Doctoral Thesis

An optical phase locked loop for coherent space communications

Author[s]:
Herzog, Frank Thomas

Publication Date:
2006

Permanent Link:
https://doi.org/10.3929/ethz-a-005132944

Rights / License:
In Copyright - Non-Commercial Use Permitted
An optical phase locked loop for coherent space communications

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Sciences

presented by
FRANK THOMAS HERZOG
Dipl. El. Ing. ETH
born 11 January 1975
citizen of Wigoltingen (TG), Switzerland

accepted on the recommendation of
Prof. Dr. W. Bächtold, examiner
Prof. Dr. H. Bölcskei, Dr. K. Kudielka, co-examiners

2006
# Contents

Abstract vii

Zusammenfassung ix

1 Introduction 1
   1.1 Focus ..................................................... 3
   1.2 Structure ................................................. 4

2 Optical free-space data transmission 5
   2.1 Optical vs. microwave transmission ....................... 6
   2.2 Transmitters .............................................. 7
      2.2.1 Laser linewidth ......................................... 8
      2.2.2 Modulation ............................................. 9
   2.3 Receivers ................................................... 10
      2.3.1 The quantum limit ....................................... 11
      2.3.2 Direct detection ......................................... 12
      2.3.3 Coherent detection ..................................... 13
      2.3.4 Homodyne vs. heterodyne detection .................... 16
      2.3.5 Synchronous vs. asynchronous processing ............. 17
      2.3.6 System comparison ..................................... 17

3 Optical phase locked loops 19
   3.1 Phase locked loops ........................................ 20
      3.1.1 Closed-loop and error transfer functions ............. 20
      3.1.2 Characteristic loop parameters ....................... 22
      3.1.3 Stability ............................................... 24
   3.2 OPLL design ............................................ 25
3.2.1 Balanced loop .................................. 25
3.2.2 Costas/Decision-driven loop .................. 27
3.2.3 SyncBit loop .................................. 28
3.2.4 Dither loop .................................. 28
3.3 System comparison ................................ 29

4 Phase locking by LO phase dithering ................................................................. 35
  4.1 Dither loop analysis .................................. 35
     4.1.1 Loop simulation and verification .............. 42
     4.1.2 Stable operating point for non-zero $\phi_e(t)$ ... 45
  4.2 Noise in a dither loop ................................ 48
     4.2.1 LO phase dithering ........................... 50
     4.2.2 Phase noise .................................. 51
     4.2.3 Shot noise .................................... 53
  4.3 Performance degradation due to noise ......................... 57
     4.3.1 Phase dither power penalty ................... 58
     4.3.2 Phase error power penalty ..................... 61
     4.3.3 Overall power penalty ......................... 63
  4.4 Dither loop synthesis ................................ 66
     4.4.1 Phase error minimization ..................... 66
     4.4.2 Dither loop design rule ....................... 68

5 Receiver measurement ................................................................. 73
  5.1 System design ..................................... 73
  5.2 Communication performance .......................... 75
     5.2.1 Digital transmission ......................... 75
     5.2.2 Analog transmission ......................... 77

6 Conclusions ................................................................. 79
  6.1 Outlook .......................................... 80

A Phase noise evaluation ................................................................. 83

B Iterative computation of $\sigma_E$ and $\epsilon$ ............................................. 85

C Bit error rate test setup ................................................................. 87

Bibliography ................................................................. 89
<table>
<thead>
<tr>
<th>Contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>99</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>101</td>
</tr>
</tbody>
</table>
Abstract

Optical free-space transmission with coherent detection is a promising technology for inter-satellite communications. It allows the use of small antennas and moderate transmitter power to build data links in the Gbit/s range over tens of thousands of km. The present thesis aims at a specific circuit of coherent reception: the optical phase locked loop.

Basic properties of optical reception are reviewed. It is shown that the sensitivity of a shot noise limited coherent receiver exceeds the quantum limited direct detection receiver by 3.5 dB (in terms of the peak optical power). Furthermore, optical heterodyning suffers from an inherent 3 dB power penalty compared to optical homodyning, making the homodyne receiver – in combination with BPSK modulation – the most sensitive in optical detection.

Several phase-locking schemes, suitable for optical communications, are presented and classified by their complexity and user signal requirements. A relatively new design, the dither loop, is identified as being superior to other concepts. In the present thesis, a detailed mathematical analysis of the dither loop is derived for the first time. The contribution of phase dithering, phase noise and shot noise to the total phase error variance is calculated. Power penalties, induced by the above mentioned noise sources, are evaluated. The knowledge gained during the analysis leads to design rules for an optimum dither loop. The design rules are based on general system specifications.

Measurement results of a homodyne receiver, employing a dither loop, are presented. The sensitivity amounts to 36 photons/bit (-55.7 dBm) for the transmission of a PRBS $2^{31} - 1$ signal at a data rate of 400 Mbit/s and a bit error rate of $10^{-9}$. The transmission system uses diode-pumped Nd:YAG lasers at a wavelength of 1.06 µm.
Zusammenfassung


Es werden Messwerte eines Homodynempfängers präsentiert, welcher einen Dither Loop verwendet. Die Empfindlichkeit beträgt 36 Photonen/bit (-55.7 dBm) für die Übertragung eines PRBS $2^{31} - 1$ Signals bei einer Datenrate von 400 Mbit/s und einer Bitfehlerrate von $10^{-9}$. Das Übertragungssystem benutzt diodengepumpte Nd:YAG Laser bei einer Wellenlänge von 1.06 µm.
Chapter 1

Introduction

Since the advent of lasers and single-mode fibers, optical communications has enabled and greatly stimulated the development of global information networks. The unprecedented combination of low loss and nearly unlimited bandwidth of the optical fiber superseded any other form of large-volume terrestrial data transmission like copper coaxial cables, hollow metallic waveguides or microwave radio relays. During the last four decades, it has been argued that optics could be advantageous for space communication applications as well [1–5]. Space communications, as it is meant in the present thesis, refers to the transmission of information between two or more satellites, and not between ground stations to satellites or vice versa.\(^1\)

For obvious reasons, it is not possible to connect satellites with optical fibers or any other form of waveguide. Hence, space communications is inherently based on the free-space propagation of an electromagnetic field.\(^2\) Absorption and dispersion, which are effects of the wave/matter interaction, can be neglected in free-space propagation. Instead, beam spreading losses and severe constraints on the pointing accuracy (i.e. the alignment of the transmitter and receiver antenna in one axis) impede the application of networking technologies in space. Since an antenna of a given size focuses an electromagnetic field stronger at

\(^1\)Although such systems are subject to research as well [6, 7].

\(^2\)In this context, the expression ‘space’ is used ambiguously. Throughout the present thesis, ‘space communications’ shall denote any form of extra-terrestrial data transmission, and ‘free-space’ represents the wave propagation in vacuum.
optical wavelengths than in the microwave regime, a free-space transmission system benefits from optical communications, given that the pointing requirements can be met.

Satellite networks could provide broadband and mobile communication channels from almost every point of the world. The narrow beam widths of an optical transmission prohibit eavesdropping, but support extensive frequency reuse. Another application can be found in the ever-increasing amount of data collected by earth or space observation satellites, which require efficient downlinks to circumvent substantial on-board storage capabilities. Satellites in low earth orbits (LEO) revolve much faster than the earth, resulting in short connection times – and therefore, in a limited average bandwidth – to a designated ground station. Instead of direct downlinks, satellites in geostationary orbits (GEO) could provide relay services, as the time in which they are in a line-of-sight constellation to the LEO satellite is greatly increased. For such an inter-satellite link (ISL), the optical data transmission could provide high bandwidths with reasonably sized antennas.

Arthur C. Clarke, the British inventor and science-fiction author best known for his novel ‘2001: A Space Odyssey’, is credited for the first proposal of geostationary satellites acting as telecommunication relay stations [8, 9]. The first operationally successful satellite network employing inter-satellite links, the Iridium system [10], was installed in 1998 and went into bankruptcy only half a year later. Since then, many satellite network constellations have been proposed, none of which turned into an operational system. Optical ISL terminals were aggressively promoted in the second half of the nineties, when – under the influence of the ever increasing internet bandwidth demand – the congestion of transatlantic networks was projected for the beginning of the next millennium [13] (a less biased analysis is given in [14]). In 2001, the first optical ISL was demonstrated by the European Space Agency (ESA), communicating between the satellites SPOT4 (LEO) and ARTEMIS in a temporary parking orbit (31000 km) [15]. This system, shown in fig. 1.1, achieved a data rate of 50 Mbit/s with diode lasers and a direct detection receiver. Other successful experiments include the classified american GeoLITE mission,

---

3The most ambitious of which, the Teledesic system, was originally planned in 1994 with 840 active satellites (924 including spare satellites), each carrying 8 ISL terminals operating at 60 GHz [11, 12].
1.1. Focus

Figure 1.1: Artist’s view of ARTEMIS (foreground) and SPOT4 (background) communicating with the optical inter-satellite link terminal SILEX (Semiconductor Inter-satellite Link EXperiment). Image: ESA.

and a japanese ground-to-satellite optical link demonstration between an OICETS terminal and ARTEMIS [16]. Recent research on optical space communications focuses on high-speed download links for scientific deep space probes, rather than on satellite networks [17,18]. These systems exploit specific properties of optical amplifiers in combination with a pulse position modulation (PPM) scheme to achieve Mbit/s data rates over inter-planetary distances.

1.1 Focus

The present thesis is concerned with an optical phase locked loop (OPLL), suited for the application in coherent optical space communications. The considered OPLL is believed to be superior over other known designs. The emphasis of the thesis is laid on the communication and control theory aspects, rather than the technical implementation of the loop. The examined phase-locking scheme has been discussed by other authors [19], but no thorough mathematical analysis of the
Chapter 1. Introduction

OPLL has been done so far. The main focus of this thesis is to derive the mathematics, and to synthesize design rules for predictable loop performance results. The terminology and notation is based on the theoretical work of Barry and Lee [20] and Kazovsky [21, 22].

1.2 Structure

The thesis is organized as follows: In Chapter 2, basic properties of optical communications, i.e. the quantum limit and the shot noise limit, are derived. An inherent 3 dB power penalty of optical heterodyning, compared to optical homodyning, is shown. In Chapter 3, the PLL theory needed for the loop analysis is discussed. Moreover, known advanced OPLL designs are explained and compared. The main contribution of the thesis is presented in Chapter 4. An expression for the phase detector gain in the considered loop is extracted. The phase error variance due to several noise sources is calculated, and a design rule for an optimum OPLL is derived. In Chapter 5, measurement results of a real world coherent optical transmission system are presented. The system is characterized by its bit error rate and signal-to-noise ratio, depending on the transmitted optical power. A conclusion and an outlook is given in Chapter 6. Auxiliary material used during the analysis or for measurements is presented in the appendices.
Chapter 2

Optical free-space data transmission

In this chapter, it is reasoned why the optical transmission of information can be advantageous for a long distance free-space inter-satellite link. It will be shown that an antenna of a given physical size focuses an electromagnetic field stronger at optical than at microwave frequencies. This results in an increased power density at the receiving satellite, and thus, in a higher detected power for a given antenna size.

Concerning the optical transmitter, an important characteristic of lasers – the laser linewidth – is established. A short introduction to electro-optical modulators is given. In the following chapters, which focus on the receiver design, it is always assumed (without reference) that the received signal has been generated by a transmitter as described in Section 2.2 and depicted in fig. 2.2 of this chapter.

Optical receiver design will be treated in more detail. An inherent limitation of direct detection, the quantum limit, is introduced. Real world direct detection receivers can never achieve the quantum limit due to various losses and noise processes. Instead, the coherent detection scheme is proposed. It will be shown that a coherent receiver requires 3.5 dB less power than predicted by the quantum limit (in terms of the peak optical power). The highest possible sensitivity of an optical transmission system can be obtained with BPSK modulation and homodyne detection.
Chapter 2. Optical free-space data transmission

![Figure 2.1: Simplified model of a free-space transmission system.](image)

### 2.1 Optical vs. microwave transmission

A simplified free-space information transmission system, depicted in fig. 2.1, could consist of a transmitter with associated power $P_{TX}$ at a wavelength $\lambda$, and a receiver at a distance $r$ away from the transmitter. Then, the power collected by the receiver antenna amounts to

$$P_S = \frac{P_{TX} G_{TX} G_S \lambda^2}{(4\pi)^2 r^2} \quad (2.1)$$

which holds true both for optical [23] and microwave systems [24]. Writing the antenna gain $G$ in terms of the aperture area $A$, i.e.

$$G = \frac{4\pi A}{\lambda^2}, \quad (2.2)$$

yields for the collected power:

$$P_S = \frac{P_{TX} A_{TX} A_S}{\lambda^2 r^2}. \quad (2.3)$$

For equal antennas of diameter $d$ on the transmitter and receiver, and assuming that the aperture area is equal to the physical antenna area, the power ratio from transmitter to receiver results in:

$$\frac{P_S}{P_{TX}} = \left(\frac{\pi}{4} \cdot \frac{d}{\lambda r}\right)^2. \quad (2.4)$$

A system designer has two options to keep the power ratio within reasonable dimensions: transmitting at short wavelengths, or using large antenna areas. The latter is not an option for inter-satellite links, because large antennas consume too much payload and volume on the
Figure 2.2: Transmitter setup for intensity modulation, carrier-suppressed amplitude modulation or BPSK modulation.

launch vehicle. Thus, the wavelength should be chosen as small as possible. Transmitting at optical (i.e. 1 μm or 300 THz) wavelengths instead of microwaves (i.e. 1 cm or 30 GHz) increases the power ratio $P_s/P_{TX}$ by 80 dB. This happens at the cost of far more stringent requirements on the beam steering mechanism and pointing accuracy [25, 26].

### 2.2 Transmitters

Since this thesis focuses on the receiver of an optical transmission system, the transmitter will be covered only superficially. A possible transmitter design is shown in fig. 2.2. It consists of a laser source, followed by an electro-optical Mach-Zehnder (MZ) modulator. A small part of the modulator output signal is fed into a bias control circuit. Depending on the power requirements of the receiver, an optical fiber amplifier (OFA) can be used to boost the transmitter output.

If the modulator is biased in quadrature (i.e. ‘half open’), it performs intensity modulation. For a bias point of total extinction, a carrier suppressed double sideband amplitude modulation is obtained [27]. For a digital input signal with a sufficiently large amplitude, this becomes a binary phase shift keying (BPSK) modulation. An optical phase modulator can be used in case of purely digital data transmission, saving the optical coupler and the bias control circuit.

The following Section 2.2.1 introduces the laser linewidth, which is a measure for the spectral impurity of the laser output. The laser linewidth is used in Chapter 4, when calculating the residual phase error of a coherent receiver. In Section 2.2.2, the design of electro-
optical modulators is briefly discussed.

2.2.1 Laser linewidth

Among the various properties of lasers, only phase noise shall be considered here. Phase noise is usually described in terms of a frequency noise process caused by the spontaneous emission of photons in the laser cavity [20, 28]. The frequency noise process $\mu(t)$ is white with a one-sided PSD of\(^1\)

$$S_\mu(f) = 2 \cdot N_0$$

(2.5)

and units of Hz\(^2\)/Hz = Hz. Since phase is the integral of frequency, the statistical phase fluctuations become

$$\phi(t) = 2\pi \int_0^t \mu(\tau) d\tau,$$

(2.6)

which is sometimes called a Wiener-Lévy process [29]. The laser lightwave perturbed by the phase noise process $\phi(t)$ can be written as

$$x(t) = \sqrt{P_{TX}} \cdot e^{i(2\pi f_0 t + \phi(t) + \theta)},$$

(2.7)

where $P_{TX}$ denotes the output power and $f_0$ the average oscillating frequency. The additional random phase expression $\theta$ in (2.7) has been included to make $x(t)$ wide-sense stationary. It has been shown in [20, 28] that the output spectrum of the laser lightwave $x(t)$ calculates to\(^2\)

$$S_x(f) = \frac{P_{TX}}{\pi^2 N_0} \cdot \frac{1}{1 + \left(\frac{f - f_0}{\pi N_0}\right)^2}.$$

(2.8)

The spectral shape of (2.8), depicted in fig. 2.3, is called ‘Lorentzian’. Solving (2.8) for the frequency where the power density has dropped by 3 dB relative to the maximum, yields

$$\pi N_0 = f_{3\text{dB}} - f_0 = \frac{\Delta \nu}{2} \rightarrow N_0 = \frac{\Delta \nu}{2\pi}$$

(2.9)

\(^1\)Unless otherwise stated, this book uses the one-sided notation (i.e. $0 \leq f \leq \infty$) of power density spectra. The difference to a two-sided notation ($-\infty \leq f \leq \infty$) is simply a multiplicative factor of 2.

\(^2\)Equation (2.8) is not a sole property of lasers, but of oscillators in general [30].
where $\Delta \nu$ (in Hz) is the laser linewidth. The PSD of the phase noise process in dependency of the linewidth amounts to \cite{21, 22}:

$$S_\phi(f) = \frac{\Delta \nu}{\pi f^2}$$

which has the dimension of rad$^2$/Hz. The laser linewidth can be measured relatively easily with heterodyning techniques \cite{31, 32}. Typical linewidths lie in the order of 10 MHz for distributed-feedback (DFB) lasers, 10 kHz for external cavity lasers (ECL) and fiber lasers, and 1 Hz for diode-pumped Nd:YAG lasers.

The output signal of a laser is not only perturbed by phase noise due to white frequency noise, but also due to $1/f$ or higher order frequency noises (‘frequency flicker noise’). These noise contributions can be suppressed by thermally stabilizing the laser cavity, and/or – in case of an optical PLL – with additional poles in the feedback loop. The present thesis considers only white frequency noise induced phase noise, with a PSD according to equation (2.10).

### 2.2.2 Modulation

Direct modulation of lasers, e.g. by changing the pump current in case of semiconductor lasers, shows a number of unwanted effects due to transient processes in the laser cavity. Therefore, high-speed optical data transmission systems usually employ external modulation. This
Figure 2.4: Electro-optical amplitude modulator, consisting of a MZI structure and a coplanar transmission line.

increases the system complexity and the optical losses, but it allows the use of (carrier-suppressed) amplitude, phase and frequency modulation formats at GHz frequencies.

