

# Continuous Stochastic Logic Characterizes Bisimulation of Continuous-time Markov Processes

Josée Desharnais\*

Département d'Informatique  
Université Laval  
Québec, Canada, G1K 7P4

Prakash Panangaden\*

School of Computer Science  
McGill University  
Montréal, Canada, H3A 2A7

July 5, 2002

## Abstract

In a recent paper Baier, Haverkort, Hermanns and Katoen [BHHK00], analyzed a new way of model-checking formulas of a logic for continuous-time processes - called *Continuous Stochastic Logic* (henceforth **CSL**) – against continuous-time Markov chains – henceforth CTMCs. One of the important results of that paper was the proof that if two CTMCs were *bisimilar* then they would satisfy exactly the same formulas of **CSL**. This raises the converse question – does satisfaction of the same collection of **CSL** formulas imply bisimilarity? In other words, given two CTMCs which are known to satisfy exactly the same formulas of **CSL** does it have to be the case that they are bisimilar? We prove that the answer to the question just raised is “yes”. In fact we prove a significant extension, namely that a subset of **CSL** suffices *even for systems where the state-space may be a continuum*. Along the way we prove a result to the effect that the set of Zeno paths has measure zero provided that the transition rates are bounded.

---

\*Research supported in part by NSERC and MITACS.

# 1 Introduction

There has been a significant maturation of the field of verification recently. Apart from tremendous advances in the theory and practice of the subject there is a growing diversification of the field into areas that were hitherto regarded as unrelated to logic and verification. This trend is most marked in the interest in continuous time and in hybrid systems. Of course, real-time systems have been part of computer science for a while but the logic and theory community have only recently become interested in it, partly through very fruitful synergy with the performance evaluation community. The authors have been active in exploring the subject of continuous state spaces [DEP02, DGJP00] but have hitherto not ventured into the realm of continuous time. It is clear however that physical systems that have continuous state spaces will often have a continuous-time dynamics. Thus there is pressure to explore the combination of the two. The present paper is a step in that direction. The immediate motivation was a question posed by J.-P. Katoen to the second author.

One of the more interesting recent papers on the subject of model checking for continuous-time systems is by Baier, Haverkort, Hermanns and Katoen [BHHK00]. They analyzed a new way of model-checking formulas of a logic for continuous-time processes [ASSB96] — called *Continuous Stochastic Logic* (henceforth **CSL**) — against continuous-time Markov chains — henceforth CTMCs. One of the important results of that paper was the proof that if two CTMCs were *bisimilar* then they would satisfy exactly the same formulas of **CSL**. This raises the converse question - does satisfaction of the same collection of **CSL** formulas imply bisimilarity? In other words, given two CTMCs which are known to satisfy exactly the same formulas of **CSL** does it have to be the case that they are bisimilar?

We prove that the answer to the question just raised is “yes.” In fact we prove a significant extension, namely that a subset of **CSL** suffices *even for systems where the state-space may be a continuum*. This extension forces us to use the apparatus of measure theory that we had used in previous work with Edalat [DEP02]. However, the technically demanding aspects of the continuous state-space situation (using special properties of analytic spaces) has been isolated into one proposition. Once one accepts this proposition, the results are not any harder for the continuous case. Of course the original question was raised for CTMCs and this is an immediate consequence of our more general result. The basic intuition behind the proof is that **CSL** is rich enough to pin down the exact transition rates. So by looking at the right formulas we can force the systems to have matching transition rates.

In the course of our analysis we had to prove that the set of so-called “Zeno” paths has measure 0. This result has been proved for discrete systems [dA97, Bai] and it is possible that easy extensions of these arguments apply to the continuous case. This proof is interesting in its own right and uses an amusing observation about Fredholm operators. However, it is not being claimed that this is an important contribution of the present paper.

This paper can be read at two levels. We expect that many readers will not want to go to the literature and refresh themselves on the details of definitions like “analytic space” so we will provide some indication of what can be skipped while still following the main thread of the argument. For the convenience of the reader we provide a very brief summary of very basic definitions.

The next section, Section 2 gives basic background on CTMCs extended to continuous state systems. The following section discusses the space paths for a CTMC and proves that Zeno paths have measure zero. Section 4 following gives a summary of the logic and its semantics. In Section 5 we prove a general theorem that is useful for establishing completeness results of the type that we address in the present paper. Finally we prove the particular completeness theorem of this paper in Section 6. The material in Section 5 requires the most background and can be skipped at a first reading without significant loss of continuity if the reader is willing to believe the statement of the theorem stated there.

## 2 Background

In this section we generalize the CTMCs and the logic **CSL** studied by Baier et al. [BHHK00] to processes with a continuous state-space. For simplicity, we follow [BHHK00] and do not consider processes with labelled transitions. The results, however, are not affected — except in trivial notational ways — if we add labels. By adding labels to transitions, one easily obtains a generalization (to continuous state-space processes) of the processes studied by Hillston in [Hil94].

In the case of a continuous state space, one cannot simply specify transition probabilities from one state to another. Instead one must work with probability densities. In the present case, we have to introduce a notion of set for which “probabilities make sense” (i.e. a  $\sigma$ -field  $\Sigma$ ) and instead of talking about probabilities of going from a state  $s$  to another state  $s'$ , we have to talk about going from a state  $s$  to a *set of states*  $X \in \Sigma$ . The key mathematical constructions require an analytic space structure on the set

of states. Thus instead of imposing an arbitrary  $\sigma$ -field structure on the set of states, we will require that the set of states be an analytic space and the  $\sigma$ -field be the Borel algebra generated by the topology. Note that it is very hard to come up with an example of a system where the state space is not analytic. Thus, any discrete set, any open or closed subset of  $\mathbf{R}^n$ , any manifold, any “fractal” or anything one is likely to imagine is analytic.

Let  $AP$  be a fixed, finite set of atomic propositions.

**Definition 2.1** *A continuous-time Markov process or CTMP is a tuple  $(S, \Sigma, R, L)$  where  $(S, \Sigma)$  forms an analytic space,  $R : S \times \Sigma \rightarrow \mathbf{R}_{\geq 0}$  is a rate function: it is measurable in its first coordinate and a measure on its second coordinate; and  $L : S \rightarrow 2^{AP}$  is the labelling (measurable) function which assigns to each state  $s \in S$  the set  $L(s)$  of atomic propositions of  $AP$  that are valid in  $s$ . One may assume an initial distribution  $\alpha$  on the state-space.*

We write  $E(s)$  for  $R(s, S)$  (the exit rate out of  $s$ ). The set of absorbing states  $R^{-1}\{0\}$  is written  $Abs$  and  $Can := S \setminus Abs$  is the set of states that can make a transition. We also write  $P(s, X) = R(s, X)/E(s)$  if  $s \in Can$ : it is the probability of jumping to a state in  $X$ , given that we start in the state  $s$ . If  $s \in Abs$ ,  $P(s, X) = 0$ . Note that  $P(\cdot, X)$  is measurable since it is a quotient of measurable functions (see [Bil95] for example).

