Math. Struct. in Comp. Science (2013), *vol.* 23, *pp.* 1–91.
DOI: 10.1017/S0960129512000795 Printed in the United Kingdom

On the concept of probability

MIOARA MUGUR-SCHÄCHTER

CeSEF (Centre pour la Synthèse d'une Epistémologie Formalisee)
and
adMCR (association pour le développement de la Méthode de Conceptualisation Relativisée),
47 Boulevard Georges Seurat, 92200 Neuilly-sur-Seine, France
Email: mms@noos.fr

Received 1 January 2012; revised 1 June 2012

In this paper, we discuss the crucial but little-known fact that, as Kolmogorov himself claimed, the mathematical theory of probabilities cannot be applied to factual probabilistic situations. This is because it is nowhere specified how, for any given particular random phenomenon, we should construct, *effectively* and without circularity, the specific and stable distribution law that gives the individual numerical probabilities for the set of possible outcomes. Furthermore, we do not even know what significance we should attach to the simple assertion that such a distribution law 'exists'. We call this problem Kolmogorov's aporia†.

We provide a solution to this aporia in this paper. To do this, we first propose a general interpretation of the concept of probability on the basis of an example, and then develop it into a non-circular and effective general *algorithm of semantic integration* for the factual probability law involved in a specific factual probabilistic situation. The development of the algorithm starts from the fact that the concept of probability, unlike a statistic, does not apply to naturally pre-existing situations but is a *conceptual artefact* that ensures, locally in space and time, a predictability that is more stable and definite than that permitted by primary statistical data.

The algorithm, which is constructed within a *method of relativised conceptualisation*, leads to a probability distribution expressed in *rational* numbers and involving a sort of quantification of the factual concept of probability. Furthermore, it also provides a definite meaning to the simple assertion that a factual probability law exists. We also show that the semantic integration algorithm is compatible with the weak law of large numbers.

The results we give provide a complete solution to Kolmogorov's aporia. They also define a concept of probability that is explicitly organised into a semantic, epistemological and syntactic whole. In a broader context, our results can be regarded as a strong, pragmatic and operational specification of Karl Popper's propensity interpretation of probabilities.

† Note added in proof: After completion of the current paper, I read Christopher Porter's contribution to this special issue, in which he called this the *problem of applicability*, which is more specific and descriptive, and thus a better name.

1. Introduction

The *factual* concept of probability is intuitive and certainly very ancient: for example, Aristotle used it as a qualification of the extension of a belief. It is a feature of common thought and speech, where, after a long and varied evolution, it now works mainly with reference to statistical features.

However, it was only relatively late in the history of thought that any mathematical expressions related to probability emerged. For example, in 1654, Blaise Pascal defined a 'triangle' describing the outcomes of games of chance. The concept itself evolved slowly: Jacob Bernoulli was the first to give a definition through a 'law of large numbers' (published posthumously in 1713). Bernoulli's law was *specifically* tied, more or less explicitly, to games of chance, and it was not until 1812 that Laplace made use of it in connection with the much larger class of factual situations where *'statistical' distributions are stable with respect to repetition of the experimental circumstances*. In this way, a quite general distinction between 'pure' statistics and probabilities crept into scientific thought.

In 1931, more than 200 years after Bernoulli's statement of the law of large numbers, Richard von Mises stressed the connection between the observable and measurable relative frequency of the outcomes of a given event and the probability assigned to that event in *any* situation that involves 'probabilities'. In this way, a general factual concept of *frequential probability* became thoroughly embedded in scientific thinking, from where it migrated more generally into everyday thought.

At the same time, a 'probabilistic syntax' was being developed within the emerging mathematical theory of abstract measures due to Borel, Lebesgue, Paul Lévy, Markov, and many others. Working along these lines, and shortly after von Mises' work, Andreï N. Kolmogorov published in 1933 the first fully worked out mathematical probabilistic syntax (Kolmogorov 1933). In this way, a *general abstract concept of probability* became established within scientific thinking.

The interpretive connection between Kolmogorov's syntactic concept of probability and the factual concept of frequential probability characterised by von Mises was then improved through reformulations of the weak law of large numbers in the form of various 'theorems' establishing 'strong' laws, with contributions from Borel, Kolmogorov, Kintchine and others. As a result, it was believed for some time that von Mises' concept of frequential probability could provide a satisfactory semantic interpretation of Kolmogorov's mathematical concept of a probability measure.

So it seemed for a time that the concept of probability had finally achieved an explicitly constructed structure that was satisfactory from all three main points of view: factual/semantic, syntactic and interpretive.

Meanwhile, between 1872 and 1877, Ludwig Boltzmann introduced the concept of *statistical entropy* into physics through the equation

$$S = k \sum_j (n(e_j)/N) \left(\log \frac{1}{(n(e_j)/N)} \right),$$

where k is a constant tied to energy, and the ratios $(n(e_j)/N)$ are the relative frequencies of a set of physical 'events' $\{e_j\}, j = 1, 2, \dots q$. In this way, he was able to root Rudolf

Clausius' phenomenological second law of thermodynamics within atomic physics, where it characterises the degree of dispersion of certain statistical distributions and their rate of change.

Much later, Shannon (Shannon 1948) published his *theory of information communication*, which was later refined in Khinchin (1957). Shannon introduced the notion of an 'alphabet' of *signs* $a_i, i = 1, 2, \dots n$, emitted by a 'source of information' for encoding and transmitting messages, which are optimised according to certain pragmatic criteria. He then just *posited* a set of stable numerical values for the individual *probabilities* $p(a_i), i = 1, 2, \dots n$, for each of the signs in this alphabet, and assumed that they obey the general conditions imposed by Kolmogorov on probability measures. Furthermore, Shannon defined, as a central concept of his theory of information communication, an entropic form

$$H(S) = k' \sum_i p_i \log(1/p_i)$$

called *the informational or probabilistic entropy of the source of the signs* $\{a_i\}, i = 1, 2, \dots n$. This expression mimics the form of Boltzmann's physical/statistical entropy, but the constant k' is not the same as Boltzmann's constant, and instead of the relative frequencies $(n(e_j)/N), j = 1, 2, \dots q$, he *inserted* the *'probabilities'* $\{p_i\}, i = 1, 2, \dots q$ (which is reasonable because the set of numbers considered by Shannon is, by hypothesis or construction, endowed with rigorous stability, provided the experimental conditions remain unchanged).

For a time it seemed that Shannon's concept of informational entropy enabled the construction of entropic measures of 'complexity', thus leading to a mathematical theory of complexity founded on the concept of probability[†].

Surprisingly, thirty years after he had constructed what was, and still is, considered almost unanimously to be a successful probabilistic syntax associated with a well-formed factual concept of probability, Kolmogorov grew dissatisfied with the factual interpretation provided by the weak laws of large numbers for a probability measure in his formal sense. Therefore, he claimed that his mathematical representation of probabilities was not, as he had previously believed, an abstract reformulation of a well-constructed factual concept of probability, but merely an interesting mathematical construct.

He also asserted that, because of this, his probabilistic syntax could not be used as a basis for Shannon's theory of communication. Nor, *a fortiori*, could it be used as a concept of informational entropy for estimating the complexities of factual entities. As a result, he initiated another approach for measuring complexities in the form of the well-known theory of the 'algorithmic complexity' of *sequences of signs*, which Per Martin Löf, Chaitin and other authors have continued to develop. However, the semantic content of the sequence of signs under consideration is entirely lost in this algorithmic representation of complexity.

It should be mentioned at this point that recent approaches to systems and organisation

[†] Unfortunately, this has led to a remarkable degree of confusion, though I believe it does contain within it the germ of an idea that may be developed rigorously and productively – see Mugur Schächter (2006, pages 261–311).

have placed increasing emphasis on the structures of what signs signify, though these approaches have so far stubbornly remained purely qualitative.

So, in the background, and almost unnoticed, the crucial concepts of factual probability, information and complexity have been undergoing a crisis. Their basic definitions appear to be flawed by vagueness and obscurities, which hinder a clear understanding of the intimate interrelationships that are felt to underlie these concepts, given their common grounding in statistical features.

Karl Popper's well-known propensity interpretation of probabilities (Popper 1967) has thrown a pale beam of light onto this fuzzy conceptual ground. However, despite having an enduring effect, it has not changed the fact that the factual concept of probability, which is so central in everyday thought, as well as in physics and numerous other scientific domains, *has simply not yet been fully worked out.* When considered globally, the concept of probability does not seem to have achieved the status of a clearly formulated epistemological/operational/syntactic structure, in the same way as happened, for instance, for the concept of 'a geometry' through the progressive integration of the Kantian conception of space as an *a priori* form of human intuition, the geometries of Euclid, Lobatchewsky and Riemann, and Henri Poincaré's analyses of the human psycho-kinetic physical actions that build the factual structure of what is called 'physical' space, which can be described syntactically using the particular geometry of Euclid alone.

However, it seems clear that epistemological factors do indeed act strongly when the concept of probability is used in practice, and that this determines the syntactic features of the concept. For example, the concept of 'local information' in Shannon's theory of information communication is basic, and is tied to what are definitely subjective considerations. The concept of informational entropy also *measures* aspects that are at the same time both subjective and objective.

The unresolved state of the concept of probability we have described constitutes a major lacuna in current scientific thinking, though it seems to be largely unrecognised. In the rest of this paper we will provide a thorough analysis of this fundamental problem and then propose a solution. To do this, we will bring together the main semantic/operational/epistemological features that explicitly or implicitly make use of the concept of probability in some way within the natural sciences, and organise these features into a coherent whole. In this way we hope to provide *a complete factual concept of probability that is fully satisfactory as an interpretation of the mathematical probabilistic syntax.* This is all we aim to achieve in the current paper, and we leave open the possibility that the creation of an improved concept of factual probability might reveal the need for some new developments of the corresponding mathematical probabilistic syntax and optimal ways to connect it with its factual interpretation.

Organisation of the paper

We shall begin in Section 2 by giving a thorough analysis and definition of the problem raised by the factual concept of probability as it now stands.

Section 3 provides an overview of the *method of relativised conceptualisation* (MRC), which we will use later in the paper to solve the problem. Space limitations mean that we will only be able to list the main features of the general method that we will need for the rest of the paper, though references are given for where full details can be found. However, as an introduction to the general method, and to provide motivation for it, we begin Section 3 with a brief sketch of a special case of the method in the form of infra-quantum mechanics (Mugur Schächter 1991; 1992a; 1992b; 1993; 2011).

In Section 4, we analyse a series of simple games based on a jigsaw puzzle to provide familiarity with the use of MRC, but also to show how we might be able to use it to construct, in any given empirical probabilistic situation, a general effective procedure for identifying a *relativised* factual probability law for that situation. The jigsaw puzzle metaphor plays an important role in the rest of the paper.

We begin Section 5 by clarifying how the concepts of a random phenomenon and a probabilistic situation are viewed within MRC, and discuss the relationship between factual probabilistic data and Kolmogorov's syntax. Then, using the analysis developed in Section 4 as a guide, we use MRC to build an *effective* algorithm that, within its domain of application, constructs the factual probability law tied to any given random phenomenon, and which thus provides a solution to Kolmogorov's aporia.

In Section 6, we consider how the algorithm developed in Section 5 is related to the weak law of large numbers, and show that they are mutually consistent. We then show how they can be combined to give a unified expression that has a similar structure to the weak law of large numbers but is also meaningful and semantically rich.

We present our general conclusions on Kolmogorov's aporia in classical probability theory in Section 7. Then, as a postscript, in Section 8, we raise a rather surprising and basic question about the possibility, or impossibility, of constructing the concept of probability *within* fundamental quantum mechanics.

2. The problem: Kolmogorov's aporia

2.1. *Kolmogorov's classical definition of a probability space*

The fundamental concept in Kolmogorov's formulation of the mathematical theory of probabilities is *a 'probability space'* $[U, \tau, p(\tau)]$ where:

— $U = \{e_i\}$, where $i \in I$ with I some index set, is a *universe of elementary events* (a set) generated by the repetition of an 'identically' reproducible procedure Π (often called an 'experiment'), which, in general, produces elementary events e_i that *vary* from one realisation of Π to the next, despite the fact that all the realisations are assumed to be identical.

— τ is an *algebra*† *of events* built on U, where an event, which we will denote by e, is a subset of U, and is defined to have occurred each time an elementary event e_i from e has occurred.

† An algebra built on a set S is a set of subsets of S, which includes both the set S itself and the empty set $\emptyset$ such that if it contains the subsets A and B, it also contains $A \cup B$ and $A - B$.

— $p(\tau)$ is a *probability measure* defined *on the algebra of events* τ (and *not* on the universe U of elementary events e_i)†.

Given a particular factual probabilistic situation, the numerical value of the individual probability $p(A)$ of an event A in the algebra of events τ considered for that particular factual situation is *not* specified by the formal concept of a probability measure $p(\tau)$. Only the *general* relations set out in the footnote below are defined by this formal concept.

A pair (Π, U) containing an identically reproducible procedure Π and the corresponding universe of elementary events U is called a *random phenomenon*.

We can define various algebras of events τ on a given universe U, and can thus create different associations of the form

$$\{[\text{random phenomenon}], [\text{corresponding probability space}]\},$$

all stemming from the same pair (Π, U).

In earlier representations of the *factual* concept of probability (in particular, those due to Bernoulli and von Mises), only the general conditions to be imposed on the structure of a factual 'probability law' were sketched out in mathematical terms. By contrast, Kolmogorov's formal concept of a probability space $[U, \tau, p(\tau)]$ gave a formal structure (*cf.* the footnotes on the previous and current pages) to all the relevant features (elementary events, events and probability measure), and was also located quite definitely within the well-developed and even more general mathematical syntax of the theory of measures. This represented a huge advance.

However, Kolmogorov's formalisation presupposed that the weak law of large numbers offered a satisfactory factual definition of the *individual* numerical probabilities of the events involved in any given empirical probabilistic situation, and thus also of their distribution. But this assumption fell apart progressively, until, finally, it became obvious that there was a gaping void in the interpretation.

† If we did define a probability measure on U, it would consist of a set of real numbers $p(A)$, each associated with an event A in U and such that:

$$\begin{aligned} 0 \leq & p(A) \leq 1 \\ p(U) &= 1 \text{ (norm)} \\ p(\emptyset) &= 0 \\ p(A \cup B) &\leq p(A) + p(B) \end{aligned}$$

where the equality in the final line only holds if A and B are mutually 'independent' in the sense of probabilities, that is, if they have no elementary event e_i in common, that is, $A \cap B = \emptyset$. In the 'frequency interpretation' of the concept of probability, the number $p(A)$ is defined as the value of the mathematical limit, which is assumed to exist, towards which any relative frequency $n(A)/N$ converges when the number of realisations N of the repeatable procedure Π is increased to infinity, $n(A)$ being the number of outcomes of A when Π is repeated N times. This definition is supposed to constitute the *factual* definition of the *individual* probabilities $p(e)$ to be assigned to the *isolated* events e in the algebra τ. But, as we stress in this paper, this 'factual' definition is deficient: in particular, this is because the concept of a mathematical limit used here *mixes* a semantic feature with a non-effective syntactic one.

2.2. *On the factual interpretation of an abstract probability measure*

The problem we have described lying at the heart of the concept of probability is still little appreciated, in part perhaps because it remains rather vague. For most physicists, communication specialists and mathematicians who use probability theory but do not have it as the main focus of their research, not to mention the man in the street, the distinctions between a syntax and its interpretation seem ill-defined, and do not appear to have any practical importance for them. Moreover, non-experts still confidently assume that all important questions concerning the concept of probability must have long since been answered in the specialist literature. Of course, this is a common situation in science, and even necessary if science is to progress without constant navel gazing. However, specialists studying the foundations of probability theory are aware that there is an interpretation problem with the current mathematical concept of a probability measure, though they may not all realise how vital this problem is. Kolmogorov himself wrote:

> I have already expressed the view ... that the basis for the applicability of the results of the mathematical theory of probability to real random phenomena must depend in some form on the *frequency concept of probability*, the unavoidable nature of which has been established by von Mises in a spirited manner... [But] The frequency concept [of probability] which has been based on the notion of limiting frequency as the number of trials increases to infinity, does not contribute anything to substantiate the applicability of the results of probability theory to real practical problems where we have always to deal with a finite number of trials.
> (Kolmogorov (1963) as quoted in Segal (2003))

Each word of this quotation merits considerable attention, and though it could not be clearer, it is worth adding a few comments at this point.

There is currently a rather ill-defined, but very potent, belief that the weak law of large numbers can establish deductively both:

— the *existence* of a factual probability 'law' for any factual random phenomenon; and
— the *numerical* distribution used in this law for the *individual* probabilities of the events involved.

However, this is not the case, since the weak law of large numbers only asserts the following, where we use the usual notation:

Given a set $\{e_j\}, j = 1, 2, \dots q$, of events e_j (or elementary events, since no distinction is made here in this respect), *if* a factual probability law $\{p(e_j)\}, j = 1, 2, \dots q$, on this set *exists*, then, for every e_j and every pair (ϵ, δ) of arbitrarily small real numbers, there exists an integer N_0 such that when the number N of 'identical' repetitions of the experiment Π related to the random phenomenon under consideration becomes equal to or greater than N_0, *then* the meta-*probability*[†]

$$\boldsymbol{\mathcal{P}}\left[(|n(e_j)/N - p(e_j)|) \leq \epsilon\right] \tag{1}$$

[†] We use the prefix 'meta' here, and throughout the paper, to mean that the definition of the considered event or probability *involves* the events e_j or the probabilities $p(e_j)$, respectively, so it is conceptually posterior to these events or probabilities.

of the meta-event defined by

[the absolute value of the difference $(n(e_j)/N - p(e_j))$ between the relative frequency $n(e_j)/N$ counted for the event e_j, and what is called the probability $p(e_j)$ of that event e_j, is *smaller* than or equal to ϵ]

becomes *greater than* or equal to $(1 - \delta)$.

This assertion can be expressed more synthetically by the following well-known expression:

$$\forall j. \forall(\epsilon, \delta). \exists N_0. \forall N. \ (N \geq N_0) \Rightarrow \boldsymbol{\mathcal{P}} \left[(|n(e_j)/N - p(e_j)|) \leq \epsilon \right] \geq (1 - \delta). \tag{2}$$

This is also sometimes expressed less precisely by saying that *if* a probability law $\{p(e_j)\}, j = 1, 2, \ldots q$, exists on the set of events $\{e_j\}, j = 1, 2, \ldots q$, *then* for any j, as N 'tends towards infinity', the absolute value of the difference between the relative frequency $(n(e_j)/N)$ and the probability $p(e_j)$ 'tends in probability' towards 0. However, the symbol $\boldsymbol{\mathcal{P}}$ in the expression

$$\boldsymbol{\mathcal{P}} \left[(|n(e_j)/N - p(e_j)|) \leq \epsilon \right]$$

only denotes a meta-*probability*, so we *cannot* be *certain* that the value of the specified difference tends to 0.

Summing up, what the weak law of large numbers proves is that: *if* an *unknown* factual probability law

$$\{p(e_j)\}, j = 1, 2, \ldots q,$$

and the corresponding meta-probability law

$$\{\boldsymbol{\mathcal{P}} \left[(|n(e_j)/N - p(e_j)|) \leq \epsilon \right]\}, j = 1, 2, \ldots q,$$

both 'exist', *then*, as the number of completed trials increases towards infinity, the mathematical tendency of each relative frequency $n(e_j)/N$ of an event e_j approaching the initially unknown numerical value $p(e_j)$ is itself very 'probable' in the sense of the other meta-probability law

$$\boldsymbol{\mathcal{P}} \left[(|n(e_j)/N - p(e_j)|) \leq \epsilon \right],$$

which we have assumed to exist.

However, the weak law of large numbers says *nothing* about what significance should be assigned to the simple assertion that a factual probability law 'exists'. All it does is correlate the two values

$$\boldsymbol{\mathcal{P}} \left[(|n(e_j)/N - p(e_j)|) \leq \epsilon \right] \qquad \text{and} \qquad p(e_j),$$

for any index j, in a way that excludes certainty and, quite essentially, allows fluctuations, which are measured by the pair of arbitrarily small real numbers (ϵ, δ).

The law of large numbers does progressively construct a definition of the *a priori* unknown numerical value of an individual probability $p(e_j)$, for any index j, in the form of the famous 'relative frequency definition'. However, this definition:

(a) is not purely factual because the concept of a mathematical limit is abstract;
(b) is *non-effective* (as noted above and stressed by Kolmogorov);

(c) introduces, through the dependence between the numerical values

$$\boldsymbol{\mathcal{P}}\left[(|n(e_j)/N - p(e_j)|) \leq \epsilon\right] \qquad \text{and} \qquad p(e_j),$$

a sort of logical regression, which develops upwards through a hierarchy of conceptualisation levels – a sort of circularity spread out along a spiral;

(d) is thought of as being constructed by the *convergence* towards the *a priori* unknown value of $p(e_j)$ on the basis of the mere assumption of the 'existence' of the two probability laws

$$\boldsymbol{\mathcal{P}}\left[(|n(e_j)/N - p(e_j)|) \leq \epsilon\right] \qquad \text{and} \qquad p(e_j),$$

without in any way specifying what sort of physical features or circumstances this 'existence' may consist of.

This convergence is defined exclusively in general mathematical terms through *a priori* postulated conditions of convergence, integrability and so on, but these purely mathematical and general assumptions have not, at least so far, been translated into a language that can adequately describe the class of physical circumstances at work when a given probabilistic situation is realised. The conceptual and physical content of the significance to be assigned to the assertion that a factual probability law 'exists' has not been made clear. In this respect, no translation has been elaborated between the mathematical language in which the weak law of large numbers is expressed and physical factuality.

Even if we ignore objections (a)–(c), point (d) alone is a very strong criticism. Indeed, the counted relative frequencies $n(e_j)/N$ from (2) can *only* be thought of as playing the role of an ideal specification by progressive materialisation of the limiting numerical value $p(e_j)$ if the process of evolution of the relative frequencies $n(e_j)/N$ while N increases is somehow *materially constrained* throughout this process by the existence of whatever it is the *a priori* unknown limit $p(e_j)$ denotes in the physical world; *otherwise why should there be any convergence towards* $p(e_j)$ *at all?* And we can ask the same question about the meta-probability $\boldsymbol{\mathcal{P}}$ for the meta-events

$$\left[(|n(e_j)/N - p(e_j)|) \leq \epsilon\right].$$

We might be able to conceive of some sort of progressive emergence of knowledge concerning the somehow pre-existing values of some qualifier† labelled '$p(e_j)$' of some unspecified material entity through the material effects of '$p(e_j)$' by following the evolution of the relative frequencies $n(e_j)/N$ considered in (2). However, to achieve a genuine understanding of such a process, and thus an epistemological command of the meaning of (2), it would be, at least, very useful for us to have an *independent definition* of the factual meaning of the assumption of the 'existence' of a 'probability law', even if this is not felt to be an essential conceptual requirement.

† Our use of the term 'qualifier' here is similar to its use in grammar, where, for example, an adjective can be described as a qualifier of a noun. More concretely, a qualifier can be thought of as the result of an experiment. We will give a precise definition of what we mean by a 'qualifier' in point (3) of Section 3.5.2. We will also use related words, such as 'qualify' and 'qualification', with corresponding meanings.

Kolmogorov's abstract concept of a probability measure cannot generate an independent factual definition of the individual numerical probability values $p(e_j)$. This is not its role since it was constructed simply as the formal representation of the *common* features of *all* conceivable factual probability laws. The individual features of each specific factual law were omitted from the start, and it was assumed that they were independently constructible. This means that they are simply not available within the pool of potential specifications encapsulated by the abstract concept of a probability measure.

It is not unusual for there to be confusion between the characters required for the primary representation of an individual specific factual instance of a given type and a purely syntactic framework for the general representation of the whole class of individual representations of that type. The power of a syntax often induces an implicit belief that it should be possible to *derive* factual data from a syntactic system, but this is never possible. The semantic conceptualisation is primary, and it has specific content, whose genesis can only consist of *direct* epistemological/operational/conceptual interactions between the mind and what we call 'reality'†. If we invert, *a posteriori*, some sequence order within the hierarchy of epistemological assumptions that emerge in the course of a process of conceptualisation, then, quite systematically, we can produce a long-lasting stagnation in the subject through the tackling of illusory problems. Even though semantic content can be located within a syntax, the *nature* of a syntax is completely different from the *nature* of a factual specification.

Other authors had already expressed some reservations about the applicability of Kolmogorov's theory of probabilities before Kolmogorov did. For instance, selecting just one amongst many possible quotations, Solomonoff wrote:

> 'Probability theory tells how to derive a new probability distribution from old probability distributions... It does *not* tell how to get a probability distribution from data in the real world.' (Solomonoff 1957)

However, it was Kolmogorov himself who finally produced a definitive veto on the application of his mathematical theory to factual problems. In particular, throughout the 1980's he refused to accept the use Shannon's central concept of 'informational entropy' for assigning numerical estimates to 'complexities'. For instance, he wrote:

(1) 'Information theory must precede probability theory and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character.'

(2) 'The applications of probability theory can be put on a uniform basis. It is always a matter of consequences of hypotheses about the impossibility of reducing in one way or another the complexity of the descriptions of the objects in question. Naturally this approach to the matter does not prevent the development of probability theory as a branch of mathematics being a special case of general measure theory.'

(3) 'The concepts of information theory as applied to infinite sequences give rise to very interesting investigations, which, without being indispensable as a basis of

† See Mugur Schächter (2011) for a striking illustration of this assertion.

probability theory, can acquire a certain value in the investigation of the algorithmic side of mathematics as a whole.' (Kolmogorov 1983)

But also, and quite radically, Kolmogorov advocated the elimination of his own formal concept of probability from *all* the representations that had been considered to be 'applications' of this formal concept.

The conceptual situation outlined above is what I call 'Kolmogorov's aporia'.

2.3. *Conclusions for Section 2*

We have seen that Kolmogorov, who was the father of the modern mathematical theory of probabilities, proposed that his own formal concept of probability, which was epistemologically profoundly rooted in concrete human experience and thought, should be isolated within an abstract cage. And Kolmogorov was a major thinker, so his view must be taken seriously, even though it represents an extreme stance. Mathematicians appear to have accepted this view without much resistance, and this has already changed the direction of research on measures of the degrees of 'complexity'. However, for many mathematicians, concepts related to factual entities are viewed merely as shadows of mathematical concepts – Platonism seems to be consubstantial with mathematics.

For a physicist, however, the following steps seem to be essential:

(1) First base the formal concept of a probability measure, which *stems from factuality*, on *an explicit and definite* meaning that can be assigned to the general assertion that a factual 'probabilistic situation' involves the 'existence' of a factual probability law.

(2) Establish how the numerical distribution of individual probabilities from the factual probability law involved in a given 'probabilistic situation' can be constructed in an effective way.

(3) Indicate how the conceptual/operational/epistemological organisation of the factual concept of probability entailed by Steps (1) and (2) can be optimally connected to Kolmogorov's probabilistic syntax – either in its present form or appropriately modified.

(4) Show precisely how the results we may get for Steps (2) and (3) can be connected coherently with the weak law of large numbers.

(5) Indicate explicitly the domain of applicability of the resulting system.

The rest of the current paper is an attempt to realise this procedure.

3. The framework for treating Kolmogorov's aporia

Kolmogorov's probabilistic syntax belongs in the realm of classical thinking, so we shall restrict ourselves to the classical domain, where the entities to be studied can be perceived directly and/or can be represented by *models*. We make this restriction clear from the start because it turns out, quite surprisingly, that in the case of fundamental quantum theory, where no models are explicitly formed, the very *definability* of the famous 'essential quantum-probabilities' raises a problem that calls for a carefully constructed and specific answer – not to mention the well-known fact that Kolmogorov's representation does not apply directly to the quantum case.

The main tool we will use is a general *method of relativised conceptualisation*, MRC, which the current author has been developing since 1984[†]. This method has been developed by synthesising and generalising results that were originally derived in the context of a study of the way in which fundamental quantum mechanics is able to signify[‡].

So, to help make the current paper more self-contained, this section provides a brief sketch of the origins and main features of the method of relativised conceptualisation.

3.1. *The genesis of MRC: infra-quantum mechanics (IQM)*

3.1.1. *A hypothesis tied to a historical fact.*

Quantum mechanics had no unique initial author – there was no equivalent of Newton, Maxwell, Carnot, Boltzmann or Einstein. It arose from a relatively large number of very different contributions by Plank, Einstein, Bohr, de Broglie, Schrödinger, Heisenberg, Born, Pauli, von Neumann, Dirac and many others, which finally led to a coherent mathematical theory of microstates in the form of *fundamental* quantum mechanics[§], which yields predictions based on a system of algorithms. However, even today, the quantum-mechanical algorithms are cryptic and raise problems of interpretation – *nobody claims to fully understand how or what quantum mechanics signifies.*

This is a very peculiar situation, and we are led to ask how it came about.

In seeking an answer to this question, we were led to the hypothesis that every time a physicist has tried to understand microstates, the cognitive demands imposed have been so radically different from all those previously encountered, and so *extreme*, that no individual mind working in isolation has been able to grasp them globally, and thus construct a coherent representation – and this *same* very peculiar cognitive situation has acted, without becoming wholly explicit, each time the problem has been addressed. So the construction of the quantum-mechanical formalism has been orchestrated by this impersonal and very peculiar cognitive situation.

Moreover, what the quantum-mechanical formalism signifies, and how it does it, have remained cryptic because each time an interpretation question has been formulated and examined, it has almost always been addressed with respect to the formalism itself rather than to the cognitive situation that determined the structure of the formalism. As a result, this cognitive situation and its consequences have *never* been characterised explicitly, thoroughly and globally.

