



Frequency Limitations of Resonant-Tunnelling Diodes in Sub-THz and THz Oscillators and Detectors

Michael Feiginov¹ 

Received: 29 October 2018 / Accepted: 26 February 2019 / Published online: 13 March 2019
© The Author(s) 2019

Abstract

The review outlines the basic principles of operation of resonant-tunnelling diodes (RTDs) and RTD oscillators followed by an overview of their development in the last decades. Further, we discuss different types of RTDs and RTD oscillators, the limitations of RTDs due to parasitics, inherent limitations of RTDs and operation of RTDs as detectors. We also give an overview of the present status of sub-THz and THz RTD oscillators and give several examples of their applications.

Keywords Resonant tunnelling · Resonant-tunnelling diodes · Terahertz sources · Terahertz detectors

1 Introduction

The concept of resonant tunnelling through the quantum states between double barriers is known since very long time, for example, it has been described in detail in a classical textbook already in 1951 [1]. The principles of operation of the devices based on the phenomenon have been discussed very long time ago as well: attainability of negative differential conductance (NDC) and even resonant-tunnelling transistors and triple-barrier resonant-tunnelling structures have been addressed as early as in 1964 [2]. However, a theoretical analysis of practically realisable semiconductor-heterostructure resonant-tunnelling diodes (RTDs) with few barriers was first presented in 1973 [3], NDC in such structures was demonstrated experimentally shortly afterwards in 1974 [4]. Semiconductor-heterostructure RTDs stay in the main focus of the research in this field until now.

✉ Michael Feiginov
michael.feiginov@tuwien.ac.at

¹ THz-Electronics Group, TU Wien (Vienna University of Technology), Gusshausstrasse 25/354, 1040 Vienna, Austria

An RTD is the most simple and basic structure to study the resonant-tunnelling phenomena. As such a toy system, it has stimulated a wealth of theoretical and experimental publications in the semiconductor physics in the last half a century. Apart from that, the prospect to use RTDs in practise has also contributed to increased interest in the devices. Specifically, tunnelling can be an extremely fast process [5], although definition of tunnelling time is a controversial subject [6]. Additionally, RTDs can exhibit NDC, which makes them active devices, i.e., they can provide amplification or gain. That has raised promise to use RTDs in high-frequency (HF) electronics, particularly, to construct mm-wave and sub-THz RTD oscillators. Indeed, the development of RTD oscillators was progressing fast in the subsequent years: 712 GHz RTD oscillators have been reported in 1991 [7]. However, in the following years, nobody could achieve higher operating frequencies or even reproduce the achieved results. Gradually, the expectation that RTDs could be used in practise has disappeared in the research community and RTDs were almost abandoned as a research subject for the following almost 20 years.

A breakthrough has been achieved in 2010 and 2011, when two groups have reported RTD oscillators at ≈ 1 THz [8] and ≈ 1.1 THz [9], respectively. In the subsequent years, the operating frequencies of RTD oscillators were continuously increasing and presently they have reached almost 2 THz [10]. It has to be mentioned, that, contrary to THz quantum-cascade lasers (QCLs) [11], RTD oscillators are room-temperature devices and that is their crucial advantage compared THz QCLs. The output power of RTD oscillators is getting close to mW level at the upper sub-THz frequencies [12]. Different types of fundamental THz and sub-THz oscillators have been demonstrated in the meanwhile [7–10, 13], some of them are extremely small [9]. High-speed wireless data transmitters based on RTDs have been also demonstrated [14]. This development in the recent years puts RTD oscillators in the position of a viable enabling technology for real-world THz applications.

The review will describe the development in sub-THz and THz RTD research and outlines the present status and major recent achievements in the field. The paper is organised as follows. Section 2 describes the basic principles of operation of RTDs and their basic properties. Section 3 describes different types of RTDs. Section 4 shortly outlines the development of RTDs and RTD oscillators in the past 50 years. Section 5 describes different types of RTD oscillators. Section 6 is concerned with the use of RTDs as detectors of THz radiation. Section 7 is devoted to discussion on HF limitations of RTDs and how they could be modelled in an accurate and simple way. An accurate linear model of the RTD admittance is crucial for the development of THz RTD oscillators and for understanding the limitations of RTDs. Section 8 is devoted to a qualitative discussion on non-linear RTD dynamics. Section 9 describes the present status of RTD oscillators and several application examples. Section 10 concludes the review.

2 Principles of Operation of RTDs

In the semiconductor heterostructure RTDs, the tunnel barriers are sandwiched between highly doped contact layers. The band diagrams together with the typical

I–V curves of a double-barrier RTD are sketched in Fig. 1. For double barriers with only weakly transparent individual barriers (with the tunnel coefficients T_1 and T_2 , where $T_1 \ll 1$ and $T_2 \ll 1$), the resonant component of the current dominates over non-resonant one [2]. In this case, the total tunnel transparency (T_{total}) of the double barriers will become large for electrons with the energy equal to or close to the energy levels of the quantised states in the quantum well (QW) between RTD barriers. For symmetric barriers, T_{total} is as high as 1 in the resonance [1]. However, T_{total} is extremely low, if the energy of electrons is not coinciding with one of the quantised states in the QW: T_{total} is $\sim T_1 T_2$. That is, the double barriers are basically only transparent for the resonant electrons. An instructive discussion of resonant tunnelling has been given in [15], where the above particularities have been addressed.

A particular N-shaped form of the RTD I–V curve with the NDC region is a consequence of the resonant tunnelling through the double barriers. The heterostructure and doping parameters of an RTD are usually chosen in such a way, that at low biases (U), the bottom of the ground subband in the QW (E_{QW}) is above the emitter Fermi level (E_F^e , Fermi level on the left-hand side of the RTD in Fig. 1a). Consequently, the resonant-tunnelling current through the RTD is close to zero at low biases. At higher applied biases (Fig. 1b), E_{QW} is shifting down below E_F^e and the resonant-tunnelling current is starting to flow. The current is roughly proportional to the number of the subband states between E_F^e and E_{QW} ; therefore, the RTD current keeps growing with further increase of bias, this is the positive-differential-conductance (PDC) region of the I–V curve. However, at some point (Fig. 1c), E_{QW} shifts below the conduction-band bottom on the emitter side of RTD (E_C^e), the elastic resonant tunnelling (the

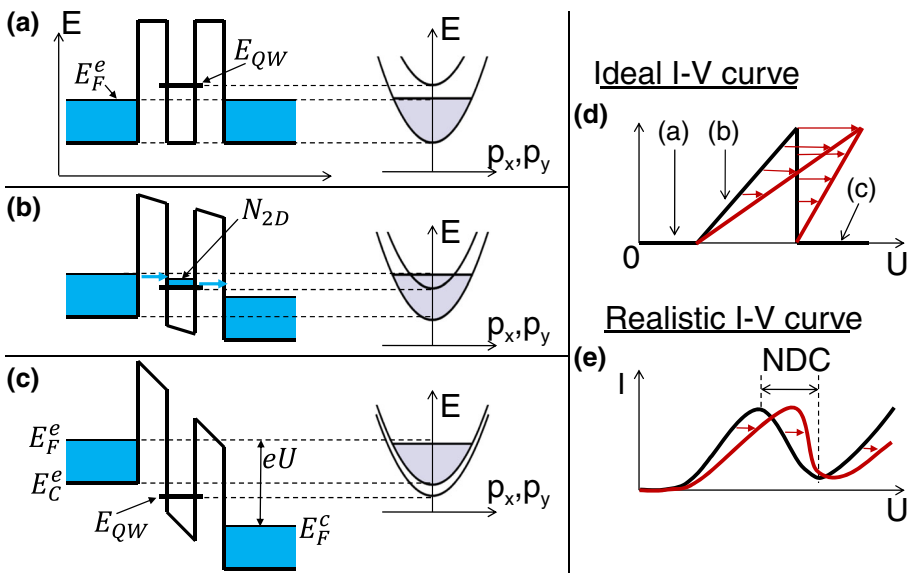


Fig. 1 Schematic RTD band diagram at different bias voltages (a–c); sketches of the RTD I–V curve (d, e), where the black and red lines illustrate I–V curves, when the space-charge effects (Coulomb interaction) are neglected and taken into account, respectively

electron momentum in the plane of the barriers and energy are conserved) becomes impossible in this case; there are no more occupied states in emitter, which are in resonance with the QW states. The RTD current suddenly drops at such biases. The I–V curve of an ideal RTD is sketched in Fig. 1d; it has a triangular shape. A realistic I–V curve is more smooth due to scattering effects in RTDs, non-zero temperature, structural imperfections, etc., but good-quality RTDs do have a distinctive N-shaped I–V curve with a prominent NDC region (see Fig. 1e) and experimental I–V curves in Fig. 2.

Implicitly, we were discussing above a coherent model of elastic resonant tunnelling: the wave function of an electron incident on the barriers extends all the way from emitter through the QW to collector. One calculates the tunnel probability of the stationary states, which are coherent in the whole structure. The model has been used to describe the properties of RTDs from the beginning [1–3]. However, later on, a sequential-tunnelling model has been suggested [17, 18], where the electron tunnelling is described as a two-stage process. First, an electron is elastically tunnelling from the emitter into the QW. Then, a scattering event is taking place, the coherency is broken, but the electron is still staying in the same QW subband, the QW electrons could be also getting thermalised. As a second step, the QW electrons are elastically tunnelling further into the collector. Although the two models seem to describe different mechanisms of resonant tunnelling, they turn out to be intimately related, e.g., the I–V curves described by both models are exactly the same [19], if we disregard difference in the broadening of the resonant states due to the scattering processes; also, the electron transit time through the barriers for both models is determined by the same electron tunnel lifetime (τ) in the QW between the barriers [20, 21].

Clearly, when the barriers are thick, the time electrons spend in the QW (τ) is long and the sequential tunnelling should dominate the tunnel electron transport through RTDs, at least at room temperatures. When the RTD barriers are getting thin, then the coherent tunnelling might play essential role, since τ might eventually become

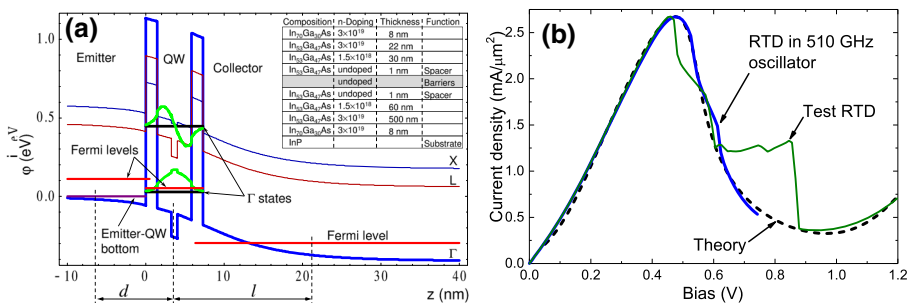


Fig. 2 Reprinted from [16], with the permission of Europhysics Letters. **a** Band diagram of an RTD. The inset shows the layer sequence of the structure, with two symmetrical 1.6 nm AlAs barriers sandwiching $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{InAs}/\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ QW with a 1.2-nm nominal thickness of each layer. **b** Measured (continuous lines) and calculated (black-dashed line) I–V curves of the RTD. The green curve corresponds to an unstabilised test RTD, the I–V curve is strongly deteriorated in the NDC region due to low-frequency parasitic oscillations. The blue line corresponds to an RTD stabilised by a parallel shunt resistor, smooth shoulder in the NDC region is due to oscillations of the diode in a resonator (slot antenna) at 510 GHz

shorter than the electron phase-breaking time. However, there are no clear evidences of coherent tunnelling at room temperatures so far, it is not even clear, what should be the signatures of coherent tunnelling to pinpoint it in experiment. The room-temperature experimental data on electron transport in RTDs are well described by the sequential-tunnelling model so far; therefore, the model is used predominantly in the analysis of RTDs.

