The concept of probability in physics: an analytic version of von Mises' interpretation

Louis Vervoort,

Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST),
University of Quebec in Montreal (UQAM),
Montreal, Canada
vervoort.louis@courrier.uqam.ca
louisvervoort@hotmail.com

Abstract. In the following we will investigate whether von Mises' frequency interpretation of probability can be modified to make it philosophically acceptable. We will reject certain elements of von Mises' theory, but retain others. In the interpretation we propose we do not use von Mises' often criticized 'infinite collectives' but we retain two essential claims of his interpretation, stating that probability can only be defined for events that can be repeated in similar conditions, and that exhibit frequency stabilization. The central idea of the present article is that the mentioned 'conditions' should be well-defined and 'partitioned'. More precisely, we will divide probabilistic systems into object, initializing, and probing subsystem, and show that such partitioning allows to solve problems. Moreover we will argue that a key idea of the Copenhagen interpretation of quantum mechanics (the determinant role of the observing system) can be seen as deriving from an analytic definition of probability as frequency. Thus a secondary aim of the article is to illustrate the virtues of analytic definition of concepts, consisting of making explicit what is implicit.

1. Introduction.

The first and simplest axiomatic system for probability, published in 1933 by Andrei Kolmogorov (1933/1956), is generally believed to cover all probabilistic systems of the natural, applied and social sciences. At the same time Kolmogorov's theory – the calculus of probabilistic or random events – does not define what a 'probabilistic / random event' is and does not provide any interpretation of the notion of probability – besides of course through the mathematical axioms it fulfills. For an intuitive understanding beyond mathematics of what probability 'is', one thus needs to resort to the ideas of other fathers of probability theory, such as Laplace, Fermat, Venn, von Mises, and to the philosophers having investigated the question (for general references on the different interpretations of probability see e.g. Fine (1973), von Plato (1994), Gillies (2000), Khrennikov (2008)).

Based on our practice as a physicist, we would think that in the broader community of (classical) physics the most popular interpretation of probability is the frequency model – especially the limiting frequency version due to Richard von Mises (1928, 1964). There seems to be a good reason for that: in the practice of physics any experimental determination or verification or measurement of a probability is *always* done by determining a relative frequency. However in philosophy and the foundations of quantum mechanics other interpretations, in particular the subjective interpretation, are increasingly popular. Suffices to consult contemporary works on the interpretation of probability in physics to realize how vivid and wide-ranging the debate is. One conclusion of a very recent review work is that for instance the objective / subjective controversy is far from being settled (see the excellent Beisbart and Hartmann 2011). At any rate, as both a philosopher and a physicist we cannot escape from noticing the relative underrepresentation of the frequency interpretation à la von Mises in the contemporary debate on probability, even in philosophy of physics. This may be due to the fact that more basic frequency interpretations have extensively been analyzed and found wanting (cf. e.g. Fine (1973), von Plato (1994), Gillies (2000) and references therein). Moreover, there are well-known *mathematical* problems with von Mises' work, notably related to the notion of 'collective'. It seems to us that all this may have unduly discredited the valuable parts of von Mises' work. In any case we believe that a revisit of von Mises' interpretation, taking these criticisms into account, is a worthwhile effort – if only to balance the debate. Specifically, we will argue that von Mises' work, if amended, can become an efficient tool for problem solving.

Thus we will not focus here on the interpretation of probability in general, i.e. on the concept of probability as it is used in general philosophy, social science, everyday life, etc. Rather, we will focus on the notion as used in physics solely. To that end we will propose an analytic definition of probability that can be seen as a modification of von Mises definition[1] (1928, 1964). We try to make explicit some notions which we believe are implicit in von Mises' theory. Moreover we claim this can be done without using much mathematical formalization, but by applying a simple but precise conceptual analysis. Also van Fraassen has provided a well-known analysis of probability as frequency (1980, pp. 190 – 194). Based

[1] In essence, von Mises interprets the probability $P(R_j)$ as $\lim_{n\to\infty} n(R_j)/n$, where $n(R_j)$ is the number of events or trials that have result R_j in a series of n trials ($j \in \{1,\ldots,N\}$).

on a study of how probability is used in physics, van Fraassen presents a logical analysis of how to link in a precise way physical experiments to probability functions. The author gives as a summary of his elaborate model following definition (1980, p. 194):

> "The probability of event A equals the relative frequency with which it would occur, were a suitably designed experiment performed often enough under suitable conditions."

Our definition will be seen to be in agreement with the above; but our starting point will be different, namely to investigate what precisely a *probabilistic physical system* is. In the following we will give in particular a further analysis of what the 'suitable conditions' in van Fraassen's definition are, and what 'often enough' might mean. At least one other very recent work concludes that von Mises' theory is highly relevant for interpretational questions (Khrennikov 2008); we will make a link also to this reference. Finally, we also found Kolmogorov's and in particular his pupil Gnedenko's standard references on the calculus highly recommendable sources[2] for studying the interpretation of probability (cf. Kolmogorov 1933/1956, Gnedenko 1967).

Anyone who tried to construct a precise definition of probability soon realizes that the topic is indeed surprisingly subtle. In a standard physics curriculum, students are offered at least two interpretations of probability, the classic interpretation of Laplace for chance games, and the frequency interpretation for natural probabilistic phenomena (cf. almost any textbook on probability calculus, e.g. Thijms 2001, Gnedenko 1967). Modern books may also provide the subjective interpretation, without linking it to the other interpretations (e.g. Thijms 2001). What we will term 'chance games' are random experiments using man-made probabilistic systems, such as dice, urns containing balls, cards, roulettes, etc. 'Natural' probabilistic phenomena are ubiquitous in nature and can be found in diffusion, population dynamics, fluid dynamics, quantum mechanics and in any branch of natural science. Scientists typically suppose that these phenomena spontaneously happen in nature according to the laws of probability theory, also when nobody looks. A first challenge we are interested in is to devise a *unified* definition that applies to both chance games and natural probabilistic phenomena – both concern physical systems, after all. Such a definition seems highly desirable, not only for foundational but also for practical reasons. First, it suffices to have a look at the texts of von

[2] We found in particular Gnedenko's book on probability calculus helpful. Boris Gnedenko did not only substantially contribute to probability calculus (especially on elaborations of the Central Limit Theorem and statistics); his work manifestly reflects his interest in the foundational issues. Also Kolmogorov insisted on the need of a philosophical definition of probability (cf. his (1933/1956) p. 9).

Mises or almost any reference text on the calculus to realize that without a precise idea of what probability *is*, beyond the formal aspect, many somewhat subtle problem is likely to be treated in a wrong manner. Von Mises provides many tens of such flawed calculations by even experts. No surprise, it has been said that “In no other branch of mathematics is it so easy to make mistakes as in probability theory” (Tijms 2004, p. 4). Needless to say, for foundational reasons it is even more important to have a clear idea of the implicit notions that the concept of probability contains. We will highlight the latter claim by exposing a (perhaps surprisingly) close link between the interpretation of probability and of quantum mechanics.

In some detail, the frequency interpretation we propose here differs from von Mises’ theory (1928, 1964) in at least two respects (some may find our model actually quite different). To start with, our model is simpler. It is essential to note that a *real-world* (or *physical*) theory (T) for probability is the union of a calculus (C) and an interpretation (I): T = C ⋃ I (Bunge (2006)). The theory ‘I’ is not part of mathematics but stipulates how to apply the theory to the real world. ‘I’ may contain e.g. the philosophical hypothesis that the theory represents things or events ‘out there’; and more importantly, it contains rules stipulating on which things / events the mathematics to apply, and how. Von Mises proposed *both* a calculus C and an interpretation I. His calculus is based on the concept of collective (an in principle infinite series of experimental results) – a calculus that strikes however by its complexity, and that probably no-one would consider using nowadays on a regular basis (see e.g. von Mises’ treatment of de Méré’s problem, von Mises (1928) p. 58ff). Therefore we will not use von Mises’ calculus (it leads of course to the same results as Kolmogorov’s). Even within the interpretational part I, we will not make use of the cumbersome concept of collective; in particular we believe that it is not necessary to resort to the notion of ‘invariance under place selection’ to characterize randomness (see next Sections and Appendix 1; Gillies (2000) p. 112 actually proves this fact). So our attitude is pragmatic: few people challenge the idea that the mathematics of probability theory is fully contained in Kolmogorov’s calculus – so our calculus (C) is Kolmogorov’s. What *is* however controversial is the following question: *what is the exact subject matter of probability theory – to what exactly to apply it?* We will push the answer further than von Mises’ classic answer: according to him probability theory treats “mass phenomena and repetitive events” (von Mises (1928 / 1981) p. v), characterized by a set of attributes that form the attribute space. (Anticipating, we will arrive at the conclusion that probability theory applies to a special type of random events, which we will term ‘p-random'.) The definition we will propose in Section 3 captures, we believe, the essence of von Mises' interpretational model (I), but in a simpler form which is not based on the often criticized notion of collective.