[†] For details, see Mugur Schächter (1984; 1991; 1992b; 1992c; 1993; 1995; 1997a; 2002a; 2002b; 2002c; 2006; 2011)

[‡] Our use of the word 'signify' here and elsewhere in the current paper is similar to its use in semiotics where a signifier (a syntactic element) is used to signify (refer to and communicate) the signified (some concept).

[§] We make a clear distinction here between fundamental quantum mechanics, where no models are explicitly formed, or even permitted in principle, and any preceding or subsequent theories of microscopic physical entities (such as atomic and nuclear physics and elementary particle theories), which quite explicitly introduce models.

3.2. *A project*

The hypothesis formulated in the previous section suggested a project: we should first make a blank slate of the mathematical formalism of quantum mechanics, and then try to construct, in strictly *qualitative* terms, some mutually communicable and agreed understanding of what we mean by 'microstates' based only on the constraints imposed by the cognitive situation under consideration and the general ways in which humans form concepts. This project has led to what I have called *infra-quantum mechanics (IQM)*, which is a sort of epistemological/physical representation of microstates, constructed *independently* of quantum theory, but through which the whole way in which fundamental quantum mechanics signifies finally becomes clear (Mugur Schächter 2011).

3.3. *Sketch of the construction of infra-quantum mechanics*

In this section, we will give a very brief summary of the construction of infra-quantum mechanics. In doing this, we will demonstrate the *radically fundamental and relative* character of the completely new and unprecedented form of description involved in the quantum mechanics formalism. Moreover, we will reveal the universality hidden in the descriptions created in this newly identified form, and the potential benefits it brings. The main aim of this sketch is to enable a better understanding of Section 3.4, which summarises the main features of the method of relativised conceptualisation, which is a *generalisation* of the method outlined here and was identified specifically for the description of microstates. Later in the paper, we shall work within this general method of relativised conceptualisation to develop our solution to Kolmogorov's aporia.

In the following we will be *obliged*, at least initially, to make use of pre-existing structures of thinking and communicating: otherwise, we could not even *begin* to communicate the results; indeed, we could not even have begun to work them out. However, despite this unavoidably classical starting point, the development process we will follow will progressively induce several non-classical concepts whose verbal formulation, though presented in familiar language, will involve some radical breaks with classical thinking. This can be viewed as one of the miracles of thought and language: the structure of the results may be quite different from the structure of the starting point, and this enables the emergence of completely new concepts. We shall make free use of this fact.

3.3.1. *Descriptions.* We say that any *knowledge* that can be communicated without restrictions (such as the restrictions involved in pointing, miming and so on) is a *'description'*. By definition, a description involves an *entity-to-be-described*, which, in general, is not necessarily an 'object' in the usual sense, together with some qualifiers of this entity. The basic entities-to-be-described in fundamental quantum mechanics are *a priori* called 'states of microsystems' or *microstates* for short†. These microstates form a class of

† The stable micro-*systems* themselves (electrons, protons, neutrons and so on) were first studied in atomic and nuclear physics, where they were characterised by specific 'particle'-constants (mass, charge, magnetic moment and so on). *Changes* to stable micro-*systems* (such as, creation or anni-

hypothetical entities, whose existence is postulated beforehand on historical and methodological grounds, but no human being could ever actually perceive them. For entities of this sort, the construction of qualifiers endowed with some kind of stability raises difficult and deep questions. Despite this, fundamental quantum mechanics does in fact deal with microstate qualifiers, which means that some *strategy of description* must have been at work, and it has succeeded in overcoming the epistemological difficulties. As mentioned above, our aim is to describe this emergent strategy using only the constraints imposed by the cognitive situation involved and general human modes of conceptualisation.

3.3.2. *Microstates as 'entities-to-be-described'.*
We will begin by considering the entities-to-be-described, *viz.* the microstates, in more detail. Since they cannot be perceived, we cannot study them by just selecting them from some ensemble of pre-existing entities. Neither can we study entities of this kind by simply examining observable marks produced spontaneously on macroscopic devices by 'naturally' pre-existing microstates, since there would be no criteria for deciding which mark is to be assigned to which microstate. The *only* possible general solution is the following:

(1) We first perform a *defined* and *repeatable* macroscopic operation, which we just assume will generate a *given*, but unknown, 'microstate'.
(2) Then, *afterwards*, we try in some way to 'know' something about this microstate we have supposedly generated.

We will see below how adventurous this approach has been.

So, consider a macroscopically defined operation that we suppose generates a 'microstate'. At this initial stage of our inquiry, we know nothing about the content of a 'microstate', which is why we have included the quotation marks. It is just an empty verbal box whose *a priori* use is determined by a general structural feature of our current modes of thought, which says that a 'thing' can only exist in some 'state': a *given* 'thing' is the *genus proximus* of all the 'states' of a 'thing', so a 'state' is an unavoidable specification of *this* 'thing'. According to the general linguistic and conceptualisation structures we use, the thing cannot be without some state, and a state without the corresponding thing is nonsense. So, according to classical thinking, a thing called a 'micro-*system*' necessarily possesses 'states'. And quantum mechanics, acting within the basic human conceptualisation structures, and through the requirement to maintain continuity with macroscopic mechanics, is concerned with the specific task of establishing knowledge about the *states* of microsystems, that is, 'microstates'. In other words, knowledge must be cast in pre-established *mechanical* terms involving what we call 'position', 'momentum', 'energy' and so on.

hilation) are studied in nuclear physics and field theory. The *states* of stable micro-*systems*, that is, the 'microstates', are specifically studied in *fundamental* quantum mechanics (for Dirac the word 'state', when it is used in relation to microscopic entities, was short for a 'way of moving' (dynamics)). Within fundamental quantum mechanics, the dynamics of microstates is characterised by distributions of values of 'dynamical state-*observables*'.

So, the *qualifier grids*† for the sorts of microstate *qualifiers* we want to use are also postulated *beforehand*, and quite *independently* of the microstate being considered. Furthermore, in general, when a microstate we wish to study emerges, it is entirely undetermined with respect to *these* qualifier grids, and it is still strictly unqualified. This radical assertion is in no way weakened by our use of the generic word microstate and the names of the qualifiers, since, *a priori*, these verbal labels will only add *newly* determined knowledge to our general, pre-existing conceptualisation structures, and do not specify any further individual content. In short:

> When the microstate generated by a macroscopically defined operation emerges, it is still strictly *unspecified* and *non-individualised* 'mechanically', or by any other sort of qualifier, within the *a priori* conceptual mould consisting of the general class of microstates.

However, and again in accord with the classical thinking we are currently using, the generated microstate *has* to be conceived of as emerging in some way *relative* to the operation used to generate it, since otherwise we would immediately be in conflict with the pre-existing causal structure of the classical conceptualisation, which we decided *from the very start* to maintain in order to ensure the intelligibility of our construction – this was for our own benefit as much as for those we want to establish communication and agreement with. So, at least to begin with, we *are obliged* to admit this relativity, though later results may mean we could decide to re-examine it critically‡.

Furthermore, causality, together with our general modes of thought, force us to think of the generated unknown microstate beginning its existence within the immediate neighbourhood of the place where the generating operation occurred (Kant postulated that the assignment of some spatial location to any perceived, or merely conceived, *physical* entity is inescapably required by 'an *a priori* form of human intuition', and this view, though many ignore it, has never been refuted).

Now, this notion that the microstate introduced by a given generating operation is somehow relative to this operation allows us to *label* it: *this* microstate is a result of *this*, known, macroscopically defined operation of state generation. Specifically, if we write G to denote the macroscopically defined generating operation under consideration, with

† The notion of a 'qualifier grid' encapsulates the requirements for determining qualifier values. A qualifier grid consists of:

(a) a semantic domain;
(b) a list/specification of possible values imposed on the semantic domain;
(c) a measurement procedure producing results that can be read directly or indirectly in some way using the human senses;
(d) a procedure to translate the results of (c) into one, and only one, of the values defined in (b), which can be communicated to and understood in a unique way by others.

See point (4) of Section 3.5.2 for more about 'qualifier grids'.

‡ However, we can say immediately that causality will have been ejected from the *final* representation we obtain as a result of the first stage of human conceptualisation we are just beginning to consider, but this will not introduce any inconsistency because the process of conceptualisation itself and the results it produces are fundamentally distinct entities – see Mugur Schächter (2011, Chapter 5).

the requirement that it be *reproducible* in some communicable way, we will write ms_G to denote the corresponding microstate.

Though at this stage the symbols G and ms_G are still devoid of any mathematical content, their introduction is of utmost importance since it allows us to *communicate* the fact that the generated microstate, though it is completely undetermined from the point of view of the specific qualifiers that we are using to investigate it, is, nevertheless, made stably available for 'study'. In this sense, it has been *captured*. From now on, by reproducing G, we can produce as many 'copies' or 'replicas' of the microstate denoted by ms_G as we want. Also, each replica can be subjected to some subsequent operation of 'examination', and we will be able to communicate clearly what we have done using words and signs. However, this does involve an *assumption*, namely, that any realisation of the operation G produces a replica of one and the same microstate ms_G:

> A microstate will be stable in the role of an entity for which qualifier values can be determined through subsequent experiments if and only if we assume a one-to-one relation $G \leftrightarrow ms_G$.

Determining the validity of this assumption is far from trivial, but it has been very thoroughly examined elsewhere†, so here we shall just assert the conclusion that, in the cognitive situation being considered, this assumption is simply *unavoidable*. Without it, we could not even *begin* to construct any knowledge about microstates. On the other hand, the consequences of accepting this assumption have been illuminating. So we shall make the assumption as a *methodological decision*, which we can re-express as follows:

> The result of any realisation of the macroscopically defined generating operation denoted G, *whatever it may be*, is called *'the' microstate corresponding to G* (note the use of the definite article 'the') and is denoted by ms_G.

In this way, we now have an *aconceptual* specification, or 'definition', of an unlimited number of replicas of the entity called '*the* microstate ms_G corresponding to G'. In other words, we have a purely operational/factual specification of an entity for which the values of any qualifiers that could be used to identify it within the whole class labelled, *a priori*, by the word 'microstate' are still strictly unknown. Note that G is not a qualifier of ms_G; it *only* tells us how to produce ms_G (for example, the fact that we know how a baby has been *produced*, does not mean we know anything about the properties of that baby *itself*). However, although G does not qualify what we have labelled ms_G, this sort of 'definition' of ms_G can be communicated, with an agreed meaning. This is very remarkable: it finally enables us to get around the lack of any *predicate* allowing us to define a microstate in the usual, classical way. Indeed, classically, a definition is usually realised verbally/conceptually using predicates that both define *and* qualify it at the same time (for example, if we look up 'cat' in a dictionary, we find (Webster, fourth edition of the Merriam series) 'carnivorous domesticated quadruped…').

† A very thorough argument was given in Mugur Schächter (2011) for microstates. The question was also examined in *general* terms, in complete detail and through all its stages, in Mugur Schächter (2006) – see also Mugur Schächter (2002a) and Mugur Schächter (2002b), and other earlier work.

So we have now completed the first step in our project on the basis of a methodological decision that introduces a radical non-classical *separation* between an entity in the role of an object-to-be-qualified and any subsequent possible operations qualifying this entity.

3.3.3. *Qualifying a microstate: the emergence of a 'primordially' statistical and 'transferred' qualifier.*

We can now begin the second step in our project, namely, the construction of knowledge about the specific microstate ms_G generated by the operation G.

3.3.3.1. *The general problems.* The microstate ms_G cannot be observed as it emerges from the operation G, so it has to be made to trigger some phenomena that can be observed through the human senses. This can only be done by means of some macroscopic apparatus that can interact with the generated microstate ms_G.

In general, however, this interaction will *change* the initial microstate ms_G.

Furthermore, the observable phenomena produced by an interaction between a replica of ms_G and some macroscopic apparatus consist purely of some observable (visible, audible and so on) marks, which are displayed by the *registering devices on the apparatus*, and not by 'ms_G' itself. The only way we can think of these marks is that they are the results of interactions between the microstate and the apparatus, and these results are then *transferred* to the registering devices on the apparatus.

Now, the observable marks resulting from an interaction between a replica of the microstate ms_G and some macroscopic apparatus never trigger in the observer's mind some *qualia* enabling us to 'feel', in a direct way, the *nature* of the qualifier for which the apparatus has been designed to register a qualitative or numerical 'value' (unlike, for example, what happens when 'red' is perceived and is directly felt to belong to the category of qualia called 'colour'). Therefore, determining what is signified by the transferred registered phenomena given in terms of a value of a given qualifying quantity has to be entirely *constructed* in some conceptual/operational way, and this is a far from trivial task.

The mathematical formalism of quantum mechanics was developed specifically as a *mechanics* applicable to microstates. Within this formalism, the classical mathematical definition of each mechanical quantity X_M has been re-expressed, again mathematically, through a formal extension of the classical definition. However, both the classical definition *and* its extension involve some *model* of the classical concept of a 'moving body'. Furthermore, within *infra*-quantum mechanics, unlike the case for fundamental quantum mechanics, as we will stress later, we have deliberately excluded mathematical representations as well as models so that we can discuss the consequences entailed exclusively by the cognitive conditions and the general ways humans form concepts that are involved when we try to construct knowledge about what we have called microstates. Nevertheless, we want to construct a representation of 'knowledge' for microstates that, though strictly qualitative in character and completely lacking any model, can in the end be compared with the mathematical representations used in quantum mechanics. So, infra-quantum mechanics must somehow allow us to refer, when desired, to 'mechanical' quantities as some sort of special case. We will now present a very brief sketch† of how this can be achieved without violating the strict conditions we have imposed on our approach.

Consider a 'test' operation **X** that is realisable on a microstate through the use of some macroscopic apparatus $A(\mathbf{X})$, and each realisation of which ends up in a 'transfer' to the registering devices of $A(\mathbf{X})$ of a set of marks $\{\mu_X\}$ that can be directly perceived by the human senses‡. We call the set of *all* such sets of transferred observable marks under consideration the *spectrum of factual data* corresponding to the test operation **X**.

We will now assume, based on conceptual and historical facts, that what we have called a microstate ms_G is such that for any *mechanical* quantity X that has been defined within classical mechanics, there exists at least one test operation $\mathbf{X}(X)$ that, *in some definite sense*, 'corresponds' to the classical mechanical quantity X, so that it can be regarded as an operational translation of this classical mechanical quantity that can be applied to microstates: this hypothesis is *only* concerned with the *existence* of an established connection between the mechanical quantity X and the test $\mathbf{X}(X)$, which is not specified any further here. Nevertheless, on the basis of this minimal existence hypothesis, the symbol $\mathbf{X}(X)$ makes us think of a previously defined mechanical quantity X. On the basis of this hypothesis, we can consider saying that $\mathbf{X}(X)$ is a *mechanical test*, and that $\mathbf{X}(X)$ can be thought of as representing a 'measurement'-interaction $M(X)$ for which the result indicates a *numerical* value X_j of the mechanical quantity X. But for this to be useful, it must also be associated with a coding rule that translates any set of observable marks $\{\mu_X\}$ produced by one realisation of the test $\mathbf{X}(X)$ on ms_G into *one* definite *numerical* value X_j from a set $\{X_j\}, j \in J$, of possible numerical values of X_j assigned to the quantity X tied to the test operation $\mathbf{X}(X)$ (here, the index set J is discrete and *finite* by construction to ensure effectiveness).

Hence, finally, *if and only if* we can actually produce an appropriate conceptual/operational/methodological construct realising such a coding, we shall indeed be able to write that $\mathbf{X}(X) \equiv M(X)$, and in that case, the finite set of all the possible numerical values X_j obtained through the numerical coding of the sets $\{\mu_X\}$ of observable marks that can be produced by the measurement interaction $M(X)$ will be called *the spectrum of the mechanical quantity X attached to test operation* $\mathbf{X}(X)$. Correspondingly, $A(X)$ will be regarded as an 'apparatus for measuring X'.

If we now suppose that all the above requirements are satisfied, then, although we have worked exclusively within the constraints entailed by the cognitive situation and the ways humans form concepts, the process of constructing a strictly qualitative consensual

† The solution to this very important problem is treated rigorously and in detail in Mugur Schächter (2011, pages 73–88) and Mugur Schächter (2013 points 2.3.2.2–2.3.2.4). The solution entails a result that, through a confrontation between infra-quantum mechanics and the mathematical formalism of quantum mechanics, is used in Mugur Schächter (2013) to do away with the central quantum-mechanical interpretation problem, namely, 'the problem of measurement'.

‡ In order to ensure effectiveness, we assume that the number of distinct sets $\{\mu_X\}$ is finite. This is the only practically realisable assumption we can make since any numerical estimate performed on these marks, even if only concerning their space–time location, introduces *units*, and thus discreteness. Furthermore, we are always confined to a finite number of tests.

knowledge about microstates, which we call infra-quantum *mechanics*, seems possible, and we can continue with this as a basis.

However, the central condition of an unambiguous numerical coding of the observable data produced by the test operation $\mathbf{X}(X)$ immediately raises a new obstacle:

> The assumption of transpositions of classical mechanical properties such as 'position' and 'momentum' that can be applied to microstates necessarily involves some model of a microstate, even if it is only a very vague model.

In the absence of *any* such model, *or any qualia* indicated by the sets $\{\mu_X\}$ of observable data, there can be no conceivable connection whatsoever between the 'mechanical' description of a microstate and the classical mechanical descriptions achieved through qualifying quantities that have been extracted by abstraction from the qualia carried by the directly observable motions of macroscopic bodies. This means that we could not justify any assertion that some given sort of measurement interaction $M(X) \equiv \mathbf{X}(X)$ corresponds precisely to some particular classical mechanical quantity X. And, indeed, a careful examination shows that, contrary to the current orthodoxy, as asserted by Bohr and others, that quantum mechanics is free from any model, de Broglie's 'wave–corpuscle' model has remained implicitly, but quite organically, incorporated in the quantum-mechanical mathematical algorithms used to represent an act of measurement (Mugur Schächter 2011, pages 77–80).

This organic connection between the definability of a measurement interaction $M(X)$ and a model of a microstate appears at first sight to be an insuperable obstacle for an approach that not only forbids mathematical representations, but also any model attached to the general concept of a microstate.

However, we have been able to get around this difficulty by using a *general frame-condition*, which enables us to code any set of observed marks using, *exclusively*, the space–time locations of the marks. In this way, we are able to tolerate a *void* at the core of our approach in the specification of the semantic contents of the observable marks produced by the measurement interactions $M(X)$ tied to a test operation $\mathbf{X}(X)$ (Mugur Schächter 2011, pages 81–87). As in the case of the measurement interactions themselves, the semantic specifications of the results of these measurement interactions are only assumed to exist in some sense and on some level of conceptualisation that will be specified later.

> The coding of the marks according to the general frame-condition mentioned above only *distinguishes* any two given sets of observable marks from each other, without specifying any semantic content, or, *a fortiori*, any numerical value.

This is sufficient for us to continue the construction.

However, in order to justify a particular mechanical *name*, such as a 'position measurement', 'momentum measurement' or 'total energy measurement', we must *inevitably* assume some model of a microstate in the same way as in quantum mechanics, even if it is done there in a non-declared and implicit way. However, we stress that:

> An acceptance of the existence of a model of a microstate, or even a specification of such a model, does *not* prevent the measurement processes from having a purely *transferred* character: quantum-mechanical measurements *are*, quite

strictly, transfer-measurements in the sense that they only produce marks that are observable on a registering device and are devoid of any qualia relatable in a definite way to the studied microstate.

Having said this, in order to achieve direct comparability with the formalism of fundamental quantum mechanics, we will consider mechanical tests $\mathbf{X}(X) \equiv M(X)$ tied to 'measurement' operations and leading to numerical values of mechanical quantities.

3.3.3.2. *The emergence of a 'primordially' statistical level of conceptualisation.* We now ask whether a numerical value X_j that codes for a set of transferred observable marks $\{\mu_X\}$ produced by a measurement interaction $M(X)$ of the kind characterised above can be thought of as qualifying the microstate itself.

The answer is obviously no. According to our general causal conceptualisation structures, the measurement interaction $M(X)$ must, in general, be thought of as changing the microstate ms_G that was initially created by the generating operation G (this is for reasons similar to those that obliged us earlier to assume that whatever is produced by a given generating operation G is somehow relative to G). So, the observable transferred marks must also be thought of as emerging relative to the mentioned change, and thus also relative to the sort of measurement interaction $M(X)$ being used. It follows that the transferred marks can *only* give a *global* characterisation of the measurement interaction, and not of the (hypothetical) microstate ms_G *separately.*

However, we can still cling on to the fact that the observable marks are *also* relative to the initially created microstate ms_G. So we have to take into account the fact that the initially produced object microstate ms_G is subject to two clearly distinct processes of change, which correspond to two clearly distinct measurement interactions $M(X)$ and $M(X')$ realised using two distinct pieces of apparatus $A(X)$ and $A(X')$ that are tied to two different mechanical quantities X and X', and that, in general, cover *two different space–time domains.* Now, when this happens, the corresponding measurement interactions $M(X)$ and $M(X')$ cannot both be achieved simultaneously for a *single* replica of a microstate ms_G. So, in *this* sense, these two measurement interactions are *mutually incompatible*†.

Furthermore, in general, a measurement evolution may destroy the microstate ms_G initially produced by the corresponding generating operation G.

It follows that if we want to obtain observable qualifiers for the microstate ms_G in terms of the values of *both* of the quantities X and X', in general, we have to generate *more* than one replica of ms_G because we have to achieve two different types of sequence of the form

$$[G.M(X)] \cong [(\text{a given operation } G \text{ generating a microstate } ms_G), \\ (\text{a measurement interaction on } ms_G)]$$

† The restriction to *one* replica of the considered microstate ms_G is not explicitly required in current presentations of the quantum-mechanical concepts of incompatibility and complementarity, though these concepts do involve it quite essentially.

specifically, the sequences $[G.M(X)]$ and $[G.M(X')]$ (the clock is re-set at the same initial time-value t_0 for each realisation of a sequence of this kind).

Furthermore, if we try to 'verify' the result by repeating the measurement interaction for just *one* quantity X with a given microstate ms_G through the corresponding sequence $[G.M(X)]$, we do *not*, in general, systematically get the *same* value X_j. If this does happen (that is, we always get the same value) for some given quantity X, then it does *not* happen for a quantity X' that is incompatible with X in the sense defined above: *this is a basic empirically observed fact.* So, in general, the results are distributed over the whole spectrum $\{X_j\}, j \in J$, of possible values of X_j of the quantity X tied to the test operation $\mathbf{X}(X)$ (where J is a discrete index set).

This means that the global observational situation that emerges from measurement interactions with microstates is essentially *statistical.* And the nature of this statistical character is 'primordial' in the sense that it marks the *very first* sort of knowledge that can be generated about microstates (see Mugur Schächter (2002c) in connection with Longo (2002)), and thus about matter. Therefore, at this primordial level of conceptualisation, the statistical character cannot be attributed to mere ignorance of some more basic conceptualisation that might have been achievable previously in individual deterministic terms, at least in principle (this assumption is always made in classical thinking with respect to any sort of statistical data). It is only through explicit models that might be constructible in the future on some *higher* level of conceptualisation than the one at which the primordially statistical transferred descriptions emerge that a fully non-statistical description of a particular microstate could be worked out.

The *first* level (chronologically) in the sequence of conceptualisation levels, when starting from aconceptual physical reality, has a non-removable, essentially primordial character.

3.3.3.3. *The peculiar descriptional form tied to primordially statistical transferred microstate qualifiers.* The previous section showed that the sort of *stability* that can be observed for a microstate (one that can be investigated at the primordial level of the conceptualisation of microstates) can only be statistical. At this primordial level, we can only study a descriptional invariant through repetition of the same sequence $[G.M(X)]$ for each given pair (G, X).

But exactly what sort of invariant can this be? It is tempting to give as a first answer:

> It is a probabilistic invariant, with a probability 'law' $p(G, X), j \in J$, tied to the pair (G, X).

But this just brings us back to the problem described in Section 2 concerning the *absence* of a factual definition of the 'probability law' to be asserted in a given factual 'probabilistic situation', in other words, a definition that is *independent* of the definition given by the weak law of large numbers in expression (2) on Page 8, which we showed to be non-effective and indefinitely recursive.

> Kolmogorov's aporia, which emerged within classical thinking, appears in its *most basic* manifestation when we try to make use of the concept of probability in connection with the study of microstates.

And it may turn out that the concept of probability cannot be constructed *within* the first level of conceptualisation, where the descriptions of microstates are originally worked out†.

So, for the moment, *all* we can factually achieve is to realise, for each pair (G, X) being studied, a finite number q of series of N repetitions of the corresponding sequence $[G.M(X)]$, with N successively taking on values drawn from some finite collection of increasingly large numbers $N_1, N_2, \dots N_k \dots N_q$, and then determine whether there is any tendency towards convergence for the relative frequencies of the corresponding sets

$$\{n(G, X_j)/N\}, \quad j \in J, \; N = N_1, N_2, \dots N_k \dots N_q.$$

Nothing guarantees, a priori, the existence of such a convergence: it is not a logical necessity, and if no convergence were found, we would be obliged to give up finally on our aim of constructing some stable observable knowledge about microstates.

In fact, it turns out that for any pair (G, X), *there is* a tendency towards convergence; of course, it is a fluctuating convergence while the integer N_q is kept definite, finite and effective. In these conditions, and assuming a restriction to effective procedures and the absence so far of any general procedure for constructing the factual probability law for a given factual situation, we can only postulate a specification for such a law. For instance, we could postulate that the relative frequencies from the set $\{n(G, X_j)/N\}$, $j \in J$, measured for the longest series of repetitions Nq of the sequence $[G.M(X)]$ will be assimilated *by convention* to the unknown factual numerical distribution of individual probabilities. This just amounts to *deciding* to write

$$\{n(G, X_j)/Nq\} \cong \{p(G, X_j)\}, \quad j \in J,$$

that is, to assigning to the ratio $n(G, X_j)/Nq, j \in J$, the role played by $p(e_j)$ in the weak law of large numbers. In this way, we introduce a sort of '*pre-probabilistic* knowledge' about the microstate ms_G, which is simply founded on a factually observed *tendency* towards convergence. This knowledge, under the cover of the dense cloud of confusion surrounding the concept of probability, is then treated by a hidden convention as a piece of factual 'probabilistic' knowledge‡. Specifically, in this case, there is a pre-probabilistic qualifier, which has *a non-removable relativity to the triad* $(G, ms_G, M(X))$. We will call this sort of transferred pre-probabilistic effective qualifier, involving some conventional choice, *the transferred description of the microstate* $ms_G \leftrightarrow G$ *through the qualifying 'mechanical' transfer-view* $V_M(X)$, and it will be denoted by the symbol

$$D/G, ms_G, V_M(X)/.$$

† Within mathematical quantum theory, it is largely accepted, more or less explicitly, that the *mathematical formalism* involves the possibility of determining the *probability* law (and not just some statistical distribution) corresponding to *any* factual situation concerning a microstate. Historically, this view stems from what is called 'Born's algorithm', and possibly also from Gleason's theorem on 'probability' measures in a Hilbert space (Gleason 1957). This view, which the current author does not share, will be discussed briefly at the end of the current paper.

‡ Indeed, something precisely like this is currently carried out systematically in any classical probabilistic situation, though often implicitly.

This symbol is written in a way that explicitly recalls the origins of the description and the relativities it involves (the 'mechanical' transfer view $V_M(X)$ is just a new name and notation for the measurement interaction $M(X)$). But we stress again that the description itself, the global qualification we have obtained, is just the partially conventional pre-probability law

$$\{n(G, X_j)/Nq\} \cong \{p(G, X_j)\}, \quad j \in J$$

introduced above.

3.3.3.4. *Characterisation of a microstate.* We have already noted that the strategy imposed by the cognitive situation we find ourselves in while constructing knowledge about microstates has led to qualifiers that must be postulated to *involve* the microstate, but cannot be assigned to it alone, in isolation. This may already seem to violate the classical concept of a description. So we will now investigate whether perhaps at least the peculiar sort of knowledge we have just constructed and denoted by the symbol $D/G, ms_G, V_M(X)/$ could be considered to be characteristic of the microstate ms_G, that is, whether it can be considered to apply exclusively to the microstate ms_G. The general answer to this question is no: if, as conceived above, $V_M(X)$ introduces only one quantity X, there can be no reason to assert that the same pre-probability law $\{p(G, X_j))\}, j \in J$, we have found for the microstate ms_G, and thus for the pair (G, X_j), could never also arise for another pair (G', X) with $G' \neq G$, but the same qualifying quantity X.

However, if we consider the case where the mechanical view V_M introduces two *mutually incompatible* measurement interactions $M(X)$ and $M(X')$ carried out on *different* replicas of a *single* microstate ms_G, then it seems safe enough for us to consider that these two distinct measurement interactions act like two distinct 'qualifier dimensions', which, in combination, through some sort of 'intersection', determine a characterisation of ms_G; that is, that no other generating operation different from G can generate a microstate for which exactly the same pair of pre-probability laws as those obtained for $ms_G \leftrightarrow G$ with $M(X)$ and $M(X')$ emerges. This is even more likely to be the case if all mutually incompatible pairs (G, X) are considered, where X runs over all the mechanical quantities redefined for a microstate: the set of all the pre-probability laws $p(G, X)$ corresponding to all these mutually incompatible pairs can be quite safely considered to express a specificity of the studied microstate ms_G. So we are led to introduce a general concept of a view V defined as a union of aspects, which are specifically mechanical aspects in our case. We will denote this general concept of a mechanical view by V_M and call it *the global mechanical qualifying view defined for microstates*, which consists of the union $V_M = \cup V_M(X)$ with X running over all the qualifying mechanical quantities defined for microstates. Hence, *'the' pre-probabilistic transferred mechanical description of the microstate ms_G* (note the use of the definite article 'the') can be denoted by the symbol $D/G, ms_G, V_M/$.

