Nonlinear dynamical model of Costas loop and an approach to the analysis of its stability in the large

G.A. Leonov a, N.V. Kuznetsov a,b,*, M.V. Yuldashev a,b, R.V. Yuldashev a,b

a Saint-Petersburg State University, Russia
b University of Jyväskylä, Finland

A R T I C L E   I N F O

Article history:
Received 3 February 2014
Received in revised form 18 June 2014
Accepted 18 August 2014
Available online 2 September 2014

Keywords:
Costas loop
BPSK
Phase-locked loop (PLL)
Nonlinear analysis
Phase detector characteristic
Phase comparator
Simulation
Stability in the large

A B S T R A C T

The analysis of the stability and numerical simulation of Costas loop circuits for high-frequency signals is a challenging task. The problem lies in the fact that it is necessary to simultaneously observe very fast time scale of the input signals and slow time scale of phase difference between the input signals. To overcome this difficult situation it is possible, following the approach presented in the classical works of Gardner and Viterbi, to construct a mathematical model of Costas loop, in which only slow time change of signal's phases and frequencies is considered. Such a construction, in turn, requires the computation of phase detector characteristic, depending on the waveforms of the considered signals. While for the stability analysis of the loop near the locked state (local stability) it is usually sufficient to consider the linear approximation of phase detector characteristic near zero phase error, the global analysis (stability in the large) cannot be accomplished using simple linear models.

The present paper is devoted to the rigorous construction of nonlinear dynamical model of classical Costas loop, which allows one to apply numerical simulation and analytical methods (various modifications of absolute stability criteria for systems with cylindrical phase space) for the effective analysis of stability in the large. Here a general approach to the analytical computation of phase detector characteristic of classical Costas loop for periodic non-sinusoidal signal waveforms is suggested. The classical ideas of the loop analysis in the signal's phase space are developed and rigorously justified. Effective analytical and numerical approaches for the nonlinear analysis of the mathematical model of classical Costas loop in the signal's phase space are discussed.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

The Costas loop [1,2] is a classical phase-locked loop (PLL) based circuit for carrier recovery. Nowadays among the applications of Costas loop there are Global Positioning Systems (see, e.g., [3–7]), wireless communication (see, e.g., [8–14] and others [15–23].

* Corresponding author at: University of Jyväskylä, Finland.
E-mail address: nkuznetsov239@gmail.com (N.V. Kuznetsov).

A PLL-based circuit behaves as a nonlinear control system and its physical model in the signal space can be described by nonlinear nonautonomous difference or differential equations. In practice, numerical simulation is widely used for the analysis of nonlinear PLL-based models (see, e.g., [24–28]). However the explicit numerical simulation of the physical model of Costas loop or its mathematical model in the signal space (e.g., full SPICE-level simulation) is rather complicated for the high-frequency signals. The problem lies in the fact that it is necessary to consider simultaneously both very fast time...
scale of the signals and slow time scale of phase difference between the signals, so one very small simulation time-step needs to be taken over a very long total simulation period [29–31].

To overcome this difficult situation it is possible, following the approach presented, e.g., in the classical works of Gardner and Viterbi, to construct a mathematical model in the signal’s phase space, in which only slow time change of signal’s phases and frequencies is considered. Such a construction, in turn, requires the computation of phase detector characteristic, which depends on PD physical realization and the waveforms of the considered signals [32–35]. Note that “understanding how phase detectors work is one of the major keys to understanding how PLLs work” [36].

Nowadays the following scheme

1. consideration of the physical model in the signal space;
2. computation of phase detector characteristic and the construction of the mathematical model in the signal’s phase space (phase-domain macromodel [29]);
3. nonlinear analysis of the transient processes of the signal’s phases adjustment and the estimation of the dependence of various important acquisition characteristics on circuit’s parameters in the mathematical model in the signal’s phase space by numerical and analytical methods,

which goes back to pioneering works on PLL and considered below in the paper, is widely used [29,37] (see, e.g., modern engineering literature [25–28,36,38–47] and others). Such an approach allows one to analyze effectively the transient processes of signal’s phases adjustment and to estimate the dependence of many important acquisition characteristics on circuit’s parameters by numerical and analytical methods. It is important to note that the construction of mathematical model and the use of results of its analysis for the conclusions on the behavior of the considered physical model are needed for rigorous mathematical foundation [35,48], but it is often ignored in engineering studies. The attempts to justify analytically the reliability of conclusions, based on simplified engineering approaches, and rigorous study of nonlinear models are quite rare (see, e.g., [49–61]).

In the present paper a general effective approach to analytical computation of phase detector characteristic is presented; the classical ideas of analysis and design of PLL-based circuits in the signal’s phase space are developed and rigorously justified; for various non-sinusoidal waveforms of high-frequency signals (see, e.g., various applications of PLL-based circuits with non-sinusoidal signals in [62–68]) its phase-detector characteristics are obtained for the first time and its dynamical model is constructed.

2. Physical model of Costas loop in the signal space

Various modifications of analog and digital Costas loops and PLL with squarer are widely used for BPSK (Binary Phase Shift Keying) and QPSK (Quadrature Phase Shift Keying) demodulation in telecommunication. Because the realization of squaring circuits can be quite difficult, the Costas loop is the preferred variant [26]. In digital circuits, the maximum data rate is limited by a speed of ADC (Analog-to-Digital Converter). In the following classical analog Costas loop, used for BPSK demodulation (similar analysis can also be done for QPSK Costas loop), is considered.

Consider the Costas loop operation (see Fig. 1) with the sinusoidal carrier and VCO (Voltage-Controlled Oscillator) signals with the same frequencies after transient processes. The input signal is BPSK signal, which is a product of the transferred data \( m(t) = \pm 1 \) and the harmonic carrier \( \sin(\omega t) \) with the high frequency \( \omega \). Since here the Costas loop in lock is considered, VCO signal is synchronized with the carrier (i.e. there is no phase difference between VCO signal and input carrier). On the lower branch (Q branch) after the multiplication of VCO signal, shifted by 90°, and the input signal by the multiplier block (\( \odot \)) one has

\[
Q = \frac{1}{2}(m(t) \sin(0) - m(t) \sin(2\omega t)) = -\frac{1}{2} m(t) \sin(2\omega t). \tag{1}
\]

From an engineering point of view, the high-frequency part \( \sin(2\omega t) \) in (1) is removed by a low-pass filter on Q branch. Thus, after filtration a signal on Q branch is zero (a constant in the general case when the initial frequencies are different).