Now, we turn to discussion on several effects, which make the measured RTD I–V curves look quite different compared to the simplistic picture shown with the black lines in Fig. 1d and e. First, we deal with the space-charge effects due to accumulation of electrons in the QW. In the sequential-tunnelling approximation, one can write the QW→collector tunnel current density (j_{wc}) in the form as follows:

$$j_{wc} = -\frac{eN_{2D}}{\tau_c}, \quad (1)$$

where $-e$ is the electron charge, τ_c is the lifetime due to electron tunnelling through the collector barrier, and N_{2D} is the 2D electron concentration in the QW. In the dc case (with “0” superscript we denote the static and quasi-static parameters), the RTD current density (j_{RTD}) is the same through each of the barriers, i.e., $j_{RTD}^0 = j_{wc}$. Discussing the ideal I–V curve in Fig. 1, we were disregarding the space-charge effects. Equation 1 shows that N_{2D} is non-zero, where j_{RTD}^0 is non-zero, and N_{2D} gets higher when the current is high. N_{2D} creates a space charge in the QW. Qualitatively, the potential of the QW shifts upwards when we switch on the space-charge effects (Coulomb interaction). To bring the QW states back to the same position relative to the conduction-band bottom in emitter (that would restore the value of j_{RTD}^0), we need to apply higher bias. As a consequence, all current points will shift to the right in the I–V curve (see Fig. 1d and e). The ideal I–V curve will get a “Z” shape with a bistability region [22, 23]. Although exotic, such I–V curves are observable in experiment at low temperatures in high-quality RTDs [24–26]. The more common RTD I–V curves sketched in Fig. 1e do not get the exotic “Z” shape, they preserve the simple “N” shape. However, the regions with higher current density shift significantly to the right, i.e., the I–V curve shape is strongly affected by the space-charge effects in this case as well. The NDC region becomes more steep and the NDC gets larger due to the space-charge effects. The effects have significant impact on the dc I–V curve of almost every RTD. We will come back to the importance of the space-charge effects, when we will be discussing the dynamic properties of RTDs.

There are other effects, which have influence on the realistic I–V curves. Particularly, LO-phonon assisted tunnelling can create a replica peak at the biases higher than those of the main peak [27]. In strongly asymmetric RTDs with thick collector barrier, the phonon-replica peak could be almost as high, as the main peak [25].

Another ubiquitous effect is parasitic oscillations in the external biasing circuit. Due to such oscillations, an averaged dc current is measured, rather than the genuine static RTD I–V curve (see Fig. 2b.) As a result, the averaged I–V curve exhibits a plateau-like behaviour and hysteresis in the NDC region [28–31]. Those plateaus could be confused with intrinsic space-charge bistability, features due to 2D subbands in the space-charge accumulation layer in emitter or other effects; sometimes, it is challenging to identify intrinsic features in the RTD I–V curves [29, 32, 33].

In general, the parasitic oscillations are almost always present and one unavoidably sees plateaus and hysteresis in the dc RTD I–V curve, if special precautions are not taken [29], e.g., one intentionally fabricates small-area RTDs to reduce their NDC (the total NDC is proportional to the RTD area) or one can also include a shunt resistor in parallel to the RTD to reduce or even eliminate the total NDC; the shunt resistor contribution to the total current should be then subtracted during processing of the measurement data (see blue curve in Fig. 2b).

3 Different Types of RTDs

Figure 3 gives an overview of different RTDs, which are relevant or could be relevant for THz applications. A most common double-barrier RTD is depicted in Fig. 3a. In former times, such RTDs were predominantly based on GaAs structures with AlAs or AlGaAs barriers. RTD oscillators with such RTDs have been working up to 420 GHz [28, 31, 34, 35]. The advantage of the material system is that it is lattice matched, the disadvantage is that the barrier height is relatively low and that leads to rather high thermionic-emission current, which is responsible for high valley current at room

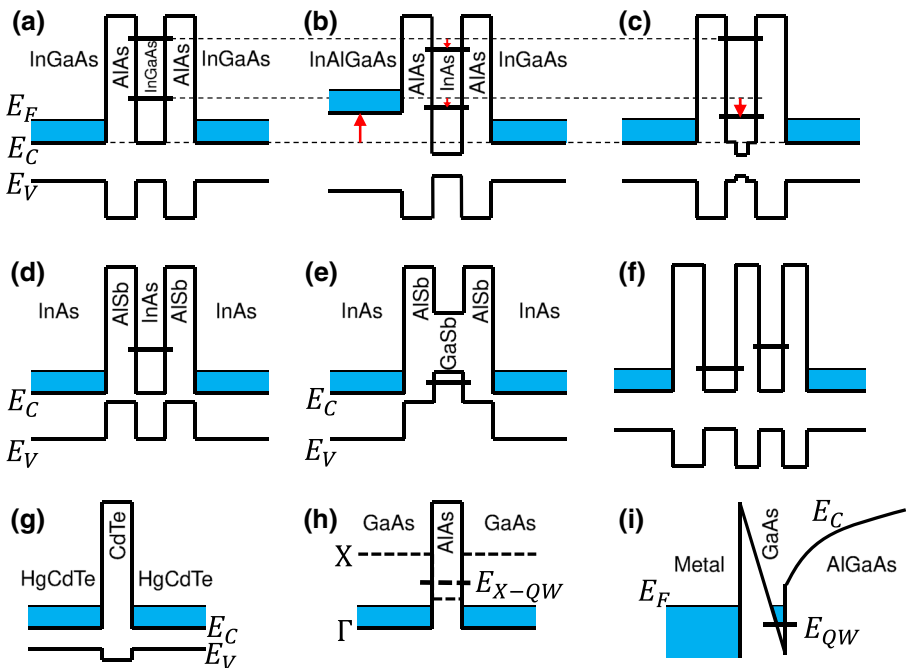


Fig. 3 Different types of RTDs. **a** Conventional double-barrier RTD. **b** An RTD with a “step emitter”. **c** An RTD with a composite QW. **d** and **e** antimonide-containing RTDs with intra-band and inter-band tunnelling, respectively; **f** triple-barrier RTD; **g** single-barrier RTD with tunnelling close to the valence band; **h** Γ -X- Γ single-barrier RTD; **i** tunnel Schottky contact with narrow 2D channel

temperature, e.g., peak-to-valley current ratio (PVCR) is around 1.5 in such sub-THz diodes [35]. To overcome the problem, InGaAs structures with AlAs barriers have been put forward, where the barrier height is ≈ 1.2 eV and the valley current could be suppressed, 650 GHz oscillators have been demonstrated with the basic InGaAs/AlAs RTD design [36]. That are strained heterostructures, therefore AlAs barrier thickness is limited to $\lesssim 4.5$ nm [37].

GaN/AlN RTDs with basic designs have been demonstrated recently [38–40]. However, the RTD band-structure profile of Fig. 3a is strongly deteriorated by the polarisation charges in this case. High current density of ≈ 4 mA/ μm^2 (although with rather limited PVCR of $\lesssim 1.5$) [39] and even oscillators at ≈ 1 GHz [40] have been reported recently with GaN/AlN RTDs. One of the challenges of GaN/AlN RTDs in view of THz applications is the relatively high contact resistance of GaN, which is so far 1–2 orders of magnitude higher than that in RTDs based in InGaAs/AlAs material system [39, 40].

Several modifications of the basic design in Fig. 3a have been used in the recent years to improve the performance of THz RTDs. To achieve high NDC, the RTD barriers are made thin (≈ 1 nm), that leads to very high current density up to ≈ 30 mA/ μm^2 [10]. As a consequence, RTDs are getting prone to thermal break down. To reduce heating, it is desirable to shift the NDC region to the lower bias. That has led to the “step”- or “graded”-emitter design of RTDs, see Fig. 3b, where an additional AlInGaAs layer is incorporated into the emitter to rise the conduction-band bottom there. Additionally, the QW is made with high indium content to shift the QW bottom and ground quantum subband down. That shifts NDC region to lower bias. Such designs have been used in the first 1 THz RTD oscillator [8] and later on in higher-frequency oscillators [41, 42]. An alternative solution is to incorporate an InAs sub-well in the middle of the QW (see Figs. 2a and 3c). As a result, the ground subband shifts down, which is shifting NDC to lower bias. However, the second subband remains almost unaffected by the sub-well in the QW, the separation between the ground and second subbands gets larger and that leads to suppression of the valley current, PVCR gets higher in such RTDs, e.g., room-temperature PVCR of 50 has been reported in [43]. Such RTDs have been used up to 1.1 THz so far [9, 16, 44].

Antimonide-containing RTDs have been also investigated in the past years. A typical design of an InAs/AlSb intra-band RTD is sketched in Fig. 3d. The advantages of such RTDs are related to InAs in the transport layers surrounding RTD, since one can make very low-resistance ohmic contacts to InAs, and to type-II band offset at the InAs/AlSb heterojunction, that leads to higher tunnel transparency of the barriers (compared to GaAs or InGaAs RTDs with AlAs barriers of the same thickness), although the barriers remain high [45]. Such RTDs have been used in RTD oscillators up to 712 GHz [7]. Making use of type-II band offsets, one can also realise inter-band RTDs sketched in Fig. 3e. Electrons are tunnelling resonantly to the quantised hole states in this case [46, 47]. In a way, the mechanism of NDC in such structures could be similar to that in Esaki tunnel diodes: the valence-band states are shifted below the conduction-band bottom with increase of bias and the current drops. The inter-band RTDs can exhibit quite high PVCR of ≈ 20 at room temperature [46, 47], but they have not been used in sub-THz or THz oscillators so far.

Triple-barrier RTDs, see Fig. 3f, have been demonstrated to work up to ≈ 500 GHz in the oscillators [13, 48]. In such RTDs, the current flows due to resonant tunnelling between subbands in the neighbouring QWs. The NDC is arising due to the shift and misalignment of the QW subbands with increase of bias. In the context of QCLs, one calls such type of active-layer designs and inter-well electron transitions as “diagonal”. The advantage of the triple-barrier RTDs is their higher flexibility in the design of the I–V curve and NDC region. That is particularly beneficial for RTD detectors (see more details in Section 6.) The disadvantage of such RTDs is that the achieved current density [13, 48] is so far approximately an order of magnitude lower than that in the double-barrier THz RTDs [10].

Several more-exotic device concepts with single tunnel barriers and NDC have been suggested, although they have never been used in sub-THz and THz devices so far. Figure 3g shows a single-barrier structure with particular band offsets. In the structure, the electrons are non-resonantly tunnelling through the barrier at energies closer to the valence, rather than to the conduction band (in the barrier). Therefore, the barrier transparency declines with increase of the energy of the incident electrons, i.e., the barrier becomes less transparent with increase of bias, that leads to NDC. NDC in such structures at cryogenic temperatures has been demonstrated in [49], the authors suggest that the mechanism should be also working at room temperature in other material systems. Strictly speaking, it is not an RTD, NDC appears due to non-resonant tunnelling.

Another type of single-barrier RTD has been suggested in [50], where the resonant current flows via the inter-valley tunnelling through the quantised X states in the barrier (see Fig. 3h). $\Gamma - X$ coupling in this case is analogous to the tunnel coupling through a single barrier in a double-barrier RTD. NDC in such single-barrier RTDs has not been demonstrated experimentally so far, but such RTDs is a very intriguing concept and there are experimental results showing that the mechanism of NDC is realistic [51].

One more exotic single-barrier RTD is sketched in Fig. 3i, where electrons are tunnelling resonantly through a triangular Schottky barrier from a metal gate into a quantum subband in a narrow 2D channel, e.g., in a high electron-mobility transistor (it will be a diode, if we use just one contact to the channel) [52, 53]. Similar to the mechanism discussed in Fig. 3g, the Schottky barrier will get higher and less transparent with increase of bias for electrons tunnelling resonantly from the gate into the channel, the mechanism leads to NDC [52, 53]. Reduction of the tunnel transparency of the Schottky barrier with bias for resonantly tunnelling electrons in such GaAs structures has been demonstrated experimentally [54], the results also indicate that NDC should be achievable in Al/InGaAs/InAlAs or Al/GaN/AlGaIn structures.

Another very intriguing RTD type demonstrated in the last years is based on graphene, where electrons are tunnelling resonantly between two graphene layers separated by boron-nitride barrier with the thickness of few atomic layers [55]. NDC in such structures appears due to misalignment of the electronic spectra in the neighbouring graphene layers with the change of bias. Such RTDs are expected to be working in the sub-THz range [56].

4 Early Development of RTDs and RTD Oscillators

An RTD is an active device, its NDC can compensate for the losses in a resonant circuit or a resonator. That leads to a simple concept of an RTD oscillator. First, one takes a resonator. Any resonator has losses, e.g., ohmic and radiation losses. In the most simple case, the resonator could be a resonant LC circuit with the losses represented by a (positive) resistor (Fig. 4a). Then, we connect an RTD to the resonator, apply a bias in the NDC region and the NDC (if large enough) will compensate for the resonator losses. The resonator will then start oscillating by itself, i.e., the RTD will turn it into an oscillator. In addition, as mentioned in the introduction, tunnelling can be an extremely fast process. These considerations have fuelled expectation, that very high frequencies could be achieved with RTD oscillators.

