RESOURCE LETTER

Roger H. Stuewer, *Editor*
School of Physics and Astronomy, 116 Church Street
University of Minnesota, Minneapolis, Minnesota 55455

This is one of a series of Resource Letters on different topics intended to guide college physicists, astronomers, and other scientists to some of the literature and other teaching aids that may help improve course content in specified fields. [The letter E after an item indicates elementary level or material of general interest to persons becoming informed in the field. The letter I, for intermediate level, indicates material of somewhat more specialized nature; and the letter A, indicates rather specialized or advanced material.] No Resource letter is meant to be exhaustive and complete; in time there may be more than one letter on some of the main subjects of interest. Comments on these materials as well as suggestions for future topics will be welcomed. Please send such communications to Professor Roger H. Stuewer, Editor, AAPT Resource Letters, School of Physics and Astronomy, 116 Church Street SE, University of Minnesota, Minneapolis, MN 55455.

Resource Letter ITP-1: Information Theory in Physics

W. T. Grandy, Jr.
Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071

(Received 25 November 1996; accepted 11 February 1997)

This Resource Letter provides a guide to the literature on the role of information theory in physics. Journal articles and books are cited for the following topics: early history, physical and mathematical connections, and a broad range of physical applications. © *1997 American Association of Physics Teachers.*

I. HISTORICAL INTRODUCTION

It has long been understood that physics and the notion of information are intimately related—indeed, information is the lifeblood of all science. In a very real sense the differential equations of physics are simply algorithms for processing the information contained in initial conditions. Data obtained by experiment and observation, sense perceptions, and communication either are, or contain information forming the basis of our understanding of nature. Yet, an unambiguous clear-cut definition of information remains as slippery as that of randomness, say, or complexity. Is it merely a set of data? Or is it itself physical? If the latter, as Einstein once commented upon the ether, it has no definite spacetime coordinates. While most physicists would agree that the only valid means of knowing the physical world is by obtaining information through observation and measurement, a general definition of the term is elusive, even though much effort has been devoted to the task without reaching any definite conclusions (Refs. 14, 15).

The difficulty is somewhat similar to that of attempting to explain the origin and meaning of inertia to beginning students. While the term can seem a bit obscure in its meaning, there is no ambiguity in defining inertial mass as its measure, and the concept becomes scientifically useful. Similarly, the general notion of information becomes a scientific one only if it is made measurable. This was first done in a tentative way by Nyquist (Ref. 1) in 1924, and then quite clearly by Hartley (Ref. 2) in 1928. Hartley was interested in the transmission of information in telegraphy and telephony, and concluded that a proper quantitative measure of the information in a message (symbol sequence) is the logarithm of the number of equivalent messages that *might* have been sent. For example, if a message consists of a sequence of k choices made from n symbols at each selection, then the number of equivalent messages is n^k and transmission of any one of these conveys an amount of information $k \log n$. Implicit here is a presumption that all messages are equally likely.

Quite appropriately, modern information theory had its origins in the theory of communication, though this is only one of the threads in the tapestry. From these heuristic beginnings there developed an elegant and complete theory by 1948, produced almost simultaneously by Norbert Wiener (Ref. 3) and Claude Shannon (Ref. 4). Wiener's contribution first appears in his book *Cybernetics*, the scope of his interests indicated by the subtitle ''control and communication in the animal and the machine.'' Influenced by John von Neumann, he introduces as a measure of the information associated with a probability density function $f(x)$ the quantity

$$\int_{-\infty}^{\infty} f(x)\ \log_2\ f(x) dx, \tag{1}$$

and applies it to a theory of messages in various systems. The similarity of this expression to some encountered in statistical mechanics did not escape Wiener's attention.

At virtually the same time, Shannon realized that the basic problem in sending and receiving messages was a *statistical* one, and he extended Hartley's ideas to situations in which the possible messages were not all equally probable (though they are presumed to constitute an exhaustive and mutually exclusive set, so that the probabilities sum to unity). If messages are composed of an alphabet A with n symbols having probabilities of transmission $(p_1,\ldots,p_n)$, the amount of information in a message is defined as

$$H(A) \equiv -K \sum_{i=1}^{n} p_i \log p_i , \qquad (2)$$

where K is a positive units-dependent constant. Shannon arrived at this expression through arguments of common sense and consistency, along with requirements of continuity and additivity. Because information is often transmitted in strings of binary digits (0's and 1's), it is conventional in communication theory to take the logarithm to the base 2 and measure H in *bits*. Thus H quantifies the average information per symbol of input, measured in bits. Note that if the symbols are equally probable then, because $\Sigma_i p_i = 1$, each $p_i = 1/n$ and we regain Hartley's result of maximum information. If, however, one symbol is transmitted with unit probability it follows that $H(A) = 0$ and no information is contained in a message whose content is known in advance.

One might object that there is indeed information in this latter event, but it is simply not useful. It is not the intent of the definition (2) to judge usefulness, however, nor is there any *meaning* to be attributed to a piece of information. Shannon originally thought of naming his measure ''uncertainty,'' but von Neumann urged him to call it *entropy* (perhaps accounting for the Greek letter H), arguing that a similar expression already existed in statistical mechanics. Thus was unleashed a flood of mischief that has yet to abate completely.

With this measure of the information required to estimate which message had been sent, Shannon laid the foundations of the modern mathematical theory of communication. If a communication channel (e.g., a telephone line) is noise-free, then one can expect the output message to reproduce faithfully the input. This is rarely the case, so one is led to introduce as well an output alphabet B with m symbols. The properties of the noisy channel can be described by conditional probabilities $p(i|j)$, in terms of which one defines the *mutual information*

$$H(B;A) \equiv \sum_i p_i \left[\sum_j p(j|i) \log_2 \left(\frac{p(j|i)}{p_j} \right) \right] \geqslant 0. \qquad (3)$$

It is this quantity, which reduces to $H(A)$ in a noiseless channel, that Shannon employed to state one of the most important results of his theory. The *capacity* C of a communication channel is the maximum rate, in bits per second, at which information can be transmitted from input to output. It is then a theorem that, with suitable coding and decoding, information can be transmitted without error at any rate up to and including C, and any attempt to transmit at a rate beyond capacity inevitably results in errors. Formally, C is proportional to the maximum of $H(B;A)$ over all possible input probability distributions $\{p\}$. As an example, for a single channel with additive white Gaussian noise having power spectrum S, bandwidth B, and signal power P, the capacity is

$$C = B \log_2 \left(1 + \frac{P}{SB} \right) \text{bits/s.} \qquad (4)$$

Finally, to send messages of the kind under discussion here it is necessary to encode them explicitly in some optimal way, such as in sequences of minimum-length bits. With the *noiseless coding theorem* Shannon and others showed that the mean length of these sequences is not only bounded below by $H(A)$, but can be brought arbitrarily close to it. In this sense, then, $H(A)$ can also be interpreted as the mean number of bits required to code the output of A with an ideal code.

A. The second thread

Prior to exploring applications to physical problems outside the realm of communication theory, it is useful to pause and examine a second developmental path toward a theory of information. The noted similarity of the Wiener–Shannon information measure to earlier expressions in statistical mechanics is much more than coincidence. Well over a century ago Ludwig Boltzmann's search for a theoretical expression to match Clausius's thermodynamic entropy $S = \oint dQ/T$ led him to relate entropy to probability. In the form later adopted by Planck in 1906, he suggested in his great paper of 1877 the well-known expression

$$S = k \log W, \qquad (5)$$

where k is Boltzmann's constant, and W is roughly the number of *a priori* equally probable microscopic states of the system compatible with the thermodynamic state. In classical mechanics it is a phase volume, and in quantum theory it is the measure of a manifold in Hilbert space. Rather than a probability, as Planck's abbreviation for *Wahrscheinlichheit* implies, W is actually a multiplicity factor, which can be a *factor* in a probability, of course. Indeed, Boltzmann took as his example the multinomial coefficient and derived the expression analogous to Eq. (2), in which p_i is replaced by the frequency of particle occupation of cells in phase space.

The point here is that the theoretical entropy provides a measure of our *lack* of information about the specific microscopic state of the system (which must be changing continually in any event). It is not certain how far Boltzmann's thoughts proceeded in linking Eq. (5) with information content, but it is quite clear that he knew something to be involved beyond the basic laws of physics. He writes (Ref. 16), ''The Second Law can never be proved mathematically by means of the equations of dynamics alone.'' Rather, conservation of information occurs only in reversible processes, whereas irreversibility reflects a loss of information and a consequent increase in entropy. It seems remarkable that what Boltzmann understood so well over a century ago is still found puzzling by some today.