On the upper branch (I) the input signal is multiplied by the output signal of VCO:

\[
I = \frac{1}{2}(m(t) \cos(0) - m(t) \cos(2\omega t)) = \frac{1}{2}(m(t) - m(t) \cos(2\omega t)). \tag{2}
\]

The high-frequency term \( \cos(2\omega t) \) is filtered by a low-pass filter. Thus, on the upper branch I after filtration one can obtain the demodulated data \( m(t) \).

Then both branches are multiplied together and after an additional filtration one gets the signal \( g(t) \) to adjust VCO frequency to the frequency of input carrier signal. After a transient process there is no phase difference and the control input of VCO is zero. In the general case when the initial

---

Fig. 1. Costas loop is in lock: \( m(t) \) is a useful information (\( \pm 1 \)); \( \omega \) is a frequency of input carrier and VCO output; \( m(t) \sin(\omega t) \) is an input signal.
frequencies are different the control input of VCO is constant: 
\[ g(t) \approx \text{const} \]  
(3)

Consider the case (see Fig. 2) when the phase of the input carrier \( \theta_1(t) \) and the phase of VCO \( \theta_2(t) \) are different. The latter means that either (1) the frequencies are different or (2) the frequencies are the same but there is a constant phase difference. For engineers it is a well-known fact [69] that in the considered case of sinusoidal signals (see details in Section 4). This allows one to consider a simplified physical open-loop model of Costas loop (Fig. 3) with only one filter at the input of VCO. In this case the transmitted data \( m(t) \) can be omitted (i.e. \( m(t) \equiv 1 \)) because after the multiplication of the upper and the lower branches at the input of filter the data are squared: \( (m(t)^2 = (\pm 1)^2 = 1) \). Here the signal \( f_1(t) = f_1(\theta_1(t)) \) represents the carrier and \( \theta_1(t) \) represents its phase. In analogy, \( f_2(t) = f_2(\theta_2(t)) \) represents the output signal of VCO, and \( \theta_2(t) \) represents its phase. The functions \( f_{1,2}(\theta) \) are called waveforms.

Consider the analysis of Costas loop for general periodic signal waveforms. Suppose that the waveforms \( f_{1,2}(\theta) \) are bounded 2\( \pi \)-periodic piecewise differentiable functions (this is true for the most considered waveforms, e.g., sinusoidal, squarewave, sawtooth, triangular, and polyharmonic). Consider the following Fourier series representation:

\[ f_p(\theta) = \sum_{i=1}^{\infty} \left( a_i \cos(i\theta) + b_i \sin(i\theta) \right), \quad \theta \geq 0 \]

where \( a_i = \frac{1}{\pi} \int_{-\pi}^{\pi} f_p(\theta) \cos(i\theta) \, d\theta \)

\[ b_i = \frac{1}{\pi} \int_{-\pi}^{\pi} f_p(\theta) \sin(i\theta) \, d\theta \]

and \( p = 1, 2 \). 

The relation between the input \( \varphi(t) \) and the output \( g(t) \) of linear filter is as follows [73]:

\[ g(t) = \alpha_0(t) + \int_0^t \gamma(t-s)\varphi(s) \, ds, \quad (4) \]

where \( \gamma(t) \) is an impulse response function of filter and \( \alpha_0(t) \) is an exponentially damped function depending on the initial state of filter at \( t=0 \).

By (4) the filter output \( g(t) \) has the form

\[ g(t) = \alpha_0(t) + \int_0^t \gamma(t-s)f_1(\theta_1(s))f_2(\theta_2(s)) \, ds. \quad (5) \]

2.1. Simplified model of Costas loop

The low-pass filters on the upper and the lower branches of Costas loop (Fig. 2) are responsible for demodulation process (see Fig. 1) and therefore they can be applied separately from the loop (see, e.g., [4]). From a point of view of the analysis of stability, the filter at the input of VCO executes their filtering functions.

Thus one can consider a simplified physical open-loop model of Costas loop (Fig. 3) with only one filter at the input of VCO. In this case the transmitted data \( m(t) \) can be omitted (i.e. \( m(t) \equiv 1 \)) because after the multiplication of

\[ \text{Fig. 2. Costas loop. Operation in the general case.} \]

\[ \text{Fig. 3. Block diagram of simplified Costas loop.} \]
A high-frequency property of signals can be reformulated in the following way. Suppose that for the frequencies
\[ \omega_{1,2}(t) = \dot{\theta}_{1,2}(t), \]
there exist a sufficiently large number \( \omega_{0\text{min}} \) such that on a fixed time interval \([0, T]\), where \( T \) is independent of \( \omega_{0\text{min}} \), the conditions
\[ \omega_{1,2}(t) \geq \omega_{0\text{min}} > 0 \]
are satisfied. The frequency difference is assumed to be uniformly bounded:
\[ |\omega(t) - \omega_0(t)| \leq \omega_{0\text{max}}^\prime, \quad \forall t \in [0, T]. \]

Requirements (7) and (8) are obviously satisfied for the tuning of two high-frequency oscillators with close frequencies. Let us introduce 
\[ \delta = \omega_{0\text{min}}^{-1/2}. \]
Consider the relations
\[ |\omega(t) - \omega_0(t)| \leq \Delta, \quad \forall t, \quad \tau \in [0, T], \]
where \( \Omega \) is independent of \( \delta \). Conditions (7)–(9) mean that the functions \( \omega_0(t) \) are almost constant and the functions \( f_\tau(\theta_\tau(t)) \) are rapidly oscillating time functions on the small intervals \([\tau, t + \delta]\).

