ldots \times \mathcal{S}_n$$

and define a function υ on $\mathcal{S}$ as follows

$$\upsilon(A) := \upsilon_1(A_1)\upsilon_2(A_2)\ldots\upsilon_n(A_n) \qquad where \quad A := A_1 \times A_2 \times \ldots \times A_n$$

Then υ is a volume on the product semiring $\mathcal{S}$; it is called a product volume.

Proof. The property $\upsilon(\varnothing) = 0$ is clear and what remains is to show the additivity of υ. Let us do it for the case $n = 2$. Consider a decomposition

$$A := \underbrace{A_1}_{\in\mathcal{S}_1} \times \underbrace{A_2}_{\in\mathcal{S}_2} = \bigsqcup_{k=1} B^{(k)} \quad B^{(i)}\cap B^{(j)} = \varnothing \ for\, i \neq j \qquad B^{(k)} := \underbrace{B_1^{(k)}}_{\in\mathcal{S}_1} \times \underbrace{B_2^{(k)}}_{\in\mathcal{S}_2}$$

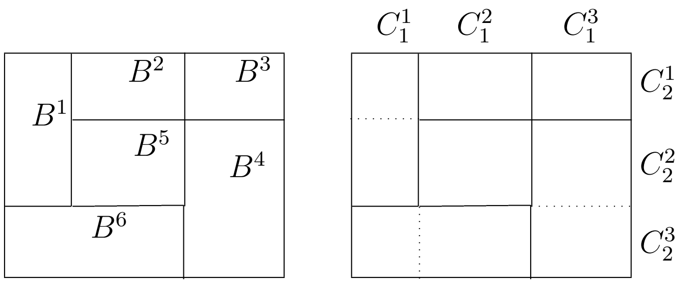
Due to Lemma 3.0.35 there exist the sequences

$$C_1^1, C_2^1, \dots C_m^1 \in \mathcal{S}_1 \qquad C_1^2, C_2^2, \dots C_n^1 \in \mathcal{S}_2$$

such that on the one hand

$$A_1 = \bigsqcup_m C_1^{(m)} \qquad A_2 = \bigsqcup_n C_2^{(n)}$$

and on the other hand $B_1^{(k)}$ are the unions of some $C_1^{(m)}$s, and $B_2^{(k)}$ are the unions of some $C_2^{(n)}$s.



Hence

$$\upsilon(A) = \upsilon_1(A_1)\upsilon_2(A_2) = \sum_n \sum_m \upsilon_1(C_1^{(m)})\upsilon_2(C_2^{(n)}) \tag{3.9}$$

$$\upsilon(B^{(k)}) = \upsilon_1(B_1^{(k)})\upsilon_2(B_2^{(k)}) = \sum_{m_k} \sum_{n_k} \upsilon_1(C_1^{(m_k)})\upsilon_2(C_2^{(n_k)}) \tag{3.10}$$

whereas at the right-hand side of (3.10) all summands of the right-hand side of (3.9) appear (for different k) exactly once.

Hence

$$\upsilon(A) = \sum_k \upsilon(B^{(k)})$$

■

Theorem 3.0.36 is proved for the product volume (i.e. for additive case). The next step would be to prove an analogous statement for the product premeasures (σ-additive case). Unfortunately we cannot do it before we get familiar with Lebesgue integral and B. Levy theorem. But for the case of n-dimensional parallelepipeds (cells) and $\upsilon := \prod_n \times\ell$ we can use our background from Chapter 2.

Denote an n-dimensional parallelepided in $\mathbb{R}^n$ with P the semiring of n-dimensional parallelepideds with $\mathcal{P}$ (in Chapter 2 for the case of $\mathbb{R}^1$ we had as counterparts, resp., I and $\mathcal{I}$). Let $\mathbf{E}$ be a ring generated by $\mathcal{P}$ (it is nothing else but the system of all sets, which can be represented as a finite disjoint union of n-dimensional parallelepideds, see Theorem 3.0.8). We know that we can uniquely extend v from $\mathcal{P}$ to $\mathbf{E}$ so that it remains additive in $\mathbf{E}$ (Theorem 3.0.9).
Now you can prove the σ-subadditivity of v in $\mathbf{E}$ by reproducing the proof of Lemma 2.2.3.5 where you formally replace I with P and $\mathcal{I}$ with $\mathcal{P}$. Finally you prove the σ-additivity of v in $\mathbf{E}$ analogous to Lemma 2.2.3.6.
So we obtain

Theorem 3.0.37. *A function $v : \mathcal{P} \to [0, \infty]$ introduced in the Definition 3.0.32 is σ-additive.*

■

So we have done a lot of necessary measure-theoretic stuff and now it is right time to see how it is applied to the theory of stochastic processes. We will do it (so far!) informally, since otherwise we have to give a formal definition of a stochastic process, which requires a knowledge of measurable functions (measurable w.r.t product σ-algebras), which... No, so much formalism *at once* is harmful for the mental digestion.
Thus we first recall that we have already encountered stochastic processes in Chapter 1. We observed how the indices values (DAX, DJIA etc) evolved in time. So it is clear the the index value is a function of time. But which function? Obviuously, had we a different market conditions in the past, the indices dynamics would be different. *Hence informally we conceive a stochastic process as a function of time, where the type of functional dependence is random and not known in advance.*
Now recall an important observation we had done in Chapter 1: returns on indices were [approximately] normally distributed and independent; for the first approach we agreed to ignore the deviations from normality. Of course our observations were done in discrete time. But we have done daily and hourly observations (indeed we could have even done them minutely) and saw that the normality in distribution of returns persists (though the parameters μ and σ certainly do depend on the discretization level between two adjacent observations).
Hence it was not too implausible to assume that whatever fine discretization level $\Delta t = t_1 - t_0$ we take, the index return from t_0 to t_1 will be $\mathcal{N}(\mu\Delta t, \sigma/\sqrt{\Delta t})$ distributed (eq. 1.1.2).
In other words we can say: we assume that we can make an observation at

any time with whatever fine discretization level and believe to obtain that the index retuns will be $\mathcal{N}(\mu\Delta t, \sigma/\sqrt{\Delta t})$ distributed. Or, if the time span between adjacent observations varies, we expect the return at time t_i to be $\mathcal{N}(\mu\Delta t_i, \sigma/\sqrt{\Delta t_i})$ distributed, where $\Delta t_i = t_i - t_{i-1}$. However, there is an important restriction: in practice we obviously can never make *infinitely* many observations. Even if we allow a highly implausible assumption of infinitely many observations, we are still in trouble due to a purely technical mathematical problem. We know how to deal with a *finite* sequence of normally distributed random variables, namely, by means of multivariate normal distribution with mean vector $\vec{\mu}$ and covariance matrix Σ, in our case

$$\vec{\mu} = \begin{bmatrix} \mu\Delta t_1 \\ \mu\Delta t_2 \\ \vdots \\ \mu\Delta t_n \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \frac{\sigma}{\Delta t_1} & 0 & \cdots & 0 \\ 0 & \frac{\sigma}{\Delta t_2} & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \vdots & \frac{\sigma}{\Delta t_n} \end{bmatrix}$$