Besides a simplification, our analytic model is intended as a clarification of the frequency interpretation, as applied to physics. We believe it helps to avoid paradoxes to which probability theory seems so sensitive. Our main claim is that one substantially gains in *partitioning* probabilistic systems into subsystems, namely test object, initiating, and probing subsystem (or 'environment' if one prefers). We will argue that this partitioning allows to solve classic problems (such as Bertrand's paradox), and bring in focus a link between classical and quantum systems. The latter link has been emphasized by other philosophers from a different angle (Szabó 1995, 2000, 2001, Rédei 2010).

As already mentioned, recently another author has come to the conclusion that von Mises' work is more than worth a reappraisal (cf. Khrennikov 2008). The author offers in his textbook a detailed and convincing analysis based on the mathematics of collectives. Khrennikov concludes in particular that well-known criticisms claiming that von Mises' theory (C) lacks mathematical rigor are not really cogent, as we argue for completeness in Appendix 1. But since we will not focus on the formalism but on a conceptual study (I) we at least are immune against the criticisms relative to C.

2. Introductory ideas: frequency stabilization and partitioned conditions. Examples.

The frequency interpretation of probability can be seen as in agreement with a general credo of philosophy of science (dating from the Vienna Circle or earlier) stating that to look for the meaning of concepts, it is a good starting point to ask how one measures / verifies instances of the concept – if this is possible. As already said, in physics any experimental probability is determined, measured, as a relative frequency, also probabilities in quantum mechanics. Of course theories may predict probabilities as numbers (such as the square of a modulus of a wave function, a transition amplitude, and values of any probabilistic property) that are not obviously ratios or frequencies, but to verify these numbers one needs to determine relative frequencies (we will illustrate this procedure in a few case studies). This is the starting point of von Mises' frequency model. A point to mention from the start is that von Mises attributes probability only to series of experimental outcomes – series of outcomes of random experiments.

In order to get a feel for the problems to be investigated, let us have a look at a typical probabilistic system, namely a die. A central question for us is: when or why is a die 'probabilistic', or 'random' (or rather the throwing events or outcomes) ? Simply stating in non-anthropocentric terms what a die throw is, seems already to bring to the fore a few key notions. In physical terms a 'die throwing event' or 'die throw' consists of the 3-dimensional movement of a (regular) die, that is 1) characterized by the time evolution of its center of mass and its three Euler angles, 2) *caused* or *initiated* by an 'initiating / initializing system' (e.g. a randomizing and throwing hand, or automat), and 3) *probed* by an 'observing / probing system' (e.g. a table) that allows to observe (in general 'measure') an outcome or result R (one up, two up, etc.). In the case of die throwing, it is easy to realize that if we want to use the die as it should, i.e. *if we want to be able to observe or measure the usual probability* for the different results of the throws, we have to repeat the experimental series (i.e. the throws) not only by using the same regular die, or similar regular dies, but also by applying the usual 'boundary conditions'. Here we obviously *not* refer to the detailed initial conditions of each individual throw, which are unknown, but to the conditions of our random experiment as they can be communicated to someone else ("throw roughly in this manner, probe on a table", etc.). These boundary conditions, then, are related to our throwing, the table, and the environment in general. *Irregular conditions in any one of these three elements may alter the probability distribution*. We can for instance not substantially alter our hand movement, e.g. by putting the die systematically ace up on the table: this could change the probability distribution for the six outcomes from the usual (1/6, 1/6, 1/6, 1/6, 1/6, 1/6) to (1, 0, 0, 0, 0, 0). Nor can we put glue on the table, position it close to our hand, and gently overturn the die on the table while always starting ace up – one of the outcomes could occur much more often than in a ratio of 1/6. Nor can we do some experiments in honey instead of air: again one can imagine situations in which the probabilities of the outcomes are altered.

As will be seen further, it will make sense to isolate the mentioned elements, and to consider a random event as involving a *random system* containing three subsystems, namely 1) the (random) *test object* itself (the die), 2) the *initiating system* (the throwing hand), and 3) the *probing* or *observing system* (the table, and let's include the human eye). We will call such a composed random system a 'p-system' or probabilistic system. Just as one can associate subsystems to the random event, one can associate (composed) *conditions* to it, so conditions under which the random experiment occurs, in particular *initiating* and *probing conditions*.

Here is another example, from darts. According to the frequency interpretation, if we want to determine the probability that a given robot hits a given section of a dartboard (suppose there are 10 sections S_1 to S_{10}), we let it throw say 500 times, and count relative frequencies: n_1 hits in S_1, so P_1 (the probability the robot hits S_1 in the given conditions) is approximately $n_1/500$, etc. Again, if we want to define probabilities for the experiment, the robot, the darts and the board need to behave well and 'operate' in well-defined 'constant' conditions. The probabilities are defined *relative to these conditions*, referring to the test object (the darts), the initiating system (the throwing robot), and the probing or observing system (the board, and we can again include some optical detection system). For instance, if during our experiment the robot's output force diminishes due to a failing component in its circuit, or if some of the darts break during flight, P_1 will be much lower than $n_1/500$, etc[3]. All this may look almost trivial; and yet it will prove highly helpful to explicitly include these observations in the definition of probability we will propose in the next Section.

At this point an essential remark is in place. Instead of 'initiating system' and 'probing system', it can be more appropriate to speak of 'environment', namely in the case of spontaneous or 'natural' probabilistic events (versus 'artificial' ones, as outcomes of chance games, which are created by human intervention). Many random events occur spontaneously, without any known cause. A spontaneously disintegrating nucleus has probabilistic properties (e.g. its half-life, i.e. the time after which it disintegrates with a probability = 1/2). Such properties are, on usual interpretations, not 'initiated', as in our die throwing, by a physical system or a person. Neither are they 'probed', except when subject to lab experiments. But the nuclei disintegrate spontaneously, according to a given well-defined probabilistic pattern, *only in a well-defined environment*. Changing the environment, e.g. by irradiating the nuclei or by strongly heating them, may very well change the probability distribution in question. In other words, if we want to determine (measure) the half-life of the nuclei, i.e. the probability of disintegration, we have to put them in well-defined ('initializing') conditions of temperature, pressure, etc. and measure their properties in well-defined ('probing') conditions, that scrupulously imitate its natural environment – if we want to know the properties 'in nature'. So also here the initial and final conditions, or environment, re-appear. (As a side-remark,

[3] Notice there is some arbitrariness in where we draw the line between the different subsystems. But that doesn't play any role in the following.

note that the half-life of ‘a’ nucleus can only be measured on a huge ensemble of ‘identical’ nuclei.)

By the above partitioning in subsystems we have thus rendered the concept of ‘conditions’ explicit. To the best of our knowledge, a somewhat detailed analysis of this notion has not yet been undertaken. It appears succinctly in Gnedenko’s reference work on probability calculus (1967). Gnedenko states (p. 21):

> “On the basis of observation and experiment science arrives at the formulation of the natural laws that govern the phenomena it studies. The simplest and most widely used scheme of such laws is the following: *Whenever a certain set of conditions C is realized, the event A occurs*.”

And a little further (1967, p. 21):

> “An event that may or may not occur when the set of conditions C is realized, is called *random*.”

Thus on a natural interpretation, and as is already implicit in Kolmogorov (1933/1956, p. 3) and explicitly stated by Gnedenko (1967, p. 67), even unconditional probabilities (P(R)) can be regarded as conditional *on the circumstances of realization*. Therefore the natural generalization in the model we will propose below is to replace, if helpful, P(R) by P(R|C) where C contains all relevant parameters that describe the initiating, probing, and ‘environmental’ experimental conditions.

The importance of what we termed ‘probing conditions’ has been emphasized by other philosophers in the specific context of quantum probabilities. According to Szabó’s ‘Kolmogorovian Censorship Hypothesis’ (Szabó 1995, Szabó 2001) there are no genuine non-classical probabilities, in the sense that ‘quantum probabilities’ are always classical conditional probabilities of the measurement outcomes, where the conditioning events are the events of choosing / performing the measurement setup to measure a certain observable. This claim can well be seen to be fully agreeing with the model we will elaborate here. The relevance of this statement for the foundations of quantum mechanics will be argued for below (and see Szabó 1995, 2000, 2001, Rédei 2010).

Now that we have a somewhat better idea about the above mentioned conditions, it seems not difficult to deduce the *physical conditions for ‘randomness’* of a well-known physical system as a thrown die (we will see further that the condition proposed above by Gnedenko is not sufficient). Suppose we throw a die 20, 30, in general n times (n >> 6), and that we note how often the 6 results R_j ($j = 1,\ldots, 6$; $R_j = 1,\ldots, 6$) occur in these n throws. We can then determine the *relative frequencies of the results R_j*, namely the 6 ratios n_j/n where n_j is the number of throws that have result R_j and n the total number of throws. Now, according to the frequency interpretation (and to any physicist's intuition and practice), in order that the die throws can be termed 'probabilistic' or 'random' or 'stochastic', it is a necessary condition that the relative frequencies of the R_j (the ratios n_j/n) *converge towards a constant number* when n grows. For a regular die thrown in regular conditions it is an *empirical matter of fact* that the six ratios n_j/n converge towards 1/6 when one increases the number of trials n; in other words the n_j/n approach this number with increasing numerical precision. If such '*frequency stabilization*' *would not occur, one cannot speak of probability*[4]. If for instance the die would sublimate during the experiment in an asymmetric manner (thus gradually losing weight on one side), or if we accompany the die's movement in a structured manner that does not lead to frequency stabilization, we can in general not attribute a probability to the outcomes. If we repeat the throws, the above ratios may not converge but erratically jump from one value to another.