To study the filtration of high-frequency signals by filter (4) it is assumed that the impulse response function of filter is a differentiable function with a bounded derivative (this is true for the most considered filters [73]). The boundedness of derivative of \( \gamma(t) \) implies that
\[ |\gamma(t) - \gamma(t)| = O(\delta), \quad |t - \tau| \leq \delta, \quad \forall t, \quad \tau \in [0, T]. \]

3. Analytical computation of phase detector characteristics for non-sinusoidal waveforms

Consider a block diagram in Fig. 4. Here PD is a non-linear block (describing the operation of all intermediate elements in Fig. 3 between inputs and filter) and its output is 2\( \pi \)-periodic function \( \varphi(\theta_2(t) - \theta_1(t)) \) (the phase detector characteristic of Costas loop); \( G(t) \) is the output of filter.

Suppose that the characteristics and the initial state of filters in Fig. 3 and in Fig. 4 coincide. By (4) the output has the form
\[ G(t) = \alpha_0(t) + \int_0^t \gamma(t - \tau) \varphi(\theta_2(\tau) - \theta_1(\tau)) \, d\tau. \]

**Theorem 1.** Consider 2\( \pi \)-periodic function \( \varphi(0) \) of the form
\[ \varphi(0) = \frac{A^2}{4} + \frac{1}{2} \sum_{m=1}^{\infty} \left( A^2 + B^2 \right) \cos(\theta) \]
\[ + \left( A^2 - B^2 \right) \sin(\theta), \]
where the coefficients \( A^2 \) and \( B^2 \) are expressed via the
\[ A^2 = \frac{a^2}{2}, \quad B^2 = \frac{b^2}{2}, \]
and \( \alpha_2 = 4p + 1, \quad \beta_2 = 4p + 2. \]

**Remark 1.** Since \( f_{1,2}(\theta) \) are piecewise-differentiable, then
\[ A^2 = O\left( \frac{1}{k} \right), \quad B^2 = O\left( \frac{1}{k} \right), \]
and \( \varphi(\theta) \) is a smooth function.

**Remark 2.** For the most considered waveforms, infinite series (12) can be truncated up to the first \( \sqrt{\omega_{0\text{min}}} \) terms. By (13) and (16), the remainder \( R_{1,2}(\omega) \) of series (12) can be estimated as
\[ |R_{1,2}(\omega)| \leq O\left( \sum_{i=1}^{\infty} \frac{1}{i^2} \right) \leq O(\delta). \]

![Fig. 4. Phase detector (PD) of Costas loop and filter.](image-url)
The theorem allows one to compute a phase detector characteristic\textsuperscript{3,4} for the following typical signals given below in the table.

4. Description of classical Costas loop in the signal’s phase space

From a mathematical point of view, linear filter (4) can also be described by a system of linear differential equations
\[
\dot{x} = Ax + b\phi(t), \quad \sigma = c^\top x, \tag{17}
\]
a solution of which has the form (4). Here A is a constant matrix, x(t) is the state vector of Filter, b and c are constant vectors. The model of VCO is usually assumed to be linear:
\[
\dot{\theta}_2(t) = \omega_{\text{free}} + Lc^\top x(t), \tag{18}
\]
Nonautonomous system (19) describes physical model of Costas loop in the signal space (see Fig. 5) and is rather difficult for the study.

Suppose that the frequency of reference signal is a constant
\[
\dot{\theta}_1(t) = \omega_1. \tag{20}
\]
Then Theorem 1 allows one to consider more simple autonomous system of differential equations (in place of nonautonomous (19)), which describes the mathematical model of Costas loop in the signal’s phase space:
\[
x = Ax + b\phi_\Delta(t), \quad \theta_\Delta = \omega_{\text{free}} - \omega_1 + Lc^\top x, \tag{21}
\]
where $\omega_{\text{free}}$ is the free-running frequency of VCO and L is the gain of VCO. Similarly, one can consider various nonlinear models of VCO (see, e.g., [37]). Therefore the initial VCO frequency is as follows:
\[
\dot{\theta}_2(0) = \omega_{\text{free}} + Lc^\top x(0).
\]

By equations of filter (17) and VCO (18) one has
\[
x = Ax + b\phi_\Delta(t) = f_1(\theta_1(t))f_2(\theta_2(t))f_1(\theta_1(t))f_2\left(\theta_2(t) - \frac{\pi}{2}\right),
\]
Nonautonomous system (19) describes physical model of Costas loop in the signal space (see Fig. 5) and is rather difficult for the study.

Suppose that the frequency of reference signal is a constant
\[
\dot{\theta}_1(t) = \omega_1. \tag{20}
\]
Then Theorem 1 allows one to consider more simple autonomous system of differential equations (in place of nonautonomous (19)), which describes the mathematical model of Costas loop in the signal’s phase space:
\[
x = Ax + b\phi_\Delta(t), \quad \theta_\Delta = \omega_{\text{free}} - \omega_1 + Lc^\top x, \tag{21}
\]
where $\omega_{\text{free}}$ is the free-running frequency of VCO and L is the gain of VCO. Similarly, one can consider various nonlinear models of VCO (see, e.g., [37]). Therefore the initial VCO frequency is as follows:
\[
\dot{\theta}_2(0) = \omega_{\text{free}} + Lc^\top x(0).
\]

By equations of filter (17) and VCO (18) one has
\[
x = Ax + b\phi_\Delta(t) = f_1(\theta_1(t))f_2(\theta_2(t))f_1(\theta_1(t))f_2\left(\theta_2(t) - \frac{\pi}{2}\right),
\]
Nonautonomous system (19) describes physical model of Costas loop in the signal space (see Fig. 5) and is rather difficult for the study.