External modulation is based on the change of the refractive index in an optically transparent substrate. The substrate, e.g. lithium niobate (LiNbO₃), contains a waveguide which is penetrated by an electrical field. A change in the field strength changes the refractive index and consequently, the phase velocity of the light in the waveguide. The electrical field has only a small effect on the refractive index, so that relatively large voltages are required for a phase change close to $\pi$. Common input power levels lie in a range of 10 to 30 dBm.

Figure 2.4 shows a possible architecture for an electro-optical amplitude modulator. The optical waveguide forms a Mach-Zehnder Interferometer (MZI) structure. A coplanar transmission line is laid over the waveguide, so that the two branches of the MZI are exposed to electrical fields of opposite direction. Broadband operation of the device can be obtained when the optical and the electrical waveguides are velocity matched, i.e. when the optical and the microwave signals propagate at the same phase velocity [33, 34]. The input impedance of velocity matched coplanar waveguides amounts to considerably less than 50 Ohm, which has to be taken into account when designing the modulator driver amplifier.

2.3 Receivers

As has been shown in Section 2.2, the design of an optical transmitter is straightforward without much opportunity for system optimiza-
2.3. Receivers

tion. Furthermore, the free-space channel provides excellent propagation conditions for the optical field, with only beam spreading losses as an impeding factor. The major noise sources in an optical transmission system are introduced during the detection process, so that most attention should be given to the receiver design. This is demonstrated in the next sections.

2.3.1 The quantum limit

A highly idealized optical information transmission system might consist of a laser, a photon counter and a decision circuit. Intensity modulation is applied, i.e. a ‘zero’ bit disables the laser, and a ‘one’ bit leads to an emitted power \( P_{TX} \). The photon counter has a quantum efficiency of 100%, so that the total number of received photons \( N \) is raised by one for each incoming photon (\( N \) is set to zero at the beginning of every bit interval). At the end of a bit interval, the decision circuit estimates the transmitted bit according to the number of received photons. If zero photons were counted, the system output becomes a logic zero. For \( N > 0 \), the output is set to a logic one.

By definition of [35], the temporal occurrence of photons in coherent light obeys a Poisson distribution. With an average photon arrival rate of

\[
\lambda_p = \frac{P_s}{hf},
\]

the probability of receiving \( N \) photons in a bit interval \( T_b \) amounts to (provided that a ‘one’ bit has been sent)

\[
Pr[N|\text{‘one’}] = \frac{(\lambda_p T_b)^N \cdot e^{-\lambda_p T_b}}{N!}.
\]

In (2.11), \( h \) is Planck’s constant \((6.626 \cdot 10^{-34} \text{ Js})\) and \( f \) denotes the frequency of the photons. In case of a transmitted ‘zero’ bit, the photon counter receives zero photons with probability 1. Assuming that ‘zero’ and ‘one’ bits are equally likely, the bit error rate (BER) computes to

\[
BER = \frac{1}{2} e^{-\lambda_p T_b},
\]

which can be solved for the number of photons in a bit interval:

\[
\lambda_p T_b = -\ln (2 \cdot \text{BER}).
\]
To achieve a bit error rate of $10^{-9}$ or less, the ideal photon counter requires at least an average of 20 photons per ‘one’ bit, or a total average of 10 photons per bit. This inherent boundary of optical reception is called the quantum limit [20, 28, 36, 37]. It is not associated to non-idealities or loss mechanisms of system components. The quantum limit follows directly from the quantized nature of light.

### 2.3.2 Direct detection

The ideal photon counter of the previous section does not exist. The sensitivity (in photons/bit) of real world direct detection receivers suffers from several losses and noise mechanisms of the detector device. For p-i-n and avalanche photodiodes, the main influences to be considered are:

- **Quantum efficiency:** Photodiodes never reach quantum efficiencies of 100%. Some photons do not generate an electron-hole pair, although they are being absorbed by the diode substrate. This is described by the photodiode responsivity $R$

$$ R = \frac{\eta q}{h\nu} \quad (2.15) $$

in A/W. In (2.15), $\eta$ denotes the quantum efficiency and $q$ the elementary charge ($1.602 \cdot 10^{-19} \text{ As}$). It should be noted in (2.15) that the responsivity increases with a decreasing photon frequency.

- **Shot noise:** The photon-induced generation of electron-hole pairs in the substrate is affected by statistical variations. Writing the photodiode current as the sum of a DC value and a purely random signal yields:

$$ i_{pd}(t) = RP_S + n(t) \quad (2.16) $$

where $n(t)$ denotes the shot noise process. For reasonable values of $P_S$, $n(t)$ follows a normal distribution with zero mean and a variance of $2qRP_S\Delta f$, where $\Delta f$ is the bandwidth of the device (in one-sided notation).

- **Dark current:** Photodiodes produce a current even in the total absence of a light source. This is due to the spontaneous gener-
2.3. Receivers

...ation of electron-hole pairs in the diode substrate. Dark current can lead to bit errors when detecting ‘zero’ bits.

- **Avalanche gain variation:** Avalanche photodiodes provide an internal current gain $M$ through an effect called impact ionization [36]. A single, photon-induced and accelerated electron may produce secondary electron-hole pairs which all contribute to the total current:

$$i_{pd}(t) = M R P_S.$$  \hspace{1cm} (2.17)

Unfortunately, the shot noise variance of an avalanche photodiode exceeds the expected $2qM^2 R P_S \Delta f$ by an excess noise factor of

$$F(M) \approx M^x$$ \hspace{1cm} (2.18)

where $x$ takes values between 0.2 and 1 depending on the substrate material [38]. In spite of the excess noise factor, avalanche photodiodes can be advantageous over p-i-n diodes, depending on the application.

- **Thermal noise:** In general, the currents produced in photodiodes are too weak for a decision circuit and need amplification. Inevitably, this adds thermal noise and reduces the signal-to-noise ratio (SNR).

The above mentioned effects degrade the sensitivity of a direct detection receiver by a substantial amount. Reported values lie within 10 to 25 dB [20, 22, 28, 36, 39].

### 2.3.3 Coherent detection

In coherent detection, the received optical signal is superimposed by a strong, local carrier before detection. It will be shown later (equation (4.8) in Chapter 4) that the amplitude of the p-i-n photodiode current amounts to

$$i_{pd}(t) = 2R \sqrt{P_S P_{LO}}$$ \hspace{1cm} (2.19)

in a coherent receiver. The expression $P_{LO}$ denotes the power of the local oscillator (LO). Deriving from equation (2.19), coherent reception provides signal amplification in the detector, so that the use of avalanche photodiodes is not needed or advantageous.
It has already been mentioned that the shot noise process is generally described as Gaussian distributed. Furthermore, it is assumed to be white with a one-sided power spectral density (PSD) of

\[ S_n(f) = 2qRP_{LO} \]  

(2.20)

where it has been used that \( P_S \ll P_{LO} \). Including a white thermal noise component, \( S_{th}(f) = N_{th} \), the signal-to-noise ratio of a coherent receiver becomes

\[ \text{SNR} = \frac{4R^2P_SP_{LO}}{B_{RF}(2qRP_{LO} + N_{th})} \]  

(2.21)

where \( B_{RF} \) denotes the receiver bandwidth. Dark current has been neglected here, since – in case of optical phase modulation – data detection occurs at peak power for ‘zero’ and ‘one’ bits. It is interesting to note in (2.21) that the signal power and the shot noise power density grow at the same rate with increasing LO power. Thus, at some point, the shot noise power exceeds the thermal noise power by far, and the SNR expression reduces to

\[ \text{SNR} = \frac{2RP_S}{qB_{RF}}. \]  

(2.22)

The condition, in which the receiver is dominated by shot noise, is referred to as the shot noise limit [20].

To determine the required LO power for the shot noise limit, a common amplifier circuit for optical receivers is considered. Figure 2.5(a) depicts a transimpedance amplifier, where the feedback resistor has been drawn as a noiseless device with an external voltage noise source \( u_{th}(t) \). The PSD of \( u_{th}(t) \) is white with a power density level of

\[ S_{u,th}(f) = 4kTR_T \]  

(2.23)

where \( k \) denotes Boltzmann’s constant \( (1.38 \cdot 10^{-23} \text{ J/K}) \) [29,40]. Transformations lead to an equivalent current source at the input with a PSD of

\[ S_{i,th}(f) = \frac{S_{u,th}(f)}{R_T^2} = \frac{4kT}{R_T} \]  

(2.24)

which is shown in fig. 2.5(b). Demanding that the shot noise power should exceed the thermal noise power by at least a factor of ten, yields

\[ \frac{S_n(f)}{S_{i,th}(f)} = 10 \rightarrow P_{LO} = 10 \cdot \frac{2kT}{qRR_T}. \]  

(2.25)
Numerical values of $T = 300$ K, $R = 0.65$ A/W and $R_T = 1000$ Ohm lead to a required LO power of around 0 dBm, which can easily be achieved with most lasers. Furthermore, such a low power level does not drive the photodiodes into saturation. It has been assumed in the foregoing analysis that the amplifier noise is dominated by the thermal noise of the feedback resistor, and not by other circuit components (e.g. transistors). For a real world coherent receiver, the power level of (2.25) should be doubled for safe operation in shot noise limited conditions.

It will be shown in Section 4.3 that the bit error rate can be calculated from the SNR through

$$\text{BER} = \frac{1}{2} \text{erfc}(\sqrt{\text{SNR}}).$$

(2.26)

Assuming that the bandwidth of the front end coincides with the first zero of the data spectrum, i.e. $T_b \approx 1/B_{RF}$, and using (2.11) and (2.15), the shot noise limited SNR computes to

$$\text{SNR} = \frac{2RP_S T_b}{q} = \frac{2\eta P_S T_b}{h_f} = 2\eta \lambda_p T_b.$$  

(2.27)

Plugging (2.27) into (2.26) and solving for $\lambda_p T_b$ yields

$$\lambda_p T_b = \frac{1}{2\eta} \left( \text{erfc}^{-1}(2 \cdot \text{BER}) \right)^2$$

(2.28)

which is the sensitivity of the coherent receiver in photons per bit. In case of PSK modulated data and a quantum efficiency of 100%, the

Figure 2.5: Transimpedance front end amplifier.
receiver sensitivity amounts to 9 photons/bit for a bit error rate of $10^{-9}$. This is 3.5 dB better than the quantum limit (in terms of the peak optical power\(^3\)), which is mostly based upon the fact that data transmission occurs at peak power for ‘zero’ and ‘one’ bits.

### 2.3.4 Homodyne vs. heterodyne detection

Generally, the term ‘coherent receiver’ indicates a receiver which uses some sort of local carrier for signal demodulation. More specifically, in a homodyne receiver, the local carrier is equal in frequency and phase (i.e. phase-locked) to the transmitted carrier. In a heterodyne receiver, an intermediate frequency of several times the data rate (i.e. $f_{if} \gg B_{RF}$) exists between the two carriers.

In classical radio-frequency (RF) receivers, heterodyning and homodyning provide the same sensitivity, so that most RF systems employ the superheterodyne concept [41]. In contrast, optical heterodyning suffers from a 3 dB power penalty compared to optical homodyning. This can easily be seen when writing equation (2.19) (which is the signal amplitude in a homodyne receiver) for the heterodyne case:

$$i_{pd}(t) = 2R\sqrt{P_s P_{LO}} \cdot \cos(\omega_{if} t).$$

(2.29)

The average signal amplitude in a heterodyne receiver amounts to

$$i_{pd,rms} = R\sqrt{2P_s P_{LO}}$$

(2.30)

and the SNR becomes

$$\text{SNR}_{\text{heterodyne}} = \frac{RP_s}{qB_{RF}}$$

(2.31)

which is 3 dB less than the SNR of the homodyne receiver given by equation (2.22) [20, 28, 42, 43]. In a RF heterodyne receiver, the major noise sources are located before the first mixer stage. Thus, both the signal power and the relevant noise power density are divided by 2, leaving the SNR on a constant level (provided that the noise signal has been filtered out at the IF frequency). In an optical receiver, the relevant noise contributions are introduced during the mixing process.

\(^3\)It can be argued whether the peak or the average optical power is more relevant. This thesis uses the terminology of [20], which is focused on peak power.
resulting in equal noise power – but different signal power – for the homodyne and heterodyne case. Thus, optical homodyning has an improved SNR compared to optical heterodyning.

2.3.5 Synchronous vs. asynchronous processing

In microwave communications, the expression ‘coherent’ is used differently than in optical communications. A coherent RF receiver is a system which considers the signal phase during the detection process, while a noncoherent RF receiver operates only on the signal envelope during data detection [40,44]. In both cases, the received signal has been down-converted using a local oscillator. In this sense, coherent detection surpassed noncoherent detection by approximately 1 dB in sensitivity.

In optics, if a receiver uses a local oscillator for down-conversion, it is called a coherent system. If the signal phase is considered during data detection, the receiver performs synchronous processing, and asynchronous processing otherwise [20]. The power penalty between synchronous and asynchronous data detection is comparable to the power penalty between coherent and noncoherent detection in the microwave case. Since this thesis focuses on a homodyne receiver, which performs synchronous processing by definition, the concept of asynchronous processing is not further elaborated here. Detailed comparisons of coherent optical receivers with synchronous and asynchronous detection are presented in [20,45].

2.3.6 System comparison

When evaluating an optical free-space link, a coherent receiver has the following advantages over a direct detection receiver:

- **Sensitivity**: The highest possible receiver sensitivity can be obtained with a BPSK modulated homodyne receiver. Other modulation types and/or heterodyne detection shows moderately less sensitivity, whereas direct detection suffers from severe power penalties.

- **Frequency selectivity**: A coherent receiver is less susceptible to background light from the sun or from stars [46]. In direct detection, the photodiode collects light from a broad wavelength range
close to one hundred nm (unless an optical pre-filter is used). In contrast to this, the equivalent optical input filter of a coherent receiver amounts to twice the front end bandwidth, which is – in optical terms – extremely narrow.

- **Front end:** Because of the shot noise limited condition, the front end amplifier of a coherent receiver has relaxed noise requirements. Thus, it can be tuned more easily for high responsivity and bandwidth.

- **Angle modulation:** A coherent receiver allows the use of phase or frequency modulation formats, so that many of the advanced RF system topologies can be translated to the optical domain.

The disadvantages of coherent reception are mainly related to an increased technological effort:

- **System complexity:** The overall system complexity of a coherent receiver is greatly enhanced. Besides the LO laser, an external phase or amplitude modulator, an optical coupler and – in case of homodyning – an optical phase locked loop is required.

- **Lasers:** The linewidth of semiconductor lasers is usually too wide to use for coherent reception. Instead, optically pumped Nd:YAG lasers have to be employed. In case of optical pumping, the number of lasers has quadrupled compared to a direct detection system. Furthermore, the low efficiency of optical pumping increases the electrical power demand.

- **Polarization:** The received signal and the LO signal produce a beat note on the detector surface only if they share the same state of polarization (SOP). Thus, a linear polarization in a designated plane has to be preserved throughout the optical setup, e.g. by using polarization-maintaining fibers.

In spite of the last three points, optical homodyne detection provides the only technologically reasonable way to build a high bandwidth, long distance inter-satellite communication system.
Chapter 3

Optical phase locked loops

Coherent detection indicates the demodulation of a weak passband signal with the aid of a strong, locally generated carrier. Demodulation – or likewise, frequency shifting – is a linear, but time-variant operation. The local carrier increases the usable range of the nonlinear detector device, thus increasing the demodulation efficiency. In other words, coherent detection provides gain during the mixing process.

In a homodyne receiver, the phase of the local carrier is rigidly coupled (or ‘locked’) to the carrier of the passband signal. This implies that the two carriers oscillate at the same frequency, and the received signal is directly mixed into the baseband. It has been shown in the previous chapter that, unlike the microwave case, optical homodyning increases the sensitivity by 3 dB compared to heterodyning. Furthermore, homodyning has a lower bandwidth demand on the front end. This happens at the expense of an increased technological effort, mainly due to the optical phase synchronization.

The circuits that perform the phase synchronization are called phase locked loops. In this chapter, some general properties of phase locked loops, which will be used throughout the rest of this thesis, are derived. Phase locking in the optical domain is discussed. A relatively new optical phase locked loop design (the dither loop) is presented, which will be analyzed in more detail in the next chapter.
3.1 Phase locked loops

A phase locked loop (PLL) is a circuit that synchronizes a local oscillator (LO) signal, called $u_{LO}(t)$, to the frequency and phase of an incoming reference signal $u_S(t)$. Figure 3.1 depicts a basic PLL design, as it is commonly described in literature [47–49]. The PLL consists of a voltage controlled oscillator (VCO)\(^1\), a phase detector and a loop filter. Associated gain factors are the frequency tuning sensitivity $G_{vco}$ in rad/s/V of the VCO and the phase detector constant $K_{PD}$ in V/rad. The loop filter, with an impulse response $f(t)$, eliminates noise and high frequency components of the phase detector output, and determines the dynamics of the feedback loop. The actual realization of an electrical PLL will not be discussed here. Instead, general frequency domain properties of the PLL shall be derived.

