IEEE Information Theory Society Newsletter

Vol. 55, No. 3, September 2005 Editor: Lance C. Pérez ISSN 1059-2362

I am writing this column from the Georgia Tech campus in Metz, France, where I have been for much of the last two years. During that time I have come to appreciate the efforts and roles of the local chapters and individuals in Europe. I have participated in events sponsored by the German Chapter and the Benelux Chapter, and had the chance to visit colleagues in Switzerland, Spain, Italy, Portugal, Ireland, Belgium, the UK, and of course France. It is very gratifying to

product IEEE sells to libraries, companies and institutions. IEEE is in the process of phasing-in a new financial model that rewards high quality content in IEL. Our revenue from IEL is tied directly to the number of papers in IEL

IEEE

IEEE I a T S c N s
(USPS 360-350) is published quarterly by the Information Theory Society of the Institute of Electrical and Electronics Engineers, Inc.

Headquarters: 3 Park Avenue, 17th Floor, New York, NY 10016-5997.

At the kind invitation of the Editor, Prof. Lance C. Pérez, I have been asked to share a few thoughts about our field of Communication Theory (CT) and its components: Information Theory (IT) [1], and "Signal Processing" (SP), on this occasion of my eighty-fifth birthday and in the sixty-third year (since 1942) of my professional activity in the field. The following few remarks are my purely personal thoughts. I hope they may be of some interest and possibly some pertinence. ([2], 2000)

First, let me "locate" what I mean by Communication Theory (CT) and its major components in the general spectrum of physical science, and from this, locate my own views of and interests in CT. The rather ad hoc sketch in figure 1 may help.

Here D+E signifies Detection and Estimation and CP, channel physics, propagation, etc. Communication Theory actually has a wider scope than the diagram indicates, since it is a component one way or another in all scientific activity.

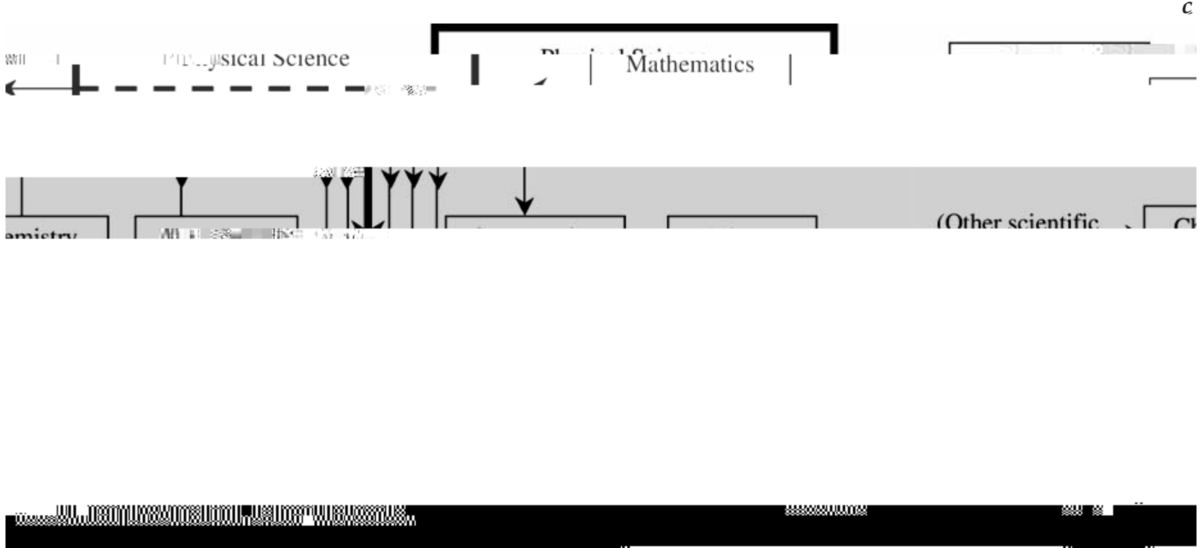
Briefly put, science is itself phenomenological: it builds models of reality, subject to the tests of experiment. Technology, broadly stated, is the application of the results of scientific exploration, i.e. verified models to public use. The models themselves have a hierarchy of refinement: from broad generality to different levels of subtlety, from the macroscopic to the microscopic, to the molecular, atomic, and subatomic; from the continuum to quantum discreteness, each having its mathematical description. At all levels, some type of Communication Theory and its accompanying technology play a central role: conveying information from one space-time point to another such point, with (ideally) minimal interference, i.e. "noise". My own interests in CT have centered largely on the middle ground where the continuum model is appropriate, on the applied physics and mathematics of most signal processing and noise modeling, usually above the quantum level.