Suppose that the frequency of reference signal is a constant
\[
\dot{\theta}_1(t) = \omega_1. \tag{20}
\]
Then Theorem 1 allows one to consider more simple autonomous system of differential equations (in place of nonautonomous (19)), which describes the mathematical model of Costas loop in the signal’s phase space:
\[
x = Ax + b\phi_\Delta(t), \quad \theta_\Delta = \omega_{\text{free}} - \omega_1 + Lc^\top x, \tag{21}
\]

\[\text{where } \omega_{\text{free}} \text{ is the free-running frequency of VCO and } L \text{ is the gain of VCO. Similarly, one can consider various nonlinear models of VCO (see, e.g., [37]). Therefore the initial VCO frequency is as follows:}
\[
\dot{\theta}_2(0) = \omega_{\text{free}} + Lc^\top x(0).
\]

By equations of filter (17) and VCO (18) one has
\[
x = Ax + b\phi_\Delta(t) = f_1(\theta_1(t))f_2(\theta_2(t))f_1(\theta_1(t))f_2\left(\theta_2(t) - \frac{\pi}{2}\right),
\]
Nonautonomous system (19) describes physical model of Costas loop in the signal space (see Fig. 5) and is rather difficult for the study.

Suppose that the frequency of reference signal is a constant
\[
\dot{\theta}_1(t) = \omega_1. \tag{20}
\]
Then Theorem 1 allows one to consider more simple autonomous system of differential equations (in place of nonautonomous (19)), which describes the mathematical model of Costas loop in the signal’s phase space:
\[
x = Ax + b\phi_\Delta(t), \quad \theta_\Delta = \omega_{\text{free}} - \omega_1 + Lc^\top x, \tag{21}
\]

\[\text{where } \omega_{\text{free}} \text{ is the free-running frequency of VCO and } L \text{ is the gain of VCO. Similarly, one can consider various nonlinear models of VCO (see, e.g., [37]). Therefore the initial VCO frequency is as follows:}
\[
\dot{\theta}_2(0) = \omega_{\text{free}} + Lc^\top x(0).
\]

By equations of filter (17) and VCO (18) one has
\[
x = Ax + b\phi_\Delta(t) = f_1(\theta_1(t))f_2(\theta_2(t))f_1(\theta_1(t))f_2\left(\theta_2(t) - \frac{\pi}{2}\right),
\]
Nonautonomous system (19) describes physical model of Costas loop in the signal space (see Fig. 5) and is rather difficult for the study.

Suppose that the frequency of reference signal is a constant
\[
\dot{\theta}_1(t) = \omega_1. \tag{20}
\]
Then Theorem 1 allows one to consider more simple autonomous system of differential equations (in place of nonautonomous (19)), which describes the mathematical model of Costas loop in the signal’s phase space:
\[
x = Ax + b\phi_\Delta(t), \quad \theta_\Delta = \omega_{\text{free}} - \omega_1 + Lc^\top x, \tag{21}
\]
pendulum-like systems can be applied (see, e.g., [79,54,80,56,81,60]). Modification of direct Lyapunov method with the construction of periodic Lyapunov-like functions, the method of positively invariant cone grids, and the method of nonlocal reduction turned out to be most effective [82,83,80,81]. The last method, which combines the elements of direct Lyapunov method and bifurcation theory, allows one to extend the classical results of Tricomi [84] and his progenies to the multidimensional dynamical systems [83,85].

5. Simulation of Costas loop

Since in the block diagram in Fig. 6 and system (21) only slow time change of signal’s phases and frequencies is considered, they can be effectively studied numerically.

For the simulation of system (21) with the function \( \psi(\cdot) \) of the form (12), in place of conditions (8) and (10) the conditions \(|\alpha_{\text{min}}| \ll \lambda_1 \ll |\lambda_2| \ll \alpha_{\text{min}} \) should be considered, where \( \lambda_1 \) is the largest (in modulus) eigenvalue of matrix \( A \). Also, it is necessary to consider \( T \ll \alpha_{\text{min}} \) to justify the transition from Eqs. (35)–(39) (see Appendix) and to use Remark 1.

The considered theoretical results are justified by the simulation\(^{13} \) of those considered in the previous section of Costas loop models in the signal and signal’s phase spaces. In Fig. 7 are shown the transient processes of VCO input in block diagrams in Figs. 5 and 6 (here it is important if and when VCO input becomes a constant, see (3)).

Here the simulation in the signal’s phase space is more than 100 times faster. Unlike the filter output in the signal’s phase space, in the signal space the filter output contains additional high-frequency oscillation. These high-frequency oscillations interfere with qualitative analysis and efficient simulation of Costas loop. The passage to the analysis of autonomous dynamical model of Costas loop (in place of the nonautonomous one) allows one to overcome the difficulties that relate to the analysis of Costas loop in the signal space.

The approach described can be adapted to digital Costas loops [4,88], where the filter and the VCO are digital unlike those in Fig. 5. If a discretization step is sufficiently small, then a digital filter acts similar to an analog filter. In this case it can be shown that the considered mathematical model in the signal’s phase space is adequate (see Fig. 8).

In conclusion it should be remarked that in the signal’s phase space a similar numerical simulation of the whole transient process (around 10 s – see Fig. 7) of Costas loop with frequencies around 1 GHz = 10^11 Hz takes less than 1 s. At the same time to perform the accurate simulation of Costas loop in the signal space for such frequencies one has to use a discretization step much less than 10^-9, what results in a very long simulation time to oversee the whole transient process: during 10 s of simulation in the signal space.

\(^{8} \) One can compare the numerical integration of systems (19) and (21) with the simulation of realization of block diagrams in Figs. 5 and 6, f.e., in Matlab Simulink (see, e.g., [28,86] and the patent application [87]).

\(^{9} \) While we consider a very simple filter, sawtooth and triangle waveforms in simulation, one can consider similar effects for lag-lead or PI filters and other waveforms.

---

\(^{5} \) Note that to consider one-dimensional stability domains, e.g., defined by \( \alpha_{\text{min}} = \omega_{\text{min}} - \omega_{\text{osc}} \), one has to assume that \( \mathbf{c}^T \mathbf{x}(0) = 0 \). In the general case one has to consider multi-dimensional stability domains taking into account the initial state of loop filter – vector \( \mathbf{x}(0) \).