3.1.1 Closed-loop and error transfer functions

Writing the reference signal and the LO signal in time domain as

$$u_S(t) = u_1 \sin (\omega_S t + \phi_S)$$  \hspace{1cm} (3.1)

$$u_{LO}(t) = u_2 \cos (\omega_{LO} t + \phi_{LO}),$$  \hspace{1cm} (3.2)

\(^1\)The expressions VCO and LO both denote the same device and will be used interchangeably. From a control theory point of view, the VCO is called after its steering mechanism. In communication theory, LO indicates the operation in regard to the demodulation process.
yields for the phase detector output (being an ideal multiplier):

\[
u_e(t) = K_{PD}u_S(t)u_{LO}(t) = \frac{K_{PD}u_1u_2}{2}\left(\sin\left((\omega_S + \omega_{LO})t + \phi_S + \phi_{LO}\right) + \sin\left((\omega_S - \omega_{LO})t + \phi_S - \phi_{LO}\right)\right). \tag{3.3}\]

In case of equal frequencies and small phase differences, (3.3) simplifies to

\[
u_e(t) = \frac{K_{PD}u_1u_2}{2}\sin(\phi_S - \phi_{LO}) \approx \frac{K_{PD}u_1u_2}{2}(\phi_S - \phi_{LO}) \tag{3.4}\]

where the expression at the sum frequency has been neglected, because it is not relevant for loop operation. The multiplier feeds the loop filter, the output signal of which computes to

\[
u_c(t) = f(t) \ast u_e(t) \tag{3.5}\]

where ‘\(\ast\)’ denotes the convolution operator. With the VCO transfer function,

\[
\phi_{LO}(t) = G_{vco} \cdot \int_0^t u_c(\tau)d\tau, \tag{3.6}
\]

equations (3.4), (3.5) and (3.6) define a feedback loop. This time domain description of the PLL does not provide much insight into the loop operation. It is therefore convenient to use the Laplace transformations:

\[
U_e(s) = \frac{K_{PD}u_1u_2}{2}\left(\Phi_S(s) - \Phi_{LO}(s)\right) \tag{3.7}
\]

\[
U_c(s) = F(s) \cdot U_e(s) \tag{3.8}
\]

\[
\Phi_{LO}(s) = G_{vco} \cdot \frac{U_c(s)}{s}. \tag{3.9}
\]

Solving the above formulas for \(\Phi_S/\Phi_{LO}\) yields the closed-loop transfer function of the PLL:

\[
H(s) = \frac{\Phi_{LO}(s)}{\Phi_S(s)} = \frac{s^{-1}G_{loop}F(s)}{1 + s^{-1}G_{loop}F(s)} \tag{3.10}
\]
where $G_{\text{loop}} = G_{vco} K_{PD} u_1 u_2 / 2$ denotes the overall loop gain. Likewise, the error transfer function can be calculated through:

$$1 - H(s) = \frac{\Phi_e(s)}{\Phi_S(s)} = \frac{\Phi_S(s) - \Phi_{LO}(s)}{\Phi_S(s)} = \frac{1}{1 + s^{-1} G_{\text{loop}} F(s)}. \quad (3.11)$$

The two transfer functions (3.10) and (3.11) are useful when evaluating noise effects on the PLL performance, as will be shown in Sections 4.2.2 and 4.2.3. To obtain simple mathematical expressions, it is necessary to specify the loop filter in more detail.

### 3.1.2 Characteristic loop parameters

To further analyze the PLL of the previous section, a first order active filter, depicted in Fig. 3.2, will be used. Assuming a large gain of the amplifier, calculations result in a transfer function of

$$F(s) = \frac{U_2}{U_1} = \frac{s R_2 C_1 + 1}{s R_1 C_1} = \frac{s \tau_2 + 1}{s \tau_1} \quad (3.12)$$

with the two time constants

$$\tau_1 = R_1 C_1 \quad (3.13)$$

$$\tau_2 = R_2 C_1. \quad (3.14)$$

This filter provides proportional/integral (PI) control, which can be seen from the infinite gain at zero frequency [50]. Plugging (3.12) into

![First order active loop filter. The second stage, an inverter, is not necessarily needed in a PLL if the phase detector is a multiplier [49].](image)
3.1. Phase locked loops

![Graph](image)

(a) Closed-loop transfer function. (b) Error transfer function.

**Figure 3.3:** Transfer functions of a critically damped ($\zeta = 1/\sqrt{2}$) second order PLL.

The closed-loop transfer function (3.10), yields:

$$H(s) = \frac{G_{loop}(s\tau_2 + 1)/\tau_1}{s^2 + sG_{loop}\tau_2/\tau_1 + G_{loop}/\tau_1} = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (3.15)$$

with

$$2\pi f_n = \omega_n = \sqrt{\frac{G_{loop}}{\tau_1}} \quad (3.16)$$

$$\zeta = \frac{\tau_2}{2} \sqrt{\frac{G_{loop}}{\tau_1}} = \frac{\omega_n \tau_2}{2} \quad (3.17)$$

In control theory, $\omega_n$ and $\zeta$ are called the natural frequency and the damping factor of the loop. A third important parameter of a PLL is the loop noise bandwidth:

$$B_n = \int_0^{\infty} |H(f)|^2 df = \pi f_n \left(\zeta + \frac{1}{4\zeta}\right) \quad (3.18)$$

in Hz, where the evaluation of the integral has been taken from [47]. The closed-loop transfer function (3.10) and the error transfer function (3.11) of a critically damped ($\zeta = 1/\sqrt{2}$) second order loop are depicted in fig. 3.3.
3.1.3 Stability

As any feedback system, the PLL is prone to instability. Unstable operation must be prevented by choosing an appropriate loop gain \( G_{\text{loop}} \). The dynamic response of the PLL in dependency of \( G_{\text{loop}} \) can be analyzed by means of a root locus plot [50], as it is depicted in fig. 3.4 for a second order system with a closed-loop transfer function as in (3.15). In this representation, the poles of the transfer function are plotted for values of \( G_{\text{loop}} \) from zero to infinity. It is stated that the system is stable if all poles are located in the left half plane (LHP), i.e. if their real part is negative. As it can be seen from fig. 3.4, the second order PLL will be stable for values of \( G_{\text{loop}} \) larger than zero. For a feedback gain of zero (i.e. open loop operation), though, it will show unstable behavior. This becomes apparent when bearing in mind that the loop filter of (3.12) contains an integrator. Once charged, the integrator output does not return to zero although its input is zero. Such a behavior is also called instability in the bounded input–bounded output (BIBO) sense. It is of less importance here, because the PLL is always operated with a feedback gain larger than zero, i.e. in closed-loop conditions. For a critically damped loop, i.e. \( \zeta = 1/\sqrt{2} \) or \( G_{\text{loop}} = 2\tau_1/\tau_2^2 \), the real and imaginary parts of the poles are equal in magnitude. This way, a good compromise between phase margin and dynamic response can be achieved.
3.2 OPLL design

With only minor modifications, the concept of the PLL can be transferred to optical frequencies. In this case, the TX and LO sources are lasers, oscillating at several hundred THz. Phase detection occurs in photodiodes, the nonlinear transfer function of which produces a beat note at the intermediate frequency of the two lasers. The photodiode output signal exists in the electrical domain, so that an ordinary analog filter (i.e. from fig. 3.2) can be used to determine the loop characteristics. Controlling the optical phase of the LO takes place through thermal and/or mechanical disturbances on the laser cavity, or through an external optical phase modulator.

Two problems, however, impede the straightforward application of the PLL concept for lasers:

- Photodiodes do not perform true multiplication. Apart from the beat note, DC currents proportional to the power of each laser are being produced. These DC currents interfere with the phase locking process unless effectively suppressed (i.e. [51]).

- In case of carrier suppressed transmission, the loop has ‘nothing to lock to’. Microwave receivers solve this problem with a Costas loop (from [52]). The Costas loop requires two LO signals with an exact 90° phase shift between them. In microwaves, this can easily be generated with a branch-line or a Lange coupler. No direct replacement exists for these couplers in integrated optics\(^2\).

Several advanced OPLL designs are known which circumvent one or both of the above mentioned problems. They will be presented in the next sections.

3.2.1 Balanced loop

The balanced loop, schematically depicted in fig. 3.5(a), employs a 180°/3-dB coupler and a dual detector front end. Since both photodiodes receive the same amount of power, their DC currents cancel out in

\(^2\)Here, the expression ‘integrated optics’ shall denote a setup without any free-space optics, i.e. the light is always guided by fibers or waveguides on LiNbO\(_3\) (or a similar) substrate.
Figure 3.5: Advanced OPLL designs.
case of phase lock. The balanced loop obtains its name from this symmetry. Further advantages of this configuration are the cancellation of LO intensity noise [53] and loss minimization by using both branches of the coupler.

A constant phase error between the two lasers leads to a deviation from the symmetrical operation and therefore to a DC current flowing into the transimpedance amplifier. This signal can be directly used for LO phase control, provided that the loop employs a DC-coupled front end. An alternative concept of the balanced loop, which omits a DC-coupling of the front end, is presented in [54]. The balanced loop requires the transmission of a residual carrier for phase locking. The modulation depth of the TX carrier determines the amount of power which is used for phase locking. A balanced loop can handle analog and digital data of any modulation format.

3.2.2 Costas/Decision-driven loop

The Costas loop, shown in fig. 3.5(b), splits the received signal into an in-phase \((I)\) and a quadrature \((Q)\) component. For technological reasons, and because it is advantageous with respect to LO intensity noise, each component is detected with a balanced front end. Multiplication of the \(I\) and \(Q\) signals yields a measure of the incident phase error. If the phase error is zero, the \(Q\) component – and thus, the multiplier output – becomes zero.

With a Costas loop, no residual carrier transmission is required, and thus, an AC coupled front end can be used. This is at the expense of a 90° coupler (with a coupling ratio different from 3 dB) instead of a 180°/3-dB coupler. The power splitting ratio of the coupler determines the amount of power which is used for phase locking. Unlike the balanced loop, where the power used for phase locking can easily be controlled by changing the modulation depth, the power demand of the phase-locking branch in a Costas loop is hard-wired into the system. The Costas loop can handle any kind of user signal.

A decision-driven loop employs the same receiver setup as the Costas loop, but with the multiplier placed after the decision circuit of the \(I\) branch. A one bit delay has to be placed in the \(Q\) branch to achieve equal transit times of the two signal components. According to [21], decision-driven loops provide smaller phase error variances than
Costas loops. The decision-driven loop is limited to purely digital data transmission.

### 3.2.3 SyncBit loop

The SyncBit loop, depicted in fig. 3.5(c), is a modification of the Costas loop which does not require a 90° coupler. Most of the time, the feedback loop is opened and the receiver acts as the in-phase branch of the Costas loop. At certain time intervals of one bit duration, the loop is closed and the LO phase is shifted by 90°, so that the receiver becomes the Q branch of the Costas loop. At the time of synchronization, the receiver needs a priori information about the transmitted bit. This can be done by placing ‘zero information’ bits into the data stream or, on a more advanced level, with coding algorithms. The power penalty due to phase locking is introduced through an increased bandwidth demand. It can be controlled by changing the frequency of occurrence of the synchronization bits.

An alternative setup of a SyncBit loop shifts the transmitter phase, rather than the LO phase, to the quadrature state (i.e. data bits \(\pm \pi/2\), sync bits 0). This reduces the complexity of the system, since the phase modulator is already required in the transmitter [55]. A further derivation of the SyncBit loop is presented in [56] with the switched residual carrier concept. The SyncBit loop performs extensive pre- and post-processing to include and extract the synchronization bits to and from the data stream.

### 3.2.4 Dither loop

The dither loop, depicted in fig. 4.1 of Chapter 4, employs a conventional balanced receiver, but with an AC coupling in the signal (and phase-locking) path. Since no residual carrier transmission takes place, the receiver output does not contain any information about the incident phase error at first. Instead, a small phase disturbance is applied to the LO, which propagates through the system. This phase disturbance, henceforth called the dither signal, can be measured at the receiver output as a fluctuation in signal power. For a non-zero phase error, the power detector output will contain a component at the dither frequency. The magnitude of this component is proportional to
the amount of phase error, and its phase (relative to the original dither signal) depends on the sign of the phase error. Thus, by synchronous demodulation, a phase error signal can be extracted. The dither signal constitutes a power penalty, because it reduces the average signal amplitude. The power penalty due to phase locking can be controlled within the receiver by changing the amplitude of the dither signal.

The dither loop, first proposed in [19] and explained in more detail in [57–59], requires a constant average power of the transmitted signal for phase locking. Otherwise, it accepts analog or digital data of arbitrary modulation.

### 3.3 System comparison

In terms of their communication performance, the above mentioned OPLL designs are all equally suited to be used in a coherent receiver. From a system engineer’s point of view, they can be distinguished by their technological complexity and by the requirements (and limitations) on the transmitted signal. The main points to be considered are specified below and summarized in table 3.1:

- **Residual carrier transmission**: The transmission of a residual carrier is undesirable, because it increases the average signal power

<table>
<thead>
<tr>
<th>Loop type</th>
<th>No residual carrier transmission required</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>180°/3-dB coupler</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AC coupled front end</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transparent link</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Variable phase-locking power</td>
<td></td>
</tr>
<tr>
<td>Balanced</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>Costas</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>Dec.-driven</td>
<td>•</td>
<td>Only digital signals</td>
</tr>
<tr>
<td>SyncBit</td>
<td>•</td>
<td>Extensive pre-/post-processing</td>
</tr>
<tr>
<td>Dither</td>
<td>•</td>
<td>Constant-envelope user signal</td>
</tr>
</tbody>
</table>

**Table 3.1**: *Comparison of different OPLL types.*
(provided that the power of the data signal is kept constant in the passband). If the transmitter employs an optical booster amplifier, the average signal power should be kept as low as possible [57,60]. This is due to the fact that optical fiber amplifiers are limited in their average power, rather than the peak power.

- **Optical coupler**: Optical 180°/3-dB couplers are commercially available devices with low excess loss. They exist in fiber coupled, integrated optics with polarization maintaining characteristics. In contrast, 90° couplers with coupling ratios different from 3 dB (e.g. 90% in-phase, 10% quadrature for a Costas loop) have never overcome an experimental state of development. Most known designs use free-space optics and/or act on the polarization of the light [61–65]. For space applications, a 90° coupler could be built by glueing together individual glass blocks, where each one accomplishes a certain function (i.e. beam splitter, quarter wave plate, etc.) [66].

- **Front end coupling**: The circuit complexity of a DC coupled front end is increased compared to an AC coupled one [67]. It is favorable if the loop does not require a DC coupled front end output.

- **Transparent link**: To achieve the highest possible flexibility regarding the transmitted signal, the loop should accept analog as well as digital data. In case of digital transmission, no system component should depend on the bit rate. A transmission system that accepts analog and digital data (at an arbitrary data rate) is called transparent.

- **Adaptively controllable power penalty due to phase locking**: In a homodyne receiver, a small amount of the received optical power has to be used to extract the phase of the transmitter laser. The fraction of the received signal which flows into the phase-locking branch is lost from data detection, and hence, it constitutes a power penalty. To minimize the power penalty, it is advantageous if the signal power which flows into the phase-locking branch can be adaptively controlled within the receiver.

As can be seen from table 3.1, only the dither loop combines a relatively simple system setup with the transparent link property. Furthermore, only in a dither loop, the amount of power used for phase
locking can be controlled within the receiver. The requirement for a constant average power of the transmitted signal is inherently achieved for BPSK modulated digital data. For analog information, the use of frequency modulated electrical subcarriers leads to a constant average signal power. Thus, the constraint of the dither loop on the transmitted signal can easily be satisfied. Due to these arguments, the dither loop is believed to be superior to other OPLL designs.

An overview of published coherent (and some direct detection) receiver experiments is given in table 3.2 and 3.3. The best reported receiver sensitivity of 20 photons/bit has been achieved with a SyncBit loop. Impressive results were demonstrated with RZ-DPSK modulated direct detection receivers. Their application for an inter-satellite link is questionable because of the critical optical input filter.
<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Laser type</th>
<th>Receiver type</th>
<th>PLL type</th>
<th>Modulation format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>Kahn</td>
<td>ECL</td>
<td>Homodyne</td>
<td>Balanced</td>
<td>PSK</td>
</tr>
<tr>
<td>1990</td>
<td>Kazovsky</td>
<td>Nd:YAG</td>
<td>Homodyne</td>
<td>Balanced</td>
<td>PSK</td>
</tr>
<tr>
<td>1990</td>
<td>Kahn</td>
<td>ECL</td>
<td>Homodyne</td>
<td>Balanced</td>
<td>PSK</td>
</tr>
<tr>
<td>1990</td>
<td>Norimatsu</td>
<td>ECL</td>
<td>Homodyne</td>
<td>Decision-driven</td>
<td>PSK</td>
</tr>
<tr>
<td>1990</td>
<td>Schöpflin</td>
<td>Solid-state</td>
<td>Homodyne</td>
<td>Costas</td>
<td>PSK</td>
</tr>
<tr>
<td>1992</td>
<td>Hornbachner</td>
<td>Nd:YAG</td>
<td>Homodyne</td>
<td>Costas</td>
<td>PSK</td>
</tr>
<tr>
<td>1992</td>
<td>Wandernoth</td>
<td>–</td>
<td>Homodyne</td>
<td>SyncBit</td>
<td>PSK</td>
</tr>
<tr>
<td>1992</td>
<td>Norimatsu</td>
<td>ECL</td>
<td>Homodyne</td>
<td>Decision-directed</td>
<td>QPSK</td>
</tr>
<tr>
<td>1993</td>
<td>Li</td>
<td>ECL</td>
<td>Homodyne</td>
<td>Balanced</td>
<td>PSK</td>
</tr>
<tr>
<td>1993</td>
<td>Guo</td>
<td>EGCL</td>
<td>Heterodyne</td>
<td>IF-PLL</td>
<td>CPFSK</td>
</tr>
<tr>
<td>1995</td>
<td>Norimatsu</td>
<td>DFB</td>
<td>Homodyne</td>
<td>Decision-driven</td>
<td>BPSK</td>
</tr>
<tr>
<td>1999</td>
<td>Atia</td>
<td>DFB</td>
<td>Direct</td>
<td>–</td>
<td>RZDPSK</td>
</tr>
<tr>
<td>2003</td>
<td>Yamakawa</td>
<td>DFB</td>
<td>Heterodyne</td>
<td>IF-PLL</td>
<td>PSK</td>
</tr>
<tr>
<td>2004</td>
<td>Sinsky</td>
<td>DFB</td>
<td>Direct</td>
<td>–</td>
<td>RZDPSK</td>
</tr>
<tr>
<td>2004</td>
<td>Spellmeyer</td>
<td>–</td>
<td>Direct</td>
<td>–</td>
<td>RZDPSK</td>
</tr>
<tr>
<td>2005</td>
<td>Herzog</td>
<td>Nd:YAG</td>
<td>Homodyne</td>
<td>Dither</td>
<td>BPSK</td>
</tr>
</tbody>
</table>

Table 3.2: Published experimental bit error rate measurement results of coherent (and some direct detection) receivers.
### Table 3.3: Table 3.2 continued. Values in parentheses are effective photoelectrons, neglecting inherent losses of a 3×3 symmetrical coupler.

