MIT Open Access Articles

The universal path integral

Massachusetts Institute of Technology

The universal path integral

Seth Lloyd[1]*, Olaf Dreyer[2]

1. Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge MA 02139 USA

2. Kimbernstieg 26, 22455 Hamburg, Germany

Abstract: Path integrals calculate probabilities by summing over classical configurations of variables such as fields, assigning each configuration a phase equal to the action of that configuration. This paper defines a universal path integral, which sums over all computable structures. This path integral contains as sub-integrals all possible computable path integrals, including those of field theory, the standard model of elementary particles, discrete models of quantum gravity, string theory, etc. The universal path integral possesses a well-defined measure that guarantees its finiteness. The probabilities for events corresponding to sub-integrals can be calculated using the method of decoherent histories. The universal path integral supports a quantum theory of the universe in which the world that we see around us arises out of the interference between all computable structures.

---

* Corresponding author: slloyd@mit.edu

Over the past few decades it has become clear that physicists must contemplate seriously the possibility that the observed laws of physics may only represent one out of a large number of possible laws. In string theory, for example, the string theory landscape identifies some 10^{500} possible vacua, each of which could give rise to different properties for the laws of elementary particles, and only one of which might correspond to our observed laws [1]. If our laws are plucked from an ensemble of possible laws, it makes sense to investigate from what types of ensembles of laws ours might arise, and how one particular set of laws is plucked from the ensemble. The theory of algorithmic information [2-4] provides a classical specification of an ensemble from which structures could arise: to each computable structure s it postulates an algorithmic probability $p(s)$ proportional the sum over all programs π that produce the structure of $2^{-|\pi|}$, where $|\pi|$ is the length of $|\pi|$ in bits. Algorithmic probability was originally proposed to make precise the notion of a universal prior probability for Bayesian inference [2]. It assigns high probability to structures that can be computed from short programs. At bottom, however, the observed universe does not arise from classical probablities but from quantum probability amplitudes. Probabilities are derived from superpositions of complex amplitudes, which can exhibit constructive and destructive interference. Accordingly, it makes sense to investigate algorithmic methods for constructing universal superpositions of quantum amplitudes, and to compare how a quantum algorithmic theory of probability differs from the classical one.

The framework on which we base this quantum algorithmic theory is that of path integrals [5-7]. We propose a universal path integral that encompasses all computable path integrals. The path integral possesses a well-defined measure that makes finite all path integrals derived from it. The method of decoherent histories provides a natural method for assigning probabilities to subsets of paths [8-13]. When the paths correspond to physical quantities, such as histories of quantum fields with a particular Lagrangian dynamics, then the probabilities that the universal path integral assigns to decohering coarse-grained subsets of probabilities can be identified with observable probabilities within a universe that obeys that dynamics. The universal path integral provides a mathematically well-defined framework for describing the ensemble of computable quantum theories, and the laws of nature as we observe them formally exist in this ensemble.

The ingredients of a conventional path integral are a set of classical configurations or 'paths' $\{\Phi\}$, for example, a set of configurations for classical fields over spacetime, together with an action $S(\Phi)$, a computable function of the classical configurations [5-6].

To perform the integral, one integrates over all configurations that are compatible with the conditions Γ that one wants to observe, e.g., configurations of fields on initial and final spacelike surfaces, yielding an amplitude for those conditions:

$$\mathcal{A}(\Gamma) = \int_{\Phi \in \Gamma} \mathcal{D}\Phi \ e^{iS(\Phi)}. \tag{1}$$

The probability of Γ is then proportional to $|\mathcal{A}(\Gamma)|^2$. There are a number of technical issues associated with evaluating such path integrals. First of all, one must define the measure of integration $\mathcal{D}\Phi$ in a sensible way. Secondly, one must actually perform the integral: the integral itself is highly oscillatory and hard to approximate. Finally, in the case of theories such as string theory or eternal inflation that attempt to provide a quantum-mechanical description of the universe as a whole, the issue arises of how to extract probabilities for observed features of the universe out of the path integral [7].

The universal path integral provides a mathematically rigorous way of addressing these issues. The classical configurations in the universal path integral are configurations of discrete sequences of symbols such as integers or bits, with actions that are computable functions of those configurations. The path integral sums over all possible computable configurations: as a result, any path integral that is itself computable – for example, a conventional path integral performed in lattice gauge theory – is contained as a sub-integral in the universal path integral. As will be shown, the universal path integral gives well-defined predictions for probabilities of events or sequences of events. In particular, the universal path integral provides a mechanism for how the classical world that we see around us arises out of constructive and destructive interference between different quantum paths.