Unfortunately, these similarities led a number of writers to jump immediately to the conclusion that Shannon's measure (the negative of Wiener's) was in fact identical to the thermodynamic entropy—a step even Boltzmann declined to take without proof. Chief among the advocates of this leap was Brillouin (Refs. 5, 17), who coined the term *negentropy* for Shannon's measure. The desire to make such an identification is understandable; but making it is lamentable, because it was not at all justified at this point on the basis of communication theory alone. The missing link was to be found several years later.

B. A third thread

In his classic thermodynamics book of 1871 Clerk Maxwell introduced his famous ''demon'' in an attempt to clarify the notions of irreversibility and the second law of thermodynamics (Ref. 19). He envisioned ''... a being whose faculties are so sharpened that he can follow every molecule in its course...,'' and inadvertantly inaugerated a vast industry in demonology that survives to the present day. The idea was that this demon could divide the volume by means of a partition containing a shutter, and then open and close the shutter so as to accumulate fast molecules on one side and slower ones on the other, thereby violating the second law. (He had actually discussed the idea in private communications as early as 1867.) Although an interesting and provocative tool at the time, the work of Boltzmann and Willard Gibbs, and its subsequent development in this century, has demonstrated that the very need and rationale for statistical mechanics is the complete lack of the kind of *microscopic* control envisioned by Maxwell's hypothetical demon. Were we able to exercise such control and follow the microscopic trajectories there would be little need for probability theory in our analysis of a many-body system. From the contextual discussion surrounding introduction of the demon, it's clear that Maxwell, too, appreciated this point.

These observations notwithstanding, the demon and its implications have been, and continue to be taken seriously, and an extensive literature has accumulated (Ref. 22). And, as might be expected from so much effort, some of the discussion has actually led to important insights, beginning with Leo Szilard's famous analysis of a one-molecule(!) gas in 1929 (Ref. 20). Briefly, Szilard (as demon) divides the volume of a cylinder into two parts by means of a partition and makes an observation as to which part the molecule occupies; the partition is now employed as a piston that is allowed to expand under the pressure of the single molecule until the gas fills the entire volume, the temperature being maintained by immersion in a heat bath; if the original partition was into equal parts, we find from Eq. (5) that the entropy decrease is just $k \log 2$, corresponding to a binary choice, and if the system is run cyclically one can continue to extract energy from it. But Szilard recognizes that there is a price for this operation in the form of acquiring the needed information to locate the molecule, and thus the entropy decrease is compensated with that represented by this information increase. (He didn't get it quite quite right, however, because it is the *discarding* of previous information at the end of each cycle that actually causes the entropy increase. In this respect Maxwell's original scenario possibly illustrates the point more transparently: after a fast or slow molecule is admitted to one side or the other that information is discarded by the demon, thereby providing an entropy increase.) This is perhaps the first *explicit* relation made between physical entropy and information. It is amusing to note that, had Szilard considered n choices rather than 2, he would have discovered Hartley's information measure.

As noted above, following Shannon's work Brillouin (Ref. 21) introduced the notion of negentropy in an attempt to cement the entropy–information relationship, but with no rigorous justification. He took the stance that the demon needed a light source to see the molecules, and it was this source that restored the entropy balance and ''exorcized'' the demon. This position was dissected and severely criticized later by Jauch and Báron (Ref. 18), although it appears that many others were also skeptical from the beginning.

II. THE PHYSICAL CONNECTION

The principal rigorous connection of information theory to physics came somewhat indirectly, with the realization by Edwin Jaynes that Shannon had actually uncovered a fundamental element of probability theory (Ref. 31). Namely, the measure of Eq. (2) can be interpreted as describing a property of *any* probability distribution. Whereas Shannon envisioned the set $\{p_i\}$ as *given* in communication theory, Jaynes turned the interpretation around to utilize available information to *determine* the probabilities. In this sense, the *entropy of a probability distribution* on an exhaustive set of mutually exclusive alternatives $(A_1,\dots,A_n)$ is defined as the functional

$$S(P_1,\dots,P_n)=-K\sum_{i=1}^{n} P_i \ln P_i, \quad K>1. \tag{6}$$

In this form S represents the uncertainty in a probability distribution as to which of the alternatives is realized. The entropy of Eq. (6) provides a quantitative measure of just how much information is required to remove this uncertainty.

A short digression is in order here to point out that Khinchin also clearly understood in 1953 that Shannon's entropy was a fundamental element of probability theory (Ref. 6). He writes, ''There is no doubt that in the years to come the study of entropy will become a permanent part of probability theory;...'' He applied information theory in this sense to Markov chains in some detail, but does not seem to have taken the probability theory connection much further.

Jaynes went on to enunciate a *principle of maximum entropy* (PME), which can be phrased as follows (Ref. 31): The distribution $\{P_i\}$ that maximizes S subject to constraints imposed by the available information is the least biased description of what we know about the set of alternatives $\{A_i\}$. The PME is a rule for rational inference that provides a variational procedure for constructing prior probabilities based on given evidence. On the one hand, if that evidence implies nothing more than the alternatives are equally probable, the only constraint is that $\Sigma_i P_i=1$ and maximization of S yields the uniform distribution $\{1/n,\dots,1/n\}$. In this event $S_{\max}=k \ln n$, the missing information is maximal, and one can make no definite predictions. On the other hand, the evidence may indicate that one alternative is certain, rendering all others impossible, in which case $S=0$ and there is no uncertainty whatsoever. The bulk of scientific inference lies somewhere in between, where one must generally cope with incomplete information. In all but the most trivial problems of science one rarely has sufficient information to construct a unique probability assignment in the same sense that declaration of an honest coin unambiguously assigns $(\frac{1}{2},\frac{1}{2})$ to the possible choices. The PME removes this ambiguity by maximizing the uncertainty subject to whatever information actually is available—it renders the distribution as uniform as possible. A direct proof that any choice of information measure other than (6) will lead to inconsistencies if pursued far enough, and that the PME is essentially unique, was subsequently provided by Shore and Johnson (Ref. 38).

As an aside, we note a slight generalization of Shannon's measure introduced by Kullback (Ref. 7):

$$H=K\sum_{i} P_i \ln(P_i/P_i^0), \quad K>0,$$

which is interpreted as the entropy of the set $\{P_i\}$ relative to the set $\{P_i^0\}$, and sometimes called the cross-entropy. It is useful when part of the initial information consists of a set of prior probabilities, and it provides for a straightforward generalization to continuous distributions, since there can be no ambiguity regarding the basic measure on the space of alternatives.

There is no logical reason at all to associate S with any physical quantity at this point, and the PME is first and foremost a rule of probability theory. But if one applies that theory to physical problems it is expected to take on physical (and maybe experimental) meaning, in the same way mathematical symbols do in any other theory. If it is applied to any other area of probable reasoning it takes on an appropriately significant meaning there. In making such applications, however, it is first necessary to express the available information in the form of mathematically well-defined constraints, and this procedure may not always be transparent.

In his 1957 papers (Refs. 31, 32) Jaynes made the compelling application to statistical mechanics and thermodynamics, having noted that the constraints provided by macroscopic information could be expressed as expectation values. He also observed that this was just the problem Gibbs had solved long ago in deriving his ensembles by minimizing his ''average index of probability of phase'' subject to constraints on average total energy, or that plus average total particle numbers (Ref. 33). Gibbs gave no explanation for why this particular function should be minimized, but this procedure is exactly what we call the PME.

With this interpretation of Shannon's information measure, along with the PME, Jaynes and others have clarified considerably the foundations of statistical mechanics, relating it ultimately to a problem of information in a way that seems to have been appreciated implicitly by the founders over a century ago. That is, S measures the amount of information about the microstate conveyed by data on macroscopic thermodynamic variables. For equilibrium systems the entropy (6) and the probabilities become equivalent to the canonical ensemble of Gibbs, with K being identified with Boltzmann's constant k. Because the canonical ensemble is known to predict experimental values, one can now safely relate the theoretical (maximum) entropy to the experimental entropy of Clausius. Quantum mechanically one employs the density matrix ρ and von Neumann's form of the entropy

$$S=-k\ (\rho \ln \rho). \tag{7}$$

Maximization of S subject to available information yields the least-biased probability assignment over the quantum states of the system. Since the theoretical function S in the form (7) is invariant under unitary transformation, it is often argued that this expression cannot describe the second law. But Jaynes (Ref. 35) later demonstrated that, in fact, it is just this property that allows one to derive the second law, which is a statement about *experimental* entropy.