Indeed, the performance of RTD oscillators was improving very fast in the 1980's last century. Fundamental oscillations at relatively low frequencies (LF) up to 9 GHz at 200 K have been reported in 1984 [28]. A room-temperature RTD oscillator at the fundamental frequency of 56 GHz has been demonstrated in 1987 [34], at 200 GHz in 1988 [31], at 420 GHz in 1989 [35] and at 712 GHz in 1991 [7]. The development of RTD oscillators was progressing very fast. However, no further progress has been

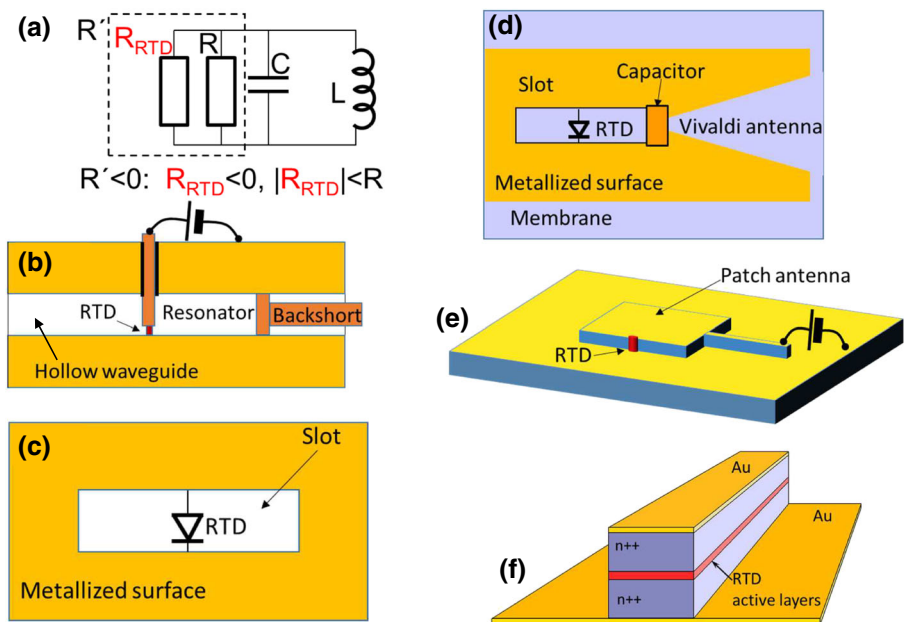


Fig. 4 Different types of RTD oscillators. **a** Schematic of a simple LC RTD oscillator, R is the resonant-circuit loss resistance, $R_{RTD} < 0$ is the RTD resistance in the NDC region. **b** Hollow-waveguide oscillators, the reported highest frequency is ≈ 0.7 THz [7]. **c** On-chip slot-antenna oscillators, the highest achieved frequency is 1.98 THz [8, 10, 36, 42]. **d** Membrane slot-antenna oscillators, the highest frequency is ≈ 1.1 THz [9]. **e** Patch-Antenna oscillators, the highest achieved frequency is ≈ 0.5 THz [13]. **f** Travelling-wave oscillators, they are expected to operate at THz frequencies [59]

achieved in the subsequent almost 20 years. RTD oscillators started to be considered as a not viable technology and they were abandoned until THz RTD oscillators have been demonstrated in 2010 and 2011 [8, 9].

In the context of the development of RTD oscillators, there was always a question standing in the background: what are the frequency limitations of RTDs? In the early years, the publication [20] was taken with much surprise, where a rectification response of an RTD was measured up to 2.5 THz. Although RTDs were not working as active devices in the experiment, operation of RTDs at so high frequencies was not expected. This puzzling result has stimulated the idea of sequential tunnelling in RTDs [17] and incorporation of tunnel lifetime in the analysis of dynamic RTD response [21, 57]. Later on, operation of RTDs as rectifiers has been demonstrated at even higher frequencies of 3.9 THz [58].

5 Different Types of RTD Oscillators

Oscillators based on diodes with NDC are investigated for many decades already. An early overview of different oscillator concepts developed for Esaki diodes can be found in [60]. Meanwhile, different types of RTD oscillators have been also demonstrated and studied. The early development has started with coaxial resonators up to 9 GHz [28], where an RTD was mounted at the bottom of a cylindrical cavity and it was contacted by a needle (with a backshort at some distance in the cavity) going through the centre of the cylinder. Hollow-waveguide resonators have been used at higher frequencies up to 0.7 THz [7, 31, 34, 35]. Such resonators consist of a piece of a hollow waveguide with a (movable) backshort at one end and a mount for an RTD with a whisker contact at the other end (see schematic in Fig. 4b). The movable backshort is used to tune the resonance frequency and to verify that the oscillator is operating at fundamental frequency. It is quite challenging to handle such resonators, therefore they are not in use presently.

While measuring the characteristics of RTD oscillators, special attention is required to check and to prove that the measured signal indeed corresponds to the fundamental oscillation frequency. Examples showing that it is a challenging and not a trivial task could be found in [7, 35]. This is one of the critical points in the analysis of the measurements of RTD oscillators. Unfortunately, this issue was disregarded in some of the subsequent publications.

Later on, the concept of slot-antenna-integrated RTD oscillators has been put forward in 1997 [36]. Slot antenna is a simple resonant type of antenna; therefore it can be used as a resonator and relatively easily integrated with an RTD in the fabrication process Fig. 4c. These are very important advantages of the slot-antenna resonators: THz RTDs usually have sub- μm dimensions and monolithic integration of RTDs in a resonator becomes crucial at THz frequencies. However, at low frequencies up to around 100 GHz, discrete RTDs fabricated separately could be also soldered onto a slot antennas [44]. The concept of slot-antenna RTD oscillators was adopted and further developed by the group of M. Asada and eventually brought to operation at frequencies up to 1.98 THz recently [8, 10, 41, 61]. Meanwhile, slot-antenna RTD oscillators became a standard type of RTD oscillators at sub-THz and

THz frequencies. Such oscillators are usually fabricated on semiconductor substrates with a high dielectric constant; therefore, the radiation is predominantly emitted into the substrate. To couple the radiation out and collimate it, the chips are mounted on top of a hemispherical Si lenses.

The slot-antenna oscillators have been also modified in different ways. The RTD capacitance is quite high, it is typically in the range of 1–10 fF/ μm^2 , it is usually dominating over the eigen slot-antenna capacitance. As a consequence, the slot antenna length in THz oscillators is becoming smaller and much smaller than the radiation wavelength. The radiation efficiency of such small slot antennas becomes poor and that limits the radiated power of the oscillators. To circumvent the problem, asymmetrical slot antennas have been used in the RTD oscillators: RTD is placed close to one end of the slot, the short end works as a resonator and defines the resonance frequency, while the long end of the slot works as an efficient radiator. The approach allows one to increase the output power of the oscillators, e.g., $\approx 400 \mu\text{W}$ at $\approx 550 \text{ GHz}$ have been achieved in such oscillators [12].

Another advantage of the slot-antenna RTD oscillators is that one can relatively easily fabricate an integrated array of such oscillators. That has been presented already in the very first publication on the slot-antenna RTD oscillators [36]. Using small number of oscillators in the array, one can achieve mutual locking of their oscillation frequencies and combine their output power coherently, e.g., two-element arrays were reported to emit 270 and 180 μW at 770 and 810 GHz, respectively, [12]. Due to variation of the parameters of the neighbouring oscillators (because of technological imperfections), it is quite difficult to achieve mutual locking of a larger number of oscillators [12].

The other modification of a simple slot-antenna resonator is a membrane RTD oscillator, when a slot antenna integrated with an RTD is fabricated on a thin dielectric membrane. Such resonators are usually also integrated with an additional planar broad-band (Vivaldi) antenna to improve the radiation efficiency of the oscillator (Fig. 4d). This kind of oscillators on a thinned $\approx 20\text{-}\mu\text{m}$ -thick InP substrate operating at 400 GHz has been reported in [62]. The highest frequency achieved with the oscillators is 1.1 THz [9], the oscillators were fabricated on few- μm thick spin-on dielectric membrane in this case. Such oscillators could be extremely tiny: $\approx 500 \times 500 \mu\text{m}^2$ [9].

The idea of integrating an RTD with a patch antenna, see Fig. 4e, which is also a resonant type of antenna, is around for relatively long time as well. In the beginning, the patch antenna was used to improve the coupling of the external radiation into an RTD [63]. Later on, an RTD oscillator at $\approx 500 \text{ GHz}$ based on patch-antenna resonator has been also reported [13]. The advantage of this approach is that the radiation is emitted by the antenna upwards from the chip and the hemispherical Si lens is not needed in such oscillators; therefore, the dimensions of the oscillators could be rather small.

Meanwhile, several types of more complex oscillators have been reported, which combine, e.g., a slot and patch antennas [64] or a slot and Yagi-Uda antennas [65] or even more complex designs, where a slot antenna is complemented by an additional array of corner-shaped slots to emit circular-polarised radiation [66]. The advantage of such oscillators is that one is forcing the slot antennas to radiate in the air, rather

than to the substrate and one can get better control over the radiation pattern of the oscillators.

The concept of travelling-wave distributed oscillators with NDC active layers has been discussed already quite long time ago [60]. A detailed theoretical analysis of travelling-wave microstrip RTD oscillators was presented in [59], see schematic of the oscillator in Fig. 4f. It turns out, that such oscillators should be working at 1–2 THz at room temperatures and perhaps even at higher frequencies. Due to larger active volume in microstrips, compared to the lumped-element oscillators, one can expect generation of higher output THz power in such oscillators. From another perspective, microstrip RTD oscillators could be seen as THz QCLs with metal-metal waveguide, where the whole cascade is replaced by its single period—an RTD. Contrary to THz QCLs, microstrip RTD oscillators do not require cryogenic cooling and they should be working at room temperature. The reason for that could be rather basic [59]: the NDC region can be preferable to achieve higher gain in the THz QCLs (with “diagonal” active transitions) [67], but QCLs cannot be operated in the region due to build up of the high-field domains. On the other hand, the QW in an RTD has direct contacts on both sides, which keep the RTD active region stable. Therefore, RTDs could be easily operated in the NDC region in their optimum regime. That is probably the main reason why RTD oscillators are working at room temperatures, although THz QCLs do not.

One more type of RTD oscillators is harmonic oscillators. Harmonic emission has been also reported long time ago, e.g., 87 GHz emission at the second harmonic was reported in [34]. The first report on THz emission from RTD oscillators was also at the third harmonic in [68]. The highest frequency harmonic emission was at 1.52 THz in triple-push oscillators (third harmonic), the output power was 1.9 μW [69].

6 RTDs as Sub-THz and THz Detectors

Experimental investigation of the maximum fundamental operating frequency of RTD oscillators is one of the ways to find out the frequency limitations of RTDs. However, RTDs are working in the NDC region in this case and they are combined with a lossy resonator in an RTD oscillator, therefore it is an investigation of the limitations of a coupled system, rather than an RTD alone. From the point of view of analysis of the frequency limitations and frequency-dependent characteristics of RTDs, the rectified (usually quadratic) response of RTDs can give extended or complementary information, compared to the RTD oscillators. Particularly, we can measure frequency-dependent detector response in the PDC region, where RTD oscillators cannot operate; we can also flexibly choose the signal amplitude applied to an RTD: the RTD relaxation time constants are, generally speaking, dependent on the signal amplitude in the non-linear regime of operation.

The frequency dependence of an RTD-detector response is determined by the interplay of several factors. The rectification response of an RTD is determined by the amplitude of the signal across the diode and by the non-linearity of the RTD characteristics. The amplitude depends on the impedance matching between the RTD and the waveguide/antenna, on the external RTD parasitics and on the

RTD impedance. The RTD impedance is frequency and, generally speaking, also signal-amplitude dependent; the impedance inherently depends on the small-signal or large-signal charge-relaxation time constants. The very same RTD impedance (in the NDC region) also determines the oscillation conditions and the output power of RTD oscillators. In the small signal operating regime of the detectors, the RTD impedance is adequately described by a linear RTD model, e.g., by a model outlined in Section 7 with the corresponding linear RTD impedance and response time. Nevertheless, the model still needs to be extended or complimented by some additional assumptions to estimate the non-linear (although small signal) response of an RTD detector. A large-signal model is required in the general case, the non-linear RTD dynamics will be qualitatively discussed in more detail in Section 8. However, we note that an accurate analysis of the non-linear HF RTD characteristics is still an open topic, it still remains to be done.

