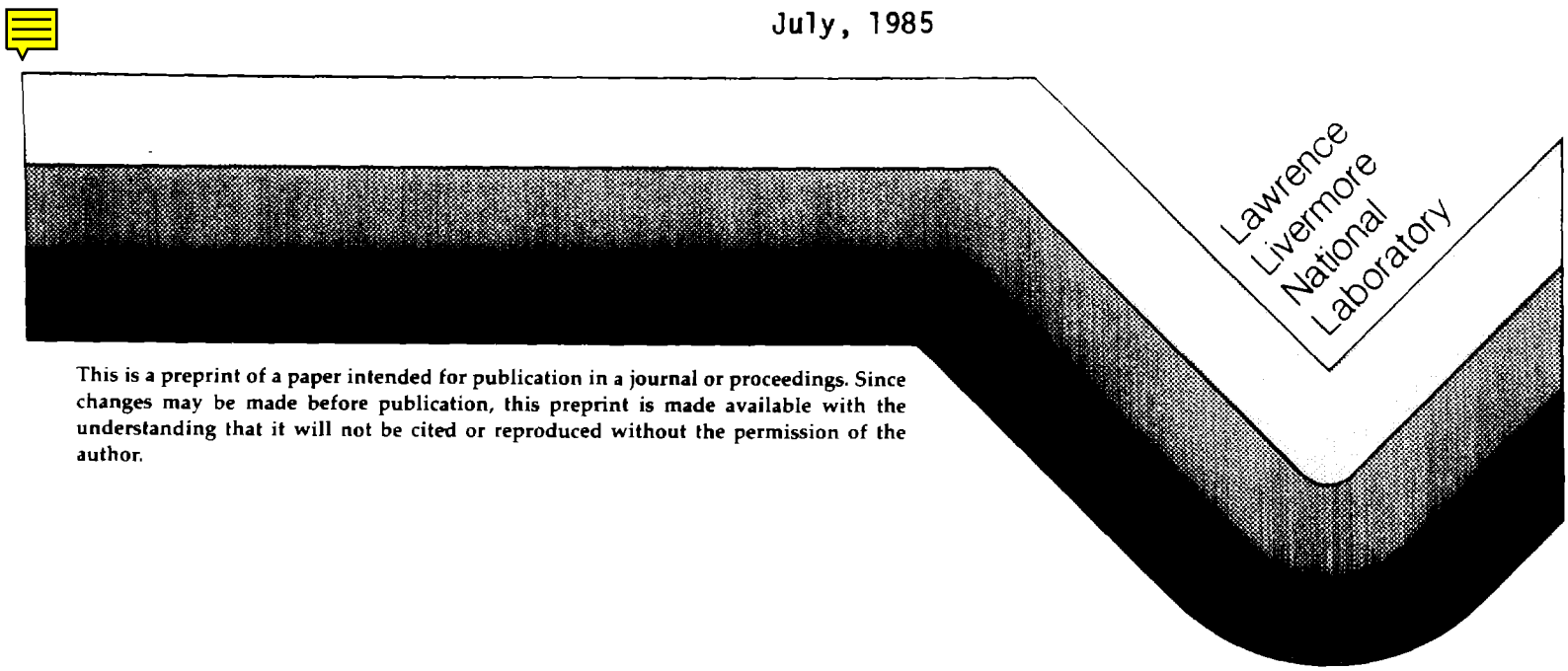
UCRL- 92414
PREPRINT

THE ULTIMATE LIMITS OF FT-IR

Tomas Hirschfeld

This paper was prepared for submittal to the
Proceedings, 1985 International Congress on
Fourier and Computerized Infrared Spectroscopy
Ottawa, Canada
June 24-28, 1985

July, 1985



The ultimate limits of Fourier transform-infrared (FT-IR)*

Tomas Hirschfeld
Chemistry & Materials Science Department, Lawrence Livermore National Laboratory
P. O. Box 808, L-322, Livermore, California 94550

Abstract

Continual improvements in FT-IR have brought us within two orders of magnitude of the theoretical limits of the technology, set by the intrinsic noise of the blackbody spectrum. Further improvements in the technology still require further optimization by step-by-step comparison between actual components performance and their theoretical limit.* Among insights from such a process, we have found a surprisingly low energy efficiency in current interferometers, substantial amounts of excess noise in the detectors and nonlinearity in their amplifiers, and an insufficient utilization of the very low aberration levels of interferometers. Beyond this, substantial improvements in spectroscopic sensitivity can be achieved by appropriate choice of the optical interaction. Here the elimination of the 100% line is a particularly desirable goal. The optics of sampling accessories also merit much attention, including throughput matching (and also throughput conservation), optimizing interaction strength, signal strength and detector noise optimization, etc. Finally, data processing can be used to substantially improve the performance capabilities of FT-IR, particularly if we are willing to depart from the conventional data presentations used in IR spectroscopy.

