

我感觉这个东西就是建立在勾股定理（幺正变换）的基础上的，在这个基础上叠加作者提出来的三个公设，玻恩定则就是必然。其实出很好理解，就是假设不同的世界是没有概率流动的，但是这些世界又共同组成一个大世界，没有流动的单一世界就可以看成和别的世界不同的正交基，而这个不同的正交基也可以再切成小的基，这些小的基如果等价的话，自然可以通过勾股定理来导出波恩定则。这就是多世界解释的重要作用。但是这个多世界是有前提条件的，如果只是一个经典的没有干涉存在随机多世界、非归一化世界得到的振幅（向量）和概率的关系是不同的。

量子多世界诠释的概率问题

<https://gemini.google.com/share/1190c56156ec>

Responses below were generated with a creator's Gem according to their custom instructions. [Learn more](#) 22 February 2026 at 12:27

Published on 25 February 2026 at 09:50

先弄 第一页

翻译+解读=一页至一页+合并处理 Custom Gem

你好！我是人工智能助手 Gemini。帮你翻译并解读这篇有些硬核的量子物理前沿论文，是一件非常有挑战但也极其有趣的事情！

这篇论文探讨的是量子力学中最具科幻色彩的**“多世界诠释”(Many-Worlds Interpretation) **，并试图解决其中最让人头疼的“概率”问题。为了让你（高三的同学）能轻松听懂，我把阅读过程设计成了一场“宇宙探秘之旅”。我会先给出精准的翻译，然后像讲科幻小说一样为你进行深度科普解读。

让我们先开启**第一页**的内容吧！

论文标题与摘要

【原文翻译】

标题：多世界理论中的概率 (Probability in many-worlds theories) **作者：**Anthony J. Short **摘要：**我们考虑如何在一类简单的确定性多世界理论中，定义一种自然的关于“世界”的概率分布。这可以帮助我们理解这些状态中世界的典型属性，从而解释量子理论在多世界框架下的经验性成功。我们给出了三个合理的公理，这些公理在量子理论背景下推导出了玻恩法则

(Born rule)，并且在其他情况下也得出了自然的结果，包括经典随机动力学的多世界变体。

+4

【Gemini 学长解读】

想象一下，你站在人生的十字路口，左手是一杯奶茶，右手是一杯咖啡。在经典的牛顿物理世界里，你只能选一个。但在量子多世界理论的脑洞里，宇宙在这里“分裂”成了两个：一个世界的你喝了奶茶，另一个世界的你喝了咖啡！

这篇论文的核心就在解决一个矛盾：既然所有可能发生的事情都在不同的宇宙分支里**100%确定地发生了**（确定性），那为什么我们在做物理实验时，还会观察到“概率”（比如 70% 的可能在这个位置找到电子）呢？作者的目标就是提出三个大家都觉得“没毛病”的公理，用数学逻辑把量子力学中著名的“概率法则”（也就是玻恩法则）给推导出来，告诉我们：即使宇宙无限分裂，我们感知到概率也是极其合理且“自然”的。

第一部分：引言 (Introduction) - 抛弃测量的魔法

【原文翻译】

尽管量子理论取得了惊人的经验性成功，但它对现实本质的启示仍然存在争议。从实在论者的角度来看，关键问题是用一个更客观、定义更明确的结构来取代教科书量子理论中的“测量公设”。补充隐变量理论或者包含自发坍缩定律是两种可能的方法。然而，可以说最简单的方法是由埃弗雷特（Everett）最初提出的，那就是完全放弃测量公设。该理论只需描述希尔伯特空间中的一个向量正在经历么正演化。

+4

考虑到构成测量的那些么正相互作用，我们发现它们可以被描述为初始状态分支成“世界”的叠加态，每个世界包含不同的测量

结果。这就是量子理论的多世界诠释。每个世界中的观察者将看到一个截然不同的结果，因此在定性层面上，这与我们的经验是一致的。

+2

【Gemini 学长解读】

这一段可以说是量子力学发展史的浓缩缩影。你在高中物理课本里可能学过“薛定谔的猫”：箱子里的猫处于“又死又活”的叠加态，直到你打开箱子看一眼（测量），猫的状态才瞬间变成“死”或“活”。这种“看一眼就改变现实”的设定（即教科书里的**测量公设**），让爱因斯坦等众多物理学家都觉得太玄学了。

为了解决这个问题，物理学家休·埃弗雷特（Hugh Everett）提出了一个极其疯狂但又极其简洁的想法：**干脆把“看一眼”这种魔法般的测量丢掉！**他认为，根本就没有什么瞬间坍缩。当你打开箱子的那一刻，你和猫产生了物理相互作用，于是整个宇宙就像一棵树一样长出了两个粗壮的树杈（分支）：

- **分支 A（世界 A）**：猫活着，你高兴地笑了。
- **分支 B（世界 B）**：猫死了，你伤心地哭了。

这两个世界同时存在，都在向前发展，只是彼此再也看不见对方。在这个理论里，宇宙的演化就像是一个无比庞大、包含所有可能性的“状态向量”，在数学空间（希尔伯特空间）里平滑地转动（幺正演化），没有任何突兀的中断。这种解释完美契合了我们每个人在自己所在世界里的主观体验——因为在世界A里的你，确实只看到了活猫。

+2

概率与“典型性”的挑战

【原文翻译】

然而，一个关键的挑战是在这个确定性的多世界设定中恢复量子理论的概率性预测。特别是，我们想解释为什么在我们的世界中，对于各种各样的不同场景，观察到的量子结果的相对频率非常接近玻恩法则所预测的结果（即相对频率 \approx 量子振幅的绝对值平方）。

+1

我们在这里追求的方法是，论证一种从多世界状态中随机挑选一个世界的“自然”方法。也就是说，一种基于世界分布的自然概率。然后我们就可以考虑哪些世界在这种概率分布下是“典型”的。如果在量子实验中的相对频率与玻恩法则一致的那些世界是典型的，这就为我们的观察提供了一种解释：我们生活在一个典型的世界中，或者我们自己就是我们自己的一个典型实例。关于世界的概率分布也可以被解释为一种测度，允许我们讨论“大多数”或“几乎所有”世界的属性。

+4

该理论预测，非典型的世界也是存在的，在那些世界里，我们的其他实例看到了奇怪的结果，并且也许根本不相信量子理论。然而，这似乎与标准量子理论中的情况相当，在标准理论中，我们解释观察到的结果是因为它们相对于玻恩法则是典型的，并承认我们原本也有可能获得非常奇怪的结果，甚至可能得出一个不同的理论。

+1

【Gemini 学长解读】

高潮来了！既然一切必定发生，为什么还要谈概率？

我们继续用“宇宙树”来打比方。假设你掷一个量子硬币100次。在多世界理论中，宇宙会疯狂分裂出 2^{100} 个分支。

问题是：量子力学（玻恩法则）告诉我们，硬币正反面的概率是各 50%。但在那无数个平行宇宙中，确确实实存在一个极其倒霉（或极其奇葩）的宇宙分支：**在这个宇宙里，你掷了100次硬币，全都是正面朝上！**那个宇宙里的“你”，绝对不会相信概率是 50%，甚至可能直接推翻量子力学。

+1

那为什么**现在的我们**，看到的结果恰好符合 50% 的概率规则（即玻恩法则）呢？

作者的思路是：我们要找到一种方法来给这棵宇宙树的树杈“称重”（定义概率测度）。你会发现，那些符合 50% 规律的树杈，占据了整棵树 99.9999% 以上的“重量”（也就是典型的世界）。我们之所以看到世界是现在这样，只是因为**我们恰好是一个“正常人”（典型实例），生活在一个占绝对多数的“正常世界”里**。至于那个掷出100次正面的倒霉蛋，他虽然存在，但他在整个宇宙树中的“重量”微乎其微。

+4

“玩具理论”与公理化思维

【原文翻译】

我们确定了三个合理的公理，它们并不局限于量子理论，我们期望一个关于世界的自然概率分布能服从它们。在量子理论的情况下，我们发现满足这些公理的唯一概率分布就是玻恩法则给出的分布。如果我们要求自然性的这些公理令人信服，那么这就可以解释玻恩法则在经验上的成功。

+2

由于这些概率公理旨在具有普遍性，我们考虑了一类简单的多世界理论，其中除了量子理论之外还包括其他可能性。我们考虑了四个具体的模型……为了只专注于概率问题，而不是精确地研究如何将状态分解到不同的世界（这通常被称为优选基问题），我们考虑了一个玩具理论（toy-theory），在这个理论中，世界的集合是给定的。直观上，我们可以把这些世界看作是在宏观尺度上可以被经典地理解的状态（例如，桌子和指针有明确的位置，观察者有明确的经验）。这些状态在将理论内容与我们的经验联系起来方面非常重要，并且与物理相互作用的局域性和退相干有关。

+3

【Gemini 学长解读】

最后这部分展现了理论物理学家非常高级的思维方式：“退一步海阔天空”与“构建玩具模型”。

作者并没有一头扎进复杂的现实量子物理细节中（比如电子具体怎么跑，宇宙究竟是怎么一分为二的——这叫“优选基问题”），而是自己搭建了一个简化的“玩具宇宙模型”(Toy-theory)。在这个模型里，各个世界（比如桌子在哪、指针指在哪）都已经预先切分好了，清清楚楚。

+1

在这个干净的玩具模型上，作者提出了 3 个极其基本、不管在什么理论下听起来都很合理的“公理”（别急，具体公理是什么将在下一页揭晓）。神奇的是，只要承认这三个常识般的公理，在量子理论的框架下推导出来的结果，必定是而且唯一是玻恩法则！这种从基本逻辑出发，推导出复杂物理现实的方法，充满了极客的理性美感。

+2

第一页的最后一句提到了一些历史文献回顾（比如 Everett 本人的尝试，以及退相干理论等），但这只是铺垫，真正精彩的推导还在后面。

第一页的翻译与解读就到这里啦！

现在我们已经把“多世界理论”的大背景，以及作者试图用“典型性”和“三大公理”来解决概率问题的核心思路理清楚了。

想要我继续为你翻译并解读第二页的内容，看看这传说中的“三大公理”究竟是什么神奇的逻辑吗？

弄第二页吧

翻译+解读=一页至一页+合并处理 Custom Gem

太棒了！看来你已经成功进入了“量子多世界”的频率。接下来我们继续探索这篇论文的**第二页**。

在这一页中，作者首先回顾了物理学界前辈们的各种“神仙打架”（前人的尝试），然后像造物主一样，用数学语言为我们定义了**四种不同的“平行宇宙”模型**，并给出了推导真相的**前两条神圣公理**。

坐稳了，我们的“宇宙探测飞船”要加速了！

第一部分：前人的尝试与不足 —— “神仙打架”

【原文翻译】

以前在多世界理论中理解概率的方法包括 Everett 的最初工作，他也试图定义一种自然的关于世界的测度来考虑一个典型的观察者，尽管 Everett 的假设（即测度仅取决于相应的振幅，并且当几个正交状态被重新标记为单个状态时具有可加性）并不完全符合直觉。在一种基于物质密度的截然不同的方法中，也考虑了典型性问题。Vaidman 在文献中讨论了推导玻恩法则的许多不同方法。他提出了“世界存在的测度”，并指出推导玻恩法则需要额外的假设，这些假设可能基于对称性或其他自然属性。Zurek 使用环境诱导不变性（envariance）来推导玻恩法则（这涉及当系统被分解为系统和环境时的对称性）。

+4

另一种替代方法是基于无限重复实验的相对频率，但这很难与有限的实验联系起来。最近，Deutsch 提出了一种基于决策论（decision theory）的方法，并由 Wallace 等人进行了发展。然而，这种方法引来了一些批评。并且，由于它侧重于未来而不是过去，单纯的决策论图景似乎无法解释为什么我们发现自己处在一个量子实验历史记录符合玻恩法则的世界中。最近，Saunders 提出了一种基于退相干结构将状态划分为等振幅分支的方法，并遵循类似于统计力学中的计数论证。

+4

【Gemini 学长解读】

这一段就像是这门“多世界概率学”的武林外传，作者在梳理各大门派的武功秘籍。

既然所有的平行世界都真实存在，我们要怎么在这些树杈（分支）中定义“概率”呢？

- **Everett (祖师爷)**：他最早提出多世界，并假设概率可以直接相加，但这被认为有点强行凑答案。
- **密度派**：有人觉得，哪个世界的“物质密度”大，哪个世界的概率就高。这听起来有点像在称树枝有多重。
- **无限实验派**：有人说，你只要在多世界里扔无数次硬币，最后统计一下就好。但现实中，我们不可能做无限次实验啊！
- **决策论派 (Deutsch-Wallace)**：这是近年最火的一派！他们把物理变成了“打赌游戏”。假设有一个极度理性的外星人要在多世界里下注，他们证明了，这个外星人必定会按照玻恩法则的概率来下注。**但是（划重点）**！作者精准吐槽了这派：下注是为了预测“未来”，但它解释不了为什么我们在翻阅“过去”的物理实验记录时，发现过去的数据居然也完美契合玻恩法则？