I. Reflections: Next, I'd like to make a few remarks about where and how we are today in CT, at least as I see it: Let me begin with

(1) *P b ca s (A s)*: Papers being authored in CT in the US today are going strong, but the authorship is very skewed toward foreign or foreign-born authors. The same is observed in the number of Ph.D. candidates (in CT) in our own graduate schools. A quick look at the reference lists of papers in the IT and SP forwards shows approximately 85-90% foreign vs. indigenous authors. Other technical journals indicate the same orders of magnitude. This is not intended as a criticism of these authors and students. Rather, they are to be commended for their hard work and expertise. It is we "natives" who seem to have largely abandoned these difficult fields of study. Our native students prefer the pursuit of money early on, and leave the science and technology to others. I think this is one major and long lasting effect of the Vietnam War – which effectively punched at least a ten-year hole in the continuity of growing up. (I myself have two children who joined the rebellion in the sixties and seventies ("Haight-Ashbury, Communes," LSD, etc.), who "dropped out" for ten years. To their ultimate credit, they did go back to school and college, as thirty-year old freshmen, and graduated as magna and summa students.)

But those ten years are still with us, expressed in the excessive consumerism of our society and the abandonment of the difficult disciplines [3]. It is the recent immigrants and temporary visitors who in the main carry on and push the progress of our science and technology. And when they leave, as many are now temporarily here? The US is losing its cutting edge, and its advantages in science and technology, which maintains our still envied economy. I hope I am unduly pessimistic, but as a child of the Great Depression ('29-'40), a young man during World

c , i , ed. age 5



Our International Symposia are always sources of intense memories. So many of us gather at remote places and our individual or common experiences are amplified and heightened in the course of the week-long interaction and coexistence. In addition to memorable technical "happenings", there are numerous personal events that frame our impressions and memories.

Let me focus today on the 1983 ISIT that took place in the vicinity of Mont Tremblant (literally, Trembling Mountain), in a beautiful but rustic resort to which winter sport enthusiasts and summer nature lovers flock every year from nearby Montreal as well as from elsewhere in Quebec, Canada, North America, and the world. The Symposium took place in late September / early October when, despite the turning fall foliage, the resort is considered "off-season" (and, hence, made available to us at logical prices). The name of the exact location was Gray Rocks and the locality is named St. Jovite. The registration fee was $110 US and the room rate $56 US per day (…those were the days!).

It was the first ISIT in Canada and the proud co-chairs were John Anderson and Ian Blake. There were three (3) vice-chairmen (the IT Society always had creative ways of elevating people to positions of … well, some prominence). They were Vijay Bhargava, Bob Gray, and G. Seguin. The 14-member (yes, you read right, fourteen) technical program committee was chaired by Lorne Campbell and included people like J. Koplowitz, T. Papantoni-Kazakos, J. Modestino, and S.

War II, of our *s* of the forties-sixties, and as a (very) amateur student of history, I see our future as a country as that of a declining empire (to be compared with that of Great Britain, for example).

However, like Henry V - "once more into the breach" - perhaps we can through understanding and wise actions avoid this historical outcome. It is not yet completely inevitable.