The universal path integral

Now define the universal path integral. The method of definition relies on the theory of algorithmic probability [2-4]. Let the action S be the output of a classical universal computer U_L that takes input programs bit strings $b = b_1 b_2 \ldots$ written in language L, producing as output the bit string $S(b)$. The bit strings can be either finite or infinite in length. We can think of such bit strings as real numbers in the interval $[0, 1]$, written in binary, e.g., $0.1011 = 11/16$. If S halts in finite time given input b, then it reads only a finite number of bits of b. Call such a set of bits a halting program p: all input bit strings that begin with the same program p yield the same output $S(p)$. That is, programs are 'prefix-free': no halting program is the prefix of another. A halting program p then

3

corresponds to a sub-interval of measure $2^{-|p|}$ where $|p|$ is the length of the program p: if one generates the bits of the input program by flipping a fair coin, the probability that the first $|p|$ bits of the input string yield the program p is $2^{-|p|}$.

Because U_L is universal, some programs never halt and give an output. Membership in the set of halting programs cannot be computed [2-4] (the halting problem). To cope with the halting problem, introduce a computer that always halts in finite time t: $S_t(b) = S(b)$ if U_L halts in time t given input b, $S_t(b) = 0$, otherwise. The universal path integral is then

$$\Sigma_L = \lim_{t \to \infty} \int_0^1 e^{2\pi i S_t(b)} db. \tag{2}$$

Note that for any finite time t, the computer only reads a finite number of bits in the program before giving an output. Accordingly, equation (2) does not suffer from the ambiguity $.1000\ldots = .0111\ldots$ that occurs for infinite bit strings. Because programs are prefix free, the universal path integral can be rewritten as a universal path sum over programs for S_t, all of which halt by time t by definition:

$$\Sigma_L = \lim_{t \to \infty} \sum_{p:|p| \leq t} 2^{-|p|} e^{2\pi i S_t(p)}. \tag{3}$$

The universal path integral Σ_L is a complex number with amplitude between zero and one.

To verify convergence of the universal path integral, note that the magnitude of the difference between the integral evaluated at time t_1 and at time $t_2 > t_1$ can be at most equal to the fraction of programs that fail to halt by time t_1 but that have halted by time t_2. If we simply define t_2 to be equal to the time at which half of the halting programs that failed to halt by time t_1 have halted, and t_3 the same relative to t_2, we note that the maximum possible magnitude of the difference between successive estimates of Σ_L at times t_k and t_{k+1} goes down by at least a factor of two at each step. Consequently, the universal path integral converges in the limit $t \to \infty$. Since at each time t the sum (3) consists of a finite number of terms, the integral is also absolutely convergent.

Because the halting set is uncomputable, the $t \to \infty$ limit in the universal path integral/sum converges more slowly than any computable sum. The amplitude for everything cannot be computed in practice. While it might sound bad at first, the uncomputability of Σ is in fact acceptable: we are interested not in the absolute amplitude for everything, but rather in the amplitudes that predict the results of experiments, given our observations. As will be seen below, such amplitudes represent computable sub-integrals of Σ.

To make the connection to conventional path integrals, we can assume that programs written in language L consist of two parts, $p = p_0 p_1$, where the first part of the program p_0 specifies a finite set of paths $W(p_0)$ and the method for computing an action for members of that set $w \in W(p_0)$. The second part of the program p_1 picks out a particular path $w(p_0 p_1) = w(p)$ from the set, and generates the action $S(p)$ for that path.

The universal path integral sums over all computable configurations of bits and actions. It includes as sub-integrals all the path integrals that physicists would like to perform. Even if the desired path integral consists of a continuous, uncountable set of paths, the universal path integral nonetheless contains sub-integrals that approximate that integral to any desired accuracy. For example, any lattice gauge theory path integral over a finite lattice with the field values truncated to a finite precision corresponds to some finite interval of the universal path integral.

Quantum computing and the universal path integral

All quantum computations are included in the universal path integral. A quantum computation can always be written as a sequence of controlled-NOT gates, Hadamard gates, and so-called $\pi/8$ gates (rotations by $\pi/4$ about the z-axis) [14]. Controlled-NOT gates flip quantum bits conditioned on the values of other bits, Hadamard gates take qubits to equal superpositions of $|0\rangle$ and $|1\rangle$, and $\pi/8$ gates apply a phase of $e^{i\pi/8}$ to $|0\rangle$ and $e^{-i\pi/8}$ to $|1\rangle$. A quantum computation can then be written as a uniform superposition of sequences of bit configurations (paths) determined by the quantum logic gates of the computation, where each Hadamard gate doubles the number of paths and each $\pi/8$ gate applies a phase to each path.