But what on earth shall we do in case of *infinite* sequence?.
Approximately this question (but of course rigourously and much more generally) was formulated and solved by A. N. Kolmogorov in his seminal paper *Grundbegriffe der Wahrscheinlichkeitsrechnung* (English translation *The foundation of probability*). Since Kolmogorov had to axiomatize the probability theory in such a way that in case of [emprically observable] finite probability space the axioms did conform to our expirience, he, when applied Carathéodory construction, worked with *algebras* (and not with *semirings*) of events[4].
Thus by the present time we frequently encounter the following situation: during a course on measure theory (in Germany it is usually a part of Analysis III, in Russia it is a separate course for the sophomores) a student learns Carathéodory construction w.r.t. semiring (usually $\mathcal{I}$). Further he takes a course in Stochastic Processes (Probability Theory II), which is usually started with Kolmogorov's extention theorem. But instead of deep analysis of this theorem, the lecturers just state that "it can be proved by means of standard tools from measure theory". So a poor student gets lost and does not see any connection to Analysis III.
That's why we will now scrutinize how to apply the Carathéodory construction to the study of stochastic processes. For simplicity, we (so far) restrict ourselves to the Wiener process. We will use semirings, since they are easier

[4] As you (should) know the system of all events of a finite probability space is an algebra, i.e. closed under all set-theoretic operations.

(in particular, easier to find) than algebras[5].

So we assume that the Wiener process W_t (a.k.a Browinian motion)[6] starts in $(0, 0)$ and has independent[7] $\mathcal{N}(0, 1/\sqrt{\Delta t_i})$-distributed increments; $\Delta t_i = t_i - t_{i-1}$, $t_0 = 0$. Moreover, we assume that the Wiener process is almost surely continuous, which means that the probability of event W_t has a discontinuity is zero. This assumption is motivated again by emprical observations (Robert Brown, observing pollen in water saw no discontinuities ☺)[8]. Continuity of the Wiener process means that it is sufficient to treat it only in the set of rational points. Indeed, since for any $T < \infty$ the interval $[0, T]$ is compact, the a.s. continuity of W_t on $[0, T]$ implies even its a.s. *uniform* continuity on $[0, T]$. It means that W_{t_a} cannot deviate from W_{t_b} too much if t_a is close to t_b. But we always can find two rational points, which are however close to each other! And the set of all rational points is countable! This means that if we can handle W_t in all rational points (or equivalently can deal with infinite sequence of independent normally distributed random variables), we could exhaustively[9] characterize the probability of all events, associated with W_t. Since in Chapter 1 the W_t was only source of randomness, we would be able to treat *all* events in financial market!!!

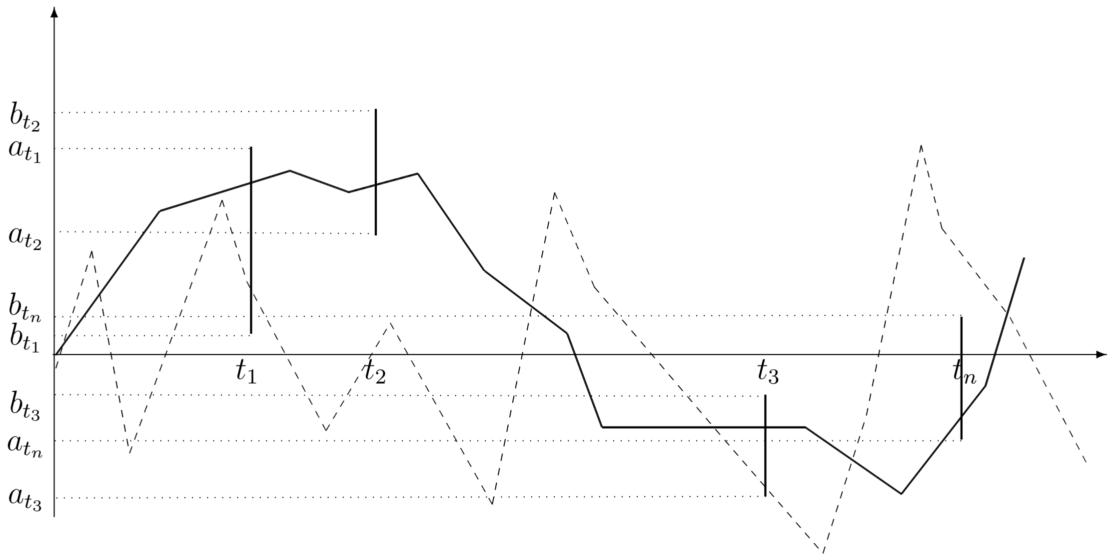
[5]However, it is worth recalling that a transition from a semiring to a ring is straightforward (Theorem 3.0.8) and an algebra is nothing but a ring with unity.

[6]From now on we, for simplicity, consider not the dynamic of the [logariphmic] price of a stock index but its random component, i.e. the Wiener process.

[7]By engaging the geometric intuition, this independence will allow us to use an *orthogonal* coordinate systes in $\mathbb{R}^n$.

[8]It turns out that the a.s. continuity of the W_t can be proved but (for now) we cannot concentrate on this proof.

[9]Upto events with zero probability.



We start as follows: for arbitrary moments of time $t_1, t_2, t_3, \dots t_n$ we consider the event[10] A that an increment of W_t belongs to the intervals

$$[a_{t_1}, b_{t_1}), [a_{t_2}, b_{t_2}), [a_{t_3}, b_{t_3}), \dots, [a_{t_n}, b_{t_n}) \qquad a_{t_i} \le b_{t_i} \quad a_{t_i}, b_{t_i} \in [-\infty, +\infty]$$

At the figure above you see that A occurs for the realization of the increments of W_t, depicted with a solid line and does not take place for the one, depicted with a dashed line.
Since we know the distribution of the increments, we can assign a probability to such events:

$$\mathbb{P}(A) = \prod_{i=1}^{n} \left[\Phi^i(b_{t_i}) - \Phi^i(a_{t_i})\right] \qquad \Phi^i(z) = \frac{1}{\sqrt{2\pi\Delta t_i}} \int_{-\infty}^{z} \exp\left(-\frac{x^2}{2\Delta t_i}\right) dx$$

As we remember, in $\mathbb{R}^1$ the length of an interval concides with its Lebesgue measure, which is σ-additive. It turns out that $\Phi^i(\cdot)$ defines another σ-additive measure of an interval[11]! This is a staightforward corollary of Radon-Nikodým theorem (which we do not yet know). But one can easily prove it as follows.
Let

$$I := \bigsqcup_{k=1}^{\infty} I_k \qquad I := [a, b) \quad I_k := [a_k, b_k) \qquad I, I_1, I_2 \in \mathcal{I}_{\mathbb{R}^1}$$

Since a_{t_i}, b_{t_i} are allowed to be $\pm\infty$, we will include the intervals of infinite length like $(-\infty, x)$, $[x, \infty)$ and $(-\infty, \infty)$ in $\mathcal{I}_{\mathbb{R}^1}$. You easily checks that $\mathcal{I}_{\mathbb{R}^1}$