Let us explain more precisely the meaning of the rate function  $R$ . Given a state  $s$  and a (measurable) set of states say  $X$ , the transition rate — transition probability per unit time — for jumping from  $s$  to a state in  $X$  — is  $R(s, X)$ . If there are two disjoint sets of states that  $s$  can jump to we will have competing transitions and there will be a *race condition* between these transitions. If at time 0 the system is known to be in state  $s$  then at time  $t$  the probability that the transition to a state in  $X$  has been triggered is

$$1 - e^{-R(s, X) \cdot t}.$$

One cannot say that this is the probability that the system will be in a state of  $X$  at time  $t$ . The probability that at time  $t$  the system reaches a state in  $X$  in one transition is 0 if  $s \in Abs$ , and

$$\frac{R(s, X)}{E(s)} [1 - e^{-E(s) \cdot t}]$$

otherwise. As a last example, the probability of leaving  $s$  within the interval of time  $[t, u]$  is

$$e^{-E(s)t} (1 - e^{-E(s)(u-t)}) = e^{-E(s)t} - e^{-E(s)u},$$

because the process must stay in  $s$  for  $t$  units of time and then leave  $s$  within  $u - t$  units of time. Note that this yields 0 if  $s \in Abs$ .

### 3 Spaces of Paths in a CTMP

In order to understand the behaviour of a CTMP one needs to understand the computations. Thus the set of possible paths and the measures on them play a significant role. In the present section we analyse these spaces. One important aspect of continuous time systems is that each transition takes some time; the number of steps by itself is not the correct way of determining elapsed time. In such a case it is possible to have a so-called Zeno path, that is, a path with infinitely many steps but in which the elapsed time intervals decrease fast enough that the total elapsed time is finite. The presence of such paths greatly complicates any calculations. Often — in semantics — one rules out such paths by fiat, usually by invoking some kind of fairness property. In probabilistic treatments the information about transition probabilities is supposed to substitute for fairness so we cannot just rule out Zeno paths. What we show instead is that the probability of Zeno paths is zero *provided that the transition rates are bounded*.

It makes sense on physical grounds to insist that the transition rates be bounded. Clearly in finite state systems the transition rates will have some maximum value. In infinite state systems if we allow the rates to grow without bound then the probability of Zeno paths is no longer zero. An explicit example due to Christel Baier shows that this is the case. It is easy to reconstruct the example by considering a countable state space and making the transition rates increase exponentially rapidly. In such systems the transition rates grow above any prescribed limit. This is as unphysical as allowing speeds that exceed any limit! The expected time in a state is the inverse of the transition rate out of that state and clearly once one reaches times like  $10^{-33}$  seconds one has an unphysical situation. This is the time unit that can be constructed from the fundamental physical constants  $G, h, c$  and it is widely accepted that our usual notions of space and times break down. In the realm of macroscopic objects that usually concern computer science and engineering these time scales are absurd.

**Definition 3.1** *A path is a finite or infinite sequence  $s_0, t_0, s_1, t_1, \dots$  where  $t_i \in \mathbf{R}^+$  for  $i \in \mathbf{N}$  and  $s_i \in Can$  for all  $i \in \mathbf{N}$  except for the last state  $s_l$  if the sequence is finite in which case  $s_l \in Abs$ . Let  $\sigma$  be an infinite path; we*

use the following notation<sup>1</sup>:

$\sigma[i] = s_i$	the $i$ -th state of $\sigma$
$\delta(\sigma, i) = t_i$	the time spent in $s_i$
$\sigma@t = \sigma[i]$	the state of $\sigma$ at time $t$
	$i$ is the least index such that $t < \sum_0^i t_j$
$Path$	the set of all paths
$Path(s)$	the set of paths starting in $s$

If  $\sigma$  is finite and ends in  $s_l$ ,  $\sigma[i]$  and  $\delta(\sigma, i)$  are defined as above for  $i \leq l$ , whereas  $\delta(\sigma, i) = \infty$ , and  $\sigma@t = s_l$  for  $t \geq \sum_0^{l-1} t_j$ .

The Borel space of paths  $\mathcal{F}(Path)$  is generated by sets of paths of the form

- $X_0 \times I_0 \times X_1 \times I_1 \times \dots \times I_{n-1} \times X_n \times (\mathbf{R} \times S)^\infty$ ,
- $X_0 \times I_0 \times X_1 \times I_1 \times \dots \times X_{n-1} \times I_{n-1} \times Y$

with  $X_i \in \Sigma$ ,  $I_i$  an interval of the reals with rational bounds and  $n \in \mathbf{N}$  and  $Y \in (\Sigma \cap Abs)$ . It is not hard to prove that these sets form a semi-ring<sup>2</sup>. This semi-ring, that we denote by  $SR(Path)$ , is countable because the intervals involved have rational bounds. We often write the tail  $(\mathbf{R} \times S)^\infty$  simply as  $\infty$  to simplify the notation.

Given a distribution  $\alpha$  on the set of states, we define a probability measure  $Pr_\alpha$  on paths as follows. Let  $X_0 \times I_0 \times X_1 \times I_1 \times \dots \times I_{n-1} \times X_n \times \infty$  be a set (of infinite paths) in  $SPath$  and let us write  $e(i)$  for the probability  $e^{-E(x_i)t_i} - e^{-E(x_i)u_i}$  of leaving  $x_i$  within the interval of time  $I_i = [t_i, u_i]$ . Then

$$Pr_\alpha(X_0 \times I_0 \times X_1 \times I_1 \times \dots \times I_{n-1} \times X_n \times \infty) = \int_{X_0} e(0) \int_{X_1} e(1) \dots \int_{X_{n-1}} (e(n-1)P(x_{n-1}, X_n)) P(x_{n-2}, dx_{n-1}) \dots P(x_0, dx_1) \alpha(dx_0).$$

The measure is defined in the same way for a set in  $SR(Path)$  of finite type by replacing  $X_n$  with  $Y$ .

This set function is easily shown to be countably additive on the semi-ring  $SR(Path)$  and hence has a unique extension to a measure on  $\mathcal{F}(Path)$ .

<sup>1</sup>The definition of  $\sigma@t$  differs from [BHHK00] where the inequality is not strict. It is a technical point to make the temporal Until operator of CSL satisfiable.

<sup>2</sup>See the appendix for the definition.

We write  $Pr_s(X)$  if we consider the state  $s$  as a starting point and hence its initial distribution.

We allow absorbing states in infinite paths because the set of infinite sequences containing absorbing states is of measure 0. In general, it is not necessary to distinguish between the two cases of paths, i.e. those involving absorbing states and those that do not because the measure handles absorbing states. In manipulating sets of paths, we usually decompose them according to the number of transitions in the paths. This way, absorbing states are taken into account. On the other hand, if we want a set of paths of finite type to get a meaningful value (i.e.  $> 0$ ) we must use a multiple integral with the right number of integrations, that is, the number of integrations must be the number of transitions of the paths in the set.

The following theorem is the main point of this section. It is well known in the finite-state case.