In sum, if there is no frequency stabilization, the die throwing is not probabilistic. On the other hand, if frequency stabilization occurs, the probability of the result R_j is given by $P(R_j) = n_j/n$ = (number of events that have result R_j) / n, where n must be large enough (we will come back to the latter condition in a moment). This is, or this is in full agreement with, the frequency interpretation of von Mises and others (von Plato (1994), Gillies (2000), Khrennikov (2008)).

At least in physics, frequency stabilization appears to be a property that any system that deserves the predicate 'probabilistic' necessarily must have, as we will illustrate below by a realistic physical case study. Von Mises writes e.g. (1928, p. 12):

> "It is essential for the theory of probability that experience has shown that in the game of dice, as in all the other mass phenomena which we have mentioned, the relative frequencies of certain attributes become more and more stable as the number of observations is increased."

Let us then illustrate the ideas of (partitioned) conditions and frequency stabilization by a realistic example from the physics lab, namely recent experimental findings in fluid mechanics and nonlinear physics, which have sparked much interest in the physics

[4] The term 'frequency stabilization' appears in von Mises' work; in our opinion it best captures the property in question.

community in the last years. A group of experimentalists have discovered that oil droplets can be made to hover over an oil film (Couder et al. 2005, 2006 and Eddi, Couder et al. 2011). To that end, the oil film is made to vibrate by an external motor. If small oil droplets are deposited on such a film they sometimes begin to horizontally walk over the film, for indefinite time; actually they bounce so fast on the film that they seem to hover[5]. However this stable walking regime only occurs in well-defined experimental conditions, i.e. for precise values of the physical parameters of the system, essentially the frequency and amplitude of the external vibration, the size of the droplet, the geometry of the oil film and bath, and the viscosities of film and droplet. If these parameters are fine-tuned and lie within precise ranges of values, well-documented by the researchers, the droplets walk horizontally; outside these value ranges the movement becomes erratic and/or the droplet is captured by the film. Now, in the walking regime it is possible to experimentally determine certain probabilities, notably the probabilities P_R that the walking droplet or 'walker' passes through a certain space region R. To that end, the physicists have counted relative frequencies (cf. e.g. Couder et al. 2006, Fig. 2-3) much as we illustrated above in the case of a dartboard: they determine n_R/n, where n_R is the number of trials in which the droplet passes through R and n the total number of trials. (Obviously they create droplets in identical conditions and measure them in identical conditions.) However before determining numerical values of probabilities as P_R, the physicists have spent weeks, months and possibly years to identify the exact conditions for which a stable probabilistic pattern (i.e. probabilities P_R) occurs. To *prove* that the system is probabilistic (and not erratic) there is only way: namely to prove that frequency stabilization occurs, i.e. to prove that the ratios n_R/n converge to a constant number with ever better precision when n is increased. When experimentalists state numerical values of probabilities in publications, they have assured that such a frequency stabilization has occurred in their experiments (note that ratios / frequencies could also be determined without first verifying that a stable regime is identified, but it would be scientifically unacceptable to call such ratios probabilities). In physics this typically manifests itself by the fact that a probability histogram (e.g. n_R/n as a function of discrete values of R) is better and better

[5] These walking droplets exhibit moreover a behavior that strikingly imitates quantum behavior, including double-slit interference, tunneling and quantization of angular momentum. Couder et al. have quite convincingly shown that the origin of such a quantum-like behavior lies in the wave-front that accompanies the hopping droplets (the wave is created by the external vibration and by the back-reaction of the bouncing droplets on the oil film). Thus they could claim that these walking droplets are the first realization of a "particle + pilot-wave" (Couder et al 2005, 2006). Here we do not focus on these intriguing quantum-like features, but on the classical probabilistic features, which were also extensively investigated by Couder et al.

approximated by a smooth curve when n is increased (see e.g. Couder et al. (2006), Fig. 2-3 where both histogram and curve are shown); often one publishes just the curve. This curve, in other words these limiting values of n_R/n, are supposed to be 'the' probabilities, or good approximations of it. Note that van Fraassen's condensed definition, "The probability of event A equals the relative frequency with which it would occur, were a suitably designed experiment performed often enough under suitable conditions" (van Fraassen 1980, p. 194), is in agreement with this operational procedure. Below we will fill in what 'often enough' might mean.

The important point that the above example highlights is that frequency stabilization, i.e. (the observation of) a probabilistic pattern or in other words probabilities, only occurs under well-defined conditions, and by no means always. Very small variations[6] in some parameter (e.g. in the external vibration frequency, the droplet size, etc.) may lead to completely different probabilities, or to erratic behavior without stable relative frequencies, or even to absorbance of the droplet by the film. Therefore if a physicist who is not an expert in the dynamics of oil droplets on films – i.e. almost anyone – is asked to determine the probabilities P_R by experimental observation, she may well come to the conclusion, even after weeks of experimentation, that the system is *not* probabilistic, or if she is more cautious that she did not identify conditions in which the system is probabilistic.

Before generalizing these ideas, a remark is in place. As is well known, von Mises uses in his theory the concept of *limiting relative frequency*, for the limit $n \to \infty$. This limit for n tending to infinity is considered problematic by some authors. For the sake of completeness, we will analyze some of these criticisms in Appendix 1 (our own model will however not use the notion of infinity). In all fairness one should remark that it is often forgotten that von Mises has dealt in detail with many of these objections (see in particular the introductions of his 1928/1981 and 1964). Among other points von Mises recalls that probability theory is not only a mathematical but also a physical theory (it applies to the physical world). And the concept of infinity is ubiquitous in theoretical physics, even if it is in many calculations obviously an idealization – a useful one when the formulas correctly predict reality to the desired precision.

[6] The extreme dependence of physical properties on certain variables is typical for nonlinear systems, such as those investigated by Couder et al. The oil droplet bouncing on the film is governed by a nonlinear equation of movement, e.g. due to the so-called 'viscous friction' between droplet and film.

On reflection, we believe that von Mises has attracted criticisms that could have been avoided had he made following observation. It is well-known that his definition of probability as $\lim_{n\to\infty} n(R_j)/n$ cannot work *in the mathematically strict sense* (cf. e.g. Richter (1978), p. 426). Mathematically,

$$P(R_j) \;=\; \lim_{n\to\infty}\frac{n(R_j)}{n} \quad\Leftrightarrow\quad \forall\varepsilon>0, \exists\ \mathrm{N}:\ \forall n>N,\ \left|\frac{n(R_j)}{n}-P(R_j)\right|<\varepsilon\,. \tag{1}$$

However, for any probabilistic event, for a given ε and for some n in the series of experimental results the last inequality in (1) may of course not be satisfied. Definition (1) works for defining the limit of *mathematical* sequences which evolve monotonically towards a limiting value. But for the sequence of *physical, experimental* results {n(Rj)}, an overwhelming amount of experimental data has indicated that the $n(R_j)$ tend towards a fixed number *but in general in an oscillating manner* – as anyone can verify by throwing dice, etc. Yet with some charity one understands what is meant with "$\lim_{n\to\infty} n(R_j)/n$": namely the "experimentally determined value of convergence (not mathematical limit in the strict sense) of the ratio if one would continue – or if one continues – determining the ratio for increasing n". In sum, it seems we can forgive von Mises' slight misuse of a mathematically precise concept; and maybe he could have avoided criticisms if he would have defined probability in terms of say "$\lim^*_{n\to\infty} n(R_j)/n$", where lim* is the *operational, experimental* notion just defined.

At any rate, this is not our problem: we will in the following not define frequency stabilization in terms of convergence in the limit $n \to \infty$, but 'for n large enough for the desired precision'. This being said, we believe that the just defined operational notion 'lim*', or frequency stabilization, reflects a particularly interesting fact of nature. At the very basis of the frequency interpretation lies the scientific hypothesis that if certain well-defined experiments (trials) *would* be repeated an infinity of times, certain ratio's (involving outcomes of these trials) *would* converge to a limit. The modal aspect of the concept of probability has been highlighted in detail by van Fraassen (1980). And experiment has shown over and over again that long series (obviously not infinite ones) do show such frequency stabilization[7].

[7] Note that these massive empirical findings corroborate the initial hypothesis (namely of frequency stabilization as a real-world phenomenon) as strongly as any hypothesis of the natural sciences can be confirmed: scientific hypotheses are never confirmed with absolute certainty. It may be that after repeated measurements of a magnetic field strength (say 0.5 Tesla) in a given point X, the next measurements unexpectedly jump to 0.8 Tesla. This would not be enough ground to reject the hypothesis that, in the given conditions, there should be a fixed field in X. Physicists would look for causes of the jump. Similarly, if in a given experiment that is known

Let us finally note that, as formulated above, random systems include deterministic systems. Indeed, in the latter case frequency stabilization will also occur. If we move a die by placing it systematically – deterministically – ace up on the table, the six $P(R_j)$ will trivially converge to a value (namely to 1, 0, 0, 0, 0, 0). We could exclude deterministic cases from our definition of random events / systems by adding that, besides exhibiting frequency stabilization, the events / systems should be unpredictable. Thus 'genuinely' random events show frequency stabilization *and* are unpredictable (or one could replace the latter by another predicate that qualifies the 'unstructured', disordered nature of randomness).

