NTT Technical Review

Feature Articles: Frontier Research on Low-dimensional Semiconductor Physics

Regular Articles

Global Standardization Activities

Practical Field Information about Telecommunication Technologies

External Awards/Papers Published in Technical Journals and Conference Proceedings

Ultimate Control of Electronic Properties in Low-dimensional Semiconductors

Akira Fujiwara

Abstract

Recent progress in semiconductor technology has enabled the development of structures that are scaled down to the nanometer level with high precision by using highly sophisticated nanofabrication and crystal growth techniques. Confining electrons to such low-dimensional nanostructures makes it possible to achieve precise control of electrons as particles as well as waves. The Feature Articles in this issue review our recent research activities involving low-dimensional semiconductor devices that will be applied to achieve innovative electronics in future science and information and communication technologies.

Keywords: low-dimensional semiconductor, single-electron device, quantum device

1. Introduction

Nowadays, we are surrounded by many electronic information processing tools such as personal computers (PCs) and smartphones that quickly process enormous amounts of information delivered through high-speed networks all over the world. All of these information and communication technology (ICT) elements have become vital lifelines to support our daily life and business. One of the key technologies for sustainable development of our ICT society is semiconductor technology. A semiconductor is a material that has electrical conductivity properties between those of conductors (metals) and insulators. They include silicon used for large-scale integrated (LSI) circuits, compound semiconductors used for laser diodes in optical fiber telecommunications, and nitride semiconductors developed for white light-emitting diodes (LEDs). One of the great features of semiconductors is that they can switch electric current or emit light at a high energy-efficiency level as well as aid in the downsizing of components.

Remarkable progress has been made in semiconductor technology in the last several decades through comprehensive research on high-quality material growth, device fabrication, and device physics aimed at improving design and performance. If we take a look at the history of the LSI circuit chip, which works similarly to the central processing unit, or *brain*, of a PC, we can see that improvements in performance have been achieved year-by-year by packaging a larger number of transistors whose sizes have been miniaturized using micro- and nano-fabrication techniques. The number of transistors per chip currently exceeds 1 billion, which is on the same order as the world population.

However, while performance has grown according to this simple scaling rule, the power consumption of the chip has also increased significantly and is becoming a serious problem in terms of the environmental impact and energy costs. As a result, there are rising technical demands and expectations for low-power devices and circuits that operate on novel principles and concepts. Semiconductors, on the other hand, have attracted much attention as a material for basic research on solid-state physics. Various structures and devices that are artificially fabricated using advanced nanofabrication processes have provided an excellent platform to advance the frontiers of physics. Therefore, one can say that semiconductor research is

a field where science and technology are most beautifully united. In this respect, if we can control new physical phenomena that manifest in novel semiconductor structures, we will be able to create new electronics that achieve the ultimate functionality and performance, which are totally different from those of simple current switches and light-emission devices. This has motivated us in the Physical Science Laboratory of NTT Basic Research Laboratories to conduct research on semiconductors to explore their applications and new physical properties in order to achieve future innovation. The main results of our recent activities are reviewed in these Feature Articles.

2. Electronic properties of low-dimensional semiconductors and their ultimate control

One of the key players in the functioning of semiconductor devices is an electron that moves around freely in the semiconductors. For example, the most fundamental function of transistors in LSI circuits is to switch electric current by controlling the flow of electrons in semiconductors. This function is often considered similar to stopping and starting the flow of water. However, this is approximately true only if the size of the semiconductors is large enough. It is now well known in quantum mechanics that an electron—which is an elementary particle with a minimum electric charge of approximately 1.6×10^{-19} Coulombs—behaves as a wave with the de Broglie wavelength. When the electron is confined in nanometer-scale structures artificially constructed through crystal growth of semiconductor heterostructures and/or nanofabrication, which are often called low-dimensional semiconductors, its quantum mechanical nature appears as a *wave*. Furthermore, even as a particle, electrons behave very differently since electrons confined in a small space have a large repulsive Coulomb force between them due to their negative charge, although the amount of the charge is small. This is in contrast to the case when electrons move around freely in a larger space. We can utilize such an electron-electron interaction to manipulate and control individual electrons. In addition, electrons have internal states called *spin* as well as the corresponding magnetic moments. Spin has been applied practically in magnetic memory constructed of ferroelectric metals, but it has also attracted attention recently as a potential information carrier in new information processes. This is because the use of the flow of spin, that is, spin current, could prevent energy dissipation and

heat generation, which are inevitable when electron current—the flow of charge—is used for information processing.

As described above, electrons in low-dimensional semiconductors show a variety of unique properties based on quantum mechanics, electron-electron interaction, spin, and their combined phenomena. Such properties are spawning research subjects from both pure physics and engineering for future science and technology. Typical structures of low-dimensional semiconductors are often categorized into quantum wells (two-dimensional planar structures), quantum wires (one-dimensional linear structures), and quantum dots (structures confined in all directions, such as cubes and spheres, thereby called zero-dimensional structures).

A zero-dimensional structure, in which electrons can be regarded as particles but are dominated by the electron-electron Coulomb interaction, is sometimes called a single-electron island because we can store electrons in it one by one. The research on low-dimensional semiconductors has a long history, and many studies have been done on their physics as well as device applications such as single-electron devices. Recently, technical progress achieved in both high-quality crystal growth and sophisticated nanodevice fabrication has been opening up the possibility of the ultimate control of electrons. The article "High-speed Single-electron Transfer toward High-accuracy Current Standards" [1] describes fast and accurate control of electrons as particles achieved with silicon nanowire devices. One of the targets of this study is to achieve electric current standards for metrology. If we can generate accurate current based on the clocked transfer of single electrons, it may be possible to complete the so-called metrological quantum triangle of electric standards. The current standards remain to be developed as the final element in addition to the already existing quantum resistance and voltage standards based on semiconductor devices and superconductor devices, respectively.

The article "Noise in Nanometer-scale Electronic Devices" [2] describes thermal noise in ultimately miniaturized electronic devices, which is investigated based on the resolution of single electrons. A silicon nanotransistor with ultra-high charge sensitivity is utilized to read out the voltage noise in an ultra-small dynamic random access memory (DRAM) and is used to measure the quantized values of fluctuating voltages resulting from the thermal fluctuation in the number of electrons in the charge node of the DRAM. We analyze the thermal noise and determine how it

deviates from that expected from the conventional electric circuit model in the tiny device.

The efforts to make silicon light emissive are introduced in "Gate Tuning of Direct/Indirect Optical Transitions in Silicon" [3]. It is well known that silicon is categorized into a material with an indirect optical transition and has low efficiency in terms of light emission. What is strikingly unique about this research is that the efficiency of the light emission is electrically increased by strengthening the direct optical transition. In fact, the reason we are able to do this is not yet clear, but the results suggest that the electron wave is strongly scattered in our quantum well structures with a specially prepared silicon/silicon dioxide interface. This is really amazing because it proves that there are still unexplored electronic properties in silicon, which already has a long history of research.

The studies above focus on silicon, but gallium arsenide (GaAs) is the focus in "Microscopic Probing of Crystalline Electrons Using Magnetic Resonance" [4]. In this work, we observe for the first time the crystallization of electrons in a two-dimensional system in a high magnetic field, which is similar to the crystal of a solid-state material in which atoms are regularly allocated in space. This takes place because electrons tend to avoid each other and find their places in an adequately spaced arrangement. In terms of physics, this corresponds to the situation where the energy of the electron-electron interaction is higher than that of the kinetic motion of electrons. Using high quality film grown based on our crystal growth technology was essential in confirming this clean electron system in two dimensions.

In "Semiconductor Quantum Structures with Single-atom Precision" [5], a unique method for fabricating low-dimensional semiconductors is introduced. It is quite different from the standard technology employing lithography (often called the top-down approach) and is used to construct an ultra-small structure by placing atoms one by one (the bottom-up approach). When a compound semiconductor is used in which the wave-like nature of electrons is prominent, the shape of the electron wave confined in a quantum-dot structure is clearly observed. This method enables us to control the shape of the electron very precisely since our structure is accurately built using the atom as a building block.

The last article, "Creating a Topological Insulator Using Semiconductor Heterostructures" [6], describes the investigation of a topological insulator, a subject of great interest in solid-state physics. Topology is originally a term used in mathematics, but a topological insulator refers to a unique material whose surfaces or edges are electrically conductive, although there are insulators in the internal parts. It is theoretically predicted that by using such a low-dimensional electron system at the edge of the material, we can generate a flow of magnetic moments of electron spin (called spin current) without any flow of electrons as charges. If this is achieved, it may lead to innovative techniques enabling information to be transmitted with almost no energy dissipation.

3. Future prospects

As indicated above, the research on semiconductors is continuously driven by both scientific and technical interests and is approaching a level where precise control of electrons and atoms will become a reality. These goals have been elusive until now. We believe that we will be able to create innovate semiconductor devices that will surpass existing devices in performance and function by further improving the controllability of electronic properties in low-dimensional semiconductors created through our advanced fabrication and crystal growth technology. The Physical Science Laboratory at NTT Basic Research Laboratories thus continues to conduct research to develop novel electronics for the future information technology society such as highly accurate and highly sensitive electronics and novel information transmitting and processing devices based on new quantum mechanical phenomena.

References

[1] G. Yamahata, K. Nishiguchi, and A. Fujiwara, "High-speed Single-electron Transfer toward High-accuracy Current Standards," NTT Technical Review, Vol. 13, No. 8, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201508fa2.html

[2] K. Nishiguchi and A. Fujiwara, "Noise in Nanometer-scale Electronic Devices," NTT Technical Review, Vol. 13, No. 8, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201508fa3.html

[3] J. Noborisaka, K. Nishiguchi, and A. Fujiwara, "Gate Tuning of Direct/Indirect Optical Transitions in Silicon," NTT Technical Review, Vol. 13, No. 8, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201508fa4.html

[4] K. Muraki and T. D. Rhone, "Microscopic Probing of Crystalline Electrons Using Magnetic Resonance," NTT Technical Review, Vol. 13, No. 8, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201508fa5.html

[5] K. Kanisawa and S. Fölsch, "Semiconductor Quantum Structures with Single-atom Precision," NTT Technical Review, Vol. 13, No. 8, 2015.

https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015
08fa6.html

[6] K. Suzuki and K. Onomitsu, "Creating a Topological Insulator Using
     Semiconductor Heterostructures," NTT Technical Review, Vol. 13,
     No. 8, 2015.
     https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015
     08fa7.html

Akira Fujiwara
Senior Distinguished Researcher, Senior Man-
ager of Physical Science Laboratory and Group
Leader of Nanodevices Research Group, NTT
Basic Research Laboratories.
He received his B.S., M.S., and Ph.D. in
applied physics from the University of Tokyo in
1989, 1991, and 1994, respectively. He joined
NTT in 1994 and has been engaged in research
on silicon nanostructures and their application to
nanodevices and single-electron devices. He was
a guest researcher at the National Institute of
Standards and Technology (NIST), Gaithersburg,
MD, USA, in 2003–2004. He was a director of
the Japanese Society of Applied Physics (JSAP)
in 2010–2011 and a visiting professor at Hok-
kaido University in 2013. He received the Inter-
national Conference on Solid State Devices and
Materials (SSDM) Young Researcher Award in
1998, the SSDM Paper Award in 1999, and JJAP
(Japanese Journal of Applied Physics) Paper
Awards in 2003, 2006, and 2013. He was award-
ed the Young Scientist Award from the Ministry
of Education, Culture, Sports, Science and Tech-
nology in 2006. He was supported by the funding
program for Next Generation World-Leading
Researchers (NEXT Program), JSPS (Japan
Society of Promotion of Science) during 2011–
2014. He is a member of JSAP and a senior
member of IEEE (Institute of Electrical and
Electronics Engineers).

High-speed Single-electron Transfer toward High-accuracy Current Standards

Gento Yamahata, Katsuhiko Nishiguchi,
and Akira Fujiwara

Abstract

Single-electron transfer is a technique for manipulating single electrons to generate an accurate electric current, which is expected to be used as a new current standard. Realization of the current standard is important for performing quantum metrology triangle experiments, with which we can confirm the accuracy of fundamental physical constants, and it could be a contributing factor in redefining the ampere. NTT Basic Research Laboratories has been studying single-electron transfer using silicon transistors. In this article, we discuss high-speed single-electron transfer via a small island electrically formed using silicon transistors and via a trap level in silicon.

Keywords: single-electron transfer, silicon transistor, current standard

1. Single-electron transfer and current standards

Single-electron (SE) transfer is a technique for conveying electrons one by one in synchronization with a clock signal. It has been extensively studied using solid devices mainly fabricated from metals or semiconductors. An accurate current flow generated by SE transfer is expected to be used as a new current standard, which corresponds to a measure of electric current. Realization of the current standard could lead to the redefinition of the ampere (**Fig. 1(a)**), which has recently attracted much attention. The definition of the ampere is presently referenced to a force produced by the same current flowing through two infinite-length conductors in a vacuum, and it is therefore difficult to achieve the exact conditions described in the definition in actual experiments. In a practical sense, the definition has the effect of fixing the value of the vacuum permeability (with a unit of $m \cdot kg \cdot s^{-2} \cdot A^{-2}$). In addition, because the ampere depends on the units of length, mass, and time, the accuracy of the ampere is strongly affected by the

change in the weight of the international prototype kilogram[*1], which is the only artifact among the standards and defines mass.

To eliminate such uncertainty originating from the artifact, the abolition of the international prototype kilogram was proposed in 2011 (and will be carried out after 2018), and, along with the abolition, the ampere will also be redefined. In the redefinition, the ampere is set by fixing the numerical value of the elementary charge e. SE transfer can generate accurate current with a value of $e \times f$ simply using an input clock signal with frequency f, which can be obtained with the highest accuracy among all standards. It is therefore expected to be used as an ultimate current standard that enables the ampere to be directly set with a single device.

In addition, an important milestone for achieving the new current standard is the quantum metrology triangle experiment (**Fig. 1(b)**). In this experiment,

---

*1 International prototype kilogram: This is an alloy of platinum and iridium, which is now used as the definition of mass. After the abolition of the international prototype kilogram, the kilogram will be set by fixing the Planck constant.
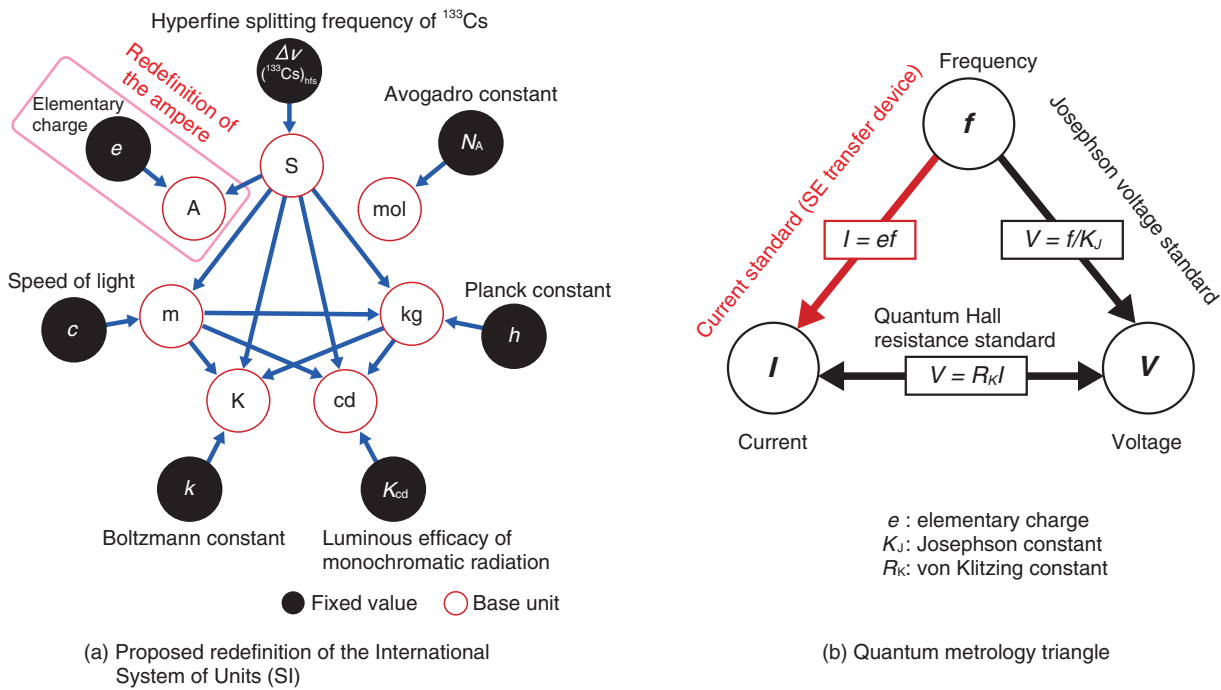
Fig. 1.   Application to current standards.

the current generated from a current standard using SE transfer is compared with that generated from a combination of the quantum Hall resistance standard[*2] and the Josephson voltage standard[*3], which are now used in the International System of Units (SI). If SE transfer with sufficiently high accuracy can be achieved, the accuracy of the von Klitzing constant $R_K = h/e^2$, where h is the Planck constant, and the Josephson constant $K_J = 2e/h$ would be verifiable through the experiment, which could lead to the realization of more accurate electrical standards.

For application to the current standard, the device should be able to generate a wide range of current. However, because e is a tiny value (~1.6×10^{-19} C), many electrons must be carried to obtain a large current level. For example, high-speed SE transfer with a gigahertz transfer frequency is necessary to perform the quantum metrology triangle experiments with high accuracy. In addition, since the accuracy of the present electrical standards is about 10^{-8}, the accuracy of SE transfer should be below 10^{-8}. SE transfer has so far been studied using various systems, but devices that can simultaneously satisfy both the speed and accuracy requirements have not yet been developed.

At NTT Basic Research Laboratories, we have been

studying SE transfer using silicon devices as part of efforts to achieve high-speed and high-accuracy SE transfer. With silicon, we should be able to stably fabricate devices that can be used for the current standard because we can use well-established device fabrication techniques accumulated through developments in the semiconductor industry. In this article, we introduce tunable-barrier SE transfer devices, which can be operated at high speed.

---

*2   Quantum Hall resistance standard: When a two-dimensional electron system is placed under a low temperature and a high magnetic field, Hall resistances are quantized due to the quantum Hall effect. The quantum Hall resistance standard is a standard using the value of the quantized Hall resistance, which is the product of the von Klitzing constant ($R_K = h/e^2$) and the reciprocal of the integer number.

*3   Josephson voltage standard: When a high-frequency signal with frequency f is applied to a structure having an insulator (or a normal conductor) sandwiched with two superconductors, the output voltages are quantized due to the alternating-current Josephson effect. The Josephson voltage standard is a standard using the value of the quantized voltage, which is the integer multiple of the product of f and the reciprocal of the Josephson constant ($K_J = 2e/h$).
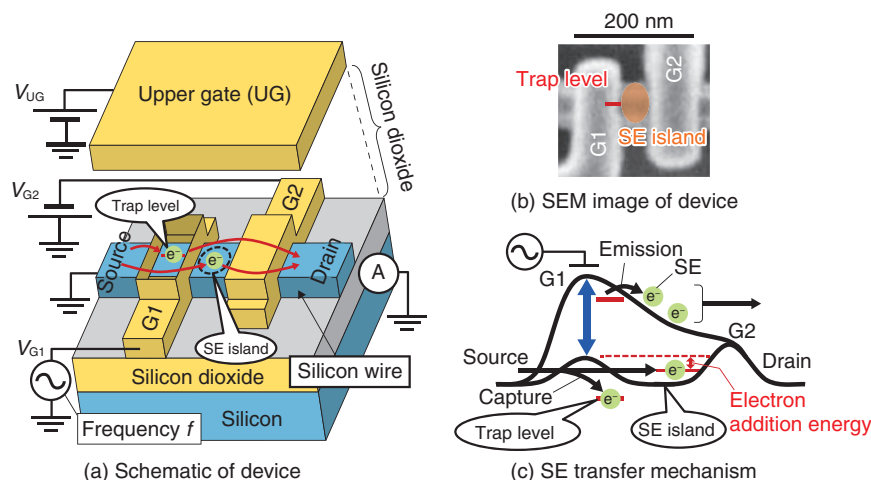
Fig. 2.   Silicon tunable-barrier SE transfer device.

2.   Silicon tunable-barrier SE transfer devices: transfer via an SE island

A schematic and an SEM (scanning electron microscope) image of the silicon tunable-barrier SE transfer device are respectively shown in **Figs. 2(a)** and **2(b)**. We used electron beam lithography to fabricate silicon transistors that have double-layer n$^+$-doped polycrystalline-silicon gate electrodes on a silicon wire with a width of a few dozen nanometers. Each lower gate (G1, G2) induces a potential barrier in the silicon wire, leading to the formation of a small region between G1 and G2 (hereafter called an SE island) in which electrons are confined. The upper gate (UG) is used to modulate the electron potential in the SE island.

To perform SE transfer via the SE island, we apply high-frequency signal V_{G1} with frequency f to G1, with fixed negative voltage V_{G2} applied to G2 (Fig. 2(a)). In this case, the barrier under G1 is periodically modulated, with a fixed barrier formed under G2, as shown in the electron potential diagram in **Fig. 2(c)**. Since G1 and the SE island are capacitively coupled, the potential in the SE island rises as the barrier under G1 rises. As a result, electrons captured by the SE island from the source when the barrier under G1 is low are eventually emitted to the drain over the barrier under G2. When an SE is transferred each cycle, the output current level is ef. To capture an accurate number of SEs in the SE island, the electron addition energy*4 in the SE island must be much larger than the thermal fluctuation energy determined by temperature. This condition can be achieved by

fabricating nanometer-scale devices. In addition, when the resistance determined by the potential barrier shape during the capture of electrons is small, the delay time determined by the product of the resistance and the island capacitance is short, resulting in fast SE transfer. In the tunable-barrier SE transfer, since the height of the potential barrier during the capture of electrons can be low, it is possible to achieve high-speed operation by reducing the resistance of the barrier.

A typical result of high-speed SE transfer using the tunable-barrier transfer device is shown in **Fig. 3(a)** [1]. By increasing the voltage applied to the UG (V_{UG}), we lower the electron potential of the SE island and change the number of electrons captured in the SE island. In the flat plateau regions (with current levels of 1ef, 2ef, and 3ef), an integer number of electrons is accurately transferred. In addition, we demonstrated high-speed single-hole transfer by fabricating a p-type source and p-type drain [2] (**Fig. 3(b)**). If radiative recombination of simultaneously transferred electrons and holes can be efficiently induced, it will be possible to realize a single-photon source with an accurate period. Furthermore, we combined the tunable-barrier SE transfer device with an SE-resolution charge sensor and evaluated the transfer accuracy by detecting the number of transferred SEs. We have so far reported that the error rate of the SE transfer is

---

*4 Electron addition energy: This is the energy required to add one electron to a certain region. The value is mainly determined by the charging energy originating from electrostatic repulsion of electrons and the quantum-mechanical energy originating from confinement of electrons.
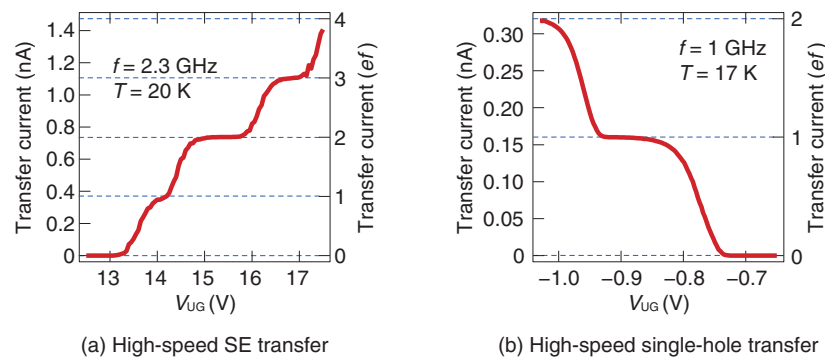
(a) High-speed SE transfer

(b) High-speed single-hole transfer

Fig. 3.　Transfer via an SE island.

(a) Comparison between
trap level and SE island

(b) Frequency dependence
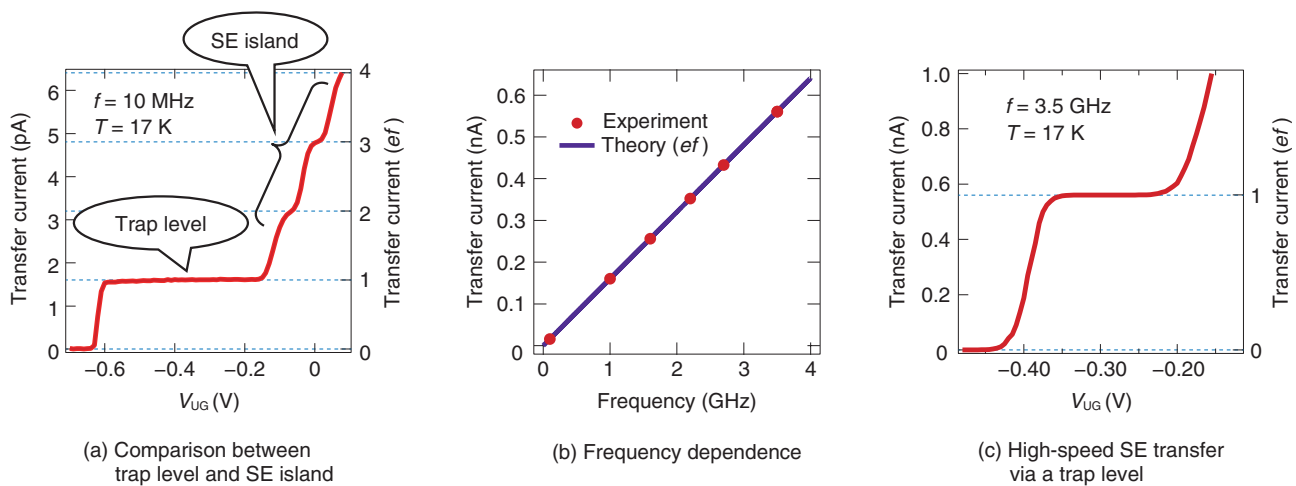
(c) High-speed SE transfer
via a trap level

Fig. 4.　SE transfer via a single trap level.

about 10^{-4} at about 100 MHz and that it is possible to reduce it to about 10^{-8} by optimizing the operating conditions [3].

3.　Silicon tunable-barrier SE transfer devices: transfer via a trap level

In the SE transfer via the SE island, increasing the electron addition energy by scaling down the device size leads to improved transfer accuracy, but device miniaturization beyond the limitation of semiconductor nanofabrication technology is not easy. In contrast, using a naturally existing trap level with an extremely fine confinement area of less than 10 nm, which is difficult to artificially fabricate, can lead to high-accuracy operation due to its large electron addition energy. Among the silicon SE transfer devic-

es like the one shown in Fig. 2(a), we selected a device that has a single trap level under the right edge of G1 and measured the transfer current via the trap level [4]. The trap level most likely originates from a trap at the interface between silicon and silicon dioxide. Although the voltages applied to G1 and G2 are similar to those for the transfer via the SE island, in the trap-mediated transfer, an SE is captured by the trap level from the source when the barrier under G1 is low. Then, when the barrier under G1 is high, the captured SE is emitted to the drain.

Current plateaus of the SE transfer at a frequency of 10 MHz are shown in **Fig. 4(a)**. The $1ef$ plateau originating from the trap-mediated transfer is much wider than the other plateaus originating from the SE-island-mediated transfer, which indicates that the trap-mediated transfer is more accurate. To perform

high-speed operation, the SE capture and emission shown in Fig. 2(c) must be sufficiently fast. Detailed measurements reveal that fast capture and emission can be achieved by lowering the barrier height under G1 during the capture phase and by applying a strong electric field at the trap level during the emission phase. This can be achieved when the high-frequency signal has a large amplitude. Under this condition, the value of the 1ef plateau increases as frequency f increases, as shown in **Fig. 4(b)**, and we achieve high-speed operation at 3.5 GHz (**Fig. 4(c)**). Moreover, high-resolution measurements of the transfer current at 3.5 GHz reveal that the transfer error rate is below the level ($\sim 10^{-3}$) that can be measured using a commercial current meter. In addition, we found theoretically that 1-GHz operation with an error rate of below 10^{-8} is possible.

4. Future work

The SE transfer via the SE island and that via the single trap level, which we have described in this article, have both advantages and disadvantages. For the trap-mediated transfer, sufficiently high-speed and high-accuracy operation is expected, but it is difficult to improve the device yield because the trap position is random. One promising way to solve this problem is to use impurity doping with a position-control technique. In contrast, the device yield of the island-mediated transfer should be better than that of the trap-mediated transfer, which is important for realizing universal standards. However, because the accuracy of the island-mediated transfer might be lower than that of the trap-mediated transfer, a temperature below 10 K is necessary. In the future, it will be necessary to evaluate the accuracy of both types of transfer at high speed by measuring the number of transferred electrons with a high-resolution charge sensor. In addition to evaluating the accuracy in more detail, we plan to conduct quantum metrology triangle experiments; we also want to apply the device to the current standards. Furthermore, we expect that the device will be able to be used for information processing based on SE manipulation. In this respect, one of the long-term goals is to achieve low-power consumption information processing by integrating the devices.

References

[1] A. Fujiwara, K. Nishiguchi, and Y. Ono, "Nanoampere Charge Pump by Single-electron Ratchet Using Silicon Nanowire Metal-oxide-semiconductor Field-effect Transistor," Appl. Phys. Lett., Vol. 92, No. 4, 042102, 2008.

[2] G. Yamahata, T. Karasawa, and A. Fujiwara, "Gigahertz Single-hole Transfer in Si Tunable-barrier Pumps," Appl. Phys. Lett., Vol. 106, No. 2, 023112, 2015.

[3] G. Yamahata, K. Nishiguchi, and A. Fujiwara, "Accuracy Evaluation and Mechanism Crossover of Single-electron Transfer in Si Tunable-barrier Turnstiles," Phys. Rev. B, Vol. 89, No. 16, 165302, 2014.

[4] G. Yamahata, K. Nishiguchi, and A. Fujiwara, "Gigahertz Single-trap Electron Pumps in Silicon," Nat. Commun., Vol. 5, 5038, 2014.
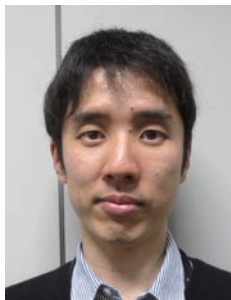
Gento Yamahata
Research Scientist, Nanodevices Research Group, Physical Science Laboratory, NTT Basic Research Laboratories.

He received his B.E., M.E., and Dr. Eng. in electrical engineering from Tokyo Institute of Technology in 2005, 2007, and 2009, respectively. In 2009–2010, he was a visiting researcher at Harvard University, MA, USA. He joined NTT Basic Research Laboratories in 2010. Since then, his research interests have focused on single-electron manipulation and detection in silicon nanometer-scale devices. He is a member of the Japan Society of Applied Physics (JSAP).

Katsuhiko Nishiguchi
Senior Research Scientist, Distinguished Researcher, Nanodevices Group, Physical Science Laboratory, NTT Basic Research Laboratories.

He received his B.E., M.E., and Ph.D. in electrical engineering from Tokyo Institute of Technology in 1998, 2000, and 2002, respectively. Since joining NTT Basic Research Laboratories in 2002, he has been researching the physics and technology of Si nanometer-scale devices for LSI applications with low power consumption and new functions. He was an invited researcher at the National Center for Scientific Research (CNRS), Lille, France, in September 2008 and also a guest researcher at Delft University of Technology, Delft, the Netherlands in 2012–2013. He received the IUPAP (International Union of Pure and Applied Physics) Young Author Best Paper Award at the International Conference on the Physics of Semiconductors 2000, the Graduate Student Award (Silver) at the Materials Research Society 2000 Fall Meeting, the Young Scientist Award at the JSAP Spring Meeting in 2000, the JSAP Outstanding Paper Award 2013, and the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology of Japan (Young Scientists' Prize) in 2013. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and JSAP.

Akira Fujiwara
Senior Distinguished Researcher, Senior Manager of Physical Science Laboratory and Group Leader of Nanodevices Research Group, NTT Basic Research Laboratories.

He received his B.S., M.S., and Ph.D. in applied physics from the University of Tokyo in 1989, 1991, and 1994, respectively. He joined NTT in 1994 and has been engaged in research on silicon nanostructures and their application to nanodevices and single-electron devices. He was a guest researcher at the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, in 2003–2004. He was a director of JSAP in 2010–2011 and a visiting professor at Hokkaido University in 2013. He received the International Conference on Solid State Devices and Materials (SSDM) Young Researcher Award in 1998, the SSDM Paper Award in 1999, and JJAP (Japanese Journal of Applied Physics) Paper Awards in 2003, 2006, and 2013. He was awarded the Young Scientist Award from the Ministry of Education, Culture, Sports, Science and Technology in 2006. He was supported by the funding program for Next Generation World-Leading Researchers (NEXT Program), JSPS (Japan Society of Promotion of Science) during 2011–2014. He is a member of JSAP and a senior member of IEEE.

Noise in Nanometer-scale Electronic Devices

Katsuhiko Nishiguchi and Akira Fujiwara

Abstract

Downsizing electronic devices in integrated circuits increases the noise-related degradation of circuit performance, and thus, it is becoming more important to analyze noise with single-electron resolution. We discuss here the use of a nanometer-scale transistor and capacitor in analyzing thermal noise, one of the most fundamental types of noise in electronic devices, with single-electron resolution. When the capacitor size is substantially reduced, the well-known model of thermal noise is no longer valid, and voltage noise is squeezed.

Keywords: noise, transistor, single electron

1. Introduction

Integrated circuits comprising silicon transistors and other electronic devices are extremely important components in our information technology society. The transistor, one of the most well-known electrical devices, controls the flow of electricity, that is, electrons, for data processing in the integrated circuits. The performance of such integrated circuits has advanced with the downsizing of transistors; for example, the 14-nm generation of transistors will be reached in 2015. However, downsizing also increases the noise-related degradation of circuit performance, and therefore, a microscopic analysis of noise and the development of countermeasures to such noise have become more important than ever before. Indeed, random telegraph noise originating from single-electron trapping in a transistor becomes prominent in small transistors and gives rise to errors in memory circuits. Shot noise originating from the discrete nature of an electron is also expected to become a serious issue in high-speed integrated circuits. Therefore, the downscaling of the transistors increases not only their performance but also the importance of understanding noise microscopically, ultimately with single-electron resolution, in small transistors as well as in other electronic devices.

In this article, we discuss the analysis of thermal noise, one of the most fundamental types of noise in electronic devices, with single-electron resolution using dynamic random access memory (DRAM) composed of a nanometer-scale transistor and capacitor [1]. Because the analysis was done by using a DRAM, one of the most common devices, we believe that our results are relevant to all electronic devices.