<table>
<thead>
<tr>
<th>$\lambda$ [nm]</th>
<th>Sequence length</th>
<th>Bit rate [Mbit/s]</th>
<th>Bit error rate</th>
<th>Required power [dBm]</th>
<th>Photons /bit</th>
<th>Power penalty [dB]</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1500</td>
<td>$2^{15} - 1$</td>
<td>1000</td>
<td>$10^{-9}$</td>
<td>-52.2</td>
<td>46</td>
<td>7</td>
<td>[68]</td>
</tr>
<tr>
<td>1320</td>
<td>$2^{15} - 1$</td>
<td>140</td>
<td>$10^{-9}$</td>
<td>-62.8 -40</td>
<td>25</td>
<td>4.4</td>
<td>15.7</td>
</tr>
<tr>
<td>1508</td>
<td>$2^{15} - 1$</td>
<td>4000</td>
<td>$10^{-9}$</td>
<td>-44.2</td>
<td>72</td>
<td>9</td>
<td>[70]</td>
</tr>
<tr>
<td>1550</td>
<td>$2^7 - 1$</td>
<td>10000</td>
<td>$10^{-9}$</td>
<td>-34.2</td>
<td>297</td>
<td>15.2</td>
<td>[71]</td>
</tr>
<tr>
<td>1320</td>
<td>$2^7 - 1$</td>
<td>140</td>
<td>$10^{-9}$</td>
<td>-66.2 -57.5</td>
<td>(11)</td>
<td>(21)</td>
<td>–</td>
</tr>
<tr>
<td>1064</td>
<td>16 bit word</td>
<td>140</td>
<td>$10^{-9}$</td>
<td>-60.4</td>
<td>35</td>
<td>5.9</td>
<td>[73]</td>
</tr>
<tr>
<td>1064</td>
<td>$2^{10} - 1$</td>
<td>565</td>
<td>$10^{-9}$</td>
<td>–</td>
<td>20</td>
<td>3.5</td>
<td>[74]</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>8000</td>
<td>$10^{-9}$</td>
<td>-27.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1551</td>
<td>‘0101’</td>
<td>80</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1546</td>
<td>‘0101’</td>
<td>2·154</td>
<td>$10^{-9}$</td>
<td>-49.2, -47.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1536</td>
<td>$2^7 - 1$</td>
<td>10000</td>
<td>$10^{-9}$</td>
<td>-38.9</td>
<td>99.6</td>
<td>10</td>
<td>[78]</td>
</tr>
<tr>
<td>–</td>
<td>$2^{31} - 1$</td>
<td>10000</td>
<td>$10^{-9}$</td>
<td>–</td>
<td>30</td>
<td>1.76</td>
<td>[79]</td>
</tr>
<tr>
<td>1550</td>
<td>$2^{31} - 1$</td>
<td>2500</td>
<td>$10^{-9}$</td>
<td>-42.5</td>
<td>174</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1550</td>
<td>$2^{31} - 1$</td>
<td>42700</td>
<td>$10^{-9}$</td>
<td>-37</td>
<td>38</td>
<td>1.2</td>
<td>[81]</td>
</tr>
<tr>
<td>1550</td>
<td>$2^{31} - 1$</td>
<td>40000</td>
<td>$10^{-9}$</td>
<td>–</td>
<td>35</td>
<td>2.2</td>
<td>[82]</td>
</tr>
<tr>
<td>1064</td>
<td>$2^{31} - 1$</td>
<td>400</td>
<td>$10^{-9}$</td>
<td>-55.7</td>
<td>36</td>
<td>6</td>
<td>[58]</td>
</tr>
</tbody>
</table>
Chapter 3. Optical phase locked loops
Chapter 4

Phase locking by LO phase dithering

This chapter focuses on a mathematical treatment of the dither loop phase-locking scheme. Formulas will be derived for various system variables including the phase detector gain. An unwanted behavior (‘false lock’) of the loop is reported. In the second part, noise processes and their effect on the receiver performance are being discussed. A design rule, based on elementary system specifications, is derived to develop an optimum dither loop. To demonstrate the design rule, dither loops are calculated for various values of the photodiode responsivity, the laser linewidth and the system bit rate. Throughout this chapter, simulations are presented to support the achieved theoretical results.

4.1 Dither loop analysis

The dither loop, schematically depicted in fig. 4.1, performs a series of signal processing steps, including phase dithering, power detection, filtering and synchronous demodulation, to extract a phase error signal. It is not immediately clear how these steps interact, and most importantly, the phase detector as such cannot easily be identified. A mathematical analysis can provide insight into the operation of the loop and lead to predictions about its performance.
Figure 4.1: Schematic overview of the dither loop. Thick lines indicate optical signals or components.
In coherent optical communications, the transmitted signal and the LO signal are typical narrowband bandpass signals, which is indicated by the fact that the carrier frequency exceeds the bandwidth by far. For such signals, the notation:

\[ E_S = \sqrt{P_S} \cdot e^{j\phi_S} \]  \hspace{1cm} (4.1)

\[ E_{LO} = \sqrt{P_{LO}} \cdot e^{j\phi_{LO}} \]  \hspace{1cm} (4.2)

is favorable, which is called the complex envelope \([40,44]\) or the complex baseband representation \([83]\) of the bandpass signal. The expression \(E_S\) denotes the transmitted signal, and \(E_{LO}\) is the light emitted from the local oscillator. \(P_S\) and \(P_{LO}\) are the power and \(\phi_S\) and \(\phi_{LO}\) the incident phases of the associated signals. The optical hybrid, being a symmetrical 180°/3-dB device, generates the sum and the difference of its input fields:

\[ E_1 = \frac{1}{\sqrt{2}}(E_S + E_{LO}) \]  \hspace{1cm} (4.3)

\[ E_2 = \frac{1}{\sqrt{2}}(E_S - E_{LO}). \]  \hspace{1cm} (4.4)

Illuminating the surface of a photodiode with an optical field \(E\) will lead to a current of

\[ i_{pd} = R \cdot |E|^2 \]  \hspace{1cm} (4.5)

flowing through the diode \([36,84,85]\). In (4.5), \(R\) is the responsivity of the photodiode in A/W. Evaluating (4.1) and (4.2) in (4.3), and (4.3) in (4.5) yields

\[ i_{pd1} = R \cdot |E_1|^2 = \frac{R}{2} \left| \sqrt{P_S}e^{j\phi_S} + \sqrt{P_{LO}}e^{j\phi_{LO}} \right|^2 \]

\[ = \frac{R}{2} \left( \sqrt{P_S}e^{j\phi_S} + \sqrt{P_{LO}}e^{j\phi_{LO}} \right) \cdot \left( \sqrt{P_S}e^{-j\phi_S} + \sqrt{P_{LO}}e^{-j\phi_{LO}} \right) \]

\[ = \frac{R}{2} \left( P_S + \sqrt{P_S P_{LO}}e^{j(\phi_S - \phi_{LO})} + \sqrt{P_S P_{LO}}e^{-j(\phi_S - \phi_{LO})} + P_{LO} \right) \]

\[ = \frac{R}{2} \left( P_S + P_{LO} + 2\sqrt{P_S P_{LO}} \cdot \cos(\phi_S - \phi_{LO}) \right), \]  \hspace{1cm} (4.6)

and accordingly

\[ i_{pd2} = \frac{R}{2} \left( P_S + P_{LO} - 2\sqrt{P_S P_{LO}} \cdot \cos(\phi_S - \phi_{LO}) \right). \]  \hspace{1cm} (4.7)
Shot noise expressions have been neglected in (4.7), they will be introduced in Section 4.2.3. With (4.6) and (4.7), a total current of

\[ i_1 = i_{pd1} - i_{pd2} = 2R \sqrt{P_s P_{LO}} \cdot \cos(\phi_S - \phi_{LO}) \]  

(4.8)

flows into the transimpedance amplifier. This amplifier has a transfer function of

\[ u_2 = -R_T i_1. \]  

(4.9)

The matching network and the RF amplifier contribute gain factors of 1/2 and \( G_{RF} \), so that the output signal of the system becomes

\[ u_{out} = -G_{RF} R_T R \sqrt{P_s P_{LO}} \cdot \cos(\phi_S - \phi_{LO}) \]

\[ = -U_0 \cos(\phi_S - \phi_{LO}). \]  

(4.10)

The expression \( G_{RF} R_T R \sqrt{P_s P_{LO}} \) is the amplitude of the output signal in V. It will be replaced by \( U_0 \) for the rest of this analysis. In a real world receiver, an automatic gain control (AGC) circuit for \( G_{RF} \) would ensure that \( U_0 \) stays independent of the received power \( P_s \).

So far, the dither loop operates like a balanced loop, e.g. as in [22]. The main difference lies in the definition of the phase expressions \( \phi_S \) and \( \phi_{LO} \):

\[ \phi_S(t) = \frac{\pi}{2} + \frac{\pi}{2} d(t) + \phi_{nS}(t) \]  

(4.11)

\[ \phi_{LO}(t) = G_{vco} \int_{-\infty}^{t} \left( u_c(\tau) + a_d \sin(\omega_d \tau) \right) d\tau + \phi_{nLO}(t) \]

\[ = G_{vco} \int_{-\infty}^{t} u_c(\tau) d\tau - \frac{a_d G_{vco}}{\omega_d} \cos(\omega_d t) + \phi_{nLO}(t). \]  

(4.12)

In (4.11), \( d(t) \) denotes the data signal taking values in \( \{-1, 1\} \), and \( \phi_{nS}(t) \) is the phase noise process of the transmitter laser. The first expression in (4.11), \( \pi/2 \), has been introduced to ease the mathematical evaluation and must not be generated in the actual system.\(^1\) It can be seen from the second term, \( \pi/2 \cdot d(t) \), that the TX laser is fully modulated, i.e. no residual carrier transmission takes place. In (4.12),

\(^1\)In phase locked condition, TX and LO lasers oscillate 90° out of phase. Consequently, the LO carrier is in-phase to the data signal [22, 47]. It is convenient for the calculation to anticipate this behavior.
4.1. Dither loop analysis

$a_d$ and $\omega_d$ are the voltage amplitude and angular frequency\(^2\) of the dither signal, $u_c(t)$ is the loop filter output and $\phi_{nLO}(t)$ the phase noise process of the LO laser. The expression $G_{vco}$ denotes the sensitivity of the frequency tuning input of the laser. Rewriting the total phase error, $\phi_E(t) = \phi_{nS}(t) - \phi_{LO}(t)$, results in

$$\phi_E(t) = \phi_e(t) + \phi_d \cos(\omega_d t)$$

(4.13)

$$\phi_e(t) = \phi_n(t) - \phi_c(t) = \phi_n(t) - G_{vco} \int_{-\infty}^{t} u_c(\tau) d\tau$$

(4.14)

$$\phi_n(t) = \phi_{nS}(t) - \phi_{nLO}(t)$$

(4.15)

where $\phi_n(t)$ refers to the combined phase noise process of the two lasers. The expression $a_d G_{vco}/\omega_d$ has been replaced by $\phi_d$, which denotes the phase amplitude of the dither signal. The total phase error $\phi_E(t)$ consists of a deterministic part (the dither signal), a stochastic part $\phi_n(t)$, and the history of the system $\phi_c(t)$ (the integrated loop filter output). The expression $\phi_e(t)$ will be called the residual phase error for the rest of this analysis. The loop is in lock when $\phi_e(t)$ equals zero.\(^3\) Evaluating (4.11)–(4.15) in (4.10) yields for the output signal of the system:

$$u_{out}(t) = -U_0 \cos \left( \frac{\pi}{2} + \frac{\pi}{2} d(t) + \phi_E(t) \right) = U_0 \sin \left( \frac{\pi}{2} d(t) + \phi_E(t) \right)$$

$$= U_0 \left( \sin \left( \frac{\pi}{2} d(t) \right) \cos \phi_E(t) + \cos \left( \frac{\pi}{2} d(t) \right) \sin \phi_E(t) \right).$$

(4.16)

Now, since $\cos (\pi/2 \cdot d(t))$ is zero and $\sin (\pi/2 \cdot d(t))$ equals $d(t)$, (4.16) becomes

$$u_{out}(t) = U_0 d(t) \cos \phi_E(t) = G_{RF} R_T R \sqrt{P_S P_{LO}} \cdot d(t) \cos \phi_E(t).$$

(4.17)

The data signal $d(t)$ determines the polarity of the receiver output. The amplitude of the output depends, among various system constants and the received power, on the total phase error $\phi_E(t)$. To have a strong output signal, it is crucial to keep the total phase error as small as possible.

---

\(^2\)The normal frequency will be denoted by $f_d = \omega_d/(2\pi)$.

\(^3\)Strictly speaking, the loop is in lock when the mean value of $\phi_e(t)$ equals zero and the variance is small. For brevity, it will be assumed that $\phi_e(t) = 0$ until Section 4.2, where noise effects are treated.
The phase-locking branch of the receiver first measures the power of the output signal:

\[ u_{\text{pwr}}(t) = G_{\text{pwr}} \frac{u_{\text{out}}^2(t)}{Z_0} = C_1 u_{\text{out}}^2(t) \]  \hspace{1cm} (4.18)

where \( G_{\text{pwr}} \) denotes the sensitivity of the power detector in V/W, and \( Z_0 \) is the impedance of the output transmission line. \( C_1 \) replaces \( G_{\text{pwr}}/Z_0 \) for brevity. Evaluating (4.17) and (4.13) in (4.18) leads to

\[ u_{\text{pwr}}(t) = C_1 U_0^2 d^2(t) \cos^2 \phi_E(t) = \frac{U_0^2 C_1}{2} \left( 1 + \cos (2\phi_E(t)) \right) \]

\[ = \frac{U_0^2 C_1}{2} \left( 1 + \cos (2\phi_e(t) + 2\phi_d \cos (\omega_d t)) \right) = \frac{U_0^2 C_1}{2} \left( 1 + \cos (2\phi_e(t)) \cdot \cos (2\phi_d \cos (\omega_d t)) \right) - \sin (2\phi_e(t)) \sin (2\phi_d \cos (\omega_d t)). \]  \hspace{1cm} (4.19)

For further simplification of (4.19), the following series expansions from [86] can be used:

\[ \cos (z \cos \theta) = J_0(z) + 2 \sum_{k=1}^{\infty} (-1)^k J_{2k}(z) \cos (2k\theta) \]  \hspace{1cm} (4.20)

\[ \sin (z \cos \theta) = 2 \sum_{k=0}^{\infty} (-1)^k J_{2k+1}(z) \cos ((2k+1)\theta) \]  \hspace{1cm} (4.21)

where \( J_n(\cdot) \) denotes the Bessel function of the first kind of order \( n \). Using (4.20) and (4.21) in (4.19) yields

\[ u_{\text{pwr}}(t) = \frac{U_0^2 C_1}{2} \left( 1 + \cos (2\phi_e(t)) \right) \left( J_0(2\phi_d) - 2J_2(2\phi_d) \cos (2\omega_d t) \right. \]

\[ + 2J_4(2\phi_d) \cos (4\omega_d t) - \ldots \left. \right) - \sin (2\phi_e(t)) \left( 2J_1(2\phi_d) \cos (\omega_d t) 
\]

\[ - 2J_3(2\phi_d) \cos (3\omega_d t) + 2J_5(2\phi_d) \cos (5\omega_d t) - \ldots \right). \]  \hspace{1cm} (4.22)

The band-pass filter following the power detector removes DC terms and any frequency components of order \( n > 4 \). Furthermore, after synchronous demodulation, beat products of the form \( \cos (n\omega_d t) \cdot \cos (\omega_d t) \)
4.1. Dither loop analysis

Do not produce any low frequency components for \( n > 1, n \in \mathbb{N} \). Thus, the relevant power detector output can be written as

\[
u_{\text{pwr}}(t) = -U_0^2 C_1 J_1(2\phi_d) \cos(\omega_d t) \sin(2\phi_e(t)). \tag{4.23}\n\]

For mathematical simplicity, the band-pass filter will not be included in the calculation. To have a negligible effect on the OPLL performance, its transfer function should be flat (in frequency and phase) across a frequency range of approximately \( f_d \pm B_n \).