Conversely, the entire universal path integral can be computed on a quantum computer in the infinite time limit. The universal path integral is simply equal to

$$\Sigma_L = \lim_{t \to \infty} \langle \Psi | V_L^t | \Psi \rangle, \tag{4}$$

where $|\Psi\rangle = \int_0^1 |b\rangle db$ is the uniform sum of all input bit strings, and V_L^t is the action of the quantum computer that implements the unitary transformation $V_L^t |p\rangle = e^{2\pi i S_t(p)} |p\rangle$. That is, the amplitude modulus squared of the universal path integral Σ_L is the probability that the quantum computer remains in its initial state. The real and imaginary parts of the amplitude can be extracted in a similar fashion. For example, start in the initial state $|\Psi\rangle \otimes (1/\sqrt{2})(|0\rangle + |1\rangle)$, and use the quantum computer to construct the state $1/\sqrt{2}(V_L^t|\Psi\rangle|0\rangle + \bar{V}_L^t|\Psi\rangle|1\rangle)$. The probability that the quantum computer remains in its initial state is then

proportional to the square of the real part of the amplitude. Sub-integrals of the universal path integral are computable by a quantum computer by restricting the input state in equation (4) to the appropriate sub-interval.

Obtaining probabilities from amplitudes

Let's see how the universal path integral assigns probabilities to events. Use the two-step description $p = p_0 p_1$ given above, where p_0 picks out a set of paths $W(p_0)$ and p_1 picks out a particular path $w(p)$ within that set. A coarse-grained path $\tilde{w}$ is some subset of W. For example, if W is a set of bit strings, $\tilde{w}$ could be the set of bit strings where the first bit takes on the value 0. The amplitude for $\tilde{w}$ is

$$\mathcal{A}(\tilde{w}) = \sum_{p:w(p)\in\tilde{w}} 2^{-|p|} e^{2\pi i S(p)}. \tag{4}$$

The probability for an event in quantum mechanics is usually taken to be proportional to the square of the magnitude of the amplitude for the event:

$$p(\tilde{w}) \propto |\mathcal{A}(\tilde{w})|^2. \tag{5}$$

We must be careful here, as the prescription that probability is proportional to amplitude squared conventionally refers to probabilities for the outcomes of measurements. It is not yet clear what a 'measurement' consists of here: any measurement apparatus must itself be somehow contained within the path integral. Fortunately, the method of consistent or decoherent histories gives us a well-established way to proceed [8-13]: even in the absence of a measurement apparatus, we can still assign probabilities as long as the probability sum rules are obeyed.

Consider two non-overlapping coarse-grained states $\tilde{w}$ and $\tilde{w}'$, e.g., $\tilde{w}' = NOT\ \tilde{w}$, the set complementary to $\tilde{w}$. We would like to assign to these states the 'probabilities'

$$p(\tilde{w}) = |\mathcal{A}(\tilde{w})|^2, \quad p(\tilde{w}') = |\mathcal{A}(\tilde{w}')|^2. \tag{6}$$

Since $\tilde{w}$, $\tilde{w}'$ correspond to mutually exclusive sets of events, to qualify as probabilities $p(\tilde{w}), p(\tilde{w}')$ should satisfy the probability sum rule:

$$p(\tilde{w}\ OR\ \tilde{w}') = p(\tilde{w}) + p(\tilde{w}'). \tag{7}$$

In other words, defining the amplitude

$$\mathcal{A}(\tilde{w}\ OR\ \tilde{w}') = \sum_{w\in\tilde{w}\cup\tilde{w}'} \mathcal{A}(w), \tag{8}$$

we require that

$$|\mathcal{A}(\tilde{w} \ OR \ \tilde{w}')|^2 = |\mathcal{A}(\tilde{w})|^2 + |\mathcal{A}(\tilde{w}')|^2. \tag{9}$$