[10]From now on I will often call it 'an event of a special type'.
[11]NB! In what follows i is fixed, though our considerations hold for any $i = 1, \dots, n$.

remains a semiring after this inclusion.
But at first we assume that I is bounded, i.e. that $\ell(I) = \mu(I) < \infty$. It holds

$$\ell(I) = \mu(I) = \sum_{k=1}^{\infty} \mu(I_k) = \mu(I) = \sum_{k=1}^{n} \mu(I_k) + \sum_{k=n+1}^{\infty} \mu(I_k)$$

The function $f^i(x) := \frac{1}{\sqrt{2\pi\Delta t_i}} e^{-\frac{x^2}{2\Delta t_i}}$ is bounded. Let $M^i := \sup_x \{f^i(x)\}$. Further

$$\sum_{k=1}^{\infty} \left[\Phi^i(b_k) - \Phi^i(a_k)\right] = \sum_{k=1}^{\infty} \int_{a_k}^{b_k} f^i(x)dx = \sum_{k=1}^{n} \int_{a_k}^{b_k} f^i(x)dx + \sum_{k=n+1}^{\infty} \int_{a_k}^{b_k} f^i(x)dx$$

It holds

$$\sum_{k=n+1}^{\infty} \int_{a_k}^{b_k} f^i(x)dx \leq M^i \underbrace{\sum_{k=n+1}^{\infty} \mu(I_k)}_{\to 0 \, as \, n \to \infty}$$

Finally we use the linearity (additivity) of integral and obtain

$$\sum_{k=1}^{n} \int_{a_k}^{b_k} f^i(x)dx = \int_{a_k}^{b_k} \sum_{k=1}^{n} f^i(x)dx = \int_{a}^{b_n} f^i(x)dx \quad - \quad \underbrace{\int_{b_n}^{b} f^i(x)dx}_{\leq M(b-b_n) \to 0 \, as \, b_n \nearrow b}$$

So letting $n \to \infty$ and hence $b_n \nearrow b$ we yield the σ-additivity, i.e. that

$$\sum_{k=1}^{\infty} \int_{a_k}^{b_k} f^i(x)dx = \int_{a}^{b} f^i(x)dx = \left[\Phi^i(b) - \Phi^i(a)\right]$$

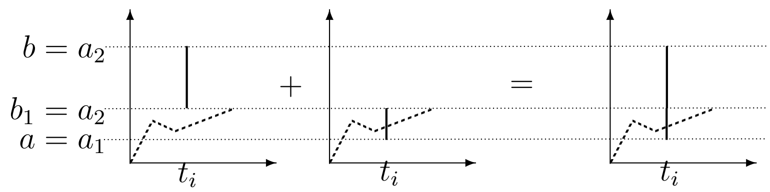
The case when I is unbounded can be easily reduced to the case of the bounded one since for any ε

$$\int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\Delta t_i}} e^{-\frac{x^2}{2\Delta t_i}} < \varepsilon \qquad \int_{b}^{\infty} \frac{1}{\sqrt{2\pi\Delta t_i}} e^{-\frac{x^2}{2\Delta t_i}} < \varepsilon$$

for sufficiently small a and large b. Thus we can consider $\widetilde{I} := [a, b)$ instead of I.

We see that $\Phi^i(\cdot)$ is σ-additive on intervals. And since $\Phi^i(\varnothing) = 0$ we convince ourselves that $\Phi^i(b) - \Phi^i(a)$ is indeed a premeasure on the semirings $\mathcal{I}_{\mathbb{R}^1}$ and $\mathcal{H}_{\mathbb{R}^1}$. It is worth reminding that now both these semirings contain $\mathbb{R}^1$ and that $\Phi^i(\cdot)$ is finite since even $\Phi^i(\mathbb{R}^1) = 1$.

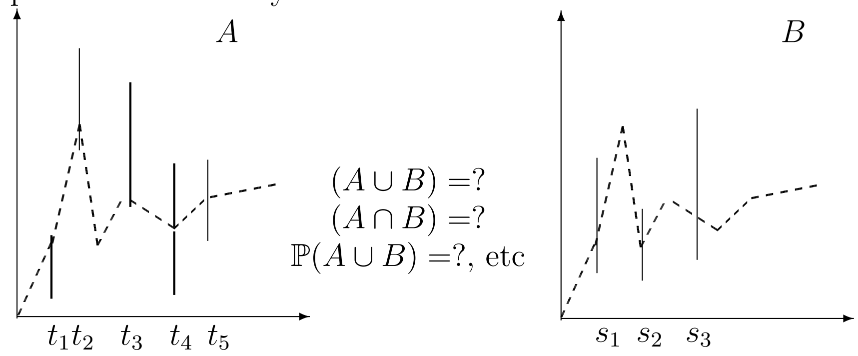
And once again, let us make clear, which σ-additivity we were just talking about. We fixed a *single* timepoint t_i and considered the *simplest* event, i.e. that the increment of W_t at time t_i will be within the range $[a, b)$ or, in other words, belongs to the interval $[a, b)$. Then we say that if $[a, b) = [a_1, b_1) \sqcup [b_1, a_1)$ then the probability of this event is equal to the probability of the event that the increment of W_t at time t_i will be within the range $[a_1, b_1)$ *OR* within the range $[a_2, b_2)$.



Thus we have reduced the consideration of the simplest events (associated with the increments of W_t to the consideration of $\mathcal{I}_{\mathbb{R}^1}$)! Thus we can straightforwadly consider slightly more compicated events, namely that at time t_i the value of the increment of W_t belongs to an element of the Borel- or Lebesgue σ-algebra on $\mathbb{R}^1$.
Indeed, the premeasure $\Phi^i(b) - \Phi^i(a)$ on the semiring $\mathcal{I}_{\mathbb{R}^1}$ can be extended to the measure on the Lebesgue σ-algebra $\mathcal{L}_{\mathbb{R}^1}$, which includes the Borel σ-algebra $\mathcal{B}_{\mathbb{R}^1}$.