A large portion of the subsequent involvement of information theory with problems of physics stems from the maximum-entropy formalism. In addition, there have been numerous other uses of information-theoretic concepts in physics not directly related to the PME, many of which are noted below. Prior to surveying these applications, though, there is another path emanating from the Wiener–Shannon formulation that requires explication.

A. An algebraic interpretation

At roughly the same time that Jaynes was developing the PME the Russian mathematician A. N. Kolmogorov realized that Shannon's information theory could be developed in several other ways (Refs. 104, 105). He noted that not all applications need pertain to events with only a probability of realization, but that the notions of entropy and information could be formulated for individual variables as well. That is, we can inquire about the information conveyed by one object about another. For example, one can consider the combinatorial aspects of binary sequences directly, and the information content of such sequences is not represented adequately in terms of Shannon's entropy. Rather, it is desirable to quantify directly the information content of a sequence recorded in the memory of a computer, say. This line of reasoning led to an algorithmic approach through the introduction of recursive functions, and eventually to a formal theory of complexity. Similar ideas were developed at almost the same time by Solmonoff (Ref. 106) and Chaitin (Ref. 107).

Briefly, the *Kolmogorov complexity*, or algorithmic information content $K(x)$ of a string x is the length l of the shortest program p executed on a universal computer U that will yield x,

$$K(x)=\min_{U(p)=x} l(p). \tag{8}$$

Subsequent developments were applied primarily to attempts to define randomness rigorously, and to a study of computable functions. A function $f(s)$ is *computable* if there is a Turing machine described by an $M(s)$ assigning finite binary strings to finite binary strings that reaches a final (or ''halting'') state such that $f(s)=M(s)$. The function $K(s)$ is uncomputable. This field of study is now known as *algorithmic information theory*, and recent years have seen a number of more direct applications to physics.

Kolmogorov initiated yet another approach to information measures in 1958 (Ref. 65), which was found independently by Sinai at about the same time (Ref. 66). The idea is to extend Shannon's entropy to the theory of dynamical systems. One considers the *dynamical* Shannon entropy per unit time h of a map, say, and defines the *Kolmogorov–Sinai entropy* S, or metric entropy as the supremum of h over all possible partitions of the phase space. The KS entropy has turned out to be very useful in nonlinear dynamics, for it can be related directly to the Lyapunov exponents of the system. In fact, it is sometimes used to *define* chaos as arising when the KS entropy is positive. Hence, the KS entropy provides a measure of the information lost per unit time as the system trajectories diverge from almost identical initial conditions. That is, information on the orbit is lost like e^{nS} as $n\rightarrow\infty$ in an iteration—or information on initial conditions is gained as the orbit makes more significant digits in these conditions important.

III. GENERAL INFORMATION THEORY

A. Journals

Almost every physics journal will contain articles with information-theoretic connections when appropriate, such as *Physical Review A* and *Journal of Statistical Physics*. The following journals are either devoted to information theory or regularly contain applications of interest:

Acta Informatica
Cybernetica
Cybernetics and Systems
IEEE Transactions on Information Theory
Information and Computation
I.R.E. Professional Group on Information Theory
Open Systems and Information Dynamics

B. Historical works

1. ‘‘Certain Factors Affecting Telegraph Speed,’’ H. Nyquist, Bell. Syst. Tech. J. **3**, 324–346 (1924). (I)
2. ‘‘Transmission of Information,’’ R. V. L. Hartley, Bell. Syst. Tech. J. **7**, 535–563 (1928). (I)
3. **Cybernetics**, N. Wiener (MIT, Cambridge, MA, 1948). (A)
4. ‘‘A Mathematical Theory of Communication,’’ C. E. Shannon, Bell. Syst. Tech. J. **27**, 379–423 (1948). Reprinted, along with a semi-popular essay by Warren Weaver, in **The Mathematical Theory of Communication**, C. E. Shannon and W. Weaver (University of Illinois Press, Urbana, 1949). (A)
5. **Science and Information Theory**, L. Brillouin (Academic, New York, 1956). (A)
6. **Mathematical Foundations of Information Theory**, A. I. Khinchin (Dover, New York, 1957). (A)
7. **Information Theory and Statistics**, S. Kullback (Wiley, New York, 1959). (A)
8. **Claude Elwood Shannon, Collected Papers**, edited by N. J. A. Sloane and A. D. Wyner (IEEE, New York, 1993). (I)

C. General theory

9. **Foundations of Information Theory**, A. Feinstein (McGraw-Hill, New York, 1958). A classic and much-quoted treatise. (A)
10. **On Measures of Information and Their Characterizations**, J. Aczel and Z. Daroczy (Academic, New York, 1975). (A)
11. **Entropy and Information in Science and Philosophy**, L. Kubat and J. Zeman (Elsevier, Amsterdam, 1975). (I)
12. **Relative Information: Theories and Applications**, G. Jumarie (Springer-Verlag, Berlin, 1990). (A)
13. **Elements of Information Theory**, T. M. Cover and J. A. Thomas (Wiley, New York, 1991). A particularly accessible introduction. (I)

The following two books are general discussions attempting to adduce the definition and meaning of information.

14. **The Meaning of Information**, D. Nauta (Mouton, The Hague, 1972). (E)
15. **The Nature of Information**, P. Young (Praeger, New York, 1987). (E)

IV. PHYSICAL APPLICATIONS

In this section we categorize the major applications of information theory to physical systems and provide the major references defining the interface. General works relating information to physics are noted, and then the major contributions stemming from utilization of the maximum entropy principle are surveyed. We next provide a sampling of the literature describing information-theoretic methods in various subfields of physics, the aim being to provide examples of how specific disciplines have adapted this tool. Finally, a few fields have been singled out for more detailed discussion because they are areas in which information-theoretic ideas are playing a principal role in very active current research.

A. General physics

16. ‘‘On Certain Questions of the Theory of Gases,’’ L. Boltzmann, Nature **51**, 413–415. (1895). A semi-popular article explaining many of his views on this subject. (E).

The following two articles provide point and counterpoint regarding physical interpretation of Shannon’s entropy.

17. ‘‘Physical Entropy and Information. II,’’ L. Brillouin, J. Appl. Phys. **22**, 338–343 (1951). (I)
18. ‘‘Entropy, Information and Szilard’s Paradox,’’ J. M. Jauch and J. G. Báron, Helv. Phys. Acta **45**, 220–232 (1972). (I)

The next four items provide rather complete coverage of the Maxwell demon issues.

19. **Theory of Heat**, J. C. Maxwell (Longmans Green, London, 1871). (I)
20. ‘‘Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen,’’ L. Szilard, Z. Phys. **53**, 840–856 (1929). Translated as ‘‘On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings,’’ A. Rapoport and M. Knoller, Behav. Sci. **9**, 301–310 (1964). (I)
21. ‘‘Maxwell’s Demon Cannot Operate: Information and Entropy. I,’’ L. Brillouin, J. Appl. Phys. **22**, 334–347 (1951). (I)
22. **Maxwell’s Demon: Entropy, Information, Computing**, edited by H. S. Leff and A. F. Rex (Princeton U.P., Princeton, and Adam Hilger, Bristol, 1990).
    Contains a very thorough survey of the literature on Maxwell’s demon. (I)

These four entries discuss a number of issues concerning the relations between information and physics as they developed from Shannon’s work.

23. **Considerations sur la theorie de la transmission de l’information et sur son applications a certains domaines de la physique**, A. Blanc-Lapierre (Institut Henri Poincaré, Paris, 1953). (I)
24. ‘‘An application of information theory: Longitudinal measurability bounds in classical and quantum physics,’’ C. D’Antoni and P. Scanzano, Found. Phys. **10**, 875–885 (1980). (A)
25. ‘‘On relations between information and physics,’’ P. Kovanic, Prob. Control Info. Th. **13**, 383–399 (1984). (I)
26. **Information and the Internal Structure of the Universe: An exploration into information physics**, T. Stonier (Springer-Verlag, London, 1990). (I)

An interesting side issue in these relations is found in

27. ‘‘Mathematical model of information communication in physics teaching process,’’ S. Liren and S. Xinping, J. Sys. Sci. Sys. Eng. **2**, 363–369 (1993). (E)

Proceedings of three recent conferences provide a broad overview of current research in the interplay between physics and information.

