Quantum Cryptography: Uncertainty in the Service of Privacy

Charles H. Bennett
IBM Research Division, T. J. Watson Research Center
Yorktown Heights, NY 10598, USA.

June 1, 1992

In general, observing a quantum system disturbs it, and prevents the observer from learning its exact state before the observation. Therefore, if quantum systems are used to carry information, an eavesdropper, or even the intended recipient, may be prevented from getting out all the information that the sender put in. This negative feature of quantum mechanics has recently been put to positive use in the arts of private and discreet communication.

The goal of privacy is illustrated by two persons wishing to communicate over a public channel such as newspaper ads or electronic mail, via coded messages no one else can understand. The subtler goal of discretion is illustrated by two singles exploring the possibility of a date, who wish to communicate in a way that protects them, not from eavesdroppers, but from *each other*. For example, if Alice likes Bob, Bob should only be able to learn this fact if he also likes Alice, and vice versa. Or Alice and Bob might seek to make some more complicated decision based on private information, while protecting its confidentiality as much as possible. Both privacy and discretion can be achieved by using quantum uncertainty to do what mathematics alone cannot do.

Mathematics offers imperfect solutions to these problems, in in the form of public key cryptography (PKC)[1]. Before the introduction of PKC in 1976, private communication over a public channel was thought to be impossible unless the two users had agreed beforehand on some random secret information that no one else knew. This information would serve as a cryptographic key, enabling Alice to scramble her messages in a way only Bob, knowing the same key, could unscramble. Privacy thus appeared to be generally unavailable to users lacking the means and foresight by which diplomats and spies share secret keys with their intended correspondents. Classical cryptography offered no purely mathematical solution to this "key distribution problem". Meanwhile the only solution to the discreet communications problem appeared to be for Alice and Bob to confide their secrets in a trusted intermediary, who might later choose to blackmail them.

All this was changed by PKC. The idea of PKC is for each user (eg Alice) to randomly choose a pair of mutually inverse transformations—a scrambling transformation and an unscrambling transformation—and to publish the directions for performing the former but not the latter. Anyone, including Bob, can then use Alice's public scrambling algorithm to prepare a message that only she can unscramble. Similarly anyone, including Alice, can use Bob's public scrambling algorithm to prepare a message that only he can unscramble. Thus Alice and Bob can exchange secret messages even though they share no secret to begin with. Subsequently Andrew Yao[2] and others[3] showed how to use public key techniques for discreet communication without the help of a third party. Unfortunately PKC rests on unproven mathematical assumptions, such as the difficulty of factoring large integers, and the entire edifice of privacy and discretion built upon it could come crashing down tomorrow.

Both the secret-key and public-key approaches described above tacitly assume that messages are in some physical form that can be accurately read and copied (though not necessarily understood) by anyone having access to them. With quantum systems, this is no longer so, because measurements on a quantum system cannot extract complete information about the state in which it was prepared. The incomplete accessibility of quantum information is illustrated by the behavior of individual photons of polarized light.

Photons can be prepared in a continuum of polarization states including in particular the two "rectilinear" states, horizontal ($\leftrightarrow$) and vertical ($\updownarrow$), and the two diagonal states 45° (⤢) and 135° (⤡). The two rectilinear states $\leftrightarrow$ and $\updownarrow$ can be distinguished by one measurement, and the two diagonal states ⤢ and ⤡ can be distinguished by another measurement; but if a rectilinear measurement is performed on a diagonal photon, the photon behaves randomly, acting half the time like $\leftrightarrow$ and half the time like $\updownarrow$, and all information about its diagonal polarization is lost. Similarly, a random result is obtained and all information is lost if a diagonal measurement is performed on a rectilinear photon. Such "conjugate" pairs of states exist for any nontrivial quantum system as a fundamental consequence of the uncertainty principle.

One might hope to learn more about a single photon's polarization by not measuring the photon directly, but rather using a device such as a laser to amplify it into a clone of many photons, then measuring these; but this hope is vain because the uncertainty principle introduces just enough randomness in the polarizations of the daughter photons to nullify any advantage gained by having more photons to measure. Limitations on the accuracy of measuring quantum states thus imply limitations on the accuracy of copying them, and vice versa. It can be seen that quantum information has a peculiar kind of conditional readability: a message consisting of rectilinear and diagonal photons can be accurately read or copied, but only by someone who knows some of the information that went into forming the message, namely which of the photons are rectilinear and which diagonal.

