Introduction to "Certain Topics in Telegraph Transmission Theory"

NORMAN C. BEAULIEU, FELLOW, IEEE

Invited Paper

Keywords**—***Digital communication, intersymbol interference, Nyquist first criterion, Nyquist rate, Nyquist second criterion, Nyquist third criterion, pulse shaping, sampling theorem.*

I. PREFACE

Harry Nyquist (see Fig. 1) holds an important place in the history of electrical engineering as a pioneer of the mathematical theory of communication. A native of Sweden, Nyquist worked his entire career at AT&T Bell Laboratories in the United States, where he applied his broad knowledge and analytic ability to the most challenging problems in his field. His studies produced fundamental results in modern communications and control engineering.[1]

In April 1928, H. Nyquist published the classic paper, "Certain Topics in Telegraph Transmission Theory" in the *Transactions of the American Institute of Electrical Engineers*. The beginnings of modern communication theory originated with this paper and an earlier companion paper, "Certain Factors Affecting Telegraph Speed" published in April 1924 in the *Bell System Technical Journal*. The former paper reprinted in this issue is the subject of this introduction. It is an elaboration and extension of the earlier paper, giving a "mathematical" in addition to an "engineering" point of view of the subject matter. The paper also comprises ideas, insights, and theories beyond the first paper. The history of digital communications has proved many of the concepts in this paper to be fundamental and they inspired and facilitated future developments in communication theory. As the title indicates, a number of important and interesting topics are considered. In this introduction, I highlight some, though not all, of the topics and results. First, a biography of Nyquist is presented.

Manuscript received November 19, 2001.

The author is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada.

Publisher Item Identifier S 0018-9219(02)01130-1.

[1]Biographical information courtesy of IEEE History Center.

Fig. 1. Harry Nyquist.

II. BIOGRAPHY

Nyquist was born in 1889 in Nilsby, Sweden. He immigrated to the United States in 1907 to pursue scientific training. After obtaining the B.A., B.S., and M.S. degrees from the University of North Dakota, Nyquist earned the Ph.D. degree in physics from Yale in 1917. He began working for AT&T, developing a metallic telegraph system and devising methods to increase the message handling capability of narrow band telegraph channels. His studies of telegraphy led to his publication of the Nyquist rate in 1928, a relation giving the maximum frequency of pulses that can be transmitted within a given bandwidth. His work in this area taught engineers how to use samples of waveforms to reconstruct complex signals, the foundation for modern pulse code modulation systems. In 1928, Nyquist also gave a mathematical analysis of the thermal noise in circuits observed by Johnson, which clarified the implications of the phenomenon for communication circuits.

In 1932, Nyquist contributed significantly to control engineering by providing the first analysis of stability in feedback amplifiers. He developed the Nyquist diagram, which gave a graphical test for stability and articulated the Nyquist criterion, which simplified circuit design by predicting the allowable limits of gain and internal phase shift in amplifier circuits. With these tools, engineers were able to enhance long distance telephone service through the use of feedback amplifiers.

Between 1921 and 1935, Nyquist devoted much time to television research. He invented a method of transmission that is used for television broadcasting today and he discovered how to correct the problem of delay distortion of television images. After 1935, Nyquist engaged in general communication research. He retired from Bell Laboratories in 1954 after reaching the position of Assistant Director of System Studies. Thereafter, he worked as a frequent consultant to several companies and to the U.S. government.

Before his death in 1976, Nyquist received many honors for his outstanding work in communications. He was the fourth person to receive the National Academy of Engineering's Founder's Medal "in recognition of his many fundamental contributions to engineering." In 1960, he received the Stuart Ballantine Medal of the Franklin Institute and IRE's Medal of Honor. The following year, he was honored as the recipient of the Mervin J. Kelly Award for his fundamental role in the evolution of modern communications and control theory.

III. Two Axioms of Digital Communications

At a time when the only data communications in use was the telegraph, Nyquist formulated the fundamental concepts and principles of digital communications. To set the stage for our appreciation of the clarity of his vision and the depth of his insights, let us recall the information transmission technology of the day. Telephone and telegraph signals were transmitted on the same wirelines. The analog telephone signals were transmitted as bandpass signals at higher frequencies than the telegraph signals which occupied the bandwidth of the wireline channel from direct current (dc) to the lower frequency limit of the bandpass-analog telephone signals. The telegraph signals were composed from three elements—dots, spaces and dashes, each of the former two occupying one unit of time and the latter occupying three units of time. In the *Preliminary Discussion*, the paper begins by considering the data communications problem intuitively. Contemplating a simple telegraph system, Nyquist arrives at the first two axioms of digital communications:

1) "The time is divided into approximately equal units which will be called time units."
2) "There is a finite number of conditions and each time unit is characterized by a single one of these conditions."