2. Techniques for analyzing thermal noise

One of the simplest and most well-known techniques for analyzing thermal noise is to measure voltage noise in a circuit composed of one resistor and one capacitor, as shown in **Fig. 1**. In this circuit, a massive number of electrons enter and exit the capacitor through the resistor in a random manner with the assistance of thermal energy. This random motion of electrons is thermal noise, and it exists absolutely at a finite temperature. When voltage noise at a node between the resistor and capacitor is measured with an oscilloscope or voltage meter, we can obtain certain information. For example, we can get a histogram of voltage-noise amplitude based on a Gaussian distribution, whose average is voltage V applied to the resistor. We can also get the variance V_{var}^2 of the distribution given by $k_B T/C$, where k_B is the Boltzmann's constant, T is temperature, and C is the capacitance of the capacitor. The fact that V_{var}^2 is proportional to T
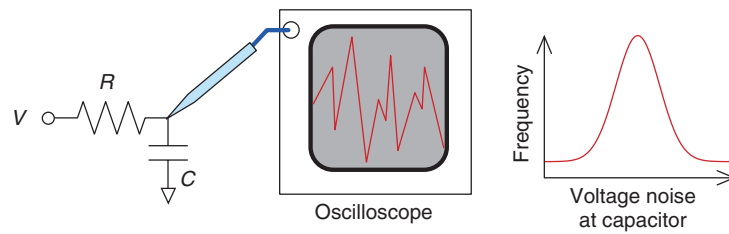
Fig. 1. Conventional technique to measure thermal noise.

(a) Equivalent circuit

(b) Scanning electron microscope image
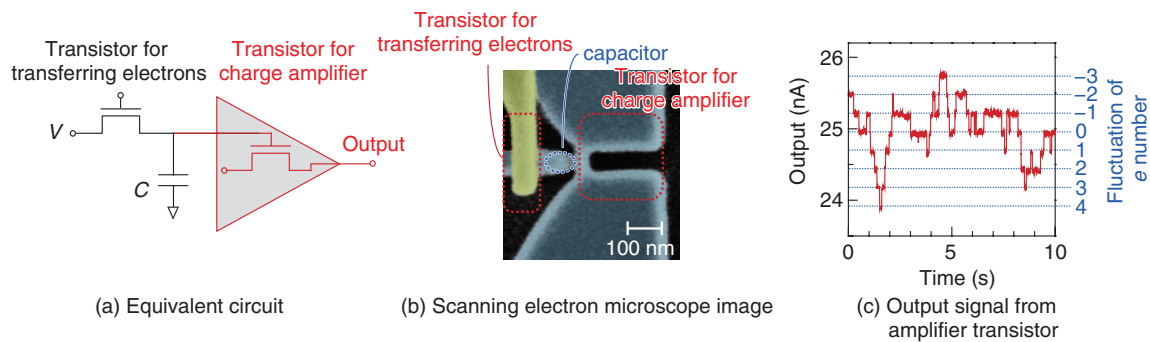
(c) Output signal from amplifier transistor

Fig. 2. Measurements of thermal noise with single-electron resolution.

means that the random motion of electrons originates from thermal energy. In contrast, V_{var}^2 increases as C decreases, regardless of the resistor. Therefore, the downsizing of electronic devices in circuits such as memory and analog circuits reduces C and thus increases noise, which leads to an undesirable hindrance of device downsizing. These features can be explained by the well-known Johnson-Nyquist model proposed in 1928 [2, 3].

In our technique for analyzing thermal noise with single-electron resolution, we use instead of a resistor a nanometer-scale transistor for transferring the electrons to the capacitor, as shown in **Fig. 2(a)**, because electrical control of the resistance of the transistor helps us to monitor the motion of individual electrons precisely in real time. In a common circuit, time intervals for the electrons to enter and exit the capacitor are too short for the electrons to be monitored: current I of 1 mA corresponds to the time interval of 1.6×10^{-16} ($=e/I$) seconds, which cannot be measured by any measurement system. However, increasing the resistance of the transistor when it transfers electrons prolongs the interval so that the electron motion can be monitored in real time. These mechanisms and the structure are the same as in a DRAM.

Since the charge of an electron is too small (1.6×10^{-19} C) to be detected by any charge sensor, the tiny signal is amplified by using another transistor [4]: The transistor for the charge amplifier has an extremely small channel (~10 nm) and is integrated with the capacitor, as shown in **Fig. 2(b)**. Consequently, the electrons in the capacitor can be detected precisely with single-electron resolution even at room temperature. This success is supported by well-established fabrication techniques for silicon transistors.

Using these features, we can monitor the random motion of individual electrons entering and exiting the capacitor through the transistor in real time, as shown in **Fig. 2(c)**: An output signal from the charge-amplifier transistor changes among discrete values, and one gap between these discrete values corresponds to the charge from one electron. To slow down the electron motion so that it can be monitored, the resistance of the transistor used for the electron transfer is adjusted to be around 10^{20} Ω. This extremely high resistance, which cannot be achieved in conventional resistors and transistors, is possible thanks to our high-quality nanometer-scale transistor.
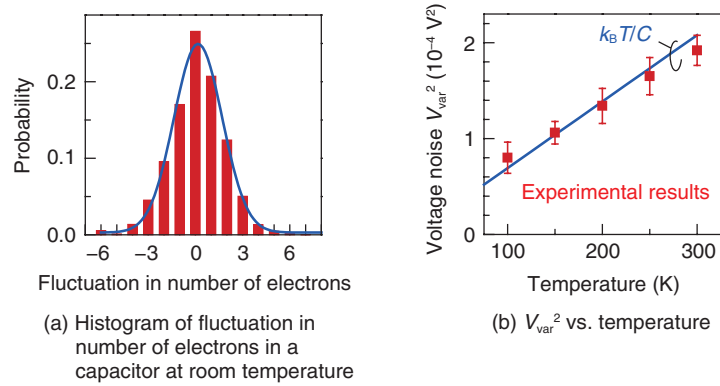
(a) Histogram of fluctuation in number of electrons in a capacitor at room temperature

(b) V_{var}^2 vs. temperature

Fig. 3.   Thermal noise with single-electron resolution.

(a) Histogram of fluctuation in number of electrons in capacitor at 60 K
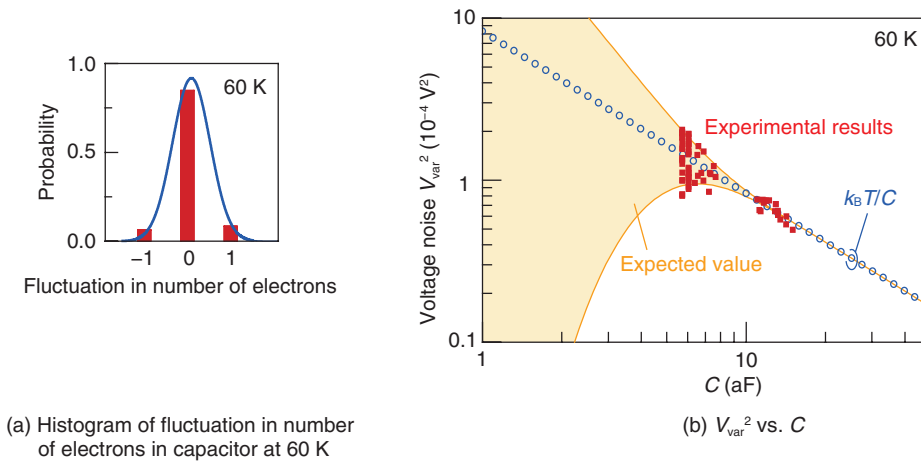
(b) V_{var}^2 vs. C

Fig. 4.   Thermal noise when charging energy is larger than thermal energy.

3.   Thermal noise with single-electron resolution

In analyzing the thermal noise with single-electron resolution, we first discuss the fluctuation, that is, the deviation from the average, of the number of electrons in the capacitor. Because single-electron injection into the capacitor increases the voltage of the capacitor by e/C ($= dV$), the fluctuation of the number of electrons in the capacitor can be converted into a voltage fluctuation, that is, voltage noise, at the capacitor. The dV can be evaluated from the change in the output signal originating from one electron, as shown in Fig. 2(c) [4].

A histogram of the number of electrons in the capacitor is shown in **Fig. 3(a)**. The histogram follows a Gaussian distribution (solid line), as in the case shown in Fig. 1. This histogram gives another

piece of information: dV^2 multiplied by the variance of the distribution shown in Fig. 3(a) corresponds to V_{var}^2, and this V_{var}^2 follows k_BT/C, as shown in **Fig. 3(b)**, which is the same as the case where thermal noise originates from the fluctuation of a massive number of electrons, as explained above and as shown in Fig. 1. Consequently, we can conclude that the well-known Johnson-Nyquist model for thermal noise is adaptable to the random motion of single electrons.

However, when charging energy E_C ($= e^2/2C$) for injecting one electron into the capacitor is larger than thermal energy k_BT, the Johnson-Nyquist model is no longer valid. A histogram of the number of electrons in the capacitor at $E_C > k_BT$ is shown in **Fig. 4(a)**. Since the distribution (solid line) becomes sharper, and the available number of electrons becomes

smaller than in the case of $E_C < k_B T$ (Fig. 3(a)), the discrepancy between the experimental results (bars) and expected values (solid line) becomes larger in the form of quantization errors in digital circuits. Indeed, V_{var}^2 (1.08×10^{-4} V^2) evaluated from the variance of the experimental distribution shown in Fig. 4(a) is smaller than $k_B T/C$ (1.38×10^{-4} V^2).

This deviation from the Johnson-Nyquist model is similar to the case of electromagnetic radiation from a black body. Energy of the electromagnetic radiation is quantized by multiples of $h\nu$, where h is Planck's constant and ν is the frequency of the electromagnetic radiation. Therefore, when a higher ν makes $h\nu$ larger than thermal energy $k_B T$, the thermal energy cannot assist the emission of electromagnetic waves from the black body. This idea, based on energy quantization proposed by Max Planck, overcomes the ultraviolet catastrophe in the low-frequency region of the spectrum of black-body radiation.

Our results for electrons can also be explained qualitatively by the same model as that for the black-body radiation. Energy for injecting N electrons into the capacitor is given by $(Ne)^2/2C = N^2 E_C$ and quantized because e is a unit charge. Therefore, as in the case of black-body radiation, when E_C is comparable or larger than $k_B T$, the thermal energy cannot assist electrons in entering and exiting the capacitor, which suppresses electron random motion and thus makes V_{var}^2 smaller than the thermal energy.

However, we can observe a unique feature in electron motion unlike in the case of black-body radiation. The average of the number of electrons in the capacitor is given by CV/e, where V is the voltage applied to the transistor as shown in Fig. 2(a). However, for example, when the average number is 0.5, the number of electrons in the capacitor switches between 0 and 1 with the same probability because an electron cannot be divided in half. As a result, the variance of the electron number becomes larger than 0.25, regardless of E_C and $k_B T$. In this sense, the fluctuation in the number of electrons depends on the average number of electrons. When the average number is an integer, the fluctuation becomes minimum; when a fractional part of the average number is 0.5, the fluctuation becomes maximum. The change in voltage noise V_{var}^2 at the capacitor as a function of C is shown in **Fig. 4(b)**. When $E_C < k_B T$ at larger C, V_{var}^2 follows $k_B T/C$ (open circles) and increases as C

decreases. When $E_C > k_B T$ at smaller C, V_{var}^2 deviates from $k_B T/C$, and the reduction in C increases the discrepancy between V_{var}^2 and $k_B T/C$, as indicated by the shaded area depicting the possible values of V_{var}^2. It should be noted that the minimum border of the shaded area in Fig. 4(b) decreases with C, behavior that is opposite to that of the conventional $k_B T C$-limited case depicted by the open circles. In the conventional case valid at $E_C < k_B T$, the reduction in C increases noise, which is a serious issue that hinders the downsizing of electronic devices. However, the case where $E > k_B T$ at much smaller C is preferable to the reduction in C due to noise reduction, which accelerates the downsizing of electronic devices.

4. Conclusion

As the downsizing of electronic devices accelerates, thermal noise will continue to deviate from its well-known behavior. However, this does not mean that any new phenomena will appear. Instead, it means that all of the results observed for single electrons follow well-established thermodynamics, for example, a Boltzmann distribution, and are thus valid for all electronic devices. Therefore, we believe that our results are important in the downsizing of devices to achieve a reduction in the number of electrons in the device. Indeed, using nanometer-scale transistors, we have developed memory and data information circuits that use one electron as one bit of information. In these circuits, errors caused by thermal noise represent one of the most serious issues. However, a wider view is that the minimum energy consumed for computation by data information circuits is governed by the thermal energy. Therefore, the analysis of thermal noise is very important for realizing low-power-consumption circuits.

References

[1] K. Nishiguchi, Y. Ono, and A. Fujiwara, "Single-electron Thermal Noise," Nanotechnology, Vol. 25, No. 27, p. 275201, 2014.
[2] J. B. Johnson, "Thermal Agitation of Electricity in Conductors," Phys. Rev., Vol. 32, No. 1, p. 97, 1928.
[3] H. Nyquist, "Thermal Agitation of Electric Charge in Conductor," Phys. Rev., Vol. 32, No. 1, p. 110, 1928.
[4] K. Nishiguchi, C. Koechlin, Y. Ono, A. Fujiwara, H. Inokawa, and H. Yamaguchi, "Single-electron-resolution Electrometer Based on Field-effect Transistor," Jpn. J. Appl. Phys., Vol. 47, No. 11R, p. 8305, 2008.

Katsuhiko Nishiguchi

Senior Research Scientist, Distinguished Researcher, Nanodevices Group, Physical Science Laboratory, NTT Basic Research Laboratories.

He received his B.E., M.E., and Ph.D. in electrical engineering from Tokyo Institute of Technology in 1998, 2000, and 2002, respectively. Since joining NTT Basic Research Laboratories in 2002, he has been engaged in research on the physics and technology of Si nanometer-scale devices for LSI applications with low power consumption and new functions. He was an invited researcher at the National Center for Scientific Research (CNRS) Lille, France, in September 2008 and also a guest researcher at Delft University of Technology, Delft, the Netherlands in 2012–2013. He received the IUPAP (International Union of Pure and Applied Physics) Young Author Best Paper Award at the International Conference on the Physics of Semiconductors 2000, the Graduate Student Award (Silver) at the Materials Research Society 2000 Fall Meeting, the Young Scientist Award at the Japan Society of Applied Physics (JSAP) Spring Meeting in 2000, the JSAP Outstanding Paper Award 2013, and the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology of Japan (Young Scientists' Prize) in 2013. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and JSAP.

Akira Fujiwara

Senior Distinguished Researcher, Senior Manager of Physical Science Laboratory and Group Leader of Nanodevices Research Group, NTT Basic Research Laboratories.

He received his B.S., M.S., and Ph.D. in applied physics from the University of Tokyo in 1989, 1991, and 1994, respectively. He joined NTT in 1994 and has been engaged in research on silicon nanostructures and their application to nanodevices and single-electron devices. He was a guest researcher at the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, in 2003–2004. He was a director of JSAP in 2010–2011 and a visiting professor at Hokkaido University in 2013. He received the International Conference on Solid State Devices and Materials (SSDM) Young Researcher Award in 1998, the SSDM Paper Award in 1999, and JJAP (Japanese Journal of Applied Physics) Paper Awards in 2003, 2006, and 2013. He was awarded the Young Scientist Award from the Ministry of Education, Culture, Sports, Science and Technology in 2006. He was supported by the funding program for Next Generation World-Leading Researchers (NEXT Program), JSPS (Japan Society of Promotion of Science) during 2011–2014. He is a member of JSAP and a senior member of IEEE.

Gate Tuning of Direct/Indirect Optical Transitions in Silicon

Jinichiro Noborisaka, Katsuhiko Nishiguchi, and Akira Fujiwara

Abstract

Silicon is one of the most important semiconductor materials in microelectronics and is known as a typical indirect bandgap semiconductor. In these semiconductors, the minimal energy state in the conduction band and the maximal energy state in the valence band appear at different momenta. This makes it difficult to obtain efficient light emission. In this article, we present a means of electrically tuning the direct/indirect optical transitions using a specially prepared Si/SiO$_2$ interface, where anomalously large valley splitting appears. This tunability is achieved by utilizing the close relationship between valley splitting and a direct optical transition in silicon and its proportional relationship with respect to a gate electric field.

Keywords: quantum well, valley splitting, light-emitting devices

1. Introduction

The continued miniaturization of silicon integrated circuits has considerably improved the performance of electronic information processors including computers and cell phones over the last decades. Similarly, optical communication technologies have improved significantly in both communication speed and capacity thanks to the development of compact and high-performance solid laser sources using compound semiconductors such as GaAs and InP. However, the recent remarkable increase in the amount of information requires further improvement in the performance of information processors. This is expected to involve the integration of optical components with electronic integrated circuits in order to meet the challenges presented by the limited bandwidth and power consumption of conventional metal lines used in processors and interconnects.

Silicon photonics is a promising technology that enables both high speed and mass-information processing with low power consumption. By combining high-performance photonic building blocks such as waveguides, resonators, and modulators fabricated in silicon with complementary metal-oxide semicon-ductor (CMOS) technologies on the same chip, some companies have been able to release commercial products that achieve ultrahigh speed optical communications. However, the light source of such systems is made from exotic semiconductors (GaAs, InP) that are incompatible with silicon processing. This is because implementing a silicon-based light source is still a significant challenge due to its indirect bandgap nature, which is a fundamental property of bulk silicon.

Therefore, laser components in the system are independently fabricated using an exotic semiconductor and then assembled in a silicon photonics system, which makes it difficult to reduce manufacturing costs. Having a silicon-based light source would drastically reduce costs and also increase performance thanks to the massive investment and existing knowledge of silicon CMOS technologies over the past decades, and therefore, the development of an efficient silicon-based light source is long-awaited.

The study of silicon-based light sources goes back a long way. In the 1990s, many reports were published on methods to make silicon emit light using low-dimensional structures such as porous silicon and nanocrystals [1], where electrons are typically

(a) Band structures of bulk silicon: Dispersion relationship between energy and wave vector. (Upper) Equi-energy surface of CB in momentum space

(b) Valley splitting due to quantum confinement and eigenfunctions in the systems
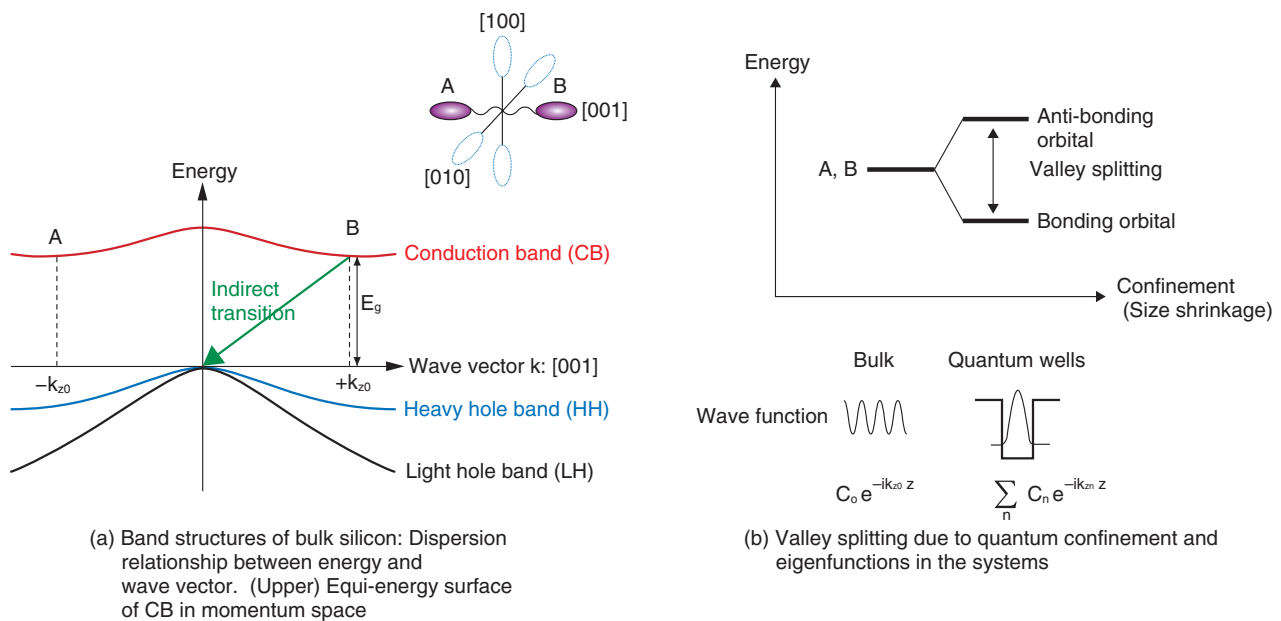
Fig. 1.   Band structures of silicon.

confined in a small space on the scale of a few nanometers. Most fundamental ideas on making silicon emit light are identical and involve confining an electron in a small space. This makes the electron's wave function sharp in real space, and thus, a wave function is dispersed in momentum space, which in turn leads to a direct optical transition in silicon, which is known as a quasi-direct optical transition. In both porous silicon and nanocrystals, strong light emission can be seen even with the naked human eye due to the strong electron confinement. However, an ideal light source for silicon photonics requires certain properties: high stability in both wavelength and power, adjustability in emission wavelength, compatibility with the present silicon manufacturing process, and high efficiency in light emission. However, these silicon light sources have some problems in that they are physically unstable and difficult to inject current into due to their extremely small size and low-dimensional structure. It is also difficult to adjust the emission wavelength because of a lack of controllable parameters.

We are working to solve these problems by applying electric tuning of direct/indirect optical transitions using a specially prepared Si/SiO$_2$ interface, where anomalously large valley splitting appears. This electric tuning is a proof of concept demonstration that is based on the close relationship between valley splitting and a direct optical transition in sili-

con and its proportional relationship with respect to a gate electric field. Our silicon light-emitting devices are driven by current injection and fabricated using fully CMOS-compatible processes. They exhibit remarkably stable and reproducible optical characteristics thanks to the use of well-established silicon CMOS processes.

2.   Electronic band and valley structures in silicon

Electronic band structures in semiconductors are uniquely determined by the constituent elements of the lattice. Silicon has a diamond-type structure as well as an indirect bandgap, where the minimal energy states of the conduction band and the maximal energy state of the valence band are misaligned in momentum space. (Hereafter in this article, we regard a momentum as having the same meaning as wavenumber (k) because it has a linear relationship with the wavenumber.) Because the electron's minimal energy state in the conduction band appears on [001] momentum axes near the X-point and there are three mutual axes, silicon has in total six energy-equivalent minima in its conduction band (**Fig. 1(a)**). These energy minima are called valleys because the states are first occupied by electrons with an increasing density, as if rain were pooling in valleys. The electrons in each valley can occupy the same space thanks

(a) Cross-sectional view of silicon light-emitting device

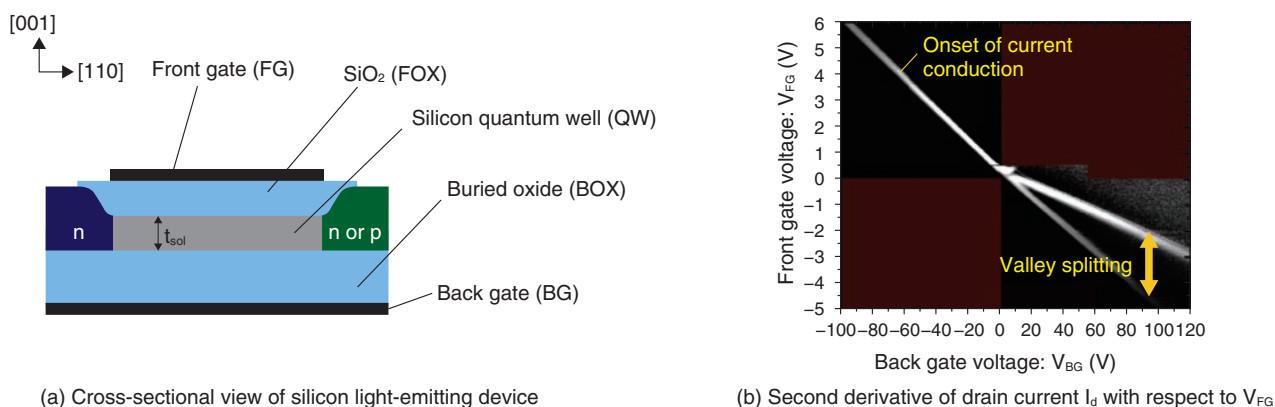(b) Second derivative of drain current I_d with respect to V_{FG}

Fig. 2.   Electrical measurement of valley splitting.

to their different momenta even if they have the same energy (they are called degenerate). Since bulk crystal has a translational symmetry, the eigenstate of the electron in the system can be explained by a plane wave with a specific wavenumber ($e^{-ik0 \cdot r}$), where the wavenumber k_0 gives the minimal energy of the conduction band in the dispersion relation. This is known as the Bloch theorem.

Suppose that spatial confinement normal to that of the [001] axes is introduced, and a quantum well (QW) is formed. Then the six-fold degenerate states split into two- and four-fold degenerate ones due to the anisotropy of electron effective mass, and accordingly, two valleys (labeled A and B in Fig. 1) lying along the confining axis become the ground states because they have heavier effective mass along the [001] axis. The eigenfunctions in the system are not a plane wave but a wave packet confined in the quantum well, as shown in **Fig. 1(b)**. If we perform Fourier transform with respect to the wave packet, it becomes clear that the electrons are largely scattered in momentum space due to the confinement because it involves higher components of a Fourier coefficient. As the degree of confinement increases, further splitting (called valley splitting) occurs, where the two-fold degenerate valley states split, and thus, degeneracy is fully lifted. This splitting can happen when the Fourier coefficient that has a wave number that is nearly equal to the inverse value of the lattice constant ($k_{z0} \sim 2\pi/a_0$) is finite. The resulting coupled states are called bonding and anti-bonding states, where the eigenfunctions are hybridized states of the $+k_{z0}$ and $-k_{z0}$ valley wave functions.

The valley splitting is tuned by the size of the QW, electric fields, or magnetic fields. In silicon metal-

oxide semiconductor field effect transistors (MOS-FETs), it can be tuned by gate electric fields. The typical size of valley splitting in general MOSFETs with gate oxides formed by the standard thermal oxidation process is at most several milli-electron volts, whereas NTT Basic Research Laboratories has reported large valley splitting up to several tens of milli-electron volts. The large valley splitting is observed in MOSFETs fabricated on a separation-by implantation-of-oxygen (SIMOX) wafer employing high-temperature thermal treatment [2], where SIMOX is one of the fabrication methods for a silicon on insulator (SOI), and a top silicon layer is separated by a buried oxide (BOX) formed by ion implantations of oxygen.

The devices used in the present work are thin SOI-MOSFETs with a double-layer gate fabricated on a [001] SIMOX wafer employing high-temperature thermal treatment (**Fig. 2**). The front gate (FG) is constructed using polycrystalline silicon with a gate oxide (FOX) formed by conventional thermal oxidation. The substrate is used as a back gate (BG). Both n- and p-type contacts are formed to inject the carriers into the QW sandwiched between the FOX and the BOX. Both FG and BG voltages (V_{FG}, V_{BG}) are applied in order to control the distribution of both electrons and holes in the QW.

When positive V_{FG} and negative V_{BG} are applied, electrons are squeezed into the FOX interface. Since the size of valley splitting is quite small at the interface, only the onset of current conduction can be detected as a white line in Fig. 2(b). In contrast, when the polarities of both gates are mutually replaced, electrons are squeezed into the BOX interface, and the white line gradually splits. This splitting is just

the valley splitting that we discuss in the following sections. Although the physical origin of the large valley splitting remains unclear and is still under debate, the notable points here are that the electrons' wave functions are strongly dispersed in momentum space when the electrons are distributed at the BOX interface, and the size of valley splitting is tunable by gate electric fields.

3. Relationship between valley splitting and dipole transitions

Next, we examine an application of large valley splitting to light emission for indirect semiconductors. Light emission is known to occur when an electron relaxes from an initial higher energy state to a final lower one and loses its energy as a photon. Since photons do not carry a significant momentum, and both the energy and momentum are preserved in any physical process, the momentum of an electron in the initial higher state is preserved in pure direct optical transition. In contrast, in bulk silicon, the minimal energy states of the conduction and valence bands have different respective momenta. This means that such a transition rarely occurs in bulk silicon. Despite this, weak light emission is observed in bulk silicon thanks to the aid of phonons. This process, which is called an indirect optical transition, can be completed only when an electron absorbs or emits a phonon with the momentum k_{z0} and also emits a photon with the energy corresponding to the value of the phonon energy ($\hbar\omega$) subtracted from the bandgap of silicon (E_g). Although this process occurs more frequently than the pure direct optical transition in indirect semiconductors, another non-radiative process such as Auger or free-carrier absorption quickly recombines an electron, and therefore, the photon emission efficiency is quite low in indirect semiconductors. This is why silicon is inappropriate for light-emitting devices.

With finite valley splitting, the wave function of electrons in the QW is composed of a broad range of components for a Fourier coefficient, as discussed in the previous section. If the amplitude of $2k_{z0}$ components does not have a zero value, then a direct optical transition can occur. This is because the variation of the confining potential formed by the gate electric field is much gentler than that at a steep Si/SiO$_2$ interface, so the wave function confined by the gate electric field has a smooth profile. The profile of such a smooth wave function gently varies as the gate electric field increases, where the amplitude of the k_{z0}

component follows that of $2k_{z0}$. Therefore, it is expected that large valley splitting will lead to significant direct optical transitions.

4. Electric tuning of direct/indirect optical transitions

We then examined light emission using finite valley splitting. In this experiment, light emission was achieved by electroluminescence (EL), where both carriers were electrons and holes and were injected into the channel. We used n- and p-type contacts as a respective source and drain. We applied forward bias and injected carriers into the QW; then we tuned both gate biases V_{FG} and V_{BG} so as to control the distribution of the electron and hole wave functions. EL spectra as a function of V_{BG} are shown in **Fig. 3(a)**, where the sizes of valley splitting are adjusted by tuning both gates (FG and BG). When we applied positive V_{FG} and negative V_{BG}, a transverse optical (TO) phonon-mediated indirect optical transition dominated the spectra. This transition is the main radiative recombination process in bulk silicon. In contrast, when the polarities of both gates were mutually replaced, a higher energy peak gradually developed with an increasing V_{BG}. The energy separation between TO and the higher energy peak are approximately 59 meV. The energy separation agrees with the energy of the TO phonon with a momentum at valley minimum (k_{z0}). When we take into consideration these peak energies with the energy separation, the higher energy peak is assigned as a non-phonon (NP) direct optical transition, where the electron in the conduction band minima directly recombines with the hole in the valence band without aid of the phonon. The intensities of the NP peak increase with an increasing $|V_{BG}|$ and they greatly increase for a positive V_{BG}. Eventually, the intensity of the NP peak becomes larger than that of the TO phonon-mediated indirect optical transition.

The efficiency of light emission is plotted in **Fig. 3(b)** taking into account the current flowing in the QW that becomes small as valley splitting increases. We assumed a condition close to the bulk silicon by applying zero gate voltages to both the FG and BG, and using a thicker QW. We define this as the bulk condition. The efficiency of the NP direct optical transition is 16 times better than that of the TO phonon-mediated indirect optical transition in the bulk condition. Weak NP direct optical transitions are known to occur even in bulk silicon thanks to an unintentionally doped slight amount of impurities that

(a) Electroluminescence (EL) spectra

(b) Emission efficiency (top) and the ratio of direct to indirect transition (NP/TO) (bottom)
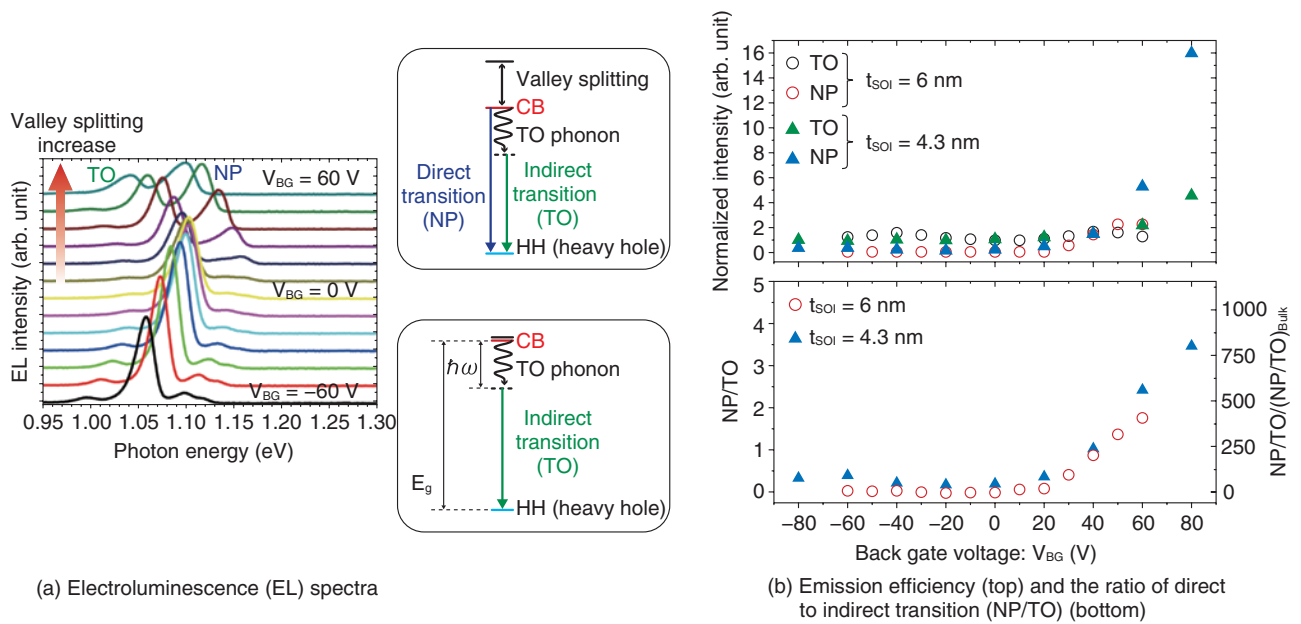
Fig. 3.   Optical measurement.

bind electrons around them and disperse the electrons' wave function. We have found that the optimal efficiency of the NP optical transition was 800 times greater than the NP direct optical transition in the bulk condition [3]. Furthermore, the ratio of NP/TO is tuned by the gate electric field. This is the world's first demonstration of electric tuning of direct/indirect optical transition in silicon by controlling the size of valley splitting at the Si/SiO$_2$ interface.

5.   Conclusion

We reported electric tuning of direct optical transitions up to 800 times greater than that in bulk silicon. However, the efficiency of such a direct optical transition is much weaker than that in GaAs or InP, and of course, this is not sufficient for implementing silicon light-emitting devices. Therefore, we will examine hybrid approaches that combine elements such as impurities and photonic crystals in order to improve the light emission efficiency.

References

[1]   L. T. Canham, "Silicon Quantum Wire Array Fabrication by Electrochemical and Chemical Dissolution of Wafers," Appl. Phys. Lett., Vol. 57, 1046, 1990.

[2]   K. Takashina, Y. Ono, A. Fujiwara, Y. Takahashi, and Y. Hirayama, "Valley Polarization in Si(100) at Zero Magnetic Field," Phys. Rev. Lett., Vol. 96, 236801, 2006.

[3]   J. Noborisaka, K. Nishiguchi, and A. Fujiwara, "Electric Tuning of Direct-indirect Optical Transitions in Silicon," Sci. Rep., Vol. 4, 6950, 2014.
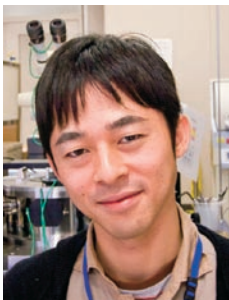
Jinichiro Noborisaka
Lead Researcher, Nanodevices Research Group, Physical Science Laboratory, NTT Basic Research Laboratories.
He received his M.S. and Ph.D. in engineering from Hokkaido University in 2005 and 2008, respectively. He joined NTT Basic Research Laboratories in 2008 and has been studying nanoscale devices and low-dimensional semiconductor physics. He is a member of the Japan Society of Applied Physics (JSAP).