In this way, the initial descriptional form

$$D/G, ms_G, V(X)/,$$

which could not be considered to fully characterise a given microstate, has been completed into a relativised description $D/G, ms_G, V_M/$ through which we can achieve such a characterisation.

We can now conclude by saying that a transferred description of a microstate ms_G consists, exclusively, of a set of one or more partially conventional 'pre-probability distributions' on sets of observable marks $\{\mu_X\}$ transferred to (in general) *various* registering devices belonging to various pieces of apparatus, and expressed through definite coding rules in terms of the values X_j from the spectra of qualifying mechanical quantities X.

> Such a description asserts strictly nothing about how the microstate ms_G 'is' itself, nor even where or when it 'is'.

So, in order to create knowledge about microstates, we have made use of the pre-existing general features of our human conceptualisation, and these, quite fundamentally, have involved, in particular, the acceptance of causality (when we postulated the relativity of ms_G to G, and the relativity of the observable marks $\{\mu_X\}$ to both ms_G and $M(X)$). Despite all this, the final result is a descriptional form $D/G, ms_G, V_M/$ *that does not even assign a connected space–time support to the microstate* ms_G *through its transferred description.* This break with classical thinking has been produced progressively, through inevitable steps that have been required by the conditions successively imposed in order to define the entity-to-be-described and the way this entity can be qualified: in other words, in order to describe it, which is, in its turn, strictly synonymous with creating communicable knowledge about this entity.

The absence of a definite and connected space–time support *for a microstate* (not the transferred observable marks tied to it), together with the required coding of these marks stripped of any semantic content tied to the described microstate, makes the concept of a transferred description $D/G, ms_G, V_M/$ completely non-classical and *unintelligible.* In this way, this sort of description triggers a strong need for qualifier values involving qualia that can be thought of as being 'possessed' by ms_G within some connected space support covered by ms_G. But for this sort of 'explanation', the primary transferred description of a microstate requires some explicit and declared model of a microstate, and this requirement takes us beyond infra-quantum mechanics. It also takes us beyond fundamental quantum mechanics. However, once we have left these primordial representations, there is nothing to prevent us from constructing such a model†.

3.3.3.5. *The global space–time tree-like structure of the transferred description of a microstate.* We will now return to the mutual incompatibility of the evolutions of two measurement processes that cover different space–time domains, which thus prevent a simultaneous realisation using a *single* replica of the microstate ms_G. Such mutual space–time incompatibilities mean that the set of *all* the physical sequences $[G.M(X)]$ that involve the *same* generating operation G is divided into subsets of *mutually incompatible classes* of *mutually compatible sequences* $[G.M(X)]$. This, by a 'geometrising'

† The so-called 'impossibility' theorems that claimed to eliminate the constructibility of any such normal, causal, space–time model were invalidated by the current author in Mugur Schächter (1964) and Mugur Schächter (1979).

process of integration, produces a *new* type of pre-probabilistic structure in the form of a *tree-like space–time structure* founded on a common 'trunk' corresponding to the space–time domain $d_G(t_G - t_o)$ covered by the realisations of the generating operation G, and possessing as many measurement interaction 'branches' as there are mutually incompatible classes of mutually compatible operations of the type considered, each branch covering a specific space–time domain and generating at its top a corresponding Kolmogorov-type pre-probability space†. We shall call this structure *the pre-probability tree of the pair* (G, V_M) *and denote it using the symbol* $T(G, V_M)$. Figure 1 shows an example with two branches corresponding to two quantities denoted by $X \cong B$ and $X \cong C$ for simplicity, with the two pre-probability spaces

$$[(C1, C2, C3, \dots Ck, \dots Cn), p(G, C)]$$

and

$$[(B1, B2, B3, \dots .Bj, \dots Bm), p(G, B)],$$

respectively, at the top of each of them. (For simplicity, the algebra on the universes of elementary events $(C1, C2, C3, \dots Ck, \dots Cn)$ is omitted, and the pre-probability law $\{p(G, C)\}$ is defined directly on the universe of elementary events; *mutatis mutandis* the same applies for the universe $(B1, B2, B3, \dots Bj, \dots Bm)$ and $p(G, B)$.)

3.3.4. *The need for a deeper and extended general theory of probabilities.*
We have shown elsewhere that the qualitative descriptional form $D/G, ms_G, V_M/$ with the tree-like space–time structure $T(G, V_M)$ encapsulating its complete integrated, 'geometrised' genesis introduces a number of features that *go beyond* Kolmogorov's classical concept of a probability space. It does this quite essentially and in several important ways‡, specifically, with respect to:

— the full *representation* of the structure of the random phenomenon being considered;
— a meta-'probabilistic' dependence between the events of the mutually *incompatible* probability spaces at the top of the branches (which involves accepting a specific mathematical representation that is new within the general concept of correlated probability spaces);
— a pre-organised sensitivity to the logical aspects of the set of all the elementary events and events involved, whether compatible or incompatible, which it turns out are *not* expressible through a lattice structure.

In this way, the concept of a pre-probability tree of the pair (G, V_M) requires an extended and deeper concept of 'probability' *unified* with a corresponding logic of all the events involved.

† These space–time specifications arising from the 'geometrising' integration of the *genesis* of a transferred description $D/G, ms_G, V_M/$ do not in the least alter the fact that there is no *intrinsic* space–time specification of the microstate ms_G *itself*.
‡ A detailed examination of the concept of the pre-probability-tree of a microstate produces several deeply non-classical results – see Mugur Schächter (2011, pages 119–131).

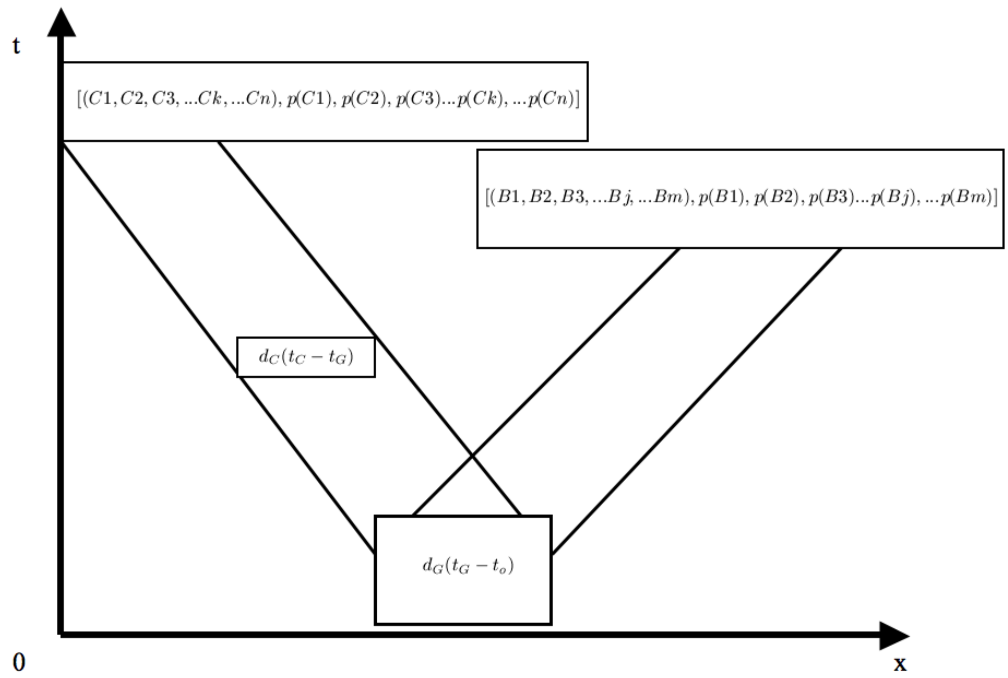

Figure 1. Pre-probability tree $T(G, V_M(B, C))$ of a pair $(G, V_M(B, C))$ where $V_M(B, C)$ involves two mechanical aspects B and C only.

This programme was partially achieved in the papers Mugur Schächter (1992a; 2002a; 2002b; 2006) in terms that are *generalised* to any relative description†.

3.3.5. *Conclusions from the discussion of infra-quantum mechanics (IQM)*
The descriptional form $D/G, ms_G, V_M/$ together with the geometrised, integrated tree-like space–time structure encapsulating its genesis, along with the consequences of this structure, lies at the heart of the strictly qualitative, physical/epistemological type of representation of microstates that we have constructed *independently* of the mathematical formalism of quantum mechanics. We have called this *infra-quantum mechanics*, or *IQM* for short‡.

In the current paper, we have only given a brief sketch of infra-quantum mechanics in an extremely simplified form. But when developed in full detail, it sheds light on the whole way in which the quantum theory manages to signify. In particular, it allows us to separate what has been introduced into quantum mechanics by an extension of classical models aimed specifically at the construction of a *mechanics* of microstates

† In the current paper, we have tried to ensure an *unrestricted* generalisation, but only within the classical domain. The problems concerned specifically with the primordial domain of conceptualisation on which the basic descriptions of microstates is constructed will be set out clearly in Section 8.
‡ We use the prefix 'infra' here to mean lying beneath the mathematical formalism, and partially encrypted in it.

from what has been induced purely on the basis of the cognitive situation, the general requirements of the ways humans form concepts and the aim of constructing knowledge about microstates.

It is striking that, notwithstanding the elements of classical models encrypted in its formalism, the actual descriptions involved in the mathematical quantum theory are *transferred* descriptions of the form $D/G, ms_G, V_M/$, which are identified within IQM and involve the tree-like space–time structure $T(G, V_M)$. These descriptions are in no way concerned overtly with the model of a microstate implied by the formalism, but are explicitly only connected with the observable marks produced by the measurement interactions. The model has been digested by the formalism, and assimilated within it. This epistemological schizophrenia exhibited by the quantum-mechanical formalism fuels its capacity to offer formal definitions of *mechanical* quantities (through eigenstate equations and the corresponding eigenvalues), and then to imagine corresponding adequate measurement operations. On the other hand, these processes, which result from a disconnected and hidden model, are strongly related to the unintelligible character of the formalism.

A systematic comparison between infra-quantum mechanics and the mathematical formalism of quantum mechanics should now allow us to deal in a *unified and coherent* manner with *all* the interpretation problems, and to achieve an organised solution of them. But this goes beyond the scope of the current paper, which only aims to identify in an effective manner the factual probability law to be asserted in any given factual probabilistic situation.

We showed in Section 2 and the earlier part of Section 3 that the problem of defining an effective factual probability law in any given probabilistic situation remains at least as open in the case of the primordially statistical transferred descriptions of microstates, which form the basis of all current physical knowledge, as in the classical domain of probabilistic thinking. And it is precisely through a generalisation of the descriptional form $D/G, ms_G, V_M/$ (which was constructed within infra-quantum mechanics for the particular case of microstates) that we shall be able to propose a solution to Kolmogorov's aporia. However, in order to introduce this generalisation, we shall first need to consider a certain universal property that is entailed by the descriptional form $D/G, ms_G, V_M/$.

3.4. *Universality and perception of the possibility of a general method of relative conceptualisation*

We have seen that in order to achieve a transferred description $D/G, ms_G, V_M/$ of a microstate, we need to:

(a) Determine the physical epistemic operation denoted by G that introduces a corresponding entity-to-be-described ms_G *independently* (in general) of any epistemic action through which this entity could be qualified.

(b) Determine the measurement interactions $M(X)$ that lead to qualifications of the entity ms_G.

(c) Realise both operations G and $M(X)$ in a new and fundamentally constructive way by *first* generating physically in space–time a new entity-to-be-described that did not already exist (as opposed to just *selecting* it from a collection of already available physical objects) and then, *afterwards*, generating, again physically, observable manifestations of ms_G (rather than just *detecting* 'properties' that are assumed to already exist and be 'possessed' by this entity).

(d) Realise the sequence $[G.M(X)]$ a large number of times for each quantity X involved in the view V_M used so that we can try to find (at the level of observable manifestations of ms_G, which inevitably have a statistical character) a set of *invariants* that constitute a sufficiently stable qualification that is characteristic of ms_G.

Steps (a)–(d) summarise a maximally *explicit* and *creative* way of producing descriptions in which all the relativities involved are made apparent in turn, and are active and obvious. Hence, the resulting final description $D/G, ms_G, V_M/$ is explicitly relative to each of the elements of the triad $(G, ms_G, V_M)^{\dagger}$.

> We believe that the descriptional form $D/G, ms_G, V_M/$ (with its inherent relativities, its development dominated by methodological decisions, and its epistemological consequences (Mugur Schächter 2002a; 2002b; 2006; 2011; 2013)) constitutes a crucial insight into the way humans generate knowledge.

It is very important to realise that the degree of explicitness and constructivity that characterises our development of the descriptional form $D/G, ms_G, V_M/$ is ignored in most current classical conceptualisations, as reflected in natural languages, classical logic, classical probabilities and classical physical theories, including Einstein's relativistic theories. In classical conceptualisations, it has always been possible to suppose, though usually implicitly, that the entities-to-be-described *pre-exist* the descriptional process and are 'defined' in advance by 'properties' that these entities 'possess' *intrinsically*, independently of any act of observation and in an already *actualised* way. Before the peculiar aim of describing microstates had been conceived, these assumptions had never led to any notable difficulties. Therefore, classically, a description is thought of as consisting exclusively in the *'detection'* of one or more of the 'properties' 'possessed' by the entity-to-be-described, which itself, is assumed to pre-exist either as an 'object' in the usual sense, or as a 'situation' or an event, and so on.

> Classically, no consideration at all is given to the question of how an entity-to-be-described is introduced.

As a result, the very deep consequences of the way in which an entity-to-be-described is *generated* are almost systematically ignored‡.

† It might seem at first sight that the relativity to ms_G can be absorbed within the relativity to G, but this is not the case: the results of the sequences $[G.M(X)]$ depend explicitly on ms_G and *cannot be derived from G*.

‡ This is even the case in fundamental quantum mechanics, where, for linguistic reasons, many physicists erroneously identify the operation G of generating a microstate with what is called the 'preparation of the state vector for registering the result for an eigenvalue', which, in fact, is involved exclusively in the operation of qualifying that microstate through *measurement* interactions. This means assuming that, like a classical 'object', the microstate to be qualified is already there. In any case, the operations that generate microstates are certainly not mathematically represented within the formalism; they are not even assigned a symbol.

Moreover, the process of examination that qualifies this entity is reduced to a single, simple, static act of detection. And this final classical reduction is the source of what are currently the most explicitly stated differences between the logic and probabilities involved in the descriptions of microstates and classical logic and probabilities.

However, it is noteworthy that, although the descriptional form $D/G, ms_G, V_M/$ does not appear in classical logic and probabilities, which form the two most fundamental classical syntactical structures, it is, nevertheless, quite obviously involved in many *current* classical epistemic procedures. Indeed, once the peculiar and very difficult cognitive situation dealt with in describing microstates has been fully appreciated, as well as the descriptional strategy that has allowed us to overcome this difficult cognitive situation, there is a sudden inversion in our perception of the issues, which is similar to the way our interpretation of a gestalt optical illusion can flip from one version to another. What at first sight had seemed to be fundamentally new and surprising in the form $D/G, ms_G, V_M/$, now suddenly appears to be endowed with a certain universality, even normality. Indeed we can immediately say that:

— Any explicit and complete account of a given process of description *must* include a specification of the action through which the entity-to-be-described is introduced, as well as a specification of the operation, physical, abstract or both, that is used to qualify this entity.
— These two actions are often mutually independent.
— The introduction of the entity-to-be-described is sometimes achieved by *creation* of this entity, while the qualifying operation, if it is a *physical* process, *always*, at least in principle, *changes* the object-entity, and sometimes it changes it radically. Hence, the consequences of the relativity to either or both of these basic epistemic actions for the resulting description must be explicitly taken into account and thoroughly analysed.

For instance, consider the actions of a detective searching for material evidence related to a crime. He usually focuses his attention on a relevant location, say the scene of the crime, and there he first extracts some samples (he cuts out fragments of cloth, he detaches a clot of coagulated blood, and so on); he might even create a complete test situation involving the suspects so that he can record their behaviour. Only afterwards does he examine the samples he has collected or the behaviours recorded during the test situation. Other examples include taking a biopsy for a medical diagnosis or extracting samples of rock using a robot on the surface of another planet, and the subsequent examination of these entities-to-be-described.

In all these cases, the investigator *generates*, completely or in part, an entity-to-be-described that did not pre-exist, in order to qualify it later using operations that are quite independent of the operation used to generate it. And in certain cases, the operation used to examine the entity-to-be-described changes it so radically that if several different examinations are required, a number of 'replicas' must be produced, which are then just *assumed* to be 'identical' to each other. Furthermore, the resulting qualifications are permanently affected by two quite distinct relativities: relativity to the way the entity-to-be-described was generated and relativity to the sort of examination carried out. Even the concept of relative existence or non-existence arises: the way in which the entity-to-be-described has been generated may simply exclude certain subsequent examinations.

These considerations lead to the following observations.

The domain and nature of communicable knowledge in classical thinking are misleadingly reduced. The whole primordial zone of conceptualisation where mind *actively constructs* the very first forms of a radically new communicable knowledge, from pure physical factuality, is so deep-set that it has remained hidden beneath the two basic building blocks of all current Western languages: namely, subjects and predicates. These both suggest available pre-existing elements to be described. Furthermore, the primordial, and always fundamentally generative, zone of conceptualisation has even remained cut off from many classical *scientific* representations. Notwithstanding the well-known analyses of Husserl, Poincaré, Einstein, Piaget and many others, which have drawn attention to the crucial role played by physical operations in the most basic conceptualisation processes, classical logic and probabilities, as well as the theory of sets, take *language* as their starting point and are, again, developed almost exclusively through the use of language. Physical operations are not considered, and factuality, through the medium of language, is widely supposed to *spontaneously* print 'information' about the already existing and real properties of pre-existing 'objects' on *passively* receptive minds. The *active* role, when it does arise, is assigned almost exclusively to the exterior reality, and not to the mind.

However, quantum mechanics, having led us through infra-quantum mechanics to the identification of the *basic, relativised descriptional form* $D/G, ms_G, V_M/$, has revealed the potentiality of a very deep-set, *general* and fundamentally operational *method of relativised conceptualisation*. Indeed, the descriptional form $D/G, ms_G, V_M/$ is paradigmatic. It encapsulates a particular embodiment of an extreme but universal epistemic situation. Namely, the situation arising when a communicable and consensual representation is investigated for some *non*-pre-existing physical entity for which, *a priori*, it is only its possible existence that can be conceived and labelled, and which, if it is generated, emerges in a non-observable state. In such extreme circumstances, we are *compelled* to adopt a fundamentally active, constructive approach, which is associated with a complete decomposition of the global process. This is because in such a situation, all the stages of the desired description have to be *built* from pure physical factuality, independently of each other, with each of them developed in full depth and extent. The strictness of these constraints has revealed the descriptional form $D/G, ms_G, V_M/$, which, although initially only concerned with the case of microstates, has a universal epistemological content that is so comprehensive and explicit that it strongly suggests the possibility of a generalisation capable of accommodating *any* form of description.

So a new target comes into focus, namely, to construct a general, consensual, canonical method of relativised conceptualisation, which we call MRC for short.

3.5. *Quick overview of the MRC approach*

3.5.1. *Preliminaries.*

It is almost systematically the case that false absolutes generate false problems and paradoxes that hinder understanding and prevent the development of knowledge. The history of thought is full of such examples. The specific goal of MRC is

> to offer a structured system of norms for conceptualising in a *relativised* way that excludes, by construction, any possibility of the emergence of false problems or paradoxes.

The seed that developed into MRC is the peculiar qualitative form of the primordial descriptions of microstates that the current author first perceived as lying *beneath* the mathematical formalism of fundamental quantum mechanics, and then constructed explicitly, quite independently of the formalism, in the epistemological/physical discipline that became *infra-quantum mechanics* (Mugur Schächter 2011).

The construction of MRC began *ab initio*, and long *before* the explicit construction of infra-quantum mechanics, where the primordial microstate descriptions mentioned above can be thought about and understood. This general method was developed in a *deductive* way, in the sense of everyday (non-formalised) logic. The peculiar form of transferred descriptions re-emerged within MRC, but only when it was at a rather advanced stage, and then directly with the status of fully general validity. The epistemological strategy perceived more or less implicitly beneath the mathematical formalism of quantum mechanics acted as a guide. Then, once constructed, MRC itself guided the explicit construction of infra-quantum mechanics.

The systematic relativisations introduced successively through the development of descriptions, from the zero-point of conceptualisation, in the form of a basic transferred description, to a more complete conceptualisation, no matter how complex, provide protection against *any* surreptitious insertion of false absolutes. And, remarkably, these relativisations reproduce, fractal-like, the same recurring basic descriptional *form*, written symbolically as $D/G, ms_G, V_M/$, at each point along these conceptualisation paths. So MRC generates hierarchical *chains* of mutually connected relativised descriptions, each having the form $D/G, ms_G, V_M/$.

These chains meet in node-descriptions and form descriptional nets.

In particular, MRC has generated a relativised reconstruction of natural logic, probabilistic conceptualisation and informational conceptualisation, and it has led to a representation of 'complexities' where the semantic content is fully preserved. It has also enabled a representation of 'time' that is drawn from atemporal elements. For further details, see Mugur Schächter (2006).

For reasons of space, we will only be able to *enumerate* the main concepts of MRC here, without showing the interconnections or giving much in the way of explanations. However, this will be sufficient for our purposes in plotting a path towards solving Kolmogorov's aporia. Unfortunately, this may give an impression of some arbitrariness since we will not be able to explain the semantic and logical considerations and requirements that drove the development of MRC along an inevitable path and unite its elements into an organic whole. Furthermore, the semantic content will be dispersed through the presentation, which hides the flow of its growth during the construction. For a fuller appreciation of how MRC has been developed into a coherent and rigorous whole combining factual content and rationality, see Mugur Schächter (2006), which is only available in French, but also Mugur Schächter (2002b), and perhaps even Mugur Schächter (2002a).

3.5.2. *Enumeration of basic MRC concepts.*

(1) *Any MRC description is explicitly relative to a given triad* $(G, œ_G, V)$ where:

  (a) G denotes the *generating operation* (which may be physical or abstract, or consist of some combination of physical and abstract operational elements) by which the entity-to-be-described is made available to be qualified. The specification of G is required to include an explicit indication of the domain of reality R_G on which G is applied.

  (b) $œ_G$ denotes the *entity-to-be-described* and introduced by G. This entity may or may *not* be directly perceptible.

  (c) We postulate a one-to-one relation $G \leftrightarrow œ_G$ between the generating operation G and the entity-to-be-described $œ_G$ it introduces.
  This relation is not a fact, it is a methodological assumption.
  Very careful analyses have shown that this postulate cannot be avoided and entails major conceptual consequences – see Mugur Schächter (2006, pages 61–66 and 213–221) and Mugur Schächter (2011).

  (d) V denotes the *view* through which the object-entity is qualified.

(2) The description relative to a *given* triad $(G, œ_G, V)$ is denoted by the *symbol*

$$D/G, œ_G, V_M/$$

where *that particular* triad is introduced.

(3) By definition, any view V is endowed with the following strictly prescribed structure:

  (a) A *view* V is a *finite* set of *aspect-views* Vg where g is an aspect index[†]:

$$V = \bigcup_g Vg, \qquad g = 1, 2 \ldots m,$$

  with m a finite integer.

  (b) An aspect-view Vg (an aspect g for short) is *a semantic dimension of qualification* (such as colour or mass) that can carry any *finite*[‡] set of ‘values’ $gk(g)$[§] of the aspect g that we wish to consider. (For instance, for ‘colour’, we could choose to consider only those ‘values of colour’ indicated by the words ‘red’,

[†] The symbol g should not be confused with the symbol G for an operation generating an object-entity $œ_G$.

[‡] By construction, every counting or numerical character involved in MRC is finite since MRC has been developed as a strictly effective method. Within MRC, any sort of infinity can only be understood in terms of the relativised *absence* of an *a priori* limitation.

[§] We write $gk(g)$ so that we can distinguish between aspect-values of different aspects, and so that we can assign to the set of value-labels $k = 1, 2, \ldots$ cardinals $w(g)$ that depend on g.

'yellow' and 'green', and we would associate a reference sample with each of them.)

The symbol $gk(g)$ functions as a *unique* index different from g alone, but in a specific case the indexes g and $gk(g)$ may be replaced by any other pair of convenient signs.

An aspect-view Vg is defined *if and only if* we have also defined all the devices (instruments, pieces of apparatus) and all the material or abstract operations on which we base the assertion that an examination of a given object-entity through the aspect-view Vg has led to a particular unique and definite value $gk(g)$ of g (or *none*).

(c) A view V is a finite qualification *filter*.

A given view V is blind with respect to aspects or values of aspects that are not contained in its initial definition – it simply does not perceive them.

(d) The qualifications of space (E) and time (T) are achieved through a very particular type of *frame view* $V(ET)$, which can be reduced, if required, to a pure space-frame view $V(E)$ or a pure time-frame view $V(T)$.

The features listed above generate the concept of a 'qualifier', which is quite unlike the 'predicates' of classical formal logic and the grammars of natural languages.

(4) Given a pair (G, Vg), the two epistemic operators G and Vg may or may not '*mutually exist*':

(a) If any examination by Vg of the entity-to-be-described $œ_G$ introduced by the generator G produces a single well-defined result (gk), then the aspect-*value* (gk) of g does exist with respect to the operation G generating an object-entity $œ_G$, that is, there is *mutual existence* between G and (gk). Hence, *a fortiori*, there is also mutual existence between the *aspect* g itself and the generating operation G. In *this* case, the pair (G, Vg) constitutes a *one-aspect epistemic referential*. This means that, in this case, if we subject the object-entity $œ_G$ introduced by G to an examination by Vg (and thus produce the operational sequence $[G.Vg]$), we *might* obtain a corresponding 'description' of $œ_G$ through the qualifier grid introduced by the aspect-view Vg. This only happens if some *invariant* result emerges through repetitions of the sequence $[G.Vg]$: this result may be an individually invariant result, some statistical stability, or a 'probabilistically invariant' result (though determining the *factual* meaning of 'probabilistically invariant' is precisely what we still need to do in the rest of the current paper).

The mutual existence of a generating operation G of an entity-to-be-described $œ_G$ and an aspect-view Vg is the MRC expression of the fact that the aspect g has emerged by abstraction from a class of entities that $œ_G$ belongs to.

(b) If, however, an examination by Vg of the entity-to-be-described $œ_G$ yields no definite result, then there is '*mutual inexistence*' between Vg and $œ_G$ (in other words, $œ_G$ does not exist relatively to Vg, and *vice versa*).

For instance, a song does not exist with respect to the qualifier grid giving the intensity values of an electrical current recorded by an ammeter, and *vice versa*.

In this case, an initial tentative pairing (G, Vg) has to be rejected *a posteriori* as being incapable of generating a relative description $D/G, œ_G, V_M/$, and it is thus unable to signify from a descriptional point of view.
Mutual inexistence between $œ_G$ and Vg is the MRC expression of the fact that the entity $œ_G$ does *not* belong to the class of entities that have contributed to the construction of Vg by a process of abstraction. So:

> The concepts of mutual existence and mutual inexistence constitute the MRC expression of the fact that *a qualification can only be applied to entities that have participated in the genesis of the qualification* (individually or in combination with others).

(c) These considerations can be extended in an obvious way to any pair (G, V) where

$$V = \bigcup_g Vg, \qquad g = 1, 2 \ldots m,$$

contains a finite number m of aspect-views Vg. In this case, we refer to the possibility, or not, of an epistemic referential (G, V).

(5) The *space–time-frame principle.*
Consider a space–time view denoted by $V(ET)$. We call it a *space–time-frame view* because of the following principle, which only applies to *physical* object-entities:

> Any *physical* entity-to-be-described exists relatively to at least one aspect-view Vg that is *different* from any space–time-frame view $V(ET)$. However, it is non-existent with respect to any space–time-frame view $V(ET)$ considered *alone* and separate from any aspect-view Vg that is different from any space–time aspect ET.

In order to ensure a place to express the space–time-frame principle and its consequences, the view V in any epistemic referential (G, V) that can generate a description of a physical entity-to-be-described includes, by convention, a space–time-frame view $V(ET)$ and at least one aspect-view Vg different from any space–time aspect†. In particular $V(ET)$ can be reduced to a space-frame-aspect $V(E)$ exclusively.

(6) Consider a pair (G, Vg) where G and Vg mutually exist. Hence, the pairing (G, Vg) constitutes an epistemic referential where it is possible to construct the relative description $D/G, œ_G$, $Vg/$ of the entity-to-be-describe $œ_G$ produced by G. Then:

(a) If after some number of repetitions N of the sequence $[G.\,Vg]$‡ only *one* and the same value (gk) of the aspect G is obtained systematically, the corresponding relative description $D/G, œ_G$, $Vg/$ is said to be an '*N-individual*' *one-aspect*

† It is possible to construct infinitely many space–time-frame views through the choice of reference axes and origin, or differential geometric reference structure (such as Riemann geometry), or space and time units.