Thus, we find in the paper an early statement that digital communication systems will clock signaling intervals and the concept that information will be conveyed by altering some property or condition of the signal in each signaling interval. Further, it is specified that there is a "finite" number of conditions and that in the case of the simple telegraph, "the total number of conditions is two, open and close. This is in sharp distinction to the case of telephony where there are neither simple numerical relations between the various time intervals, nor a finite number of possible current values."

Following the establishment of the two axioms, it is identified that the distortion of the wireline will smear a signaling waveform, which originally occupies a given time period into other time periods overlapping signaling waveforms corresponding to other signaling times. Nyquist goes on to say that in the case of overlapping signals, it is important to consider the full distorted waveform including the part outside of its own time unit. The term "signal element" is introduced to specify the full waveform ascribed to a given symbol period and in a hint of things to come, it is proposed that signal shaping may be used to purposely extend the signal element into adjacent time units. Thus, the overlapping of the signal elements turns from being an implied detriment to being an implied benefit. Subsequent discussion suggests that the signal elements should be the same, differing only by a factor dictated from the time unit corresponding to the signal element. "In the general case any signal element may be expressed by the product

$$a_h f(t) \tag{1}$$

where a_h is a real factor which may vary from one signal element to another and where $f(t)$ is a function of time. The origin of t is, of course, a fixed instant with respect to the signal element. The function $f(t)$ will be called the wave form. It is determined by the wave form at the sending end and by the transmission characteristics of the system between the sending end and the point under consideration. It is not affected by the particular signal or form of intelligence being transmitted over the system. The factor a_h will be called the magnitude factor. It differs from one signal element to another but is the same in all parts of the system and, in fact, in all systems transmitting the same signals." Equation (1) defines pulse amplitude modulation (PAM).

The *Preliminary Discussion* concludes prophetically, asserting that "In the illustrations given, the sent waves were voltage waves. This is not necessarily the case in order for the discussion to apply. The sent wave may be a current wave, or it need not even be electrical. By providing suitable coupling the sent wave may be … in the form of variations of a light beam." Interestingly, the word variations perhaps suggests that Dr. Nyquist was thinking of more than a simple on-off light beam.

IV. The Nyquist Rate

The section *Proportionality Between Speed of Signaling and Transmitted Frequency Band*, written in qualitative terms, introduces the mathematical technique used throughout the sequel. Detailed examinations of the frequency spectra associated with digital signals will be undertaken by constructing periodic extensions of the entire transmitted message and using the mathematics of Fourier series. "While this is convenient, it constitutes no real limitation on the generality of the analysis, because the intervals of repetition may be made as large as desirable. There is nothing to prevent us from making the interval very great, say, an hour or a year." Leaving the detailed mathematics for later sections and using physical reasoning it is "concluded that for any given deformation of the received signal the transmitted frequency range must be increased in direct proportion to the signaling speed… The conclusion is that the frequency band is directly proportional to the speed." This section concludes by introducing the next section with the question, "What is the limiting value of this factor under ideal conditions?"

It is in the section *Analysis of D-c. Wave* and Appendix I that the famous Nyquist rate is established and Nyquist's sampling result is found. Using a spectral analysis of a Fourier series of a periodic extension of the total message signal, Nyquist establishes that the maximum signaling rate that can be supported over a baseband (dc) telegraph channel of given bandwidth W is 2W pulses per second. The intuitive notion behind the mathematical reasoning presented is the following. Each frequency component has two free parameters associated with it, the amplitudes of the corresponding cosine and sine quadrature components. Thus, "it follows that a number of components approximately equal to one-half the number of signal elements should be sufficient to determine the a's completely, provided that the components are chosen to be mutually independent." Nyquist's development of the maximum signaling rate is interesting in that it initially assumes an arbitrary signal and, therefore, Nyquist's result is equivalent to a version of the sampling theorem for bandlimited signals later given precisely and rigorously by Shannon.

