Adaptive dynamical systems and the EPR–chameleon experiment

Luigi Accardi[1], Kentaro Imafuku[2], Massimo Regoli[3]

Centro Vito Volterra, Università degli Studi di Roma "Tor Vergata"
WEB-page: http://volterra.mat.uniroma2.it,

Abstract

We prove that the structure of the classical dynamical system, constructed in [AcImRe01] to reproduce the EPR correlations, is a natural consequence of the combination of the general theory of classical dynamical systems with the extension of von Neumann measurement theory, proposed in [Ac93] to include in it the requirements of locality and causality. We then prove some a priori estimates which guarantee the convergence of the simulation procedure.

The program to run the experiment is available from the WEB-page: http://volterra.mat.uniroma2.it.

1 Introduction

In the paper [AcImRe01] we constructed a family of classical deterministic (macroscopic) dynamical systems which, performing purely local and independent choices, reproduces the EPR correlations hence violates the Bell inequality. In the same paper an experimental implementation of this theoretical construction using three separated and independent computers was described.

As evident from Theorem (1) below and its proof, no arbitrary ingredients or artificial selection procedures are inserted by hands in our model: everything is pre–determined in the usual sense of classical deterministic dynamical systems.

The above mentioned results give an experimental and theoretical confirmation of the thesis, first advocated by quantum probability, that there is absolutely no contradiction between a realistic view of the world, locality and the predictions of quantum theory.

It also explains, by means of a concrete and easily understandable example, that this apparent contradiction has its roots in a too naive view of the notion of classical dynamical system. This naive view overlooks the existence of **adaptive dynamical systems**, i.e. systems whose dynamical evolution depends on the environment (in the case of measurements – on the observable to be measured). The chameleon metaphora gives a nice intuitive picture of such systems.

In section (2) of the present paper we recall the statement and the proof of the main result in [AcImRe01]. The proof presented here corrects some notational misprints in that paper.

In section (3) we prove how the structure of the mathematical model described in section (2) naturally follows from the combination of the standard theory of classical dynamical systems with the local causal extension of von Neumann's measurement theory proposed in [Ac93]. This combination can be considered as a natural formulation of the theory of **adaptive dynamical systems**.

Finally in section (8) we describe the estimates that guarantee the validity of the simulation procedure.

[1]accardi@volterra.mat.uniroma2.it
[2]imafuku@volterra.mat.uniroma2.it
[3]massimo@volterra.mat.uniroma2.it

2 A local, deterministic, reversible, classical dynamical system realizing the EPR correlation

In this section we recall the definition of the EPR–chameleon dynamical system constructed in [AcImRe01] and we prove that it reproduces the EPR correlations. The notions of dynamical system, local dynamics, ... will be used here and will be introduced formally in the following sections.

Definition 1 *Let be given:*
– a measurable (state) space

$$(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$$

– for each $a, b \in [0, 2\pi]$ two ± 1–valued maps (observables)

$$S_a^{(1)} : \omega_1 \in \Omega_1 \to S_a^{(1)}(\omega_1) \in \{+1, -1\}$$

$$S_b^{(2)} : \omega_2 \in \Omega_2 \to S_b^{(2)}(\omega_2) \in \{+1, -1\}$$

– for each $a, b \in [0, 2\pi]$ two maps (local dynamics)

$$T_{1,a} : \Omega_1 \to \Omega_1 \quad ; \quad T_{2,b} : \Omega_2 \to \Omega_2$$

The local, deterministic, reversible, classical, adaptive (cf section (6)) dynamical system

$$\{(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, P) \ , \ T_{1,a} \otimes T_{2,b}\}$$

is said to **realize the EPR correlations** *if*

$$\langle S_a^{(1)} S_b^{(2)} \rangle := \int \int_{\Omega_1 \times \Omega_2} S_a^{(1)}(\omega_1) S_b^{(2)}(\omega_2) dP \circ T^{-1}(\omega_1, \omega_2) = -\cos(b - a)$$

We will consider a classical deterministic local dynamical system composed by two spatially separated sub–systems $(1, M_1)$ and $(2, M_2)$ interpreted as follows:
– 1 and 2 are particles
– M_1 and M_2 are measurement apparata
– the state space of $(1, M_1)$ is $[0, 2\pi] \times \mathbb{R}$ and also the state space of $(2, M_2)$ is $[0, 2\pi] \times \mathbb{R}$
– therefore the **state space** of the whole system $(1, M_1, 2, M_2)$ is $[0, 2\pi]^2 \times \mathbb{R}^2$

The **local adaptive dynamics** $T_{1,a}, T_{2,b}$ where $1, 2$ are labels for particles and a, b are labels for apparata, are defined by

$$T_{1,a}, T_{2,b} \ : \ [0, 2\pi] \times \mathbb{R} \to [0, 2\pi] \times \mathbb{R} \tag{1}$$

$$T_{1,a}(\sigma_1, \lambda_1) := (s_{1,a}(\sigma_1, \lambda_1), m_{1,a}(\sigma_1, \lambda_1))$$

$$T_{2,b}(\sigma_2, \lambda_2) := (s_{2,b}(\sigma_2, \lambda_2), m_{2,b}(\sigma_2, \lambda_2))$$

$$s_{1a}(\sigma_1, \lambda_1) := \sigma_1, \quad ; \quad m_{1,a}(\sigma_1, \lambda_1) := \lambda_1 \frac{1}{T'_{1,a}(\sigma_1)} \tag{2}$$

$$s_{2,b}(\sigma_2, \lambda_2) := \sigma_2 \quad ; \quad m_{2,b}(\sigma_2, \lambda_2) := \lambda_2 \frac{1}{T'_{2,b}(\sigma_2)} \tag{3}$$

where the functions

$$T'_{1,a}, T'_{2,b} \ : \ [0, 2\pi] \to [0, 2\pi]$$

are defined by

$$T'_{1,a}(\sigma_1) = \frac{\sqrt{2\pi}}{4} |\cos(\sigma_1 - a)| \qquad , \qquad T'_{2,b}(\sigma_2) = \sqrt{2\pi} \qquad , \ \sigma_1, \sigma_1 \in [0, 2\pi] \tag{4}$$

In other words our adaptive dynamics (1) are given, for $\sigma_1, \sigma_1 \in [0, 2\pi]; \lambda_1, \lambda_2 \in \mathbb{R}$, by:

$$T_{1,a}(\sigma_1, \lambda_1) := (\sigma_1 \ , \ \frac{\lambda_1 4}{\sqrt{2\pi}} |\cos(\sigma_1 - a)|) \quad , \quad T_{2,b}(\sigma_2, \lambda_2) := (\sigma_2 \ , \ \frac{\lambda_2}{\sqrt{2\pi}})$$

For $(\sigma_1, \sigma_2 \in [0, 2\pi] \ , \ \lambda_1, \lambda_2 \in \mathbb{R})$ we define the local measures

$$p_S(d\sigma_1, d\sigma_2) = \frac{1}{2\pi} \delta(\sigma_1 - \sigma_2) d\sigma_1 d\sigma_2$$

$$p_{1,a}(\sigma_1, d\lambda_1) = \delta(m_{1,a}(\sigma_1, \lambda_1) - m_a) d\lambda_1$$

$$p_{2,b}(\sigma_2, d\lambda_2) = \delta(m_{2,b}(\sigma_2, \lambda_2) - m_b) d\lambda_2$$

where m_a, m_b are arbitrary real numbers.

Remark (1) $p_S(d\sigma_1, d\sigma_2)$ is the initial preparation of the composite system $(1, 2)$ and, by the causality principle, it cannot depend on the setting of the apparatus.

In fact at time $t = 0$ the particles cannot know which will be the setting of the apparatus at time $t = 1$ (the first time of interaction with it).

