Diachronic Coherence and Radical Probabilism

Brian Skyrms†

The question of *diachronic coherence*, coherence of degrees of belief across time, is investigated within the context of Richard Jeffrey's radical probabilism. Diachronic coherence is taken as fundamental, and coherence results for degrees of belief at a single time, such as additivity, are recovered only with additional assumptions. Additivity of probabilities of probabilities is seen to be less problematic than additivity of first-order probabilities. Without any assumed model of belief change, diachronic coherence applied to higher-order degrees of belief yields the martingale property.

1. Introduction. Richard Jeffrey advocated a flexible theory of personal probability that is open to all sorts of learning situations. He opposed what he saw as the use of a conditioning model as an epistemological straitjacket in the work of Clarence Irving Lewis (1946). Lewis's dictum "No probability without certainty" was based on the idea that probabilities must be updated by conditioning on the evidence. Jeffrey's *probability kinematics*—now also known as Jeffrey conditioning—provided an alternative (see Jeffrey 1957, 1965, 1968).

It was not meant to be the only alternative. Jeffrey articulated a philosophy of *radical probabilism* that held the door open to modeling all sorts of epistemological situations. In this spirit, I will look at diachronic coherence from a point of view that embodies minimal epistemological assumptions and then add constraints little by little.

2. Arbitrage. There is a close connection between Bayesian coherence arguments and the theory of arbitrage (see Shin 1992).

Suppose we have a market in which a finite number[1] of assets are bought

†To contact the author, please write to: Department of Logic and Philosophy of Science, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697–5100; e-mail: bskyrms@uci.edu.

1. We keep things finite at this point because we want to focus on diachronic coherence and avoid the issues associated with the philosophy of the integral.

Philosophy of Science, 73 (December 2006) pp. 959–968. 0031-8248/2006/7305-0042$10.00

and sold. Assets can be anything: stocks and bonds, pigs and chickens, apples and oranges. The market determines a unit price for each asset, and this information is encoded in a price vector $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$. You may trade these assets today in any (finite) quantity. You are allowed to take a short position in an asset; that is to say, you sell it today for delivery tomorrow. Tomorrow, the assets may have different prices, $\mathbf{y}_1, \ldots, \mathbf{y}_m$. To keep things simple, we initially suppose that there are a finite number of possibilities for tomorrow's price vector. A *portfolio*, **p**, is a vector of real numbers that specifies the amount of each asset you hold. Negative numbers correspond to short positions. You would like to *arbitrage the market*, that is, to construct a portfolio today whose cost is negative (you can take out money) and such that tomorrow its value is nonnegative (you are left with no net loss), no matter which of the possible price vectors is realized.

According to the *fundamental theorem of asset pricing*, you can arbitrage the market if and only if the price vector today falls outside the convex cone spanned by the possible price vectors tomorrow.[2]

There is a short proof that is geometrically transparent. The value of a portfolio, **p**, according to a price vector, **y**, is the sum over the assets of quantity times price: the dot product of the two vectors. If the vectors are orthogonal, the value is zero. If they make an acute angle, the value is positive; if they make an obtuse angle, the value is negative. An arbitrage portfolio, **p**, is one such that $\mathbf{p} \cdot \mathbf{x}$ is negative and $\mathbf{p} \cdot \mathbf{y}_i$ is nonnegative for each possible $\mathbf{y}_i$; **p** makes an obtuse angle with today's price vector and is orthogonal or makes an acute angle with each of the possible price vectors tomorrow. If **p** is outside the convex cone spanned by the $\mathbf{y}_i$'s, then there is a hyperplane that separates **p** from that cone. An arbitrage portfolio can be found as a vector normal to the hyperplane. It has zero value according to a price vector on the hyperplane, a negative value according to today's prices, and a nonnegative value according to each possible price tomorrow. On the other hand, if today's price vector is in the convex cone spanned by tomorrow's possible price vectors, then (by Farkas's lemma) no arbitrage portfolio is possible.