3. Generalization, definitions.

Before proposing our definitions, which will essentially be straightforward generalizations of what we learned in the preceding Section, two philosophical problems deserve to be highlighted.

We argued with von Mises that the essential feature qualifying physical events or systems as random, or probabilistic, is frequency stabilization. At this point an obvious, but neglected, question pops up: are all physical random systems really random – in the sense of probabilistic ? In the case of systems *studied in physics*, one often declares that physical systems are either random in the probabilistic sense, or deterministic. This classic dichotomy is for instance investigated by Bell's theorem (see e.g. Vervoort [2013]). However the above question becomes more interesting in the case of 'physical systems' in a larger sense, especially if the systems in question 'contain' human beings. If frequency stabilization is indeed the essential feature of real-world random systems and events, then it is clear that *not* all so-called random systems are random in the sense of probabilistic or stochastic. *Frequency stabilization is by no means guaranteed for an arbitrary disordered event, system or parameter*. Consider a mountain torrent, and define a parameter characterizing its flow, such as the height of its waves or its rate of flow through a given surface. Such a torrent may show such erratic behavior over time that it may be impossible to define a parameter of which the frequency stabilizes. In the former Section we saw a concrete example from physics (fluid dynamics) that illustrates the daunting practical difficulty for finding conditions of probabilistic stabilization. Arguably a majority of physical systems are so complex and non-

to be p-random frequency stabilization would not occur, one would not reject frequency stabilization but look for causes to explain the strange behavior.

linear that they are extremely sensitive to variations in the parameters that describe them. As illustrated above this leads often or ‘normally’ to the impossibility of determining conditions of frequency stabilization. Sure, this is a statement of pragmatic nature – maybe one could argue that an omniscient being could know such conditions. As a next example, atmospheric temperatures in a certain city look random, they even look statistically random, but they arguably don’t show frequency stabilization over long periods, e.g. because of climate changes[8]. What is even less likely to exhibit frequency-stabilized features, is 1) any ‘random’ system that is subject to human actions, e.g. the chaos on my table, or the random-looking features of a city-plan, or the duration of my Sunday walks etc., and 2) any ‘composite’ random-looking system, like a car and the house in front of it, or a tree, a stone underneath, and a cloud above.

In sum, it seems that frequency stabilization happens often[9], but certainly not always. Then not all randomness is the randomness that is the object of probability calculus. This observation justifies the introduction of the concept of ‘p-randomness’, to be contrasted with ‘randomness’ and ‘non-p-randomness’: p-randomness, or probabilistic or structured randomness, is randomness (chance) that is characterized by frequency stabilization (at infinity, or ‘for large n’). Before one applies probability calculus to a physical event, one conjectures or has evidence that it is p-random – a feature that can only be confirmed by experiment (however never with absolute logical certainty, but with the certainty or uncertainty that characterizes empirical hypotheses). Note that the technically familiar way to characterize a parameter or variable R as p-random, is to say that it has a ‘probability density function’, a parameter often symbolized by ‘ρ’ in physics ($\rho = \rho(R)$).

Another philosophical problem that deserves attention is the following. In Section 2 we have emphasized the role of the conditions under which experiments disclosing the p-random nature of physical events need to be performed. Such experiments involve repeated tests, in which the object under study (e.g. the die) should be subjected to repeatable, ‘similar’ or ‘identical’ conditions before being probed. It is a well-known problem of the foundations

[8] This may however be seen as a limiting case, in which frequency stabilization in experiments occurs ‘for large n’, even if not ‘at infinity’. Indeed, modern meteorologists heavily use statistical talk: they assume that the conditions at a given time in a given place are sufficiently similar to those for which they have a large set of data which allow to deduce (approximate) probabilities. Since these data sets are very large, one can accept that statistics is still applicable here and leads to ‘reasonable’ precision. But not the kind of precision one can attain in more typical physical systems, for which the sampling size can be virtually infinite.

[9] A particularly beautiful mathematical argument for the ubiquity of probabilistic randomness is provided by the Central Limit Theorem.

of probability to identify what ‘similar’ / ‘identical’ exactly means. How similar is similar (we will propose an answer in C4 below) ? Now we all know by experience that we do not need a lot of precautions when shaking a true die so as to launch it in a ‘regular’ manner – leading to the normal probability distribution of a true die. In other words, it seems that in general we have a good idea what these ‘similar conditions’ are; but that is certainly not always the case – in particular in physics, and even more so in quantum mechanics. At this point the idea of partitioning we introduced in the preceding Section proves useful. Indeed, similar tests on similar systems imply, in general: similar initiating system (or environment), similar test object, and similar probing system. These *three* subsystems should ‘act’, in the repeated test sequence, in *similar (enough) ways* in order that a probability can be defined. Alterations or divergences in any of these three subsystems can lead to different probabilities (probability distributions) – the latter may not even be defined, i. e. existing. Remember the sublimating die, or the inadequate randomization, which can lead to an undefined (i.e. changing, unstable) probability distribution of the die throws. Or recall the experiments in fluid mechanics of Couder and collaborators (Couder et al. 2005, 2006 and Eddi, Couder et al. 2011). This just to put the remainder, and in particular the question what ‘similar’ means, in perspective.

We now propose a definition that can be derived from von Mises’ theory, but that avoids the concept of collective. The move we make is essentially the following. 1) Von Mises starts by defining a collective as an (in principle infinite) series of *random* results, and defines probability with respect to a given collective. But we have just argued that probability cannot be attributed to any random series, but only to a p-random series. We therefore first define the concept of frequency stabilization or p-randomness: it is logically prior to that of probability. 2) We *partition* the probabilistic system, as explained in Section 2. *We emphasize that we do so essentially for pragmatic reasons* – i.e. for addressing concrete problems as those presented in this and the next Section. So our aim is quite minimal and we refrain from strongly metaphysically tainted claims – in any case if they have no verifiable function of problem-solving[10]. The task is then to construct a definition that makes the notion of ‘experimental conditions’ explicit, *and* that applies to both chance games and natural probabilistic phenomena. If we say ‘identical’ in DEF1 and DEF2, one should read ‘identical or similar enough’, a predicate which will be specified in C4 below. We have explained in

[10] For instance, von Mises’ claim that probability is relative frequency ‘at infinity’, or Popper’s claim that probabilities are propensities (cf. next footnote) might be seen as such less concrete statements.

Section 2 how a stochastic system or experiment in which stochastic events occur can be partitioned in subsystems (test object, initiating and probing subsystems, or environment for natural phenomena).

DEF1. A system (or event occurring in the system) possesses the property of frequency stabilization IFF

(i) it is possible to repeat n identical experiments (with n a number sufficiently large for the desired precision) on identical test objects by applying n times identical initial and final actions or conditions on the object (realized by the initiating and probing subsystems, or more generally, by the environment); and

(ii) in this experimental series the relative frequencies of the outcomes R_j (j = 1,…, J) of the corresponding experimental events (detected by the probing subsystem) converge towards a constant number when n grows; more precisely, the ratios { (number of events in the series that have result R_j) / n } converge towards a constant number when n grows, for all j.

DEF2: Only for an event (or system, or outcome, or variable) that shows frequency stabilization according to DEF1, as it can be tested by an experiment in well-defined conditions, the probability of the result or outcome R_j (j = 1,…, J) of the event is defined, and given by:

$P(R_j)$ = (number of events that have result R_j) / n, for n large enough for the desired precision, and for all j.

Following von Mises we conjecture that the above definitions allow to characterize all physical probabilistic phenomena; i.e. that they cover the physical events or systems described by probability calculus. To some believers of the frequency interpretation this claim may seem acceptable or even obvious. Others may find that a heavy burden of providing arguments awaits us. Clearly, we cannot review all probabilistic systems and show how the definitions apply in all these cases. But this article will provide several examples, and the cited works of von Mises offer a wealth of these (von Mises 1928, 1964). Moreover, under C1 – C6 we provide those arguments and completing notes that seem the most important to us in rendering above claim defendable.

C1. It is well-known that the frequencies or ratios defining probability in a model like the above fulfill the axioms of Kolmogorov's probability calculus (cf. e.g. van Fraassen (1980), Gillies (2000), Khrennikov (2008)).

C2. If one reduces the definitions to their essence, they state that probability *only* exists for systems or events that *can be* subjected to massively repeated physical tests, occurring under well-defined (and composed) conditions; and that in that case probability is given by a simple ratio or 'frequency'. The numerical precision of that ratio grows when the number of experimental repetitions grows. Of course, probabilities of events may 'exist' even if no experiments have been or will be done; but it must be physically possible that such an experiment be done – as stipulated in DEF1. The intimate link between probability and physical experiments, leading to the idea that probability is a property not of objects but of repeatable events, or even better of *repeatable experiments*, is obviously built-in in von Mises' concept of collective (a series of experimental results), as in our model. In the case of physics, this link has been further emphasized and analyzed by philosophers as Popper (1957, p. 67)[11] and van Fraassen (1980, Ch. 6) (see also e.g. von Plato (1994)).