\(^{6} \) Note that the derivation of dynamical model (21) and the rigorous justification of its adequacy for the analysis of stability are possible here only under condition (20), while formula (12) is obtained for PD characteristic in the general case without condition (20).

\(^{7} \) See also counterexamples to the filter hypothesis, Aizerman’s and Kalman’s conjectures on the absolute stability of nonlinear control systems [77], and the Perron effects of the largest Lyapunov exponent sign inversions [78], etc.

---

5. Simulation of Costas loop

Since in the block diagram in Fig. 6 and system (21) only slow time change of signal’s phases and frequencies is considered, they can be effectively studied numerically.

For the simulation of system (21) with the function \( \psi(\cdot) \) of the form (12), in place of conditions (8) and (10) the conditions \( |\alpha_{\text{min}}| \ll \alpha_{\text{min}} \ll |\lambda_1| \ll \alpha_{\text{min}} \) should be considered, where \( \lambda_1 \) is the largest (in modulus) eigenvalue of matrix \( A \). Also, it is necessary to consider \( T \ll \alpha_{\text{min}} \) to justify the transition from Eqs. (35)–(39) (see Appendix) and to use Remark 1.

The considered theoretical results are justified by the simulation\(^{13} \) of those considered in the previous section of Costas loop models in the signal and signal’s phase spaces. In Fig. 7 are shown the transient processes of VCO input in block diagrams in Figs. 5 and 6 (here it is important if and when VCO input becomes a constant, see (3)).

Here the simulation in the signal’s phase space is more than 100 times faster. Unlike the filter output in the signal’s phase space, in the signal space the filter output contains additional high-frequency oscillation. These high-frequency oscillations interfere with qualitative analysis and efficient simulation of Costas loop. The passage to the analysis of autonomous dynamical model of Costas loop (in place of the nonautonomous one) allows one to overcome the difficulties that relate to the analysis of Costas loop in the signal space.

The approach described can be adapted to digital Costas loops [4,88], where the filter and the VCO are digital unlike those in Fig. 5. If a discretization step is sufficiently small, then a digital filter acts similar to an analog filter. In this case it can be shown that the considered mathematical model in the signal’s phase space is adequate (see Fig. 8).

In conclusion it should be remarked that in the signal’s phase space a similar numerical simulation of the whole transient process (around 10 s – see Fig. 7) of Costas loop with frequencies around 1 GHz = 10^11 Hz takes less than 1 s. At the same time to perform the accurate simulation of Costas loop in the signal space for such frequencies one has to use a discretization step much less than 10^-9, what results in a very long simulation time to oversee the whole transient process: during 10 s of simulation in the signal space.

\(^{8} \) One can compare the numerical integration of systems (19) and (21) with the simulation of realization of block diagrams in Figs. 5 and 6, f.e., in Matlab Simulink (see, e.g., [28,86] and the patent application [87]).

\(^{9} \) While we consider a very simple filter, sawtooth and triangle waveforms in simulation, one can consider similar effects for lag-lead or PI filters and other waveforms.
domain voltage/current waveforms is often difficult and inaccurate.”

6. Conclusion

The approach, proposed in this paper, allows one to compute analytically PD characteristics for the general case of periodic waveforms and to construct the nonlinear mathematical model in the signal’s phase space for classical Costas loop and, ultimately, to apply numerical simulation and analytical methods (various modifications of absolute stability criteria for pendulum-like systems) for the effective analysis of its stability in the large.

Acknowledgment

This work was supported by Saint Petersburg State University, Russian Foundation of Basic Research (Russia), and the Academy of Finland. The authors would like to thank R.E. Best (Best Engineering company, Oberwil, Switzerland) for valuable discussions on practical implementations of Costas loops.

Appendix A

A.1. Averaging method

Application of averaging methods [75,76,89,90] requires the consideration of constant data signal and constant frequency of input carrier (20):

$$\theta_1(t) = \omega_1 t + \theta_1(0).$$

In this case (19) is equivalent to

$$\dot{x} = \Lambda x + b f_1(\omega_1 t + \theta_1(0)) f_2(\omega_1 t + \theta_1(0) + \theta_\Delta)$$

$$f_1(\omega_1 t + \theta_1(0)) f_2(\omega_1 t + \theta_1(0) + \theta_\Delta - \frac{\pi}{2}),$$

$$\theta_\Delta = \omega_\Delta + L(c^\tau x).$$

(22)

Assuming that input carrier is a high-frequency signal ($\omega_1$ is large), one can consider small parameter

$$\varepsilon = \frac{1}{\omega_1}\quad (23)$$

Denote

$$\tau = \omega_1 t.\quad (24)$$

Then system (22) can be transformed as

$$\frac{dx}{d\tau} = \varepsilon(\Lambda x + b f_1(\tau + \theta_1(0)) f_2(\tau + \theta_1(0) + \theta_\Delta)$$

$$f_1(\tau + \theta_1(0)) f_2(\tau + \theta_1(0) + \theta_\Delta - \frac{\pi}{2}),$$

$$\frac{d\theta_\Delta}{d\tau} = \varepsilon(\omega_\Delta + L(c^\tau x)),\quad (25)$$

and represented in such a way that

$$\frac{dz}{d\tau} = \varepsilon F(z, \tau),\quad z = (x, \theta_\Delta)^\tau.\quad (26)$$

In the classical averaging theory such a form of system is

space only $3.5 \times 10^{-8}$ s of 10 s of transient process was obtained (see Fig. 9).

These difficulties are described in [37]: “Direct time-domain simulation of PLLs at the level of SPICE circuits is typically impractical because of its great inefficiency. PLL transients can last hundreds of thousands of cycles, with each cycle requiring hundreds of small time steps for accurate simulation of the embedded voltage-controlled oscillator (VCO). Furthermore, extracting phase or frequency information, one of the chief metrics of PLL performance, from time-

Fig. 7. $f_{\text{rev}} = 100.5$ Hz, $f_1 = 100$ Hz, $L = 27$, filter transfer functions $1/(s + 2)$, sawtooth and triangle waveforms.