$p_{1,a}(\sigma_1, d\lambda_1)$ and $p_{2,b}(\sigma_2, d\lambda_2)$ are the initial preparations of the local apparata. They are typical "response–type" preparations and must be interpreted in the adaptive sense, i.e.:

if, at time $t = 1$, the particle will arrive to me in state $\sigma_x(= \sigma_1, \sigma_2)$, then I will set myself in the state $p_{1,x}(\sigma_x, d\lambda_x)$ $(x = 1, 2 \ )$ (cf. sections (4), (5) where it is described how the theory of adaptive dynamical systems naturally leads to measures of this type. As explained in section (6), the above scheme is the discrete idealization of a continuous time process of adaptation.

Combining the above local measures we define the measure (**initial distribution**) on $[0, 2\pi]^2 \times \mathbb{R}^2$:

$$p_S(\sigma_1, \sigma_2) p_{1,a}(\sigma_1, \lambda_1) p_{2,b}(\sigma_2, \lambda_2) d\sigma_1 \sigma_2 d\lambda_1 d\lambda_2 \tag{5}$$

Finally, for a, b as above, define the ± 1–valued maps (**observables**)

$$S_a^{(1)}, S_b^{(2)} \ : \ [0, 2\pi] \times \mathbb{R} \to \{\pm 1\} \tag{6}$$

by

$$S_a^{(1)}(\sigma, \mu) = S_a^{(1)}(\sigma) = \text{sgn}(\cos(\sigma - a)) \quad , \quad S_b^{(2)} = -S_b^{(1)} \qquad ; \qquad \sigma \in [0, 2\pi], \mu \in \mathbb{R} \tag{7}$$

Theorem 1 *In the above notations, the measure (5) is a probability measure on* $[0, 2\pi]^2 \times \mathbb{R}^2$. *Moreover*

$$\int S_a^{(1)} (s_{1,a}(\sigma_1, \lambda_1), m_{1,a}(\sigma_1, \lambda_1)) S_b^{(2)} (s_{2,b}(\sigma_2, \lambda_2), m_{2,b}(\sigma_2, \lambda_2)) \tag{8}$$

$$\cdot p_S(\sigma_1, \sigma_2) p_{1,a}(\sigma_1, \lambda_1) p_{2,b}(\sigma_2, \lambda_2) d\sigma_1 d\sigma_2 d\lambda_1 d\lambda_2 = -\cos(a - b)$$

Proof. The positivity is obvious. The normalization condition

$$\int p_S(\sigma_1,\sigma_2)p_{1,a}(\sigma_1,\lambda_1)p_{2,b}(\sigma_2,\lambda_2)d\sigma_1\sigma_2 d\lambda_1 d\lambda_2 = 1 \tag{9}$$

follows from (8) by choosing $a = b$. To prove (8) notice that, with the choices (7), (5), the correlations (8) become

$$\int\int\int\int S_a^{(1)}(s_{1,a}(\sigma_1,\lambda_1), m_{1,a}(\sigma_1,\lambda_1))S_b^{(2)}(s_{2,b}(\sigma_2,\lambda_2), m_{2,b}(\sigma_2,\lambda_2)) \tag{10}$$

$$\delta(m_{1,a}(\sigma_1,\lambda_1) - m_a)\delta(m_{2,b}(\sigma_2,\lambda_2) - m_b)p_S(\sigma_1,\sigma_2)d\lambda_1 d\lambda_2 d\sigma_1 d\sigma_2$$

Changing variables

$$m_{1,a}(\sigma_1,\lambda_1) = \mu_1 \quad ; \quad m_{2,b}(\sigma_2,\lambda_2) = \mu_2$$

$$m'_{1,a}(\sigma_1,\lambda_1)d\lambda_1 = d\mu_1 \quad ; \quad m'_{2,b}(\sigma_2,\lambda_2)d\lambda_2 = d\mu_2$$

and noting that given a, $b \in [0, 2\pi]$ for almost all $a, b, \sigma_1, \sigma_2 \in [0, 2\pi]$ the functions

$$m_{1,a}(\sigma_1,\cdot),\ m_{2,b}(\sigma_2,\cdot) : \mathbb{R} \to \mathbb{R} \tag{11}$$

given by (2), (3), (4), are invertible, one finds

$$d\lambda_1 = \frac{1}{m'_{1,a}(\sigma_1, m_{1,a}^{-1}(\sigma_1,\mu_1))} d\mu_1 =: T'_{1,a}(\sigma_1,\mu_1)d\mu_1 \tag{12}$$

$$d\lambda_2 = \frac{1}{m'_{2,b}(\sigma_2, m_{2,b}^{-1}(\sigma_2,\mu_2))} d\mu_2 =: T'_{2,b}(\sigma_2,\mu_2)d\mu_2 \tag{13}$$

and, after the change of variables, (10) becomes

$$\int\int\int\int S_a^{(1)}(s_{1,a}(\sigma_1, m_{1,a}^{-1}(\sigma_1,\mu_1)),\mu_1)S_b^{(2)}(s_{2,b}(\sigma_2, m_{2,b}^{-1}(\sigma_2,\mu_2)),\mu_2)$$

$$T'_{1,a}(\sigma_1,\mu_1)T'_{2,b}(\sigma_2,\mu_2)\delta(\mu_1 - m_a)\delta(\mu_2 - m_b)p_S(\sigma_1,\sigma_2)d\mu_1 d\mu_2 d\sigma_1 d\sigma_2 \tag{14}$$

Because of our choice (2), (3), (4) of the functions $s_{1,a}$, $s_{2,b}$ and of the choice (7) of $S_a^{(1)}$, $S_b^{(2)}$, these have the form

$$S_a^{(1)}(\sigma_1) := S_a^{(1)}(s_{1,a}(\sigma_1, m_{1,a}^{-1}(\sigma_1, m_a),\mu_1)$$

$$S_b^{(2)}(\sigma_2) := S_b^{(2)}(s_{2,b}(\sigma_2, m_{2,b}^{-1}(\sigma_2, m_b),\mu_2) \tag{15}$$

in the sense that the right hand side depends only on the variables written on the left hand side. Therefore (14) becomes

$$\int\int S_a^{(1)}(\sigma_1)S_b^{(2)}(\sigma_2)T'_{1,a}(\sigma_1)T'_{1,b}(\sigma_2)p_S(\sigma_1,\sigma_2)d\sigma_1 d\sigma_2 \tag{16}$$

Finally, since

$$p_S(\sigma_1,\sigma_2) = \frac{1}{2\pi}\delta(\sigma_1 - \sigma_2) \tag{17}$$

using (4) we arrive at

$$\int S_a^{(1)}(\sigma)S_b^{(2)}(\sigma)T'_{1,a}(\sigma)T'_{1,b}(\sigma)\frac{d\sigma}{2\pi} = -\frac{1}{4}\int_0^{2\pi} \cos(\sigma - a)\text{sgn}(\cos(\sigma - b))d\sigma = -\cos(b - a)$$

which is the thesis.

3 (Passive) Dynamical systems

The scheme of a usual (passive, deterministic, conservative) dynamical system is the following:

– at time 0 the system is in a well defined state (i.e. it has a well defined probability distribution which, in the totally deterministic case, may be a measure concentrated on a single atom).

– the system evolves with a dynamics which depends on **its own initial state and not on the initial state of the measurement apparatus**

– the initial state of the measurement apparatus **does not depend on any state of the system**

– each observable of the system has a definite value at any time (dynamics)

– at time 1 the experimentalist reads, i.e. measures, the value of an observable of the system

The mathematical model of such dynamical systems is defined as follows.