28. **Symposium on the Foundations of Modern Physics**, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1993). (A)
29. **Complexity, Entropy and the Physics of Information**, edited by W. H. Zurek (Addison-Wesley, Redwood City, CA, 1990). (A)
30. **Physical Origins of Time Assymetry**, edited by J. J. Halliwell, J. Pérez Mercador, and W. H. Zurek (Cambridge U.P. Cambridge, 1994). (A)

B. Maximum entropy

The principle of maximum entropy originated in the last century with Gibbs and was re-constructed in its modern, broader form by Jaynes. Because of their interests it was natural that the first major application was to statistical mechanics and the derivation of classical thermodynamics. The original papers in the modern sequence are

31. ‘‘Information Theory and Statistical Mechanics,’’ E. T. Jaynes, Phys. Rev. **106**, 620–630 (1957). (A)
32. ‘‘Information Theory and Statistical Mechanics. II,’’ E. T. Jaynes, Phys. Rev. **108**, 171–190 (1957). (A)
33. **Elementary Principles in Statistical Mechanics**, J. W. Gibbs (Ox Bow, Woodbridge, CT 1981; first published, in 1902). (I)
34. ‘‘Foundations of Probability Theory and Statistical Mechanics,’’ E. T. Jaynes, in **Delaware Seminar in the Foundations of Physics**, edited by M. Bunge (Springer-Verlag, Berlin, 1967), pp. 77–101. (A).
35. ‘‘Gibbs vs. Boltzmann Entropies,’’ E. T. Jaynes, Am. J. Phys. **33**, 391–398 (1965). (I)

These last four papers are reprinted, along with others, in

36. **E. T. Jaynes: Papers on Probability, Statistics and Statistical Physics**, edited by R. D. Rosenkrantz (Reidel, Dordrecht, Holland, 1983). (A)

A twentieth-anniversary conference was held in 1978:

37. **The Maximum Entropy Formalism**, edited by R. D. Levine and M. Tribus (MIT, Cambridge, MA, 1979). (A)
38. ‘‘Axiomatic Derivation of the Principle of Maximum Entropy and the

Principle of Minimum Cross-Entropy,'' J. E. Shore and R. W. Johnson, IEEE Trans. Inf. Th. **IT-26**, 26–37 (1980). A consistency proof of the PME. (A)

A number of textbooks and monographs developing statistical mechanics based on Jaynes's ideas have appeared since 1957, the following group being somewhat comprehensive.

39. Thermostatics and Thermodynamics, M. Tribus (Van Nostrand, Princeton, 1961). (I)

40. Concepts in Statistical Mechanics, A. Hobson (Gordon and Breach, New York, 1971). (I)

41. Atoms and Information Theory, R. Baierlein (Freeman, San Francisco, 1971). (I)

42. Foundations of Statistical Mechanics, Volume I: Equilibrium Theory, W. T. Grandy, Jr. (Reidel, Dordrecht, Holland, 1987). (A)

43. Foundations of Statistical Mechanics, Volume II: Nonequilibrium Phenomena, W. T. Grandy, Jr. (Reidel, Dordrecht, Holland, 1988). (A)

Representative applications to more specific problems in statistical mechanics are given in the following articles.

44. ``Dissipative evolution, initial conditions, and information theory,'' A. N. Proto, J. Aliaga, D. R. Napoli, D. Otero, and A. Plastino, Phys. Rev. A **39**, 4223–4229 (1989). (A)

45. ``Maximum-entropy approach to classical hard-sphere and hard-disk equations of state,'' D. Wang and L. R. Mead, J. Math. Phys. **32**, 2258–2262 (1991). (A)

In his 1975 Ph.D. thesis John Burg introduced the first application of maximum-entropy techniques into data analysis, in the context of geophysical time series (Ref. 46), though he had reported the idea about 8 years earlier. This opened an entirely new and rich area for the use of information-theoretic methods applied to physical problems. A short time later there followed an imaginative adaptation to image reconstruction by Gull and Daniell (Ref. 47), and various authors began applying these methods to general spectral analysis.

46. ``Maximum Entropy Spectral Analysis,'' J. P. Burg, Ph.D. thesis, Stanford University, 1975. (A)

47. ``Image reconstruction from incomplete and noisy data,'' S. F. Gull and G. J. Daniell, Nature **272**, 686–690 (1978). (A)

48. ``On the Rationale of Maximum-Entropy Methods,'' E. T. Jaynes, Proc. IEEE **70**, 939–952 (1982). This contains a lucid explication of Burg's method for applying the PME to time series analysis. (A)

49. Nonlinear Maximum Entropy Spectral Analysis Methods for Signal Recognition, C. H. Chen (Research Studies Press, Chichester, England, 1982). (A)

50. Nonlinear Methods of Spectral Analysis, edited by S. Haykin (Springer-Verlag, New York, 1983). (A)

For the past 15 years annual international workshops have been conducted on maximum-entropy methods, primarily but not exclusively in data analysis, and the proceedings volumes constitute an excellent source for these and numerous other applications. Here is the first, along with the most recent (the entire list can be found on Kluwer's WWW homepage, http://kapis.www.wkap.nl/).

51. Maximum Entropy and Bayesian Methods in Inverse Problems, edited by C. R. Smith and W. T. Grandy, Jr. (Reidel, Dordrecht, Holland, 1985). (A)

52. Maximum Entropy and Bayesian Methods, Santa Fe, New Mexico, 1995, edited by K. M. Hanson and R. N. Silver (Kluwer, Dordrecht, Holland, 1997). (A)

In addition, an excellent tutorial volume in the application of maximum entropy methods is

53. Maximum Entropy in Action, edited by B. Buck and V. A. Macauley (Clarendon, Oxford, 1991). (A)

C. Physics subfields

Maximum-entropy methods have found application in almost every subfield of physics, and in many other areas of science. For example, the Colorado Alliance of Research Libraries (CARL) database *UnCover* scans some 17 000 journals, and since 1988 cites almost 500 articles applying maximum-entropy techniques in more than 100 different areas of research. While it is not possible to list all those articles here, the following references provide examples of the use of general information-theoretic ideas in various subfields of physics, as well as a few earlier works incorporating maximum entropy.

1. Acoustics

54. ``Structural Information Theory of Sound,'' T. W. Barrett, Acustica **36**, 271–281 (1976). (A)

2. Atmospheric physics

55. ``A Statistical Description of Coagulation,'' J. M. Rosen, J. Colloid Interface Sci. **99**, 9–19 (1984). (A)

3. Chemistry and chemical physics

56. ``Studies in Chemical Dynamics: Information Theory and the Franck-Condon Model,'' C. L. Vila, Ph.D. thesis, Massachusetts Institute of Technology, 1978. (A)

57. ``Application of Information Theory in Chemical Physics,'' S. B. Sears, Ph.D. thesis, University of North Carolina, 1980. (A)

58. Information Theory in Analytical Chemistry, K. Echschlager (Wiley, New York, 1994). (A)

59. ``An Information-Theoretical Estimate of the Exchange Parameter in *X* Alpha Theory,'' K. B. K. Raju, P. S. V. Nair, and K. D. Sen, Chem. Phys. Lett. **170**, 89–93 (1990). (A)

4. Condensed matter

60. ``Maximum Entropy in Condensed Matter Theory,'' D. Drabold and G. Jones, in **Maximum Entropy and Bayesian Methods, Laramie, Wyoming, 1990**, edited by W. T. Grandy, Jr. and L. H. Schick (Kluwer, Dordrecht, Holland, 1991), pp. 79–92. (A)

61. ``Statistical Geometry. I. A Self-Consistent Approach to the Crystallographic Inversion Problem,'' S. W. Wilkens, J. N. Varghese, and M. S. Lehmann, Acta Cryst. A **39**, 47–60 (1983). (A)

5. Geophysics

62. ``The Maximum Entropy Approach to Inverse Problems,'' E. Rietsch, J. Geophys. **42**, 489–506 (1977). (A)