Akira Fujiwara
Senior Distinguished Researcher, Senior Manager of Physical Science Laboratory and Group Leader of Nanodevices Research Group, NTT Basic Research Laboratories.
He received his B.S., M.S., and Ph.D. in applied physics from the University of Tokyo in 1989, 1991, and 1994, respectively. He joined NTT in 1994 and has been engaged in research on silicon nanostructures and their application to nanodevices and single-electron devices. He was a guest researcher at the National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA, in 2003–2004. He was a director of JSAP in 2010–2011 and a visiting professor at Hokkaido University in 2013. He received the International Conference on Solid State Devices and Materials (SSDM) Young Researcher Award in 1998, SSDM Paper Award in 1999, and JJAP (Japanese Journal of Applied Physics) Paper Awards in 2003, 2006, and 2013. He was awarded the Young Scientist Award from the Ministry of Education, Culture, Sports, Science and Technology in 2006. He was supported by the funding program for Next Generation World-Leading Researchers (NEXT Program), JSPS (Japan Society of Promotion of Science) during 2011–2014. He is a member of JSAP and a senior member of IEEE.

Katsuhiko Nishiguchi
Senior Research Scientist, Distinguished Researcher, Nanodevices Group, Physical Science Laboratory, NTT Basic Research Laboratories.
He received his B.E., M.E., and Ph.D. in electrical engineering from Tokyo Institute of Technology in 1998, 2000, and 2002, respectively. Since joining NTT Basic Research Laboratories in 2002, he has been engaged in research on the physics and technology of Si nanometer-scale devices for LSI applications with low power consumption and new functions. He was an invited researcher at the National Center for Scientific Research (CNRS) Lille, France, in September 2008 and also a guest researcher at Delft University of Technology, Delft, the Netherlands in 2012–2013. He received the IUPAP (International Union of Pure and Applied Physics) Young Author Best Paper Award at the International Conference on the Physics of Semiconductors 2000, the Graduate Student Award (Silver) at the Materials Research Society 2000 Fall Meeting, the Young Scientist Award at the JSAP Spring Meeting in 2000, the JSAP Outstanding Paper Award 2013, and the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology of Japan (Young Scientists' Prize) in 2013. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and JSAP.

Microscopic Probing of Crystalline Electrons Using Magnetic Resonance

Koji Muraki and Trevor David Rhone

Abstract

Electrons in semiconductors are often referred to as free electrons, meaning that they can move freely. They are also spatially extended like a wave. On the other hand, electrons are negatively charged particles and thus repel each other. Theory predicts that at low temperatures and in high magnetic fields, electrons—in an effort to maintain a distance from one another—organize themselves. They form a regular array like the atoms in a crystal, a state known as the Wigner crystal. This article describes experiments that exploit nuclear magnetic resonance (NMR) to probe the microscopic structure of a Wigner crystal. These experiments demonstrate the capability of NMR to resolve the spatial variation of electron waves on the nanometer scale.

Keywords: two-dimensional electron system, low-temperature physics, nuclear magnetic resonance

1. Two-dimensional electron system in a semiconductor

Metal-oxide-semiconductor field-effect transistors, which form the basis of today's electronics, operate by controlling the motion of charge carriers (electrons, holes) induced at the interface between silicon (a semiconductor) and silicon dioxide (an insulator). The electrons (or holes) carry electrical current by moving along the interface in response to the voltage applied between the source and drain electrodes. Since these charge carriers are confined near the interface by the electric field perpendicular to it, they can only move in the plane of the interface. Such a system is often referred to as a two-dimensional electron gas. As the term *gas* suggests, electrons in semiconductors can be considered to move independently between inter-particle collisions. This situation is analogous to that of electrons in a vacuum tube traveling from the cathode to the anode without collisions. However, in the solid state, the mobility of electrons depends greatly on the material and the concentration of impurities. Mobility is a quantity that describes the ability of electrons to move through a medium in response to an electric field.

As a result of the continual miniaturization of electronic devices, the distance electrons travel in a transistor (the gate length) is now on the order of tens of nanometers. Research on new materials with high carrier mobility (for example, graphene or silicon germanium) is under way. This is aimed at achieving high-speed operation of electronic devices—a situation that draws parallels to Olympic sprinters on the 100-meter track.

2. Two-dimensional electron system at low temperature and high magnetic field

At room temperature, where electronic devices operate, the mobility of electrons is governed primarily by lattice vibrations. At low temperatures, reduced lattice vibrations generally lead to an increase in mobility. A heterostructure comprising GaAs and $Al_xGa_{1-x}As$ is a system in which the highest electron mobility is attained at low temperatures. The electron mobility of the sample used in this study reaches 10^7 cm^2/Vs at a cryogenic temperature. This implies that the mean free path of electrons is as long as 100 μm. Such ultrahigh mobility requires the purity of the host semiconductor crystal to be exceedingly high; the residual impurity concentration in a material with 10^7 cm^2/Vs electron mobility is about 10^{14} cm^{-3}

(~2 ppb).

Under conditions of high material purity and low temperature, electrons are able to move freely without being scattered by impurities or lattice vibrations. On the other hand, electrons are negatively charged particles and thus repel each other. Nevertheless, since electrons occupy low-energy states with long wavelengths, they tend to be spatially spread, so the repulsive interaction does not usually produce visible effects.

In contrast, when a two-dimensional electron system is subjected to a strong perpendicular magnetic field, under the influence of the Lorentz force, the electrons make a circular motion (cyclotron motion) with the radius becoming even smaller than the inter-electron distance (for a sufficiently strong field). It is widely believed that in such cases, electrons—in an effort to maintain a distance from one other—organize themselves into a regular array, like the atoms in a crystal. Such a state is referred to as a Wigner crystal, named after the physicist who first predicted it. Thus far, experimental studies on Wigner crystals have relied on rather indirect information such as microwave absorption.

3. Nuclear magnetic resonance (NMR)

NMR is a phenomenon in which nuclei in a magnetic field absorb electromagnetic radiation at a frequency specific to the nuclear species. It is widely used as a highly sophisticated method of spectroscopic analysis in various fields including chemistry, physics, and the medical sciences. Its principle is based on the property of nuclei (referred to as nuclear spin) that make them behave like tiny magnets. The resonance occurs at a frequency determined by the magnetic field strength and the nuclear species. In addition, the resonance frequency undergoes a tiny shift whose magnitude reflects the details of the chemical bonds and the state of the electrons surrounding the nuclei. Useful information can be obtained by measuring this tiny frequency shift (Knight shift). The reason the resonance of nuclei is affected by the surrounding electrons is that electrons also have the property (referred to as electron spin, or simply, spin) that makes them behave like tiny magnets. Electron spin can point up or down with respect to the external magnetic field. When the numbers of up-spin and down-spin electrons are unequal, the imbalance acts on the nuclei as an effective magnetic field, which shifts their resonance frequency. In NMR, measuring this Knight shift allows one to obtain information on the electrons in contact with the nuclear spins.

4. Resistively detected NMR

In standard NMR, a pick-up coil is used to detect signals produced by nuclear spins processing in an external magnetic field. However, the low sensitivity of the inductive detection makes the application of the standard NMR technique to a single-layer two-dimensional electron system unfeasible. Resistively detected NMR (RD-NMR) provides a way to overcome the issues of sensitivity, thereby enabling us to perform NMR on a two-dimensional electron system [1]. In RD-NMR, we measure the change in the electrical resistance of the sample that occurs when the frequency of the applied radio-frequency (rf) wave matches the resonance frequency of the nuclei. It also allows us to selectively detect the signal from those nuclear spins in contact with the two-dimensional electron system.

The structure of the $GaAs/Al_xGa_{1-x}As$ ($x = 0.25$) quantum well sample used in this study is schematically shown in **Fig. 1(a)**. The sample was grown by molecular beam epitaxy. Electrons are supplied from the Si delta doping layer in the upper $Al_xGa_{1-x}As$ barrier to the GaAs quantum well. The electron density n_e is varied using the n^+-GaAs substrate as a back gate. The measurement setup is shown in **Fig. 1(b)**. The sample was processed into a 100-µm-wide Hall bar, mounted on a chip carrier, wrapped with a coil used for applying rf radiation, and cooled to 0.01 K in a dilution refrigerator. The resistance of the sample was measured using a lock-in technique with alternating current (72 Hz) in a magnetic field applied perpendicular to the sample.

The energy level diagram of electrons in a magnetic field is schematically shown in **Fig. 2**. When a two-dimensional electron system is subjected to a strong perpendicular magnetic field B, the cyclotron motion of electrons is quantized, and their energy spectrum splits into a set of equally spaced discrete levels (Landau levels) designated by the orbital quantum number N (= 0, 1, 2, ...). Each Landau level is further split into spin-up (↑) and spin-down (↓) levels. The number of electrons that each spin-split Landau level can accommodate per unit area is given by $n_\varphi = eB/h$, where h is Planck's constant and e is the elementary charge. The filling factor is a quantity that represents the number of occupied (spin-split) Landau levels, and is given by $\nu = n_e/n_\varphi = nh/eB$. As the figure shows, a filling factor of 2.2 (2.1) means that
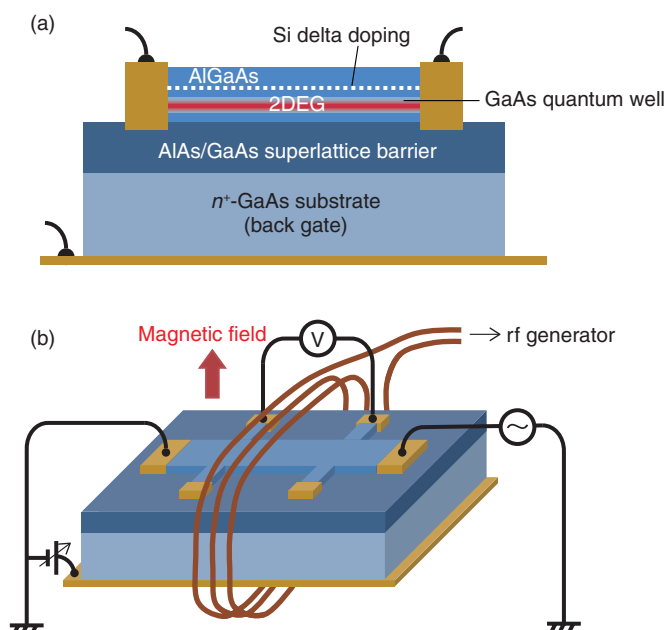
Fig. 1.   (a) Cross-sectional view of sample and (b) measurement setup for resistively detected NMR.
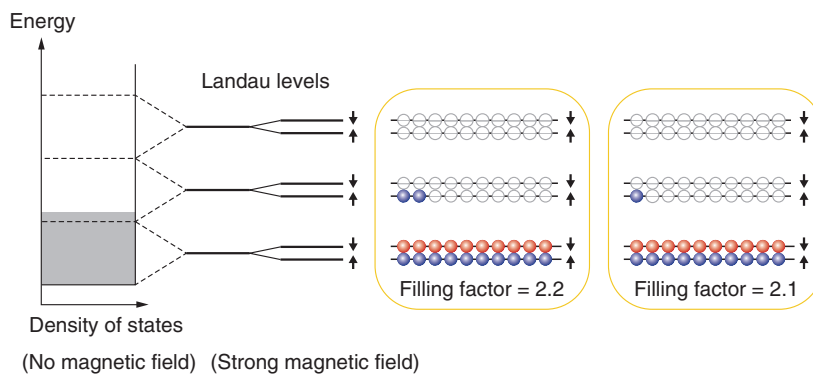

Fig. 2.   Energy level diagram of two-dimensional electron system in a magnetic field.

both spin branches of the lowest Landau level are fully occupied, and the lower spin branch of the first excited Landau level is 20% (10%) occupied.

The resonance spectra of ^{75}As taken at a magnetic field of 6.4 T are shown in **Fig. 3(a)**. For these measurements, the filling factor is tuned in the range from 2.20 to 2.33 using the back gate. The dashed line in the figure indicates the resonance frequency of ^{75}As nuclei observed when the quantum well is depleted of electrons. The NMR spectra taken in the presence of electrons are shifted to lower frequencies. The asymmetric line shape of the spectra with a tail on the high-

frequency side reflects the variation of the probability density of electrons along the direction normal to the quantum well (**Fig. 3(b)**). That is, ^{75}As nuclear spins near the center of the GaAs well and those near the $Al_xGa_{1-x}As$ barriers experience different effective magnetic fields from the electrons. Simulations that take into account the variation of the electron probability density along the direction normal to the quantum well, shown by the solid lines in Fig. 3(a), reproduce the experimentally observed spectra.

The resonance spectrum of ^{75}As taken at the same magnetic field of 6.4 T, but with the filling factor
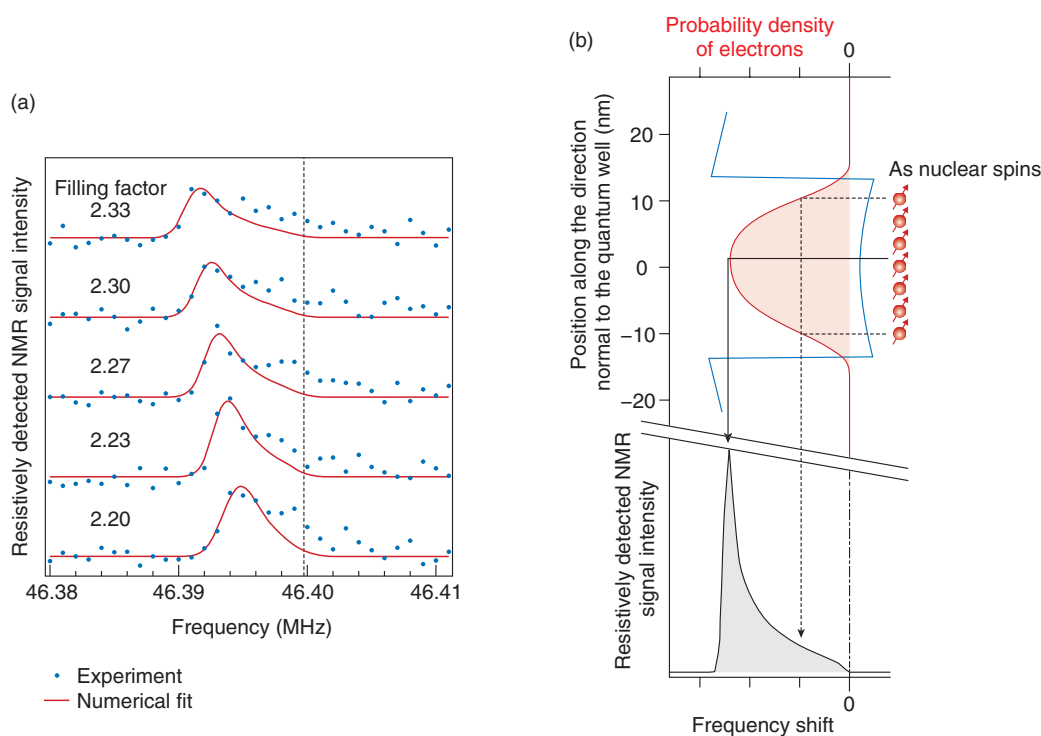
Fig. 3. (a) Resistively detected NMR spectra of ^{75}As at filling factor 2.20–2.33 and (b) effects of finite thickness of two-dimensional electron system on the NMR lineshape.

reduced to 2.1, is shown in **Fig. 4(a)**. (Here, the horizontal axis is taken to be the frequency shift with respect to the zero Knight shift bare resonance frequency.) Obviously, the measured spectrum does not match the simulation shown in **Fig. 4(b)**, which is based on the assumption that the electrons are uniformly distributed in the plane of the quantum well. In contrast, simulations carried out assuming that the electrons in the first excited Landau level form a Wigner crystal (solid line in Fig. 4(a)) reproduce the experimentally measured spectrum very well.

The spatial distribution of the probability density of electrons used for the above simulation is shown in **Fig. 4(c)**. Each of the mounds arranged in a triangular lattice corresponds to a single electron. As this figure shows, at a filling factor of 2.1, the inter-electron distance becomes greater than the spatial extent of the electron wave. Consequently, it becomes energetically more favorable for the electrons to be in the crystalline state. When a Wigner crystal is formed, the probability density of electrons is no longer uniform in the plane of the quantum well, as **Fig. 4(d)** shows. Regions appear where the local probability density is higher or lower than that expected for the

homogeneous system. In the spectrum shown in Fig. 4(a), the tail on the low-frequency side corresponds to the top of the mounds where the electron probability density is highest and the peak on the high-frequency side to the valleys where the probability density is lowest. The dip at the center of the mounds in Fig. 4(c) is a characteristic feature of the first excited Landau level wave function. The line shape of the measured NMR spectrum captures such details of the microscopic structure of the electron waves.

5. Conclusion

The results presented here demonstrate that NMR is a powerful means for probing the spatial variation of electron waves in a semiconductor. In particular, it offers a significant advantage of spatial resolution provided by nuclear spins, allowing one to obtain microscopic information about the electron wave on the nanometer scale. This opens the way to studies of other exotic phases with nontrivial spatial order. It also enables us to quantify the spatial inhomogeneity of electron density induced by randomly distributed impurities. This will be a useful technique for
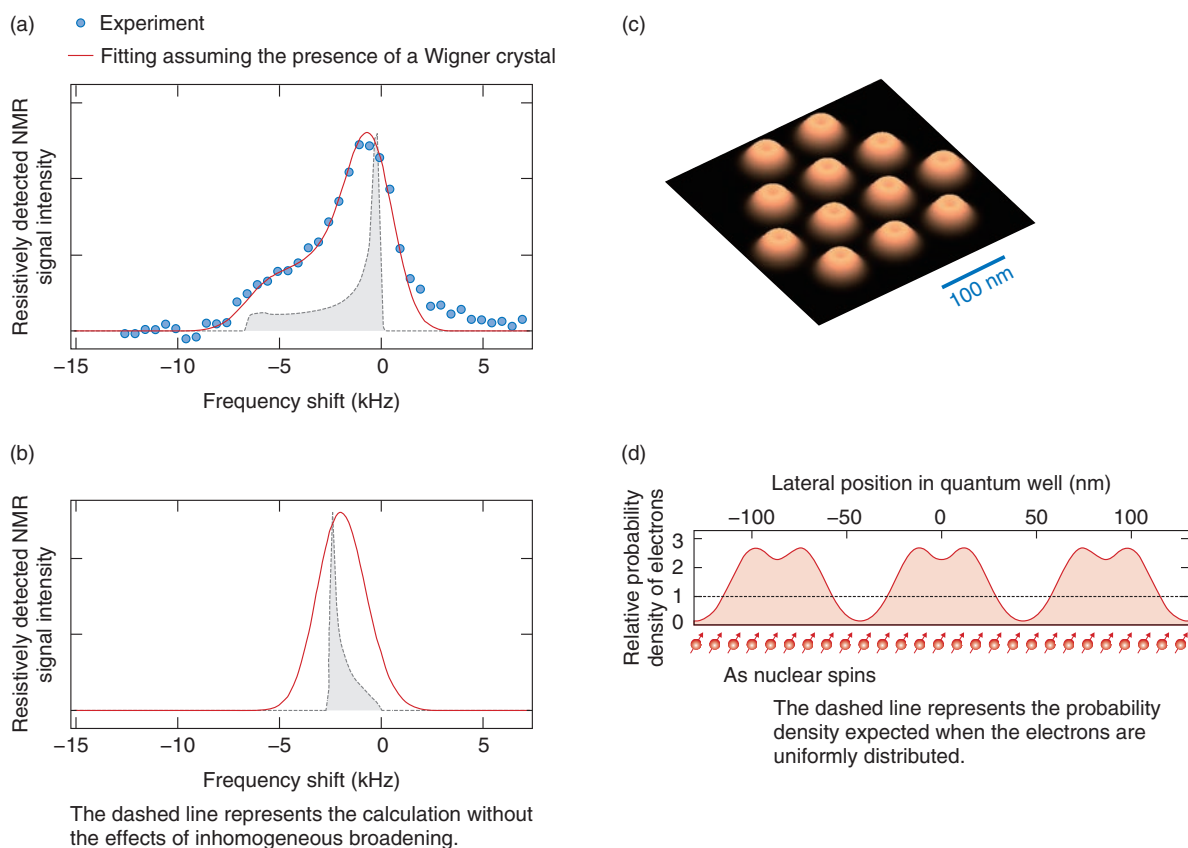
Fig. 4. (a) Resistively detected NMR spectra of ^{75}As at filling factor 2.1, (b) simulation assuming that electrons are distributed uniformly in the two-dimensional plane, (c) probability density of electrons in a crystalline state (simulation), and (d) profile of probability density of electrons in a crystalline state.

characterizing electronic devices facing nanometer-scale miniaturization, where eliminating device-to-device variability is a critical issue.

These experiments were performed as a joint research effort between NTT and the Japan Science and Technology Agency (JST).

Reference

[1] L. Tiemann, T. D. Rhone, N. Shibata, and K. Muraki, "NMR Profiling of Quantum Electron Solids in High Magnetic Fields," Nature Physics, Vol. 10, No. 9, pp. 648–652, 2014.

Koji Muraki
Senior Research Scientist (Distinguished Researcher) and Group Leader of the Quantum Solid State Physics Research Group, NTT Basic Research Laboratories.
He received his B.E., M.E., and Ph.D. in applied physics from the University of Tokyo in 1989, 1991, and 1994, respectively. He joined NTT Basic Research Laboratories in 1994. From 2001 to 2002, he was a visiting researcher at the Max Planck Institute for Solid State Research, Stuttgart, Germany. His research interests are focused on many-body effects in low-dimensional semiconductor structures. He is a member of the Physical Society of Japan and the Japan Society of Applied Physics.

Trevor David Rhone
Research Associate (Postdoctoral Research Fellow), Quantum Solid State Physics Research Group, NTT Basic Research Laboratories.
He received his B.A. from Macalester College, MN, USA, in 2005, and his M.A., M.Phil., and Ph.D. in physics from Columbia University, NY, USA, in 2007, 2008, and 2012, respectively. He joined the Quantum Solid State Physics Research Group, NTT Basic Research Laboratories, in 2012. From 2012 to 2013, he was a postdoctoral fellow at JST. His research interests are focused on correlated electronic phases in low-dimensional semiconductor structures. He is a member of the American Physical Society, the National Society of Black Physicists, and the New York Academy of Sciences.

Semiconductor Quantum Structures with Single-atom Precision

Kiyoshi Kanisawa and Stefan Fölsch

Abstract

Atom manipulation is a technology to assemble microstructures atom-by-atom by controlling the number of component atoms and their configuration. Combining this technology with the epitaxial growth of high quality semiconductor crystals makes it possible to uniformly manufacture and integrate quantum structures with single-atom precision. This technological innovation is expected to enable the next generation of electronics to overcome the limits of conventional silicon technology. In this article, atom manipulation of quantum structures at a compound semiconductor surface is reviewed, and a typical example is described—the fabrication of an ultimately precise quantum dot and quantum dot molecules.

Keywords: atom manipulation, STM, quantum structure

1.   Introduction

We have access to enormous amounts of information on the Internet these days. We can casually observe street views of arbitrary places almost anywhere on Earth. In the bygone era of telephone cards and message dial services, such a ubiquitous and interactive service was only a dream. Microfabrication is one of the technologies that has played a leading role in making this dream come true in our daily lives.

Semiconductor microfabrication technology has steadily progressed following Moore's law, which predicts that the number of transistors in an integrated circuit (IC) will double approximately every two years. However, this performance improvement of ICs is expected to come to an end in the near future because of quantum mechanical leakage current manifested by the miniaturization. Extensive research has been ongoing to find a way to overcome this limitation, and some studies have explored novel electric circuits that exploit evident quantum mechanical phenomena. To achieve such quantum mechanical electric circuits, microstructures called quantum structures are used in order to take advantage of the quantum mechanical behavior of electrons. A nano-structure[*1], one such microstructure, is considered to be a fundamental structure in which quantum behavior becomes the driving mechanism of novel electric circuits. The kinetics of a quantum shows a wave-particle duality, a concept that consists of the properties of a particle (kinetic momentum $p = mv$, where m and v are the mass and the velocity of the quantum) characterized by Newton's equation of motion, and the properties of a wave (wavelength λ) characterized by the Schrödinger equation. This λ is called the de Broglie wavelength[*2] and is estimated as $\lambda = h / mv$ (h: Planck constant). The de Broglie wavelength of an electron in a semiconductor is roughly a few tens of nanometers (1 nm = 10^{-9} m).

In contrast, the distance between the nearest-neighbor atoms in solids is of the order of 0.25 nm, which corresponds to about 1% of λ. The impact of a microfabrication error becomes more critical as the

---

*1 Nanostructure: A nanostructure is a microscopic structure formed or processed to have an exact size smaller than 100 nm in length, width, or height. In a wide sense, an object that includes a microscopic structure fabricated with these geometric specifications is referred to as a nanostructure.

*2 de Broglie wavelength: As light propagates as a wave, the motion of matter also behaves as a wave in the quantum mechanical picture. The de Broglie wavelength is the wavelength of the matter wave.
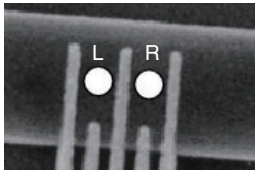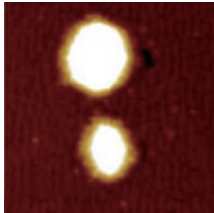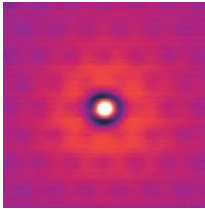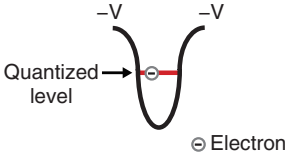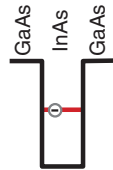
Microstructures	Lithography	Self-organization	Deposited atom (ion)
	300 nm	10 nm	1 nm
	L R		
Quantum confinement	External gate voltage	Band discontinuity of heterostructures	Coulomb potential
	−V −V Quantized → level ⊝ Electron	GaAs InAs GaAs	
Size error	5–10%	10–20%	0%
Positioning error	Nanometer scale	Nanometer to micrometer scale	Nanometer to micrometer scale or larger

Fig. 1. Comparison of nanostructures constructed using conventional microfabrication techniques with focused atom deposition.

manufactured device becomes smaller. This means that the electronic states in quantum structures are highly sensitive to a structural fabrication error on the order of a few atoms. The present error status of microfabricated nanostructures is shown in **Fig. 1**. It is clear that conventional microfabrication methods such as lithography and self-organization result in unavoidable variations comparable to or larger than the required precision in size and positioning. The effects of the inevitable processing errors make the device characteristics considerably inaccurate and problematic. Thus, to precisely control quantum mechanical electrons, microfabrication technology for quantum structures needs to have a degree of accuracy with the precision of a natural atom, which is the fundamental unit of matter.

In natural atoms, the Coulomb potential caused by the nucleus determines the quantum states of confined electrons. Since the potential form is determined by the atomic number, the electronic states are identical in each atom. Therefore, in the deposition process of atoms targeting a specific position on a solid crystal surface, the size error of the confinement structure by atoms becomes zero (Fig. 1), except for the positioning errors induced by focusing accuracy

and thermodynamic uncertainty [1]. Atom manipulation is a method of flexible positioning of atoms on a surface. The practicability was demonstrated [2] soon after the invention of the scanning tunneling microscope (STM) [3]. If the positioning uncertainty of the focused atom beam process can be corrected by atom manipulation after the deposition process, we can solve the positioning error problem by utilizing the size-error-free feature of natural atoms. In other words, atom manipulation is a key microfabrication technique with ultimate precision.

2. Experimental method

As a template for the atom manipulation, we used a (111)A-oriented surface of indium arsenide (InAs) crystal. InAs is a compound semiconductor composed of indium (In) and arsenic (As). It is one of the most important materials in high-speed electronic and infrared photonic devices. A high-quality InAs single-crystal thin film was grown on the (111)A-oriented substrate with atomically controlled thickness by using the crystal growth technique molecular beam epitaxy (MBE), as shown in **Fig. 2(a)**. Epitaxy refers to single-crystal growth with one particular

Fig. 2.   Collaborative semiconductor crystal growth and atom manipulation.

crystallographic orientation on a substrate. MBE is a vapor growth method, which performs the epitaxy by exposing the substrate to beams of atomic or molecular gases of high purity materials to the substrate under an ultra-high vacuum of 10^{-8} Pa or less. The MBE method can grow semiconductor crystals with very high purity and crystallinity. After the growth, the grown InAs surface was covered by a protective film (amorphous As). Then the sample was sent from NTT to the Paul-Drude-Institut für Festkörperelektronik (PDI) in Germany.

At PDI, the protective film was removed in an ultrahigh vacuum to recover the clean (111)A surface, and the sample was loaded into an STM operated at a sample temperature of 5 K. An STM is an electron microscope in which an atomically sharpened conductive needle (tip) is scanned above the conductive sample surface. The tip probes a tunneling current by applying a voltage difference between the tip and the sample surface at a distance of 1 nm or less, as shown in **Fig. 2(b)**. Through signal processing of the measured tunneling current, the STM visualizes the local electronic states to observe the surface atoms.

NTT and PDI have collaboratively developed atom manipulation at semiconductor surfaces using a low-temperature STM as a new approach to surpass conventional nanostructure fabrication technologies. Our achievements demonstrate that an international collaborative research project that combines individual skills and experience is effective for advancing science and technology.

3.   Atom manipulation on semiconductor surfaces

Surface atoms are arranged periodically on a

(a) Surface structure of substrate crystal used for atom manipulation

(b) Concept of atom manipulation and models of atomic rows

(c) Atom manipulation process and STM images of atomic rows
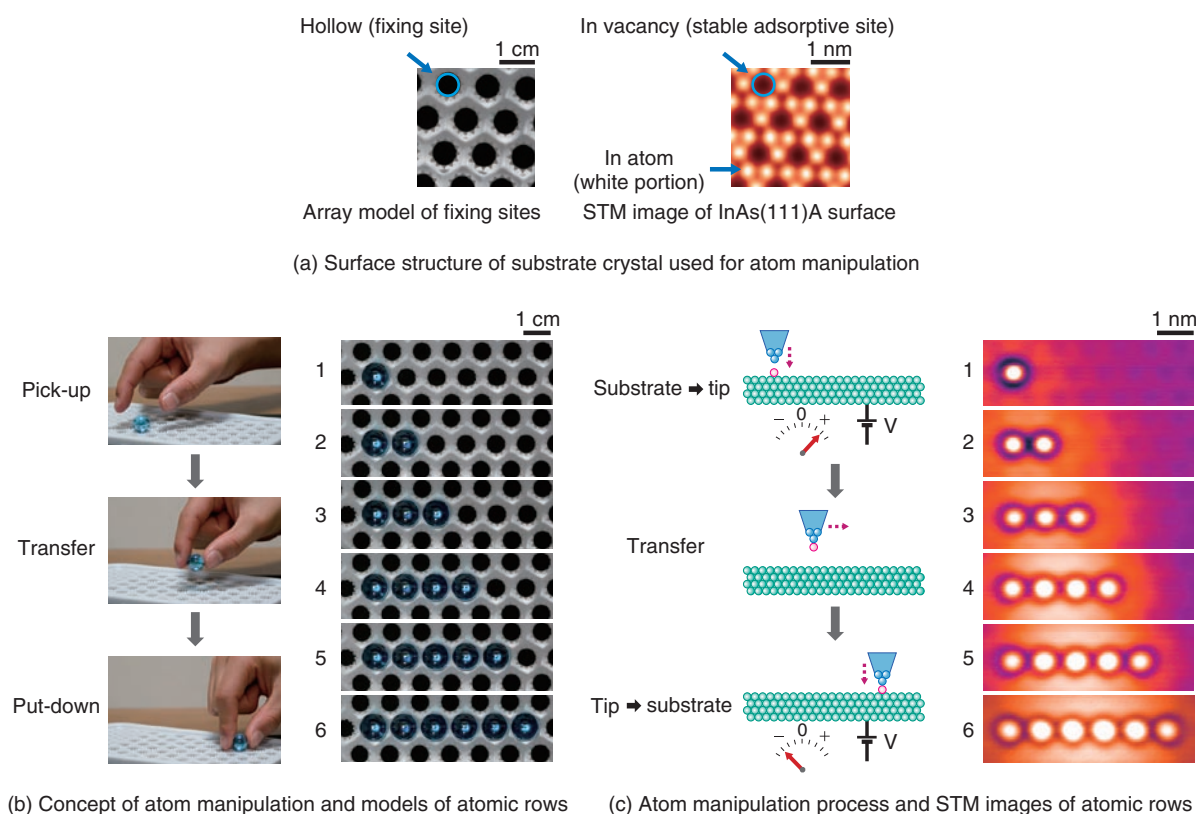
Fig. 3.   Atom manipulation to construct In atomic rows using STM.

semiconductor surface. Therefore, stable adsorptive sites of adsorbates also form periodic arrays. By using this periodicity, we can use the crystal substrate as a template to precisely control the structure formation by placing individual atoms at specific well-defined adsorptive sites. After the protective layer is removed, the (111)A surface recovers intrinsic hollow sites caused by periodic vacancies of surface In atoms at a distance of 0.857 nm. This surface structure is equivalent to that at the MBE-grown surface, as shown in **Fig. 3(a)**. These hollow sites work as stable adsorptive sites.

The concept of atom manipulation and some models of atomic chains are shown in **Fig. 3(b)**. In the atom manipulation process, the STM tip is used to pick up target atoms from the surface one by one and put them down in the designated position as the building blocks to form well-tailored nanostructures of atoms. The manipulated In adatoms are those that remain at the surface after the MBE growth [4]. The *pick-up* process and the *put-down* process are controlled by the applied sample bias voltage *V*.

At the experimental substrate temperature of 5 K,

thermal diffusion of adatoms at the sample surface is suppressed, and individual manipulated atoms are fixed stably at the positioned sites. We manufactured atomic chains ($1 \leq N$ (the number of atoms) ≤ 6) by arranging each In atom one-by-one in a line at the (111)A surface, as shown in **Fig. 3(c)** [5]. Precise and reproducible atomic configurations were possible because of the periodic and symmetric arrangement of stable adsorptive sites.

