Eur. Phys. J. D **38**, 409–413 (2006)
DOI: 10.1140/epjd/e2006-00045-y

THE EUROPEAN PHYSICAL JOURNAL D

A fundamental threat to quantum cryptography: gravitational attacks

R. Plaga[a]

Federal Office for Information Security (BSI), 53175 Bonn, Germany

Received 25 November 2005 / Received in final form 23 January 2006
Published online 7 March 2006 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2006

Abstract. An attack on the "Bennett-Brassard 84" (BB84) quantum key-exchange protocol in which Eve exploits the action of gravitation to infer information about the quantum-mechanical state of the qubit exchanged between Alice and Bob, is described. It is demonstrated that the known laws of physics do not allow to describe the attack. Without making assumptions that are not based on broad consensus, the laws of quantum gravity, unknown up to now, would be needed even for an approximate treatment. Therefore, it is currently not possible to predict with any confidence if information gained in this attack will allow to break BB84. Contrary to previous belief, a proof of the perfect security of BB84 cannot be based on the assumption that the known laws of physics are strictly correct, yet. A speculative parameterization that characterizes the time-evolution operator of quantum gravity for the gravitational attack is presented. It allows to evaluate the results of gravitational attacks on BB84 quantitatively. It is proposed to perform state-of-the-art gravitational attacks, both for a complete security assurance of BB84 and as an unconventional search for experimental effects of quantum gravity.

PACS. 03.67.Dd Quantum cryptography – 04.60.-m Quantum gravity – 03.65.Ta Foundations of quantum mechanics; measurement theory

QICS. 22.70.+s Security proofs

1 Introduction

Quantum key-distribution (QKD) protocols, often collectively called "quantum cryptography", exploit the principles of quantum mechanics to enable the secure distribution of information [1]. It is a common belief that the perfect secrecy of keys exchanged by such protocols is guaranteed if the "known laws of physics"[1] are assumed to be strictly correct [2,3]. This would be a major advantage of quantum cryptography because an analogous security guarantee for classical cryptography — based on the correctness of proven, or at least highly plausible, mathematical theorems[2] — is not possible, yet [4].

Section 2 presents a novel attack procedure against the first and best known QKD protocol, the "Bennett-Brassard 84" (BB84) protocol [1], in which the attacker exploits the action of gravity. I demonstrate in Section 3 that this attack cannot be modelled — not even to any approximation — on the basis of the known laws of physics without making assumptions that are not based on broad consensus.

Even though its security proof is shown to be incomplete, BB84 retains its great value because it rests on completely different foundations than its classical counterparts. However, for a complete security assurance one needs to attack the protocol experimentally. In Section 4, I propose a framework in which the results of gravitational attacks on BB84 can be evaluated quantitatively. In this framework Eve breaks BB84 via gravitationally cloning a qubit, Section 5 studies if this indirectly violates special relativity. Section 6 concludes.

2 The "gravitational-attack" protocol

In the BB84 protocol the honest party ("Alice") encodes a bit of the key to be distributed by preparing a qubit "Q" either in one of the four quantum-mechanical states $|\Psi\rangle = |0\rangle$, $|\Psi\rangle = |1\rangle$, $|\Psi\rangle = |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ or $|\Psi\rangle = |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. She then sends Q to its designated receiver ("Bob"). Rigorous proofs of the security of BB84 [5,6]

[a] e-mail: rainer.plaga@bsi.bund.de

[1] Defined here as an expression that was derived from a consistent mathematical framework (a "theory of physics") and has been confirmed by repeated scientific experiments.

[2] The analogues to laws of physics.

are based on the assumption that the laws of quantum physics are correct. However, these proofs ignore gravitation. Implicitly they assume that attackers only employ the resources of quantum physics in flat space time. However, it seems overly optimistic to "require" eavesdroppers to avoid the profound difficulties that still beset any attempt to definitely answer the question: "What gravitational field corresponds to a given quantum state?"