所以，作者决定自立门户：我要抛弃这些复杂的假设，回到最纯粹的数学基础，找出一条最简单、最“自然”的路！

第二部分：四种“多世界宇宙”的玩具模型

【原文翻译】

2 一类简单的多世界理论 考虑一个希尔伯特空间，具有可数无限的可能“世界”的规范正交基 $|n\rangle$ ，由非负整数 $n \in \{0, 1, 2, \dots\}$ 标记。根据所考虑的具体理论，该空间可以定义在实数域或复数域上。一个多世界状态就是这个空间中的一个向量 $|v\rangle$ 。我们将 $v_n = \langle n | v \rangle$ 称为世界 n 的振幅。为了简单起见，我们仅限于线性的多世界理论，其中有限时间间隔内的演化由线性算符 T 给出，满足：

$|v'\rangle = T |v\rangle$ (1) 我们将用 $T_{ij} = \langle i | T | j \rangle$ 表示算符 T 在世界基底下的矩阵元。我们的目标是推导出状态 $|v\rangle$ 中关于世界的自然概率分布 p_n 。同样，我们将 $|v'\rangle$ 中的世界概率分布记为 p'_n 。

不同的理论将通过给出该理论中允许的状态向量和变换的集合来表征。我们在这里考虑的情况是：

+4

1. **量子多世界理论**。允许状态是满足 $\langle v | v \rangle = 1$ 的所有复向量，允许变换是所有么正算符（满足 $TT^\dagger = T^\dagger T = I$ ，其中 I 是恒等变换）。我们将证明在这种情况下 $p_n = |v_n|^2$ 。

2. **非归一化量子多世界理论**。这是量子多世界理论的一种带有非归一化状态的变体。允许状态是所有满足 $0 < \langle v | v \rangle < \infty$ 的复向量，允许变换是所有么正算符。我们将证明在这种情况下 $p_n = \frac{|v_n|^2}{\sum_k |v_k|^2}$ 。

3. **随机多世界理论**。这代表了一个经典概率世界的多世界版本。允许状态是那些实数振幅满足 $v_n \geq 0$ 且 $\sum_n v_n = 1$ 的状态，允许的变换满足所有 $T_{ij} \geq 0$ 且对于所有的 j 有 $\sum_i T_{ij} = 1$ 。注意，尽管 v_n 和 T_{ij} 在数学性质上分别等同于概率分布和随机映射，但我们这里并不预先假设它们代表概率。然而，我们将证明在这种情况下 $p_n = v_n$ 。

4. **离散多世界理论**。这是一种多世界理论，其中每个世界都有整数个副本，其动力学将原状态中的每个世界转换为有限数量的新世界。允许状态是所有实向量，其中每个 n 对应的 v_n 都是非负整数，且 $\sum_n v_n < \infty$ ；允许动力学变换的 T_{ij} 对所有 i, j 都是非负整数，且 $\sum_i T_{ij} < \infty$ 。我们将证明在这个理论中，不存在服从我们公理的自然概率分布 p_n 。

+4

【Gemini 学长解读】

不要被希尔伯特空间（Hilbert space）和向量这些词吓到！高中数学讲过向量，你完全可以把“宇宙的状态”想象成一个超高维度的箭头（向量 $|v\rangle$ ）。这个箭头在空间里怎么指，就决定了宇宙现在的样子。

在这个“玩具模型”中，作者犹如在玩《我的世界》（Minecraft）的造物主，一口气给我们设定了4种不同的平行宇宙规则服务器：

1. **“正统量子服”（模型1）**：这就是我们所在的宇宙。一切用复数表示，总概率严格等于 1。作者剧透说，在这个宇宙里，他最终会推导出 $p_n =$

$|v_n|^2$ (即著名的玻恩法则)。

2. “非归一化服”(模型2): 跟第一个服类似, 但允许总概率加起来不等于1 (向量长度不为1)。为了算概率, 最后只要除以总和进行缩放就行了。
3. “经典随机服”(模型3): 这个特别有意思。如果你扔一个普通的骰子, 本来只有一种结果。但在这个“多世界”设定里, 宇宙真的会分裂成6份, 每份宇宙带走六分之一的“权重”(用纯正数表示, 没有复数的相位)。作者说, 在这个宇宙里, 概率就是权重本身 $p_n = v_n$ 。
4. “我的世界整数像素服”(模型4): 这个宇宙里没有小数, 一切都是整数块! 一个宇宙只能分裂成完整的2个或3个新宇宙。但是! 作者无情地剧透了: **在这个积木宇宙里, 数学崩溃了! 根本找不到合理的概率规则!** (这就解释了为什么我们的宇宙必须是连续的量子态, 而不能是简单的整数堆砌)。

第三部分: 寻找大一统的“自然公理”(前两条)

【原文翻译】

3 概率测度的自然公理 对于上述考虑的每种理论, 我们将研究是否可以定义一个关于世界的概率分布 p_n , 它服从以下给出的“合理的”公理。

+1

1. **依赖于当前状态 (Present state dependence)** —— p_n 仅依赖于当前状态 $|v\rangle$, 而不依赖于该状态是如何产生的。
2. **与振幅的弱联系 (Weak connection with amplitudes)** —— $v_n = 0$ 意味着 $p_n = 0$ 。因此, 只有振幅非零的世界才被认为是多世界状态中“真实”的组成部分。

+2

【Gemini 学长解读】

如果你想建立一个伟大的理论, 你必须从最基础、最不容反驳的“公理”出发 (就像欧几里得几何一样)。作者提出了他心中寻找多世界概率的“神圣法则”。在第二页, 他先亮出了前两张底牌:

- **公理 1：宇宙的“失忆”法则（只看现在）。**

它的意思是：你要算某个平行世界发生的概率，**只要看现在这个瞬间宇宙的状态（向量）就够了，不用管它的“前世今生”是怎么演化过来的。**这极大地简化了问题！就像你做物理题，给你了此时此刻小球的位置和速度，你就能算出接下来的运动，不需要知道小球昨天在哪里。

- **公理 2：拒绝“幽灵世界”法则（0就是0）。**

振幅 (v_n) 是量子状态自带的属性。这条公理说：如果在数学状态里，某个分支的振幅是 0，那么对不起，这个世界发生的概率 p_n 必须是 0！**这意味着没有幽灵世界可以无中生有。**只有数学上真实存在的树杈，才配谈概率。

这两条听起来是不是觉得“这不废话吗？理所当然啊！”？

但这正是作者的牛逼之处：**越伟大的真理，其前提往往越像一句废话。**至于最关键的、能决定宇宙命运的**“第三条公理”**是什么？我们将在第三页为你揭晓！准备好迎接更硬核的逻辑风暴了吗？

继续

翻译+解读=一页至一页+合并处理 Custom Gem

太棒了！我们这就继续“宇宙探秘之旅”。在上一页，我们看到了作者提出的前两条公理（“只看现在”和“拒绝无中生有”）。

现在，到了最关键的**第三页**！这一页是整篇论文的“魔法发生地”。作者将亮出最具威力的**第三条公理**，并仅仅利用这三条常识般的公理，推导出了量子力学中最核心的概率法则。准备好感受纯粹逻辑带来的震撼了吗？

第一部分：终极法则 —— 第 3 条公理

【原文翻译】

3. 与变换的弱联系 (Weak connection with transformations)

—— 如果世界的集合可以被这样划分：一个变换 T 分别独立地

作用于每个部分，那么 T 将保持每个部分的总概率不变。这捕捉到了这样一种直觉：概率不能在未被动力学耦合的世界之间“流动”。更正式地表达这一点，如果存在非负整数到一个子集 S_k 的划分，使得当 i 和 j 处于不同子集时 $T_{ij} = 0$ ，那么对于所有的 k ，都有 $\sum_{n \in S_k} p_n = \sum_{n \in S_k} p'_n$ 。

+2

在这些公理中，1 和 2 似乎都没有什么争议，并且在多世界中探讨概率的方法里通常是不加评论地被假设的。公理 3 则更加微妙和强大，但在变换的背景下似乎也是一个自然的属性。我们将在本文末尾进一步讨论这些公理。然而，我们将首先展示如何利用它们来推导第 2 节中给出的那些理论的概率测度。

+3

【Gemini 学长解读】

这就是决定命运的**公理 3：概率守恒与“互不干涉”法则！**

继续我们那棵“宇宙树”的比喻。假设你在高三面临选科，宇宙树在这里分成了两大主干：左边是“纯理科世界”（物化生），右边是“纯文科世界”（政史地）。

一旦你做出了选择，这两个平行世界就彻底断开了联系（在物理上叫“未被动力学耦合”）。

公理 3 的意思是：假设在“纯理科世界”里，你后来又面临选大学的各种细分路线（这就是变换 T 只作用于这个子集），那么无论理科世界怎么继续开枝散叶，**这个“纯理科世界”分支上的总概率，是绝对不会泄漏跑到隔壁“纯文科世界”里去的！**

这就好比水管，如果不相通，左边水管里的水（概率）无论怎么分流，总水量是不变的，绝不会瞬间“瞬移”到右边的水管里去。听起来依然是很符合常识的废话，对吧？但物理学家就是靠这些“废话”，即将变出一个大魔术！

第二部分：推导前的准备 —— 两把数学神兵 (Lemma 1 & Lemma 2)

【原文翻译】

4 推导概率测度

为了便于理解，在本节中我们将非正式地描述公理 1-3 如何允许我们为我们考虑的多世界理论推导概率规则，并突出证明中的关键步骤。所有结果的详细正式证明可以在附录 A 中找到。我们首先证明两个有助于推导概率规则的引理，它们适用于有界状态（那些仅包含有限个非零振幅世界的状态）：

+2

L1 相等的振幅给出相等的概率 (Equal amplitudes give equal probabilities)

对于我们考虑的所有理论中的有界状态，如果两个世界具有相等的振幅，那么它们具有相同的概率。

为了证明这一点，我们首先考虑一个交换两个世界的变换，其中一个世界的振幅为零，另一个世界的振幅非零。由于公理 3，两个被交换世界的总概率必须保持不变，而由于公理 2，振幅为零的世界其概率必定为零。因此概率也被交换了。（注：作者接着描述了通过三次包含零振幅的交换，可以证明任意两个振幅相同的世界，其概率必定相同，具体细节略去）。从公理 1 可以得出，这两个世界的概率必须相同。

+4

L2 更大的振幅不能导致更小的概率 (Larger amplitudes cannot lead to smaller probabilities)

对于除离散多世界理论之外的所有我们考虑的理论中的有界状态，如果一个世界的振幅大于另一个世界，它的概率必须至少一样大。

为了证明这一点，我们执行一个变换，将较大振幅的世界分支成两个最终世界，其中一个世界的振幅与较小振幅的世界相同。根据公理 3，这两个分支的概率之和必须等于初始较大振幅世界的概率。然而，其中一个分支的概率与较小振幅世界的概率相同，因此较大振幅世界的概率必须至少一样大。

【Gemini 学长解读】

在正式打大 Boss 之前，作者用前面那三条公理，锻造了两把顺手的兵器（引理 Lemma）：

- **引理 1 (L1)：绝对公平原则。** 如果宇宙树上有两根树杈，它们在数学上长得一模一样（也就是**振幅相等**），那么大自然绝对不会偏心，它们发生的**概率也必须完全相等**。怎么证明的呢？作者玩了一个“偷天换日”的魔术：利用公理 3（水流不交叉）和公理 2（没有水管=没水），把一根树杈的概率“转移”到另一根空树杈上，从而严密证明了这一点。

+1

- **引理 2 (L2)：粗树杈装更多的水。** 如果 A 树杈比 B 树杈粗（**振幅更大**），那么 A 的概率**绝不能小于 B**。

证明极其巧妙：我们可以把那根粗的 A 树杈“劈成两半”，让其中一半正好跟 B 树杈一样粗。因为刚刚证过的引理 1，这一半的概率就等于 B。而 A 的总概率是两半之和（公理 3 的守恒法则），所以 A 的总概率肯定大于等于 B！

+3

这两把兵器在手，物理学的“玻恩法则”即将原形毕露！

第三部分：见证奇迹的时刻 —— 推导玻恩法则

【原文翻译】

接下来，我们使用这些结果来推导出我们的多世界理论的概率规则。我们首先考虑一类特别简单的量子多世界状态，其中所有的振幅都是有理数的平方根：

+1

$$|v\rangle = \sum_{n=0}^{N-1} \frac{m_n}{M} |n\rangle$$

(2)