Suppose, for example, that the market deals in only two goods, apples and oranges. One possible price vector tomorrow is \$1 for an apple, \$1 for an orange. Another is that an apple will cost \$2, while an orange is \$1. These two possibilities generate a convex cone, as shown in Figure 1a. (We could add lots of intermediate possibilities, but that wouldn't make any difference to what follows.) Let's suppose that today's price vector lies outside the convex cone, say apples at \$1, oranges at \$3. Then

2. If we were to allow an infinite number of states tomorrow, we would have to substitute the *closed* convex cone generated by the possible future price vectors.

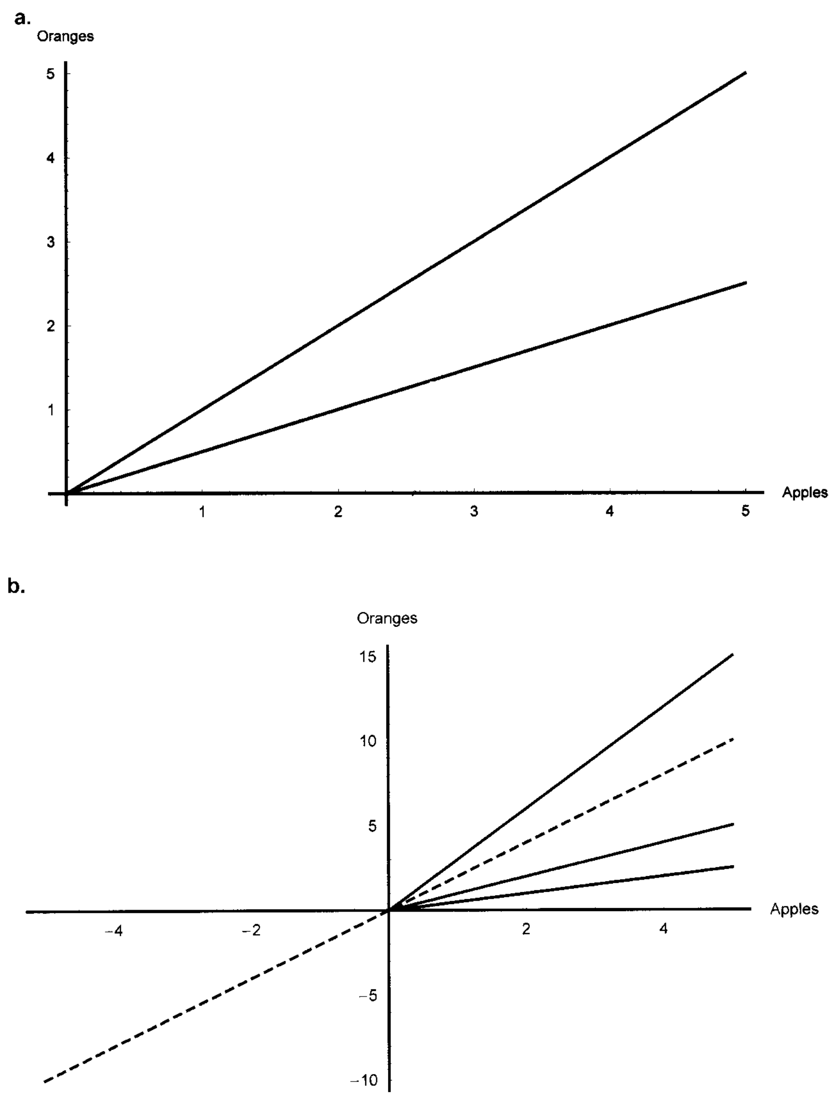
a.
Oranges
5
4
3
2
1
1
2
3
4
5
Apples
b.
Oranges
15
10
5
−4
−2
2
4
Apples
−5
−10

Figure 1.