Fig. 8. $f_{\text{rev}} = 105$ Hz, $f_1 = 100$ Hz, $L = 100$, analog filter transfer function $1/(0.1 s + 1)$, discrete filter in $z$-domain $10/(1 - \exp^{-10}z^{-1})$, sample time $10^{-3}$, sawtooth and triangle waveforms.

Fig. 9. $f_{\text{rev}} = 1000.5$ Hz, $f_1 = 1000$ Hz, $L = 27$, filter transfer functions $1/(s + 2)$, sawtooth and triangle waveforms.
called a standard form. Consider an averaged equation
\[
\frac{dy}{d\tau} = \varepsilon F(y),
\]
where
\[
F(y) = \frac{1}{\varepsilon} \int_0^\rho F(y, \tau) \, d\tau.
\]
Suppose \( D \) is a bounded domain containing the point \( z_0 = (x(0), \theta(0)) \). Consider solutions \( z(\tau, \varepsilon) \) and \( y(\tau, \varepsilon) \) with the initial data \( z_0 = y_0 \). In this case there exists a constant \( T \) such that \( z(\tau, \varepsilon) \) and \( y(\tau, \varepsilon) \) remain in the domain \( D \) for \( 0 \leq \tau \leq T/\varepsilon \). Define \( \rho_{\max} = 1/\rho_{\min} \).

Consider system of differential equations
\[
\frac{dz}{d\tau} = F(z, z, \lambda),
\]
where \( z, F \) are points of \( n \)-dimensional Euclidean space \( E_n \), \( \lambda \) is a parameter.

Let function \( F(\tau, z, \lambda) \) be real measurable function in \( \tau \in [0, T/\varepsilon], z \in D \) for any \( \lambda \in \Lambda \), \( \lambda_0 \in \Lambda \), where \( \Lambda \) is a domain of \( E_n \).

**Theorem 2.** Consider the following system:
\[
\frac{dz}{d\tau} = \varepsilon F(\tau, z).
\]
Assume that the right-hand side \( F(\tau, z) \) is uniformly bounded and integrals
\[
\int_0^\tau \int_c^\tau F(\tau, x) \, dx \, d\tau, \quad 0 \leq \tau < \infty, \quad y \in D
\]
are smooth for any fixed \( c \leq D; \) the following limit
\[
\lim_{t \to \infty} \int_0^\tau F(\tau, z) \, d\tau = \mathcal{F}(z)
\]
exists (uniformly with respect to \( z \in D \)); uniformly with respect to \( y, r_m \):
\[
\lim_{t \to \infty} \int_0^\tau \int_0^r F_\delta(\tau, z_1, \ldots, z_n) \, d\tau \, dx = 0,
\]
where \( r_m \) is a decreasing sequence \( r_m \to 0 \) while \( m \to \infty \), and \( F_\delta = F(\tau, z) - F(z) ; \mathcal{F}(z) \) are \( k \)-Lipschitz functions; solution \( y(\tau) \) of the averaged equation \((27)\) for any \( \tau \) from \( 0 \leq \tau < \infty \) belongs to the domain \( D \) together with its \( \rho \)-neighborhood, and solution of the \((30)\) with initial conditions \( z(0) = y(0) \) is unique.

Then for any \( \eta > 0 \), \( T > 0 \) there is an \( \varepsilon_0 > 0 \), such that for \( 0 < \varepsilon < \varepsilon_0 \) the solution \( z(\tau) \) of \((30)\) satisfies
\[
|z(\tau) - y(\tau)| < \eta, \quad \tau \in \left[ 0, \frac{T}{\varepsilon} \right]
\]

**A.2. Proof of the main theorem**

Suppose, \( t \in [0, T] \). Consider a difference
\[
g(t) - G(t) = \int_0^t \gamma(s) \left[ f_1(\theta_1(s)) f_2(\theta_2(s)) - \varphi(\theta_2(s) - \theta_1(s)) \right] ds.
\]
Let \( m \in \mathbb{N} \) such that \( t \in [m\delta, (m+1)\delta] \). By definition of \( \delta \),
one has \( m < T/\delta + 1 \). The continuity condition implies that \( \gamma(t) \) is bounded on \([0, T]\) and \( f_1(\theta_1), f_2(\theta_2) \) are bounded on \( \mathbb{R} \). Since \( f_{1,2}(\theta_1) \) are piecewise differentiable, one gets
\[
\begin{align*}
\int_0^{(m+1)\delta} \gamma(t-s) f_1(\theta_1(s)) f_2(\theta_2(s)) \, ds & = O(\delta), \\
\int_0^{(m+1)\delta} \gamma(t-s) \varphi(\theta_2(s) - \theta_1(s)) \, ds & = O(\delta).
\end{align*}
\]
It follows that \((35)\) can be represented as
\[
g(t) - G(t) = \sum_{k=0}^m \gamma(t-k\delta) - \varphi(\theta_2(s) - \theta_1(s)) \, ds + O(\delta).
\]
Prove now that on each interval \([k\delta, (k+1)\delta)\) the corresponding integrals are equal to \( O(\delta^2) \).