**Theorem 3.2** *The set of paths having a sequence of time that converges is of measure zero provided that the rates are bounded; i.e.  $\sup_{s \in S} R(s, S)$  exists.*

**Proof .** The goal is to show that given  $T \in \mathbf{R}, s \in S$ ,

$$Pr_s(\sigma = s, t_0, s_1, t_1, \dots : \sum_{i=1}^N t_i < T) \xrightarrow{N \rightarrow \infty} 0.$$

First we show that the set

$$\{\sigma = s, t_0, s_1, t_1, \dots : \sum_{i=1}^N t_i < T\}$$

is measurable. It is easier to manipulate if we write this as

$$\{\sigma \in Paths(s) : \sum_{i=0}^N \delta(\sigma, i) < T\}.$$

We proceed by induction on  $N$ . The base case reduces to the set

$$\{s\} \times [0, T] \times S \times \infty$$

which is clearly measurable. For the inductive case we have

$$\begin{aligned} \{\sigma \in Paths(s) : \sum_{i=0}^{N+1} \delta(\sigma, i) < T\} = \\ \bigcup_{\alpha < T} \{\sigma : \sum_{i=0}^N \delta(\sigma, i) < T - \alpha \wedge \delta(\sigma, N+1) < \alpha\} \text{ where } \alpha \text{ is rational.} \end{aligned}$$

Each set in the above union is the intersection of two measurable sets (the first is measurable by the inductive hypothesis and the second by the definition of the  $\sigma$ -field on the path space). Since we have restricted  $\alpha$  to be rational we have a countable union of (not necessarily disjoint) measurable sets, hence the above set is measurable for all  $N$ .

Let  $g_N(s, T)$  be the measure of each of the sets of paths described above. Then  $g_0(s, T) = 1 - e^{-E(s)T}$  the probability of leaving  $s$  before time  $T$ . In order to derive an equation for  $g_N(s, T)$  we proceed as follows. By definition,  $g_N(s, T)$  is the probability that the system starts in state  $s$  and makes  $N$  transitions before time  $T$ . The conditional probability density that the system makes  $N$  transitions in time less than  $T$  *given that* it makes its first transition at time  $t$  is  $g_{N-1}(s_1, T - t)$ . Integrating this density with the measure  $R(s, ds_1)$  over the possible intermediate states gives

$$\int_S R(s, ds_1) g_{N-1}(s_1, T - t) dt$$

for the conditional probability that the system makes  $N$  transitions in time less than  $T$  *given that* it makes its first transition at time  $t$ . Now we integrate over all possible times  $t$  less than  $T$  to get

$$g_N(s, T) = \int_0^T e^{-E(s)t} \int_S R(s, ds_1) g_{N-1}(s_1, T - t) dt.$$

The intuition is that you stay in  $s$  for  $t$  units of time (with  $t \in [0, T]$ ) and then you jump to some state  $s_1 \in S$  from which you start a path with the property that the sum of the  $N - 1$  first time parameters is less than  $T - t$ . We must prove that this sequence converges to 0.

If we take the limit as  $N \rightarrow \infty$  we get the recursive equation

$$G(s, T) = \int_0^T e^{-E(s)t} \int_S R(s, ds_1) G(s_1, T - t) dt.$$

In order to get the above we have to move the limit on the right hand side inside the integral. To justify this we note that the integrand is dominated by  $\rho \int g_n(s_1, T - t)$  where  $\rho = \sup_{s \in S} R(s, S)$ . Thus we can use the dominated convergence theorem [Rud66] to commute the limit and the integration. We know that  $G(s, 0) = 0$  for all  $s$  trivially from the form of the outer integral. If we differentiate with respect to  $T$  (taking care to take into account that  $T$  occurs as an upper limit of an integral as well as a parameter explicitly inside the integral) we get

$$\frac{\partial G}{\partial T} = e^{-E(s)T} \int_S R(s, ds_1) G(s_1, 0) dt + \int_0^T e^{-E(s)t} \int_S R(s, ds_1) \frac{\partial G}{\partial T}(s_1, T - t) dt.$$

For short we write  $G'$  for  $\frac{\partial G}{\partial T}$ . The first term is zero because  $G(s_1, 0)$  vanishes. This gives

$$G'(s, T) = \int_0^T e^{-E(s)t} \int_S R(s, ds_1) G'(s_1, T - t) dt.$$

The explicit form of the exponentials was necessary for this to turn out this way. In short we see that  $G'$  satisfies the same recursive equation as  $G$  and has the same boundary condition at  $T = 0$ . Inductively we see right away that  $G^{(n)}$ , the  $n$ th derivative of  $G$  with respect to  $T$ , also satisfies the same equation and hence that all the derivatives of  $G$  vanish at  $T = 0$ . If we had a differential equation for  $G$  we would be done because the uniqueness theorem would say that the trivial solution obtained by inspection, i.e.  $G \equiv 0$  is the only solution.

We have an integral equation for  $G$  instead of a differential equation. However, the integral equation is of a very special type, it is called an integral equation of Fredholm type, see for example chapter 3 of Courant and Hilbert's text "Methods of Mathematical Physics" [CH53]. The requirement for a Fredholm operator is that the integral involve continuous functions of  $t$ , this is manifestly the case here (of course  $R$  may be far from continuous as a function of  $s$  but that is irrelevant). The main theorem in this subject is that there are only finitely many linearly independent solutions to a Fredholm equation. Thus, since all the  $G^{(n)}$  satisfy the same equation there can only be finitely many linearly independent solutions. Thus for  $k$  large enough  $G^{(k)}$  must be given by a linear combination of lower derivatives of  $G$ . In short there has to be a linear differential equation for  $G$ . This completes the proof. ■

## 4 The Logic CSL

In this section, we recall the logic introduced in [BHHK00] and its semantics. The version of **CSL** [ASSB96] originally introduced by Aziz et. al. does not have the next state formula, it has the Until construct which allows one to express most path modalities. For our purposes the next state formula is very useful and is very standard in such logics. The version of Baier et. al. is essentially like ours except that they also talk about rewards. In [BHHK00], the logic **CSL** was augmented with a very important formula  $S_{\text{exp}}(\phi)$  to represent steady-state properties. We do not need it for the purposes of the present paper so we omit it. It is not clear to us that it is well defined in the continuous-state space case but we did not pursue that very actively

since it is not needed for the completeness proof of the present paper. It would be interesting to see if this operator could be defined properly for continuous-state systems.

**Definition 4.1** *Our version of the logic **CSL** [ASSB96] has the following syntax. Let  $a \in AP$ ,  $p \in [0, 1] \cap \mathbb{Q}$  and  $\bowtie \in \{<, \leq, \geq, >\}$ . State formulas  $\phi$  are defined by*

$$\phi := \top \mid a \mid \neg\phi \mid \phi \wedge \phi' \mid P_{\bowtie p}(\psi)$$

where  $\psi$  is a path formula constructed by

$$\psi := X\phi \mid X^{[t,u]}\phi \mid \phi U \phi' \mid \phi U^{[t,u]}\phi';$$

for  $t, u$  rational.