Definition 2 *A classical (discrete time) deterministic dynamical system is a quadruple:*

$$\{(\Omega, \mathcal{F}) \ , \ \mathcal{O} \ , \ P \ , \ T\}$$

where:

– $(\Omega, \mathcal{F})$ *is a measurable space (the state space)*

– $\mathcal{O}$ *is a set of measurable maps from* $(\Omega, \mathcal{F})$ *to* $\mathbb{R}$ *(the observables)*

– P *is a probability measure (the preparation of the experiment) called the initial distribution of the system.*

– $T : \Omega \to \Omega$ *is an* $\mathcal{F}$*–measurable map (called the dynamics)*

The system is called reversible if T^{-1} *exists and is measurable (*P*–a.e. invertibility is sufficient).*

For such a system the probability that at time 1 the state of the system is in a subset $A \subseteq S$ is given by

$$Pr\{x \in S : Tx \in A\} = Pr\{x \in S : x \in T^{-1}A\} = P(T^{-1}A) =: P \circ T^{-1}(A)$$

Equivalently we can write

$$P \circ T^{-1}(\chi_A) = P(\chi_{T^{-1}A}) = P(\chi_A \circ T)$$

By taking linear combinations and limits, we arrive to the identity

$$P \circ T^{-1}(f) = P(f \circ T)$$

for any measurable function $f : S \to \mathbb{C}$. More explicitly, if $\langle \cdot \rangle$ denotes the expectation value at time 1, then

$$\langle f \rangle = \int_S f(Tx) dP(x) = \int_S f(y) dP \circ T^{-1}(y) \tag{18}$$

4 Dynamical measurements

Definition 3 *A dynamical measurement is a dynamical system of the form*

$$\{(S \times M, P_1), T\}$$

where S is called the system space and M the apparatus space.

In such a system the dynamics is uniquely determined by 2 maps

$$s : (\sigma, \lambda) \in S \times M \to s(\sigma, \lambda) \in S$$

$$m : (\sigma, \lambda) \in S \times M \to m(\sigma, \lambda) \in M$$

through the identity

$$T(\sigma, \lambda) = (s(\sigma, \lambda), m(\sigma, \lambda)) \tag{19}$$

The expectation value of any observable

$$f : S \times M \to \mathbb{R}$$

of the composite system can be evaluated using (18) and (19). One finds

$$\langle f \rangle = \int_S \int_M f(\sigma, \lambda) d(P_1 \circ T^{-1})(\sigma, \lambda) = \int_S \int_M f(T(\sigma, \lambda)) dP_1(\sigma, \lambda) =$$

$$= \int_S \int_M f(s(\sigma, \lambda), m(\sigma, \lambda)) dP_1(\sigma, \lambda)$$

In the following we will assume that both S and M are subsets of the real line $\mathbb{R}$ and we will write all measures as densities with respect to the Lebesgue measure (distribution densities being allowed). If f is chosen of the form

$$f = F \otimes G$$

where $F : S \to \mathbb{R}$ and $G : M \to \mathbb{R}$ are observables of the system and the apparatus respectively, then

$$\langle f \rangle = \int_S \int_M F(\sigma) G(\lambda) d(P_1 \circ T^{-1})(\sigma, \lambda) = \int_S \int_M F(s(\sigma, \lambda)) G(m(\sigma, \lambda)) dP_1(\sigma, \lambda) =$$

$$= \int_S \int_M F(s(\sigma, \lambda)) G(m(\sigma, \lambda)) p(\sigma, \lambda) d\sigma d\lambda$$

In general we want to measure F (an observable of the system) and to this goal we must always associate to it an observable $G = G_F$ of the measurement apparatus. We call this observable the form factor of the apparatus. With no loss of generality we can restrict our attention to those form factors of the apparatus, i.e. G, which satisfy the conditions

$$G(\lambda) \geq 0$$

$$\int_S \int_M G(\lambda) d(P \circ T^{-1})(\sigma, \lambda) = 1$$

i.e. that G is a $(P\circ T^{-1})$–probability density. Equivalently:

$$G(m(\sigma,\lambda))p(\sigma,\lambda)\geq 0$$

$$\int_S\int_M G(m(\sigma,\lambda))p(\sigma,\lambda))d\sigma d\lambda = 1$$

In real life measurements only the values of F are considered so that the values of G are averaged over. This means that the quantity $G(m(\sigma,\lambda))p(\sigma,\lambda))d\sigma d\lambda$ behaves like a single object: an effective probability measure. For example, if for P–almost all $\sigma\in S$ the map

$$m_\sigma:\lambda\in M\to m_\sigma(\lambda):=m(\sigma,\lambda)\in M$$

is invertible, then we can write

$$\langle F\otimes G\rangle = \langle F\rangle_G = \int_S\int_M F(s(\sigma,\lambda))p_G(\sigma,m(\sigma,\lambda),\lambda)d\lambda$$

where the effective probability measure p_G is given by

$$p_G(\sigma,m(\sigma,\lambda),\lambda)d\lambda := G(m(\sigma,\lambda))p(\sigma,m_\sigma^{-1}m(\sigma,\lambda))d\lambda$$

5 Independent preparations

Suppose that the initial preparations of the system and of the apparatus are independent:

$$dP(\sigma,\lambda) = dP_S(\sigma)dP_M(\lambda)$$

or equivalently, for the density:

$$p(\sigma,\lambda) = p_S(\sigma)p_M(\lambda)$$

In this case, and in the above notations,

$$G(m(\sigma,\lambda))p(\sigma,\lambda) = G(m(\sigma,\lambda))p_S(\sigma)p_M(\sigma,\lambda)$$

As explained in the previous section, the details of the function G are very seldom known explicitly and what one can experimentally control is the combined action of G and p_M, i.e. one introduces the effective probability density

$$\hat{p}_M(\sigma,m(\sigma,\lambda),\lambda) := G(m(\sigma,\lambda))p_M(\sigma,\lambda)$$

With these notations the expectation value $\langle F\rangle$ becomes

$$\langle F\rangle = \int_S\int_M F(s(\sigma,\lambda))p_S(\sigma)\hat{p}_M(\sigma,m(\sigma,\lambda),\lambda)d\sigma d\lambda \tag{20}$$

In other terms, the effective measure $\hat{p}_M(\sigma,m(\sigma,\lambda),\lambda)d\sigma$ includes the information on the evolution of the apparatus observable. A large class of natural examples of apparatus observables G, satisfying the conditions listed above can be constructed as follows.

It is clear that, in any real measurement, the apparatus is prepared in such a way that the system will not interact with all its possible (micro–)states, but only with a well defined subset of them. This means that one fixes a set $A \subseteq M$ of states of the apparatus (say – macrostates) and the measurement is conditioned on this set.

For example the set A can be specified by the space location of the apparatus, its location and extension, its orientation, a definite range of energies,...

Therefore, if F is any observable of the system S, its experimentally obtained mean value will be of the form

$$\langle F \rangle = \int_S \int_M F(\sigma) \frac{\chi_A(\lambda)}{P_1(S \times A)} \, d(P_1 \circ T^{-1})(\sigma, \lambda) \tag{21}$$

Using the general formula (18) and (19) we can write this expectation value in the form

$$\langle F \rangle = \int_S \int_M F(s(\sigma, \lambda)) \frac{\chi_A(m(\sigma, \lambda))}{P_1(S \times A)} \, dP_1(\sigma, \lambda) \tag{22}$$

The detailed description of the set $A \subseteq M$ is in general impossible so what one does in practice is to consider globally the measure

$$\frac{\chi_A(m(\sigma, \lambda))}{P_1(S \times A)} \, dP_1(\sigma, \lambda) \tag{23}$$

as an effective probability measure

$$dP(\sigma, m(\sigma, \lambda), \lambda) \tag{24}$$

and to model (phenomenologically) this effective measure.