63. ``Detection of the 11-Year Sunspot Cycle Signal in Earth Rotation,'' R. G. Currie, Geophys. J. R. Astron. Soc. **61**, 131–140 (1980). (A)

6. Mathematical physics

64. ``Maximum Entropy in the Problem of Moments,'' L. R. Mead and N. Papanicolaou, J. Math. Phys. **25**, 2404–2417 (1984). (A)

7. Nonlinear dynamics

65. ``A new metric invariant of transitive dynamical systems,'' A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **119**, 861–864 (1958). (A)

66. ``On the concept of entropy for a dynamic system,'' Ya. G. Sinai, Dokl. Akad. Nauk SSSR **124**, 768–771 (1959). (A)

67. ``Kolmogorov entropy and numerical experiments,'' G. Benettin, L. Galgani, and J. M. Strelcyn, Phys. Rev. A **14**, 2338–2345 (1976). An application to the Hénon–Heiles model that explicates its properties based on numerical studies. (A)

68. ``Short Paths and Information Theory in Quantum Chaotic Scattering: Transport Through Quantum Dots,'' H. U. Baranger and P. A. Mello, Europhys. Lett. **33**, 465–470 (1996). (A)

69. Chaos in Dynamical Systems, E. Ott (Cambridge U.P., Cambridge, 1993). (A)

70. **Thermodynamics of Chaotic Systems**, C. Beck and F. Schlögl (Cambridge U.P., Cambridge, 1993). (A)

8. Nuclear physics

71. ''Information and estimation in nuclear measurements,'' J. K. Vaurio, Nucl. Instrum. Methods **99**, 373–378 (1972). (A)

72. ''Information Theory and Statistical Nuclear Reactions, I. General Theory and Applications to Few-Channel Problems,'' P. A. Mello, P. Bereyra, and T. H. Seligman, Ann. Phys. (N.Y.) **161**, 254–275 (1985). (A)

73. ''Information Theory and Statistical Nuclear Reactions. II. Many-Channel Case and Hauser–Feshbach Formula,'' W. A. Friedman and P. A. Mello, Ann. Phys. (N.Y.) **161**, 276–302 (1985). (A)

9. Optics

74. ''Information Theory in Holography,'' D. Gabor, in **Optical and Acoustical Holography**, edited by E. Camatini (Plenum, New York, 1972), pp. 23–40. (I)

75. **Optics and Information Theory,** F. T. S. Yu (Wiley, London, 1976). (A)

76. ''Information Theory Applied to Solar Radiation Concentrators,'' R. P. Patera, Ph.D. thesis, University of Miami, 1979. (A)

10. Quantum mechanics

77. ''The Information Gain by Localizing a Particle,'' V. Majernik, Acta Phys. Acad. Sci. Hung. **25**, 331–340 (1983). (A)

78. ''Uncertainty in Quantum Measurements,'' D. Deutsch, Phys. Rev. Lett. **50**, 631–633 (1983). (A)

79. ''Entropic Formulation of Uncertainty for Quantum Measurements,'' M. H. Partovi, Phys. Rev. Lett. **50**, 1883–1885 (1983). (A)

80. ''Information and quantum nonseparability,'' B. W. Schumacher, Phys. Rev. A **44**, 7047–7052 (1991). (A)

81. ''Quantum Measurements and Information Theory,'' K. E. Hellwig, Int. J. Theor. Phys. **32**, 2401–2412 (1993). (A)

11. Spacetime physics

82. **Information Theory Applied to Space–Time Physics**, H. F. Harmuth (World Scientific, Singapore, 1992). (I)

D. Physics of computation

Information processing by computers has become one of the hallmarks of our age. Because computation, no matter how abstract, is fundamentally a physical process, it is inevitably governed by the laws of physics, and these relationships have been studied by physicists and computer scientists in a number of contexts over the past 20 years or more.

Principal concern has focused on energy consumption in the computational process, and on questions of reversible and irreversible computation. In particular, the question of minimal energy requirements has produced a lively debate that continues at present. The following references summarize the developments over the past two decades, and contain references to all the original papers.

83. ''Physics and Computation,'' T. Toffoli, Int. J. Theor. Phys. **21**, 165–175 (1982). (I)

84. ''Conservative Logic,'' E. Fredkin and T. Toffoli, Int. J. Theor. Phys. **21**, 219–253 (1982). (I)

85. ''The Thermodynamics of Computation—a Review,'' C. H. Bennett, Int. J. Theor. Phys. **21**, 905–940 (1982). (I)

86. ''Information is Physical,'' R. Landauer, Phys. Today **44**, 23–29, May (1991). (I)

87. ''Minimal Energy Requirements in Communication,'' R. Landauer, Science **272**, 1914–1918 (1996). (A)

The following articles take issue with the arguments above by Bennett and Landauer. Each article is followed by rejoinders from these authors and others.

88. ''Dissipation in Computation,'' W. Porod, R. O. Grondin, D. K. Ferry, and G. Porod, Phys. Rev. Lett. **52**, 232–235 (1984). (A)

89. ''The Computer and the Heat Engine,'' O. Costa de Beauregard, Found. Phys. **19**, 725–727 (1989). (A)

90. ''Letter to the Editor,'' E. Biedermann, Phys. Today **43** (11), 122 (1990). (A)

E. Black hole physics

As is common knowledge, gravitationally collapsing objects of sufficient mass are doomed to form black holes (BHs), defined by an event horizon within which resides the singularity of the general relativistic equations. All information about the initial state of the object is radiated away during the collapse and, remarkably, the general stationary solution depends on only three externally observable parameters: mass M, angular momentum J, and charge Q of the BH. This scenario is encapsulated in John Wheeler's phrase that ''A black hole has no hair.''

Building upon a general proof by Stephen Hawking that the BH surface area cannot decrease in any process (Ref. 91), Jacob Bekenstein recognized the similarity to the mandated increase of entropy in thermodynamics, and the relation to Shannon's information measure (Ref. 92). A BH can be created in a number of ways, leading to a number of possible internal configurations corresponding to the same set of external parameters. One then defines the BH entropy as a measure of the inaccessibility of this information. Note carefully that this entropy refers to an equivalence class of BHs, and has nothing to do with thermal entropy inside the BH. After careful consideration Bekenstein found this entropy to be

$$\begin{aligned} S_{\mathrm{BH}} &= (\tfrac{1}{2}\ln 2)(\kappa c^3/4\pi\hbar G)A \\ &\simeq (1.46\times 10^{48}\ \mathrm{erg\ K^{-1}\ cm^{-2}})A, \end{aligned} \tag{9}$$

where κ is Boltzmann's constant, G is the gravitational constant, and A is the surface area of the BH. This is an enormous number, but appropriate to the maximum entropy of a massive collapsing object.

All this is relatively straightforward and provides an interesting example of the role of information theory in general relativity. To an outside observer the original information is not missing, it simply resides inside the BH and can be described by a pure state. But in 1974 Hawking made the theoretical discovery, by means of an appropriate blending of quantum mechanics and general relativity, that BHs can radiate away their energies *thermally* (Refs. 93, 94). One can think of this as pair creation in the presence of a strong gravitational field, with one member going down the hole and the other moving off to infinity. Consequently, as the BH evaporates two related contradictions emerge: the final thermal state is a *mixed* state, in contradiction of the quantum theorem that a pure state cannot evolve to a mixed state; and, all the information encapsulated within the BH somehow is lost forever when the BH finally disappears. This is the BH *information paradox*.

Attempts at a resolution now constitute a very active area of research in general relativity and quantum field theory. At this time there are essentially three separate views: (1) gravitational effects introduce an additional uncertainty over and above Heisenberg's into quantum physics; (2) the Hawking radiation may not be completely thermal, but actually carries away the information; (3) it is possible that the BH does not evaporate completely and the information remains within a Planck-scale ($\sim 10^{-33}$-cm) remnant. These three lines of thought are explored in the references below.

Subsequently Bekenstein developed BH thermodynamics a bit further (Ref. 96) by utilizing the principle of maximum entropy to verify a generalized, intrinsically quantum second law. This asserts that BH entropy plus ordinary entropy exterior to BHs never decreases. Note that this is a theoretical statement of a statistical law that goes beyond ordinary thermodynamics.

The following papers provide some of the original literature connecting BHs to information theory.