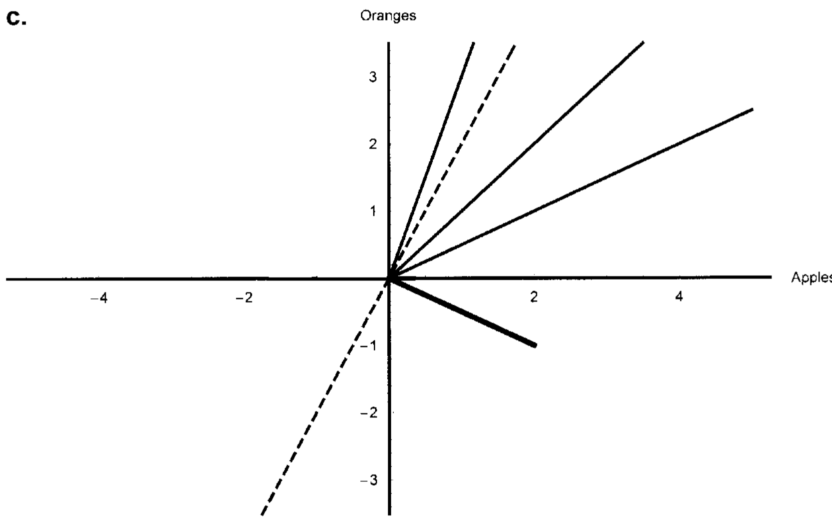

Figure 1. (Continued)

it can be separated from the cone by a hyperplane (in two dimensions, a line), for example, the line oranges = 2 apples, as shown in Figure 1b. Normal to that hyperplane we find the vector ⟨2 apples, −1 orange⟩, as in Figure 1c. This should be an arbitrage portfolio, so we sell one orange short and use the proceeds to buy two apples. But at today's prices, an orange is worth \$3; so we can pocket a dollar, or—if you prefer—buy three apples and eat one.

Tomorrow we have to deliver an orange. If tomorrow's prices were to fall exactly on the hyperplane, we would be covered. We could sell our two apples and use the proceeds to buy the orange. But in our example, things are even better. The worst that can happen tomorrow is that apples and oranges trade one-to-one, so we might as well eat another apple and use the remaining one to cover our obligation for an orange.

The whole business is straightforward: sell dear, buy cheap. Notice that at this point there is no probability at all in the picture.

3. Degrees of Belief. In the foregoing, assets could be anything. As a special case they could be tickets paying \$1 if p, nothing otherwise, for various propositions, p. The price of such a ticket can be thought of as the market's collective *degree of belief* or *subjective probability* for p. We have not said anything about the market except that it will trade arbitrary quantities at the market price. The market might or might not be imple-

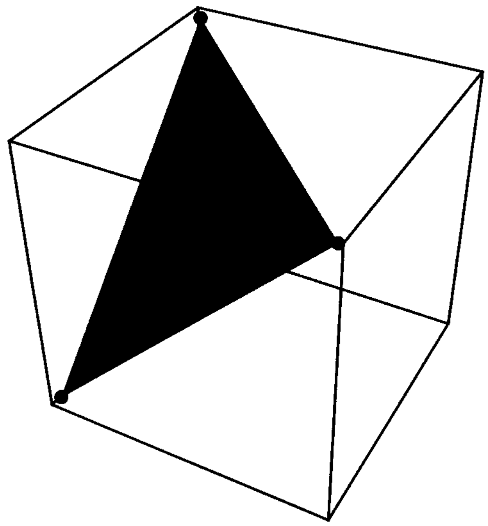

Figure 2.

mented by a single individual—the bookie of the familiar Bayesian metaphor.

Without yet any commitment to the nature of the propositions involved, the mathematical structure of degrees of belief, or the characteristics of belief revision, we can say that arbitrage-free degrees of belief today must fall within the convex cone spanned by the degrees of belief tomorrow. This is the fundamental diachronic coherence requirement. Convexity is the key to everything that follows.

4. Probability. Suppose, in addition, that the propositions involved are true or false and that tomorrow we learn the truth. We can also assume that we can neglect discounting the future. A guarantee of getting \$1 tomorrow is as good as getting \$1 today. Then tomorrow a ticket worth \$1 if p, nothing otherwise, would be worth either \$1 or \$0 depending on whether we learn whether p is true or not.