In terms of the effective measure (24) the expectation value (22) becomes

$$\langle F \rangle = \int_S \int_M F(s(\sigma, \lambda)) dP(\sigma, m(\sigma, \lambda), \lambda) \tag{25}$$

and for simplicity we will further restrict our attention to the case in which $P(\sigma, m(\sigma, \lambda), \lambda)$ does not explicitly depend on λ. In this case the expectation value (22) becomes

$$\langle F \rangle = \int_S \int_M F(s(\sigma, \lambda)) dP(\sigma, m(\sigma, \lambda)) \tag{26}$$

Notice the difference with the expectation value (18) for usual (passive) dynamical systems, which in this case would look like

$$\langle F_1 \rangle = \int_S \int_M F_1(s(\sigma, \lambda)), m(\sigma, \lambda)) dP(\sigma, \lambda) \tag{27}$$

This difference is only apparent because, as explained in the above discussion, (26) is obtained from (27) with the particular choice $F_1 = F \otimes G$.

If the preparations of the system and of the apparatus are independent:

$$dP_1(\sigma, \lambda) = p(\sigma, \lambda) d\sigma d\lambda = dP_S(\sigma) dP_M(\lambda) = p_S(\sigma) p_M(\lambda) d\sigma d\lambda$$

then the effective probability measure (24) has the form

$$dP(\sigma, m(\sigma, \lambda), \lambda) = \frac{\chi_A(m(\sigma, \lambda))}{P_1(S \times A)} \, dP_1(\sigma, \lambda) =$$

$$= \frac{\chi_A(m(\sigma,\lambda))}{P_M(A)}\, dP_S(\sigma)dP_M(\lambda) = dP_S(\sigma)\hat{P}_M(m(\sigma,\lambda),d\lambda)$$

where we have introduced the notation

$$\hat{P}_M(m(\sigma,\lambda),d\lambda) := \frac{\chi_A(m(\sigma,\lambda))}{P_M(A)}\, dP_M(\lambda) = p_S(\sigma)d\sigma \frac{\chi_A(m(\sigma,\lambda))}{P_M(A)}\, p_M(\lambda)d\lambda$$

The local, causal measure constructed in [AcIR01] has precisely this form.

Notice that in general the positive measures $\hat{P}_M(m(\sigma,\lambda),d\lambda)$ will not be normalized, i.e. it will not be true that for all $\sigma \in S$, up to a set of P_S–measure zero, one has

$$\int_M \hat{P}_M(m(\sigma,\lambda),d\lambda) = 1$$

The normalization condition will hold in general only for the global measure:

$$\int_S dP_S(\sigma) \int_M \hat{P}_M(m(\sigma,\lambda),d\lambda) = 1$$

For such a measure the expectation value (26) becomes

$$\langle F\rangle = \int_S \int_M F(s(\sigma,\lambda))dP_S(\sigma)\hat{P}_M(m(\sigma,\lambda),d\lambda) \tag{28}$$

and, in terms of densities

$$\langle F\rangle = \int_S \int_M F(s(\sigma,\lambda))p_S(\sigma)d\sigma\hat{p}_M(m(\sigma,\lambda),\lambda)d\lambda \tag{29}$$

The fact that the normalization condition does not hold for $\hat{P}_M$ is precisely the new mathematical ingredient which has allowed to overcome the multiplicity of no–go theorems which have proliferated in the past 40 years, starting from the Bell inequality. In fact it has been proved in the appendix of the paper [AcImRe01] that, if the positive measures $\hat{P}_M(m(\sigma,\ \cdot\ )$ are normalized, then it is impossible to violate Bell's inequality.

6 Adaptive deterministic dynamical systems

The above considerations suggest to study the properties of the probability measures which define expectations values of the form (26). In order to carry out this investigation, we have to make more precise what is meant by a classical, adaptive, local, causal dynamical system.

Adaptive deterministic dynamical systems describe the following typical situation:

– a system is in a box at time 0

– the system evolves with a dynamics which depends on:

(i) its own initial state at time 0

(ii) the state of the measurement apparatus at the first time the system interacts with the measurement apparatus, say, time 1

– the state at time 1 of the measurement apparatus depends on the state of the system at time 1

– each observable of the system has a definite value at any time (dynamics)

– the experimental measure takes place at time 1

The von Neumann measurement scheme starts at time 1. Therefore for this scheme:

– the initial state of the system is at time 0

– the initial state of the measurement apparatus is at time 1 hence it depends on the initial state of the system at time 1

This is an idealization of the following physical situation:

– at time 0 the initial state of both system and apparatus are independent

– gradually each system begins to feel the interaction of the other

– this interaction cumulates until, at time 1, it triggers a macroscopic reaction which activates the detector

The mathematical model of an adaptive dynamical systems reflects the intuitive idea that, in an adaptive dynamical system, we measure the response to an apparatus pre–disposed to measure an observable A.

Therefore the preparation of our system depends on A. Moreover also the dynamical evolution of our system will depend on A (chameleon effect).

Definition 4 *A classical deterministic adaptive dynamical system is a quadruple:*

$$\{(\Omega, \mathcal{F}) \ , \ \mathcal{O} \ , \ \{P_A\}_{A\in\mathcal{O}} \ , \{T_A\}_{A\in\mathcal{O}}\}$$

where:

– $(\Omega, \mathcal{F})$ *is a measurable space (the state space)*

– $\mathcal{O}$ *is a set of measurable maps from* $(\Omega, \mathcal{F})$ *to* $\mathbb{R}$ *(the observables)*

– for each $A \in \mathcal{O}$*:*

(i) P_A *is a probability measure (the preparation of the experiment)*

(ii) $T_A \ : \ \Omega \to \Omega$ *is an* $\mathcal{F}$*–measurable map (the adaptive dynamics)*

The system is called reversible if T_A^{-1} *exists and is measurable for each* $A \in \mathcal{O}$*.*

7 Local causal probability measures for composite systems

Consider a classical dynamical system composed of two particles $(1, 2)$ with state space S_1, S_2 respectively and two apparata A_1, A_2 with state spaces M_1, M_2 respectively. The state space of the composite system will then be

$$\Omega = S_1 \times S_2 \times M_1 \times M_2 \tag{30}$$

According to von Neumann's measurement theory a measurement of the system $(1, 2)$ by means of the apparatus (A_1, A_2) will be described by a reversible dynamical system

$$T_t : \Omega \to \Omega$$

If time is discretized such a dynamics is described by a single map

$$T : \Omega \to \Omega$$

The preparation of the experiment is described by a probability measure P on Ω

$$P \in \text{Prob } (\Omega)$$

The two additional requirements of:

(i) locality

(ii) causality

were not discussed by von Neumann. Following [Ac93], let us introduce these notions in von Neumann's measurement scheme.

Definition 5 *A dynamics T on the state space (30) is called local if it has the form*

$$T = T_1 \otimes T_2$$

where

$$T_1 : S_1 \times M_1 \to S_1 \times M_1$$

$$T_2 : S_2 \times M_2 \to S_2 \times M_2$$

are dynamics.

Definition 6 *A probability measure P on the space (30) is called local and causal if it has the form*

$$P(d\sigma_1, d\sigma_2, d\lambda_1, d\lambda_2) = P_S(d\sigma_1, d\sigma_2) P_1(\sigma_1, d\lambda_1) P_2(\sigma_2, d\lambda_2) \tag{31}$$

where $P_S(d\sigma_1, d\sigma_2)$ is a probability measure on $S_1 \times S_2$. For all $\sigma_1 \in S_1$, $P_1(\sigma_1, d\lambda_1)$ is a positive measure on M_1.