Note that the definitions can be used both for artificial systems (chance games, typically) and natural systems (behaving in a p-random manner without human intervention). In other words our phrasing allows to see how the same concept applies to such widely different events as a die thrown by a human, a molecule in a gas, and a quantum measurement. Indeed, in the case of chance games the initializing and probing subsystems are easily identified as truly separate systems acted upon, or made, by humans (see also C3). In the case of natural probabilistic phenomena (quantum phenomena, diffusion phenomena, growth phenomena, etc.) the initializing and probing systems coincide with – are – the environment. A detailed application of DEF1+2 to the case of gas molecules is given in Appendix 2, as well as other examples.

DEF1 of p-randomness is stated in terms of 'possible' repetitions, or of 'possible' trials, not necessarily actual ones, which highlights the modal aspect of probability (van Fraassen 1980). As already stated, natural, spontaneous random events are normally not initiated nor measured by humans in repeated experiments. Still, according to the model we advocate, they only 'have' a probability if they occur, or can occur, on a massive scale as could be imitated in the lab, or studied under repeatable conditions. The probabilities of such

[11] Popper (1957, p. 67) famously interprets probability as the *propensity* that a certain experimental set-up has to generate certain frequencies.

natural events can only be revealed by experiments imitating the environment; they can sometimes be hypothesized by calculations based on experimental data / assumptions, and representing such experiments[12] (cf. examples in Appendix 2); at any rate such probabilities can only be verified / proven by real experiments. But this seems to be nothing more than the scientific paradigm.

C3. In physics most p-random phenomena or events are characterized by outcomes that are (described by) *continuous* variables, rather than discrete ones. In other words, R will range over an interval of the real numbers, rather than taking discrete values R_j, j=1,…, J as in die throws, coin tosses and other chance games. One could observe that these objects used in games are constructed on purpose by humans in order to 'produce' a limited number of discrete outcomes – in contrast to natural systems. The latter, like diffusion phenomena, gas kinetics, biological growth phenomena, etc. have outcomes that are continuous, or quasi-continuous. (A notorious exception are certain quantum systems, notably when they are carefully 'prepared' – we would say 'initiated'. Still many or most quantum properties have a continuous spectrum.) In the case of a continuous p-random variable x (we used R before), one defines in probability theory the probability density function ρ(x), the meaning of which is however defined via the concept of probability. Indeed, in the continuous case $P(R_j) = \rho(R_j).dR$, defining ρ via P as characterized in DEF2. One can thus for the present purpose safely treat probability density functions and probability on the same foot. A model that interprets probability for discrete variables also does the job for the ubiquitous continuous variables. (Also note in this context that one can formally describe discrete cases by a density function $\rho(R) = \Sigma_j P_j.\delta(R-R_j)$, where δ(–) is the Dirac delta-function: the discrete case is a special case of the more general continuous case.)

C4. DEF1 of p-randomness relies on the notion of 'identical' or 'similar' conditions and objects. It may thus look suspicious in the eyes of a philosopher. But we will now argue that the definition is sound in the sense that it allows to define p-randomness and probability in a scientifically sound manner. First, notice that it seems impossible to define 'similar' or

[12] Von Mises found it important to emphasize that when one applies probability theory to a concrete problem, one always starts from assuming a certain probability distribution. One of his best known maxims is "Probability in, probability out". E.g., when one calculates the chances in urn picking by combinatorics, one typically (and almost always implicitly) starts from assuming that the chance that a given ball is picked is equal for all balls. Then one proceeds to the problem posed (such as: what is the probability that two successively picked balls are white ?). In theoretical physics too, one often guesses or has evidence or derives by using physical theory that the probability distribution of a certain property is of a certain form (say a Gaussian); based on this hypothesis one then uses probability calculus to derive other probabilistic predictions which can be tested. Examples of this are given in Appendix 2.

‘identical’ in a more explicit manner. One could try in following way: “to the best of common knowledge, or of expert knowledge, one should repeat the ‘same’ event in a manner that allows to speak about the ‘same’ event.” But does this help ? It seems however there is no real problem. The important point is that, as defined, *frequency stabilization can be tested for by independent observers; and these observers can come to the same conclusions*. The conditions for doing the ‘frequency stabilization test’ can be communicated for any particular system: “do such-and-such (‘identical’) initial and final actions on such-and-such (‘identical’) systems - and the probabilities P_j will emerge. I found frequency stabilization and the probabilities P_j, so if you do the experiment in the ‘right, identical’ conditions, you should find them to.” It would seem that the problematic term ‘identical / similar’ of probability theory can thus be defined in an operationally fully consistent manner.

We believe the above shows that one could, or should, speak of ‘p-identical’, or rather ‘ρ-identical’, events / systems, where ρ is the probability density of the outcomes of the events under study. To test whether a system is probabilistic, and to identify the probability of the corresponding probabilistic events, one needs to perform a series of experiments on ρ-identical systems (including object, environment, initiating / probing subsystem) – systems that lead to the probability distribution ρ(R) of the results R of the events ‘defined’ or ‘generated’ by the system. Upon this view, then, *‘identical’ is ‘identical’ insofar a probability distribution ρ emerges*. It will also be clear that in an analogous way one could meaningfully define a ρ-system = ρ-{object, environment, initiating /probing subsystem}. We therefore believe that in the context of probability theory, one should in principle speak of ρ-systems, or ρ-identical systems[13].

C5. Does the frequentist model cover the classical interpretation of probability, traditionally used to tackle chance games, urn pulling and the like ? Von Mises (1928/1981, p. 66 ff.) and many modern texts on probability calculus come to this conclusion. We will therefore only succinctly present what we believe to be the essential arguments. Notice, first, that our definitions can at least in principle be applied to such chance games. The initiating /

[13] At this point one might be tempted to include a speculative note. In effect, there seems to exist such a strong link between probability (ρ) and the concept of (identical) ‘system’ – or ‘thing’ –, that one may indeed wonder whether we conceive of something being a (simple) ‘object’ just because it shows frequency stabilization. We saw that composed objects, like a house and the tree in front, or a car and a cloud above, are very unlikely to exhibit frequency stabilization – and indeed we do not conceive of them as simple objects, but as composed ones. (Remember in this context that deterministic systems are just a type of probabilistic systems; obviously they are the first candidates for being associated with ‘things’.)

initializing subsystem is most of the time a randomizing hand (tossing a coin or a die, pulling a card or a colored ball from an urn, etc.); the probing subsystem is often simply a table (plus a human eye).

According to the famous expression of Pierre-Simon Laplace, for calculating a probability of a certain outcome, one should consider "events of the same kind" one is "equally undecided about"[14]; within this set, the probability of an outcome is the ratio of "favorable cases to all possible cases". In all reference books on probability calculus, the latter definition is the basis of probability calculations in chance games. Notice now that dice, card decks, urns containing balls, roulette wheels, etc. are constructed so that they can be used to produce *equiprobable* (and mutually exclusive and discrete) *basic outcomes*, i.e. *having all* P_j *equal, and given by 1/J* (J = the number of basic[15] outcomes). Equiprobability is at any rate the assumption one starts from for making mathematical predictions, and for playing and betting; and indeed Laplace's "events of the same kind one is equally undecided about" would now be termed equiprobable events. Along the lines exposed above, a chance game can thus be seen to correspond to an (artificially constructed) ρ-system with $\rho(R) = \Sigma_j (1/J)\, \delta(R-R_j)$.

It is at this point straightforward to show that in the special case of equiprobable and mutually exclusive events, the frequentist and classical interpretation lead to the same numerical values of probability. Indeed, within the frequentist model, $P(R_j) \rightarrow \dfrac{n.\frac{1}{J}}{n} = 1/J$ (the numerator n / J = the number of R_j-events among n (>>J) exclusive and equiprobable events each having a probability 1/J). Thus the result, 1/J, is equal to the prediction given by Laplace's formula (1 favorable case over J possible cases). Let us immediately note, however, that Laplace's formulation is not superfluous. It allows *in the special case of chance games* for calculation, i.e. theoretical prediction: 'favorable cases' and 'possible cases' can conveniently be calculated by the mathematical branch of combinatorics – a theory of counting, initiated by the fathers of probability theory.

In sum, it appears that the classical interpretation is only applicable to a small subset of all probabilistic systems, namely chance games, in general artefacts having high degrees of symmetry leading to easily identifiable, discrete basic outcomes. For this subset Laplace's

[14] The events thus fulfill the 'principle of indifference' introduced by Keynes.
[15] The 'basic events' or 'basic outcomes' of coin tossing are: {heads, tails}, of die throwing: {0, 1, …, 6}, etc. The probability of 'non-basic', or rather composed, events (two consecutive heads in two throws, etc.) can be calculated by using probability calculus and combinatorics.

interpretation can be seen as a *formula*, handy for calculation, rather than an interpretation of probability. As we often stated, after calculation the only way to verify the prediction is by testing for frequency stabilization. It is among others for this reason we believe the latter property is the natural candidate for a basis of an overarching interpretation.