‡ In general, after a sequence $[G.\,Vg]$, the replica of the object-entity $œ_G$ involved in that sequence is either changed by the examination using Vg or it is destroyed (for example, it may be absorbed in a device). So, in general, *repetitions* of $[G.\,Vg]$ will also require repetitions of the generation operation G to create a new replica of $œ_G$.

description (or an 'individual description' relative to N repetitions of $[G.\,Vg]$, with N *finite*). So, within MRC, in order to include the case of entities to be described that are 'consumed' by an examination using Vg, an 'individual description' requires repetitions of the operational sequence $[G.\,Vg]$, and is relative to the number of these repetitions.

(b) If, however, the value (gk) varies, in general, from one realisation of the sequence $[G.\,Vg]$ to another, the corresponding relative description $D/G, œ_G,\ Vg/$ is said to be a *non*-individual description. If in this case, using a very large, but *finite*, number N' of *series* of N repetitions of $[G.\,Vg]$, we can (with respect to some explicitly defined criteria of 'precision') discern some '(N, N')-*stability*', we will say that $D/G, œ_G,\ Vg/$ is an (N, N')-*stable statistical description*†‡.

(c) If G and Vg were initially found to mutually *exist* but no sort of individual or statistical stability is finally found, we say that a description $D/G, œ_G,\ Vg/$ corresponding to this pair does not 'exist', and the epistemic referential $(G,\ Vg)$ is discarded *a posteriori*.

(d) All the preceding statements can be generalised to the case where the view V we use contains more than one aspect-view Vg. We then have to carry out (in general, separately) repetitions of *all* the sequences of operations $[G.\,Vg]$ for *all* the aspect-views Vg in V.
In that case, the set of *all* the final *qualifications* thus obtained will be said to constitute the resulting description $D/G, œ_G,\ Vg/$ itself: by definition, the triad $(G, œ_G, V)$ appearing in the symbol for the resulting description is not included in that description, but is just a reminder of its genesis. Moreover, and again by definition, the description itself 'exists' only if some stability is manifested with respect to *all* of the aspect-views involved. However, the degree of stability can vary with the considered aspect-view Vg, so it is relative to Vg. So, like a description $D/G, œ_G,\ Vg/$, a description $D/G, œ_G, V/$ may also be found to be either an individual relative description or a statistical relative description, and in the latter case it will be endowed with some (N, N')-*stabilities*.

(e) Now consider a description for which the generating operation G creates an entity-to-be-described that has never been examined before, and for which the observable phenomena cannot be *directly* observed for some reason (for instance, the chemical structure of a piece of rock sampled by a robot on the moon that is equipped with some analytical apparatus that can identify chemical structure and transmit the result to a computer screen in a laboratory on the Earth). De-

† Hence, a 'statistical' description in MRC is, by definition, endowed with some (N, N')-*stability*. This distinguishes it from the standard notion of a 'statistic', which does not involve repetitions, or stability of any sort. This should be borne in mind throughout the following.

‡ An (N, N')-*statistical description* can at most 'point towards' a 'probabilistic' description $D/G, œ_G,\ Vg/$ (Mugur Schächter 2006, Proposition π13). However, the specification of the conditions under which a *factual* 'probabilistic' invariant associated with the epistemic referential $(G,\ Vg)$ does 'exist' and, furthermore, can be identified by some effective procedure is precisely the aim of the current paper. Until we have identified such a procedure, we shall only refer to *statistical* descriptions endowed with (N, N')-*stability*.

scriptions of this sort form the primordial stratum of human conceptualisations of physical reality. The qualifications produced by a description from this primordial stratum consist exclusively of observable marks ‘transferred’ through ‘measurement interactions’ to the registration devices belonging to pieces of measurement apparatus. A description of this kind is called a *basic transferred description*†.

(f) Within a relative description $D/G, œ_G, V/$, the ‘generator’ G of the ‘entity-to-be-described and the view are *not* fixed entities but are descriptional *roles* freely assigned by the observer/conceiver according to his own descriptional *aims* with respect to some available physical or conceptual element. So an entity that in one description performs the role of the view can in another relative description play the role of the entity-to-be-described or the generating operation. This sort of freedom, which is characteristic of MRC, is one of the sources of the general applicability of this method to any process of conceptualisation subject to the constraint of excluding false absolutes by construction.

(7) Recall that a view V is, by definition, a union of a finite number m of aspect-views Vg, that is,

$$V = \bigcup_g Vg, \qquad g = 1, 2 \ldots m.$$

Each aspect-view Vg introduces its own *semantic g-axis* carrying the ‘values’ $gk(g)$, $k = 1, 2, \ldots w(g)$, chosen for consideration on G, where *w(g)* is the cardinal of the set of values chosen for consideration on g. So V introduces, by construction, the abstract *representation space* defined by the set of its m semantic g-axes. It follows that:

> Any relative description $D/G, œ_G, V/$ consists of a finite cloud structure, *viz.* a finite ‘points-form’ of (gk)-value-points with $g = 1, 2 \ldots m, k = 1, 2, \ldots w(g)$, contained in the m-dimensional representation space of the view V introduced by $D/G, œ_G, V/$.

If the object-entity $œ_G$ is *physical* in nature, we must include a 4-dimensional discrete space–time view $V(ET)$ within V. The relative description $D/G, œ_G, V/$ then becomes a finite cloud structure or a ‘form’ of *(space–time-*(gk)*-value)-points* with $g = 1, 2 \ldots m$, $k = 1, 2, \ldots w(g)$, and with x, y, z, t some *finite* space–time grid on which the units of space and time impose a discrete set of possible space–time val-

† Recall that when a microstate is the entity-to-be-described, the observer can *never* perceive the entity itself. Although the *observer* perceives the observable marks that comprise this basic transferred description on a device endowed with a space–time support, the marks carry no information whatsoever about the space–time location of the physical entity that is being qualified. Therefore, the basic transferred descriptions of microstates *fundamentally* disobey the space–time-frame principle. This is tied to the strange and unintelligible character of the descriptions of microstates, which generates an overwhelming need for some normal space–time representation (a model) that can ‘explain’ the registered marks: the basic transferred descriptions of microstates form the most extreme example of the concept of basic transferred descriptions, but this concept can also apply to macroscopic or even cosmic entities.

ues. This whole form is contained in the $(m+4)$-dimensional representation-space introduced by the view V.

(8) We can create *chains* of relativised descriptions that are connected by common elements in either their respective entities-to-be-described $œ_G$ (and thus somehow connected through the relevant operations of generation G) or through the structures of their views V. There will then be a *descriptional hierarchy* or *ordering* along such a chain. By convention, the first description in the chain is generally assigned the order 1; the description connected directly to it is then of order 2 *with respect* to the first description, and is called a *meta-description*† with respect to the first description; the next description is assigned the order 3 and is a meta-description with respect to the description of order 2 and a *meta-meta*-description with respect to the first description in the chain; and so on. So, in general, the order of a description within a given chain is relative to the process of construction of that chain.

However, in the case of a chain of descriptions that starts with a basic, first-stratum transferred description:

> The initial basic transferred description determines an *absolute*‡ starting point for a particular process of knowledge construction. To show this, we assign the order 0 to basic transferred descriptions.

(9) The passage from a given description in a chain of descriptions to the next one is controlled by the *methodological 'principle of separation'*, or PS for short. However, we will need a little preparation before we state this principle.

Each relative description $D/G, œ_G, V/$ is achieved within an epistemic referential (G, V) where G (in consequence of the methodological assumption of a one-to-one relation $G \leftrightarrow œ_G$) is tied to a *single* entity-to-be-described $œ_G$ and the view V consists of a *given finite* set of aspect-views Vg, each of which carries a *finite* set of aspect-values (gk). Furthermore, the relative description $D/G, œ_G, V/$ is achieved through some finite number of realisations of sequences $[G.Vg]$. So a relative description $D/G, œ_G, V/$ is, by construction, *a finite 'cell of conceptualisation'*. If *all* the aspect-views of the global view V have been taken into account, *each* with *all* its values gk, then, after the realisation of some arbitrarily large but *finite* number of sequences $[G.Vg]$ performed for *all* the aspect-views Vg in V, if we find a descriptional invariant, then the description $D/G, œ_G, V/$ has been achieved and the descriptional resources of the epistemic referential (G, V) will have been completely *exhausted*. If, however, we want to obtain some new knowledge that is connected with $D/G, œ_G, V/$, but has not been produced within $D/G, œ_G, V/$, we will need to introduce *another* convenient epistemic referential (G', V') that is different from (G, V) because either $G' \neq G$ or $V' \neq V$, or both, and then construct the new

† In logic, the prefix 'meta' indicates an *embedding* language, so it is thought of as placed 'under' the object language as a support. Here, however, we use 'meta' specifically to mean *'after'-and-connected-with*.

‡ This is not a *false* absolute since it is a factual datum, and is thus allowed by MRC (in the same way as it allows definitions, principles and conventions that are absolute within the method).

relative description $D/G', œ_G, V'/$ inside (G', V'), which will be connected with $D/G, œ_G, V/$, but will correspond to the new descriptional aim.

> The *principle of separation* (PS) means that a *new* description $D/G', œ_G, V'/$ in a chain is always achieved by a process that is *explicitly and entirely separate* from the descriptional process that led to $D/G, œ_G, V/$.

In this way, we can systematically avoid any uncontrolled merging or confusion between the aims and origins of two distinct but connected relative descriptions.

(10) When a chain starts with a basic transferred description of order 0, it is often the case that this initial transferred description of order 0 taken as a whole plays the role of the new entity-to-be-described for the immediately following description, which will be of order 1, so that it can be qualified by a certain peculiar sort of view that assigns it 'values' of an 'aspect' with a *definite (and usually connected) space–time support.* In this way, the unintelligible transferred description of order 0 becomes intelligible in the sense that it conforms with the space–time-frame principle – see point (5) earlier in this section.

We call a view that generates such conformity an *intrinsically modelling view.* The final result of such an explanatory description of order l can then be *detached* from its origin. This leaves us with a *model* of the basic, transferred description with order 0 in the chain being considered. Later in the same chain, we can construct a meta-description of higher order that introduces the classical concepts of 'cause' and 'locality', and we thus enter the domain of validity of 'determinism' in the sense of classical physics.

In this way, there is in MRC a *divide* within the pool of relativised descriptions achieved at any given time between the very first relative descriptions in this pool (which are basic, transferred descriptions with absolute order 0), which form *a primordial stratum of conceptualisation* and all the others. The classical models corresponding to the transferred descriptions from the primordial stratum have increasingly complex forms through their inclusion in nets of more complex conceptual structures and constitute an evolving classical *'volume'* of conceptualisation of indefinitely growing thickness.

In this way, MRC incorporates the famous 'quantum–classical cut', and explains it as a special case of the following *universal transition* concept:

$$\text{(transferred descriptions)} \quad \Rightarrow \quad \text{(classical descriptions)}.$$

Note that we write 'transition' here rather than 'cut' because MRC gives a detailed definition of the connection between a basic transferred description and the models that 'explain' it.

(11) According to MRC, any *knowledge* that can be communicated in a *non*-restricted way[†] is a *description.* In other words, descriptions are the only *unrestrictedly* com-

[†] For example, the action of 'pointing towards something' is restricted because it requires a real or virtual co-presence within some delimited space–time domain. Mimes, emotional sounds and so on also require similar restrictions in some way.

municable *knowledge*. 'Facts' that are exterior to any psyche or psychic facts (emotions, desires and so on) that are not manifested through some more or less consensually perceptible expression (verbal or of some other constitution) are not 'descriptions', they are not unrestrictedly communicable knowledge. The statement 'I know this house' is illusory, and can only be made through a lack of awareness, or perhaps as a sort of shorthand. The only rigorous way of stating what is presumably meant is to say 'I know *descriptions* of this house'.

(12) Finally, but crucially in the present context:

> When the concept of probability is reconstructed inside MRC, the events, whether elementary or not, acquire the conceptual status of relativised *descriptions*.

The status of a probabilistic event in MRC is *not* that of an entity-to-be-described $œ_G$ but of a relative description of some entity-to-be-described $œ_G$ involved in the situation, and the entity $œ_G$ itself must be clearly distinguished from any of its individual descriptions, whether realised or potential (Mugur Schächter 2006). If the entity remains unchanged, its descriptions can be varied freely and indefinitely through the use of convenient views. This, as will become clear shortly, is an essential step forward as it provides a considerable increase in expressive and discriminatory power and avoids quite a lot of dead-ends.

This concludes our enumeration of the main concepts of the kernel of MRC.

4. A key example – games with a chopped up painting

The discussion throughout this section will be confined to the *classical* level of conceptualisation, so the entities-to-be-described will be 'objects' in the classical sense† (dice, apparatus, tables, vehicles, roads and so on).

4.1. *Introduction*

In this section we will consider a sequence of examples to gain familiarity with the use of MRC. However, in doing so, and by going through a series of small and obvious steps, it will also become clear how, in any given empirical probabilistic situation, a general effective procedure might be constructed that would enable us to identify a *relativised* factual probability law to be associated with that situation.

4.1.1. *Relativised partitioning and notation*

Consider a square painting P that represents, for instance, a landscape, and for which we have any desired number of replicas.

Now consider a spatial grid of 100 squares σ that fits exactly over the painting P. The location of each square σ is identified by two space coordinates (x_k, y_h) where x_k is an element of a set of 10 successive equidistant coordinates $\{x_k\}$, $k = 1, 2 \ldots 10$, marked on

† In other words, modelling constructs in the sense of Husserl.

a horizontal space axis ox at the lower edge of the grid and y_h is an element of a set of 10 successive equidistant coordinates $\{y_h\}$, $h = 1, 2 \dots 10$, marked on a vertical space axis oy at the left-hand edge of the grid. When the grid is placed over the painting P, the label (x_1, y_1) indicates the square at the bottom left corner of P; the bottom left corner of this square is itself the origin 0 of the plane Cartesian reference system xoy attached to the grid placed over P. So the pair (x_{10}, y_{10}) indicates the square at the top right corner of P.

Consider an epistemic referential (G_P, V) where the generator G_P of the entity-to-be-described is a 'selector' that selects the painting P to be described and V is a view consisting of three aspect-views defined as follows:

— A frame-aspect-view $V(El)$ of *spatial location* (E: space; l: location) for which the possible values are the 100 pairs of spatial coordinates (x_k, y_h), $k = 1, 2, \dots 10, h = 1, 2, \dots 10$.
— A colour aspect-view Vc endowed with a given set of colour values (c: colour).
— A two-dimensional frame-aspect-view $V(E)$ endowed with a very large number of *length*-values (this amounts to the introduction of a very small unit of length)[†].

So we have

$$V \equiv V(El) \cup V(E) \cup Vc.$$

In this case, we can combine the aspect-views $V(E)$ and Vc into a single *view of colour-forms* $Vc\phi \equiv V(E) \cup Vc$. The frame-aspect-view $V(E)$ is endowed with a very small unit of length, so, if the colour aspect-view Vc is sufficiently richly endowed with colour values, the view $Vc\phi \equiv V(E) \cup Vc$ will perceive patterns of colour that will reproduce as 'satisfactorily' as desired those perceived by a normal human eye.

With these definitions and this notation, the description of the painting P achieved within the epistemic referential (G_P, V) is written as

$$D/G_P, P, V(El) \cup Vc\phi/.$$

We will now consider a 'local' epistemic referential $(G(\sigma), V)$ where $G(\sigma)$ selects an entity-to-be-described consisting of only *one* of the squares σ demarcated by the superimposed grid, while the view V is the same as in the referential (G_P, V). Hence, a relative description corresponding to $(G(\sigma), V)$ can be denoted by the symbol

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

and it consists of some 'colour-form' covering the selected square σ, for which the global location is indicated by its 'value' x_k, y_h of spatial location detected by the aspect-view $V(El)$.

If we now *suppress* the global spatial location defined by the frame-aspect-view $V(El)$

[†] Note that $V(E)$ does not include the aspect of the square's *global* spatial *location* that can be perceived using $V(El)$: here, the only role played by $V(E)$ is to satisfy the general space–time-frame principle, which says that in order to examine a *physical* entity, there must be at least one aspect-view that exists with respect to that entity *together with* a space- and/or time-frame view – see point (5) in Section 3.5.2.

in the local relative description

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

of a square σ, we are left with a local description

$$D/G(\sigma), \sigma, Vc\phi/$$

of σ that is achieved within the referential $(G(\sigma), Vc\phi)$. In this description, any *direct* indication of the spatial location of σ has been filtered out. Nevertheless, there still remain some *indications* of the location of the selected square σ in P, namely, the *indirect* indications contained in its colour-*form* since this form reaches the *borders* of the square σ, which allows us to discern a continuity or discontinuity with respect to the patterns of colour reaching the border of some other square σ in P. This leads to a sort of 'attraction by semantic continuity' between the border of one square and the border of the square next to it and a sort of 'repulsion by semantic discontinuity' between non-matching borders.

> This means that if we cut up a replica of the painting P into the 100 squares σ defined by the superimposed grid, we can use the resulting pieces as a jigsaw puzzle (albeit an unusual one that has non-interlocking pieces that are all the same size and shape).

Now suppose we progressively *impoverish* the set of colour values carried by the aspect-view of colour Vc in the view of colour-forms $Vc\phi \equiv Vc \cup V(E)$ by using smaller sets of colour values $j = 1, 2 \ldots q$. This amounts to assuming that the colour-forms view

$$Vc\phi \equiv Vc \cup V(E)$$

becomes blind with respect to many of the colours composed into the 'forms' perceived in the local descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

on P. Hence, we move progressively to a view $V(E) \cup Vcs$ that only perceives areas covered by a *single* approximately uniform *colour-shade* denoted *cs*. We can denote a global version of this new, impoverished view as, say, a view of approximate colours

$$Vac \equiv V(E) \cup Vcs.$$

So how will the local description

$$D/G_\sigma, \sigma, Vac/$$

of a given individual square σ appear to an observer using the view Vac? The answer will obviously depend on certain relationships, which we will have to specify.

So we suppose that the global dimensions of the picture P, the distance between two successive values of the x_k or y_h coordinates, and the spatial distribution of colours in P are *such* that:

(a) The spatial extension of a square σ is small enough to 'exist' (in the sense of point (4a) in Section 3.5.2) with respect to a *single* approximate-colour value j.

Hence, the relative description

$$D/G_\sigma, \sigma, Vac/$$

of *any* given square σ using the uniform approximate-colour view Vac consists of a *single* uniform approximate-colour 'value' j, which no longer forms any outline on the surface of σ that is distinct from the square outline of σ itself. So we now have

$$\{D/G(\sigma), \sigma, Vac/\} \equiv \{D\} \equiv \{j\}, \quad j = 1, 2, \dots q.$$

(b) Any given partial description $Dj \equiv j$ is realised within P on a number of squares σ that is *much* greater than 1 and these squares are, in general, at quite different locations within P. Hence, by construction, the cardinal q of the set of *mutually different* relative descriptions

$$\{Dj\} \equiv \{j\}, \quad j = 1, 2 \dots q,$$

is much less than 100.

Under these conditions, a relative description $\{Dj\} \equiv \{j\}$, $j = 1, 2 \dots q$, instead of allowing us to say

'on this square I perceive that *form*-value'

as the description

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

did, will only allow us to say

'on this square I perceive this uniform shade of red, while on that square I perceive that uniform shade of blue, and so on'.

The colour-*form* that, through the colour-forms view $Vc\phi \equiv Vc \cup V(E)$, appeared to cover a square σ has now been filtered out in its turn, and within the descriptions

$$D/G(\sigma), \sigma, Vac/\,,$$

we have lost the ability to perceive the pattern of colours. Therefore, the descriptions from the set $\{Dj\} \equiv \{j\}$, $j = 1, 2 \dots q$ have now been *cut off* from the global description

$$D/G_P, P, V(El) \cup Vc\phi/$$

of the painting P. We can no longer perceive any 'semantic attractions' or 'semantic repulsions' at the borders of the squares.

Therefore, we will not be able to use the 100 squares σ as a jigsaw.

So, we now have three distinct views for describing a square σ:

— the frame-aspect-view $V(El)$ giving the spatial location of a square;
— the aspect-view $Vc\phi$ giving the colour-form; and
— the aspect-view Vac giving the uniform approximate-colour.

Recall that we have any desired number of replicas of P at our disposal. We will now define a sequence of 'games' with squares σ from chopped up replicas of P that will lead us to the result mentioned at the beginning of Section 4.

4.1.2. *Game illustrating the power of reconstruction contained in a space (or a space–time) order*

For our first game, we will cut up a single replica of P into squares σ, mix them up and put them into a bag. We will then take one square σ at a time from the bag and examine it using the frame-aspect-view $V(El)$ of spatial location *alone*. By construction, this gives us a description for each square consisting purely of the square's coordinates x_k, y_h, but this is enough for us to place the square in its place on the reference grid given by the system of axes *xoy*. Since a view acts as a filter, we have been able to do this *without* taking into account the colour-form carried by the square, or the uniform approximate-colour j defined on it by the view *Vac*. Nevertheless, after we have taken exactly 100 squares from the bag, the complete painting P will have been reconstructed. Although the order in which the squares are removed from the bag is random, each step in reconstructing the global painting P is accomplished in a way marked by *certainty*, and the global process is *finite*. Hence, the spatial frame reference grid, taken by itself, possesses a *topological organising power* that is *independent* of the 'semantic content' of the squares.

The above remarks can be extended in an obvious way to the case of an '*evolving* picture' that is broken up into space–time cubes.

4.1.3. *Jigsaw made from a single replica of P*

In this game, we will again cut up a single replica of P into 100 squares, mix them up and put them into a bag. We will also again remove the squares one at a time from the bag, but this time, we will use the colour-form view $Vc\phi$ exclusively. So each square will be perceived through its relative description

$$D/G(\sigma), \sigma, Vc\phi/.$$

The space location label (x_k, y_h) and the uniform approximate-colour label j will be ignored – they are filtered out.

Under these conditions, the global painting P will have again been reconstructed after we have taken exactly 100 squares from the bag. However, in general, in order to find the correct place for a given square, we will have had to use a lot of trial and error to find where to put the square, or may even have had to put a square aside for a while in order to take out and place other squares to help us identify the correct place for it. Nevertheless, guided by the structure of the patterns of colour it carries, we will eventually be able to identify where it goes. And we will have used the structure of the patterns of colour on each square mainly through its content near the *borders* of the square, where it determines a sort of *neighbourhood coherence* with the patterns of colour near *one* of the borders of *one* of the other squares through the 'attraction by semantic continuity' and 'repulsion by semantic discontinuity' mentioned earlier.

In this game, the content-independent topological organising power of the space coordinates has been filtered out and *replaced* by these 'attractions by semantic continuity' and 'repulsions by semantic discontinuity'. And again, everything is finite, and, notwithstanding the presence of trials and errors and the randomness in the order of taking the squares out of the bag, there is nothing random *in the reconstruction process.*

This game, like the game in the previous section, can be extended in an obvious way to the case of an '*evolving* picture' broken up into space–time cubes. But this time the space–time labels are ignored, and we need to use the attractions by continuity and repulsions by discontinuity along the borders entailed by some other descriptional content instead. (For example, during a criminal investigation, the detective tries, in essence, to complete a generalised space–time jigsaw puzzle.)

4.1.4. *Jigsaw with several replicas of the painting P*

In this game, we will use $1,000$ replicas of the chopped-up painting P and proceed in the same way as we did for a single replica in the previous section. We mix together all $100,000$ squares we now have and put them in a large bag. We then remove the squares one at a time, and, ignoring the space and approximate-colour labels printed on it, search for an appropriate place to put each of them on one of the $1,000$ space–time grids placed in front of us.

What will happen? After we have removed $100,000$ squares from the bag, we shall *certainly* have entirely reconstructed all $1,000$ replicas of the painting P. However, this will only have been achieved after quite a lot of trials and errors. And we will not have completed each of the replicas separately in succession, we will have completed all the replicas together, which involves jumping from one replica to another to place pieces. In general, it is only at the end of the process that all $1,000$ replicas will have become completely separate from one another.

In principle, no essentially new features would be introduced by using 10^N replicas, where N is some finite whole number, instead of $1,000$ replicas. And, as in the previous games, we can also extend this game to a set of 'evolving paintings' in an obvious way. And again, everything is finite in all these cases, and, notwithstanding the presence of trials and errors and the randomness in the order of taking the squares out of the bag, there is, once more, nothing random *in the reconstruction process* for the multiple replicas of P.

The only randomness involved in a jigsaw puzzle, no matter how big or complex it may be, is in the order we take the squares from the bag.

> The attractions by semantic continuity and the repulsions by violation of semantic continuity along the borders of the squares exclude all randomness from the process of reconstructing any number of replicas of the global entity that has been partitioned.

4.1.5. *Probability game with one replica of the painting P*

We will now consider a game showing how randomness can appear within the reconstruction process itself.

In this game we will follow a similar procedure to that used in the previous games, except that when we have finished with each piece, we will put it back in the bag rather than place it in the reconstruction. Although this may, at first sight, seem to be an insignificant change to the procedure, the effect will be that all the characteristics of *probabilistic randomness* will come flooding in, and we will be faced with: unending sequences of events; the corresponding statistical relative frequencies; probabilistically estimated randomness in the evolution of these relative frequencies; and probabilistically estimated convergence. Our apparently insignificant change will turn out to have been a radical conceptual jump.

Our new game will use a single replica of the painting P cut into squares, but this time, instead of using them as a jigsaw, we will play the following 'probabilistic game':

(1) As in the earlier games, the first step is to mix up the squares and put them in a bag.
(2) Again as in the earlier games, take one square out of the bag.
(3) Use the uniform approximate colour view *Vac* (and only that view) to determine and record the value of the index j appearing in the corresponding relative description

$$Dj, \qquad j = 1, 2, \ldots q.$$

Since we do not use the colour-form aspect $Vc\phi$ (or the spatial location frame-aspect $V(El)$), *a fortiori*, the semantic continuities on the borders of the square remain inactive.
(4) Unlike in the earlier games, drop the square back into the bag.
(5) Shuffle the squares in the bag.
(6) Repeat the same procedure from Step (2) an arbitrarily large number of times.

The changes to the procedure mean that this is now a standard 'probabilistic situation'. In this game, unlike the earlier ones, just *before* we take each square from the bag, a certain set of *invariant* conditions is reconstituted, and this set defines, in the sense of the usual factual probabilistic language, a 'reproducible procedure' or 'experiment' and a *stable* set of events $\{j\}, j = 1, 2, \ldots q$. Since, according to MRC, any communicable knowledge is a description (see point (11) in Section 3.5.2), we can explicitly rewrite this set of 'events' as the set of relative descriptions

$$\{j\} \equiv \{Dj\}, \qquad j = 1, 2, \ldots q.$$

So, can we *predict* what will happen in these new conditions?

If the number of times we take a square from the bag and replace it is very much bigger than the number q of elements in U (that is, the number of approximate colours), we can make the following rather *obvious* remarks:

(R1) Since the initial contents of the bag are reconstituted each time we remove and replace a square, all the descriptional values $j = 1, 2 \ldots q$ possible when we are about to take out a square will be equally possible when we are about to take out the next square. Unlike the earlier games, no possibility is irreversibly 'consumed' when we take a square from the bag (since we put it back again). This entails the stability of the global factual situation.
(R2) Correlatively, the contents of the bag are never exhausted. There is no longer a natural conclusion to the sequence of results obtained by repetition of the taking out and replacing procedure. The sequence has arbitrary length, and can grow 'towards infinity'. This is the source of any potential non-effectiveness.

We can also ask two questions, which have less obvious answers:

(Q1) If we continue the game indefinitely, will we observe all q possible values of the index j of uniform approximate colour?

(Q1) If we continue the game indefinitely, how will the relative frequency $n(j)/N$ of the outcomes of a given value j of the index of uniform approximate colour evolve as N increases?

However, despite the fact that we have said the answers are not obvious, after short reflection, almost all people skilled in *current* probabilistic thinking will give the following answers to these questions:

(A1) It is *nearly certain* that if the number of repetitions N is large enough, then all q values of the index j of uniform approximate colour will show up.

(A2) If the number of repetitions N is increased without *a priori* limitation, then, sooner or later, but *nearly certainly*, and for *any* j, the relative frequency $n(j)/N$ of the outcomes of a given value j of the index of uniform approximate colour will exhibit a certain convergence. Specifically, the value of the relative frequency $n(j)/N$ will tend to reproduce the value of the ratio $n_{P(j)}/100$, where $n_{P(j)}$ is the number of squares in the cut-up replica of the painting P that carry the value j of uniform approximate colour, and 100 is the total number of squares that the replica of P was cut up into.

But *why* should there be *any* convergence? And *why should it be precisely towards the ratio* $n_{P(j)}/100$ *defined on* P? And why, in both formulations (A1) and (A2), is it 'nearly certain' and not 'completely certain'? To answer these questions, we need to look into the minds of those giving the answers (A1) and (A2), where we will find, more or less explicitly, some equivalent of the following reasoning:

Reasoning for answer (A1):

Since each time we take a square out of the bag and record the j-value for it, we then return the square to the bag, and since this process can be repeated indefinitely, there is, given a sequence of arbitrarily long length N, no *a priori* basis for strictly excluding:

— any specific individual possible outcome $j = 1, 2, \ldots q$;
— any specific individual sequence of j-values;
— any specific individual global statistical distribution of relative frequencies

$$n(j)/N, \qquad j = 1, 2, \ldots q,$$

with

$$\sum_j n(j)/N = 1,$$

that can be constructed for a given N, with j-values belonging to the universe $\{Dj\}$, $j = 1, 2, \ldots q$, of relative descriptions.

