




Iterated AGM Revision Based on Probability Revision

Sven Ove Hansson^{1,2} 

Accepted: 26 July 2023 / Published online: 20 August 2023
© The Author(s) 2023

Abstract

Close connections between probability theory and the theory of belief change emerge if the codomain of probability functions is extended from the real-valued interval $[0, 1]$ to a hyperreal interval with the same limits. Full beliefs are identified as propositions with a probability at most infinitesimally smaller than 1. Full beliefs can then be given up, and changes in the set of full beliefs follow a pattern very close to that of AGM revision. In this contribution, iterated revision is investigated. The iterated changes in the set of full beliefs generated by repeated revisions of a hyperreal probability function can, semantically, be modelled with the same basic structure as the sphere models of belief change theory. The changes on the set of full beliefs induced by probability revision satisfy the Darwiche–Pearl postulates for iterated belief change.

Keywords AGM model · Darwiche–Pearl postulates · Hyperreal probabilities · Infinitesimal probabilities · Iterated revision · Jeffrey conditionalization · Probability revision

1 Introduction

One of the most important issues in formal epistemology is how to combine full beliefs and beliefs held to lower degrees in one and the same formal model. In our everyday lives, we have both types of beliefs, and we have no difficulty in shifting our epistemic attitude to a particular proposition between the two types. The full beliefs are those which we currently do not doubt. For instance, when walking in the morning from my home to the nearest metro station, I have no doubt whatsoever which is the shortest way to it. But one morning I encountered large roadwork barriers, which made me uncertain whether I could at all reach the metro via the usual route. Such shifts from

✉ Sven Ove Hansson
soh@kth.se

¹ Department of Philosophy, Uppsala University, Box 627, 751 26 Uppsala, Sweden

² Division of Philosophy, Royal Institute of Technology (KTH), 100 44 Stockholm, Sweden

certainty (full belief) to uncertainty (belief to a lower degree) are common, and we do not conceive them as problematic.

The standard formal representation of beliefs coming in degrees makes use of probability functions. However, although probability-based models provide highly useful accounts of beliefs coming in different degrees, these models are not adapted to also represent full beliefs. The most obvious mode of representation is to assign probability 1 to full beliefs. This has the devastating disadvantage that full beliefs cannot be given up. In standard probability theory, there are no means for reducing the probability of a proposition from 1 to a lower number. Other demarcations of full beliefs, such as identifying them as represented by sentences with a degree of belief above some real-valued limit lower than 1, have other difficulties, as shown in the lottery and preface paradoxes (Kyburg 1961, p. 197; Makinson 1965; Hansson 2018a; Schurz 2019).

This does not mean that we lack formal models of full beliefs. A whole area of formal epistemology, the theory of (dichotomous) belief change, is devoted to the representation of full beliefs (Fermé and Hansson 2018). The problem is that we need different types of models, probability theory and belief change models, to represent the two types of belief. The fact that we combine the two types without problem in our everyday lives, and shift between them without effort, attests to the urgency of finding ways to combine them in one and the same formal model. We need to do this both for theoretical reasons and in order to develop forms of knowledge representation that mirror human reasoning.

Recent work has shown that fairly small changes in standard probability theory are sufficient to allow for a more realistic representation of full beliefs (Hansson 2020; 2022b). The two crucial changes are as follows: First, the codomain of probability functions has to be extended from the real-valued interval $[0, 1]$ to a hyperreal-valued interval with the same limits. Instead of assigning probability 1 to empirical full beliefs, they are then assigned a probability that is infinitesimally smaller than 1. Secondly, revision by a sentence a is performed as a Jeffrey conditionalization that assigns probability $1 - \delta$ to a and probability δ to $\neg a$, for some infinitesimal δ . This makes it possible to give up full beliefs, and it allows the probability function to contain the information needed to construct its successor after revision by the negation of a full belief.

In this article, the hyperreal model of probability revision will be applied to one of the major topics in studies of dichotomous belief change, namely how to perform a series of revisions by different sentences (“iterated revision”). We will investigate the changes in full beliefs that result from a series of revisions of a hyperreal probability function. After some formal preliminaries have been dealt with in Sect. 2, our model for probability revision will be presented in Sect. 3, with an emphasis on the changes in full beliefs that it gives rise to. Section 4 reports our main results, namely strong connections between the hyperreal model of probability revision and standard constructions and axioms from the literature on iterated dichotomous belief change. All formal proofs are deferred to an Appendix.

2 Formal Preliminaries

\mathcal{L} is the object language. Its elements are sentences used to express beliefs. They are represented by lowercase letters (a, b, \dots), and sets of sentences by capital letters (A, B, \dots). The object language is formed from atomic sentences with the usual truth-functional connectives: negation (\neg), conjunction ($\&$), disjunction (\vee), implication (\rightarrow), and equivalence (\leftrightarrow). \top is a tautology and \perp a logically contradictory sentence.

A Tarskian consequence operation Cn expresses the logic. It satisfies the standard conditions: inclusion ($A \subseteq Cn(A)$), monotony (If $A \subseteq B$, then $Cn(A) \subseteq Cn(B)$) and iteration ($Cn(A) = Cn(Cn(A))$). Furthermore, Cn is supraclassical (if a follows from A by classical truth-functional logic, then $a \in Cn(A)$) and compact (if $a \in Cn(A)$, then there is a finite subset A' of A such that $a \in Cn(A')$), and it satisfies the deduction property ($b \in Cn(A \cup \{a\})$ if and only if $a \rightarrow b \in Cn(A)$). $Cn(\emptyset)$ is the set of tautologies. $X \vdash a$ is an alternative notation for $a \in Cn(X)$ and $\vdash a$ for $a \in Cn(\emptyset)$.

A set A of sentences is a (consistent) *belief set* if and only if it is consistent and logically closed, i.e. $A = Cn(A) \neq Cn(\{\perp\})$. K denotes a belief set. The conjunction of all elements of a finite set A of sentences is denoted $\&A$. For all sets A of sentences and all sentences a , the *remainder set* $A \perp a$ is the set of inclusion-maximal subsets of A not implying a . The maximal consistent sets of the language, i.e. the elements of $\mathcal{L} \perp \perp$, will be called *maxisets* (thus avoiding the common but misleading term “possible worlds”). For all sentences a , $|a|$ is the set of a -containing maxisets, i.e. $|a| = \{X \mid a \in X \in \mathcal{L} \perp \perp\}$.

We will use a system of hyperreal numbers for modelling purposes. A hyperreal number system is an extension of the set R of real numbers. It consists of both finite and infinite numbers, but our focus will be on the finite numbers. The finite hyperreal numbers consist of (1) the real numbers and (2) other numbers on an extended number line, which are posited between the real numbers. The positive hyperreal numbers that are larger than 0 but smaller than all positive real numbers are the positive infinitesimals. The negative infinitesimals are the numbers that are smaller than 0 but larger than all negative real numbers. The construction of the hyperreal number system to be used here is presented in Definition 2 and Postulate 1.

The letters s, t, u, v, x, y , and z represent hyperreal numbers (which may be real). The letters δ and ε represent numbers that are either 0 or infinitesimal.

Each finite hyperreal number is infinitely close to exactly one real number, which is called its standard part. The standard part of the number t is denoted $st(t)$. Standard parts satisfy the following rules.

$$\begin{aligned} st(-s) &= -st(s) \\ st(s + t) &= st(s) + st(t) \\ st(s - t) &= st(s) - st(t) \\ st(s \times t) &= st(s) \times st(t) \\ \text{If } st(t) \neq 0 &\text{ then } st(s/t) = st(s)/st(t) \end{aligned}$$

The symbols \approx and \ll are used as follows:

$$\begin{aligned} a \approx b &\text{ if and only if } st(a) = st(b) \\ a \ll b &\text{ if and only if } st(a) < st(b) \end{aligned}$$

The set of hyperreal numbers satisfies the same algebraic laws as the real numbers. Let δ and ϵ be infinitesimals, and let s and t be finite numbers that are not infinitesimals. Then:

$\delta + \epsilon$ and $\delta\epsilon$ are infinitesimals. If $s \neq 0$, then $s\delta$ and δ/s are infinitesimals.

$s + \epsilon$ and st are finite and not infinitesimal. If $s \neq 0 \neq t$, then $s/t \neq 0$.