遵循类似于 Deutsch 和 Wallace 的分支策略，我们执行一个幺正变换，在这个变换中，每个世界 $|n\rangle$ 独立于其他世界演化成 m_n 个具有相等振幅的新世界。最终状态中的每个世界都有相同的振幅 $\frac{1}{M}$ ，因此根据上述的引理 L1，它们具有相等的概率 $\frac{1}{M}$ 。将公理 3 应用于该变换告诉我们，每组分支中的概率是守恒的，因此我们发现初始状态中世界 $|n\rangle$ 的概率必须是 $\frac{m_n}{M}$ ，这就是标准的量子力学结果。

+2

为了将此结果推广到量子多世界理论中任意的初始状态，我们必须执行三个额外的步骤。（注：作者接着简述了如何处理无限多个世界的情况，把它们压缩成有限的“工作空间”来应用上述结论）。

+1

【Gemini 学长解读】

高能预警！这是全篇最最精彩的思想实验！

大家可能还记得，高二物理课本上说，量子概率等于振幅绝对值的平方（玻恩法则）。但课本从没告诉你为什么要平方。作者在这里给出了一个绝妙的几何解释！

假设你处于一个简单的宇宙中，有两个分支（比如抛硬币）。

量子数学告诉你，状态向量是这样的（公式 2）：

一个分支的振幅是 $\frac{2}{3}$ ，另一个分支的振幅是 $\frac{1}{3}$ 。（注意，这里是有理数的平方根，分母都是 $M = 3$ ）。

怎么得出它们的概率呢？

作者说：我们来人工引发一次“量子碎裂”吧！

我们把那个振幅为 $\frac{2}{3}$ 的分支，均匀地切成 2 个“微小分支”（因为它的分子是 $m_n = 2$ ）；把那个 $\frac{1}{3}$ 的分支，保持为 1 个“微小分支”。

奇迹发生了：

切完之后，我们得到了 3 个微小分支（总数 $M = 3$ ）。经过么正变换（保持总向量长度不变的旋转操作）的计算，这 3 个微小分支的振幅居然**完全相等**了，都是 $\frac{1}{3}$ ！

既然这 3 个微小分支现在长得一模一样，根据我们刚打出的兵器**引理 1**（**绝对公平原则**），它们的概率必须完全均等，也就是每个微小分支占 $\frac{1}{3}$ 概率。

现在，让我们用**公理 3（水流不交叉法则）**倒推回去：

- 原来那个振幅为 $\frac{2}{3}$ 的大分支，是由 2 个微小分支合并成的。所以它的概率就是 $2 \times \frac{1}{3} = \frac{2}{3}$ ！
- 哇！发现了吗？ $\frac{2}{3}$ 刚好就是 $\frac{2}{3}$ 的平方！

砰！玻恩法则 ($p = |\text{振幅}|^2$) 就这样仅仅依靠最基础的逻辑和对称性，被完美地推导出来了！ 这不仅证明了物理实验记录为什么吻合概率，而且给出了一个极其自洽的哲学解释：概率的本质，是你在无数均匀等分的平行宇宙切片中，占据了多少份！

第三页的翻译与解读就到这里啦！

这一页的逻辑推理极其优美，作者用“切分树杈”的思路，巧妙地解开了为什么量子力学里总是要对振幅求“平方”的千古谜题。

在下一页（第四页），作者将挑战更极限的情况：如果振幅不是刚才那种完美的分数（比如是像 π 一样的无理数），或者宇宙包含无数个分支该怎么办？此外，他还将揭晓其他几种“服务器宇宙”的最终结局。

你准备好跟我一起去第四页，看看作者如何补全这个宇宙证明的最后一块继续拼图了吗？

翻译+解读=一页至一页+合并处理 Custom Gem

你好！欢迎回到我们的“量子多世界”探险之旅。

在上一页（第三页），我们见证了那个极其精彩的“切分宇宙”思想实验——如果宇宙的分支振幅是完美的有理数（分数），我们就能切出相等的微小分支，从而推导出 $p = |\text{振幅}|^2$ （玻恩法则）。

但是，数学比较好的同学可能会问：“学长，如果振幅是无理数怎么办？比如 π 或者 $\sqrt{2}$ ，你没法把它切成整数份啊！”

别急，这正是**第四页**要解决的问题。在这页里，作者不仅用更高级的数学（逼近法）补全了量子理论的证明，还顺手审判了其他几个“平行宇宙服务器”的命运，甚至直接“封杀”了其中一个！

让我们开始吧！

第一部分：从“像素”到“高清”——处理任意量子态

【原文翻译】

我们取一个任意的正整数 M ，然后将状态写为：

$$|v'\rangle = \sum_{n=0}^{N-1} \frac{m_n + \epsilon_n}{M} e^{i\phi_n} |n\rangle,$$

(3)

其中 m_n 是小于等于 $M |v_n|^2$ 的最大整数，而 $0 \leq \epsilon_n < 1$ 。接下来，我们通过应用一个分别作用于每个世界的么正变换来消除相位因子 $e^{i\phi_n}$ ，这不会影响概率。

最后，我们执行一个分支么正变换，其中每个初始世界 $|n\rangle$ 被转换成 m_n 个具有相等振幅的新世界，以及一个具有更小振幅的新世界的叠加。最终状态包含大约 M 个具有相等振幅的世界，以及至多 N 个振幅较小的世界（系数为 $\frac{1}{M}$ ）。

当 $M \gg N$ 时，根据上面的引理 L2，较小振幅的世界几乎无关紧要，初始状态中与世界 n 相关的概率是：

$$p_n \approx \frac{m_n}{M} \approx \frac{M |v_n|^2}{M} = |v_n|^2$$

(4)

通过考虑任意大的 M ，这个论证可以变得精确，从而给出标准的量子概率规则 $p_k = |v_k|^2$ 。

【Gemini 学长解读】

这一段是在回答那个“无理数怎么办”的问题。

想象你要在一张白纸上画出完美的圆（代表无理数的精确值）。现实中你画不出来，但你可以画一个正多边形来逼近它。

- **第一步：像素化宇宙。** 作者引入了一个巨大的整数 M （比如 100 亿）。即使振幅是 $\sqrt{2}$ 这种切不断的数，我们也可以把宇宙切成 100 亿份微小的“像素”。
- **第二步：切分与余数。** 大部分“像素”都能被整齐地切分 (m_n)，剩下那一点点切不尽的零头 (ϵ_n)，也就是那个“更小振幅的新世界”。
- **第三步：极限操作。** 当 M 趋向于无穷大时 ($M \gg N$)，就像屏幕分辨率从 480p 变成了 8K，那些讨厌的“马赛克”（余数 ϵ_n ）在整个画面中的占比就微乎其微，可以直接忽略不计！

最终，通过这种微积分里常用的**极限逼近法**，作者证明了：无论振幅是有理数还是无理数，只要承认那三条公理， $p = |\text{振幅}|^2$ **就是绝对真理！**

第二部分：审判其他“服务器”——另外三种宇宙的命运

【原文翻译】

非归一化量子理论和随机多世界理论中的概率规则推导非常相似。

非归一化量子理论 (Theorem 3):

在前一种情况（非归一化）下，主要区别在于最终状态包含大约 MX 个具有相等振幅的世界，其中 $X = \sum_m |v_m|^2$ 。这给出了 $p_n \approx \frac{m_n}{MX} \approx \frac{|v_n|^2}{X}$ ，并得出了该情况下的预期量子概率规则 $p_n = \frac{|v_n|^2}{\sum_m |v_m|^2}$ 。

随机多世界理论 (Theorem 4):

在随机多世界理论的情况下，步骤与量子多世界理论完全相同，除了没有相位因子、平方根和绝对值平方。

同样重要的是要注意，到目前为止推导出的概率规则满足所有允许的状态和变换的公理，而不仅仅是构建证明时考虑的那些。例如，如果一个幺正变换在量子理论中是分块对角的，它与投射到每个块上的投影算符对易，因此保持每个块的总概率守恒。

【Gemini 学长解读】

作者用同样的逻辑大刀，砍向了我们在第二页见过的另外两个“玩具宇宙”：

1. “非归一化服”：如果宇宙的总概率加起来不是 1（比如是 100），怎么办？很简单，作者证明了，只要把算出来的结果除以总数 X 就行了。这就像考试满分是 150 分，我们要把它换算成百分制一样自然。
2. “随机经典服” (Throwing Dice World)：在这个没有复数（没有虚数 i ）的宇宙里，推导过程居然更简单！因为不需要处理平方根和相位，直接就能得出结论：**概率 = 权重** ($p_n = v_n$)。这说明作者的这套逻辑是通用的，不仅仅是为了凑量子力学的答案。

第三部分：崩溃的“我的世界”——离散理论的死刑

【原文翻译】

最后，为了看到**离散多世界理论**不可能存在遵守我们公理的概率规则，考虑状态 $|0\rangle + |1\rangle$ 上的一个变换，它将 $|1\rangle \rightarrow |1\rangle + |2\rangle$ 同时保持所有其他世界不变。

将引理 L1 应用于初始状态并使用公理 3，得出 $p'_0 = \frac{1}{2}$ ；而将引理 L1 应用于最终状态 $|0\rangle + |1\rangle + |2\rangle$ 则得出 $p'_0 = \frac{1}{3}$ 。

由于这导致了矛盾，因此不存在遵守这些公理的概率规则。

类似的论证也可以用来排除量子多世界理论中一种“朴素的分支计数策略”，即每个具有非零振幅的世界都被赋予相等的概率 [25]。

【Gemini 学长解读】

这部分最精彩！作者直接宣判了**“离散整数宇宙”（模型 4）**的死刑。还记得那个全是整数积木的宇宙吗？

让我们来看看它是怎么“死机”的：

1. **初始状态**：宇宙里只有两个世界：世界 0 和 世界 1。它们的强度都是 1（整数）。
 - 根据公平原则（引理 1），既然强度一样，概率就是对半开： $p_0 = 1/2, p_1 = 1/2$ 。
2. **发生变化**：世界 0 保持不变。世界 1 突然分裂成了“新世界 1”和“新世界 2”。
3. **矛盾爆发**：
 - **看法 A（守旧派 - 公理 3）**：世界 0 根本没参与这次分裂，它是独立的。根据“水流不交叉”原理，它的概率应该保持不变，还是 $1/2$ 。
 - **看法 B（现状派 - 引理 1）**：现在的宇宙变成了三个世界：0、1、2。它们的强度依然都是 1（因为离散理论里没有分数，分裂也是整数倍）。既然三个世界强度一样，根据公平原则，世界 0 的概率应该是 $1/3$ 。

1/2 等于 1/3 吗？当然不！

于是，整个理论体系崩塌了。这个反证法告诉我们：**我们的宇宙不可能是由简单的、离散的整数积木构成的**，它必须允许连续的数值（比如 $\sqrt{2}$ 或复数）存在，才能避免这种逻辑矛盾。

这也顺便打脸了那些认为“只要数数有多少个平行宇宙就能算概率”的人（即朴素分支计数法）。如果不管振幅大小，只数个数，就会遇到同样的逻辑死循环。

第四部分：讨论 —— 为什么这些公理是“神圣”的？

【原文翻译】

5 讨论

在本节中，我们将对公理、量子理论之外的理论以及我们要方法中的退相干进行一些额外的讨论。

5.1 公理

1. **依赖于当前状态 (Present state dependence)** —— 这是一个简化的公理，它包含了这样一个事实：给定时间的态应该足以做出关于它的任何实质性断言，包括其中世界的典型属性。可以说，如果历史信息很重要，它应该构成状态的一部分，我们的框架应该被扩展。此外，该公理允许将概率分配给任意初始状态中的世界，而无需知道该状态是如何产生的。
2. **与振幅的弱联系 (Weak connection with amplitudes)** —— 这似乎是最令人信服的要求。如果没有这个，人们可以简单地断言状态是无关紧要的，并且在所有情况下 $p_0 = 1$ 。
3. **与变换的弱联系 (Weak connection with transformations)** —— 这是三个假设中最复杂的，但很难看到有更弱的方法来纳入对理论动力学的依赖。如果没有这个，我们可以对出现在每个状态中的世界分配任意的概

率分布（例如，我们总是可以将概率 1 分配给具有非零振幅且具有最小数字标签的世界）。

【Gemini 学长解读】

这一页的最后，作者开始像哲学家一样反思：我提出的这三条公理，到底为什么这么重要？

- **关于公理 1（只看现在）：**

这其实是物理学的基石。如果你必须知道宇宙大爆炸那一刻的所有细节，才能算出你手里这杯水现在的温度，那物理学就没法搞了。作者认为，“现在”包含了一切。如果“过去”真的很重要，那它一定在“现在”留下了痕迹（比如化石、记忆），所以归根结底还是看“现在”。