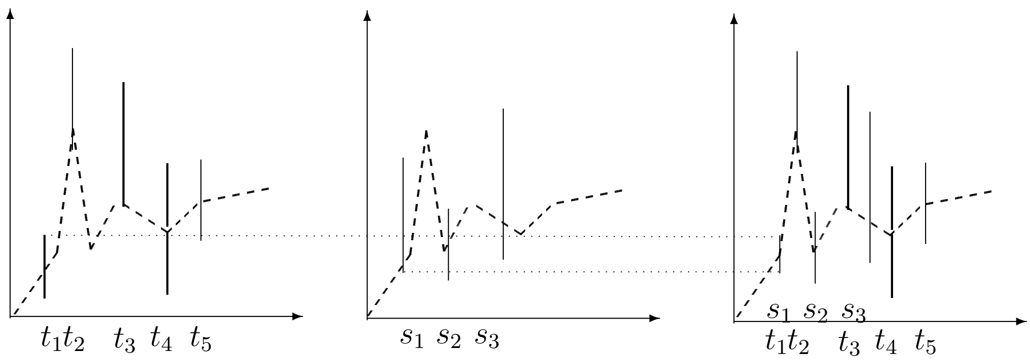
Our next objective is to be able to treat more complex events, when we fix several timepoints $t_1, t_2, \ldots t_n$ and consider the event A that the increments of W_t at $t_1, t_2, \ldots t_n$ belong, respectively, to $[a_1^t, b_1^t), [a_2^t, b_2^t), \ldots, [a_n^t, b_n^t)$. Moreover, there can be another event B that the increments of W_t at $s_1, s_2, \ldots s_n$ belong, respectively, to $[a_1^s, b_1^s), [a_2^s, b_2^s), \ldots, [a_m^s, b_m^s)$, whereas in general $t_1 \neq s_1$, $t_2 \neq s_2$ and moreover $n \neq m$. Actually, we have already defined the probability of such events, it is $\mathbb{P}(A) = \prod_{i=1}^n [\Phi^i(b_{t_i}) - \Phi^i(a_{t_i})]$. We must also be able to calculate the probability of the sum, difference and intersection of A and B, which may though be cumbersome, but reduces to elementary probability calculus. Much more important is to recognize the "remarkable properties" of the class of such events, this will give us power to treat the

Wiener process exhaustively.



So first of all we can define a semiring of such more complex events (yes, it will be, in a sence, a Cartesian product of $\mathcal{I}_{\mathbb{R}^1}$ but do not hurry up so far).

We start with an intersection of A and B. Obviously the event $A \cap B$ means that the increments of W_t belong at $t_1, t_2, \ldots t_n$ to $[a_1^t, b_1^t), [a_2^t, b_2^t), \ldots, [a_n^t, b_n^t)$ *AND* at $s_1, s_2, \ldots s_n$ to $[a_1^s, b_1^s), [a_2^s, b_2^s), \ldots, [a_m^s, b_m^s)$.



Graphically it means that we lay one chart on another and if for some i and j $t_i = s_j$ (in our case $t_1 = s_1$) we take an intersection of $[a_i^t, b_i^t)$ and $[a_j^s, b_j^s)$. So we (informally, inuitively) see: considering a finite number of events like A and B and taking their intersection we still obtain an event of the same type. It is very promising, since $\cap$-stability is a must for a semiring.

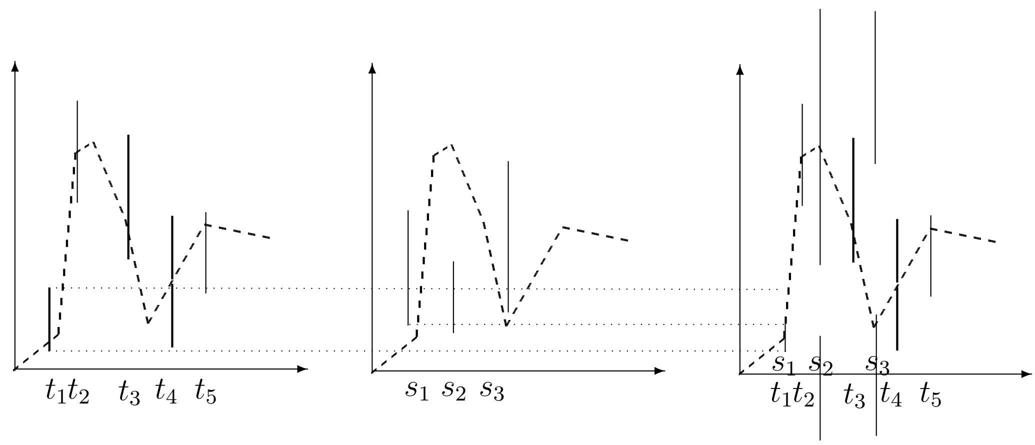
Next we try to represent this way an event $A \setminus B$, which means that the increments of W_t belong at $t_1, t_2, \ldots t_n$ to $[a_1^t, b_1^t), [a_2^t, b_2^t), \ldots, [a_n^t, b_n^t)$ *BUT* at $s_1, s_2, \ldots s_n$ *NOT* to $[a_1^s, b_1^s), [a_2^s, b_2^s), \ldots, [a_m^s, b_m^s)$.

It means the following: the increments of W_t belong at $t_1, t_2, \dots t_n$ to

$$[a_1^t, b_1^t), [a_2^t, b_2^t), \dots, [a_n^t, b_n^t)$$

AND at $s_1, s_2, \dots s_n$ to

$$[-\infty, a_1^s) \sqcup [b_1^s, \infty), [-\infty, a_2^s) \sqcup [b_2^s, \infty), \dots, [-\infty, a_m^s) \sqcup [b_m^s, \infty)$$



Or, equivalently, $A \backslash B$ means that the increments of W_t belong at $t_1, t_2, \dots t_n$ to

$$[a_1^t, b_1^t), [a_2^t, b_2^t), \dots, [a_n^t, b_n^t)$$

AND at $s_1, s_2, \dots s_n$ to

$$[-\infty, a_1^s), [-\infty, a_2^s), \dots, [-\infty, a_m^s)$$

$$OR$$

at $t_1, t_2, \dots t_n$ to

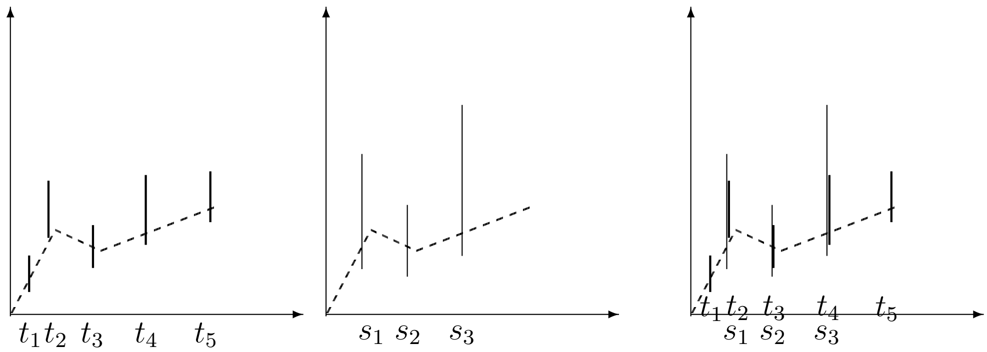
$$[a_1^t, b_1^t), [a_2^t, b_2^t), \dots, [a_n^t, b_n^t)$$

AND at $s_1, s_2, \dots s_n$ to

$$[b_1^s, \infty), [b_2^s, \infty), \dots, [b_m^s, \infty)$$

It looks better since our operands are again the events of the same type as A and B are. But there is a new operator OR, which is clearly to accociate with "$\cup$" just like we have already associated AND with "$\cap$". We need as well an idea of what "$A \subset B$" is, in order to show Definition2.2.2.6(smR3) holds.