In addition to enabling microfabrication of quantum devices with single-atom precision, atom manipulation is advantageous in producing novel assemblies of integrated atoms, which do not exist in nature and are also not possible by means of chemical reaction. This is a very attractive subject that warrants further study. We expect that atom manipulation on semiconductor substrates will provide a novel aspect to our understanding of the physical and chemical properties of materials.
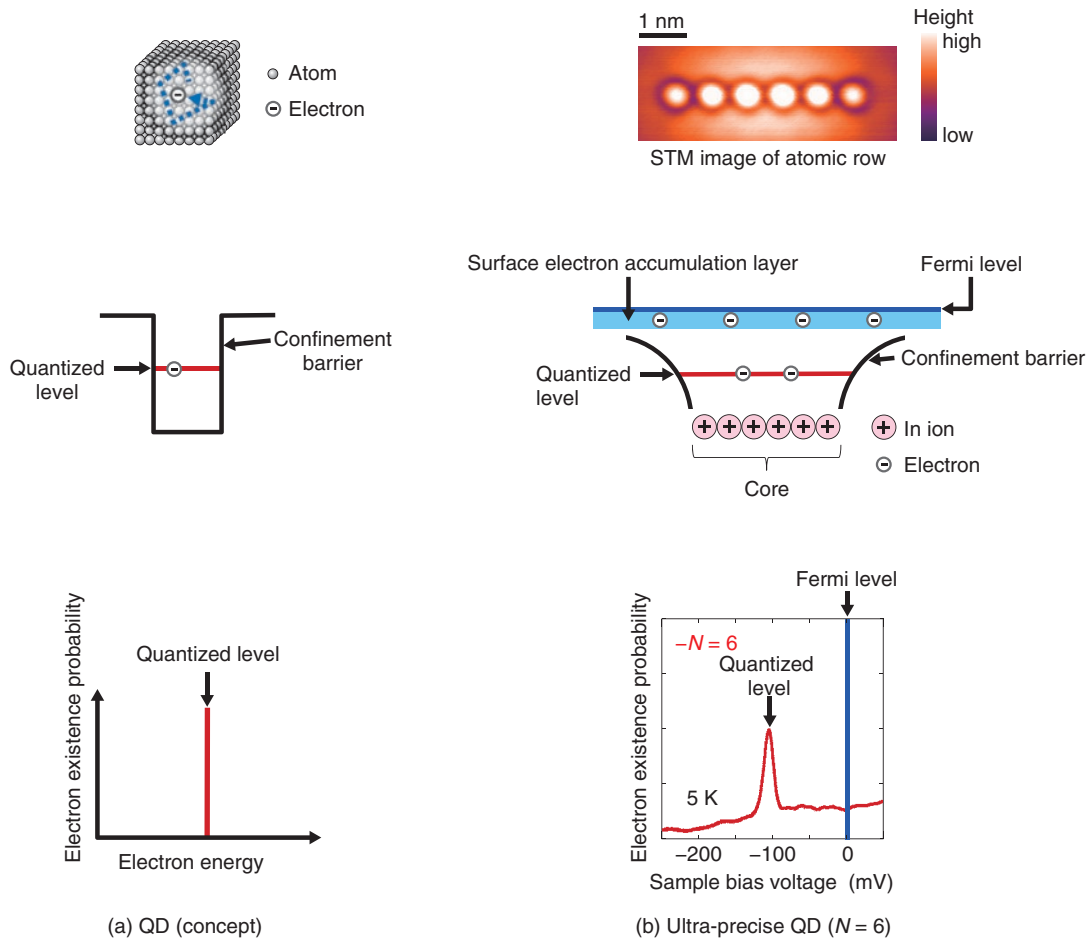
Fig. 4. Quantum dot (QD) structure, energy diagram, and density of states spectrum.

4. Quantum structures with ultimate precision

4.1 Ultra-precise quantum dots

The current atom manipulation technology enables us to flexibly implement quantum dots (QDs) with properties identical to those of natural atoms at the semiconductor substrate surface with single-atom precision of position and configuration. At the InAs surface, the indium (In) adatom is self-ionized to become a +1 charged ion that releases an electron to the substrate. Thus, a row of such ions behaves as a core, which confines electronic states at the semiconductor surface to the induced local Coulomb potential. This potential is well-defined, like that of a natural atom. These adatom-assembled structures work as ultra-precise QDs with excellent uniform properties. A QD is a microscopic structure that confines electrons in a narrow, nanometer-scale space with length, width, and height all comparable to or smaller than

the de Broglie wavelength of free electrons (1–100 nm), making quantum mechanical effects evident. The concept of the QD is shown in **Fig. 4(a)**. The QD is an essential component of electronic and optical semiconductor devices and has been extensively studied in the information technology field as a key element for developing a quantum computer. The electrons confined to a QD obey quantum mechanical statistics and form discrete energy levels similar to those of natural atoms. Moreover, QDs reproduce intrinsic properties of natural atoms such as the electron shell structure and Hund's rules [6]. Thus, a semiconductor QD is often called an artificial atom (which is not to be confused with an artificial element).

The experimental results for a QD with $N = 6$ are shown in **Fig. 4(b)**. Electrons are attracted to the core of ionized adatom rows. The spectral peak indicates that the free electronic states at the semiconductor
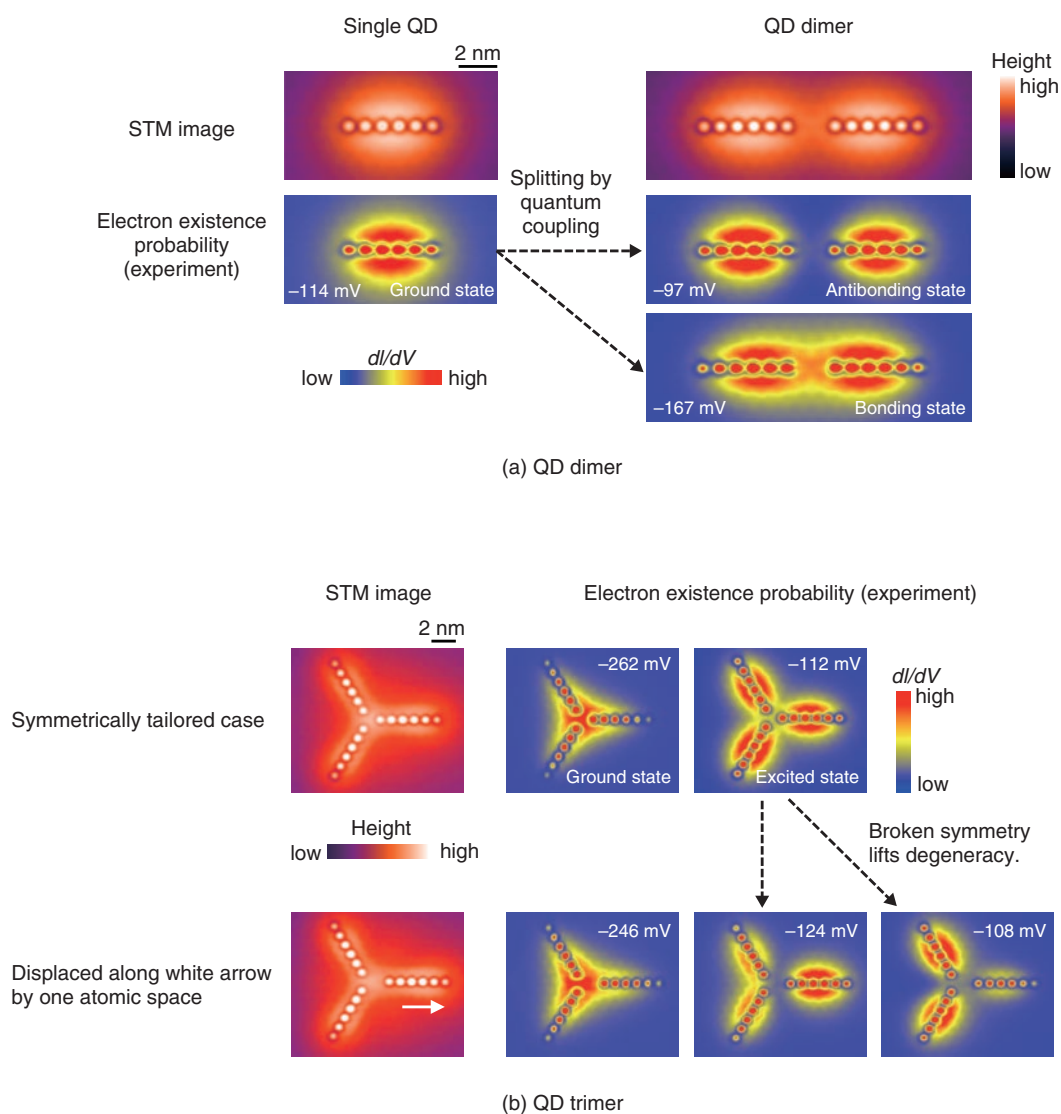
(a) QD dimer

(b) QD trimer

Fig. 5.   Artificial molecules of ultra-precise QDs.

substrate are confined, and a quantized level is obtained below the Fermi level[*3]. The discrete peak spectra of QDs ($6 \leq N \leq 25$) were characterized experimentally [7].

4.2   Integration of ultra-precise quantum dots

An artificial molecule, a microscopic structure made of artificial atoms, shows properties similar to those of a natural molecule. When two ultra-precise QDs ($N = 6$) are coupled as shown in **Fig. 5(a)**, molecular electronic states appear as a result of the interaction between individual quantized states in the QDs [6]. The observed states correspond to a bonding state and an antibonding state, similar to a natural

dimer molecule. Here, the electron existence probability maps (spatial distributions of the local density of states) were visualized by measuring the differential conductance (dI / dV) signal of the tunneling current I at a given sample bias voltage V. When exactly the same artificial molecule structures were fabricated, the standard deviation of the splitting energy between the bonding and antibonding levels was about 1 meV. This means that the fluctuation of the quantization characteristics is equivalent to the

---

*3   Fermi level: The Fermi level is an energy position at which the quantum mechanical electron distribution function has exactly 50% probability of occupancy by an electron.

unavoidable thermal fluctuation limit at the substrate temperature of 5 K ($3.5k_BT / 2 = 0.8$ meV; k_B: Boltzmann constant). This confirms that the reproducibility of quantization characteristics is excellent.

It should be emphasized that it would not be possible to substantially demonstrate the characteristics of the ultra-precise quantum structures fabricated using the atom manipulation technique if there were any fluctuations due to defects or impurities. We were able to experimentally confirm the excellent fabrication accuracy by atom manipulation only because the quality of the MBE-grown semiconductor crystal was sufficiently high. This indicates that very minute semiconductor quantum structures require the synergistic effect of combining the extremely high processing accuracy at atomic-scale precision and the high crystallinity of the semiconductor crystal. The availability of such ultra-precise quantum structures enables us to manufacture ultimate quantum devices with atomic-level reproducibility like an array of quantum bits with uniform functions, which have not yet been achievable because of statistical errors in structural fluctuations. Furthermore, integration and control of these precise nanostructures will help us to develop a quantum computer and the next generation of technology to overcome the limits of conventional silicon technology.

When three ultra-precise QDs ($N = 6$) are combined as shown in **Fig. 5(b)**, they form a nanostructure corresponding to a QD trimer molecule [6]. When one of the component QDs was displaced at a distance of one adsorptive site (0.857 nm), and thereby the rotational symmetry was broken, tunneling spectroscopy measurements revealed that the doubly degenerated excited level split into two individual levels. This infers that the electronic structure of the symmetric QD trimer consists of molecular levels, as is expected from a triatomic molecule of threefold symmetry. It also shows that an atomic-level error of the quantum structure has a crucial effect on the quantum characteristics.

From the viewpoint of device integration, this QD trimer is regarded as a semiconductor structure made of three QDs integrated in a 10×10 nm square area. If the degree of integration is estimated by assuming that one QD corresponds to one transistor, the degree is about 1000 times as high as that of conventional ICs, though it is limited to a local area. This indicates that the QD trimer molecule in Fig. 5(b) can be considered a mock-up of an ultimately integrated semiconductor microstructure. The thermal de Broglie wavelength λ_D (defined by $\lambda_D = h / \sqrt{2\pi mk_BT}$) gives a good estimate of the maximum length scale of the quantum mechanical features in the surroundings at temperature T. For semiconductor electrons with the typical effective mass $m^* = 0.1m_0$ (m_0: the rest mass of an electron in vacuum), λ_D is 105 nm at the substrate temperature ($T = 5$ K). At room temperature ($T = 300$ K), λ_D is 13.6 nm. This suggests that nanostructures comparable in size to the QD trimer shown in Fig. 5(b) will operate as functional elements of quantum devices even at room temperature. This remarkable fact greatly encourages advances in nanostructure research.

The design and analysis of the artificial atoms and molecules were performed by doing theoretical first-principles calculations at the Naval Research Laboratory (NRL) in the USA in collaboration with PDI. This theoretical analysis enabled us to confirm that the confined states in QDs originate from the intrinsic electronic states of the semiconductor substrate, not from the atomic orbitals of In adatoms.

5. Future perspectives

At present, STM-based atom manipulation is the sole technology for fabricating ultra-precise quantum structures with respect to position and configuration. We expect that the present achievements will open the door to developing new electronics technology by combining atomic and molecular electronics with semiconductor thin film technology. However, the throughput of atom manipulation, namely the efficiency of manufacturing nanostructures per unit time, is inferior to that of conventional lithography technologies [8]. When this drawback is overcome and the fabrication and characterization of quantum structures with atomic precision become possible at the semiconductor substrate surface, we will be a giant step closer toward realizing a new IC technology through the fusion of wafer-level semiconductor technology and atomic and molecular electronics. By exploring novel properties of many integrated atoms and their interaction with semiconductor heterostructures, we plan to develop architectures for quantum computers and high-performance semiconductor devices composed of well-defined semiconductor nanostructures with robust fidelity. Further study will bring many benefits to a broad range of science and technology fields.

References

[1]   T. Shinada, S. Okamoto, T. Kobayashi, and I. Ohdomari, "Enhancing

Semiconductor Device Performance Using Ordered Dopant Arrays," Nature, Vol. 437, No. 7062, pp. 1128–1131, 2005.

[2] D. M. Eigler and E. K. Schweizer, "Positioning Single Atoms with a Scanning Tunnelling Microscope," Nature, Vol. 344, No. 6266, pp. 524–526, 1990.

[3] G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, "Surface Studies by Scanning Tunneling Microscopy," Phys. Rev. Lett., Vol. 49, No. 1, pp. 57–61, 1982.

[4] K. Kanisawa and T. Fujisawa, "Mechanism of Electron Accumulation Layer Formation at the MBE-grown InAs(111)A Surface," Hyomen Kagaku, Vol. 29, No. 12, pp. 747–757, 2008 (in Japanese).

[5] S. Fölsch, J. Yang, C. Nacci, and K. Kanisawa, "Atom-By-Atom Quantum State Control in Adatom Chains on a Semiconductor," Phys. Rev. Lett., Vol. 103, No. 9, pp. 096104-1–4, 2009.

[6] S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, and L. P. Kouwenhoven, "Shell Filling and Spin Effects in a Few Electron Quantum Dot," Phys. Rev. Lett., Vol. 77, No. 17, pp. 3613–3616, 1996.

[7] S. Fölsch, J. Martínez-Blanco, J. Yang, K. Kanisawa, and S. C. Erwin, "Quantum Dots with Single-atom Precision," Nature Nanotechnol., Vol. 9, No. 7, pp. 505–508, 2014.

[8] R. Garcia, A. W. Knoll, and E. Riedo, "Advanced Scanning Probe Lithography," Nature Nanotechnol., Vol. 9, No. 8, pp. 577–587, 2014.

Kiyoshi Kanisawa
Senior Research Scientist, Quantum Solid-State Physics Research Group, Physical Science Laboratory, NTT Basic Research Laboratories.
He received his B.S., M.S., and Dr.Sci. in physics from Tokyo University of Science in 1986, 1988, and 1998, respectively. He joined NTT LSI Laboratories in 1988. He has been conducting research on semiconductor surface science at NTT Basic Research Laboratories since 1995. His current research interest is the physics of semiconductor surfaces and epitaxial growth. He is a member of the Physical Society of Japan (JPS), the Japan Society of Applied Physics (JSAP), the Surface Science Society of Japan (SSSJ), and the American Vacuum Society (AVS).
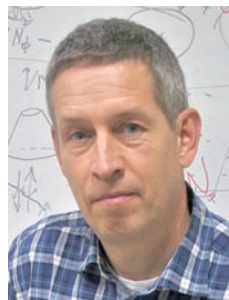
Stefan Fölsch
Senior Scientist, Leader of the Low-temperature STM Research Group, Paul-Drude-Institut für Festkörperelektronik (PDI).
He received his B.S. and Ph.D. from the University of Hannover in 1987 and 1991, respectively. He was a research associate at NTT from 1992 to 1994 and at the Free University of Berlin from 1994 to 2002. He received Habilitation (academic qualification) at the Free University of Berlin in 2004. He joined PDI in 2002. His current research interests are surface physics-related aspects of atom and molecule manipulation by STM. He is a member of the Deutsche Physikalische Gesellschaft (DPG, German Physical Society).

Creating a Topological Insulator Using Semiconductor Heterostructures

Kyoichi Suzuki and Koji Onomitsu

Abstract

Topological insulators are new states of matter that cannot be classified into any existing categories of materials. They have attracted much attention for their potential use in electronic and quantum computing devices because of their specific electron transport properties. In this article, we describe how we created an artificial topological insulator using a heterostructure comprising common semiconductor materials.

Keywords: semiconductor heterostructure, topological insulator, edge channel

1.   Introduction

We have recently seen a new movement in materials science to classify materials from the viewpoint of topology. In the process, new states of matter have been found, which cannot be classified into existing categories of materials. One such state is the topological insulator (TI), which has an insulating phase with a topology different from that of normal insulators. Mathematically, a state cannot continuously transit between different topologies and is therefore necessarily broken at the boundary (**Fig. 1**). In an analogous way, the insulating state in TIs cannot continue to the normal insulating state outside, and it is therefore necessary to break the insulating states. As a result, a conductive channel arises at the boundary of a TI. This topologically protected conductive channel is expected to have low power consumption or specific transport properties, and thus, TIs have attracted a great deal of attention as a new type of electronic material.

Several TIs have been confirmed experimentally, for example, Bi_2Se_3 and HgTe. However, these materials are not common in industry, and specific know-how is necessary to achieve their crystal growth. They have a distinct energy band structure by nature, which is what realizes a TI (discussed later). In contrast, we created an artificial TI by using a semiconductor heterostructure comprising indium arsenide (InAs) and gallium antimonide (GaSb) [1], which are commonly used industrially. Taking advantage of the highly developed semiconductor technology will make it possible to carry out detailed clarifications and to implement various device applications of TIs.

2.   Topological insulator

When atoms are condensed and crystallized, the electron orbits for each atom are hybridized. As a result, electronic states called energy bands are formed due to the periodicities of the crystal. The energy band accumulates electrons from the lower energy side. When one energy band is completely filled by the electrons and the upper energy band is empty, the state of the matter becomes an insulator (**Fig. 2**). The filled energy band is called the valence band, and the upper empty energy band is called the conduction band. The gap between these energy bands is called the band gap. This is the situation in an intrinsic semiconductor. At low temperatures, thermal excitation of electrons from the valence band to the conduction band is suppressed. This state results in an insulator.
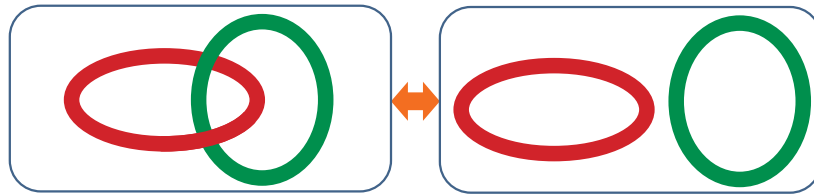
In TIs, the conduction and valence bands overlap

Fig. 1. Topologically different states.

(a) Insulator (intrinsic semiconductor)
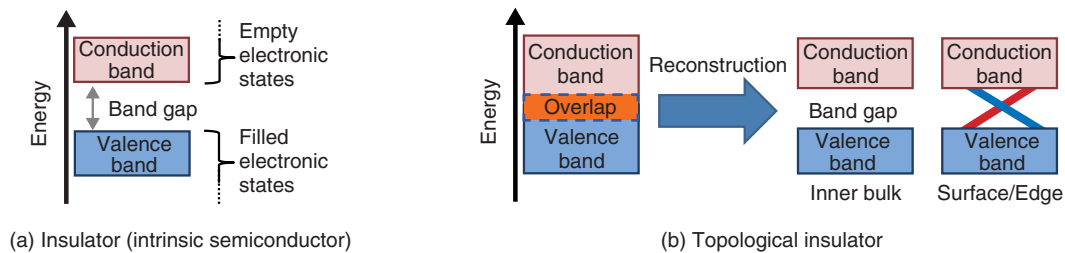
(b) Topological insulator
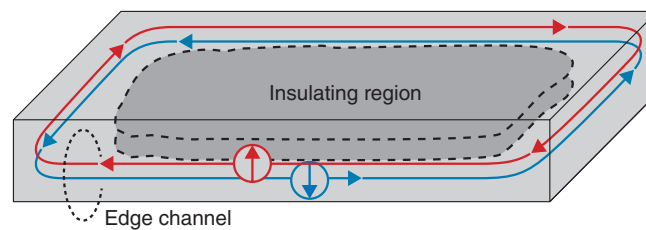
Fig. 2. Energy band structures.

Fig. 3. 2D TI and edge channel.

energetically. The wave functions in each band are hybridized based on the relativistic interaction. As a result, the energy gap is reopened and the conduction and valence bands are reconstructed. When the reconstructed valence band is completely filled with electrons and the conduction band is empty, an insulator state exists. However, the topology of this state is different from that outside. Therefore, the band gap should be closed at the boundary. Electronic states connecting the valence and conduction bands appear as a conductive channel (red and blue lines in Fig. 2(b)). For a three-dimensional (3D) TI, this conductive channel corresponds to the two-dimensional (2D) surface state. For a 2D TI, this channel corresponds to the edge channel. Singular electronic transport is expected due to the restriction of electron scattering.

We have been investigating 2D TIs whose edge

channel is composed of counter-propagating electron currents with opposite spin directions, which are called helical spin currents (**Fig. 3**). The topologically robust protection of the edge channel prevents the back scattering of electrons between counter-propagating currents in the edge channel with spin flip. As a result, dissipationless electron transport, that is, quantized conductance, should be achieved. In practice, a value very close to the quantized conductance has been observed in an HgTe/HgCdTe heterostructure, which is a representative 2D TI. In addition, at the junction between the helical spin current and the superconductor, a curious quasiparticle state called the Majorana fermion, which is simultaneously a particle and its antiparticle, is expected to be found. The Majorana fermion is promising as a quantum bit for fault-tolerant quantum computing.
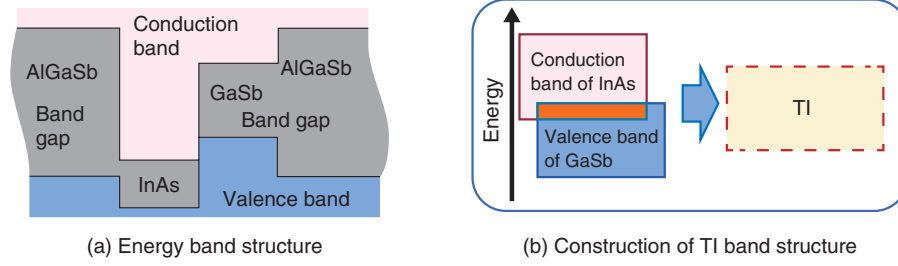
(a) Energy band structure

(b) Construction of TI band structure

Fig. 4.   InAs/GaSb heterostructure.

(a) Cross-sectional microscopic image
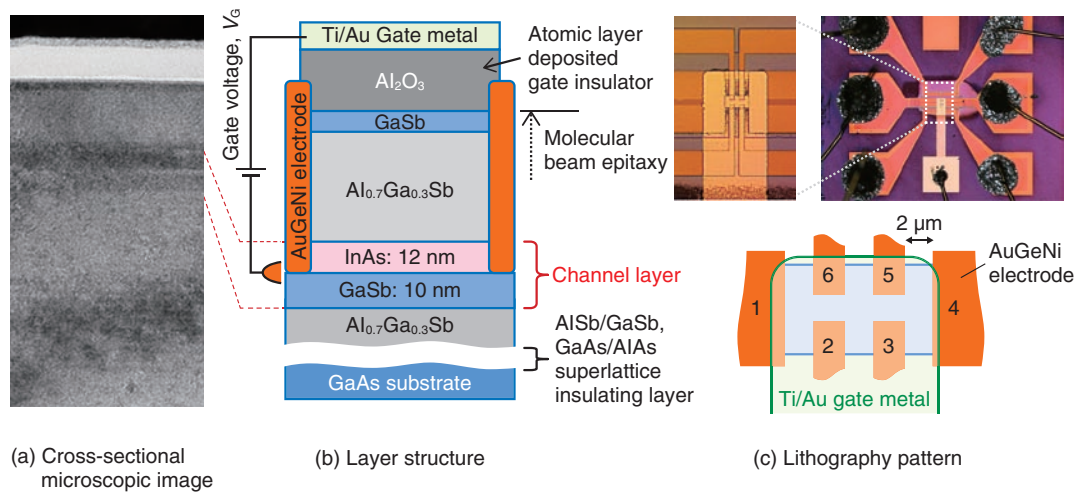
(b) Layer structure

(c) Lithography pattern

Fig. 5.   InAs/GaSb heterostructure sample.

As described above, 2D TIs are promising for use in low-power-consumption devices, quantum computing, and spintronic devices.

3.   InAs/GaSb heterostructure

Both InAs and GaSb are common and widely used semiconductors. When the heterojunction is made, the conduction band in InAs and the valence band in GaSb overlap energetically. After the layer thicknesses are optimized and the InAs/GaSb heterostructure is sandwiched between AlGaSb barrier layers, the energy band structure of TIs is artificially realized through the hybridization of the adjacent conduction and valence bands (**Fig. 4**). This heterostructure becomes a 2D TI because of the layer structure.

The structure of our sample is shown in **Fig. 5**. We grew the semiconductor layers using molecular beam epitaxy, which enables us to control each layer thickness to one atomic layer. We used photolithography and etching to fabricate a small device pattern, as shown in Fig. 5(c). Then, six electrodes were formed by evaporating AuGeNi alloy. An Al_2O_3 gate insulating layer was deposited by atomic layer deposition, and a Ti/Au top gate was evaporated. The electron density can be controlled by applying voltage between the gate and the InAs/GaSb channel layer. The TI is realized when the valence band is completely filled with electrons and the conduction band is empty.

4.   Low-temperature measurement

Similar to common insulators and intrinsic semiconductors, the insulating phase in a TI is broken by thermal electron excitation from the valence band to the conduction band. Therefore, to obtain the distinct transport properties of a TI, an extremely low
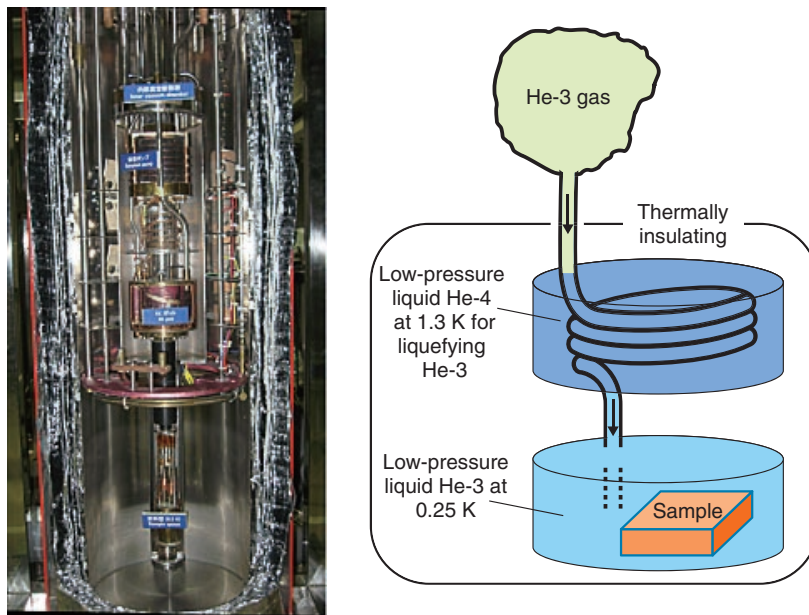
Fig. 6.   He-3 refrigerator.

$R_{14,23} = h/(2e^2)$

(a) Longitudinal resistance
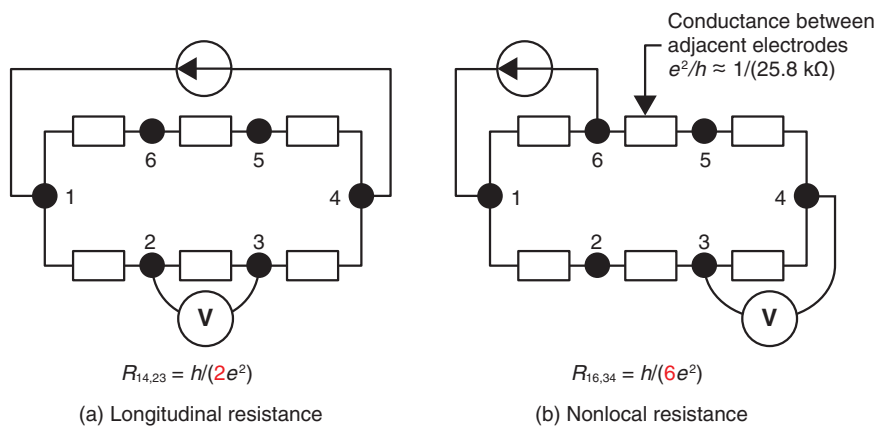
$R_{16,34} = h/(6e^2)$

(b) Nonlocal resistance

Fig. 7.   Equivalent circuits for 2D TI with six electrodes.

temperature is necessary. Here, we used a He-3 (an isotope of helium) refrigerator (**Fig. 6**). In this refrigerator, He-3 gas is liquefied by low-pressure liquid He-4 (common helium), whose temperature is about 1.3 K. Then the pressure of the liquid He-3 is reduced. As a result, the lowest possible temperature of 0.25 K is obtained.

5.   Confirmation of TI

In the 2D TI, the inner bulk part is insulating. The electron transport takes place through the edge channel between adjacent electrodes. Ideally, the conductance between the adjacent electrodes should be quantized conductance, e^2/h, and the resistance should be its inverse, $h/e^2 \approx 25.8$ kΩ. The equivalent circuit models of the sample with six electrodes are shown in **Fig. 7**. Based on a simple resistance calculation, longitudinal resistance $R_{14,23} = h/(2e^2) \approx 12.9$ kΩ and, for example, nonlocal resistance $R_{16,34} = h/(6e^2) \approx 4.3$ kΩ are expected. In practice, however, electron scattering occurs for various reasons.
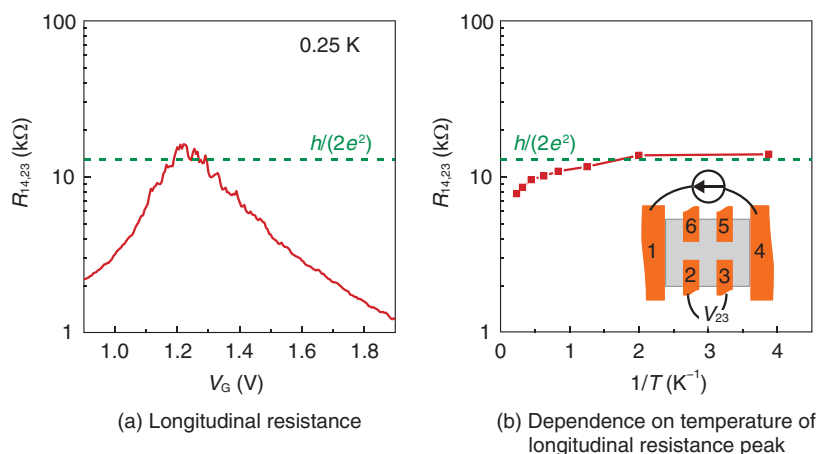
Fig. 8. Longitudinal resistance and its dependence on temperature.

Therefore, the conductance between the adjacent electrodes deviates from the quantized value.

Longitudinal resistance ($R_{14,23}$) experimentally obtained as a function of the gate voltage (V_G) is shown in **Fig. 8**. The electron density is low on the low V_G side, and there are unoccupied electron states (holes) in the valence band. The holes act as carriers with a positive charge and contribute to the electronic transport. Therefore, the inner bulk region is conductive, and $R_{14,23}$ becomes low. In contrast, on the high V_G side, electrons partially accumulate in the conduction band and act as conductive electrons, which also lowers $R_{14,23}$. Around $V_G = 1.25$ V, the valence band is completely filled with electrons and the conduction band is empty. There are no conductive electrons or holes in the inner bulk region. Therefore, $R_{14,23}$ increases. The peak resistance of $R_{14,23}$ is about 20 kΩ. If the sample were an intrinsic semiconductor, the resistance peak would become infinite. The finite resistance indicates that the sample has a conductive channel, and this resistance peak value is not far from that expected as quantized transport, $h/(2e^2)$. This suggests that the sample becomes a TI at this V_G.

The dependence of the $R_{14,23}$ peak on temperature is shown in Fig. 8(b), where the horizontal axis is the inverted temperature. On the high-temperature side (left side) the resistance peak is low. This is because the thermally excited electrons from the valence band to the conduction band and simultaneously generated holes in the valence band contribute to the electronic transport in the inner bulk region. The estimated energy gap from the slope in the high-temperature region is 0.18 meV (≈ 2.1 K). On the low-temperature

side, the peak resistance approaches $h/(2e^2)$.

In an attempt to obtain clear evidence for the TI, we measured nonlocal resistances. $R_{16,23}$ and $R_{65,23}$ as a function of V_G are shown in **Fig. 9(a)**. Here, $R_{ij,kl}$ (= V_{kl}/I_{ij}) indicates that electrodes i and j are used for current (I_{ij}) injection and ejection, and electrodes k and l are used for voltage (V_{kl}) measurement. Nonlocal resistances $R_{16,34}$ and $R_{65,34}$ are shown in **Fig. 9(b)**. The resistance peaks deviate from the expected value $h/(6e^2)$ in the quantized transport. We could not confirm the realization of a TI from these results.

Next, the nonlocal resistance ratios (voltage ratios between the adjacent electrode pairs, V_{23}/V_{34}) are plotted in **Fig. 9(c)**. For the pink curve, the current is injected into electrode 1 and ejected from electrode 6. For the purple one, it is injected into 3 and ejected from 4. Around the resistance peak, these two curves completely agree with each other independently of the current injection/ejection path. We observed such agreement of the nonlocal resistance ratios between the adjacent voltage electrode pairs for all combinations of electrode arrangements. This indicates that the transport occurs only between adjacent electrodes; that is, the inner bulk region is insulating, and the transport is governed by the edge channel. Here, we can confirm the realization of a TI in the InAs/GaSb heterostructure. When the temperature is increased to 4.3 K, the agreement of the nonlocal resistance ratios disappears, as shown in **Fig. 9(d)**. This indicates that the thermally excited carriers contribute to the transport in the inner bulk region.
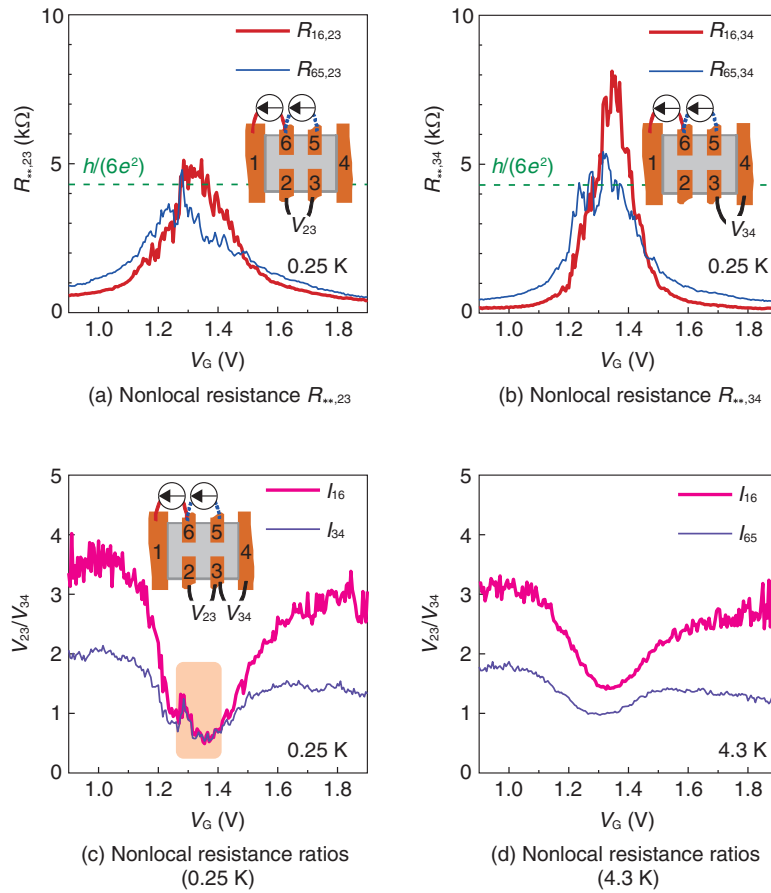
(a) Nonlocal resistance $R_{**,23}$

(b) Nonlocal resistance $R_{**,34}$

(c) Nonlocal resistance ratios
(0.25 K)

(d) Nonlocal resistance ratios
(4.3 K)

Fig. 9.   Nonlocal resistances and their ratios.