For all $\sigma_2 \in S_2$, $P_2(\sigma_2, d\lambda_2)$ is a positive measure on M_2.

Definition 7 *A local and causal probability measure on the space (30)*

$$P_S(d\sigma_1, d\sigma_2) P_1(\sigma_1, d\lambda_1) P_2(\sigma_2, d\lambda_2) \tag{32}$$

is called trivial if, denoting

$$\int_{M_1} P_1(\sigma_1, d\lambda_1) := p_1(\sigma_1) \tag{33}$$

$$\int_{M_2} P_2(\sigma_2, d\lambda_2) := p_2(\sigma_2) \tag{34}$$

one has:

$$p_1(\sigma_1) p_2(\sigma_2) = 1 \quad ; \qquad P_S - a.e. \tag{35}$$

Theorem 2 *The local causal probability measure (32) is trivial if and only if both P_1 and P_2 are conditional probabilities, i.e. if and only if*

$$p_j(\sigma_j) = 1 \quad ; \qquad P_{S,j} - \forall\, \sigma_j \in S_j \quad ; \qquad j = 1, 2 \tag{36}$$

where $P_{S,1}$, $P_{S,2}$ denote the marginal measures of P_S.

Proof. Suppose that (32) is trivial. Then by condition (35) there exists a set $N \subseteq$ such that

$$P_S(N) = 0$$

and $\forall\, (\sigma_1, \sigma_2) \in N^c (= S_1 \times S_2 \backslash N)$

$$p_1(\sigma_1) p_2(\sigma_2) = 1$$

If $N^c = S_1 \times S_2$ this clearly implies that both p_1 and p_2 are constant (hence $= 1$) P_S–a.e. In fact $\forall\, \sigma_1 \in S_1$

$$p_1(\sigma_1) = \frac{1}{p_2(\sigma_2)} \quad ; \qquad \forall\, \sigma_2 \in S_2$$

which implies that p_2 is constant and therefore also p_1.

In the general case the argument is similar: denote N_j^c ($j = 1, 2$) the projection of N^c on S_j. Then by Lemma (1) below N_1^c, N_2^c have measure 1 for the marginal measures of P_S. Moreover $\forall\, \sigma_1 \in N_1^c$

$$p_1(\sigma_1) = \frac{1}{p_2(\sigma_2)} \quad ; \qquad \forall\, \sigma_2 \in N_2^c$$

hence p_2 is constant on N_2^c and similarly p_1 is constant on N_1^c.

But then $p_1 \otimes p_2$ is constant on $N_1^c \times N_2^c$ which has measure 1 because $N_1^c \times N_2^c \supseteq N^c$

Lemma 1 *Let $(S_1 \times S_2, P)$ be a probability space let*

$$\pi_j : S_1 \times S_2 \to S_j \quad ; \qquad j = 1, 2$$

denote the canonical projections:

$$\pi_j(s_1, s_2) = s_j$$

and let

$$P_j := P \circ \pi_j^{-1} \quad ; \qquad j = 1, 2$$

denote the marginal measures of P. Then, if $U \subseteq S_1 \times S_2$ has measure 1, also $\pi_j(U)$ ($j = 1, 2$) has P_j–measure 1.

Proof. Since $\pi_j^{-1}(\pi_j(U)) \supseteq U$

$$P_j(\pi_j(U)) = P(\pi_j^{-1}(\pi_j(U)) \supseteq P(U) = 1$$

Lemma 2 *On the space (30) there exist non trivial, local, causal probability measures.*

Proof. Because of Theorem (2) our problem is to construct fields of positive measures

$$\sigma_1 \in S_1 \mapsto P_1(\sigma_1, \cdot) \quad ; \qquad \sigma_2 \in S_2 \mapsto P_2(\sigma_2, \cdot)$$

on M_1 and M_2 respectively, enjoying the following property: denoting

$$p_j(\sigma_j) := \int_{M_j} P_j(\sigma_j, d\lambda_j) \quad ; \qquad j = 1, 2$$

the (positive) functions $p_1(\sigma_1), p_2(\sigma_2)$ are not P–a.e. equal to 1, but $p_1(\sigma_1)p_2(\sigma_2)$ is P–integrable and its integral is equal to 1.

To this goal it is sufficient to consider arbitrary positive measures μ_j on M_j ($j = 1, 2$) and arbitrary non constant, bounded, positive functions $\hat{p}_j \in L^\infty(S_j, P_{S,j})$ with different mean values. Then $\hat{p}_1(\sigma_1)\hat{p}_2(\sigma_2)$ is bounded hence P–integrable. Denoting Z the integral $P(\hat{p}_1 \otimes \hat{p}_2)$, it is clear that the fields of measures

$$\sigma_j \in S_j \mapsto P_j(\sigma_j, \cdot) := \frac{\hat{p}_j(\sigma_j)}{\sqrt{Z}} \mu_j(\ \cdot\ ) \qquad ;\ (j = 1, 2)$$

have the required properties.

By taking convex combinations and limits thereof one constructs more general examples.

8 The role of non trivial local causal measures in the simulations of EPR type experiments

If one calculates the EPR correlations with a local causal measure, one finds (in obvious notations)

$$\int_{S_1\times S_2}\int_{M_1}\int_{M_2} S_a^{(1)}(\sigma_1)S_b^{(2)}(\sigma_2)P_S(d\sigma_1,d\sigma_2)P_{1,a}(d\sigma_1,d\lambda_1)P_{2,b}(\sigma_2,d\lambda_2)$$
$$=\int_{S_1\times S_2} P_S(d\sigma_1,d\sigma_2)\left[\int_{M_1} S_a^{(1)}(\sigma_1)\frac{P_{1,a}(\sigma_1,d\lambda_1)}{p_{1,a}(\sigma_1)}\right]$$
$$\left[\int_{M_2} S_b^{(2)}(\sigma_2)\frac{P_{2,b}(\sigma_2,d\lambda_2)}{p_{2,b}(\sigma_2)}\right]p_{1,a}(\sigma_1)p_{2,b}(\sigma_2)$$

Notice that the integrals in square brackets are functions with values in $[-1,+1]$, however the probability measure

$$p_{1,a}(\sigma_1)p_{2,b}(\sigma_2)P_S(d\sigma_1,d\sigma_2)$$

depends on a and b unless

$$p_{1,a}(\sigma_1)p_{2,b}(\sigma_2)=1 \quad ; \qquad P_S(d\sigma_1,d\sigma_2)-\text{a.e.}$$

i.e. unless P is trivial and in this case the EPR correlations will always satisfy the Bell inequality.

However if the measure is non trivial, the dependence on a and b of the distribution at the source prevents the possibility to apply the Bell inequality.

In other words: the EPR correlations with respect to a local causal measure will not, in general, satisfy the Bell inequality. Therefore the problem of the local violation of the Bell inequality can be reformulated as follows:

can one construct a local causal probability measure on a space of the form

$$S_1\times S_2\times M_1\times M_2$$

and $\{\pm1\}$–valued random variables $S_x^{(j)}$ such that the pair correlations of these random variables reproduce the EPR correlations?

In the paper [AcIR01] such a measure was first constructed.

9 The measurement and the simulation problem

Definition 8 *The problem of measurement consists in approximating the expectation value (18) by the empirical average of a sequence* $(f_j)_{j=1,\dots,N}$ *of measured values of* f*:*

$$\frac{1}{N}\sum_{j=1}^{N} f_j \tag{37}$$

Definition 9 *The problem of simulation consists in producing a sequence of data* (x_j)*, the initial states, such that the empirical average*

$$\frac{1}{N}\sum_{j=1}^{N} f(Tx_j)$$

is a good approximation of the expected value (18).