Well, "$A \subset B$" means that $n \geq m$ and for every s_j $(j = 1,..,m)$ there exists t_i $(i = 1,..,n)$ such that $s_j = t_i$ and moreover $a^s_j \leq a^t_i, b^t_i \leq b^s_j$.



Note that it is really $n \geq m$ and not only $n = m$ as you might at first suppose (the picture above should make it clear).

In what follows I will also use a more compact notation for event, e.g.

$$A := \{[a^t_1, b^t_1), [a^t_2, b^t_2), \ldots, [a^t_n, b^t_n)\ at\ t_1, t_2, \ldots, t_n\}$$

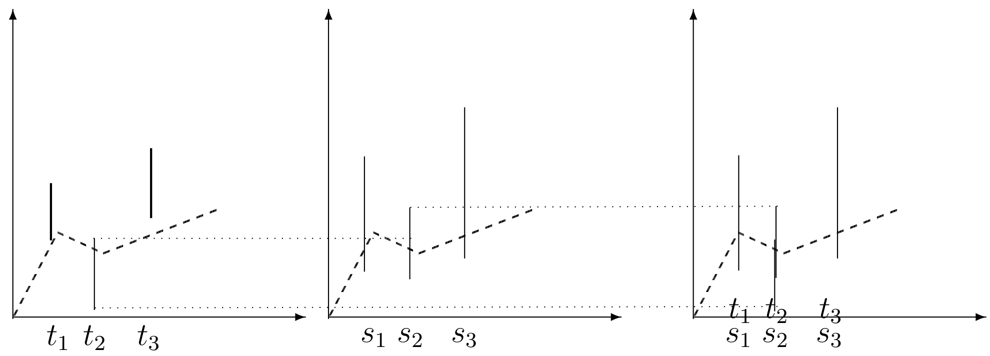
But now, as we are going to proceed with $A \cup B$, we are in danger, since the graphical intuition we exploit can make a malicious joke to us!...
We know that the union of two elemenens of a semiring *does not have to* belong to this semiring, so it not necessarily try to depict graphically *every* event (we sketch graphically only a special type of events, like A and B). But, however, $A \cup B$ *may* be in our semiring, and if it is, we have to be able to sketch it too.
And by "inertness of mind" we might try to assert (*in*correctly) the following : if A is an event that the increments of W_t belong at $t_1, t_2, \ldots t_n$ to $[a^t_1, b^t_1), [a^t_2, b^t_2), \ldots, [a^t_n, b^t_n)$ and B is an event that of increments W_t belong at $s_1, s_2, \ldots s_m$ to $[a^s_1, b^s_1), [a^s_2, b^s_2), \ldots, [a^s_m, b^s_m)$ and, moreover,

$$n = m, \quad t_1 = s_1, t_2 = s_2, \ldots, t_n = s_m$$

and finally

$$[a^t_1, b^t_1) \cup [a^s_1, b^s_1) \in \mathcal{I}_{\mathbb{R}^1},\ [a^t_2, b^t_2) \cup [a^s_2, b^s_2) \in \mathcal{I}_{\mathbb{R}^1},\ \ldots,\ [a^t_n, b^t_n) \cup [a^s_m, b^s_m) \in \mathcal{I}_{\mathbb{R}^1}$$

then $A \cup B$ belongs to the semiring of events. Graphical intuition at first supports this assertion:



And moreover, (smR3) seems to hold true ... but *hold true for what*?! If we had considered a semiring of *graphical sketches* at which vertical intervals are drawn at timepoints $t_1, t_2, \ldots, t_2$ then we did construct a semiring (in purely abstract algebraic sence). But these graphical sketches are just a tool to depict the events, associated with increments of W_t. AND A SEMIRING ITSELF (WITHOUT A PREMEASURE ON IT) DOES NOT BRING US ANYTHING! We also see that the additivity of $\mathbb{P}(\cdot)$ fails on this semiring and, what is even worse, this semiring does not conform to the class of our events. Indeed, there is a simple (counter)example. Consider

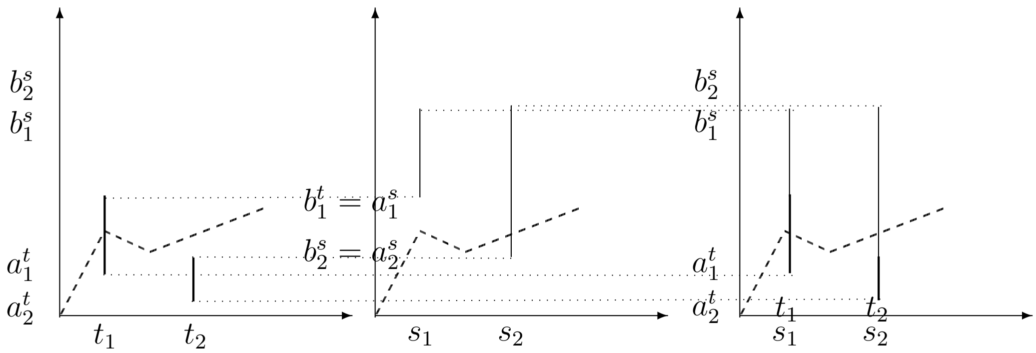
$$A = \{[a_1^t, b_1^t), [a_2^t, b_2^t) \text{ } at \text{ } t_1, t_2\} \qquad B = \{[a_1^s, b_1^s), [a_2^s, b_2^s) \text{ } at \text{ } s_1, s_2\}$$

and let

$$t_1 = s_1, \quad t_2 = s_2 \quad and \quad b_2^t = a_2^s, \quad b_1^t = a_1^s$$

Then $A \cap B = \varnothing$ and

$$A \cup B = A \sqcup B = \{[a_1^t, b_1^s), [a_2^t, b_2^s) \text{ } at \text{ } t_1, t_2\}$$



But the picture above contradicts the logic of events, indeed we see that for this realization of the increments of W_t *neither* A *nor* B occur but $A \sqcup B$ does!!

So does it mean we have to through our geometric intuition into a waste bin? Not at all! We just have to find an appropriate geometrical objects ☺. And it is a good time to return to the n-dimensional right-open cells. Now, however, thing will be slightly complicated. Namely, we have previously considered n-dimensional parallelepipeds and right-open cells in $\mathbb{R}^n$. But now it is more plausible to conceive the following construction: consider two events (we start with two and then by analogy can extend the construction to an arbitrary but finite number of events).

$$A = \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t) \; at \; t_1, \ldots, t_n\}$$

$$B = \{[a_1^s, b_1^s), \ldots, [a_m^s, b_m^s) \; at \; s_1, \ldots, s_m\}$$

Let d be the number of *unique* elements of the sequence $\{t_1, \ldots, t_n, s_1, \ldots, s_m\}$
So if $t_i \neq s_j$ for all pairs $(i, j) \quad i = 1, ..n, \; j = 1..m$ then obviously $d = n+m$.
If there are k pairs (i, j) such that $t_i = s_j$ then[12] $d = n + m - k$.
So $\mathbb{R}^d$ will be our underlying space and we assocciate events A and B with, respectively, the following d-dimensional cells

$$\widehat{A} := [a_1^t, b_1^t) \times \ldots \times [a_n^t, b_n^t) \times \underbrace{\mathbb{R}^1 \ldots \mathbb{R}^1}_{d-n \, times} = [a_1^t, b_1^t) \times \ldots \times [a_n^t, b_n^t) \times \mathbb{R}^{(d-n)}$$

$$\widehat{B} := [a_1^s, b_1^s) \times \ldots \times [a_m^s, b_m^s) \times \underbrace{\mathbb{R}^1 \ldots \mathbb{R}^1}_{d-m \, times} = [a_1^t, b_1^t) \times \ldots \times [a_m^t, b_m^t) \times \mathbb{R}^{(d-m)}$$