6.   Future prospects

Both InAs and GaSb can achieve good electronic connections with ferromagnetics or superconductors. This is expected to lead to their application for spintronic or quantum information devices. In addition, our artificial TIs created using heterostructures between non-topological materials create the possibility of combining various materials, including insulators or metals, to achieve new TIs.

Reference

[1]   K. Suzuki, Y. Harada, K. Onomitsu, and K. Muraki, "Edge Channel Transport in the InAs/GaSb Topological Insulating Phase," Phys. Rev. B, Vol.87, No. 23, 235311, 2013.
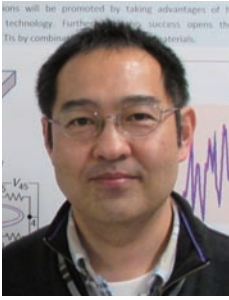
Kyoichi Suzuki

Senior Research Scientist, Quantum Solid State Physics Research Group, Physical Science Laboratory, NTT Basic Research Laboratories.

He received his B.E., M.E., and Ph.D. in applied physics from Tohoku University, Miyagi, in 1991, 1993, and 2004, respectively. He joined NTT Basic Research Laboratories in 1993. He has studied magneto-transport, electronic, and optical properties in semiconductor low-dimensional structures. He is a member of the Physical Society of Japan (PSJ) and the Japan Society of Applied Physics (JSAP). In 2008, he received the JSAP Outstanding Paper Award.

Koji Onomitsu

Senior Research Scientist, Low-Dimensional Nanomaterials Research Group, Materials Science Laboratory, NTT Basic Research Laboratories.

He received a B.E. and M.E. in electrical engineering from Aoyama Gakuin University, Tokyo, in 1998 and 2000, respectively, and a Ph.D. in electrical engineering from Waseda University, Tokyo, in 2004. He joined NTT Basic Research Laboratories in 2006. His current interests are crystal growth and device applications of two-dimensional layered materials such as MoSe2. He is a member of JSAP.

Two-mode Squeezing in an Electromechanical Resonator

Imran Mahboob, Hajime Okamoto, and Hiroshi Yamaguchi

Abstract

A mechanical resonator integrated with piezoelectric transducers enables mechanical nonlinearities to be dynamically engineered to emulate non-degenerate parametric down-conversion. In this configuration, millions of phonons are simultaneously generated in pairs in two macroscopic vibration modes, which results in the amplification of their motion by more than 20 dB. Mechanical two-mode squeezed states are also created in parallel, which exhibit fluctuations 5 dB below the thermal level of their constituent modes and they harbor correlations between the modes that become perfect as their amplification is increased. This remarkable observation of correlations between two massive phonon ensembles establishes the means to create an entangled macroscopic mechanical system at the single phonon level.

Keywords: electromechanical resonator, parametric down-conversion, squeezed-states, phonons

1. Introduction

Electromechanical systems consist of a single degree of mechanical freedom hosting a spectrally pure resonance that is embedded in an electrical transduction circuit, and they have emerged as a versatile platform for a range of technological applications and to study fundamental science [1, 2]. For instance, their small inertial mass and high quality resonance have enabled the development of ultra-precise sensors that can detect a single electron spin and even the mass of a proton leading to a new class of medicinal diagnostic technology [3–5]. The electrical transduction circuit can also enable the underlying potential energy landscape of the mechanical resonator to be dynamically engineered, yielding a range of nonlinear motional dynamics that can be harnessed for both information storage and processing, which brings forth the concept of mechanical computation offering both ultra-low power consumption and the capacity for unprecedented parallel data processing [6–9].

Most tantalizingly, a high frequency mechanical resonator operated at sufficiently low temperatures can even condense into its quantum ground state, where its low mass and spectral purity yield quantum zero point fluctuations that are large enough to be observed, enabling a macroscopic quantum system to be studied [10–12]. In this instance, although the mechanical resonator is composed of billions of atoms, only the phonons sustained by the fundamental resonance mode are quantized, and it corresponds to a tangible vibration of the mechanical element. This phonon picture, namely the collective excitation of atoms in the fundamental vibration mode, enables concepts from atomic molecular optical (AMO) physics developed for photons, that is, a quanta of electromagnetic radiation, to be exploited. Indeed, it is the laser cooling techniques pioneered in AMO physics that have been utilized most successfully to cool the macroscopic mechanical oscillator so that on average, its fundamental mode is occupied by much less than one phonon [13, 14].

Although the notion of a macroscopic mechanical resonator composed of billions of atoms being considered only in terms of the number of phonons sustained by its fundamental mode might seem counterintuitive, this paradigm has successfully been exploited
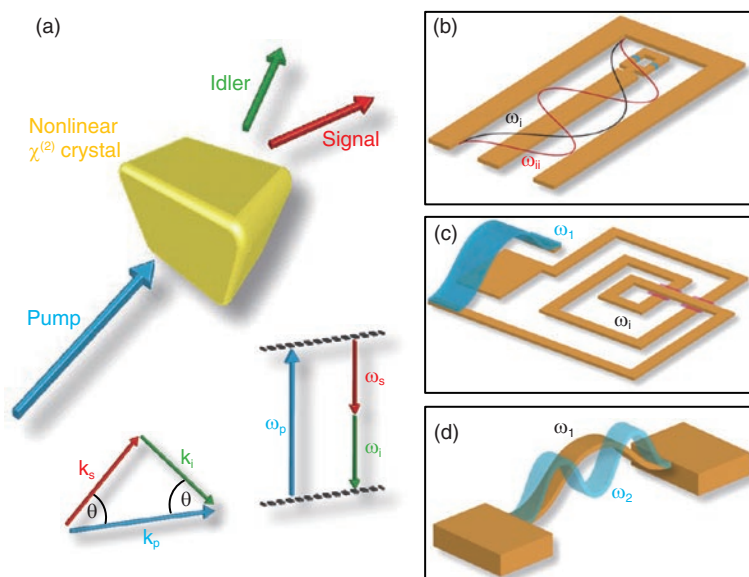
Fig. 1. (a) Schematic illustration of the spontaneous parametric down-conversion process used to generate non-classical light where the signal and idler photons are entangled in a two-mode squeezed state. This process occurs at the expense of the pump photons where the signal and idler photons conserve both momentum $k_p = k_s + k_i$, where k is the wavevector, and energy $\omega_p = \omega_s + \omega_i$. The nonlinear crystals typically used are BBO (beta-barium borate) for Type II down-conversion (where the signal and idler photons have perpendicular polarization) and KDP (potassium dihydrogen phosphate) for type I down-conversion (where the signal and idler photons have parallel polarization). (b) Schematic depiction of a microwave cavity composed from a superconducting metal (orange) and terminated by a SQUID with two Josephson junctions (blue) that can sustain electromagnetic standing waves, for instance ω_i and ω_{ii}, which are coupled via the SQUID. (c) Cavity electromechanical system with a microwave resonator of frequency ω_i consisting of a spiral inductor and a capacitor and a mechanical resonator of frequency ω_1 which forms a compliable capacitor. Light and sound in these two subsystems are nonlinearly coupled via the oscillating electrical energy stored in the capacitor. (d) Phonon-cavity electromechanical system consisting of two localized mechanical vibration modes ω_1 and ω_2 that are coupled via strain. In all three systems, if the respective flux penetrating the SQUID, the electric field stored in the capacitor, or the mechanical strain are sinusoidally pumped at the sum frequency of their electromagnetic and/or mechanical modes, signal and idler photons and/or phonons are generated that conserve energy but the momentum conservation constraint is relaxed due to their *stationary wave* nature.

to demonstrate the signature feature of photonics, namely lasing but with the localized phonons in the mechanical resonator [15–18]. Consequently, the ability to cool a solid-state macroscopic mechanical resonator into its ground state opens up the possibility of generating an all-mechanical macroscopic entangled state [10, 19], which would enable foundational aspects of quantum mechanics to be queried such as the nature of the quantum to classical divide and the absence of quantum phenomena in our everyday classical world [20, 21].

2. Two-mode squeezed states

Entangled photons in the guise of two-mode squeezed states were among the first non-classical states of light to be generated in the lab and nowadays are routinely siphoned from spontaneous parametric

down-conversion [22]. This typically involves a $\chi^{(2)}$ nonlinear crystal that is exposed to a strong laser pump beam whose dielectric polarization responds nonlinearly to the pump's electric field. In this process, a pair of photons—the signal and the idler—are generated at the expense of the pump photons, which conserve both energy and momentum, as depicted in **Fig. 1(a)**. Remarkably, the quantum fluctuations in the amplitude and phase of the generated photon pair are correlated, as encapsulated by the Einstein, Podolsky, and Rosen (EPR) paradox, where their individual fluctuations are amplified, while their relative fluctuations are reduced below the vacuum noise level [23–26]. The EPR paradox recast in the Bell inequality has been routinely violated via the entangled signal and idler photons whose fluctuations remain correlated even when they are separated over distances of kilometers [27, 28]. The non-classical

light generated from spontaneous parametric down-conversion has subsequently engendered a whole host of quantum-enabled technologies including quantum communication [29], optical quantum computing [30], quantum teleportation [31], quantum enhanced measurements and metrology [32, 33].

This unprecedented success has led to two-mode squeezed states being translated to other parts of the electromagnetic spectrum, for instance microwaves utilizing superconducting circuits [34]. Specifically, a superconducting resonator terminated with a nonlinear element such as a superconducting quantum interference device (SQUID), as generically depicted in **Fig. 1(b)**, can enable parametric down-conversion [35, 36]. Pumping the SQUID with flux in these so-called Josephson parametric amplifiers (JPAs) yields a nonlinear variation in the boundary conditions of the microwave resonator which creates a frequency modulation in the standing waves it hosts. If the pump frequency coincides with the sum of two modes (say ω_i and ω_{ii} as depicted in Fig. 1(b)) in the microwave resonator, it non-degenerately amplifies their vacuum noise, generating signal and idler photons in a process reminiscent to that described above where these photons are entangled [37, 38]. Indeed, JPA-like devices have now opened the door to many of the concepts pioneered with optical two-mode squeezed states but in an on-chip microwave circuit that could even be yoked into a quantum computer.

A variation on these microwave circuits has recently emerged in the form of cavity electromechanical systems where the cavity, namely the microwave resonator, is composed of an inductor and capacitor in series, and the mechanical resonator forms a mechanically compliant element of the parallel plate capacitor, thus capacitively coupling the mechanics to the microwaves as generically depicted in **Fig. 1(c)** [39–41]. Electric fields in the forms of microwaves can be pumped into this hybrid circuit via a capacitively coupled transmission line that nonlinearly modulates the energy stored in the capacitor and hence the frequencies of the mechanics and the cavity. If the pump coincides with the sum frequency of the subsystems (i.e. ω_i and ω_1 as depicted in Fig. 1(c)), it behaves like a non-degenerate parametric amplifier and it creates photons and phonons in pairs that are correlated in a two-mode squeezed state [42, 43]. This remarkable demonstration illustrates that two vastly dissimilar systems, namely light and sound, can be entangled in a macroscopic context, providing a new regime in which quantum physics can be explored [14].

The cavity electromechanical system detailed in Fig. 1(c) straddles the two extremes of light and sound, and it naturally suggests the possibility of a purely mechanical analogue of a two-mode squeezed state in a *phonon-cavity* electromechanical system as depicted in **Fig. 1(d)**. As in the microwave case, the mechanical resonator can also sustain multiple modes, say ω_1 and ω_2, as shown in Fig. 1(d), which if nonlinearly coupled, could yield non-degenerate parametric amplification. In contrast to the microwave circuit that needs to be engineered with a nonlinear element, namely the SQUID, to enable the modes to couple, the modes in the mechanical resonator can dispersively couple naturally via the strain induced from the motion of a given mode that modifies the restoring potential, that is, the spring constant and hence, the natural frequency of the other modes [44]. Consequently, if the strain in the mechanical element is sinusoidally pumped, it will nonlinearly modulate the frequencies of all the localized modes, and once this modulation coincides with a pair of modes, it will non-degenerately amplify their motion and in the process generate a two-mode squeezed state [45]. Critically, however, an entanglement will only be generated if the constituent modes are initialized in their quantum ground state [10–12], but the ability to harness this interaction, even if the modes in question are thermalized, would lay a pivotal marker on the road to generating an all-mechanical macroscopic entanglement [46].

3. Electromechanical resonator

Although degenerate parametric amplification was demonstrated in the early nineties, it could not be used to disentangle correlations between the signal and idler phonons due to their having identical frequencies and being spatially localized to the same mode in the mechanical resonator [47]. Alternatively, a non-degenerate variant of this process in the first two modes of a beam resonator yielded only a modest phonon generation rate, which was insufficient to significantly amplify the thermomechanical fluctuations before they were dissipated, resulting in statistically insignificant correlations [45].

To that end, the electromechanical resonator shown in **Fig. 2(a)** was developed, which consists of two mechanical elements that are strongly coupled via the exaggerated overhangs between them [48]. This results in the two lowest-order vibration modes shown in **Figs. 2(b)** and **2(c)** that are extracted from a finite element calculation and henceforth labeled
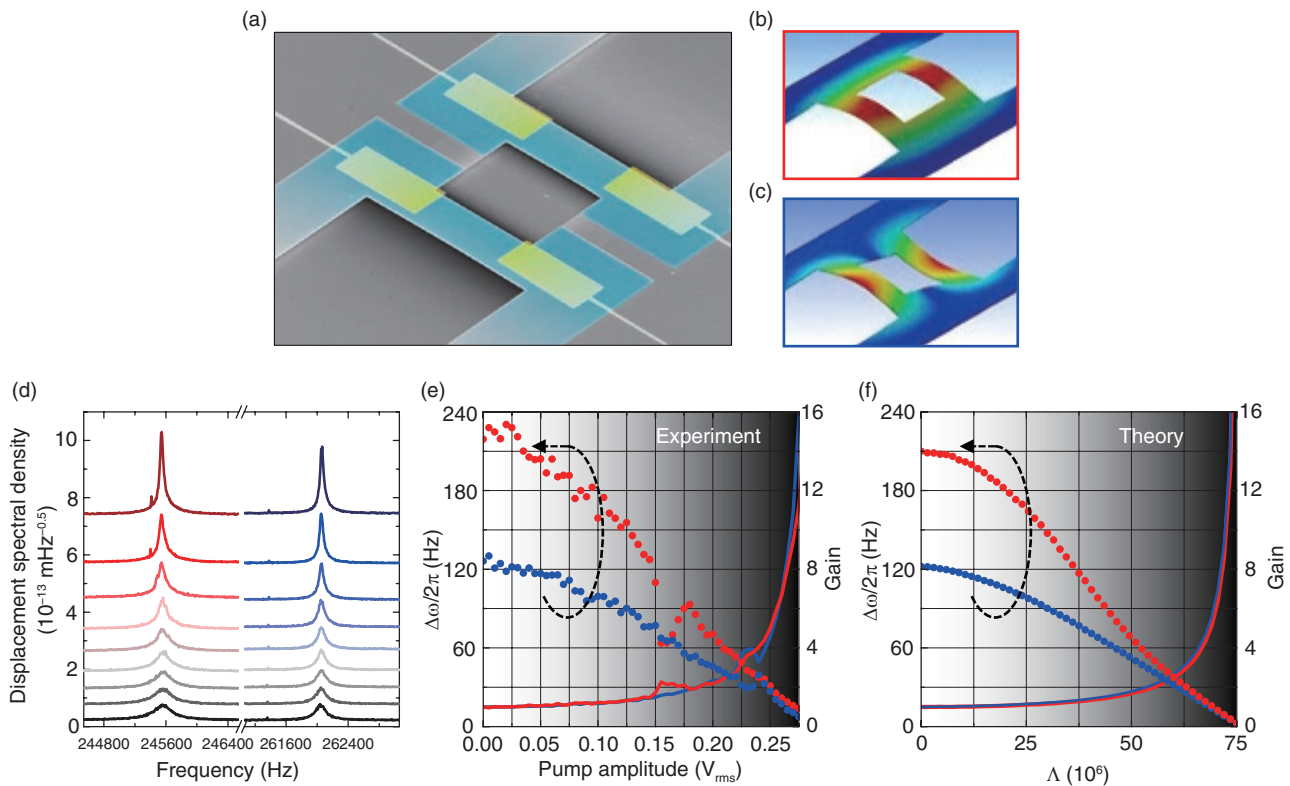
Fig. 2. (a) Scanning electron micrograph of the electromechanical resonator. Each mechanical element is 80 μm long, 20 μm wide, 800 nm thick, and is integrated with two piezoelectric transducers formed from a doped GaAs layer located 300 nm below the surface (blue) and a gold gate electrode (yellow). (b) and (c) The electromechanical resonator sustains symmetric (red) and asymmetric (blue) vibration modes composed from both mechanical elements that are strongly coupled via the two large overhangs between them. (d) Thermomechanical vibrations of both modes when driven by the Langevin force (black) and demodulated in a spectrum analyzer. Activating the non-degenerate parametric down-conversion via the piezoelectric transducers on the left mechanical element from 0 to 0.24 V_{rms} in 0.03 V_{rms} increments results in the thermomechanical fluctuations of both modes being amplified. (e) The corresponding amplification gain and spectral power bandwidth ($\Delta\omega$) with pump amplitude increments of 5 mV_{rms}. (f) This experimental response can be faithfully reproduced by numerically solving the Hamiltonian in eq. (1).

symmetric (S) and asymmetric (A), which are closely spaced in frequency. In addition, the spatial profiles of both modes have a large overlap yielding greater strain-mediated dispersive coupling between them. This combination of enhanced dispersive coupling and a smaller frequency separation, almost tending towards a degenerate configuration, is designed to yield larger non-degenerate parametric amplification when the strain in this system is modulated at the sum frequency of both modes.

The piezoelectric effect is utilized in order to modulate the strain in the mechanical modes. By design, the electromechanical system is fabricated from gallium arsenide (GaAs), which is piezoelectrically active [6, 49]. The piezoelectric transducer is formed from a GaAs conducting layer and a gold gate electrode sandwiching an insulating GaAs layer, which is integrated directly into the mechanical element, as shown in Fig. 2(a). Application of an electric field across the transducer generates in-plane strain that can nonlinearly modulate the spring constants and hence, the frequencies of the modes supported by the mechanical system [6].

The Hamiltonian for the electromechanical system in this configuration can then be expressed as:

$$H = \sum_{n=S}^{A} \left(P_n^2 / 2m_n + m_n \omega_n^2 Q_n^2 / 2 \right) + \Lambda Q_S Q_A \cos(\omega_p t). \quad (1)$$

Here, P and Q are the canonical momentum and conjugate position of the symmetric and asymmetric modes with mass m and natural frequency ω_n, where the summation expresses their kinetic and potential

energies. The last term describes the piezoelectrically activated non-degenerate parametric down-conversion with amplitude Λ and frequency $\omega_p = \omega_S + \omega_A$, that is, at the sum frequency of the modes in question. Ostensibly, this Hamiltonian is classical, which superficially suggests the unavailability of two-mode squeezing. However, the last term is analogous to $g_0\sqrt{n}\,(ab + a^\dagger b^\dagger)$ for non-degenerate parametric down-conversion in cavity electromechanical systems in the frame rotating at the pump frequency, where a, b and $a^\dagger$, $b^\dagger$ are the annihilation and creation operators for the mechanical and microwave resonators, respectively, and g_0 is the intrinsic coupling rate between them that is enhanced by n pump photons [41–43]. Consequently, Λ encapsulates the intrinsic coupling rate between the symmetric and asymmetric modes that is enhanced by the piezoelectric pump, which results in phonon pairs being simultaneously generated in them [50]. Remarkably, these phonons should still be correlated, even if the two modes are thermalized and far from their quantum ground state, as long as their generation rate exceeds the rate at which they are dissipated from both modes [46].

4. Results

4.1 Non-degenerate parametric amplification

To ascertain if this latter expectation is satisfied, the thermomechanical fluctuations of both modes are spectrally measured via optical interferometry with a 3-µW helium neon laser probing the right element of the electromechanical system at room temperature and in a high vacuum [46]. This measurement reveals the modes with natural frequencies $\omega_S/2\pi \approx 246$ kHz and $\omega_A/2\pi \approx 262$ kHz with quality factors $Q_n = \omega_n/\Delta\omega_n$ of 1300 and 2200, respectively, from their spectral power bandwidth $\Delta\omega_n$, and it corresponds to both modes sustaining $> 10^7$ phonons. Next non-degenerate parametric down-conversion is activated, and the thermal fluctuations from both modes are monitored. The results of this measurement shown in **Fig. 2(d)** indicate that as the piezoelectric pump voltage is increased, the thermomechanical fluctuations of both modes are enhanced. This amplification can be referenced to the bare thermal fluctuations, and it yields gains of more than 20 dB that are accompanied by a narrowing of the power spectral bandwidth from both modes, as detailed in **Fig. 2(e)**. At the largest pump amplitudes (>0.275 V$_{rms}$) $\Delta\omega_n \to 0$, resulting in both modes undergoing non-degenerate parametric resonance [51, 52].

To confirm these experimental observations, the equations of motion for both modes are extracted from the Hamiltonian in eq. (1) and reformulated in their rotating frames with $Q_n = X_n \cos(\omega_n t) + Y_n \sin(\omega_n t)$, where X_n and Y_n are the slowly varying in-phase and quadrature components. The resultant equations are then numerically solved, as detailed in a previous study [46], and they faithfully reproduce the experimental response as shown in **Fig. 2(f)**, thus verifying that the amplification can be ascribed to non-degenerate parametric down-conversion.

The temporal dynamics of this amplification can also be acquired by mixing the output from the optical interferometer with local oscillators locked onto the resonances of both modes and then demodulated in a phase sensitive detector (PSD). The PSD samples the random displacement fluctuations of the mechanical modes driven by the thermal Langevin force at a rate of 50 ms over a period of 300 s. This yields four time series for the in-phase and quadrature components of both modes, enabling their phase portraits to be constructed as shown in **Figs. 3(a)** and **3(b)**. This measurement reveals that the thermal fluctuations of both modes are random and uncorrelated as evidenced by their circular distribution in phase space, indicating all vibration phases are equally likely (red and blue points in Figs. 3(a) and 3(b)). Repeating this measurement with a pump amplitude of 0.25 V$_{rms}$ confirms that the fluctuations in both modes are enhanced via the non-degenerate parametric down-conversion while retaining their random nature; namely, this amplification is phase insensitive [53].

The observed amplification arises from phonons generated in both modes from the same process; hence, their fluctuations should be correlated. To that end, the cross quadratures are constructed in phase space from the in-phase component of the symmetric mode and the quadrature component of the asymmetric mode and vice versa, as shown in **Figs. 3(c)** and **3(d)**. This unveils *squeezed* distributions where a particular phase orientation is amplified while the perpendicular phase is de-amplified. The narrowness of this distribution implies the existence of correlations between the symmetric and asymmetric modes (if no correlations between the modes existed, the cross-quadratures would yield circular and therefore random distributions) and is the signature feature of a two-mode squeezed state [54, 55]. However, in order to quantitatively verify the existence of correlations between the modes, two criteria need to be satisfied. First, the two-mode squeezed distributions should exhibit smaller fluctuations than the bare distributions; otherwise, the correlations would be washed
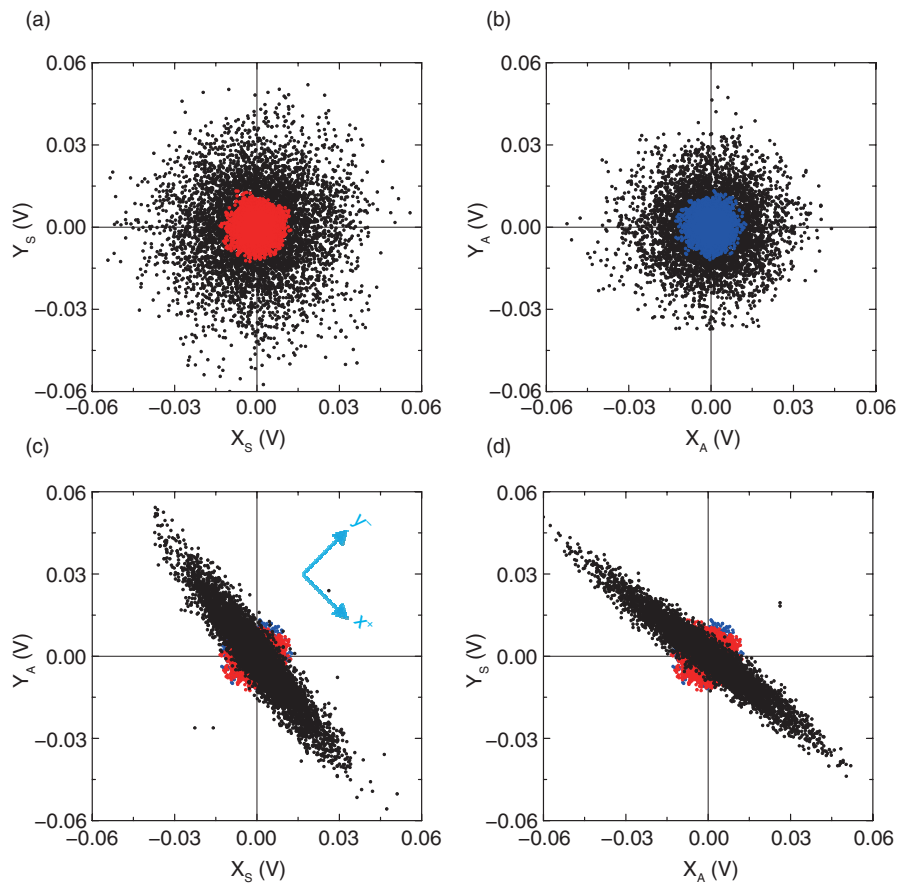
Fig. 3. (a) and (b) Bare thermal fluctuations of symmetric (red points) and asymmetric (blue points) modes respectively acquired temporally on-resonance and projected in phase space via the in-phase (X) and quadrature (Y) axes. Activating the non-degenerate parametric down-conversion with a pump amplitude of 0.25 V_{rms} amplifies the fluctuations of both modes while preserving their phase (black points). (c) and (d) The cross quadratures of the pumped measurement X_S: Y_A and X_A: Y_S respectively reconstructed in phase space (black points) reveal two-mode squeezed states have been generated that exhibit fluctuations in a particular phase which are below the bare thermal fluctuations from both modes (red and blue points), while fluctuations in the orthogonal phase are amplified.

out by the thermomechanical noise from both modes. Second, the correlations should be statistically evident from the temporal data used to construct the phase portraits.

4.2 Squeezing below the thermal noise

The former condition can in fact be visibly confirmed by examining the data in Figs. 3(c) and 3(d), which indicate that the de-amplified phases encompass a narrower distribution than the bare thermal motion from both modes. Quantitatively, new axes x^+ and y^- are introduced, as shown in the inset to Fig. 3(c), onto which the counts in the squeezed distributions are projected, as shown in **Figs. 4(b)** and **4(c)** [46]. This reveals Gaussian profiles with a zero mean displacement, as expected from random fluctuations

driven by the thermal Langevin force, whose full width at half maximum can be used to extract their standard deviation σ. This analysis is repeated as a function of non-zero pump amplitude for the cross quadratures X_S: Y_A and X_A: Y_S, where the standard deviations for both the amplified and de-amplified phases corresponding to the x^+ and y^- axes are determined as shown in **Fig. 4(a)**. Naturally, as the pump amplitude is increased, the fluctuations along the x^+ axis are amplified, while concurrently they are reduced along the y^- axis where entropy is conserved in this process. The gain can then be extracted by normalizing with the narrowest standard deviation from the bare thermal distribution (corresponding to the quadrature component of the asymmetric mode). While this slightly overestimates the amplification, it
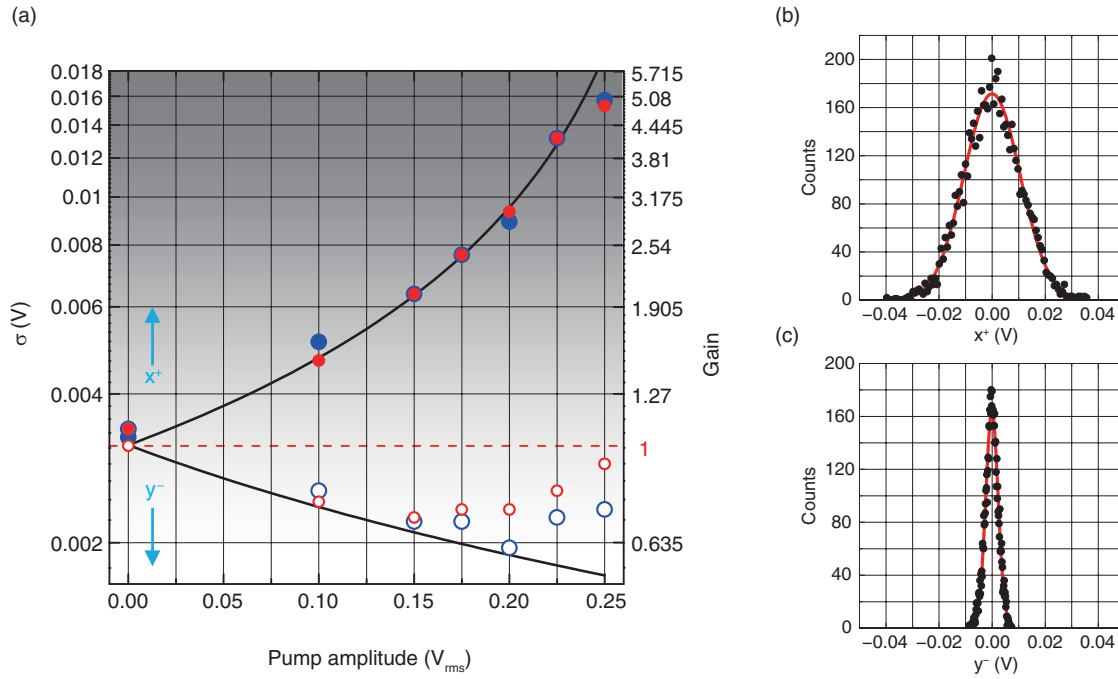
(a)

(b)

(c)

Fig. 4.   (a) Standard deviation (σ) of the fluctuation distribution from the two-mode squeezed states in phase space when projected onto the rotated axes detailed in Fig. 3(c) as a function of pump amplitude. The amplification along the x^+ axis is concurrently accompanied with a de-amplification along the y^- axis. The red (blue) points correspond to the X_S: Y_A (X_A: Y_S) cross-quadrature reconstruction, and the solid black lines denote the ideal theoretical response extracted from the Hamiltonian in eq. (1). Also shown is the gain normalized to the narrowest bare thermal distribution highlighted by the dashed red line. (b) and (c) The phase portrait from X_A: Y_S with a pump amplitude of 0.2 V_{rms} projected onto the rotated x^+ and y^- axes reveals Gaussian distributions (points), and the corresponding least squares Gaussian fit (line) enables their standard deviation to be extracted as quantified above.

is consistent with the spectral response detailed in Fig. 2(e). However, this normalizing reference underestimates the de-amplification, yielding a conservative 5-dB suppression of the mechanical fluctuations below the thermal level where further reduction is inhibited by noise from the piezotransducers at large pump voltages [46].

To confirm these observations, the thermal Langevin force, in the equations of motion extracted from the Hamiltonian in eq. (1), is decomposed into in-phase and quadrature components [16, 45, 46]. This leads to a new set of composite variables for the squeezed distributions given by $X_S \pm Y_A$ and $X_A \pm Y_S$ that naturally correspond to the rotated axes x^+ and y^- introduced in Fig. 3(c) [37, 46]. The pump-induced gain in the fluctuations of these composite variables can then be extracted from the Langevin correlation function, yielding the solid lines in Fig. 4(a). This verifies that the noise reduction below the thermal level arises from the non-degenerate parametric down-conversion.

4.3   Correlation coefficient

The statistical correlations between the time series for the in-phase and quadrature components from both modes can be analyzed via the absolute correlation coefficient $|cov(Z_iZ_i) / \sigma_{Z_i}\sigma_{Z_i}|$, where $Z_i \in \{X_S, Y_S, X_A, Y_A\}$ and the numerator describes the covariance. Analyzing the bare thermal fluctuations (shown by the red and blue points in Figs. 3(a) and 3(b)) enables a correlation coefficient matrix to be constructed, yielding 16 permutations. However, only 10 combinations are relevant due to their symmetry, as shown in **Fig. 5(a)**. This reveals that the diagonal elements corresponding to the auto-correlations of Z_i yield a coefficient of 1, which indicates that they are perfectly correlated with themselves, as one would expect. However, all the off-diagonal elements yield a coefficient of 0, indicating an absence of correlations; for instance $|cov(X_SY_S) / \sigma_{X_S}\sigma_{Y_S}| \approx 0$, which is unsurprising, as this maps the circular distribution in Fig. 3(a) corresponding to the random uncorrelated fluctuations driven by the thermal Langevin force.
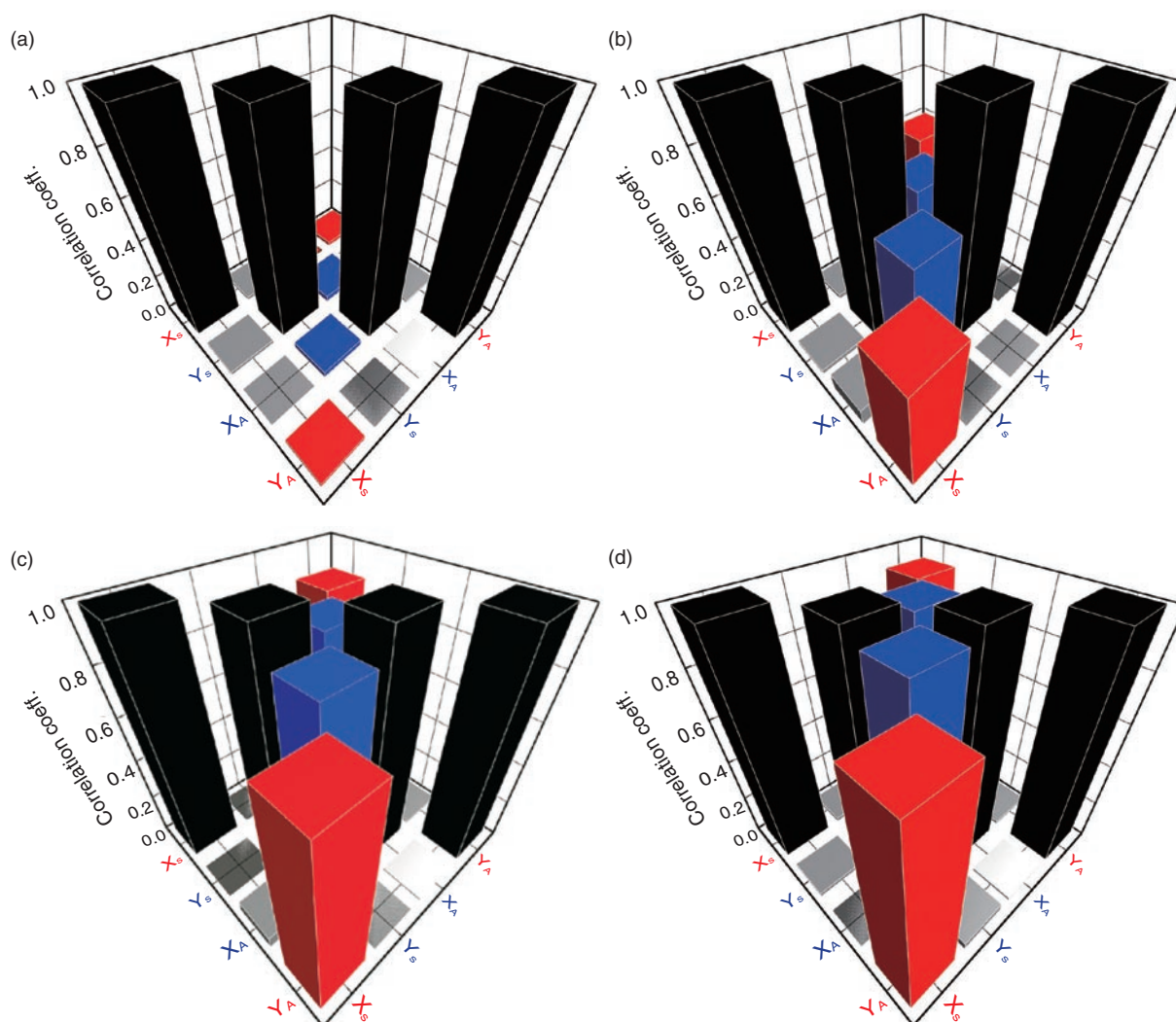
Fig. 5. (a)–(d) Correlation coefficient matrix as a function of pump amplitude from 0, 0.1, 0.2, and 0.25 V_{rms}, respectively. The diagonal elements (black) correspond to auto-correlations, which naturally yield a coefficient of 1. The off-diagonal elements from X_S: Y_S and X_A: Y_A yield a coefficient of 0, which corresponds to the uncorrelated random fluctuations corresponding to the circular distributions in phase space as shown in Figs. 3(a), 3(b) and are independent of the pump amplitude. In contrast, the off-diagonal elements from the two-mode squeezed states X_S: Y_A (red) and X_A: Y_S (blue) converge towards a coefficient of 1 as the pump amplitude is increased, implying that the two mechanical vibration modes have become perfectly correlated via the simultaneous generation of signal and idler phonons in pairs from the non-degenerate parametric down-conversion.