Given the indefinite repeatability we have assumed here, any outcome of any feature that cannot be *a priori* excluded on the basis of some specified reason has to be *a* *priori* admitted as possible. These two formulations have the same significance, so any distinction between them would amount to a contradiction. For instance, nothing allows us to *a priori* strictly exclude the maximally unbalanced statistical distribution $n(j')/N = 1$ for some given N and j', that is:

$$\begin{aligned} n(j') &= N \\ n(j) &= 0 \qquad \text{for any } j \neq j' \end{aligned}$$

(for example, with $j' = 2$, the sequence would be 2222222222... of length N). Indeed, if it is possible to get a square carrying $j = 2$ as the first square taken from the bag, since that square is put back in the bag before the second square is taken out, the same possibility also holds for the second square, and so on, indefinitely. However, nothing excludes getting a square with $j \neq 2$ either. These considerations lead us to the answer (A1).

Reasoning for answer (A2):

We know that:

— the number of squares in the bag and the number of possible approximate-colour values j are both finite;
— all the squares come from the cut-up replica of the complete painting P; and
— *all* the squares from the replica are placed in the bag.

Under these conditions, *before* each square is taken out of the bag, it is natural for us to have a greater expectation of finding on that square a j-value (the approximate colour) that is repeated, say, on 10 different squares of the painting P, than a j-value that is only repeated on 2 different squares of P.

The result we get *after* taking a square out of the bag has no effect on the reasonableness of what we expected *before* taking it out – we must avoid any confusion between *a priori* and *a posteriori* as well as between what is 'possible' and what is 'probable'. So, since we know that each time before we take a square out of the bag it contains *exclusively* the 100 squares that make up one replica of the painting P, it is natural for us to expect *a priori* that in a sufficiently long sequence of j-values, each possible j-value will be obtained a number of times that is approximately proportional to the number of squares on which this j-value is realised in the complete painting P. It is also natural to expect that as the number or repetitions N increases, the relative frequency $n(j)/N$ will be found to tend to converge, for each given j-value, towards the ratio

$$n_P(j)/100$$

where $n_P(j)$ is the number of squares in one complete replica of the painting P that give that j-value. More generally,

$$n_P(j)/n_{PT}$$

where n_{PT} is the total number of squares, which is 100 in this example.

Given the conditions, it is impossible to justify any other assumption, while this assumption, in a certain sense, simply *follows*.

Indeed, the global form of a single replica of P is contained in the bag, even though it has been cut up. So, in the long term, it *must* manifest itself through any view that is not entirely blind with respect to it. Now, under the conditions of our probability game, the only active view is the approximate colour view *Vac* with the possible values $j = 1, 2, \dots q$, and this view is not entirely blind with respect to the form of P. Moreover, under the conditions of our probability game, the *unique* possible manifestation of the global colour-form aspect of P we can get through the approximate-colour view *Vac* consists of a set of relative frequencies

$$\{n(j)/N\}, \qquad j = 1, 2, \dots q,$$

which reproduces the set of ratios

$$\{n_P(j)/n_{PT}\}, \qquad j = 1, 2, \dots q$$

from the global colour-form aspect of P. So, by default, such a set of relative frequencies is what has to be expected, and this amounts to the convergence mentioned in answer (A2).

However, this can only be expected with near certainty, not complete certainty. This is entailed by the conditions we have imposed: these conditions simply *exclude* the assertion that each relative frequency $n(j)/N$ will certainly converge towards the corresponding ratio

$$n_P(j)/n_{PT} = n_P(j)/100,$$

and thus, all the more, the assertion that it will strictly reproduce this ratio.

Indeed, we have already pointed out that any sequence of N results j is possible, even a sequence $kkkkkkkkk\dots$ of N results with $j = k$. However, we are reasoning within the abstract framework of the concept of probability[†], so there is a *normalisation* condition for the probabilities on a universe of events of any sort: the sum of all the probabilities assigned to the events from the universe being considered must be equal to 1. So, consider a sequence $\sigma_\omega(N, j)$ of N results j, where ω is an index of some *statistical structure*

$$\{n(j)/N\}, \qquad j = 1, 2 \dots q$$

and N is any whole number. For any give N, there exists a corresponding *finite* set

$$\{\sigma_\omega(N, j)\}, \qquad j = 1, 2, \dots q, \quad \omega = 1, 2, \dots \nu,$$

of N mutually distinct statistical structures that can be constructed with N. These constitute a universe of events (meta-events, with respect to the events from $\{j\}, j = 1, 2, \dots q$). And the probabilities

$$p(\sigma_\omega(N, j)), \qquad \omega = 1, 2, \dots \nu,$$

[†] This statement does *not* introduce any circularity here since it is only the concept of a *factual* probability law that is still undefined; we do *accept* the abstract mathematical probabilistic syntax introduced by Kolmogorov, at least as an initial basis.

assigned to these new (meta-)events are also subject to the normalisation condition

$$\sum_{\omega} p(\sigma_\omega(N,j)) = 1.$$

So any sequence $\sigma_\omega(N,j)$ is possible *a priori*, and thus 'consumes' a certain 'quantity of probability' within this normalisation condition. This forbids us from assigning *a priori* complete certainty (that is, with probability equal to 1) to any specific sequence $\sigma_\omega(N,j)$. If we did assign complete certainty to one of the sequences $\sigma_\omega(N,j)$, then, contrary to the initial assumption that *any* sequence is *a priori* possible, we would *a priori* exclude the possibility of any of the other sequences $\{\sigma_\omega(N,j)\}$ occurring. This would be a contradiction. So a certain and strict convergence towards *all* the ratios $n_P(j)/100$ is excluded by the rules of our probabilistic game.
However, *nothing* in the rules of the probability game prevents us from holding the intuitive notion that with sufficiently large numbers N, *each* relative frequency $n(j)/N$ would, nearly certainly, come arbitrarily close to the corresponding ratio

$$n_P(j)/n_{PT} = n_P(j)/100.$$

This is precisely the answer (A2) given for question (Q2).

So we have now made the reasoning and motivation underlying the answers (A1) and (A2) explicit. However, as we said earlier, this is not the reasoning and motivation in any mind, but the reasoning and motivation in the minds of people skilled in current probabilistic thinking, and those minds are conditioned, precisely, by a deep understanding of the law of large numbers. So we will now compare our discussion of the special case of our probability game with the general law of large numbers.

4.2. *An effective definition of the factual probability law in the case of the 'probability game' with the painting P*

At first sight, the reasoning and motivation we described in the previous section for the answers to questions (Q1) and (Q2) may seem trivial, but, in fact, they lead us to a conclusion that is far from trivial. Indeed, the answers (A1) and (A2) will finally provide us, for the particular case of the probability game with the picture P, with an *effective* definition, founded on 'real facts', of the elusive concept of a factual probability law. And we will arrive at this definition by reference to the law of large numbers:

$$\forall j.\,\forall(\epsilon,\delta).\,\exists N_0.\,\forall N.\ (N \geq N_0) \Rightarrow \boldsymbol{\mathcal{P}}\left[(|n(e_j)/N - p(e_j)|) \leq \epsilon\right] \geq (1-\delta). \qquad (3)$$

Indeed, by simply identifying and substituting terms, we will be able to see that the expression (3) of the law of large numbers can be viewed as a rigorous mathematical translation of precisely the partially intuitive and partially 'reasoned' answers (A1) and (A2). So we set

$$\begin{aligned} e_j &\equiv Dj \\ \{p(Dj)\} &\equiv \{n_p(j)/n_{PT}\} \equiv \{n_P(j)/100\} \end{aligned} \qquad (4)$$

for $j = 1, 2, \ldots q$, in (3) to give

$$\forall j.\, \forall(\epsilon, \delta).\, \exists N_0.\, \forall N.\ (N \geq N_0) \Rightarrow \boldsymbol{\mathcal{P}}\,[(|n(Dj/N) - n_P(j)/100|) \leq \epsilon] \geq (1 - \delta).$$

where the numbers $\{n_P(j)/100\}, j = 1, 2, \ldots q$, satisfy all the conditions required for a probability law (*cf.* the footnote on Page 6), whether formal or factual. For example, they are real positive numbers (here they are *rational* numbers) and they obey the normalisation condition

$$\sum_j n_P(j)/100 = 1.$$

So, in the case of the probability game with the picture P, the set of ratios

$$\{n_P(j)/n_{PT}\} \equiv n_P(j)/100\}, \qquad j = 1, 2 \ldots q,$$

defines on the set of events $\{Dj\}, j = 1, 2, \ldots q$, the quite definite and *effective* factual probability law

$$\{p_F(Dj)\} \equiv \{n_P(j)/n_{PT}\} \equiv \{n_P/100\}, \quad j = 1, 2, \ldots q, \qquad \text{F: factual} \tag{5}$$

It is striking that the definition (5) amounts to our using the intuitive, primitive concept of the probability of an 'outcome' as

$$\frac{[\text{the number of 'favourable cases'}]}{[\text{the total number of possible 'cases'}]}.$$

Indeed:

The weak law of large numbers implies this intuitive definition!

But the law of large numbers is *non-effective* and *absolute*, and these are both forbidden in MRC. However, the construction that led us to the definition (5) was developed using the descriptional *relativities* forced on us by the *discrete and finite* treatment required by MRC. And despite this, no incompatibility has emerged: we shall have to decipher this remarkable disagreement/agreement relation between the law of large numbers and MRC.

4.3. *On the significance in this case of the mere 'existence' of a factual probability distribution*

The belief in the existence of a factual probability law in the answers (A1) and (A2) was founded on the fact that before each square was taken out of the bag, it contained a complete cut-up replica of the painting P:

> The systematically repeated presence of a cut-up complete replica at the same point (before taking a square from the bag) in the sequence of random operations (taking a square from the bag) suggests that this presence *had* to manifest itself in some way, if possible. And in the circumstances under consideration, the only possible way it could manifest itself was in the emergence of a probability law.

If we had started directly with the probabilistic game in Section 4.1.5, without considering the earlier games in Sections 4.1.2–4.1.4, the significance of this remark would have remained hidden. The presence of a cut-up whole inside the bag would have manifested itself only in cryptic terms: specifically, in terms of the values j of the impoverished aspect-view Vac, which filter out any trace of the local patterns of colour initially carried by each piece of the picture, thus producing the set of events $\{Dj\}, j = 1, 2, \ldots q$. Any observable hint of participation in a more complete structure endowed with any inherent global 'significance' going beyond the descriptions $Dj \equiv j$ has vanished if all we consider is this set of events. But, thanks to *the explicit successive relativisations carried out through the use of MRC*, the vague intuitive definition of the probability of an outcome j has acquired a *traceable* connection to the semantic notion of the 'complete form' of a painting P involved in the random phenomenon being considered.

> We have released the direct perception of random outcomes from the bonds imprisoning it exclusively within the probabilistic level of conceptualisation by building awareness of an essential connection to another, higher level of conceptualisation. And on *this* higher level, we can finally and clearly perceive a *precise* meaning to the statement that 'a probability law must exist'.

4.4. *Conclusion for Section 4*

In the probability game we have considered in this section, knowledge of the colour-form carried by the complete replica of the painting P, the way we cut it up and the view defined on the pieces determined in *effective* terms, both, the *significance* to be assigned to the simple assertion of the existence of a factual probability law acting on the set of events $\{Dj\}, j = 1, 2, \ldots q$, and the *numerical structure* of this factual probability law.

Furthermore, the way these emerged makes the weak law of large numbers intelligible and reveals the *source* (in this case) of its non-effective character, namely, the *impossibility* of making relativised use of the finiteness of the whole (the complete painting) entailing the probability law

$$\{p(e_i)\} \equiv \{p_F(Dj)\} \equiv \{n_P(j)/100\}, \qquad j = 1, 2, \ldots q,$$

because we do not *know* what this 'whole' is, or even that it *exists*. This is because we have been trained to rush into absolute formulations that lead us into a dead-end and leave us there indefinitely. Given such ignorance of the 'whole', we have no alternative but to make use of:

— An *indefinitely* increasing integer N that counts the total number of achieved realisations of the relevant experiment and allows us to compare the relative frequencies of the various outcomes.

— A *non*-effective definition of the concept of probability in terms of the *mathematical virtual limit* of an infinite sequence of values of the relative frequency registered for the outcomes of any given possible event.

In order to compensate for the drawbacks resulting from our ignorance with respect to the 'whole' and the absolute forms of expression we are trapped in, we have to offer conceptual *room* for *any* possible sequence of relative frequencies, and this does indeed require *virtual limits* and *real* numbers to represent a factual probability: when we ignore why and how the observed facts emerge, we are forced to over-represent in order to ensure that nothing essential escapes the representation. All the various 'laws of large numbers', weak or strong, involve semantically empty mathematical conditions of *convergence* and *integrability* that, in particular, also cover the possibility of a global form placed on a meta-level, from which the observed random phenomenon issue in some way†. But they do not explicitly identify such a meta-form, and in consequence of this, they introduce descriptional looseness and non-effectiveness.

While a definition founded on *knowledge* of a definite delimited whole divided into a finite number of pieces can be achieved in terms of *rational* numbers, and these numbers can be found by just counting the pieces making up the whole, provided the pieces are clearly distinguished by convenient relativisations.

So we now have many suggestions for how to proceed, but we are also left with an obvious question. In the case of the probabilistic game with the picture P, we simulated ignorance of the complete picture that was, in fact, both pre-existing and known. But then, in order to identify the definition (5) of the factual probability law involved, we *made use of it as a reference.* Now, we need to ask how far the conclusions we reached on this basis can be explicitly generalised to the case most commonly found in practice, where the pre-existence of an integrated meta-form corresponding to a given random phenomenon is not ensured, and for which, *a fortiori*, no knowledge of such a form is available.

5. Construction of the factual effective probability law tied to any given random phenomenon

5.1. *Karl Popper's propensity interpretation of probabilities*

Up until now, Popper had been the only thinker to have proposed a definite meaning (an interpretation) for the assertion of the existence of a factual probability law, which he did through his famous 'propensity interpretation':

> 'Take for example an ordinary symmetrical pin board, so constructed that if we let a number of little balls roll down, they will (ideally) form a normal distribution curve. This curve will represent the *probability distribution* for each single experiment, with each single ball, of reaching a possible resting place.
> Now let us "kick" this board; say, by slightly lifting its left side. Then we also kick the propensities – the probability distribution – ...

† The French Wikipedia article on the law of large numbers says, in French:

> 'Pour Andreï Kolmogorov, la valeur épistémologique de la théorie des probabilités est fondée sur le fait que les phénomènes aléatoires engendrent à grande échelle une régularité stricte, où l'aléatoire a, d'une certaine façon, disparu.' (`http://fr.wikipedia.org/wiki/Loi_des_grands_nombres`)

This remark shows that Kolmogorov himself perceived that the definitions of a factual probability law and of the abstract concept of a probability measure involve, on a *meta*-level of conceptualisation, a 'globality' or a 'unique' entity that, on that level, behaves as a coherent whole (this might be connected with mathematical conditions of integrability).

> Or let us, instead, remove *one pin.* This will alter the probability for every single experiment with every single ball, *whether or not the ball actually comes near the place from which we removed the pin.* ...
>
> ... we may ask: "How can the ball 'know' that a pin has been removed if it never comes near the place?" The answer is: the ball does not "know"; but the board as a whole "knows", and changes the probability distribution, or the *propensity*, for *every* ball; a fact that can be tested by statistical tests.' (Popper 1967)

According to this interpretation, the global experimental situation introduced by any given random phenomenon, with all the material objects and all the actions it involves, determines a specific corresponding numerical law giving the distribution of the probabilities of the elementary events involved.

The current author has for many years held that this interpretation is both deep and important (Mugur Schächter 1992c; Mugur Schächter 2002b; Mugur Schächter 2006). However, it does not provide any pragmatic guidance for the effective construction of the factual probability law asserted to apply to the experimental situation. It is obvious that if *this* is what we require, Popper's propensity interpretation, though it constitutes a highly suggestive insight, is in need of some rigorous specifications.

We should first say that it has already become clear that it is not appropriate to speak of 'the' probability law, in any absolute sense. But even if we translate this concept into MRC terms so that we always speak of the probability law *relative* to some epistemic referential, is it appropriate to assume in advance that such a relative probability law *does* in fact 'exist' for *any* random phenomenon? Also, is it enough to raise this question in such ontological terms, or do we need to consider it from the start, and quite fundamentally, in purely methodological terms when our aim is just to construct a corresponding probability law? And, furthermore, is it always possible to connect this constructive aim to some individual global meta-form?

In thinking about these questions, our 'probabilistic game' with the painting P suggests the following conjecture:

> Any given factual 'probabilistic situation' can be connected in some way to a constructible *'global form'* that re-expresses in a geometrised way, with an integrated simultaneous structure, the progressively revealed semantic content of the observable statistical manifestations of the situation, thereby enabling us to identify by counting a *finite* set of *rational* numbers constituting a factual probability law that is valid in that particular probabilistic situation.

In the following, we will begin by assuming this conjecture and then develop a *purely methodological* and *constructive* approach. We will leave ontological questions of existence or truth open to begin with, but if our approach succeeds, we shall have transformed our initial conjecture into a proof by construction.

The following points will guide the development:

— We will begin by assuming that any random phenomenon, in consequence of the stabilities involved in the 'procedure' Π, even though these are only relative to a set of macroscopic parameters, nevertheless entails the *constructibility* of a corresponding relative global form that, if available, will allow us to specify, in finite terms, the applicable probability law for the random phenomenon.

— For conceptual consistency with the approach we took in Section 4, we shall work systematically within the framework of MRC, so the global form and the resulting probability law will have to be specified in MRC relative terms.
— According to Kolmogorov's probabilistic syntax, any probability law is tied to a probability space $[U, \tau, p(\tau)]$. Since our aim is to assign a domain of effective factual applicability to Kolmogorov's syntax, we shall have to show how this syntax can be connected to our construction of factual probability laws, which will then ensure that we have produced an interpretation of the formal concept of a probability measure.

5.2. *The MRC concepts of a random phenomenon and a 'probabilistic situation'*

5.2.1. *The MRC definition of a random phenomenon*†

Definition 5.1 (random phenomenon). Consider an epistemic referential (G, V) where:

— G is a factual generating operation that can be repeated 'identically' (with respect to some given set of parameters) and for which each realisation is methodologically assumed to introduce the 'same' *entity-to-be-described* $œ_G$ (see point (1) in Section 3.5.2).
(In many cases, an MRC analysis leads us to define G as simply the identically repeated choice of the material initial conditions and the operations required to carry out the 'experiment'.)
— V is, in general, a union of aspect-views

$$V \equiv \bigcup_g Vg, \qquad g = 1, 2, \dots m,$$

with m a finite integer. This is a factual *operational* view that acts upon $œ_G$ and generates the complete action leading to a qualification of $œ_G$ through aspect-values of the aspect-views $Vg \in V$ (see point (3) in Section 3.5.2; point (4a) is also assumed).

The *random phenomenon* corresponding to the epistemic referential (G, V) consists of the pair (Π, U), where $\Pi \equiv [G.V]$, considered globally, constitutes the 'identically' repeatable procedure (the 'experiment') and U denotes the universe of outcomes of a large number of repetitions of $\Pi \equiv [G.V]$.

5.2.2. *The MRC descriptional status of a random phenomenon*

By construction, a random phenomenon is certainly some type of MRC relative description, but what type exactly? For clarity, we will give a very detailed and explicit answer to this question.

According to MRC, *any* relative description $D/G, œ_G, V/$ realised with an epistemic

† See Mugur Schächter (2002; 2006, pages 193–202) for full details.

referential (G, V) involves a large number of *repetitions* of the sequence $[G.V]$ (see point (6) in Section 3.5.2). Then:

— If the procedure $\Pi \equiv [G.V]$ is repeated N times and we get the *same* collection† of gk-values for the aspect-views Vg in V for each of the N repetitions, then $D/G, œ_G, V/$ is an 'N-*individual*' description (see point (6a) in Section 3.5.2).

— Otherwise:

- If some (N, N')-stability manifests itself, it is a statistical description in the sense of MRC.
- Otherwise, if no such stability is observed, the epistemic referential (G, V) is discarded *a posteriori*, and there is no relative description corresponding to the pair (G, V).

  In this case, the entity-to-be-described $œ_G$ under consideration, although it exists with respect to V (in the sense of point (4a) in Section 3.5.2), appears, *a posteriori*, to be unsuitable for qualification using V (the resulting qualifications are 'too variable' to justify any labelling).

Among these possibilities, agreement with the usual understanding of the concept of a random phenomenon requires us to assume that the corresponding epistemic referential (G, V) produces a *statistical* description $D/G, œ_G, V/$ (see point (6b) in Section 3.5.2). When the sequence $\Pi \equiv [G.V]$ is repeated N times, where N is a large number, it produces a set of mutually distinct outcomes, which constitute the universe U. Since everything in MRC is finite by construction, the universe U is a finite universe.

We will now consider the outcomes in U, which we will assume are *physical* entities. So, in accordance with the *frame principle* (see point (5) in Section 3.5.2), the view V should involve at least one aspect G that is not a space–time aspect *and* at least one space–time-frame view $V(ET)$ or purely space-frame view $V(E)$. In the case of a random phenomenon, we only require a space-frame view $V(E)$‡. In general, this space-frame view will involve *several* space aspects (the three-dimensional 'position', parameters defining an orientation, and so on). So the view involved in the procedure $\Pi \equiv [G.V]$ has the structure

$$V = \bigcup_g Vg \cup V(E), \qquad g = 1, 2, \ldots m,$$

with $m \geq 1$. Note that this view is *not* chosen so that it exhausts all the aspects g and all the space-frame aspects that are 'naturally' observable on an outcome of the experiment Π (with some 'naturally' perceptible interval of values for each given aspect). Indeed, this view generally does not involve *all* the aspects that are naturally observable

† Note that we deliberately avoid the use of the word 'set' here to emphasise the fact that the gk-values do not form a simple set of elements that are just assumed to exist, but are relative descriptions resulting from definite interactions between 'reality', whose nature cannot be directly perceived, and a defined active view V.

‡ A stable random phenomenon usually divides up the time dimension into temporal extensions for which no ordering is required since only counting will be performed. Because of this, and because we are assuming stability, the time dimension is not active in the usual sense.

on an outcome of the experiment Π. In particular, the spatial qualifications are almost systematically ignored. In this sense, the relative descriptions realised for the epistemic referential (G, V) are just *labelling* descriptions[†].

Putting this briefly, the definition of any possible *single* outcome of a realisation of $\Pi \equiv [G.V]$ will consist of a description that only makes use of a *smaller collection of aspect-views*

$$[\{gk(g)\}, g = 1, 2, \dots m, k(g) = 1, 2, \dots w(g)]$$

(see point (3) in Section 3.5.2) selected from amongst all the observable qualifications available within the epistemic referential (G, V).

In order to simplify the notation, we will now associate a unique *global* index r with each of the possible collections of aspect-values associated with one particular outcome, as defined above. We write $r = 1, 2, \dots \rho$, where ρ denotes the cardinal of the set of these global values (the number of mutually distinct combinations of values $k(g)$ of the aspects G in V that have been selected to define the outcomes in U). We will write Dr to denote the particular outcome of this kind tied to a given value of the index r, and we will write Vr to denote the view that contains the selected aspects[‡].

So, finally, we can say:

In MRC, the concept of the random phenomenon (Π, U) amounts to the general concept of a 'statistical description' $D/G, œ_G, Vr/$ in the sense of point (6) in Section 3.5.2, that is, it is a relative description that can be represented by[§]

$$D/G, œ_G, Vr/ \equiv [(\Pi, U)]$$
$$\equiv [\{[G.Vr]_n, \{Dr\}\}, \quad n = 1, 2, \dots N, \quad r = 1, 2, \dots \rho]$$

where the number of labels N is the number of successive repetitions of the pro-

[†] In the case of a microstate, the aspect-views Vg in V will not, in general, all be mutually compatible (they cannot be realised simultaneously on a single replica of the microstate), so, in general, a 'single' full realisation of $\Pi \equiv [G.V]$ will consist of a union of distinct and mutually incompatible sequences $[G.Vg]$. However, we are working here under classical conditions, so, in accordance with the classical theory of probabilities, all the aspect-views will be mutually *compatible*, so all the effects of the realisation of a single sequence $[G.V]$ will be obtained simultaneously.

[‡] Since any description $D/G, œ_G, V/$ in the sense of MRC involves a large number of outcomes N, a single outcome 'Dr' is not itself a 'description of $œ_G$' in the sense of MRC. Nevertheless, it is a relative qualification of the entity-to-be-described $œ_G$, which can be regarded as a sort of limiting relative description corresponding to the limiting case $N = 1$. In this sense, the letter D in the symbol Dr is justified.

[§] The expression shows explicitly that MRC incorporates the fact, which was revealed in the study of *infra-quantum mechanics* (Mugur Schächter 2011; 2013), that any strictly basic, primary, *scientific* knowledge emerges with an inherently statistical character, notwithstanding the fact that this knowledge is generated in a way that produces observable results with maximal stability through repetitions of the same sequence of operations $[G.V]$. This, in turn, underlines the fact that MRC is a *method* for generating *consensual* knowledge. In particular, MRC is not concerned with the 'natural' ways of describing a phenomenon, since 'natural' statistics that are free of any stability constraints play no part. Such statistics *can* be represented within MRC as *a priori* pairings (G, V), where the entity-to-be-described generated by G exists relatively to V. However, these descriptions will be eliminated *a posteriori* because they do not produce descriptional results with any stability.

cedure Π, which is large but finite, and we have used the definitions

$$\Pi \equiv [G.Vr]$$
$$U \equiv \{Dr\}$$

where $r = 1, 2, \dots \rho$.

The descriptions Dr, like all MRC descriptions, can only include finite features. However, in the case of a factual random phenomenon, we will also make the following explicit assumption:

> *Every* description Dr, as well as the whole universe U of these descriptions, is *confined* within a definite *space domain*.

5.2.3. *'Random phenomenon': a conceptual artefact*

So, in MRC, a random phenomenon is assumed, by construction, to be endowed with some definite stability entailed by a *repeatable* procedure $\Pi \equiv [G.Vr]$ involving a stable epistemic referential, and so on. Even though this may only be a conventionally limited 'partial' (N, N')-stability, it *is*, nonetheless, a sort of stability. It is clear that factual statistical situations endowed with this sort of stability do *not* pre-exist naturally. The naturally occurring situations that are standardly called 'statistical' do not require any 'identical' repeatability of some definite procedure Π, so they do not involve any explicitly defined stable content. They are not organised in advance to ensure predictability, not even some local (and generally short lived) domain of predictability. They are just pieces of information, though they are often gathered precisely with the aim of achieving some predictability at a *later* time.

Hence, a random phenomenon appears to be a *conceptual/operational artefact* that is deliberately conceived and created to extract, from the mass of natural fluctuations surrounding us, a sample, subject to *local constraints*, that exhibits some predictability, which, although leaving some *uncertainty*, is more stable than the phenomena usually described as being 'naturally statistical'. However, it is also less precise than an 'individual' relative description in the sense of MRC (see point (6a) in Section 3.5.2), or, *a fortiori*, a 'deterministic' description in the classical sense.

There may be some temptation to think that factual random phenomena may also occur naturally, but a little reflection should make it clear that it is impossible to unambiguously identify a random phenomenon in nature. This is even true in the case of cosmological regularities, where humans can discern not only (N, N')-statistical descriptions in the sense of MRC but also 'individual' N-descriptions, since, if we consider a sufficiently long period, any statements about stability assigned to phenomena would be invalid because they cannot be controlled artificially (Henri Poincaré developed this idea of the non-permanence of natural physical laws). In the mathematical discipline of 'statistics', there is no fundamental or systematic distinction between the natural variation in the measurement of observed outcomes and the variation connected with a random phenomenon. However, MRC is, by construction, a *methodological* representation of *verifiable* 'scientific' assertions, and this means we must distinguish between stable and unstable statistics.

5.2.4. *'Probabilistic situation'.*
A game of chance consists of a random phenomenon with a deliberately constructed probability law, and thus an *a priori known* probability law. This is why the theory of 'probabilities' was based on the study of games of chance, and from there, the concept of probability spread into the field of scientific research, and then into the fields of engineering and manufacturing, where, again, the probability laws tend to be pre-established. However, note that:

> Kolmogorov's aporia vanishes when the probability law is known by construction, since in that case the numbers $p(e_i)$ in (2) on Page 8 do not need to be defined frequentially. Hence, the aim of the law of weak numbers is then only to show that, in the long run, the known factual probability law *does* arise observably, with near certainty.

However, the probability laws used in *scientific* research, and even in engineering, are just *assumed* to be involved as a consequence of the deliberately created stabilities, and are never actually known in advance. This means they have to be *identified*, so the aporia will occur.

There is also much confusion of various sorts about the factual probability laws themselves†.

These subtleties call for a more explicit classification of the 'factual situations' that can be involved in the concept of a random phenomenon. We will conclude Section 5.2 with the following *definition.*

Definition 5.2 (probabilistic situation). The factual situation generated by the realisation of a random phenomenon will be called a *probabilistic situation.*

And since a random phenomenon is a conceptual/operational artefact, a probabilistic situation in the sense just defined is also a conceptual/operational artefact: *a probabilistic situation never exists naturally; it is always a result of design and realisation.*

We believe that this is a non-trivial conclusion.

† In the course of a private exchange, Professor Jean-Marie Fessler, the 'Advisor to the President' of the MGEN (the mutual organisation providing social and health insurance for all public-sector workers in France), drew my attention to 'the "unreasonable" proliferation of assertions of the existence of Gaussian distributions in the statistics connected with health issues' (Fessler 2009). He thinks that these assertions are due to a confusion between the variation due to *errors* in the *numerical* values obtained through *measurements* of some quantity performed on some event in the universe of possible outcomes, and the distribution of the probabilities of those events (which, in general, are not numbers).