C6. We have defined frequency stabilization and probability by means of the notion of 'convergence' of a certain ratio when n, the number of trials or repetitions of an experiment, grows. It is clear that this phrasing is close to the original definition by von Mises of probability as $P(R_j) = \lim_{n\to\infty} n(R_j)/n$. However, our phrasing "$P(Rj) = n(Rj)/n$ for a number of trials n that is large enough for the desired precision" avoids the notion of infinity; it may therefore avoid problems of mathematical rigor (see discussion in Appendix 1 and Section 2). Note that from a pragmatic point of view, our definition allows to derive, if one would use it to experimentally determine a probability, numbers that are equal to those identified by von Mises' definition to any desired precision. At least operationally there is no difference in the definitions: they lead to the same results. (Note that some may find von Mises' definition more satisfactory in view of their metaphysical aspirations: one could say that 'the' probability is the limiting relative frequency 'at infinity'. An 'ideal' number that one can seldom calculate with infinite precision (except, for instance, if one can use combinatorics), but that one can always measure with a precision that grows when n grows.)

These notes conclude the basic description of our model. Needless to say, they are a first introduction to the model; we are well aware that questions will remain unanswered. The only way to validate and strengthen a model is to show that it applies to non-controversial cases, and that it may solve problems left open by other models. Concerning non-controversial cases, the examples we have presented and will present below will allow the reader to verify that DEF1+2 work for these cases (more examples can be found in von Mises' works). It is maybe expedient to verify these definitions explicitly in a slightly more subtle case of a natural probabilistic property as the velocity of gas molecules. For completeness, we have done this exercise in Appendix 2. It illustrates the essential role of the 'environment' in the case of natural probabilistic phenomena. Next, Beisbart has analyzed interesting examples of probabilities as they are generated by physical models (Beisbart and Hartmann 2011, pp. 143 – 167). We will also briefly compare these to our model in Appendix 2.

Since the above model condensed in DEF1+2 is a direct elaboration of von Mises' interpretation, it should be able to tackle the probabilistic systems that the latter can tackle. However it is much simpler: we do not need the concept of collective, nor its calculus; our calculus is Kolmogorov's.

4. Applications.

The aim of this Section is to show that the model of Section 3 allows to interpret controversies and solve paradoxes of probability theory, as exposed in general works as e.g. Fine (1973), von Plato (1994), Gillies (2000). We will investigate here problems R1 – R4. We emphasize that we can only provide first arguments. The first thing we should do in this article is to show that the model allows to address a rather wide scope of problems. But we need some charity of the reader: we cannot go in full detail into problems that have been addressed since decades nay centuries by countless philosophers.

R1. According to frequency interpretations it makes no sense to talk about the probability of an event that cannot be repeated. In condition (i) of the definition of p-randomness, repeatability is an explicit requirement. For instance, according to such models it makes no sense to talk about the 'probability' of a dictator starting a world-war and the like: no experiments can be repeated here, and even less experiments in well-defined conditions. Popper's propensity interpretation of probability was an attempt to accommodate such probabilities of single events – quantum events were his inspiration, like the disintegration of one atom (Gillies (2000) Ch. 6). But again, *any quantum property of any single quantum system* can only be attributed a probability if the property can be measured on an ensemble of such systems – as in the case of macroscopic systems. Measurement (verification) of a probability is always done on an ensemble of similar systems and events, whether quantum or classical. Therefore the added value of the propensity interpretation is not obvious to us. For a recent discussion and overview of the intense controversies related to 'single-case probabilities', see Beisbart and Hartmann (2010) p. 6-7.

According to the above model, probability is not a property of an object on its own; it is a property of certain systems under certain repeatable human actions; or, more generally, of certain systems in dynamical evolution in well-defined and massively repeatable conditions. The following slogan captures a part of this idea: probability is a property of certain composed systems. Similarly, probability is not a property of an event *simpliciter*, but of an event in well-defined conditions, in particular initializing and probing conditions. An even more precise way to summarize these ideas would be, it seems, to attribute probability to experiments (van Fraassen 1980, Ch. 6), or experimental conditions (Popper 1957). The advantage of the term 'experiment' is that it only applies to 'composed events' for which the conditions are defined in a scientifically satisfying manner – exactly as we believe is the case for the use of the term probability in physics. Note that a scientific experiment is also, by definition, repeatable.

R2. A classic paradox of probability theory is Bertrand's paradox. It goes as follows: "A chord is drawn randomly in a circle. What is the probability that it is shorter than the side of the inscribed equilateral triangle ?" Bertrand showed in 1888 that apparently three valid answers can be given. A little reflection shows however that the answer depends on exactly how the initial randomization is conceived. Indeed, in the real world there are many ways to 'randomly draw a chord', which may not be obvious upon first reading of the problem. One can for instance randomly chose two points (homogeneously distributed) on the circle by using a spinner[16]; a procedure that leads to the probability 1/3, as can be measured and calculated. But other randomization procedures are possible, leading in general to different outcomes. In other words, *Bertrand's problem is not well posed*, a conclusion that seems now widely accepted (see e.g. Marinoff (1994)). Now, this conclusion is a direct consequence of our model, according to which probability is only defined for experiments in well-defined conditions, *among others initiating conditions*. More precisely (see DEF1+2), probability exists only for events which show frequency stabilization in experiments under precise conditions, initial, final and 'environmental'. For a situation that is experimentally ambiguous no unique probabilities can be defined.

R3. In the following we show how our detailed model allows to counter subjectivist threats – at least in the context of physics. It is both a very popular and very tempting idea that, somehow, "probability depends on our knowledge, or on the subject". This is a key ingredient of subjective interpretations of probability, or subjective Bayesianism, associating probability with strength of belief (see detailed expositions in e.g. Fine (1973), von Plato (1994), Gillies (2000)). When I throw a regular die, I consider the probability for any particular throw to show a six to be 1/6. But what about my friend Alice who is equipped with a sophisticated camera allowing her to capture an image of the die just before it comes to a

[16] More precisely, by fixing a spinner at the center of the circle; a pointer on the spinner and two independent spins generate two such independent random points.

halt ? For her the probability seems to be 0 or 1. Or what about following case: imagine Bob, waiting for a bus, only knowing that there is one bus passing per hour. He might think that the probability he will catch a bus in the next five minutes is low (say 5/60). Alice, who sits on a tower having a look-out over the whole city, might have a much better idea of the probability in case (she might quickly calculate, based on her observations, that it is close to 1). Are these not patent examples of the idea that a same event can be given different probabilities, depending on the knowledge (or strength of belief) of the subject ? And is in that case probability not a measure of the strength of belief of the subject who attributes the probability ?

Examples as these are unlimited, but the above claims can be countered by referring to the boundary conditions that are explicitly part of our definition of probability. In the above model we claimed that probability is only defined – only exists – for repeatable experiments on well-defined composed systems, including among others well-defined conditions of observation. Doing a normal die throw, and observing the result on a table as is usually done, corresponds to a well-defined p-system, with a specific initiating system, probing system etc. In the example, our second observer Alice does not measure the same probability: she does not measure the probability of finding a six on a table after regular throwing and regular observing, but of finding a six after measurement of whether a six will land or not on the table. The latter is a very different, and indeed fully deterministic, experiment; *at any rate, the observing subsystem (including a high-speed camera) is very different*. A similar remark holds for the bus-case; the measurement system (even if just the human eye in a given location) is part of the p-system; one cannot compare observer Alice and observer Bob if their means of observation are widely different. Thus according to our model Alice and Bob do not measure the probabilities of the same event or system: that is why they measure or predict different probabilities.

It thus seems that DEF1+2 allow to safeguard the objective interpretation of probability of physical events. A more modest claim is the following: it seems that cases for which the subjective interpretation might be helpful (in physics), might be re-integrated into the frequency interpretation along above lines.

In sum, based on DEF1+2, if one includes in the experimental series all boundary conditions and especially the probing subsystem, the probability of a given outcome can be seen as an objective measure. True, ‘objective’ (and ‘observer-independent’ even more) is a tricky word here: it means '*identical (and mind-independent) for all observers performing identical experiments*', so *objective in the scientific sense* – even if the observer, or rather the observing subsystem, is in a sense part of the system ! (Remember that the observing subsystem is part of the p-system.) Stated more precisely, the probability of an event is an 'objective' property of the p-system that generates the event, in that independent observers can determine or measure the same probability for the same event.

We emphasize that we believe that this is not in disagreement with the idea – bracketed here – that there is also a subjective element in what someone calls a 'probabilistic system'. Indeed, we saw in Section 2 that it is often extremely difficult to identify conditions of probabilistic behavior, for instance in the experiments of Couder and collaborators. What these experts consider to be a *probabilistic* system (a droplet in fine-tuned conditions of vibration, etc.) will for almost any other person appear to be an extremely complex, erratic, hermetic or why not deterministic system – not a probabilistic one. So in this sense it seems that the attribution of the term 'probabilistic system' does depend on the knowledge of the attributor ! But we believe this is definitively *not* in contradiction with the idea that for a well-defined p-systemprobabilities can be defined (as in DEF1) in an objective or if one prefers inter-subjective manner – as argued above. Needless to say, the mentioned subjective element and its precise link with an objective definition deserves to be investigated in much more detail than we can do here (we have planned to do so elsewhere). Finally, let us note that for physical systems it is not a very shocking claim to consider probability as an objective measure: after all, different physicists all over the world do measure the same probabilities if they do the same experiment.