Recently Fourier transform infrared spectrometers have achieved signal-to-noise values in excess of one million, and have thus become the second highest signal-to-noise instrument in the analytical laboratory, right after the analytical balance. Under these circumstances, it seems reasonable to ask ourselves: do we really need any further improvement? and, is this really possible?

To answer the first question, many of the techniques that are at the leading edge of Fourier transform infrared spectroscopy today (such as trace analysis, microanalysis, thin film analysis, GC-IR, infrared microprobing, and the 1001 advanced computer tricks of infrared spectroscopy) have an unlimited demand for signal-to-noise; and they often demand it under difficult conditions, at low throughput, short measurement times, restricted area or transmission samples, etc. Other techniques, also popular, such as absorbance substract, diffuse reflectance, and trace analysis, are limited, on one hand, by instrumental accuracy and stability; and, on the other hand, by our ability to trade in excess signal-to-noise ratio for mathematical corrections to known error sources.

Given this open-ended need, can we meet it? To study this, we must consider the capabilities of an ultimate instrument--an instrumental analog of that so useful fiction, the ideal gas. This will deliver at its output the full signal-to-noise available, that is, the shot noise limited SNR of the blackbody source. Given reasonable assumptions, such as 1 mm^2 throughput and 1% efficiency (!), Fig. 1 shows the capabilities of such an instrument for a 1.0 sec measurement at 2 cm^{-1} resolution. The performances shown are well in excess of anything achieved in FT-IR today, and tell us that there is substantial room between current instrument performance and its theoretical limits. Even if we bring into the analysis the real world limitations of

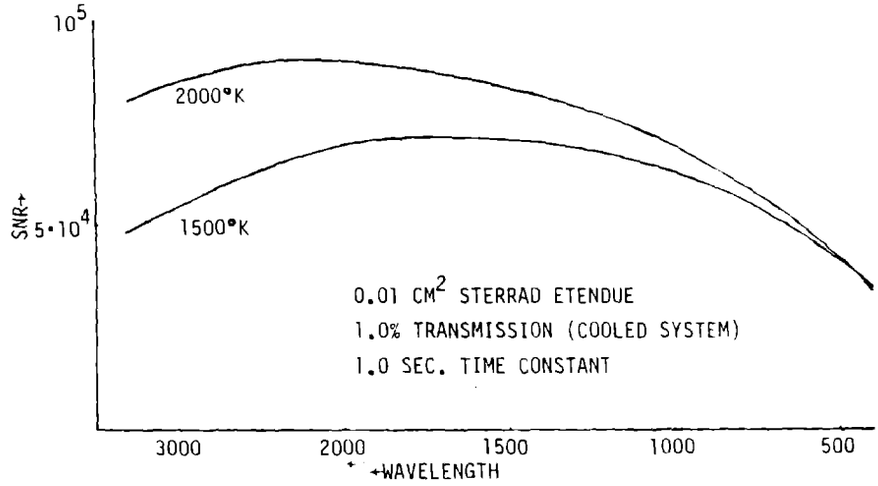


Fig. 1. Shot noise limited signal-to-noise ratio in blackbody source.

*Work performed under the auspices of the U. S. Department of Energy by the Lawrence Livermore National Laboratory under contract number W-7405-Eng-48.

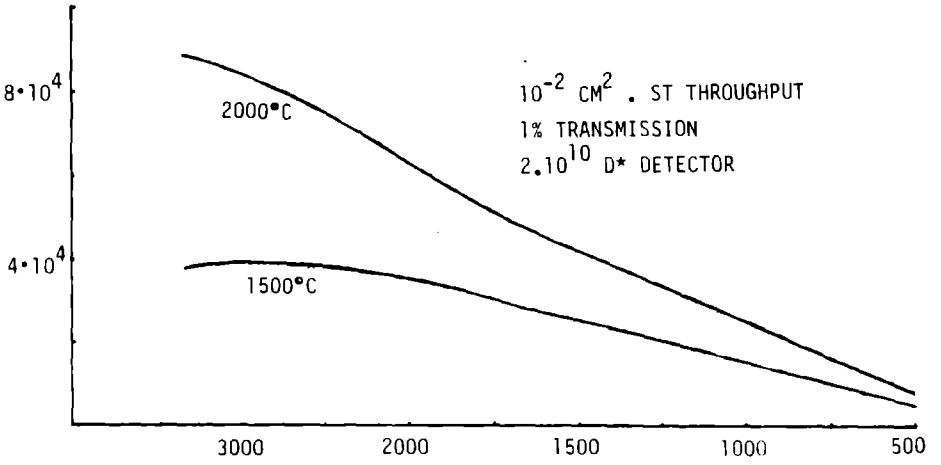


Fig. 2 Detector noise limited signal-to-noise ratio in blackbody emission.

detector performance (Fig. 2), predicted SNRs of the order of 80,000 are well in advance of current instrument performance.