The non-linear characteristics are also of high interest for practical application of RTDs in THz detectors. There are two regions with large non-linearity in the I–V curve of a double-barrier RTD (see Fig. 5a). First, it is the region (A) of the onset of the resonant-tunnelling current, where the current starts flowing through the RTD due to resonant tunnelling of electrons in the tail of the electron energy distribution function in the RTD emitter. The RTD current density in the region is approximately described by the thermal exponent:

$$j_{RTD} \propto \exp\left(\frac{\alpha e U_{RTD}}{kT}\right), \tag{2}$$

where U_{RTD} is the bias applied to the RTD, k is the Boltzmann constant, T is the temperature, the coefficient $\alpha = d/(d + l)$ describes which part of the RTD bias falls between the emitter and QW, d and l are the effective thicknesses of the emitter-QW and QW-collector regions, respectively; the thicknesses include the barrier thicknesses, depletion region in collector and screening-lengths (in emitter and collector) (see Fig. 2a). Except for the factor α , this thermal exponent is similar to that in Schottky diodes. Correspondingly, the RTD non-linearity in the region is weaker (by the

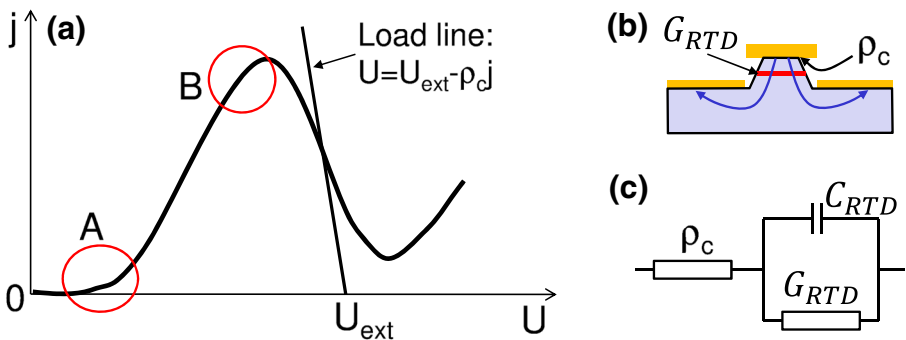


Fig. 5 a A sketch of an RTD I–V curve with highly non-linear regions (A and B) in the PDC range, which are usable for detection. b Schematic of an RTD with top and bottom contacts to n++ layers. c A simplified equivalent circuit of an RTD with a contact resistance

factor α) than that in the Schottky diodes. Additionally, the space-charge effects were neglected in Eq. 2, the effects should weaken the RTD non-linearity even further.

The second region (B in Fig. 5a) with significant non-linearity is close to the current peak, when the resonant current is starting to switch off with increase of bias. The non-linearity is not due to the thermal exponent in this case; it is determined by the broadening of the QW resonant subband instead. Since the broadening can be much smaller than $kT \approx 25$ meV (at room temperature), the responsivity of an RTD detector can be much larger than that of a typical Schottky diode detector. That gives a fundamental advantage to RTD detectors, which can be elucidated by defining the current responsivity of a detector in a usual way as $I''/2I'$ (where $'$ denotes the derivative of $I(U)$). In a Schottky diode, the maximum responsivity is as follows:

$$\frac{I''}{2I'} = \frac{e}{2kT}, \quad (3)$$

which is equal to ≈ 20 A/W at room temperature. An alternative type of common detectors gaining ground in the last decade is based on rectification in the channel of field-effect transistors (FETs) operated above cut-off frequencies. The current responsivity of FETs turns out to be twice less than the value given by Eq. 3 [70]. These numbers determine the fundamental limitations of those common detectors. However, since the RTD non-linearity in the region (B) in Fig. 5a is determined by the broadening of the resonant subband, RTD detectors can overcome the limitation. For example, a factor of ≈ 12 dB higher sensitivity of RTDs compared to Schottky diodes has been experimentally determined in [71].

The triple-barrier RTDs could be also used in detectors. Their advantage is that the non-linearity at the onset of the resonant current is not due to the thermal tail in the electron distribution function; it is due to the alignment of the 2D subbands in the neighbouring QWs, i.e., the non-linearity can be strong and the current responsivity of such RTDs can exceed the thermal limit of Eq. 3 not only in the region (B) but also in the region (A) in Fig. 5a. Such diodes could be even designed in a way, that the region (A) is shifted to zero bias [72]. The superior performance of such triple-barrier RTDs is not demonstrated experimentally so far but some preliminary results are published recently in [48, 72]. The disadvantage of such RTDs is that their current density is lower than that of double-barrier RTDs; their impedance is higher and that results in difficulties to achieve an optimum matching with the detector antenna [72].

The non-linearity of the RTD I–V curve could be easily used in the PDC region, although, the NDC region should be avoided, since RTD becomes unstable in the region and it starts to oscillate by itself. However, the RTD non-linearity is the largest near the peak of the RTD I–V curve, i.e., quite close to the NDC region. That imposes a limitation on the amplitude of the detected signal: the oscillating RTD bias should be small, it should not go into NDC region. As a consequence, the dynamic range of non-linear RTD detectors is more limited than that of a Schottky diode detector, although their sensitivity could be much higher.

RTD detectors can operate at very high frequencies. 2.5 THz detectors (rectification signal) have been reported already very long time ago [20]. The highest reported operating frequency of RTD detectors is 3.9 THz [58], RTD was whisker contacted

and mounted in a corner-cube reflector in this case. A number of antenna-integrated RTD detectors have been reported more recently both with triple- and double-barrier RTDs [48, 71]. Such detectors have been used in imaging [73], high-speed data transmission at 9 Gbit/s [74] and spectrometer [75] experiments at sub-THz frequencies.

The RTD detector could be constructed in a similar way as an RTD oscillator. That is an important advantage of the RTD detectors: both detectors and RTD oscillators could be fabricated in the very same process and easily combined together in the transmitter/receiver systems. Even the very same RTD can be operated as a gain media in an oscillator or a detector, depending whether it is biased in the NDC or non-linear PDC region of the I–V curve, respectively, [48]. However, the optimum designs of RTD oscillators and detectors are different: in an oscillator, an RTD must be combined with a resonator or a resonant antenna, whereas it is preferable to combine an RTD with a non-resonant broad-band antenna in the case of a detector.

7 High-Frequency Characteristics of RTDs: Linear Model

The operating frequencies of RTD oscillators are limited due to losses in resonators (e.g., ohmic or radiation losses) and due to limitations of the RTDs themselves, e.g., RTD cannot provide enough NDC to compensate for the resonator losses or the RTD cannot provide any gain at all beyond certain limiting frequency, since NDC turns into PDC with increase of frequency. The properties of resonators are discussed in [41]. Here, we are concerned with the analysis of the RTD limitations, which could be divided into internal and external.

The external limitations arise due to RTD parasitics. The most important one is the specific (per unit area) resistance (ρ_c) of the top RTD contact (see Fig. 5b). Its typical value is 1–10 $\text{Ohm}\cdot\mu\text{m}^2$ in RTDs with InGaAs contact layers. However, it might turn out to be above 100 $\text{Ohm}\cdot\mu\text{m}^2$, if special care is not taken to fabricate good contacts. We can recognise the importance of ρ_c already looking at the load line and the dc I–V curve of an RTD (see Fig. 5a). THz RTDs usually have thin tunnel barriers, high current density and high NDC. The working point on an RTD is in the NDC region and the RTD must be dc stable there, i.e., $\rho_c < -1/G_{\text{RTD}}^0$, where G_{RTD}^0 is the specific static differential conductance of an RTD ($G_{\text{RTD}}^0 < 0$ in the NDC region). For example, an RTD with not the record-high peak current density of $j_{\text{RTD}} \approx 25 \text{ mA}/\mu\text{m}^2$ used in 1.46 THz oscillator in [42] had $1/G_{\text{RTD}}^0 \approx -15 \text{ Ohm}\cdot\mu\text{m}^2$. Correspondingly, the value of ρ_c should be less than $-1/G_{\text{RTD}}^0$ to be able to bias the RTD in the NDC region.

The bottom RTD contact has larger area and it is usually less critical, but its contribution to the total contact resistance of RTD can still be considerable and should be also taken into account, as well as the spreading resistance of the n++ layers (see Fig. 5b). That would mean that ρ_c must be even less than the value in the above assessment. We note at this point, that an elegant solution to the contact-resistance problem has been suggested in [36], where a Schottky collector has been built into an RTD, that allows one to eliminate the top contact ρ_c .

Of course, the contact resistance and other parasitics impose also limitations on the operating frequencies of RTDs: due to parasitics, the negative total conductance of an RTD will reduce and become positive at high frequencies. For example, if we disregard all other parasitics except for ρ_c and assume that RTD layers could be represented by a simple RC model with G_{RTD} and C_{RTD} (specific RTD capacitance) (see Fig. 5c), then the total conductance will be negative only below the maximum angular frequency ω_{max} :

$$\omega_{\text{max}} = \frac{1}{C_{\text{RTD}}} \sqrt{-G_{\text{RTD}} \left(\frac{1}{\rho_c} + G_{\text{RTD}} \right)}. \quad (4)$$

At frequencies $\omega \gtrsim \omega_{\text{max}}$, the total RTD conductance is positive and the RTD cannot provide any amplification/gain. The limitation of Eq. 4 will be somewhat relaxed due to the contact capacitance, which is $\sim 20 \text{ fF}/\mu\text{m}^2$ in typical InGaAs THz RTDs. However, the frequency limitation is getting more severe, when other parasitics are not negligibly small, particularly, bottom-contact resistance, spreading resistance in the n++ layers, additional parasitic capacitances, inductance of the air-bridge, skin effect, etc. Although parasitics are playing a crucial role in the frequency limitations of RTDs, their accurate estimation and assessment could be done in a straightforward manner, in the same way as that is done for Schottky diodes and transistors at sub-THz and THz frequencies.

Internal limitations of RTDs is a more involved problem. During resonant tunnelling, the electrons are getting trapped to and released from the resonant states, i.e., the process is inherently associated with a certain time delay and it might be slow. Therefore, the RTD NDC should decrease with frequency and it can even become positive beyond some limiting frequency. Such effects can impose inherent frequency limitations on RTDs. The way to tackle the problem is to develop physics-based models for the electron transport through RTDs, which should describe the complex RTD admittance or, in other terms, frequency dependence of RTD conductance and capacitance. Such models could be analytical [21, 37, 76–78] or fully based on numerical simulations [79, 80]. Although numerical simulations could be quite general, the interpretation of their results is far from being straightforward. Often, it is difficult to use them, when one wants to single out main effects responsible for particular behaviour of RTDs. Numerical simulations are not easy to use in practise in the analysis and design of RTDs. They are quite helpful though, when the main effects are identified, qualitatively clear and one needs to get more accurate estimates of them. Some analytical models [37, 76, 77], although accurate from the physics point of view, turn out to be quite complicated as well, involve many parameters and could be more appropriate for numerical calculations in practise.

An intuitive, simple and not necessary less accurate description of HF properties of RTDs could be achieved with analytical and semi-analytical models. Meanwhile, several types of such models have been suggested. The most simple one is an RC circuit, where RTD is modelled as a resistor (with the conductance G_{RTD}^0) and capacitance (C_{RTD}^0) connected in parallel (see Fig. 6a). Now and further, we will be referring to the specific (per unit area) parameters (conductance, capacitance, etc.) in the RTD

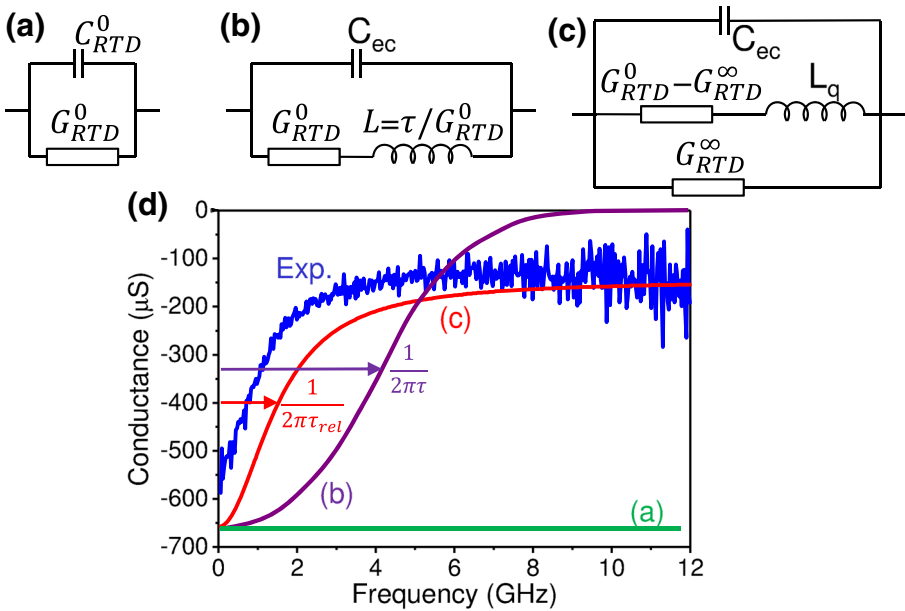


Fig. 6 RTD equivalent circuits. **a** The most simple RC model. **b** RLC model [21]. **c** RLCR model [78, 81], where $L_q = \tau_{rel}/(G^0_{RTD} - G^\infty_{RTD})$. **d** Comparison of the frequency dependence of RTD conductance according to **a** RC, **b** RLC (schematic behaviour), **c** RLCR models and experimental data reported in [82]. Small mismatch at low frequencies is because the theoretical value of G^0_{RTD} is based on a derivative of the calculated I–V curve, which is overlapping quite good, but not ideally with the experimental curve

models. The model is widely used due to its simplicity, but it obviously disregards the electron accumulation in the QW and the time delay due to the resonant tunnelling.