Condition \((10)\) implies that on each interval \([k\delta, (k+1)\delta)\) the following relation
\[
\gamma(t-s) = \gamma(t-k\delta) + O(\delta), \quad t > s, \; s, t \in [k\delta, (k+1)\delta)
\]
is satisfied. Here \( O(\delta) \) is independent of \( k \) and the relation is satisfied uniformly with respect to \( t \). By \((37), (38)\), and the boundedness of functions \( f_1(\theta_1), f_2(\theta_2), \varphi(\theta_1) \) it can be obtained that
\[
g(t) - G(t) = \sum_{k=0}^m \gamma(t-k\delta)
\]
\[
\int_{[k\delta,(k+1)\delta]} \left[ f_1(\theta_1(s)) f_2(\theta_2(s)) f_1(\theta_1(s)) f_2(\theta_2(s) - \frac{\pi}{2}) - \varphi(\theta_2(s) - \theta_1(s)) \right] ds + O(\delta).
\]
Denote
\[
\theta_k(s) = \theta_k(s) + \theta_0(s) - \theta_k(s) = \theta_0(s) + O(\delta).
\]
Then for \( s \in [k\delta, (k+1)\delta) \), condition \((9)\) yields
\[
\theta_k(s) = \theta_k^T(s) + O(\delta).
\]
From \((8)\) and the boundedness of \( \varphi(\theta_1) \) on \( \mathbb{R} \) it follows that
\[
\int_{[k\delta,(k+1)\delta]} \left[ \varphi(\theta_2(s) - \theta_1(s)) - \varphi(\theta_0(s) - \theta_1(s)) \right] ds = O(\delta^2).
\]
If \( f_1(\theta_1) \) and \( f_2(\theta_2) \) are continuous on \( \mathbb{R} \), then for \( \int_{[k\delta,(k+1)\delta]} f_1(\theta_1(s)) f_2(\theta_2(s)) f_1(\theta_1(s)) f_2(\theta_2(s) - \frac{\pi}{2}) \) the relation
\[
\int_{[k\delta,(k+1)\delta]} f_1(\theta_1(s)) f_2(\theta_2(s)) f_1(\theta_1(s)) f_2(\theta_2(s) - \frac{\pi}{2}) \, ds
\]
\[
= \int_{[k\delta,(k+1)\delta]} f_1(\theta_1^T(s)) f_2(\theta_2^T(s))
\]
\[
+ f_1(\theta_1^T(s)) f_2(\theta_2^T(s) - \frac{\pi}{2}) \, ds + O(\delta^2)
\]
is satisfied. Consider why this estimate is valid for the considered class of piecewise differentiable waveforms. Since conditions \((7)\) and \((9)\) are satisfied and the functions \( \theta_{1,2}(s) \) are differentiable and satisfy \((8)\), for all \( k = 0, \ldots, m \) there exist sets \( E_k \) (the union of sufficiently small neighborhoods of discontinuity points of \( f_{1,2}(t) \)) such that the relation \( \int E_k \, ds = O(\delta^2) \) is valid, in which case this relation is satisfied uniformly with
respect to $k$. Then the piecewise differentiability and boundedness of $f_{1,2}(\vartheta)$ imply relation (43).

By (42) and (43), relation (39) can be rewritten as

$$g(t) - G(t) = m \sum_{k=0}^{\infty} \gamma(t - k\delta) \int_{[k\delta,(k+1)\delta]} \left[ \left( \sum_{j=1}^{n} a_j^1 \cos (j\theta_1(s)) + b_j^1 \sin (j\theta_1(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^2 \cos (j\theta_2(s)) + b_j^2 \sin (j\theta_2(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^3 \cos (j\theta_3(s)) + b_j^3 \sin (j\theta_3(s)) \right) \right. $n \left. - \varphi(q_1(s) - \varphi(q_1(s))) \right] ds + O(\delta).$$

By (44)

$$g(t) - G(t) = m \sum_{k=0}^{\infty} \gamma(t - k\delta) \int_{[k\delta,(k+1)\delta]} \left[ \left( \sum_{j=1}^{n} a_j^1 \cos (j\theta_1(s)) + b_j^1 \sin (j\theta_1(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^2 \cos (j\theta_2(s)) + b_j^2 \sin (j\theta_2(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^3 \cos (j\theta_3(s)) + b_j^3 \sin (j\theta_3(s)) \right) \right. $n \left. - \varphi(q_1(s) - \varphi(q_1(s))) \right] ds + O(\delta).$$

Since conditions (7)-(9) are satisfied, it is possible to choose $O(1/\delta)$ the sufficiently small time intervals of length $O(\delta)$ such that outside this interval the functions $f_1(\theta_1(t))$ and $f_2(\theta_2(t))$ are continuous.

It is known that on each interval, there is no discontinuity points, Fourier series of functions $f_1(\theta)$ and $f_2(\theta)$ converge uniformly. Then there exists a number $M = M(\delta) > 0$ such that outside sufficiently small neighborhoods of discontinuity points of $f_1(\theta_1(t))$ and $f_2(\theta_2(t))$ the sum of the first $M$ terms of series approximates the original function with an accuracy to $O(\delta)$. In this case by relation (45) and the boundedness of $f_1(\theta)$ and $f_2(\theta)$ on $\mathbb{R}$ it can be obtained:

$$g(t) - G(t) = m \sum_{k=0}^{\infty} \gamma(t - k\delta) \int_{[k\delta,(k+1)\delta]} \left[ f_1(\theta_1(s)) f_2(\theta_2(s)) f_1(\theta_1(s)) f_2(\theta_2(s) - \frac{\pi}{2}) \right. $n \left. - \varphi(q_1(s) - \varphi(q_1(s))) \right] ds + O(\delta) = m \sum_{k=0}^{\infty} \gamma(t - k\delta) \int_{[k\delta,(k+1)\delta]} \left[ \left( \sum_{j=1}^{n} a_j^1 \cos (j\theta_1(s)) + b_j^1 \sin (j\theta_1(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^2 \cos (j\theta_2(s)) + b_j^2 \sin (j\theta_2(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^3 \cos (j\theta_3(s)) + b_j^3 \sin (j\theta_3(s)) \right) \right. $n \left. - \varphi(q_1(s) - \varphi(q_1(s))) \right] ds + O(\delta).$$

Thus,

$$g(t) - G(t) = m \sum_{k=0}^{\infty} \gamma(t - k\delta) \int_{[k\delta,(k+1)\delta]} \left[ \left( \sum_{j=1}^{n} a_j^1 \cos (j\theta_1(s)) + b_j^1 \sin (j\theta_1(s)) \right) \right. $n \left. + \left( \sum_{j=1}^{n} a_j^2 \cos (j\theta_2(s)) + b_j^2 \sin (j\theta_2(s)) \right) \right. $n \left. - \varphi(q_1(s) - \varphi(q_1(s))) \right] ds + O(\delta).$$

(47)

