

SUBJECTIVE EXPECTED UTILITY THEORY

by
R. Duncan Luce
The Institute for Advanced Study
Princeton, N.J.

This paper was prepared for delivery at the 7.^o Colóquio Brasileiro de Matemática. Poços
Caldas, Brasil. July 6-26, 1969

1. The Problem

The goal of this class of theories is to formulate what it means to make a “rational” choice from among uncertain alternatives and to find which, if any, numerical maximization is equivalent to it. The task has two major conceptual aspects: we must say exactly what comprises the domain of uncertain alternatives and we must describe what we mean by a rational choice in that domain.

Roughly, an “uncertain alternative” is a gamble (ou lottery) in which chance determines the exact consequence to the decision maker. Somewhat more specifically, if A_1, \dots, A_n are pairwise disjoint events and c_1, \dots, c_n are entities that can be delivered to a person and among which he has preferences, then an uncertain alternative of the form $\langle A_1, c_1; \dots; A_n, c_n \rangle$ is interpreted to mean that you receive c_i if A_i occurs. (I am being purposely non-committal about the extent of $\bigcup_{i=1}^n A_i$.)

The simplest rationality postulate is the transitivity of preference. That it is rational to be transitive can be argued as follows: If you prefer alternative f to g , then you should be willing to pay me a sum of money, however small, to guarantee f rather than g . Similarly, if you prefer g to h you will pay to get g rather than h ; and if you are intransitive, you will pay to get h rather than f . And so I can relieve you of all your assets without ever giving you anything in return. A somewhat more subtle condition of rational behavior can be formulated as follows. Suppose that c and c' are two consequences and that you prefer c to c' , then provided that A has a non-negligible chance of occurring you should prefer $\langle A, c; \bar{A}, d \rangle$ to $\langle A, c'; \bar{A}, d \rangle$. The argument is that if A occurs you will receive a more preferred consequence from the former gamble, and if A does not occur you receive the same consequence from either gamble. This is known, for obvious reasons, as the “sure-thing principle.”

The sort of numerical representation that seems to arise naturally is this: There exists a probability measure P over the events and a numerical (utility) function v over the consequences with the property that the expectation of the random variable induced by v over the uncertain alternatives is order preserving. More explicitly, $\langle A_1, c_1; \dots; A_n, c_n \rangle$ is preferred or indifferent (written \succsim) to $\langle B_1, d_1; \dots; B_m, d_m \rangle$ iff

$$\sum_{L=1}^n v(c_L) P(A_L) \geq \sum_{L=1}^m v(d_L) P(B_L).$$

This representation surely suggests the phrase “expected utility theory” of the title. The adjective “subjective” is appended to emphasize that the probability measure P may very well be unique to the decision maker and that, in particular, it need not agree with a relative frequency definition of probability, when that exists.

2. Some History

In practice, little disagreement exists over the conditions of rationality, even though most expositions of expected utility theory place the main emphasis on these postulates. What really distinguishes the existing theories is the exact definition of the collection of uncertain alternatives. As an indirect consequence, proofs of the representation theorem differ considerably. I will attempt to make clear how the several theories characterize the class of alternatives, but, with one exception, I will not state their exact assumptions. For the latest and most general theory, the assumptions and representation theorem are stated fully and a brief outline of the proof is supplied.

The first modern and mathematically complete discussion of an expected utility theory appeared in an appendix to the 1947 edition of von Neumann and Morgenstern’s classic *Theory of games and economic behavior*. If, for simplicity, we denote the two-component gamble $\langle A, c; \bar{A}, d \rangle$ by cAd , then they assumed that a probability measure P is given and that cAd is indifferent (written \sim) to cBd when $P(A) = P(B)$. Therefore, the symbol for a gamble can be altered to cPd , meaning that you receive consequence c with probability P and d with probability $1 - P$. In addition to such objects, they included those where c and d are replaced by simple two-component gambles: $(cQd)P(c'Q'd')$. Their axioms imply that

$$(cQd)Pd \sim cQPd,$$

which makes sense only if the events underlying P and Q are independent. Finally, although my discussion may have suggested that the uncertain

alternatives are defined over a fixed sample space, this is not the only possible interpretation of cPd . If $P = m/n$, then the gamble cPd can be realized by spinning a fair roulette wheel that is divided into n equal sectors, m of which are assigned to c and the remaining $n - m$ to d . Alternatively, one can have $n + 1$ different coins with the probabilities $0/n, 1/n, \dots, m/n, \dots, n/n$ of heads. In this case, $c(m/n)d$ is realized by selecting and tossing a particular coin. The gamble is conditional on that coin; whereas, with the roulette wheel all gambles are conditional on the same sample space. Without taking much note of the choice being made, experimentalists usually assume an analogue of the coin interpretation and statisticians almost always assume the formal analogue of the roulette wheel. As we shall see, this makes a difference.