A more accurate and still simple model has been suggested in [21], where the finite tunnel electron lifetime (τ), which electrons spend in the QW during resonant tunnelling, is taken into account. The time delay can be represented by an inductance (L) connected in series with G^0_{RTD} in the RLC model (see Fig. 6b). Qualitatively, the current through an RTD does not follow instantaneously the variation of the applied bias, it takes time τ for the current to adjust to the bias variations. The time delay is analogous to the effect of an inductance with the value of $L = \tau/G^0_{RTD}$ in series with a resistor with conductance G^0_{RTD} . C_{ec} in Fig. 6b is the specific geometrical emitter-collector RTD capacitance. The model predicts, that the RTD conductance is decreasing with increase of frequency, the rate of the decrease is related to the time constant τ (see line (b) in Fig. 6d). Lower RTD conductance at high frequencies together with account of contact resistance ρ_c leads to a lower bound for ω_{max} , where RTD conductance becomes positive [21].

However, the space-charge (Coulomb interaction) effects have been neglected in the above RLC model. The effects have a strong impact on the electron transport in two ways. First, the variation of the electron charge in the QW shifts the QW potential the same way, as we have seen that in the discussion on the static I–V characteristics of RTDs in Section 2. It turns out, that the (dynamic) shift has a profound effect on

the rate of charge relaxation processes: as a result, the QW-charge relaxation process will be described by a relaxation time constant (τ_{rel}), rather than τ , and τ_{rel} could be much shorter or much longer than τ [78, 83]. Second, the account of Coulomb interaction effects is switching on the displacement-current mechanism of the current flow through an RTD and through its individual barriers [78, 81]. For example, the real electron current might be flowing only between the emitter and QW, but the displacement current in the collector region makes it flowing through the whole structure and appear in the current through the external contacts of RTD. It is a sort of additional instantaneous (no delays) current channel through the RTD. It turns out, that both effects could be included into a slightly modified RLCR equivalent-circuit model of an RTD [78, 81] (see Fig. 6c). The change in the relaxation time modifies the equivalent “quantum” inductance (L_q) as follows:

$$L_q = \frac{\tau_{\text{rel}}}{G_{\text{RTD}}^0 - G_{\text{RTD}}^\infty} \quad (5)$$

and the displacement-current mechanism adds an additional parallel current-flow channel, which appears as an additional “HF” parallel resistor (G_{RTD}^∞). That leads to the following equation for the admittance ($Y_{\text{RTD}}(\omega)$) of the RTD, which could be represented as RLCR circuit in Fig. 6c [78, 81, 83]:

$$Y_{\text{RTD}}(\omega) = i\omega C_{\text{ec}} + G_{\text{RTD}}^\infty + \frac{G_{\text{RTD}}^0 - G_{\text{RTD}}^\infty}{1 + i\omega\tau_{\text{rel}}}. \quad (6)$$

As mentioned above, the role of the inductance in the equivalent RTD circuit is the delay of the current with respect to the bias variation. The delay is equal to L/R in a serial LR circuit. In the RLCR model, the RTD current has both an instantaneous (G_{RTD}^∞) and delayed (by τ_{rel}) current components, represented by the second and the last terms in Eq. 6, respectively and the corresponding branches in the equivalent circuit in Fig. 6c. The inductance L_q in Eq. 5 is positive in the PDC region, but it becomes negative, when $G_{\text{RTD}}^0 - G_{\text{RTD}}^\infty < 0$. However, note, the delay τ_{rel} is always positive. The negative sign of L_q is related to the change in sign of the conductance ($G_{\text{RTD}}^0 - G_{\text{RTD}}^\infty$) in the nominator of the last term in Eq. 6, i.e., a positive change in the bias leads to a negative change in the current, which is delayed by τ_{rel} .

Figure 6d illustrates the difference in the frequency dependence of the RTD conductance ($G_{\text{RTD}} = \text{Re}(Y_{\text{RTD}})$) for different models and compares them with the experimental data from [82]. In this example, we were intentionally studying an RTD with very thick barriers, so that $\tau_{\text{rel}} \approx 100$ ps. The RC model predicts frequency-independent behaviour, which is oversimplified to describe accurately the experimental data. In this example, it would be applicable only below ~ 0.5 –1 GHz. RLC and RLCR models exhibit roll-off of the RTD conductance with frequency, however, at different frequencies of $1/2\pi\tau$ and $1/2\pi\tau_{\text{rel}}$, respectively. RLC model overestimates the roll-off frequency, since it is ascribed to τ , rather than τ_{rel} in the case of RLCR model. The later one demonstrates essentially better agreement with the experimental data, since $\tau_{\text{rel}} > \tau$ in the NDC region of the I–V curve and the roll-off frequency is getting lower [78, 81, 83]. It is also important to notice, that RLC model predicts that the RTD conductance should tend to zero at high frequency, although the RLCR model predicts asymptotic convergence of the RTD conductance

towards a finite value of G_{RTD}^∞ , which is in agreement with experiment: RTD conductance indeed can stay negative at $\omega\tau_{rel} \gg 1$. RLCR model has been compared to other experimental data in [81] and, further on, it has been used extensively in the analysis of different RTD oscillators in [9, 16, 44], where very good agreement with experimental data has been demonstrated.

The RLCR circuit has been derived in [78, 81] based on sequential-tunnelling model [17] of electron transport in RTDs with account of Coulomb interaction effects. It turns out, that one can give a simple and intuitive picture for the mechanism, describing how Coulomb interaction affects the relaxation time. The mechanism is illustrated in Fig. 7. First, we assume that we have a QW separated by a tunnel barrier from a contact (Fig. 7a). We assume, that there are many equidistant quantum states in the QW separated by the energy interval ΔE , each of the states is characterised by the same tunnel lifetime τ . Next, we fill the states in the QW and the contact by electrons up to a certain Fermi level (E_F); this is a stationary state of the system (Fig. 7b). Further, we introduce a perturbation, we move an electron from the contact into the QW, Coulomb interaction is switched off (Fig. 7c). The electron will occupy a state in the QW above the E_F in the contact (temperature is assumed to be zero) and it will tunnel out from the QW with the time constant τ (Fig. 7c). However, when we switch on the Coulomb interaction, the picture is changing (Fig. 7d). The perturbing electron we have moved into the QW will add some charge to the QW, that will shift the QW potential upwards due to charging of the contact-QW capacitance

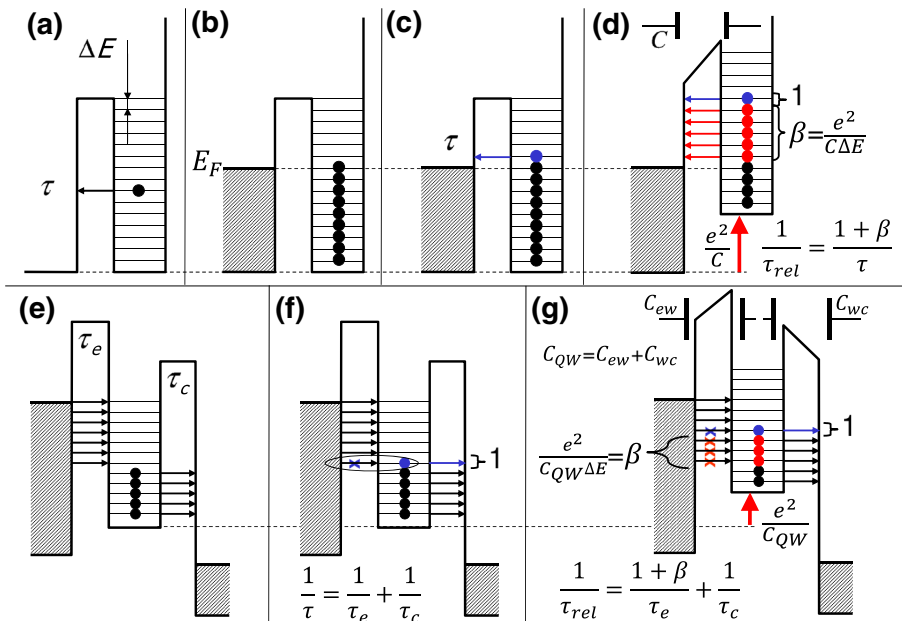


Fig. 7 Mechanism of Coulomb acceleration of the charge relaxation in **a–d** single- and **e–f** double-barrier structures with a series of equidistant quantum states in the QW. **b, e** Steady states of the systems. **c, d, f, g** Perturbation due to an extra electron in the QW with the Coulomb interaction switched off (**c, f**) and switched on (**d, g**)

(C). The corresponding energy shift is e^2/C . That means that not only the electron we have placed initially as a perturbation into the QW, but also $\beta = e^2/C\Delta E$ additional electrons will contribute to the tunnel relaxation current, since the states occupied by the electrons will be shifted above the contact E_F . Noticing that to bring the system back to the initial stationary state, only one electron should tunnel through the barrier, the relaxation time of such process will be $(1 + \beta)$ times shorter than τ , i.e., $\tau_{rel} = \tau/(1 + \beta)$. The mechanism leads to acceleration of the relaxation process by the factor of $(1 + \beta)$.

The same mechanism works in a similar way also in the double-barrier tunnel structures, that is illustrated in Fig. 7e–g. Now, we sandwich QW between two tunnel barriers and two contacts. The emitter and collector barriers are characterised by the tunnel lifetime τ_e and τ_c , respectively. When we apply a bias to the system, a dc current will be flowing through the barriers. This is a stationary state of the system, Fig. 7e. In the same way as above, we introduce a perturbation, we add an electron into the QW and look at the rate of relaxation of the perturbation. First, the Coulomb interaction of the perturbing electron is switched off (Fig. 7f). The electron can tunnel to the collector with the tunnel rate $1/\tau_c$. The electron in the QW also blocks an emitter→QW tunnel channel. The emitter→QW electron flow will reduce and we can represent that as if we are adding a compensating current flow QW→emitter with the value e/τ_e . Adding up these relaxation current flows from QW to emitter and collector, we find that the relaxation time constant is equal to the tunnel electron lifetime τ , where,

$$\frac{1}{\tau} = \frac{1}{\tau_e} + \frac{1}{\tau_c}. \quad (7)$$

This is the relaxation rate, which is taken into account in the RLC model (Fig. 6b). However, when we switch on the Coulomb interaction, the relaxation time is changing (Fig. 7g). The potential of the QW shifts upwards, and now the additional $\beta = e^2/C_{QW}\Delta E$ states will be blocked for tunnelling from emitter into QW, where C_{QW} denotes a combined capacitance between QW and emitter and collector. In result, the relaxation time will be accelerated according to as follows:

$$\frac{1}{\tau_{rel}} = \frac{1 + \beta}{\tau_e} + \frac{1}{\tau_c} = \frac{1}{\tau} + \frac{\beta}{\tau_e}. \quad (8)$$

We did not specify the parameter ΔE so far. In RTDs with 2D QW, electrons are quantised in the QW in one direction, but they can freely move in two other directions and form a subband. Then, ΔE in such structures is equal to $1/\rho_{2D}S$, where ρ_{2D} is the 2D density of states in the QW subband and S is the RTD area. Substituting this expression into the above equation for β , we get the following expression for an RTD with 2D QW:

$$\beta = \frac{e^2\rho_{2D}}{C_{QW}}, \quad (9)$$

where as C_{QW} we denote further the sum of QW-emitter and QW-collector capacitances per unit area.