In a "gravitational attack" the eavesdropper (malicious "Eve") employs a beam splitter to evolve Q into a state:

$$|\Psi\rangle = (|\Psi(x_1)\rangle + |\Psi(x_2)\rangle)/\sqrt{2} \quad (1)$$

consisting of two spatially separated components at the spatial positions x_1 and x_2, respectively. She then measures the state of $|\Psi(x_1)\rangle$ in one of the bases employed in the BB84 protocol ($|0\rangle, |1\rangle$) and the state of $|\Psi(x_2)\rangle$ in the other base ($|+\rangle, |-\rangle$). Depending on which of the four possible measurement results $|s\rangle$ (with $s = 0$, $s = 1$, $s = +$ or $s = -$) is obtained, a macroscopic test mass M, initially at the spatial position $x(1)$, is automatically moved to (or left at) one of four separated spatial positions $x(0)$, $x(1)$, $x(+)$, $x(-)$. Immediately thereafter Eve experimentally determines the gravitational field surrounding these four positions, e.g. with the help of a Cavendish setup. Can we derive a definite prediction for the results of Eve's field-strength determination, based on the known laws of physics?

3 The attack cannot be described by the known laws of physics — not even approximately

The only "known law of physics" that describe gravitation are classical: they derive from Einstein's theory of general relativity [7]. The only theoretical ansatz to describe the attack that keeps both general relativity and quantum physics unchanged is "semiclassical gravity". It proposes that the source of the gravitational field is the quantum expectation value of the energy tensor of matter [8]:

$$G_{\mu\nu} = 8\pi G/c^2 \langle M|T_{\mu\nu}|M\rangle \quad (2)$$

here $G_{\mu\nu}$ is the classical Einstein tensor, G is Newton's constant of gravitation, c the speed of light, $T_{\mu\nu}$ the stress-energy tensor and $|M\rangle$ the quantum mechanical state of the gravitating body. However, this expression cannot be considered to be even an established approximation to a law of nature. If there is no wave-function collapse and standard quantum physics allows a complete description of nature, i.e. if the "many-worlds" interpretation (MWI) of quantum mechanics [10,11] is correct, equation (2) predicts a nonlinear coupling of quantum-mechanical state components [9] (see Sect. 4.1 for further explanation). Page and Geilker [9] presented experimental data that rule out such a coupling at the strength expected from equation (2) with a high confidence level. The nonlinear coupling does not vanish in the low-energy or weak-field limit of equation (2). Within the MWI equation (2) is wrong even to the approximation that general relativity describes gravitation. Page and Geilker drew the conclusion that there must be as yet unknown laws of physics, beyond semiclassical general relativity, that describes their experiment.

The quantum-information community is currently not in a state of agreement whether the MWI is correct, but some of its eminent members advocated this idea [12–14], and many specialists at least admit the principal possibility that it might be correct [15]. The assumption that this interpretation is wrong clearly would be not based on a broad consensus and can therefore serve neither as a basis for a prediction of the outcome of the attack nor for any sound security proof.

More complicated schemes to couple a classical gravitational field to the quantized matter field might be possible, but have not been proposed, yet, to my knowledge. It is generally considered much more likely that general relativity will turn out to be the limit of $\hbar \to 0$ of a theory of "quantum gravity", that remains to be discovered. However, due to various technical and conceptual difficulties, all candidate theories of quantum gravity [8] still fall far short of a reliable basis for deriving "known law of physics". Moreover none of the nonperturbative approaches to this problem have obtained a definite classical limit, yet. Thus, even if one of them were a correct theory of physics, it would not be possible to derive predictions for the attack, yet. In particular there is no basis on which certain properties of the known laws of quantum mechanics, like e.g. its linearity, can be assumed to hold for quantum gravity.

Summarizing, in the possible case that the "many-worlds interpretation" is correct, even for a qualitative prediction of the result of Eve's measurement a theory of quantum gravity is needed. All proofs of the security of BB84 remain incomplete because a definite theoretical basis to address the question "What information can Eve extract from the exchanged qubit in a "gravitational attack"? does not exist presently. With other words: our failure to understand quantum gravity prevents a basic condition for any security assurance to be met for QKD, yet: the target of evaluation must be thoroughly understood, also when being under attack.