The following is finite and may or may not be infinitesimal: $s + t$

The following may be infinite or finite, and in the latter case it may or may not be infinitesimal: δ/ϵ .

There is a considerable literature on infinitesimal probabilities. Most of it has employed infinitesimals to attack the problems arising when classical probability theory is applied to infinite domains (event spaces). For instance, a fair lottery with an infinite number of tickets cannot be modelled with real-valued probabilities, but it can be modelled by assigning the same infinitesimal probability to all tickets. In this article, infinitesimals will instead be used for two other purposes, both of which are known from previous literature: (1) we will identify the set of full beliefs with the set of propositions whose probabilities are infinitesimally close to 1, and (2) we will use propositions with infinitesimal probabilities as “memory tracks” of beliefs that have been given up.

For a more extensive introduction to the arithmetic of hyperreals and infinitesimals, readers are referred to Keisler (1986 or 2022). Key references to the literature on infinitesimal probabilities can be found in Hansson (2020, p. 1009).

3 Probability Revision

By a *probability revision* is meant an operation that takes us from a probability function and an input to a new probability function. Standard probability theory contains no specifically defined operation of probability revision, but the role is filled by conditionalization, with the protasis (a in $\mathfrak{p}(d \mid a)$) serving as input. When we conditionalize a probability function \mathfrak{p} by a sentence a such that $\mathfrak{p}(a) \neq 0$, we construct a new probability function \mathfrak{p}' such that for all sentences d : $\mathfrak{p}'(d) = \mathfrak{p}(a \& d)/\mathfrak{p}(a)$.

The standard notation $\mathfrak{p}(\mid a)$ for conditionalization (revision) of \mathfrak{p} by a sentence a does not show clearly that it refers to an operation. This notation is impractical for repeated operations of change, such as revising first by a_1 and then by a_2 . We will therefore use the notation for operations of change that is used in the literature on dichotomous belief change, where the revision of a belief set K by a proposition a_1 is denoted $K * a_1$, its revision by first a_1 and then a_2 is denoted $K * a_1 * a_2$, etc. Generalizing this to other types of inputs, we can define probability revision as follows:

Definition 1 An operation of *probability revision* is an operation \circ such that for some set \mathcal{I} of inputs, for all probability functions \mathfrak{p} and inputs $i \in \mathcal{I}$, $\mathfrak{p} \circ i$ is a probability function.

Thus, $\mathfrak{p} \circ i_1 \cdots \circ i_n$ is the outcome of revising \mathfrak{p} in turn by each of i_1, \dots, i_n in the given order. The input set \mathcal{I} can be a set of sentences, for instance (as in dichotomous

belief change) the whole object language. However, there are other options, such as Jeffrey conditionalization inputs, which are pairs consisting of a sentence and its new probability.

To further clarify the notation, we use boldface brackets around compositely notated probability functions; thus we write $(\mathbf{p} \circ i_1 \circ i_n)(d)$ instead of $\mathbf{p} \circ i_1 \circ i_n(d)$.

As mentioned in the Introduction, all probability functions are assumed to have a hyperreal interval $[0, 1]$ as their codomain.¹ Infinitesimal probabilities represent possibilities that are currently treated as negligible, but kept in memory for possible future use. For instance, let a_1 denote that Shakespeare wrote *Hamlet* and a_2 that Francis Bacon did so, and let $\mathbf{p}(\neg a_1) = \delta$ and $\mathbf{p}(a_2) = \delta/2$. Then we have $\mathbf{p}(\neg a_1 \& a_2)/\mathbf{p}(\neg a_1) = 0.5$, which at least for Bayesians can be a reason to conclude that after revising by $\neg a_1$, the epistemic agent would assign a probability around 0.5 to a_2 . If we instead had $\mathbf{p}(\neg a_1) = 0$, then the probability function would not carry any information useful for constructing a new probability function after revising by $\neg a_1$.

We will have use for infinitesimals of different magnitudes, but we will not need a transfinitely extended hierarchy of infinitesimals. The following definitions and postulate limit the the hierarchies of infinitesimals to what is needed:

Definition 2 (Robinson 1973, pp. 88–89; Hammond 1994, pp. 46–47) Let $\bar{\epsilon}$ be a hyperreal number such that $0 < n\bar{\epsilon} < 1$ for all positive integers n . \mathfrak{F} is the set of fractions of the form

$$\frac{s_0 \times \bar{\epsilon}^0 + s_1 \times \bar{\epsilon}^1 + s_2 \times \bar{\epsilon}^2 + \dots + s_k \times \bar{\epsilon}^k}{t_0 \times \bar{\epsilon}^0 + t_1 \times \bar{\epsilon}^1 + t_2 \times \bar{\epsilon}^2 + \dots + t_n \times \bar{\epsilon}^n}$$

within the closed hyperreal interval $[0, 1]$, such that s_0, \dots, s_k and t_0, \dots, t_n are finite series of real numbers and at least one of t_0, \dots, t_n is non-zero.

Definition 3 (Hansson 2022a) A hyperreal number $y \in \mathfrak{F}$ is an infinitesimal of the first order (in \mathfrak{F}) if and only if $0 \neq y \approx 0$ but there is no $z \in \mathfrak{F}$ such that $0 \neq z \approx 0$ and $y/z \approx 0$.

¹ There is a small literature on how conditional probabilities can be based on standard conditionalization on one-place probability functions with a hyperreal codomain (Skyrms 1980; Lewis 1980; McGee 1994; for an overview see Makinson 2011). This is a way to obtain definability for conditionals with an antecedent that is too implausible to be assigned a positive real-valued probability. However, many authors have avoided this solution. For instance, Spohn (1986, p. 69) mentioned nonstandard analysis as one of four possible solutions to the problem of conditionalizing on propositions with the probability 0, but he chose not to pursue this approach since nonstandard analysis is “a very intricate matter”. Two other solutions of the problem are more popular, namely primitive conditional probabilities (“Popper functions”; Rényi 1955; Popper 1959) and lexicographic models in which there is a sequence of probability functions $\mathbf{p}_1, \mathbf{p}_2 \dots$ such that $\mathbf{p}(a \mid d) = \mathbf{p}_k(a \& d)/\mathbf{p}_k(d)$ for the first \mathbf{p}_k with $\mathbf{p}_k(d) \neq 0$ (Blume et al. 1991a, b; Van Fraassen 1995; Halpern 2010). Variants of the lexicographic approach have been shown to be equivalent with important classes of Popper functions (Hammond 1994; Spohn 1986; Hawthorne 2014). McGee (1994) has shown a similar correspondence with models based on non-standard one-place probabilities. With few exceptions (Gärdenfors 1986; Lindström and Rabinowicz 1989; Makinson 2011) this literature on conditionals does not discuss problems of belief change beyond the simplistic identification of belief change with conditionalization. The requirement of contingency retainment (below, Definition 5), which is essential for iterated revision, is not satisfied in belief revision directly based on the conditionals of the major systems proposed in the literature.

An infinitesimal $y \in \mathfrak{F}$ is an infinitesimal of the n th order, for some $n > 1$, if and only if:

- (1) There is a series z_1, \dots, z_{n-1} of non-zero elements of \mathfrak{F} , such that $z_1 \approx 0$, $z_k/z_{k-1} \approx 0$ whenever $1 < k \leq n - 1$ and $y/z_{n-1} \approx 0$, and
- (2) There is no series z'_1, \dots, z'_n of non-zero elements of \mathfrak{F} , such that $z'_1 \approx 0$, $z'_k/z'_{k-1} \approx 0$ whenever $1 < k \leq n$ and $y/z'_n \approx 0$.

An infinitesimal is *finite-ordered* if and only if it is of the n th order for some positive integer n .

Observation 1 $\bar{\epsilon}$ is a first-order infinitesimal.

Observation 2 If $x \in \mathfrak{F}$ and $0 \neq x \approx 0$, then x is finite-ordered.

Postulate 1 Probability functions have the codomain \mathfrak{F} .

This postulate puts rather strict restrictions on the set of hyperreal numbers that we use in our model. For instance, square roots of first order infinitesimals are left undefined. This may be unsatisfactory from a more general mathematical point of view. However, it is necessary to obtain the hierarchical structure required for our purpose, and it does not hamper the calculus of probabilities.