Remark that the addends in (47) consist of the product of four coefficients and four trigonometric functions. Apply the formulas of product of sines and cosines to each addend and use Lemma 1 (assertions of Lemmas 1 and 2 are at the end of Appendix), taking into account the conditions of high-frequency property (7)-(9) and the introduced notion (40). Note that the addends in (47) are similar and the types of functions (sin or cos) and coefficients ($a, b, \alpha, \beta$) in no way affect the conclusion of Lemma 1 and its proof. Consider, for example, the following addend of sum (47):

$$\sum_{j,k,l,r=1}^{M} a_j^1 \cos (j\theta_1(s)) a_k^1 \cos (k\theta_1(s))$$

$$a_l^2 \cos (l\theta_2(s)) a_r^2 \cos (r\theta_2(s)).$$

(48)

By the relation

$$\cos (\theta_1) \cos (\theta_2) = \frac{1}{2} \left( \cos (\theta_1 + \theta_2) + \cos (\theta_1 - \theta_2) \right)$$

(49)

one obtains

$$S = \sum_{j,k,l,r=1}^{M} a_j^1 \cos (j\theta_1(s)) a_k^1 \cos (k\theta_1(s)) a_l^2 \cos (l\theta_2(s)) a_r^2 \cos (r\theta_2(s))$$

$$= \frac{1}{8} \left( \cos ((i-j)\theta_1(s)) \right)$$

$$+ \cos ((i+j)\theta_1(s)) \theta_2(s) + \cos ((i-j)\theta_1(s)) \theta_2(s) + \cos ((i+j)\theta_1(s)) \theta_2(s)$$

$$+ \cos ((i-j)\theta_1(s)) \theta_2(s) + \cos ((i+j)\theta_1(s)) \theta_2(s)$$

$$+ \cos ((i-j)\theta_1(s)) \theta_2(s) + \cos ((i+j)\theta_1(s)) \theta_2(s)$$

$$+ \cos ((i-j)\theta_1(s)) \theta_2(s) + \cos ((i+j)\theta_1(s)) \theta_2(s).$$

(50)

Consider an integral of this expression over the interval $[k\delta,(k+1)\delta]$. By Lemma 1 and (40) one has

$$\int_{[k\delta,(k+1)\delta]} \cos (\theta_2(s)) ds = O(\delta^p),$$

$p = 1, 2$.  


The use of relations (7) gives the estimate
\[
\int_{[k\delta,(k+1)\delta]} a_2^* \cos\left(j\theta_2^*(s)\right) ds = O\left(\delta^2\right), \quad p = 1, 2. \tag{51}
\]
Then for the integral of the first addend of (50) over the interval \([k\delta,(k+1)\delta]\) one obtains
\[
\int_{[k\delta,(k+1)\delta]} \sum_{j,l,j,l} \frac{a_1^*a_1^*a_2^*a_2^*}{8} \cos\left((i+j)\theta_1^*(s) + (l+r)\theta_2^*(s)\right) ds
= \sum_{j,l,j,l} O\left(\delta^2\right) \frac{1}{ijlr \max(i+j,l+r)} \tag{52}
\]
Since the series \(\sum_{i,j,l,r=1}^\infty 1/ijlr \max(i+j,l+r)\) converges \((i+j \geq 2\sqrt{ij} \text{ and } l+r \geq 2\sqrt{ir})\), the integration over (50) gives
\[
\int_{[k\delta,(k+1)\delta]} S ds = \int_{[k\delta,(k+1)\delta]} \sum_{j,l,j,l} \frac{a_1^*a_1^*a_2^*a_2^*}{8} \cos\left((i+j)\theta_1^*(s) + (l+r)\theta_2^*(s)\right) ds + O\left(\delta^2\right). \tag{53}
\]
From (8) and Lemma 2 (see below) it follows that
\[
\int_{[k\delta,(k+1)\delta]} \sum_{j,l,j,l} \frac{a_1^*a_1^*a_2^*a_2^*}{8} \cos\left((i+j)\theta_1^*(s) + (l+r)\theta_2^*(s)\right) ds = \sum_{j,l,j,l} O\left(1/ijlr(i+j+l+r)\right) = O\left(\delta^2\right). \tag{54}
\]
Similarly,
\[
\int_{[k\delta,(k+1)\delta]} \sum_{j,l,j,l} \frac{a_1^*a_1^*a_2^*a_2^*}{8} \cos\left((i+j)\theta_1^*(s) - (l+r)\theta_2^*(s)\right) ds = O\left(\delta^2\right), \tag{55}
\]
The rest of the addends in (53) enter into definition (12) of \(\varphi(s)\).

Note that relations (51) and (52) remain true if \(c\) is replaced by \(s\). Then
\[
\int_{[k\delta,(k+1)\delta]} \sum_{j,l,j,l} \frac{a_1^*a_1^*a_2^*a_2^*}{8} \sin\left((i+j)\theta_1^*(s) + (l+r)\theta_2^*(s)\right) ds = O\left(\delta^2\right), \tag{56}
\]
Obviously, the changes from \(a^2\) to \(b^2\) and from \(a^1\) to \(b^1\) remain unchanged relations (56). Thus, some addends from (47) satisfy the relations similar to (56) and the rest of the addends enter into \(\varphi(s)\). Theorem is proved.

**Lemma 1.** For sufficiently large frequencies \(\omega_{\text{min}}\) the following relations hold:
\[
\int_{[k\delta,(k+1)\delta]} \cos\left(j(\omega_{\text{min}}s + \varphi)\right) ds = O\left(\delta^2\right), \tag{57}
\]
\[
\int_{[k\delta,(k+1)\delta]} \sin\left(j(\omega_{\text{min}}s + \varphi)\right) ds = O\left(\delta^2\right), \quad j \in \mathbb{N}, \quad k \in \mathbb{N}_0.
\]
where \(\delta^2 = \omega_{\text{min}}^{-1}\), are satisfied.

**Lemma 2.** The series \(\sum_{i,j,l,r=1}^\infty 1/ijlr(i+j+l+r)\) converge.

References