Remark. In the problem of measurement one needs not to know the dynamics T or the initial states (x_j). In the problem of simulation, this is necessary.

Similarly for the probability P_1: in the problem of measurement one does not need to know the probability measure P_1, but this is reproduced according to the following scheme

In the average (37) one knows that each measured value corresponds to an (unknown) final state Tx_j. Now, among the Tx_j some might be equal among themselves. Denote

$$\{s_1, \ldots, s_M\}$$

the set of all mutually different Tx_j and, for each $h = 1, \ldots, M$, denote

$$F_h := \{j \in \{1, \ldots, N\} : Tx_j = s_h\} = \{j \in \{1, \ldots, N\} : x_j = T^{-1}s_h\}$$

with these notations

$$\frac{1}{N}\sum_{j=1}^{N} f_j = \frac{1}{N}\sum_{f}(Tx_j) = \frac{1}{M}\sum_{h=1}^{M} f(s_h)\frac{|F_h|}{N} \sim \frac{1}{M}\sum_{h=1}^{M} f(s_h)p(T^{-1}s_h)$$

$$\sim \int_S f(s)d(P \circ T^{-1})(s)$$

where, here and in the following, the symbol $\sim$ denotes a good approximation and $|F|$ denotes the cardinality of the set F. In the simulation problem we have to:

(i) know the probability distribution $P_1 \circ T^{-1}$

(ii) produce a sequence of points $(y_j)_{j=1}^{N}$ such that

$$\frac{1}{N}\sum_{j=1}^{N} f(y_j) \sim \int_S f(y)d(P \circ T^{-1})(y) \tag{38}$$

In the following lemma we prove that, for a general class of functions f and of measures P_1, sequences (y_j) satisfying (38) always exist. Moreover from now on we suppose that all measures are regular, i.e. the measure of the total space is the sup of the measures of compact subsets.

Moreover the proof of the lemma will provide an algorithm to produce such sequences.

Lemma 3 *In the above notations, suppose that*

(i) $S = \mathbb{R}^m$, $M = \mathbb{R}^k$ *for some* m, $k \in \mathbb{N}$

(ii) *the set of points of discontinuity for* f *has* P_1*–measure zero and* f *is bounded on* S

(iii) *for any* $\varepsilon >$ *and* $x \in S$ *which is a point of discontinuity for* f*, there is an open ball* $B(x)$ *containing* x *such that*

$$P_1(B(x)) < \varepsilon \tag{39}$$

Then for any $\varepsilon > 0$*, there exists a natural integer* $N \in \mathbb{N}$ *and a sequence of point y* (y_j) *in* S *such that*

$$\left|\frac{1}{N}\sum_{j=1}^{N} f(y_j) - \int_S f(y)d(P \circ T^{-1})(y)\right| \leq \varepsilon$$

Proof. Given $\varepsilon > 0$, fix a compact set S_ε such that

$$P(S \backslash S_\varepsilon) \leq \frac{\varepsilon}{3\|f\|'_\infty}$$

$$\|f\|'_\infty := \min\{1, \|f\|_\infty\}$$

Such a set S_ε always exists.

Now fix a finite open cover $(B_j)_{j=1}^{M_0}$ with the property that, for any $j = 1, \ldots, M$

$$|f(x) - f(y)| < \frac{\varepsilon}{3} \quad ; \qquad \forall\, x, y \in B_j$$

From $(B_h)_{h=1}^{M_0}$ form a disjoint partition $(Q_h)_{h=1}^{M}$ of S_ε and, for each $j = 1, \ldots, M_0$, fix a point

$$Tx_h^0 \in Q_h$$

Having fixed M, (x_h^0), (Q_h) find natural integers K_h, N $(h = 1, \ldots, M)$ such that

$$\left| \frac{K_h}{N} - P_1(Q_h) \right| \leq \frac{\varepsilon}{3M\|f\|'_\infty}$$

The existence of such numbers K_h, N will be proved in Lemma (4).

Notice that in this case

$$1 - \frac{\varepsilon}{3\|f\|'_\infty} \leq \sum_h P_1(Q_h) = \left(\sum_h \left\{ P_1(Q_h) - \frac{K_h}{N} \right\} \right) + \frac{\sum_h K_h}{N} \leq 1$$

and

$$\left| \sum_h (P_1(Q_h) - \frac{K_h}{N}) \right| \leq \frac{\varepsilon M}{3M\|f\|'_\infty} \leq \frac{\varepsilon}{3\|f\|'_\infty}$$

therefore

$$1 - \frac{2\varepsilon}{3\|f\|'_\infty} \leq \frac{\sum_{h=1}^{M} K_h}{N} \leq 1 + \frac{\varepsilon}{3\|f\|'_\infty} \tag{40}$$

which means that

$$\sum_{h=1}^{M} K_h \sim N$$

In the following we assume that

$$\sum_{h=1}^{M} K_h = N$$

The precise meaning of this approximation will be explained in Lemma (4) below.

Having defined N and (K_h) form a new sequence

$$x_1, \ldots, x_N$$

with the property that $\forall\, h = 1, \ldots, M$

$$|\{j : x_j = x_h^0\}| = K_h$$

For such a sequence

$$\left|\frac{1}{N}\sum_{j=1}^{N} f(Tx_j) - \sum_{h=1}^{M} f(Tx_h^0)P(Q_h)\right| = \left|\sum_{h=1}^{M} f(Tx_h^0)\frac{K_h}{N} - \sum_{h=1}^{M} f(Tx_h^0)P(Q_h)\right| \leq$$

$$\leq \|f\|'_\infty \sum_{h=1}^{M}\left|\frac{K_h}{N} - P(Q_h)\right| \leq \frac{\varepsilon}{3}$$

Therefore

$$\left|\frac{1}{N}\sum_{j=1}^{N} f(Tx_j) - \int_S f(Tx)dP(x)\right| \leq$$

$$\leq \frac{\varepsilon}{3}\left|\sum_{h=1}^{M} f(Tx_h^0)P(Q_h) - \int_{\cup_{h=1}^{M} Q_h} f(Tx)dP(x)\right| + \int_{S\setminus(\cup_{h=1}^{M} Q_h)} |f(Tx)|dP(x) \leq$$

$$\leq \frac{2\varepsilon}{3} + \left|\sum_{h=1}^{M}\int_{Q_h} f(Tx_h^0)dP(x) - \sum_{h=1}^{M}\int_{Q_h} f(Tx)dP(x)\right|$$

$$\leq \frac{2\varepsilon}{3} + \sum_{h=1}^{M}\int_{Q_h} |f(Tx_h^0) - f(Tx)|dP(x) \leq \varepsilon$$

Lemma 4 *In the notations and assumptions of the previous Lemma, let us denote*

$$N_0 := \sum_{h=1}^{M} K_h$$

then, for any $h = 1, \ldots, M$

$$\left|\frac{K_h}{N_0} - \frac{K_h}{N}\right| \leq \frac{3\varepsilon}{\|f\|'_\infty}$$

Proof. For any $h = 1, \ldots, M$

$$\left|\frac{K_h}{N_0} - \frac{K_h}{N}\right| = \left|\frac{N}{N_0} - 1\right|\frac{K_h}{N}$$

But from (40) we now that

$$\frac{1}{1+\frac{\varepsilon}{3\|f\|'_\infty}} \leq \frac{N}{N_0} \leq \frac{1}{1-\frac{2\varepsilon}{3\|f\|'_\infty}}$$

Therefore

$$1 - \frac{1}{1+\frac{\varepsilon}{3\|f\|'_\infty}} \geq 1 - \frac{N}{N_0} \quad ; \quad \frac{N}{N_0} - 1 \leq \frac{1}{1-\frac{2\varepsilon}{3\|f\|'_\infty}} - 1$$