91. ''Gravitational Radiation from Colliding Black Holes,'' S. W. Hawking, Phys. Rev. Lett. **26**, 1344–1346 (1971). (A)

92. ''Black Holes and Entropy,'' J. D. Bekenstein, Phys. Rev. D **7**, 2333–2346 (1973). (A)

93. ''Black hole explosions?,'' S. W. Hawking, Nature **248**, 30–31 (1974). (A)

94. ''Breakdown of predictability in gravitational collapse,'' S. W. Hawking, Phys. Rev. D **14**, 2460–2473 (1976). (A)

95. ''Particle Creation by Black Holes,'' S. W. Hawking, Commun. Math. Phys. **43**, 199–220 (1975). (A)

96. ''Statistical black-hole thermodynamics,'' J. D. Bekenstein, Phys. Rev. D **12**, 3077–3085 (1975). (A)

97. ''Black-hole thermodynamics,'' J. D. Bekenstein, Phys. Today **33** (1), 24–31 (1980). (I)

A selection of articles representing current research on the information paradox follows.

98. ''How Fast Does Information Leak Out from a Black Hole?,'' J. D. Bekenstein, Phys. Rev. Lett. **70**, 3680–3683 (1993). (A)

99. ''Quantum mechanics, common sense, and the black hole information paradox,'' U. H. Danielsson and M. Schiffer, Phys. Rev. D **48**, 4779–4784 (1993). (A)

100. ''Spacetime information,'' J. B. Hartle, Phys. Rev. D **51**, 1800–1817 (1995). (A)

101. ''Lectures on black holes and information loss,'' T. Banks, Nucl. Phys. (Proc. Suppl.) **41**, 21–65 (1995). A review article. (A)

102. ''Black hole evolution,'' L. Thorlacius, Nucl. Phys. (Proc. Suppl.) **41**, 245–275 (1995). A review article. (A)

103. ''Black holes, Hawking radiation, and the information paradox,'' G. 't Hooft, Nucl. Phys. B (Proc. Suppl.) **43**, 1–11 (1995). A review article. (A)

F. Algorithmic information theory

Kolmogorov's development of an information measure for problems of symbol-sequence type was followed by numerous mathematically oriented applications in computer science and to models of randomness, many of these by Chaitin (Ref. 110). During the past decade, however, various models of physical systems have been analyzed with these tools, and are noted below. Efforts to relate such ''microscopic'' entropies to Shannon's measure and thermodynamic entropy have been made, and remain an area of current research. These approaches to the many-body problem are close in philosophy to that of Boltzmann's H function and H theorem. It is not yet clear whether such microscopic functions will suffer the same fate as H—namely, that they become unrelated to thermodynamic entropy in any system with substantial potential energy (Ref. 109).

104. ''Three Approaches to the Quantitative Definition of Information,'' A. N. Kolmogorov, Probl. Inf. Trans. **1**, 3–11 (1965). (A)

105. ''Logical Basis for Information Theory and Probability Theory,'' A. N. Kolmogorov, IEEE Trans. Inf. Th. **IT-14**, 662–664 (1968). (I)

106. ''A Formal Theory of Inductive Inference. I, II.,'' R. J. Solmonoff, Inform. Control **7**, 1–22, 224–254 (1964). (A)

107. ''On the length of programs for computing binary sequences,'' G. J. Chaitin, J. Assoc. Comput. Mach. **13**, 547–569 (1966). (A)

108. ''Microscopic and macroscopic entropy,'' K. Lindgren, Phys. Rev. A **38**, 4794–4798 (1988). (A)

109. ''Violation of Boltzmann's H-theorem in real gases,'' E. T. Jaynes, Phys. Rev. A **4**, 747–750 (1971). (I)

110. Algorithmic Information Theory, G. J. Chaitin (Cambridge U.P., Cambridge, 1987). (A)

111. ''Thermodynamic cost of computation, algorithmic complexity and the information metric,'' W. H. Zurek, Nature **341**, 119–124 (1989). (A)

112. ''Algorithmic treatment of the spin-echo effect,'' S. Lloyd and W. H. Zurek, J. Stat. Phys. **62**, 819–839 (1991). (A)

113. ''Complexity in quantum systems,'' A. Crisanti, M. Falcioni, and A. Vulpiani, Phys. Rev. E **50**, 138–144 (1994). An application of information complexity to a spin-$\frac{1}{2}$ particle in a magnetic field, where the Shannon entropy vanishes. (A)

114. ''Information entropy, chaos and complexity of the shell-model eigenvectors,'' V. Zelevinsky, M. Horoi, and B. A. Brown, Phys. Lett. B **350**, 141–146 (1995). (A)

115. ''Algorithmic Complexity of a Schwarzschild Black Hole,'' V. D. Dzhunnshaliev, Russ. Phys. J. **38**, 317–319 (1995). (A)

116. ''Algorithm complexity of a protein,'' D. T. Gregory, Phys. Rev. E **54**, R39–R41 (1996). (A)

117. An Introduction to Kolmogorov Complexity and Its Applications, M. Li and P. Vitányi (Springer-Verlag, Berlin, 1993). (A)

G. Quantum information theory

While von Neumann surely had some inkling of the potential interrelations between quantum mechanics and information [witness his expression (7)], the first direct connection with the new information theory appears to be that given by Jerome Rothstein in 1951 (Ref. 118). He envisions future development of an intimate relation between communication and measurement theories, which only began to be realized in the last decade.

Subsequently, von Weizsäcker proposed a theory of *ur objects* (in the German sense of ''primitive''), information atoms characterized as one bit of potential information (Ref. 119). Its adherents consider this to be a quantum theory of information and, though it has not been developed very far, it is indeed a precursor of more recent work in quantum theory. We also finds traces of information-theoretic concepts on a fundamental level in Hugh Everett's many-worlds interpretation of quantum mechanics (Ref. 120), wherein the notion of an operator is utilized in conjunction with an analog of Shannon's information measure to explicate the measurement process.

Quantum generalizations of Shannon's expression (4) for channel capacity began to appear in the early sixties, and shortly later quantum bounds on information storage capacity were obtained. Much of this work is summarized in the review article by Bekenstein and Schiffer (Ref. 122), who also obtain new results on channel capacity.

In 1973 Kholevo proved a truly fundamental theorem in quantum communication theory (Ref. 121), which establishes an upper bound on transmitted information in terms of the quantum entropy (7) when expressed in terms of bits. Armed with this theorem, Caves and Drummond (Ref. 123) have provided a general proof of the quantum-mechanical wideband, single-channel capacity upper bound for a linear bosonic channel, as well as presented an up-to-date review of this field of quantum limits. For convenient reference we state their result:

$$C_{\mathrm{WB}}=\frac{\pi}{\ln 2}\sqrt{\frac{2P}{3h}}\quad \text{bits/s}, \tag{10}$$

in terms of input power P and Planck's constant. We hasten to add, however, that this result was derived as early as 1963 by Lebedev and Levitin (Ref. 124), who provide a thermodynamic derivation of the narrowband capacity as well. An excellent introduction to the various issues of channel capacity is provided by Pendry (Ref. 125), who also gives a rather general thermodynamic derivation of Eq. (10).

The preceding results are basically straightforward generalizations of classical information theory in a direction to have been expected. Only in recent years have the full implications of a purely quantum theory of information started to emerge, based on the magic of coherent superposition.

Classically, the two Boolean states 0 and 1 can be sent down a channel one bit at a time. A similar representation can be created with a quantum-mechanical 2-state system employing a fixed pair of orthogonal states—for example, a spin-$\frac{1}{2}$ system with states $|\uparrow\rangle$, $|\downarrow\rangle$. These systems are appropriately referred to as *qubits*, for they can also exist as superpositions, a state that might be considered a ''random'' bit. Things become more interesting when a *pair* of qubits is considered. Possible basis vectors are the direct-product states $|\uparrow\rangle_1|\uparrow\rangle_2$, $|\downarrow\rangle_1|\downarrow\rangle_2$, for example. But multiparticle superposition can also lead to *entangled states*, as Schrödinger called them, such as the singlet

$$|\psi\rangle=(|\uparrow\rangle_1|\downarrow\rangle_2-|\downarrow\rangle_1|\uparrow\rangle_2)/\sqrt{2}. \tag{11}$$

Some striking features of entangled states were first discussed in the famous Einstein–Podolsky–Rosen paper of 1935, arising because these states cannot be factored into direct products of two single-particle states in *any* representation.

