Title	The role of synchronization in digital communications using chaos - part I: fundamentals of digital communications.
Author(s)	Kolumbán, Géza; Kennedy, Michael Peter; Chua, Leon O.
Publication date	1997-10
Original citation	Kolumban, G.; Kennedy, M.P.; Chua, L.O.; , 1997. The role of synchronization in digital communications using chaos. I . Fundamentals of digital communications. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 44(10), pp.927-936. doi:10.1109/81.633882
Type of publication	Article (peer-reviewed)
Link to publisher's version	http://dx.doi.org/10.1109/81.633882 Access to the full text of the published version may require a subscription.
Item downloaded from	http://hdl.handle.net/10468/185

The Role of Synchronization in Digital Communications Using Chaos—Part I: Fundamentals of Digital Communications

Géza Kolumbán, *Member, IEEE,* Michael Peter Kennedy, *Senior Member, IEEE,*
and Leon O. Chua, *Fellow, IEEE*

***Abstract*— In a digital communications system, data is transmitted from one location to another by mapping bit sequences to symbols, and symbols to sample functions of analog waveforms. The analog waveform passes through a bandlimited (possibly time-varying) analog channel, where the signal is distorted and noise is added. In a conventional system the analog sample functions sent through the channel are weighted sums of one or more sinusoids; in a chaotic communications system, the sample functions are segments of chaotic waveforms. At the receiver, the symbol may be recovered by means of coherent detection, where all possible sample functions are known, or by noncoherent detection, where one or more characteristics of the sample functions are estimated. In a coherent receiver, synchronization is the most commonly used technique for recovering the sample functions from the received waveform. These sample functions are then used as reference signals for a correlator. Synchronization-based receivers have advantages over noncoherent ones in terms of noise performance and bandwidth efficiency. These advantages are lost if synchronization cannot be maintained, for example, under poor propagation conditions. In these circumstances, communication without synchronization may be preferable. The main aim of this paper is to provide a unified approach for the analysis and comparison of conventional and chaotic communications systems. In Part I, the operation of sinusoidal communications techniques is surveyed in order to clarify the role of synchronization and to classify possible demodulation methods for chaotic communications. In Part II, chaotic synchronization schemes are described in detail and proposed chaotic communications techniques are summarized. In Part III, examples of chaotic communications schemes with and without synchronization are given, and the performance of these schemes is evaluated in the context of noisy, bandlimited channels.**

***Index Terms*— Chaos, chaotic communications, chaotic synchronization.**

I. INTRODUCTION

THE observation by Pecora and Carroll [1] that two chaotic systems can be synchronized has generated tremendous interest in transmitting information from one location to another using a wideband chaotic signal.

Many modern communications applications, such as mobile or indoor radio, are susceptible to multipath propagation effects arising from interaction between signals at the receiver which travel along different propagation paths. By contrast with a conventional digital modulation scheme, where the transmitted symbols are mapped to a finite set of periodic waveform segments for transmission, every transmitted symbol in a chaotic modulation scheme produces a different nonperiodic waveform segment. Because the cross correlations between segments of a chaotic waveform are lower than between pieces of periodic waveforms, chaotic modulation ought to offer better performance under multipath propagation conditions. Thus, chaotic modulation offers a potentially simple solution for robust wideband communications.

The chaotic communication schemes which have been proposed to date have been developed using heuristic arguments that make it impossible to compare them with conventional communications systems. Using the language of communications theory, this paper extends the basis function approach to proposed chaotic communications systems in order to provide a unified framework in which to compare and contrast conventional and chaotic communications techniques.

The principal difference between conventional and chaotic systems is that segments of chaotic waveforms, rather than sinusoids, are used as basis functions in chaotic communications. Because of the nonperiodic property of chaotic signals, there is a fundamental difference between conventional and chaotic systems. The parameters of the received chaotic waveform that are required in order to recover the transmitted information must be estimated from sample functions of finite length. Even in the noise-free case, this estimation has a nonzero variance; therefore, the symbol duration cannot be reduced below a certain lower bound.

Most of the research in the field of chaotic communications to date has assumed that the transmitter is connected to the receiver by an ideal channel [2]. Since the principal source

Manuscript received February 10, 1997; revised June 18, 1997. This work was supported in part by the National Scientific Research Foundation of Hungary (OTKA) under Grant T-020522, by the Information Technologies RTD Programme of the European Commission (Esprit Project 21103 INSPECT), and by the U.S. Office of Naval Research under Grant N00014-96-1-0753. This paper was recommended by Guest Editor M. J. Ogorzałek.

G. Kolumbán is with the Department of Measurement and Instrument Engineering, Technical University of Budapest, H-1521 Budapest, Hungary (e-mail: kolumban@mmt.bme.hu).

M. P. Kennedy is with the Department of Electronic and Electrical Engineering, University College Dublin, Dublin 4, Ireland (e-mail: Peter.Kennedy@ucd.ie).

L. O. Chua is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA (e-mail: chua@diva.eecs.berkeley.edu).

Publisher Item Identifier S 1057-7122(97)07316-9.