Note that there are countably many formulas since  $AP$  is countable and  $p, t, u$  are rationals.

**Meaning of formulas** Given a CTMP  $\mathcal{S} = (S, \Sigma, R, L)$  and  $a \in AP$ , the definition of the satisfaction relation  $\models$  over state-formulas is given by induction:

$$\begin{aligned} s \models \top & \quad \text{for all } s \in S; \\ s \models a & \quad \text{iff } a \in L(s); \\ s \models \phi \wedge \phi' & \quad \text{iff } s \models \phi \text{ and } s \models \phi'; \\ s \models \neg\phi & \quad \text{iff } s \not\models \phi; \\ s \models P_{\bowtie p}\psi & \quad \text{iff } Pr_s\{\sigma \in Path(s) : \sigma \models \psi\} \bowtie p. \end{aligned}$$

Note that the set  $\{\sigma \in Path(s) : \sigma \models \psi\}$  is measurable; this will be shown in the next lemma below. To be more formal, we could have considered at this point the greatest measurable set contained in it. The semantics of path-formulas is defined as follows.

$$\begin{aligned} \sigma \models X\phi & \quad \text{iff } \sigma[1] \text{ is defined and } \sigma[1] \models \phi; \\ \sigma \models X^{[t,u]}\phi & \quad \text{iff } \sigma \models X\phi \text{ and } \delta(\sigma, 0) = t_0 \in [t, u]; \\ \sigma \models \phi U \phi' & \quad \text{iff } \exists k \geq 0. \sigma[k] \models \phi' \text{ and } \forall 0 \leq i < k, \sigma[i] \models \phi \\ \sigma \models \phi U^{[t,u]}\phi' & \quad \text{iff } \exists t^* \in [t, u]. \sigma @ t^* \models \phi' \text{ and } \forall 0 \leq t' < t^*, \sigma @ t' \models \phi. \end{aligned}$$

We write  $\llbracket \phi \rrbracket_{\mathcal{S}}$  for the set  $\{s \in S \mid s \models \phi\}$ . We often omit the subscript when no confusion can arise. Similarly, we write  $\llbracket \psi \rrbracket_s = \{\sigma \in Path(s) : \sigma \models \psi\}$ . The following lemma shows that these sets are measurable in  $\mathcal{S}$ .

Note that the logic **CSL** could be extended to witness *labelled* transitions, by replacing  $X$  with  $\langle a \rangle$ . In this case, the formula  $\langle a \rangle_p \phi$  from [DEP02] would be represented in **CSL** by the formula  $P_{>p}(\langle a \rangle \phi)$ .

**Lemma 4.2** *Let  $S = (S, \Sigma, R, L)$  be a CTMP. Then*

1. *for every formula  $\phi$  of CSL,  $\llbracket \phi \rrbracket \in \Sigma$ ;*
2. *for every path formula  $\psi$  of CSL and every  $s \in S$ ,  $\llbracket \psi \rrbracket_s \in \mathcal{F}(\text{Path})$ .*

**Proof .** We prove the two statements in parallel using structural induction. Trivially  $\llbracket \top \rrbracket = S \in \Sigma$ . For every  $a \in AP$ ,  $\llbracket a \rrbracket \in \Sigma$  because  $L$  is measurable. Conjunction and negation are trivial since finite intersections and complements of measurable sets are measurable. To prove that  $\llbracket P_{\bowtie p}(\psi) \rrbracket$  is measurable for every path-formula  $\psi$ , we will prove that the function  $Pr_{(\cdot)}(\llbracket \psi \rrbracket_{(\cdot)}) : S \rightarrow [0, 1]$  is measurable.

Now assume  $\llbracket \phi \rrbracket \in \Sigma$ . Consider the path formula  $X^{[t,u]}\phi$  and fix  $s \in S$ . We have

$$\llbracket X^{[t,u]}\phi \rrbracket_s = \{s, t_0, s_1, \dots : t_0 \in [t, u], s_1 \models \phi\} = \{s\} \times [t, u] \times \llbracket \phi \rrbracket \times \infty$$

and hence  $\llbracket X^{[t,u]}\phi \rrbracket_s \in \mathcal{F}(\text{Path})$ . Now  $Pr_{(\cdot)}(\llbracket X^{[t,u]}\phi \rrbracket_{(\cdot)}) : S \rightarrow [0, 1]$  satisfies

$$Pr_s(\llbracket X^{[t,u]}\phi \rrbracket_s) = (e^{-E(s)t} - e^{-E(s)u}) \frac{R(s, \llbracket \phi \rrbracket)}{E(s)}.$$

Now products, differences and compositions of measurable functions are measurable by standard results in measure theory [Ash72, Bil95, Rud66, Hal74]. Thus our function above is measurable in  $s$  since it is constructed as a combination (product, difference, quotient, composition) of the measurable functions  $E(s)$  and  $R(s, \llbracket \phi \rrbracket)$ . Consequently,  $\{s : Pr_s(\llbracket X^{[t,u]}\phi \rrbracket_s) \bowtie p\} = \llbracket P_{\bowtie p}(X^{[t,u]}\phi) \rrbracket \in \Sigma$  because it is the inverse image under a measurable function of a measurable subset of  $[0, 1]$ , namely the subset defined by  $\bowtie p$ .

We now prove that the path formula  $\phi U^{[t,u]}\phi'$  describes a measurable set of paths. In the following we write out the set of paths satisfying the Until formula as a union, indexed by the number of transitions it makes before satisfying  $\phi'$ . Each such union is itself written as the union over the possible *rational* times at which the transitions occur. It suffices to use intervals with rational end points since every possible transition time will be represented by some interval. Note that we are not claiming that we have a disjoint union. We specify in each line the time at which the paths satisfy the second condition of the Until formula and use the notation  $\sigma = s, t_0, s_1, t_1, \dots$ . Note that if  $s$  does not satisfy  $\phi$  then it cannot satisfy

the Until formula so we assume that  $s \models \phi$  in the following.

$$\begin{aligned}
& \llbracket \phi U^{[t,u]} \phi' \rrbracket_s \\
&= \cup_i \{ \sigma \in \text{Path}(s) : \exists T \in [t, u] \sigma @ t = s_i \models \phi' \text{ and } \forall t' < T, \sigma @ t' \models \phi \} \\
&= (\{s\} \cap \llbracket \phi' \rrbracket) \times (t, \infty) \times \infty \quad s \models \phi' \text{ and } T \text{ in } [t, \min\{t_0, u\}) \\
&\cup \{s\} \times [t, u] \times \llbracket \phi' \rrbracket \times \infty \quad s_1 \models \phi' \text{ and } T = t_0 \\
&\cup \bigcup_{u_1 < u_2 < t} \{s\} \times [u_1, u_2] \times \llbracket \phi \wedge \phi' \rrbracket \times (t - u_1, \infty) \times \infty \\
&\quad \quad \quad s_1 \models \phi \wedge \phi' \text{ and } T \in [t, \min\{t_0 + t_1, u\}) \\
&\cup \bigcup_{\substack{u_1 < u_2 < u \\ u_2 - u_1 < u - t}} \{s\} \times [u_1, u_2] \times \llbracket \phi \rrbracket \times [t - u_1, u - u_2] \times \llbracket \phi' \rrbracket \times \infty \\
&\quad \quad \quad s_2 \models \phi' \text{ and } T = t_0 + t_1 \\
&\cup \bigcup_{\substack{u_1 < u_2 \\ u_3 < u_4 \\ u_2 + u_4 < t}} \{s\} \times [u_1, u_2] \times \llbracket \phi \rrbracket \times [u_3, u_4] \times \llbracket \phi \wedge \phi' \rrbracket \times (t - u_1 - u_3, \infty) \times \infty \\
&\quad \quad \quad s_2 \models \phi \wedge \phi' \text{ and } T \in [t, \min\{t_0 + t_1 + t_2, u\}). \\
&\quad \quad \quad \vdots
\end{aligned}$$