The probability measures of A and B are

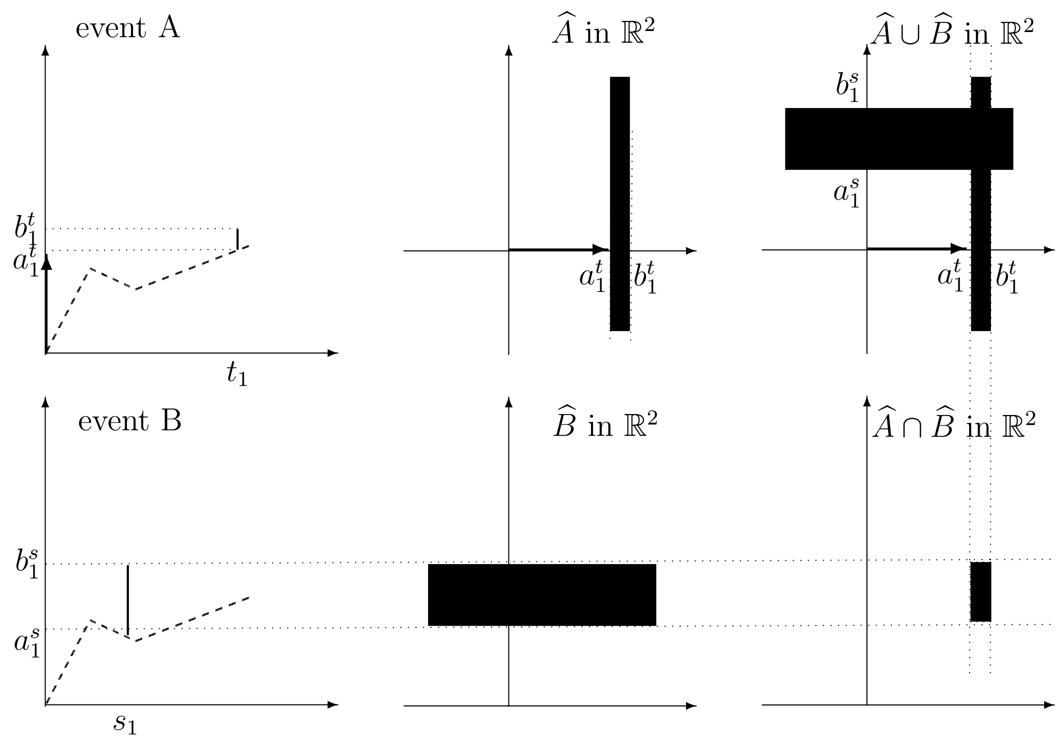
$$\mathbb{P}(A) = \prod_{i=1}^{d} \left[\Phi^i(b_{t_i}) - \Phi^i(a_{t_i})\right] = \prod_{i=1}^{n} \left[\Phi^i(b_{t_i}) - \Phi^i(a_{t_i})\right]$$

$$\mathbb{P}(B) = \prod_{i=1}^{d} \left[\Phi^i(b_{s_i}) - \Phi^i(a_{s_i})\right] = \prod_{i=1}^{m} \left[\Phi^i(b_{s_i}) - \Phi^i(a_{s_i})\right]$$

since if the i-th parallelepiped edge equals to $\mathbb{R}^1$ we have

$$\Phi^i(\infty) - \Phi^i(-\infty) = 1$$

[12]You readily check that $n \leq d$ and $m \leq d$.



The picture shows that in the simplest case A and B are associated with infinite strips $\widehat{A}$ and $\widehat{B}$ in $\mathbb{R}^2$ parrallel to the respective coordinate axes. There union is not a parallelepiped in $\mathbb{R}^2$ (well, need not to be) but there intersection is, as required from a semiring.

Below we consider the case $d = 2 + 1 - 0 = 3$. Though I did my best to make a 3D sketch, it might be easily to treat the task just algebraically ☺. Indeed,

$$\widehat{A \cap B} = \{(x, y, z) \in \{[a_1^t, b_1^t), [a_2^t, b_2^t), [a_1^s, b_1^s)\}\}$$

is just a cell in $\mathbb{R}^3$.

Note that it does not matter whether we first consider $A \cap B$ and then its geometrical counterpart $\widehat{A \cap B}$ or we first work with $\widehat{A} \cap \widehat{B}$ and then map it back to $A \cap B$.
In other words

$$\widehat{A \cap B} = \widehat{A} \cap \widehat{B}$$

I am sure you see it from the pictures but there is a simple formal proof. Let

$$A := \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t) \; at \; t_1, \ldots, t_n\}$$

$$B := \{[a_1^s, b_1^s), \ldots, [a_m^s, b_m^s) \text{ at } s_1, \ldots, s_m\}$$

Then

$$\widehat{A} = \{(x_1, \ldots, x_n) \in \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t)\}\}$$

$$\widehat{B} = \{(x_1, \ldots, x_m) \in \{[a_1^s, b_1^s), \ldots, [a_m^s, b_m^s)\}\}$$