(2) *Sc a s* : What does this mean to me, at least in the context of our published work? It means primarily a serious attempt at adequate referencing (and its use), of original sources and subsequent (significant) contributions based on them. I note, consulting the aforementioned journals, for example, that most references (i.e., at least 80%) go back in time only about 10 years for papers and books. But no subject is so new that it doesn't have a key "old" paper or two or three for the "new wave" hot topic. With the search engines available (Google, AOL, and others), there surely must be worthy material preceding the citations listed. (However, see my cautionary comments under "Computers" following below.) There is the danger of "reinventing the wheel" here.

I attribute this effect to youth. (A youth I am no longer, and I recognize that 50-60 years ago the flood of papers was a trickle compared to the present. Adequate referencing was then much easier (but not, of course, perfect). Also, my teachers were older (I call them the Zeroth Generation), and more careful about this kind of thing. Moreover, the rush to publish was more restrained, in fact, downright leisurely!) Finally, in any field (particularly in ours), the research generation is about 15 years - from Ph.D. to some form of professorial life. Thus, there are four generations (including mine, the First) of teachers, between me and the current crop of Assistant Professors. No one can grasp all the relevant material in any one generation, particularly, with the increasing avalanche of papers coming along. Much of the valued past is being lost, inevitably, but experience can be at least a partial corrective - one gets to know many of the critical papers.

(3) *W S + T*: This is hardly an original topic now but it deserves attention. We badly need to encourage more women into S + T I realize that the critical times of encouragement are in the early teens, but at least we can be encouraging at the undergraduate level. More young women in S + T would go a long way to repairing the gap produced by the male drop-out noted above in (1). This sounds like PC but that it is not my intent: an original mind is always welcome. This is a valuable resource which needs to be greatly developed.

(4) *T C - S R a s*: Let me begin this complaint (?) by stating the obvious up front: the computer, in a very important way, has saved science. It is a prodigious instrument for solving problems which, though formulated, could never be handled otherwise. It also makes it possible to translate (read: obtain numbers for) the "macro-algorithms" of theoretical and applied physics, technology, etc. into specific, quantitative results. In short, it is a magnificent tool, a magnificent symbiote for every scientific discipline. (I am not mentioning its misuses, trivial and unfortunate applications for games, trivial messages, gossip, invasion of privacy, identity theft, and so on.)

But, the computer is not a substitute for thought [4] — [8], not only logical thought but intentional, inspirational, and other forms of human mental creativity. (This gets into Philosophy, unsolvability (Godel's theorem, etc.), paradoxes, Platonism vs. inspirationism, etc. All I will say is that computers may be developed to have consciousness, but it will always be a machine consciousness, not a human one, with an ultimate Turing test which still does not distinguish any difference. (I refer the reader to Penrose [8] for a fascinating discussion.))

My practical complaint is one involving the human-machine interface with regard to finding key references. I maintain that in the finite time available to a human being it is not possible to obtain and select all (finite no.) of the key references in a scientific field, or for that matter, in a reasonably-sized subfield, for instance D + E in "Signal-Processing" (SP). Even with a sufficiently detailed interrogation of the computer (say, updated Google, etc.), which also takes time to compose, one will miss key papers. Why? Because it takes a comparatively long time to read and understand the abstract (the title is only enough to wet one's appetite). One can end up, in a finite number of cases, with say a couple of thousand abstracts, of which maybe only a hundred (to an experienced interrogator) appear relevant. Winnowing these down to, say ten takes time (which is always at a premium), which then gets the reader's full attention. I'll wager that the whole process can take a week [for 2000 $\rightarrow$ 100 $\rightarrow$ 10 $\rightarrow$ (?) scientific papers], possibly longer. In most cases this becomes a waste of time (unless, possibly, one is a technical historian, who usually does not have that high level of expertise to make discerning decisions). In the interest of overall efficiency, one picks one or some of the papers, without a guarantee of selecting the one (or two) really important ones. Being knowledgeable in the field is a very great help, but one always runs the risk of missing key material. (This is my a priori apology for the missing ones.)