R4. At this point it seems an interesting and important step to quantum mechanics can be made. Indeed, it seems our conceptual analysis has brought in focus a striking similarity between classical and quantum systems. It belongs to the key elements of the standard or Copenhagen interpretation that "the observer belongs to the quantum system", or at least that the measurement detectors belong to it, and influence it. Suffices here to cite Bohr in his famous debate with Einstein on quantum reality (Bohr 1935, see also Gauthier 1983 on the role of the 'observer' in quantum mechanics). The debate concerned the completeness of quantum mechanics, questioned by Einstein, Podolsky and Rosen in the case of two non-commuting observables of 'EPR electrons'. The key point of Bohr's counterargument is summarized in following quote: "The procedure of measurement has an essential influence on the conditions on which the very definition of the physical quantities in question rests" (Bohr 1935, p. 1025). Bohr invokes here the 'quantum of action' linked to Plank's constant h: any measurement of any property needs the physical interaction of a detector / observing system with the test object under study, an interaction that carries minimally one quantum of energy. Bohr's phrase is arguably one of the most famous quotes of the Copenhagen interpretation (cf. e.g. our [YYY], [ZZZ]), corresponding to one of its most basic ingredients. However, even experts of the foundations of quantum theory as John Stuart Bell, author of Bell's theorem, have complained that precisely this quoted phrase is incomprehensible[17] (cf. Bell (1981) p. 58). But it seems we are now armed to make Bohr's phrasing transparent. According to our model Bohr's words can well be understood as meaning that *the definition of quantum properties depends in a fundamental way on the measurement conditions*. Quantum systems are probabilistic systems; and we have argued throughout this article that the numerical value of a probability depends in a fundamental way on the observing subsystem or conditions. In classical systems one has to look a bit more carefully for examples to exhibit this in-principle fact (we gave several examples), but in the quantum realm it apparently becomes basic. As an example, the probability that an x-polarized photon passes an y-polarizer obviously depends on x and y (the angles x and y are the parameters that describe the initiating and probing conditions).

In conclusion, we believe Bohr's quote is understandable within our model, and that a careful inspection of the concept of probability shows that in *all* probabilistic systems, quantum or classical, the measurement system plays an essential role. We thus come to the perhaps surprising conclusion that, in this respect, quantum systems are not as exceptional as often thought. This claim can be much elaborated, as we have done elsewhere (Vervoort [2012]). Let us also note that other philosophers have come to a similar conclusion from a different angle (Szabó 1995, 2000, 2001, Rédei 2010).

Thus results R1 – R4 are all derived from putting to the fore the experimental 'conditions' mentioned in DEF1+2. Now, that a precise definition of initial and final conditions or subsystems is necessary for defining probability may have – in hindsight – something obvious about it. But if that is true, our model would show nothing more than that analytically defining concepts, making implicit notions explicit, is an efficient method for addressing problems also in the philosophy of physics.

[17] We could, for that matter, not make sense of the phrase before we did the present study on the notion of probability.

5. Conclusion.

We have argued here that maybe not all efforts have been done to save the frequency interpretation of probability as applied to physics. The philosophy literature abounds with criticisms against more basic frequency interpretations (cf. e.g. Fine (1973), von Plato (1994), Gillies (2000)), to the point that von Mises' work might have unduly suffered from it. We have argued that frequency models à la von Mises, if reworked along the lines indicated here, might be a viable interpretation of probability in physics. To make our point we proposed an analytic definition, and showed that it has a certain unifying capacity and that it allows to address problems of the philosophy of probability. We do not claim that all aspects of what 'probability' is are captured by the model. For instance, as exposed in Section 4 (R3) we are very sympathetic to the idea that there are also subjective elements related to the notion of probabilistic system – which we bracketed here, for the moment. We leave it to the reader to judge whether our model is still close to von Mises' theory or not: we retained von Mises' notion of frequency stabilization, but rejected the infinite collective and did not use his calculus.

In some detail, we proposed a model in which it only makes sense to define probability for systems or events (or better, experiments) that exhibit frequency stabilization. This is according to our interpretation (DEF1+2) the essential characteristic of (physical) probabilistic systems; frequency stabilization ultimately provides the only empirical criterion to judge whether a system is probabilistic or not – in agreement with the practice of physics. Next we argued it is useful, if not necessary, to partition probabilistic systems into three parts: if natural, in object and environment; if artificial, in object, initiating subsystem, and probing subsystem. In particular, we claimed that including the probing subsystem into the probabilistic system allows to define physical probability in an objective manner. Only if the probing is defined, the probability is. By the same token we argued that there is an essential parallel between quantum and classical systems: in order to be able to define a probability for an event, in both cases one needs to specify the 'observer', or rather the probing subsystem. Including the initiating subsystem into the probabilistic system also allows to solve paradoxes, such as Bertrand's paradox.

We thus hope to spark a renewed debate on the frequency interpretation in physics. We would welcome questions aiming at challenging the model proposed here. Ultimately the convincing way to challenge the model would be to expose typical cases in which physicists use ‘probability’ in a sense that is not captured by the model. Such cases might exist; maybe the model can be amended to include them.

Acknowledgements. For detailed discussion of the issues presented here I would like to thank Yvon Gauthier, Yves Gingras, Henry E. Fischer, Andrei Khrennikov, and Jean-Pierre Marquis. I also thank Guido Bacciagaluppi for discussion of specific points. Funding was by a scholarship of the Philosophy Department of the University of Montreal.

Appendix 1. Alleged mathematical problems of von Mises’ theory.

Let us have a look at some criticisms of von Mises’ mathematical theory, which is based on the concept of ‘collective’. One criticism, stating that von Mises’ probability as a limit for $n \rightarrow \infty$ is not always well defined (see e.g. Richter (1978)), was dealt with in the main text. It may further become transparent when one realizes that von Mises’ theory is a physical theory, not a strictly mathematical one (see von Mises’ clear arguments in his (1928/1981), e.g. p. 85). A tougher critique concerns the ‘condition of randomness’ that von Mises imposes on collectives. According to him, the randomness of a collective can be characterized as ‘invariance under place selection’ (roughly, the limiting frequencies of a collective should remain invariant in subsequences obtained under ‘place selections’, certain functions defined on the original collective). But which and how many place selections are required ? – von Mises’ critics ask. An important result was obtained by A. Wald, who showed that for any denumerable set of functions performing the subsequence selection, there exist infinitely many collectives à la von Mises: a result that seems amply satisfying for proponents of von Mises’ theory (see the reviews in von Plato (1994), Gillies (2000) and Khrennikov (2008) p. 25). Even the famous objection by J. Ville can be shown not to be a real problem (see Khrennikov (2008) p. 27 and also Ville’s own favorable conclusion reproduced in von Plato (1994) p. 197).

As a side remark, let us note that von Mises’ attempts to mathematically describe randomness led to interesting developments in the mathematics of string complexity, as produced by no-one else than his ‘competitor’ Kolmogorov, and mathematicians as Martin-Löf. A general result of these developments can broadly be stated as follows: real physical randomness seems to defy full mathematical characterization (Khrennikov (2008) p. 28). In our view, this is not really surprising: if a series of experimental results can be generated by an algorithm, one would think it is not random by definition (it may of course look random). A real series of outcomes of coin tosses cannot be generated (predicted) by an algorithm; but – and this is close to magic – it does show frequency stabilization.

This observation allows to counter a curious critique by Fine (1973), who derives a theorem (p. 93) that is interpreted by the author as showing that frequency stabilization is nothing objective, but "the outcome of our approach to data". However, closer inspection shows that Fine's argument only applies to complex mathematical sequences (of 0's and 1's) that can be generated by computer programs. But if a sequence can be generated by a computer algorithm, it is by definition not random, but deterministic, even if it looks random. The essential point of real series of coin tosses is that they cannot be predicted by numerical algorithms… *and* that they show frequency stabilization, as can easily be shown by experimentation. From this perspective, mathematical randomness or complexity, as discussed in relation to the complexity of number strings, has little or nothing to do with physical randomness[18].

To end our succinct review of critiques of von Mises' theory, the decisive point is the following. It is of course possible to withhold essential elements of von Mises' interpretation (I) of probability (as a frequency), even if his calculus (C) would have shortcomings in mathematical strictness - the present article only needs the interpretational part, as often emphasized. On top of that, the sometimes criticized randomness condition appears to be not necessary for the theory, in the sense that the equivalence with Kolmogorov's measure-theoretic approach can be proven without using that condition (see an explicit proof in Gillies (2000) p. 112). This is again, we believe, a clear indication of the fact that the essential feature of probabilistic randomness is frequency stabilization, not invariance under place selections.