And suppose that we have three assets being traded that have a logical structure. There are tickets worth \$1 if p, nothing otherwise; \$1 if q, nothing otherwise; and \$1 if p or q, nothing otherwise. Furthermore, p and q are incompatible. This additional structure constrains the possible price vectors tomorrow, so that the convex cone becomes the two-dimensional object $z = x + y$ (x, y nonnegative), as shown in Figure 2.

Arbitrage-free degrees of belief must be additive. *Additivity of subjective probability* comes from the *additivity of truth value* and the fact that *ad-*

ditivity is preserved under convex combination. One can then complete the coherence argument for probability by noting that coherence requires a ticket that pays \$1 if a tautology is true to have the value \$1.

Notice that from this point of view, the synchronic Dutch books are really special cases of diachronic arguments. You need the moment of truth for the synchronic argument to be complete. The assumption that there is such a time is a much stronger assumption than anything that preceded it in this development.

An intuitionist, for example, may have a conception of proposition and of the development of knowledge that does not guarantee the existence of such a time, even in principle. Within such a framework, coherent degrees of belief need not obey the classical laws of the probability calculus.

5. Probabilities of Probabilities. Today the market trades tickets that pay \$1 if p_i, nothing otherwise, where the p_i's are some "first-order" propositions. All sorts of news come in, and tomorrow the price vector may realize a number of different possibilities. (We have not, at this point, imposed any model of belief change.) The price vector for these tickets tomorrow is itself a fact about the world, and there is no reason why we could not trade in tickets that pay off \$1 if tomorrow's price vector is $\mathbf{p}$ or if tomorrow's price vector is in some set of possible price vectors, for the original set of propositions. The prices of these tickets represent subjective probabilities today about subjective probabilities tomorrow.

Some philosophers have been suspicious about such entities, but they arise quite naturally. And in fact, they may be less problematic than the first-order probabilities over which they are defined. The first-order propositions, p_i, could be such that their truth value might or might not ever be settled. But the question of tomorrow's price vector for unit wagers over them is settled tomorrow. Coherent probabilities of tomorrow's probabilities should be additive, no matter what.

6. Diachronic Coherence Revisited. Let us restrict ourselves to the case in which we eventually do find out the truth about everything and all bets are settled (perhaps on Judgment Day), so degrees of belief today and tomorrow are genuine probabilities. We can now consider tickets that are worth \$1 if the probability tomorrow of $p = a$ and p, nothing otherwise, as well as tickets that are worth \$1 if the probability tomorrow of $p = a$.

These tickets are logically related. Projecting to the two dimensions that represent these tickets, we find that there are only two possible price vectors tomorrow. Either the probability tomorrow of p is not equal to a, in which case both tickets are worth nothing tomorrow, or the prob-

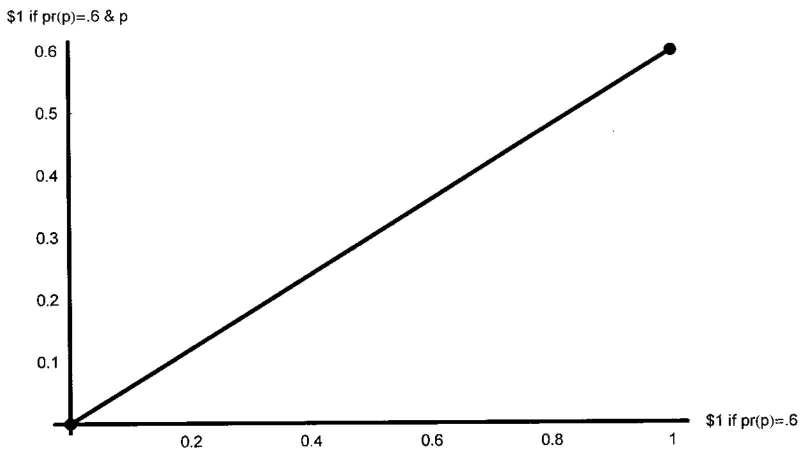

Figure 3.

ability tomorrow of p is equal to a, in which case the former ticket has a price of $\$a$ and the latter has a price of \$1. The cone spanned by these two vectors is just a ray as shown in Figure 3. So today, the ratio of these two probabilities (provided that they are well defined) is a. In other words, today the conditional probability of p, given that the probability tomorrow of $p = a$, is a. It then follows that to avoid a Dutch book, the probability today must be the expectation of the probability tomorrow (see Goldstein 1983; van Fraassen 1984). Since convexity came on stage, it has been apparent that this expectation principle has been waiting in the wings. The introduction of probabilities of probabilities allows it to be made explicit.