But for any $0 < \delta < 1/2$

$$\frac{1}{1-\delta} - 1 = \frac{\delta}{1-\delta} \leq 2\delta$$

Therefore

$$\left|\frac{N}{N_0} - 1\right| \leq m_0\left\{\frac{4\varepsilon}{3\|f\|'_\infty}, \frac{8\varepsilon}{3\|f\|'_\infty}\right\} \leq \frac{3\varepsilon}{\|f\|'_\infty}$$

Lemma 5 *Let P be a Borel probability measure on $\mathbb{R}^d$ and let $f : \mathbb{R}^d \to \mathbb{R}$ be a function which is continuous on the complement of a set of measure zero and P–integrable. Then, for any $\varepsilon > 0$ there exists a finite set*

$$I_1, \dots, I_K \quad ; \qquad K \in \mathbb{N} \tag{41}$$

of mutually disjoint subsets of $\mathbb{R}^d$ with the following property: for any choice of the points

$$x_h^0 \in I_h \quad ; \qquad h = 1, \dots, K$$

$$\left| \int_{\mathbb{R}^d} f(x)P(dx) - \int_{\mathbb{R}^d} \sum_{h=1}^{K} f(x_h^0)\chi_{I_h}(x)P(dx) \right| \leq \varepsilon \tag{42}$$

Proof. Since f is P–integrable, denoting $\|f\|_1$ its $L^1(P)$–norm, one has, for any $L > 0$,

$$\|f\|_1 \geq \int_{\{x \in \mathbb{R}^d : |f(x)| > L\}} |f(x)|P(dx) \geq LP(\{x \in \mathbb{R}^d : |f(x)| > L\})$$

Therefore the left hand side of (42) is less or equal than

$$\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \left| \int_{S_0} f(x)P(dx) - \sum_{h=1}^{K} f(x_h^0)P(I_h) \right|$$

where S_0 is a compact set. Therefore we can find a finite partition

$$I_1, \dots, I_h \quad ; \qquad I_j \cap I_h = \phi \quad ; \qquad j \neq 0$$

$$\bigcup_{j=1}^{K} I_j = S_0$$

such that the above integral is equal to

$$\frac{2\varepsilon}{3} + \left| \sum_{h=1}^{K} \int_{I_h} (f(x) - f(x_h^0))P(dx) \right| \leq \frac{2\varepsilon}{3} + \sum_{h=1}^{K} \int_{I_h} |f(x) - f(x_h^0)|P(dx) \leq$$

$$\leq \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} \sum_{h=1}^{K} P(I_h) = \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} P(S_0) \leq \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

Now choose arbitrary points

$$x_h^0 \in I_h \qquad ; \qquad h = 1, \dots, K$$

Then, if L is chosen so that

$$\frac{2\|f\|_1}{L} \leq \frac{\varepsilon}{3}$$

one has

$$\left| \int_{S} f(x)P(dx) - \sum_{h=1}^{K} f(x_h^0)P(I_h) \right| \leq$$

$$\int_{S(L) \cup B} |f(x)|P(dx) + \left| \int_{(S(L) \cup B)^c} f(x)P(dx) - \sum_{h=1}^{K} f(x_h^0)P(I_h) \right| =$$

$$= \int_{S(L)} |f(x)|P(dx) + \int_{B\setminus S(L)} |f(x)|P(dx) + \left| \int_{S_0} f(x)P(dx) - \sum_{h=1}^{K} f(x_h^0)P(I_h) \right|$$

we conclude that, for any $L > 0$, there is an open set $S(L) \subseteq \mathbb{R}^d$ such that

$$\int_{S(L)} |f(x)|P(dx) < \frac{2\|f\|_1}{L} \tag{43}$$

$$|f(x)| \le L \quad ; \qquad \forall\, x \in S(L)^c \tag{44}$$

By assumption f is continuous outside a set N of P–measure zero. Since P is a Borel measure on $\mathbb{R}^d$ and $P(N) = 0$, there is an open set B containing N and such that

$$P(B) \le \varepsilon/3^L \tag{45}$$

It follows that the set

$$S_0 := S(L)^c \cap B^c$$

is compact and therefore f is uniformly continuous on S_0. Hence there exists $\delta > 0$ such that

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon/3$$

Cover S_0 by a finite set of balls of radius $\delta/2$ and extract from partition of S_0:

Lemma 6 *In the notations and assumptions of Lemma (5), if either*

(i) *f is bounded*

(ii) *there exists one $\overline{x}$ in the domain of f such that $|f(\overline{x})| \le 1$*

Then the family of sets (41) can be chosen to be a partition of $\mathbb{R}^d$ if one replaces, in (42), f by the function f_0, equal to f on S_0 and to $f(\overline{x})$ on S_0^c where, in case (i) $\overline{x}$ is an arbitrary point in the domain of f.

Proof. If f is bounded then one can choose, in Lemma (5),

$$S(L) = \phi \quad ; \qquad L := \max\{1, \|f\|_\infty\}$$

Then (43), (44) hold and defining

$$I_0 := S_0^c = B^c \quad ; \qquad x_0^0 = \overline{x}$$

($\overline{x}$ is an arbitrary point in the domain of f) one has

$$\left| \int_{\mathbb{R}^d} f(x)P(dx) - \int_{\mathbb{R}^d} \sum_{h=0}^{K} f_0(x_h^0)\chi_{I_h}(x)P(dx) \right| \tag{46}$$

$$\le \left| \int_{\mathbb{R}^d} f(x)P(dx) - \int_{\mathbb{R}^d} \sum_{h=1}^{K} f(x_h^0)\chi_{I_h}(x)P(dx) \right| + \int_B |f_0(\overline{x})|P(dx)$$

Using (42) and (45) this is

$$\le \varepsilon + |f(\overline{x})|P(B) \le \varepsilon + \frac{\varepsilon|f(\overline{x})|}{3L} \le 2\varepsilon$$

and, since ε is arbitrary, the thesis follows. In case (ii) one defines

$$I_0 = S_0^c \quad ; \qquad x_0^0 = \overline{x}$$

with $\overline{x}$ as in the statement of the theorem. Then, with the same argument as above, the quantity (46) is majorized by

$$\left| \int_{\mathbb{R}^d} f(x)P(dx) - \int_{\mathbb{R}^d} \sum_{h=1}^{K} f(x_h^0)\chi_{I_h}(x)P(dx) \right| + \int_{S_0^c} |f_0(\overline{x})|P(dx)$$

$$\leq \varepsilon + |f(\overline{x})|P(S_0^c) \leq \varepsilon + P(S(L)) + P(B)$$

$$\leq \varepsilon + \frac{\varepsilon}{3} + \frac{\varepsilon}{3L} < 2\varepsilon$$

and again the thesis follows from the arbitrariness of ε.

Remark. Notice that, in the notations of (6), one has

$$\text{Range } (f_0) \subseteq \text{ Range } (f)$$

i.e., replacing f by f_0, we are not "adding new values" to the function f. In the simulation of physical phenomena this condition is important because the $f(x_j)$ correspond to measured (or measurable) values of an observable f and, in the empirical averages used in the simulations, we want that only these values appear.

Lemma 7 *Let (q_n) be an increasing sequence of numbers in $[0,1]$. On the space of sequences in $[0,1]$*

$$\Omega := \prod_{\mathbb{N}} [0,1]$$

define the probability measure

$$P_0(A) = \int_{\pi_0(A)} dx$$

where

$$\pi_0 : (x_n) \in \prod_{\mathbb{N}} [0,1] \to x_0 \in [0,1]$$

is the projection onto the 0–th coordinate.