It has been possible for a number of years now to produce and study entanglement in the laboratory using photon polarizations (Ref. 126), but only relatively recently have the properties of these states been exploited in the development of quantum information theory. For example, in a scheme first suggested by Bennett and Wiesner (Ref. 129), a pair of qubits can be employed to communicate between two parties in such a way that 2 bits of information are transmitted by manipulating only *one* of the two particles. The past year has seen the first experimental realization of this form of quantum communication (Ref. 128), in which data were encoded as 0's, 1's, and 2's because the photon pair can actually represent three states. This unit is called a *trit* (~1.58 bit).

Entanglement has also been employed to prove the possibility of *quantum teleportation*, another example of quantum nonlocality (Ref. 130). An unknown quantum state is teleported from one place to another by clever interaction with an entangled EPR-type pair. For the moment this operation appears beyond present technology.

Rather than bits, the fundamental units of quantum information theory are qubits, and we might expect the quantum entropy $S(\rho)$ of Eq. (7) to appear in a *quantum coding theorem* analogous to that of classical information theory. Schumacher (Ref. 131) has shown that this is indeed the case, and that $S(\rho)$ describing an ensemble of states is just the mean number of qubits required to encode these states in an ideal coding scheme. One might expect this coding theorem to play a significant role in the potential applications of quantum information theory described below.

118. ''Information, Measurement, and Quantum Mechanics,'' J. Rothstein, Science **114**, 171–175 (1951). (I)

119. Die Einheit der Natur, C. F. Von Weizsäcker (Hanser, München, 1971). (A)

120. The Many-Worlds Interpretation of Quantum Mechanics, edited by B. S. DeWitt and N. Graham (Princeton U.P., Princeton, 1973). (A)

121. ''Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel,'' A. S. Kholevo, Probl. Inf. Trans. **9**, 177–183 (1973). (A)

122. ''Quantum Limitations on the Storage and Transmission of Information,'' J. D. Bekenstein and M. Schiffer, Int. J. Mod. Phys. **1**, 355–422 (1990). (A)

123. ''Quantum limits on bosonic communication rates,'' C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 481–537 (1994). (A)

124. ''The Maximum Amount of Information Transmissible by an Electromagnetic Field,'' D. S. Lebedev and L. B. Levitin, Sov. Phys. Dokl. **8**, 377–379 (1963). (A)

125. ''Quantum limits to the flow of information and entropy,'' J. B. Pendry, J. Phys. A: Math. Gen. **16**, 2161–2171 (1983). (A)

126. ''Multiparticle interferometry and the superposition principle,'' D. M. Greenberger, M. A. Horne, and A. Zeilinger, Phys. Today **46** (8), 22–29 (1993). (I)

127. ''Quantum Information and Computation,'' C. H. Bennett, Phys. Today **48** (10), 24–30 (1995). (I)

128. ''Dense Coding in Experimental Quantum Communication,'' K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger, Phys. Rev. Lett. **76**, 4656–4659 (1996). (A)

129. ''Communication via One- and Two-Particle Operators on Einstein–Podolsky–Rosen States,'' C. H. Bennett and S. Wiesner, Phys. Rev. Lett. **69**, 2881–2884 (1992). (A)

130. ''Teleporting an Unknown Quantum State via Dual Classical and Einstein–Podolsky–Rosen Channels,'' C. H. Bennett, G. Brassard, C. Crespeau, R. Jozsa, A. Peres, and W. Wooters, Phys. Rev. Lett. **70**, 1895–1899 (1993). (A)

131. ''Quantum coding,'' B. Schumacher, Phys. Rev. A **51**, 2738–2747 (1995). (A)

Next, and finally, we consider two specific potential applications of quantum information theory. They are exciting to contemplate for no other reason than the stimulus they give to experimental work at the very foundations of the quantum theory. Those who continually struggle to make sense of quantum mechanics—which, to one degree or another, includes most physicists—can but encourage these activities. It is also possible that the work will eventually enhance our technology.

1. Quantum cryptography

Throughout history the use of secret codes in both government and commerce has attracted substantial interest and resources. Most of the security of any code resides in its *key*, which in its absolutely secure realization is a pseudo-random sequence of bits as long as the message itself, and is discarded after a single use. This is the so-called *one-time pad*. But the key is also the achilles heel of the code, for it must be transmitted from sender to receiver without being compromised. Modern cryptographers have generally circumvented the issue. The present paradigm for secure encryption is public key cryptography, which is a 2-key system, one for enciphering and one for deciphering. The first is made public by the potential receiver, who keeps secret the latter, and both keys are needed for deciphering an encrypted message. The present standard realization of this scheme is the so-called RSA algorithm based on keys that are products of large prime numbers ($\gtrsim$200 digits). Security is provided by the (present) extreme difficulty in factoring large numbers in a reasonable time. While the scheme has never been broken, one difficulty is that it has never been proven to be unbreakable. Another is that in practice the scheme has now been shown to be vulnerable to timing attacks. That is, by measuring the amount of time a computer takes to perform private key operations it is possible to find the secret key, and in a computationally inexpensive way. Thus, public key distribution is subject to both technological and mathematical advances, one extreme example of which is noted below. For this reason, new methods of secure information transfer are being explored, and one very promising scheme is that of quantum cryptography, which in turn employs some fundamental features of quantum information theory.

In 1984 Bennet and Brassard, building on an earlier idea of Wiesner, proposed an alternative to public key cryptography that re-introduces the one-time pad, but provides an absolutely secure means of distributing a key (Ref. 132). The scheme relies on transmitting polarized photons, and on the uncertainty principle. That is, anyone eavesdropping on transmission of the key bits does not know the polarization in advance, so must obtain precise values of two non-commuting observables to find out. But any such attempt at eavesdropping (interfering with the system) can be detected, in which case the transfer is abandoned and re-attempted until a secure channel is obtained. Transmission of polarized photons via an optical fiber has progressed to the point that an effort this past year was successful over a distance of 22.7 km (Ref. 139), so that the scheme can no longer be thought impractical. Similar programs have also been advocated using entangled states to transfer the key (Ref. 136), encoding by two non-orthogonal states (Ref. 137), and finally with two orthogonal states (Ref. 138). All of these schemes have been proved secure only for *noiseless* channels. Just very recently has a protocol been developed and proved secure in the presence of both noise and an eavesdropper (Ref. 141).

This field is changing so rapidly that we include here, for the most part, only papers and reviews that are relatively recent. The first three articles are excellent reviews and contain complete references to the (short) historical path, whereas the remainder provide a selection of key current research efforts.

132. ‘‘Quantum Cryptography,’’ C. H. Bennett, G. Brassard, and A. K. Ekert, Sci. Am. **267**, (4), 50–57 (1992). (I)

133. ‘‘Quantum cryptography,’’ R. J. Hughes, D. M. Alde, P. Dyer, G. G. Luther, G. L. Morgan, and M. Schauer, Contemp. Phys. **36**, 149–163 (1995). (A)

134. ‘‘Quantum cryptography: How to beat the code breakers using quantum mechanics,’’ S. J. D. Phoenix and P. D. Townsend, Contemp. Phys. **36**, 165–195 (1994). (A)

135. ‘‘Information theoretic limits to quantum cryptography,’’ S. M. Barnett and S. J. D. Phoenix, Phys. Rev. A **48** (4), R5–R8 (1993). (A)

136. ‘‘Quantum Cryptography Based on Bell’s Theorem,’’ A. K. Ekert, Phys. Rev. Lett. **67**, 661–663 (1991). (A)

137. ‘‘Quantum Cryptography Using Any Two Nonorthogonal States,’’ C. H. Bennett, Phys. Rev. Lett. **68**, 3121–3124 (1992). (A)

138. ‘‘Quantum Cryptography Based on Orthogonal States,’’ L. Goldenberg and L. Vaidman, Phys. Rev. Lett. **75**, 1239–1243 (1995). (A)

139. ‘‘Quantum cryptography over 23 km in installed under-lake telecom fibre,’’ A. Muller, H. Zbinden, and N. Gisin, Europhys. Lett. **33** (4), 335–339 (1996). (A)

140. ‘‘Security against eavesdropping in quantum cryptography,’’ N. Lutkenhaus, Phys. Rev. A **54**, 97–111 (1996). (A)

141. ‘‘Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels,’’ D. Deutsch, A. Ekert, R. Jozsa, C. Machiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818–2821 (1996). (A)

2. Quantum computing

Less imminent of realization, but equally fascinating to contemplate is a computer operating on quantum-mechanical principles—possibly the ultimate information processor. Richard Feynman appears to be the first to have entertained the utilization of such a machine, in 1982, while pondering how to simulate quantum processes computationally (Refs. 142, 143). At about the same time both Paul Benioff (Ref. 144) and David Deutsch (Ref. 145) laid out the principles and studied models of quantum-mechanical computers, and in the past several years possible realizations have been described (Refs. 146, 148).