As a first try, one might use standard Bayesian conditionalization for probability revision of hyperreal probabilities. However, this does not work satisfactorily, since after revision by a sentence a_1 , the probability of $\neg a_1$ will be 0, and it will remain so after any subsequent series of revisions. (Let \star_0 denote Bayesian revision. Then it holds for any hyperreal probability function p and any series of sentences a_1, \dots, a_n with non-zero probabilities that $(p \star_0 a_1 \dots \star_0 a_n)(a_1) = 1$.) To avoid this problem we can use Jeffrey conditionalization, and always leave an infinitesimal probability that the input does not hold. In this way, the information needed to remove the input sentence in some later revision is retained in the new probability function:

Definition 4 (Hansson 2022b) The *hyperreal Bayesian probability revision* in a language \mathcal{L} is the operation \star such that for all probability functions p with \mathcal{L} as domain and all δ with $0 < \delta \approx 0$:

$$(p \star_\delta a)(d) = \begin{cases} p(d) & \text{if } p(a) = 0 \text{ or } p(a) = 1 \\ (1 - \delta) \times \frac{p(a \& d)}{p(a)} + \delta \times \frac{p(\neg a \& d)}{p(\neg a)} & \text{if } 0 \neq p(a) \neq 1 \end{cases}$$

The index notation has been chosen for convenience. In the terminology of Definition 1, \star is an operation with inputs of the form $\langle \delta, a \rangle$, with $0 < \delta \approx 0$ and $a \in \mathcal{L}$. The requirement that $0 < \delta$ has the effect of disallowing the formation of unremovable contingent full beliefs. This limitation can be removed by replacing the condition $0 < \delta \approx 0$ by $0 \leq \delta \approx 0$.

The following definition and observation confirm that the use of Jeffrey conditionalization instead of standard conditionalization has the desired effect.

Definition 5 A probability revision \circ satisfies *contingency retainment* if and only if it holds for all probability functions p , all inputs $i \in \mathcal{I}$, and all sentences d that if $0 \neq p(d) \neq 1$, then $0 \neq (p \circ i)(d) \neq 1$.

This definition makes no assumptions about the formal structure of the set \mathcal{I} of potential inputs. Sentences in the object language may be elements of \mathcal{I} (as in conventional probability theory), they may be components of the inputs (as in Jeffrey conditionalization), or they may be totally unrelated to the set of inputs (Hansson 2017, pp. 45–52).

Observation 3 Hyperreal Bayesian probability revision (\star) satisfies contingency retainment.

The following definition and observation introduce the crucial property of hyperreal probability functions that makes them useful for combining probabilistic and full beliefs in one and the same formal framework, namely that the sentences whose probability is at most infinitesimally smaller than 1 form a logically closed belief set.

Definition 6 (Hansson 2020) Let p be a hyperreal probability function. Then $\llbracket p \rrbracket = \{a \mid \text{st}(p(a)) = 1\}$.

Observation 4 (Hansson 2020) Let p be a hyperreal probability function. Then: $\llbracket p \rrbracket = \text{Cn}(\llbracket p \rrbracket)$ (top closure).

It follows from Observation 4 that our operation \star of probability revision gives rise to a sentential revision (revision of a belief set with sentences as inputs) that only records the effects of \star on the belief sets associated with probability functions. This can be called the “top revision” associated with \star , since it only revises the top-ranked (full) beliefs.²

Definition 7 (Hansson 2022b) Let $*$ be a sentential revision on a belief set K in a language \mathcal{L} . Then $*$ is a *local top revision* on K if and only if there is a probability function p on \mathcal{L} and some δ with $0 < \delta \approx 0$, such that $\llbracket p \rrbracket = K$ and $\llbracket p \star_\delta a \rrbracket = K * a$ for all $a \in \mathcal{L}$.

It was shown in (Hansson 2022b) that for finite languages, the top revisions of local (single-step) hyperreal Bayesian probability revisions almost coincide with AGM revision, which is the standard operation of revision in the belief change literature (Alchourrón et al. 1985; Fermé and Hansson 2018). More precisely, they coincide with AGM revision except in the limiting case of revision by an inconsistent sentence.

Definition 8 (Hansson 2022b) Let $*$ be a sentential operation on a consistent belief set K . Then $*$ is an *AGM^C revision* (consistent AGM revision) on K if and only if there is an AGM operation $*'$ on K such that:

- (a) $K * a = K *' a$ if $a \not\vdash \perp$, and
- (b) $K * a = K$ if $a \vdash \perp$.

² The term “top” for $\{a \mid p(a) = 1\}$ was introduced by Lindström and Rabinowicz (1989).

Theorem 1 (Hansson 2022b) *Let K be a consistent belief set in a finite language \mathcal{L} , and let $*$ be a sentential operation on K . The following two conditions are equivalent:*

- (1) *$*$ is a local top revision on K , based on a (hyperreal) probability function \mathfrak{p} such that $\mathfrak{p}(a) = 1$ if and only if $\vdash a$.*
- (2) *$*$ is an AGM^C revision on K .*

In the limiting case that constitutes the difference between AGM and AGM^C , AGM revision follows the maxim “if instructed to believe in an inconsistency, believe in everything”, whereas AGM^C revision adheres to the arguably more plausible maxim “if instructed to believe in an inconsistency, ignore that instruction”.

4 Iterative Revision

An operation on belief states is *local* if it is only applicable to a specific belief state. It is *iterative*, or with a more common but less adequate term, “iterated”, if it can be applied to all outcomes that it gives rise to. It is *global* if it can be applied to all belief states available within its framework. In dichotomous belief change with belief sets serving as belief states, these definitions can be further specified as follows:

- Definition 9** (1) \circ is a *local sentential operation* on a belief set K in \mathcal{L} if and only if it holds for all sentences $a \in \mathcal{L}$ that $K \circ a$ is a belief set in \mathcal{L} .
- (2) \circ is an *iterative sentential operation* on a belief set K in \mathcal{L} if and only if for all belief sets K in \mathcal{L} and sentences $a_1, a_2 \in \mathcal{L}$: If $K \circ a_1$ is a belief set in \mathcal{L} , then so is $K \circ a_1 \circ a_2$.
- (3) \circ is a *global sentential operation* on belief sets in \mathcal{L} if and only if it holds for all belief sets K in \mathcal{L} and all sentences $a \in \mathcal{L}$ that $K \circ a$ is a belief set in \mathcal{L} .

In probability theory, only global operations of revision appear to have been discussed. For instance, both standard conditionalization and Jeffrey conditionalization—including our variant in Definition 4—are global operations. In contrast, the literature on dichotomous belief change usually treats local operations as the basic case. A standard AGM operation of revision applies only to one specific belief set. It can be extended to iterative and global revision with the help of additional constructions, such as a function that assigns an operation of revision to each belief set. The focus on local operations as the basic case seems to depend largely on the history of this tradition; the operations introduced in the original AGM framework were all local (Alchourrón et al. 1985).³ For overviews of the major approaches to iterative and global dichotomous belief revision, see Rott (2009), Peppas (2014) and Fermé and Hansson (2018, pp. 59–64).

We now turn to our main issue, namely whether the close connection between our operation \star of hyperreal probability revision and local AGM revision of belief sets that

³ This was largely unnecessary, since the original AGM operation, partial meet contraction, can in fact straightforwardly be generalized to a global operation. This is because it holds for all beliefs set K_1 and K_2 and all non-tautologous sentences $p_1 \in K_1$ and $p_2 \in K_2$ that if $K_1 \perp p_1 = K_2 \perp p_2$, then $K_1 = K_2$ and p_1 and p_2 are logically equivalent. See Hansson (2012, p. 152 and p. 160).

was reported in Theorem 1 can be extended to global (and thus also iterative) AGM revision. For that purpose, we are going to use the semantic framework for AGM that has been most used in studies of iterative revision, namely ordered partitionings of the maxisets (“possible worlds”) of the object language. The most common variant of this approach is the spheres model, in which the partitioning is represented by a series of concentric sets (“spheres”), such that the intersection of the elements of the smallest sphere is equal to the current belief set, and the largest sphere contains all the maxisets (Lewis 1973; Grove 1988). For formal developments it is convenient to focus on the rings of sphere systems. A ring consists of those elements of some sphere that are not elements of any of the spheres that are its proper subsets (Hansson 2017, p. 201). This approach was pioneered by Katsuno and Mendelzon (1992), whose system of “faithful assignments” lacks the geometrical connotations of sphere systems. The following definition dispenses both with the geometrical connotations of spheres and the metaphysical connotations of possible worlds:

Definition 10 An ordered maxiset partitioning (OMP) for the finite language \mathcal{L} is a vector $\langle R_0, \dots, R_n \rangle$ of maxisets such that:

- (1) $R_0 \cup \dots \cup R_n = \mathcal{L} \perp \perp$, and
- (2) $R_k \cap R_m = \emptyset$ whenever $0 \leq k < m \leq n$.