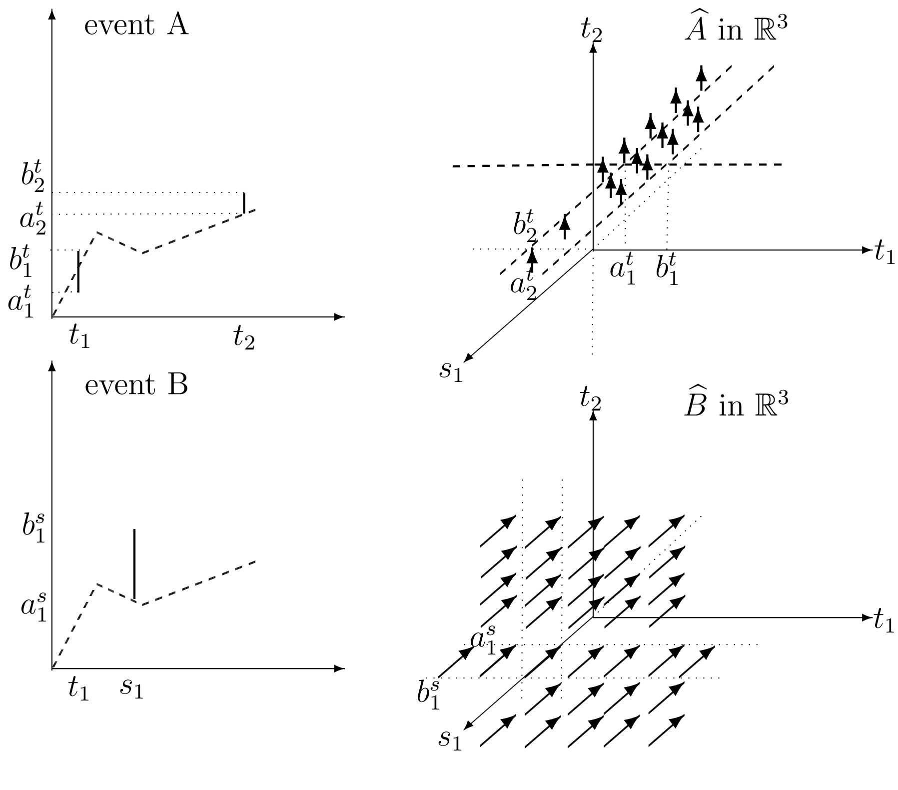
So

$$\widehat{A} \cap \widehat{B} = \{(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{m+n}) \in \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t), [a_1^s, b_1^s), \ldots, [a_m^s, b_m^s)\}\}$$

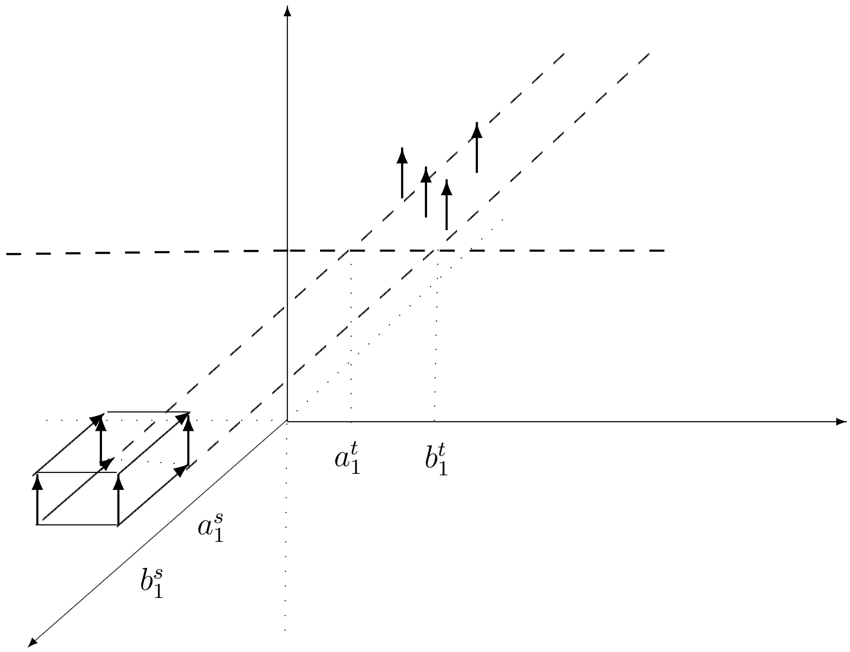
But

$$A \cap B = \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t), [a_1^s, b_1^s), \ldots, [a_m^s, b_m^s) \text{ at } t_1, \ldots, t_n\, s_1, \ldots, s_m\}$$

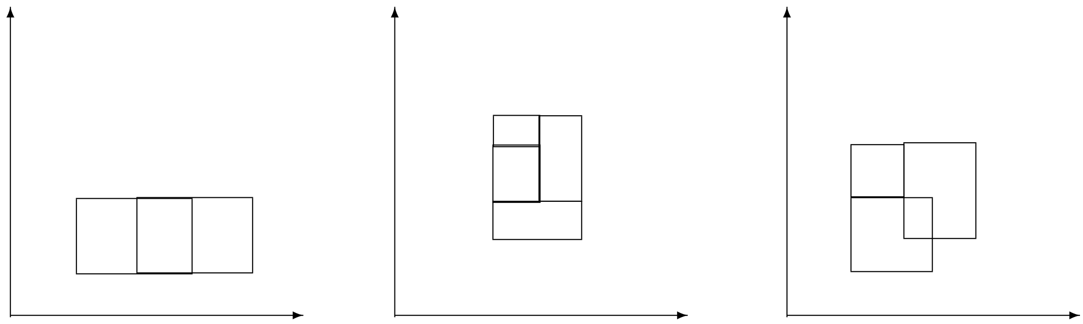
$$\widehat{A \cap B} = \{(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{m+n}) \in \{[a_1^t, b_1^t), \ldots, [a_n^t, b_n^t), [a_1^s, b_1^s), \ldots, [a_m^s, b_m^s)\}\}$$



$\widehat{A} \cap \widehat{B}$ in $\mathbb{R}^3$



So what remains is to find how to deal with "$\cup$"-operation, whereas we need to consider not only finite unions but countably-infinite too. We start with the case of the finite union.
Consider a finite sequence of events $A_1, A_2, \ldots A_p$. There is a *bijection* between them and the cells $\widehat{A_1}, \widehat{A_2}, \ldots \widehat{A_p}$ in $\mathbb{R}^d$.
If we unite a finite number of d-dimensional parallelepipeds, the union $\widehat{A} := \bigcup_{l=1}^{p} \widehat{A}_l$ may or may not be again a d-dimensional parallelepiped.