4 An "insecure-BB84" scenario: nonlinear quantum gravity

To illustrate how laws of quantum gravity could render BB84 brittle, to motivate experimental attacks on this protocol, and to supply a framework for their analysis, I characterize a speculative time-evolution operator of quantum gravity for the gravitational attack in Section 4.3. This is not meant as a serious proposal for a theory of quantum gravity, but merely as parameterization to allow a quantitative analysis of "gravitational attacks". Nonlinearity was chosen only as an example. There might be other characteristics of quantum gravity that render QKD vulnerable.

Section 4.1 reviews semiclassical gravity, already discussed in Section 3, more formally. Under the assumptions discussed in Section 3, this theory predicts a strongly nonlinear evolution during and after the "gravitational attack". As the opposite extreme Section 4.2 presents a completely linear form of the time-evolution operator of quantum gravity. In Section 4.3, I propose a "general" time-evolution operator that interpolates between these two cases.

For illustration let us always assume below that initially Alice prepares the exchanged qubit in the BB84 protocol in the state $|\Psi\rangle=|1\rangle$.

4.1 Semiclassical gravity

Let us first assume that the gravitational field remains a classical field even at the fundamental level, i.e. that equation (2) is a law of physics. As in Section 3 we assume the MWI. The initial "state"[3] of the system of qubit Q and test mass M is given as:

$$|\phi_{\text{semiclassical gravity}}\rangle(t=0) = |1\rangle \otimes |M(1)\rangle G_{\mu\nu}(1). \quad (3)$$

Here and in the following $|s\rangle$ denotes the state of a qubit exchanged in BB84, and $|M(s)\rangle$ the one of the macroscopic test mass. $G_{\mu\nu}(s)$ is the classical Einstein tensor, that characterizes the structure of space time with an isolated macroscopic test mass $M(s)$ at the spatial position $x(s)$. s denotes the state of the exchanged qubit Q according to the attack protocol (see Sect. 2). According to equation (2):

$$G_{\mu\nu}(s) = 8\pi G/c^2 \langle M(s)|T_{\mu\nu}|M(s)\rangle. \quad (4)$$

The exchanged qubit is neglected in this expression because of its usually very small mass energy. The quantum-mechanical state after the gravitational attack (Sect. 2) is given as:

$$|\phi_{QM}\rangle(t=t_f) = \frac{1}{\sqrt{2}}|1\rangle \otimes |M(1)\rangle + \frac{1}{2}(|+\rangle \otimes |M(+)\rangle + |-\rangle \otimes |M(-)\rangle). \quad (5)$$

Including the gravitational field one obtains:

$$|\phi_{\text{semiclassical gravity}}\rangle(t=t_f) = V_{\text{sg}}|\phi\rangle(t=0) = \frac{1}{\sqrt{2}}|1\rangle \otimes |M(1)\rangle + \frac{1}{2}(|+\rangle \otimes |M(+)\rangle + |-\rangle \otimes |M(-)\rangle) G_{\mu\nu}(\phi_{QM}). \quad (6)$$

The classical Einstein tensor $G_{\mu\nu}(\phi_{QM})$ characterizes the gravitational field exerted by all three mass components $M(1)$, $M(+)$ and $M(-)$. It can be evaluated by inserting equation (5) into equation (2). Because cross terms rapidly vanish due to decoherence, the source of this gravitational-field are the expectation values of the energy tensor of the three masses and one obtains:

$$G_{\mu\nu}(\phi_{QM}) = \frac{1}{2}G_{\mu\nu}(1) + \frac{1}{4}G_{\mu\nu}(+) + \frac{1}{4}G_{\mu\nu}(-). \quad (7)$$

The further evolution of this state is strongly nonlinear due to the gravitational coupling, i.e. V_{sg} cannot be a linear operator, but must be some different nonlinear operator of quantum gravity.