The belief set associated with an OMP is the intersection of its top-ranked maxisets:

Definition 11 Let $\langle R_0, \dots, R_n \rangle$ be an ordered maxiset partitioning. Then $\|\langle R_0, \dots, R_n \rangle\| = \bigcap R_0$ is the belief set associated with $\langle R_0, \dots, R_n \rangle$.

A particularly interesting type of operations of revision on OMPs are those that, when revising by a sentence a , move up a -containing maxisets and/or move down $\neg a$ -containing maxisets, but keep the internal order and distance within each of those two groups of maxisets unchanged. In order for the revision to be successful, it must be the case that if a is consistent, then a is an element of all top-positioned maxisets in the new OMP that results from the revision.⁴ This brings us to the following operation of iterative revision for OMPs:

Definition 12 The operation of *translational revision* on ordered maxiset partitionings is the operation \otimes such that for any ordered maxiset partitioning $\langle R_0, \dots, R_n \rangle$, positive integer v and consistent sentence a :

$$\langle R_0, \dots, R_n \rangle \otimes_v a = \langle R'_0, \dots, R'_n \rangle,$$

such that for all $X \in \mathcal{L} \perp \perp$:

- (1) If $a \in X \in R_k$, then $X \in R'_{k-m}$ where m is the lowest number such that $\neg a \notin \bigcap R_m$, and
- (2) If $\neg a \in X \in R_k$, then $X \in R'_{k-\bar{m}+v}$ where \bar{m} is the lowest number such that $a \notin \bigcap R_{\bar{m}}$.

⁴ AGM^C does not satisfy the *success* postulate of AGM ($a \in K * a$), but it satisfies the postulate of *consistent success* (If $\not\vdash \neg a$ then $a \in K * a$).

Just as for \star , the index notation for \otimes was chosen for convenience. Its inputs have the form $\langle v, a \rangle$, where v is a positive integer and a is a consistent sentence. By specifying v in different ways, various operations of revision for OMPs can be constructed that have the consistent sentences in \mathcal{L} as their input set.⁵

Each OMP corresponds to a sphere system $\langle S_0, \dots, S_n \rangle$, such that $S_k = \bigcup \{R_0, \dots, R_k\}$ for all k with $0 \leq k \leq n$. It follows directly from Definition 12 that

$$\|\langle R_0, \dots, R_n \rangle \otimes_v a\| = \bigcap (|a| \cap R_m) = \bigcap (|a| \cap S_m),$$

where R_m is the lowest-numbered element of $\langle R_0, \dots, R_n \rangle$ whose intersection does not contain $\neg a$, and S_m is the lowest-numbered element of $\langle S_0, \dots, S_n \rangle$ whose intersection does not contain $\neg a$. Since $\bigcap (|a| \cap S_m)$ is the outcome of revision by a in a sphere system, it follows that the local restriction (one-step usage) of translational revision coincides with both AGM and AGM^C for consistent inputs.

The following theorem connects our (global) hyperreal probability function \star with (global) translational revision in essentially the same way as Theorem 1 connects the local restriction of \star with the local operation AGM^C:

Theorem 2 *Let \star be the hyperreal Bayesian probability revision and \otimes the translational revision on OMPs, both in the finite language \mathcal{L} .*

- (1) *Let \mathfrak{p} be a (hyperreal) probability function with \mathcal{L} as domain. Then there is an OMP $\langle R_0, \dots, R_m \rangle$ in \mathcal{L} , such that for all consistent sentences a_1, \dots, a_n in \mathcal{L} and finite-order infinitesimals $\delta_1, \dots, \delta_n$ there are positive integers v_1, \dots, v_n with:*

$$\llbracket \mathfrak{p} \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n \rrbracket = \|\langle R_0, \dots, R_m \rangle \star_{v_1} a_1 \dots \star_{v_n} a_n\|$$

- (2) *Let $\langle R_0, \dots, R_m \rangle$ be an OMP in \mathcal{L} . Then there is a (hyperreal) probability function \mathfrak{p} with \mathcal{L} as domain, such that for all consistent sentences a_1, \dots, a_n in \mathcal{L} and positive integers v_1, \dots, v_n there are finite-order infinitesimals $\delta_1, \dots, \delta_n$ with:*

$$\llbracket \mathfrak{p} \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n \rrbracket = \|\langle R_0, \dots, R_m \rangle \star_{v_1} a_1 \dots \star_{v_n} a_n\|$$

The four postulates introduced by Darwiche and Pearl (1997, p. 11) have a central role in the discussion on iterated belief change⁶:

- If $a_1 \vdash a_2$, then $K * a_2 * a_1 = K * a_1$ (DP1)
- If $a_1 \vdash \neg a_2$, then $K * a_2 * a_1 = K * a_1$ (DP2)
- If $K * a_1 \vdash a_2$, then $K * a_2 * a_1 \vdash a_2$ (DP3)
- If $K * a_1 \not\vdash \neg a_2$, then $K * a_2 * a_1 \not\vdash \neg a_2$ (DP4)

The following observation shows that close analogues of these four postulates hold for the top revision obtained from our probability revision \star .

Observation 5 *Let \mathfrak{p} be a probability function, a_1 and a_2 consistent sentences and δ, δ' and ε finite-order infinitesimals. Then:*

- *If $a_1 \vdash a_2$, then $\llbracket \mathfrak{p} \star_{\delta} a_2 \star_{\delta'} a_1 \rrbracket = \llbracket \mathfrak{p} \star_{\varepsilon} a_1 \rrbracket$ (DP1)*

⁵ Revision by inconsistent sentences, left undefined in Definition 12, can be included by a convention, such as setting $\langle R_0, \dots, R_n \rangle \otimes_v a = \langle R_0, \dots, R_n \rangle$ whenever $\vdash \neg a$.

⁶ See Aravanis et al. (2019, pp. 1510–1511) for background information on these postulates.

- If $a_1 \vdash \neg a_2$, then $\llbracket \mathbf{p} \star_{\delta} a_2 \star_{\delta'} a_1 \rrbracket = \llbracket \mathbf{p} \star_{\varepsilon} a_1 \rrbracket$ (DP2)
- If $a_2 \in \llbracket \mathbf{p} \star_{\varepsilon} a_1 \rrbracket$, then $a_2 \in \llbracket \mathbf{p} \star_{\delta} a_2 \star_{\delta'} a_1 \rrbracket$ (DP3)
- If $\neg a_2 \notin \llbracket \mathbf{p} \star_{\varepsilon} a_1 \rrbracket$, then $\neg a_2 \notin \llbracket \mathbf{p} \star_{\delta} a_2 \star_{\delta'} a_1 \rrbracket$ (DP4)

Observation 5 serves to strengthen the connections between probability revision and the dominant approach to (iterative) changes in full beliefs. However, this connection does not prove that the models thus connected with each other adequately reflect the properties of either actual or ideally rational patterns of belief change. The Darwiche–Pearl postulates have the status of a gold standard for iterated revision, and they have been shown to be compatible with requirements of syntax splitting that limit the effects of a revision on beliefs that are irrelevant for the input sentence (Kern-Isberner and Brewka 2017). However, there is also much criticism of these postulates, largely based on counter-examples (Konieczny and Pérez 2000, p. 352; Jin and Thielscher, 2007, p. 6 and p. 14; Stalnaker 2009, pp. 205–206; Hansson 2017, pp. 39–42; Aravanis et al. 2019). The search for plausible models of epistemic change should continue. The results presented here indicate that a stronger focus should be put on the construction of models of iterative change that account for changes both in full beliefs and in beliefs that come in lower degrees.

Acknowledgements Thanks are due to two anonymous referees for unusually valuable comments.

Funding Open access funding provided by Uppsala University. This work was supported by the Swedish Research Council, Grant 2020-01460.

Availability of data and materials Not applicable.

Declarations

Conflicts of interest Not applicable.