In the expression following the last equality sign above, the first line is for the case where  $s$  already satisfies  $\phi'$  and the system stays in this state until at least  $t$ . In the second line we have the situation where at time  $t_0$  the system jumps to  $s_1$  which satisfies  $\phi'$ . In the third line we have the case where the jump to  $s_1$  occurs between  $u_1$  and  $u_2$  but before  $t$ . In this case we must have  $s_1$  satisfying both  $\phi$  and  $\phi'$  and staying in this state until time  $t$ . The subsequent lines follow the same pattern.

Since the set in question is expressed as a countable union of measurable sets, it is measurable.

Now we have to show that  $Pr_s(\llbracket \phi U^{[t,u]} \phi' \rrbracket_s)$  viewed as a function of  $s$  is a measurable function. We have expressed the set  $\llbracket \phi U^{[t,u]} \phi' \rrbracket_s$  as a countable but not disjoint union. Let us consider what the intersections of two of these sets can look like. If the number of jumps before  $\phi'$  is satisfied is different in the two sets then the intersection will be empty. In other cases, if we intersect two sets from the union occurring on the same line, we will be intersecting two sets where the time intervals overlap. In this case the intersection will be in the form of a basic measurable set of paths. If we apply the  $Pr_s(\cdot)$  to such a set we clearly get a measurable function of  $s$ . Thus when we form finite intersections we get measurable functions of  $s$ . Now when we have a finite union - say of the family  $U_1, U_2, \dots, U_n$  - we can

write the probability as<sup>3</sup>

$$\sum_{i=1}^n Pr_s(U_i) - \sum_{i \neq j} Pr_s(U_i \cap U_j) + \dots$$

But each such term defines a measurable function of  $s$  so the combination, being constructed from sums and products is also measurable. When we have a countable union we can construct the function  $Pr_s(\cdot)$  as the sup of the functions for the finite unions. Since the sup of a family of measurable functions is a measurable function we are done. ■

## 5 A General Technique for Relating Bisimulation and Logic

In a previous paper [DEP02], we have defined the notion of probabilistic bisimulation and logics that were compatible with this notion. In various unpublished work we have done other closely related proofs with variations of the logic and the notion of bisimulation. The equivalence proofs have all used similar techniques. In this section, we formulate - in a more general and useful form - the theorems that were used implicitly in all of our proofs. In particular, we use it for the proof of the present paper. This general result should be reusable for several situations.

Before we give this general form we need some notation. Let  $\mathcal{F}$  be a family of measurable sets, i.e.  $\mathcal{F} \subseteq \Sigma$ . There is an induced equivalence relation, written  $\equiv_{\mathcal{F}}$ , defined by

$$x \equiv_{\mathcal{F}} y \leftrightarrow \forall A \in \mathcal{F}. (x \in A \leftrightarrow y \in A).$$

In other words two states are equivalent if they belong to exactly the same sets of  $\mathcal{F}$ . We write  $cl(\mathcal{F})$  for the collection of measurable sets closed under the equivalence relation  $\equiv_{\mathcal{F}}$  (i.e. measurable unions of equivalence classes). The general theorem can now be stated.

**Theorem 5.1** *Let  $(S, \Sigma)$  be an analytic space. Let  $\mathcal{F} \subseteq \Sigma$  be countable and closed under intersection and let  $S \in \mathcal{F}$ . Then if two measures agree on  $\mathcal{F}$  then they agree on  $cl(\mathcal{F})$ .*

It is worth summarizing how to use this result. Typically, the set  $\mathcal{F}$  will contain the meaning of basic formulas and the measures will be the

---

<sup>3</sup>This purely combinatorial fact is called the “principle of inclusion-exclusion”.

transition probabilities from two equivalent states. Consequently, to prove that logical equivalence implies bisimilarity, we only have to prove that two logically equivalent states have the same value on transitions to the set of states that satisfy some formula - the definable sets of the logic. If the logic is indeed rich enough to characterize bisimulation then the transition probabilities to the definable sets will force the transition probabilities to agree on all sets and thus be bisimilar. We see that we need our logic to have conjunction and to have only countably many formulas - both are very mild restrictions - and to be rich enough to encode the transition probabilities (or rates) to definable sets. If the logic has these basic properties then a completeness proof can now be produced routinely. Our main Theorem 6.3 is precisely such an application of this result.

One major step towards proving Theorem 5.1 is given by a classical result, the  $\lambda$ - $\pi$  theorem. This result gives a condition under which two measures are equal. It is a standard textbook result, for example, it appears as Theorem 10.4 of [Bil95].

**Theorem 5.2** [*The  $\lambda$ - $\pi$  theorem*] *Let  $X$  be a set and  $\mathcal{A}$  a family of subsets of  $X$ , closed under finite intersections, and such that  $X$  is a countable union of sets in  $\mathcal{A}$ . Let  $\sigma(\mathcal{A})$  be the  $\sigma$ -field generated by  $\mathcal{A}$ . Suppose that  $\mu_1, \mu_2$  are finite measures on  $\sigma(\mathcal{A})$ . If they agree on  $\mathcal{A}$  then they agree on  $\sigma(\mathcal{A})$ .*

One can see that the two theorems are very similar and that the only result that we need to prove in order to get Theorem 5.1 from the  $\lambda$ - $\pi$  theorem is that  $\sigma(\mathcal{F}) = cl(\mathcal{F})$ . This is exactly the statement of Lemma 5.5. However, this lemma requires some nontrivial results. The proof of Lemma 5.5 hinges on an important theorem about analytic spaces and one of its corollaries - see, for example, Theorem 3.3.5 of [Arv76]. These are

**Theorem 5.3** [*Unique structure theorem*] *Let  $(S, \Sigma)$  be an analytic space and let  $\Sigma_0$  be a countably generated sub- $\sigma$ -field of  $\Sigma$  that separates points in  $S$ . Then  $\Sigma_0 = \Sigma$ .*

We say that a  $\sigma$ -field separates points if every pair of points is separated by a set in the  $\sigma$ -field, that is, there is some set in the  $\sigma$ -field that contains only one of these points.