This conditional readability was first put to use by Wiesner[4] in the impractical invention shown in Fig 1: quantum money that is physically impossible to copy. Quantum money would not need the warning printed on French money, of life imprisonment for counterfeiters, because the crime would be impossible to commit. A quantum banknote would contain twenty polarized photons in a secret random sequence of polarization states ($\updownarrow$, $\leftrightarrow$, $\swarrow\!\!\!\nearrow$, or $\nwarrow\!\!\!\searrow$), stored in individual perfectly reflective mirror-lined boxes capable of holding the photons indefinitely without loss (possible in principle, but not in practice). Each banknote would also have a serial number printed in ordinary ink, which would be recorded, along with that banknote's secret list of polarizations, in a book made available to banks but not the general public. The sequence of polarizations can thus be accurately read and checked by banks, who know the secret information, but can neither be read nor copied reliably by a would-be counterfeiter, who does not.

Quantum money is impractical because of the difficulty of storing photons. We now show how an invention very similar to quantum money, but using photons in flight, offers a practical solution to the key distribution problem mentioned earlier, and thus a means for two parties who share no secret initially to be assured of the privacy of their subsequent communications [5]. To perform quantum key distribution, Alice and Bob use a quantum channel, through which they send polarized photons, in conjunction with a classical public channel, through which they send ordinary messages. The eavesdropper, Eve, is free to tamper with the photons in the quantum channel and can read, but not alter, the messages in the public channel.

To begin the key distribution Alice sends Bob a random sequence of the four kinds of photons, $\leftrightarrow$, $\updownarrow$, $\swarrow\!\!\!\nearrow$, and $\nwarrow\!\!\!\searrow$, and Bob decides randomly for each photon whether to measure its rectilinear or diagonal polarization. Of course this spoils half the data, because half the time Bob will have made the wrong type of measurement. But now Bob and Alice use their public channel to locate and discard the bad data, without compromising the remaining good data. To do this, Bob announces publicly which types of measurements, rectilinear or diagonal, he made, but not the measurement outcomes; Alice replies, telling him which of his measurements were of the correct type; and both parties agree to keep only those instances, discarding the others. If no eavesdropping has occurred, the result should be a shared secret, which can be interpreted as a binary key by letting $\leftrightarrow$ and $\swarrow\!\!\!\nearrow$ represent 0 and $\updownarrow$ and $\nwarrow\!\!\!\searrow$ represent 1.

If Eve attempts to intercept one or more of the photons in flight, measure it, and resend a forged copy of it to Bob, she faces the same problem as the would-be counterfeiter of quantum money: not knowing at the time whether the photon in her hands is rectilinear or diagonal, she cannot measure it without running a risk of disturbing it in such a way as to introduce an error later in Bob and Alice's supposedly shared data. The final step of quantum key distribution is therefore for Alice and Bob to test their data for discrepancies, eg by publicly comparing some of it. If they find discrepancies, they have reason to suspect eavesdropping and should discard all their data and start over with a fresh batch of photons. Otherwise, they can be reasonably sure that those parts of their data not disclosed in the public comparison are indeed a shared secret, and can therefore be used as a secret key to encrypt subsequent meaningful messages.

An improved version of the above scheme, able to cope with practical problems such as detector noise, has been implemented over short distances with a light beam[6], and there is no obstacle other than expense to implementing it over arbitrarily long distances with a light beam, or over a moderate length of optical fiber (a long fiber would require amplifying the quantum signal in transit, which is equivalent to eavesdropping). Although it requires a special and somewhat inconvenient physical channel, quantum key distribution offers a solution to the private communication problem based on fundamental laws of physics, rather than on unproved mathematical assumptions.

During the eighties computer scientists[2][3] showed that any two-party problem discreetly solvable with the help of an intermediary could be discreetly solved by the two parties alone using public key cryptography. They also showed [7][8] that all such problems could be reduced to a simple, almost pointless-seeming primitive called *oblivious transfer*[9] which consists of Alice sending Bob a one-bit messages in such a way that it has exactly a half chance of arriving, and only he will know whether it did. This task involves discretion (eg Alice is not supposed to learn whether the message arrived) and it could of course be performed by an intermediary; but if Alice and Bob could perform it by themselves, they could solve any other two-party problem by themselves with maximum discretion.