4.2 Linear quantum gravity

Alternatively the hypothetical gravitational quantum field could obey an equation of motion, that is precisely linear — like all other known quantum fields do. The initial state of the setup before the attack is then written as:

$$|\phi\rangle(t=0) = |1\rangle \otimes |M(1)\rangle \otimes |G_{\mu\nu}(1)\rangle \quad (8)$$

$|G_{\mu\nu}(s)\rangle$ symbolizes a hypothetical "quantum state of the gravitational field" that is characterized by a space-time structure described by the classical Einstein tensor $G_{\mu\nu}(s)$ (Eq. (4)) that describes space time for an isolated test mass $M(s)$ at spatial position $x(s)$.

U_{lqg} be a linear unitary operator. The final state at time t_f after the attack described in Section 2 is then given as:

$$\begin{aligned}|\phi_{\text{linear quantum gravity}}\rangle(t_f) &= U_{\text{lqg}}|\phi\rangle(t=0) \\ &= \frac{1}{\sqrt{2}}|1\rangle \otimes |M(1)\rangle \otimes |G_{\mu\nu}(1)\rangle \\ &+ \frac{1}{2}(|+\rangle \otimes |M(+)\rangle \otimes |G_{\mu\nu}(+)\rangle \\ &+ |-\rangle \otimes |M(-)\rangle \otimes |G_{\mu\nu}(-)\rangle). \end{aligned} \quad (9)$$

The further evolution of this state will be linear.

4.3 General quantum gravity

If the MWI interpretation is correct, it is experimentally excluded that equation (6), that is initially relatively well motivated theoretically[4], is correct (Sect. 3). On the other hand, the assumption of strictly linear U_{lqg} in equation (9), that is in agreement with all available data, lacks any theoretical basis. It has indeed been recently speculated that quantum gravity is nonlinear [16].

Nonlinear effects might not be negligible even if they occur only near the Planck energy scale M_{Planck}. Phenomenological effects at familiar energies would are typically suppressed by a factor $s = (m_p/M_{\text{Planck}})^2$, where m_p is the proton mass. Recently string theories with large

[3] This is not a quantum mechanical state in the usual sense but a juxtaposition of quantum-mechanical and classical fields.

[4] Since it only combines "known laws of physics".

extra dimensions, in which the Planck scale M_{Planck} might be as small as 1 TeV, have been developed [17]. The suppression factor s might thus be of respectable magnitude for energies commonly encountered in the laboratory.

Clearly a plausible general phenomenological ansatz for time evolution in quantum gravity must allow for the possibility of nonlinearity. Let us assume the initial state of equation (8). As the final state I propose a combination of the linear equation (9) and the semiclassical equation (6):

$$\begin{aligned}|\phi_{\text{general quantum gravity}}\rangle(t=t_f) &= V_{\text{gqg}}|\phi\rangle(t=0)\\ &= \frac{1}{\sqrt{2}}|1\rangle\otimes|M(1)\rangle\otimes|G_{\mu\nu}^{\text{general}}(1)\rangle\\ &+\frac{1}{2}(|+\rangle\otimes|M(+)\rangle\otimes|G_{\mu\nu}^{\text{general}}(+)\rangle\\ &+|-\rangle\otimes|M(-)\rangle\otimes|G_{\mu\nu}^{\text{general}}(-)\rangle)\end{aligned}\qquad(10)$$

with the classical Einstein tensor:

$$G_{\mu\nu}^{\text{general}}(s)=G_{\mu\nu}(s)+be^{-\lambda\Delta t}G_{\mu\nu}(\phi_{QM})\qquad(11)$$

$G_{\mu\nu}(s)$ is given by equation (4) and $G_{\mu\nu}(\phi_{QM})$ by equation (7). b and λ are both purely phenomenological constants. $b<1$ is the amplitude of a "nonlinear component" and $1/\lambda$ a time scale on which the nonlinear component of the gravitational field is assumed to decay spontaneously after it first appears due to some mass movement. $\Delta t=t-t_f$ is the time since moving the masses to their respective spatial positions, i.e. after the end of the attack. The evolution of this state is nonlinear due to a gravitational coupling with an amplitude $b\ e^{-\lambda\Delta t}$, i.e. V_{gqg} can be a linear operator only to some approximation.