And in the extreme case of 'primordial' random phenomena, such as the 'transferred descriptions' in the sense of Section 3.3.3, which are connected with microstates and studied in quantum mechanics, it may turn out that the factual probability laws, which we now just assume always 'exist', simply cannot be constructed, in general, *within fundamental quantum mechanics*: it may be that in order to construct these probability laws systematically, we will need to make explicit use of a mutually accepted model of a microstate, which is very likely to be some improved version of the de Broglie–Böhm model.

5.3. *Constructive principles*

In this section, we shall show how we can construct knowledge of the factual probability law involved in any given random phenomenon.

5.3.1. *The appropriate connection between factual probabilistic data and Kolmogorov's syntax*
Consider a random phenomenon

$$(\Pi, U) \equiv \{[G.\,Vr]_n, \{Dr\}\}, \qquad n = 1, 2, \ldots N, \quad r = 1, 2, \ldots \rho.$$

In Kolmogorov's approach, the probability space $[U, \tau, p(\tau)]$ tied to this random phenomenon is founded on the universe U generated by the Π, which is identified with the universe of *'elementary'* events (*cf.* Section 2.1). However, this is *not* relevant to our present aim. In order to specify how we can construct the factual probability law for (Π, U), we will need to find some other way of making the connection between the factual data produced by (Π, U) and Kolmogorov's syntactic concept of a probability space, as we will now explain.

In the example of the probabilistic game with the painting P, the universe of *labelling* elementary-event-descriptions

$$U \equiv Dj, \qquad j = 1, 2, \ldots q,$$

consisted of a set of q mutually distinct types of piece (pieces with the same approximate colour), with $q \ll 100$. These had been *extracted* from the initial set of 100 distinct 'local' relative descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi$$

comprising the pieces of P by 'simplifying' these initial local descriptions to give the universe

$$\begin{aligned} U &\equiv \{D/G(\sigma), \sigma, Vac/\} \\ &\equiv \{Dj\}, \qquad j = 1, 2, \ldots q. \end{aligned} \tag{6}$$

In this way, any perceptible connection to the description

$$D/G_P, P, V(El) \cup Vc\phi/$$

of the complete painting P has been effaced from the elements of U; these elements have been *perceptually cut off from* P by the 'simplifications' applied to the global view. It is because of this cut that we could not use the elements of the simplified universe (6) as a jigsaw. It is also the reason we could not specify the factual numerical probability distribution

$$\{p_F(Dj)\}, \qquad j = 1, 2, \ldots q,$$

on U for the probabilistic game using these elements.

However, in that particular case, the theorem of large numbers, together with *knowledge* of the complete painting P, enabled us to identify the numerical factual probability distribution. This was because each labelling elementary event description Dj had remained *materially* immersed in (though perceptually cut off from) an *already available* and previously perceived more complex 'local' relative description

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

(note that the generating operation $G(\sigma)$ is the same, and only the aspect-view applied to it has been changed) that was itself a *material fragment* of the *known* description

$$D/G_P, V(El) \cup Vc\phi/$$

of the complete painting P. This very particular circumstance compensated for the cut: namely, it enabled us to *refer* the 'simplified' descriptions in $U \equiv \{Dj\}$ to the more complex local relative descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

for the jigsaw of P, and then to count the occurrences of each given Dj in the description

$$D/G_P, P, V(El) \cup Vc\phi/$$

of the complete picture P. So it was our *knowledge* of the integral whole P and the reference structure entailed by this knowledge that enabled us to specify the factual probability distribution

$$\{p_F(Dj)\} \equiv \{n_P(j)/n_{PT}\}, \qquad j = 1, 2 \ldots q,$$

through a comparison with the law of large numbers. However, real life is rarely as co-operative as this, so, in general, the data provided by the universe

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

of (more or less) conventionally determined label-descriptions of the outcomes of a random phenomenon $\Pi \equiv [G.\,Vr]$, as defined in Section 5.2.2, cannot guide us towards the construction of the corresponding factual probability law.

In general, however, the syntax of a given domain of physical facts should help us in solving problems about that domain. So, could some connection between $\Pi \equiv [G.\,Vr]$ and a corresponding abstract Kolmogorov probability space be useful in constructing the corresponding factual probability law?

In Kolmogorov's formalism, the probability space associated with a random phenomenon (Π, U) is based directly on the universe U of outcomes to which we have assigned the role of *elementary* events. But in the case of the probability game with the painting P, a connection of this sort would have been entirely useless without any knowledge of P, that is, if the *only* information we had was the label-descriptions

$$\{Dj\}, \qquad j = 1, 2, \ldots q.$$

Indeed, as we will observe in Section 5.3.2, in a probability space founded on $U \equiv \{Dj\}$, both Laplace's principle and MRC would have required us to assume an *a priori* uniform probability distribution on $U \equiv \{Dj\}$, which would have initiated a non-effective and indefinite alternation between *a priori* and *a posteriori* considerations. However, the probability law on an algebra constructed on U (and thus concerned with event-descriptions with qualifications *less* specific than those from U) is not the problem we had to solve, since such an algebra is not the universe produced by the random phenomenon under consideration: our aim in the case of the probability game with the picture P was to specify the probability law on precisely the set of events

$$\{Dj\}, \qquad j = 1, 2 \ldots q,$$

introduced by the random phenomenon (Π, U), with

$$U \equiv \{Dj\}, \qquad j = 1, 2 \ldots q.$$

Now, this would have led us to locate the set

$$\{Dj\}, \qquad j = 1, 2 \ldots q,$$

in the *algebra* from the space associated with this random phenomenon, and to base the space upon the *initial* set

$$\{D/G(\sigma, \sigma, V(El) \cup Vc\phi/\}$$

of 'local' descriptions, which is more complex than the descriptions Dj, and thus *offers a reference* for counting how many initial local descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

'realise' any given description Dj. Hence, an appropriate procedure would be:

(1) Assign to the members of the set

$$U^c \equiv \{D/G(\sigma), \sigma, V(El) \cup Vc\phi/\}$$

the role of *elementary* events and use the index 'c' to indicate that the descriptions in this universe are *more* 'complex', *relative* to the descriptions in the universe

$$U \equiv \{Dj\}, \qquad j = 1, 2, \ldots q,$$

generated by the random phenomenon in the probabilistic game with the picture P.

(2) *Then* construct on U^c an algebra that can contain the descriptions in the set

$$U \equiv \{Dj\}, \qquad j = 1, 2, \ldots q.$$

For instance, the total algebra τ_T^c on U^c certainly contains $U \equiv \{Dj\}$ since it contains, by definition, all the events that can be constructed from the elementary events in U^c. However, the algebra based on the *classification* of the descriptions in

$$U^c \equiv \{D/G(\sigma), \sigma, V(El) \cup Vc\phi/\}$$

according to the values j of the simplified view Vac† (which we will denote by $\tau^c(j)$) also

† In Kolmogorov's classical representation, the syntactic elements used to accommodate specific factual elementary events, or events, do not have a structure that is definite and detailed enough for us to define an algebra based on a classification of the elementary events.

By contrast, MRC does. Indeed, since MRC says that any knowledge that is communicable without restriction is a *description*, an elementary event, or an event, has to be specified as a description. Now, a relative description is referred to a triad $(G, œ_G, V)$, and any individual aspect-view Vg in V introduces *two* sorts of index (formal specifications): one indicating a semantic dimension g; and a satisfies this condition. Hence, in both of the algebras τ_T^c and $\tau^c(j)$, any description Dj for the random phenomenon (Π, U),

$$U \equiv \{Dj\}, \qquad j = 1, 2 \ldots q,$$

would have reappeared as an *event*-description that is 'realised' by a complete set of elementary-event-descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

in U^c. Furthermore, since the probability measure in a Kolmogorov probability space is defined on the algebra on the space, and not on the universe of elementary events, a space where the role of the universe of elementary events is assigned to U^c instead of U would have been more coherent with our aim of specifying the factual distribution of probabilities on $U^\dagger$. These remarks suggest that, in *general* terms:

> Given a given random phenomenon (Π, U),
>
> $$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$
>
> involving an unknown probability distribution, in order to specify this distribution, it might be useful to introduce an abstract Kolmogorov space denoted by
>
> $$(\Pi, U) \rightsquigarrow [U^c, \tau_T^c, p_F(\tau_T^c)]$$
>
> *where U is contained in τ_T^c* (and the sign $\rightsquigarrow$ does not have any standard mathematical or logical meaning).

In this discussion, we have reached this conclusion simply on the basis of an example, but it can be given quite general foundations within MRC, where it can be *proved* that information about the structure of the factual probability law on the universe U cannot be obtained *within* the epistemic referential (G, *Vr*) where U is itself represented, so reference must be made *outside* it‡.

second index, k, involving a value index of g denoted by (gk) (or $(gk(g))$. (Note that we cannot just use the set notation e_i here without severely restricting what we can express by including the *descriptions* of the elementary events, since there is no specific syntactic place for them in the flat set notation e_i.)

Similar considerations concerning the denotation e for the events in the algebra τ show that this notation does not offer a formal location allowing us to express any 'logical' classificatory character for this algebra – see Mugur Schächter (2002b; 2006). Indeed, *it is this fact that has hindered the unification of the classical theory of probabilities and classical logic.*

More generally, in applications of the classical theory of probabilities, the impoverished set-theoretical notation has obstructed the *expression* of the observable features involved. Formal languages, just like natural ones, can have different degrees of expressivity. As illustrated by the example with the painting P, using MRC to move from a set-theoretical notation to the relative descriptional notation provides an invaluable increase in expressivity, which allows us to see features and criteria that can guide us towards a whole class of conclusions that were previously hidden from us.

† Notice how MRC highlights the well-known but often ignored *relativity* of the 'elementary' character of an event (in the probabilistic sense): whether a probabilistic outcome is elementary or not is 'only' determined by mutual *reference* within a corresponding probability space.

‡ See the Appendix for a formal statement and proof.

The substitution of

$$(\Pi, U) \rightsquigarrow [U^c, \tau_T^c, p_F(\tau_T^c)]$$

for the usual

$$(\Pi, U) \rightsquigarrow [U, \tau(U), p_F(\tau(U))]$$

could well have consequences for many probabilistic investigations.

5.3.2. *Laplace's principle versus MRC*

We shall now examine *under what conditions* the introduction of a universe U^c of descriptions that are more complex than those in the universe

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho$$

generated by a given random phenomenon (Π, U) could be useful in specifying the factual probability law for U.

We will begin by observing that *if* we just *assumed* that the probability of an event Dr from $\tau(r)$ is *proportional* to the number of mutually distinct elementary events of the 'complexified' universe U^c, the required probability law on

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

would follow trivially. But this assumption would be equivalent to assuming that the distribution of probabilities on the universe U^c is always *uniform.* Indeed this assumption would precisely entail a factual distribution of probabilities of the form (5) on Page 50 because of the general formal requirement

$$P(A \cup B) \leq p(A) + p(B),$$

where there is equality if and only if A and B are mutually 'independent' (*cf.* footnote † on Page 6).

However, we need to ask: *why* should we postulate a uniform factual distribution of probabilities on the universe of elementary events U^c in the general case?

This question draws attention to Laplace's well-known classical principle, which requires the *a priori* assignment of a uniform probability distribution on the universe U of outcomes generated by any random phenomenon. For Laplace, however, such an *a priori* assignment was just an *initial bet* that is expected, *in general*, to be invalidated by *a posteriori* measurements of the relative frequencies recorded for the elementary outcomes in U. So Laplace's principle induces a back and forth alternation between *a priori* and *a posteriori* considerations involving the non-effective law of large numbers[†‡], and thus *must be avoided.*

So Laplace's principle[§] does not help us identify the form of a *true*, though always relative, factual probability distribution on the universe of elementary events in a probability space. So, finally, we can say:

A probability space

$$[U^c, \tau_T^c, p_F(\tau_T^c)]$$

based on a 'complexified' universe U^c *might* offer a useful medium of reference for constructing the unknown factual probability law for a random phenomenon (Π, U) with

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho.$$

However, this is true if and only if the factual distribution of probabilities on the reference universe U^c (not an *a priori* postulated law) is 'truly' uniform[¶].

Now this is possible only if an empirically true, a factually established, 'satisfactory' degree of uniformity of the probability law on U^c is ensured *by construction*, and thus through factual data concerning the outcomes Dr^c in the complexified universe U^c: as with any syntax, the probabilistic syntax can work in a pragmatically useful way, but only if it is fuelled with factual data that has a nature and organisation that fits conveniently into the syntax.

This gives even more support for the notion that a 'probabilistic situation' is a conceptual artefact subject to pragmatic aims.

5.3.3. *Conclusions for the constructibility of a generalisation of the cut-up picture games*

So we have now established that in order to identify, in an effective way, a factually true numerical distribution of probabilities on a universe of outcomes induced by a given factual random phenomenon, we must rely *entirely* on data that are produced and controlled factually.

We must now investigate what such data could be.

† However, many authors ignore the non-effective *a priori*/*a posteriori* alternation required by a rigorous use of the Laplace's principle. Instead, they just make a definitive *a priori* assumption of a uniform probability distribution of the relevant universe of elementary events, either on the basis of considerations of symmetry (as in Boltzmann's statistical theory of gases and, presumably, in many quantum-mechanical investigations where symmetries are invoked), or without any explicit justification. We stress that the fact that in an abstract probability space $[U, \tau_X, p(\tau_X)]$, the abstract probability measure $p(\tau_X)$ is defined directly on the *algebra* of events τ_X, and not on the universe of elementary events U (*cf.* Section 2.1), indicates that Kolmogorov considered that the distribution of probabilities on U is always to be decided on the basis of factual data external to the probabilistic syntax. This is consistent with what we have called Kolmogorov's aporia.

‡ The proofs of the two laws of large numbers (by Kintchine and others) did not exist at the time of Laplace.

§ Similar reasoning applies to the well-known principle due to Jaynes.

¶ Tarski said 'the snow is white' is true if and only if the snow is white. Here we say 'the factual probability distribution on U is uniform' is true if and only if the factual probability distribution on U is uniform.

In the case of the games with the painting P, effectiveness was entailed by the possibility of *counting* both elementary events and events inside any replica of the finite closed whole denoted by P. The semantic attractions along the borders of the 'complex' and 'local' descriptions

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

were essential in carrying out such counting when starting from the fragments of the picture P because:

— They enabled us to build the finite closed whole P.
— They enabled us to determine when we had completed a replica of P. In this way we were able to avoid the ever-increasing number N of recorded outcomes required by the theorem of large numbers, but instead had a sequence of sets of elements of equal cardinal n_{PT}, which correspond to mutually separated replicas of the completed puzzle of P. Though these replicas *emerged* together in a mixed way, each replica could always be distinguished from the others and, as soon as it was finished, it became completely separate and was *sufficient* for us to determine the ratios $n_P(j)/n_{PT}$ within it[†] that determined the required probability distribution.

Because the painting P already existed and was known in the case of the painting games, when we played the probabilistic game with the label-descriptions

$$\{Dj\}, \qquad j = 1, 2 \ldots q,$$

we knew that the probability law on the more 'complex' universe U^c for the 'local' elementary-event-descriptions

$$\{D/G(\sigma, \sigma, V(El) \cup Vc\phi/\}$$

was *truly* uniform *by construction.* However, in the general case, this is not guaranteed in advance, and we cannot assume it will always happen spontaneously.

In conclusion, the discussion in Section 5.3 suggests that a generalisation of the probabilisitic picture game must introduce a definite procedure that:

— is based on the connection

$$(\Pi, U) \rightsquigarrow [U^c, \tau_T^c, p_F(\tau_T^c)];$$

— enables us to build, using knowledge of the random phenomenon *alone*, an indefinite number of replicas of a *defined and demarcated 'whole'*;
— enables us to estimate *when* the internal content of one such replica has been *completed*;
— enables us to ensure that the number of contributions, within the above-mentioned whole, made by each of the *elementary* events making up the 'whole' have a *factually uniform* distribution (that is, the distribution is not just unspecified or *a priori* assumed to be uniform).

[†] Note that the *ratios* here are not the 'relative frequencies' mentioned in the law of large numbers, but, as in the painting games, ratios comparing the number of contributions of a given 'element' to the total number of components making up the 'whole' structure.

5.4. *Construction of the factual probability law tied to a given random phenomenon*

5.4.1. *Outline*

In this section, we will lay out some guidelines for the development of our general procedure.

Consider an epistemic referential

$$(G, Vr), \qquad r = 1, 2, \dots \rho,$$

inside which a random phenomenon (Π, U) with $U \equiv \{Dr\}$ is described, where r is a *global* label-index and each value of r consists of some set of definite observable values of aspect-views $Vg \in Vr$. Quite generally, r does *not* exhaust† the directly observable aspects on the outcome denoted by Dr (for instance, the space aspects will be systematically ignored); *a fortiori*, aspects hidden from immediate observation (such as atomic and chemical structures) are rarely considered. So, the descriptions Dr are always poorer than other possible descriptions. Moreover, when the descriptions $Dr \in U$ are explicitly introduced purely as *label*-descriptions, they are cut off, by construction, from any possible integrated meta-form tied to (Π, U), so they certainly cannot provide any hints for how to construct a representation of such a meta-form. In order to obtain such hints, we will need to *expand* or 'complexify' each label-description $Dr \in U$ from the definition of a random phenomenon (Π, U) with

$$U \equiv \{Dr\}, \qquad r = 1, 2, \dots \rho,$$

to get a set of 'convenient' relative descriptions that are more complex than the *label*-description Dr. This means we need a complexifying view $V^c \supset Vr$ and a corresponding complexified representation space that includes the representation space of Vr.

We also need to know what the word 'convenient' means. In the present context, it means precisely: *such* that each description $Dr \in U$ is extended into a local form of space-and-gk-values that reaches the *borders* of that local description‡ so that they are sensitive to semantic attractions/repulsions with respect to other local forms developed from descriptions of other events $Dr \in U$. This amounts to us requiring that the view Vr involved in the label-descriptions is enriched in such a way that we can use the resulting complexified descriptions to make an abstract sort of jigsaw puzzle within the complexified representation space V^c.

Now, according to MRC, this is *always* possible in the classical domain:

— Any event $Dr \in U$ involves a *physical* entity. So the space–time-frame principle (see point (5) in Section 3.5.2) means we can introduce space specifications that *identify the volume of 'physical' space containing that physical entity.* This already enables us

† According to MRC, 'exhaust' is a *forbidden* false absolute since 'all' is a false absolute, and it is a huge false absolute, though a very common one.

‡ By the definition of a random phenomenon in the classical sense, the event-descriptions Dr are localised in space, and thus have 'borders'. In other words, we have an equivalent of the local form

$$D/G(\sigma), \sigma, V(El) \cup Vc\phi/$$

in the probabilistic game with the painting P.

to considerably enrich the initial label-description of any event $Dr \in U$, but it also allows us to *specify 'physical' borders for the label-descriptions* Dr.

— There are, in general, very many (or even a virtually unlimited number of) aspect-views Vg that 'exist' (in the sense of point (4) in Section 3.5.2) with respect to the entity involved in a label-description $Dr \in U$, and this considerably increases the possibilities for enriching the initial label-description Dr.

— Apart from the limits arising from the requirements of mutual existence (see point (4) in Section 3.5.2), MRC imposes no limits on the number of aspect-*values* $gk(g)$ that can be introduced within a given interval of the semantic axis assigned to an aspect G. So, in general, the density of an axis can be greatly increased, if desired.

— According to MRC, inside the representation space of the view that complexifies the label-descriptions $Dr \in U$, each complexified description constitutes a *discrete* local points-form of space and gk aspect-values (see point (7) in Section 3.5.2). Nevertheless, if the *density* of the relevant aspect-values is sufficiently large, border characteristics may emerge from this local points-form, notwithstanding its discreteness, that are sensitive to semantic attractions/repulsions with respect to the border characteristics of other local points-forms.

For these reasons, within the causal/deterministic *classical* domain:

> An assertion that the label-descriptions $Dr \in U$ *cannot* always be transformed through complexification into a structure that allows us to use them as a sort of abstract jigsaw puzzle would amount to a denial of the principle of causality[†].

In this way, using *trial and error*, we will be able to identify[‡] an individual integrated relative points-form tied to

$$\{Dr\}, \qquad r = 1, 2, \ldots \rho.$$

Of course, in any given case, the details of the complexification procedure will have to be developed specifically for that case using the relevant available scientific knowledge and the factual data that can be established. But this is true for any technological construction process[§¶].

[†] This radical statement is made here in the weak sense of everyday logic and it may not be amenable to a formal proof – this idea deserves further attention.

[‡] The observable semantic attractions/repulsions may not emerge when the random phenomenon under consideration is involved in a game of chance (think of a dice game). However, we are not interested in such cases here since the factual probability distribution of the outcomes is then known by construction. Also, the weak law of large numbers is free from circularity in such cases, and the non-effective character of the frequential definition of the factual probability law is not a problem since the factual law is known in advance. Hence, it seems safe to assert that it is very likely that the law of weak numbers has only been established for games of chance.

[§] This approach is often applied, more or less explicitly, in many important procedures involving the integration of components, such as in anthropological reconstructions, scanning techniques, meteorology and the transmission of television programmes – the novelty here lies in the association with the concept of probability.

[¶] In the case of microstate descriptions, which are basic transferred descriptions in the sense of Section 3.3 and point (6e) in Section 3.5.2, the classical view is *invalid* because there is no worked out *model* for the emergence of the 'transferred' outcomes for which a *primordial* basic description is made, and, in general, the transferred outcomes themselves are just observable marks produced by

5.4.2. *Notation*

5.4.2.1. *General notation.* In this section we shall introduce some general notation we will use in the following rigorous development.

Consider the epistemic referential (G, V^c) where the generator G is the same as in (G, Vr) and $V^c \supset Vr$ is a complexified view that contains the initial view Vr and introduces *supplementary* aspects g and can also introduce supplementary *values* for each of the aspects brought into play. Let

$$U^c \equiv \{Dr^c\}, \qquad r^c = 1, 2, \dots \rho^c,$$

be the universe of *all* the mutually distinct complexifications of the elementary-event-descriptions

$$\{Dr\} \equiv U, \qquad r = 1, 2, \dots \rho\,,$$

that could, *in principle*, emerge relative to the view V^c, where $r^c = 1, 2, \dots \rho^c$, is a new complexified index, each value of which refers globally to *one* of the ρ^c possible mutually distinct structures of space–time-gk-values in V^c. By construction, ρ^c is much larger than ρ, and *one* value of the global index r^c points to a complete points-form of space–time-gk-values (without specifying it explicitly).

The set U^c defined above is not a factual datum, it is only an *a priori*, abstract, *combinatorial*, theoretical universe, in the following sense:

> We do not *know* in advance which complexified descriptions, from amongst all the complexified descriptions $Dr^c \in U^c$ that can, in principle, be constructed with the help of V^c, will be factually realised when the experiment Π is repeated a very large number of times and the outcomes are examined using the complexified view V^c.

Note that our use here of the words 'all' and 'each' does not introduce false absolutes since they just refer to the maximal content given by all the *a priori* possibilities assigned, by construction, to the particular well-defined reference universe U^c, and to which they remain *relative*. This content will be reduced *a posteriori* by the *factually* observed outcomes, but we are obliged to begin in this all-inclusive way.

Consider now, as described in Section 5.4.1, the probability space

$$[U^c, \tau_T^c, p_F(\tau_T^c)]$$

defined on the complexified universe U^c, where τ_T^c is the total algebra on U^c. According to our discussion in Section 5.3.1, the label-descriptions $Dr \in U$ for the random phenomenon (Π, U), with $U \equiv \{Dr\}, r = 1, 2, \dots \rho$, are now located in the algebra τ_T^c. And

measurements, which, though they are located in some space-time domain, are, nevertheless, void of any qualia (that is, of any gk-aspect-view values with $g \neq E$ and $g \neq T$). Furthermore, the postulate of causal determinism cannot be constructed *within* the level of conceptualisation containing the transferred basic descriptions of microstates (Mugur Schächter 2011; 2013). We will return to the consequences of this remarkable fact in Section 8. As we pointed out earlier, this is why we have restricted consideration here to the classical domain.

in that algebra, when compared with the complexified view V^c, they appear as 'simplified' '*event*'-descriptions (or 'label'-descriptions) that have been reduced to qualifications through the coarser label aspect-view $Vr \subset V^c$ only. So now, with respect to the new complexified view V^c, an outcome $Dr \in U$ appears to be realisable through a complete set $\{Dr^c(r)\}$ of relative 'elementary'-event-descriptions $Dr^c(r) \in U^c$, where r is *fixed* and $r^c(r) = 1, 2, \ldots \rho^c(r)$.

However, this too is only an *a priori*, *theoretical*, oversized set of possibilities: since we do not know beforehand which of the elements of U^c will factually emerge when the experiment Π is repeated, we cannot know which elements in the subset of elements of U^c that arise from a given value of r can factually emerge. However, we do know that:

> Two outcomes in $Dr^c(r)$ and $Dr^c(r')$ with index-values $r \neq r'$ are, by construction, mutually exclusive as results of a *single* realisation of Π.

Note that U^c and τ_T^c are just used as a medium for embedding the initial label-descriptions $Dr \in U$ for conceptual development and reference. Without any exterior reference structure, the label-descriptions give an illusory appearance of being absolute, which prevents any further cognitive development.

5.4.2.2. *The representation space for* V^c*.* We will now explicitly construct the representation space for the complexified view V^c. Since each elementary event $Dr^c \in U^c$ describes a physical entity, according to the MRC frame principle (see point (5) in Section 3.5.2), the complexified view V^c necessarily includes a space–time-frame aspect-view $V(ET)$, which can be endowed, in particular, with some definite length-aspect and a corresponding 'length unit' (the smallest perceptible distance); but we are also free to introduce other aspects such as angles.

By construction, the random phenomenon (Π, U) confines the set of all factually possible *space*-coordinates tied to the elementary outcomes $Dr^c \in U^c$ within some given demarcated space domain (or zone) Z. Inside Z, the complexified elementary events $Dr^c \in U^c$ can be as mutually distinguishable as we want because of the practically unlimited precision that can be assigned to spatial qualifications through a virtually free choice of units.

As required by point (7) in Section 3.5.2, *all* the semantic axes are endowed with aspect-values, which may or may not be measurable. We assume that the order of the aspect-values $gk(g)$ (see Section 5.2.2 for the notation) on each semantic dimension g is specified, even if it is only done arbitrarily.

5.4.2.3. *The points-grid of* V^c*'s representation space.* The discrete sets of $gk(g)$ values carried by the semantic axes g mean that V^c's representation space defines a multi-dimensional finite grid of multi-dimensional points $\gamma(V^c)$. The oversized character of the complexified universe U^c ensures that this grid $\gamma(V^c)$ can *enclose* the spatial zone Z with a *margin*: here we assume there is a margin around Z.

By construction, the representation on $\gamma(V^c)$ of a given complexified description $Dr^c \in U^c$ consists of a local points-form, and a *single* value of the complexified index

$$r^c = 1, 2, \ldots \rho^c,$$

points to *one* of these ρ^c possible points-forms in $\gamma(V^c)$†.

5.4.3. *Algorithm to identify factually the probability law associated with a random phenomenon*

Consider the representations on $\gamma(V^c)$ of the local points-forms labelled by the values of the complexified index $r^c = 1, 2, \ldots \rho^c$.

We have shown that if we choose aspect-views g and space–time-gk-values appropriately, the semantic border-characters of these local points-forms will display 'semantic attractions/repulsions' that are sufficiently definite to allow us to make a sort of abstract jigsaw puzzle on $\gamma(V^c)$. However, if the border-characteristics of the points-forms in the grid $\gamma(V^c)$ do not allow us to make a jigsaw puzzle, then we must choose other supplementary aspects g and/or space–time-gk-values (the aspects, units for the aspect-values or just a qualitative density of the aspect-values, and so on). So, finally, through trial and error, we will reach a position where we can make an abstract jigsaw puzzle on $\gamma(V^c)$.

In the following, we will assume that we have established the possibility of making an abstract jigsaw puzzle on $\gamma(V^c)$, and will construct an algorithm based on this possibility.

This algorithm consists of two steps, which also both involve trial and error.

5.4.3.1. *Step 1: semantic integration of replicas of a relative 'points-form' $\phi(\Pi, U)$ tied to (Π, U).* We can use as many replicas of the points-grid $\gamma(V^c)$ as we need, and will denote each of them by $\gamma k(V^c)$, where k is an integer from a finite index-set $k = 1, 2, \ldots K$, with cardinal K (note that this k is *not* the same as the index 'k' used in connection with a value (gk) of an aspect g).

We begin by performing the experiment Π. Using the complexified view V^c, the first outcome $Dr \in U$ is observable as a local *complexified* description $Dr^c \in U^c$ that, inside τ^c_T, 'realises' that particular $Dr \in U$ (in the sense defined in Section 2.1). *By construction, this complexified description Dr^c is* **unique** *among the mutually exclusive ρ^c possible elements Dr^c in U^c.*

The 'local points-form' corresponding to the outcome Dr^c of the first realisation of the experiment Π is also unique by construction. It is also bounded within any replica of the points-grid $\gamma(V^c)$, and it can *certainly* be represented on the *first* replica of $\gamma 1(V^c)$ of $\gamma(V^c)$, since *none* of the replicas have been used yet.