**Theorem 5.4** *Let  $(S, \Sigma)$  be an analytic space and let  $\sim$  be an equivalence relation on  $S$ . Assume that there is a sequence  $f_1, f_2, \dots, f_n, \dots$  of measurable real-valued functions on  $S$  such that for any pair of points  $x, y$  in  $S$  one has  $x \sim y$  if and only if  $\forall n. f_n(x) = f_n(y)$ . Then  $S/\sim$  is also an analytic space and the trivial quotient map  $q: S \rightarrow S/\sim$  is measurable.*

Both Theorem 5.3 and Theorem 5.4 rely on properties of analytic spaces. This is the main reason why analytic spaces appear in the subject. Normally one works with Polish spaces but they do not have the key quotient property that we needed for these proofs.

Now with these two properties of analytic spaces in hand we can prove the following key lemma.

**Lemma 5.5** *Let  $(S, \Sigma)$  be an analytic space. Let  $\mathcal{F} \subseteq \Sigma$  be countable and assume  $S \in \mathcal{F}$ . Then  $cl(\mathcal{F}) = \sigma(\mathcal{F})$ .*

**Proof .** It is easy to prove that  $\sigma(\mathcal{F}) \subseteq cl(\mathcal{F})$  by showing that  $cl(\mathcal{F})$  is a  $\sigma$ -field containing  $\mathcal{F}$ . To prove equality, note that the equivalence relation  $\equiv_{\mathcal{F}}$  can be defined by saying that  $x$  and  $y$  are equivalent if the characteristic functions of the sets in  $\mathcal{F}$  all agree on  $x$  and  $y$ . Since  $\mathcal{F}$  is countable and the characteristic functions are measurable and  $(S, \Sigma)$  is analytic, by Theorem 5.4, the quotient  $S/\mathcal{F}$  is analytic and the trivial quotient map  $q : S \rightarrow S/\mathcal{F}$  is measurable. Recall that the Borel  $\sigma$ -field of  $S/\mathcal{F}$  is given by  $\Sigma/\mathcal{F} = \{A \subseteq S/\mathcal{F} : q^{-1}(A) \in \Sigma\}$ .

We now prove that  $q(\sigma(\mathcal{F})) = q(cl(\mathcal{F}))$ . Note that  $q^{-1}q\sigma = \sigma$  if  $\sigma \in cl(\mathcal{F})$ . Thus

$$q(\sigma(\mathcal{F})) \subseteq q(cl(\mathcal{F})) \subseteq \Sigma/\mathcal{F}.$$

Since  $q(\mathcal{F})$  separates points, then  $q(\sigma(\mathcal{F}))$  also does. We now show that  $q(\sigma(\mathcal{F}))$  is a  $\sigma$ -field. Since  $\sigma(\mathcal{F})$  is one, we have that  $q(\sigma(\mathcal{F}))$  contains  $S/\mathcal{F}$  and is closed under countable unions. Moreover, since  $\sigma(\mathcal{F}) = q^{-1}q\sigma(\mathcal{F})$ , we have that two states that belong to different sets of  $\sigma(\mathcal{F})$  cannot have the same image under  $q$ . This implies that the complement of the image (under  $q$ ) of a set is the image of its complement. This establishes that  $q(\sigma(\mathcal{F}))$  is a  $\sigma$ -field. Consequently, we can apply Theorem 5.3 and conclude that the inclusions above are in fact equalities. Hence the images of  $\sigma(\mathcal{F})$  and  $cl(\mathcal{F})$  agree.

Finally, we have to show that the sets  $\sigma(\mathcal{F})$  and  $cl(\mathcal{F})$  also agree. Observe that if  $q(A) = q(B)$  for  $A \in \sigma(\mathcal{F})$  and  $B \in cl(\mathcal{F})$  then  $A \subseteq B$ . To see this, let  $a \in A$  and  $q(a) = x$ . Then there is a  $b \in B$  such that  $q(b) = x$  which means that  $a \equiv_{\mathcal{F}} b$  so  $a \in B$  since  $B$  is closed under  $\equiv_{\mathcal{F}}$ . Similarly, if  $b \in B$  there there is an  $a \in A$  with  $q(a) = q(b)$  and hence  $a \equiv_{\mathcal{F}} b$ . But since  $\sigma(\mathcal{F}) \subseteq cl(\mathcal{F})$ ,  $A \in cl(\mathcal{F})$  thus  $b \in A$  and  $B \subseteq A$ , i.e.  $A = B$ . Thus  $\sigma(\mathcal{F}) = cl(\mathcal{F})$ . ■

This lemma establishes Theorem 5.1 immediately.

## 6 Bisimilarity and CSL

In this section, we introduce the definition of bisimulation for CTMPs and prove that it coincides with the equivalence induced by the logic. We prove this result by applying Theorem 5.1 from the preceding section to the model of this paper.

**Notation** Let  $F$  be a set of formulas of **CSL**; then  $L_F(s)$  is the set of formulas of  $F$  that are satisfied by  $s$ . Let  $\equiv$  be an equivalence relation on  $S$ ; then  $cl(\equiv)$  contains all the closed sets w.r.t.  $\equiv$ , that is,

$$cl(\equiv) = \{\sigma \in \Sigma : \text{if } s \in \sigma \text{ and } s \equiv s' \text{ then } s' \in \sigma\}.$$

The following definition of bisimulation generalizes the definition of  $F$ -bisimulation for discrete CTMC's introduced in [BHHK00].

**Definition 6.1** *Let  $F$  be a set of formulas. An equivalence relation  $\equiv$  is an  $F$ -bisimulation if whenever  $s \equiv s'$ , we have  $L_F(s) = L_F(s')$  and for every  $C \in cl(\equiv)$ ,  $R(s, C) = R(s', C)$ .*

$F$  is intended to be the set of observable formulas. If we take  $F = AP$  then  $F$ -bisimulation is standard bisimulation. If we take  $F = \{\top\}$  and hence ignore atomic propositions, we get the analogue of probabilistic bisimulation as defined by Larsen and Skou for discrete probabilistic processes (with only one label on the transitions). This parametrization allows us a more flexible treatment of bisimulation.

Obviously, satisfying only formulas in  $F$  would be a far too weak condition for two states to be bisimilar: for example, if  $F = \{\top\}$ . Bisimulation involves matching transitions and it is necessary that the set of formulas that defines the equivalence between states be closed under the constructor  $P_{\bowtie p}(X^{[0,t]}(\cdot))$ . One must keep in mind that  $F$  is just a restriction on bisimulation that we can impose with the help of atomic propositions. The bigger  $F$  is, the fewer states are going to be bisimilar.