7. Coherence and Conditioning. In accord with Jeffrey's philosophy of radical probabilism, we have imposed no restrictive model of belief change. A conditioning situation is allowed, but not required. That is to say, there may be first-order propositions, $e_1, \ldots, e_n$, that map one-to-one to possible degrees of belief tomorrow, $\mathbf{q}_1, \ldots, \mathbf{q}_n$, such that, for our degrees of belief today, $\mathbf{p}$, and for all propositions under consideration, s, $\mathbf{q}_i(s) = \mathbf{p}(s \text{ given } e_i)$, in which case we have a conditioning model. But there need not be such propositions, which is the case that radical probabilism urges us not to ignore. In this case, convexity still provides an applicable test of diachronic coherence.

On the other hand, with the introduction of second-order probabilities, coherence *requires* belief change by conditioning, that is to say, conditioning on propositions about what probabilities will be tomorrow (see

Skyrms 1980; Good 1981). These are, of course, quite different from the first-order sense-data propositions that C. I. Lewis had in mind.

8. Probability Kinematics. Where does Richard Jeffrey's probability kinematics fit into this picture? Belief change by kinematics on some partition is not sufficient for diachronic coherence. The possible probability vectors tomorrow may have the same probabilities conditional on p and on its negation as today's probability without today's probability of p being the expectation of tomorrow's. Diachronic coherence constrains probability kinematics.

In a finite setting, belief change is always by probability kinematics on *some* partition, the partition whose members are the atoms of the space. But, as Jeffrey always emphasized, coherent belief change need not consist of probability kinematics on some nontrivial partition. That conclusion follows only from stronger assumptions that relate the partition in question to the learning situation.

Suppose that between today and tomorrow we have a learning experience that changes the probability of p, but not to zero or one. And suppose that then, by the day after tomorrow, we learn the truth about p. We can express the *assumption* that we have gotten information only *about* p on the way through tomorrow to the day after tomorrow by saying that we move from now to then by conditioning on p or on its negation. This is the assumption of *sufficiency* of the partition $\{p, -p\}$. Then one possible probability tomorrow has the probability of p as one and the probability of p & q as equal to today's $\Pr(q \text{ given } p)$ and the other possible probability tomorrow has the probability of p as zero and the probability of $-p$ & q as equal to today's $\Pr(q \text{ given } -p)$. This is shown in Figure 4 for $\Pr(q \text{ given } p) = .9$ and $\Pr(q \text{ given } -p) = .2$. Diachronic coherence requires that tomorrow's probabilities must fall on the line connecting the two points representing possible probabilities the day after tomorrow and thus must come from today's probabilities by kinematics on $\{p, -p\}$. This is the basic diachronic coherence argument that in my earlier work (Skyrms 1987, 1990) was cloaked in concerns about infinite spaces.

As Jeffrey always emphasized, without the assumption of sufficiency of the partition, there is no coherence argument. But equally, if there is no assumption that we learn just the truth of p, there is no argument for conditioning on p in the case of certain evidence.