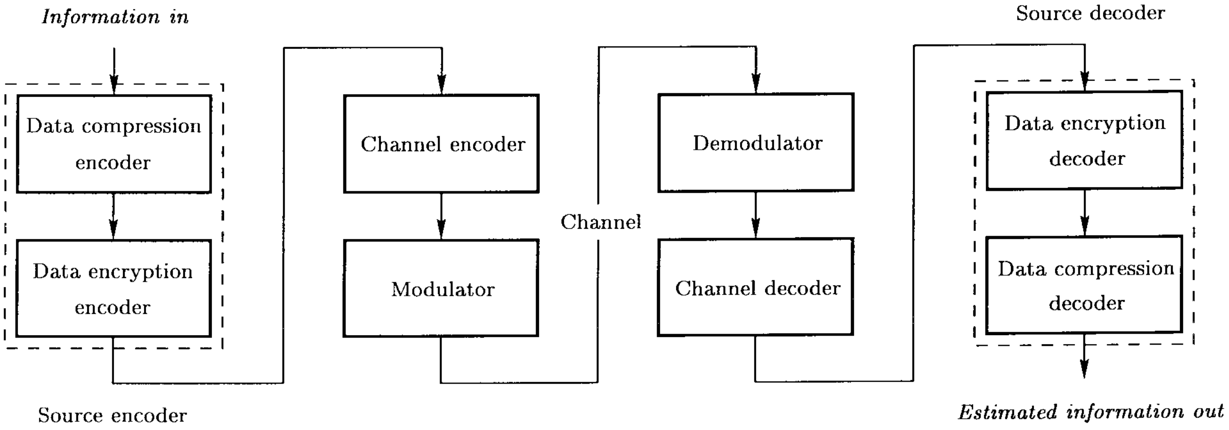

Fig. 1. Digital communications system showing source and channel coding, modulation, and channel.

of errors in a digital communications system is the channel, it is impossible to quantify the performance of a chaotic modulation technique by assuming an ideal channel. A realistic channel model must at least include additive noise and linear filtering.

Recent studies of chaotic synchronization, where significant noise and filtering are introduced in the channel, suggest that synchronization of chaos is not yet sufficiently robust for practical applications in communications [3].

In this three-part tutorial paper, we explain the role of synchronization in a digital communications system and evaluate the performance of chaotic modulation schemes.

In Section II of this part, we describe the major components of a digital communications system and show that the primary source of errors is the analog channel. We explain why a realistic channel model must include at least additive white Gaussian noise and band-limiting. We review the notion of bit error rate as a way of comparing digital modulation schemes.

In Section III, we show how a signal set may be constructed from a limited set of orthonormal basis functions and explain the advantages of this choice.

In Section IV, we show that the primary motivation for carrier synchronization is to permit coherent detection, the benefits of which are improved noise performance and bandwidth efficiency.

Under poor propagation conditions, where synchronization cannot be maintained, the advantages of coherent detection are lost. In such circumstances, a noncoherent receiver offers a more robust and less complex solution, as shown in Section V.

The potential advantages of using a chaotic carrier signal are explained in Section VI by highlighting the disadvantages of narrowband (sinusoidal) communications when propagation conditions are poor.

In Part II of the paper, we consider the state-of-the-art in synchronization of chaotic systems in the context of digital communications; in addition, proposed chaotic communications techniques are surveyed.

In Part III, performance targets for chaotic communications techniques are summarized and examples (CSK with synchronization, noncoherent CSK and DCSK correlation receiver) are given. Finally, the performance of these systems is evaluated in the context of a noisy and bandlimited channel.

II. OVERVIEW OF DIGITAL COMMUNICATIONS

A. Basic Structure of a Digital Communications System

Communication system theory is concerned with the transmission of information from a source to a receiver through a channel [4], [5].

The goal of a digital communications system, shown schematically in Fig. 1, is to convey information from a digital information source (such as a computer, digitized speech or video, etc.) to a receiver as effectively as possible. This is accomplished by mapping the digital information to a sequence of symbols which vary some properties of an analog electromagnetic wave called the carrier. This process is called *modulation*. Modulation is always necessary because all practical telecommunications channels are bandlimited analog systems which cannot transmit digital signals directly.

At the receiver, the signal to be received is selected by a channel filter, demodulated, interpreted, and the information is recovered.

Conversion of the digital information stream to an analog signal for transmission may be accompanied by encryption and coding to add end-to-end security, data compression, and error-correction capability.

Built-in error-correction capability is required because real channels distort analog signals by a variety of linear and nonlinear mechanisms: attenuation, dispersion, intersymbol interference, intermodulation, PM/AM and AM/PM conversions, noise, interference, multipath effects, etc.

A *channel encoder* introduces algorithmic redundancy into the transmitted symbol sequence that can be used to reduce the probability of incorrect decisions at the receiver.

Modulation is the process by which a symbol is transformed into an analog waveform that is suitable for transmission. Common digital modulation schemes include *amplitude shift keying* (ASK), *phase shift keying* (PSK), *frequency shift keying* (FSK), *continuous phase modulation* (CPM), and *amplitude phase keying* (APK), where a one-to-one correspondence is established between amplitudes, phases, frequencies, phase and phase transitions, and amplitudes and phases, respectively, of a sinusoidal carrier and the symbols.

The *channel* is the physical medium through which the information-carrying analog waveform passes as it travels between the transmitter and receiver.

The transmitted signal is invariably corrupted in the channel. Hence, the receiver never receives exactly what was transmitted. The role of the *demodulator* in the receiver is to produce from the received corrupted analog signal an estimate of the transmitted symbol sequence. The role of the *channel decoder* is to reconstruct the original bit stream, i.e., the information, from the estimated symbol sequence. Because of disturbances in real communications channels, error-free transmission is never possible.

Nonlinear dynamics has potential applications in several of the building blocks of a digital communications system: data compression, encryption, and modulation [6]. Data compression and encryption are potentially reversible, error-free digital processes. By contrast, the transmission of an analog signal through a channel and its subsequent interpretation as a stream of digital data are inherently error-prone.

In this paper, we focus on the application of chaos as a modulation scheme. In order to compare the use of a chaotic carrier signal with that of a conventional sinusoidal carrier, we must consider a realistic channel model and quantify the performance of each chaotic modulation scheme using this channel.

In this section, we introduce the minimum requirements for a realistic channel model and the performance measures by which we will compare conventional and chaotic modulation schemes.