Code availability Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: Proofs

Proof of Observation 1 Clearly, $0 \neq \bar{\varepsilon} \approx 0$. It remains to show that there is no $z \in \mathfrak{F}$ with $0 \neq z \approx 0$ and $\bar{\varepsilon}/z \approx 0$. Suppose to the contrary that there is such a z . According to Definition 2, there is some positive real number s and some positive integer k such that $(s \times \bar{\varepsilon}^k)/z \approx 1$, thus $(\bar{\varepsilon}/z) \times s \times \bar{\varepsilon}^{k-1} \approx 1$, contrary to $\bar{\varepsilon}/z \approx 0$. This contradiction concludes the proof. \square

Lemma 1 (Hansson 2022a) *If y and y' are both n^{th} order infinitesimals, then $0 \lll y/y'$.*

Proof of Lemma 1 Suppose that this is not the case. Then $y/y' \approx 0$ and the following holds:

$z_1 \approx 0, z_2/z_1 \approx 0, \dots, y'/z_{n-1} \approx 0, y/y' \approx 0$
 so that y is of at least $(n + 1)^{\text{th}}$ order, contrary to the assumption. □

Lemma 2 Let \hbar be an infinitesimal of the first order. Then t is an infinitesimal of the n^{th} order if and only if there is a real number x such that:

$$\frac{t}{\hbar^n \times x} \approx 1.$$

Proof of Lemma 2 From left to right: Since t is an infinitesimal of the n^{th} order, there are hyperreal numbers z_1, \dots, z_{n-1} such that:

$$\begin{cases} z_1 \approx 0 \\ z_k/z_{k-1} \approx 0 \text{ for } 1 < k \leq n - 1 \\ t/z_{n-1} \approx 0 \end{cases}$$

Suppose that z_1 is an infinitesimal of order m , with $m > 1$. Then there are y_1, \dots, y_{m-1} such that:

$$\begin{cases} y_1 \approx 0 \\ y_k/y_{k-1} \approx 0 \text{ for } 1 < k \leq m - 1 \\ z_1/y_{m-1} \approx 0 \end{cases}$$

so that t is of at least order $n + m - 1$, contrary to the conditions. In the same way it follows that z_k/z_{k-1} is of order 1 whenever $1 < k \leq n - 1$ and that t/z_{n-1} is of order 1. Thus, each of $z_1, z_2/z_1, \dots, z_{n-1}/z_{n-2}, t/z_{n-1}$ is of order 1. It follows from Lemma 1 that there are non-zero real numbers x_1, \dots, x_n such that:

$$\begin{cases} z_1/\hbar \approx x_1 \\ z_k/(z_{k-1} \times \hbar) \approx x_k \text{ whenever } 1 < k \leq n - 1 \\ t/(z_{n-1} \times \hbar) \approx x_n \end{cases}$$

Let $x = x_1 \times \dots \times x_n$. We then have:

$$\begin{aligned} t &= z_1 \times \frac{z_2}{z_1} \times \dots \times \frac{z_{n-1}}{z_{n-2}} \times \frac{t}{z_{n-1}} \\ \frac{t}{\hbar^n} &= \frac{z_1}{\hbar} \times \frac{z_2}{z_1 \times \hbar} \times \dots \times \frac{z_{n-1}}{z_{n-2} \times \hbar} \times \frac{t}{z_{n-1} \times \hbar} \\ &\approx x_1 \times \dots \times x_n = x \end{aligned}$$

Thus, $\frac{t}{\hbar^n \times x} \approx 1$.

From right to left: Let $t/(\hbar^n \times x) \approx 1$. To see that clause (1) of Definition 3 is satisfied, let $z_1 = x \times \hbar$ and $z_k = \hbar^k$ when $1 < k \leq n - 1$. Then $z_1 \approx 0, z_k/z_{k-1} \approx 0$ when $1 < k \leq n - 1$, and $t/z_{n-1} \approx (\hbar^n \times x)/\hbar^{n-1} \approx 0$

To see that clause (2) is satisfied, suppose to the contrary that it is not. Then there are infinitesimals $\delta_1, \dots, \delta_m$ with $n \leq m$, such that $\hbar^{n-1}/(\delta_1 \times \dots \times \delta_m) \approx 1$. It follows that $(\hbar/\delta_1) \times \dots \times (\hbar/\delta_{n-1}) \approx \delta_1 \times \dots \times \delta_m$, thus $\hbar/\delta_k \approx 0$ for at least one δ_k , thus $0 \neq \delta_k \approx 0$ and $\hbar/\delta_k \approx 0$, contrary to the assumption that \hbar is a first-order infinitesimal (cf. Definition 3). □

Proof of Observation 2 Let s_a be the first non-zero coefficient in the numerator of x and t_b the first non-zero coefficient in its denominator. It follows from $0 \approx x$ that $b < a$. Dividing both numerator and denominator by $\bar{\varepsilon}^b$, we obtain:

$$x = \frac{s_a \times \bar{\varepsilon}^{a-b} + s_{a+1} \times \bar{\varepsilon}^{a-b+1} + \dots + s_k \times \bar{\varepsilon}^k}{t_b \times \bar{\varepsilon}^0 + t_{b+1} \times \bar{\varepsilon}^1 + \dots + t_n \times \bar{\varepsilon}^n} \approx \frac{s_a}{t_b} \times \bar{\varepsilon}^{a-b}$$

If $a - b = 1$, then it follows from Observation 1 that x is a first-order infinitesimal. If $a - b > 1$, then it follows from Lemma 2 that x is an infinitesimal of order $a - b$. \square

Proof of Observation 3 Excluding trivial limiting cases, we assume that $0 \neq p(a) \neq 1$. Let $0 \neq p(d) \neq 1$. We have:

$$(p \star_\delta a)(d) = (1 - \delta) \times \frac{p(a \& d)}{p(a)} + \delta \times \frac{p(\neg a \& d)}{p(\neg a)}$$

First suppose that $(p \star_\delta a)(d) = 0$. Since $\delta \neq 0$ we then have $p(a \& d) = p(\neg a \& d) = 0$, thus $p(d) = 0$, contrary to the conditions. Next suppose that $(p \star_\delta a)(d) = 1$. Since $\delta \neq 0$ we then have $p(a \& d) = p(a)$ and $p(\neg a \& d) = p(\neg a)$, thus $p(d) = p(a \& d) + p(\neg a \& d) = p(a) + p(\neg a) = 1$, contrary to the conditions. \square

Proof of Observation 4 See Hansson (2020, pp. 1034–1035). \square

Proof of Theorem 1 See Hansson (2022a). \square

Definition 13 Two hyperreal probability functions p and p' are *order-equivalent* if and only if it holds for all sentences a that either (1) $p(a) = p'(a) = 0$, (2) $0 \ll p(a)$ and $0 \ll p'(a)$, or (3) $p(a)$ and $p'(a)$ are infinitesimals of the same order.

Lemma 3 Let p and p' be order-equivalent, let δ and ε be infinitesimals of the same (finite) order, and let a be a consistent sentence. Then $p \star_\delta a$ and $p' \star_\varepsilon a$ are order-equivalent.

Proof of Lemma 3 Due to Definition 4, the cases when $p(a) = 0$ and $p(a) = 1$ are trivial. In all other cases, the same definition yields:

$$(p \star_\delta a)(d) = (1 - \delta) \times \frac{p(a \& d)}{p(a)} + \delta \times \frac{p(\neg a \& d)}{p(\neg a)}$$

$$(p' \star_\varepsilon a)(d) = (1 - \varepsilon) \times \frac{p'(a \& d)}{p'(a)} + \varepsilon \times \frac{p'(\neg a \& d)}{p'(\neg a)}$$

Step 1: We are going to show that $(1 - \delta)(p(a \& d)/p(a))$ and $(1 - \varepsilon)(p'(a \& d)/p'(a))$ are either both positive non-infinitesimal numbers or infinitesimals of the same order. Clearly, $(1 - \delta)(p(a \& d)/p(a))$ and $p(a \& d)/p(a)$ are either both positive non-infinitesimal numbers or infinitesimals of the same order. It follows from Lemma 1 and the order-equivalence of p and p' that there are positive non-infinitesimal numbers x_1 and x_2 such that $p'(a \& d) = x_1 \times p(a \& d)$ and $p'(a) = x_2 \times p(a)$, thus $p'(a \& d)/p'(a) = (x_1/x_2) \times (p(a \& d)/p(a))$, thus $p'(a \& d)/p'(a)$ and $p(a \& d)/p(a)$ are either both positive non-infinitesimal numbers or infinitesimals of the same order, thus so are $(1 - \delta)(p(a \& d)/p(a))$ and $(1 - \varepsilon)(p'(a \& d)/p'(a))$.