Synchronous demodulation occurs in a multiplier with a reference voltage \( K_{\text{dmd}} \). The reference voltage is necessary to have reasonable output units, i.e. V instead of V\(^2\). Feeding the multiplier with (4.23) and the phase shifted dither signal, \(-A_d \cos(\omega_d t)\), results in

\[
u_{\text{dmd}}(t) = \frac{U_0^2 C_1 A_d}{K_{\text{dmd}}} \cdot J_1(2\phi_d) \cos^2(\omega_d t) \sin(2\phi_e(t)). \tag{4.24}\n\]

The \( \cos^2(\cdot) \) expression contains a DC-term of 1/2 and an expression at twice the dither frequency, which will be suppressed by the closed-loop transfer function. Thus, the multiplier output becomes

\[
u_{\text{dmd}}(t) = \frac{U_0^2 C_1 C_2}{2} \cdot J_1(2\phi_d) \sin(2\phi_e(t)) \tag{4.25}\n\]

where \( C_2 \) replaces \( A_d / K_{\text{dmd}} \) for brevity. Further simplifications can be made using the approximations for small arguments \( z \):

\[
J_{\nu}(z) \approx \frac{(z/2)^{\nu}}{\Gamma(\nu + 1)} \tag{4.26}\n\]

from [86], and

\[
\sin(z) \approx z \tag{4.27}\n\]

which follows from the series expansion of the sine function. In (4.26), \( \Gamma(\cdot) \) denotes the gamma function, related to the factorial by \( \Gamma(n) = (n - 1)! \) if \( n \) is an integer. Finally, with (4.26) and (4.27) in (4.25), a measure for the residual phase error \( \phi_e(t) \) has been extracted:

\[
u_{\text{dmd}}(t) = U_0^2 \phi_d C_1 C_2 \cdot \phi_e(t) = K_{PD} \cdot \phi_e(t) = \frac{G_v^2 R_F^2 R_L^2 P_S P_{LO} g_d G_{vco} G_{pwr} A_d}{\omega_d Z_0 K_{\text{dmd}}} \cdot \phi_e(t). \tag{4.28}\n\]
Table 4.1: Feasible parameter values for a real world dither loop receiver.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_S$</td>
<td>W</td>
<td>$1 \cdot 10^{-8}$</td>
<td>Received power (-50 dBm)</td>
</tr>
<tr>
<td>$P_{LO}$</td>
<td>W</td>
<td>$1 \cdot 10^{-2}$</td>
<td>LO power (10 dBm)</td>
</tr>
<tr>
<td>$R$</td>
<td>A/W</td>
<td>0.65</td>
<td>Photodiode responsivity</td>
</tr>
<tr>
<td>$R_T$</td>
<td>$\Omega$</td>
<td>1000</td>
<td>Transimpedance</td>
</tr>
<tr>
<td>$G_{RF}$</td>
<td>scalar</td>
<td>11</td>
<td>RF amplifier gain</td>
</tr>
<tr>
<td>$U_0$</td>
<td>V</td>
<td>$7.1 \cdot 10^{-2}$</td>
<td>Output voltage</td>
</tr>
<tr>
<td>$a_d$</td>
<td>rad/s</td>
<td>$7 \cdot 10^{-3}$</td>
<td>Dither voltage amplitude</td>
</tr>
<tr>
<td>$\omega_d$</td>
<td>rad/s</td>
<td>$2\pi \cdot 10^5$</td>
<td>Dither frequency</td>
</tr>
<tr>
<td>$G_{vco}$</td>
<td>rad/s/V</td>
<td>$2\pi \cdot 10^6$</td>
<td>VCO sensitivity</td>
</tr>
<tr>
<td>$\phi_d$</td>
<td>rad</td>
<td>$7 \cdot 10^{-2}$</td>
<td>Dither phase amplitude</td>
</tr>
<tr>
<td>$G_{pwr}$</td>
<td>V/W</td>
<td>$1.43 \cdot 10^5$</td>
<td>Power detector gain</td>
</tr>
<tr>
<td>$Z_0$</td>
<td>$\Omega$</td>
<td>50</td>
<td>Transmission line impedance</td>
</tr>
<tr>
<td>$C_1$</td>
<td>V$^{-1}$</td>
<td>2870</td>
<td>$G_{pwr}/Z_0$</td>
</tr>
<tr>
<td>$A_d$</td>
<td>V</td>
<td>10</td>
<td>Amplitude for demodulation</td>
</tr>
<tr>
<td>$K_{dmd}$</td>
<td>V</td>
<td>10</td>
<td>Multiplier reference voltage</td>
</tr>
<tr>
<td>$C_2$</td>
<td>scalar</td>
<td>1</td>
<td>$A_d/K_{dmd}$</td>
</tr>
<tr>
<td>$K_{PD}$</td>
<td>V/rad</td>
<td>1</td>
<td>Phase detector gain</td>
</tr>
</tbody>
</table>

The fraction on the right side of (4.28) can be considered the phase detector gain $K_{PD}$ of the OPLL in V/rad. As an example, table 4.1 shows feasible values for the various parameters of $K_{PD}$. In a dither loop, the phase detector is a distributed device, extending over the entire system. Phase sensing occurs in the photodiodes, but since no optical carrier has been transmitted, the receiver output does not contain a direct phase error signal. By phase dithering and synchronous demodulation, it is determined in which direction the phase error $\phi_c(t)$ decreases, and the LO phase is tracked accordingly.

### 4.1.1 Loop simulation and verification

A Simulink model has been developed to verify the calculations of the previous and the following sections. The model uses the values of table 4.1, but with $a_d$ (and therefore $\phi_d$) changed for some simulations. In
4.1. Dither loop analysis

Figure 4.2: Simulink OPLL model.

In this case, $G_{pwr}$ (and therefore $C_1$) have been changed as well to achieve a phase detector gain of $K_{PD} = 1 \text{ V/rad}$ for all simulations. The Simulink model is depicted in fig. 4.2.

The loop dynamics are determined by a first order active filter as described in Section 3.1.2. The time constants of the filter have been chosen so that the loop has a natural frequency of $f_n = 10 \text{ kHz}$, a damping factor of $\zeta = 1/\sqrt{2}$ and thus, a noise bandwidth of $B_n = 33.3 \text{ kHz}$. The dither frequency is ten times the natural frequency, i.e. $f_d = 100 \text{ kHz}$. The overall loop gain amounts to $G_{loop} = K_{PD} \cdot G_{vco}$. The loop filter provides proportional/integral control, i.e. it will track out any phase deviation until $\phi_x(t)$ becomes zero. Since the system contains two integrators - one in the loop filter and one in the VCO - the loop has a second order transfer function.

A fixed step solver calculates the internal states of the integrator and the filters. The step size $\Delta t$ amounts to $(100 \cdot f_d)^{-1}$. The maximum simulated frequency for noise spectra (see Section 4.2) is $1/(2\Delta t) = 5 \text{ MHz}$. The required frequency resolution $\Delta f$ determines the simulation length $T_{sim} = 1/\Delta f$.

The band-pass filter following the power detector consists of a 20 kHz high-pass and a 500 kHz low-pass filter, both of the first order. Inevitably, the band-pass filter introduces small phase deviations which
Figure 4.3: Simulated open loop demodulator output voltage in time and frequency domain, for a constant phase error of 0.1 rad. The continuous-type spectrum, rather than a discrete Fourier series, is a simulation artifact (e.g. due to the finite simulation time and the limited accuracy of the integration algorithm).

can deteriorate the error suppression capability of the OPLL. In that sense, the high-pass/low-pass combination showed better performance in simulations than a regular first order band-pass filter. The effect of the band-pass filter on the closed-loop transfer function can be neglected.

During the first 200 µs of a simulation, the system is in open loop condition to wait for the band-pass filter transient to decay. In the following 200 µs, phase locking occurs. It is then determined whether the loop locked to $\phi_e(t) = 0$ or $\phi_e(t) = \pi/2$. In the latter case, the LO phase is shifted by 180°. Another 400 µs is awaited for the system to reach stable operation conditions. Any signal measurement considers only values calculated after this initial sequence.

Figure 4.3 shows the demodulator output voltage and its Fourier transform for a constant phase error of 0.1 rad.\(^4\) For this simulation, the system is kept in open loop condition. As can be expected for a phase detector gain of 1 V/rad and phase error of 0.1 rad, the DC output voltage amounts to 0.1 V. The demodulator signal contains major high
\(^4\)In the ordinate of fig. 4.3(b) and throughout this thesis, $\log(\cdot)$ indicates the logarithm to the base 10. The unit of V refers to the argument of the logarithm.
4.1. Dither loop analysis

![Graphs showing phase error and VCO phase](image)

(a) Phase error. (b) VCO phase.

**Figure 4.4:** Dither loop step response. Solid: Simulink OPLL model. Dash-dotted: Calculated from the error and closed-loop transfer function of a generalized PLL model. The time axis has been shifted so that the step occurs at $t = 0$ s.

frequency components – particularly at twice the dither frequency – which will be suppressed by the closed-loop transfer function.

The closed-loop step responses of the phase error and the VCO phase are depicted in fig. 4.4. The step response of a generalized PLL model (see Section 3.1) with equal characteristic parameters ($\zeta$, $f_n$, $K_{PD}$) and the same loop filter as the Simulink OPLL model is plotted for comparison. Except for the superimposed dither signal, the step responses match well.

### 4.1.2 Stable operating point for non-zero $\phi_e(t)$

A stable operating point shall be defined as a condition where the loop filter output remains on a constant and finite level.\(^5\) For a second order PLL, the only stable operating point is expected to exist for a loop filter output voltage of zero. Only then, the LO phase remains unchanged, preserving the phase-lock state ($\phi_e(t) = 0$) if it existed before. Likewise, phase locking is essential for a phase detector output of zero, and thus, for a non-changing loop filter output.

\(^5\)Stability, as it is usually described in control theory [50], is not an issue here. This section describes a phenomenon sometimes called ‘false lock’ [47].
Simulations with the previously described Simulink OPLL model have shown that the dither loop can reach a stable operating point although the loop filter output (and likewise, $\phi_c(t)$) is not equal to zero. The following analysis investigates this phenomenon. A heuristic explanation will be given later.

First, it is assumed that the loop filter feeds a DC value of $\omega_d/(2 \cdot G_{vco})$ into the LO frequency tuning input. Neglecting phase noise expressions, the phase output of the VCO then becomes:

$$\phi_{LO}(t) = G_{vco} \int_{-\infty}^{t} \left(u_C + a_d \sin (\omega_d \tau)\right) d\tau$$

$$= \omega_d t/2 - \phi_0 - \phi_d \cos (\omega_d t) \tag{4.29}$$

where $\phi_0$ denotes the initial phase of the LO. Arbitrarily, $\phi_0$ shall amount to

$$\phi_0 = \pi/4 + \Delta \phi, \tag{4.30}$$

where $\Delta \phi$ is small. With (4.29), the receiver output (4.17) oscillates at half the dither frequency and, with much smaller amplitudes, at the dither frequency and higher order harmonics:

$$u_{out}(t) = U_0 d(t) \cos \phi_E(t)$$

$$= U_0 d(t) \cos \left(\omega_d t/2 - \pi/4 - \Delta \phi - \phi_d \cos (\omega_d t)\right). \tag{4.31}$$

The power detector output upon an input signal (4.31) is

$$u_{pwr}(t) = \frac{U_0^2 C_1}{2} \left(1 + \sin \left(\omega_d t - 2\Delta \phi - 2\phi_d \cos (\omega_d t)\right)\right)$$

$$= \frac{U_0^2 C_1}{2} \left(1 + \sin \left(\omega_d t - 2\Delta \phi\right) \cos \left(2\phi_d \cos (\omega_d t)\right)\right)$$

$$+ \cos \left(\omega_d t - 2\Delta \phi\right) \sin \left(2\phi_d \cos (\omega_d t)\right). \tag{4.32}$$

It can be reasoned with the Bessel series (4.20) and (4.21) that the major component of (4.32) at the dither frequency amounts to

$$u_{pwr}(t) \approx \frac{U_0^2 C_1}{2} \cdot \sin (\omega_d t - 2\Delta \phi) J_0(2\phi_d)$$

$$\approx \frac{U_0^2 C_1}{2} \cdot \sin (\omega_d t - 2\Delta \phi) \tag{4.33}$$
4.1. Dither loop analysis

\[ u_{dmd}(t) \approx \frac{U_0^2 C_1 C_2}{2} \cdot \Delta \phi. \]  

(4.34)

If \( \Delta \phi \) is zero, the phase detector output is zero and the loop filter output remains at \( \omega_d/(2 \cdot G_{vco}) \). The loop has reached a stable operating point for \( \phi_e(t) = \omega_d t/2 \). When comparing (4.34) with (4.28), it is interesting to note that the dither amplitude \( \phi_d \) has disappeared in the demodulator output voltage expression. Feeding the dither signal into the LO is not needed to maintain this type of phase lock.

In this operating condition, the TX and LO lasers in combination act as the VCO, oscillating at half the dither frequency. The power detector doubles this frequency to \( \omega_d \). Phase detection occurs in the multiplier, where the power detector output is referred to the dither source. When the laser intermediate frequency amounts to \( \omega_d/2 \) and the input signals at the multiplier are 90° out of phase, then the loop is locked in its secondary stable operating point, \( \phi_e(t) = \omega_d t/2 \).

Whether the OPLL reaches its intended state of operation (\( \phi_e(t) = 0 \)) or the one described in this section, depends on the initial LO phase \( \phi_0 \) and on the amplitude of the dither signal \( \phi_d \). Figure 4.5 shows two simulated phase-locking events, where only the initial LO phase has been altered. In 4.5(a), the initial phase is \( \pi/8 \) and the OPLL locks to

**Figure 4.5:** Simulated loop filter output signal. Left: Initial LO phase \( \phi_0 = \pi/8 \). Right: \( \phi_0 = \pi/4 \). The loop is closed at \( t = 0.2 \) ms.

which, after synchronous demodulation, leads to a value at low frequencies of
its intended state of $\phi_e(t) = 0$. With an initial phase of $\pi/4$, shown in fig. 4.5(b), the loop locks to $\phi_e(t) = \omega_d t/2$. The DC value of the loop filter output amounts to 0.05 V, which, with the given values, is equal to $\omega_d/(2 \cdot G_{vco})$.

The dependency of the phase-locking behavior on the dither amplitude and the initial LO phase can be seen in fig. 4.6. To achieve reliable phase locking to $\phi_e(t) = 0$, either the initial phase $\phi_0$ should be close to 0 or $\pi$, or the dither amplitude $\phi_d$ has to be chosen larger than 3.8°. In the intermediate region, the loop behaves somewhat chaotic. It is important to note that the phase detector gain $K_{PD}$ has been constant for all simulations, i.e. increasing $a_d$ (to increase $\phi_d$) implied decreasing $G_{pwr}$.

In a real world dither loop receiver, it might be difficult to control the exact initial phase $\phi_0$ at which phase locking occurs. Therefore, the system should be designed with a dither amplitude of around 4° or more to obtain reliable phase locking to $\phi_e(t) = 0$.

### 4.2 Noise in a dither loop

Noise affects a coherent receiver in two ways. Shot noise directly flows into the data detection branch, generating bit errors at low levels of the signal-to-noise ratio (SNR). Phase noise, and other noise signals which flow into the phase-locking branch, lead to a residual phase error.
4.2. Noise in a dither loop

The phase error inevitably reduces the signal amplitude, and therefore deteriorates the SNR.

According to the closed-loop transfer function of a PLL, the feedback loop can track out low frequency variations of the phase input (i.e. phase noise). Likewise, due to the error transfer function (rejection at high frequencies), high frequency components of internal noise sources (i.e. shot noise) are suppressed. The amount of noise suppression is mainly determined by the loop noise bandwidth $B_n$.

The following list identifies the major noise sources that contribute to the total phase error $\phi_E(t)$ in a dither loop:

- **LO phase dithering:** The dither signal is not noise in a strict sense, due to its deterministic nature. Furthermore, it is intentionally applied to the loop. Still, for mathematical coherence, it is convenient to treat the dither signal like an internal noise source of the receiver. The phase error due to LO phase dithering can be considered as the power penalty one has to pay for phase locking.

- **Phase noise:** Phase noise has been introduced in Section 2.2.1. It is generated by the spontaneous emission of photons in the laser cavity. It will be shown that the phase error variance due to phase noise is inversely proportional to the loop noise bandwidth $B_n$.

- **Shot noise:** Due to the shot noise limit, introduced in Section 2.3.3, shot noise contributions can become considerably large. The shot noise signal generates low frequency components at the loop filter input, and therefore modulates the LO phase $\phi_{LO}(t)$. The phase error variance due to shot noise is proportional to $B_n$.

- **1/f frequency noise:** The laser phases are not only affected by white frequency noise, but also by 1/f frequency noise (‘frequency flicker noise’). The 1/f frequency noise can be suppressed by thermally stabilizing the laser cavity and/or by additional poles in the loop transfer function. The following analysis will not consider 1/f or higher order frequency noises.

- **LO intensity noise:** It has been shown in [53] that LO intensity noise can be greatly reduced by using a balanced receiver.
Therefore, it will not be considered here. To achieve good noise suppression over the whole frequency range of the receiver, it is important to have equal path lengths (with tolerances in the millimeter range) between the coupler and the two photodiodes.

The influence of LO phase dithering, white frequency noise induced phase noise and shot noise on the total phase error $\phi_E(t)$ is analyzed in the next sections. To ease the mathematics, the dither loop is assumed to be a linear time-invariant (LTI) system. The linearity assumption implies that the variance $\sigma_E$ of the total phase error $\phi_E(t)$ remains small, i.e. $\sigma_E \leq 10^6$. As shown in table 4.2 of Section 4.4.2, this can be satisfied for well designed loops.

### 4.2.1 LO phase dithering

Generally written, the phase error variance is defined as

$$\sigma_E^2 = E[\phi_E^2(t)] - E^2[\phi_E(t)]$$

where the expected or mean value of $\phi_E(t)$ is

$$E[\phi_E(t)] = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \phi_E(t) dt.$$  

(4.36)

Strictly speaking, (4.36) is only valid for ergodic noise processes. For the dither signal to become ergodic, the initial phase of the dither source must be random. This can be safely assumed without affecting the validity of the analysis.

Neglecting any noise sources except the dither signal, the phase error signal can be written as

$$\phi_E(t) = \phi_d \cos(\omega_dt).$$

(4.37)

Now, since $E[\cos(\cdot)]$ is zero and $E[\cos^2(\cdot)]$ equals $1/2$, evaluating (4.37) in (4.35) using (4.36) yields

$$\sigma_{E,dither}^2 = \frac{\phi_d^2}{2} = \frac{1}{2} \cdot \left( \frac{a_d G_{vco}}{\omega_d} \right)^2$$

(4.38)

with units of rad$^2$. Formula (4.38) describes the phase error variance due to LO phase dithering in the absence of any other noise sources.
4.2. Noise in a dither loop

Figure 4.7(a) depicts the Fourier transformation of a simulated phase error signal in closed-loop mode. Integrating the spectrum yields a value which deviates less than 4% from the predicted value of (4.38). The origin of the ‘noise floor’ at low frequencies lies in simulation artifacts.

4.2.2 Phase noise

When transmitting a stationary random process $x(t)$ with PSD $S_x(f)$ over a linear time-invariant (LTI) system with a transfer function $G(f)$, then the PSD of the output signal $y(t)$ becomes [40, 44]6

$$S_y(f) = S_x(f)|G(f)|^2.$$  \hfill (4.39)

With (4.39), the phase error variance due to laser phase noise can be calculated through

$$\sigma_{E, pn}^2 = \int_0^\infty S_{pn}(f)|1 - H(f)|^2 df$$ \hfill (4.40)

where $|1 - H(f)|$ is the error transfer function of the loop. In (4.40), $S_{pn}(f)$ denotes the one-sided power spectral density (PSD) of the phase noise process:

$$S_{pn}(f) = \frac{2\Delta \nu}{\pi f^2}$$ \hfill (4.41)

with units of rad$^2$/Hz, and the laser linewidth $\Delta \nu$ in Hz. The factor of 2 in (4.41) (compared to equation (2.10) in Section 2.2.1) is due to the fact that the total phase noise process is generated by two individual lasers. The solution of the integral (4.40) is presented in appendix A, the result has already been published for other loop types [21, 22, 87]:

$$\sigma_{E, pn}^2 = \frac{3\pi \Delta \nu}{4B_n}.$$  \hfill (4.42)

Formula (4.42) describes the phase error variance due to laser phase noise, in the absence of any other noise sources, and under the assumption that the dither loop is an LTI system. The phase error variance depends only on the laser linewidth and the noise bandwidth of the loop.