I will discuss the other theories in terms of their attempts to cope with one or more of the difficulties of the von Neumann-Morgenstern theory: (1) the assumption that gambles can be described in terms of a given probability measure rather than just in terms of events; (2) the restriction to simple (= two-component) gambles and simple compounds of them; (3) the assumption that there are many independent events; and (4) the uncertainty about the assumed domain of the gambles.

Provided that nothing else is changed, it is nearly trivial to generalize the theory to gambles with an arbitrary, but finite, number of components. Examples can be found in Blackwell & Girshick (1954), Luce & Raiffa (1957), and Samuelson (1952).

All other generalizations involve dropping the assumption that a probability measure is given. The first, Ramsey (1931), clearly pre-dates von Neumann and Morgenstern, but it was virtually without influence at the time and it was not rendered fully rigorous until the papers by Suppes & Winet (1955) and Davidson & Suppes (1956). In this theory there are only two-component gambles and comparisons are made on the same partition, e.g., aAb versus cAd . For this reason, there is no reason to be clear about whether all gambles have a common sample space or not. The key idea is to postulate the existence of an event E^* with the property that for all a and b , $aE^*b \sim bE^*a$. It is easy to see that if an expected utility representation exists, $P(E^*) = P(\bar{E}^*)$. Thus, again assuming that the representation exists, $aE^*b \succsim cE^*d$ iff $v(a) + v(b) \geq v(c) + v(d)$. Axioms were stated that yield this additive representation over consequences, and then further assumptions were introduced which permit the construction of a probability measure P from v and from the orderings of aAb versus cAd . This theory is seriously flawed not only by the restriction to two-component gambles, but by the fact that aAb is not compared with cBd .

The first major attempt to get rid of all the difficulties is in Savage’s

1954 book *The foundations of statistics*. He took as his primitives a sample space \mathcal{S} (which he called the set of states of nature), a set \mathcal{C} of consequences, and a set \mathcal{D} of possible decisions of the form $f : \mathcal{S} \rightarrow \mathcal{C}$. (Uncertain alternatives of this generality I shall call “decisions”; Savage called them “acts.”) He restricted his representation theorem to those decisions that have a finite range – gambles. Between any two decisions, he assumed a judgment of preference or indifference. With this relation \succsim given, he stated axioms which permitted him to define a natural ordering of preference over the consequences and another of “qualitatively more probable than” over the events. The one over \mathcal{C} arises from the ordering of the constant decisions [$f(s) = c$ for all $s \in \mathcal{S}$], and the one over the events arises from the ordering of gambles of the form cAd versus cBd , where c is preferred to d . Additional postulates were made about the probability ordering so that, following de Finetti (1937), an orderpreserving probability measure could be constructed. Savage was then in a position comparable to the starting point of von Neumann and Morgenstern’s theory, and the remainder of his proof parallels theirs.

Savage’s system has two main drawbacks. First, it is awkward to state the axioms directly in terms of the primitives because a number of intermediate definitions and proofs are needed. Second, the simultaneous assumptions that decisions are defined over the whole sample space and that all the constant decisions exist are, from a realistic point of view, virtually contradictory. Although one’s first impulse is to try to restrict the domain of decisions to exclude the constant ones, two facts ultimately lead one to consider altering the definition of a decision. First, Savage’s proof is crucially dependent upon the existence of constant acts, and so if they are dropped a wholly different approach must be taken. Second, our natural formulation of decisions is never over the whole sample space. For example, when you consider traveling from Rio to São Paulo and the decision is between flying and driving, you evaluate each alternative conditional on that mode of travel. The travel by plane involves the restrictive event of you in the airplane, and the drive involves the restrictive event of your driving your car from Rio to São Paulo (which you surely will not do if you fly there). You – the decision maker – determine, by your choice, which of these two events will occur; this conditional nature of decisions is, in my opinion, one of their most basic characteristics.