B. Minimum Requirements for a Channel Model

The definition of the telecommunications channel depends on the goal of the analysis performed. In the strict sense, the channel is the physical medium that carries the signal from the transmitter to the receiver. If the performance of a modulation scheme has to be evaluated, then the channel model should contain everything from the modulator output to the demodulator input. Even if the physical medium can be modeled by a constant attenuation, the following effects have to be taken into account:

1) In order to get maximum power transfer, the input and output impedances of the circuits of a telecommunications system are matched. This is why thermal noise modeled as *additive white Gaussian noise* (AWGN)[1] is *always present* at the input to an RF receiver.
2) The bandwidth of the channel has to be limited by a so-called channel (selection) filter in order to suppress the unwanted input signals that are always present at the input of a radio receiver and that cause interference due to the nonlinearities of the receiver.

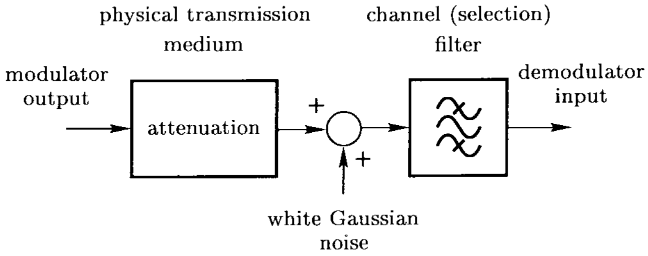

Fig. 2. Model of an additive-white-Gaussian-noise channel including the frequency selectivity of the receiver.

The simplest channel model that can be justified when evaluating the performance of a modulation scheme is shown in Fig. 2. Note that the channel filter is used only to select the desired transmission frequency band at the receiver and not to model any frequency dependence of the physical transmission medium. If, in addition to noise and attenuation, other nonidealities of the physical transmission medium (such as frequency dependence, selective fading, interferences, etc.) are to be taken into account, then these should be included in the first block in Fig. 2.

In the model shown in Fig. 2, we have assumed that the received signal is corrupted by AWGN. In a real telecommunications system, the noise may not be exactly white or Gaussian. The reasons for assuming AWGN are that

1) it makes calculations tractable;
2) thermal noise, which is of this form, is dominant in many practical communications systems; and
3) experience has shown that the relative performance of different modulation schemes determined using the AWGN channel model remains valid under real channel conditions, i.e., a scheme showing better results than another for the AWGN model also performs better under real conditions [4], [5].

C. Performance Measures

The primary source of errors in a digital communications system is the analog channel. The fundamental problem of digital communications is to maximize the effectiveness of transmission through this channel.

The performance of a digital communications system is measured in terms of the *bit error rate* (BER) at the receiver. In general, this depends on the coding scheme, the type of waveform used, transmitter power, channel characteristics, and demodulation scheme. The conventional graphical representation of performance in a linear channel with AWGN, depicted in Fig. 3, shows BER versus E_b/N_0, where E_b is the energy per bit and N_0 is the power spectral density of the noise introduced in the channel.

For a given background noise level, the BER may be reduced by increasing the energy associated with each bit, either by transmitting with higher power or for a longer period per bit. The challenge in digital communications is to achieve a specified BER with *minimum energy per bit*. A further consideration is *bandwidth efficiency*, defined as the ratio of data rate to channel bandwidth [4].

---

[1] The definition for the Gaussian process is given in [4] and [7]. The autocorrelation of white noise is a Dirac delta function multiplied by $N_0/2$ and located at $\tau = 0$, where N_0 is the power spectral density of the noise.

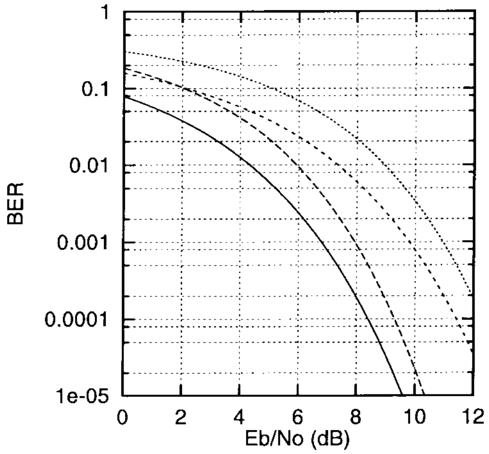

Fig. 3. Comparison of the noise performances of digital modulation schemes. From left to right: coherent binary phase shift keying (solid), differential phase shift keying (long dash), coherent (short dash) and noncoherent (dot) binary orthogonal frequency shift keying.

III. FACTORS AFFECTING THE CHOICE OF MODULATION SCHEME

For a given BER and background noise level, the main goal in the design of a digital communications system is to minimize the energy required for the transmission of each bit. The second goal is the efficient utilization of channel bandwidth. These design requirements affect the choice of the modulation scheme to be used.

While modulation is a relatively straightforward process of mapping symbols to analog waveforms (elements of the so-called "signal set") in a deterministic manner, demodulation, which is concerned with mapping samples of a corrupted stochastic analog signal back to symbols, is a more difficult and error-prone task.

In this section, we consider ways of designing the signal set to maximize the bandwidth efficiency and minimize the probability of making incorrect decisions at the receiver.

A. M-ary Modulation Schemes for Bandwidth Efficiency

In binary modulation schemes, where the bit stream is mapped to two possible signals, bandwidth efficiency is poor since the required channel bandwidth is proportional to the bit rate.

The bandwidth efficiency can be improved by using M-ary modulation schemes, where the signal set contains M possible signals. In almost all applications, the number of possible signals M is 2^n, where n is an integer. The symbol duration is given by $T = nT_b$, where T_b is the bit duration.

In conventional digital communications systems, the elements of the signal set are sinusoids, where, for example, the amplitude, phase or frequency of the transmitted signal is varied among M discrete values in the case of M-ary ASK (MASK), M-ary PSK (MPSK) and M-ary FSK (MFSK), respectively. In M-ary APK (MAPK), of which M-ary *quadrature amplitude modulation* (QAM) [8] is an example, both the amplitude and the phase of the reference sinusoid are varied.