---

6 Known as the ‘Wiener-Lee Beziehung’ in German literature.
Figure 4.7: Fourier transform of the phase error signal due to various noise sources. 4.7(a): Power spectrum. 4.7(b)–4.7(d): Power spectral density. Solid: Simulink OPLL model. Dash-dotted: Calculated from error and closed-loop transfer functions and the PSD’s of phase noise and shot noise processes. Noise spectra are averaged 1000 times. Reference value for the dB scales is $1 \text{ rad}^2$. 
4.2. Noise in a dither loop

Verification of (4.42) in a simulation requires to shut down any noise sources except phase noise. Unfortunately, without the dither signal, the OPLL cannot acquire or maintain phase lock. Assuming linearity between individual noise sources, phase noise can be simulated in the presence of the dither signal. This is shown in fig. 4.7(b). Integrating over the spectrum and subtracting the phase error variance due to phase dithering (i.e. fig. 4.7(a)), yields a value which deviates by 11% from the prediction of (4.42). For frequencies between $f_n$ and $f_d$, the simulated phase error lies slightly above the calculated trace, indicating that the dither loop is not exactly behaving like a second order linear PLL. This is generated by a small phase delay which is introduced at the edges of the transmission band of the band-pass filter. To consider this phase lag in the calculation, a delay function could be introduced in the linearized loop model. As in the previous section, the noise floor at low frequencies is a simulation artifact. It has only a minor effect on the accuracy of the simulation.

Despite the limited accuracy, formula (4.42) is considered to be a sufficiently good approximation of the phase error variance due to laser phase noise. For the sake of simplicity, the effect of the band-pass filter on the closed-loop transfer function is neglected.

4.2.3 Shot noise

When taking shot noise into account, the receiver output (4.17) becomes

$$u_{out}(t) = U_0 d(t) \cos \phi_E(t) - R_T G_{RF}/2 \cdot n(t),$$

(4.43)

where $n(t)$ denotes the combined shot noise process of the two photodiodes with a one-sided PSD of

$$S_n(f) = 2qR_{LO}$$

(4.44)

in $\text{A}^2/\text{Hz}$. The shot noise signal flows through the power detector, band-pass filter and demodulator into the loop filter. The shot noise signal at the loop filter input will be denoted by $n_{sn}(t)$ with an associated PSD of $S_{sn}(f)$. The following calculation focuses on an expression for $S_{sn}(f)$. It is kept in mind (but not explicitly written) that the shot noise PSD $S_n(f)$ is limited by the front end cut-off frequency $B_{RF}$. The system bit rate, which will be used in the following calculation, shall be written as $R_b$ in bit/s, and the bit interval as $T_b = 1/R_b$ in s.
When fed with (4.43), the power detector output yields

\[
upwr(t) = C_1 u^2_{out}(t) = C_1 \left( U_0 d(t) \cos(\phi_E) - R_T G_{RF} / 2 \cdot n(t) \right)^2
\]

\[
= C_1 \left( U_0^2 d^2(t) \cos^2(\phi_E(t)) - U_0 R_T G_{RF} d(t) n(t) \cos(\phi_E(t)) \right)
\]

\[
\quad + \frac{R_T^2 G_{RF}^2}{4} \cdot n^2(t). \tag{4.45}
\]

The first expression in (4.45) is further processed to extract the phase error, as developed in Section 4.1. The rest of (4.45) are shot noise components that possibly affect the loop operation. For the second expression, called the signal×noise beat note, the following approximations are made:

\[
\cos(\phi_E(t)) \approx 1 \tag{4.46}
\]

\[
S_{[d(t)n(t)]}(f) \approx S_n(f) \tag{4.47}
\]

\[
S_{[n(t)\cos(\omega_d t)]}(f) \approx S_n(f) / 2. \tag{4.48}
\]

The first line holds true when the total phase error remains small. This is assumed throughout this chapter. To justify (4.47), it has to be kept in mind that \(d(t)\) is a random binary NRZ signal which takes only values in \(\{-1, 1\}\). According to [88], the PSD of \(d(t)\) computes to

\[
S_d(f) = T_b \cdot \frac{\sin(\pi f T_b)}{\pi f T_b} \tag{4.49}
\]

with the main power content in frequencies below \(R_b\). The convolution of \(S_d(f)\) with \(S_n(f)\) yields approximately \(S_n(f)\) as long as \(R_b \leq B_{RF}\). The third approximation, equation (4.48), states that a frequency shift does not affect the power spectral density of a white noise process, except for a multiplicative factor of 1/2. This becomes apparent when bearing in mind that the power spectrum of \(\cos(\cdot)\) is a delta function with area 1/2. Applying (4.46)–(4.48) to the signal×noise beat note, the relevant multiplier output becomes

\[
n_{sn1}(t) = \frac{U_0 C_1 C_2 R_T G_{RF}}{\sqrt{2}} \cdot n(t) \tag{4.50}
\]
with a PSD of

\[ S_{sn1}(f) = \left( \frac{U_0 C_1 C_2 R_T G_{RF}}{\sqrt{2}} \right)^2 \cdot S_n(f). \] (4.51)

This is one of two shot noise expressions that flow into the OPLL and possibly affect its performance. The second expression, called the noise\times noise beat note, will be calculated shortly.

If \( x(t) \) denotes a stationary normal process with zero mean and a PSD of \( S_x(f) \), then, according to [29], the PSD of

\[ y(t) = x^2(t) \]

computes to

\[ S_y(f) = R_x^2(0) \delta(f) + 2S_x(f) \star S_x(f). \] (4.52)

For further processing, it is assumed that \( S_x(f) \) is white up to a maximum frequency \( f_{\text{max}} \), and zero beyond \( f_{\text{max}} \). Then, the first expression in (4.52), \( R_x^2(0) \), denotes the square of the total power of \( x \), i.e. \( (S_x f_{\text{max}})^2 \). The convolution in (4.52) yields a ramp with a maximum value of \( S_x^2 f_{\text{max}} \) at \( f = 0 \) and a minimum value of 0 at \( f = 2f_{\text{max}} \). Thus:

\[ S_y(f) = \begin{cases} (S_x f_{\text{max}})^2 \delta(f) + 2S_x^2 (f_{\text{max}} - \frac{f}{2}) & f \leq 2f_{\text{max}}, \\ 0 & \text{else.} \end{cases} \] (4.53)

Equation (4.53) can be used to calculate the noise contribution of the noise\times noise beat note in (4.45). The PSD of the squared shot noise process, \( n^2(t) \), becomes

\[ S_{[n^2(t)]}(f) = \begin{cases} (S_n B_{RF})^2 \delta(f) + 2S_n^2 (B_{RF} - \frac{f}{2}) & f \leq 2B_{RF}, \\ 0 & \text{else} \end{cases} \] (4.54)

with units of \( \text{A}^4/\text{Hz} \). The band-pass filter removes the DC-term and the high frequency part of the ramp function. Furthermore, it has to be recalled that \( B_{RF} \gg f_d \), so that \( (B_{RF} - f/2) \approx B_{RF} \) for frequencies around \( f_d \). Thus, the PSD of the relevant shot noise signal results in

\[ S_{sn2}(f) = \left( \frac{C_1 C_2}{\sqrt{2}} \left( \frac{R_T G_{RF}}{2} \right)^2 \right)^2 \cdot 2S_n^2(f) B_{RF}. \] (4.55)
This is the second of two shot noise expressions that possibly affect the loop operation.

Dividing the two PSD’s (4.51) and (4.55) (and using 4.44) can help to find the dominant shot noise expression:

$$\frac{S_{sn1}(f)}{S_{sn2}(f)} = \frac{4RP_S}{qB_{RF}}$$

which is plotted in fig. 4.8. It can be seen that, for a received power of -50 dBm and a front end bandwidth of 10 GHz, $S_{sn1}$ dominates $S_{sn2}$ by more than a factor of ten. Thus, it is safe to assume that the relevant shot noise PSD is generated by the signal×noise beat note:

$$S_{sn}(f) = S_{sn1}(f) = \left( \frac{U_0C_1C_2R_TG_{RF}}{\sqrt{2}} \right)^2 \cdot S_n(f).$$

At significantly lower levels of $P_S$, though, $S_{sn1}$ and $S_{sn2}$ are comparable in size and both expressions have to be taken into account.

If $x(t)$, with associated Laplace transform $X(s)$, denotes an additive signal to the loop filter input, then it can be reasoned with basic systems analysis that the transfer function of $x(t)$ to the phase error $\phi_E(t)$ computes to

$$\frac{\Phi_E(s)}{X(s)} = -\frac{H(s)}{K_{PD}}$$

where $H(s)$ is the closed-loop transfer function of the loop, as defined in Section 3.1. With (4.39) and (4.58), the phase error variance due to
shot noise can be calculated through

\[ \sigma^2_{E,sn} = \int_0^\infty \frac{S_{sn}(f)}{K_{PD}^2} |H(f)|^2 df. \]  \hfill (4.59)

Using (4.28) and (4.57) and remembering that, by definition,

\[ B_n = \int_0^\infty |H(f)|^2 df \]  \hfill (4.60)

(4.59) becomes

\[ \sigma^2_{E,sn} = \frac{q}{RP_S\phi_d^2} B_n = \frac{q\omega_d^2}{RP_Sa_d^2G_{vco}^2} B_n. \]  \hfill (4.61)

Formula (4.61) describes the phase error variance due to the photodiode shot noise process, in the absence of any other noise sources.

A simulation of a shot noise perturbed dither loop is shown in fig. 4.7(c). Integrating over the spectrum, and subtracting the phase error variance due to phase dithering, yields a value which lies within a 2% accuracy range to the predicted value of (4.61). It can be assumed from this good agreement that the dither loop is a linear system concerning shot noise injection. For completeness, fig. 4.7(d) depicts a loop simulation with all three noise sources switched on. The result matches well with the calculated trace.

### 4.3 Performance degradation due to noise

In the preceding Sections 4.2.1 – 4.2.3, expressions for the phase error variance due to various noise sources have been developed. It has been stated before that a non-zero phase error will reduce the SNR and therefore increase the bit error rate. This could be compensated by increasing the transmitted power, so that a certain specified bit error rate can be sustained. The power penalty shall be defined as the ratio of the deteriorated SNR to the SNR of an ideal (undisturbed) receiver. The next sections focus on expressions for the power penalty induced by phase dithering, phase noise and shot noise.
4.3.1 Phase dither power penalty

The photodiode current (4.8), considering the phase expressions (4.11) – (4.15) and including shot noise, can be written as

\[ i_1(t) = 2R \sqrt{P_s P_{LO}} \cdot d(t) \cos \phi_E(t) + n(t) \]
\[ = I_0 d(t) \cos \phi_E(t) + n(t). \]  (4.62)

If the optical signals on the photodiode surface are strong, the shot noise process \( n(t) \) is usually described as Gaussian distributed and white with a PSD of (4.44) \([20, 36]\). In a coherent receiver, the condition for strong optical signals can easily be satisfied. For further calculation, it is assumed that the signal \( i_1(t) \) passes a low-pass filter with a bandwidth \( R_b \), where \( R_b \) denotes the system bit rate in bit/s. A slicer estimates the data signal \( d(t) \) at constant time intervals \( T = 1/R_b \). The system bit rate has to be be much greater than the dither frequency \( f_d \). Phase noise or shot noise contributions to the total phase error \( \phi_E(t) \) are not considered in the following analysis.

According to [40], the input signal of the slicer at the sampling time \( t_s \), \( y_d(t_s) \), is a random variable with a mean value \( I_0 d(t_s) \cos \phi_E(t_s) \) and a variance of \( N = 2qR P_{LO} R_b \). For a logic one, this random variable has a probability density function (PDF) of\(^7\)

\[ p_{y_1}(x) = \frac{1}{\sqrt{\pi N}} \exp \left(-\left(x - I_0 \cos \phi_E(t_s)\right)^2/N\right). \]  (4.63)

For a logic zero, only the sign of the mean value changes in (4.63). With the PDF, the bit error rate can be written as

\[ \text{BER} = \Pr [y_1(t_s) < 0] = \Pr [y_0(t_s) > 0] \]
\[ = \int_{-\infty}^{0} p_{y_1}(x)dx = \int_{0}^{\infty} p_{y_0}(x)dx \]  (4.64)

which is the probability that, at the sampling time, the slicer input becomes less than zero although a one bit has been transmitted, or

\(^7\)The PDF follows from the probability function \( P_{y_1}(x) \), which denotes the probability that \( y_1(t_s) \) is smaller or equal to a certain threshold \( x \): \( P_{y_1}(x) = \Pr [y_1(t_s) \leq x] \). Taking the derivative yields the PDF: \( p_{y_1}(x) = dP_{y_1}(x)/dx \).
4.3. Performance degradation due to noise

vice versa. Evaluating (4.63) in (4.64), and substituting \( \xi = (x - I_0 \cos \phi_E(t_s))/\sqrt{N} \), yields

\[
\text{BER}(t_s) = \frac{1}{2} \text{erfc} \left( \frac{I_0 \cos \phi_E(t_s)}{\sqrt{N}} \right) = \frac{1}{2} \text{erfc} \left( \sqrt{\frac{2RPS \cos^2 \phi_E(t_s)}{qR_b}} \right) \tag{4.65}
\]

where \( \text{erfc}(\cdot) \) denotes the complementary error function defined as

\[
\text{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_{z}^{\infty} e^{-\xi^2} d\xi. \tag{4.66}
\]

Some publications [20, 22] rather use the Gaussian Q function than the \( \text{erfc}(\cdot) \) function:

\[
Q(z) = \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} e^{-\xi^2/2} d\xi = \frac{1}{2} \text{erfc} \left( \frac{z}{\sqrt{2}} \right) \tag{4.67}
\]

so that (4.65) becomes

\[
\text{BER}(t_s) = Q \left( 2 \sqrt{\frac{RPS \cos^2 \phi_E(t_s)}{qR_b}} \right). \tag{4.68}
\]

The bit error rate in a dither loop depends on the time, because of the time varying LO phase. The rest of this analysis focuses on an average BER over a single dither period, and likewise, an average signal amplitude that determines the average bit error rate.

The \( \cos^2 \phi_E(t) \) expression in (4.65), with \( \phi_E(t) = \phi_d \cos(\omega_d t) \) and using (4.20), can be calculated to

\[
\cos^2 \phi_E(t) = \frac{1}{2} \left( 1 + \cos 2\phi_E(t) \right) = \frac{1}{2} \left( 1 + \cos \left( 2\phi_d \cos(\omega_d t) \right) \right) = \frac{1}{2} \left( 1 + J_0(2\phi_d) - 2J_2(2\phi_d) \cos(2\omega_d t) + \ldots \right) \tag{4.69}
\]

with a time average of

\[
E \left[ \cos^2 \phi_E(t) \right] = \frac{1}{2} \left( 1 + J_0(2\phi_d) \right). \tag{4.70}
\]
Using the relationship

\[ J_\nu(z) = (z/2)^\nu \sum_{k=0}^{\infty} \frac{(-z^2/4)^k}{k!\Gamma(\nu + k + 1)} \]  \hspace{1cm} (4.71)

from [86], (4.70) becomes

\[ E \left[ \cos^2 \phi_E(t) \right] = \frac{1}{2} \left( 1 + 1 - \phi_d^2 + \frac{\phi_d^4}{4} - \ldots \right) \approx 1 - \frac{\phi_d^2}{2}. \]  \hspace{1cm} (4.72)

With (4.72), the average bit error rate during a single dither period is approximated by

\[ \text{BER} \approx \frac{1}{2} \text{erfc} \left( \sqrt{\frac{2RP_s(1 - \phi_d^2/2)}{qR_b}} \right). \]  \hspace{1cm} (4.73)

The expression

\[ \epsilon_{\text{dither}}(\phi_d) = 1 - \phi_d^2/2 = 1 - \sigma_{E,\text{dither}}^2 \]  \hspace{1cm} (4.74)

can be considered the power penalty one has to pay for phase locking. Likewise, \( \phi_d^2/2 \) is the fraction of the received power that is fed into the phase locking path. Equation (4.74) is plotted in fig. 4.9(a). For example, a dither amplitude of \( \phi_d = 10^5 \) results in a power penalty of 0.07 dB.

It has been assumed in the foregoing analysis that the system bit rate \( R_b \) exceeds the dither frequency \( f_d \) by several orders of magnitude. From this it follows that for a consecutive trail of a certain number of bits, the signal amplitude, and therefore the bit error rate, does not change significantly. Thus, the signal amplitude during one dither period can be discretized with a sampling rate \( M \), where \( f_d < M \ll R_b \). By calculating the BER for each sample individually and averaging over all samples, equation (4.73) can be verified numerically. The relative error of (4.73) from the numerical calculation is shown in fig. 4.9(b). For any reasonable value of \( \phi_d \), the relative error stays within an acceptable limit. In Section 4.3.3.3, a more accurate way to calculate the power penalty due to phase dithering will be presented. The method of this section has the advantage of a very low computational effort.
4.3. Performance degradation due to noise

Figure 4.9: Left: The power penalty due to LO phase dithering. Right: Relative error of the bit error rate approximation (4.73), compared to a numerical simulation. The BER level given in the figure legend determines the SNR of the undisturbed receiver ($\phi_d = 0$).

The power penalty due to phase locking is a unique property of a PLL and can be used for comparison with other designs. In a balanced loop, it amounts to

$$\epsilon_{\text{balanced}} = \sin^2 \phi_{\text{mod}} = 1 - \cos^2 \phi_{\text{mod}}, \quad (4.75)$$

where $\phi_{\text{mod}}$ denotes the modulation depth of the optical carrier [22]. The power penalty in a Costas loop is determined by the power splitting ratio $\kappa$ of the 90° coupler [21]:

$$\epsilon_{\text{Costas}} = 1 - \kappa. \quad (4.76)$$

The similarity of equations (4.74), (4.75) and (4.76) further support the evaluations made thus far.