The first attempt to take into account the conditional nature of decisions is Pfanzagl (1957, 1968). Earlier Fishburn (1964) discussed this tack informally, but he did not propose an axiomatic theory. Pfanzagl’s theory is flawed in two major ways. First, as in the von Neumann

– Morgenstern theory, it deals only with two-component gambles and simple compounds of them. Second, it includes a strong assumption about the existence of independent events (defined in terms a qualitative probability ordering) which implies that the underlying sample space must be atomless. A generalization by Luce & Krantz (1969) avoids both difficulties; I shall describe it in some detail.

Table 1 summarizes how the various theories differ in their notions of an uncertain alternative.

3. Conditional Expected Utility

This theory has five primitives:

- \mathcal{E} is an algebra of subsets of a sample space X ;
- \mathcal{N} is a subset of \mathcal{E} (which will turn out to be events having probability 0);
- \mathcal{C} is an arbitrary set (whose elements are interpreted as consequences);
- \mathcal{D} is a subset of $\{f_A \mid A \in \mathcal{E} - \mathcal{N} \text{ and } f_A : A \rightarrow \mathcal{C}\}$;
- \succsim is a binary relation on \mathcal{D} .

A typical decision in this theory is an assignment of consequences to the sample points of some non-null event on which the decision can be said to be conditional. We do not take \mathcal{D} to be all such functions, but a subset of them that is characterized by our assumptions.

The first assumption includes, in part, the first rationality postulate mentioned earlier.

1. \succsim is a weak order, i.e., it is connected and transitive. As usual, \sim is defined to mean that both \succsim and \precsim hold and \succ to mean that \succsim and not \precsim .

The next four postulates are structural in the sense that they place restrictions on \mathcal{N} and \mathcal{D} . Throughout, it is assumed that $A, B \in \mathcal{E} - \mathcal{N}$ and $f_A, g_B, h_A^{(i)}$, etc. $\in \mathcal{D}$. The restriction of a function f_A to the domain B is denoted $(f_A)_B$.

2. i) If $B \subset A$, then $(f_A)_B \in \mathcal{D}$;
- ii) if $A \cap B = \phi$, then $f_A \cup g_B \in \mathcal{D}$.
3. i) $\mathcal{E} - \mathcal{N}$ has three pairwise disjoint events;
- ii) \mathcal{D}/\sim has at least two elements.
4. i) If $f_A \in \mathcal{D}$ and $B \in \mathcal{E} - \mathcal{N}$, then there exists $g_B \in \mathcal{D}$ such that $g_B \sim f_A$;
- ii) if $A \cap B = \phi$ and

$$h_A^{(1)} \cup g_B \succsim f_{A \cup B} \succsim h_A^{(2)} \cup g_B,$$

then there exists $h_A \in \mathcal{D}$ such that $h_A \cup g_B \sim f_{A \cup B}$.

5. i) If $R \in \mathcal{N}$ and $S \subset R$, then $S \in \mathcal{N}$;
 ii) $R \in \mathcal{N}$ iff, for all $A \in \mathcal{E} - \mathcal{N}$ and all $f_{A \cup R} \in \mathcal{D}$,

$$f_{A \cup R} \sim (f_{A \cup R})_A.$$

Assumption 2 is a form of closure for \mathcal{D} ; 3 simply avoids trivialness; 4 insures that \mathcal{D} is adequately rich; and 5 characterizes \mathcal{N} as the set of events over which the associated consequences do not matter.

The remaining four assumptions are, in a sense, all rationality ones.

6. If $A \cap B = \phi$ and $f_A \sim g_B$, then $f_A \cup g_B \sim f_A$.
 7. If $A \cap B = \phi$, then $f_A^{(1)} \succ f_A^{(2)}$ iff

$$f_A^{(1)} \cup g_B \succ f_A^{(2)} \cup g_B.$$

To state the last two assumptions, it is convenient to introduce a definition. Let N be an "interval" of integers, then

$$\{f_A^{(i)} | N\} \equiv \{f_A^{(i)} | f_A^{(i)} \in \mathcal{D} \text{ and } i \in N\}$$

is a *standard sequence* if there exist $g_B^{(1)} \succ g_B^{(0)}$, $A \cap B = \phi$, such that for all i , $i + 1 \in N$,

$$f_A^{(i+1)} \cup g_B^{(0)} \sim f_A^{(i)} \cup g_B^{(1)}.$$

8. (Archimedean) Every bounded standard sequence is finite.

