

Perturbation theory for generalized van der Pol-type resonant-tunneling-diode oscillators

Michael Feiginov^{✉*} and Petr Ourednik[✉]

Department of Electrical Engineering and Information Technology, *TU Wien*, 1040 Vienna, Austria



(Received 15 January 2026; accepted 4 May 2026; published 28 May 2026)

We generalize the van der Pol oscillator model to account for a realistic nonlinearity of the resonant-tunneling-diode (RTD) I - V curve. In the case of weak nonlinearity, perturbation theory is applied, and analytical expressions are derived for first-order corrections to the oscillation frequency, quadrature component (phase shift), and higher harmonics. Applicability requirements for perturbation theory are discussed. These requirements are always satisfied at sufficiently high frequencies, but they can fail already at frequencies as low as a few hundred GHz for RTDs with thicker barriers and at slightly above 500 GHz for thin-barrier RTDs, where oscillators transition to a strongly nonlinear regime. Calculation results for the above corrections for several RTD wafers, with typical parameters for sub-THz and THz oscillators, are presented. In particular, our calculations indicate that first-order correction terms play an essential role in the mechanism of variation of the RTD-oscillator frequency with bias. Taking into account the realistic nonlinearity of the RTD I - V curve leads to significant quantitative and even qualitative differences in the behavior of RTD oscillators, compared to the conventional van der Pol oscillator model. The presented generalized model lays the groundwork for a more accurate analysis of nonlinear phenomena in RTD oscillators, it can serve as a reference for numerical simulations, it is also essential for qualitative and semiquantitative analysis of RTD oscillators.

DOI: [10.1103/hcqb-44hy](https://doi.org/10.1103/hcqb-44hy)

I. INTRODUCTION

Currently, the THz frequency range is still not in use because of the lack of THz sources suitable for practical applications. That drives efforts to develop compact, room temperature, sufficiently powerful, and simple THz sources [1,2]. One of the most promising contenders for practical use is resonant-tunneling-diode (RTD) oscillators, which have seen renewed interest over the past two decades [3,4]. RTD oscillators show promise in satisfying the above requirements in the frequency range of up to 2 THz [3–8].

Perhaps, an even more scientifically useful aspect of the THz frequency range is that it offers opportunities and incentives to gain a deeper understanding of the operational principles and limitations of different devices. Conventional electronic or photonic devices have to operate at and sometimes beyond their limits at THz frequencies, while unconventional devices are also intensively explored

[1,2]. This scientific aspect plays a pivotal role in the research on RTDs and RTD oscillators [3,4].

Any oscillator and RTD oscillator in particular is an essentially nonlinear system: the oscillation amplitude is determined by the nonlinearity of the system. Typically, the amplification (gain) of an active element in the oscillator decreases with the increase of the amplitude, and the oscillation amplitude assumes a steady-state value at some point. More subtle oscillator parameters are also determined or affected by the system's nonlinearity, e.g., phase of the oscillator [9,10], its frequency [7], generation of higher harmonics [11], etc. The mechanisms of the oscillator control, specifically, injection locking [9,12,13], rely on the system nonlinearity as well. In this paper, we aim to refine and generalize the nonlinear model for an RTD oscillator.

So far, the nonlinearity in RTD oscillators has been analyzed by averaging of the large-amplitude RTD conductance [14], by applying a third-order polynomial approximation to the RTD I - V curve (van der Pol oscillator) [12], or by numerical methods [15]. There are several issues with these techniques. First, the applicability range and conditions for some of the models are not defined. Second, approximation of the RTD I - V curve by a third-order polynomial can lead to quantitatively and even qualitatively different results for the oscillator characteristics [14,16]. Third, for qualitative and semiquantitative

*Contact author: michael.feiginov@tuwien.ac.at

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

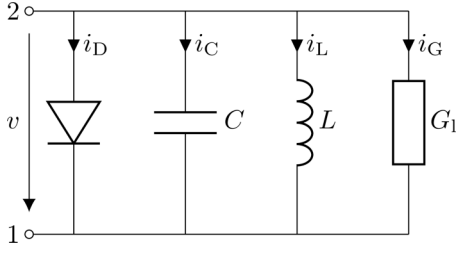


FIG. 1. Equivalent circuit of an RTD oscillator.

analysis, it is desirable to have an accurate but relatively simple oscillator model.

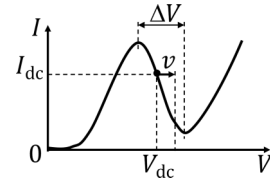
Here, we present a generalized van der Pol oscillator model, where the real nonlinearity of the RTD I - V curve is taken into account. The perturbation series for the model is constructed up to the first order. There is a large number of phenomena, which go beyond the presented model—nonlinearity of the RTD capacitance, variation of the RTD admittance with frequency, accurate account of nonlinear RTD dynamics [17,18], more accurate model for the resonator, rather than a simple LC circuit, etc.—nevertheless, the presented model describes and predicts a number of phenomena that go beyond the van der Pol oscillator. The presented model would also be helpful as a reference for more complicated models and for numerical simulations.

The paper is organized as follows. First, we outline the model in Sec. II. Then, we describe the analytical calculations and perturbation theory up to the first order in Sec. III. In Sec. IV, we summarize the calculation results to facilitate their practical use. In Sec. V, we show consistency of the calculation results with the van der Pol oscillator model, when the RTD I - V curve is approximated by a third-order polynomial. In Sec. VI, we apply the perturbation theory to two typical RTD wafers used at sub-THz and THz frequencies and illustrate the behavior of the oscillators in the framework of the presented model. Finally, we summarize our results in Sec. VII.

II. DEFINITION OF THE MODEL

The objective of this paper is to analyze the impact of the nonlinearity of the RTD I - V curve on the oscillation frequency and other characteristics of an RTD oscillator.

An RTD oscillator is represented by an equivalent circuit shown in Fig. 1 (notations are similar to Ref. [12]), where the capacitance C is a sum of the RTD and resonator capacitances; the conductance G_l represents the ohmic and radiation losses of the resonator and L is its inductance; i_D and v are the RTD nonlinear ac current and voltage, respectively; a positive sign of v corresponds in Fig. 1 to a higher potential at point 2 than at point 1.

FIG. 2. dc RTD I - V curve, quiescent point, and ac voltage (v).

The currents of different elements of the oscillator circuit are defined as follows:

$$i_C = C \partial_t v, \quad (1)$$

$$i_L = \frac{1}{L} \int v dt, \quad (2)$$

$$i_G = G_l v, \quad (3)$$

$$i_D(v) = I_{dc}(V_{dc} + v) - I_{dc}(V_{dc}), \quad (4)$$

where $I_{dc}(V_{dc})$ is the dc RTD I - V curve, see Fig. 2. In Eq. (4), we assume that the RTD operates in a quasistatic regime, i.e., its characteristics are independent of frequency. The RTD capacitance, which is included in C , is assumed to be a constant. That implies that the nonlinearity of the RTD capacitance is ignored, i.e., the only nonlinearity taken into account in Fig. 1 is that of the quasistatic RTD I - V curve (4).

From Kirchhoff's law, the sum of all currents in Fig. 1 is zero, i.e.,

$$i_D(v) + C \partial_t v + \frac{1}{L} \int v dt + G_l v = 0. \quad (5)$$

Taking the time derivative of the above equation, we get

$$LC \partial_t^2 v + v + L[G_l \partial_t v + \partial_t i_D(v)] = 0 \quad (6)$$

or

$$LC \partial_t^2 v + v + L[G_l + \partial_v i_D(v)] \partial_t v = 0. \quad (7)$$

Introducing a nonlinear RTD conductance $G_D(v)$ as

$$G_D(v) = \partial_v i_D(v), \quad (8)$$

we rewrite Eq. (7) as

$$LC \partial_t^2 v + v = -L[G_l + G_D(v)] \partial_t v. \quad (9)$$

Our aim now is to recast this equation in a unitless form, common in the mathematical literature on nonlinear equations (e.g., van der Pol equation).

As a first step, we split $i_D(v)$ and $G_D(v)$ into their linear and nonlinear (denoted by the preceding symbol δ) parts:

$$i_D(v) = G_D(0)v + \delta i_D(v), \quad (10)$$

$$G_D(v) = G_D(0) + \delta G_D(v), \quad (11)$$

where $G_D(0) = G_D(v)|_{v \rightarrow 0}$ is the linear RTD conductance at a given dc bias point V_{dc} and $\delta G_D(0) = 0$. Then, we can rewrite Eq. (9) as

$$LC\partial_t^2 v + v = -L[G_D(0) + G_l + \delta G_D(v)]\partial_t v. \quad (12)$$

In the following, we will need to evaluate the linear and nonlinear terms on the right-hand side of Eq. (12). For this purpose, we define the maximum value ($-G_0$) of the RTD negative differential conductance (NDC)

$$G_0 = \max(-G_D(0)). \quad (13)$$

If the amplitude of the RTD oscillations is large, then the RTD will be sweeping twice through the whole NDC region each oscillation period and G_0 gives a good estimate for the variation scale of the expression in the square brackets on the right-hand side of Eq. (12). The condition for large-amplitude oscillations is $G_l \ll G_0$, see Refs. [14,16]. In the small-amplitude regime, when $G_l \approx G_0$, the expression in the square brackets can vary much less than G_0 . Nevertheless, G_0 gives a good estimate of how large the variation scale of the square-bracket expression can in principle be. Subsequently, we normalize the square brackets in Eq. (12) on G_0 and introduce the parameter ξ :

$$LC\partial_t^2 v + v = LG_0 \left[\xi - \frac{\delta G_D(v)}{G_0} \right] \partial_t v, \quad (14)$$

$$\xi = -\frac{G_D(0) + G_l}{G_0}. \quad (15)$$

ξ is a constant for a given bias point and G_l . $\xi \approx 1$, if RTD is biased at the maximum of its NDC and $G_l \ll G_0$, otherwise $\xi \lesssim 1$.