In trying to understand this considerable gap, we may suspect our theoretical understanding of the behavior of the FT-IR instruments is accurate enough. However, a comparison of calculated and actual interferograms shows that, while theoretical analysis does not give exact predictions, it comes quite close. At the present time, large differences between the theoretical and the actual performance of instrumentation are not satisfactorily understood.

In looking for possible improvements, the assumed and realistic 1% efficiency is an obvious candidate (Table I). Clearly, there is need for improved optics and beamsplitter designs. However, an even larger improvement is possible by improving the source. A very high temperature blackbody, such as the hafnium tantalum carbide blackbody source (Fig. 3), operates at 2000° K with high emissivity. The actual performance improvement is even higher than what we could calculate from here, due to a substantial reduction in noise. Apparently, the geometry of this source suppresses air scintillations, an, up to now, unsuspected source of noise in high signal-to-noise FT-IR abosrption measurements. The noise improvement disappears in vacuum instruments, supporting this interpretation.

Table I. Overall efficiency of FT-IR Spectrometer.

Source emissivity	90%
Aperture	95%
Intrinsic efficiency	25%
Beamsplitter efficiency	60%
Angular aperture	86%
Compensator	61%
Surface figure	95%
Mirror loss	51%
Detector efficiency	40%
Electronic efficiency	90%
Total efficiency	1.15%

Another improvement in instrumentation is to tailor the throughput of the instrument to that required for the application. The signal in absorption spectroscopy, the dominant use of FT-IR, is not the light that hits the detector, but the light that fails to get there because it was absorbed by the sample. Therefore, optimum performance is achieved when the geometry of the sampling system is adjusted to make the transmission of the sample approximate 37%.

In many forms of spectroscopy, this can be done by using reduced throughput spectroscopies. Reducing the throughput of the instrument will buy us improvement in such techniques as diffuse reflectance, attenuated total reflectance, microsampling, microprobing, and GC-IR (high throughput instruments are optimal in absorbance substract, low transmission samples, and in photo-acoustic spectroscopy).

We see then that the techniques of FT-IR in which SNR is dominant are all techniques favoring low throughput instruments. Such a low throughput instrument, matched to the throughput of an optimized geometry

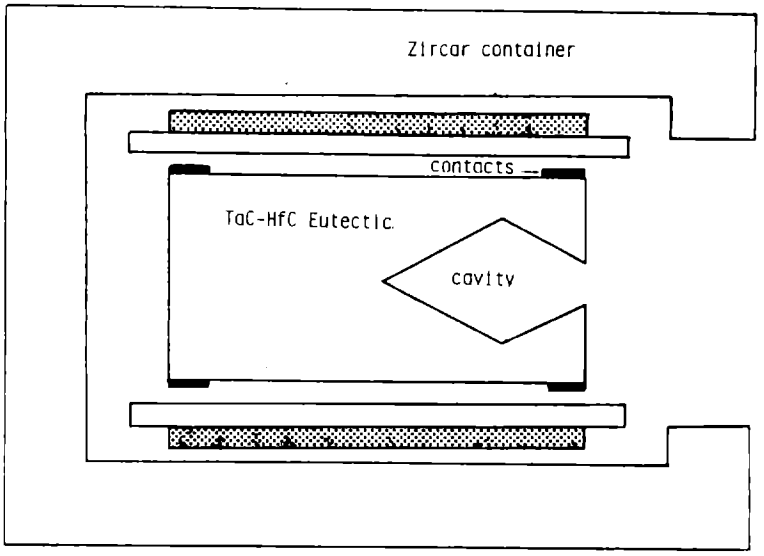


Fig. 3. High temperature blackbody source.

sample, will, of course, see less light. It will, however, see the same or more absorbed light, which is what matters. But the smaller instrument will require a smaller detector. Then, since smaller detectors have less noise, the low throughput instrument is actually the best. The detector can be further reduced by using immersion lenses in front of it. Such a reduced throughput instrument has a number of other advantages. It can use smaller, easier to make, usually higher quality beam splitters. It has far less problems with mirror parallelism. It is much more readily compatible with high power blackbody sources, and much easier to purge. And, last but not least, because it uses very small detectors on which it imposes a small heat load, it can use detectors refrigerated with a Joule-Thompson cooler. This does not require liquid nitrogen (hurrah!) but only 30 cubic centimeters a minute of high pressure nitrogen for cooling to cryogenic temperatures (Fig. 4).