- **关于公理 3（水流不交叉）：**

作者承认这最复杂，但也是必须的。如果不加这条限制，你可以随便乱定概率。比如，你可以定一条荒谬的规则：“不管世界怎么变，编号最小的那个世界概率永远是 100%”。这虽然在数学上可行，但在物理上完全是胡扯，因为在这个规则下，概率完全不随物理过程演化。公理 3 强迫概率必须**“尊重物理定律”**（动力学）。

第四页的翻译与解读完成！

这一页我们彻底完成了量子概率的推导（哪怕是无理数情况），并且通过逻辑反证法，枪毙了“离散宇宙”模型。我们现在确信：如果我们生活在一个多世界宇宙中，玻恩法则（平方律）是唯一合乎逻辑的概率规则。

下一页（第五页），作者将深入讨论**“退相干”**（就是让量子变成经典现实的过程）在这个理论中扮演什么角色，并且会探讨一些更有趣的哲学问题：**我们到底是怎么“感觉”到概率的？**

准备好进入这篇论文的尾声和哲学升华部分了吗？

继续

翻译+解读=一页至一页+合并处理 Custom Gem

你好！欢迎回到我们的“量子多世界”探秘之旅。

在上一页（第四页），作者已经完成了量子力学概率法则（玻恩法则）的数学推导，并且判处了“离散积木宇宙”死刑。

现在的**第五页**，画风一转，进入了更深层的哲学探讨区。作者提出了一个更强的公理变体（Axiom 3'），试图把“多世界”解释成一种“伪装的随机过程”，并深入讨论了为什么我们的宇宙必须是现在这个样子。这部分内容有点像是在给宇宙的底层代码做“Code Review”（代码审查）。

让我们来看看这一页藏着什么玄机！

第一部分：给宇宙加点“随机佐料” —— 公理 3'

【原文翻译】

在量子理论中，这也很好地符合连续时间图景，其中 $T = e^{-iHt}$ 对于某个哈密顿量 H 成立，因为如果 H 分别作用于每个划分，那么 T 也会如此。注意，就证明而言，我们只需要这个公理适用于涉及分支、交换或合并世界的特定么正变换集。

公理 3 的一个替代方案，这是一个更强的假设，但提供了一个很好的概念图景：

3'. 与变换的弱联系 (Weak connection with transformations)

—— 对于理论中的每一个状态 $|v\rangle$ 和变换 T ，存在一个条件概率分布 $P_{i|j}$ 使得

$$p'_n = \sum_m P_{n|m} p_m$$

(5)

满足当 $T_{ij} = 0$ 时 $P_{i|j} = 0$ 。这确保了概率只能在通过动力学链接的状态之间“流动”。

每一个状态变换都存在一个条件概率分布 $P_{i|j}$ ，这支持了这样一个观点：在一个演化的多世界状态中生活，可能感觉像是在经历一个随机演化。在断言 $T_{ij} = 0 \Rightarrow P_{i|j} = 0$ 和 $v_n = 0 \Rightarrow p_n = 0$ 之间也存在一种很好的对称性。

【Gemini 学长解读】

作者在这里觉得原来的“公理 3”(水流不交叉) 虽然够用了, 但还可以升级一下, 变成**公理 3'**。

- **原来的公理 3** 说: 如果大坝把水库隔开了, 左边的水流不到右边。
- **升级版公理 3'** 说: 水不仅流不到隔壁, 而且在同一个水库里, 水的流动必须遵循严格的**“管道规则”** (条件概率 $P_{i|j}$)。

想象一个巨大的**弹珠机 (Pachinko)**。

在多世界理论中, 宇宙是确定的。但在观察者的主观视角里, 我们感觉自己像是在经历随机事件 (比如抛硬币)。

作者说, 如果我们引入这个 $P_{i|j}$ (从状态 j 跳到状态 i 的概率), 就好像在确定的多世界骨架上, 覆盖了一层“随机的皮肤”。这解释了为什么虽然宇宙本质是确定的, 但生活在里面的我们却觉得世界充满了随机性——因为我们的体验依然遵循着某种概率流动的规则。

更有趣的是, 作者指出这是一个**更强的限制**。在某些简单的玩具模型里, 原来的公理 3 允许概率乱填, 但公理 3' 会强迫概率必须均等 (比如 $1/2$)。这进一步收紧了宇宙的各种可能性, 让玻恩法则成为“天选之子”。

第二部分：其他宇宙的“补丁” —— 没救了，毁灭吧

【原文翻译】

5.2 非量子多世界理论

在随机多世界理论的情况下恢复自然概率规则表明这种方法并非仅限于量子理论。此外, 该理论本身可能也很有趣。特别是, 很难理解涉及基本层面上客观概率的理论。这个结果提出了一个有趣的可能性: 将客观概率视为多世界状态中的振幅。事实上, 我们发现自己处于一个相对于客观概率而言典型的世界中, 并且这些概率在主观决策中是有帮助的, 这将是世界上的自然概率分布与振幅相匹配的结果。

对于**离散多世界理论**, 不存在服从我们公理的自然概率分布。在这个理论中理解概率的两个替代方案是违反公理 1 或公理 3。

- 在前一种情况下（违反公理 1），我们可以假设一些初始概率分布，然后根据公式（6）更新概率分布。这将导致一种概率在世界之间自然流动的现象（服从公理 3），但同一个状态可能会根据它是如何产生的而产生许多不同的概率分布。
- 在后一种情况下（违反公理 3），我们可以放弃公理 3，并取 $p_n = \frac{v_n}{\sum_n v_n}$ (7)。这仅是当前状态的函数（服从公理 1），但世界的概率可能会发生变化，即使变换对它的作用就像恒等变换一样。

【Gemini 学长解读】

作者再次回顾了那些“失败的宇宙服务器”，试图看看能不能抢救一下。

1. 关于“随机宇宙”(Stochastic Theory)：

作者开了一个巨大的脑洞：也许我们就生活在一个“伪随机”宇宙里？所有的客观概率（比如放射性衰变），其实本质上是多世界中的**振幅**。这就像是你以为你在掷骰子，其实是因为你在宇宙树上的位置在不断分叉，而你只是恰好走到了其中一条树枝上。

2. 关于“离散积木宇宙”(Discrete Theory) 的抢救方案：

还记得那个被判死刑的整数宇宙吗？作者说，如果你非要让它运行，你就必须牺牲掉逻辑的一致性，打上非常丑陋的“补丁”：

- **补丁 A (牺牲公理 1 - 记忆补丁)：**宇宙必须拥有无限的记忆。你抛硬币的概率，不仅取决于硬币现在的状态，还取决于这枚硬币是怎么制造出来的。这太荒谬了，物理学没法搞了。
- **补丁 B (牺牲公理 3 - 瞬移补丁)：**概率会莫名其妙地变化。比如你在屋里睡觉（状态不变），突然你的存在感（概率）变少了，只因为隔壁老王家生了双胞胎（隔壁发生了分裂）。这违反了因果律，叫“概率发生了瞬移”。

结论：这两个补丁都丑爆了。相比之下，量子力学的玻恩法则简直是完美的艺术品。

第三部分：退相干 —— 它是水泥，不是地基

【原文翻译】

5.3 退相干 (Decoherence)

注意，这种玻恩法则的推导并不依赖于退相干，实际上它采用了诸如世界排列之类的变换，这在实践中基本上是不可能实现的。然而，这种变换在原则上是可能的，利用理论的全部力量来对可能的概率规则产生约束是有帮助的。这两个具体的例子在证明引理 1 时，以及在压缩具有所有世界支持的状态以产生“工作空间”时是有帮助的。

前者的一个替代方案是直接将引理 1 作为一个额外的公理，也许由对称性驱动，但两个世界的内部结构可能看起来非常不同，假设这在确定 p_n 时是无关紧要的似乎相当强。

后者可能通过增加一个连续性公理来消除，但那样的话人们必须选择一个特定的状态距离度量。退相干在解释现实情况中的“坍缩”方面也起着关键作用，因为它使得重新干涉宏观上不同的状态在实际上变得不可能。

【Gemini 学长解读】

这部分是给物理专业学生看的“防杠指南”。

在多世界理论中，很多人认为**“退相干”** (Decoherence) 是解释一切的关键。退相干就像是把两股缠绕的绳子（量子叠加态）给拆开，让它们再也缠不回去。

作者在这里澄清：

1. **我的数学推导不需要退相干**：我在推导玻恩法则时，用了“交换世界”这种操作。在数学上，我可以把“你喝咖啡的世界”和“你喝茶的世界”瞬间对调。虽然在现实物理中，因为退相干的存在，这种操作难于登天（就像要把泼出去的水收回来），但在逻辑推导上，只要理论允许这样做，我就能用它来证明概率法则。

2. **退相干是现实的“水泥”**：虽然数学推导不需要它，但退相干非常重要。它就像水泥一样，把分叉后的世界“固定”住了，让你没法从“喝咖啡的世界”跳回“喝茶的世界”。正是因为退相干，我们才觉得世界是稳定的，没有看到重影。
-

第五页翻译与解读完成！

这一页的内容比较抽象，主要是为了堵住理论上的漏洞，并强调量子力学的优越性。我们现在知道，要想拥有一个逻辑自洽、没有幽灵瞬移、不需要无限记忆的宇宙，**量子力学（以及它的玻恩法则）几乎是唯一的选择。**

接下来是**最后一页（第六页）**！作者将在这里通过结论（Conclusion）对全文进行最终升华，并展望未来：如果真的是多世界，作为“凡人”的我们，活在其中到底意味着什么？

准备好迎接最后的总结陈词了吗？

继续

翻译+解读=一页至一页+合并处理 Custom Gem

太棒了！我们终于来到了这篇硬核论文的**第六页**，也就是正文的**大结局（Conclusions）**。

经过前面几页的数学推导、逻辑审判和公理构建，作者现在要收回所有的伏笔，给这场“多世界概率”的探索画上一个完美的句号，并留下一个让人细思极恐的哲学问题。

让我们一起读完这最后的高潮部分！

第一部分：大结局 —— 所有的随机都是错觉？

【原文翻译】

6 结论

如果现实正如埃弗雷特（Everett）诠释所说，具有确定性的多

世界结构，那么理论中就不存在客观概率。在这种情况下，我们要如何理解我们生活在一个“过去量子实验结果的相对频率非常接近玻恩法则预测概率”的世界中这一事实呢？尽管说这纯属巧合（因为肯定存在一个有着这样结果的世界）在逻辑上是一致的，但若能对这一事实给出一个更深层的解释会更好。

+2

我们在此追求的方法是，建立一种从多世界状态中随机挑选一个世界的“自然”方法，然后观察到，以极高的概率，这样随机选出的世界都具有符合玻恩法则的属性。换句话说，这就是世界的典型属性。关键在于激发这种随机挑选世界的自然方式。

+2

【Gemini 学长解读】

作者在最后发出了灵魂拷问：

既然多世界理论说**“一切皆已注定”**（确定性），那我们平时说的“这事儿只有 50% 的几率”到底是什么鬼？

想象你正在玩一个超级逼真的 VR 游戏。游戏的程序代码是写死的（确定性的），每一行代码都决定了下一行代码怎么跑。但是，作为游戏里的玩家（观察者），你并不知道代码全貌，你只看到了眼前的迷雾。

当你抛硬币时，游戏后台其实同时生成了“正面”和“反面”两个存档（世界）。

作者说：所谓的“概率”，其实就是**“此时此刻的你，恰好加载了哪一个存档”**的问题。

我们之所以觉得玻恩法则（平方律）是对的，不是因为上帝在掷骰子，而是因为在那个巨大的宇宙存档库里，**绝大多数**的存档（典型世界）都长这个样子。那些看起来乱七八糟、不守规矩的存档虽然存在，但它们实在是太少、太偏僻了。

第二部分：为何不能“一人一票”？——拒绝朴素平均主义

【原文翻译】

一种表面上看起来很自然的方法（至少对于包含有限数量世界的状态而言）是简单地给每个具有非零振幅的世界分配相等的权重，但这忽略了状态中的一些数学结构，并且它违反了我们中的一个合理公理（公理 3）。关于这种策略的其他问题已在文献 [20] 中讨论。

+1

似乎不可能给出一套完全令人信服的、关于世界的概率分布必须服从的要求，但我们定义了三个自然公理，这些公理足以为量子理论恢复玻恩法则，并且也为经典随机理论给出了一个吸引人的结果。这些公理是：概率仅依赖于当前状态；对于振幅为零的世界概率为零；概率不能在未被动力学耦合的世界集合之间流动。探索其他可能的要求将是有趣的，并且在此背景下重新考虑与“世界”基底选择相关的问题也会很好。