As a next step, we introduce a dimensionless voltage. A natural characteristic scale for the RTD ac voltage is the voltage separation (ΔV) between the peak and valley points of the RTD I - V curve, see Fig. 2. Therefore, we define a unitless ac voltage ($\tilde{v}$) as

$$\tilde{v} = \frac{v}{\Delta V}, \quad (16)$$

and rewrite Eq. (14) as

$$LC\partial_t^2 \tilde{v} + \tilde{v} = LG_0 \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \partial_t \tilde{v}. \quad (17)$$

Now, we define the eigenfrequency of an ideal linear oscillator, if the attenuation (gain) and nonlinearity are

ignored:

$$\omega_0 = 2\pi f_0 = \frac{1}{\sqrt{LC}}. \quad (18)$$

Substituting that into Eq. (17) and also expressing L as $1/C\omega_0^2$, we can rewrite Eq. (17) as

$$\frac{\partial_t^2 \tilde{v}}{\omega_0^2} + \tilde{v} = \frac{G_0}{\omega_0 C} \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \frac{\partial_t \tilde{v}}{\omega_0}. \quad (19)$$

Thereafter, we define a perturbation parameter

$$\epsilon = \frac{G_0}{C\omega_0}, \quad (20)$$

and rewrite Eq. (19) as

$$\frac{\partial_t^2 \tilde{v}}{\omega_0^2} + \tilde{v} = \epsilon \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \frac{\partial_t \tilde{v}}{\omega_0}. \quad (21)$$

If $\epsilon = 0$, then the right-hand side of Eq. (21) disappears and we get a linear harmonic equation with the oscillation frequency ω_0 . Our aim is to construct a perturbation theory for the weakly nonlinear solutions of Eq. (21), considering that ϵ determines the smallness of the subsequent terms in a perturbation series. We will treat ϵ as a small parameter ($\epsilon \ll 1$) in Sec. III. Indeed, $\epsilon \rightarrow 0$, when ω_0 grows. However, we will relax this condition later on, and show that ϵ does not necessary need to be small. The applicability conditions for the perturbation theory are discussed in Sec. IV.

Equation (21) has the advantage, that it has a structure of a generalized van der Pol equation. The perturbation parameter and the nonlinear term are clearly defined in the equation. However, it is convenient to rewrite the equation in a different form, which will eventually simplify the following calculations. In terms of δi_D , Eq. (21) could be written down in the form

$$\frac{\partial_t^2 \tilde{v}}{\omega_0^2} + \tilde{v} = \frac{\epsilon}{\omega_0} \left[\xi \partial_t \tilde{v} - \frac{\partial_t \delta i_D(\tilde{v})}{\Delta V G_0} \right], \quad (22)$$

as it follows from Eqs. (6) and (10).

III. PERTURBATION THEORY

For the construction of the perturbation theory, we use the Poincaré-Lindstedt method, see Refs. [19,20]. We are looking for a periodic solution with the period (T) and the initial conditions for $\tilde{v}(t)$ and $\partial_t \tilde{v}(t)$ dependent on the parameter ϵ :

$$T = T(\epsilon), \quad \tilde{v}(0) = a(\epsilon), \quad \partial_t \tilde{v}(0) = 0. \quad (23)$$

The last equation in Eq. (23) does not impose a restriction, since we are looking for a solution for an autonomous

equation (not explicitly dependent on t) with the trajectories in the phase plane $(\tilde{v}, \partial_t \tilde{v})$ encircling the zero point; i.e., a trajectory at some point in time must cross the axis $\partial_t \tilde{v} = 0$ and we take this point as the initial time ($t = 0$).

A. Fixed-period representation

It is convenient to deal with an equation with a fixed (independent of ϵ) period. To this end, we define a fundamental frequency ω of the solution as

$$\omega = \frac{2\pi}{T} \quad (24)$$

and a unitless time τ as

$$\tau = t\omega. \quad (25)$$

Then, the period is fixed and equal to 2π . The time derivative will become

$$\partial_t = \omega \partial_\tau. \quad (26)$$

Equation (21) can be rewritten as

$$\frac{\omega^2}{\omega_0^2} \partial_\tau^2 \tilde{v} + \tilde{v} = \epsilon \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \frac{\omega}{\omega_0} \partial_\tau \tilde{v}. \quad (27)$$

Introducing further the unitless frequency

$$\tilde{\omega} = \frac{\omega}{\omega_0}, \quad (28)$$

we rewrite Eq. (27) as

$$\tilde{\omega}^2 \partial_\tau^2 \tilde{v} + \tilde{v} = \epsilon \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \tilde{\omega} \partial_\tau \tilde{v}. \quad (29)$$

Since we expect $\omega \approx \omega_0$, then $\tilde{\omega} \approx 1$, and we can write it in the form

$$\tilde{\omega} = 1 + \epsilon\kappa, \quad (30)$$

where $|\epsilon\kappa| \ll 1$ and κ is a parameter describing the frequency deviation.

B. Perturbation series

Now, we collect all small terms [$O(\epsilon)$ and smaller] on the right-hand side of Eq. (29) and rewrite the equation as

$$\partial_\tau^2 \tilde{v} + \tilde{v} = \frac{\epsilon}{\tilde{\omega}} \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \partial_\tau \tilde{v} + \frac{\tilde{\omega}^2 - 1}{\tilde{\omega}^2} \tilde{v}. \quad (31)$$

Thereafter, we introduce the function g :

$$\begin{aligned} g(\tilde{v}, \partial_\tau \tilde{v}, \epsilon, \kappa) \\ = \frac{1}{1 + \epsilon\kappa} \left[\xi - \frac{\delta G_D(\tilde{v})}{G_0} \right] \partial_\tau \tilde{v} + \kappa \frac{2 + \epsilon\kappa}{(1 + \epsilon\kappa)^2} \tilde{v} \end{aligned} \quad (32)$$

and rewrite Eq. (31) as

$$\partial_\tau^2 \tilde{v} + \tilde{v} = \epsilon g(\tilde{v}, \partial_\tau \tilde{v}, \epsilon, \kappa). \quad (33)$$

Equation (33) should be solved for $\tilde{v}(\tau)$ with the initial conditions (23). We also need to find such parameters a and κ that the solution is 2π periodic.

The solution of Eq. (33) could be written in the form

$$\tilde{v}(\tau) = a \cos(\tau) + \epsilon \int_0^\tau g(\tilde{v}, \partial_s \tilde{v}, \epsilon, \kappa) \sin(\tau - s) ds, \quad (34)$$

where the first term on the right-hand side is a homogeneous solution adjusted for the last initial condition in Eq. (23). The last term in Eq. (34) is a particular solution [that could be verified directly by substitution into Eq. (33)].

If $\tilde{v}$ is a periodic function of τ , then $g(\tau)$ is also periodic. However, the integral appearing in the last term of Eq. (34) is not necessarily periodic, although $g(\tau)$ itself is periodic. Therefore, we need to impose explicitly the periodicity condition:

$$\int_\tau^{\tau+2\pi} g(\tilde{v}, \partial_s \tilde{v}, \epsilon, \kappa) \sin(\tau - s) ds = 0. \quad (35)$$

This condition could be rewritten in an equivalent form as two equations

$$\int_0^{2\pi} g(\tilde{v}, \partial_s \tilde{v}, \epsilon, \kappa) \sin(s) ds = 0, \quad (36)$$

$$\int_0^{2\pi} g(\tilde{v}, \partial_s \tilde{v}, \epsilon, \kappa) \cos(s) ds = 0, \quad (37)$$

which we use to fix the values of a and κ . Essentially, Eqs. (36) and (37) require that the fundamental (first) harmonics of $g(\tau)$ are zero. That eliminates the secular terms at the eigenfrequency of the left-hand side of Eq. (33); if

present, the terms would lead to linearly growing in time (not-periodic) solutions of Eq. (33).

We look for a solution of Eq. (34) in the form of a perturbation series:

$$\tilde{v}(\tau) = \sum_{n=0}^{\infty} \epsilon^n u_n(\tau), \quad (38)$$

$$a = \sum_{n=0}^{\infty} \epsilon^n a_n, \quad (39)$$

$$\kappa = \sum_{n=0}^{\infty} \epsilon^n \kappa_n. \quad (40)$$

Substituting Eqs. (38)–(40) into Eq. (32), we can group different terms together and write down also g as a series

$$g = g_0 + \epsilon g_1 + \epsilon^2 g_2 + O(\epsilon^3). \quad (41)$$

The perturbation series is constructed in the following way. One starts with Eq. (33). In the zero-order approximation, it is a homogeneous equation. Its solution contains free constants, which are adjusted by imposing the requirement, that $\epsilon g(u_0)$ [the right-hand side of Eq. (33)] does not contain fundamental-harmonic terms. This requirement could be applied to $g(u_0)$ either directly by inspection or by applying the conditions (36) and (37). Essentially, this requirement implies that the fundamental harmonics are zero in the next-order right-hand side of Eq. (33), since ϵ stands in front of g there. In the calculation of the next orders of the perturbation series, we repeat the procedure, keeping the proper-order terms in the right-hand side of Eq. (33).

Now, we rewrite Eq. (32) in the terms of $\delta i_D(v)$

$$\begin{aligned} g(\tilde{v}, \partial_\tau \tilde{v}, \epsilon, \kappa) \\ = \frac{1}{1 + \epsilon \kappa} \left[\xi \partial_\tau \tilde{v} - \frac{1}{\Delta V} \frac{\partial_\tau \delta i_D(\tilde{v})}{G_0} \right] + \kappa \frac{2 + \epsilon \kappa}{(1 + \epsilon \kappa)^2} \tilde{v}. \end{aligned} \quad (42)$$

Equation (42) has the advantage that it does not contain a product of two functions, like $\delta G_D(\tilde{v}) \partial_\tau \tilde{v}$ in Eq. (32), which would lead to a product of multiple terms, when we use a Fourier-series representation for the functions.

C. Fourier-series representations

For the following calculations, we will need the Fourier-series representation of the RTD current and conductance, if a harmonic bias oscillation $a_0 \cos(\tau)$ with the zero-order amplitude a_0 is applied to the RTD. The RTD current and conductance are periodic functions of τ with a period of

2π . We define the representations as

$$\delta i_D(a_0 \cos(\tau)) = \sum_{n=0}^{\infty} \delta i_D^{(0,n)} \cos(n\tau), \quad (43)$$

and

$$\delta G_D(a_0 \cos(\tau)) = \sum_{n=0}^{\infty} \delta G_D^{(0,n)} \cos(n\tau), \quad (44)$$

where

$$\delta i_D^{(0,n)} = \frac{2}{\pi} \int_0^\pi \delta i_D(a_0 \cos(\tau)) \cos(n\tau) d\tau \quad (45)$$

and

$$\delta G_D^{(0,n)} = \frac{2}{\pi} \int_0^\pi \delta G_D(a_0 \cos(\tau)) \cos(n\tau) d\tau \quad (46)$$

or $n \neq 0$, and

$$\delta i_D^{(0,0)} = \frac{1}{\pi} \int_0^\pi \delta i_D(a_0 \cos(\tau)) d\tau \quad (47)$$

and

$$\delta G_D^{(0,0)} = \frac{1}{\pi} \int_0^\pi \delta G_D(a_0 \cos(\tau)) d\tau \quad (48)$$

for $n = 0$. Due to the symmetry of the functions $\delta i_D(a_0 \cos(\tau))$ and $\delta G_D(a_0 \cos(\tau))$, the series (43) and (44) contain only cos harmonics.