An obvious quantum way to obliviously transfer a bit is for Alice to encode it in a photon and send it to Bob, randomly deciding whether to use a rectilinear photon (where $\leftrightarrow = 0$ and $vpol = 1$) or a diagonal one (where ⤢ $= 0$ and $qpol = 1$). Bob then chooses randomly whether to measure the photon rectilinearly or diagonally. If he makes the right measurement, he learns Alice's bit; otherwise he gets a random result. Finally Alice tells him whether the photon she sent was rectilinear or diagonal, thereby allowing him (but not her) to learn whether the transfer succeeded. This simple scheme described has various flaws (eg Bob can get too much partial information by measuring the polarization along an intermediate axis such as $22.5°$) which can overcome at the expense of making the scheme more complicated and using more photons.

A fully practical version of quantum oblivious transfer been described[10], and can be implemented with apparatus similar to that used for quantum key distribution. Quantum oblivious transfer has the advantage of being useful over short distances (discreet decisions are often sought by parties occupying the same room), and the disadvantage of being rather inefficient mathematically, with known practical schemes requiring thousands of photons to be sent and received to reach even simple decisions.

We have not yet mentioned the most famous quantum phenomenon involving information: the Einstein-Podolsky-Rosen (EPR) effect, which generates correlated random outcomes simultaneously in two remote places, in a way that cannot be explained ("Bell's inequality") by hypothesizing a common random cause in the past. Although the experimental verification of violations Bell's inequality are the most celebrated evidence of the correctness of quantum mechanics, the EPR effect has had surprisingly modest consequences so far for quantum cryptography. Some practical quantum key distribution schemes[11][12] use EPR states as a matter of convenience, and, if it were possible to store photons (or other conjugate states), EPR schemes would offer a secure way to store secret keys *after* distribution, which non-EPR schemes cannot[13]. Thus it appears that in the adversarial setting of cryptography, where one wishes to prove what cannot be done, the negative uncertainty principle has more useful consequences than the positive EPR effect, whose greatest consequence–an observable quantum phenomenon with no classical explanation—in turn is chiefly useful in establishing the validity of quantum mechanics.

References

[1] Diffie, W. and M.E. Hellman, "New Directions in Cryptography," *IEEE Trans. Info. Theory* **IT–22**, 644-654 (1976).

[2] Yao, A.C., "How to Generate and Exchange Secrets," *Proc. 27'th IEEE Symposium on Foundations of Computer Science*, 162-167 (1986).

[3] Goldreich,O., S. Micali, and A. Wigderson, *Proc 27'th IEEE Symposium on Foundations of Computer Science*, 174-187 (1986).

[4] Wiesner, S., "Conjugate coding", manuscript written *circa* 1970, unpublished until it appeared in *Sigact News*, Vol. 15, no. 1, 1983, pp. 78 – 88.

[5] Bennett, C.H. and G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, (IEEE, New York 1984), Bangalore, India, pp. 175.

[6] Bennett, C.H., F. Bessette, G. Brassard, L. Salvail and J. Smolin, *J. Cryptology* **5**, 3 (1992).

[7] Kilian, J., "Founding cryptography on oblivious transfer", *Proceedings of the 20th Annual ACM Symposium on Theory of Computing*, May 1988, pp. 20 – 31.

[8] Crépeau, C., "Verifiable disclosure of secrets and application", *Advances in Cryptology: Proceedings of Eurocrypt '89*, April 1989, Springer-Verlag, pp. 181 – 191.

[9] Rabin, M.O., "How to exchange secrets by oblivious transfer", Technical Memo TR–81, Aiken Computation Laboratory, Harvard University, 1981.

[10] Bennett, C.H., G. Brassard, C. Crépeau, and M.-H. Skubiszewska "Practical Quantum Oblivious Transfer" *Advances in Cryptology—Crypto–91 proceedings*, edited by J. Feigenbaum, Lecture Notes in Computer Science vol. 576, pp. 351-366 (Springer, Berlin Heidelberg, 1992).

[11] A. Ekert, *Phys. Rev. Lett.* **67**, 661 (1991).

[12] J. Rarity and P. Tapster *Phys. Rev. A*, **45**, 2052-2056 (1992).

[13] C.H. Bennett, G. Brassard, and N.D. Mermin, Phys. Rev. Lett., **68**, 557 (1992).