Step 2: We are going to show that $\delta(p(\neg a \& d)/p(\neg a))$ and $\varepsilon(p'(\neg a \& d)/p'(\neg a))$ are either both positive non-infinitesimal numbers or infinitesimals of the same order. Due to Lemma 1, there are positive non-infinitesimal numbers y_1 , and y_2 , and y_3 such that $\varepsilon = y_1 \times \delta$, $p'(\neg a \& d) = y_2 \times p(\neg a \& d)$, and $p'(\neg a) = y_3 \times p(\neg a)$. It follows that $\varepsilon(p'(\neg a \& d)/p'(\neg a)) = (y_1 y_2 / y_3) \times \delta(p(\neg a \& d)/p(\neg a))$, thus $\varepsilon(p'(\neg a \& d)/p'(\neg a))$ is an infinitesimal of the same order as $\delta(p(\neg a \& d)/p(\neg a))$.

Step 3: First: $0 \ll (p \star_\delta a)(d)$ iff $0 \ll (1 - \delta)(p(a \& d)/p(a))$, due to step 1 iff $0 \ll (1 - \varepsilon)(p'(a \& d)/p'(a))$, iff $0 \ll (p' \star_\varepsilon a)(d)$.

Secondly: If $(p \star_\delta a)(d)$ is an infinitesimal, then it has the lowest order that either $(1 - \delta)(p(a \& d)/p(a))$ or $\delta(p(\neg a \& d)/p(\neg a))$ has, which according to steps 1 and 2 is the lowest order that either $(1 - \varepsilon)(p'(a \& d)/p'(a))$ or $\varepsilon(p'(\neg a \& d)/p'(\neg a))$ has, which is the same order that $(p' \star_\varepsilon a)(d)$ has. □

Lemma 4 *Let p be a hyperreal probability function such that $p(a) = 1$ if and only if $\vdash a$. Let a_1, \dots, a_n be consistent non-tautologous sentences. Let $\delta_1, \dots, \delta_n$ be infinitesimals, and such that for all k with $1 \leq k \leq n$, δ_k has the order v_k , with v_k a positive integer. Let \hbar be a first order infinitesimal. Then $(p \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n)$ and $(p \star_{\hbar^{v_1}} a_1 \dots \star_{\hbar^{v_n}} a_n)$ are order-equivalent.*

Proof of Lemma 4 From repeated applications of Lemmas 2 and 3. □

Definition 14 Let p be a probability function and $\langle R_0, \dots, R_n \rangle$ an ordered maxiset partitioning (OMP), both in the finite language \mathcal{L} . Then p and $\langle R_0, \dots, R_n \rangle$ are *order-equivalent* if and only if it holds for all $X \in \mathcal{L} \perp \perp$ that

- (1) $X \in R_0$ if and only if $0 \ll p(\&X)$, and
- (2) If $0 < k$, then $X \in R_k$ if and only if $p(\&X)$ is a k th order infinitesimal.

Lemma 5 *Let \mathcal{L} be a finite language. Then:*

- (1) *For each probability function p on \mathcal{L} that assigns either a positive non-infinitesimal or a finite-order infinitesimal probability to all consistent elements of \mathcal{L} there is an ordered maxiset partitioning $\langle R_0, \dots, R_n \rangle$ such that p and $\langle R_0, \dots, R_n \rangle$ are order-equivalent.*
- (2) *For each ordered maxiset partitioning $\langle R_0, \dots, R_n \rangle$ on \mathcal{L} there is a probability function p on \mathcal{L} such that p and $\langle R_0, \dots, R_n \rangle$ are order-equivalent.*

Proof of Lemma 5 *Part 1:* For all $X \in \mathcal{L} \perp \perp$, let $X \in R_0$ if and only if $0 \ll p(\&X)$, and if $0 < k$, then let $X \in R_k$ if and only if $p(\&X)$ is a k^{th} order infinitesimal.

Part 2: Let \hbar be a first-order infinitesimal. For all $X \in \mathcal{L} \perp \perp$, if $X \in R_k$ and $0 < k$, let

$$p(\&X) = \frac{\hbar^k}{\text{numb}(R_k)},$$

where $\text{numb}(R_k)$ is the number of elements in R_k . If $X \in R_0$, let

$$p(\&X) = \frac{1 - \sum\{p(\&Y) \mid Y \in \mathcal{L} \perp \perp \text{ and } p(\&Y) \approx 0\}}{\text{numb}(R_0)} \quad \square$$

Lemma 6 Let \hbar be a first-order infinitesimal, let v be a positive integer and let a be a logically consistent sentence. Let \mathfrak{p} and $\langle R_0, \dots, R_n \rangle$ be order-equivalent. Then $\mathfrak{p} \star_{\hbar^v} a$ and $\langle R_0, \dots, R_n \rangle \otimes_v a$ are order-equivalent.

Proof of Lemma 6 Let $\langle R_0, \dots, R_n \rangle \otimes_v a = \langle R'_0, \dots, R'_n \rangle$. Consider a maxiset X . There are three cases.

First case, $a \in X \in R'_0$:

$X \in R'_0$ iff $a \in X \in R_k$, where R_k is the lowest-numbered level holding an a -containing maxiset.

Definition 12

iff $a \in X$ and $\mathfrak{p}(\&X)$ and $\mathfrak{p}(a)$ are either both positive non-infinitesimals or same-ordered infinitesimals.

Each level holds a finite number of maxisets

iff $a \in X$ and $0 \ll \mathfrak{p}(\&X)/\mathfrak{p}(a)$

Lemma 1

iff $0 \ll \mathfrak{p}(\&X\&a)/\mathfrak{p}(a)$

$a \in X$

iff $0 \ll (1 - \hbar^v) \times \frac{\mathfrak{p}(\&X\&a)}{\mathfrak{p}(a)} + \hbar^v \times \frac{\mathfrak{p}(\&X\&\neg a)}{\mathfrak{p}(\neg a)}$ Only infinitesimal terms added

iff $0 \ll (\mathfrak{p} \star_{\hbar^v} a)(\&X)$

Definition 4

Second case, $a \in X \notin R'_0$: We need to show that for all $t > 0$: $X \in R'_t$ iff $(\mathfrak{p} \star_{\hbar^v} a)(\&X)$ is a t^{th} order infinitesimal.

Let $a \in X \in R'_t$. Due to Definition 12 there are k and m such that $t = k - m$ and that $X \in R_k$ and that R_m is the lowest-numbered level holding an a -containing maxiset.

It follows that $\mathfrak{p}(\&X)$, i.e. $\mathfrak{p}(\&X\&a)$, is an infinitesimal of the k^{th} order and $\mathfrak{p}(a)$ an infinitesimal of the m^{th} order. Due to Lemma 2, $\frac{\mathfrak{p}(\&X\&a)}{\mathfrak{p}(a)}$ is an infinitesimal of

order $k - m$. Thus so is

$$(1 - \hbar^v) \times \frac{\mathfrak{p}(\&X\&a)}{\mathfrak{p}(a)} + \hbar^v \times \frac{\mathfrak{p}(\&X\&\neg a)}{\mathfrak{p}(\neg a)}$$

to which we have just added the two terms $-\hbar^v \times (\mathfrak{p}(\&X\&a))/(\mathfrak{p}(a))$, which is an infinitesimal of higher order than $k - m$, and $\hbar^v \times (\mathfrak{p}(\&X\&\neg a))/(\mathfrak{p}(\neg a))$, which is equal to 0 since $\mathfrak{p}(\&X\&\neg a) = 0$. Thus according to Definition 4, $(\mathfrak{p} \star_{\hbar^v} a)(\&X)$ is an infinitesimal of order $k - m$, i.e. of order t .