Furthermore, integrating Eq. (45) by parts and using Eq. (8), we derive the following recurrent relations for the Fourier coefficients. For $n \geq 2$,

$$\delta G_D^{(0,n-1)} - \delta G_D^{(0,n+1)} = \frac{2n}{\Delta V a_0} \delta i_D^{(0,n)}, \quad (49)$$

and for $n = 1$,

$$\delta G_D^{(0,0)} - \frac{1}{2} \delta G_D^{(0,2)} = \frac{1}{\Delta V a_0} \delta i_D^{(0,1)}. \quad (50)$$

D. Zero-order approximation

In the zero-order approximation, Eq. (33) is

$$\partial_\tau^2 u_0 + u_0 = 0, \quad (51)$$

and its solution, satisfying the initial conditions (23), is

$$u_0 = a_0 \cos(\tau). \quad (52)$$

As a next step, we need to analyze g_0 . Collecting the zero-order terms in Eq. (42), we write them down as

$$g_0 = \xi \partial_\tau u_0 - \frac{\partial_\tau \delta i_D(u_0)}{\Delta V G_0} + 2\kappa_0 u_0. \quad (53)$$

Substituting Eqs. (52) and (43) into Eq. (53), we get

$$g_0 = -\xi a_0 \sin(\tau) + \frac{1}{\Delta V G_0} \sum_{n=1}^{\infty} n \delta i_D^{(0,n)} \sin(n\tau) + 2\kappa_0 a_0 \cos(\tau). \quad (54)$$

The first-harmonic terms in this equation are

$$g_0|_{\text{first harm.}} = -a_0 \left[\xi - \frac{\delta i_D^{(0,1)}}{a_0 \Delta V G_0} \right] \sin(\tau) + 2\kappa_0 a_0 \cos(\tau). \quad (55)$$

Now, we need to impose the condition that the cos and sin first harmonics are zero, and that leads to the equations for a_0 and κ_0 :

$$\kappa_0 = 0, \quad (56)$$

$$\xi = \frac{\delta i_D^{(0,1)}}{a_0 \Delta V G_0}. \quad (57)$$

Equation (56) means that the oscillation frequency is equal to ω_0 in the zero order of the perturbation theory. Using Eq. (15), Eq. (57) could be rewritten in the form

$$G_l = -G_D(0) - \frac{\delta i_D^{(0,1)}}{a_0 \Delta V}. \quad (58)$$

The solution to this equation defines the amplitude a_0 , since the Fourier coefficients, see Eq. (45), depend on a_0 . Equation (58) is equivalent to Eqs. (7) and (19) in Ref. [14]:

$$G_l = -\frac{i_D^{(0,1)}}{a_0 \Delta V}, \quad (59)$$

where the first-harmonic RTD current is defined as

$$i_D^{(0,1)} = \frac{2}{\pi} \int_0^\pi i_D(a_0 \cos(\tau)) \cos(\tau) d\tau. \quad (60)$$

To complete the calculations, we write down g_0 . After elimination of the first-order harmonics, Eq. (54) takes the

form

$$g_0 = \frac{1}{\Delta V G_0} \sum_{n=2}^{\infty} n \delta i_D^{(0,n)} \sin(n\tau). \quad (61)$$

E. First-order approximation

In the first-order approximation, we write Eq. (33) as

$$\partial_\tau^2 u_1 + u_1 = g_0(\tau), \quad (62)$$

where g_0 is given by Eq. (61). The solution of this equation is

$$u_1(\tau) = a_1 \cos(\tau) + b_1 \sin(\tau) - \frac{1}{\Delta V G_0} \sum_{n=2}^{\infty} \frac{n}{n^2 - 1} \delta i_D^{(0,n)} \sin(n\tau). \quad (63)$$

The coefficient b_1 is determined by the initial condition ($\partial_\tau u_1(0) = 0$) on the time derivative (23):

$$b_1 = \frac{1}{\Delta V G_0} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} \delta i_D^{(0,n)}. \quad (64)$$

As outlined above, the parameters a_1 and κ_1 are determined by the requirement, that g_1 does not contain the first-order harmonics. As it follows from Eq. (42), g_1 is

$$g_1 = \xi \partial_\tau u_1 - \frac{\partial_\tau [\delta G_D(u_0) u_1]}{G_0} + 2\kappa_1 u_0. \quad (65)$$

As a next step, we need to collect all the first-harmonic terms in Eq. (65) and set them to zero. We start with the second term. From Eqs. (44) and (63), it follows that the first-harmonic terms of $\delta G_D(u_0) u_1$ are

$$\begin{aligned} & \delta G_D(u_0) u_1|_{\text{first harm.}} \\ &= a_1 \cos(\tau) \left[\delta G_D^{(0,0)} + \frac{1}{2} \delta G_D^{(0,2)} \right] \\ &+ b_1 \sin(\tau) \left[\delta G_D^{(0,0)} - \frac{1}{2} \delta G_D^{(0,2)} \right] \\ &- \frac{1}{2} \frac{\sin(\tau)}{\Delta V G_0} \\ &\times \sum_{n=2}^{\infty} \frac{n}{n^2 - 1} \delta i_D^{(0,n)} \left[\delta G_D^{(0,n-1)} - \delta G_D^{(0,n+1)} \right]. \quad (66) \end{aligned}$$

Using the identities (49) and (50), we can rewrite them as

$$\begin{aligned} \delta G_D(u_0)u_1|_{\text{first harm.}} &= a_1 \cos(\tau) \left[\frac{\delta i_D^{(0,1)}}{a_0 \Delta V} + \delta G_D^{(0,2)} \right] \\ &+ b_1 \sin(\tau) \frac{\delta i_D^{(0,1)}}{a_0 \Delta V} \\ &- \frac{1}{a_0} \frac{\sin(\tau)}{\Delta V^2 G_0} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} \left[\delta i_D^{(0,n)} \right]^2. \end{aligned} \quad (67)$$

Substituting that into Eq. (65), we get

$$\begin{aligned} g_1|_{\text{first harm.}} &= -a_1 \sin(\tau) \left[\xi - \frac{1}{a_0 \Delta V} \frac{\delta i_D^{(0,1)}}{G_0} - \frac{\delta G_D^{(0,2)}}{G_0} \right] \\ &+ b_1 \cos(\tau) \left[\xi - \frac{1}{a_0 \Delta V} \frac{\delta i_D^{(0,1)}}{G_0} \right] \\ &+ \frac{1}{a_0} \frac{\cos(\tau)}{(\Delta V G_0)^2} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} \left[\delta i_D^{(0,n)} \right]^2 \\ &+ 2\kappa_1 a_0 \cos(\tau). \end{aligned} \quad (68)$$

According to Eq. (57), the expression in the square brackets in Eq. (68) in the term with b_1 is zero. Additionally, the expression in the square brackets in the term with a_1 becomes $-\delta G_D^{(0,2)}/G_0$. Since this is the only term with $\sin(\tau)$, it must be zero and then

$$a_1 = 0, \quad (69)$$

if we ignore the special case $\delta G_D^{(0,2)} = 0$. Note, the condition $\delta G_D^{(0,2)} = 0$ is fulfilled at the outer edges of the oscillator hysteresis regions (where the derivative over a_0 from the right-hand side of Eq. (57) is zero; use Eqs. (50), (45), and (8) in the derivation), when the oscillation amplitude becomes unstable and, therefore, a_1 becomes undefined; see the hysteresis examples in Sec. VI.

Setting the other terms with $\cos$ equal to zero, we get the following equation for κ_1 :

$$\kappa_1 = -\frac{1}{2(a_0 \Delta V G_0)^2} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} \left[\delta i_D^{(0,n)} \right]^2. \quad (70)$$

IV. SUMMARY OF THE PERTURBATION-THEORY RESULTS AND DISCUSSION

In the beginning, we make several comments on the structure of the generalized van der Pol equation (21). Its

left-hand side is linear and it describes a lossless harmonic oscillator with the eigenfrequency ω_0 . The right-hand side of the equation describes the linear gain (or loss) and nonlinearity of the oscillator.

It is instructive to note that the parameter $\epsilon\xi$ is an instability gain (or loss) parameter of a linear (small-signal) oscillator equation, since it stands in front of the first-order time derivative $\partial_t \tilde{v}$. The nonlinear term on the right-hand side of Eq. (21) disappears in the small-signal limit. The product $\epsilon\xi$ is

$$\epsilon\xi = -\frac{G_D(0) + G_I}{C\omega_0}, \quad (71)$$

as it is seen from Eqs. (15) and (20). G_0 is introduced in the definitions of ϵ and ξ , see Eqs. (15) and (20), to give a common scale to different conductance-related parameters and to keep them unitless, but eventually it cancels out in $\epsilon\xi$. If RTD is biased at the maximum of its NDC ($G_D(0) = -G_0$) and also $G_I \ll G_0$, then $\xi \approx 1$, see Eq. (15). In this case, the parameter ϵ has a straightforward physical meaning: it determines a small-signal (linear) instability rate. If $\epsilon \ll 1$, the instability evolves slowly on the time scale of the period of the eigenfrequency oscillations. On the other hand, if $\epsilon \gtrsim 1$, the instability develops faster or on the timescale of the oscillation period. At other bias points and/or at larger values of G_I , the linear instability rate is reduced by ξ (note, $\xi \leq 1$). This dynamic behavior is relevant to the modulation speed of RTD oscillators, particularly for on/off keying. In the case of a relaxation oscillator ($\epsilon \gg 1$), the oscillation amplitude indeed sets up within one cycle.

The nonlinear term in the square brackets on the right-hand side of Eq. (21) defines the scale of the variable $\tilde{v}$, when the nonlinearity becomes large enough to compensate for the linear gain in the system. That happens when the nonlinear term, averaged in a certain way over the oscillation period, becomes approximately equal to ξ (or 1, if the RTD is biased at the NDC maximum and G_I is small). Here, we can face two situations. First, the slightly nonlinear regime, when the oscillating voltage is sinusoidal (zero-order approximation), its shape is determined by the harmonic equation on the left-hand side of Eq. (21). Note, the nonlinearity and the small-signal instability are decoupled in this case: the gain-compensation condition is independent of ϵ , and ϵ does not influence the zero-order oscillation amplitude.