9. If $\{f_A^{(i)} | N\}$ and $\{h_A^{(i)} | N\}$ are standard sequences and, for some k , $k + 1 \in N$,

$$f_A^{(k)} \sim h_A^{(k)} \text{ and } f_A^{(k+1)} \sim h_A^{(k+1)},$$

then for all $i \in N$, $f_A^{(i)} \sim h_A^{(i)}$.

Of the rationality postulates, 7 is the most controversial. Note that it is a natural generalisation of the sure-thing principle, which itself is not controversial. Examples have been presented by Allais (1953) and Ellsberg (1961) which convince many people that they would violate 7. Moreover, MacCrimmon (1966) has performed an experiment with business executives, a considerable fraction of whom violated 7 and, what is more impressive, when confronted with their violations many argued that they had chosen sensibly. This contrasts with their embarrassed desire to revise their choices when confronted with violations of transitivity. Nonetheless, as Savage (1954) and Raiffa (1961) have argued convincingly, 7 is compelling as a normative principle of rational behavior.

Theorem 1. If assumptions 1-9 hold, then there exist $u : \mathcal{D} \rightarrow \mathbb{R}$ and $P : \mathcal{E} \rightarrow [0, 1]$ such that

- i) $\langle X, \mathcal{E}, P \rangle$ is a finitely additive probability space;
 ii) $R \in \mathcal{N}$ iff $P(R) = 0$;

- iii) $f_A \succ g_B$ iff $u(f_A) \geq u(g_B)$;
 iv) if $A \cap B = \phi$, then

$$u(f_A \cup g_B) = u(f_A)P(A | A \cup B) + u(g_B)P(B | A \cup B).$$

In contrast to the other theories I have mentioned, it can be shown by example that \mathcal{D} need not include any constant decisions whatsoever. As a result, the utility function u is defined on \mathcal{D} , not on \mathcal{C} , and no expectation result is established. Part iv merely says that u has a major property of an expectation, but it does not say that it is one. To get an expectation we need both to add assumptions and to restrict our attention to a subclass of decisions. We say that f_A is a *gamble* if range of f_A is a finite subset of \mathcal{C} , we denote by c_A the constant decision with $c_A(x) = c$ for $x \in A$. Obviously, a gamble is a finite union of disjoint constant decisions.

Theorem 2. Assume, in addition to 1-9,

a) for $c \in \mathcal{C}$, there exists $A(c) \in \mathcal{E} - \mathcal{N}$ such that the constant decision $c_{A(c)} \in \mathcal{D}$, and

b) if $c_A, c_B \in \mathcal{D}$, then $c_A \sim c_B$.

Then there exists $v : \mathcal{C} \rightarrow \mathbb{R}$ such that for all gambles f_A

$$u(f_A) = E[v(f_A) | A],$$

where u is defined in Theorem 1 and E is the expectation relative to P of Theorem 1.

Given that Theorem 1 is true, the proof of Theorem 2 is nearly trivial.

Note that, in contrast to Savage's theory, the domain of the postulated constant decisions can be very restricted; for example, the toss of a coin will do. Assumption b is probably not a very good one. Certainly it does not describe behavior accurately and it can only be defended as a principle of rationality if we assume that all chance events have the same utility. As this seems silly, much interest attaches to the other representations that are possible when (b) is dropped. This problem is open. One particularly interesting case is to discover conditions sufficient to show that there exist $v : \mathcal{C} \rightarrow \mathbb{R}$ and $w : \mathcal{E} \rightarrow \mathbb{R}$ such that for any gamble f_A ,

$$u(f_A) = E[v(f_A) | A] + w(A).$$

The natural interpretation is that the utility of a decision is decomposed into the sum of the expectation of the utility of the consequences and the utility of the underlying event itself. Jeffrey (1965) Studied such functions W .