It is worth pointing out, finally, that preservation of desired material by computers is dubious. (Paper (or papyrus) is much longer lived!) There is also the problem for the observer / user — one does not have the old computers to read the old stuff? (Ask the Librarians!)

(5) *S F R s a c T cs*: I would like to suggest a number of areas worthy of attention — some of which have already been mentioned (see Verdu, IT Newsletter of December, 1994, on Shannon Theory and Imai's "President's Column," IT Newsletter of December 2004.) They are:

- (i) More physics of the channel and EM environment (see (6) below). Space-time extensions of present MIMO applications as well as spatial diversity generally;
- (ii). Nongaussain noise (physical models) as important interference in dealing with the coding process for transmission and reception, etc.;
- (iii). Shannon Theory in nongaussian noise: limits on channel capacity, finite time effects, etc.;
- (iv). Hardware and software design for the above.

And many others I'm sure the reader can think of.

(6). *T B* . On a personal note, let me briefly describe my book-in-progress, now half-done. The tentative titles are, at this point:

(A I d c) N Ga ss a C ca T , N Ga ss a C ca P s cs, (P a a , N s , a d S a P c ss Ca ca C a),

subject to some possible modulation – a small volume of some 1200 pages (!).

The emphasis here is on discrete sampling of space-time signal and noise fields, jointly coupled detection and estimation for improved performance, four-dimensional matched filters, and physical non-gaussian noise models. Scattering Theory – classical and now probabilistic treatments, as well as Ambient Noise models. Considerable attention is also given to Threshold Signal Theory, i.e. the detection and estimation of weak signals in (generally), nongaussian noise, and to some aspects of “learning”, with sequences of decisions (i.e. tracking). Doppler effects, fading, and other channel modifications are included. These, and other examples, occupy three of the four parts of the book. A fourth part is devoted to special topics: noise signals, path integrals, some elements of optical communications and a brief excursion into quantum effects, with applications to astrophysics, matched field processing, among others. Communication Theory has become such an all-inclusive field that it cannot be technically described in one volume, even by the present volume, which treats Signal Processing and the Canonical Channel.

The Book is necessarily incomplete: many important topics are unavoidably omitted due to space and knowledge constraints. Hence, the necessity for the word “Introduction”. This volume is intended to complement my earlier book and the many important works by others on Information Theory (i.e. Coding etc.): see Figure 1 above. It is conceptually related to the coding world through the generality of the signal waveforms chosen (but not otherwise specified here in detail), and by the physical character of the prototypical channel specified therein. [Finally, I expect to finish this work in a couple more years, then off to the Publisher (IEEE Press + Wiley) by ‘07 (or ‘08).]

(7) *R sc c s*: My final topic here is a personal note, in memoriam for the many professional acquaintances, teachers, colleagues, and friends, whom I worked with or have know personally and who at this writing are no longer with us. They are among the ones (the Zeroth Generation) who largely created the Communication Theory field [9], during and since World War II, and thereafter by the succeeding 0th generations.† I list them more or less chronologically (with approximate dates and places of my initial meeting)

(Later >’45):	S.O. Rice (1943 — Bell Labs)	J.A. Stratton (1948 — MIT); Pres. (MIT)
(Nobel L.)	J.H. VanVleck (1943 — Harvard + RRL)	W.B. Davenport (1946 — MIT)
(Nobel L.)	Felix Bloch (1944 —" ") N. Wiener (1946 — MIT) C. Shannon (1947 — MIT — Bell Labs.)	W. Root (1946 _ MIT) Kolmogoroff (1973, Moscow), Math.
(Nobel L.)	W. Brattain (1948 — Bell Labs.) A. Siegert (1943 —MIT — Rad. Lab.)	Pinsker (1973, 76, Moscow) *Brekhovskii (1973, Moscow), Acoustic Inst.