Second, the regime with stronger nonlinearity, when ϵ becomes sufficiently large, and higher order(s) of the perturbation theory should be taken into account (or the perturbation theory becomes not applicable). The time dependence of the oscillating voltage becomes more (or essentially) nonsinusoidal and the impact of the nonlinear

terms becomes dependent on the shape of the oscillating voltage. Then, the dynamic and nonlinear parameters become intertwined.

Now, we summarize the perturbation-theory results. The solution of the oscillator equation (33) is determined by the perturbation parameter ϵ , see Eq. (20). The solution of Eq. (33) in the zero-order approximation is

$$\tilde{v}(\tau) = a_0 \cos(\tau) + O(\epsilon), \quad (72)$$

where a_0 is determined by the solution of Eq. (59); we recall that $i_D^{(0,1)}$ is dependent on a_0 . Note that Eq. (59) and the amplitude a_0 are independent of ϵ , as mentioned above. The oscillation frequency stays equal to the eigenfrequency of the LC resonator:

$$\tilde{\omega} = 1 + O(\epsilon^2). \quad (73)$$

As a next step, we summarize the first-order approximation results. Since $\delta i_D^{(0,n)} = i_D^{(0,n)}$ for $n \neq 1$, we can drop δ in the sums in Sec. III E and write the results as

$$\begin{aligned} \tilde{v}(\tau) &= a_0 \cos(\tau) + \epsilon b_1 \sin(\tau) \\ &+ \epsilon \sum_{n=2}^{\infty} b_1^{(n)} \sin(n\tau) + O(\epsilon^2), \end{aligned} \quad (74)$$

$$b_1 = \frac{1}{\Delta V G_0} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} i_D^{(0,n)}, \quad (75)$$

$$b_1^{(n)} = -\frac{1}{\Delta V G_0} \frac{n}{n^2 - 1} i_D^{(0,n)}, \quad \text{for } n \geq 2, \quad (76)$$

$$\tilde{\omega} = 1 + \epsilon^2 \kappa_1 + O(\epsilon^3), \quad (77)$$

$$\kappa_1 = -\frac{1}{2(a_0 \Delta V G_0)^2} \sum_{n=2}^{\infty} \frac{n^2}{n^2 - 1} [i_D^{(0,n)}]^2. \quad (78)$$

Now, we will make several general comments on the results that were obtained.

First, $\tilde{v}$ in the zero-order approximation is purely harmonic, see Eq. (72). Nonfundamental harmonics appear only in the first order of the perturbation theory, see Eq. (74). Their amplitudes decrease as $\propto i_D^{(0,n)}/n$, see Eq. (76).

Second, a nonzero correction to the oscillation frequency appears starting at the first order of the perturbation theory. That is true for the van der Pol equation, but this also holds true in the generalized case of an arbitrary RTD I - V curve. The first-order frequency correction is of the second order in the perturbation parameter ($\propto \epsilon^2$), see Eq. (77). Nevertheless, it originates in the first order of the perturbation theory and, following the mathematical literature (Poincaré-Lindstedt method), we denote it as a first-order perturbation term.

Third, the zero-order amplitude a_0 and first-order coefficients b_1 , $b_1^{(n)}$, and κ_1 are determined exclusively by the dc RTD I - V curve and the load conductance G_L . The dynamic RTD parameters (C and ω_0) appear only in the perturbation parameter ϵ , see Eq. (20), which further determines the strength of the first-order terms in Eqs. (74) and (77).

Fourth, we introduced the maximum NDC G_0 to give a scale to the nonlinear variation of the RTD conductance and to the linear instability and perturbation parameter ϵ and to eventually write the equations in a unitless form. However, amplitudes of the first-order terms are determined by the parameters ϵb_1 , $\epsilon b_1^{(n)}$, and $\epsilon^2 \kappa_1$, see Eqs. (74)–(78), which are independent of G_0 , since $\epsilon \propto G_0$ and $b_1 \propto 1/G_0$, $b_1^{(n)} \propto 1/G_0$, $\kappa_1 \propto 1/G_0^2$, and G_0 cancels out.

Fifth, a subtle point needs to be discussed. When is the perturbation theory applicable? Obviously, it is applicable when $\epsilon \ll 1$. However, this condition could be relaxed. We can define the validity range of the perturbation theory by imposing the condition that the first-order corrections are small, since $a_0 \sim 1$. That means that we need to require

$$|\epsilon b_1| \ll 1, \quad |\epsilon b_1^{(n)}| \ll 1, \quad |\epsilon^2 \kappa_1| \ll 1. \quad (79)$$

Essentially, as it follows from Eqs. (74)–(78), we can define the small parameter in the problem as

$$\epsilon \frac{i_D^{(0,n)}}{a_0 \Delta V G_0} \ll 1, \quad \text{for } n \geq 2, \quad (80)$$

or we can write it in the form

$$\frac{1}{C \omega_0} \frac{i_D^{(0,n)}}{a_0 \Delta V} \ll 1, \quad \text{for } n \geq 2. \quad (81)$$

In practical terms, we can replace the strong requirement $\ll 1$ in the inequalities (79) by $\lesssim 0.3$. The above inequalities are dependent on ϵ , which is $\propto 1/\omega_0$. Therefore, these inequalities are always satisfied at sufficiently high oscillator frequencies. However, they always fail at sufficiently low frequencies. At those low frequencies, the perturbation theory (including its zero-order approximation) is not applicable anymore, the RTD oscillators are entering a strongly nonlinear oscillation regime, and they work essentially as relaxation oscillators. In this regime, the RTD oscillator should be analyzed using different mathematical or numerical methods. In particular, the oscillation frequency of the RTD oscillator should drop significantly below its eigenfrequency ω_0 , as indicated by the negative sign of κ_1 in Eq. (78).

By applying conditions (79), we assume that the higher-order terms in the perturbation series will keep decreasing as the higher orders of the parameter (80). Although intuitively expectable, that is not guaranteed and remains to be proved or verified by analysis of the high-order terms or

with numerical simulations. We leave this question open and beyond the scope of this paper.

Lastly, it can be shown from the parity symmetry of g that only odd orders of the perturbation affect κ and generate sinusoidal harmonics, while even orders perturb the coefficient a_0 and cos harmonics. Consequently, the next correction to κ beyond the leading term scales with ϵ^4 . Therefore, for small ϵ , the first-order perturbation provides a very accurate prediction of κ .

Sixth, one additional point needs to be addressed: under what conditions does the quasi-static approximation for the I - V curve (4) hold? This approximation is valid when the charge-relaxation processes inside the RTD can be considered fast. This condition can be expressed as $\omega\tau_{\text{rel}} \ll 1$, where τ_{rel} is the small-signal charge-relaxation time constant [21–24]. However, this condition becomes less stringent (i.e., the effective averaged τ_{rel} decreases) when RTDs operate in a large-amplitude regime [18]. Since one usually tries to operate RTD oscillators in this regime, the applicability condition of the presented theory may be approximately reformulated for simplified assessments as $\omega\tau_{\text{rel}} \lesssim 1$, where the largest value of τ_{rel} at the NDC maximum is used [21–24].

Beyond this approximation, the analysis of RTD oscillators in the form of a generalized van der Pol oscillator, as presented in this paper, remains valid when applied to a “flattened” high-frequency RTD I - V curve, which becomes frequency independent when $\omega\tau_{\text{rel}} \gg 1$ [17]. Otherwise, more accurate oscillator models based on general large-signal dynamic RTD models [18] are required to describe RTD oscillators at intermediate frequencies.

V. VAN DER POL EQUATION

To verify the model, we apply it to an RTD with its I - V curve approximated by a third-order polynomial. This approximation leads to the van der Pol equation.

In this case, we approximate the nonlinear RTD I - V curve as the following third-order polynomial

$$i_D(v) = -G_0 v + B v^3, \quad (82)$$

where B is the third-order nonlinearity coefficient. In Eq. (82), we assume that the RTD is biased at the centre of its NDC region, i.e., at the point of its maximum NDC ($G_D(0) = -G_0$). We adjust the parameters G_0 and B in such a way that the peak and valley points of the real and the approximated RTD I - V curves are coinciding. Then, the parameters in Eq. (82) take the following values:

$$G_0 = \frac{3}{2} \frac{\Delta I}{\Delta V} \quad (83)$$

and

$$B = \frac{4}{3} \frac{G_0}{\Delta V^2}, \quad (84)$$

where ΔI and ΔV are the current and bias separation between the peak and valley points of the real RTD I - V curve.

Rewriting (82) in terms of the normalized voltage (16), we get

$$i_D(\tilde{v}) = -G_0 \Delta V \tilde{v} + \frac{4}{3} G_0 \Delta V \tilde{v}^3 \quad (85)$$

$$G_D(\tilde{v}) = -G_0 + 4G_0 \tilde{v}^2. \quad (86)$$

Correspondingly,

$$\delta i_D(\tilde{v}) = \frac{4}{3} G_0 \Delta V \tilde{v}^3 \quad (87)$$

$$\delta G_D(\tilde{v}) = 4G_0 \tilde{v}^2. \quad (88)$$

Applying the zero-order solution leads to the following Fourier series representation of the RTD current

$$\begin{aligned} i_D(a_0 \cos(\tau)) \\ = G_0 \Delta V a_0 \left[(-1 + a_0^2) \cos(\tau) + \frac{1}{3} a_0^2 \cos(3\tau) \right]. \end{aligned} \quad (89)$$

Taking the first-harmonic term in the above equation (which is $i_D^{(0,1)}$) and substituting it into Eq. (59), we get an equation for the amplitude a_0

$$a_0^2 = \xi = \frac{G_0 - G_I}{G_0}. \quad (90)$$

This completes the solution in the zero order of the perturbation theory, according to Eqs. (59), (72), and (73).