**Definition 6.2** *If  $F$  is a set of **CSL** formulas then the closure of  $F$  under conjunction  $\wedge$  and the operator  $P_{\bowtie p}(X^{[0,t]}(\cdot))$  is written  $\overline{F}$ .*

**Theorem 6.3** *Let  $F$  be a set of formulas that contains the trivial formula  $\top$ . If two states of a CTMP satisfy the same formulas of  $\overline{F}$  then they are  $F$ -bisimilar.*

**Proof .** Let  $F$  be a set of formulas; then  $\mathcal{F} = \{\llbracket \phi \rrbracket : \phi \in \overline{F}\}$  is countable and closed under intersection. We write  $s \equiv s'$  if  $s$  and  $s'$  satisfy the same formulas of  $\overline{F}$ . We show that  $\equiv$  is an  $F$ -bisimulation. Let  $s \equiv u$ ; then  $L_F(s) = L_F(u)$  trivially. We want to prove that  $R(s, C) = R(u, C)$  for every  $C \in cl(\equiv)$ . By Theorem 5.1, we only have to prove that it is true for  $C \in \mathcal{F}$ . We first prove that  $E(s) = E(u)$ . Since  $s \equiv u$ ,  $s$  and  $u$  satisfy the same formulas of the form  $P_{\triangleright\triangleright p}(X^{[0,t]}\phi)$  where  $\phi$  is a state formula constructed from formulas in  $\overline{F}$ ; consequently, we have

$$Pr_s(\{\sigma \in Path_s : \sigma \models X^{[0,t]}\phi\}) = Pr_u(\{\sigma \in Path_u : \sigma \models X^{[0,t]}\phi\}). \quad (1)$$

Consider the case where  $\phi = \top \in \overline{F}$ . Then

$$1 - e^{-E(s)t} = 1 - e^{-E(u)t}$$

which implies that  $E(s) = E(u)$ .

We now prove that  $R(s, \llbracket \phi \rrbracket) = R(u, \llbracket \phi \rrbracket)$  for every formula  $\phi \in F$ . We get from Equation 1 that

$$Pr_s(\{\sigma \in Path_s : \sigma \models X^{[0,t]}\phi\}) = (1 - e^{-E(s)t}) \frac{R(s, \llbracket \phi \rrbracket)}{E(s)}.$$

Then

$$(1 - e^{-E(s)t}) \frac{R(s, \llbracket \phi \rrbracket)}{E(s)} = (1 - e^{-E(u)t}) \frac{R(u, \llbracket \phi \rrbracket)}{E(u)}$$

which implies that  $R(s, \llbracket \phi \rrbracket) = R(u, \llbracket \phi \rrbracket)$ .

By Theorem 5.1, we have that  $R(s, A) = R(u, A)$  for every  $A \in cl(\equiv)$  as wanted. ■

Not all the operators of the logic are needed in the preceding proof. In particular, there is no use of constant, negation and no use of Until. We have observed this phenomenon before [DEP02], namely that one does not need very many formulas to get a logic rich enough to get a characterization of bisimulation. Of course the proof is rendered harder by using fewer formulas. Of particular interest is the fact that no negation - or even any kind of negative construct - is needed for our proof.

Let  $\mathbf{CSL}_F$  denote the smallest set of formulas of  $\mathbf{CSL}$  containing  $F$  and closed under  $\mathbf{CSL}$  operators. The following theorem has been proven for CTMCs in [BHHK00].

**Theorem 6.4** *If two states are  $F$ -bisimilar, then they satisfy the same formulas of  $\mathbf{CSL}_F$ .*

**Proof .** The proof is an easy induction on formulas. The strategy is to show that if  $\equiv$  is an  $F$ -bisimulation, then  $\llbracket \phi \rrbracket$  is in  $cl(\equiv)$ , for every state-formula  $\phi$ . If  $s \equiv s'$  and  $\llbracket \phi \rrbracket \in cl(\equiv)$  by induction, then definition of bisimulation implies that  $R(s, \llbracket \phi \rrbracket) = R(s', \llbracket \phi \rrbracket)$  and also that  $R(s, S) = R(s', S)$  since  $S$  is certainly in  $cl(\equiv)$ . Then we use the equations developed in Lemma 4.2 for  $\llbracket X^{[t,u]} \phi \rrbracket_s$  and  $\llbracket \phi U^{[t,u]} \phi' \rrbracket_s$ . Measurable functions representing the probabilities to these sets are in terms of  $R(s, \llbracket \phi \rrbracket)$  and  $R(s, S)$ ; manipulations of measurable functions and sets complete the proof. ■

## 7 Conclusions

Our contributions in this paper are:

- a completeness proof that shows that a subset of **CSL** gives a logical characterization of bisimulation for CTMPs;
- a general strategy for such proofs, in fact the arguments that we went through using the unique structure theorem and other aspects of analytic spaces are now modularized so that it should be only a combinatorial question to apply this method to other situations (of course, the combinatorics could be tricky);

In previous work we have developed a theory of approximation for continuous state systems [DGJP00]. What this theory provides is a family of *finite-state* approximants for every continuous-state Markov process. These approximants converge - in a suitable metric - to the original process. We also showed that labelled Markov processes could be given the structure of a continuous domain. Then the approximants form a directed set with a supremum that gives back the original process. Most importantly we were able to show that any formula - of a certain weak modal logic - that is satisfied by the process being approximated is satisfied by one of the finite-state approximants. Though the logic is weak it is strong enough to characterize bisimulation.

It would be very appealing to develop such results for continuous-time systems. If we had a good theory of finite approximation we could imagine using model checking techniques developed for finite state systems for physical or hybrid systems. The logical characterization result of the present paper is a reasonable starting point. Our metrics for discrete time processes [DGJP99] were inspired by the earlier logical characterization results,

thus it seems reasonable that we could develop such metrical notions for continuous time and follow that up with an approximation theory.

## Acknowledgments

We are very grateful to J.-P. Katoen for bringing this question to our attention and to Radha Jagadeesan and Holger Hermans for useful discussions. We thank Christel Baier for finding - and Nilima Nigam for fixing - a flaw in our earlier proof of the theorem about Zeno paths. We are grateful to the referees for their helpful comments and thorough refereeing job. Finally we thank MITACS and NSERC for support.

## References

- [Arv76] W. Arveson. *An Invitation to  $C^*$ -Algebra*. Springer-Verlag, 1976.
- [Ash72] R. B. Ash. *Real Analysis and Probability*. Academic Press, 1972.
- [ASSB96] Adnan Aziz, Kumud Sanwal, Vigyan Singhal, and Robert K. Brayton. Verifying continuous time markov chains. In *Proc. of Conference on Computer-Aided Verification, 1996*.
- [Bai] C. Baier. Zeno paths occur with probability zero. Private communication.
- [BHHK00] C. Baier, B. Haverkort, H. Hermans, and J.-P. Katoen. Model checking continuous-time Markov chains by transient analysis. In *CAV 2000: Proceedings of the 12th Annual Symposium on Computer-Aided Verification*, number 1855 in Lecture Notes In Computer Science, pages 358–372. Springer-Verlag, 2000.
- [Bil95] P. Billingsley. *Probability and Measure*. Wiley-Interscience, 1995.
- [CH53] R. Courant and D. Hilbert. *Methods of Mathematical Physics I*. J. Wiley, 1953.
- [dA97] L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997. Technical Report STAN-CS-TR-98-1601.
- [DEP02] J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labeled Markov processes. *Information and Computation*, 2002.