Appendix 2. Probability of the velocity of molecules, and other 'theoretical' probabilities.

[18] Except if the number strings are generated by a real physical number generator, which is, as far as I know, never or almost never the case.

To further clarify DEF1+2, it is instructive to verify these definitions in the case of a natural probabilistic property as the velocity of gas molecules. This is a case to which von Mises devotes attention himself (von Mises 1928/1981 p. 20), since it is slightly subtle.

First, recall that in (mathematical) physics a probabilistic ‘event’ formally corresponds to a stochastic physical variable assuming a certain value. Then, the probability that a gas molecule has a certain velocity ‘exists’ according to DEF1+2. Indeed, the system (the gas molecule in its environment), or more precisely the event consisting in the molecule having a certain velocity in the given environment, exhibits frequency stabilization: physics has shown that it is – at least in principle – possible to repeat velocity measurements on the same or on similar molecules, and that in long enough series of such experiments the relative frequencies of the measured velocities converge to fixed values. In actuality these values are for many types of gases given by the Maxwell-Boltzmann distribution, stating that the probability P(v) that the gas molecule has a velocity in an interval dv around v is:

$$\mathrm{P}(v) \;=\; 4\pi .\ \sqrt{\left(\frac{m}{2\pi kT}\right)^{3}}\ .\, v^{2}.\exp\left(\frac{-mv^{2}}{2kT}\right).\, dv\,, \tag{2}$$

where m is the mass of the molecule, k Boltzmann’s constant and T the temperature of the environment. (This of course implies that the probability density $\rho(v) = \mathrm{P}(v)/\mathrm{d}v$.)

This case is a bit subtle because the probabilities (2) are known to be a good description of reality since more than a century, even if in practice they were not directly determined by velocity measurements (at least not until recently). So due to technical limitations in instrumentation, in von Mises’ time it was practically impossible to measure velocities of individual molecules and verify frequency stabilization for these velocities. How then did physicists come to accept (2) as the right formula ? First Maxwell and Boltzmann had derived (2) based on theoretical arguments. But the essential point is that (2) could subsequently be tested in numerous manners in the following *indirect* way: if one assumes (2) as the correct probability for velocities, one can use physical theory to derive various predictions for other variables (say kinetic energy) that are functionally related to velocity (kinetic energy = $mv^2/2$) – and it is these predictions which were tested so often that the starting hypothesis (i.e. (2)) got the status of ‘scientifically generally accepted’, and indeed one of the pillars of statistical physics.

So when we say in DEF1 (i) “it is possible to repeat n identical experiments”, we maybe should add the proviso “direct or indirect”; or better if we say in DEF1 (ii) “the relative frequencies of the outcomes R_j (j = 1,…, J) of the corresponding experimental events [converge]", we might say "the relative frequencies of the outcomes R_j (j = 1,…, J) of the corresponding experimental events, or of functionally related events [converge]".

The above is a somewhat more detailed analysis of von Mises' words (1928/1981 p. 20):

> "It is true that nobody has yet tried to measure the actual velocities of all the single molecules in a gas, and to calculate in this way the relative frequencies with which the different values occur. Instead, the physicist makes certain theoretical assumptions concerning these frequencies (or, more generally, their limiting values), and tests experimentally certain consequences, derived on the basis of these assumptions. Although the possibility of a direct determination of the probability does not exist in this case, there is nevertheless no fundamental difference between it and the other […] examples treated. The main point is that in this case too, all considerations are based on the existence of constant limiting values of relative frequencies […]".

Note that nowadays we likely have means to measure individual atomic speeds. And if one would use these to directly verify (2) one would determine relative frequencies, again in agreement with DEF1+2. But for our concern the most relevant point illustrated by this example is that probability (2) heavily depends on the 'environment' – the essential parameter describing the environment being the temperature T of the gas.

It seems that, mutatis mutandis, the same can be said for the interesting cases of 'theoretical' probabilities that are considered by Beisbart (Beisbart and Hartmann 2011, Ch. 6). Beisbart investigates the interpretation of probabilistic properties as they are predicted by physical models, notably in the case of Brownian motion and the spatial distribution of galaxies. Without entering in any detail, it seems DEF1+2 can well be understood to also apply to these cases, as the reader will easily verify. (For instance, a Brownian particle / system can be subject to repeated tests; what is more, on usual physical intuition, experimental verification of stochastic properties of Brownian particles, say dwell-time or displacement radius, could be done (and doubtlessly is done) via the determination of frequencies as in DEF1+2.) Somewhat more subtle are probabilistic parameters physicists use for describing spatial patterns of e.g. galaxies. If such a model describes the spatial distribution of galaxies in the probabilistic sense, via e.g. the probability $P_R(N)$ that a given radius R contains N galaxies, then it should be possible to apply the recipe of DEF1+2. One can pick 'many' (say n) different regions with radius R on the map, and count within each of these regions the number of galaxies. Then the experimental ratio "(number of regions containing N galaxies) / n" should increasingly approximate $P_R(N)$ if one increases n; if not physicists would not accept the model as probabilistic. Now here the same remark as we already made for meteorological 'probabilities' might apply: it does not necessarily make sense to consider the limit for infinite n (cf. footnote 8).

References

BELL, John S., 1981, 'Bertlmann's Socks and the Nature of Reality', Journal de Physique, 42, Complément C2, pp. C2-41 – C2-62.

BEISBART, Claus and HARTMANN, Stephan, 2011, 'Probabilities in Physics' (Editors), Oxford University Press, Oxford.

BEISBART, Claus, 2011, 'Probabilistic modeling in physics', in Beisbart and Hartmann (Eds.), Probabilities in Physics, Oxford University Press, Oxford, pp. 143 – 167.

BOHR, Nils, 1935, 'Quantum Mechanics and Physical Reality', Nature, 136, pp. 1025-1026.

COUDER, Y., S. PROTIÈRE, E. FORT, A. BOUDAOUD, 2005, Nature 437, p. 208.

COUDER, Y., E. FORT, 2006, Physical Review Lett. 97, p. 154101.

BUNGE, Mario, 2006, Chasing Reality: Strife over Realism, Univ. Toronto Press, Toronto.

EDDI, A. et al., 2011, J. Fluid Mechanics, vol. 674, pp. 433–463

FINE, Terrence, 1973, Theories of Probability, Academic Press, New York.

GAUTHIER, Yvon, 1983, 'Quantum mechanics and the local observer', Intl. J. Theor. Phys., 22 (12) pp. 1141-1152.

GILLIES, Donald, 2000, Philosophical Theories of Probability, Routledge, London.

GNEDENKO, Boris, 1967, Theory of Probability, Chelsea Publishing Co., New York.

KHRENNIKOV, Andrei, 2008, Interpretations of Probability, de Gruyter, Berlin.

KOLMOGOROV, Andrei, 1956, Foundations of the Theory of Probability (2nd ed.), Chelsea, New York.

MARINOFF, Louis, 1994, 'A resolution of Betrand's paradox', Phil. of Science, 61, pp. 1 – 24.

POPPER, Karl, 1957, ‘The Propensity Interpretation of the Calculus of Probability, and Quantum Mechanics’, pp. 65-70, in S. Körner (ed.), Observation and Interpretation, Academic Press, New York.

RÉDEI, M., 2010, ‘Kolmogorovian Censorship Hypothesis for general quantum probability theories’, Manuscrito – Revista Internacional de Filosofia, 33, pp. 365-380.

RICHTER, Egon, 1978, Höhere Mathematik für den Praktiker, Johann Ambrosius Barth, Leipzig.

SZABÓ, L. E., 1995, ‘Is quantum mechanics compatible with a deterministic universe ? Two interpretations of quantum probabilities’, Foundations of Physics Letters 8, p. 421.

SZABÓ, L. E., 2000, ‘Attempt to resolve the EPR–Bell paradox via Reichenbachian concept of common cause’, International Journal of Theoretical Physics 39, p. 901.

SZABÓ, L. E., 2001, ‘Critical reflections on quantum probability theory’, in M. Rédei, M. Stoeltzner (Eds.), John von Neumann and the Foundations of Quantum Physics, Vienna Circle Institute Yearbook, Vol. 8, Kluwer.

TIJMS, Henk, 2004, Understanding Probability: Chance Rules in Everyday Life, Cambridge University Press, Cambridge.

VAN FRAASSEN, Bas, 1980, The Scientific Image, Clarendon Press, Oxford.

VERVOORT, Louis, 2013, ‘Bell's Theorem: Two Neglected Solutions’, Foundations of Physics (2013) 43, pp. 769-791.

VERVOORT, Louis, 2012, ‘The instrumentalist aspects of quantum mechanics stem from probability theory’, American Institute of Physics Conf. Proc., FPP6 (Foundations of Probability and Physics 6, June 2011, Vaxjo, Sweden), Ed. M. D’Ariano et al., p. 348.

VON MISES, Richard, 1928, Probability, Statistics and Truth, 2nd revised English edition, Dover Publications, New York (1981).

VON MISES, Richard, 1964, Mathematical Theory of Probability and Statistics, Academic Press, New York.

VON PLATO, Jan, 1994, Creating Modern Probability, Cambridge University Press, Cambridge.