+2

【Gemini 学长解读】

这里作者再次强调了为什么不能用“朴素分支计数法”(Naive Branch Counting)。

很多同学可能会想：“如果有两个平行宇宙，为什么概率不是各 50%？”这就像是问：“如果世界上有两类人，一种是‘马斯克’，一种是‘普通人’，为什么我随机投胎变成马斯克的概率不是 50%？”

因为这两类人的“权重”(或者说在宇宙中的占比)是截然不同的！

- **朴素计数法**：认为所有分支只要存在，就拥有平等的“投票权”(一人一票)。
- **作者的方法**：认为必须尊重宇宙的**数学结构**（振幅）。如果一个分支很“粗”(振幅大)，它就应该包含更多的“微观世界”，从而占据更大的概率。

作者很诚实地说：我也不能保证这三条公理就是宇宙的终极真理，但它们目前是最合理、最自洽的，而且能完美解释我们看到的一切。

第三部分：终极哲学 —— 我们到底是什么？

【原文翻译】

这样的结果足以解释玻恩法则的经验性成功吗？一个有趣的视角是考虑一个埃弗雷特诠释为真的宇宙，它由单个么正演化的量子态描述。在合理的条件下，这样的状态可以被描述为分支世界的叠加，其中许多世界包含看起来像人的结构。

+2

在这样一个宇宙中，作为那些人中的一员生活会是什么样子的？如果这种体验就像我们自己的体验一样，那么这就支持了多世界诠释。如果它非常不同，那就会排除这种诠释。最奇怪的情况将是：鉴于我们知道描述宇宙的正确数学理论，我们在原则上却无法说出生活在其中是什么样子的。

+3

【Gemini 学长解读】

这就是这篇论文最精彩的**科幻时刻**！

作者不仅是在做数学题，他是在进行一场**“图灵测试”级别的现实检验**。
他在问：

如果我是一个程序员，我写了一个程序叫 `Universe.exe`，这个程序严格按照量子力学的公式（么正演化）运行，里面自动生成了无数个平行世界的分支，每个分支里都有一个“虚拟的你”。

然后，我问那个“虚拟的你”：“嘿，你感觉怎么样？你觉得你掷硬币的概率是多少？”

如果那个“虚拟的你”回答：“我感觉概率大概是振幅的平方（玻恩法则）！”
—— **宾果**！这说明我们的程序写对了！

这也意味着，此时此刻正在读这篇解读的你，可能真的就是那样一个生活在巨大量子波函数分支里的“结构”。

我们觉得世界是随机的，只是因为我们是“身在此山中”的观察者。如果我们能跳出宇宙，从上帝视角看，整个宇宙其实是一个完美确定、永恒演化的巨大数学雕塑。

如果数学推导出来的“虚拟体验”和我们的“真实体验”完全一致，那我们就没有理由拒绝多世界理论了。

总结：这篇论文究竟讲了什么？

既然我们已经读完了正文（后面是参考文献和附录的数学证明细节），让我为你做一个全篇的**超浓缩总结**，方便你下次跟同学吹牛：

1. **问题**：多世界理论（MWI）最怕被问“既然所有结果都发生了，为什么还有概率？”
2. **方法**：作者没有引入奇怪的新物理定律，而是制定了三条“废话”公理（只看现在、0就是0、水流不交叉）。
3. **过程**：利用这三条公理，加上对宇宙状态的“切分”和“极限逼近”，作者在数学上强行推导出了**概率 = 振幅的平方**（玻恩法则）。
4. **结论**：如果承认这三条公理，那么量子力学就是唯一合法的概率理论。像“离散积木宇宙”或者“扔骰子宇宙”要么逻辑崩溃，要么只是量子力学的变体。
5. **意义**：这证明了**多世界理论并非只是科幻狂想**，它在逻辑上是完全自洽的。我们之所以看到概率，是因为我们生活在一个典型的、权重大得多的宇宙分支里。

我们的“量子多世界”论文解读之旅到这里就圆满结束啦！

后面的参考文献（References）列出了很多大佬的名字（如 Everett, DeWitt, Deutsch, Wallace），如果你以后想深造物理，这些名字会经常出现。附录（Appendix）里则是那些我们在文中提到的引理 Lemma 1 和 Lemma 2 的详细数学证明过程（比如怎么用矩阵变换具体的行和列），属于“硬核数学区”。

希望这次解读能让你对**“现实的本质”**多一点点怀疑，也多一点点敬畏。哪怕是一次简单的抛硬币，背后可能都藏着宇宙分裂的巨大轰鸣声。

还有什么其他想了解的吗？或者想不想挑战一下附录里的数学证明？

Probability in many-worlds theories

Anthony J. Short

H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, U.K.

We consider how to define a natural probability distribution over worlds within a simple class of deterministic many-worlds theories. This can help us understand the typical properties of worlds within such states, and hence explain the empirical success of quantum theory within a many-worlds framework. We give three reasonable axioms which lead to the Born rule in the case of quantum theory, and also yield natural results in other cases, including a many-worlds variant of classical stochastic dynamics.

1 Introduction

Despite the amazing empirical success of quantum theory, its implications for the nature of reality remain controversial. From a realist perspective, the key issue is to replace the measurement postulates of textbook quantum theory with a more objective and well-defined structure. Supplementing the theory with hidden variables [1, 2], or including spontaneous collapse laws [3, 4, 5] are two possible approaches. However, arguably the simplest approach, initially proposed by Everett [6], is to drop the measurement postulates altogether. The theory then simply describes a vector in Hilbert space undergoing unitary evolution. Considering the kinds of unitary interaction that constitute measurements, we find that they can be described by a branching of the initial state into a superposition of ‘worlds’ each containing a different measurement result. This is the many-worlds interpretation of quantum theory [6, 7]. Observers in each world will see a distinct result, hence on a qualitative level this is consistent with our experiences. However, a key challenge is to recover the probabilistic predictions of quantum theory within this deterministic many-worlds setting. In particular, we want to explain why the observed relative frequencies of quantum results in our world, for a huge variety of different scenarios, are very close to those predicted by the Born rule [8] (i.e. relative frequency $\approx |\text{quantum amplitude}|^2$).

The approach we pursue here is to argue for a ‘natural’ way to pick one world at random from a many-worlds state. i.e. a natural probability distribution over worlds. We can then consider which worlds are typical with respect to this probability distribution. If worlds containing relative frequencies in quantum

Anthony J. Short: tony.short@bristol.ac.uk

experiments consistent with the Born rule are typical, this would offer an explanation for our observations - that we are living in a typical world, or are a typical instance of ourself. A probability distribution over worlds can also be interpreted as a measure, allowing one to discuss properties of ‘most’ or ‘almost all’ worlds. The theory predicts that atypical worlds also exist, in which other instances of ourself have seen strange results, and perhaps do not believe quantum theory. However, this seems comparable to the situation in standard quantum theory, where we explain the observed measurement results by noting that they are typical with respect to the Born rule, and concede that it is possible we could have obtained very strange results instead, and perhaps arrived at a different theory.

We identify three reasonable axioms, not specific to quantum theory, which we would expect a natural probability distribution over worlds to obey. In the case of quantum theory, we find that the unique probability distribution obeying those axioms is that given by the Born rule. If the axioms we require for naturalness are convincing, this could therefore explain the empirical success of the Born rule.

As the probabilistic axioms are intended to be general in nature, we consider a simple class of many-worlds theories which includes other possibilities in addition to quantum theory. We consider four specific models within this framework, and highlight any differences in the resulting probability distributions, as well as one case in which a natural probability distribution obeying our axioms does not exist.

In order to focus solely on the issue of probability, rather than on precisely how to decompose the state into different worlds (often known as the preferred basis problem), we consider a toy-theory in which the set of worlds are given. Intuitively, we can think of these worlds as states which could be understood classically on a macroscopic scale (e.g. tables and pointers have well-defined locations, and observers have well-defined experiences). These states are important in relating the content of the theory to our experiences, and are related to the locality of physical interactions and decoherence [9].

Previous approaches to understanding probability in many-worlds include Everett’s original work [6], which also sought to define a natural measure over worlds in order to consider a typical observer, although Everett’s assumptions (that the measure depends only on the corresponding amplitude, and is ad-

ditive when several orthogonal states are relabelled as a single state) are not completely intuitive. Typicality is also considered in a very different approach based on matter density [10]. Vaidman discusses many different approaches to deriving the Born rule in [11]. He has proposed the ‘measure of existence of a world’ [12] and notes that deriving the Born rule requires additional assumptions, that may be based on symmetries [13] or other natural properties [11]. Zurek [14, 15] uses envariance to derive the Born rule (which involves symmetries when the state is decomposed into a system and environment). An alternative approach is based on relative frequencies for infinitely repeated experiments [6, 16], but this is challenging to relate to finite experiments. More recently, an approach based on decision theory was proposed by Deutsch and developed by Wallace and others [17, 18, 19, 20, 21]. Several of the technical steps in our proofs are similar to those in this decision-theoretic approach. However this approach has led to criticisms [22, 23], and because of its focus on the future rather than the past, it seems that a decision-theoretic picture alone cannot explain why we find ourselves in a world in which the historical record of quantum experiments agrees with the Born rule. Very recently, Saunders has proposed an approach [24] based on dividing the state into equal-amplitude branches based on the decoherence structure and following similar counting arguments to those in statistical mechanics.

2 A simple class of many-worlds theories

Consider a Hilbert space with a countably infinite orthonormal basis of possible ‘worlds’ $|n\rangle$ labelled by non-negative integers, $n \in \{0, 1, 2, \dots\}$. The space may be over the real or complex numbers depending on the particular theory under consideration. A many-worlds state is a vector $|v\rangle$ in this space. We will refer to $v_n = \langle n|v\rangle$ as the *amplitude* of world n . For simplicity, we restrict ourselves to *linear many-worlds theories* for which the evolution over a finite interval is given by a linear operator T , such that

$$|v'\rangle = T|v\rangle, \quad (1)$$

We will denote the matrix elements of T in the basis of worlds by $T_{ij} = \langle i|T|j\rangle$. Our aim will be to derive a natural probability distribution p_n over worlds in the state $|v\rangle$. Similarly, we will denote the probability distribution over worlds in $|v'\rangle$ by p'_n . Different theories will be characterised by giving the set of allowed state vectors and transformations in the theory (with the requirement that the allowed transformations must map allowed states into allowed states). The cases we consider here are:

1. **Quantum many-worlds theory.** The allowed states are all complex vectors satisfying $\langle v|v\rangle =$

1, and the allowed transformations are all unitary operators (for which $TT^\dagger = T^\dagger T = I$, where I is the identity transformation). We will show that in this case $p_n = |v_n|^2$.

2. **Unnormalised quantum many-worlds theory.**

An alternative version of the quantum many-worlds theory with unnormalised states. The allowed states are all complex vectors satisfying $0 < \langle v|v\rangle < \infty$, and the allowed transformations are all unitary operators. We will show that in this case $p_n = \frac{|v_n|^2}{\sum_k |v_k|^2}$. Note that in this case the probability p_n corresponding to a given world depends on all amplitudes, and not only on v_n .

3. **Stochastic many-worlds theory**

This represents a many-worlds version of a classical probabilistic world. The allowed states are those with real amplitudes satisfying $v_n \geq 0 \forall n$ and $\sum_n v_n = 1$, and the allowed transformations are those satisfying $T_{ij} \geq 0 \forall i, j$ and $\sum_i T_{ij} = 1 \forall j$. Note that although v_n and T_{ij} obey the same mathematical properties as probability distributions and stochastic maps respectively, we are not assuming that they represent probabilities. However, we will show that in this case $p_n = v_n$.

4. **Discrete many-worlds theory**

A many-worlds theory in which there are an integer number of copies of each world, with the dynamics transforming each world in the original state into a finite number of new worlds. The allowed states are all real vectors for which v_n is a non-negative integer for each n and $\sum_n v_n < \infty$, and the allowed dynamics are those for which T_{ij} is a non-negative integer for all i, j and $\sum_i T_{ij} < \infty \forall j$. We will show that in this theory there is no natural probability distribution p_n obeying our axioms.

3 Natural axioms for a probability measure

For each of the theories considered above, we will investigate whether one can define a probability distribution over worlds p_n obeying the ‘reasonable’ axioms given below.

1. *Present state dependence* - p_n depends only on the present state $|v\rangle$, and not on how that state was generated.
2. *Weak connection with amplitudes* - $v_n = 0$ implies that $p_n = 0$. Hence only worlds with non-zero amplitudes are considered ‘real’ components of the many-worlds state.