But, Feynman’s idea aside, what are the possibly practical reasons for building a quantum computer? Speed for one, incredibly massive parallel processing for another. That is, such a computer can accept input states representing a coherent superposition of many different possible inputs and evolve them into a superposition of outputs—computation is simply a sequence of unitary transformations. The first realistic example was provided by Peter Shor in 1994 (Ref. 153), who developed an algorithm that exploits the quantum multiplicity of paths to factor large numbers (n digits) in polynomial time ($\sim n^2$), which classical computers cannot do ($\exp n^{1/3}$). Following this result, Lov Grover has recently constructed a search algorithm for quantum computers that requires $\sqrt{n}$ steps to search n entries (Ref. 160).

While these results have stimulated a great deal of theoretical work in this field, perhaps more important is the associated experimental effort toward creating the gates and circuitry a quantum computer would require. By studying simple gate operations we are starting to learn much more about quantum mechanics itself, via experiments with trapped ions (Ref. 148, 151), single-atom-photon interactions in small cavities (Ref. 150), and Rydberg atoms in superconducting cavities (Ref. 159).

As important as this work is, however, the difficulties in actually realizing a quantum computer remain enormous, if not overwhelming. Recent experimental work has surely made the construction of quantum-logic gates feasible. But combining a large number of gates requires maintenance of quantum coherence on a very large scale throughout a computer. Macroscopic quantum effects such as superfluidity and superconductivity involve only a *single* quantum state, whereas quantum computation involves coherent superposition of huge numbers of states. While this decoherence problem is immense, some progress has been booked by developing appropriate error-correcting codes. These concerns are spelled out in more detail in two recent articles (Refs. 152, 156).

All of these problems are under intense scrutiny and constitute an exciting area of current research that is described in the following references.

142. ‘‘Simulating Physics with Computers,’’ R. P. Feynman, Int. J. Theor. Phys. **21**, 467–488 (1982). (A)

143. ‘‘Quantum Mechanical Computers,’’ R. P. Feynman, Found. Phys. **16**, 507–531 (1986). Originally appeared in Optics News (February 1985), pp. 11–20. (A)

144. ‘‘Quantum-Mechanical Models of Turing Machines that Dissipate No Energy,’’ P. Benioff, Phys. Rev. Lett. **48**, 1581–1585 (1982). (A)

145. ‘‘Quantum Theory: The Church-Turing Principle and the Universal Quantum Computer,’’ D. Deutsch, Proc. R. Soc. London, Ser. A **400**, 97–117 (1985). (A)

146. ‘‘A Potentially Realizable Quantum Computer,’’ S. Lloyd, Science **261**, 1569–1571 (1993). (A)

147. ‘‘Quantum-Mechanical Computers,’’ S. Lloyd, Sci. Am. **273**(4), 140–145 (1995). (I)

148. ‘‘Quantum Computations with Cold Trapped Ions,’’ J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091–4094 (1995). (A)

149. ‘‘Maintaining coherence in quantum computers,’’ W. G. Unruh, Phys. Rev. A **51**, 992–997 (1995). (A)

150. ‘‘Measurement of Conditional Phase Shifts for Quantum Logic,’’ Q. A. Turchette, C. J. Hood, W. Lange, H. Mabushi, and H. J. Kimble, Phys. Rev. Lett. **75**, 4710–4713 (1995). (A)

151. ‘‘Demonstration of a Fundamental Quantum Logic Gate,’’ C. Monroe,

D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, Phys. Rev. Lett. **75**, 4714–4717 (1995). (A)

152. ‘‘Is Quantum Mechanics Useful?,’’ R. Landauer, Philos. Trans. R. Soc. A **353**, 367–376 (1995). (A)

153. ‘‘Algorithms for Quantum Computation: Discrete Logarithms and Factoring,’’ P. W. Shor, in **Proceedings of the 35th Annual Symposium on the Foundations of Computer Science**, edited by S. Goldwasser (IEEE Computer Society, Los Alamitos, CA, 1994), pp. 124–134. An expanded version of this paper entitled ‘‘Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer,’’ is available on the Los Alamos National Laboratory e-print archive: http://xxx.lanl.gov/archive/quant-ph/?9508027. Very recently it has been shown that this algorithm must be supplemented with an exponentially efficient error-correction algorithm (Ref. 158). (A)

154. ‘‘Semiclassical Fourier Transform for Quantum Computation,’’ R. B. Griffiths and C.-S. Nin, Phys. Rev. Lett. **76**, 3228–3231 (1996). Presents an improved method for performing Fourier transforms in Shor’s algorithm. (A)

155. ‘‘Information, Physics, and Computation,’’ S. C. Kak, Found. Phys. **26**, 127–137 (1996). Questions the notion that quantum computers as currently conceived can simulate quantum physics. (A)

156. ‘‘Quantum Computing: Dream or Nightmare?,’’ S. Haroche and J.-M. Raimond, Phys. Today **49**(8), 51–52 (1996). A thoughtful critique questioning the feasibility of constructing a viable quantum computer in the foreseeable future. A response by C. Monroe and D. Wineland, along with a reply by Haroche and Raimond, appears as a Letter to the Editor in Phys. Today **49**(11), 107–108 (1996). (I)

157. ‘‘Quantum computation and Shor’s factoring algorithm,’’ A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733–753 (1996). An excellent and up-to-date review. (A)

158. ‘‘Quantum computers and dissipation,’’ G. M. Palma, K. A. Suominen, and A. K. Ekert, Proc. R. Soc. London, Ser. A **452**, 567–584 (1996). (A)

159. ‘‘From Lamb Shift to Light Shifts: Vacuum and Subphoton Cavity Fields Measured by Atomic Phase Sensitive Detection,’’ M. Brune, P. Nussenzveig, F. Schmidt-Kaler, F. Bernadot, A. Maali, J. Raimond, and S. Haroche, Phys. Rev. Lett. **72**, 3339–3342 (1988). (A)

160. ‘‘A fast quantum mechanical algorithm for database search,’’ L. K. Grover, **Proceedings of The 28th ACM Symposium on Theory of Computing (STOC)** (ACM, Philadelphia, 1996), pp. 212–218. (Available on the LANL WWW e-print archive: http://xxx.lanl.gov/archive/quant-ph/?9605043. Similarly, ‘‘A fast quantum mechanical algorithm for estimating the median,’’ L. K. Grover, September 1996, is also available as e-print quant-ph/?9607024.) (A)

3. Quantum information theory on the World Wide Web

The field has become so active that there are a number of Web sites devoted exclusively to topics in quantum information theory. Be warned, though, that individual and group homepages are not always current. The preprint archives, however, *are* kept up-to-date.

Quantum Computation/Cryptography at Los Alamos
http://qso.lanl.gov/qc/
Quantum Computation and Cryptography at Oxford
http://eve.physics.ox.ac.uk/QChome.html
Laboratory for Theoretical and Quantum Computing, Univerité de Montreal
http://www.iro.umontreal.ca/labs/theorique/index_en.html
Quantum Computation at IBM
http://www.research.ibm.com/xw-quantuminfo
Tutorial on Quantum Computation
http://chemphys.weizmann.ac.il//~schmuel/comp/comp.html
Quantum Computing at Australia National University
http://aerodec.anu.edu,au/~qc/index.html
Quantum Information Page
http://vesta.physics.ucla.edu/~smolin
Quantum Computation Archive
http://feynman.stanford.edu/qcomp/

In addition, many current preprints in the field can be found at the Los Alamos preprint archive:
http://xxx.lanl.gov/archive/quant-ph/

Numerous other preprint servers can be accessed from the ICTP ‘‘one-shot’’ server in Trieste:
http://www.ictp.trieste.it/indexes/preprints.html