Appendix III concludes by asking and answering an intriguing question. It was shown that a bandwidth of W is necessary to signal at a rate of 2W symbols per second without intersymbol interference and using pulses shaped according to Nyquist's first criterion. "It might be questioned whether the proposition is true when this number is small. Consider the case where there are just two, distinct" signaling levels, say, $+1$ and -1. Since only two signaling levels are involved and they are of opposite sign, we may consider it necessary and sufficient that the received signal at the sampling instants have the correct sign, regardless of the magnitudes of the samples. By considering the spectral components of particular data sequences, it is shown that this cannot be guaranteed when the signaling bandwidth is less than the Nyquist bandwidth.

V. Nyquist's First Criterion

The paper examines a number of conditions for signal waveform shaping in the section *Distortionless Transmission* and Appendixes II and III. The first criterion for distortionless transmission of a data signaling waveform is that "a wave will be said to be nondistorting when the value at the mid-instant of any time unit is proportional to the magnitude factor for the corresponding element." That is, the contribution from other signaling elements must be zero at the midpoint of the signaling interval. "Now an important property of the… wave is that its zeros are located at equal intervals, which is precisely the property required of a nondistorting wave under the criterion discussed… the interference due to all previous signals elements is zero because the interference due to each one of them is zero. It follows that any receiving device, which is made to function by the current at that point will be distortionless as far as the interference from adjacent signal elements is concerned." Solving simultaneous linear equations, the pulse $f(t) = (\sin 2\pi Wt)/2\pi Wt$ is derived as a minimum bandwidth pulse that has no intersymbol interference at the sampling instants. It is then proved that there is an infinite number of "shape factors," which may be added to the frequency spectrum of this pulse while maintaining the intersymbol interference-free property. In particular, any real shape factor having odd symmetry about the point at the bandedge of the frequency spectrum preserves the intersymbol interference-free property. Such shape factors increase the bandwidth required for the pulse transmission. However, the tails of the minimum bandwidth pulse decay as the inverse first power of time; this is problematic for practical filter design and for symbol timing recovery. The time rate of decay of the pulse tails can be made greater (higher inverse powers of time) by using appropriate shape factors that expand the bandwidth required. Subsequent design practice has traded off this additional bandwidth against symbol timing recovery jitter. It is also shown that the addition of an imaginary shape factor that has even symmetry about the bandedge point retains the intersymbol interference-free property. This fact has achieved less application as high rate transmission systems use both quadrature carriers.

VI. Nyquist's Second Criterion

Nyquist's second "criterion for distortionless transmission is that the interval, between the instants when the received wave passes through the mean value, shall be the same as the corresponding interval at the transmitting end. A receiving device which responds to the values of the wave at the ends of the time units, instead of the middle, will give distortionless transmission… This is the criterion in ordinary land-line telegraphy." In this case, a time sample taken at the end of the symbol period, rather than at the center, is proportional to $(a_h + a_{h+1})/2$. That is, there is controlled intersymbol interference at the sampling instant since the sample is determined by a_h and a_{h+1} and has zero intersymbol interference from signaling elements other than the two corresponding to a_h and a_{h+1}. As was the case for the first criterion, solution of simultaneous linear equations is used to determine a simple frequency spectrum satisfying the criterion giving $S(f) = 2\cos(\pi f/2W)$.

This signal frequency specification later appeared at the core of much research on partial response signaling. The term partial response is a reference to the fact that the receiver filter response to an input symbol at the sampling instants is spread over two symbol periods rather than being restricted to one. Since the tails of the minimum bandwidth pulse decay as t^{-1}, it is desirable to use inband shaping (i.e., shaping that does not increase the bandwidth of the signaling element) to increase the rate of decay of the time pulse tails. This is achieved by partial response signaling. Since the intersymbol interference is controlled, the data is still recoverable.

Nyquist also gives shape factors that preserve the second condition for distortionless transmission. Namely, any real shape factor possessing even symmetry about the bandedge point and any imaginary shape factor having odd symmetry about the bandedge point can be added to the basic cosine frequency shape preserving the criterion while expanding the bandwidth.

VII. Nyquist's Third Criterion

The third Nyquist criterion for distortionless digital transmission is the least well known. It "is that the area under the received wave shall have the same value as that under the sent wave during each time unit" or symbol period. "This case may be of interest in some methods of picture transmission where the integrated exposure over a small interval of time may be important." Unanticipated by Nyquist, this criterion gained unexpected relevance 50 years later in the context of digital frequency modulation schemes, in particular, tamed frequency modulation. In such systems, it is desirable to avoid sharp edges in the phase path and the phase is the integral of the frequency. A minimum bandwidth frequency shape factor that satisfies the third criterion is derived $S(f) = \pi f/(\sin(\pi f/2W))$, $0 \leq f < W$; $S(f) = \pi/2$, $f = W$; $S(f) = 0$, elsewhere. Discussion of the determination of shape factors that can be added to the basic frequency shape factor while preserving the third criterion is also found in Appendix III.