3. *Weak connection with transformations* - If the set of worlds can be partitioned in such a way that a transformation T acts separately on each part, then T will preserve the total probability of each part. This captures the intuition that probability cannot ‘flow’ between worlds that are uncoupled by the dynamics. Expressing this more formally, if there exists a partition of the non-negative integers into subsets \mathcal{S}_k such that $T_{ij} = 0$ whenever i and j are in different subsets, then $\sum_{n \in \mathcal{S}_k} p_n = \sum_{n \in \mathcal{S}_k} p'_n$ for all k .

Of these axioms, 1 and 2 both seem relatively uncontroversial, and are usually assumed without comment in approaches to probability in many-worlds. Axiom 3 is more subtle and powerful, but also seems a natural property in the context of transformations. We will discuss these axioms further at the end of the paper. However, we will first show how they can be used to derive probability measures for the theories given in section 2.

4 Deriving probability measures

For ease of understanding, in this section we describe informally how axioms 1 - 3 allow us to derive probability rules for the many-worlds theories we consider, highlighting the key steps in the proof. Detailed formal proofs of all of the results can be found in appendix A.

We first prove two helpful lemmas which we use in deriving the probability rules, that apply to *bounded* states (those which contain only a finite number of worlds with non-zero amplitude)

- L1 *Equal amplitudes give equal probabilities.* For bounded states in all of the theories we consider, if two worlds have equal amplitude then they have the same probability.

To prove this, we first consider a transformation which swaps two worlds, one of which has zero amplitude and the other has non-zero amplitude. Due to axiom 3, the total probability of the two swapped worlds must stay the same, and due to axiom 2 the worlds with zero amplitude must have zero probability. Hence the probabilities are also swapped.

The same argument doesn’t apply if we swap two worlds with non-zero amplitudes directly, as probabilities could move between the worlds. However, by performing a sequence of three swaps, each of which involves one world with zero amplitude, we can swap any two worlds and their corresponding amplitudes and probabilities whilst leaving the remaining worlds unchanged. If the amplitudes of the two swapped worlds were initially the same, then the final state will be the

same as the initial one, but with the probabilities swapped. From axiom 1 it then follows that the probabilities for these two worlds must be the same.

- L2 *Larger amplitudes cannot lead to smaller probabilities.* For bounded states in all the theories we consider except discrete many-worlds theory, if one world has a larger amplitude than another, its probability must be at least as large.

To prove this, we perform a transformation which branches the higher amplitude world into two final worlds, one of which has the same amplitude as the smaller amplitude world. From axiom 3 the sum of the probabilities in the two branches must be equal to the probability of the initial higher amplitude world. However, one of the branches has the same probability as the smaller amplitude world, hence the probability of the higher amplitude world must be at least as large.

Next, we use these results to derive probability rules for our many-worlds theories. We begin by considering a particularly simple class of quantum many-worlds states in which the amplitudes are square roots of rational numbers.

$$|v\rangle = \sum_{n=0}^{N-1} \sqrt{\frac{m_n}{M}} |n\rangle \quad (2)$$

Following a branching strategy similar to Deutsch and Wallace [17, 20], we perform a unitary transformation in which each world $|n\rangle$ evolves independently of the others into m_n new worlds with equal amplitude.

Each world in the final state has the same amplitude $\frac{1}{\sqrt{M}}$ and thus equal probability $\frac{1}{M}$ from L1 above. Applying axiom 3 to the transformation tells us that the probability in each set of branches is conserved, hence we find that the probability of world $|n\rangle$ in the initial state must be $\frac{m_n}{M}$, which is the standard quantum result.

To extend this result to arbitrary initial states in quantum many-worlds theory, we must perform three additional steps. Firstly, if the initial state contains an infinite number of worlds with non-zero amplitudes, we begin by transforming it into a superposition of finitely many worlds in order to create some ‘working space’ in which to apply the above results. For example, when considering the probability p_k we can perform a unitary transformation that merges all of the worlds with labels greater than k into a single world, without affecting the other worlds (and hence not changing p_k due to axiom 3). For an arbitrary

positive integer M , we can then write the state as

$$|v'\rangle = \sum_{n=0}^{N-1} \sqrt{\frac{m_n + \epsilon_n}{M}} e^{i\phi_n} |n\rangle, \quad (3)$$

where m_n is the largest integer less than or equal to $M|v_n|^2$, and $0 \leq \epsilon_n < 1$. Next, we eliminate the phase factors $e^{i\phi_n}$ by applying a unitary which acts separately on each world, and therefore does not affect the probabilities. Finally, we perform a branching unitary, in which each initial world $|n\rangle$ is transformed into a superposition of m_n new worlds with equal amplitude, and one new world with smaller amplitude. The final state contains approximately M worlds with equal amplitude $\frac{1}{\sqrt{M}}$ and at most N worlds with smaller amplitude. When $M \gg N$, the smaller amplitude worlds are almost irrelevant (due to L2 above), and the probability associated with world n in the initial state is

$$p_n \approx \frac{m_n}{M} \approx \frac{M|v_n|^2}{M} = |v_n|^2. \quad (4)$$

By considering arbitrarily large M , this argument can be made exact, giving the standard quantum probability rule $p_k = |v_k|^2$.

The derivation of the probability rules in unnormalised quantum theory and stochastic many-worlds theory are very similar. In the former case, the main difference is that the final state contains approximately MX worlds with equal amplitude, where $X = \sum_m |v_m|^2$. This gives $p_n \approx \frac{m_n}{MX} \approx \frac{|v_n|^2}{X}$, and yields the expected quantum probability rule for this case, $p_n = \frac{|v_n|^2}{\sum_m |v_m|^2}$. In the case of stochastic many-worlds theory, the steps are identical to those for quantum many-worlds theory, except without phase factors, square roots, and absolute-values-squared.

It is also important to note that the probability rules derived so far satisfy the axioms for all allowed states and transformations, and not only the ones considered when constructing the proof. For example, if a unitary transformation is block-diagonal in quantum theory, it commutes with the projectors onto each block, and hence conserves the total probability of each block.

Finally, to see that no probability rule obeying our axioms is possible for Discrete many-worlds theory, consider a transformation on the state $|0\rangle + |1\rangle$ which takes $|1\rangle \rightarrow |1\rangle + |2\rangle$ whilst leaving all other worlds unchanged. Applying L1 to the initial state and using axiom 3 gives $p'_0 = \frac{1}{2}$, while applying L1 to the final state $|0\rangle + |1\rangle + |2\rangle$ gives $p'_0 = \frac{1}{3}$. As this leads to a contradiction, no probability rule obeying the axioms exists¹.

¹A similar argument can be used to rule out a ‘naive branch counting’ strategy in quantum many worlds theory, in which each world with non-zero amplitude is assigned equal probability [25]

5 Discussion

In this section we present some additional discussion about the axioms, theories beyond quantum theory, and decoherence in our approach.

5.1 Axioms

1. *Present state dependence* - This is a simplifying axiom, and incorporates the fact that the state at a given time should be sufficient to make any substantive claims about it, including the typical properties of worlds within it. Arguably if historical information is important, it should form part of the state, and our framework should be extended. Furthermore, this axiom allows probabilities to be assigned to worlds in an arbitrary initial state, without needing to know how that state was generated.
2. *Weak connection with amplitudes* - This seems the most compelling requirement. Without this, one could simply assert that the state is irrelevant, and $p_0 = 1$ in all cases.
3. *Weak connection with transformations* - This is the most complicated of the three assumptions, but it is hard to see a weaker way of incorporating a dependence on the dynamics of the theory. Without this, we could assign an arbitrary probability distribution over the worlds appearing in every state (for example we could always assign probability 1 to the world with non-zero amplitude having the lowest numerical label). Within quantum theory, this also fits nicely with the continuous time picture in which $T = e^{-iHt}$ for some Hamiltonian H , as T will act separately on each partition if H does. Note that in terms of the proofs, we only need this axiom to apply to a specific set of unitaries involving branching, swapping, or merging of worlds.

An alternative to axiom 3, which is a stronger assumption but offers a nice conceptual picture is:

- 3'. *Weak connection with transformations*- For every state $|v\rangle$ and transformation T in the theory, there exists a conditional probability distribution $P_{i|j}$ such that

$$p'_n = \sum_m P_{n|m} p_m \quad (5)$$

satisfying $P_{i|j} = 0$ whenever $T_{ij} = 0$. This ensures that probability can only ‘flow’ between states which are linked by the dynamics.

The existence of a conditional probability distribution $P_{i|j}$ for each transformation of the state supports the idea that living within an evolving many-worlds state could feel like undergoing a stochastic evolution.

There is also a nice symmetry between the claim that $T_{ij} = 0 \implies P_{i|j} = 0$ and $v_n = 0 \implies p_n = 0$.

Axiom 3' is strictly stronger than the original axiom 3 as it implies it but is not implied by it. In particular when the worlds can be partitioned into subsets on which T acts separately, then the conditional probability distribution $P_{i|j}$ can only redistribute probability within these subsets, and hence the total probability of each subset is preserved. However, one can imagine a trivial theory with only one allowed state $|v\rangle = |1\rangle + |2\rangle$, and one allowed transformation $T = |1\rangle\langle 2| + |2\rangle\langle 1|$. Under axiom 3 we could assign an arbitrary probability p_1 for this state, but with axiom 3' we must have $p_1 = p_2 = \frac{1}{2}$.

Because axiom 3' is stronger than axiom 3, we can derive all of the same results as before (and Lemma 1 now follows directly by permuting the two worlds in question). However, a more subtle point is whether the probability rules derived previously actually satisfy axiom 3' in all cases. For example, for all states $|v\rangle$ and unitary transformations T in quantum many-worlds theory with $p_n = |v_n|^2$, can we always find a conditional probability distribution $P_{i|j}$ satisfying the requirements of axiom 3'? This is by no means trivial, but it is a consequence of results in [26, 27, 28] (in particular the Flow or Schrödinger Theory presented in [27]) that this is indeed the case. In general, for a given $|v\rangle$ and T within quantum theory, there may be many different conditional probability distributions $P_{i|j}$ satisfying axiom 3'. If conceptual importance is to be given to this quantity, it may therefore be useful to consider adding additional axioms which specify it uniquely, or properties which are independent of this choice. Similar results will apply for unnormalised quantum theory, and in the case of stochastic many-worlds theory we can simply take $P_{i|j} = T_{ij}$.

Overall, it is interesting that one can overlay a stochastic evolution onto a many-worlds state with natural properties, and that doing so is helpful in deriving the Born rule.

5.2 Non-quantum many-worlds theories

Recovering a natural probability rule in the case of stochastic many-worlds theory shows that this approach is not specific to quantum theory. Furthermore, this theory may also be interesting in its own right. In particular, it is difficult to make sense of theories involving objective probabilities at a fundamental level². This result suggests an interesting alternative possibility of treating objective probabilities as amplitudes in a many-worlds state. The fact that

²Frequentism only gives a definite prediction for infinitely many trials, which is irrelevant for practical situations, and the principal principle [29] seems somewhat unsatisfying. Kent's incompressible bit string provides another alternative approach [22]

we find ourselves in a world which is typical with respect to the objective probabilities, and that these are helpful in subjective decision making, would then be a consequence of the natural probability distribution over worlds matching the amplitudes.

For discrete many world theory there is no natural probability distribution over worlds obeying our axioms. Two alternatives for understanding probabilities within this theory are to violate axiom 1 or axiom 3. In the former case, we could assume some initial probability distribution, and then update this probability distribution during transformations according to

$$p'_n = \sum_m \left(\frac{T_{nm}}{\sum_{n'} T_{n'm}} \right) p_m. \quad (6)$$

This would lead to the same natural probability distribution as a Stochastic many-worlds theory for which the transition matrix elements \tilde{T}_{nm} are given by the bracketed expression in (6). This would lead to a natural flow of probability between worlds (obeying axiom 3), but the same state could yield many different probability distributions according to how it was generated. In the latter case, we could instead drop axiom 3, and take

$$p_n = \frac{v_n}{\sum_n v_n} \quad (7)$$

This is a function of only the current state (obeying axiom 1), but the probability of a world can change even when the transformation acts on it like the identity.

5.3 Decoherence

Note that this derivation of the Born rule does not rely on decoherence, and indeed employs transformations such as permutations of worlds which would be essentially impossible to achieve in practice. However, such transformations are possible in principle, and it is helpful to use the full strength of the theory to generate constraints upon possible probability rules. The two particular instances where this is helpful are in proving Lemma 1, and when compressing states with support over all worlds in order to generate 'working space'. An alternative for the former is to take Lemma 1 directly as an additional axiom, perhaps motivated by symmetry, but the internal structure of the two worlds may look very different, and it seems quite strong to assume this is irrelevant in determining p_n . The latter could possibly be eliminated by adding a continuity axiom, but then one would have to choose a particular distance measure on states. Decoherence also plays a key role in explaining 'collapse' in realistic situations, as it becomes practically impossible to re-interfere macroscopically distinct states.