9. Tomorrow and Tomorrow and Tomorrow. Consider not two or three days, but an infinite succession of days. Assume that degrees of belief are all probabilities (e.g., Judgment Day comes at time $\omega + 1$). The probability of p tomorrow, the probability of p the day after tomorrow, the probability

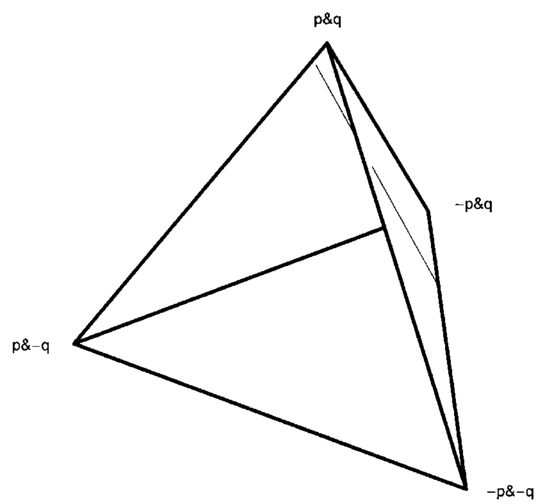

Figure 4.

of p the next day, and so forth are a sequence of random variables. Diachronic coherence requires that they form a *martingale* (see Skyrms 1996; Zabell 2002). Richer cases lead to vector-valued martingales.

10. Diachronic Coherence Generalized. Looking beyond the scope of this paper, suppose that you throw a point dart at a unit interval, and the market can trade in tickets that pay $1 if it falls in a certain subset, $0 otherwise. This is something of an idealization to say the least, and the question arises as to how coherence might be applied. One natural idea might be to idealize the betting situation so as to allow a countable number of bets, in which case coherence requires countable additivity, a restriction of contracts to measurable sets, and, in general, the orthodox approach to probability. Then the martingale convergence theorem applies: Coherence entails convergence.

An approach more faithful to the philosophy of de Finetti would allow a finite number of bets at each time. This leads to *strategic measure*, a notion weaker than countable additivity but stronger than simple finite additivity (see Lane and Sudderth 1984, 1985). Orthodox martingale theory uses countable additivity, but there is a finitely additive martingale theory built on strategic measures (see Purves and Sudderth 1976). A version of the "coherence entails convergence" result can be recovered, even on this more conservative approach (see Zabell 2002).

REFERENCES

Goldstein, Michael (1983), "The Prevision of a Prevision," *Journal of the American Statistical Association* 78: 817–819.

Good, Irving John (1981), "The Weight of Evidence Provided from an Uncertain Testimony or from an Uncertain Event," *Journal of Statistical Computation and Simulation* 13: 56–60.

Jeffrey, Richard (1957), *Contributions to the Theory of Inductive Probability*. PhD dissertation, Princeton University.

——— (1965), *The Logic of Decision*. New York: McGraw-Hill; 3rd rev. ed. Chicago: University of Chicago Press, 1983.

——— (1968), "Probable Knowledge," in I. Lakatos (ed.), *The Problem of Inductive Logic*. Amsterdam: North-Holland.

Lane, David A., and William Sudderth (1984), "Coherent Predictive Inference," *Sankhya*, ser. A, 46: 166–185.

——— (1985), "Coherent Predictions Are Strategic," *Annals of Statistics* 13: 1244–1248.

Lewis, Clarence Irving (1946), *An Analysis of Knowledge and Valuation*. LaSalle, IL: Open Court.

Purves, Roger, and William Sudderth (1976), "Some Finitely Additive Probability," *Annals of Probability* 4: 259–276.

Shin, Hyun Song (1992), "Review of *The Dynamics of Rational Deliberation*," *Economics and Philosophy* 8: 176–183.

Skyrms, Brian (1980), "Higher Order Degrees of Belief," in D. H. Mellor (ed.), *Prospects for Pragmatism*. Cambridge: Cambridge University Press, 109–138.

——— (1987), "Dynamic Coherence and Probability Kinematics," *Philosophy of Science* 54: 1–20.

——— (1990), *The Dynamics of Rational Deliberation*. Cambridge, MA: Harvard University Press.

——— (1996), "The Structure of Radical Probabilism," *Erkenntnis* 45: 285–297.

van Fraassen, Bas (1984), "Belief and the Will," *Journal of Philosophy* 81: 235–256.

Zabell, Sandy (2002), "It All Adds Up: The Dynamic Coherence of Radical Probabilism," *Philosophy of Science* 69 (Proceedings): S98–S103.