The above effect might look similar to the Coulomb blockade, but it is essentially different. First, the Coulomb-blockade effect blocks the dc current, although

the above effect describes Coulomb acceleration of the tunnel relaxation process, i.e., it is a dynamic effect. Second, the Coulomb-blockade effect relies on the discretisation of the electrons charges, although for the above effect the discretisation of charges is not important. We could have put an arbitrary number N of electrons into the QW as an initial perturbation in the above examples, then $(1 + \beta)N$ electrons would contribute to charge relaxation and τ_{rel} would stay unchanged: τ_{rel} is a time constant in a *linear* relaxation model. Third, the effect is determined by the parameter $\beta = e^2/C\Delta E$, rather than e^2/CkT in the case of Coulomb blockade. β is temperature independent, therefore, low temperature is not required and the above effect works at room temperatures as well. Fourth, we can notice, that the parameter β in Eq. 9 is independent of the RTD area, i.e., the effect is present in arbitrary large RTDs, since it is determined by the ratio of the 2D electron density in the QW and the specific QW capacitance. On the contrary, extremely small capacitance is required to observe the Coulomb-blockade effect ($e^2/C \gg kT$). We can see from the above discussion, that the Coulomb acceleration of the relaxation process is an ubiquitous effect in RTDs, it is determined by the parameter given by Eq. 9, with the typical values of β for InGaAs or GaAs RTDs in the range ≈ 2 –10. The effect is by far not small.

The above simple mechanism of Coulomb acceleration of relaxation processes works well in the PDC region of the RTD I–V curve: $\beta > 0$ in the PDC region and $\tau > \tau_{\text{rel}}$, as it follows from Eq. 8. In this sense, RTD is a fast device in the PDC region, its relaxation time is shorter than τ . However, an additional effect starts to play an essential role in the NDC region: τ_e is changing rapidly with the Coulomb shift of the QW potential. A detailed analysis of β in the NDC region [78, 83] shows, that $\beta < 0$ there and, as a consequence of Eq. 8, $\tau < \tau_{\text{rel}}$, i.e., RTD is a slow device in the NDC region, its relaxation time is longer than τ .

Both the deformation of the dc I–V curve (discussed in Section 2) and modification of τ_{rel} are related to the QW-potential shift due to Coulomb interaction. Both effects are intimately related and it turns out that one can derive an equation [78, 81, 83] relating the static differential conductance of RTD (G_{RTD}^0) and the ratio τ/τ_{rel} :

$$G_{\text{RTD}}^0 = \frac{C_{\text{wc}}}{\tau_c} \left(1 - \frac{\tau_{\text{rel}}}{\tau} \right), \quad (10)$$

where C_{wc} is the QW-collector capacitance of RTD. The equation turns out to be very helpful for calculation and analysis of τ_{rel} , since G_{RTD}^0 is directly measurable and C_{wc} , τ_c and τ could be assessed rather easily. At this point, we also note that the RLCR model, represented by Eq. 8 and Fig. 6c, is very simple; it contains only four parameters. Out of them, G_{RTD}^0 is directly measurable, G_{RTD}^∞ could be also measured directly, at least in the relatively low-frequency RTDs [82], or evaluated (see below), C_{cc} is easy to calculate relying on the doping profile of the RTD layers. Relying on Eq. 10, we can also evaluate τ_{rel} , then we know all 4 parameters of the model. That makes the RCLR model easy to use in practise.

The impact of the displacement currents on the HF properties of RTDs could be illustrated also in a quite intuitive way with the help of the Shockley-Ramo theorem [84, 85]. According to the theorem, the current density (j_{RTD}) at the external

boundaries of the RTD layers, i.e., in the conducting regions of RTD emitter and collector, is as follows:

$$j_{\text{RTD}} = \frac{d}{d+l} j_{\text{ew}} + \frac{l}{d+l} j_{\text{wc}} + C_{\text{ec}} \frac{\partial U_{\text{RTD}}}{\partial t}, \quad (11)$$

where d and l are defined earlier and illustrated in Fig. 2a; j_{ew} and j_{wc} are the emitter-QW and QW-collector current densities, respectively; generally speaking, $j_{\text{ew}} \neq j_{\text{wc}}$ at high frequencies; U_{RTD} is the bias across RTD layers. Equation 11 is exact, it takes into account both the real electron and displacement currents flowing through the RTD layers. The first two terms in Eq. 11 describe contributions of real currents j_{ew} and j_{wc} to j_{RTD} , where the currents are multiplied by the coupling coefficients containing only geometrical parameters d and l . The mechanism of coupling is via displacement currents. The last term in Eq. 11 describes the capacitive current due to accumulation of charges at the external boundaries of the RTD layers.

The HF behaviour of RTD could be explained with the help of Eq. 11 [78, 81]. At high frequencies, when $\omega\tau_{\text{rel}} \gg 1$, N_{2D} becomes a time independent constant, it cannot follow bias variations at so high frequencies. j_{wc} has non-resonant nature, it changes only due to the variation of the collector-barrier tunnel transparency in this regime. The mechanism gives a positive contribution of j_{wc} to RTD conductance, i.e., RTD collector current contributes to loss at high frequencies. On the other hand, the mechanism of j_{ew} is resonant tunnelling and j_{ew} still gives negative contribution (amplification and gain) to the RTD conductance also at high frequencies. If RTD has a thick depletion region on the collector side ($d \ll l$), then the (lossy) contribution of j_{wc} will be dominant and (gain) contribution of j_{ew} will be suppressed (its contribution in Eq. 11 will have a small pre-factor $d/(d+l)$). As a consequence, RTD conductance will be positive (lossy) at high frequencies in such RTDs. However, in RTDs with highly doped collector $l \approx d$, the pre-factor in front of j_{ew} in Eq. 11 will get higher (≈ 0.5) and the contribution of j_{wc} will get lower (the pre-factor is ≈ 0.5 compared to ≈ 1 in the previous case). As a result, such RTDs should have NDC (gain) in the regime $\omega\tau_{\text{rel}} \gg 1$ [78, 81], i.e., such RTDs have a wide operating-frequency range.

The general RLCR model we have discussed here could be reduced to a simple RC models in the LF and HF limits, although the RC parameters are quite different in these two regimes [78, 81, 83]. The general Eq. 6 is reduced to a LF RC circuit with the admittance:

$$Y_{\text{RTD}}^{\text{LF}}(\omega) \approx i\omega \left(C_{\text{ec}} + \tau_{\text{rel}} \left(G_{\text{RTD}}^{\infty} - G_{\text{RTD}}^0 \right) \right) + G_{\text{RTD}}^0. \quad (12)$$

The circuit is applicable only in the limit: $\omega\tau_{\text{rel}} \ll 1$. Note, that the LF capacitance ($C_{\text{RTD}}^0 = C_{\text{ec}} + \tau_{\text{rel}}(G_{\text{RTD}}^{\infty} - G_{\text{RTD}}^0)$) is not equal to the geometrical RTD capacitance C_{ec} . As described in [78, 81, 83], the additional term can lead to decrease or increase of the capacitance in the PDC region of the I–V curve, but it gives a spike in the LF capacitance in the NDC region measured in [86, 87]. The capacitance spike could be attributed to the negative sign of L_q ; however, the additional term in C_{RTD}^0 indicates that the origin of the spike is the negative sign of the conductance term ($G_{\text{RTD}}^0 - G_{\text{RTD}}^{\infty}$) in Eq. 6.

On the other hand, the RTD admittance is reduced to the following:

$$Y_{\text{RTD}}^{\text{HF}}(\omega) \approx i\omega C_{\text{cc}} + G_{\text{RTD}}^{\infty} \quad (13)$$

in the HF limit ($\omega\tau_{\text{rel}} \gg 1$), with the geometrical capacitance C_{cc} and the conductance replaced by G_{RTD}^{∞} . We note that in the HF regime in RTDs with heavily doped collectors ($l \sim d$), the contribution of j_{wc} to the HF component of j_{RTD} is getting negligibly small compared to the contribution of j_{ew} , since $N_{2D} \approx \text{const}$ and the HF part of j_{wc} tends to zero (see Eq. 1) that leads to a simple equation for G_{RTD}^{∞} in such RTDs [78, 81]:

$$G_{\text{RTD}}^{\infty} = -\frac{d}{l+d} \frac{C_{\text{wc}}}{\tau}. \quad (14)$$

Accurate validity conditions for the approximation are defined in [78, 81]; to be precise, Eq. 14 defines the maximum achievable value of G_{RTD}^{∞} for given RTD geometry (d and l) and τ [78, 81]. Using Eq. 14, one can easily evaluate G_{RTD}^{∞} for further use in the RLCR model, Eq. 6 and Fig. 6c. We also note, that the transport time through the depleted collector region does not play any role in RTDs with heavily doped collector in the HF regime, since HF part of j_{wc} is negligibly small in this regime.

A detailed theory of the above effects has been developed in [78, 81, 83]. Later on, it was demonstrated experimentally that $\tau > \tau_{\text{rel}}$ in the PDC and $\tau < \tau_{\text{rel}}$ in the NDC regions of the I–V curve [82]. NDC in the regime $\omega\tau_{\text{rel}} \gg 1$ and $\omega\tau \gg 1$ up to 12 GHz in RTDs with heavily doped collector has been also demonstrated experimentally [82]. Further on, operation of RTD oscillators in the regime $\omega\tau_{\text{rel}} \gg 1$ and $\omega\tau \gg 1$ has been demonstrated experimentally up to ≈ 550 GHz [16, 44]; simultaneously, these works give direct proof that the RTDs do exhibit NDC at the operating frequencies of oscillators, i.e., also in the regime $\omega\tau_{\text{rel}} \gg 1$ and $\omega\tau \gg 1$. Our analysis of the experimental data in [9] indicates that RTD oscillators should be operating in the regime $\omega\tau_{\text{rel}} > 1$ at ≈ 2 THz. This series of works proves that RTDs inherently can operate far beyond the tunnel lifetime and relaxation-time limits.

We have also investigated RTDs, where the doping in collector is so high, that the bottom of the QW subband stays immersed not only under the emitter, but also under the collector Fermi levels [42]. The electron back injection into the QW from collector is comparable to the injection from emitter in this case. It is a quite unusual regime of operation for RTDs. We have shown experimentally that such RTDs can work at ≈ 1.5 THz in oscillators [42].

8 Non-linear RTD Dynamics: Qualitative Discussion

The above linear RTD model is sufficient for the analysis of instability (oscillation) conditions of the oscillators, to determine the frequency limitations of the oscillators, to describe the frequency dependence of the small-signal RTD impedance and the frequency roll-off of its NDC (or gain). However, the amplitude of the steady-state oscillations is limited by the RTD non-linearity, i.e., a non-linear RTD model is required for the analysis of the output power of an oscillator. One can distinguish between several regimes of operation of the RTD oscillators. If an RTD oscillator is approaching its frequency limit, then its output power tend to zero, the amplitude of

the oscillations is small and the RTD should be working in a quasi-linear regime in the vicinity of its operating point. Particularly, the time constant τ_{rel} and the above linear model for the admittance should give an adequate description of the dynamic behaviour of the RTD in this limiting regime. Nevertheless, incorporation of non-linearity into the RTD model is still required also in this quasi-linear regime, if we need to estimate the output power in addition to the oscillation conditions, limiting frequency of an RTD, etc.

However, when RTD oscillators are operated significantly below their limiting frequency, RTDs are working in essentially non-linear regime, which might have a profound impact on their dynamic characteristics and on RTD time constants. As a qualitative example, assume that the RTD is bouncing between some PDC (close to peak) and valley points of its I–V curve. During half a cycle, the RTD QW has to become fully depleted, the process might be associated with the time constant τ_c , when RTD is pushed into the valley region of the I–V curve, where the emitter-barrier current is switched off and only the collector-barrier current can deplete the QW. During the second half cycle, the RTD with an empty QW is pushed back into the PDC region of the I–V curve, where the QW has to be filled in by electrons; the relevant time constant then could be τ_{rel} in the PDC region. The time constants are not necessary long, e.g., τ_{rel} in the PDC region and τ_c could be short as compared to τ_{rel} in the NDC region, as we have discussed that in Section 7. These qualitative arguments indicate, that different time constants might be relevant to the non-linear RTD dynamics, it might be also a multi-time-constant dynamics and the time constants are not necessary longer than NDC τ_{rel} (describing the linear dynamics of an RTD). An accurate non-linear dynamic analysis of RTDs still remains to be done.

Apart from the dynamic time constants, another aspect one needs to take into account in the analysis of the non-linear RTD characteristics is the shape of the RTD I–V curve, which depends on the operating frequency. At low frequencies, when ω is small compared to $1/\tau_{\text{rel}}$ and other time constants, which might be relevant for the non-linear RTD dynamics, RTD I–V curve remains close to its dc shape. This is a quasi-static regime of operation. Analysis of RTD oscillators is straightforward in this case: the dc RTD I–V curve (which is directly measurable) could be used in the non-linear analysis of the amplitude of the oscillations. For example, approximating the non-linear RTD I–V curve by a third-order polynomial, one can derive a simple equation for the maximum power generated by an RTD: $\Delta I \Delta V^3/16$, see [88], where ΔI and ΔV are the peak-to-valley current and voltage swings, respectively.