In M-ary modulation schemes, one *symbol* is transmitted for every n bits in the data stream. In particular, the incoming bit stream is transformed into a sequence of symbols and every symbol is mapped to an element of the signal set. Since symbols are transmitted once per n bits, instead of once per bit, the required channel bandwidth is proportional to the *symbol rate* rather than the *bit rate* (except in the case of MFSK [4]), and the bandwidth efficiency is improved considerably.

B. Orthonormal Basis Functions and Correlation Receivers

A coherent receiver performs demodulation by comparing the incoming signal with *all* elements of the signal set. If a linear AWGN channel model is assumed, the most effective way to accomplish this is by correlating the received signal with every element of the signal set and selecting the one with the largest correlation.

Therefore, a coherent receiver must in principle know all M elements of the signal set. Since M can be as large as 256 in modern modulation schemes, this seems like a difficult task. However, the large number of signals which must be known at the receiver can be reduced by introducing the idea of orthonormal basis functions.

1) Orthonormal Basis Functions: Let $s_i(t), i = 1, 2, \cdots, M$ denote the elements of the signal set. Our goal is to minimize the number of special signals, called *basis functions*, that have to be known at the receiver. Let the elements of the signal set be represented as a linear combination of N real-valued orthonormal basis functions

$$g_j(t), \quad j = 1, 2, \cdots, N,$$

where

$$\int_0^T g_i(t)\, g_j(t)\, dt = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{elsewhere.} \end{cases}$$

Then the elements of the signal set can be expressed as a linear combination of basis functions

$$s_i(t) = \sum_{j=1}^{N} s_{ij} g_j(t), \quad \begin{cases} 0 \le t \le T \\ i = 1, 2, \cdots, M \end{cases} \tag{1}$$

where $N \le M$. In conventional digital telecommunications systems, sinusoidal basis functions are used; the most common situation involves a quadrature pair of sinusoids.

2) Signal Set Generation: The coefficient s_{ij} in (1) may be thought of as the jth element of an N-dimensional *signal vector* $\boldsymbol{s}_i$. The incoming bit stream is first transformed into a symbol sequence; the elements of the signal vector are then determined from the symbols. The signals $s_i(t)$ to be transmitted are generated as a weighted sum of basis functions, as given by (1).

3) Recovery of the Signal Vector by Correlation: Because the basis functions are orthonormal, the elements of the signal vector can be recovered from the elements of the signal set, i.e., from the received signal, if every basis function is known in the receiver. In particular

$$s_{ij} = \int_0^T s_i(t)\, g_j(t)\, dt, \quad \begin{cases} i = 1, 2, \cdots, M \\ j = 1, 2, \cdots, N. \end{cases} \tag{2}$$

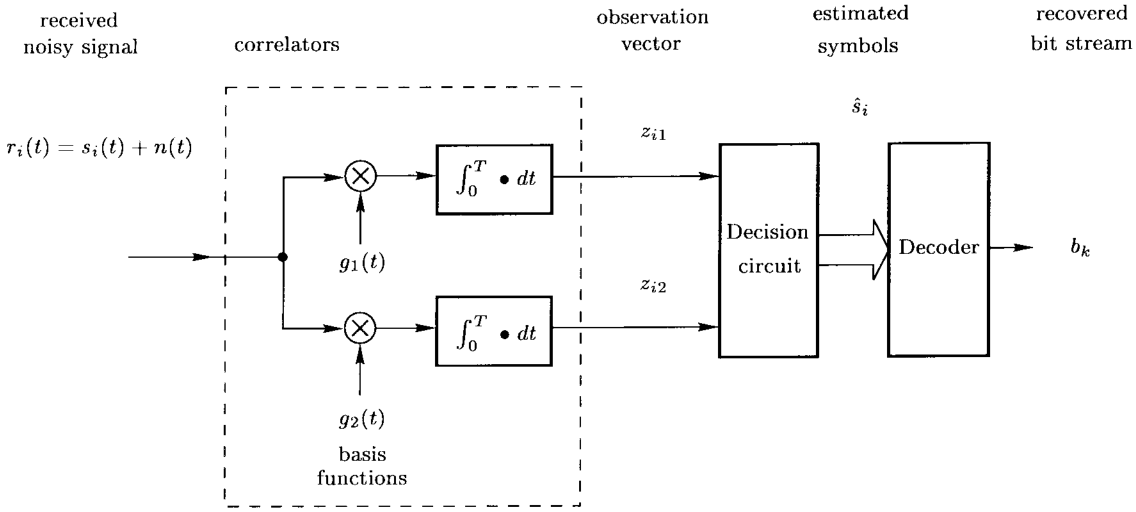

Fig. 4. Block diagram of correlation receiver for $N = 2$.

Thus, a demodulator consists of a bank of N correlators, each of which recovers the weight s_{ij} of basis function $g_j(t)$. Since there exists a one-to-one mapping between signal vectors and symbols, the transmitted symbols can be recovered by post-processing the outputs of the correlators, and the demodulated bit stream can thus be regenerated.

C. Orthonormal Basis Functions for Bandwidth Efficiency

The main advantage of using orthonormal basis functions is that a huge signal set can be generated from a small number of basis functions. Typically, a pair of quadrature sinusoidal signals (a cosine and a sine) is used as the set of basis functions. Since quadrature sinusoidal signals can be generated using a simple phase shifter, it is sufficient to know (or recover) only one sinusoidal signal in the receiver.

An example showing modulator and demodulator circuits for binary PSK (BPSK) is given in Section V-A.

IV. DETECTION OF A SINGLE SYMBOL IN NOISE: THE BASIC RECEIVER CONFIGURATIONS

The receiver must recognize the symbols sent via the channel in order to recover the information which has been transmitted. For the sake of simplicity, only the detection of a single isolated symbol is considered in this section, the effect of *intersymbol interference* (ISI), i.e., the interference between successive symbols is neglected [9].

Our goal is to minimize the average probability of symbol errors, i.e., to develop an optimum receiver configuration. For an AWGN channel and for the case when all symbols to be transmitted are equally likely, *maximum likelihood* (ML) detection has to be used in order to get an optimum receiver [8]. The ML detection method can be implemented by either correlation or matched filter receivers [4].