The gravitational field described by the second term in equation (11) is determined by all three components of the test-mass state even after Eve measured Q. From equation (7) one reads that the component with the largest tensor amplitude in the second term (in our example $|1\rangle$) corresponds to the state in which Alice prepared the qubit. Via experimentally determining the exact structure of the second term, Eve can thus infer the state of the exchanged qubit. She is then able to construct a clone of the exchanged qubit and sends it to Bob. BB84 is now broken, because Eve disposes of the same resources as Bob who cannot detect her eavesdropping. In the scenario Eve's attack exploits an EmSec vulnerability: Q can be cloned due to the uncontrolled emission of static gravitational fields.

I constructed the framework of Section 4.3 wearing the hat of a security specialist, not the one of a research scientist. The latter would tend to make assumptions that allow a consistent understanding of the attack: either the standard interpretation of quantum mechanics or a quantum theory of gravity that is strictly linear. The former tries to endanger the security of BB84 with ideas that are reasonably plausible and are clearly not in conflict with the known laws of physics: the "many-worlds interpretation" of quantum mechanics and nonlinear quantum gravity.

I propose to perform the attack described in Section 2 as sensitive and on a time scale as short as possible with state-of-the-art equipment. The results of such an experiment can be used to set an upper limit on b and λ in equation (10), respectively.

For $\lambda=0$ the experimental results of Page and Geilker [9] limit b to be smaller than about 0.1. However, an attacker who exploits state-of-the-art methods could explore magnitudes of b several orders of magnitude smaller.

Sensitive limits on b and λ would be an empirical assurance that BB84 is secure against gravitational attacks. Our trust in BB84 could then be analogous to the one conferred to classical cryptographic procedures by dedicated but unsuccessful attempts of highly qualified personnel to break them. In both cases there is no guarantee that an attacker might not find some creative, unexpected way to break the protocol.

5 A successful attack does not need to violate special relativity

The illustrative successful attack option described in Section 4.3 involved the cloning of a quantum state. The "no-cloning" theorem forbids this, but its proof [1] assumes the linearity of temporal evolution that is guaranteed by the laws of conventional quantum mechanics but might not hold in quantum gravity.

More generally it was argued that any successful cloning of quantum states would necessarily enable superluminal signaling [18]. If that were true, a successful attack would appear to be ruled out under the usually stated assumptions for quantum cryptography, because superluminal signaling contradicts the "no-signaling theorem" a known law of physics that can be derived from special relativity. However, Kent [19] has recently argued that a procedure that allows the cloning of pure, localized states, but not the cloning of subsystems of "non-local" mixed states, avoids the argument above. Moreover Polchinski [20] has shown that if the MWI is correct, universal cloning leads to the possibility of communication between macroscopic components of the total wavefunction, rather than superluminal signaling. Such an "Everett phone" would neither be in obvious contradiction with any known law of physics nor would it lead to counterintuitive effects if the time scale $1/\lambda$ in equation (11) is sufficiently short.

6 Summary and outlook

It is well-known that the security of quantum cryptography could be compromised if the laws of quantum mechanics are not strictly correct[5], e.g. if the usual quantum-mechanical operator "U", describing temporal evolution, would contain a small nonlinear term. However,

[5] This realization has led to recent proposals for QKD protocols that are claimed to remain secure even if the laws of quantum mechanics are not strictly correct [21].

the assumption of its strict linearity is a law of physics that

a. derives from quantum mechanics (a mathematically consistent theory of physics) and
b. has been verified experimentally to great precision [22].

Therefore, the security of quantum cryptography was thought to rest on very solid foundations.