Uhlenbeck (1943 – MIT – U. Mich. – Rad Lab.)	*Lysanov (1973, Moscow), “
D.O. North (1943 – RCA)	*Ol’Shevskii (1973, Moscow), “
J. Van Newmann (1947 — Princeton)	
L. Brillouin (1947 — Harvard)	
J.B. Wiesner (1947 — Rad. Lab. — MIT); Pres.	
MIT M. Hammermesh (1944 — Harvard, RRL)	

(and a few more, whom I can’t remember without my archives, many of which have been sent to MIT). The presence of the large number of physicists stems from the fact that it was mainly a physicists’ war [93] with the engineers taking over predominantly since about 1950. (I was Research Assistant to Van Vleck during the war (late ‘42 — ‘44) and did my Ph.D. thesis (1947) under his direction [2]. I still have swell memories of my time at RRL and MIT when my World was young. “O Brave New World, that had much creatures in it”.

† Among the living colleagues of my generation, are Irving Reed and Nelson Blackman, and possibly others of the 1st generation.

* These may be still alive, among my Russian friends.

References:

[1]. L. Brillouin, 1962. Science and Information Theory, Academic Press (New York),

[2]. D. Middleton, "A Brief Personal History in Science and Information Theory: April 19, 1920" IEEE Information Theory Soc. Newsletter, June, 2000. See also, ibid, Dec. 1995, "A Conversation with David Middleton" and "Some Personal Reminiscences: Communication Theory and Nobel Prize," IEEE Communications Society Mag., July 1978, pp. 9, 10.

[3]. T. Friedman, "It's a Flat World, After All," NY Times Magazine, 4/03/05, p. 33-.

[4]. E. Rothstein, "Deciphering the Grammer of Mind, Music, and Math," p. A19, N.Y. Times, Sat. 6/19/04.

[5]. K. Hafner, "Digital Memories, Piling Up, May Prove Fleeting," N.Y. Times, Wednesday, Nov. 10, 2004.

[6]. E. Ullman, "The Boss in the Machine," N.Y. Times, 2/19/05.

[7]. J. Fallows, "Enough Keyword Searches. Just Answer My Question," N.Y. Times, Techno Files, 6/12/05.

[8]. R. Penrose, The Emperor's New Mind, (Concerning Computers, Minds, and the Laws of Physics), Oxford Univ. Press (New York), 1989.

[9]. P. Baxter, Scientists Against Time

Glossary:

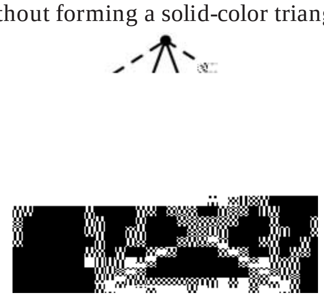
CT	Communication Theory
IT	Information Theory
SP	Signal Processing
D+E	Detection and Estimation
C.P. (C. Phys.)	Channel Physics
S+T	Science & Technology
EM	Electromagnetic
US	United States
PC	Political Correctness

Dr. Norman C. Beaulieu, Department of Electrical and Computer Engineering, has received the Thomas W. Eadie Medal awarded by the Royal Society of Canada in recognition of major contributions to Engineering and Applied Science. The Royal Society of Canada, the Canadian Academy of the Sciences and Humanities, presents medals and awards to Canadians for extraordinary achievement in the social sciences, humanities, and pure and applied sciences. The Thomas W. Eadie Medal, in recognition of major contributions to Engineering of Applied Science, with preference given to those having an impact on communications, in particular the development of the internet, is awarded thanks to the generous financial support of Bell Canada.