Third case, $\neg a \in X$: Then $X \notin R'_0$. Let $\neg a \in X \in R'_t$. Then there are k and \bar{m} such that $t = k - \bar{m} + v$ and $X \in R_k$ and $R_{\bar{m}}$ is the lowest-numbered level holding a $\neg a$ -containing maxiset. It follows that $\mathfrak{p}(\&X)$ is an infinitesimal of order k and $\mathfrak{p}(\neg a)$ an infinitesimal of order \bar{m} . Thus:

$$(\mathfrak{p} \star_{\hbar^v} a)(\&X) = (1 - \hbar^v) \times \frac{\mathfrak{p}(\&X\&a)}{\mathfrak{p}(a)} + \hbar^v \times \frac{\mathfrak{p}(\&X\&\neg a)}{\mathfrak{p}(\neg a)} \quad \text{Definition 4}$$

$$= \hbar^v \times \frac{\mathfrak{p}(\&X\&\neg a)}{\mathfrak{p}(\neg a)} \quad \text{since } \mathfrak{p}(\&X\&a) = 0 \text{ due to } \neg a \in X$$

$$= \hbar^v \times \frac{\mathfrak{p}(\&X)}{\mathfrak{p}(\neg a)} \quad \text{since } \neg a \in X$$

Since \hbar^v , $\mathfrak{p}(\&X)$, and $\mathfrak{p}(\neg a)$ are infinitesimals or orders v , k and \bar{m} , respectively, it follows that $(\mathfrak{p} \star_{\hbar^v} a)(\&X)$ is an infinitesimal of order $k - \bar{m} + v$, i.e. of order t . \square

Lemma 7 If \mathfrak{p} is order-equivalent with $\langle R_0, \dots, R_n \rangle$, then $\llbracket \mathfrak{p} \rrbracket = \|\langle R_0, \dots, R_n \rangle\|$.

Proof of Lemma 7 $d \in \llbracket \mathfrak{p} \rrbracket$

iff $\mathfrak{p}(d) \approx 1$

iff $\sum \{\mathfrak{p}(\&X) \mid d \in X \in \mathfrak{L} \perp \perp\} \approx 1$

iff $\sum \{\mathfrak{p}(\&X) \mid d \in X \in \mathfrak{L} \perp \perp \text{ and } 0 \ll \mathfrak{p}(\&X)\} \approx 1$ A finite number of infinitesimal terms excluded

iff $\sum \{\mathfrak{p}(\&X) \mid d \in X \in R_0\} \approx 1$ Definition 14

iff $d \in X$ for all $X \in R_0$ $0 \ll \mathfrak{p}(\&X)$ for all $X \in R_0$

iff $d \in \bigcap R_0$

iff $d \in \|\langle R_0, \dots, R_n \rangle\|$ Definition 1 □

Proof of Theorem 2 Part 1: Let \mathfrak{p} be a probability function. It follows from Lemma 5, part 1, that there is some OMP $\langle R_0, \dots, R_m \rangle$ such that \mathfrak{p} and $\langle R_0, \dots, R_m \rangle$ are order-equivalent. Let \hbar be a first order infinitesimal, and for each δ_k , let v_k be the order of infinitesimals that δ_k belongs to. It follows from Lemma 4 that $\mathfrak{p} \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n$ and $\mathfrak{p} \star_{\hbar^{v_1}} a_1 \dots \star_{\hbar^{v_n}} a_n$ are order-equivalent. It follows by repeated application of Lemma 6 that $\mathfrak{p} \star_{\hbar^{v_1}} a_1 \dots \star_{\hbar^{v_n}} a_n$ and $\langle R_0, \dots, R_m \rangle \otimes_{v_1} a_1 \dots \otimes_{v_n} a_n$ are order-equivalent. Thus $\mathfrak{p} \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n$ and $\langle R_0, \dots, R_m \rangle \otimes_{v_1} a_1 \dots \otimes_{v_n} a_n$ are order-equivalent, thus according to Lemma 7, $\llbracket \mathfrak{p} \star_{\delta_1} a_1 \dots \star_{\delta_n} a_n \rrbracket = \|\langle R_0, \dots, R_m \rangle \otimes_{v_1} a_1 \dots \otimes_{v_n} a_n\|$.

Part 2: Let $\langle R_0, \dots, R_m \rangle$ be an OMP. It follows from Lemma 5, part 2, that there is some probability function \mathfrak{p} such that \mathfrak{p} and $\langle R_0, \dots, R_m \rangle$ are order-equivalent.

It follows by repeated application of Lemma 6 that $\mathfrak{p} \star_{\hbar^{v_1}} a_1 \dots \star_{\hbar^{v_n}} a_n$ and $\langle R_0, \dots, R_m \rangle \otimes_{v_1} a_1 \dots \otimes_{v_n} a_n$ are order-equivalent, thus according to Lemma 7, $\llbracket \mathfrak{p} \star_{\hbar^{v_1}} a_1 \dots \star_{\hbar^{v_n}} a_n \rrbracket = \|\langle R_0, \dots, R_m \rangle \otimes_{v_1} a_1 \dots \otimes_{v_n} a_n\|$. □

Proof of Observation 5 Due to Theorem 2 we can equivalently prove the following properties of translation revisions on ordered maxiset partitionings, for consistent sentences a_1 and a_2 and positive integers v, v_1 , and v_2 :

If $a_1 \vdash a_2$, then $\|\langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2 \otimes_{v_1} a_1\| = \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$ (DP1)

If $a_1 \vdash \neg a_2$, then $\|\langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2 \otimes_{v_1} a_1\| = \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$ (DP2)

If $a_2 \in \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$, then $a_2 \in \|\langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2 \otimes_{v_1} a_1\|$ (DP3)

If $\neg a_2 \notin \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$, then $\neg a_2 \notin \|\langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2 \otimes_{v_1} a_1\|$ (DP4)

The following abbreviations will be used:

- $\langle R'_0, \dots, R'_n \rangle = \langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2$
- $\langle R''_0, \dots, R''_n \rangle = \langle R_0, \dots, R_n \rangle \otimes_{v_2} a_2 \otimes_{v_1} a_1$
- $\langle R^\dagger_0, \dots, R^\dagger_n \rangle = \langle R_0, \dots, R_n \rangle \otimes_v a_1$

Part 1, DP1: Let R_k be the lowest-numbered a_2 -containing element in $\langle R_0, \dots, R_n \rangle$ and R_m its lowest-numbered a_1 -containing element.⁷ It follows from $a_1 \vdash a_2$ that $k \leq m$. It holds for all maxisets X that contain a_2 (and thus for all that contain a_1) that if $X \in R_s$, then $X \in R'_{s-k}$. Since R_m is the lowest-numbered a_1 -containing element in $\langle R_0, \dots, R_n \rangle, R'_{m-k}$ is the lowest-numbered a_1 -containing element of $\langle R'_0, \dots, R'_n \rangle$.

All a_1 -containing maxisets X are moved leftwards from $\langle R'_0, \dots, R'_n \rangle$ to $\langle R''_0, \dots, R''_n \rangle$, so that the leftmost of them is moved from R'_{m-k} to R''_0 . This means that for all a_1 -containing X : If $X \in R'_t$, then $X \in R''_{t-(m-k)}$. We have $t - (m - k) = (s - k) - (m - k) = s - m$. Thus, if $a_1 \in X \in R_s$ then $X \in R''_{s-m}$.

⁷ R_k is a -containing if and only if $\neg a \notin \bigcap R_k$.

Since R''_0 consists only of a_1 -containing maxisets, it follows that $X \in R''_0$ if and only if $a_1 \in X \in R_m$.

Next, note that for all maxisets X , if $a_1 \in X \in R_s$, then $X \in R^{\dagger}_{s-m}$. Since R^{\dagger}_0 consists only of a_1 -containing maxisets, it follows that $X \in R^{\dagger}_0$ if and only if $a_1 \in X \in R_m$. Thus:

$$\|\langle R''_0, \dots, R''_n \rangle\| = \bigcap R''_0 = \bigcap (R_m \cap |a_1|) = \bigcap R^{\dagger}_0 = \|\langle R^{\dagger}_0, \dots, R^{\dagger}_n \rangle\|.$$

Part 2, DP2: Let R_k be the lowest-numbered a_1 -containing element in $\langle R_0, \dots, R_n \rangle$, and let R'_m be the lowest-numbered a_1 -containing element in $\langle R'_0, \dots, R'_n \rangle$. Since all a_1 -containing maxisets contain $\neg a_2$, they are moved in parallel in the transition from $\langle R_0, \dots, R_n \rangle$ to $\langle R'_0, \dots, R'_n \rangle$, thus $|a_1| \cap R_k = |a_1| \cap R'_m = R''_0$. It follows directly that $R^{\dagger}_0 = |a_1| \cap R_k$, thus $R''_0 = R^{\dagger}_0$, thus $\|\langle R''_0, \dots, R''_n \rangle\| = \|\langle R^{\dagger}_0, \dots, R^{\dagger}_n \rangle\|$.