As a next step, we calculate the solution up to the first order of the perturbation theory, Eqs. (74)–(78). As it follows from Eq. (89), only the third harmonic is nonzero in the sums in Eqs. (74), (75), and (78), with

$$i_D^{(0,3)} = \frac{1}{3} G_0 \Delta V a_0^3. \quad (91)$$

Substituting that into Eqs. (74)–(78), we get

$$\begin{aligned} \tilde{v}(\tau) = a_0 \cos(\tau) + \epsilon b_1 \sin(\tau) + \epsilon b_1^{(3)} \sin(3\tau) \\ + O(\epsilon^2), \end{aligned} \quad (92)$$

$$b_1 = \frac{3}{8} a_0^3, \quad (93)$$

$$b_1^{(3)} = -\frac{1}{8} a_0^3, \quad (94)$$

$$\kappa_1 = -\frac{1}{16} a_0^4. \quad (95)$$

For completeness, we rewrite Eq. (21) in the usual form of the van der Pol equation. First, we substitute Eq. (88) into

Eq. (21), and it takes the form

$$\frac{\partial_t^2 \tilde{v}}{\omega_0^2} + \tilde{v} = \epsilon \xi \left[1 - 4 \frac{\tilde{v}^2}{\xi} \right] \frac{\partial_t \tilde{v}}{\omega_0}. \quad (96)$$

Second, we rescale $\tilde{v}$ as

$$\tilde{\tilde{v}} = 2 \frac{\tilde{v}}{\sqrt{\xi}} = 2 \frac{\tilde{v}}{a_0}, \quad (97)$$

then Eq. (96) takes the usual form of the van der Pol equation

$$\frac{\partial_t^2 \tilde{\tilde{v}}}{\omega_0^2} + \tilde{\tilde{v}} = \epsilon \xi \left[1 - \tilde{\tilde{v}}^2 \right] \frac{\partial_t \tilde{\tilde{v}}}{\omega_0} \quad (98)$$

with the perturbation parameter $\epsilon \xi$. Applying Eq. (97) to Eqs. (92)–(95), keeping $\epsilon \xi$ as the perturbation parameter, and using Eq. (90), we get

$$\tilde{\tilde{v}} = 2 \cos(\tau) + \epsilon \xi \frac{3}{4} \sin(\tau) - \epsilon \xi \frac{1}{4} \sin(3\tau) + O((\epsilon \xi)^2), \quad (99)$$

$$\tilde{\tilde{\kappa}}_1 = -\frac{1}{16}. \quad (100)$$

The coefficients in Eqs. (99) and (100) correspond exactly to those for the zero- and first-order solutions of the van der Pol equation, see Refs. [19,20].

VI. EXAMPLES

For illustration, we apply the perturbation theory to the analysis of oscillators based on two RTD wafers with rather typical parameters. Those are InGaAs/AlAs RTDs with the main parameters listed in Table I. One RTD (“1.6-RTD”) has relatively thick AlAs barriers of 1.6 nm. This RTD has a peak current density of $6.2 \text{ mA}/\mu\text{m}^2$ and $G_0 \approx 44 \text{ mS}/\mu\text{m}^2$, see Table I. This RTD wafer was designed for sub-THz frequencies. It was extensively used by us in a number of experimental studies in the past [14,16,25,26]. The other RTD (“1.0-RTD”) has very thin barriers of 1.0 nm, the peak current density of $24 \text{ mA}/\mu\text{m}^2$ and $G_0 \approx 105 \text{ mS}/\mu\text{m}^2$, see Table I. It was designed for THz frequencies, and used in a number of experimental studies [7,16,27–29]. The RTD layer sequences of both wafers are provided in the above references. Theoretically calculated dc RTD I - V curves are shown in Fig. 3. These I - V curves match very well the measured ones, see the above references. Figure 4 shows the differential conductances (G_D) of the RTD wafers. τ_{rel} in Table I denotes the small-signal charge-relaxation time constant at the NDC maximum [21–24]. As discussed in Sec. IV, the quasistatic

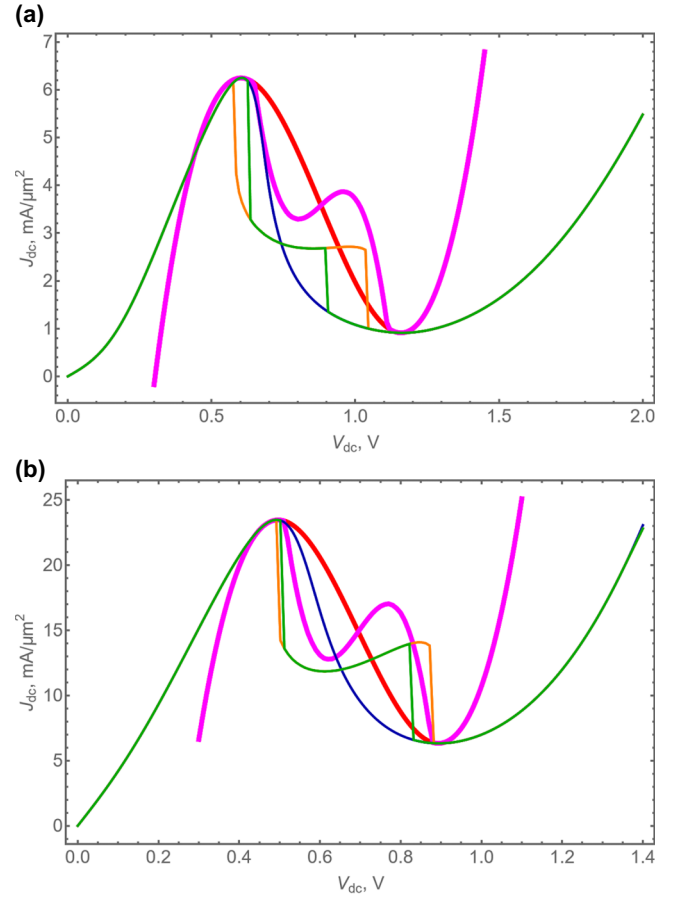


FIG. 3. Theoretical I - V curves for the 1.6-RTD (a) and 1.0-RTD (b) wafers. Blue lines show stable dc I - V curves. Orange and green I - V curves are the dc I - V curves of an oscillating RTD in an oscillator. The red lines show third-order approximations of the real I - V curves. The magenta lines show the third-order polynomial I - V curves modified by the oscillations.

approximation for the I - V curve (4) is expected to hold approximately up to the frequency defined by the condition $\omega \tau_{\text{rel}} \approx 1$, which corresponds to approximately 2 THz for the 1.0-RTD and approximately 400 GHz for the 1.6-RTD.

TABLE I. Main parameters of the considered RTD wafers.

RTD wafer	1.6-RTD	1.0-RTD
Barrier thickness	1.6 nm	1.0 nm
ΔV	0.55 V	0.4 V
Peak current density	$6.2 \text{ mA}/\mu\text{m}^2$	$24 \text{ mA}/\mu\text{m}^2$
G_0	$44 \text{ mS}/\mu\text{m}^2$	$105 \text{ mS}/\mu\text{m}^2$
G_0 (third order)	$14 \text{ mS}/\mu\text{m}^2$	$65 \text{ mS}/\mu\text{m}^2$
C	$6 \text{ fF}/\mu\text{m}^2$	$8.9 \text{ fF}/\mu\text{m}^2$
τ_{rel} (at max NDC)	380 fs	85 fs

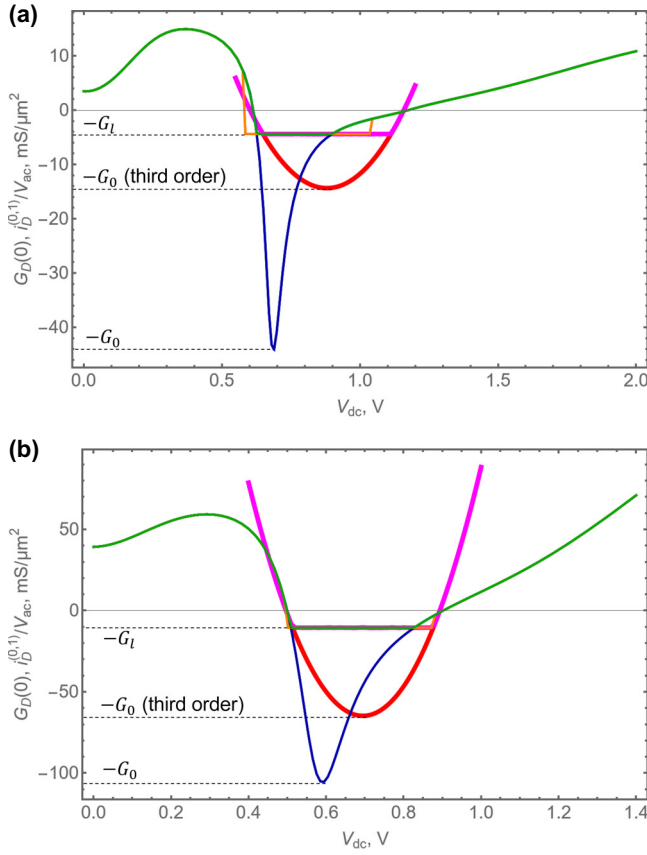


FIG. 4. Differential RTD conductances (blue lines) for the 1.6-RTD (a) and 1.0-RTD (b) wafers. The orange and green lines show the zero-order RTD conductances, modified by the large-amplitude oscillations; these conductances are defined as $i_D^{(0,1)}/a_0 \Delta V$. Red and magenta lines correspond to third-order approximations of the real RTD I - V curves, see notations in the subscript to Fig. 3.

A. Zero-order solutions and first-order perturbation terms

Our aim is to illustrate the RTD-oscillator characteristics in the regime of nearly maximal oscillation amplitudes. Therefore, we choose fixed small values of the load conductance (different for different wafers): $G_l = 0.1 G_0$. The solutions of Eq. (59) define the zero-order oscillation amplitudes a_0 . The calculation results are shown in Fig. 5. The figure shows that $a_0 \sim 1$ for both wafers and a_0 are dependent on the dc operating bias (V_{dc}). On both sides of the oscillation regions, one can see hysteresis regions (between the green and orange curves). The corresponding effective large-signal RTD conductances are shown in Fig. 4. As expected, they exhibit plateaus in the oscillation region at the level $-G_l$, as defined by Eq. (59). The I - V curves with steplike behavior due to the self-rectification effect in and near the NDC region, where RTD oscillates, are shown in Fig. 3; the I - V curves also exhibit hysteresis regions. These I - V curves are calculated as $I_{dc}(V_{dc}) + i_D^{(0,0)}$.