The citation of the award reads, in part, "Norman Beaulieu, FRSC, is Professor, iCORE Research Chair and Canada Research Chair in the Department of Electrical and Computer Engineering at the University of Alberta. He is a scientific leader in the analysis and modelling of wireless communications systems. He has discovered ingenious mathematical solutions and models for a wide range of digital communications components and applications, including prediction of coverage areas and outage rates of cellular telephony systems, error rate performances of interference hampered receivers, error propagation in decision feedback equalizers, efficient and statistically accurate simulation methods and tools, and novel electrical pulse shapes for data modems. International researchers have widely used his methods, models and results to design wireless communication components and systems, and to predict the quality of service experienced by users of wireless networks."

Professor Beaulieu is also internationally recognized for his lead-

The original **Ramsey Theorem** is often stated as follows: In any collection of six people, there will always be either three people mutually acquainted, or three mutually unacquainted. In graph theory terms, this says that if K_6 is the "complete graph" on 6 points (i.e.~there is an edge between each pair of points, for a total of $\binom{6}{2} = 15$ edges), if two colors are used to color the 15 edges there must always be a solid-color triangle (3 points connected by 3 edges of the same color). In contrast, the $\binom{5}{2} = 10$ edges of K_5 *ca* be 2-colored without forming a solid-color triangle, as shown:



(Here the two colors are represented by solid or dotted lines.)

1. How many of the 15 edges of K_6 must be deleted so that the remaining edges can be 2-colored without forming a solid-color triangle?

2. How many of the $\binom{10}{2} = 45$ edges of K_{10} must be deleted so that the remaining edges can be 2-colored without forming a solid-color triangle?

3. It is known that if the $\binom{17}{2} = 136$ edges of K_{17} are colored using 3 colors (i.e.~3-colored), a solid-color triangle must be formed, but that it is possible to 3-color the $\binom{16}{2} = 120$ edges of K_{16} without forming a solid-color triangle. How many edges of K_{17} must be deleted so that the remaining edges can be 3-colored without forming a solid-color triangle?

4. Let $= (c)$ be the smallest positive integer such that, if the $\binom{}{2}$ edges of K are colored using c colors, then there must be a solid-color triangle. How many of the edges of K must be deleted so that the remaining edges can be c-colored without forming a solid-color triangle anywhere? (Surprisingly, the answer to this question does not depend on knowing the value of for the given value of c.)

Following the appeal in the President's Column from March this year, stressing out the importance of doing a better job reaching out to other communities, and in the hope of pleasing our Society's historian A. Ephremides , we developed an information theoretical method helping historians clarify disputed authorship attribution of texts, e.g. of the so called "Federalist papers".

The Federalist Papers were written and published during the years 1787 and 1788 in several New York State newspapers. Their purpose was to persuade New York voters to ratify the proposed constitution. In total, the Federalist Papers consist of 85 essays outlining how the new government would operate under the new constitution and why this type of government was the best choice for the United States of America. All of the essays were written by A. Hamilton, J. Madison, and J. Jay under the pseudonym Publius. Madison, widely recognized as the Father of the Constitution, would later go on to become President of the United States. Hamilton would serve in the Cabinet and Jay would become the first Chief Justice of the US Supreme Court. Altogether Hamilton wrote 52 of these essays, Madison wrote 16, and Jay contributed 5. The authorship of the remaining 12 is disputed. These are essays No. 49-58, 62 and 63. The majority of historians believe that they were all written by Madison and the results obtained by our method support this claim. The Federalist Papers discuss very similar topics and are written in an almost identical style typical for political discourse of that time. It is therefore considered a very challenging task for purely algorithmic approaches to correctly determine the authorship attribution of these essays.