VIII. Carrier Systems

Following the establishment and examination of the three criteria for distortionless transmission of digital signals, which were developed for baseband (dc) signaling, the section *Analysis of Carrier Wave* and Appendixes IV and V investigate digital signaling at carrier frequencies considering the spectrum of the modulated carrier and single sideband transmission. The frequency spectral analysis shows the need "to relate the speed of signaling and the carrier frequency." Importantly, when the upper and lower sidebands are symmetrical, "the total frequency range required… for a given speed of signaling is just twice that required in the direct-current case. Except for this, the results as to bandwidth required and ideal shape factors for the received wave are the same as obtained above for the direct-current case… It is obvious that this type of carrier telegraph is relatively inefficient in its utilization of the available frequency range." The paper then describes that, as had been previously proposed, by using two quadrature carriers $\sin \omega t$ and $\cos \omega t$ at the carrier frequency, "the total amount of intelligence transmitted for a given bandwidth is twice as great… and is, in fact, the same as for the direct-current case."

IX. Conclusion

I conclude this introduction by pointing out that there are more ideas and results in the 26 pages that I have not highlighted here. Tribute to the communications contributions of electrical engineering pioneer Harry Nyquist and to this paper was paid by the father of information theory, Claude E. Shannon, in his famous paper "A Mathematical Theory of Communication." In the first paragraph, Shannon states that a basis for a general theory of communication is contained in the important papers of Nyquist and Hartley on this subject. Nyquist's important results are standard material in undergraduate and graduate curricula in electrical engineering. They will continue to inspire and influence engineers and scientists for at least as long as they already have.

Acknowledgment

The author would like to thank Professor P. J. McLane for his generous and constructive feedback on this and other works.

References

[1] H. Nyquist, "Certain topics in telegraph transmission theory," *Trans. AIEE*, vol. 47, pp. 617–644, Apr. 1928.

[2] ——, "Certain factors affecting telegraph speed," *Bell Syst. Tech. J.*, vol. 3, pp. 324–346, Apr. 1924.

[3] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[4] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communications Principles*. New York: Plenum, 1992.

[5] P. C. Mabon, *Mission Communications—The Story of Bell Laboratories*. Murray Hill, NJ: Bell Telephone Labs, 1975.

[6] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, July 1948.

Norman C. Beaulieu (Fellow, IEEE) received the B.A.Sc. (honors), M.A.Sc., and Ph.D. degrees in electrical engineering from the University of British Columbia, Vancouver, BC, Canada, in 1980, 1983, and 1986, respectively.

He was a Queen's National Scholar Assistant Professor with the Department of Electrical Engineering, Queen's University, Kingston, ON, Canada, from September 1986 to June 1988, Associate Professor from July 1988 to June 1993, and Professor from July 1993 to August 2000. In September 2000, he became the iCORE Research Chair in Broadband Wireless Communications at the University of Alberta, Edmonton, AB, Canada, and in January 2001, the Canada Research Chair in Broadband Wireless Communications. His current research interests include broadband digital communications systems, fading channel modeling and simulation, pulse shaping, interference prediction and cancellation, decision-feedback equalization and communication theory.

Dr. Beaulieu received the University of British Columbia Special University Prize in Applied Science in 1980 and the Natural Science and Engineering Research Council (NSERC) E.W.R. Steacie Memorial Fellowship in 1999. He is a Member of the IEEE Communication Theory Committee and served as its Representative to the Technical Program Committee of the 1991 International Conference on Communications and as Co-Representative to the Technical Program Committee of the 1993 International Conference on Communications and the 1996 International Conference on Communications. He was General Chair of the Sixth Communication Theory Mini-Conference in association with GLOBECOM'97 and Co-Chair of the Canadian Workshop on Information Theory 1999. He has been Editor of Wireless Communication Theory of the IEEE Transactions on Communications since January 1992, an Associate Editor for Wireless Communication Theory of the IEEE Communications Letters since November 1996, Editor-in-Chief of the IEEE Transactions on Communications since January 2000, and on the Editorial Board of the Proceedings of the IEEE since November 2000.