The opposite limit, HF case, is more involved; there is no straightforward way, how to measure the HF I–V curve at THz frequencies. In the HF regime, ω is large compared to the inverse of τ_{rel} and of other non-linear time constants. Charge screening of the electric field by the charge accumulation in the QW is suppressed at so high frequencies. As a result, HF RTD I–V curve deviates significantly from the dc one. Analysis of RTD operation in such HF regime has been done in [89] and illustrated in Fig. 8. The dc and HF I–V curves are shown in Fig. 8a. In the HF regime, the electron concentration in the QW is “frozen”, it cannot follow the external bias oscillations; therefore, the tunnel currents through the emitter ($j_{\text{ew}}^{\text{HF}}$) and collector ($j_{\text{wc}}^{\text{HF}}$) barriers become quite different. $j_{\text{wc}}^{\text{HF}}$ does not show the typical RTD I–V signatures

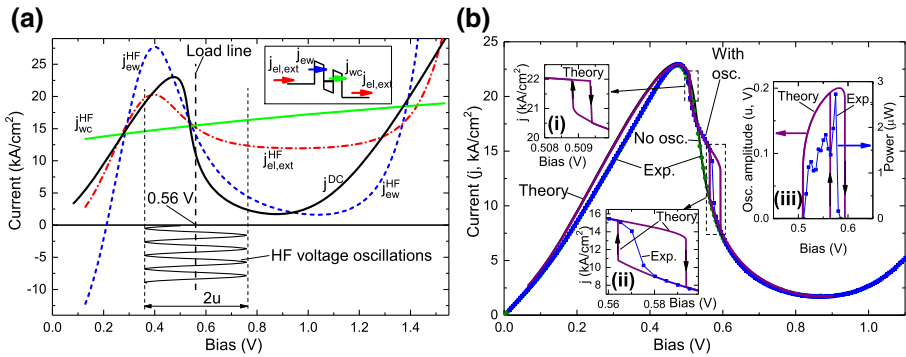


Fig. 8 Non-linear HF characteristics of RTDs. Reprinted from [89], with the permission of AIP Publishing. **a** The figure shows calculated dc (j^{DC} , black curve) and HF electron-current densities, flowing through the emitter (j^{HF}_{ew} , blue curve) and collector (j^{HF}_{wc} , green curve) barriers and through the external contacts ($j^{HF}_{el,ext}$, red curve) of an RTD. One can notice strong discrepancy between j^{DC} and $j^{HF}_{el,ext}$. **b** Relying on $j^{HF}_{el,ext}$ distortions of the dc RTD I–V curve due to oscillations are calculated (violet curve). The measured dc RTD I–V curves with (green curve) and without (blue curve) oscillations in RTD are also shown. The insets (i) and (ii) zoom the hysteresis regions at the onset of oscillations. Hystereses are not seen in the measured curves due to relatively high instability of the dc voltage source used for RTD biasing. The inset (iii) shows calculated (violet curve) and measured (blue curve) RTD oscillation amplitude and the output power, respectively

at all: j^{HF}_{wc} is determined only by the variation of the tunnel transparency of the collector barrier with bias; as a function of bias, j^{HF}_{wc} is a smooth function and it has a positive slope, i.e., its contribution to the external RTD conductance is positive. On the other hand, j^{HF}_{ew} does exhibit NDC and PDC regions, although j^{HF}_{ew} has a smaller slope than the dc I–V curve, that is due to suppression of screening, as we discussed that in Section 2. However, j^{HF}_{ew} shows a larger peak current, since QW electron concentration is “frozen” at a lower level than that at the peak of the dc I–V curve and that leads to lower back-flow current from the QW to emitter. After proper averaging of j^{HF}_{ew} and j^{HF}_{wc} , applying Shokley-Ramo theorem, see Eq. 11, one gets the external electron RTD current ($j^{HF}_{el,ext}$) shown in Fig. 8a. One can notice, that the HF I–V curve is quite different from the dc one; it has different values and bias positions for the peak and valley currents, the amplitude of the current swing is different, the HF I–V is also dependent on the RTD working point. We note, that in the close vicinity to the RTD operating point, the slope of $j^{HF}_{el,ext}$ corresponds to G^{∞}_{RTD} derived in the framework of the above linear RTD model.

Non-linear $j^{HF}_{el,ext}$ curve was further used in the analysis of the oscillation characteristics of a HF RTD oscillator. Figure 8b shows the measured and calculated (relying on the outlined HF non-linear model) RTD I–V curves with and without distortions due to oscillations (oscillations change the dc RTD current, e.g., due to quadratic non-linearity). The insets also show the oscillation onset regions, amplitude of the voltage swing and the output power of the RTD oscillator. The measurement and simulation data were in perfect agreement with each other. The example illustrates that the

HF non-linear model can describe well the characteristics of RTD oscillators. However, accurate description of the non-linear RTD characteristics in the intermediate (between quasi-static and HF) regime is much more challenging and still remains to be done.

9 Characteristics of RTD Oscillators and Their Applications

In the meanwhile, the operating frequencies of RTD oscillators have reached 1.98 THz (see an overview in Fig. 9). The output power of single-element RTD oscillators has reached 1 mW below 300 GHz, $\approx 400 \mu\text{W}$ at 550 GHz; it drops to few tens of μW at frequencies above 1 THz and to $\approx 0.3 \mu\text{W}$ close to 2 THz. RTD oscillators with a two-element RTD array could provide $\approx 600 \mu\text{W}$ at ≈ 600 GHz and $\approx 200 \mu\text{W}$ at ≈ 800 GHz [12].

The output power and operating frequencies outlined above could be already useful for certain applications. However, mW output-power level is usually considered as desirable for practical use of RTD oscillators. The level is not reached so far at frequencies near 1 THz, but based on the analysis of studied RTD oscillators, we know that there is much room for further improvement of the RTD-oscillator parameters, and mW output power at THz frequencies seems to be a realistic. Specifically, the ohmic losses in RTD resonators could be reduced by a better design of the resonators;

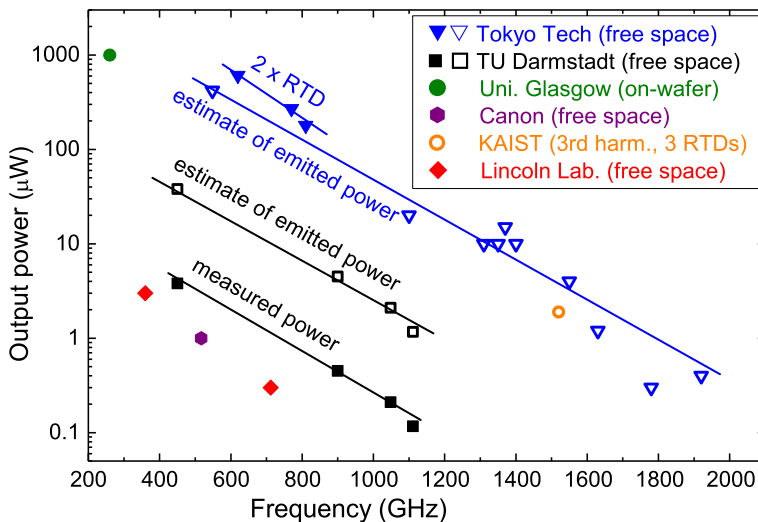


Fig. 9 Operating frequencies and output power of sub-THz and THz coherent (single-frequency) RTD oscillators. It has to be noted, that the output power is not always specified consistently in the publications. Sometimes, it is the power coupled into a detector. However, more often, the reported power refers to an estimate of the emitted oscillator power based on several assumptions correcting for the back reflections at the surface of Si lens, mismatch of the radiation pattern and the parabolic mirrors, probe and waveguide losses, etc. The data in the figure are taken from the following publications: [10, 12, 42] (Tokyo Institute of Technology), [9] (Technical University of Darmstadt), [90] (University of Glasgow), [13] (Canon), [69] (KAIST, an oscillator was combining three RTDs emitting at the third harmonic), [7] (Lincoln Laboratory)

an optimum coupling of the RTD to resonator still needs to be achieved to increase the output power, the radiation efficiency of the emitting antenna of an oscillator still needs to be improved, and RTD parameters could be also improved in the terms of an optimum width of the NDC region, suppression of the valley current, etc. Eventually, some unconventional concepts of RTD oscillators still could be invented.

For practical applications, not only the output power and operating frequencies are relevant but also other parameters of RTD oscillators are important. The typical linewidth of a free-running RTD oscillator is in the range of ~ 10 MHz (see, e.g., [44, 91]). The linewidth is too large for many foreseen applications. To overcome the problem, phase-lock-loop stabilisation of an RTD oscillator has been demonstrated recently, where the linewidth was reduced to ≈ 1 Hz [92]. Another limitation of the simplest RTD oscillators is the frequency-tuning range; it is usually around 1–2%. The frequency is changing due to the variation of the RTD capacitance, when the bias is changing within NDC region. Recently, more complex RTD oscillators have been demonstrated, where an additional varactor diode has been built into an oscillator. The tuning range of a single oscillator was ≈ 10 –20% and $\approx 40\%$ for an integrated four-element array of oscillators [41, 93, 94].

The RTD oscillators could be used for high-speed wireless data transmission. The biasing circuit of the oscillator should allow for HF modulation of the RTD for such applications. For example, 34 Gbit/s data transmission based on amplitude shift keying at 500 GHz has been demonstrated in [14]. Two-channel transmission with the data rate of 2×28 Gbit/s has been demonstrated relying on the frequency division multiplexing in two channels at 500 and 800 GHz and, alternatively, relying on polarisation division multiplexing (2 orthogonal linear polarisations) [95]. RTD oscillators were mounted on Si lens in the above experiments. An on-chip RTD oscillator with an array of (passive) circular slot antennas around an RTD oscillator designed to emit a circularly polarised beam upwards from the chip (no Si lens) has been also demonstrated in data transmission experiments with the data rate of 1 Gbit/s at ≈ 500 GHz [96]. Schottky diodes have been used as high-speed detectors in the above experiments. However, RTDs can perform also well as detectors in the wireless communication systems: 9 Gbit/s at 286 GHz have been demonstrated in [74]. Meanwhile, RTD oscillators have been used also in an imaging demonstrator [73] and in a high-resolution photonic-crystal spectrometer [75]; both are at ≈ 300 GHz. The recent development shows that RTD oscillators are getting mature enough for many practical applications.

10 Conclusions

We could see in the last years that room-temperature RTDs and RTD oscillators can operate at THz frequencies, their output power is steadily increasing and getting close to mW level at frequencies close to 1 THz, RTD oscillators could be extremely tiny and they consume low dc power. Typical examples of THz applications have been also demonstrated recently, where RTDs have been used in oscillators and detectors. On the other hand, we see that there is much room for further development of RTD oscillators: RTDs are still not optimised, the RTD resonators and antennas are still

very simplistic and not efficient, even the basic concept of the contemporary RTD oscillators might need to be reconsidered and reinvented to achieve the performance RTDs are capable of. This recent development in the field demonstrates that RTDs is a promising viable technology for real-world THz applications, and they have large potential for further development.