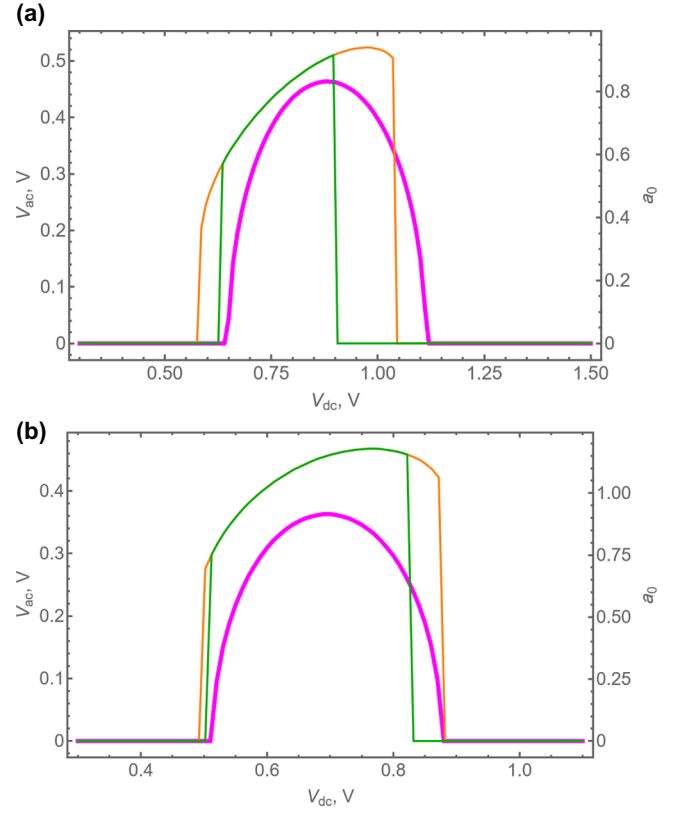


FIG. 5. Oscillation amplitudes with $G_l = 0.1 G_0$ for the 1.6-RTD (a) and 1.0-RTD (b) wafer, see notations in the subscript to Fig. 3. Magenta lines show the oscillation amplitudes for the third-order approximations of the I - V curves.

To calculate the first-order terms, we need to define the parameter ϵ , which is dependent on the RTD capacitance, see Eq. (20). In the calculations, we assume that the oscillator capacitance C is equal to the RTD capacitance. We set the values of the RTD capacitance as given in Table I, see discussion later. The corresponding values of ϵ as a function of frequency are shown in Fig. 6.

Now, we need to check the applicability conditions for the perturbation theory. As pointed out in Sec. IV, the inequalities (79) should be satisfied. First, we look at the data for the 1.6-RTD wafer. The calculation results for the first-order terms at 200 GHz are shown in Figs. 7(a) and 8(a). Figure 7(a) shows that the first-order terms nearly satisfy the requirement $\lesssim 0.3$ at this frequency. The coefficient κ_1 is strongly dependent on the operating bias point, as one can see in Fig. 8(a). The condition $\lesssim 0.3$ is valid for the frequency-deviation term $\epsilon^2 \kappa_1$ in the bias range $V_{dc} \approx 0.75$ – 1.0 V, see Fig. 8(a). However, at lower dc biases of < 0.75 V, this term becomes significantly larger and the perturbation theory stops to be applicable there already at frequencies of around 300 GHz. The data for the

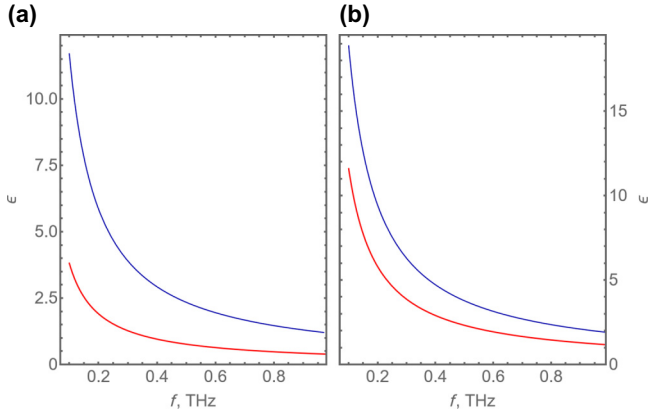


FIG. 6. Coefficients ϵ (blue lines) for the 1.6-RTD (a) and 1.0-RTD (b) wafer. Red lines show the results for third-order polynomial approximations of the RTD I - V curves (G_0 for the same wafers are correspondingly adjusted, see Table I and Fig. 4).

1.0-RTD wafer, shown in Figs. 7(c) and 8(b), look similar, but the lower-bound frequency for applicability of the perturbation theory is approximately 700 GHz in this case.

To illustrate how much the oscillator frequency drops compared to its eigenfrequency, we plot the parameter $\epsilon^2 \kappa_1$ vs frequency in Fig. 9. For the plots, we choose the dc bias points, where the RTD oscillators have nearly maximal oscillation amplitude (see Fig. 5) and also a small value of κ_1 (see Fig. 8), i.e., the frequency deviation is minimal at these points. These dc bias points are close to $V_{dc} = 0.9$ V for the 1.6-RTD wafer and to $V_{dc} = 0.7$ V for the 1.0-RTD wafer. For the 1.6-RTD wafer, the oscillation frequency should drop by 30% at $f_0 = 100$ GHz (that is already nearly strongly nonlinear regime); at 200 GHz, the frequency drop is approximately 7%, which is a significant value; at higher frequencies, the oscillator should operate closer to its eigen frequency. For the 1.0-RTD wafer, the frequency drop is approximately 8% at $f_0 = 700$ GHz and it is lower at higher frequencies.

As indicated in Sec. IV, beyond the specified range of biases and at lower frequencies, the perturbation theory is not applicable, the RTD oscillators go into a relaxation-oscillator regime, and the oscillation frequency should drop significantly below the resonator eigenfrequency ω_0 . With respect to the 1.6-RTD wafer, these considerations indicate why we had to choose (empirically) a somewhat larger value of the RTD capacitance of $C = 9$ fF/ μm^2 in Ref. [14] (contrary to 6 fF/ μm^2 in Table I), where this wafer was used for RTD oscillators at approximately 100 GHz: it turns out, that those oscillators were probably working in the nearly strongly nonlinear regime with a reduced oscillation frequency. The frequency reduction was mimicked (unknowingly) by the increased RTD capacitance.

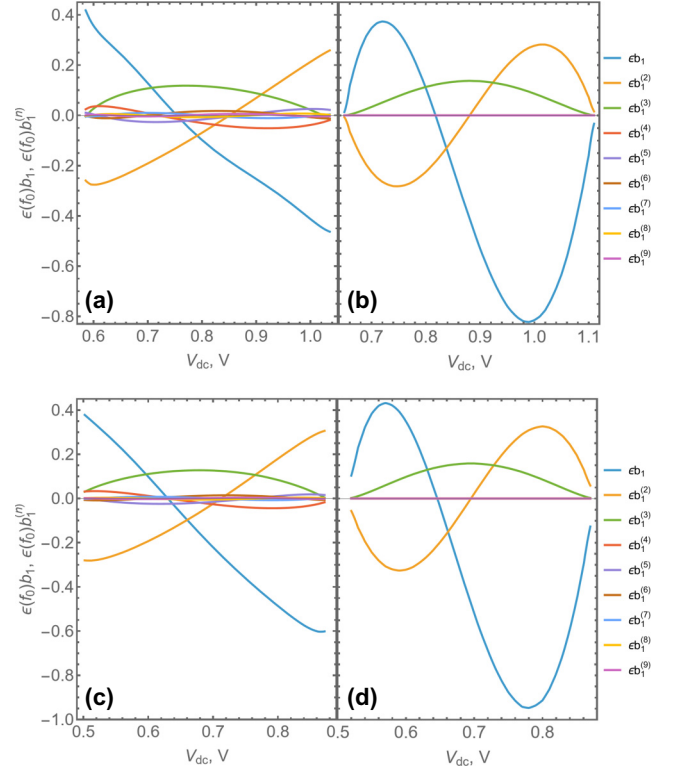


FIG. 7. first-order terms ϵb_1 and $\epsilon b_1^{(n)}$ for the 1.6-RTD (a),(b) and 1.0-RTD (c),(d) wafers at $f_0 = 200$ GHz and $f_0 = 700$ GHz, respectively. (a),(c) Calculation results for the real I - V curves; (b),(d) similar results for the third-order polynomial approximations of the RTD I - V curves.

For comparison, we now apply the third-order polynomial approximations to the RTD I - V curves and calculate the same zero- and first-order parameters of the perturbation theory. We keep the values of G_l for each of the wafers the same as before (to have a fair comparison), but G_0 will have modified values, corresponding to the approximated

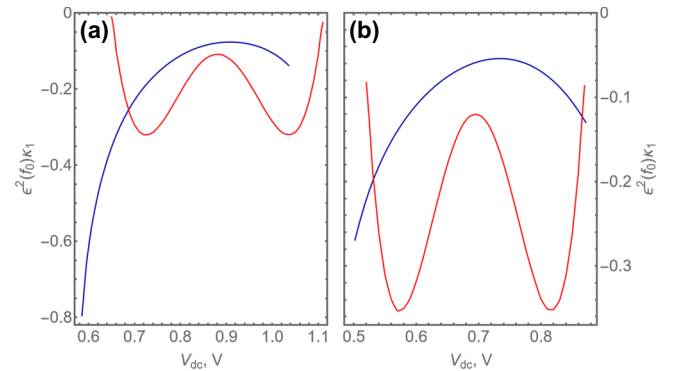


FIG. 8. First-order terms $\epsilon^2 \kappa_1$ (blue lines) for the 1.6-RTD (a) and 1.0-RTD (b) wafers at $f_0 = 200$ GHz and $f_0 = 700$ GHz, respectively. Red lines: similar results in the case of the third-order polynomial approximation of the RTD I - V curve.

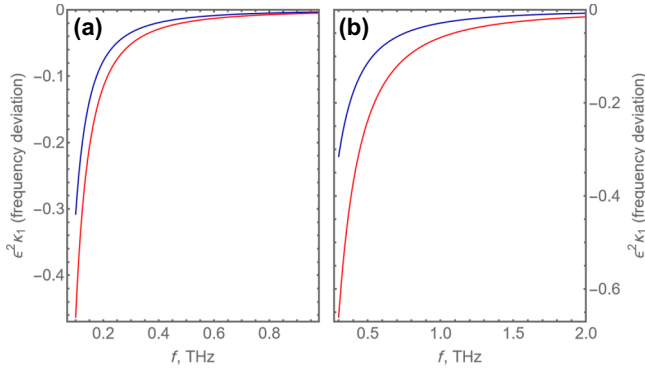


FIG. 9. Frequency reduction (blue lines) compared to the eigen oscillator frequency for the 1.6-RTD (a) and 1.0-RTD (b) wafers at $V_{dc} = 0.9$ V and $V_{dc} = 0.7$ V, respectively. Red lines: similar results for the third-order polynomial approximations of the RTD I - V curves.