Here I proposed a practical attack procedure that cannot be described without a theory of quantum gravity even approximately in general. It breaks BB84 if gravitational nonlinearities exist. However, neither

a. do we know a consistent theory of quantum gravity nor
b. was the linearity of evolution in the presence of gravitational fields checked with the precision that can be achieved with state-of-the-art equipment.

Therefore presently the security of quantum cryptography against this attack can be guaranteed neither by recourse to general principles nor by evaluating results of sensitive experimental tests.

The latter gap could be quickly closed: experimental attacks on BB84 could assure at least the practical (if not theoretical) security of this protocol. Such a test receives additional justification as an unconventional search for experimental clues to quantum gravity.

A complete theoretical treatment of the fundamental security of quantum cryptography will only be possible when the correct theory of quantum gravity is found. This raises a considerable practical interest in the most fundamental subject of contemporary physics. If the security of quantum cryptography can be proved in the absence of a full theory of quantum gravity, perhaps for other protocols than BB84, is an important question for further research [23].

I sincerely thank Don Page for a very helpful extended correspondence on his seminal experiment [9], that was the model for the attack suggested here and Claus Kiefer for useful comments on a manuscript draft. Discussions with Jonathan Barrett and Adrian Kent about the security of quantum cryptography in the presence of nonlinearities were crucial to my understanding of this issue. Two anonymous referees helped to improve the manuscript with critical comments.

References

1. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002)
2. H. Lo, abstract of talk: *From quantum cheating to quantum security*, April 8, 2004, Los Alamos National Laboratory; "... quantum cryptography can come to the rescue by allowing perfectly secure communication guaranteed by the laws of physics"
3. M. Christandl, R. Renner, A. Ekert, *A Generic Security Proof for Quantum Key Distribution*, e-print `arXiv:quant-ph/0402131` (2004); "Quantum-key distribution provides perfect security because, unlike its classical counterpart, it relies on the laws of physics rather than on ensuring that successful eavesdropping would require excessive computational effort"
4. O. Goldreich, S. Goldwasser, *On the possibility of basing Cryptography on the assumption that $P \neq NP$*, Cryptology e-print `Archive1998/005` (1998)
5. H. Lo, H.F. Chau, Science **283**, 2050 (1999)
6. D. Mayers, JACM **48**, 351 (2001)
7. A. Einstein, *Grundzüge der Relativitätstheorie* (Vieweg & Sohn, Braunschweig, 1956)
8. C. Kiefer, *Quantum Gravity* (Clarendon Press, Oxford, 2004)
9. D. Page, C. Geilker, Phys. Rev. Lett. **47**, 979 (1981)
10. H. Everett III, Rev. Mod. Phys. **29**, 454 (1957)
11. H.D. Zeh, Found. Phys. **1**, 69 (1970)
12. D. Deutsch, *The Fabric of Reality: The Science of Parallel Universes - And Its Implications* (Penguin, London, 1997)
13. J. Preskill, Lecture Notes for Physics 229: *Quantum Information and Computation*, Caltech, September 1998
14. B.S. DeWitt, Phys. Today **58**, 32 (2005)
15. Private communication with participants of the workshop "Quantum information, computing and logic", July 2005 (Perimeter Institute, Waterloo)
16. C.H.-T. Wang, Class. Quantum Grav. **22**, 33 (2005)
17. M. Cavaglia, Int. J. Mod. Phys. A **18**, 1843 (2003)
18. N. Gisin, Helv. Phys. Acta **62**, 363 (1989)
19. A. Kent, Phys. Rev. A **72**, 012108 (2005)
20. J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991)
21. J. Barrett, A. Kent, L. Hardy, Phys. Rev. Lett. **95**, 010503 (2005)
22. J.J. Bollinger, D.J. Heinzen, W.M. Itano, S.L. Gilbert, D.J. Wineland, Phys. Rev. Lett. **63**, 1031 (1989)
23. J. Barrett, A. Kent, R. Plaga, in preparation