Part 3, DP3: Let R_m be the lowest-numbered a_1 -containing element in $\langle R_0, \dots, R_n \rangle$. It follows from $a_2 \in \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$ and $\|\langle R^{\dagger}_0, \dots, R^{\dagger}_n \rangle\| = \bigcap R^{\dagger}_0 = \bigcap (R_m \cap |a_1|)$ that $a_2 \in \bigcap (R_m \cap |a_1|)$.

Let X be one of the leftmost maxisets in $\langle R_0, \dots, R_n \rangle$ such that $a_1 \& a_2 \in X$ and let X' be one of the leftmost maxisets in $\langle R_0, \dots, R_n \rangle$ such that $a_1 \& \neg a_2 \in X'$. It follows from $a_2 \in \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$ that X is in a position to the left of that of X' in $\langle R_0, \dots, R_n \rangle$. Thus $R_m \cap |a_1 \& a_2| = R_m \cap |a_1|$. In the operation that takes us from $\langle R_0, \dots, R_n \rangle$ to $\langle R'_0, \dots, R'_n \rangle$, X' will be moved to the right, whereas X will either not be moved or moved to the left. Thus X is positioned to the left of X' in $\langle R'_0, \dots, R'_n \rangle$. Thus $X \in R''_0$ but $X' \notin R''_0$. Consequently $a_2 \in \bigcap R''_0$, i.e. $a_2 \in \|\langle R''_0, \dots, R''_n \rangle\|$.

Part 4, DP4: Let R_m be the lowest-numbered a_1 -containing element in $\langle R_0, \dots, R_n \rangle$.

Since $R^{\dagger}_0 = R_m \cap |a_1|$, it follows from $\neg a_2 \notin \|\langle R_0, \dots, R_n \rangle \otimes_v a_1\|$ that there is some X' such that $a_2 \in X' \in R_m \cap |a_1|$. In other words, X' is one of the lowest-numbered a_1 -containing maxisets in $\langle R_0, \dots, R_n \rangle$. It follows from $a_2 \in X'$ that X' is one of the leftmost maxisets containing a_1 in $\langle R'_0, \dots, R'_n \rangle$ as well. Therefore X' is an element of R''_0 . It follows from $a_2 \in X' \in R''_0$ that $\neg a_2 \notin R''_0$, i.e. $\neg a_2 \notin \|\langle R''_0, \dots, R''_n \rangle\|$. \square

References

- Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50, 510–530.
- Aravanis, T. I., Peppas, P., & Williams, M. A. (2019). Observations on Darwiche and Pearl’s approach for iterated belief revision. In Kraus, S. (ed.) *Proceedings of the twenty-eighth international joint conference on artificial intelligence (IJCAI-19)* (pp. 1509–1515). <https://www.ijcai.org/proceedings/2019/209>.
- Blume, L., Brandenburger, A., & Dekel, E. (1991). Lexicographic probabilities and choice under uncertainty. *Econometrica: Journal of the Econometric Society*, 59, 61–79.
- Blume, L., Brandenburger, A., & Dekel, E. (1991). Lexicographic probabilities and equilibrium refinements. *Econometrica: Journal of the Econometric Society*, 59, 81–98.
- Darwiche, Q., & Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence*, 89, 1–29.
- Fermé, E., & Hansson, S. O. (2018). *Belief change: Introduction and overview*. Berlin: Springer.
- Gärdenfors, P. (1986). The dynamics of belief: Contractions and revisions of probability functions. *Topoi*, 5, 29–37.
- Grove, A. (1988). Two modellings for theory change. *Journal of Philosophical Logic*, 17, 157–170.
- Halpern, J. (2010). Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, 68(1), 155–179.

- Hammond, P. J. (1994). Elementary non-Archimedean representations of probability for decision theory and games. In P. Humphreys (Ed.), *Patrick Suppes: Scientific Philosopher* (Vol. 1, pp. 25–61). Dordrecht: Kluwer.
- Hansson, S. O. (2012). Global and iterated contraction and revision: An exploration of uniform and semi-uniform approaches. *Journal of Philosophical Logic*, 41, 143–172.
- Hansson, S. O. (2017). *Descriptor revision. Belief change through direct choice*. Berlin: Springer.
- Hansson, S. O. (2018). Review of Hannes Leitgeb, The stability of belief. How Rational belief coheres with probability. *Journal of Philosophy*, 115(5), 276–280.
- Hansson, S. O. (2020). Revising probabilities and full beliefs. *Journal of Philosophical Logic*, 49(5), 1005–1039.
- Hansson, S. O. (2022a). A characterization of probability-based dichotomous belief revision. *Studia Logica*, 110, 511–543.
- Hansson, S. O. (2022b). A basis for AGM revision in Bayesian probability revision. *Journal of Philosophical Logic* (in press).
- Hawthorne, J. (2014). A primer on rational consequence relations, Popper functions, and their ranked structures. *Studia Logica*, 102(4), 731–749.
- Jin, Y., & Thielscher, M. (2007). Iterated belief revision, revised. *Artificial Intelligence*, 171(1), 1–18.
- Katsuno, H., & Mendelzon, A. O. (1992). On the difference between updating a knowledge base and revising it. In Gärdenfors, P. (Ed.), *In* (pp. 183–203). Cambridge: Cambridge University Press.
- Keisler, H. J. (1986). *Elementary calculus: An infinitesimal approach* (2nd ed.). Boston: Prindle, Weber & Schmidt.
- Keisler, H. J. (2022). *Foundations of infinitesimal calculus*. Madison: University of Wisconsin. <https://people.math.wisc.edu/~keisler/foundations.pdf>.
- Kern-Isberner, G., & Brewka, G. (2017). Strong syntax splitting for iterated belief revision. In: *Proceedings of the twenty-sixth international joint conference on artificial intelligence (IJCAI-17)* (pp. 1131–1137).
- Konieczny, S., & Pérez, R. P. (2000). A framework for iterated revision. *Journal of Applied Non-Classical Logics*, 10(3–4), 339–367.
- Kyburg, H. E. (1961). *Probability and the logic of rational belief*. Middletown, CT: Wesleyan University Press.
- Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.
- Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. II, pp. 263–293). Berkeley: University of California Press.
- Lindström, S., & Rabinowicz, W. (1989). On probabilistic representation of non-probabilistic revision. *Journal of Philosophical Logic*, 18, 69–101.
- Makinson, D. (1965). The paradox of the preface. *Analysis*, 25(6), 205–207.
- Makinson, D. (2011). Conditional probability in the light of qualitative belief change. *Journal of Philosophical Logic*, 40, 121–153.
- McGee, V. (1994). Learning the impossible. In E. Eells & B. Skyrms (Eds.), *Probability and conditionals: Belief revision and rational decisions* (pp. 179–199). Cambridge: Cambridge University Press.
- Peppas, P. (2014). A panorama of iterated revision. In S. O. Hansson (Ed.), *David Makinson on classical methods for non-classical problems* (pp. 71–94). Berlin: Springer.
- Popper, K. (1959). *The logic of scientific discovery* (2nd ed.). New York: Basic Books.
- Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiae Hungaricae*, 6, 268–335.
- Robinson, A. (1973). Function theory on some nonarchimedean fields. *American Mathematical Monthly*, 80(6,2), 87–109.
- Rott, H. (2009). Shifting priorities: Simple representations for twenty-seven iterated theory change operators. In D. Makinson, J. Malinowski, & H. Wansing (Eds.), *Towards Mathematical Philosophy* (pp. 269–296). Springer.
- Schurz, G. (2019). Impossibility results for rational belief. *Noûs*, 53(1), 134–159.
- Skyrms, B. (1980). *Causal necessity: A pragmatic investigation of the necessity of laws*. New Haven: Yale University Press.
- Spohn, W. (1986). The representation of Popper measures. *Topoi*, 5(1), 69–74.
- Stalnaker, R. (2009). Iterated belief revision. *Erkenntnis*, 70(2), 189–209.
- Van Fraassen, B. C. (1995). Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic*, 24(4), 349–377.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.