In this section, we demonstrate the connection between correlation and matched filter receivers, and consider the relative merits of coherent and noncoherent detection.

A. Correlation and Matched Filter Receivers

1) Correlation Receiver: Equation (2) shows how the signal vector can be recovered from a received signal by correlators if the basis functions $g_j(t)$ are orthonormal and are known at the receiver. Note that, in addition to the basis functions, both the symbol duration T and the initial time instant of symbol transmission have to be known at the receiver. The latter data are called *timing information*. The idea suggested by (2) is exploited in the correlation receiver shown in Fig. 4.

In any practical telecommunications system, the received signal is corrupted by noise, i.e., the input to each of the correlators is the sum of the transmitted signal $s_i(t)$ and a sample function $n(t)$ of a zero-mean, stationary, white, Gaussian noise process. The elements of the signal vector can still be estimated using correlators, although the estimates may differ from their nominal values, due to corruption in the channel.

The outputs of the correlators, called the *observation vector*, are the inputs of a decision circuit. The decision circuit applies the ML detection method, i.e., it chooses the signal vector from among all the possibilities that is the closest to the observation vector. Estimates of the symbols are determined from the signal vector and finally the demodulated bit stream is recovered from the estimated symbol sequence.

Note that in a correlation receiver *all the basis functions* and *the timing information* are required; these must be recovered from the (noisy) received signal.

2) Matched Filter Receiver: The observation vector can be also generated by a set of *matched filters* [4]. In a matched filter receiver a bank of matched filters is substituted for the correlators in Fig. 4. In this case the basis functions are stored locally as the impulse responses of the matched filters, i.e., only the timing information must be recovered from the received signal.

B. Coherent and Noncoherent Receivers

1) Coherent Receivers: Receivers in which exact copies of all the basis functions are known are called *coherent receivers*.

In practice, coherent correlation receivers are used almost exclusively to demodulate ASK, PSK, and its special case of *quadrature phase shift keying* (QPSK), MPSK, and M-ary QAM (MQAM) signals.

The required impulse response of a matched filter can at best be approximated by a physically realizable analog filter. Any deviation from the ideal impulse response results in a large degradation of performance. Therefore, coherent matched filter receivers are not used in radio communications.

2) Noncoherent Receivers: In applications where the propagation conditions are poor, the basis functions $g_j(t)$ cannot be recovered from the received signal. In these cases, the conventional solution is to use MFSK ($M \geq 2$) modulation and a noncoherent receiver.

The basis functions $g_j(t)$ or the elements $s_i(t)$ of the signal set are not known in a noncoherent receiver, but one or more robust characteristics of $s_i(t),\ i = 1, 2, \cdots, M$ can be determined. Demodulation is performed by evaluating one or more selected characteristics of the received signal.

For example, M different *signaling frequencies* are used in MFSK. In a noncoherent FSK receiver, a bank of bandpass filters is applied to recognize the different signaling frequencies. The observation vector is generated by envelope detectors and the decision circuit simply selects the "largest" element of the observation vector [4].

C. Relative Merits of Coherent and Noncoherent Receivers

It is often claimed that the main advantage of coherent receivers over noncoherent ones is that their performance in the presence of additive noise in the channel is better than that of their noncoherent counterparts. Let us estimate the size of this advantage for the selected application domain: digital communications.

In a practical digital communications system, communication is not possible if the BER becomes worse than 10^{-3} or 10^{-2}, so we only consider operation below this range. For example, the average value of "raw" BER for terrestrial microwave radio systems varies from 10^{-7} to 10^{-6}; with error correction, this can be reduced to below 10^{-9} [9].

The noise performance of coherent and noncoherent binary FSK receivers is shown in Fig. 3. At $E_b/N_0 = 10^{-2}$, the E_b/N_0 required by the noncoherent FSK receiver is only 1.6 dB greater than the corresponding value for the coherent one. Moreover, at high values of E_b/N_0, noncoherent FSK receivers perform almost as well as coherent ones for the same E_b/N_0.

The real advantage of the coherent technique is that huge signal sets can be generated by means of very few orthonormal basis functions. For example, in terrestrial digital microwave radio systems, 256 signals are typically generated using a pair of quadrature sinusoidal signals. This huge signal set results in excellent *bandwidth efficiency*. Moreover, the receiver must recover *just one* sinusoidal signal from the incoming signal.

For their part, noncoherent techniques offer two advantages over coherent detection:

1) When *propagation conditions are poor*, the basis functions cannot be recovered from the received signal because $r_i(t)$ differs too much from $s_i(t)$. In this case, a noncoherent receiver is the only possible solution.
2) Noncoherent receivers can, in principle, be implemented with *very simple circuitry*, because the basis functions do not need to be recovered.

V. THE ROLE OF SYNCHRONIZATION IN DIGITAL COMMUNICATIONS

In this section, we consider two fundamental synchronization problems: *timing recovery*, which is an essential part of digital communications, and *carrier recovery*, which is necessary only for coherent detection. We illustrate these issues in the context of BPSK.

A. Example: Coherent Detection of BPSK

The block diagram of a coherent BPSK transceiver can be developed from (1) and Fig. 4. As shown in Fig. 5, the most important operations performed in the receiver are

1) recovery of the basis function $\hat{g}_1(t)$ and timing information;
2) determination of the elements z_{ij} of the observation vector; and
3) decision making.

In the case of BPSK, two symbols are used to transmit the bit stream b_k. Thus, the signal set contains two sinusoidal signals $s_1(t)$ and $s_2(t)$. The binary symbols 0 and 1 are mapped to the signals

$$s_1(t) = \sqrt{\frac{2E_b}{T}}\cos(\omega_c t)$$

and

$$s_2(t) = -\sqrt{\frac{2E_b}{T}}\cos(\omega_c t)$$

respectively, where $0 \leq t \leq T$, $T = T_b$ and E_b is the transmitted energy per bit.