Funding Information Open access funding provided by TU Wien (TUW).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, N.J., 1951)
2. L. V. Iogansen, Zh. Eksp. Theor. Fiz. **47**, 270 (1964). [L. V. Iogansen, Sov. Phys.-JETP **20**, 180 (1965)]
3. R. Tsu, L. Esaki, Appl. Phys. Lett. **22**, 562 (1973)
4. L.L. Chang, L. Esaki, R. Tsu, Appl. Phys. Lett. **24**, 593 (1974)
5. M. Büttiker, R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982)
6. R. Landauer, T. Martin, Rev. Mod. Phys. **66**, 217 (1994)
7. E.R. Brown, J.R. Söderström, C.D. Parker, L.J. Mahoney, K.M. Molvar, T.C. McGill, Appl. Phys. Lett. **58**, 2291 (1991)
8. S. Suzuki, M. Asada, A. Teranishi, H. Sugiyama, H. Yokoyama, Appl. Phys. Lett. **97**, 242102 (2010)
9. M. Feiginov, C. Sydlo, O. Cojocari, P. Meissner, Appl. Phys. Lett. **99**, 233506 (2011)
10. R. Izumi, S. Suzuki, M. Asada, in: *42nd Int. Conf. on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (2017), pp. 1–2
11. B.S. Williams, Nature Photon. **1**, 517–525 (2007)
12. S. Suzuki, M. Shiraishi, H. Shibayama, M. Asada, IEEE J. Sel. Top. Quantum Electron. **19**, 8500108 (2013)
13. R. Sekiguchi, Y. Koyama, T. Ouchi, Appl. Phys. Lett. **96**, 062115 (2010)
14. N. Oshima, K. Hashimoto, S. Suzuki, M. Asada, Electron. Lett. **52**, 1897 (2016)
15. B. Ricco, M.Y. Azbel, Phys. Rev. B **29**, 1970 (1984)
16. M. Feiginov, C. Sydlo, O. Cojocari, P. Meissner, EPL **97**, 58006 (2012)
17. S. Luryi, Appl. Phys. Lett. **47**, 490 (1985)
18. H. Morkoç, J. Chen, U.K. Reddy, T. Henderson, S. Luryi, Appl. Phys. Lett. **49**, 70 (1986)
19. T. Weil, B. Vinter, Appl. Phys. Lett. **50**, 1281 (1987)
20. T.C.L.G. Sollner, W.D. Goodhue, P.E. Tannenwald, C.D. Parker, D.D. Peck, Appl. Phys. Lett. **43**, 588 (1983)
21. E.R. Brown, C.D. Parker, T.C.L.G. Sollner, Appl. Phys. Lett. **54**, 934 (1989)
22. V.J. Goldman, D.C. Tsui, J.E. Cunningham, Phys. Rev. B **35**, 9387 (1987)
23. F.W. Sheard, G.A. Toombs, Appl. Phys. Lett. **52**, 1228 (1988)
24. A. Zaslavsky, V.J. Goldman, D.C. Tsui, J.E. Cunningham, Appl. Phys. Lett. **53**, 1408 (1988)
25. L. Eaves, M. Leadbeater, D. Hayes, E. Alves, F. Sheard, G. Toombs, P. Simmonds, M. Skolnick, M. Henini, O. Hughes, Solid-State Electron. **32**, 1101 (1989)
26. A.D. Martin, M.L.F. Lerch, P.E. Simmonds, L. Eaves, Appl. Phys. Lett. **64**, 1248 (1994)
27. V.J. Goldman, D.C. Tsui, J.E. Cunningham, Phys. Rev. B **36**, 7635 (1987)
28. T.C.L.G. Sollner, P.E. Tannenwald, D.D. Peck, W.D. Goodhue, Appl. Phys. Lett. **45**, 1319 (1984)
29. T.C.L.G. Sollner, Phys. Rev. Lett. **59**, 1622 (1987)
30. J.F. Young, B.M. Wood, H.C. Liu, M. Buchanan, D. Landheer, A.J. Spring Thorpe, P. Mandeville, Appl. Phys. Lett. **52**, 1398 (1988)
31. E.R. Brown, W.D. Goodhue, T.C.L.G. Sollner, J. Appl. Phys. **64**, 1519 (1988)
32. V.J. Goldman, D.C. Tsui, J.E. Cunningham, Phys. Rev. Lett. **58**, 1256 (1987)

33. V.J. Goldman, D.C. Tsui, J.E. Cunningham, *Phys. Rev. Lett.* **59**, 1623 (1987)
34. E.R. Brown, T.C.L.G. Sollner, W.D. Goodhue, C.D. Parker, *Appl. Phys. Lett.* **50**, 83 (1987)
35. E.R. Brown, T.C.L.G. Sollner, C.D. Parker, W.D. Goodhue, C.L. Chen, *Appl. Phys. Lett.* **55**, 1777 (1989)
36. M. Reddy, S.C. Martin, A.C. Molnar, R.E. Muller, R.P. Smith, P.H. Siegel, M.J. Mondry, M.J.W. Rodwell, H. Kroemer, S.J. Allen, *IEEE Electron Device Lett.* **18**, 218 (1997)
37. J.P. Mattia, A.L. McWhorter, R.J. Aggarwal, F. Rana, E.R. Brown, P. Maki, *J. Appl. Phys.* **84**, 1140 (1998)
38. T.A. Growden, D.F. Storm, W. Zhang, E.R. Brown, D.J. Meyer, P. Fakhimi, P.R. Berger, *Appl. Phys. Lett.* **109**, 083504 (2016)
39. T.A. Growden, W. Zhang, E.R. Brown, D.F. Storm, K. Hansen, P. Fakhimi, D.J. Meyer, P.R. Berger, *Applied Physics Letters* **112**, 033508 (2018)
40. J. Encomendero, R. Yan, A. Verma, S.M. Islam, V. Protasenko, S. Rouvimov, P. Fay, D. Jena, H.G. Xing, *Appl. Phys. Lett.* **112**, 103101 (2018)
41. M. Asada, S. Suzuki, *J. Infrared Millimeter Terahertz Waves* **37**, 1185 (2016)
42. M. Feiginov, H. Kanaya, S. Suzuki, M. Asada, *Appl. Phys. Lett.* **104**, 243509 (2014)
43. J.H. Smet, T.P.E. Broekaert, C.G. Fonstad, *J. Appl. Phys.* **71**, 2475 (1992)
44. M. Feiginov, C. Sydlo, O. Cojocari, P. Meissner, *EPL* **94**, 48007 (2011)
45. J.R. Söderström, E.R. Brown, C.D. Parker, L.J. Mahoney, J.Y. Yao, T.G. Andersson, T.C. McGill, *Appl. Phys. Lett.* **58**, 275 (1991)
46. J.R. Söderström, D.H. Chow, T.C. McGill, *Appl. Phys. Lett.* **55**, 1094 (1989)
47. J. Schulman, D. Chow, T. Hasenberg, *Solid-State Electron.* **37**, 981 (1994)
48. K. Arzi, A. Rennings, D. Erni, N. Weimann, W. Prost, S. Suzuki, M. Asada, in: *First Int. Workshop on Mobile Terahertz Systems (IWMTS)* (2018), pp. 1–4
49. D.H. Chow, T.C. McGill, I.K. Sou, J.P. Faurie, C.W. Nieh, *Appl. Phys. Lett.* **52**, 54 (1988)
50. H.C. Liu, *Appl. Phys. Lett.* **51**, 1019 (1987)
51. A.R. Bonnefoi, T.C. McGill, R.D. Burnham, *Phys. Rev. B* **37**, 8754 (1988)
52. M.N. Feiginov, *Appl. Phys. Lett.* **81**, 930 (2002)
53. M. Feiginov, *Physica E: Low-dim. Syst. Nanostructures* **17**, 643 (2003)
54. M.N. Feiginov, I.N. Kotel'nikov, *Appl. Phys. Lett.* **91**, 083510 (2007)
55. L. Britnell, R.V. Gorbachev, A.K. Geim, L.A. Ponomarenko, A. Mishchenko, M.T. Greenaway, T.M. Fromhold, K.S. Novoselov, L. Eaves, *Nature Commun.* **4**, 1794 (2013)
56. J. Gaskell, L. Eaves, K.S. Novoselov, A. Mishchenko, A.K. Geim, T.M. Fromhold, M.T. Greenaway, *Appl. Phys. Lett.* **107**, 103105 (2015)
57. D.D. Coon, H.C. Liu, *Appl. Phys. Lett.* **49**, 94 (1986)
58. J.S. Scott, J.P. Kaminski, M. Wanke, S.J. Allen, D.H. Chow, M. Lui, T.Y. Liu, *Appl. Phys. Lett.* **64**, 1995 (1994)
59. M. Feiginov, *Appl. Phys. Lett.* **107**, 123504 (2015)
60. M.E. Hines, *Bell Syst. Techn. J.* **39**, 477 (1960)
61. S. Suzuki, A. Teranishi, K. Hinata, M. Asada, H. Sugiyama, H. Yokoyama, *Appl. Phys. Express* **2**, 054501 (2009)
62. K. Urayama, S. Aoki, S. Suzuki, M. Asada, H. Sugiyama, H. Yokoyama, *Appl. Phys. Express* **2**, 044501 (2009)
63. M. Asada, K. Osada, W. Saitoh, *Solid-State Electronics* **42**, 1543 (1998)
64. K. Kasagi, S. Suzuki, M. Asada, in: *40th Int Conf Infrared, Millimeter, and Terahertz waves (IRMMW-THz)* (2015), pp. 1–2
65. K. Kasagi, S. Fukuma, S. Suzuki, M. Asada, in: *41st Int. Conf. Infrared, Millimeter, and Terahertz waves (IRMMW-THz)* (2016), pp. 1–2
66. S. Suzuki, M. Asada, in: *11th European Conf. Antennas and Propagation (EUCAP)* (2017), pp. 1615–1617
67. J. Faist, G. Scalari, *Electron. Lett.* **46**, s46 (2010)
68. N. Orihashi, S. Suzuki, M. Asada, *Appl. Phys. Lett.* **87**, 233501 (2005)
69. J. Lee, M. Kim, K. Yang, *IEEE Trans. THz Science and Technol.* **6**, 336 (2016)
70. S. Boppel, A. Lisauskas, M. Mundt, D. Seliuta, L. Minkevicius, I. Kasalynas, G. Valusis, M. Mitterdorff, S. Winnerl, V. Krozer, H.G. Roskos, *IEEE Trans. Microwave Theory and Techn.* **60**, 3834 (2012)
71. T. Shiode, T. Mukai, M. Kawamura, T. Nagatsuma, in: *Asia-Pacific Microwave Conf.* (2011), pp. 1122–1125

72. G. Keller, A. Tchegho, B. Münstermann, W. Prost, F. Tegude, M. Suhara, in: *Int. Conf. on Indium Phosphide and Related Materials (IPRM)* (2013), pp. 1–2
73. T. Miyamoto, A. Yamaguchi, T. Mukai, *Jap. J. Appl. Phys.* **55**, 032201 (2016)
74. S. Diebold, K. Nishio, Y. Nishida, J.. Kim, K. Tsuruda, T. Mukai, M. Fujita, T. Nagatsuma, *Electron. Lett.* **52**, 1999 (2016)
75. K. Okamoto, K. Tsuruda, S. Diebold, S. Hisatake, M. Fujita, T. Nagatsuma, *J. Infrared, Millimeter, and Terahertz Waves* **38**, 1085 (2017)
76. J. Genoe, C. Van Hoof, W. Van Roy, J. Smet, K. Fobelets, R. Mertens, G. Borghs, *IEEE Trans. Electron Dev.* **38**, 2006 (1991)
77. H.P. Joosten, H.J.M.F. Noteborn, K. Kaski, D. Lenstra, *J. Appl. Phys.* **70**, 3141 (1991)
78. M.N. Feiginov, *Nanotechnology* **11**, 359 (2000)
79. A.P. Jauho, N.S. Wingreen, Y. Meir, *Phys. Rev. B* **50**, 5528 (1994)
80. W.R. Liou, P. Roblin, *IEEE Trans. Electron Dev.* **41**, 1098 (1994)
81. M.N. Feiginov, *Appl. Phys. Lett.* **78**, 3301 (2001)
82. M.N. Feiginov, D. Roy Chowdhury, *Appl. Phys. Lett.* **91**, 203501 (2007)
83. M.N. Feiginov, *Appl. Phys. Lett.* **76**, 2904 (2000)
84. W. Shockley, *J. Appl. Phys.* **9**, 635 (1938)
85. S. Ramo, *Proc. IRE* **27**, 584 (1939)
86. T. Wei, S. Stapleton, E. Berolo, *J. Appl. Phys.* **73**, 829 (1993)
87. J. Jo, H.S. Li, Y.W. Chen, K.L. Wang, *Appl. Phys. Lett.* **64**, 2276 (1994)
88. M. Asada, S. Suzuki, N. Kishimoto, *Jap. J. Appl. Phys.* **47**, 4375 (2008)
89. M. Feiginov, C. Sydlo, O. Cojocari, P. Meissner, *Appl. Phys. Lett.* **99**, 133501 (2011)
90. A. Al-Khalidi, K. Alharbi, J. Wang, E. Wasige, in: *19th Int. Radar Symp. (IRS)* (2018), pp. 1–6
91. K. Karashima, R. Yokoyama, M. Shiraishi, S. Suzuki, S. Aoki, M. Asada, *Jap. J. Appl. Phys.* **49**, 020208 (2010)
92. K. Ogino, S. Suzuki, M. Asada, in: *42nd Int. Conf. Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (2017), pp. 1–2
93. S. Kitagawa, S. Suzuki, M. Asada, *IEEE Electron Dev. Lett.* **35**, 1215 (2014)
94. S. Kitagawa, S. Suzuki, M. Asada, *Electron. Lett.* **52**, 479 (2016)
95. N. Oshima, K. Hashimoto, S. Suzuki, M. Asada, *IEEE Trans. Terahertz Sci. Technol.* **7**, 593 (2017)
96. D. Horikawa, Y. Chen, T. Koike, S. Suzuki, M. Asada, *Semicond. Sci. Technol.* **33**, 114005 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.