When X is finite, it is easy to describe how to pass from a conditional representation of decisions to an equivalent Savage-type unconditional formulation. Set $\mathcal{S} = \prod_{A \in \mathcal{E} - \mathcal{N}} A$ and define a probability measure over \mathcal{S} in terms of that over \mathcal{E} in the natural way. From this we see that comparatively simple conditional situations can lead to impractically complex unconditional ones. For example, if $\mathcal{E} - \mathcal{N}$ consists of all non-empty subsets of a five element sample space, then \mathcal{S} has 5×12^{10} elements!

4. Sketch of the Proof of Theorem 1.

The following outline of the proof is (with trivial changes in the numbering of assumptions) a direct quotation from Luce & Krantz (1969).

“A result from the theory of additive, n-dimensional conjoint measurement is used to prove Theorem 1; it is similar to, but more general than, a theorem of Debreu (1967). Suppose that \succsim is a binary relation on $\prod_{i=1}^n \mathbf{a}_i$, where $n \geq 3$ and \mathbf{a}_i are non-empty sets. We say that the relation is *independent* if, for every $M \subset N = \{1, 2, \dots, n\}$, the ordering induced by \succsim on $\prod_{i \in M} \mathbf{a}_i$ for a fixed element in $\prod_{i \in N-M} \mathbf{a}_i$ is independent of that choice. The relation satisfies *restricted solvability* if whenever

$$(b_1, \dots, b_i^{**}, \dots, b_n) \succsim (a_1, \dots, a_i, \dots, a_n) \succsim (b_1, \dots, b_i^*, \dots, b_n),$$

then there exists $b_i \in \mathbf{a}_i$ such that

$$(b_1, \dots, b_i, \dots, b_n) \sim (a_1, \dots, a_i, \dots, a_n).$$

Let I be a sequence of integers (positive or negative or both, finite or infinite), then a set $\{a_j^i \in \mathbf{a}_i | j \in I\}$ is a *standard sequence* if there exist $b_k^0, b_k^1 \in \mathbf{a}_k$, $k \in N - \{i\}$, such that for all $j, j+1 \in I$,

$$(b_1^1, \dots, a_j^i, \dots, b_n^1) \sim (b_1^0, \dots, a_j^{i+1}, \dots, b_n^0).$$

And, finally, when \succsim is a weak order and independent, we say that component \mathbf{a}_i is *essential* if \mathbf{a}_i / \sim has at least two equivalence classes.

Then the following result can be proved (Krantz, et al, in preparation):

Theorem 3. If \succsim on $\prod_{i=1}^n \mathbf{a}_i$ is an independent weak ordering for which restricted solvability holds, every bounded standard sequence is finite, and at least three components e are essential, then there exist real-valued functions φ_i on \mathbf{a}_i , $i \in N$, such that for all $a_i, b_i \in \mathbf{a}_i$,

$$(a_1, \dots, a_n) \succsim (b_1, \dots, b_n) \text{ if and only if } \sum_{i=1}^n \varphi_i(a_i) \geq \sum_{i=1}^n \varphi_i(b_i).$$

Moreover, if φ_i^* is another family of functions having the same property, then there exist numbers $\alpha > 0$ and β_i such that for all $i \in N$, $\varphi_i^* = \alpha \varphi_i + \beta_i$.

“We now turn to Theorem 1. Let A_i , $i \in N = \{1, 2, \dots, n\}$, $n \geq 2$, be any set of pairwise disjoint events from $\mathcal{E} - \mathcal{N}$. Let \mathbf{a}_{A_i} denote the set of decisions conditional on A_i and define \succsim on $\prod_{i \in N} \mathbf{a}_{A_i}$ by $f_1 \cup \dots$

$\cup f_n \succsim g_i \cup \dots \cup g_n$. For $n \geq 3$, one proves that the hypothesis of Theorem 3 are fulfilled, and for $n = 2$, one uses Axiom 3i to convert the problem into a 3-component case. The conclusion is that functions φ_i exist that are additive over the components. A variety of functions on \mathbf{a}_A will arise from various partitions of X . One next shows that these are all related to one another by positive linear transformations.