- [DGJP99] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labeled Markov systems. In *Proceedings of CONCUR99*, Lecture Notes in Computer Science. Springer-Verlag, 1999.
- [DGJP00] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Approximation of labeled markov processes. In *Proceedings of the Fifteenth Annual IEEE Symposium On Logic In Computer Science*, pages 95–106. IEEE Computer Society Press, June 2000.
- [Dud89] R. M. Dudley. *Real Analysis and Probability*. Wadsworth and Brookes/Cole, 1989.
- [Hal74] P. Halmos. *Measure Theory*. Number 18 in Graduate Texts in Mathematics. Springer-Verlag, 1974. Originally published in 1950.
- [Hil94] J. Hillston. *A Compositional Approach to Performance Modelling*. PhD thesis, University of Edinburgh, 1994. Published as a Distinguished Dissertation by Cambridge University Press in 1996.
- [KT66] J. F. C. Kingman and S. J. Taylor. *Introduction to Measure and Probability*. Cambridge University Press, 1966.
- [Rud66] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1966.

## A A Summary of Measure Theory

For completeness, we give the relevant definitions from measure theory in this section. We assume that the reader knows the basic ideas of measure theory and probability as expounded in, for example “Probability and Measure” by Billingsley [Bil95] or “Real Analysis and Probability” by Ash [Ash72] or the book with the same title by Dudley [Dud89] or “Introduction to Measure and Probability” by Kingman and Taylor [KT66].

**Definition A.1** *A  $\sigma$ -field on a set  $X$  is a family of subsets of  $X$  which includes  $X$  itself and which is closed under complementation and countable unions.*

A set equipped with a  $\sigma$ -field is called a *measurable space*. Given a topological space  $(X, \mathcal{T})$ , we can define the  $\sigma$ -field, often written  $\mathcal{B}$ , generated by the open sets (or, equivalently, by the closed sets). This is usually called the *Borel algebra*.

One is often interested in  $\sigma(\mathcal{A})$ , the  $\sigma$ -field generated by a family of sets  $\mathcal{A}$ . It is easy to prove that the intersection of any number of  $\sigma$ -fields is again a  $\sigma$ -field. Then one can define the  $\sigma$ -field generated by  $\mathcal{A}$  to be the intersection of all  $\sigma$ -fields that contain  $\mathcal{A}$ . This collection is not empty since the powerset of a set always forms a  $\sigma$ -field and this will certainly contain  $\mathcal{A}$ .

**Definition A.2** *Given a  $\sigma$ -field  $(X, \Sigma)$ , a **subprobability measure** on  $X$  is a  $[0, 1]$ -valued set function,  $\mu$ , defined on  $\Sigma$  such that*

- $\mu(\emptyset) = 0$ ,
- for a pairwise disjoint, countable collection of sets,  $\{A_i | i \in I\}$ , in  $\Sigma$ , we require

$$\mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i).$$

*In addition, for probability measures we require  $\mu(X) = 1$ .*

One often wants to define measures on  $\sigma$ -fields by defining the measure on a more restricted family that generates the  $\sigma$ -field. Then one can hope to extend the measure to the whole  $\sigma$ -field. Unfortunately one cannot always do this and even if one can the resulting measure may not be canonical. The usual way is to take a generating set with nice algebraic properties.

**Definition A.3** *A family  $\mathcal{F}$  of subsets of  $X$  is called a **semi-ring** if*

1.  $\emptyset \in \mathcal{F}$ ,
2.  $A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$  and,
3. if  $A \subseteq B$  are in  $\mathcal{F}$  then there are finitely many pairwise disjoint subsets  $C_1, \dots, C_k \in \mathcal{F}$  such that  $B - A = \bigcup_{i=1}^k C_i$ .

This is not the form of the definition that one is used to in algebra because of the strange last condition but this is precisely the property that holds for “hyperrectangles” in  $\mathbf{R}^n$ .

**Theorem A.4** *Suppose that  $\mathcal{F}$  is a semi-ring on  $X$  and  $\mu : \mathcal{F} \rightarrow [0, \infty]$  satisfies*

1.  $\mu(\emptyset) = 0$ ,
2.  $\mu$  is finitely additive and
3.  $\mu$  is countably subadditive.

Then  $\mu$  extends uniquely to a measure on the  $\sigma$ -field generated by  $\mathcal{F}$ .

Countably subadditive means that  $\mu(\cup_n A_n) \leq \sum_n \mu(A_n)$ . The proof of this theorem may be found in a standard text on probability and measure, for example the book by Kingman and Taylor [KT66] or the one by Billingsley [Bil95] or the one by Ash [Ash72]. The intervals on the reals form a typical example of a semi-ring and the length function satisfies the conditions of the theorem so this extension theorem gives another construction of Lebesgue measure.

**Definition A.5** A function  $f : (X, \Sigma_X) \rightarrow (Y, \Sigma_Y)$  between measurable spaces is said to be **measurable** if  $\forall B \in \Sigma_Y. f^{-1}(B) \in \Sigma_X$ .

In older books like Halmos [Hal74] and Rudin [Rud66] a different definition of measurable function is used. That definition is somewhat more general than the one we are using (which is the definition one sees in modern texts) but has the bad feature that it does not compose; i.e. the composite of two measurable functions need not be measurable according to the Halmos definition.

The next several definitions and results pertain to analytic spaces.

**Definition A.6** A **Polish** space is the topological space underlying a complete, separable metric space; i.e. it has a countable dense subset.

**Definition A.7** An **analytic** space is the image of a Polish space under a continuous function from one Polish space to another.

The following proposition [Dud89] gives equivalent definitions of analytic set.

**Proposition A.8** Suppose that  $X$  and  $Y$  are Polish spaces and  $f$  is a function from  $X$  to  $Y$ . The following are equivalent:

- $f$  is continuous and  $A$  is the image of  $X$  under  $f$ ,
- $f$  is measurable and  $A$  is the image of  $X$  under  $f$ ,
- $f$  is continuous and  $A$  is the image of a Borel subset  $B$  of  $X$ ,

- *$f$  is measurable and  $A$  is the image of a Borel subset  $B$  of  $X$ ,*
- *$g : \mathbf{N}^\infty \rightarrow Y$  is continuous and  $A$  is the image of  $\mathbf{N}^\infty$  and*
- *$g : \mathbf{N}^\infty \rightarrow Y$  is measurable and  $A$  is the image of  $\mathbf{N}^\infty$ .*

Thus in this definition it turns out to be equivalent to say “measurable” image and it makes no difference if we take the image of the whole Polish space or of a Borel subset of the Polish space.

Analytic spaces are more general than Polish spaces in that they need not be complete. Any discrete space is analytic. Any space that one is likely to meet in physical applications is Polish. When one takes the quotient of a Polish space by “reasonable” equivalence relations one may not get a Polish space but one will get an analytic space. It is the construction of these quotients that force us to work with analytic spaces.