When it comes to assessing relatedness, mutual information is the intuitive tool of choice for an information theorist. Mutual information precisely describes the amount of information shared by stochastic processes and can thus be used to derive distance measures quantifying the similarity of these processes. Different authors writing essays, modeled as sources generating messages, can be regarded as such stochastic processes. The task of authorship attribution is a content recognition type classification problem, trying to assign each disputed essay (message generated by an unknown source) to one of the authors characterized by their respective essays. Mutual information is an absolute measure of information common to both sources whose relatedness is to be quantified. It can be transformed into a bounded distance through normalization. For content recognition (CR), we normalize by the maximum possible mutual information $I(S\ ;S\ )$ the two sources $S\ , S$ can share, which corresponds to the minimum of their respective entropy rates,

$$d_{CR}(S\ ,S\ ) = 1 - \frac{I(S\ ;S\ )}{\min(H(S\ ),H(S\ ))} \leq 1.$$

The defined distance measure can be reformulated in terms of entropy rates. To determine the entropy rate of a source from a message it generated, we make use of Shannon's compression theorem. In terms of compression the distance measure becomes

$$d_{CR} \approx \frac{|c\quad (s\ ,s\ )| - |c\quad (s\ )|}{|c\quad (s\ )|}.$$

The sequence s with larger compressed size $|c\quad (s\ )|$, where $|.|$ denotes the size in bits or symbols, is used as training for the compressor. Thus, the distance can be regarded as the ratio of the compressed size with training to the compressed size without training. For more details, please refer to [1].

The Federalist Papers were obtained from the website of the Gutenberg Project. As content, several concatenated essays by all three authors were used: Madison=(41,…,48), Hamilton=(59,…,61,65,…,69), Jay=(2,…,5,64). From all the different kinds of investigated compression algorithms Prediction by Partial Matching (PPM) performed best for content recognition of linguistic data. A PPM compressor was used to generate the distances presented in Table 1. The term "best" marks the author who most likely wrote the essay. The % value is the relative difference between the distance to the current author and to the best matching author. It is a good indicator for the reliability of the assumed attribution, e.g., essay No. 57 might also have been written by Hamilton judging by the small relative distance.

Our results coincide with the findings of others [2] based on word

Handbook of Elliptic and Hyperelliptic Curve Cryptography,

by Henri Cohen, Gerhard Frey, Roberto M. Avanzi, Christophe

1. The matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ are similar, since $P^{-1}AP = B$ with $= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ and $P^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. However, $AB = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ and $BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ are not similar, since clearly $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is similar only to itself.

2. *T* . If, for a given complex matrix M, there exists a unitary matrix U such that $U^{-1}MU = \Lambda$, where Λ is a diagonal matrix, then M is normal.

   P . From $U^{-1}MU = \Lambda$, we have $\Lambda^* = \Lambda^H = (U^{-1}MU)^H = U^H M^H (U^{-1})^H = U^{-1}M^H U$. Now $\Lambda\Lambda^* = \Lambda^*\Lambda$, because if Λ is the diagonal matrix with $\lambda_1, \lambda_2, \ldots, \lambda$ as its diagonal elements, then Λ^* is the diagonal matrix with $\lambda_1^*, \lambda_2^*, \ldots, \lambda^*$ as its diagonal elements, and both $\Lambda\Lambda^*$ and $\Lambda^*\Lambda$ are diagonal matrices with $|\lambda_1|^2, |\lambda_2|^2, \ldots, |\lambda|^2$ as their diagonal elements. Now $\Lambda\Lambda^* = (U^{-1}MU)(U^{-1}M^H U) = U^{-1}(MM^H)U$, $\Lambda^*\Lambda = (U^{-1}M^H U)(U^{-1}MU) = U^{-1}(M^H M)U$, and since these are equal, $MM^H = M^H M$, so M is normal.

3. “If N_1 and N_2 are normal $\times$ 9 0 0 9 9Tw(n)Tj/F8 1 Tf0.7552 0 TD5 TD0 Tc(N)Tj6Tw475.847 553.294 Tmf0.755244 TD0 9 169.3719 503.999

(ITW'2006)

of the fields of information theory (including application to biological sciences), communication networks, signal, image and video processing, systems and control, learning and statistical inference.

Submission dead...

Dept. of Electrical Engineering

Princeton University