To simplify the recovery of the basis function, each transmitted symbol is designed to contain an integral number of cycles of the sinusoidal carrier wave. Given that there is just one basis function of unit energy

$$g_1(t) = \sqrt{\frac{2}{T_b}}\cos(\omega_c t), \quad 0 \leq t < T$$

we recover the elements of the signal vector from (2) as

$$s_{11} = +\sqrt{E_b}$$

and

$$s_{21} = -\sqrt{E_b}.$$

Let us assume equally likely symbols; then ML detection yields an optimum receiver [8]. The decision rule is simply to make the decision in favor of symbol 0 if the received signal is "closer" to $s_1(t)$, i.e., if the output of the correlator is greater

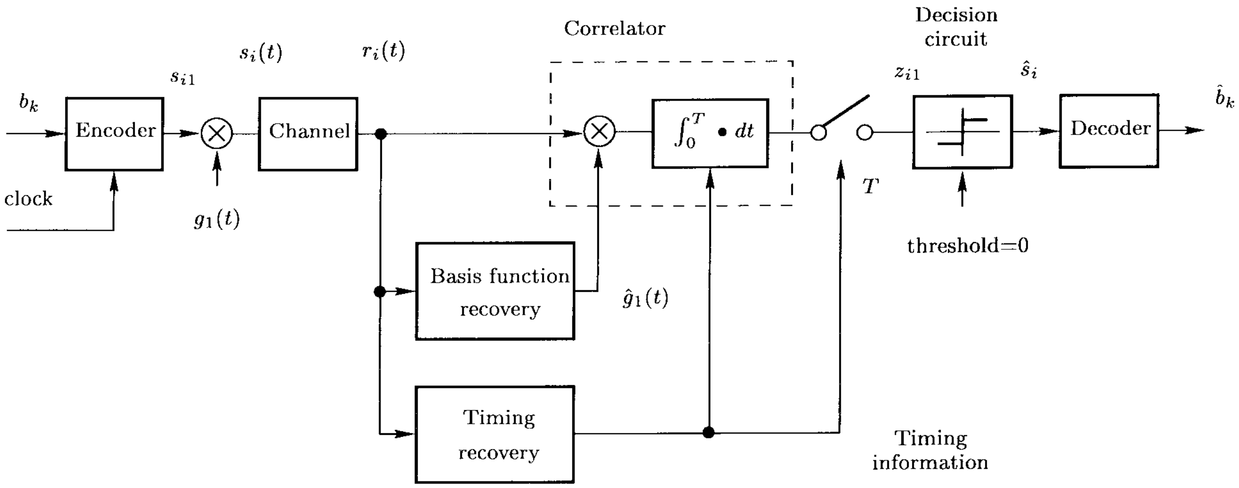

Fig. 5. Block diagram of a coherent BPSK receiver.

than zero at the decision time instant. If the observation signal z_{i1} is less than zero, then the receiver decides that a symbol 1 has been transmitted. The decision circuit is simply a level comparator with zero threshold.

B. Carrier Recovery and Timing Recovery

1) Carrier Recovery and the Need for Synchronization in Coherent Detection: In general, coherent reception requires knowledge of the basis functions at the receiver. Because matched filter receivers cannot be implemented in the analog signal domain, only correlation receivers can be used for coherent detection, and synchronization must be used to recover the basis functions.

In the special case of sinusoidal basis functions, knowledge of both the frequency and phase of a carrier is required. The basis functions are typically recovered from the received noisy signal by means of a suppressed carrier *phase-locked loop* (PLL). In conventional systems, estimation of the frequency and phase of the carrier is called *carrier recovery* [9].

2) Timing Recovery and the Need for Symbol Synchronization in Digital Communications: A second and more important type of synchronization also arises in digital communications. In any practical system, not only an isolated single symbol, but a sequence of symbols, has to be transmitted. To perform demodulation, the receiver has to know precisely the time instants at which the modulation can change its state. That is, it has to know the start and stop times of the individual symbols in order to assign the decision time instants and to determine the time instants when the initial conditions of the correlators have to be reset to zero in the receiver. Determination of these time instants is called *timing recovery* or *symbol synchronization.*

In contrast with carrier recovery, which is an optional step that is required only by coherent receivers, timing recovery is a *necessary* function in digital communications. The decision times at the receiver must be aligned in time (synchronized) with those corresponding to the ends of symbol intervals T at the transmitter. Symbol synchronization must be achieved as soon as possible after transmission begins, and must be maintained throughout the transmission.

This paper is aimed at providing a clear exposition of the important issues in both conventional and chaotic modulation/demodulation techniques. Although symbol synchronization has to be solved in every digital communications system, it belongs to the decision circuit and not to the demodulation process. Therefore, the details of the timing recovery problem are not discussed in this paper. The interested reader can find excellent expositions of timing recovery in [8] and [11], for example.

In the next section, we discuss the advantages and disadvantages of synchronization for basis function recovery in coherent receivers.

C. Advantages and Disadvantages of Synchronization

The main advantage of synchronization is that it makes the implementation of coherent receivers possible. As mentioned in Section IV-C, the most significant feature of a coherent receiver when used with a sinusoidal carrier is that by recovering just one signal, the carrier, (and regenerating the quadrature basis function by means of a simple phase-shifter) a pair of orthonormal basis functions can be generated and therefore a huge signal set can be used (256-QAM, for example). In addition, a coherent receiver has marginally better noise performance than its noncoherent counterpart.

However, there are significant costs associated with synchronization, in terms of synchronization time, circuit complexity, and severe penalties associated with loss of synchronization. In this section, we discuss these issues.

1) Costs Associated with Achieving Synchronization in a Coherent Receiver: In conventional digital communications systems, various types of PLL's are used to perform synchronization [10].