Repeating this analysis as a function of pump amplitude ranging from 0.1, 0.2, and 0.25 V_{rms} as shown in **Figs. 5(b)–(d)** indicates that the auto-correlations in the diagonal elements remain perfect. However, the off-diagonal elements corresponding to the cross-quadrature phase portraits in Figs. 3(c) and 3(d), namely, X_S: Y_A and X_A: Y_S, emerge and converge towards 1, which implies that the two modes have become perfectly correlated from the simultaneous generation of phonon pairs via non-degenerate parametric down-conversion.

5. Implications and outlook

Although both mechanical modes sustain more than 10 million phonons at room temperature, the simultaneous generation of phonon pairs from the non-degenerate parametric down-conversion occurs at a rate that renders their fluctuations indistinguishable. Thus, remarkably, these thermalized macroscopic

mechanical vibration modes become perfectly entwined in a two-mode squeezed state. Ultimately, if the modes can be operated with phonon populations of much less than one, then these correlations will manifest themselves in a macroscopic all-mechanical entanglement [10, 42, 56]. Access to such a state would be extremely tantalizing, as it would provide an invaluable platform to investigate the absence of quantum mechanical phenomena in our everyday classical world [20].

From a technological point of view, these results herald the emergence of quantum optics to phononics, and thus, concepts such as quantum cryptography and optical quantum computing could potentially be harnessed with sound in a microchip. More immediately, the possibility of even greater squeezing becomes available via more strongly coupled modes in conjunction with a more efficient piezoelectric pump [57]. This enhanced squeezing would not only offer detectors that could operate below the quantum limit, yielding unprecedented sensitivities for metrological applications [33] and fundamental science [58], but it could even be utilized to create room temperature entanglements, thus bringing *quantum sound* into a more technologically accessible regime [59].

References

[1]   M. L. Roukes, "Nanoelectromechanical Systems Face the Future," Physics World, Vol. 14, No. 2, pp. 25–31, 2001.

[2]   K. L. Ekinci and M. L. Roukes, "Nanoelectromechanical Systems," Rev. Sci. Instrum., Vol. 76, No. 6, p. 061101, 2005.

[3]   D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui, "Single Spin Detection by Magnetic Resonance Force Microscopy," Nature, Vol. 430, No. 6997, pp. 329–332, 2004.

[4]   J. Chaste, A. Eichler, J. Moser, G. Ceballos, R. Rurali, and A. Bachtold, "A Nanomechanical Mass Sensor with Yoctogram Resolution," Nature Nanotech., Vol. 7, No. 5, pp. 301–304, 2012.

[5]   J. L. Arlett, E. B. Myers, and M. L. Roukes, "Comparative Advantages of Mechanical Biosensors," Nature Nanotech., Vol. 6, No. 4, pp. 203–215, 2011.

[6]   I. Mahboob and H. Yamaguchi, "Bit Storage and Bit Flip Operations in an Electromechanical Oscillator," Nature Nanotech., Vol. 3, No. 5, pp. 275–279, 2008.

[7]   R. L. Badzey and P. Mohanty, "Coherent Signal Amplification in Bistable Nanomechanical Oscillators by Stochastic Resonance," Nature, Vol. 437, No. 7061, pp. 995–998, 2005.

[8]   I. Mahboob, E. Flurin, K. Nishiguchi, A. Fujiwara, and H. Yamaguchi, "Interconnect-free Parallel Logic Circuits in a Single Mechanical Resonator," Nature Commun., Vol. 2, p. 198, 2011.

[9]   M. L. Roukes, "Mechanical Computation, Redux?", IEEE IEDM Technical Digest, pp. 539–542, 2004.

[10] A. D. O'Connell, M. Hofheinz, M. Ansmann, R. C. Bialczak, M. Lenander, E. Lucero, M. Neeley, D. Sank, H. Wang, M. Weides, J. Wenner, J. M. Martinis, and A. N. Cleland, "Quantum Ground State and Single-phonon Control of a Mechanical Resonator," Nature, Vol. 464, No. 7289, pp. 697–703, 2010.

[11] J. D. Teufel, T. Donner, D. Li, J. W. Harlow, M. S. Allman, K. Cicak, A. J. Sirois, J. D. Whittaker, K. W. Lehnert, and R. W. Simmonds, "Sideband Cooling of Micromechanical Motion to the Quantum

[12] J. Chan, T. P. Mayer Alegre, A. H. Safavi-Naeini, J. T. Hill, A. Krause, S. Gröblacher, M. Aspelmeyer, and O. Painter, "Laser Cooling of a Nanomechanical Oscillator into Its Quantum Ground State," Nature, Vol. 478, No. 7367, pp. 89–92, 2011.

[13] T. J. Kippenberg and K. J. Vahala, "Cavity Optomechanics: Back-Action at the Mesoscale," Science, Vol. 321, No. 5893, pp. 1172–1176, 2008.

[14] M. Aspelmeyer, P. Meystre, and K. Schwab, "Quantum Optomechanics," Phys. Today, Vol. 65, No. 7, pp. 29–35, 2012.

[15] I. S. Grudinin, H. Lee, O. Painter, and K. J. Vahala, "Phonon Laser Action in a Tunable Two-level System," Phys. Rev. Lett., Vol. 104, No. 8, p. 083901, 2010.

[16] I. Mahboob, K. Nishiguchi, A. Fujiwara, and H. Yamaguchi, "Phonon Lasing in an Electromechanical Resonator," Phys. Rev. Lett., Vol. 110, No. 12, p. 127202, 2013.

[17] J. B. Khurgin, "Viewpoint: Phonon Lasers Gain a Sound Foundation," Physics, Vol. 3, p. 16, 2010.

[18] J. T. Mendonça, "Viewpoint: Lasers of Pure Sound," Physics, Vol. 6, p. 32, 2013.

[19] T. A. Palomaki, J. D. Teufel, R. W. Simmonds, and K. W. Lehnert, "Entangling Mechanical Motion with Microwave Fields," Science, Vol. 342, No. 6159, pp. 710–713, 2013.

[20] W. H. Zurek, "Decoherence and the Transition from Quantum to Classical," Phys. Today, Vol. 44, No. 10, pp. 36–44, 1991.

[21] W. Marshall, C. Simon, R. Penrose, and D. Bouwmeester, "Towards Quantum Superpositions of a Mirror," Phys. Rev. Lett., Vol. 91, No. 13, p. 130401, 2003.

[22] D. C. Burnham, and D. L. Weinberg, "Observation of Simultaneity in Parametric Production of Optical Photon Pairs," Phys. Rev. Lett., Vol. 25, No. 2, pp. 84–87, 1970.

[23] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-mechanical Description of Physical Reality Be Considered Complete?", Phys. Rev., Vol. 47, No. 10, pp. 777–780, 1935.

[24] M. D. Reid and P. D. Drummond, "Quantum Correlations of Phase in Nondegenerate Parametric Oscillation," Phys. Rev. Lett., Vol. 60, No. 26, pp. 2731–2733, 1988.

[25] Z. Y. Ou and L. Mandel, "Violation of Bell's Inequality and Classical Probability in a Two-photon Correlation Experiment," Phys. Rev. Lett., Vol. 61, No. 1, pp. 50–53, 1988.

[26] Y. H. Shih and C. O. Alley, "New Type of Einstein-Podolsky-Rosen-Bohm Experiment Using Pairs of Light Quanta Produced by Optical Parametric Down Conversion," Phys. Rev. Lett., Vol. 61, No. 26, pp. 2921–2924, 1988.

[27] J. S. Bell, "On the Einstein Podolsky Rosen Paradox," Physics, Vol. 1, No. 3, pp. 195–200, 1964.

[28] P. R. Tapster, J. G. Rarity, and P. C. M. Owens, "Violation of Bell's Inequality over 4 km of Optical Fiber," Phys. Rev. Lett., Vol. 73, No. 14, pp. 1923–1926, 1994.

[29] A. K. Ekert, "Quantum Cryptography Based on Bell's Theorem," Phys. Rev. Lett., Vol. 67, No. 6, pp. 661–663, 1991.

[30] E. Knill, R. Laflamme, and G. J. Milburn, "A Scheme for Efficient Quantum Computation with Linear Optics," Nature, Vol. 409, No. 6816, pp. 46–52, 2000.

[31] D. Bouwmeester, J. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, "Experimental Quantum Teleportation," Nature, Vol. 390, No. 6660, pp. 575–579, 1997.

[32] A. N. Boto, P. Kok, D. S. Abrams, S. L. Braunstein, C. P. Williams, and J. P. Dowling, "Quantum Interferometric Optical Lithography: Exploiting Entanglement to Beat the Diffraction Limit," Phys. Rev. Lett., Vol. 85, No. 13, pp. 2733–2736, 2000.

[33] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum-enhanced Measurements: Beating the Standard Quantum Limit," Science, Vol. 306, No. 5700, pp. 1330–1336, 2004.

[34] J. Q. You and Franco Nori, "Atomic Physics and Quantum Optics Using Superconducting Circuits," Nature, Vol. 474, No. 7353, pp. 589–597, 2011.

[35] B. Yurke, P. G. Kaminsky, R. E. Miller, E. A. Whittaker, A. D. Smith, A. H. Silver, and R. W. Simon, "Observation of 4.2-K Equilibrium-

noise Squeezing via a Josephson-parametric Amplifier," Phys. Rev. Lett., Vol. 60, No. 9, pp. 764–767, 1998.

[36] M. A. Castellanos-Beltran, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, "Amplification and Squeezing of Quantum Noise with a Tunable Josephson Metamaterial," Nature Phys., Vol. 4, No. 12, pp. 929–931, 2008.

[37] C. Eichler, D. Bozyigit, C. Lang, M. Baur, L. Steffen, J. M. Fink, S. Filipp, and A. Wallraff, "Observation of Two-mode Squeezing in the Microwave Frequency Domain," Phys. Rev. Lett., Vol. 107, No. 11, p. 113601, 2011.

[38] E. Flurin, N. Roch, F. Mallet, M. H. Devoret, and B. Huard, "Generating Entangled Microwave Radiation over Two Transmission Lines," Phys. Rev. Lett., Vol. 109, No. 18, p. 183901, 2012.

[39] C. A. Regal, J. D. Teufel, and K. W. Lehnert, "Measuring Nanomechanical Motion with a Microwave Cavity Interferometer," Nature Phys., Vol. 4, No. 7, pp. 555–560, 2008.

[40] T. Rocheleau, T. Ndukum, C. Macklin, J. B. Hertzberg, A. A. Clerk, and K. C. Schwab, "Preparation and Detection of a Mechanical Resonator Near the Ground State of Motion," Nature, Vol. 463, No. 7277, pp. 72–75, 2009.

[41] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, "Cavity Optomechanics," Rev. Mod. Phys., Vol. 86, No. 4, pp. 1391–1452, 2014.

[42] T. A. Palomaki, J. D. Teufel, R. W. Simmonds, and K. W. Lehnert, "Entangling Mechanical Motion with Microwave Fields," Science, Vol. 342, No. 6159, pp. 710–713, 2013.

[43] S. G. Hofer, W. Wieczorek, M. Aspelmeyer, and K. Hammerer, "Quantum Entanglement and Teleportation in Pulsed Cavity Optomechanics," Phys. Rev. A, Vol. 84, No. 5, p. 052327, 2011.

[44] H. J. R. Westra, M. Poot, H. S. J. van der Zant, and W. J. Venstra, "Nonlinear Modal Interactions in Clamped-Clamped Mechanical Resonators," Phys. Rev. Lett., Vol. 105, No. 11, p. 117205, 2010.

[45] I. Mahboob, K. Nishiguchi, H. Okamoto, and H. Yamaguchi, "Phonon-cavity Electromechanics," Nature Phys., Vol. 8, No. 5, pp. 387–392, 2012.

[46] I. Mahboob, H. Okamoto, K. Onomitsu, and H. Yamaguchi, "Two-mode Thermal-noise Squeezing in an Electromechanical Resonator," Phys. Rev. Lett., Vol. 113, No. 16, p. 167203, 2014.

[47] D. Rugar and P. Grütter, "Mechanical Parametric Amplification and Thermomechanical Noise Squeezing," Phys. Rev. Lett., Vol. 67, No. 6, pp. 699–702, 1991.

[48] H. Okamoto, N. Kitajima, K. Onomitsu, R. Kometani, S. Warisawa, S. Ishihara, and H. Yamaguchi, "High-sensitivity Charge Detection Using Antisymmetric Vibration in Coupled Micromechanical Oscilla-tors," Appl. Phys. Lett., Vol. 98, No. 1, p. 014103, 2011.

[49] S. C. Masmanidis, R. B. Karabalin, I. De Vlaminck, G. Borghs, M. R. Freeman, and M. L. Roukes, "Multifunctional Nanomechanical Systems via Tunably Coupled Piezoelectric Actuation," Science, Vol. 317, No. 5839, pp. 780–783, 2007.

[50] H. Okamoto, A. Gourgout, C. Chang, K. Onomitsu, I. Mahboob, E. Y. Chang, and H. Yamaguchi, "Coherent Phonon Manipulation in Coupled Mechanical Resonators," Nature Phys., Vol. 9, No. 8, pp. 480–484, 2013.

[51] J. D. Teufel, J. W. Harlow, C. A. Regal, and K. W. Lehnert, "Dynamical Backaction of Microwave Fields on a Nanomechanical Oscillator," Phys. Rev. Lett., Vol. 101, No. 19, p. 197203, 2008.

[52] F. Massel, T. T. Heikkilä, J. M. Pirkkalainen, S. U. Cho, H. Saloniemi, P. J. Hakonen, and M. A. Sillanpää, "Microwave Amplification with Nanomechanical Resonators," Nature, Vol. 480, No. 7377, pp. 351–354, 2011.

[53] N. Bergeal, F. Schackert, M. Metcalfe, R. Vijay, V. E. Manucharyan, L. Frunzio, D. E. Prober, R. J. Schoelkopf, S. M. Girvin, and M. H. Devoret, "Phase-preserving Amplification Near the Quantum Limit with a Josephson Ring Modulator," Nature, Vol. 465, No. 7294, pp. 64–68, 2010.

[54] C. C. Gerry and P. L. Knight, "Introductory Quantum Optics," Cambridge University Press, Cambridge, England, 2005.

[55] D. F. Walls and G. J. Milburn, "Quantum Optics," Springer, 2008.

[56] J. R. Johansson, N. Lambert, I. Mahboob, H. Yamaguchi, and F. Nori, "Entangled-state Generation and Bell Inequality Violations in Nanomechanical Resonators," Phys. Rev. B, Vol. 90, No. 17, p. 174307, 2014.

[57] H. Vahlbruch, M. Mehmet, S. Chelkowski, B. Hage, A. Franzen, N. Lastzka, S. Goßler, K. Danzmann, and R. Schnabel, "Observation of Squeezed Light with 10-dB Quantum-noise Reduction," Phys. Rev. Lett., Vol. 100, No. 3, p. 033602, 2008.

[58] The LIGO Scientific Collaboration, "Enhanced Sensitivity of the LIGO Gravitational Wave Detector by Using Squeezed States of Light," Nature Photon., Vol. 7, No. 8, pp. 613–619, 2013.

[59] F. Galve, L. A. Pachón, and D. Zueco, "Bringing Entanglement to the High Temperature Limit," Phys. Rev. Lett., Vol. 105, No. 18, p. 180501, 2010.

Imran Mahboob
Senior Research Scientist, Distinguished Researcher, Hybrid Nanostructure Physics Research Group, NTT Basic Research Laboratories.
He received an MPhys. in theoretical physics from the University of Sheffield, UK, in 2001 and a Ph.D. in physics from the University of Warwick, UK, in 2004. He joined NTT Basic Research Laboratories in 2005 and has been researching electromechanical systems to harness their nonlinear dynamics to develop phonon-based information technologies. He is a member of the Institute of Physics (IoP) and the American Physical Society (APS).

Hajime Okamoto
Senior Research Scientist, Hybrid Nanostructure Physics Research Group, NTT Basic Research Laboratories.
He received a B.E., M.E., and a Ph.D. in materials science and engineering from Waseda University, Tokyo, in 1998, 2000, and 2004, respectively. He joined NTT Basic Research Laboratories in 2004 to research and develop semiconductor opto/electro-mechanical devices. He received an Outstanding Paper Award from the Japan Society of Applied Physics (JSAP) (2010), SSDM2009 Paper Award (2010), and the Young Scientists' Prize (2014) from the Minister of Education, Culture, Sports, Science and Technology (MEXT). He is a member of JSAP.

Hiroshi Yamaguchi
Executive Manager of Quantum and Nano Device Research, Group Leader of Hybrid Nanostructure Physics Research Group, Senior Distinguished Researcher, NTT Basic Research Laboratories.
He received a B.E. and a M.S. in physics and a Ph.D. in engineering from Osaka University in 1984, 1986 and 1993, respectively. He joined NTT Basic Research Laboratories in 1986 and studied compound semiconductor surfaces using electron diffraction and scanning tunneling microscopy. His current interests are micro/nanomechanical devices using semiconductor heterostructures. He was a visiting research fellow at Imperial College, University of London, UK, during 1995–1996 and a visiting researcher at the Paul Drude Institute, Germany, in 2003. He has been a guest professor at Tohoku University since 2006 and has served on more than 40 committees of academic societies and international conferences. He has received numerous awards including a JSAP Fellowship (2013), Commendation for Science and Technology by MEXT (2013), Inoue Prize for Science (2012), IoP Fellowship (2011), SSDM2009 Paper Award (2010), MNC2008 Outstanding Paper Award (2009), and Paper Awards from JSAP (1989, 2004, 2010). He is a member of JSAP, the Physical Society of Japan, IoP, APS, and the Institute of Electrical and Electronics Engineers.

Regular Articles

Improving User Capacity and Disaster Recovery Time in IP Telephone Service Systems

Hiroshi Shibata, Kouki Minamida, Hiroshi Miyao, Takashi Nambu, and Toru Takahashi

Abstract

The NTT Group is working on improving Internet protocol (IP) telephone networks and solving certain problems associated with them. Specifically, efforts are underway to improve the accommodation rate, reduce network operating costs when the number of network users increases, and reduce service recovery time when a network is damaged because of a natural disaster. To resolve these problems, we developed a system architecture that simplifies the operation of IP telephone networks and uses network equipment much more efficiently. This is possible with a subscriber data management server. We describe here our new network architecture, the mechanism to solve these problems, and the effect it has on IP telephone networks.

Keywords: IP telephone, disaster recovery, call session control function server, home subscriber server

1. Introduction

Essential social infrastructures such as electric power systems, water systems, and communication systems were destroyed over a huge area after the Great East Japan Earthquake struck on March 11, 2011, and these systems could not be used for a long time after that. Specifically, 29,000 mobile communications base stations and 1.9 million fixed telephone lines could not be used. It took about three days to recover telephone services provided by the NTT Group [1]. As a social infrastructure provider, the NTT Group learned from this experience that network operators must ensure consistent availability of services and early recovery of services even after wide-area disasters.

The devices used for Internet protocol (IP) telephone service range from smartphones to conventional fixed telephones. Therefore, IP telephone services must accommodate frequent terminal location registrations and high-frequency use of each terminal.

We describe here the new IP telephone network architecture that will improve the flexibility of subscriber accommodation, and we explain the results of a study on how the new network architecture reduces the time to recover from a disaster.

2. Characteristics of and problems with IP telephone network architecture

Currently, the NTT Group offers a variety of IP telephone services such as Hikari Denwa and the 050 number voice over IP (VoIP) service. The IP Multimedia Subsystem (IMS) architecture [2–4] is partially adopted in these IP telephone services.

The IP telephone service was initially available for fixed telephones via VoIP routers for private use. Office use, such as the pilot telephone number service, also began in conventional fixed telephone networks. Therefore, we adopted a network architecture that efficiently performs a call control process; the correspondence between the call control server (the call session control function (CSCF)) and the user
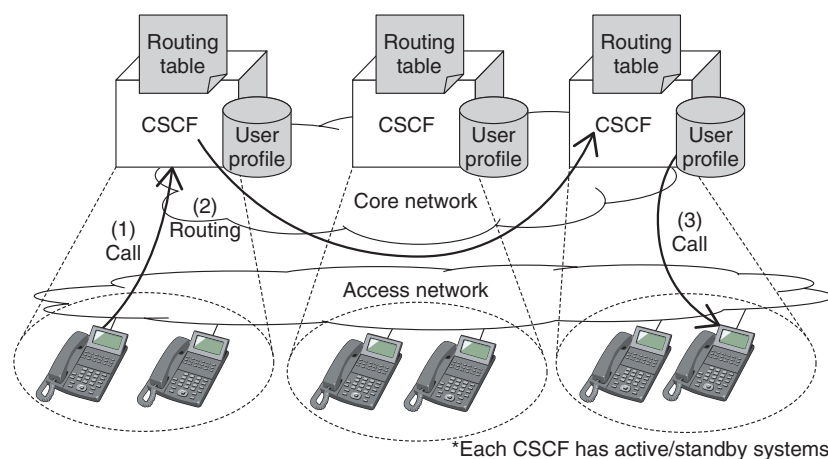
Fig. 1. Old IP telephone network architecture.

(telephone number) is fixed, and the CSCF is designed to accommodate each number range, the same as for the traditional public switched telephone network (PSTN). The CSCF holds the subscriber data (service contract information and authentication information) related to the telephone number to be accommodated [5]. Routing information to identify the CSCF that accommodates the number range is managed by each CSCF. Therefore, the CSCF that receives a call request from a terminal (indicated as (1) in **Fig. 1**) determines the destination CSCF that accommodates the terminal by using its own routing information (2), and transfers the call request from the terminal to the destination CSCF (3). In addition, to maintain reliability and non-interrupted service during system changes, each CSCF adopts an active and standby (ACT/SBY) dual system as a redundant system.

In extending the system to cope with an increase in the number of users and the amount of traffic, it is conventionally assumed that there is a proportional relationship between users and traffic; therefore, resources are designed after calculating the load from the traffic rate and average holding time to install additional equipment in accordance with the increase in traffic. However, because softphone users, for example, include many non-active users* who install applications on their terminal but do not use them frequently, there are cases in which the number of users is large, but the load in the IP telephone network is low. Therefore, the CPU (central processing unit) working rate is low at the maximum amount of subscriber data to be accommodated, and vice versa, so equipment resources cannot be fully used. As a result,

equipment utilization efficiency is low compared to the older IP telephone network (Problem 1).

For load-balancing, however, we need to change the routing information of each CSCF to identify the destination CSCF of incoming users, which is a significant operational burden (Problem 2).

In addition, with the current redundant configuration, both ACT/SBY systems could be damaged in the event of a large-scale disaster, so there is a problem concerning the availability of the social infrastructure (recovery time for both system failures) (Problem 3). We explain the details of this problem in the next section.

Our solution to these problems is to separate the subscriber data and routing information from the CSCF. The home subscriber server (HSS) manages subscriber data and routing information, so the relationship between telephone numbers and the CSCF is flexible. Therefore, the CSCF-HSS separation architecture can be adopted. It should be noted that upon separation, we take into account the impact on services specific to the conventional fixed telephone networks such as pilot telephone number and call-pick-up services (Problem 4).

3. Issues regarding service recovery time after natural disaster

We now explain the details of Problem 3. NTT Network Service Systems Laboratories has been working

---

* A non-active user is an IP telephone service subscriber who has not used an IP telephone recently for a given period (for example, a month). The other subscribers are active users.
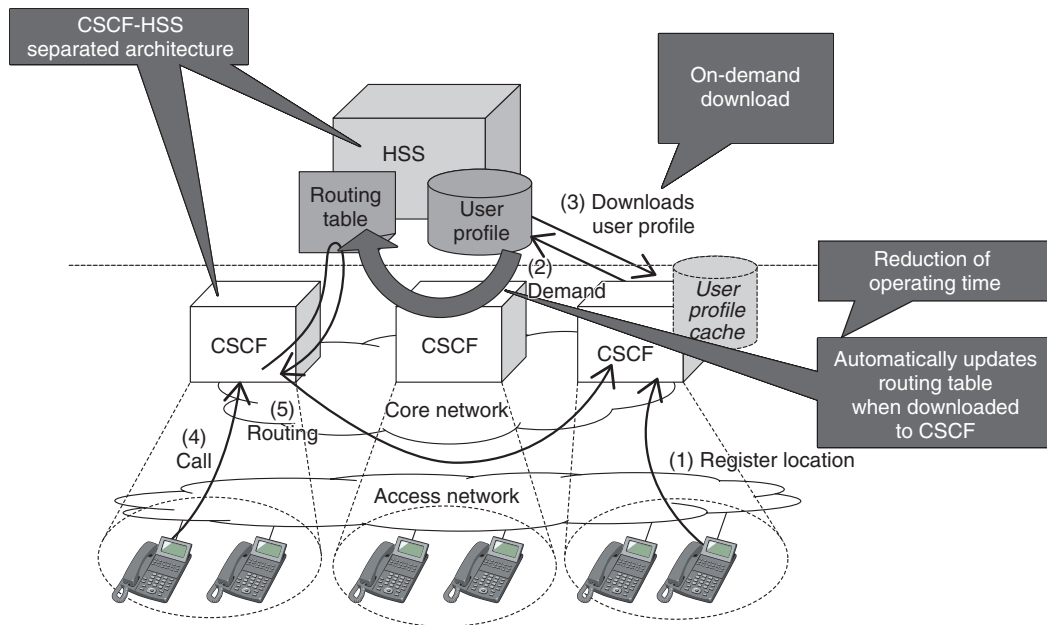
Fig. 2.   CSCF-HSS separated architecture.

to improve reliability during serious disasters by building on knowledge and experience gained in the aftermath of the 2011 Japan earthquake. With the conventional network configuration, a simultaneous failure of redundant configuration systems must be assumed after a serious disaster because the subscriber data contained in the CSCF are fixed. Therefore, we need to rebuild equivalent devices.

Specifically, the following procedure is required.
- Replace afflicted CSCF hardware.
- Restore operating system (OS) and applications.
- Restore subscriber data.

It takes several days to complete this procedure, which is itself an issue. (Note that the recovery time depends on the manpower and equipment available for the recovery work.)

4.   CSCF-HSS separated architecture

We introduce here the CSCF-HSS separated architecture to solve the above-mentioned problems.

4.1   Effective use of server resources for managing users

In the CSCF-HSS separated architecture (**Fig. 2**), HSSs manage the master data of user profiles, and the CSCF downloads user profile data on demand from the HSSs when the user equipment registers its loca-

tion information via an IP telephone. After that, the CSCF retains the user profile data as a cache during the validity period of the location information. With this mechanism, the CSCF does not retain the user profiles for non-active users who have not used an IP telephone for a certain period of time. Therefore, the CSCF can manage more user profiles of active users because it does not consume as much memory as the old architecture. Thus, we can use the CSCF's resources effectively to solve Problem 1.

4.2   Automatic updating of routing information

To update the routing information in order to detect the CSCF that manages a certain user when user profile data are moved from one CSCF to another, we introduce a mechanism in which an HSS automatically updates the routing table when a CSCF downloads user profile data from the HSS (Fig. 2). Therefore, it is not necessary for each CSCF to have a routing table because the CSCFs can detect the destination CSCF that manages the user profile data. Therefore, service operators do not need to update the routing information, and we can solve Problem 2.

4.3   Maintaining communication quality for fixed telephone service

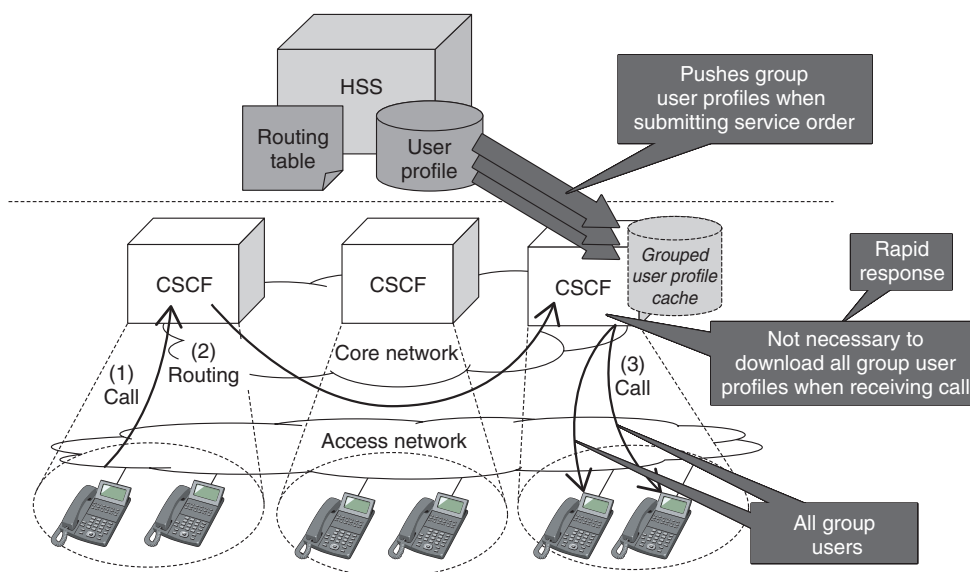We provide fixed IP telephone services for businesses and user-grouping services such as pilot

Fig. 3.   Rapid response by pushing all group user profiles to CSCFs.

telephone number or call pick-up services. In these services, we need to update the cache of user profile data in CSCFs when the master data are updated in the HSS by submitting or changing the service order. In the IMS standard models, a CSCF downloads user profile data from an HSS on demand when receiving a call; however, this incurs delays due to the downloading of all group member profile data. Thus, communication quality decreases.

Therefore, we extended an HSS's function to push user profile data to CSCFs defined in the IMS standard [6]. Our HSS simultaneously pushes all user profile data in a group. In our new architecture, CSCFs do not need to download all group member profile data on demand and can reduce the number of times the HSSs need to be accessed because the CSCFs already have all profile data in the group due to the extended push function and can respond rapidly. Thus, the communication quality of user-grouping services is maintained (**Fig. 3**) and we can solve Problem 4.

5.   New recovery procedure and effectiveness

We solve Problem 3 by creating recovery procedures for the new architecture in which HSSs are separated from CSCFs. We introduce two new recovery procedures. First, we describe the recovery procedure (pattern 1), in which the time to restore user profile data is reduced compared to the old procedure.

Second, we describe the procedure (pattern 2) in which we use server resources of other surviving CSCFs to recover the failed CSCFs. We describe these procedures and their effectiveness below.

Pattern 1: Other surviving CSCFs do NOT have enough unused server resources.
(1)   Recovery procedure
   Step 1: Replace completely failed CSCF hardware with new hardware.
   Step 2: Restore the OS and application software.
   After step 2 is completed, the rebuilt CSCFs can serve users that were prevented from using IP telephone services due to failed CSCFs by downloading user profile data on demand from the HSSs (**Fig. 4**).
(2)   Effectiveness
   We can reduce the recovery time by skipping unnecessary steps such as restoring user profile data in the rebuilt CSCFs.