We now repeat the experiment a number of times. Each new outcome Dr^c will be represented:

(a) On the first replica $\gamma 1(V^c)$ of the grid $\gamma(V^c)$ *if* the *unique* domain reserved there for the new local points-form corresponding to this new outcome is still *unoccupied.*

In this case, the position on $\gamma 1(V^c)$ where the new local points-form has to be located will be identified through the semantic-attractions/repulsions of its border-characteristics with respect to the border-characteristics of the other local points-

† Like the label-index $r = 1, 2, \ldots \rho$, the index $r^c = 1, 2, \ldots \rho^c$ is just a *label*-index, and it is only when the local form appears displayed on the grid $\gamma(V^c)$ that the value of the index r^c can point to it; otherwise it remains hidden in (or folded into) the value of r^c.

forms already present on $\gamma 1(V^c)$. This identification involves trial and error, and sometimes *waiting for new clues*, exactly as in the jigsaw puzzle game.

(b) Otherwise, on the *next* (in k order) replica $\gamma k(V^c)$ that we can find with a vacancy in the reserved domain where the new outcome Dr^c can be placed.

Now, by construction, the grid $\gamma(V^c)$ encloses the domain Z of space where, by the definition of the random phenomenon being studied, the *factual* outcomes $Dr \in U$ can appear. So, the grid $\gamma(V^c)$ can hold the representations of *all* the unique and mutually exclusive outcomes $Dr^c \in U^c$ that are *in principle* (that is, on purely combinatorial grounds) possible within the epistemic referential (G, V^c). In consequence of this, the postulated unknown complete 'points-form' $\phi(\Pi, U)$ is entirely contained (potentially) within $\gamma(V^c)$, and within some confined boundary. The procedure defined by points (a) and (b) above leads to the *progressive and observable* emergence of this boundary on the first representation $\phi_1(\Pi, U)$ of $\phi(\Pi, U)$, which is realised on the first replica $\gamma 1(V^c)$ of $\gamma(V^c)$.

Moreover, a series of other, increasingly *incomplete*, reproductions of the same unknown points-form $\phi(\Pi, U)$ will emerge and become observable on subsequent replicas

$$\gamma 2(V^c), \gamma 3(V^c), \ldots \gamma k(V^c), \ldots \gamma K(V^c)$$

of the points-grid $\gamma(V^c)$. So, at any time, we shall be able to perceive a series of K increasingly incomplete reproductions $\phi_K(\Pi, U)$ of the points-form $\phi(\Pi, U)$.

Sooner or later, but in a *finite* time (because everything in MRC is finite by construction), we will observe that the first representation $\phi_1(\Pi, U)$ of $\phi(\Pi, U)$ appears to have stopped evolving, while the number N of repetitions of the procedure Π continues to grow, and continue to feed the growth of the points-forms labelled by $k = 2, 3, \ldots K$. In this way, we shall *know* that $\phi_1(\Pi, U)$ almost certainly 'completely' materialises the unknown form $\phi(\Pi, U)$ in an observable way.

At this point we will pause briefly to explain exactly and explicitly what we mean by 'almost certainly completely', and why we only say that '$\phi_1(\Pi, U)$ of $\phi(\Pi, U)$ "appears" to have stopped evolving'.

We employ these verbal hedges because nothing can absolutely exclude the possibility that at some future time, as N increases, there will be an outcome Dr^c for which the position (and the corresponding neighbourhood) on $\gamma 1(V^c)$ is still free (in particular, this may be on the boundary of $\phi_1(\Pi, U)$), and thus an entirely unexpected new element will be added to the unknown points-form $\phi(\Pi, U)$. All our considerations are affected by the possibility of this occurring.

However, *all* natural laws are affected by similar possibilities. And it is always possible that there may be even more radical future modifications of $\phi(\Pi, U)$, in the same way as for *any* 'natural law'.

Probability distributions and natural laws are only local conceptual constructs, which are only asserted as 'certainties' on the basis of an implicit pragmatic assumption of 'stable conditions', which are never valid everywhere and for all time (Mugur Schächter 2002c, pages 291–303). A certain, definitive, absolutely stable *factual* truth can never be constructed conceptually. Only syntactic 'truths' can be absolutely stable, but they are just the conclusions of deductions and are thus, in fact, merely expressions of the *logical internal consistency of the syntactic system* and not observational factual truths†. Moreover, the sort of stability of a syntactic 'truth' in this sense, that is, a formal internal consistency (a misused word, in such a context), even though it can *last* indefinitely, nevertheless is as inherently relative as a factual truth is: it is relative to the necessarily restrictive set of axioms and definitions that define the syntax in which the formal truth has been established. And the factual truth of these varies in the long run, and in a way that cannot be avoided. The basic, primary data lie beyond the realm of pure reason, and are inherently statistical.

5.4.3.2. *Step 2: controlling the uniformity of the distribution of the elements Dr^c of U^c.* Suppose we have continued to increase the number N of realisations of the experiment Π, and have now produced a large number, say K, of points-forms that are *all identical* to $\phi_1(\Pi, U)$ – these will be followed by a series of increasingly incomplete emerging points-forms of representations of outcomes Dr^c connected through semantic border continuities. We will now denote *any* of these K mutually identical points-forms by

$$\sim_K \phi(\Pi, U)$$

where '$\sim_K$' is to be read as saying 'almost certainly, on the basis of K mutually identical points-forms'. We will also call the concept denoted by $\sim_K \phi(\Pi, U)$, *'the' K-representation of the unknown points-form* $\phi(\Pi, U)$ (note the definite article).

Suppose now that while N is being increased still further, and despite the 'almost certainty' of completeness expressed above, we observe, very unexpectedly, a new point-event Dr^c for which the location in $\gamma 1(V^c)$ is still *unoccupied*, even though the content of $\gamma 1(V^c)$ has remained unchanged for a very long sequence of repetitions of Π, thus contradicting the assumption represented by $\sim_K \phi(\Pi, U)$. How should we react?

Our first remarks will continue the discussion of Step 1 of the procedure. This new observation reveals that the new point-event $Dr^c \in U^c$, despite being unexpected, appears to be *factually* possible in the conditions produced by the experiment Π. In other words, its factual probability, whatever it may be, is *non-zero*. But saying that the new Dr^c has a non-zero probability is equivalent to saying that if we greatly increase the number of repetitions N of the experiment Π, then Dr^c will *almost certainly* reappear from time to time, so that, over a sufficiently long period, it will progressively occupy its place in *all* K previously constructed identical replicas of $\phi(\Pi, U)$ used as the basis for the concept '$\phi_K(\Pi, U)$'. This amounts to a retroactive modification of the concept '$\phi_K(\Pi, U)$'. But we prepared for this possibility by specifying the *relativity to* K of the concept represented by $\sim_K \phi(\Pi, U)$. On the basis of this reasoning, *any* event Dr^c that finds its place free in $\phi_1(\Pi, U)$, even after a very long period of apparent saturation, can be immediately reproduced on all other registered replicas of $\phi(\Pi, U)$. This mutability

† We define an 'observational truth' to be an observed datum that is confirmed by comparison with natural circumstances, and not just with the internal consistency requirements of a purely formal construct, which requires its own definition tied to formal internal consistency (consider Gödel's theorem).

of the representations of $\phi(\Pi, U)$ is entailed by the fact that, unlike the particular case of the probabilistic game with the painting P, in the general case treated here, not only is the factual distribution of probabilities tied to the presupposed meta-form $\phi(\Pi, U)$ unknown at the start, but the *meta-form itself* is also unknown.

Now that we have completed the discussion of Step 1, we will move on to consider the specific objective of Step 2, namely, the possibility of ensuring the uniformity of the numerical distribution of the elements $Dr^c \in U^c$ of the universe U^c of elementary events, which we have used as the basis for the Kolmogorov probability space

$$[U^c, \tau_T^c, p_F(\tau_T^c)]$$

associated with the random phenomenon under investigation.

The fact that the new point-event $Dr^c \in U^c$ was 'unexpected' draws attention to the fact, already mentioned in Section 5.3.2, that a generalisation of the approach taken in Section 4 is only conceivable if, on the basis of *factual* data, the elementary events of U^c (that is, the 'elements' thought of as being mutually distinguished within the complete form $\phi(\Pi, U)$) can be assumed to be themselves *uniformly* distributed within that form. In the case of the probability game with the painting P, this distribution was *known* by construction to be uniform, and this was a necessary condition for asserting the factual probability given by (5) on Page 50, since it enabled us to know the total number N_{PT} of puzzle pieces in P. But in the general case we are now considering, we know nothing about the distribution of the complexified descriptions Dr^c in U^c. And if a point-event Dr^c that still has an empty location in $\phi_1(\Pi, U)$ emerges after a long sequence

$$\phi_1(\Pi, U) \equiv \phi_2(\Pi, U) \equiv \ldots \equiv \phi_K(\Pi, U)$$

of representations of $\phi(\Pi, U)$ has already been completed, then this point-event is almost certainly *less frequent* in U^c than the other local complexified descriptions $Dr^c \in U^c$ making up the K preceding representations of $\phi(\Pi, U)$ and used as the basis for the concept $\sim_K \phi(\Pi, U)$.

Now, the general relative and constructive features of the framework we are using suggest that the factual procedure for dealing with this problem of *true* uniformity is as follows:

> In the above circumstances, where we observe a new 'unexpected' outcome after a long period, we shall interpret it as *a manifestation of the possibility that the structure assigned to the complexifying view V^c that led us to the considered universe U^c is not yet fully pertinent to our aim*†.

So we shall need to correct for this through the tentative use of a modified or enriched complexifying view V^c. In this way (by trial and error, and with respect to a number K that, according to whatever requirements we have chosen for the degree of stability, and thus the degree of certainty, is deemed to be satisfactory in the search for a pertinent complexifying view V^c, and relatively to this number K), we will finally reach an

† Another possibility is that there is some long-term 'fluctuation' in the order of the outcomes, but, for simplicity, we will not consider this possibility here.

acceptable assumption of 'K-uniformity' of the distribution of the elements Dr^c in U^c, in terms of which the complete unknown form $\phi(\Pi, U)$ is conceived[†]. Since this process does not involve any infinite operations, it is *neither* non-effective nor arbitrary – it is just an empirical procedure.

There is no more 'correct' or more 'exact' way to deal with this problem. When we reach the limits of the syntax we are using and are faced with factuality, we have to import semantic material, which inherently contains fluctuations and uncertainty.

> When we try to construct a concept of factual *probability*, any requirement of strict deductibility everywhere, and, in particular, a requirement for a smooth stable progression in the features of the sequences of factual results, would simply be a *contradiction*.

We will call the procedure outlined above the *K-relative algorithm of semantic integration of the form* $\phi(\Pi, U)$ that we have assumed to be associated with the random phenomenon being studied.

5.4.4. *The relative factual K-probability law to be associated with a random phenomenon*
We can now use the same reasoning as in Sections 4.1.5 and 4.2 for the probabilistic game with the painting P to try to identify the factual numerical probability distribution $\{p_F(U)\}$ to be asserted on the universe of events.

In the same way as we did for the completed picture P, we can now *count* the *total* number of complexified outcomes Dr^c that have been factually realised in $\phi_K(\Pi, U)$. We will denote this number by $n^c_{\phi(K)T}$. We can also *count*, for any *given* description $Dr \in U$ viewed as an event in the algebra of events τ^c_T, the number, say $n^c_{\phi(K)T}(r)$, of 'realisations' of *that* Dr *inside* $\phi_K(\Pi, U)$: namely, a specific complexified factually realised outcome in the set of all the factual outcomes corresponding to a fixed value r. We will denote this set by

$$\{Dr^c_{\phi(K)}(r)\},$$

with r fixed and

$$r^c_{\phi(K)}(r) = 1, 2, \ldots \rho^c_{\phi(K)}(r)$$

an index in an index-set of cardinal $\rho^c_{\phi(K)}(r)$. Note that

$$\rho^c_{\phi(K)}(r) \neq \rho^c(r).$$

> Unlike the cardinal $\rho^c(r)$ of the set $\{Dr^c(r)\}$, with r *fixed* and
>
> $$r^c(r) = 1, 2, \ldots \rho^c(r),$$
>
> in the universe U^c on which we founded the Kolmogorov space $[U^c, \tau^c_T, p_F(\tau^c_T)]$ associated with the random phenomenon (Π, U), the cardinal $\rho^c_{\phi(K)}(r)$ does not

[†] This form itself will *never* be *known* with 'definitive accuracy and absolute certainty'. Indeed, according to MRC, the concept of definitively true and certain knowledge of a physical entity is a false absolute (Mugur Schächter 2002a; Mugur Schächter 2002b; Mugur Schächter 2006). However, remember that it is only the (strictly non-qualified) existence of a form $\phi_K(\Pi, U)$ that is methodologically *assumed* in the present approach.

count purely theoretical combinatorial possibilities, but counts elementary-event-descriptions that are *factually* realised within the meta-form $\phi_K(\Pi, U)$ and endowed *there* with a *factually* uniform numerical distribution.

To stress this important fact, we will write

$$\rho^c_{\phi(K)}(r) \equiv_F n_{\phi(K)T}(r),$$

where the sign '$\equiv_F$' is to be read as saying 'factually identical to'.

By construction, a complexified description Dr^c that is an element of the set

$$\{Dr^c_{\phi(K)}(r)\}$$

cannot also be an element of a set

$$\{Dr^c_{\phi(K)}(r')\}$$

where $r' \neq r$. Furthermore, we have

$$n_{\phi(K)T} = \sum_r n_{\phi(K)}(r)$$

$$\sum_r n_{\phi(K)T}(r)/n_{\phi(K)T} = 1$$

where $r = 1, 2, 3, \ldots, \rho$. Therefore, using reasoning that is strictly analogous to that developed in Section 4.2 for the probabilistic game with the picture P, but now quite *generally* based, we can assert that, on the basis of a probability game with the elements $Dr^c_{\phi(K)}$ of the finite and confined representation $\phi_K(\Pi, U)$ of the unknown 'points-form' $\phi(\Pi, U)$, the factual numerical definition of the probability of an event Dr is the *rational* number

$$p_F(r, K) \equiv n_{\phi(K)}(r)/n_{\phi(K)T}.$$

So we will call the set, relative to K, of *rational* and *factually* identified numbers with normalisation 1,

$$\{p_F(r, K)\} \equiv \{n_{\phi(K)}(r)/n_{\phi(K)T}\}, \qquad r = 1, 2, \ldots s, \tag{7}$$

the *factual K-representation of the distribution of probabilities associated with the random phenomenon* (Π, U)†.

5.4.5. *Conclusions for Section 5*

The procedure that led to the definition in (7) was developed entirely within the MRC framework. This procedure is effective, non-circular and *strictly factual in its specific content*: all components of (7) that are specific to the random phenomenon (Π, U) under investigation, though expressed in terms of a preconstructed general representational framework, are drawn exclusively from factually observed data produced by (Π, U). Any necessity for a possibly unending confrontation between an *a priori* assertion of a uniform

† Computer simulations might enable us to rapidly organise the whole substratum required for the calculations in (7).

distribution of the elements of U^c and an *a posteriori* non-effective 'verification' of this assertion has been removed. We have generated a unique finite sequence of outcomes of the experiment $\Pi \equiv [G.V]$, which is controlled by the 'natural' reasoning inherent in everyday thought.

We will call this general procedure the *algorithm of factual determination of the probability law to be tied to a given random phenomenon*, and we will denote it by the symbol

$$\mathcal{A}\left[\{p_F(r,K)\} \leftrightarrow (\Pi, U)\right].$$

Note that the form of this symbol shows explicitly that this algorithm includes within it the K-relative algorithm of semantic integration of the unknown form $\phi(\Pi, U)$.

According to MRC, *any* communicable and consensual knowledge is a description. So it is precisely a points-form of space–time and aspect-values endowed with some invariance (see points (5), (6) and (11) in Section 3.5.2). Moreover, any relative description, whether it be 'individual' or 'statistical' in the sense of MRC, involves, by definition, *many* reproductions of the sequence $[G.V]$ of the two epistemic operations G and V in the corresponding epistemic referential (G, V). These conceptual uniformities in MRC tend to obscure the *clear* outlines of the concept of an 'individual' description.

However, the way we have constructed the algorithm $\mathcal{A}\left[\{p_F(r,K)\} \leftrightarrow (\Pi, U)\right]$ raises a number of issues, which we will discuss in the following section.

5.4.5.1. *On intelligibility.* When a description is individual (in other words, immediately global, like a description of an apple perceived on a table), the corresponding form of space–time aspect-values is usually immediately 'understood' (so, for example, it can be immediately inserted in 'causal chains'). Whereas, in the case of a description tied to a natural statistical situation or a random phenomenon (Π, U) in the sense of MRC, the global physical situation involved generally *evades* any direct and complete perception because it involves features that are partly hidden or too small, too large or too complex to be immediately grasped by a human mind and composed into 'one' organised structure. Only fragments of this complete structure are directly perceived and apprehended. And for these fragments, certain relative connecting features (semantic continuities, *distances, angles, relative durations*† and so on) are hidden, that is, they are *filtered out from direct perception.* The features that can be directly observed on these fragments are massively cut off from any conceivable global source structure. What we have denoted by $Dr \in U$ and called a 'label-description' is precisely such a fragment carrying features that appear to be cut off from any known or immediately guessed global structure. These cuts destroy *intelligibility.* They even destroy the capacity to merely *imagine* the possibility of the existence of a related factual 'whole' that might ensure intelligibility.

> This is because the fragments denoted $Dr \in U$ belong to a level of conceptualisation that is different from the level where the global form $\phi(\Pi, U)$ is postulated to exist.

† For simplicity, we have generally restricted the discussion to spatial features, but everything can be straightforwardly extended to space–time features.

Indeed, generally speaking, we are mentally immersed exclusively in the level of conceptualisation where we perceive the fragments from which we construct statistical/probabilistic descriptions. And on that level, we are blind to any relative representation of a complete $\phi(\Pi, U)$ tied to the random phenomenon $(\Pi, U)^{\dagger}$.

The relative outcome frequencies of the observed events $Dr \in U$, through their observable tendency towards a certain degree of *stability* when the number of repetitions N of the procedure Π increases, construct, but only a bit at a time through a sequence of disjointed random effects, a fundamentally cryptic and purely numerical representation of the unknown physical whole that can be thought of as the source of the events $Dr \in U$. These relative frequencies cannot carry the meta-qualia tied to non-perceptible meta-contours: they do not *exist* (in the sense of point (4) in Section 3.5.2) with respect to a view where the relational aspects between the various events $Dr \in U$ are not defined. They can only generate a sort of coded, random and asymptotic 'reading', which can only convey fragmented signals that suggest the relevance of a meta-view sensitive to relational aspects. Any tendency in these signals to show regularities may then induce a feeble and fluctuating idea of the possibility of a unique source that somehow acts on the evolution of the observed relative frequencies. However, even with this idea, and even *a posteriori*, once the 'reading' process has been completed (which it must be at some time since, like any human action, it is finite), if the collection of relative frequencies provided by the process is reconsidered for all the events Dr at the same time, the values of the relative frequencies being now known to have reached a certain degree of stability relative to the parameter K, the factual probability distribution

$$\{p_F(r, K)\}, r = 1, 2, \ldots \rho,$$

defined by use of the recorded relative frequencies will *still* only yield a *meaningless* numerical expression.

It is at this point that we see the remarkable role played by the *'complexifications'* $Dr^c(r) \in U^c$ constructed for the directly perceived event-descriptions $Dr \in U$. These complexifications (through embedding, reference and local semantic-attractions/repulsions at their borders) are what enabled us to climb to a meta-level of conceptualisation where we could gain access to a representation $\phi_K(\Pi, U)$ of the postulated global form $\phi(\Pi, U)$.

But in order to benefit from this embedding and reference, the usual connection, for which we will use the symbol

$$(\Pi, U) \rightsquigarrow [U, \tau, p_F(\tau)],$$

has had to be replaced by the connection

$$(\Pi, U) \rightsquigarrow [U^c, \tau^c_T, p_F(\tau^c_T)]$$

$\dagger$ We have evolved on the level of our direct perceptibility like ants, for whom the tiniest clods of earth obliterate the view, and for whom the tiniest crevice acts as an abyss. We cannot perceive, and do not imagine, the meta-contours of large entities that would become clear from a more comprehensive point of view. In order to become aware of such meta-contours, we need to take off in an appropriate conceptual vehicle that can bring us up to a level where these contours can be perceived and defined.

defined in Section 5.3.1. And in order to make use of the local semantic-attractions/repulsions at the borders, we have had to add *definite factual data* to the basic universe U^c and the algebra τ_T^c.

Of course, the various conventions and descriptional relativities involved in the construction of any meta-form $\phi_K(\Pi, U)$ mean that such a meta-form is only one element of a large class of various possible representations of $\phi(\Pi, U)$: the choice of the meta-view V^c is never unique. Moreover, any element in this class *still* only offers a coded, though now integrated, representation of the postulated unknown form $\phi(\Pi, U)$. And, again of course, the qualia implied by a representation $\phi_K(\Pi, U)$ of the unknown form $\phi(\Pi, U)$ are only represented by symbols: in general, they too are *not* perceivable.

Nevertheless, it remains the case that when we combine

— the mutual individualisation of the various complexified outcomes $Dr^c(r) \in U^c$ of any given $Dr \in U$,
— the possibility of representing the semantic location of each of these complexified outcomes, and, above all,
— the possibility of, finally, *counting* these outcomes inside a *closed* relative representation of a whole that contains a *finite* number of outcomes $Dr^c(r)$,

we can specify, in a non-circular way, the relative factual probability law

$$\{p_F(r, K)\}, r = 1, 2, \ldots \rho,$$

tied to (Π, U), and, moreover, we can gain, through $\phi(\Pi, U)$, some global *understanding* of the general concept of a factual probability law, and some clear unambiguous *significance* that can be assigned to a conveniently selected meta-form $\phi(\Pi, U)$.

We are confident that over time the procedure in Section 5.4.3 will be simplified and improved, and its current complexity is just a temporary flaw resulting from its novelty.

As for the factual 'truth' of the existence of the postulated unknown form $\phi(\Pi, U)$, we can regard it as consisting of the fact that, while it generates intelligibility, its conceptual and factual consequences can be 'verified' *consensually.* However, this is not a problem, since factual truth is never more than consensual for the physical theories constructed by humans; not even when what we qualify is directly observable, because even then, it is relative to the language we use, which is itself consensual, and which in its turn is relative to the biological processes through which we perceive and reason.

6. The weak law of large numbers versus the algorithm $\mathcal{A}\left[\{p_F(r, K) \leftrightarrow (\Pi, U)\right]$

In this section, we shall examine the relationship between the algorithm

$$\mathcal{A}\left[\{p_F(r, K) \leftrightarrow (\Pi, U)\right]$$

and the weak law of large numbers.

6.1. *Mutual characterisation*

We will begin by restating some characteristic features of the weak law of large numbers:

$$\forall r.\, \forall(\epsilon, \delta).\, \exists N_0.\, \forall N.\; (N \geq N_0) \Rightarrow \boldsymbol{\mathcal{P}}\left[(|n(e_r)/N - p(e_r)|) \leq \epsilon\right] \geq (1 - \delta). \qquad (8)$$

In (8), the symbol $p(e_r)$ presupposes:

(a) The *existence* of a numerically *unspecified* mathematical limit, denoted by $p(e_r)$, towards which the measured relative frequencies $n(e_r)/N$ converge in consequence of some assumed mathematical conditions, together with a 'frequential *definition*' of the abstract concept of the 'probability of the event e_r' that consists of precisely that limit, for any fixed value of the index r.

(b) A characterisation (see the footnote on Page 6) of, exclusively, the well-known abstract concept of a 'probability measure', which only requires *general* and global conditions on the set of mathematical limits $\{p(e_r)\}, r = 1, 2, \ldots s^{\dagger}$: however, the elements of this set are left numerically unspecified, and thus completely lacking any definite connection to any particular factual probabilistic situation.

Mutatis mutandis, the same holds for the abstract meta-probability $\boldsymbol{\mathcal{P}}$.

Points (a) and (b) reflect a non-effective point of view, which is located on a general level of abstract conceptualisation that has been deliberately emptied of any individual factual semantic content to ensure maximal generality so that it can act as a mathematical vehicle for *any* factual probabilistic data.

Considered globally, (8) represents the syntactic evolution of the relationship between the mathematical limit $p(e_r)$, which is initially completely unspecified numerically, and the relative frequency $n(e_r)/N$, as the integer N counting the number of completed repetitions of the experiment Π under consideration tends towards infinity in uniform and ordered steps of one unit. Each step progressively injects some (fluctuating and globally uncertain) numerical content into the initially numerically unspecified sign $p(e_r)$. This uniform progression of N is assumed to extend over the whole abstract interval from 1 to ∞. It does not distinguish between *a priori* hypothetical factual assertions and *a posteriori* factual findings; it just goes on towards ∞ in identical steps. The fluctuations in the way the value of the ratio $n(e_r)/N$ evolves are free of any semantic regulation. The only control over fluctuations within the framework of (8) is expressed in purely syntactic terms, and consists exclusively of the meta-*probable* restrictions imposed by the pair of arbitrarily small real numbers (ϵ, δ).

We will now compare this with the algorithm $\mathcal{A}\left[\{p_F(r,K)\} \leftrightarrow (\Pi, U)\right]$.

Our assumption postulating the existence of a global meta-form $\phi(\Pi, U)$ tied to any given random phenomenon (Π, U) induced a peculiar sort of semantic K-*quantification* into the evolution of the number of repetitions N of the experiment Π. From a purely numerical point of view, the semantic integration sub-algorithm giving $\phi_K(\Pi, U)$ representations of $\phi(\Pi, U)$ maintains the uniform, ordered and unlimited progression of the integer N introduced by equation (8). However, this algorithm also takes into account the semantic content of all the elementary complexified outcomes

$$Dr^c(r), r^c = 1, 2, \ldots \rho(r), \qquad r = 1, 2, \ldots \rho,$$

and on the basis of these, it selects the *place* for each outcome Dr^c on the appropri-

† We speculate that the integrability conditions required for the laws of large numbers may be connected with an intuitive perception of a meta-whole reflected in a probability measure.

ate replica of the grid $\gamma(V^c)$, quite *independently* of when it appears in the sequence of $1, 2, 3, \dots N$ repetitions of the experiment Π. In this way, through the semantically based positioning of the outcomes, the sub-algorithm for identifying the $\phi_K(\Pi, U)$ representations of $\phi(\Pi, U)$ organises, out of the steadily increasing sequence of repetitions of the procedure Π, the perceptibility of *'K-quanta' of significance* with respect to the definability of the factual probabilities $p_F(r, K)$. Given our aim of determining such significance, it focuses attention on the way the successive representations

$$\phi_1(\Pi, U), \phi_2(\Pi, U), \phi_3(\Pi, U), \dots \phi_K(\Pi, U)$$

of $\phi(\Pi, U)$ emerge through a *semantic treatment* of the outcomes. This treatment, in turn, punctuates the featureless flow of results from the uniformly increasing number N of realisations of Π: specifically, N tends to be split progressively into a sequence of mutually exclusive sets with identical cardinals $n_{\phi(K)T}$.

Even though, at any given time after we have achieved what we believe to be a set of completed and stabilised K-representations of $\phi(\Pi, U)$, we cannot know with 'absolute certainty' (which is a non-effective concept) that the set of elementary complexified outcomes Dr^c will not acquire any new elements if N continues to be increased, the K completed and thus far apparently stable, representations

$$\phi_1(\Pi, U), \phi_2(\Pi, U), \phi_3(\Pi, U), \dots \phi_K(\Pi, U)$$

already mark a quantification of N that, *retroactively*, obscures the order in which the outcomes Dr^c emerged†.

> The semantically controlled positioning of the representation of each result Dr^c produced by the uniform increase in the number N of repetitions of the procedure Π, *removes the need for a never-ending (a priori)/(a posteriori) dialog*‡.

6.2. *Compatibility between the weak law of large numbers and $\mathcal{A}\left[\{p_F(r, K)\} \leftrightarrow (\Pi, U)\right]$, and their unification*

In this section we will *not* give any *formal proofs* but just an informal, though explicit, account of the syntactic and semantic compatibility existing between the weak law of large numbers, as given by (8), and the algorithm $\mathcal{A}\left[\{p_F(r, K)\} \leftrightarrow (\Pi, U)\right]$. On the basis of this compatibility, we will combine them into a unified expression in mathematical terms, which will be associated with the semantic/syntactic connection

$$(\Pi, U) \rightsquigarrow [U^c, \tau_T^c, p_F(\tau_T^c)]$$

† Note added in proof: The contribution by Christopher Porter to this Special Issue of the Journal describes some work of Kolmogorov in which, under some definite constraints, which are tied to a different objective, certain finite sequences are selected *a posteriori* from an indefinitely long sequence of signs. Given the similarities, it would be interesting to know whether there is some non-trivial methodological connection between his procedure and ours.

‡ To date, the necessity, at least in principle, for such an endless dialog has not been removed. In fact, the question has often just been ignored by simply assuming from the start that the elements of the universe produced by the random phenomenon under consideration are uniformly distributed. However, this does not resolve the problem, which remains in principle, and largely unaddressed.

involved in $\mathcal{A}\left[\{p_F(r,K) \leftrightarrow (\Pi,U)\right]$.

This unified expression will provide a complete formalised global representation of the essential syntactic *and* semantic features of the concept of probability we have proposed in the current paper. The rather unusual operations in this section may be regarded as just a fresh manifestation of the general constructive approach used throughout the current paper.