4.3.2 Phase error power penalty

It follows from (4.65) that the highest possible SNR in a coherent receiver amounts to

$$\text{SNR}_{\text{max}} = \frac{2RP_S}{qR_b} = \rho_0^2 \quad (4.77)$$
which is achieved for perfect phase locking, i.e. $\phi_E(t) = 0$. If the carrier phase recovery is deteriorated by the residual phase error $\phi_e(t)$, then the bit error rate computes to

$$\text{BER} = \frac{1}{2} \text{erfc} \left( \rho_0 \cdot \cos \phi_e(t) \right). \quad (4.78)$$

For (4.78), it has been assumed that the dither signal is switched off and the lasers are locked by some other technique. It is known from [89] that (4.78) can be written as

$$\text{BER} = \frac{1}{2} \text{erfc} (\rho_0) + \sum_{i=0}^{\infty} (-1)^i H_i B_i \quad (4.79)$$

with

$$H_i = \frac{\rho_0 e^{-\rho_0^2/2}}{\sqrt{\pi} (2l + 1)} \left( I_l(\rho_0^2/2) + I_{l+1}(\rho_0^2/2) \right) \quad (4.80)$$

$$B_i = 1 - E \left[ \cos \left( (2l + 1) \phi_e(t) \right) \right] \quad (4.81)$$

where $I_l(\cdot)$ denotes the modified Bessel function of the first kind of order $l$. Equation (4.81) is valid for an arbitrary distribution function of $\phi_e(t)$. In an OPLL, $\phi_e(t)$ contains a linear combination of the shot noise and the phase noise process. These processes are presumed to be statistically independent and Gaussian distributed. Then, according to [40], the sum process has a Gaussian distribution function, zero mean, and a variance of

$$\sigma_e^2 = \sigma_{E,\text{noise}}^2 = \sigma_{E,\text{pn}}^2 + \sigma_{E,\text{sn}}^2. \quad (4.82)$$

For a Gaussian distributed random variable $x$ with mean $\eta$ and variance $\sigma^2$, the following relationship holds true [29]:

$$E[e^{j\omega x}] = e^{j\eta\omega - \sigma^2 \omega^2/2}. \quad (4.83)$$

The left side of (4.83) is called the characteristic function of $x$. Likewise, it is the Fourier transform of its distribution function.
4.3. Performance degradation due to noise

With \( \cos(z) = \frac{1}{2} \cdot (e^{jz} + e^{-jz}) \), (4.82) and (4.83), (4.81) becomes

\[
B_l = 1 - \exp\left(-\frac{(2l + 1)^2 \sigma_e^2}{2}\right). \tag{4.84}
\]

The formulas (4.79), (4.80) and (4.84) determine the bit error rate \( \text{BER}(\rho_0, \sigma_e) \) in a coherent receiver with imperfect carrier phase recovery. The associated signal-to-noise ratio amounts to

\[
\text{SNR} = \rho^2 = \left(\text{erfc}^{-1}\left(2 \cdot \text{BER}(\rho_0, \sigma_e)\right)\right)^2 \tag{4.85}
\]

where \( \text{erfc}^{-1}(\cdot) \) denotes the inverse of the complementary error function, i.e. when \( y = \text{erfc}(x) \) then \( x = \text{erfc}^{-1}(y) \). The power penalty due to phase noise and shot noise contributions to the total phase error \( \phi_E(t) \) becomes

\[
\epsilon_{\text{noise}}(\rho_0, \sigma_e) = \frac{\left(\text{erfc}^{-1}\left(2 \cdot \text{BER}(\rho_0, \sigma_e)\right)\right)^2}{\rho_0^2} \tag{4.86}
\]

where the bit error rate has been calculated through (4.79), (4.80) and (4.84). In (4.86), It should be noted that \( \epsilon_{\text{noise}} \) depends only on the SNR \( \rho_0^2 \) of the ideal (i.e. zero phase error) receiver and the residual phase error standard deviation \( \sigma_e \).

4.3.3 Overall power penalty

In a dither loop, the total phase error \( \phi_E(t) \) consists of a stationary random process \( \phi_e(t) \) and the deterministic dither signal:

\[
\phi_E(t) = \phi_e(t) + \phi_d \cos(\omega_dt). \tag{4.87}
\]

Consequently, equation (4.81) from the previous section results in

\[
B_l = 1 - \mathbb{E}\left[\cos\left((2l + 1)\phi_E(t)\right)\right]
\]

\[
= 1 - \mathbb{E}\left[\cos\left((2l + 1)(\phi_e(t) + \phi_d \cos(\omega_dt))\right)\right]. \tag{4.88}
\]
Transformations on the argument of the expected value in (4.88) yield
\[
\cos ((2l + 1)\phi_e(t) + (2l + 1)\phi_d \cos (\omega_d t)) = \cos ((2l + 1)\phi_e(t)) \\
\cdot \cos ((2l + 1)\phi_d \cos (\omega_d t)) - \sin ((2l + 1)\phi_e(t)) \sin ((2l + 1)\phi_d \cos (\omega_d t)) \\
= \cos ((2l + 1)\phi_e(t)) \left( J_0((2l + 1)\phi_d) - J_2((2l + 1)\phi_d) \cos (2\omega_d t) + \ldots \right) \\
- \sin ((2l + 1)\phi_e(t)) \left( J_1((2l + 1)\phi_d) \cos (\omega_d t) - \ldots \right). \tag{4.89}
\]
In (4.89), all \(\cos (n\omega_d t)\) expressions have a mean value of zero for \(n \geq 1\), \(n \in \mathbb{N}\). Thus, (4.88) reduces to
\[
B_l = 1 - J_0((2l + 1)\phi_d) \quad \mathbb{E}\left[ \cos ((2l + 1)\phi_e(t)) \right] \\
= 1 - J_0((2l + 1)\phi_d) \exp \left( -(2l + 1)^2 \sigma_e^2/2 \right) \tag{4.90}
\]
which defines, together with equations (4.79) and (4.80), the bit error rate \(\text{BER}(\rho_0, \sigma_e, \phi_d)\) in a dither loop. The overall power penalty due to phase locking and incomplete carrier phase recovery then amounts to
\[
\epsilon(\rho_0, \sigma_e, \phi_d) = \left( \frac{\text{erfc}^{-1}(2 \cdot \text{BER}(\rho_0, \sigma_e, \phi_d))}{\rho_0^2} \right)^2 \tag{4.91}
\]
where the bit error rate has been calculated through (4.79), (4.80) and (4.90). Results of (4.91) are presented in fig. 4.10 for SNR levels \(\rho_0^2\) of 10.5, 11.3, 12.0, 12.5, 13.1 and 13.9 dB. These SNR levels are equal to bit error rates of \(10^{-6}\), \(10^{-7}\), \(10^{-8}\), \(10^{-9}\), \(10^{-10}\) and \(10^{-12}\), respectively, of the ideal (i.e. zero phase error) homodyne receiver.

With the theory derived in this section, the simplified power penalty expression (4.74) can be verified. This is done by evaluating (4.91) for a residual phase error standard deviation of zero, so that the loop is only affected by phase dithering:
\[
\epsilon(\rho_0, 0, \phi_d) \approx \epsilon_{\text{dither}}(\phi_d). \tag{4.92}
\]
A comparison of figs. 4.9(a) and 4.10(d) for a power penalty value of 0.05 dB reveals a deviation of less than 2% for the two calculation methods. This is well within a reasonable accuracy range.
4.3. Performance degradation due to noise

![Graphs showing performance degradation due to noise](image)

(a) $\rho_0^2 = 10.5$ dB.
(b) $\rho_0^2 = 11.3$ dB.
(c) $\rho_0^2 = 12.0$ dB.
(d) $\rho_0^2 = 12.5$ dB.
(e) $\rho_0^2 = 13.1$ dB.
(f) $\rho_0^2 = 13.9$ dB.

**Figure 4.10:** Dependency of the power penalty from the dither amplitude $\phi_d$ and the residual phase error standard deviation $\sigma_e$. 
4.4 Dither loop synthesis

In the preceding sections, the relevant noise expressions, their effect on the phase error variance and the associated power penalties have been derived. This knowledge can be used to develop design rules for an optimized loop, with minimum phase error variance and hence, a minimum power penalty due to incomplete carrier phase recovery. The two main design parameters are the dither amplitude $\phi_d$ and the loop noise bandwidth $B_n$.

4.4.1 Phase error minimization

The total phase error variance is the sum of the variances due to the individual noise sources (equation (4.38), (4.42) and (4.61)):

$$\sigma_E^2 = \sigma_{E,dither}^2 + \sigma_{E,pn}^2 + \sigma_{E,sn}^2$$

$$\sigma_E^2(\phi_d, B_n) = \frac{\phi_d^2}{2} + \frac{3\pi \Delta \nu}{4B_n} + \frac{qB_n}{RP_S \phi_d^2}$$

(4.93)

provided that the noise processes are statistically independent. For (4.93), the central-limit theorem has been used [29,40]. Furthermore, it is assumed that the loop does not perform nonlinear processing between the noise signals, i.e. it does not create mixing products of two or more noises. It has been shown in Sections 4.2.1 – 4.2.3 that this assumption is, at least up to a certain degree, fulfilled.

Equation (4.93) has the two partial derivatives:

$$\frac{\partial \sigma_E^2}{\partial \phi_d} = \phi_d - \frac{2qB_n}{RP_S \phi_d^2}$$

(4.94)

$$\frac{\partial \sigma_E^2}{\partial B_n} = \frac{q}{RP_S \phi_d^2} - \frac{3\pi \Delta \nu}{4B_n^2}$$

(4.95)

Setting both to zero and solving for $\phi_d$ and $B_n$, respectively, yields:

$$\phi_d = \left(\frac{2qB_n}{RP_S}\right)^{1/4}$$

(4.96)

$$B_n = \sqrt{\frac{3\pi RP_S \Delta \nu \phi_d^2}{4q}}.$$
Plugging (4.96) into (4.97) and solving for $B_n$ results in:

$$B_{n,\text{opt}} = \frac{1}{2} \left( \frac{(3\pi)^2 R P_S \Delta \nu^2}{q} \right)^{1/3}$$  \hspace{1cm} (4.98)

and (4.98) in (4.96):

$$\phi_{d,\text{opt}} = \left( \frac{3\pi q \Delta \nu}{R P_S} \right)^{1/6}.$$  \hspace{1cm} (4.99)

Strictly speaking, the above analysis only shows that $\sigma_E^2(\phi_d, B_n)$ has some kind of extreme value at $(\phi_{d,\text{opt}}, B_{n,\text{opt}})$. A more thorough analysis, e.g. as in [90], would be required to determine the type of extreme value (minimum, maximum or saddle point). Instead, it is stated (and it has been numerically verified) that $(\phi_{d,\text{opt}}, B_{n,\text{opt}})$ defines the global minimum of $\sigma_E^2(\phi_d, B_n)$.

Evaluating the optimum values (4.98) and (4.99) in the variances of the individual noise sources yields

$$\sigma_{E,dither}^2 = \sigma_{E,\text{pn}}^2 = \sigma_{E,\text{sn}}^2 = \frac{1}{2} \left( \frac{3\pi q \Delta \nu}{R P_S} \right)^{1/3}$$  \hspace{1cm} (4.100)

so that each of the three expressions in (4.93) contribute the same amount of phase error. The total phase error variance becomes

$$\sigma_E^2 = \frac{3}{2} \left( \frac{3\pi q \Delta \nu}{R P_S} \right)^{1/3}$$  \hspace{1cm} (4.101)

and the variance due to the residual phase error $\phi_e(t)$ can be written as

$$\sigma_e^2 = \sigma_{E,\text{pn}}^2 + \sigma_{E,\text{sn}}^2 = 2\sigma_{E,dither}^2 = \phi_d^2.$$  \hspace{1cm} (4.102)

It has been shown in Section 4.3.3 (equation (4.91)) how to compute the power penalty in a dither loop for an arbitrary combination of $\phi_d$ and $\sigma_e$. Equation (4.102) puts $\phi_d$ and $\sigma_e$ into a fixed relation, so that the power penalty can be calculated as a function of the phase error standard deviation $\sigma_E$ alone:

$$\epsilon(\rho_0, \sigma_e, \phi_d) \rightarrow \epsilon(\rho_0, \sigma_E).$$  \hspace{1cm} (4.103)
Figure 4.11: The overall power penalty $\epsilon$ in an optimally designed dither loop as a function of the phase error standard deviation $\sigma_E$ and the bit error rate.

This can be done by changing equation (4.90) to

$$B_l = 1 - J_0\left(\sqrt{2/3} \cdot (2l + 1)\sigma_E\right) \exp\left(-2(2l + 1)^2\sigma_E^2/6\right), \quad (4.104)$$

or by reading the power penalty value from fig. 4.10 for $\sigma_e = \phi_d$. An example of (4.103) is depicted in fig. 4.11.

4.4.2 Dither loop design rule

The coherent receiver has two distinct requirements on the received power $P_S$. It follows from (4.101) that the phase locking branch demands an optical power (at the receiver input) of

$$P_S = \frac{81\pi q\Delta\nu}{8R\sigma_E^2} \quad (4.105)$$

for specified values of the laser linewidth $\Delta\nu$ and the acceptable phase error standard deviation $\sigma_E$. The power requirement of the data detection branch can be calculated through

$$\text{BER} = \frac{1}{2} \text{erf}\left(\sqrt{\rho_0^2\epsilon}\right) = \frac{1}{2} \text{erf}\left(\sqrt{\frac{2RP_S\epsilon}{qR_b}}\right) \quad (4.106)$$

$$\rightarrow P_S = \frac{qR_b}{2R\epsilon} \left(\text{erfc}^{-1}(2 \cdot \text{BER})\right)^2 \quad (4.107)$$
for given values of \( R_b \), BER and the acceptable power penalty \( \epsilon \). The receiver reaches its optimum operating condition when both power requirements (4.105) and (4.107) are equal. This leads to the computation of the optimum phase error standard deviation \( \sigma_E \) for given values of \( R_b \), \( \Delta \nu \) and BER:

\[
\sigma_{E,\text{opt}} = \left( \frac{81 \pi \epsilon_{\text{opt}}}{4 \text{BRLR}} \right)^{1/6} \left( \text{erfc}^{-1}(2 \cdot \text{BER}) \right)^{-1/3},
\]

where

\[
\text{BRLR} = \frac{R_b}{\Delta \nu}
\]

(4.109) is the bit-rate-to-linewidth ratio (BRLR). It is important to note that (4.108) cannot be solved explicitly, because the computation of \( \epsilon_{\text{opt}} \) through (4.103), (4.91), (4.79), (4.80) and (4.104) does not have an inverse function. An iterative algorithm to solve (4.108) is presented in appendix B. An example of (4.108) is shown in fig. 4.12.

The theory derived until here describes a complete design rule for a dither loop, based on specified values of \( R_b \), \( R \), \( \Delta \nu \) and BER. The phase error standard deviation \( \sigma_E \) and the power penalty \( \epsilon \) follow from (4.108), through the algorithm described in appendix B. The required power \( P_S \) can be calculated with (4.105) or (4.107) (both equations yield the same value). Equations (4.98) and (4.99) determine the optimum loop noise bandwidth \( B_n \) and phase dither amplitude \( \phi_d \). Choosing a damping factor \( \zeta \) and solving (3.18) yields the natural frequency
of the loop \( f_n \). The dither frequency \( f_d \) has to be chosen much larger than \( f_n \), e.g. \( f_d = 10 \cdot f_n \).

In table 4.2, several dither loop designs with optimum performance measures are presented. For all six designs, the overall power penalty due to the phase error \( \phi_E(t) \) remains very low. This is because with the specified values of \( R_b \) and \( \Delta \nu \), the power demand is determined by the data detection unit, and not by the phase-locking branch. Moreover, it is interesting to note that the system with the smaller laser linewidth requires higher optical input powers, because of an inferior photodiode responsivity. The advantage of the narrow-linewidth system lies in much smaller values of the optimum noise bandwidth, which eases the design of the feedback loop.

During the analysis, various restrictions have been made to simplify the mathematics. For the linearity assumption, it has been stated that the phase error standard deviation should not exceed 10°. It can be seen in table 4.2 that this is achieved for all loops. The system with the smallest bit-rate-to-linewidth ratio is at the limit of the allowed phase error standard deviation. A further assumption has been made during the calculation of the relevant shot noise contribution. It has been found in Section 4.2.3 that the signal×noise beat note exceeds the noise×noise beat note by at least a factor of ten for most receiver setups. Reading the values of \( R_b \) and \( P_S \) from table 4.2 and comparing it with fig. 4.8 (where \( R_b \approx B_{RF} \)), reveals that the signal×noise beat note is indeed the dominant shot noise expression.
### Table 4.2: Optimum dither loop designs for two laser types and three different values of the system bit rate (damping factor $\zeta = 1/\sqrt{2}$).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit error rate</td>
<td>scalar</td>
<td>$\eta$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantum efficiency</td>
<td></td>
<td>$\lambda$</td>
<td>1064</td>
<td>nm</td>
<td></td>
</tr>
<tr>
<td>Laser type</td>
<td></td>
<td>$\Delta \nu$</td>
<td>10</td>
<td>Hz</td>
<td></td>
</tr>
<tr>
<td>Wavelength</td>
<td></td>
<td>$R$</td>
<td>0.64</td>
<td>A/W</td>
<td></td>
</tr>
<tr>
<td>Linewidth</td>
<td></td>
<td>$R_b$</td>
<td>10</td>
<td>GHz</td>
<td></td>
</tr>
<tr>
<td>Bit-rate-to-line. ratio</td>
<td></td>
<td>$B_{RLR}$</td>
<td>5</td>
<td>10</td>
<td>10^6</td>
</tr>
<tr>
<td>Phase error std. dev.</td>
<td></td>
<td>$\sigma_E$</td>
<td>3.3</td>
<td>deg</td>
<td>2.2</td>
</tr>
<tr>
<td>Power penalty</td>
<td></td>
<td>$\epsilon$</td>
<td>-0.02</td>
<td>dB</td>
<td>-0.009</td>
</tr>
<tr>
<td>Required power</td>
<td></td>
<td>$P_S$</td>
<td>-56.5</td>
<td>dBm</td>
<td>-0.34</td>
</tr>
<tr>
<td>Loop noise bandwidth</td>
<td></td>
<td>$B_n$</td>
<td>21.6</td>
<td>kHz</td>
<td>-57.8</td>
</tr>
<tr>
<td>Phase dither amplitude</td>
<td></td>
<td>$\phi_d$</td>
<td>2.7</td>
<td>deg</td>
<td>-51.0</td>
</tr>
<tr>
<td>Natural frequency</td>
<td></td>
<td>$f_n$</td>
<td>6.5</td>
<td>kHz</td>
<td>-48.0</td>
</tr>
<tr>
<td>Dither frequency</td>
<td></td>
<td>$f_d$</td>
<td>65</td>
<td>kHz</td>
<td>-46.7</td>
</tr>
</tbody>
</table>

Note: The system bit rates for the two laser types are 150 MHz and 1.1 GHz, respectively.
Chapter 4. Phase locking by LO phase dithering
Chapter 5

Receiver measurement

In this chapter, measurement results are presented for a real world coherent transmission system employing a dither OPLL. The system has been developed and built by Contraves Space AG, Zurich. Special care has been taken to guarantee a space qualifiable system design. The receiver is dominated by losses from optical components (mainly the photodiodes) and connectors, so that the power penalty due to phase dithering and incomplete carrier phase recovery, as calculated in Section 4.3.3 of the previous chapter, is only of minor importance.