Pattern 2: Other surviving CSCFs HAVE enough unused server resources and can serve users supported by failed CSCFs.
(1)   Recovery procedure
   Step 1: Reassign the affected users to other surviving CSCFs instead of the failed CSCFs.
   Specifically, we change the settings in the HSSs so that the surviving CSCFs serve the affected users instead of the failed CSCFs. After that, the routing tables in the HSS are updated automatically, so we do
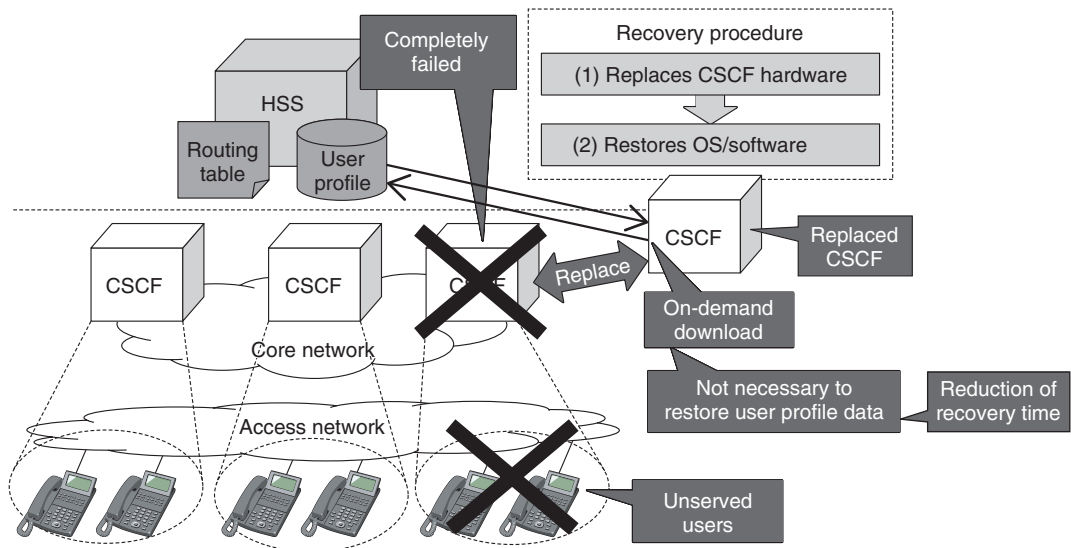
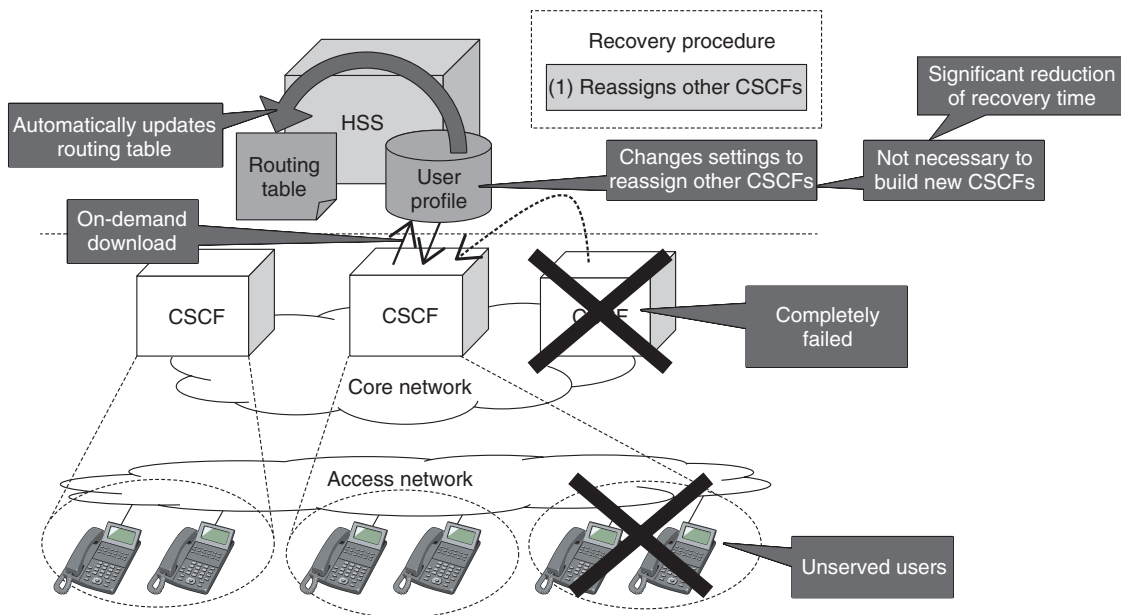Fig. 4.   CSCF recovery procedure in new architecture (pattern 1).

Fig. 5.   CSCF recovery procedure in new architecture (pattern 2).

NOT need to change the routing tables of all surviving CSCFs (**Fig. 5**).

(2)   Effectiveness

We can recover IP telephone services for affected users without building new CSCFs to replace the failed CSCFs, so the disaster recovery time drastically decreases.

The difference between the old and new IP telephone network architectures with respect to recovery time when redundant CSCFs fail completely is shown in **Fig. 6**. Specifically, recovery is very fast in pattern 2 because we can recover the services within several hours. It takes several days to recover services with the old architecture. Thus, we can solve Problem 3.
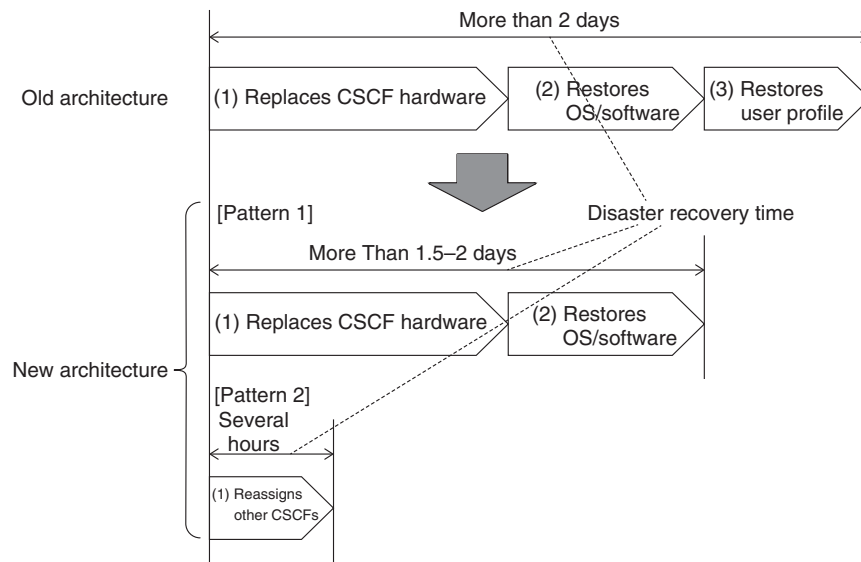
Fig. 6.   Improved disaster recovery time in new architecture.

6.   Future work

We plan to work on developing a much faster recovery method and will therefore investigate a recovery procedure that involves not only servers such as HSSs and CSCFs but also user equipment, transport network elements, and business support systems.

References

[1]   Ministry of Internal Affairs and Communications, "Part 1: Information and Communications in the Aftermath of the Great East Japan Earthquake," in The 2011 White Paper on Information and Communications in Japan.
http://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2011/part1.pdf

[2]   3GPP, "TS 23.228: IP Multimedia Subsystem (IMS); Stage 2 (Release 7)," 2007.

[3]   ETSI/TISPAN, "TS 182.006: IP Multimedia Subsystem (IMS); Stage 2 description," 2008.

[4]   ETSI/TISPAN, "ES 282 001: NGN Functional Architecture," 2009.

[5]   M. Miyasaka, N. Horikome, and K. Kishida, "Bandwidth Management and Control Technology in the NGN," NTT Technical Review, Vol. 7, No. 7, Jul. 2009.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200907sf5.html

[6]   3GPP, "TS 29.229: Cx and Dx Interfaces based on the Diameter Protocol; Protocol details (Release 5)," 2007.

Hiroshi Shibata
Research Engineer, NTT Network Service Systems Laboratories.
He received his B.S. and M.S. from Osaka University in 1994 and 1996, respectively. He joined NTT Network Service Systems Laboratories in 1996. His current R&D area is software engineering. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Takashi Nambu
Senior Research Engineer, NTT Network Service Systems Laboratories.
He received his B.E. and M.E. from the University of Tokyo in 1998 and 2000, respectively. He joined NTT Network Service Systems Laboratories in 2000 and engaged in VoIP system development. He is now engaged in R&D on improving software quality and productivity.

Kouki Minamida
Research Engineer, NTT Network Service Systems Laboratories.
He received his B.S. and M.S. from the University of Tokyo in 1992 and 1994, respectively. He joined NTT in 1994. His current R&D area is software engineering. He is a member of IEICE.

Toru Takahashi
Senior Research Engineer, Supervisor, NTT Network Service Systems Laboratories.
He received his B.E. from Tokyo University of Science in 1991 and joined NTT Communication Switching Laboratories the same year. His current R&D area is session control technologies. He is a member of IEICE.

Hiroshi Miyao
Senior Research Engineer, NTT Network Service Systems Laboratories.
He received his B.S. in physics from Ibaraki University in 1988 and his M.S. in physics from Hokkaido University in 1990. He joined NTT Transmission Systems Laboratories in 1990. His current R&D area is software engineering.

Development of High-capacity Protocol for M2M Services and Its Application to a Pallet Management System

Nobuaki Mochizuki, Masashi Shimizu, Kenji Suzuki, Kohei Mizuno, Akihiro Yamagishi, Mitsuru Harada, and Shuichi Yoshino

Abstract

NTT Network Innovation Laboratories has developed a high-capacity protocol for the 920-MHz band in order to enhance the effectiveness and accelerate the adoption of machine-to-machine (M2M) services. This protocol can detect the movement of hundreds of terminals and manage more than 10,000 terminals. The protocol is applied to physical distribution systems in which very large inventories need to be controlled. It offers efficient management of loading/unloading activities. The protocol was commercialized as a pallet management system in fiscal year 2015. In this article, we describe the protocol and explain how we verified its performance in tests conducted in an actual goods distribution center.

Keywords: M2M, high-capacity, physical distribution

1. Introduction

Machine to machine (M2M) communication systems have recently been attracting a lot of attention. The concept of the Internet of Things (IoT) involves connecting everything around us to the Internet. The information that comes from these multiple components will be collected and analyzed in a network. This information will then be utilized as big data and is expected to lead to the creation of new value services [1]. In addition, the infrastructure for M2M services is being improved. For example, the Association of Radio Industries and Businesses (ARIB) issued standard STD-T108 [2] for 920-MHz band use in 2012.

We have developed a high-capacity protocol that supports ARIB STD-T108. As a precursor of IoT/M2M, we applied the protocol to a pallet manage-

ment system and completed initial trials.

2. M2M service requirements

M2M services create different wireless loads from those of personal communication devices such as cellular phones. Most M2M service terminals cannot move independently and thus remain stationary for long periods. M2M systems must therefore be designed in consideration of these attributes. The typical communication environment desired for M2M services is depicted in **Fig. 1**. We assume the environment has the following characteristics.

(1) Wide area: Even if the wireless environment is inadequate, a terminal cannot be moved easily. For example, the access point and terminal may not have a direct line of sight. Thus, communication must be established even in
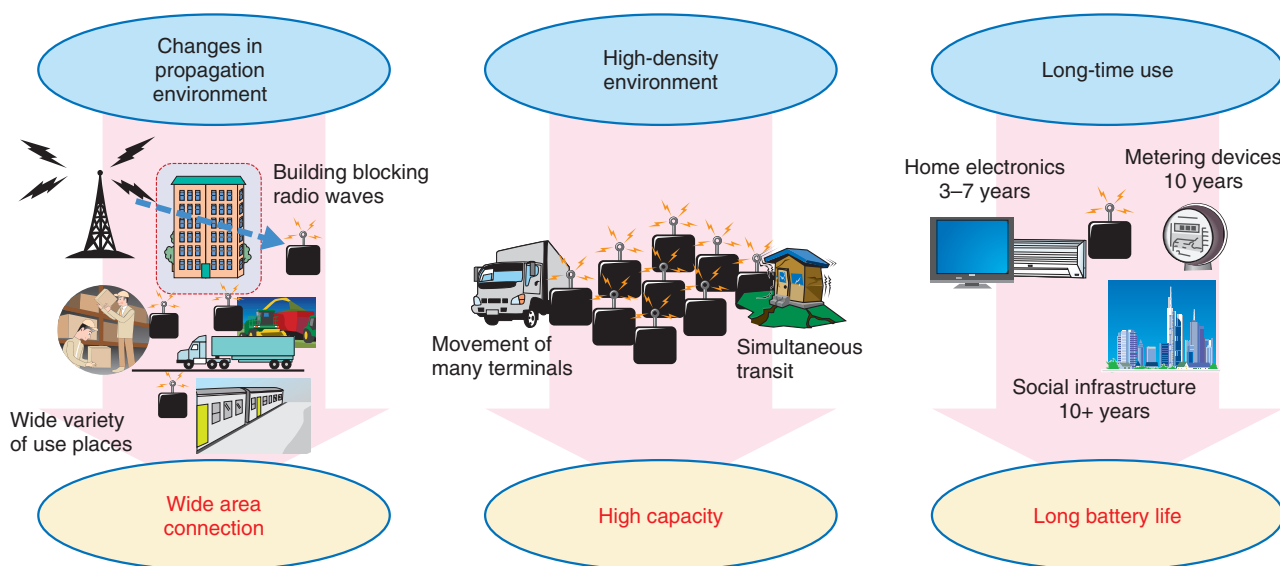
Fig. 1. Communication attributes desired for IoT/M2M service.

difficult circumstances.

(2) High capacity: It is assumed that many terminals are concentrated at one access point or that they transmit simultaneously after a disaster. Therefore, it is necessary to be able to communicate with many terminals.

(3) Long battery life: It is assumed that a terminal has no AC (alternating current) power supply and that some terminals will be installed in the social infrastructure, for example, on bridges and water pipes, and that their operation will be managed without changing the battery. Therefore, a long battery life must be achieved while maintaining the low power consumption.

Various wireless systems have been developed for M2M services in consideration of these requirements. The characteristics of wireless systems that can be applied to M2M services are shown in **Fig. 2**. The IMS (Industrial, Scientific and Medical) bands are in very heavy use, for example, for passive radio frequency identification (RFID) tags, Bluetooth*, and IEEE (Institute of Electrical and Electronics Engineers) 802.15.4/4g [3]/4e [4], the standards specifying the physical layer for low-rate wireless personal area networks.

Passive RFID tags have a short transmission distance, about 1 m, but they require no batteries. Thus, the usage scenario is to apply them as a substitute for barcodes in managing the logistics of goods. The Bluetooth transmission distance is about 10 m, and higher communication speeds are possible compared with other wireless systems used for M2M services, although the number of connected terminals is limited to 10. Bluetooth LE (low energy) specializes in low power consumption. Usage scenarios include applying it in wearable and healthcare related devices. IEEE 802.15.4/4g/4e specify a transmission speed of about 100 kbit/s, a maximum transmission distance of 1 km, and a capacity of about 1000 terminals. Unfortunately, the battery life is only about one year. Thus, the usage scenarios include metering devices, as stated in Wi-SUN (Wireless Smart Utility Network) alliance standards.

In contrast, the dotted line in Fig. 2 delineates the requirements of the pallet management system assumed here. This system must accommodate tens of thousands of pallets at a depot with only a few access points. Thus, the wireless distance is hundreds of meters, and the system capacity is more than 10,000 terminals. Moreover, the battery lifetime should be at least 10 years, roughly the lifetime of a pallet. In addition, rapid recognition of the loading and unloading of hundreds of pallets carried by trucks must be achieved, as indicated in **Fig. 3**. A new innovative wireless system is necessary to meet these requirements.

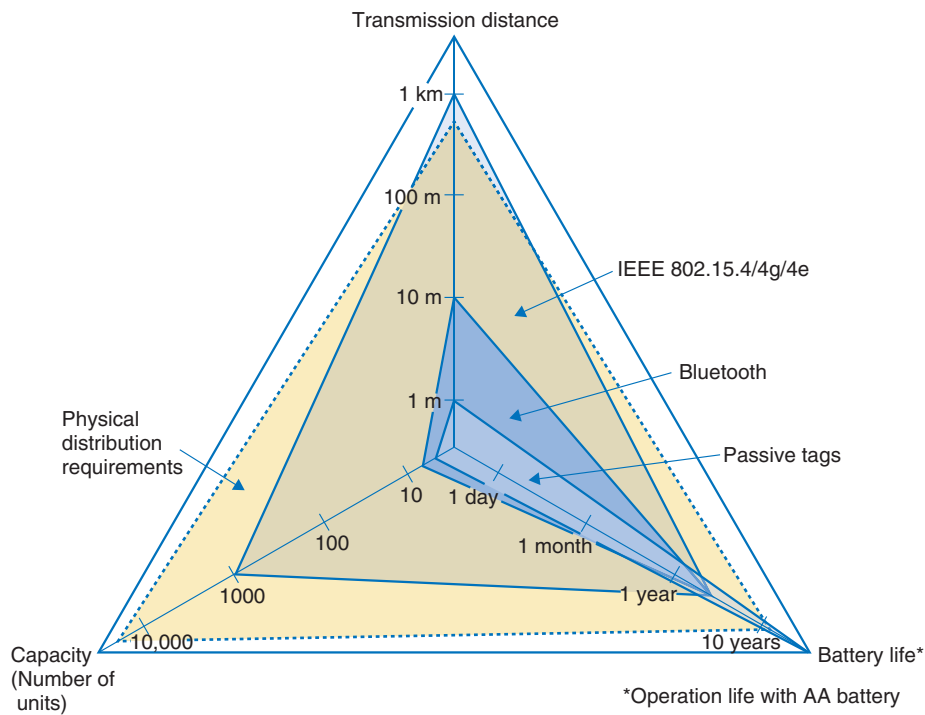* Bluetooth is a registered trademark of Bluetooth SIG Inc.

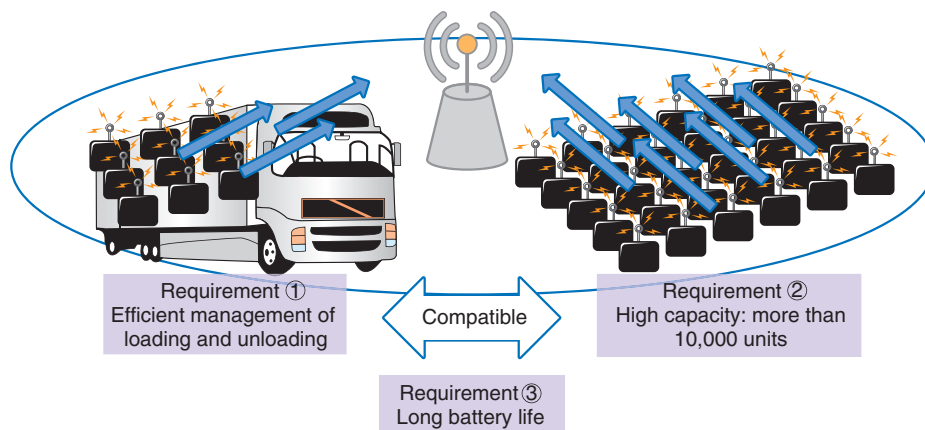Fig. 2.   Wireless system characteristics.

Fig. 3.   Requirements of pallet management system.

3.   High-capacity protocol technology

Our solution is a new high-capacity protocol that can connect tens of thousands of terminals and manage the movement of hundreds of terminals. Moreover, it enables a battery lifetime that can exceed 10 years. The protocol is outlined in **Fig. 4**.

The terminal has two states, a resting state and a movement state. The resting state indicates that the terminal is stationary. The movement state is triggered by the terminal's vibration sensor and indicates that the pallet is being transferred by a forklift or truck. The terminal transmits beacon packets consisting of a terminal ID, sequence number, and state information of the terminal. The transmission interval of the beacon packet is controlled by the state of the
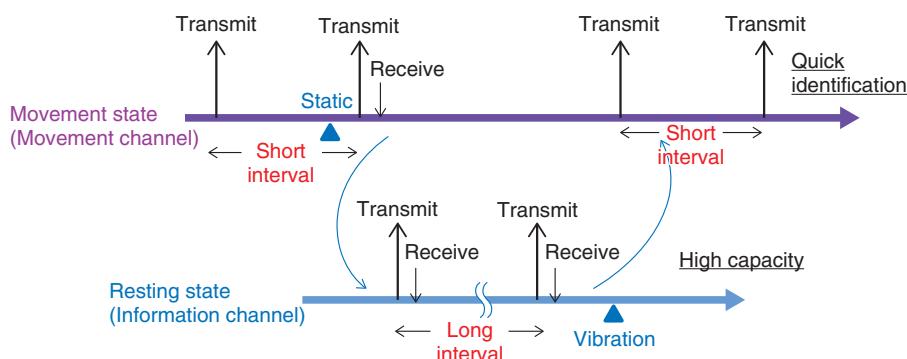
Fig. 4.   Outline of high-capacity protocol.

terminal. The beacon transmission interval is set very short in order to quickly identify the movement state. The resting state uses long time periods when transmitting beacon packets in order to accommodate many terminals.

In addition, the movement and resting states use different frequencies. For this reason, even if many resting state terminals are present, collision of beacon packets is minimized, and rapid recognition of terminals in the movement state becomes possible.

The protocol can be applied to a variety of M2M services with the use of different or additional sensors. For example, if the temperature exceeds or falls below a set threshold, rapid notification is sent. Furthermore, notification of sudden changes in water level, of a river, for example, can be sent quickly. Thus, our approach supports various M2M services by changing the sensor(s) and the beacon transmission interval.

4.   Evaluation of pallet management system in practical use

We developed equipment that hosts the aforementioned high-capacity protocol. The key specifications of the wireless equipment are listed in **Table 1**. The frequency band, transmission power, and other radio specifications are based on STD-T108, the ARIB standard specifications for the 920-MHz band. The transmission speed and modulation method are based on the IEEE 802.15.4g specifications. The access point can handle multiple frequencies through the use of multiple wireless modules and antennas. The terminal is equipped with components including a wireless module, built-in antenna, vibration sensor, and a battery and is roughly the same size as a pack of cigarettes, so it is easy to mount on a pallet.

Table 1.   Main specifications of wireless equipment.

Item	Specification
Frequency	920-MHz band
Transmission power	20 mW
Transmission speed	50 kbit/s
Modulation method	Frequency-shift keying
Equipment size	AP: 280 × 280 × 130.5 mm Terminal: 130 × 40 × 25 mm
Electric power supply	AP: AC power Terminal: battery

AP: access point

We applied the protocol in a pallet management system and conducted tests at a depot of UPR Corporation, a collaborative partner. The system configuration is depicted in **Fig. 5**. Several terminals were installed on pallets, and the other terminals were put in the depot to replicate the resting state. The access point was placed in the depot and was able to connect to the terminals in all locations and at the depot entrance. The access point transferred the beacon packets to a cloud server for pallet management.

Through these tests, we confirmed that the prototype system was able to accommodate more than 10,000 terminals. In addition, we confirmed that the reception probability reached 100% within 3 minutes when the terminal was in a movement state and within 100 m from the access point. Note that the number of terminals that one access point handles can be increased by enlarging the beacon packet interval in a resting state.

As a result of these successful tests, commercial operation of a pallet management system using this protocol was commenced at UPR Corporation in April 2015 [5, 6].
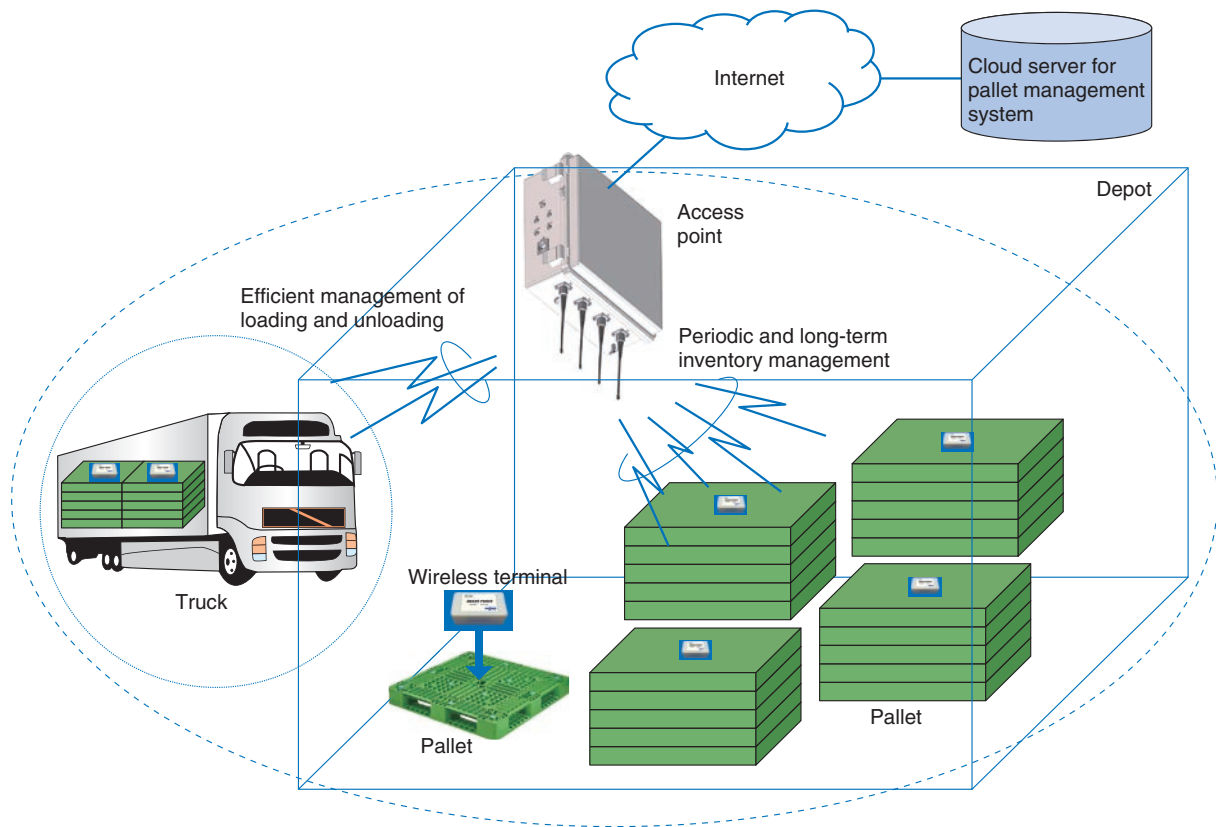
Fig. 5.   System configuration of pallet management system.

5.   Conclusion

In the future, we will work on advancing our research to achieve higher capacity and lower power consumption. Our goal is to achieve an M2M service that offers higher capacity and longer battery life.

References

[1]   World Economic Forum: Personal Data: The Emergence of a New Asset Class, Jan. 2011.

[2]   ARIB STD-T108: 920MHz-band Telemeter, Telecontrol and Data Transmission Radio Equipment, Feb. 2012.
http://www.arib.or.jp/english/html/overview/doc/1-STD-T108v1_0.pdf

[3]   IEEE Std 802.15.4g^{TM}-2012: Part 15.4-Low-Rate Wireless Personal Area Networks (LR-WPANs), Apr. 2012.

[4]   IEEE Std 802.15.4e^{TM}-2012: Part 15.4-Low-Rate Wireless Personal Area Networks (LR-WPANs), Apr. 2012.

[5]   Press release issued by UPR Corporation, Sept. 2014 (in Japanese).
http://www.upr-net.co.jp/pdf/news_20140908_02.pdf

[6]   Press release issued by NTT Electronics, Sept. 2014 (in Japanese).
http://www.ntt-electronics.com/new/information/2014/9/upr-rfid-smart-pallet.html

Nobuaki Mochizuki

Senior Research Engineer, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received his B.E. and M.E. from Tohoku University, Miyagi, in 1992 and 1994, respectively. He joined NTT in 1994 and engaged in R&D of a modulation/demodulation scheme for a broadband 5-GHz wireless access system and synchronization scheme for coded orthogonal frequency-division multiplexing. He is currently developing M2M wireless access systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Masashi Shimizu

Senior Research Engineer, Supervisor, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received his B.E. and M.E. in mechanical engineering from Keio University, Yokohama, in 1986 and 1988, respectively. He joined NTT in 1988 and began researching pointing control for deployable space antennas and surface error compensation through feed distribution control. His recent interests include active RFID and its applications. He is a member of IEICE.

Kenji Suzuki

Senior Research Engineer, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received his B.E. and M.E. in electrical and electronic engineering from Tokyo Institute of Technology in 1999 and 2001, respectively. In 2001, he joined NTT Telecommunications Energy Laboratories, where he worked on a wireless transceiver LSI architecture for low power dissipation. His interests include analog and RFIC (radio frequency integrated circuit) design for wireless communications. He is currently developing M2M systems. He is a member of IEICE and IEEE.

Kohei Mizuno

Senior Research Engineer, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received his B.E. and M.E. in electrical engineering from Keio University, Kanagawa, in 1997 and 1999, respectively. In 1999, he joined NTT Network Innovation Laboratories, where he worked on a wireless multihop network, active RFID tag, and home information and communication technology network. He is currently working on the design of 5G and beyond networks. He is a member of IEICE and IEEE.

Akihiro Yamagishi

Senior Research Engineer, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received a B.E. and M.E. in electrical engineering from Toyama University in 1986 and 1988, respectively, and a Ph.D. in electrical engineering from Tohoku University, Miyagi, in 2005. In 1988, he joined NTT LSI Laboratories, where he engaged in R&D of frequency synthesizer ICs. From 2000 to 2004, he studied 1-V-operation CMOS (complementary metal-oxide-semiconductor) RF circuits for wireless systems. He is currently studying a wireless terminal for M2M communication services. He is a member of the Institute of Image Information and Television Engineers of Japan.

Mitsuru Harada

Senior Research Engineer, Supervisor, Group Leader of Wireless Communication Circuits Research Group, Smart Devices Laboratory, NTT Microsystem Integration Laboratories.

He received a B.S. and M.S. in physics from Tsukuba University, Ibaraki, in 1985 and 1987, and a Ph.D. in electronic engineering from Tohoku University, Miyagi, in 2003. In 1990, he joined NTT Atsugi Electrical Communication Laboratories, where he engaged in research on thin-film semiconductor-on-insulator devices. Since 1997, he has been researching high-speed/low-power CMOS circuits for wireless terminals.

Shuichi Yoshino

Senior Research Engineer, Supervisor, Wireless Systems Innovation Laboratory, NTT Network Innovation Laboratories.

He received his B.E. and M.E. in mechanical engineering from Kanazawa University, Ishikawa, in 1990 and 1992, respectively. He joined NTT in 1992 and worked on satellite Internet system and wireless networking technology development. He is currently engaged in R&D on wireless technology for M2M wireless access services.

Global Standardization Activities

Reports of 4th ITU-T Review Committee and 1st FG-DFS (Digital Financial Services) Meetings

Hideo Imanaka

Abstract

Agreement was reached in the ITU-T (International Telecommunication Union, Telecommunication Standardization Sector) to establish a Review Committee (RevCom) to discuss a new standardization structure for the next study period starting in 2017. This agreement occurred at the WTSA (World Telecommunication Standardization Assembly) held in 2012. Additionally, it was agreed to establish a new Focus Group on Digital Financial Services (FG-DFS) at the TSAG (Telecommunication Standardization Advisory Group) meeting held in June 2014. This article reports on the key topics discussed at two meetings of these committees, the fourth meeting of RevCom, held in Tunis, Tunisia, in January 2015, and the first meeting of FG-DFS, held in Geneva, Switzerland, in December 2014.

Keywords: ITU-T, Review Committee, digital financial services

1. Introduction

At the World Telecommunication Standardization Assembly (WTSA) held in Dubai, UAE, in November 2012, it was agreed to establish a Review Committee (RevCom) in order to discuss restructuring the International Telecommunication Union, Telecommunication Standardization Sector (ITU-T) standardization structure in order to carry out more efficient standardization activities. The RevCom has held four meetings so far, and the discussion topics have included reconstruction ideas to implement starting in the year 2017. The fourth meeting hosted by Tunis Telecom was held January 19–21, 2015, and included a discussion on the proposal of a Standardization Strategy Team.

The establishment of a new Focus Group on Digital Financial Services (FG-DFS) was also agreed as a new work item in ITU-T. This was proposed by the Bill & Melinda Gates Foundation, which is a member of ITU-T, in the USA at the Telecommunication Standards Advisory Group (TSAG) meeting in June 2014. The first meeting was held in Geneva on December 5, 2014, in which the direction of study and the working

structure were discussed.

This article presents an overview of both the fourth RevCom and first FG-DFS meetings.

2. Fourth ITU-T RevCom meeting

2.1 Brief meeting summary

The fourth RevCom meeting was attended by 84 participants—including those participating remotely—from 15 countries. From Japan, five people including Mr. Yoichi Maeda, the chairman of RevCom and CEO of the Telecommunication Technology Committee (TTC), and other members from the MIC (Ministry of Internal Affairs and Communications), NICT (National Institute of Information and Communications Technology), NEC, and NTT (the author of this article) attended on site. Additionally, several people including the Study Group 3 (SG3) chairman (from KDDI) and the SG16 chairman (from Mitsubishi Electric) joined in remotely via an e-meeting system. The venue was located in the Gammarth area, a seaside resort about 20 km north of Tunis.

An ITU-T TSAG rapporteur meeting, which focused on a discussion of collaboration issues

between ITU-T and organizations inside/outside ITU, was held after the RevCom meeting.

2.2 Discussion on ITU-T reconstruction

The structuring and management team of RevCom is listed in **Fig. 1**. Mr. Maeda from TTC serves as chairman of RevCom, and there are six vice chairmen from six ITU regions. Issues on organizational restructuring in ITU-T were discussed in the rapporteur group under RevCom. The major results of the meeting are summarized as follows.

(1) Establishment of Standardization Strategy Team

Japan proposed establishing a new group called a Standardization Strategy Team to develop standardization strategies for future ITU-T activities such as standardization areas and timings. The strategies should be developed by taking into consideration the expectations from industries throughout the world, as well as future market and technology trends. This team would be led by the ITU-T director, and the members are expected to be the chairmen of SGs and FGs in ITU-T and key persons from industry.

The USA and France argued to clarify the relationship between TSAG and the proposed team, and Germany brought up a membership issue, which involved whether or not non-ITU-T members would be able to attend meetings of the proposed team. As a result of the discussion, the importance of developing standardization strategies was recognized, and the establishment of a strategic team was agreed, along with an inclusion policy and regulations. The discussion period and concrete framework were discussed further at the subsequent RevCom meeting held in June 2015.

(2) Establishment of vertical Study Group

Korea proposed to establish a new SG called a vertical SG, which would study new standardization issues that are not easily allocated to existing SGs. Some examples of new issues were: ITS (intelligent transport systems), IoT (Internet of Things), and 5G (fifth-generation) mobile networks. The main topic in the discussion was whether the vertical SG was actually necessary, since ITS is already discussed in Collaboration on ITS Communication Standards meetings with the ISO (International Organization for Standardization), and because the IoT-JCA (Joint Coordination Activity) in ITU-T has been studying IoT-related issues for several years. As a result, the discussion on the proposal of a vertical SG will continue at the next TSAG meeting.
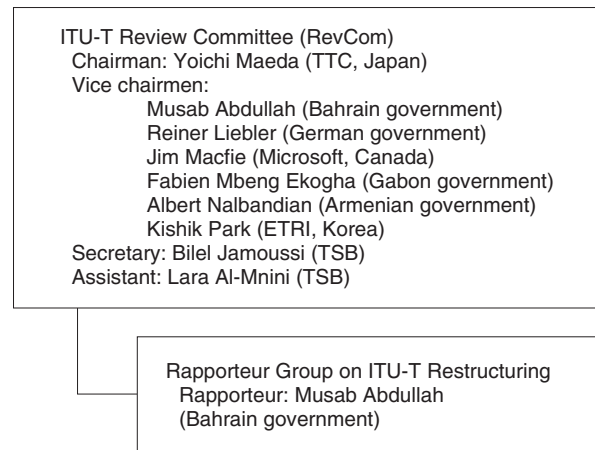
Fig. 1. Management team of ITU-T RevCom.

(3) SG reconstruction

There are currently 10 SGs in ITU-T, and a UK contribution proposed a discussion on merging some SGs, specifically, merging SG2 and SG3, and SG9 and SG16. Because the study items differ in each SG, it was suggested at the meeting that the measurement criteria of activities be examined. The proposal was basically agreed, and this topic will be studied further.

2.3 Discussion of working method

On the basis of the report of the CTO (Chief Technology Officers') meeting, it was agreed to make a work plan for working items such as IoT and SDN (software-defined networking) in RevCom. In addition, each ITU-T SG chairman reported their recent activities.

2.4 Future plans

At WTSA-16 to be held at the end of 2016, the restructuring of ITU-T SGs will be discussed. To make proposals for SG reconstruction from TSAG to WTSA-16, it is necessary to summarize the outcome of RevCom meetings by mid-2016. The next RevCom meeting may be the last chance to bring up new proposals, so it is important to contribute to the restructuring discussion to achieve efficient standardization activities for Japanese industry.