“Select any two decisions with $f^1 \succ f^0$ to serve as unit and zero (they exist by Axiom 3ii). Choose that normalization u_A of the “additive” functions defined over \mathbf{a}_A such that, for $f_A^1 \sim f^1$ and $f_A^0 \sim f^0$ (which exist by Axiom 4i), $u_A(f_A^1) = 1$ and $u_A(f_A^0) = 0$. For any $A, B \in \mathcal{E} - \mathcal{N}$ with $A \cap B = \phi$, we know that the additive decomposition

$$u_{A \cup B} = \varphi_{A, B} + \varphi_{B, A}$$

exists. But we also know that $\varphi_{A, B}$ is linearly related to u_A and $\varphi_{B, A}$ to u_B , i.e.,

$$u_{A \cup B} = P(A | \mathcal{A} | B) u_A + P(B | \mathcal{A} | B) u_B + \beta_{A, B} + \beta_{B, A}.$$

The proof is completed by showing that the β terms sum to 0, that $P(A | \mathcal{A} | B)$ can be written as $P(A)/P(\mathcal{A} | B)$, where the unconditional P is a finitely additive probability measure, and that the functions u_A are order preserving. The last part is by no means obvious and requires an extended argument. Axiom 9 is used to show that standard sequences behave appropriately.”

REFERENCES

- ALLAIS, M. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, 1953, **21**, 503-546.
- BLACKWELL, D. & Girshick, M. A. *Theory of games and statistical decisions*. New York: John Wiley & Sons, 1954.
- DAVIDSON, D. & Suppes, P. A finitistic axiomatization of subjective probability and utility. *Econometrica*, 1956, **24**, 264-275.
- DEBRUE, G. Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.) *Mathematical methods in the social sciences*, 1959. Stanford: Stanford Univer. Press, 1950 Pp. 16-26.
- DE FINETTI, B. La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 1937, **7**, 1-68. Translated into English in H. E. Kyburg, Jr. & H. E. Smokler (Eds.) *Studies in subjective probability*. New York: John Wiley & Sons, 1964. Pp. 93-158.
- ELLSBERG, D. Risk, ambiguity, and the Savage axioms. *Q. J. Econ.*, 1961 **75**, 643-669.
- FISHBURN, P. C. *Decision and value theory*. New York: John Wiley & Sons, 1964.
- JEFFREY, R. C. *The logic of decision*. New York: McGraw-Hill, 1965.
- KRANTZ, D. H., Luce R. D., Suppes, P., & Tversky, A. *Foundations of measurement*. In preparation, 1969.
- LUCE, R. D. & Krantz, D. H. Conditional expected utility. *Econometrica*, in press, 1969.
- LUCE, R. D. & Raiffa, H. *Games and decisions*. New York: John Wiley & Sons, 1957.
- MACCRIMMON, K. R. *Descriptive and normative implications of the decision theory postulates*. Western Management Sci. Inst., Univer. of Calif., Los Angeles. Working paper 102, 1966.
- PFANZAGL, J. Subjective probability derived from the Morgenstern-von Neumann utility concept. In M. Shubik (Ed.) *Essays in mathematical economics in honor of Oskar Morgenstern*. Princeton: Princeton Univer. Press, 1967. Pp. 237-251.
- PFANZAGL, J. *Theory of measurement*. New York: John Wiley & Sons, 1968.
- RAIFFA, H. Risk, ambiguity, and the Savage axioms: comment. *Q. J. Econ.*, 1961, **75**, 690-694.
- RAMSEY, F. P. Truth and probability. In F. P. Ramsey. *The foundations of mathematics and other logical essays*. New York: Harcourt, Brace, 1931, Pp. 156-198. Also in H. E. Kyburg, Jr. & H. E. Smokler (Eds.) *Studies in subjective probability*. New York: John Wiley & Sons, 1964. Pp. 61-92.
- SAMUELSON, P. Probability, utility, and the independence axiom. *Econometrica*, 1952, **20**, 670-678.
- Savage, L. J. *The foundations of statistics*. New York: John Wiley & Sons, 1954.
- SUPPES, P. & Winet, Muriel An axiomatization of utility based on the notion of utility differences. *Management. Sci.*, 1955, **1**, 259-270.
- VON NEUMANN, J. & Morgenstern, O. *Theory of games and economic behavior*. Princeton: Princeton Univer. Press, 1944, 1947, 1953.