I - V curves, see Table I. Figure 3 shows the third-order polynomial approximations of the RTD I - V curves. The polynomial parameters are defined as described in Sec. V. Figures 3–9 show the calculation results applied to the I - V curves given by the third-order polynomial approximations. Figure 5 shows that the oscillation amplitudes retain the same scale as for the real I - V curves. However, there are quantitative and even qualitative differences. First, the third-order polynomial approximations do not reproduce hysteresis regions; higher-order polynomial approximation better matching the real RTD I - V curves are required for that. Correspondingly, the behavior of the RTD conductance and I - V curves for the oscillating RTDs also do not exhibit hysteresis, see Figs. 3 and 4. Second, one can notice certain quantitative deviations in the positions and shapes of the oscillating zero-order characteristics in Figs. 3–5.

Further on, the first-order characteristics also exhibit a number of quantitative and qualitative differences. The value of ϵ drops by approximately 3 times for the wafer 1.6-RTD, although the reduction is not that large for wafer 1.0-RTD, see Fig. 6. The differences are due to lower peak NDC for the third-order polynomial approximation, see Fig. 4. In Fig. 7, we see that the third-order polynomial approximation results (as they should) contain only the fundamental, second- (when bias is shifted from the center of NDC) and third-order harmonics, although the real RTD exhibits much more rich output spectrum with many harmonics with nonzero amplitudes. Significantly larger value of ϵb_1 in the case of third-order polynomial approximation is also apparent; this component is responsible for a phase shift of the fundamental harmonic (with respect to $t = 0$ defined by the initial conditions (23), the phase shift accompanies deformation of the trajectories in the phase plane $(\tilde{v}, \partial_t \tilde{v})$ with the variation of the dc bias. One can also clearly notice the qualitative differences in the behavior of the first-order terms in Fig. 7: all terms

converge to zero at the edges of the oscillation region in the case of the third-order polynomial approximation, although some terms remain large and exhibit an abrupt jump at the edges of the oscillation region for the real RTD I - V curves. This behavior is related to the hysteresis at the edges of the oscillation region and sharp changes of the oscillation amplitude there, see Fig. 5.

The variations of the frequency deviations with dc bias look different for the real I - V curves and their third-order polynomial approximations, see Fig. 8. The former ones are quite smooth, although the latter ones are wavy. Such nonmonotonous dependence on the dc bias has never been observed experimentally, which indicates that the real RTD I - V curves are not well approximated by the third-order polynomials, as far as the frequency deviations in RTD oscillators are concerned. Significant quantitative differences between these two curves are also apparent. If we choose the dc operating point close to the NDC maximum or a maximum amplitude (a_0), in our case it is $V_{dc} \approx 0.9$ V for 1.6-RTD wafer, then the frequency deviation may become quite similar, see Fig. 8(a) and also Fig. 9(a), at least for this example RTD wafer. However, at a similar bias of $V_{dc} \approx 0.7$ V for 1.0-RTD wafer, the difference is quite large, it is a factor of approximately 2, see Figs. 8(b) and 9(b). Again, the third-order polynomial approximation can lead to significant quantitative errors.

B. Variation of the oscillation frequency with bias

It is known that the RTD-oscillator frequency is changing with dc bias [30]. However, the mechanism of this change and its quantitative and qualitative behavior were not thoroughly investigated. Here, we try to elucidate certain characteristic features of the frequency variation in RTD oscillators with bias.

We start with several experimental observations. There are experimental works reporting or mentioning the frequency change in RTD oscillators with bias; the frequency variation is usually rather small (approximately 1% and less), but sometimes it can become quite large (approximately 5–7%) [10,28–31]. The frequency typically increases with the increase of RTD bias [30]. However, the shape of this frequency change is sometimes [28,29] very peculiar: first, the frequency grows with bias, but then it reaches a maximum and later somewhat (smoothly) drops until the opposite side of the oscillation region is reached.

It is intuitively clear why the oscillation frequency is rising with bias: the RTD collector region gets more depleted and longer with the increase of bias; that leads to the reduction of the RTD capacitance and, as a consequence, to the increase of the oscillation frequency. However, that does not explain why the frequency starts to decrease at high biases. Something essential is missing in this simple explanation.

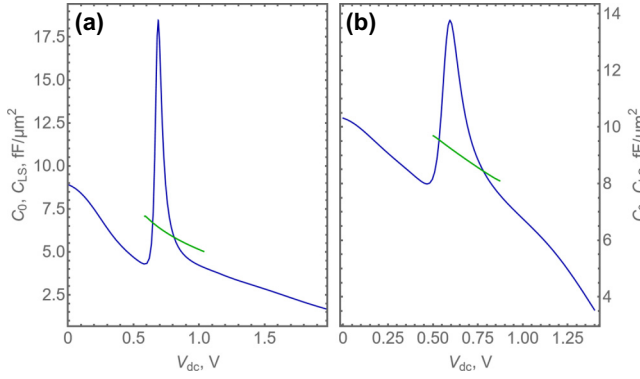


FIG. 10. Small-signal RTD capacitances (blue lines) for the 1.6-RTD (a) and 1.0-RTD (b) wafers. Green lines: large-signal RTD capacitances ($C_{LS}(V_{dc})$) in the oscillation region.

In an attempt to make these capacitive arguments more accurate, we include effects related to the large-signal capacitance. This goes beyond the van der Pol equation, which is based on the assumption of a constant capacitance in the oscillator. So, first, we calculate a small-signal RTD capacitance (C_0), see Fig. 10. The calculation procedure follows [21–23]. The small-signal capacitance exhibits a characteristic peak in the NDC region of the RTD I - V curve. Then, we calculate a large-signal RTD capacitance (C_{LS}): when large-amplitude oscillations are applied to an RTD, its effective capacitance is averaged in a certain way over the oscillation cycles. The calculation procedure follows that described in Ref. [7]:

$$C_{LS}(V_{dc}) = \frac{2}{\pi} \int_0^\pi C_0(V_{dc} + \Delta V a_0 \cos(\tau)) \sin^2(\tau) d\tau. \quad (101)$$

As the oscillation amplitudes, we take the zero-order amplitudes shown in Fig. 5; the large-signal capacitances (C_{LS}) for both wafers calculated in such a way are shown in Fig. 10. For 1.6-RTD wafer, C_{LS} changes smoothly in the range 5–7 fF/μm². Therefore, we take an average value of $C = 6$ fF/μm² in our calculations in Sec. VI A, see also Table I. In a similar way, we estimate an average capacitance of $C = 8.9$ fF/μm² for the 1.0-RTD wafer.

As a next step, one can evaluate the frequency change ($\Delta\omega$) due to the change in C_{LS} with bias [$\Delta C_{LS}(V_{dc})$]; according to Eq. (18), it is

$$\frac{\Delta\omega}{\omega_0} \approx -\frac{1}{2} \frac{\Delta C_{LS}(V_{dc})}{C}. \quad (102)$$

The corresponding frequency deviations with bias are shown in Fig. 11. They exhibit a smooth monotonous increase of frequency with bias, i.e., the capacitance change with bias does not fully explain the effect.

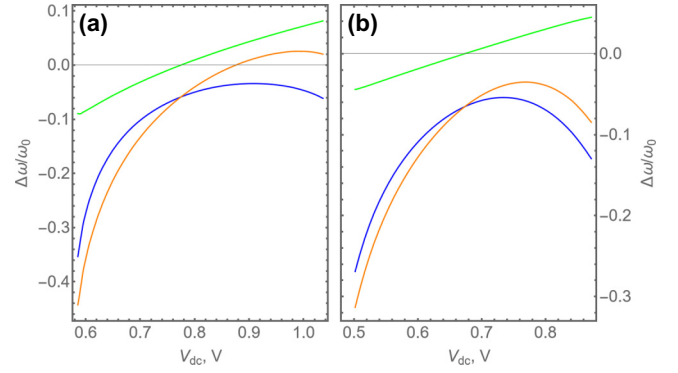


FIG. 11. Frequency deviations for the 1.6-RTD (a) and 1.0-RTD (b) wafers. Green lines: variation of the oscillation frequency due to the bias dependence of $C_{LS}(V_{dc})$. Blue lines: frequency deviation due to the first-order corrections of the perturbation theory at the oscillation frequency of $f_0 = 300$ GHz (compare with Fig. 8) for the 1.6-RTD wafer and $f_0 = 700$ GHz for the 1.0-RTD wafer. Orange lines: the sums of both contributions.

Now, let us include the first-order frequency corrections of the perturbation theory into our analysis. As Fig. 8 indicates, first-order frequency corrections have qualitatively the right shape. We choose the oscillation frequency of 300 GHz for 1.6-RTD wafer, so that the corrections are smaller than those in Fig. 8 and roughly fall within the validity range of the perturbation theory. For the 1.0-RTD wafer, we choose the frequency of 700 GHz, where the perturbation theory is still roughly applicable. The calculation results are shown in Fig. 11. We also show the sum of both contributions of $C_{LS}(V_{dc})$ and the first-order frequency corrections. The combined contributions have qualitatively correct shapes for both wafers in agreement with experimental observations: the frequency grows with bias, reaches a maximum and then slightly drops.

In such a way, we demonstrate that the first-order corrections of the perturbation theory are, probably, playing an essential role in determining the frequency variation of the oscillation frequency with bias. Note that these corrections could also be quantitatively larger than those related to $C_{LS}(V_{dc})$, as the given examples show.

We need to address one more point in this context. The 1.0-RTD wafer was used in the experiments in Refs. [28,29], where the frequency variation of 5.5% with dc bias at approximately 650 GHz was measured. Instead, the scale of the frequency variation shown in Fig. 11 is approximately 30% at 700 GHz, although the shape of the frequency variation agrees well with the experimental data. Where is this large quantitative discrepancy coming from? There could be several factors responsible for that. First, in the above calculations, we assumed that the RTD C_{LS} is equal to the total oscillator capacitance. In reality, the RTD capacitance constitutes a part of the total capacitance C , the other part is coming from the antenna

and resonator. Therefore, the frequency variation due to $C_{LS}(V_{dc})$ —around 10%, see the green line in Fig. 11—will be scaled down as $\propto C_{LS}/C$ with an increase in the total capacitance. Additionally, the first-order frequency correction is $\propto \epsilon^2$, and it will decrease quadratically with the increase of C . Second, we assumed that the RTD oscillator is working in nearly ultimately large-amplitude regime: G_I was assumed to be equal to a very small fraction of the RTD NDC ($G_I = 0.1 G_0$), this led to approximately 20% of frequency variation due the first-order corrections, see blue line in Fig. 11. In reality, G_I likely constitutes a larger fraction of G_0 ; then, the RTD oscillation amplitude is smaller and, as a consequence, the scale of the frequency variation with bias due to this mechanism should be smaller. Third, the RTD was assumed to be working in a quasistatic regime in the presented analysis. In reality, high-frequency effects might be already leading to some change in the RTD conductance (reduction of NDC) at 700 GHz; those effects might also reduce frequency variation with bias. Fourth, there are several effects that go beyond the generalized van der Pol oscillator model we discuss in this paper. For example, such effects are as follows: account of the nonlinearity of the RTD capacitance in the oscillator equation, variation of the resonator parameters (L , C , G_I) with frequency (resonator is not a simple LC circuit), etc. Those effects may also change quantitatively and qualitatively the bias dependence of the corrections to the oscillation frequency. Fifth, the measurements were not performed over the entire oscillation region due to difficulties in precise biasing closer to the current peak (due to large NDC) and hysteresis region, where the frequency variation is expected to be the steepest. As a result, a considerable portion of the frequency tuning range may be not measured. Presumably, careful incorporation of these factors into the oscillator model and measurements can lead to a good quantitative agreement between the theory and experimental data, but that goes beyond the scope of this paper.