Two basic operation modes have to be distinguished for a PLL. Under normal operating conditions, the phase-locked condition has been achieved and is maintained. The PLL simply follows the phase of an incoming signal; this is called the *tracking mode*.

Before the phase-locked condition is achieved, the PLL operates in a highly nonlinear *capture mode*. The time required for the PLL to achieve the phase-locked condition is called the *pull-in time*. The *transient time*, which is associated with the tracking mode, is always significantly shorter than the pull-in time.

In a digital communications system where the synchronization is lost at the beginning of every new symbol, the received symbol can be estimated from the noisy received signal only after the basis functions have been recovered, i.e., only after the phase-locked condition has been achieved. In this case, the total detection time associated with each bit is the sum of the pull-in time and the estimation time. In particular, a long pull-in time results in a very low symbol rate. This is why *synchronization is always maintained* in the carrier recovery circuits of conventional digital communications systems.

2) Penalties for Failing to Achieve Synchronization in a Coherent Receiver: The block diagram of a coherent correlation receiver is shown in Fig. 4. The received signal is always a stochastic process due to additive channel noise. The observation vector, i.e., the output of the correlators, is an estimation, where the mean value of estimation depends on the bit energy and the "goodness" of recovery of the basis functions. In conventional receivers, the variance of this estimation is determined by the noise spectral density N_0 [4]. The probability of error, i.e., the probability of making wrong decisions, depends on the mean value and variance of estimation. The main disadvantage of synchronization follows from the sensitivity of noise performance to the "goodness" of recovery of the basis functions.

The most serious problem is caused by the cycle slips in PLL's used for recovering a suppressed carrier. Due to the noise, the phase error is a random process in the PLL. If the variance of the phase error is large, cycle slips appear with high probability due to the periodic characteristic of the phase detector [12]. This means that the VCO phase, i.e., the phase of the recovered carrier, slips one or several cycles with respect to the reference phase. Every cycle slip results in a symbol error. For high *signal-to-noise ratio* (SNR), the probability of cycle slips is low, but it increases steeply with increasing noise power. This phenomenon results in a large degradation in noise performance.

Even if the physical transmission medium can be characterized by a pure attenuation, a small error in the phase of the recovered carrier may be present due to nonideal properties of the synchronization circuit. This error generally causes a large degradation in the noise performance of a coherent receiver [5], [13].

A real telecommunications channel always causes some transformation of the basis functions. This problem is especially hard to overcome in coherent receivers if the transmission medium is time-varying. The time-dependent channel transformation requires adaptive control of the recovered basis functions. This can be accomplished only by means of a wideband PLL circuit. However, in this case the recovered basis functions are corrupted by the channel noise passed by the PLL transfer function. This also results in a large degradation in the system's noise performance.

Disturbances and interferences passed by the channel filter may cause a transient in the carrier recovery circuit. The result of this transient is that the recovered basis functions deviate from their ideal values. This also results in performance degradation.

Synchronization can also be lost from time to time due to deep and/or selective fading [14]. Loss of synchronization automatically initiates a pull-in process which means that all symbols received during the pull-in time are lost. Once again, the result is a degradation in BER.

The conclusion is that coherent receivers exploiting synchronization offers the best system performance *if synchronization can be maintained.* However, they do not offer optimum performance if the SNR is low, the propagation conditions are poor, the properties of the channel are time-varying, or if the probability of deep fading is relatively high. In these cases, a more robust modulation scheme such as FSK with a noncoherent receiver must be used.

A further disadvantage of a coherent receiver from an implementation point of view is that it generally requires more complicated circuitry than its noncoherent counterpart.

VI. WHY USE A CHAOTIC CARRIER?

In the previous sections, we have shown that the use of sinusoidal signals as basis functions in conventional digital modulation techniques offers excellent bandwidth efficiency. Moreover, these basis functions can be reconstructed easily by recovering a single sinusoidal carrier at the receiver. Why, then, is it necessary even to consider a modulation scheme which uses anything other than a sinusoidal carrier?

When a sinusoidal carrier is used, the transmitted power is concentrated in a narrow band, thereby resulting in high power spectral density. This has a number of serious drawbacks.

1) Multipath propagation is always present in many important radio applications such as mobile telephony and wireless LAN. It results in very high attenuation over narrow frequency bands. This means that the SNR may become very low or even a dropout may occur in a narrowband communications system. Low SNR results in symbol errors due to the cycle slips in the carrier recovery circuit. The extremely high attenuation causes not only a dropout in reception, but also loss of synchronization. Recall that when synchronization is lost in a coherent receiver, all symbols transmitted during the pull-in time of the receiver's synchronization circuitry are also lost.
2) Due to the high transmitted power spectral density, narrowband communications cause *high levels of interference* with other users. Therefore, they are not suitable for unlicensed radio applications.
3) Narrowband signals are sensitive to narrowband interference.

4) Because of the high transmitted power spectral density, the probability of interception of narrowband communications is high.
5) The reception of messages by an unauthorized receiver is very simple because limited *a priori* knowledge is required for demodulation.

The difficulties summarized above can be overcome by using *spread-spectrum* (SS) techniques, where, in addition to a conventional digital modulation scheme, a pseudorandom spreading sequence is used to spread the spectrum of the transmitted signal [14]. The benefits of spreading can be achieved only if the pseudorandom sequences in the transmitter and receiver are *synchronized*.

Spread-spectrum communications using spreading sequences has two major disadvantages: it is not possible to achieve and maintain synchronization under poor propagation conditions, and the spreading and despreading processes require additional circuitry.

Chaotic signals are wideband signals that can be generated using very simple circuitry. A potentially cost-effective solution for wideband communications is to use a wideband chaotic carrier. In this approach, sample functions of chaotic waveforms are used as basis functions or as the elements of the signal set.

VII. CONCLUDING REMARKS

Much of the recent research in chaotic communications has focused on synchronization. Our objectives in this paper are

1) to provide a theoretical context in which the performance of modulation schemes based on chaotic synchronization can be evaluated;
2) to develop a unified framework for discussing and comparing conventional and chaotic communications systems; and
3) to highlight the special problems that arise when chaotic basis functions are used.