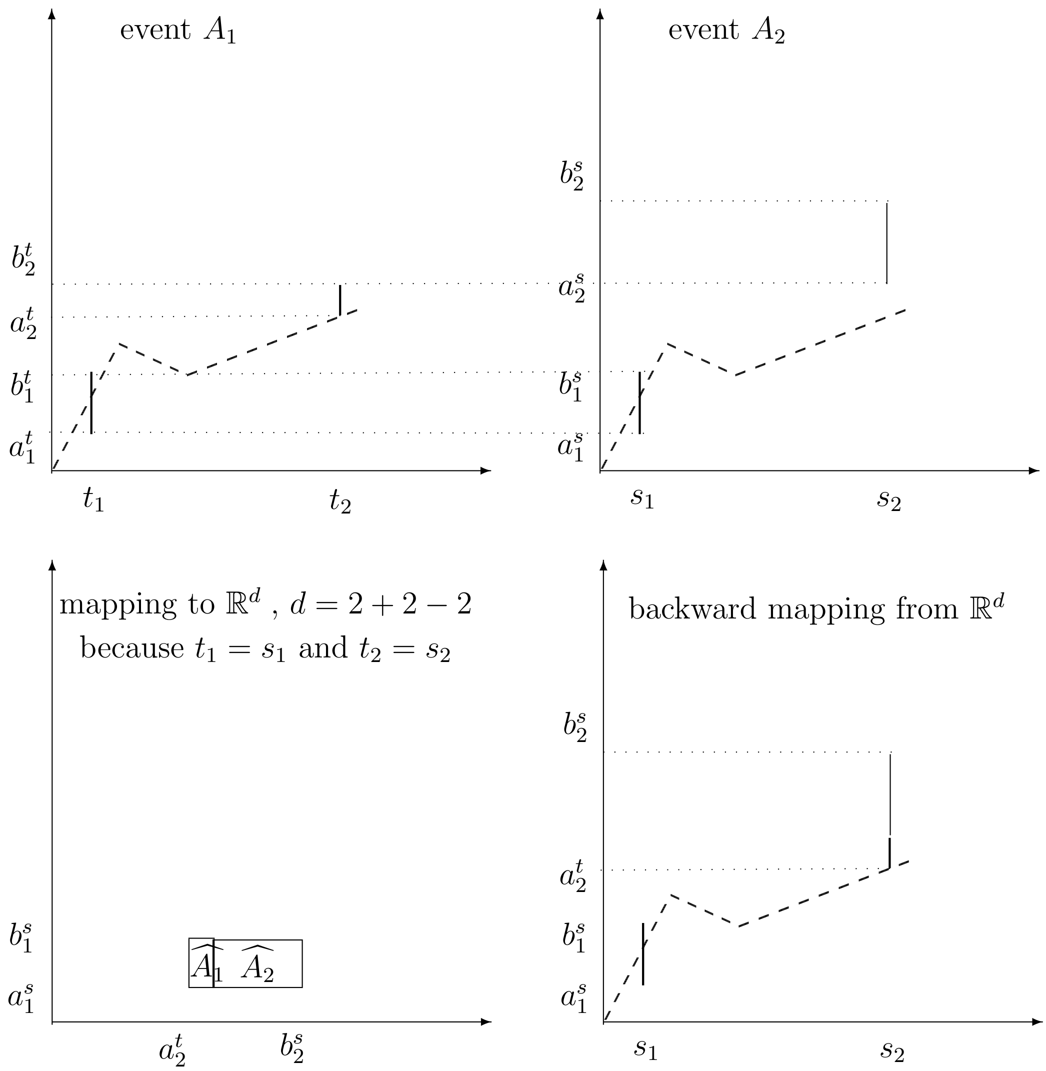


If it is, we map it back to the system of events ($\widehat{A} \to A$) and *define* A to be a union of events $A_1, A_2, \ldots A_p$. If it is not, we still say that if $\widehat{A} = \bigcup_{l=1}^{p} \widehat{A}_l$ then $A := \bigcup_{l=1}^{p} A_l$. But of course in this case A will not be an event of the special type we consider.

So we obtain a semiring of events (of special type) via a semiring of d-dimensional cells[13]! But what about a σ-additive measure on it?! We have it, the σ-additivity of $\mathbb{P}(\cdot) = \prod_{i=1}^{d} [\Phi^i(b_{t_i}) - \Phi^i(a_{t_i})]$ follows from the σ-additivity of $\Phi^i(\cdot)$ on $\mathbb{R}^1$ for all $i = 1, ..d$; so $\mathbb{P}(\cdot)$ is a product-premeasure of σ-additive premeasures.

Note that since the [probability] premeasure was actually defined on d-dimensional cells (although of course it makes interest to us in context of events). Thus we can be sure that "internal logic" of events remains consistent w.r.t. definition of union of events through the union of their images in $\mathbb{R}^d$, i.e. it will never happen that $\mathbb{P}(A)$ contradicts our practical experience.

[13]Recall how we defined for events "$B \subset A$", "$A \setminus B$" and check yourself that for the definition of union of events we just have given, (smR3) holds true.



Finally we have to consider the case of countably infinite number of events $A_1, A_2, \ldots$ and $A = \bigcup_{l=1}^{\infty} A_l$. Now we have to deal definitely with more that two events, so we must slightly change notation and denote an event A_l as

$$A_l := \{[a_1^{t(l)}, b_1^{t(l)}), [a_2^{t(l)}, b_2^{t(l)}), \ldots, [a_{n(l)}^{t(l)}, b_{n(l)}^{t(l)}) \; at \; t(l)_1, t(l)_2, \ldots, t(l)_{n(l)}\}$$

It turns out that this case can be straightforwadly reduced to the case of the finite union. Indeed, *if* A is in our semiring, it means that there is only a *finite* sequence of *unique* $t(l)_j$ (where $l = 1, 2, \ldots \; j = 1, .., n(l)$).
In words it means that only a finite number of events are related to the time-points of observation, to which no other event is related.

Were it not so, then $\widehat{A}$ would be an infinitely-dimensional cell (i.e. d would be infinite). But of course it cannot be in a semiring of finitely-dimensional cells.

So we have a semiring of events (of a special type) and a σ-additive [probability] premeasure on it. Thus by Carathéodory construction we can extend it to a σ-algebra of events of much more general type. Among them will be even such events, which we never observe in practice, but it is not a shortcoming of a mathematical model (as long as it conforms to the events which we do observe).

In particular, this σ-algebra of events *will* contain all *countably-infinite* union of the following events

$$A_l := \{[a_l^t, b_l^t) \ at \ t_l\}$$

And this time t_l *can* be unique for infinite number of events, because now we deal with a σ-algebra and not with a semiring.

But this means that we can model the behaviour of W_t in all rational timepoints. But the trajectories of W_t are almost surely continuous. This, in turn, means that we can treat all events, generated by W_t (upto events of zero probability measure).

So, as I hope, you have seen that the measure-theoretic stuff does bring probabilistic fruits. Now we are going to introduce the Lebesgue integral. Afterthat we will make some steps back to the measure theory, e.g. prove the existence of a Lebesgue-measurable set, which does not belong to the Borel σ-algebra.
The Lebesgue integral will be also a bridge between the measure theory and the functional analysis, as we consider the function spaces L^p ($1 \leq p \leq \infty$, the cases L^1, L^2 and maybe L^∞ are the most important).

But before we move to the next chapter I would like to make some short notes. As you remember we considered the cases where the underlying space $\mathbb{X}$ was $(0, 1)$ and $\mathbb{R}^1$ and, respectively, had finite and infinite measure. In case of $\mathbb{R}^1$ the measure was *sigma-finite*, i.e. $\mathbb{R}^1$ can be represented as *countable* union of the sets of finite measure.
Dealing with a probability measure $\mathbb{P}$ we always have that $\mathbb{P}(\mathbb{X}) = 1$ and

thus do not bother. But the "σ-finiteness" is important in real analysis.
Here is an example: we considered the [σ-finite and translation-invariant] Lebesgue measure in $\mathbb{R}^1$ and then in $\mathbb{R}^n$, n was finite. An *infinite* counterpart[14] of $\mathbb{R}^n$ is L^2 (the space of the Lebesgue square-integrable functions, which we are going to study later).
So if we have a Lebesgue measure in $\mathbb{R}^n$, why not to try to introduce it in L^2 ?! But it turns out that there is no σ-finite and translation-invariant measure in L^2 ! The proof is very concise and elegant but it is prematurely to discuss it now. So far I will just mention that instead there is a so-called *Gaussian measure* in L^2, which has some properties - nice enough to build an interesting theory w.r.t. the stochasic processes.

[14] The analogy with $\mathbb{R}^n$ is achieved through the idea of a scalar product.