VII. CONCLUSION

Generalized van der Pol oscillator models are essential for qualitative and semiquantitative analysis of nonlinear oscillators, in particular, RTD oscillators. Here, we develop such a model to account for a realistic nonlinear RTD I - V curve. To describe a weak nonlinear operation regime of the oscillators, we construct a perturbation theory and derive analytical expressions for the relevant perturbation terms up to the first order. The derived analytical expressions can serve as a reference for more complicated models and numerical simulations. We show that the nonzero corrections to the oscillation frequency ($\propto \epsilon^2$), higher harmonics ($\propto \epsilon$), and quadrature fundamental-frequency component ($\propto \epsilon$, which is responsible for the oscillator phase shift) appear only in the first order of the

perturbation theory, where (ϵ) is the perturbation parameter. That was true for the van der Pol oscillator model with a third-order polynomial approximation for the RTD I - V curve, but that remains valid also for the generalized model. We also indicate that the applicability conditions for the perturbation theory could be relaxed and they can require smallness of the first-order corrections [Eqs. (79)–(81)], rather than a much stronger condition $\epsilon \ll 1$.

For two typical RTD wafers—one with relatively thick 1.6 nm and the other one with quite thin 1.0-nm AlAs barriers, designed for sub-THz and THz frequencies, respectively—we compare perturbation-theory calculation results for a real and approximated (third-order polynomial) RTD I - V curves. We show that the account of nonlinearity of a real RTD I - V curve is responsible for qualitatively different features in the behavior of the oscillator parameters, like the appearance of harmonics of the order higher than 3, hysteresis features, smooth non-monotonous variation of the oscillator frequency with RTD bias. Furthermore, our calculations indicate that the first-order corrections to the oscillation frequency play an essential role in the mechanism of the variation of the oscillator frequency with bias. Apart from qualitative features listed above, quantitative differences between the calculated parameters for real and approximated I - V curves are also apparent and illustrated.

Furthermore, we show that RTD oscillators in a large-amplitude operation regime—this regime is aimed in high-output-power RTD oscillators—can be at the edge of the applicability region of the perturbation theory already at the frequencies of a few hundred GHz for RTDs with thicker barriers and slightly above 500 GHz for thin-barrier RTDs. At lower frequencies, these RTD oscillators should be operating in a strongly nonlinear regime of the relaxation oscillators.

If, along with RTD capacitance, the parasitic or resonator capacitances also contribute significantly to the total oscillator capacitance C , this would lead to the increase in C , to the reduction of $\epsilon \propto 1/C$, and to a corresponding reduction of the first-order corrections. This would also extend the applicability range of perturbation theory to lower frequencies.

ACKNOWLEDGMENTS

This research was funded in part by the Austrian Science Fund (FWF) [10.55776/P30892; 10.55776/DFH3543724]. Helpful discussions with Gabriele Picco are gratefully acknowledged. M.F. is grateful to Zoltán Jéhn for indicating the absence of hysteresis in the case of RTD I - V curves approximated by a third-order polynomial.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

- [1] G. Carpintero, E. Garcia-Munoz, H. Hartnagel, S. Preu, and A. Raisanen, *Semiconductor Terahertz Technology: Devices and Systems at Room Temperature Operation* (John Wiley & Sons, Chichester, 2015).
- [2] A. Leitenstorfer *et al.*, The 2023 terahertz science and technology roadmap, *J. Phys. D: Appl. Phys.* **56**, 223001 (2023).
- [3] M. Feiginov, Frequency limitations of resonant-tunnelling diodes in sub-THz and THz oscillators and detectors, *J. Infrared Millimeter Terahertz Waves* **40**, 365 (2019).
- [4] S. Suzuki and M. Asada, Fundamentals and recent advances of terahertz resonant tunneling diodes, *Appl. Phys. Express* **17**, 070101 (2024).
- [5] E. R. Brown, J. R. Söderström, C. D. Parker, L. J. Mahoney, K. M. Molvar, and T. C. McGill, Oscillations up to 712 GHz in InAs/AlSb resonant-tunneling diodes, *Appl. Phys. Lett.* **58**, 2291 (1991).
- [6] 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, and S. J. Allen, Monolithic Schottky-collector resonant tunnel diode oscillator arrays to 650 GHz, *IEEE Electron Device Lett.* **18**, 218 (1997).
- [7] D. Tuan Nguyen, P. Ourednik, and M. Feiginov, Island THz on-chip slot-antenna resonant-tunneling-diode oscillators, *Appl. Phys. Lett.* **123**, 043508 (2023).
- [8] R. Izumi, S. Suzuki, and M. Asada, in *42nd Int. Conf. Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (IEEE, Cancún, Mexico, 2017), pp. 1–2.
- [9] R. Adler, A study of locking phenomena in oscillators, *Proc. IRE* **34**, 351 (1946).
- [10] Y. Suzuki, T. Van Mai, X. Yu, S. Suzuki, and M. Asada, Phase control of terahertz waves using injection-locked resonant tunneling diode oscillator, *IEEE Trans. Terahertz Sci. Technol.* **12**, 481 (2022).
- [11] N. Orihashi, S. Suzuki, and M. Asada, One THz harmonic oscillation of resonant tunneling diodes, *Appl. Phys. Lett.* **87**, 233501 (2005).
- [12] M. Asada, Theoretical analysis of subharmonic injection locking in resonant-tunneling-diode terahertz oscillators, *Japanese J. Appl. Phys.* **59**, 018001 (2020).
- [13] K. Arzi, S. Suzuki, A. Rennings, D. Erni, N. Weimann, M. Asada, and W. Prost, Subharmonic injection locking for phase and frequency control of RTD-based THz oscillator, *IEEE Trans. Terahertz Sci. Technol.* **10**, 221 (2020).
- [14] C. Spudat, P. Ourednik, G. Picco, D. T. Nguyen, and M. Feiginov, Limitations of output power and efficiency of simple resonant-tunneling-diode oscillators, *IEEE Trans. Terahertz Sci. Technol.* **13**, 82 (2023).
- [15] E. R. Brown, C. D. Parker, S. Verghese, M. W. Geis, and J. F. Harvey, Resonant-tunneling transmission-line relaxation oscillator, *Appl. Phys. Lett.* **70**, 2787 (1997).
- [16] G. Picco, P. Ourednik, D. T. Nguyen, and M. Feiginov, Generalized analysis of output-power limitations of resonant-tunnelling-diode oscillators with symmetrical slot antennas, *IEEE Trans. Terahertz Sci. Technol.* **14**, 621 (2024).
- [17] M. Feiginov, C. Sydlo, O. Cojocari, and P. Meissner, High-frequency nonlinear characteristics of resonant-tunnelling diodes, *Appl. Phys. Lett.* **99**, 133501 (2011).
- [18] P. Ourednik, G. Picco, D. Tuan Nguyen, C. Spudat, and M. Feiginov, Large-signal dynamics of resonant-tunneling diodes, *J. Appl. Phys.* **133**, 014501 (2023).
- [19] R. Grimshaw, *Nonlinear Ordinary Differential Equations* (Blackwell Scientific Publications, Oxford, 1990), p. 173.
- [20] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems* (Springer, Berlin Heidelberg, 1996), p. 122.
- [21] M. N. Feiginov, Effect of the Coulomb interaction on the response time and impedance of the resonant-tunneling diodes, *Appl. Phys. Lett.* **76**, 2904 (2000).
- [22] M. N. Feiginov, Does the quasibound-state lifetime restrict the high-frequency operation of resonant-tunnelling diodes? *Nanotechnology* **11**, 359 (2000).
- [23] M. N. Feiginov, Displacement currents and the real part of high-frequency conductance of the resonant-tunneling diode, *Appl. Phys. Lett.* **78**, 3301 (2001).
- [24] P. Ourednik and M. Feiginov, Unified admittance model for resonant tunneling diodes: Incorporating the space-charge dynamics of the quantum well and collector, *Phys. Rev. Appl.* **25**, 024033 (2026).
- [25] P. Ourednik, T. Hackl, C. Spudat, D. Tuan Nguyen, and M. Feiginov, Double-resonant-tunneling-diode patch-antenna oscillators, *Appl. Phys. Lett.* **119**, 263509 (2021).
- [26] D. T. Nguyen, G. Picco, P. Ourednik, C. Spudat, and M. Feiginov, Impact of slot width on performance of symmetrical-slot-antenna resonant-tunneling-diode oscillators, *IEEE Trans. Terahertz Sci. Technol.* **14**, 29 (2024).
- [27] P. Ourednik and M. Feiginov, Double-resonant-tunneling-diode bridge-less patch-antenna oscillators operating up to 1.09 THz, *Appl. Phys. Lett.* **120**, 183501 (2022).
- [28] P. Ourednik, G. Picco, and M. Feiginov, in *2022 47th Int. Conf. Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)* (IEEE, Delft, Netherlands, 2022), pp. 1–2.
- [29] P. Ourednik and M. Feiginov, in *2022 15th UK-Europe-China Workshop on Millimetre-Waves and Terahertz Technologies (UCMMT)* (IEEE, Tonsberg, Norway, 2022), pp. 1–3.
- [30] N. Orihashi, S. Hattori, S. Suzuki, and M. Asada, Voltage-controlled sub-terahertz oscillation of resonant tunnelling diode integrated with slot antenna, *Electron. Lett.* **41**, 872 (2005).
- [31] M. Feiginov, C. Sydlo, O. Cojocari, and P. Meissner, Operation of resonant-tunnelling oscillators beyond tunnel lifetime limit, *Europhys. Lett.* **94**, 48007 (2011).