Then for a set of sequences of P_0–probability 1 one has, $\forall\, \alpha \in \mathbb{N}$:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \chi_{[q_\alpha, q_{\alpha+1})}(x_j) = q_{\alpha+1} - q_\alpha \tag{47}$$

Proof. It is known that, in the probability space $([0,1], dx)$ there exist ergodic transformations. Let $T : [0,1] \to [0,1]$ be any of then. Then $\forall\, \alpha \in \mathbb{N}$, there exists a set $N_\alpha \subseteq [0,1]$ of Lebesgue measure zero, such that, for $x_j = Tx$ and

$$x \in [0,1] \backslash N_\alpha$$

the limit (47) holds. Therefore on the set of measure 1

$$[0,1] \backslash \bigcup_{\alpha \in \mathbb{N}} N_\alpha$$

the limit (47) will hold for any $\alpha \in \mathbb{N}$. From this the statement easily follows.

Corollary 1 *If* $(p_j)_{j=1}^K$ *is any finite probability sequence:*

$$p_j \geq 0 \quad ; \qquad \sum_{j=1}^{K} p_j = 1$$

Then for each $\varepsilon > 0$*, there exists a sequence of integers* (f_α) ; $f_\alpha \in \mathbb{N}$; $\alpha = 1, \dots, K$ *such that*

$$\left| \frac{f_\alpha}{\sum_{\beta=1}^{K} f_\beta} - p_\alpha \right| \leq \varepsilon \quad ; \qquad \forall\, \alpha = 1, \dots, \alpha$$

Moreover the sum $\sum_{\beta=1}^{K} f_\beta =: N$ *can be made arbitrarily large.*

Proof. For $\alpha \in \{0, 1, \dots, K-1\}$ define:

$$q_0 := 0 \quad ; \qquad q_{\alpha+1} - q_\alpha := p_\alpha$$

and let (x_j) be a sequence in $[0,1]$ satisfying (47). Then there exists N_ε such that, if $N \geq N_\varepsilon$, then for any $\alpha = 0, \dots, K-1$

$$\varepsilon \geq \left| \frac{1}{N} \sum_{j=1}^{N} \chi_{[q_\alpha, q_{\alpha+1})}(x_j) - (q_{\alpha+1} - q_\alpha) \right| = \left| \frac{f_\alpha}{N} - p_{\alpha+1} \right|$$

where f_α denotes the cardinality of the set

$$\{ j \in \{0, 1 \dots, K-1\} : x_j \in [q_\alpha, q_{\alpha+1}) \}$$

It is clear that

$$\sum_{\alpha=1}^{K} f_\alpha = N$$

and this proves the statement.

Theorem 3 *Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a bounded function satisfying the conditions of Lemma (5).*

Then, for any $\varepsilon > 0$, there exist an integer $N \in \mathbb{N}$ and a sequence of points $x_1, \dots, x_N \in \mathbb{R}^d$ such that

$$\left| \int_{\mathbb{R}^d} f(x) P(dx) - \frac{1}{N} \sum_{j=1}^{N} f(x_j) \right| \leq \varepsilon \tag{48}$$

Proof. Given $\varepsilon > 0$, by Lemma (6) (and in its notations) there exist a partition $I_0, \dots, I_k$ of $\mathbb{R}^d$ and points $x_h^0 \in I_h$ such that

$$\left| \int_{\mathbb{R}^d} f(x) P(dx) - \sum_{h=0}^{K} f_0(x_h^0) P(I_h) \right| \leq \frac{\varepsilon}{2}$$

According to Corollary (1) there exist integers $f_0, f_1, \dots, f_K \in \mathbb{N}$ such that, denoting

$$N = \sum_{h=1}^{K} f_h$$

one has

$$\left|\frac{f_\alpha}{N} - P(I_\alpha)\right| \leq \frac{\varepsilon}{2\|f\|'_\infty K} \; ; \quad \forall\, \alpha = 0, \dots, K$$

Therefore

$$\left|\int_{\mathbb{R}^d} f(x)P(dx) - \sum_{h=0}^{K} f_0(x_h^0)\frac{f_\alpha}{N}\right| \leq \frac{\varepsilon}{2} + \left|\sum_{h=0}^{K} f_0(x_h^0)\frac{f_\alpha}{N} - \sum_{h=0}^{K} f_0(x_h^0)P(I_h)\right| \leq \tag{49}$$

$$\leq \frac{\varepsilon}{2} + \sum_{h=0}^{K} |f_0(x_h^0)| \cdot \left|\frac{f_\alpha}{N} - P(I_h)\right| \leq \frac{\varepsilon}{2} + \|f\|_\infty K \cdot \frac{\varepsilon}{2\|f\|'_\infty K} \leq \varepsilon$$

Now let $x_0, \dots, x_N$ be a sequence of points in $\mathbb{R}^d$ such that the cardinality of the set

$$F_\alpha := \{j \in \{0, \dots, N\} : x_j = x_\alpha^0\}$$

is equal to f_α, $\forall\, \alpha = 0, \dots, K$. Then we have

$$\frac{1}{N}\sum_{j=1}^{N} f_0(x_j) = \frac{1}{N}\sum_{h=0}^{K} f_0(x_h^0) \sum_{\{j : x_j = x_h^0\}} 1 = \sum_{h=0}^{K} f_0(x_h^0)\frac{f_\alpha}{N} \tag{50}$$

But, according to Lemma (6) in the present case $f_0(x_j) = f(x_j)$. Combining this with (49) and (50), (48) follows.

10 Bibliography

[AcImRe01] L.Accardi, K.Imafuku, M.Regoli: On the physical meaning of the EPR–chameleon experiment, Infinite dimensional analysis, quantum probability and related topics, 5 N. 1 (2002) 1–20 quant-ph/0112067 Volterra Preprint N. 494 December (2001) N. 494

[AcRe00b] Luigi Accardi, Massimo Regoli, Locality and Bell's inequality. Preprint Volterra, N. 427 (2000), quant-ph/0007005.

[AcRe01a] Luigi Accardi, Massimo Regoli, The EPR correlations and the chameleon effect, discussed during the "Japan-Italy Joint workshop on Quantum open systems and quantum measurement", Waseda University, 27–29 September 2001, quant-ph/0110086, Volterra preprint No. 487 (2001).

[Ac81a] Accardi L., Topics in quantum probability, Phys. Rep. 77 (1981) 169-192

[Ac93] L. Accardi, Einstein-Bohr: one all, in: The interpretation of quantum theory: where do we stand? Acta Enciclopedica, Istituto dell'Enciclopedia Italiana (1994) 95–115, Volterra preprint No. 174 (1993).

[Be64] Bell J.S, "On the Einstein Podolsky Rosen Paradox", *Physics* **1** no. 3 (1964) 195-200.

][Tar98] Angelo Tartaglia: Is the EPR paradox a paradox? Eur. J. Phys., 19 (1998) 307

][Pea70] Phillip M. Pearle, Phys. Rev. D **2**, 1418 (1970),

][CHSH] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23** 880 (1969),

[WJSWZ98] Gregor Weih, Thomas Jennewein, Cristoph Simon, Harald Weinfurter and Anton Zeillinger: Violation of Bell's inequalitiy under strict Einstein locality conditions, Phys. Rev. Lett. **81** (1998) 5039–5043

[Muynck] de Muynck W.M., De Baere W., Martens H.: Interpretations of quantum mechanics, joint measurement of incompatible observables, and counterfactual definiteness, Found. of Phys. **24**, (1994) 1589–1663.

[Khrennikov] A.Yu. Khrennikov: Non-Archimedean analysis: quantum paradoxes, dynamical systems and biological models. Kluwer Acad.Publl., Dordreht (1997).