5.1 System design

The measurement system setup employs only fiber-coupled optics, and all fibers and components (i.e. modulator, coupler) maintain the polarization of the light. No free-space transmission takes place from transmitter to receiver. Instead, a fiber-coupled optical attenuator simulates the beam spreading losses.\(^1\) The electrical circuits mostly consist of high reliability (HiRel) components, which is an essential requirement for space applications.

The transmitter has been built according to fig. 2.2, but without optical amplification. The light source is a diode-pumped Nd:YAG

\(^1\)Although a successful TV signal transmission has been demonstrated with a fully operational free-space telescope [91].
laser\textsuperscript{2}, emitting at 1.06 µm [92]. The laser frequency can be tuned with a heater (slow, but wide tuning range) and with a piezoelectric transducer (fast, but small tuning range) bonded to the laser cavity. Modulation occurs in an electro-optical amplitude modulator, as described in Section 2.2.2. The modulator bias control circuit is based on a similar concept as the dither loop. A small bias point variation is applied and synchronously demodulated in the detected optical power. In case of minimum modulator output power, the optical carrier is suppressed.

For the front end of the receiver, the photodiodes and the transimpedance amplifier have been integrated on a custom made indium phosphide (InP) chip. The photodiodes achieve quantum efficiencies of around 55\%, and the transimpedance amounts to approximately 900 Ohm. In a future design, attempts should be made to increase the photodiode quantum efficiency. The front end is one of the few non-HiRel grade electrical components. It has a flat frequency response from zero to around 4 GHz.

Besides the main dither loop circuitry as depicted in fig. 4.1, the receiver employs additional logics to control the initial frequency acquisition. The loop can achieve phase-lock for laser intermediate frequencies far beyond the loop natural frequency $f_n$. This is done by thermally scanning the LO frequency on a random basis, and detecting periodic fluctuations in the power of the receiver output signal.

The dither amplitude and frequency of the OPLL amount to 10° and 110 kHz, respectively. The loop natural frequency is 10 kHz. The loop filter has a second pole to suppress $1/f$ frequency noise components in the laser phases. This results in a third order closed-loop transfer function. The loop filter acts on the piezo tuning input of the laser. A secondary feedback is applied to the heater input, so that slow frequency variations are thermally tracked.

A single phase-locking event is shown in fig. 5.1. In open loop condition, the receiver output power oscillates according to the intermediate frequency of the TX and LO laser. As soon as the acquisition logic detects a beat frequency of 10 kHz or less, the feedback loop is closed. Phase locking occurs immediately and without cycle slip. The decreasing power in closed-loop operation is due to an AC-coupling of the power detector signal.

\textsuperscript{2}Neodymium-doped yttrium aluminum garnet laser.
5.2 Communication performance

The communication performance of the previously described transmission system has been measured in terms of the bit error rate and the SNR, which is described in the next two sections. In advance, it had been verified that the receiver operates in shot noise limited conditions. With an optical power of 5 dBm launched into the coupler, the LO adds 11 dB to the noise power density at the receiver output. This is sufficiently large for the shot noise limit.

5.2.1 Digital transmission

The RF section of the receiver has a lower cut-off frequency of around 200 MHz. Because NRZ pseudo-random bit sequences (PRBS) have major components at low frequencies, the bit error rate cannot be measured directly. Instead of a baseband transmission, an electrical subcarrier needs to be used. In appendix C, the test setup for the digital transmission test is explained in more detail. The transmitted signal is a 400 Mbit/s PRB sequence of length $2^{31} - 1$ on a 600 MHz subcarrier. The relevant content of this signal spans over a frequency range of 200 to 1000 MHz.

Figure 5.1: Single phase-locking event. Channel 1: AC-coupled power detector signal. Channel 4: Loop control signal. Image: Contraves Space AG.
Figure 5.2: Receiver eye diagram at different optical input power levels ($R_b = 400$ Mbit/s).
Figure 5.3: Measured bit error rates in dependency of the received optical power. The power penalty, referred to the shot noise limit, amounts to 6 dB (for a PRBS $2^{31} - 1$ signal at 400 Mbit/s).

In fig. 5.2, the eye diagram of the receiver output is depicted for several levels of the optical input power. The signal amplitude remains on a constant level due to the AGC, but the SNR deteriorates when the received power decreases. The eye diagrams show that the system is free of nonlinearities or saturation effects. The result of the BER measurement is presented in fig. 5.3. The receiver requires an average of 36 photons/bit for a bit error rate of $10^{-9}$. This is 6 dB more than predicted by the shot noise limit.

5.2.2 Analog transmission

For the digital transmission test, a relatively complex test setup was employed. The results of the previous section can be verified with an analog transmission test, which requires far less test equipment. This is done by feeding a single 800 MHz carrier directly into the transmission system. At the system output, the SNR can easily be measured with a spectrum analyzer. The bit error rate follows from the SNR by equation (4.106). The resulting BER levels have been included in fig. 5.3. They
match well with the results from the digital transmission test. From the total loss of 6 dB, 2.5 dB are due to the photodiodes ($\eta = 56\%$). Coupler loss and asymmetry is accounted for 1.5 dB, and 1 dB is lost in two optical connectors. On the electrical side, 0.5 dB of loss is generated by thermal noise, and 0.5 dB by the residual phase error.
Chapter 6

Conclusions

The highest possible sensitivity in an optical free-space transmission system can be obtained with BPSK modulation and homodyne reception. A homodyne receiver requires an OPLL, for which the dither loop is believed to be superior to other designs. In the present thesis, the dither loop is mathematically described for the first time. The first step of the analysis was to calculate the phase detector gain, which is disguised in a series of signal processing operations. Furthermore, a ‘false-lock’ behavior of the dither loop was reported, where the two laser were locked to a non-zero intermediate frequency. Precautions, based on simulation results, were proposed to circumvent this condition.

The phase error variance due to LO phase dithering, white frequency noise induced phase noise and shot noise was evaluated, and power penalties, generated by the aforementioned noise sources, were calculated. This knowledge was used to synthesize design rules for an optimum dither loop with respect to the system specifications. The optimization variables are the dither amplitude $\phi_d$ and the loop noise bandwidth $B_n$. All other system parameters (especially $f_n$, $f_d$ and $K_{PD}$) follow from $B_n$, $\phi_d$ and $\zeta$. Simulations, performed on a Simulink model, support the achieved results.

A real-world dither loop receiver, using diode-pumped Nd:YAG lasers at a wavelength of 1.06 µm, was characterized by its BER and SNR performance. The achieved sensitivity amounts to 36 photons/bit for a PRBS signal of length $2^{31} - 1$ at a data rate of 400 Mbit/s and a
bit error rate of $10^{-9}$. This is a 6 dB power penalty compared to the shot noise limit, of which the photodiodes account for by 2.5 dB. The SNR results, determined with a much simpler setup and recalculated to an equivalent bit error rate, validated the BER measurement.

### 6.1 Outlook

The dither loop analysis could be enhanced with a $1/f$ frequency noise induced phase noise expression, and a low frequency pole in the loop filter to suppress it. This is due to the fact that most lasers exhibit a thermally induced (and therefore slow) frequency drift of large magnitude. An alternative dither loop design rule could be developed which improves the phase-locking reliability and/or reduces the loop noise bandwidth, at the cost of moderately higher power penalties.

It has been shown in Section 4.3.1 that the power penalty due to phase locking depends merely on the dither amplitude $\phi_d$. A highly optimized receiver design could adaptively control $\phi_d$ — and therefore, the amount of power used for phase locking — according to the receiving conditions. This is an advantage of the dither loop, since, in most other OPLL designs, the phase-locking power is hard-wired into the implementation, e.g. by the power splitting ratio of the $90^\circ$ coupler in a Costas loop.

An attempt should be made to shift the wavelength of the real-world transmission system to 1550 nm. This promises the following advantages:

- **Responsivity:** According to equation (2.15), the photodiode responsivity is proportional to the wavelength of the light. An optical source with a certain output power in a specified time interval generates more photons, each carrying less energy, at longer wavelengths. A higher photodiode responsivity reduces the power demand of the receiver to achieve a certain bit error rate.

- **Components:** Terrestrial fiber-optic communications makes wide use of 1550 nm wavelengths. Hence, components at 1550 nm exhibit better availability and performance due to a higher degree of technological evolution. This eases and accelerates the system design, and increases the overall performance.
6.1. Outlook

It must be stated, however, that no appropriate narrow-linewidth optical source ($\Delta \nu < 100$ kHz) exists in the 1550 nm regime at present. The use of fiber lasers [93], which seem to meet the wavelength and linewidth requirements, should be evaluated.

Whether coherent optical space communications will become a reality, remains an open question. Today's systems, with bulky Nd:YAG lasers at exotic wavelengths, might well be too complex and too expensive for commercial success. If a compact and efficient narrow-linewidth semiconductor source were available instead, the feasibility and cost effectiveness of optical communications would probably outperform microwave systems. The OPLL of a coherent receiver should be a dither loop, designed according to the theory derived in the present thesis.
Chapter 6. Conclusions
Appendix A

Phase noise evaluation

When calculating the phase error variance due to phase noise, the following integral has to be solved:

$$\sigma_{E,pm}^2 = \int_0^\infty S_{pn}(f)|1 - H(f)|^2 df. \quad (A.1)$$

Plugging (3.11) and (4.41) into (A.1), and substituting $s = j2\pi f$, yields

$$\sigma_{E,pm}^2 = \frac{2\Delta \nu}{\pi} \cdot \int_0^\infty \left| \frac{1}{f^2} \right|^2 \frac{1}{1 + (j2\pi f)^{-1}G_{loop}F(f)} \cdot \left| \frac{1}{1 + (j2\pi f)^{-1}G_{loop}F(f)} \right|^2 df. \quad (A.2)$$

Transformations on the integrand of (A.2), using the filter transfer function (3.12), result in

$$\frac{1}{f^2} \left| \frac{1}{1 + (j2\pi f)^{-1}G_{loop}F(f)} \right|^2 = \frac{1}{f^2} \left| \frac{1}{1 + (j2\pi f)^{-1}G_{loop}j2\pi f \tau_2 + 1} \right|^2$$

$$= \left| -4\pi^2 f \tau_1/G_{loop} \right|^2. \quad (A.3)$$

With the following substitutions from (3.16) and (3.17):

$$2\pi f_n = \sqrt{G_{loop}/\tau_1} \rightarrow \frac{1}{f^2} = \frac{4\pi \tau_1}{G_{loop}} \quad (A.4)$$

$$\zeta = \frac{\tau_2}{2} \sqrt{G_{loop}/\tau_1} \rightarrow \pi \tau_2 = \frac{\zeta}{f_n} \quad (A.5)$$
equation (A.3) becomes

\[
\left| \frac{-4\pi^2 f\tau_1/G_{\text{loop}}}{-4\pi^2 f^2\tau_1/G_{\text{loop}} + j2\pi f\tau_2 + 1} \right|^2 = \left| \frac{-f/f_n^2}{-f^2/f_n^2 + j2\zeta f/f_n + 1} \right|^2 = \frac{f^2/f_n^4}{f^4/f_n^4 + 2(2\zeta^2 - 1)f^2/f_n^2 + 1}.
\]  

(A.6)

With (A.6), and assuming a critically damped (i.e. \( \zeta = 1/\sqrt{2} \)) loop, (A.2) amounts to

\[
\sigma_{E,\text{pn}}^2 = \frac{2\Delta\nu}{\pi} \int_0^\infty \frac{f^2/f_n^4}{f^4/f_n^4 + 1} df = \frac{2\Delta\nu}{\pi f_n} \int_0^\infty \frac{z^2}{z^4 + 1} dz
\]  

(A.7)

where \( z \) substitutes \( f/f_n \). For the above integral, the identity

\[
\int_0^\infty \frac{z^{p-1}}{z^q + 1} dz = \frac{\pi}{q} \cdot \frac{1}{\sin(\pi p/q)}
\]  

(A.8)

can be used (from [94], Section 3.241). With (A.8), equation (A.7) results in

\[
\sigma_{E,\text{pn}}^2 = \frac{2\Delta\nu}{\pi f_n} \cdot \frac{\pi}{4\sin(3\pi/4)} = \frac{\Delta\nu}{\sqrt{2}f_n}
\]  

(A.9)

or, in terms of the loop noise bandwidth:

\[
\sigma_{E,\text{pn}}^2 = \frac{3\pi\Delta\nu}{4B_n}
\]  

(A.10)

where \( f_n = 4B_n/(3\sqrt{2}\pi) \) has been calculated from (3.18). Equation (A.10) denotes the phase error variance due to phase noise, depending only on the laser linewidth and the loop noise bandwidth.
Appendix B

Iterative computation of $\sigma_E$ and $\epsilon$

Equation (4.108) for $\sigma_E$ cannot be solved analytically, because the power penalty $\epsilon$ depends on $\sigma_E$ through equations (4.103), (4.91), (4.79), (4.80) and (4.104). This dependency can not be inverted. The following iterative algorithm is proposed to compute $\sigma_E$ for given values of $R_b$, $\Delta\nu$ and BER:

1. Choose an $\epsilon_1$ which is expected to be smaller than the true power penalty $\epsilon$, i.e.:

$$\epsilon_1 < \epsilon.$$  

As it can be seen from table 4.2, the power penalties are usually very small. A good choice for $\epsilon_1$ could therefore be a power penalty of 0.89 (-0.5 dB).

2. Solve (4.108):

$$\sigma_{E,n} = \left( \frac{81\pi \epsilon_{1,n}}{4BR\Lambda} \right)^{1/6} \left( \text{erf}^{-1}(2\text{BER}) \right)^{-1/3}$$

Here, $\epsilon_{1,n}$ and $\sigma_{E,n}$ denote $\epsilon_1$ and $\sigma_E$ of the n-th iteration, respectively.
3. Calculate the SNR $\rho_{0,n}^2$ of the ideal receiver (i.e. from (4.106)):

$$
\rho_{0,n}^2 = \frac{(\text{erfc}^{-1}(2\text{BER}))^2}{\epsilon_{1,n}}.
$$

4. Compute the power penalty $\epsilon_2$ through (4.103), (4.91), (4.79), (4.80) and (4.104) for the given combination of $\sigma_{E,n}$ and $\rho_{0,n}$:

$$
\epsilon_{2,n} = \epsilon(\rho_{0,n}, \sigma_{E,n}).
$$

With a good choice in step 1), $\epsilon_{2,n}$ is larger than $\epsilon_{1,n}$.

5. Calculate a new $\epsilon_1$ between $\epsilon_{1,n}$ and $\epsilon_{2,n}$:

$$
\epsilon_{1,n+1} = \frac{\epsilon_{1,n} + \epsilon_{2,n}}{2}.
$$

6. Repeat with step 2). An exit criterion can be defined as

$$
\frac{|\epsilon_{2,n} - \epsilon_{1,n}|}{\epsilon_{1,n}} < 10^{-6}
$$

or any other level of the desired accuracy.

When the algorithm terminates, $\epsilon_{1,n}$ and $\sigma_{E,n}$ contain values of $\epsilon$ and $\sigma_E$ which solve (4.108) with a sufficiently good accuracy, under consideration of the dependency of $\epsilon$ from $\sigma_E$ through (4.103), (4.91), (4.79), (4.80) and (4.104).
Appendix C

Bit error rate test setup

The test setup for measuring the bit error rate is depicted in fig. C.1. A pattern generator produces the PRBS $2^{31} - 1$ test signal. This signal is modulated onto a 600 MHz subcarrier and, after level shifting, fed into the coherent optical transmission system. The transmitted power is adjusted with an optical attenuator. An optical power meter, attached to a 3 dB coupler in front of the receiver, observes the transmitted power. The receiver output signal is amplified and demodulated with the original subcarrier. The phase of the subcarrier is adjusted with a variable delay line. A low-pass filter removes the spurious carrier signal and high frequency noise. With a limiting amplifier, the received signal is boosted to a power level suitable for the error detection unit. The bit error rate is measured during a 60 s interval, or 300 s for small error probabilities.
Figure C.1: Bit error rate measurement setup.
Bibliography


Acknowledgments

I would like to thank my PhD supervisor Werner Bächtold for the opportunity to conduct a dissertation at the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH). I am grateful to Helmut Bölskei and Klaus Kudielka for joining the thesis committee, and to Daniel Erni for the support during the whole period of the thesis. I greatly appreciate the work of the IFH staff members, who provided and maintained the necessary infrastructure for successful research.

Very special thanks go to my colleagues at IFH, IFE and Contraves Space, who created a cheerful and inspiring working atmosphere. I would like to acknowledge the help of Matthias Frey during the Conference in Japan, and the assistance of Christine Fleischmann, who revised the manuscript of the thesis.

Sometimes, if stuck in an unsolvable calculation, or while reading an inscrutable paper, it is best to take a break. In this sense, I am deeply grateful to all my friends outside of work for the time we have spent together. Last but not least, I would like to thank my parents for their continuous support during my studies.

Zurich, November 2005

Frank Herzog
Curriculum Vitae

Frank Thomas Herzog

born 11 January 1975
in Münsterlingen, Switzerland.

Doctor of Sciences:

2005    Dissertation at the Swiss Federal Institute of Technology Zurich (ETH Zurich) on the subject of coherent optical space communications.

2001–2005 Research and teaching assistant at the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH) at ETH Zurich.

Diploma in electronics engineering:


1998    Internship at NF Engenharia de Automação Ltda., Criciúma, Brazil.

Research interests: microwave electronics, optoelectronics, optics, communication theory and acoustics.
Publications:


