PROC. INT. CONG. OF MATH. – 2018
Rio de Janeiro, Vol. 1 (211–224)

FROM BLACKBOARD TO BEDSIDE – GAUß PRIZE LECTURE

DAVID DONOHO

Abstract

(I) *Aide Mémoire.* We briefly review the contents of the 2018 Gauss Lecture ;
(II) *Esprit de l'Escalier.* Some mathematical work that ought to have been mentioned.
(II) *Meta Remarks.* Comments about the ICM and its awards.

1 *Aide Mémoire*

The 2018 Gauss Lecture had two parts.

Part I. Here, I review a June 2017 Congressional Briefing in Washington, organized by the American Mathematical Society and the Mathematical Sciences Research Institute. That briefing has been described at length in an article published in the January 2018 Donoho [2018]. The interested reader might consult that article before proceeding.

The Congressional Briefing took place in an atmosphere of uncertainty for US Federal policy, where there was some possibility that funding for US Mathematics research might see noticeable declines under the budgets soon to be put forward by the new administration and congressional majority.

David Eisenbud, director of MSRI, had in mind that I review for an audience of Congresspersons, staff and science officials, a recent transition from research to practice, showcasing the contribution that Mathematical research can make to issues of national concern.

I told the audience on Capitol Hill how the US Food and Drug Administration had very recently (Spring 2017) given bioequivalence certification to a new generation of Magnetic Resonance Imaging (MRI) scanners that the major manufacturers[1] were each bringing to market. Depending on the application, the new generation of scanners can get equivalent images in much less time – in some cases $1/10$ the time. Applications I mentioned included:

- Imaging fidgety children in much less time, without sedation and without ionizing radiation[2].

[1] There are 3 major manufacturers of MRI scanners for clinical use: GE, Philips, and Siemens. Siemens and GE obtained FDA bioequivalence certification in Spring 2017. Philips in Spring 2018.

[2] I showed examples developed by radiologist Shreyas Vasanawala (MD) Lucille Packard Children's Hospital (LCPH, Stanford), in collaboration with Michael Lustig (PhD) (EECS Prof. UC Berkeley) and John Pauly (Elec Eng Prof. Stanford).

- Commercializing new diagnostics, for example dynamic images (movies) of the beating heart.[3]

- Enabling new clinical procedures - for example MRI guided biopsy[4] - and advanced surgical interventions[5].

The individual examples that I showed really only scratch the surface; the industrial developers of the new technology envision it spreading, over time, throughout the whole range of MRI applications[6]. MRI imaging is today a large field: there are 40 million MRI scans annually in the US alone, and 80 million globally. We can envision that as the world grows more wealthy and developed there will eventually be billions of MRI scans yearly, all of which are made much less expensive and much more convenient by these speedups. While many of these scans might indeed be looking at knees, hips, and ankles of aging wealthy Americans and Europeans, many others will be peering deep into the brains of needy younger patients, as MRI imaging is also essential for modern Neurology and Neurosurgery.

The interested viewer might watch the documentary *The English Surgeon* about the UK Neurosurgeon Henry Marsh's travels to perform surgery in less advanced countries. In a wrenching pivotal scene, Marsh diagnoses what will inevitably become permanent blindness in a young girl who did not receive an MRI scan when it would still have been helpful. Marsh remarks that the young girl will live forever after with the consequences of lack of availability of prompt MRI scanning in her country. As MRI becomes more available and cheaper globally, we can expect such human tragedies to diminish in frequency.

In marketing the new generation of scanners, the manufacturers all say they are using *Compressed Sensing*, a term of art [7] introduced by mathematicians a little over 10 years ago.

---

[3]I showed movies illustrating commercial bioequivalence at 15X speedup, provided by Edgar Muller of Siemens Healthineers; and movies from the team led by Dr. Shreyas Vasanawala at Stanford LCPH showing the beating heart overlayed with measured flow fields visualizing flow through the arteries and veins supplying the beating heart. Data acquisition for this last example was sped up from more than an hour to under 10 minutes.

[4]I mentioned MRI-guided prostate biopsy, which might have saved my own father from aggressive prostate cancer, had it been available ten years ago. Dr. Faina Shtern of ADMeTech and her collaborators have shown that MRI-guided biopsies can have a 5x better success rate in cancer detection

[5]I mentioned MRI-guided surgery, giving examples from my own son's work as a Neurosurgeon at LA County Hopsital in Los Angeles, where a procedure conducted today with pre-existing technology takes an hour and a half of MRI scanning, during which time more than a half dozen skilled surgeons, anesthesiologists and nurses are sitting idle.

[6]For this, I quoted Jason Polzin of GE Research and Edgar Muller of Siemens.

[7]Siemens' corporate website used the tagline 'Compressed Sensing: Beyond Speed' in Spring 2017. Philips published an article in their 'FieldStrength' MRI publication, January 2018, with the title 'Faster MRI throughout the whole body with Compressed SENSE'; GE published an article in GESignaPulse magazine Autumn 2016 with the title 'HyperSENSE gives shorter scan times without compromising image quality'. The second sentence of the GE article reads '*Compressed Sensing (CS) is a technique similar to image compression ... [that] .. acquires less k-space data and uses a special reconstruction algorithm to recover the missing information without an appreciable impact on image quality*' and later says '*GE Healthcare has developed a CS technology called HyperSense which obtains randomly undersampled raw data thereby reducing the overall scan time with compromising overall image quality.*'.

In a series of papers submitted starting in 2004[8], mathematicians put forward theorems showing (under a variety of assumptions) that one does not need to actually make one million measurements to reconstruct a million-pixel image.

Such theorems inspired efforts by many engineers, doctors and physicists to develop new algorithms in some way related to the mathematics, to run experiments testing the new algorithms, and to present their work at major conferences. For an early example, see Lustig, Donoho, and Pauly [2007]. For a recent review article reflecting the evolving state-of-the-art, see Feng, Benkert, Block, Sodickson, Otazo, and Chandarana [2017]. In some individual research hospitals, such as Lucille Packard Children's Hospital (Vasanawala and Lustig [2011]), faster MRI scanning entered regular use within a few years. Rolling out a technology for widespread us in the US medical market entails regulatory scrutiny from the US Food and Drug Administration. In this present case, the transition to the marketplace has happened in not much more than 10 years.

In my talk, I made the point that in this instance, the Federal research funding system worked exactly as lawmakers had intended. I specifically mentioned Federal funding of individual investigators that supported the many mathematicians and engineers who worked to penetrate the mysteries of randomly-undersampled measurements; I pointed to Michael Lustig, who in 2004 was a Federally-supported graduate student in EE at Stanford but is today a Federally-supported EECS Professor at UC Berkeley; and to Emmanuel Candès, Justin Romberg, and Terence Tao, who in 2004 collaborated on some of the early and inspiring results, while Candès had an office on the UCLA campus at IPAM, one of the NSF-funded Mathematics Institutes, during an IPAM long program on multiscale geometric analysis.

I also tried to make the point that 'serious' mathematics was involved. I formulated the intellectual question in the following terms: *Consider a million-dimensional vector which has 10,000 entries scattered at random among many zeros.*[9] *Can one reconstruct this vector from only 100,000 measurements rather than 1,000,000?* If the measurements are in an appropriate sense random, and we recinstruct by ℓ_1-minimization of the reconstructed vector, this is the same thing as asking if a random 900,000-dimensional linear subspace typically intersects a certain simplicial cone transversely.

An instructive but very elementary version of the same problem is to ask if a random 1-dimensional linear subspace of R^N has nontrivial intersection with a simplicial cone. In the $N = 3$-D dimensional case, this amounts to asking if a point randomly distributed on the surface of a sphere lies inside a given spherical triangle. This of course is the ratio of the area of the triangle to the surface area of the sphere.

Gauss himself gave us the formula that solves this elementary problem; he showed that the area of a spherical triangle is given in terms of the sum of the 3 interior angles, which will not be 180 degrees owing to the curvature of the sphere[10].

Today there is a field called *combinatorial geometry* which develops generalizations of Gauss' formula, involving interior and exterior angles of polyhedral cones, Schneider and Weil [2008], and which allows us to solve the more ambitious problem involving

[8]Examples include: Candès, Romberg, and Tao [2006a,b] and Donoho [2006a,b]

[9]We don't know the positions of the nonzeros, nor do we know the values of the nonzeros.

[10]In my lecture I pointed out that yes, I really had invoked Gauss' theorem in the Congressional Briefing, even illustrating it with computer graphics.

$1,000,000$-dimensional simplicial cones, mentioned above. Such mathematical results allow us to see that, for the combination of parameters outlined above, the probability is effectively 1.0 that the subspace and cone intersect transversely! See for example: Donoho and Tanner [2005], Donoho [2006c], and Donoho and Tanner [2009] In short, with overwhelming probability, we can recover such a vector from 10x fewer measurements than one might have naïvely supposed.

Part II. Here we discuss issues raised by Part I. I mentioned two questions

- *Why did the transition from Theory to Practice happen so quickly?*
  Indeed, it seems that a dozen years is very fast for such a transition, especially in a regulated industry (Medicine).

- *What areas of mathematics contributed?*
  Indeed, attendees of the ICM2018 would no doubt be more interested in an understanding of which fields of mathematics contributed to speeding up MRI, than in knowing about the specific clinical applications (Pediatrics, Cardiology, Prostate Cancer Diagnosis) mentioned in Part I. Although in Part I, I was able to mention combinatorial geometry, in reality *many* subfields of the mathematical sciences played important roles.

1.0.1 Why did the transition from Theory to Practice happen so quickly? In my lecture, I pointed to several ingredients that combined to cause a rapid transition from Theory to Practice.

Everyone wanted MRI speedups. Everyone wants MRIs to go faster. Sitting confined in a tube is no fun, even when one is healthy and strong[11]

Experimental efforts were already succeeding. In my talk I pointed to many prior experimental works exploiting undersampling to in some way speed up MRI[12] In the lecture I called out the 2001 Leiden PhD Thesis of Frank Wajer, see Wajer [2001] who used random undersampling and nonlinear reconstruction to speed up MRI[13].

The MRI industry has assembled a massive collection of talent . The industry is servicing 80 million MRI scans annually. It has assembled a talent pool of thousands of MR physicists, electrical engineers, and computer scientists; this pool was available to be pressed into service at any sign of exciting new opportunities.

[11] I personally have been aware of the problem of speeding up MRI's since the early 1990's. I directed a *stage* by two students of *École Polytechnique* in the summer of 1995, hoping we could speed up MRI's by adaptive sampling. Stéphane Mallat was the examiner. And if even *I* knew of this problem you can be sure that this was a known and important problem in the MRI research community.

[12] In the cognate field of MR spectroscopy, Jeffrey Hoch, a Physical Chemist now at University of Connecticut Health Sciences Center, and co-authors showed already in the early/mid 1990's that random undersampling combined with nonlinear reconstruction could lead to notable speedups in MR spectroscopy.

[13] Thanks to Michael Lustig for this reference.

What, then, was preventing progress?

In lecture, I explained that experimental evidence of successful undersampling certainly existed, but many people didn't lend it much weight.

In general, purely computational results face an uphill battle to acceptance, partially because *computational results have historically not been not fully reproducible*. If someone told you of a computational result which required novel algorithms and data, it might involve a great deal of work for you to try the idea yourself. Under such circumstances you might simply find it easier to imagine that a reported success was due to a mistake in coding or in data preparation; and you might therefore view the result very warily.

In contrast, if someone tells you that there are mathematical theorems spelling out precise conditions under which certain effects can be observed, there is much less room for corrosive doubt.

On top of that, I mentioned in Lecture another factor: Terence Tao was awarded a Fields Medal in 2006, close to the beginning of the 'Compressed Sensing phenomenon'.

While in general the Fields Medal is not instrumental to any larger purpose, in this specific instance, at least, the Fields medal must have played an enabling role, lending charisma and intellectual authority to those striving for change in MRI.

It seems that mathematics broke a conceptual logjam, after which the MRI community really mobilized its talent to implement and study compressed sensing.

In lecture I mentioned that Michael Lustig - initially a graduate student at Stanford, now a Professor of EECS at UCBerkeley - in addition to being the first to develop and present many very influential compressed sensing applications to MRI, also developed a compelling approach to explaining Compressed Sensing for MRI researchers, in which he streamlined and recast some of the mathematician's assumptions and language in ways that MRI researchers could better understand, and developed a successful intuitive and heuristic framework that non-mathematicians found persuasive. Lustig himself points to decisive influences by many other MRI researchers[14].

1.0.2 What areas of Mathematics Contributed? I next pointed to several subfields of mathematics that had contributed substantially to the formulation and development of compressed sensing.

Computational Harmonic Analysis. In the 1980's and 1990's a variety of new tools for digital image processing - wavelet transforms and associated time-frequency transforms - were developed by applied mathematicians including Ingrid Daubechies and Yves Meyer, as well as Raphy Coifman, Stéphane Mallat, Michael Unser, and Martin Vetterli. The wavelet transforms developed by these researchers, when applied to digital signals and images, revealed a previously hidden sparsity – images were sparse in the wavelet domain; this sparsity can be helpful for data compression, and ultimately explains the adoption of wavelets as part of JPEG-2000 and in MPEG-4 standards.

[14] Lustig wrote me: "I would emphasize the contributions of Tobias Block, Uecker, and Frahm [2007] and Ricardo Otazo, Kim, Axel, and Sodickson [2010]. These guys have taken the clinical translation of Compressed Sensing orders of magnitude forward." He also mentioned key contributions of Zhi-Pei Lian of UIUC, Joshua Trzasko of Mayo Clinic, and of Alexey Samsonov and Julia Velikina of UW Madison

A more ambitious 'sparsity-promoting' program arose from this: the thesis that all natural signals and images can be sparsely represented in *some* transform domain – we simply construct the appropriate transform. Attempts to realize this sparsity-promoting program led to Ridgelets, and from their to curvelets and shearlets and other 'X'-lets[15].

The idea of sparsity as a property to be sought after and enhanced led to the identification of ℓ_p balls with $0 < p < 1$ as mathematical models of sparsity[16]. While generations of mathematics analysis students have been taught to consider $p \geq 1$ as natural, the demands of real world signal processing made it now seem natural to consider cases which would have previously seemed pathological[17].

Combinatorial Geometry. We earlier mentioned the role that Combinatorial Geometry can play in justifying compressed sensing. Key concepts in combinatorial geometry can be learned from Branko Grunbaum's 50-year old classic *Convex Polytopes*, Grünbaum [1967] or Guenther Ziegler's more recent *Lectures on Polytopes*, Ziegler [1995]. In particular, these books define the face lattices of convex polytopes and properties of those lattices, such as k-neighborliness.

Combinatorial questions around the face lattices of random point sets (how many points are on the convex hull of a random point set? How many edges? How many faces?) were already discussed in the 1960's by Renyi and Sulanke and by Bradley Efron.

In the early 1990's, Anatoly Vershik and Pyotr Sporyshev, contemporaneous with Fernando Affentranger and Rolf Schneider, published exact formulas for the expected face numbers of various dimensions of random polytopes.

By duality, these quantities can be related to the so-called Grassmann angles defined by Branko Grunbaum, which measure the probability that linear subspaces transversely intersect a convex cone.

These diverse ingredients set the stage for the high-dimensional geometric calculations needed to determine the success of Compressed Sensing. Knowing them, and knowing that compressed sensing can be posed in terms of questions about Grassmann angles, one can calculate the probability that was mentioned in the Congressional Briefing: namely, the chance that a 900,000-dimensional random linear subspace is transverse to a certain cone with a 10,000-dimensional field of apices.

Those are merely two of the many fields that contributed to our knowledge of compressed sensing. In my lecture I mentioned four others. I can only very briefly allude to some contributions.

[15] See for example the article documenting in part my 2002 ICM Plenary address 'Multiscale Geometric Analysis' at Donoho [2002].

[16] See for example the article documenting Emmanuel Candès' 2014 ICM Plenary address 'Mathematics of Sparsity' at Candès [2014].

[17] We mention the pioneers of $p < 1$, such as Jaak Peetre and Vladimir Popov, who realized its significance in nonlinear approximation theory long ago, and in more recent times, Ron DeVore and Vladimir Temlyakov.

High-dimensional Analysis. Geometric functional analysis has played a key role in Compressed Sensing theory. Boris Kashin, in his solution of Kolmogorov's problem on the ϵ-entropy of Sobolev spaces, calculated the Gel'fand n-widths of ℓ_1^N balls and uncovered surprising structure, showing that they are controlled as $O(\sqrt{\log(N)/n})$. Today, by work in information-based complexity, we know that this implies that compressed sensing works for objects which are sparse in the sense that they belong to $\ell_{1,N}$ balls. More recently, Holger Rauhut and co-authors had pushed the calculation to encompass $\ell_{p,N}$, with $0 < p < 1$, which is really the case of interest for CS, Foucart, Pajor, Rauhut, and Ullrich [2010].

Over the last 50 years, random matrix theory has grown up into a powerful tool, and it too had its impact on compressed sensing. Concentration of measure results for eigenvalues were used by Candés and Tao to establish the important Restricted Isometry Property in their analyses of CS.

Optimization. At the center of CS is the solution of high dimensional optimization problems which are in a sense non-smooth and nonregular. And without algorithms there are *no* applications.

Pioneers in efficient high dimensional algorithms include Arkady Nemirovksii and Yuri Nesterov, whose broad influence is nearly staggering.

At the time that CS was first being developed, Ingrid Daubechies, Michel DeFrise and Christine DeMol had just written an influential paper in CPAM about provably solving ℓ_1 minimization problems in high dimension by iterative soft thresholding, Daubechies, Defrise, and De Mol [2004]. Today, the algorithm FISTA by Amir Beck and Marc Teboulle became very popular for these problems, see Beck and Teboulle [2009].

Mathematical Statistics. With Arian Maleki and Andrea Montanari, I published a paper in 2009 that exposed an unsuspected connection between Mathematical Statistics and Compressed Sensing, Donoho, Maleki, and Montanari [2009].

It showed that the phase transition boundary calculated by combinatorial geometry was identical to the minimax mean squared error in a signal denoising problem, an unrelated problem that Iain Johnstone and I had worked on 15 years earlier, and where our calculations used unrelated methods. Later, Johnstone and Montanari and I put forward a reason for this connection and validated it empirically in many other compressed sensing-like problems, Donoho, Johnstone, and Montanari [2013]. It is simply spooky to come across an echo of an earlier result in a long-ago period of one's life – with no known reason for the echo.

Information Theory. The full story about the connection between denoising and compressed sensing has since been developed to a very fine point in modern information theory. We now understand that certain 'master conjectures' in spin glass theory would imply that connection as a special case, and important new techniques have developed to prove the conjectures in important situations. Here some of the relevant work has been done by Andrea Montanari and collaborators, other work by Florent Krzkala and Lenka Zdeborova and co-authors.

An amazing fact about the above list is: *all this important work is very recent.* One might have thought to study compressed sensing 30 years ago, but *the mathematics to study it simply did not exist*!

2 *Esprit de L'Escalier*

Denis Diderot was a renowned *encyclopediste* of the French enlightenment. Most mathematicians would know of him via E.T. Bell's *Men of Mathematics*, which recounts an amusing story of Diderot supposedly debating Euler at Catherine the Great's court[18].

Diderot recorded for posterity the notion of *Esprit de L'Escalier* - in English, literally 'staircase wit' - which refers to the experience of finding the right words to say only after leaving an important meeting or event.

Yves Meyer, the 2010 IMU Gauss Prize winner, once told me about a pilgrimage he made to Auxerre, the hometown of Fourier. Thinking about this - after the Gauss lecture - it occurred to me that I should really have mentioned some of my mathematical heroes. Here is my little attempt to make amends.

2.1 Constantin Carathéodory. Books have been written about Carathéodory's life and times, Georgiadou [2004], but still not enough of our young generation know of him. He made several lasting contributions, including what is today known as dynamic programming in optimal control. Relevant to our story, Carathéodory discovered what would today be called the *neighborliness of the moment curve* and of the trigonometric moment curve. Such neighborliness properties can be rephrased as follows (see also Donoho and Tanner [2005]: If a function $f(t)$ is a nonnegative combination of at most s complex sinusoids, $f(t) = \sum_{k=1}^{s} c_k \exp(i\omega_k t)$, it can be uniquely recovered from measurements of any $2s+1$ points $f(t_i)$, $i = 1,\dots,2s+1$, where $t_i \neq t_j$, $i \neq j$[19]. Also, if a function $g(t)$ is a nonnegative combination of at most s decaying exponentials, $g(t) = \sum_{k=1}^{s} c_k \exp(-\lambda_k t)$, it can be uniquely recovered from measurements of any $2s+1$ points $g(t_i)$, $i = 1,\dots,2s+1$, where $t_i \neq t_j$, $i \neq j$. In our telling, such results contain some of the key features of compressed sensing: for example, a sparsity measure – in this case, s – and a sparsity-dependent sampling limit – in this case, $2s$.

2.2 Arne Beurling. As a young assistant professor, I purchased the collected works of Arne Beurling and, for a time, read them daily, see Beurling [1989b]. These works contain many beautiful results, even some important unpublished ones, but I found especially inspiring his work on interpolation, balayage, and minimal extrapolation. Here is an example most relevant to the Gauss Lecture. Beurling considered the problem: given a finite measure $\nu(dt)$ on the real line, suppose we observe the the Fourier transform $y(\omega) = \hat{\nu}(\omega) = \int \exp(i\omega t)\nu(dt)$ - but only over frequencies $|\omega| \leq \Omega$. He defined

[18] This story was engagingly related to my Honors Multivariable Calculus class many years ago, by our professor, Charles Fefferman. Unfortunately, E.T. Bell's story is now said to have been a fabrication of Augustus DeMorgan.

[19] We don't know the frequencies ω_k, nor do we know the values c_k of the coefficients.

Beurling minimal extrapolation, Beurling [1989a], i

$$\nu^{\#} = argmin_{\eta} \int |\eta|(dt) : \text{ subject to: } \{\hat{\eta}(\omega) = \hat{\nu}(\omega),\ |\omega| \le \Omega\}.$$

From our current perspective we can see that he is doing a form of L^1 minimization[20]. I don't recall if Beurling wrote this down or not, but a key point about his notion of minimal extrapolation is that (with the right interpretations) it perfectly recovers sparse probability measures: *if ν is a finitely supported nonnegative measure and yet the argmin runs over all nonnegative measures, then the solution perfectly recovers ν: $\nu^{\#} = \nu$.* In short, we have perfectly recovered a discrete measure from only low frequency information. This is a strengthening of Carathéodory's result above, since it gives an algorithm. Of course, Beurling did much more than just define important concepts; I learned a great deal of very specific ideas both in soft and hard analysis from his papers.

2.3 Benjamin Logan. Ben Logan worked at Bell Labs during the glory days, and teamed up with Larry Shepp for many important and well-known results, for example on tomography during the immediate explosion of computer-assisted tomography during the 1970's, and on counting of Young Tableaux. But Ben Logan had an independent career developing interesting and still very surprising results in nonlinear signal processing, long before the rest of the world understood that signal processing would soon be a ubiquitous latent presence in human life, for the rest of time[21].

Some of Logan's I think very interesting results concerned signals designed for recovery after arbitrary clipping, Logan Jr. [1984a,b]. He designed a signal processing scheme that allowed a bandlimited signal (i.e. an entire function of exponential type) to be distorted by the signum nonlinearity $sgn()$ and yet still be recovered. He even designed a practical recovery scheme and published it in the Bell System Technical Journal.

A result from Ben's Ph.D. thesis Logan Jr. [1965] bears special mention. Suppose that g is a function zero outside an interval $(-T/2, T/2)$. Suppose that f is a bandlimited signal, i.e. $f(t) = \int_{-\Omega/2}^{\Omega/2} \exp\{i\omega t\} \hat{f}(\omega) d\omega$. We are interested in the L^1 distance between f and g. Ben's thesis showed that, if $T\Omega < 1/2$, and $f \neq 0$,

$$\|f - g\|_{L_1(\mathbf{R})} > \|g\|_{L_1(\mathbf{R})}.$$

The simplicity of the statement and even the proof are deceptive. No such statement could possibly be true for the L^2 norm. This expresses in a simple but profound way the nonlinearity and nonsmoothness of the L^1 norm, and the interaction with sparsity, the same phenomenon we exploit in compressed sensing. The main difference in approach is simply that in Logan's earlier result, sparsity was replaced by the more restrictive condition of narrowness of support.

[20] AKA total variation minimization, *pace* Stanley Osher (2014 IMU Gauss Medalist).

[21] Logan had interesting pursuits outside mathematics. He was an outstanding bluegrass fiddler (violinist) called 'Tex' Logan in the bluegrass music community, and served as a studio musician for important country music stars.

3 Meta Remarks

3.1 The ICM. With the 2018 ICM, I have now attended 4 ICM's spanning 32 years (1986, 1994, 2002, 2018). During that period, this event has changed utterly. The 1986 and 1994 ICM's were held on university campuses and so had the flavor of traditional academic meetings. The opening ceremonies were straightforward affairs. The 2018 Rio ICM was held in a convention center away from any university, and the opening ceremony was polished in a way that would have been inconceivable in 1986[22]. The ceremony had extensive professional video presentations of the Fields Medalists and even entertainment with Samba Dancers from Brazil. The new persona is more fitting to today's situation. The Fields Medals are now a global brand.

What has stayed constant, in my view, is the quality of the ICM presentations. I remember very clearly some of the talks from the 1986 Congress; they were given by deservedly prominent figures and for me were era-defining. The talks I saw at the 2018 ICM met the same standard.

3.2 The Gauss Medal, and the Gauss Prize. At the 1986 ICM, I remember hallway discussions about the fact that Applied Mathematics was not traditionally recognized at events like the ICM, for example no aApplied Mathematician had ever been awarded a Fields Medal. The question young people were asking at the time was whether a young person should go into applications when there was no recognition. Apparently young people are very attuned to cues provided by who gets recognized.

This supposed situation of Applied Mathematics was partially addressed in 1994, with the award of the Fields Medal to Pierre-Louis Lions. It has now been directly addressed with the creation of the Gauss prize of the IMU, that directly recognizes impactful work in applications.

During the Rio Congress, I learned that Martin Grötschel had the idea and took the initiative to found the Gauss Prize, using the support of the German Mathematical Society, DMV. In a sense, the previous situation – of the supposed lack of importance of Applied Mathematics – was not some inherent fact about mathematics. It was simply due to the fact that no-one previously had Grotschel's organizational skills, energy, and clear vision. Take note.

The Gauss Prize medal has a story relevant to the 2018 Lecture. The artist Jan Arnold actually undersampled the image of Gauss! At the IMU website we read

> *"Dissolved into a linear pattern, the Gauss effigy is incomplete. It is the viewer's eye which completes the barcode of lines and transforms it into the portrait of Gauss."*

This really caught my attention, because the same scheme of undersampled vertical lines is *precisely* the k-space sampling pattern that is often used with compressed sensing[23]. Here art anticipates science!

[22] The essential book Curbera [2009], has many interesting photos. They show that earlier ICMs had much simpler opening ceremonies than the 2018 iteration.

[23] Monajemi, H. and Donoho, D. L., 2017. Sparsity/Undersampling Tradeoffs in Anisotropic Undersampling, with Applications in MR Imaging/Spectroscopy, arXiv 1702.03062.

3.3 My Collaborators. I have had wonderful collaborators. In the context of the ICM, note that my two close collaborators Iain Johnstone and Emmanuel Candès, have each given Plenary Addresses at ICM (2006 and 2014, respectively). They are each very serious and powerful scientists in their own right. I am very lucky to have been able to work with them.

This is only the start. At the 2018 ICM itself, my co-author Raphy Coifman (Yale) gave a plenary lecture and my co-author Andrea Montanari (Stanford) gave an invited talk, as did my former PhD student Noureddine El Karoui (UC Berkeley). It is my great good fortune to have been able to interact with, and learn from, such leading figures[24]. The broad range of work these different scientists presented was very striking.

3.4 The Long Tail. The professional writers working with the ICM to prepare profiles of the Fields, Nevanlinna, Chern and Gauss awardees are inclined to portray these otherwise fallible individuals as heroes. But in the case of the Gauss Award this idea is poorly conceived.

An award for applications, one that specifically recognizes real-world impact, ought to also recognize that applications only arise from a 'long tail' of mathematical scientists who, while not professional mathematicians themselves, are very mathematically minded and talented, while being much more applications facing than any mathematician would be. Such a long tail needs to be inspired and to be creative and driven in its own right, or no impact will result.

In this case, clearly Michael Lustig and John Pauly have distinguished themselves as electrical engineers with exceptional feel for mathematics and its proper deployment into MRI, and exceptional capacity for being inspired by mathematics. Lustig mentions several other MRI researcher colleagues who did major work translating Compressed Sensing into practice, including Tobias Block, Zhi-Pei Lian, Ricardo Otazo, Alexey Samsonov, Daniel Sodickson, Joshua Trzasko, and Julia Velikina. Doubtless there are many others that could also be mentioned.

In my own case, I would like to mention particularly two other mathematical scientists who have always been impressed with their ability to take mathematical inspiration and deliver practical results: Jean-Luc Starck (Centre Européen d'Astronomie) and Michael Elad (Technion). Their successful empirical results have consistently, over decades, gone far beyond what I would have dared to suggest based on mathematical analysis alone. In some sense they have always been followers of Wigner's 'unreasonable effectiveness of mathematics' while in many cases I have been too pessimistic.

[24] I mentioned earlier the names of several other co-authors: Benjamin F. Logan, Jr. (deceased, ATT Labs), Michael Lustig (UC Berkeley), Philip Stark (UC Berkeley), Jared Tanner (Oxford). I should also have mentioned co-authors whose work was deeply connected to this lecture but not easy to work into the limited space here: Xiaoming Huo (Georgia Tech), Jeffrey Hoch (U. Conn), Gerard Kerkyacharyan, (Université de Paris), Brenda MacGibbon, Alan Stern (Rowland Institute), Dominique Picard (Université de Paris).

References

Amir Beck and Marc Teboulle (2009). "A fast iterative shrinkage-thresholding algorithm for linear inverse problems". *SIAM J. Imaging Sci.* 2.1, pp. 183–202. MR: 2486527 (cit. on p. 217).

Arne Beurling (1989a). "Interpolation for an interval in $\mathbb{R}^1$". In: *The collected works of Arne Beurling. Vol. 2*. Contemporary Mathematicians. Harmonic analysis, Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer. Birkhäuser Boston, Inc., Boston, MA, pp. 351–365. MR: 1057613 (cit. on p. 219).

– (1989b). *The collected works of Arne Beurling. Vol. 2*. Contemporary Mathematicians. Harmonic analysis, Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer. Birkhäuser Boston, Inc., Boston, MA. MR: 1057613 (cit. on p. 218).

Kai Tobias Block, Martin Uecker, and Jens Frahm (2007). "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint". *Magnetic Resonance in Medicine* 57.6, pp. 1086–1098 (cit. on p. 215).

Emmanuel J. Candès (2014). "Mathematics of sparsity (and a few other things)". In: *Proceedings of the International Congress of Mathematicians, Seoul 2014*. Ed. by Sun Young Jang, Young Rock Kim, Dae-Woong Lee, and Ikkwon Yie. Vol. 1. Kyung Moon SA, pp. 235–258 (cit. on p. 216).

Emmanuel J. Candès, Justin K. Romberg, and Terence Tao (2006a). "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information". *IEEE Trans. Inform. Theory* 52.2, pp. 489–509. MR: 2236170 (cit. on p. 213).

– (2006b). "Stable signal recovery from incomplete and inaccurate measurements". *Comm. Pure Appl. Math.* 59.8, pp. 1207–1223. MR: 2230846 (cit. on p. 213).

Guillermo P. Curbera (2009). *Mathematicians of the world, unite!* The International Congress of Mathematicians—a human endeavor, With a foreword by Lennart Carleson. A K Peters, Ltd., Wellesley, MA, pp. xviii+326. MR: 2499757 (cit. on p. 220).

Ingrid Daubechies, Michel Defrise, and Christine De Mol (2004). "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint". *Comm. Pure Appl. Math.* 57.11, pp. 1413–1457. MR: 2077704 (cit. on p. 217).