The semantic features introduced by the algorithm $\mathcal{A}\left[\{p_F(r,K) \leftrightarrow (\Pi,U)\right]$, together with the effects they create and the results they produce, will allow us to rewrite (8) in a new form. To do this, we will begin by introducing some new notation:

$$e_r \equiv Dr \equiv r \tag{9}$$

and

$$\begin{aligned} N &\equiv Kn_{\phi(K)T} + N', \\ n(r) &\equiv Kn_{\phi(K)T}(r) + n'(r) \end{aligned} \tag{10}$$

In (10), the term $Kn_{\phi(K)T}$ represents, for a given integer K, the *number* of unit steps (whatever their results) in N that contributed to the construction of K mutually identical representations $\phi_K(\Pi,U)$ of the postulated unknown whole form $\phi(\Pi,U)$. These unit steps will have introduced $Kn_{\phi(K)}(r)$ complexified realisations $Dr^c(r)$ of any given particular event $Dr \equiv r$. The term N' is the number of other unit steps in N, which will have contributed to the emergent, but incomplete, new representations of $\phi(\Pi,U)$. These N' supplementary steps will have introduced some supplementary unknown global number $n'(r)$ of complexified elementary realisations $Dr^c(r)$ of the same event $Dr \equiv r$ considered above. These supplementary N' unit steps are not yet endowed with the semantic significance given by the definition (7) of a probability $p_F(r,K)$, which is only valid for an already (nearly certainly) completed replica of the form $\phi(\Pi,U)$.

We will now also make the tentative assumption of a 'nearly constant and exact equality' between the two numbers $p(e_r)$ and $p_F(r,K)$ when K is 'sufficiently' large. It is assumed that in each specific case there will be some declared criteria (denoted by c) defining the meaning of the term 'sufficiently' large – such criteria are similar to those used to define the required 'precision' of measurements.

We can write this assumption in the form

$$p(e_r) \approx_c p_F(r,K), \qquad \text{for sufficiently large } K, \tag{11}$$

where the sign '$\approx_c$' is to be read as saying 'equal in the sense of the criteria denoted by c'.

The *syntactic* concept of probability in (8) has a non-effective and circular 'frequential' definition given (exclusively) by its numerical value in the form of the *real* numerical value *limit* $p(e_r)$. By contrast, (11) provides a *factual* definition, which is only relative, and may not be strictly stable, but is, at any specific time, effective (finite) and non-circular, and possesses a definite significance in the form of the *ratio*

$$n_{\phi(K)}(r)/n_{\phi(K)T},$$

which, inside *any* already completed representation $\phi_K(\Pi,U)$ of a 'whole' $\phi(\Pi,U)$ associated with the random phenomenon under investigation, refers to the counted number $n_{\phi(K)}(r)$ of outcomes $Dr^c(r)$ and the counted total number $n_{\phi(K)T}$ of outcomes Dr^c contained in $\phi_K(\Pi, U)$ (note the clear distinction between the ratio $n_{\phi(K)}(r)/n_{\phi(K)T}$ and the relative frequencies $n(r)/N$ appearing in (8)).

We would now like to substitute (11) into (8), but, at first sight, it is not obvious that we can do this without introducing some numerical inconsistencies since, unlike $p(e_r)$, the number $p_F(r, K)$ is not an unknown but fixed limit, but has a known and definite numerical value, which, in principle, might *evolve*. However, we can say:

— The only direct effect of substituting (11) into (8) is on the numerical value of the absolute difference $|n(e_r)/N - p(e_r)|$.

— Moreover, the numerical difference between the absolute value $|n(e_r)/N - p(e_r)|$ in (8) and the absolute value

$$\left|\frac{n(r)}{N} - p_F(r, K)\right|$$

we get by substituting (11) (and (9)) in (8) can always be absorbed in the conditions specified on the left-hand side of formula (8) and the fact that, formally, (8) only asserts two *inequalities* on the right-hand side of the implication (one of which, moreover, involves a meta-probability).

Indeed, (8) is not an equation, and it is unaffected by any possible numerical discrepancies arising from the substitution (11) because of the (ϵ and δ) *approximations* and the (meta)-*probabilistic* character given by $\boldsymbol{\mathcal{P}}$, which allow it to support the *a priori* possible fluctuations of what we have called 'the almost certain completeness' of the representation $\phi_K(\Pi, U)$ of $\phi(\Pi, U)$. Therefore, it can equally support the possible numerical fluctuations of the value of $p_F(r, K)$, as well as the existence, in general, of fluctuations in the values of the numbers $n'(r)$ and N' while K is increased.

Hence, we *can* now use (9)–(11) to rewrite the absolute difference $|n(e_r)/N - p(e_r)|$ appearing in (8) as:

$$\left|\frac{n(r)}{N} - p_F(r, K)\right| = \left|\frac{Kn_{\phi(K)}(r) + n'(r)}{Kn_{\phi(K)T} + N'} - \frac{n_{\phi(K)}(r)}{n_{\phi(K)T}}\right|. \tag{12}$$

This yields:

$$\begin{aligned}&\forall r.\,\forall(\epsilon, \delta).\,\forall K.\,\exists N_0.\,\forall(Kn_{\phi(K)T} + N').\ ((Kn_{\phi(K)T} + N') \geq N_0) \Rightarrow \\ &\qquad \boldsymbol{\mathcal{P}}\left(\left|\frac{Kn_{\phi(K)}(r) + n'(r)}{Kn_{\phi(K)T} + N'} - \frac{n_{\phi(K)}(r)}{n_{\phi(K)T}}\right| \leq \epsilon\right) \geq (1 - \delta).\end{aligned} \tag{13}$$

Now, when the number K (and thus N by (10)), is increased indefinitely,

$$\left|\frac{Kn_{\phi(K)}(r) + n'(r)}{Kn_{\phi(K)T} + N'} - \frac{n_{\phi(K)}(r)}{n_{\phi(K)T}}\right|$$

tends in probability towards zero†. This is because the numbers $n'(r)$ and N' can be considered to be constant *in the mean*, even though, in general, they do fluctuate when

† Note that, in principle, both of the terms in the absolute difference are now variable.

K increases from some given value to the next value $K + 1$. Indeed, by construction, the general physical conditions expressed by the experiment $\Pi \equiv [G.\,Vr]$ remain *invariant with respect to any such increase in the value of* K, so, as K increases, the numerator of the term

$$\frac{Kn_{\phi(K)}(r) + n'(r)}{Kn_{\phi(K)T} + N'}$$

tends towards $Kn_{\phi(K)}(r)$ and the denominator tends towards $Kn_{\phi(K)T}$, so the ratio tends towards

$$\frac{n_{\phi(K)}(r)}{n_{\phi(K)T}}.$$

This means that the absolute difference appearing in the argument of $\boldsymbol{\mathcal{P}}$ in (13) tends to 0, so the inequality with respect to ϵ is almost certainly fulfilled when K becomes sufficiently large. This also entails the almost certain fulfilment of the inequality involving δ.

Note that this does *not* mean that (13) is just a variant of the original form (8) since (13) and (8) are quite essentially different from both *logical and semantic points of view* because:

(a) The expression (13) is logically *non-recursive* in the sense that the definition (7) of a factual probability

$$p_F(r, K) = [n_{\phi(K)}(r)/n_{\phi(K)T}]$$

makes *no* changes to the numerical values of the relative frequencies $n(r)/N$ in (8), or to the way they evolve†. A 'semantic' definition like (7) does not affect the numerical values of the relative frequencies $n(r)/N$ of the events involved. Moreover, a definition like (7) can be constructed for *any* probability, and thus, in particular, for the meta-probability $\boldsymbol{\mathcal{P}}$ in (8) (for which the 'events' consist of the numerical values progressively acquired by the absolute difference in (13) for each choice of a set of values $(r, (\epsilon, \delta), K)$). This then means that:

> The expression (13) no longer *defines* one probability $p(e_r)$ through another probability $\boldsymbol{\mathcal{P}}$. Indeed, the probabilities $p(e_r)$ now have their *own* factual numerical redefinition in the form of (11), *viz.*
>
> $$p(e_r) \approx p_F(r, K),$$
>
> and this is *independent* of the definition of $\boldsymbol{\mathcal{P}}$, and can be constructed independently. This removes the circularity introduced by (8).

(b) Unlike (8), the form of (13) shows explicitly that any non-zero value of the absolute difference involved arises exclusively from the *semantically non-significant excess*

† We have already stressed, and will repeat here, that the *only* effect of the translation of the relative frequencies $n(r)/N$ into the form

$$n(r)/N = [(Kn_{\phi(K)}(r) + n'(r))/(Kn_{\phi(K)T} + N')]$$

is to identify the *semantic locations* on the grid $\gamma 1(Vc)$ of the outcomes Dr^c of N already completed successive repetitions of the experiment Π: there are no changes to the pure counting of the number N of these repetitions or to their outcomes Dr^c.

values (denoted by $n'(r)$ and N', and introduced by a relative frequency $n(r)/N$) with respect to the peaks of exact definability of the *factual* probability $p_F(r, K)$ through the algorithm of semantic integration.

So:

> *It is only the 'quanta of significance' $\phi(K)$ that play a decisive role in an effective mathematical definition of the semantic concept of factual probability.*

(c) A *factual* probability law is defined by a set of *rational* numbers. However, a probability measure is conveniently represented by real numbers because its own descriptional role is to express and formally control *any* conceivable factual probability law.

We believe that the non-effective definition of a formal probability as a real numerical (probable) limit of rational numbers can be preserved within the probabilistic *syntax* without inducing any circularity in the various theorems of large numbers, since we can now construct the factual concept of probability independently.

Finally, if the essence of the mathematical expression (13) is translated into standard terminology, it just says:

> When K is increased sufficiently, the relative frequency $n(r)/N$ of the complexified realisations $Dr^c(r)$ of the event $Dr \equiv r$ tends to become identical to the relative factual probability $p_F(r, K)$ assigned to $Dr \equiv r$ in (7) through the algorithm $\mathcal{A}\left[\{p_F(r, K)\} \leftrightarrow (\Pi, U)\right]$.

And this shows that:

> The central idea in the weak law of large numbers, *viz.* that the relative frequency $n(r)/N$ of outcomes of an event can be regarded as generating a numerical measure of the factual value of the probability of that event, despite being non-effective, is, nevertheless, fully *compatible* with the algorithm $\mathcal{A}\left[\{p_F(r, K)\} \leftrightarrow (\Pi, U)\right]$ for the identification of the factual probability law to be associated with a given random phenomenon.

In short, when we remove the circularity and non-effectiveness that blight the weak law of large numbers, this law can be regarded as a *mathematical confirmation* that the postulated existence of an unknown 'integral whole' $\phi_K(\Pi, U)$ corresponding to any random phenomenon fits coherently within a global concept of probability that is explicitly worked out with respect to *all* relevant points of view: factual/conceptual/operational, methodological and syntactic. Moreover, the numerical value of the individual syntactic probability value $p(e_r)$ involved in (8) can be factually materialised safely using the algorithm $\mathcal{A}\left[\{p_F(r, K)\} \leftrightarrow (\Pi, U)\right]$ in the way we did in (11).

7. General conclusions on Kolmogorov's aporia

We can now conclude that in the definition of the *factual* concept of probability associated with a given random phenomenon, we can still use the primitive, naive definition of the concept of probability, in other words, 'the number of favourable cases divided by the total number of all possible cases', but *these two numbers must be counted inside some* *rigorously relativised MRC representation of a meta-whole associated with the random phenomenon under investigation*, where their definitions are no longer flawed by any arbitrariness arising from a mass of vague and implicit assumptions.

Given our earlier discussion, this conclusion may come as a surprise.

In the current paper, we have explicitly reconstructed the naive definition with full generality and rigour. This solves Kolmogorov's aporia and provides us with a concept of probability that is both effective and coherently worked out from all relevant points of view, *viz.* factual/conceptual/operational, methodological and syntactic.

The result can be regarded as a pragmatic, and very strong, specification of Popper's propensity interpretation of probabilities.

In the future, the newly organised concept of probability elaborated here may well produce new consequences, questions and solutions, as well as other yet to be discovered features.

The current author has already established a major consequence of this concept, namely, the definability of *effective* measures of the various sorts of relativised complexities tied to the various descriptional roles associated with the method of relativised conceptualisation: specifically, the complexity measures of any relative MRC description and any MRC view, and even the complexity measure of any MRC object-to-be-qualified *relative to a given set of MRC-views* (Mugur Schächter 2006). Since any communicable knowledge is a description, and any description involves a generator of the entity-to-be-described and a view, this list is exhaustive. Moreover, it does not ignore the semantic features, but, on the contrary, incorporates them quite essentially. We will not pursue this any further here, but in the next section we will give a very brief outline of a surprising problem concerning quantum mechanics' famous 'essential probabilities'.

8. Quantum-mechanical 'probabilities'

It is well known that so-called 'quantum-mechanical probabilities' go beyond both classical logic and the classical concept of probability. This was clearly established in Mackey (1963), and the probability tree of the set of random phenomena corresponding to the generation of a single microstate offers an explicit and detailed conceptual/operational understanding of Mackey's mathematical discussion. Moreover, the fact that the algorithm $\mathcal{A}\left[\{p_F(r,K) \leftrightarrow (\Pi, U)\right]$ cannot be applied to quantum-mechanical random phenomena produces a fundamental divide between classical probabilities and the conceptualisation achieved within fundamental quantum mechanics.

Indeed, the results of measurements performed on a microstate almost always consist exclusively of observable marks (point-like impacts on a sensitive medium, sounds, and so on) that do not produce any *qualia* in the observer's mind that carry any *significance* tied to the microstate or the measured mechanical quantity, or, *a fortiori*, to the values of this quantity derived from translations of the observable marks carried out using conceptual/mathematical formulations based on the existing classical conceptualisation. This all means that *the aspect-views required for complexifications are absent.*

Similarly, when we look carefully at the essence of the quantum theory of measurement, it appears to offer no resources for coding the registered marks in terms of the eigenvalues of measured mechanical quantity other than the *space–time location* of the observed marks, together with the conceptual/mathematical formulations mentioned above – see Mugur Schächter (2011), Sections 3.3.2–3.3.5 and, especially, Mugur Schächter (2013).

Furthermore, fundamental quantum mechanics contains no explicit model of a microstate, and because of this profoundly non-classical situation:

> *Within quantum mechanics*, the space–time locations of the observable marks produced by measurements (being isolated, naked and with no significance that can be *directly* connected to the microstate) cannot in any way be expanded into more complex observable events.

Therefore, it is impossible to construct a *reference* structure that will allow us to give a factual specification of a probability law that can be associated with the random phenomena tied to a given microstate.

On the other hand, it is well known that according to the quantum-mechanical postulates, the quantum-mechanical probability law corresponding to any given case can be calculated formally from the state vector of the microstate involved. Now, quantum mechanics' first successes were for simple bound microstates and for free microstates that do not encounter any non-Hamiltonian 'obstacles', and in these cases it was possible to write down the state vector. However, in general, it is impossible to identify the state vector for a particular microstate under investigation, although, despite this, the investigation can generally be carried out in a purely empirical way. Indeed, in order to identify the state-vector of a microstate†:

(1) The experimental situation must be a 'Hamiltonian situation'.
(2) The relevant Hamiltonian operator must be defined.
(3) The corresponding time-independent Schrödinger equation must be solved to give an infinite set of formally possible solutions.
(4) Finally, in order to extract from the set of possible solutions the one corresponding to the case being considered, we have to specify the relevant boundary conditions mathematically.

In practice, this procedure is impossible for *any* microstate that *can* actually be generated and then studied through measurements.

The possibility of producing a microstate and performing measurements on it *does not entail* the possibility of determining a corresponding state vector.

Furthermore, it does not allow us to apply the algorithm constructed in the current paper for identifying a factual probability law.

Hence, it may even be the case that the famous 'essential' quantum-mechanical concept of probability, which it is said always generates such precise and precisely verified

† For example, consider Schrödinger's calculations in determining the state vector of the single electron in an atom of hydrogen, which required a number of successive stages, mathematical tricks, ingenious pieces of reasoning and appropriate approximations. And then imagine a slightly more complex case. As another example, consider a free state of a microsystem that encounters some irregular obstacle, and then try to define the Hamiltonian for it (*assuming* it is a Hamiltonian case; but what do you do if it isn't?).

predictions, cannot even be constructed in an effective way *inside quantum mechanics itself.*

However, the view very roughly outlined above is a personal one, and has only been raised here to submit it for discussion in the physics community at large.

The opposite view, that the quantum-mechanical factual probability law corresponding to any definite case can be *derived* inside the formalism of quantum mechanics, has been upheld by various authors – see Destouches-Février (1946), Ballentine (1973), Deutsch (1999) and Anandan (2002), and no doubt others as well.

However, the issue is far from settled, and I doubt whether those who have commented on it were aware of Kolmogorov's aporia, and of the fuzziness it reveals in the *classical* concept of probability, which was itself the starting point for the quantum-mechanical concept. Such commentaries all seem to assume that we know how to construct the particular factual probability law involved in *any* probabilistic physical theory, but that inside the quantum-mechanical formalism, it might *also* be possible to formulate a general procedure for *deriving* the factual probability laws from the *mathematical representation* of a microstate. I believe that this manifests an astonishing lack of awareness of the fact that formal representational elements *cannot* be used as a pool of factual data that have not first been independently obtained empirically and *then* translated into representational formal elements using a convenient coding – this confusion is the result of a misinterpretation of Gleason's theorem (Gleason 1957).

In short, I believe that the hope that the quantum-mechanical probability laws can be derived is completely opposed to the conceptual situation revealed in the current paper. Indeed, given the results presented here, it may finally turn out that, purely on the basis of uncontestable arguments of principle, the concept of probability is quite essentially a *classical* concept that can only be effectively constructed inside the domain of classical thinking where we work with *models* and a postulate of causal determinism (Mugur Schächter 2006, pages 118–127).

If this is the case, the quantum-mechanical descriptions would have to be recognised as being simply *'primordially' statistical transferred descriptions* (see point (6b) in Section 3.5.2). Of course, we could just *assert* the existence of an *ideal* mathematical limit, and this assertion could, more or less, be checked as an observed *tendency* to converge, but this would mean we were again subject to Kolmogorov's aporia.

Nevertheless, nothing prevents us from thinking that the postulate of existence of a global form $\phi_K(\Pi, U)$ may also remain valid for the random phenomena connected with a microstate, but that this global form can only be *represented* on some higher level of conceptualisation of the microstates, where, unlike the case for fundamental quantum mechanics, *a generic finite model* of microstates is explicitly introduced (such as an improved version of the de Broglie–Böhm model that is always contained within a definite finite space domain). This would amount to believing that it will *only* be after we have developed some entirely satisfactory consensual interpretation of the primordial transferred descriptions produced by fundamental quantum mechanics that we could justify applying the concept of probability to the random phenomena related to the quantum-mechanical descriptions of microstates. Such an interpretation would then provide us with a sort of conceptual iceberg whose small probabilistic tip would only appear within our new interpretation of fundamental quantum mechanics, while the much larger statistical part lying under the sea would be expressed through the purely transferred descriptions of microstates offered by fundamental quantum mechanics.

9. General conclusions

The concepts of statistics and probability are currently in need of thorough reconstruction if they are to cover, exhaustively and coherently, both the classical domain and modern microphysics. The current paper represents a basic effort in this direction, but at the same time has been written as an illustration of the way in which the general method of relativised conceptualisation can be used to clarify concepts and problems, and then construct solutions to the problems. Our overarching conclusion is that a general and explicitly constructed *method* of *relativised* conceptualisation seems to be an unavoidable requirement in the current phase of scientific thought.

Appendix A.

Consider a random phenomenon

$$(\Pi, U), U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

within an epistemic referential (G, Vr) where the descriptions Dr are *maximally* specified with respect to all the aspect-views in the view Vr. Also consider the probability space

$$[U, \tau_T(U), p(\tau_T)]$$

based on the universe

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

and where $\tau_T(U)$ is the *total* algebra on U, which, in particular, contains all the elementary events in U. Now consider the unknown factual numerical probability distribution $p(U)$ on the universe U of elementary events. By construction, this distribution is *contained* in the probability law $p(\tau_T)$ (which is assumed to exist) on the total algebra τ_T.

Proposition A.1. *Nothing* can be asserted about $p(U)$ *within* the scope of possible knowledge demarcated by the epistemic referential (G, Vr): not the *uniform* distribution

$$\begin{aligned} p(U) &\equiv \{p(Dr)\} \\ &\equiv \{1/\rho\}, \ \rho \text{ times} \end{aligned}$$

(where ρ is the cardinal of the index set $r = 1, 2, \ldots \rho$) or any other form.

Proof. By hypothesis, each description Dr in the universe

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

is maximally qualified with respect to all the aspect-views in the view V working within the epistemic referential (G, Vr). So, once the relative descriptions Dr have all been fully worked out inside (G, Vr), all the qualifying powers of (G, Vr) will have been *exhausted.*

Now suppose we want to qualify the resulting set of descriptions

$$\{Dr\}, \qquad r = 1, 2, \ldots \rho,$$

using two *new* aspect-views for numerical qualifications leading to the specification of the factual probability law on U:

(1) in terms of the set of the numerical values

$$\{n(Dr)/N\}, \qquad r = 1, 2, \ldots \rho,$$

observed for the meta-aspect $n(Dr)/N$ of relative frequency of an outcome of $Dr \in U$ in a sequence of N repetitions of the experiment $\Pi \equiv [G.\,Vr]$;

(2) in terms of some meta-meta way of estimating the set of numerical values

$$\{p(r, (N, N')\}, \qquad r = 1, 2, \ldots \rho,$$

towards which, in N' repetitions of a sequence of N repetitions of the experiment $[G.\,Vr]$, the set of relative frequencies $\{n(Dr)/N\}$ shows a tendency to convergence, which is *assumed to 'exist'*.

In order to determine such meta-estimations and meta-meta-estimations, we have to introduce, respectively:

(1) An epistemic referential (G', V') where:

— $G' \neq G$ introduces as [entity-to-be-described] $œ_{G'}$ (which is different from $œ_G$) the whole universe $U \equiv \{Dr\}$ of the descriptions Dr previously achieved *inside* the initial epistemic referential (G, Vr), considered globally;

— $V' \neq Vr$ consists of a *single* aspect-view, which is a 'statistical' aspect-view of relative frequency that qualifies, using numerical values, the elements of the set $\{n(Dr)/N\}$ (generically written $n(Dr)/N'$), but *leaves unchanged the semantic content of the descriptions Dr in U that were determined inside* (G, Vr).

(2) An epistemic referential (G'', V'') where:

— G'' introduces as [entity-to-be-described] $œ_{G''}$ (which is different from both $œ_G$ and $œ_{G'}$) the set of all the numerical qualifications of the relative frequencies in the set $\{n(Dr)/N\}$, considered globally;

— the view V'' (which is different from V' and Vr) qualifying each element of this new set through estimates involving the limiting number towards which it is assumed to converge, while *leaving unchanged* the semantic content of the descriptions previously achieved inside (G, Vr) and (G', V').

Now, according to the principle of separation PS (see point (9) in Section 3.5.2), each of the three successive descriptions specified above has to be carried out exclusively within its own epistemic referential and in a way that is strictly *separated* from the descriptional processes performed within other epistemic referentials. Hence, no indication of any sort can be found *inside* the initial epistemic referential (G, Vr) concerning the form of a *factually* true probability law

$$\{p(Dr)\}, \qquad r = 1, 2, \ldots \rho,$$

on

$$U \equiv \{Dr\}, \qquad r = 1, 2, \ldots \rho. \qquad \square$$

Acknowledgements

I would like to express my considerable gratitude to Professor Giuseppe Longo, the Editor-in-Chief of *Mathematical Structures in Computer Science*, for offering me the opportunity to organise this special issue of the journal, through which I have been able to put forward, and submit to debate by experts, some ideas that are not often perceived, but might be fertile, and to which I am attached.

References

Anandan, J. (2002) Causality, Symmetries and Quantum Mechanics. `http://arxiv.org/abs/quant-ph/0112020`.

Bailly, F. and Longo, G. (2007) Randomness and determination in the interplay between the continuum and the discrete. *Mathematical Structures in Computer Science* **17** 289–307.

Ballentine, L. E. (1973) Can the statistical postulate of quantum theory be derived? *Foundations of Physics* **3** 229–240.

Born, M. (1935) *Atomic Physics*, third edition, Blackie and Son.

Destouches-Février, P. (1946) Signification profonde du principe de décomposition spectrale. *Comptes Rendus de l'Académie des Sciences* **222** 866–68.

Deutsch, D. (1999) Quantum Theory of Probability and Decisions. *Proceedings of the Royal Society* **A455** 3129–3197.

Fessler, J-M. (2009) Private communication.

Gleason, A. M. (1957) Measures on the Closed Subspaces of a Hilbert Space. *Journal of Mathematics and Mechanics* **6** 885–93.

Khinchin, A. I. (1957) *Mathematical Foundations of Information Theory*, Dover Publications.

Kolmogorov, A. N. (1933) *Grundbegriffe der Wahrscheinlichkeitrechnung*, Ergebnisse der Mathematik. (English translation: Kolmogorov, A. N. (1950) *Foundations of the Theory of Probabilities*, Chelsea Publishing Company.)

Kolmogorov, A. N. (1963) On tables of random numbers. *Sankhyā Series A* 176–183. (Reprinted in Shiryayev, A. N. (ed.) *Selected works of A. N. Kolmogorov Volume III: Information Theory and the Theory of Algorithms*, Kluwer.)

Kolmogorov, A. N. (1983) Combinatorial foundations of information theory and the calculus of probabilities. *Russia Mathematical Surveys* **38** 29–40.

Longo, G. (2002) Laplace, Turing et la géométrie impossible du 'jeu de l'imitation' : aléas, déterminisme et programmes dans le test de Turing. *Intellectica* **2** (35). (English translation: Longo, G. (2007) Laplace, Turing and the 'imitation game' impossible geometry: randomness, determinism and programs in Turing's test. In: Epstein, R., Roberts, G. and Beber, G. (eds.) *The Turing Test Sourcebook*, Springer-Verlag.)

Mackey, G. (1963) *Mathematical Foundations of Quantum Mechanics*, Benjamin.

Mugur Schächter, M. (1964) *Étude du caractère complet de la mécanique quantique*, Gauthiers Villars.

Mugur Schächter, M. (1979) Study of Wigner's Theorem on Joint Probabilities. *Foundations of Physics* **9** (5-6) 389–404.

Mugur Schächter, M. (1984) Esquisse d'une représentation générale et formalisée des descriptions et le statut descriptionnel de la mécanique quantique. *Epistemological Letters Lausanne* **36** 1–67.

Mugur Schächter, M. (1991) Spacetime Quantum Probabilities I. *Foundations of Physics* **21** 1387–1449.

Mugur Schächter, M. (1992a) Toward a Factually Induced Space–time Quantum Logic. *Foundations of Physics* **22** 963–994.

Mugur Schächter, M. (1992b) Quantum Probabilities, Operators of State Preparation, and the Principle of Superposition. *International Journal of Theoretical Physics* **31** 1715–1751.

Mugur Schächter, M. (1992c) Spacetime Quantum Probabilities II: Relativized Descriptions and Popperian Propensities. *Foundations of Physics* **22** 269–303.

Mugur Schächter, M. (1993) From Quantum Mechanics to Universal Structure of Conceptualization and Feedback on Quantum Mechanics. *Foundations of Physics* **23** 37–122.

Mugur Schächter, M. (1995) Une méthode de conceptualisation relativisée. *Revue Internationale de Systémique* **9**.

Mugur Schächter, M. (1997a) Les leçons de la mécanique quantique (vers une épistémologie formelle). *Le Débat* **94** (2) 169–192.

Mugur Schächter, M. (1997b) Mécanique quantique, réel et sens. In: *Physique et réalité, un débat avec Bernard d'Espagnat*, Frontières.

Mugur Schächter, M. (2002a) Objectivity and Descriptional Relativities. *Foundations of Science* **7** 73–180.

Mugur Schächter, M. (2002b) Quantum Mechanics versus Relativised Conceptualisation. In: Mugur-Schächter, M. and Van der Merwe, A. (eds.) *Quantum Mechanics, Mathematics, Cognition and Action*, Kluwer 109–307.

Mugur Schächter, M. (2002c) En marge de l'article de Giuseppe Longo sur Laplace, Turing et la géométrie impossible du 'jeu d'imitation'. *Intellectica* **2** 163–174.

Mugur Schächter, M. (2006) *Sur le tissage des connaissances*, Hermès Science Publications.

Mugur Schächter, M. (2011) *L'infra-mécanique quantique*. `arXiv:0903.4976v2[quant-ph]`.

Mugur Schächter, M. (2013) Principles of a 2nd quantum mechanics: Construction of the foundations of an intelligible Hilbert–Dirac formulation. `arXiv:1310.1728[quant-ph]`

Petitot, J. (2002) Debate with J. Petitot on mathematical physics and a formalized epistemology. In: Mugur-Schächter, M. and Van der Merwe, A. (eds.) *Quantum Mechanics, Mathematics, Cognition and Action*, Kluwer 73–102.

Popper, K (1967) Quantum mechanics without the observer. In: Bunge, M. (ed.) *Quantum Theory and Reality*, Springer 7–44.

Putnam, H. (1981) *Reason, Truth and History*, Cambridge University Press.

Quine, W. V. O. (1977) *Ontological Relativity and Other Essays*, Columbia University Press.

Segal, J. (2003) *Le zéro et le un*, Syllepse.

Shannon, E. C. (1948) The Mathematical Theory of Communication. *Bell System Technical Journal* **27** 379–423 and 623–656.

Solomonoff, R. J. (1957) An inductive inference machine. *IRE National Record* **5**.