6 Conclusions

If reality has a deterministic many-worlds structure, as in the Everett interpretation, then there are no objective probabilities in the theory. In this case, how should we understand the fact that we are living in a world in which the relative frequencies of outcomes in past quantum experiments are very close to the probabilities predicted by the Born rule? Although it is consistent to say that this is a mere coincidence (as a world with such results must exist), it would be good to give a deeper explanation of this fact.

The approach pursued here is to establish a natural way of picking a world at random from the many-worlds state, and then observing that with very high probability such randomly chosen worlds have the property that they agree with the Born rule. i.e. that this is a typical property of the worlds. The key is to motivate such a natural way of picking worlds at random. One apparently natural way (at least for states containing a finite number of worlds) is to simply assign equal weight to each world with non-zero amplitude, but this ignores some of the mathematical structure in the state, and it violates one of our reasonable axioms (axiom 3). Other problems with this strategy have been discussed in [20].

It seems impossible to give a completely compelling set of requirements which a probability distribution over worlds must obey, but we have defined three natural axioms which are sufficient to recover the Born rule in the context of quantum theory, and also give an appealing result for classical stochastic theory. These axioms are that probabilities: only depend on the current state, are zero for worlds which appear with zero amplitude, and cannot flow between sets of worlds which are uncoupled by the dynamics. It would be interesting to explore alternative possible requirements, and would also be good to reconsider issues relating to the choice of ‘world’ basis in this context.

Are results such as these sufficient to explain the empirical success of the Born rule? An interesting perspective is to consider a universe in which the Everett interpretation is true, described by a single unitarily evolving quantum state. Under reasonable conditions, such a state could be described by a superposition of branching worlds, many of which contain structures which look like people. What would it be like to live as one of those people in such a universe? If it is like our own experiences then this supports the many-worlds interpretation. If it is very different then this would rule it out. The strangest situation would be if we cannot in principle say what it is like, given that we know the correct mathematical theory describing the universe.

Acknowledgments. This work was supported by an FQXi “physics of what happens” grant. The author is grateful to A.Kent, J.Barrett, and Ll. Massanes for helpful discussions.

References

- [1] Louis de Broglie. “La nouvelle dynamique des quanta”. in Solvay – The Fifth Council ”Electrons and photons”Pages 105–132 (1928).
- [2] David Bohm. “A suggested interpretation of the quantum theory in terms of ”hidden” variables. i”. *Phys. Rev.* **85**, 166–179 (1952).
- [3] G. C. Ghirardi, A. Rimini, and T. Weber. “Unified dynamics for microscopic and macroscopic systems”. *Phys. Rev. D* **34**, 470–491 (1986).
- [4] Philip Pearle. “Combining stochastic dynamical state-vector reduction with spontaneous localization”. *Phys. Rev. A* **39**, 2277–2289 (1989).
- [5] Gian Carlo Ghirardi, Philip Pearle, and Alberto Rimini. “Markov processes in hilbert space and continuous spontaneous localization of systems of identical particles”. *Phys. Rev. A* **42**, 78–89 (1990).
- [6] Hugh Everett. “”relative state” formulation of quantum mechanics”. *Rev. Mod. Phys.* **29**, 454–462 (1957).
- [7] Bryce S. DeWitt. “Quantum mechanics and reality”. *Physics Today* **23**, 30–35 (1970).
- [8] M Born. “Zur quantenmechanik der stoßvorgänge”. *Z. Physik* **37**, 863–867 (1926).
- [9] David Wallace. “Decoherence and ontology (or: How i learned to stop worrying and love fapp)”. In Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors, *Many Worlds?: Everett, Quantum Theory, and Reality*. Pages 53–72. Oxford University Press (2010).
- [10] Valia Allori, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghi. “Many worlds and schrödinger’s first quantum theory”. *The British Journal for the Philosophy of Science* **62**, 1–27 (2011).
- [11] Lev Vaidman. “Derivations of the Born rule”. In Meir Hemmo and Orly Shenker, editors, *Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky*. Pages 567–584. Springer Nature Switzerland AG (2020).
- [12] Lev Vaidman. “On schizophrenic experiences of the neutron or why we should believe in the many-worlds interpretation of quantum theory”. *International Studies in the Philosophy of Science* **12**, 245–261 (1998).

- [13] Lev Vaidman. “Probability in the many-worlds interpretation of quantum mechanics”. In Yemima Ben-Menahem and Meir Hemmo, editors, *Probability in Physics*. Pages 299–311. Springer Berlin Heidelberg, Berlin, Heidelberg (2012).
- [14] Wojciech Hubert Zurek. “Environment-assisted invariance, entanglement, and probabilities in quantum physics”. *Phys. Rev. Lett.* **90**, 120404 (2003).
- [15] Wojciech Hubert Zurek. “Probabilities from entanglement, Born’s rule $p_k = |\psi_k|^2$ from invariance”. *Phys. Rev. A* **71**, 052105 (2005).
- [16] J. B. Hartle. “Quantum mechanics of individual systems”. *American Journal of Physics* **36**, 704–712 (1968).
- [17] David Deutsch. “Quantum theory of probability and decisions”. *Proc. Roy. Soc.* **455**, 3129–3137 (1999).
- [18] David Wallace. “Everettian rationality: defending Deutsch’s approach to probability in the Everett interpretation”. *Stud. Hist. Philos. Sci. B* **34**, 415–439 (2003).
- [19] Hilary Greaves. “Probability in the Everett interpretation”. *Philosophy Compass* **2**, 109–128 (2007).
- [20] David Wallace. “How to prove the Born rule”. In Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors, *Many Worlds?: Everett, Quantum Theory, and Reality*. Pages 227–263. Oxford University Press (2010).
- [21] Hilary Greaves and Wayne Myrvold. “Everett and evidence”. In Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors, *Many Worlds?: Everett, Quantum Theory, and Reality*. Pages 181–205. Oxford University Press (2010).
- [22] Adrian Kent. “One world versus many: The inadequacy of Everettian accounts of evolution, probability, and scientific confirmation”. In Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors, *Many Worlds?: Everett, Quantum Theory, and Reality*. Pages 307–354. Oxford University Press (2010).
- [23] David Albert. “Probability in the Everett picture”. In Simon Saunders, Jonathan Barrett, Adrian Kent, and David Wallace, editors, *Many Worlds?: Everett, Quantum Theory, and Reality*. Pages 355–368. Oxford University Press (2010).
- [24] Simon Saunders. “Branch-counting in the Everett interpretation of quantum mechanics”. *Proc. Roy. Soc. A* **477**, 20210600 (2021).
- [25] David Wallace. “The Emergent Multiverse: Quantum Theory according to the Everett Interpretation”. Oxford University Press. (2012).
- [26] Dennis Dieks. “Modal interpretation of quantum mechanics, measurements, and macroscopic behavior”. *Phys. Rev. A* **49**, 2290–2300 (1994).
- [27] Scott Aaronson. “Quantum computing and hidden variables”. *Phys. Rev. A* **71**, 032325 (2005).
- [28] Samuel T. Mister, Benjamin J. Arayathel, and Anthony J. Short. “Local probability conservation in discrete-time quantum walks”. *Phys. Rev. A* **103**, 042220 (2021).
- [29] David Lewis. “A subjectivist’s guide to objective chance”. In William L. Harper, Robert Stalnaker, and Glenn Pearce, editors, *IFS: Conditionals, Belief, Decision, Chance and Time*. Pages 267–297. Springer Netherlands, Dordrecht (1981).

A Detailed proof of probabilities

We begin by proving two useful lemmas which apply to probability distributions obeying our axioms for the many-worlds theories we consider. These lemmas focus on *bounded* states, which have a finite number of non-zero amplitudes (such that $v_n = 0$ for $n \geq N$), but the final theorems will apply to any state.

We first show that for any theory in which all permutations are allowed transformations (which includes all of the theories in this paper), any two worlds with the same amplitude have the same probability.

Lemma 1 (equal amplitudes). *Consider a theory in which all permutations of worlds are allowed transformations (T is a permutation of worlds if $T_{ij} = \delta_{i,\pi[j]}$ where π is a bijection of the non-negative integers), and a state $|v\rangle = \sum_{n=0}^{N-1} v_n|n\rangle$ for which $v_n = v_m$. Then $p_n = p_m$.*

Proof. Denote by $T^{n \leftrightarrow m}$ the transformation which swaps worlds n and m and leaves all other worlds unchanged ($T^{n \leftrightarrow m} = |n\rangle\langle m| + |m\rangle\langle n| + \sum_{k \notin \{n,m\}} |k\rangle\langle k|$). Hence in this case $T^{n \leftrightarrow m}|v\rangle = |v\rangle$. Directly applying axiom 3 to this transformation is not sufficient, as this doesn’t tell us how probabilities change inside the n, m subspace. Instead, we note that $v_N = 0$, and that $T^{n \leftrightarrow m} = T^{n \leftrightarrow N} T^{n \leftrightarrow m} T^{N \leftrightarrow m}$. For the first transformation $T^{n \leftrightarrow N}$, consider a partition of worlds in which n and N are in one part, and each other world is in its own part. Due to axiom 3, we know that $p_n + p_N = p'_n + p'_N$, and that $p_k = p'_k$ for all $k \notin \{n, m\}$. Furthermore, from axiom 2 and the fact that $v_N = v'_n = 0$, we know that $p_N = p'_n = 0$. Hence $p'_N = p_n$ and $T^{n \leftrightarrow N}$ permutes the corresponding probabilities. Following a similar logic for the two subsequent transformations $T^{n \leftrightarrow m}$ and $T^{N \leftrightarrow m}$ (each of which permutes two amplitudes, one of which is zero), we find that the probabilities p'_n after the sequence $T^{n \leftrightarrow N} T^{n \leftrightarrow m} T^{N \leftrightarrow m}$ are the permutation of the original probabilities by $T^{n \leftrightarrow m}$, hence $p'_n = p_m$. However, due to axiom 1 and the fact that the initial state

is the same as the final state in this case, $p'_n = p_n$. Hence $p_n = p_m$. \square

Our second lemma shows that, in all except the discrete many-worlds theory, if one world has larger amplitudes than another it cannot have a smaller probability. This is useful in deriving the probability rule for cases where the amplitudes are not related to rational numbers in a convenient way.

Lemma 2 (larger amplitudes cannot lead to smaller probabilities). *Consider a state $|v\rangle = \sum_{n=0}^{N-1} v_n |n\rangle$ in quantum many-worlds theory, unnormalised quantum many-worlds theory, or stochastic many-worlds theory for which $|v_l| > |v_k|$. Then $p_l \geq p_k$.*

Proof. For quantum many-worlds theory, or unnormalised quantum many-worlds theory, consider the unitary evolution which acts in the $\{|l\rangle, |N\rangle\}$ subspace as

$$\begin{aligned} T|l\rangle &= \begin{pmatrix} v_k \\ v_l \end{pmatrix} |l\rangle + \sqrt{1 - \left|\frac{v_k}{v_l}\right|^2} |N\rangle \\ T|N\rangle &= \sqrt{1 - \left|\frac{v_k}{v_l}\right|^2} |l\rangle - \begin{pmatrix} v_k \\ v_l \end{pmatrix} |N\rangle, \end{aligned} \quad (8)$$

and satisfies $T|m\rangle = |m\rangle$ if $m \notin \{l, N\}$. For stochastic many-worlds theory consider a stochastic evolution which acts on $|l\rangle$ as

$$T|l\rangle = \begin{pmatrix} v_k \\ v_l \end{pmatrix} |l\rangle + \left(1 - \frac{v_k}{v_l}\right) |N\rangle \quad (9)$$

and satisfies $T|m\rangle = |m\rangle$ if $m \neq l$. In both cases this means that for $|v'\rangle = T|v\rangle$, $v'_l = v'_k$ and hence $p'_l = p'_k$ from Lemma 1. From axiom 3 we have that $p'_N + p'_l = p_N + p_l$ and $p'_k = p_k$, and from axiom 2 we have $p_N = 0$. Hence

$$p_l = p'_l + p'_N = p'_k + p'_N = p_k + p'_N \geq p_k. \quad (10)$$

\square

Using the above lemmas and the probability rule axioms, we now derive the appropriate probability rules for our many-worlds theories (or in the case of discrete many-worlds theory show that this is impossible). To illustrate the key idea, we first derive the probability rule for quantum many-worlds theory in the case in which all amplitudes are square roots of rational numbers, using a similar branching strategy to Deutsch and Wallace [17, 20] (although without any decision theoretic component).