3. First ITU-T FG-DFS meeting

3.1 Overview

The first meeting of FG-DFS focused on a discussion about digital financial services and was held in Geneva on December 5, 2014; additionally, a workshop on digital financial services was held the previous day. These meetings were attended by around 100 persons from 30 countries, including several African countries.

3.2 Background of establishing FG-DFS

The digital financial services studied in FG-DFS include micro-finance in developing countries, mobile payments using cell phones, and financial inclusion, which improve financial services in developing countries. Some areas of digital financial services were also studied in TC 68 (Technical Committee 68) in ISO and JTC 1 (Joint Technical Committee 1) in ISO/IEC (International Electrotechnical Commission). This FG was established with the support of many countries, mainly developing countries, in order to clarify the telecommunication functions required to securely transfer money electronically all over the world, including developing countries, and to rapidly standardize telecommunication technologies related to financial services to promote its utilization in developing countries in ITU-T. Because the scope of FG-DFS may intersect with several SGs,

TSAG was selected to be the parent group, and the duration of the FG will be two years.

The management team includes Mr. Sacha Polverini from the Bill & Melinda Gates Foundation, who is serving as the chairman of the FG-DFS, and 12 vice chairmen, listed in **Table 1**, who were elected at the first meeting. In addition, four Working Groups (WGs) were agreed to be established, and 2–4 persons were appointed as the leader and sub-leader for each WG.

3.3 Overview of DFS workshop

The day before the first FG-DFS meeting, the workshop on digital financial services was held; it was chaired by the FG-DFS chairman [1]. At the workshop, 22 presentations were given on topics such as digital financial services, mobile payment services, and financial inclusion, primarily from African and European countries.

The Minister of Posts and Telecommunications of Somalia, H.E Dr. Mohamed Ibrahim, and the Director-General of the Communications Authority of Kenya, Mr. Francis Wangusi, delivered keynote speeches. In Somalia, the demand is strong for digital financial services using mobile devices such as cell phones, and there is great interest among Somalians in the concept of mobile money. However, Dr. Ibrahim pointed out that they needed a legal framework for providing such services. In Kenya, mobile money

Table 1. WG structure and management of FG-DFS.

WGs	Objectives	Leaders of WG (Vice chairmen of FG-DFS)
DFS Ecosystem	To identify and understand key elements of a digital financial services ecosystem that enables financial inclusion	- Carol Coye Benson, Glenbrook Partners (USA) - Yury Grin, Intervale (USA) - Bruno Antunes, UNCTAD (UN) - Zhao Ping, China Telecom (China)
Technology, Innovation & Competition	To identify evolving technologies and market dynamics related to digital financial services, and to assess the implications for financial inclusion	- Leon Perlman, Lethan Consulting (USA) - Madeleine Scherb, Health and Environment Program (Switzerland)
Interoperability	To assess models and policy frameworks to achieve interoperable systems, products and networks, and to develop a framework that promotes financial inclusion	- Thomas Lammer, The World Bank Group (USA) - Mark McCullagh, Learo Consulting (UK) - Yétondji Houyetognon, ARCEP (Benin)
Consumer Experience & Protection	To define minimum requirements for quality of service for reliable digital transactions and data protection safeguards in order to build consumer confidence	- Sumit Jamuar, KYCTrust Limited (UK) - Bedoui Adel, Tunisie Télécom (Tunisia) - Nicola O'Reilly, Consumers International (UK)

* The above list is as of December 5, 2014.

services were initiated in 2007, and 64% of the population now use the M-Pesa service, a kind of mobile money service that handles US$23 billion annually. Kenya is the leading country in terms of total amount of transactions in mobile money and micro-finance in the African region.

3.4 Discussion topics of first FG-DFS meeting
(1) Terms of Reference

Terms of Reference, ToR, was discussed, and the major opinions were as follows:
- It is necessary to clarify the terminology and definitions for mobile money, mobile finance, digital money, electronic money, and digital finance.
- The digital financial service will have a new financial structure and a new distribution model.
- Although the administration area of the central bank and telecommunication regulation office differs in each government, it is necessary to establish cooperation between the central bank and telecommunication regulation office. In addition to the services provided by banks, related financial services provided by insurance companies and non-bank organizations are also involved in digital financial services.
- One of the objectives was to include use cases for women; thus, although the concept is mainly focused on the poor, the existence of a gender gap was acknowledged.
- Security and availability are required as the standards for basic digital financial services. In addition, it is also necessary to consider ensuring accessibility, minimizing the initial cost, and reducing transaction costs.

(2) WG structure

The FG chairman proposed establishing WGs, and it was agreed to establish four WGs and to nominate 12 leaders as a result of the discussion. The WG structure is indicated in Table 1. Several opinions were raised in the discussion; for example, one member said it would be hard to participate if several WGs were running in parallel. However, because the discussion themes of each WG are different, and the interests of participants may also differ, it was confirmed that holding several WGs in parallel would not be a problem. Another opinion was that terminology should be unified at an early stage, and it was agreed to define the terminology at an early stage in each WG.

(3) Discussion in each WG

FG participants were divided into four WGs, and a one-hour brainstorming session was held in each WG to define the main issues that should be discussed. The discussion results were presented orally, and no output documents were produced.

3.5 Subsequent work of FG-DFS
The second FG meeting was held in Washington, D.C., USA, in April of 2015, the same timing of the spring conference of the World Bank.

4. Conclusion

4.1 ITU-T RevCom
The task force of the TTC Global Collaboration Advisory Group (GCAG), led by the author of this article, handles Japan's policies and contributions for RevCom and TSAG. To achieve an efficient ITU-T standardization structure for Japanese industry, it is necessary to proactively contribute to RevCom by continuously aggregating domestic opinions in the task force and by collaborating with foreign countries.

4.2 ITU-T FG-DFS
There are now two directions for international standards of digital financial services; the first is achieving universal and easy deployment in developing countries. Standardization of Japanese mobile wallet services would make traveling more convenient for foreigners who come to Japan and for Japanese travelers who visit foreign countries. The second direction is that international standards require a safer and more economical way to deploy mobile payment systems such as M-Pesa in Kenya. Utilizing FG-DFS may not only be effective for promoting Japanese solutions to the world, but also for clarifying the role of Japanese companies in achieving financial inclusion in developing countries.

4.3 Consideration of organization for new study items
This article briefly reported on the current status of RevCom and FG-DFS. In the near future, international standardization activities for required functions will be more important for non-telecom industries that utilize the telecommunication infrastructure (vertical industries), for purposes such as digital financial services. ITU standards have played an important role in the rapid growth of information and communication technology (ICT). Currently, ICT has been entering a second stage; that is, ICT supports vertical industries, which are indicated as "OO×ICT."

Contributing to RevCom is expected to lead to achieving an efficient standardization structure to enable progress to be made in anything supported by ICT.

Reference

[1] ITU Workshop on Digital Financial Services and Financial Inclusion, http://www.itu.int/en/ITU-T/Workshops-and-Seminars/ifds/Pages/Programme.aspx

Hideo Imanaka
Senior Manager, NTT R&D Planning Department*.
He received his B.E., M.E., and Ph.D. in electrical engineering from Mie University in 1985, 1987, and 2001, respectively. After joining NTT Telecommunication Network Laboratories in 1987, he conducted research on a fiber optic access network architecture and network operation process reengineering methods. From 1996 to 2003, he worked on enterprise resource planning (ERP) systems integration as a consultant in the Solutions Business Division of NTT Communications. Since 2004, he has been involved in NGN (Next Generation Network) standardization work at ITU-T. He was the Rapporteur of Question 1 of SG13 from 2007 to 2010. He has also played an active role in IP (Internet protocol) TV standardization work at ITU-T. He is currently in charge of standardization strategies in the NTT Group. He received the ITU-AJ Award from the ITU Association of Japan in 2009. He is a member of the Institute of Electronics, Information and Communication Engineers and the Society of Instrument and Control Engineers.

* He moved to NTT Advanced Technology Corporation in April 2015.

Practical Field Information about Telecommunication Technologies

Case Studies of Faults and Countermeasures in Access Telephone Office

Abstract

This article reports on recent faults in optical access communication facilities that occurred in a telephone office. We review (1) a communication facilities problem caused by removal of a LAN (local area network) cable, (2) an accidental cable breakage that occurred while fixing cables on a distribution frame, and (3) a service interruption that occurred due to deterioration of an optical branch module used for testing optical lines. Cases (1) and (2) involved breakdowns that occurred during construction, whereas case (3) was an unforeseen breakdown related to long-term deterioration. We also present the appropriate countermeasures to these faults that were applied in order to maintain continuous operation of the communication facilities. This article is the thirtieth in a bimonthly series on practical field information on telecommunication technologies. This month's contribution is from the Access Engineering Group, Technical Assistance and Support Center, Maintenance and Service Operations Department, Network Business Headquarters, NTT EAST.

Keywords: access communication facilities, fault cases, fault countermeasures

1. Introduction

Fiber-to-the-home (FTTH) systems support today's information society and are capable of transmitting vast amounts of data. They are primarily designed with PON (passive optical network) topology [1]. When services supplied through the FTTH systems cease because of communication failures or maintenance work related to network operation, a large number of users are disadvantaged.

The Technical Assistance and Support Center at NTT EAST grapples with faults that occur in access communication facilities that support the FTTH systems, and for which solutions are difficult, and develops countermeasures to the faults to improve the quality of communication facilities. Previous reports centered on actual fault cases of optical communication facilities consisting of outdoor offices and customer buildings [2–4]. This article presents recent faults in access communication facilities that occurred at a telephone office.

2. Case studies

2.1 Case 1: Fault caused by cable removal

This case concerns damage to functioning cables that occurred while removing disused cables in a telephone office.

(1) Fault description

Multiple damage traces were discovered on the outer sheaths of optical cables and local area network (LAN) cables installed on an aerial rack. Although communication services had not yet been interrupted at the time of the discovery, the damage was extensive enough that the interior of the cables was exposed.

(2) Verification results (Replication experiment)

The fault location and the state of cable damage are shown in **Fig. 1**. The burred lines indicate the damage, and as shown, the burrs are all oriented in the same direction. In addition, each mark corresponds to an arc-shaped groove about equal in width to the diameter of the existing LAN cables. In light of the above, we inferred that the cause of the fault was an accident involving friction generated when removing
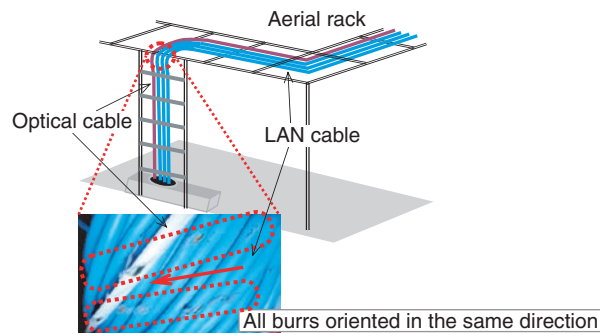
Fig. 1.   Location of fault and cable damage.

(b) Replication experiment sample
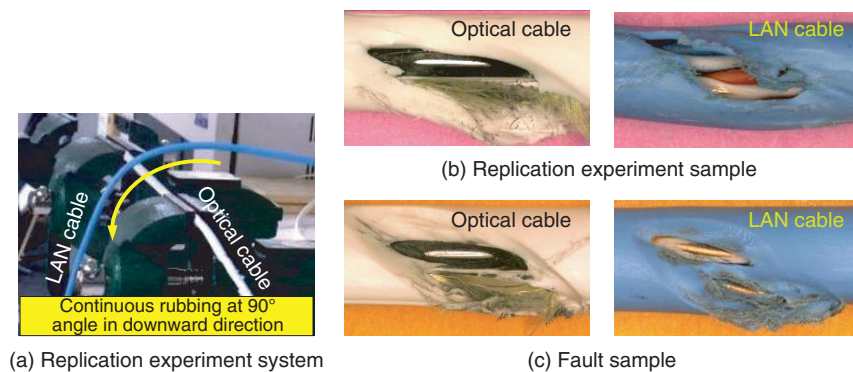
(a) Replication experiment system

(c) Fault sample

Fig. 2.   Experiment to replicate cable damage.

LAN cables from the aerial rack. We therefore conducted a replication experiment in which we made two cables rub up against each other. Specifically, we prepared an optical cable and a LAN cable the same as those in use, and as shown in **Fig. 2(a)**, we pulled the LAN cable over the optical cable that was fixed in place in an attempt to simulate actual cable removal work. We found that a crack began to form on the outer sheath of the fixed cable after pulling the LAN cable about 2 m and that the tension member of the cable came to be exposed after pulling the cable 10–20 m (**Fig. 2(b)**). We confirmed that this damage was similar to that of the original fault (**Fig. 2(c)**). Similar results were obtained for a LAN cable (Fig. 2(b)).

(3)   Countermeasure

In view of the experimental results presented above, we surmised that this fault was caused by friction generated on an optical cable or LAN cable while removing another LAN cable. We therefore considered that the use of a lubricant or anti-friction sheet at the point where the cables come in contact and gener-

ate friction would be an effective countermeasure to the friction that can accompany cable laying/removal work.

2.2   Case 2: Fault at CTF-cable binding location

This case concerns a cut in an optical cable fixed in a cable termination frame (CTF) and the resulting break in communications.

(1)   Fault description

A deep cut was found in a double-core optical cable laid between a CTF and on-premise equipment near the point where the cable was bound with a binding strap to other cables at the lower end of the CTF (**Fig. 3**). No reasons as to why the cable should be damaged near the lower end of the CTF could be found, and there was no place where the fault location could have easily come into contact with the floor.

(2)   Verification results (Replication experiment)

On dismantling the damaged cable and inspecting its interior, we found that constituent optical cords had been cut and that the tension-member covering
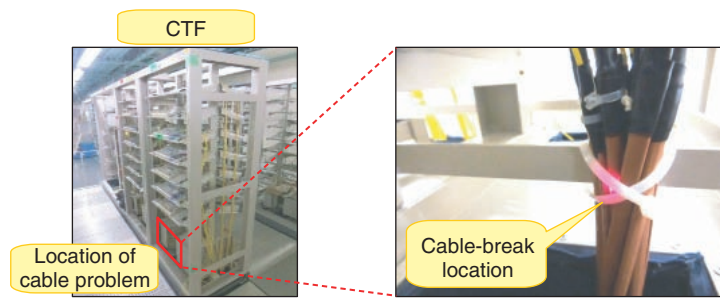
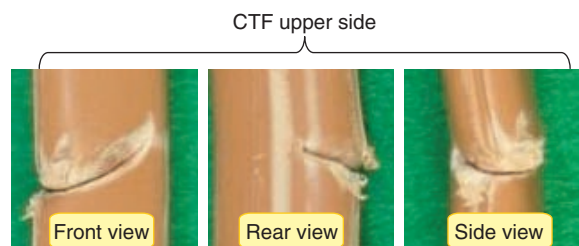Fig. 3.   Location of cable problem and image of cable damage.

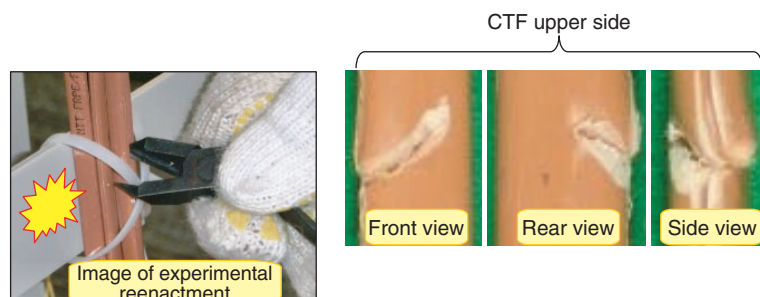Fig. 4.   Damage to outer sheath of optical cable.

Fig. 5.   Verification of reenactment results.

was damaged (**Fig. 4**). Consequently, we suspected that the nippers and cable clippers frequently used in fixing and removing on-premise cables might be the problem, so we examined what form of damage such tools could cause. We found that the fracture caused by nippers resembled the actual fault more closely than that caused by cable clippers. The replication setup using nippers and the damage formed in this experimental reenactment are shown in **Fig. 5**. Through this experiment, we found that orienting the blade tips of the nippers toward the cables when cutting the X-shaped binding strap could damage a cable

as deeply as its internal optical cords. Then, on examining the fiber cross sections at the damaged location using an electronic microscope, we found that they resembled those of the fault sample as shown in **Fig. 6**.

(3)   Countermeasure

On the basis of our observations of the damaged cable and the results of the replication experiment, we surmised that the tips of the nippers had penetrated the cable's outer sheath when cutting the binding strap and had reached the optical-cord jacket, and that the optical fiber was sliced by the lateral pressure
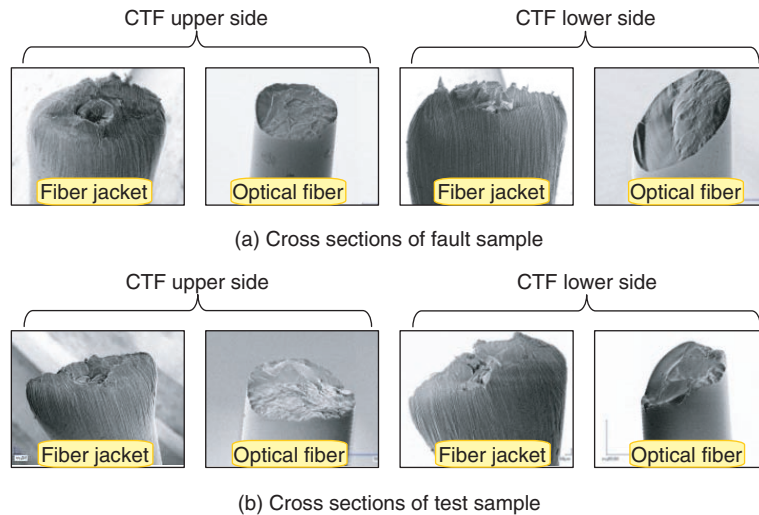
CTF upper side   CTF lower side

Fiber jacket   Optical fiber   Fiber jacket   Optical fiber

(a) Cross sections of fault sample

CTF upper side   CTF lower side

Fiber jacket   Optical fiber   Fiber jacket   Optical fiber

(b) Cross sections of test sample

Fig. 6. Comparison of fiber jacket and cross sections of optical fiber between fault sample and test sample.

OLT: optical line terminal
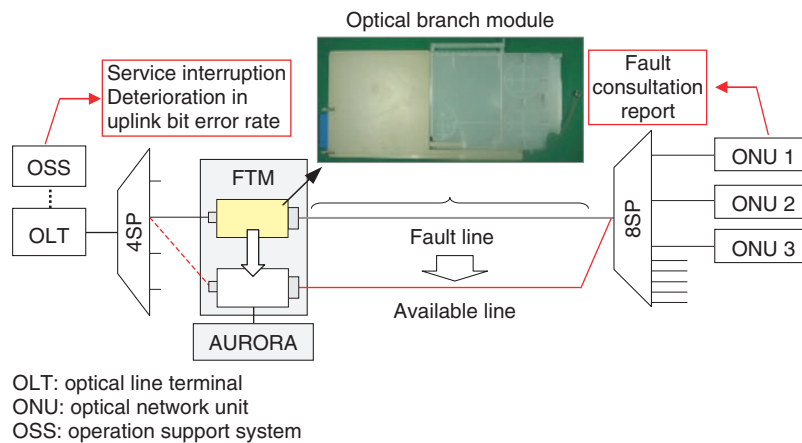ONU: optical network unit
OSS: operation support system

Fig. 7. Facility configuration and optical core switching.

applied at that time. We considered that the main cause of this accident was a working environment with tight spaces and blind spots. It is therefore important that maintenance personnel make a conscious effort to orient blade tips away from cables when using nippers. In addition, cutting the binding strap from the backside of a fixing location could also be an effective measure to prevent faults of this type.

2.3 Case 3: Fault in optical branch module

This case concerns a fault in an optical branch module used for injecting test signals along an optical path.

(1) Fault description

A report of a service interruption was received from a customer using the FLET'S Hikari optical broadband service. On troubleshooting the fault, an increase in loss along the outside optical path that included an optical branch module was suspected, but the fault location could not be isolated. Although the location of the original fault had still not been clarified, it was eventually decided to switch the fault line reported as being faulty to a circuit on another optical branch module. This action effectively restored services to the customer (**Fig. 7**).
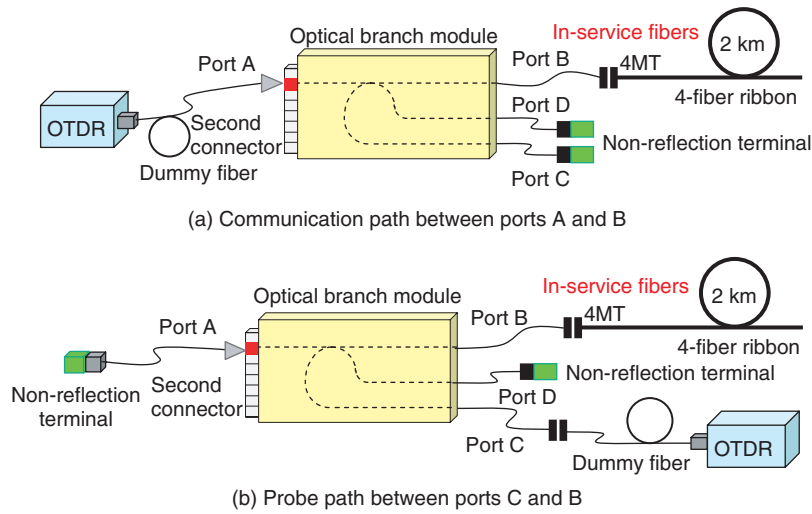
(a) Communication path between ports A and B

(b) Probe path between ports C and B

Fig. 8.   Transmission loss measurement system.

(a) Actual faulty module

—○— : Between ports A-B ⎤ Specified
—■— : Between ports C-B ⎦ value exists

(b) Normal module for comparison

▮ : Specified values between ports A-B
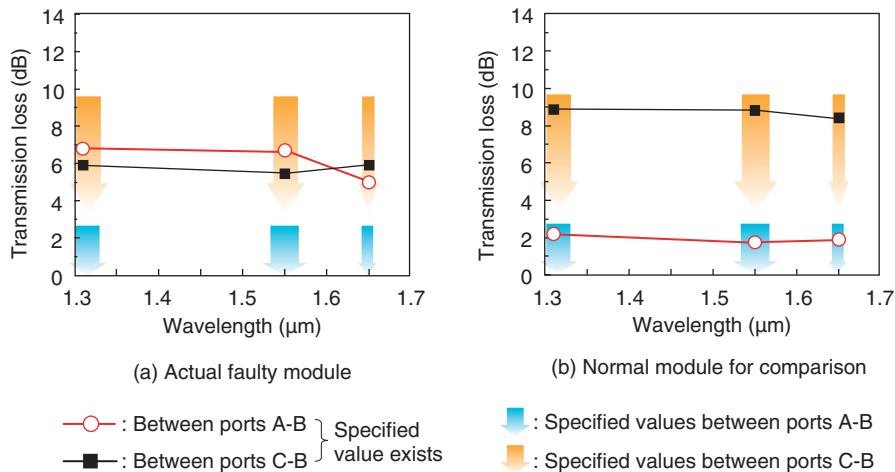▮ : Specified values between ports C-B

Fig. 9.   Measurement results.

(2)   Verification results (On-site inspection)

Because the optical branch module targeted for inspection was accommodating other core fibers in use, we proceeded to inspect their current state in the fiber termination module (FTM) rack. First, on visually inspecting the module from the outside, we found no obvious problems in the rack or connector section. Next, we measured transmission loss between ports A-B related to the second connector cited in the fault report and also between ports C-B using an optical time-domain reflectometer (OTDR), as shown in **Fig. 8**. Transmission loss at the second connector of the faulty module and that of a normal one as reference are shown in **Fig. 9**. The transmission loss of the faulty module satisfied the specified values at standard wavelengths (1.31, 1.55, and 1.65 μm) between ports C-B but greatly exceeded the specified values between ports A-B. Furthermore, in comparison with the normal product examined in Fig. 9(b), it was found that the difference in transmission loss between port-interval A-B and port-interval C-B in the faulty module was smaller.

(3)   Inference of fault location

The on-premise optical branch module uses a wavelength-independent optical coupler that has an optical transmission loss difference of about 6 dB

between port-intervals A-B and C-B, as shown in Fig. 9(b). However, this difference is only about 1 dB in the faulty module, as shown in Fig. 9(a), so we assumed that a fault exists for some reason at the second connector of the module. We were not able to isolate the fault in this optical branch module through fault maintenance work at the site during this investigation. We are therefore recommending that an OTDR test connecting a dummy fiber to port A of the optical branch module be conducted the next time a similar event occurs. In this way, we should be able to clarify the transmission loss between ports A-B from the difference in levels of backscattering light from the dummy fiber and outside path. This difference could then be used as a criterion for assessing problems in the module.

3. Conclusion

In this article, we introduced case studies of faults that occurred in a telephone office and explained their countermeasures. Optical communications services come in many forms, and the facilities used to support them are likewise many and varied. We can therefore expect maintenance operations for telephone offices to be just as diverse. The Technical Assistance and Support Center is committed to responding to all requests for technical consultation and assistance and to disseminating practical telecommunication technologies while leveraging the knowledge and experience gained to date from the maintenance of diverse facilities.

References

[1] ITU-T Recommendation G.983.1, "Broadband optical access systems based on Passive Optical Networks (PON)," 1998.
[2] NTT EAST, "Case Studies of Faults and Countermeasures in a Passive Optical Network System," NTT Technical Review, Vol. 10, No. 7, 2012.
     https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201207fa2.html
[3] NTT EAST, "Fault Cases and Countermeasures for Field Assembly Connectors in Optical Access Facilities," NTT Technical Review, Vol. 9, No. 7, 2011.
     https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201107fa2.html
[4] NTT EAST, "Failure Cases in Optical Access Facilities (Fiber Breakage in Optical Cabinets) and Countermeasures," NTT Technical Review, Vol. 11, No. 4, 2013.
     https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201304pf1.html

External Awards

The 37th Awaya Prize Young Researcher Award
Winner: Daichi Kitamura, Sokendai (The Graduate University for Advanced Studies); Nobutaka Ono, National Institute of Informatics; Hiroshi Sawada, NTT Service Evolution Laboratories; Hirokazu Kameoka, NTT Communication Science Laboratories; and Hiroshi Saruwatari, the University of Tokyo
Date: March 17, 2015
Organization: The Acoustical Society of Japan (ASJ)

For "Efficient Multichannel Nonnegative Matrix Factorization with Rank-1 Spatial Model."
Published as: D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Efficient Multichannel Nonnegative Matrix Factorization with Rank-1 Spatial Model," Proc. of ASJ Autumn Meeting, pp. 579–582, Sept. 2014.

ITU-AJ Accomplishment Award
Winner: Makoto Murakami, NTT Network Service Systems Laboratories
Date: May 15, 2015
Organization: The ITU Association of Japan (ITU-AJ)

For his contribution to implementation of regulatory revisions and the creation of a handbook and tutorial relating to a group of recommendations covering optical transport network technology in the Telecommunication Standardization Sector of International Telecommunication Union (ITU-T) Study Group 15 (SG15). As head of the Japanese delegation, he is committed to resolving the international controversy over MPLS-TP (multiprotocol label switching-transport profile) standardization by leading the way in a WTSA (World Telecommunication Standardization Assembly) joint proposal with other member states in APT (Asia-Pacific Telecommunity). He also contributed as a liaison officer in settling a dispute over specifications between IEC (International Electrotechnical Commission) and SG15. In addition, he participated in the ITU Focus Group on disaster relief systems as editor and contributed to finalizing deliverables.

Interop 2015 Best of Show Award (SDI Special Prize)
Winner: O3 project*
* NTT researchers involved are Hirokazu Takahashi, Yoshihiro Nakajima, Hitoshi Masutani, Takeshi Kinoshita, Tomoya Hibi, Sakiko Kawai, Atsushi Yamamoto, and Atsushi Taniguchi, NTT Network Innovation Laboratories.
Date: June 10, 2015
Organization: Interop Tokyo Steering Committee

For "SDN/OpenFlow Software Switch Lagopus."
Published as: O3 project, "SDN/OpenFlow Software Switch Lagopus," Proc. of Interop Tokyo 2015, Chiba, Japan, Jun. 2015.

The Meritorious Award on Radio presented by the Chairman of the Board of ARIB
Winner: Masashi Shimizu, NTT Network Innovation Laboratories; Yoshiya Sakata, UPR Corporation; Takumi Watanabe, NTT Electronics Corporation; and Akihiko Taniya, Nagano Japan Radio Co., Ltd.
Date: June 16, 2015
Organization: Association of Radio Industries and Businesses (ARIB)

For development of a logistics pallet management system using active radio-frequency identification techniques.

TTC Distinguished Service Award
Winner: Takefumi Yamazaki, NTT Service Evolution Laboratories
Date: June 22, 2015
Organization: Telecommunication Technology Committee (TTC)

For his contribution to the promotion of standardization of home network systems.

TTC Distinguished Service Award
Winner: Makoto Murakami, NTT Network Service Systems Laboratories
Date: June 22, 2015
Organization: TTC

For his contribution to the promotion of standardization of inter-network transmission.

Papers Published in Technical Journals and Conference Proceedings

Large-scale Collection and Analysis of Personal Question-answer Pairs for Conversational Agents

H. Sugiyama, T. Meguro, R. Higashinaka, and Y. Minami

Proc. of IVA 2014 (the 14th International Conference on Intelligent Virtual Agents), pp. 420–433, Boston, MA, USA, August 2014.

In conversation, a speaker sometimes asks questions that relate to another speaker's detailed personality, such as his/her favorite foods and sports. This behavior also appears in conversations with conversational agents; therefore, agents should be developed that can respond to such questions. In previous agents, this was achieved by creating question-answer pairs defined manually. However, when a small number of persons create the pairs, we cannot know what types of questions are frequently asked. Therefore, such essential question-answer pairs for conversational agents are possibly overlooked. This study analyzes a large number of question-answer pairs for six personae created by many question-generators, with one answer-generator for each persona. A comparison with questions appearing in conversations between humans shows that 50.2% of the questions were contained in our question-answer pairs, and the coverage rate was almost saturated with the 20 recruited question-generators.

Material Transmission Loss Modeling for Indoor Propagation Modeling

M. Inomata, T. Ogawa, and S. Yoshino

Proc. of IEEE PIMRC 2014 (the 25th International Symposium on Personal Indoor and Mobile Radio Communications), pp. 800–804, Washington, DC, USA, September 2014.

We are studying a method to minimize mutual radio wave interference by coordinating wireless home network access points (APs). To control the radio wave range so as to reduce interference with neighboring terminals or APs, precise techniques to estimate range are important. A commonly used method to estimate radio wave reaching ranges is site surveys, where the user or operator measures propagation characteristics. However, the problem with this method is that a large burden is placed on the user because it is necessary to densely collect the radio data over the entire area to precisely estimate the radio wave reaching ranges, especially in indoor radio propagation environments where there is considerable propagation variability. We therefore propose a method that uses sparse data to estimate material parameters for precisely modeling indoor radio propagation. Experimental evaluations demonstrated it decreased the number of data items needing to be collected.

Open-domain Utterance Generation Using Phrase Pairs Based on Dependency Relations

H. Sugiyama, T. Meguro, R. Higashinaka, and Y. Minami

Proc. of SLT 2014 (2014 IEEE Spoken Language Technology Workshop), pp. 60–65, South Lake Tahoe, NV, USA, December 2014.

The development of open-domain conversational systems remains difficult since user utterances vary too widely for such systems to respond appropriately. To address this issue, previous research has retrieved sentences from the web as system utterances by applying shallow sentence matching with user utterances. However, since the retrieved sentences include the inherent contexts of the document in which the sentences originally appeared, the retrieved sentences may possibly contain information that is irrelevant to user utterances. We propose combining two strongly related *semantic units* (phrase pairs with dependency relations) to create a system utterance. Here, the first semantic unit is the one found in the user utterance, and the second semantic unit is the one that has a dependency relation with the first one in a large text corpus. This way, we can guarantee that the generated utterance is related to the input user utterance. Our experiments, which examine the appropriateness of response sentences, show that our proposed method significantly outperforms other retrieval and rule-based approaches.

Comparison of Photo Degradation Behavior of LDPE Using Accelerated Weathering Instruments

T. Miwa, Y. Takeshita, Y. Akage, M. Watanabe, M. Takaya, and T. Sawada

Zairyo-to-Kankyo, Vol. 64, pp. 139–144, April 2015.

Samples of low-density polyethylene (LDPE) were photodegraded using accelerated weathering instruments and outdoor exposure. The physical properties and chemical structures of the photodegraded samples were studied through a tensile test, gel chromatography, and infrared spectroscopy.

The molecular weight distribution of a photodegraded sample by using a fluorescent UV lamp at a high black panel temperature (80°C) was more similar to that of an outdoor-degraded sample than that of other artificially photodegraded samples by using a Xenon lamp at the standard black panel temperature (63°C). It is estimated that accelerated weathering tests at a high sample temperature could accelerate cross-linking more than chain scission, consequently recreating molecule-enlargement similar to the outdoor-degraded sample.

Path Loss Model for the 2 to 37 GHz Band in Street Microcell Environments

M. Inomata, W. Yamada, M. Sasaki, M. Mizoguchi, K. Kitao, and T. Imai

IEICE Communications Express, Vol. 4, No. 5, pp. 149–154, May 2015.

Path loss characteristics are analyzed on the basis of measurement results obtained using the 2 to 37 GHz band in street microcell environments. By taking dependencies on frequency and distance from transmitter to intersection into account, the proposed model can decrease the root mean square error of prediction results to within about 5 dB in the 2 to 37 GHz band.

A Tracing Technique for Understanding the Behavior of Large-scale Distributed Systems

Y. Bando

Proc. of LinuxCon Japan, Tokyo, Japan, June 2015.

Debugging or troubleshooting large-scale distributed systems is difficult due to its complexity; a single request may trigger the

execution of hundreds of components running in parallel on many different machines. To help developers or operators gain deeper knowledge about the behavior of their distributed systems, I proposed a tracing method that can be applied to their applications simply by slightly modifying an existing RPC library. This tracing helps them know the flow of processing and find performance bottlenecks. I implemented it in Eventlet, an RPC library widely used in OpenStack projects, and also started discussing it with the Eventlet community in order to have this feature included in Eventlet.

In this talk, I demonstrated the tracking of swift, OpenStack object storage, as an example. In addition, the visualization of trace data and the overhead will be reported.

Variable-length Lossy Source Code Using a Constrained-random-number Generator

J. Muramatsu

IEEE Transactions on Information Theory, Vol. 61, No. 6, pp. 3574–3592, June 2015.

A variable-length lossy source code is introduced with a rate-distortion pair close to the rate-distortion function. Random numbers that satisfy a condition specified by a function and its value are used to construct a stochastic encoder. The proof of the theorem is based on the balanced-coloring property of an ensemble of functions. Since an ensemble of systematic sparse matrices has this property, we can construct a tractable code for a memoryless source. Some algorithms for implementing the code are introduced and compared by simulation.

Smooth Motion Parallax Autostereoscopic 3D Display Using Linear Blending of Viewing Zones

M. Date, T. Kawakami, M. Sasai, and H. Takada

Proc. of the SID (Society for Information Display) Display Week 2015, pp. 983–986, San Jose, CA, USA, June 2015.

A new autostereoscopic three-dimensional (3D) display is proposed. Using only a small number of projectors, it produces smooth and exact motion parallax by applying the visual effects of dual edge perception in a depth-fused 3D (DFD) display. It provides a breakthrough in overcoming the trade-off between 3D image reality and the number of video sources.