David L. Donoho (2002). "Emerging applications of geometric multiscale analysis". In: *Proceedings of the International Congress of Mathematicians, Beijing 2002*. Ed. by L. I. Daquian. Vol. 1. Higher Education Press, pp. 209–233 (cit. on p. 216).

– (2006a). "Compressed sensing". *IEEE Trans. Inform. Theory* 52.4, pp. 1289–1306. MR: 2241189 (cit. on p. 213).

– (2006b). "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution". *Comm. Pure Appl. Math.* 59.6, pp. 797–829. MR: 2217606 (cit. on p. 213).

– (2006c). "High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension". *Discrete Comput. Geom.* 35.4, pp. 617–652. MR: 2225676 (cit. on p. 214).

– (Jan. 2018). "From Blackboard to Bedside: High-dimensional Geometry is Transforming the MRI Industry – Aide Mémoire for a Congressional Briefing". *Notices of the AMS* 65 (1) (cit. on p. 211).

David L. Donoho, Iain Johnstone, and Andrea Montanari (2013). "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising". *IEEE Trans. Inform. Theory* 59.6, pp. 3396–3433. MR: 3061255 (cit. on p. 217).

David L. Donoho, Arian Maleki, and Andrea Montanari (2009). "Message-passing algorithms for compressed sensing". *Proc. Natl. Acad. Sci. USA* 106.45, pp. 18914–18919 (cit. on p. 217).

David L. Donoho and Jared Tanner (2005). "Sparse nonnegative solution of underdetermined linear equations by linear programming". *Proc. Natl. Acad. Sci. USA* 102.27, pp. 9446–9451. MR: 2168715 (cit. on pp. 214, 218).

– (2009). "Counting faces of randomly projected polytopes when the projection radically lowers dimension". *J. Amer. Math. Soc.* 22.1, pp. 1–53. MR: 2449053 (cit. on p. 214).

Li Feng, Thomas Benkert, Kai Tobias Block, Daniel K. Sodickson, Ricardo Otazo, and Hersh Chandarana (2017). "Compressed sensing for body MRI". *Journal of Magnetic Resonance Imaging* 45.4, pp. 966–987 (cit. on p. 213).

Simon Foucart, Alain Pajor, Holger Rauhut, and Tino Ullrich (2010). "The Gelfand widths of ℓ_p-balls for $0 < p \leq 1$". *J. Complexity* 26.6, pp. 629–640. MR: 2735423 (cit. on p. 217).

Maria Georgiadou (2004). *Constantin Carathéodory*. Mathematics and politics in turbulent times, With a foreword by George Papanicolaou. Springer-Verlag, Berlin, pp. xxviii+651. MR: 2033701 (cit. on p. 218).

Branko Grünbaum (1967). *Convex polytopes*. With the cooperation of Victor Klee, M. A. Perles and G. C. Shephard. Pure and Applied Mathematics, Vol. 16. Interscience Publishers John Wiley & Sons, Inc., New York, pp. xiv+456. MR: 0226496 (cit. on p. 216).

B. F. Logan Jr. (1965). "Properties of high-pass signals". PhD thesis. Columbia University, New York (cit. on p. 219).

– (1984a). "Signals designed for recovery after clipping – I. Localization of infinite products". *AT&T Bell Laboratories Technical Journal* 63 (2), pp. 261–285 (cit. on p. 219).

– (1984b). "Signals designed for recovery after clipping – II. Fourier transform theory of recovery". *AT&T Bell Laboratories Technical Journal* 63 (2), pp. 287–306 (cit. on p. 219).

Michael Lustig, David L. Donoho, and John M. Pauly (2007). "Sparse MRI: The application of compressed sensing for rapid MR imaging". *Magnetic Resonance in Medicine* 58.6, pp. 1182–1195 (cit. on p. 213).

Ricardo Otazo, Daniel Kim, Leon Axel, and Daniel K. Sodickson (2010). "Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI". *Magnetic Resonance in Medicine* 64.3, pp. 767–776 (cit. on p. 215).

Rolf Schneider and Wolfgang Weil (2008). *Stochastic and Integral Geometry* (cit. on p. 213).

Shreyas S. Vasanawala and Michael Lustig (Aug. 2011). "Advances in pediatric body MRI". *Pediatric Radiology* 41.2, p. 549 (cit. on p. 213).

Frank T. A. W. Wajer (2001). "Non-Cartesian MRI scan time reduction through sparse sampling". PhD thesis. Technische Universiteit Delft (cit. on p. 214).

Günter M. Ziegler (1995). *Lectures on Polytopes*. Vol. 152. Graduate Texts in Mathematics. Springer-Verlag New York, pp. XI+370 (cit. on p. 216).

Received 2018-08-30.

David Donoho
Department of Statistics
Stanford University