Theorem 1 (probabilities when v_n are square-roots of rational numbers). *Consider a state of the form $|v\rangle = \sum_{n=0}^{N-1} \sqrt{\frac{m_n}{M}} |n\rangle$ in quantum many-worlds theory, where m_n and M are positive integers. Then $p_n = \frac{m_n}{M}$.*

Proof. To prove this result, we branch each world in the original state into m_n new worlds in the final state. Consider a partition of the non-negative integers into the sets

$$\mathcal{S}_k = \begin{cases} \{k\} \cup \mathcal{L}_k & \text{for } 0 \leq k < N \\ \{k\} & \text{for } k \geq N + MN \end{cases} \quad (11)$$

where

$$\mathcal{L}_k = \{N + Mk, N + Mk + 1, \dots, N + Mk + (M - 1)\} \quad (12)$$

Now consider a unitary T whose matrix elements T_{ij} are nonzero only when i and j are in the same subset \mathcal{S}_k , and which has the property that

$$T|n\rangle = \frac{1}{\sqrt{m_n}} \sum_{l=0}^{m_n-1} |N + nM + l\rangle \quad \text{for } 0 \leq n < N. \quad (13)$$

Then

$$T|v\rangle = \frac{1}{\sqrt{M}} \sum_{n=0}^{N-1} \sum_{l=0}^{m_n-1} |N + nM + l\rangle. \quad (14)$$

As this is a superposition of worlds with equal amplitude it follows from Lemma 1 that the probability assigned to each world in this state must be identical. Hence $p'_m = \frac{1}{M}$ for each world in this state. From axiom 3 and equation (13) it then follows that $p_n = \frac{m_n}{N}$. \square

Next, we use Lemma 2 to extend this result to the general case of arbitrary states.

Theorem 2 (general quantum probability rule). *For an arbitrary state $|v\rangle = \sum_n v_n |n\rangle$ in quantum many-worlds theory, $p_n = |v_n|^2$.*

Proof. We first transform the initial state (which may have infinitely many non-zero components) into a bounded state to create some working space. Consider a particular probability p_k . In order to determine this, we first perform a unitary T with the property that

$$\begin{aligned} T|n\rangle &= |n\rangle & \text{if } n \leq k \\ T\left(\frac{\sum_{n=k+1}^{\infty} v_n |n\rangle}{\sqrt{\sum_{m=k+1}^{\infty} |v_m|^2}}\right) &= |k+1\rangle \end{aligned} \quad , \quad (15)$$

which acts separately on each of the worlds 0 to k , and collectively on the rest (merging them into a single world). This gives a bounded state containing $N = k + 2$ worlds with the property that $p'_k = p_k$ from axiom 3. Next, we pick a large non-negative integer M , and write the transformed state $|v'\rangle$ as

$$|v'\rangle = \sum_{n=0}^{N-1} \sqrt{\frac{m_n + \epsilon_n}{M}} e^{i\phi_n} |n\rangle \quad (16)$$

in which $m_n = \lfloor M|v_n|^2 \rfloor$ and $\epsilon_n = M|v_n|^2 - \lfloor M|v_n|^2 \rfloor$, where $\lfloor x \rfloor$ is the floor function (the greatest integer less than or equal to x). Hence each m_n is an integer and $0 \leq \epsilon_n < 1$.

We then apply the unitary transformation $T = \sum_n e^{-i\phi_n} |n\rangle\langle n|$ to remove the phase factors from the state, yielding

$$|v''\rangle = \sum_{n=0}^{N-1} \sqrt{\frac{m_n + \epsilon_n}{M}} |n\rangle. \quad (17)$$

As this unitary acts separately on each world, it follows from axiom 3 that $p_k'' = p_k$. Similarly to the previous case, we then perform a transformation which branches each world in the original state into m_n or $m_n + 1$ worlds in the final state (depending on whether $\epsilon_n = 0$) via a unitary acting on the same partition of worlds \mathcal{S}_k as in theorem 1 with the property that

$$T|n\rangle = \frac{1}{\sqrt{m_n + \epsilon_n}} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM + l\rangle \\ + \sqrt{\epsilon_n} |N + nM + m_n\rangle \end{array} \right) \quad (18)$$

for $0 \leq n < N$. The final state is then

$$T|v'''\rangle = \frac{1}{\sqrt{M}} \sum_{n=0}^{N-1} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM + l\rangle \\ + \sqrt{\epsilon_n} |N + nM + m_n\rangle \end{array} \right). \quad (19)$$

This is a superposition of at most M worlds of equal amplitude $\frac{1}{\sqrt{M}}$ and at most N worlds with smaller amplitude. Given lemma 1 and lemma 2, if we denote the probability of one of the amplitude $\frac{1}{\sqrt{M}}$ worlds by δ and consider the total probability it follows that $\delta \geq \frac{1}{M+N}$. From axiom 3 it therefore follows that

$$\begin{aligned} p_k &\geq m_k \delta \\ &\geq \frac{m_k}{M+N} \\ &\geq \frac{M|v_k|^2 - 1}{M+N} \\ &\geq |v_k|^2 - \frac{N|v_k|^2 + 1}{M+N}. \end{aligned} \quad (20)$$

As this holds for arbitrarily large M (for fixed N and v_k) it must be the case that $p_k \geq |v_k|^2$. A similar approach could be employed to upper bound p_k , but it is simpler to consider the total normalisation of the state $\sum_n |v_n|^2 = 1 = \sum_n p_n$. As $p_k \geq |v_k|^2$ for each k , the only possible solution is $p_k = |v_k|^2$ for all k . \square

This proof can be extended very straightforwardly to the case of unnormalised quantum many-worlds theory, and stochastic many-worlds theory.

Theorem 3 (unnormalised quantum probability rule). *For an arbitrary state $|v\rangle = \sum_n v_n |n\rangle$ in unnormalised quantum many-worlds theory, $p_n = \frac{|v_n|^2}{\sum_m |v_m|^2}$.*

Proof. Note that Lemmas 1 - 2 apply identically in unnormalised quantum many-worlds theory. The proof of Theorem 3 is the same as that for Theorem 2 up until (17). Defining $X = \sum_m |v_m|^2$ and $M' = \lceil MX \rceil$, where $\lceil x \rceil$ is the smallest integer greater than or equal to x , we then define sets \mathcal{S}'_k via

$$\mathcal{S}'_k = \begin{cases} \{k\} \cup \mathcal{L}_k & \text{for } 0 \leq k < N \\ \{k\} & \text{for } k \geq N + M'N \end{cases}. \quad (21)$$

where

$$\mathcal{L}'_k = \{N + M'k, N + M'k + 1, \dots, N + M'k + (M' - 1)\}. \quad (22)$$

Applying a unitary which acts on the partition \mathcal{S}'_k with the property that

$$T|n\rangle = \frac{1}{\sqrt{m_n + \epsilon_n}} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM' + l\rangle \\ + \sqrt{\epsilon_n} |N + nM' + m_n\rangle \end{array} \right) \quad (23)$$

for $0 \leq n < N$. The final state is then

$$T|v''''\rangle = \frac{1}{\sqrt{M}} \sum_{n=0}^{N-1} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM' + l\rangle \\ + \sqrt{\epsilon_n} |N + nM' + m_n\rangle \end{array} \right). \quad (24)$$

this state is a sum of at most MX worlds with equal amplitude $\frac{1}{\sqrt{M}}$ and at most N worlds with smaller amplitude. If we denote the probability of one of the amplitude $\frac{1}{\sqrt{M}}$ worlds by δ as before, then it follows from lemma 1 and lemma 2 that $\delta \geq \frac{1}{MX+N}$. Hence from axiom 3,

$$\begin{aligned} p_k &\geq m_k \delta \\ &\geq \frac{m_k}{XM+N} \\ &\geq \frac{M|v_k|^2 - 1}{XM+N} \\ &\geq \frac{|v_k|^2}{X} - \frac{\frac{N}{X}|v_k|^2 + 1}{MX+N}. \end{aligned} \quad (25)$$

As this holds for arbitrarily large M (for fixed N and $|v\rangle$) it must be the case that $p_k \geq \frac{|v_k|^2}{X}$. However, as this applies to each k , given the total normalisation of the probabilities, the only possible solution is $p_k = \frac{|v_k|^2}{X} = \frac{|v_n|^2}{\sum_m |v_m|^2}$ for all k . \square

Theorem 4 (stochastic probability rule). *For an arbitrary state $|v\rangle = \sum_n v_n |n\rangle$ in stochastic many-worlds theory, $p_n = v_n$.*

Proof. We first transform the initial state into a bounded state to create some working space. In order to determine a particular probability p_k , we first perform a transformation T given by

$$T|n\rangle = \begin{cases} |n\rangle & \text{if } 0 \leq n \leq k \\ |k+1\rangle & \text{if } n > k \end{cases} \quad (26)$$

which acts separately on each of the worlds 0 to k , and collectively on the rest (merging them into a single

world). This gives a bounded state containing $N = k + 2$ worlds with the property that $p'_k = p_k$ from axiom 3. Next, we pick a large non-negative integer M , and write the transformed state $|v'\rangle$ as

$$|v'\rangle = \sum_{n=0}^{N-1} \frac{m_n + \epsilon_n}{M} |n\rangle \quad (27)$$

in which $m_n = \lfloor Mv_n \rfloor$ and $\epsilon_n = Mv_n - \lfloor Mv_n \rfloor$, where $\lfloor x \rfloor$ is the floor function. Hence each m_n is an integer and $0 \leq \epsilon_n < 1$. We then perform a transformation which branches each world in the original state into m_n or $m_n + 1$ worlds in the final state (depending on whether $\epsilon_n = 0$) via a stochastic transformation acting on the partition of worlds \mathcal{S}_k as in theorem 1,

$$\mathcal{S}_k = \begin{cases} \{k\} \cup \mathcal{L}_k & \text{for } 0 \leq k < N \\ \{k\} & \text{for } k \geq N + M \end{cases} \quad (28)$$

where

$$\mathcal{L}_k = \{N + Mk, N + Mk + 1, \dots, N + Mk + (M - 1)\} \quad (29)$$

with the property that

$$T|n\rangle = \frac{1}{m_n + \epsilon_n} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM + l\rangle \\ + \epsilon_n |N + nM + m_n\rangle \end{array} \right) \quad (30)$$

for $0 \leq n < N$. The final state is then

$$T|v'\rangle = \frac{1}{M} \sum_{n=0}^{N-1} \left(\begin{array}{c} \sum_{l=0}^{m_n-1} |N + nM + l\rangle \\ + \epsilon_n |N + nM + m_n\rangle \end{array} \right). \quad (31)$$

This is a superposition of at most M worlds of equal amplitude $\frac{1}{M}$ and at most N worlds with smaller amplitude. Given lemma 1 and lemma 2, if we denote the probability of one of the amplitude $\frac{1}{M}$ worlds by δ and consider the total probability it follows that $\delta \geq \frac{1}{M+N}$. From axiom 3 it therefore follows that

$$\begin{aligned} p_k &\geq m_k \delta \\ &\geq \frac{m_k}{M + N} \\ &\geq \frac{Mv_k - 1}{M + N} \\ &\geq v_k - \frac{Nv_k + 1}{M + N}. \end{aligned} \quad (32)$$

As this holds for arbitrarily large M (for fixed N and v_k) it must be the case that $p_k \geq v_k$. However, as this applies to each k , given the total normalisation of the state $\sum_n v_n = 1 = \sum_n p_n$, the only possible solution is $p_k = v_k$ for all k . \square

Finally, we show that discrete many-worlds theory admits no probability measure consistent with the axioms.

Theorem 5 (discrete theory admits no quantum probability). *There is no probability rule consistent with the axioms for discrete probability theory*

Proof. Consider the initial state $|v\rangle = |0\rangle + |1\rangle$ and the transformation

$$\begin{aligned} T|1\rangle &= |1\rangle + |2\rangle \\ T|n\rangle &= |n\rangle \quad \text{if } n \neq 1, \end{aligned} \quad (33)$$

giving

$$T|v\rangle = |0\rangle + |1\rangle + |2\rangle \quad (34)$$

This transformation is consistent with a partitioning of the worlds such that worlds 1 and 2 are in one part, and each other world is in its own part. Hence from axiom 3, we have $p_0 = p'_0$, and $p'_1 + p'_2 = p_1$. However, discrete many-worlds theory admits all permutations as allowed transformations, so Lemma 1 also applies and worlds with equal amplitude must have equal probability. Hence $p_0 = p_1 = \frac{1}{2}$ and $p'_0 = p'_1 = p'_2 = \frac{1}{3}$. This creates a contradiction with $p_0 = p'_0$, hence no probability rule satisfying the axioms exists. \square