This requirement to obey the probability sum rule is equivalent to demanding that

$$\text{Re}\mathcal{A}(\tilde{w})\bar{\mathcal{A}}(\tilde{w}') = 0. \tag{10}$$

More generally, two non-overlapping coarse-grained states obey the probability sum rule to accuracy ϵ if

$$\frac{(\text{Re}\mathcal{A}(\tilde{w})\bar{\mathcal{A}}(\tilde{w}'))^2}{|\mathcal{A}(\tilde{w})|^2|\mathcal{A}(\tilde{w}')|^2} \leq \epsilon. \tag{11}$$

This is the usual requirement for destructive interference from conventional quantum mechanics: an observer trying to determine whether the sequences of events $\tilde{w}$ and $\tilde{w}'$ obey the probability sum rule would have to repeat the experiment of seeing how many times $\tilde{w}$ or $\tilde{w}'$ occured $O(1/\epsilon^2)$ times in order to discern deviations from the probability sum rule due to the effects of quantum interference. For example, if $\tilde{w}$ represents a particle showing up in one region of the screen in a double slit experiment, and if $\tilde{w}'$ represents a particle showing up in a non-overlapping part of the screen, the particle must be sent through the slits $O(1/\epsilon^2)$ times to detect the effects of quantum interference.

The method for obtaining probabilities for events from the path integral reveals a crucial difference between quantum [15] and classical algorithmic descriptions of the universe [16-17]. The classical algorithmic description simply assigns algorithmic probabilities $2^{-|p|}$ to programs and to the bit strings that they create. In the classical case, all computable structures are represented, but they do not interfere with each other [16-17]. In the quantum case, by contrast, the world that we see around is arises out of the interference between different computable structures [15]. Computable structures effectively conspire with each other via constructive and destructive interference to create the observable world.

Observable probabilities are language independent

The universal path integral allows us to assign probabilities to coarse grained histories that obey the probability sum rules. As defined so far, these probabilities depend on the language L used to define the universal path integral. The absolute probabilities defined so far are not the same as the probabilities as measured by observers 'living in' the path integral, however: the measured probabilities are not absolute but are conditioned on the fact of the observer's existence, and on the dynamics of the sub-integral that the observer

7

inhabits. We now show the probabilities for observables as measured by observers are in fact independent of the L.

First of all, note that the universal path integral defined according to one language contains the universal path integrals defined according to every other language as a non-zero measure sub-integrals. For any two fully recursive languages L, L', there is a program $p_{L \to L'}$ written in L, that instructs the computer to interpret what follows as a series of instructions in language L', where the symbols recognized by L' have been suitably encoded as bit strings of finite length. The universal path integral $\Sigma_{L'}$ is then equal to the universal path integral Σ_L restricted to the interval in which the initial $|p_{L \to L'}|$ bits of the input program in Σ_L are $p_{L \to L'}$. That is Σ_L contains $\Sigma_{L'}$ as a sub-integral. Note that this means that Σ_L also contains itself as a sub-integral (and does so an infinite number of times). It also means that the amplitudes for events determined by the universal path integral defined by L' differ by at most a multiplicative constant from those defined by L.

Now consider the probabilities for events as observed by someone 'living' in the path integral. The 'life' of such an observer is nothing more or less than a coarse-grained set of events, whose joint probabilities are those for a system that is gathering and processing information about other systems to which it has access. In Gell-Mann's nomenclature [10], an observer is an information gathering and using system, or IGUS, embedded in the path integral. This observer sees events occurring with probabilities that depend on what part of the path integral it occupies. The key point is that the observer has no access to the absolute probabilities of events: it only has access to the probabilities of events conditioned on its existence and on the laws in its sector of the path integral. In other words, this observer, like all the rest of us, is subject to the weak anthropic principle: we only have access to the part of the universe that supports our existence.

The universal path integral allows us to make this conditional nature of probabilities precise: every observer occupies a sub-integral of the path integral governed at bottom by a particular language L, which is typically only partly known to that observer. For example, in our case, the language L specifies the part of the path integral that governs the laws of elementary particles and quantum gravity as they figure in our particular sub-integral, acting in a spacetime that obeys a particular set of initial conditions. Within the part of the path integral specified by L, further sub-integrals pick out the quantum accidents that specify our observed laws of chemistry, biology, economics, etc. [10]. That is, L is the 'language of nature' for the part of the universal path integral accessible to us as

observers. The theory of algorithmic inference [2] then implies that Bayesian updating of the probabilities of theories based on repeated observation, will lead us closer and closer to a full knowledge of the language of nature that specifies our part of the path integral. The nested nature of the universal path integral implies that knowledge of our 'local' language of nature is just as good as knowledge of the 'global' language.