In Section II, we described the major components of a digital communications system and emphasized that the primary source of errors is the analog channel. We identified the minimum requirements for a realistic channel model (additive white Gaussian noise and band-limiting) and illustrated the performance measures by which modulation schemes are judged.

The signal set in a conventional digital communications system is constructed from orthonormal sinusoidal basis functions. In Section III, we explained the advantages of this choice in terms of bandwidth efficiency and easy recovery of basis functions.

In Section IV, we demonstrated the equivalence of correlation and matched filter receivers and explained the ideas of coherent and noncoherent detection. The primary motivation for carrier synchronization is to permit coherent detection. Coherent detection, in turn, offers greater potential bandwidth efficiency.

Under poor propagation conditions, where synchronization cannot be maintained, the advantages of coherent detection are lost. In such circumstances, a noncoherent receiver offers a more robust and less complex solution.

Synchronization can be exploited in two ways in a digital communications system. Synchronization is required for timing recovery to establish the starts and ends of symbols in the transmitted sequence. If coherent detection is used, synchronization is also required to recover the carrier, typically by means of a PLL. In Section V, we showed when and how coherent detection fails.

We motivated the use of a chaotic carrier signal in Section VI by highlighting the disadvantages of narrowband communications when propagation conditions are poor.

In Part II of the paper, we consider the state-of-the-art in synchronization of chaotic systems in the context of digital communications, and current chaotic communications techniques are surveyed.

In Part III, performance targets for chaotic communications techniques are summarized and examples (CSK with synchronization, noncoherent CSK and DCSK correlation receiver) are given. Finally, the performance of these systems is evaluated in the context of a noisy and bandlimited channel.

ACKNOWLEDGMENT

The authors have benefited significantly from extensive interaction with their colleagues in the Chaos Communications Collective, particularly during the "First Little Workshop on Spread Spectrum Communication and Chaos" at the EPFL, Lausanne, on April 2, 1996, and the "Second Little Workshop on Spread Spectrum Communication and Chaos" at UC Berkeley on August 7, 1996. Last, but not least, the authors wish to thank the reviewers for their constructive criticism of this paper.

REFERENCES

[1] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, no. 8, pp. 821–824, 1990.
[2] M. Hasler, "Engineering chaos for secure communication systems," *Phil. Trans. R. Soc. Lond. A*, vol. 353, no. 1701, Oct. 16, 1995, pp. 115–126.
[3] G. Kolumbán, H. Dedieu, J. Schweizer, J. Ennitis, and B. Vizvári, "Performance evaluation and comparison of chaos communications systems," in *Proc. 4th Int. Workshop in Nonlinear Dynam. Electron. Syst.*, Seville, Spain, June 27–28, 1996, pp. 105–110.
[4] S. S. Haykin, *Communication Systems*, 3rd ed. New York: Wiley, 1994.
[5] J. G. Proakis, *Digital Communications.* Singapore: McGraw-Hill, 1983.
[6] M. P. Kennedy, "Communicating with chaos: State of the art and engineering challenges," in *Proc. 4th Int. Workshop in Nonlinear Dynam. Electron. Syst.*, Seville, Spain, June 27–28, 1996, pp. 1–8.
[7] J. S. Bendat and A. G. Piersol, *Measurement and Analysis of Random Data*, 2nd ed. New York: Wiley, 1996.
[8] E. A. Lee and D. G. Messerschmitt, *Digital Communications*, 2nd ed. Boston, MA: Kluwer Academic, 1993.
[9] I. Frigyes, Z. Szabó, and P. Ványai, *Digital Microwave Transmission.* Amsterdam, The Netherlands: Elsevier Science, 1989.
[10] F. M. Gardner, *Phaselock Techniques*, 2nd ed. New York: Wiley, 1979.
[11] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques: Signal Design and Detection.* Englewood Cliffs, NJ: PTR Prentice-Hall, 1995
[12] G. Ascheid and H. Meyr, "Cycle slips in phase-locked loops: A tutorial survey," *IEEE Trans. Commun.*, vol. COM-30, pp. 2228–2241, Oct. 1982.
[13] R. W. Lucky, J. Salz, and E. J. Weldon, *Principles of Data Communication.* New York: McGraw-Hill, 1968.
[14] R. C. Dixon, *Spread Spectrum Communication Systems with Commercial Applications*, 3rd ed. New York: Wiley, 1994.

Géza Kolumbán (M'92) received the M.S. and Ph.D. degrees from the Technical University of Budapest, Hungary, and the C.Sc. degree from the Hungarian Academy of Sciences in 1976, 1990, and 1990, respectively.

After graduation, he was with Fine Mechanical Enterprise, Hungary, as a research engineer where he developed local generators, microwave transistor power amplifiers and VCO circuits for high-capacity microwave analog radio relay systems. He joined the Research Institute for Telecommunications, Hungary, in 1980, where he was involved in many system engineering projects like SCPC-type satellite telecommunication system, microwave satellite up- and down-converter, low-capacity microwave digital radio system, etc. There, he headed a group of engineers, whose duty was to develop frequency synthesizers and local generators for satellite and frequency hopping spread spectrum systems. He spent one year with the Bilkent University, Turkey (1991–1992), and another year with the Eastern Mediterranean University, Cyprus (1992–1993) as an Associate Professor. He returned to the Technical University of Budapest in 1993, where he has been an Associate Professor with the Department of Instrument and Measurement Engineering. His current research interests are in nonlinear dynamics of different-type phase-locked loops, frequency synthesis by sampling phase-locked loop, chaotic communications, and application of chaotic signals in measurement engineering.

Michael Peter Kennedy (S'84–M'91–SM'95), for a photograph and biography, see this issue, p. 854.

Leon O. Chua (S'60–M'62–SM'70–F'74), for a biography, see this issue, p. 904.