Discussion

The universal path integral contains as sub-integrals all computable path integrals, including the path integrals for lattice gauge theories. It supports quantum computation and can be computed by a quantum computer in the infinite time limit. It supplies quantum amplitudes for fine-grained sequences of events, and predicts probabilities for coarse-grained events. The probabilities for events as viewed by observers 'living in' the universal path integral are independent of the language by which the path integral is defined. The world that we see around us arises out of quantum interference between all possible computable structures.

As in all theories where the observed laws of Nature are just some instance of possible laws, e.g., the string theory landscape, the universal path integral theory of nature is not as satisfying as a theory that predicts the exact laws that we see from one fundamental principle. Because of its intrinsic connection to Occam's razor, however, the universal path integral does not discourage us from looking for ever simpler laws: on the contrary, it exhorts us to carry on in the search for simplicity, in the hope that we will discover new regularities in the language of Nature, and to use those regularities to predict the results of future observations.

References

1. L. Susskind, 'The anthropic landscape of string theory,' in B. Carr, *Universe or multiverse*, Cambridge University Press, 2007; arXiv: hep-th/0302219.

2. R.J. Solomonoff, 'A formal theory of inductive inference Part I,' *Information and Control* **7**, 1-22 (1964); 'A formal theory of inductive inference Part II,' *ibid.* 224-254.

3. G.J. Chaitin, *Algorithmic Information Theory,* Cambridge University Press, Cambridge, (1987).

4. A.N. Kolmogorov, 'Three approaches to the quantitative definition of information,' *Problems of Information Transmission* **1**, 1-7, (1965).

5. K.G. Wilson, 'The renormalization group and critical phenomena,' *Rev. Mod. Phys.* **55**, 583-600 (1983).

6. H.J. Rothe, *Lattice Gauge Theories: an Introduction*, World Scientific, Singapore (2005).

7. A. Linde, 'Sinks in the landscape, Boltzmann brains and the cosmological constant problem,' *J. Cosm. Astro. Phys.* **22**, 1-22, (2007).

8. R. Griffiths, 'Consistent Histories and the Interpretation of Quantum Mechanics ' *J. Stat. Phys.* **36**, 219-272 (1984).

9. R. Omnés, 'Logical reformulation of quantum mechanics I-III,' *J. Stat. Phys.* **53**, 893-932, 933-955, 957-975 (1988); 'Logical reformulation of quantum mechanics IV,' *J. Stat. Phys.* **57**, 359-382 (1989); *The Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, 1994.

10. M. Gell-Mann, J.B. Hartle, 'Classical equations for quantum systems,' *Phys. Rev. D* **47**, 3345-3382 (1993); arXiv: gr-qc/9210010.

11. J.J. Halliwell, 'Decoherent histories and hydrodynamic equations,' *Phys. Rev. D* **58**, 105015 (1998), arXiv: quant-ph/9805062; 'Somewhere in the universe: Where is the information stored when histories decohere?' *Phys. Rev. D* **60** 105031 (1999), arXiv: quant-ph/9902008; 'Decoherent histories and the emergent classicality of local densities,'

Phys. Rev. Lett. **83** 2481-2484 (1999), arXiv: quant-ph/9905094; 'Decoherent Histories for Spacetime Domains,' in *Time in Quantum Mechanics,* edited by J.G.Muga, R. Sala Mayato and I.L.Egususquiza (Springer, Berlin, 2001), arXiv: quant-ph/0101099.

12. J.B. Hartle, 'Quantum pasts and the utility of history,' *Phys. Scripta T* **76**, 67-77 (1998), arXiv: gr-qc/9712001.

13. F. Dowker and A. Kent, 'Properties of consistent histories,' *Phys. Rev. Lett.* **75**, 3038-3041 (1995).

14. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge (2000).

15. S. Lloyd, 'Universe as quantum computer,' *Complexity*, **3/1**, 32-35 (1997); arXiv: quant-ph/9912088.

16. M. Tegmark, 'Is "the theory of everything" merely the ultimate ensemble theory?' *Annals of Physics* **270**, 1-51 (1998); arXiv: gr-qc/9704009.

17. J. Schmidhuber, 'A computer scientist's view of life, the universe, and everything,' in *Foundations of Computer Science: Potential - Theory - Cognition,* Lecture Notes in Computer Science, C. Freksa, ed. 201-208, Springer, (1997); arXiv: quant-ph/9904050.