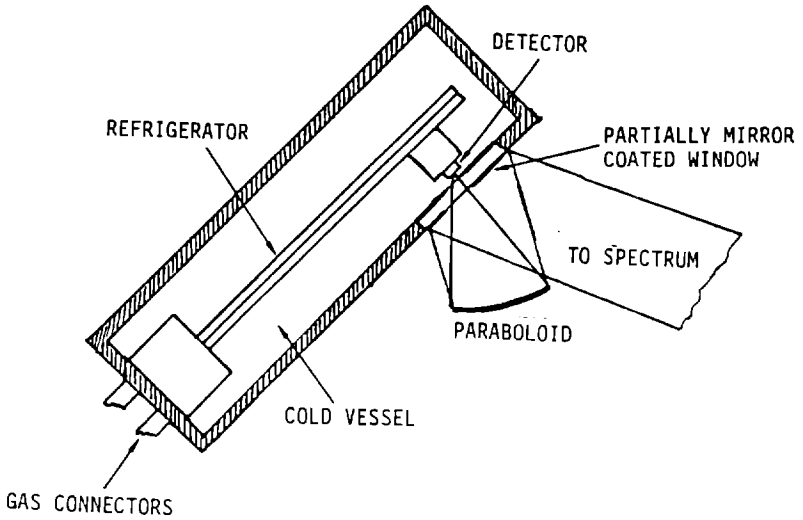


Fig. 4. Miniature cryogenic detector assembly requiring no liquid nitrogen.

When we use such a high efficiency system, the detector signal is quite large. It is well-known that mercury cadmium telluride detectors are nonlinear when they are strongly illuminated. Not only do they lose signals, but they also produce a severely distorted spectrum. While this is well know, it is not true. Much of this apparent nonlinearity in the detector is not a property of the detector at all, but one of its preamplifier. Indeed, by using a properly designed preamplifier (Fig. 5), it is possible to recover just about all the lost linearity. Such a preamplifier serves loops impedance matching to the detector, has a bias source with enough reserve power to avoid droop during the centerburst, and corrects the last residual

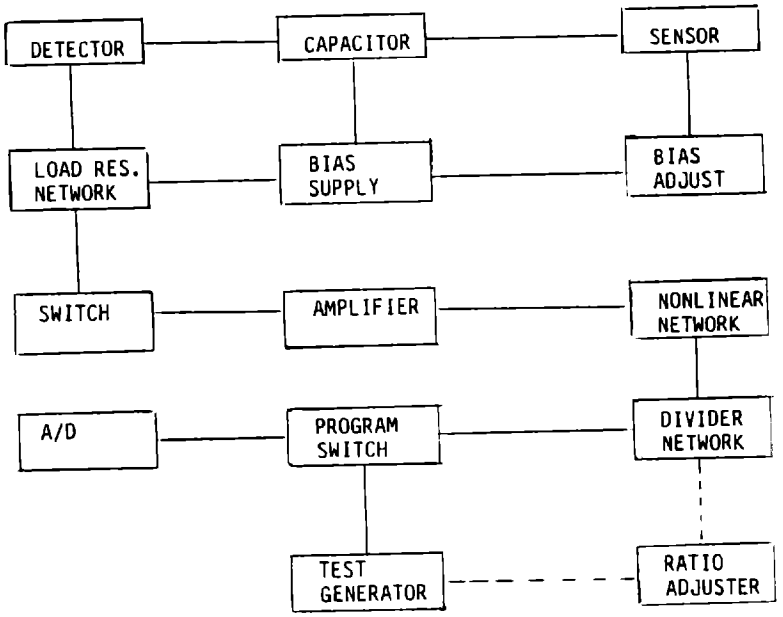


Fig. 5. Block diagram of high dynamic range preamplifier for HgCdTe detector.

Fig. 6. Baseline rectification by superchipping.

nonlinearity (very small) with a nonlinear network. Such systems have achieved from 19 to 21 bits of single-scan detector SNR. Of course, the data system cannot handle 21 bits, and we need to use gain ranging. This is a somewhat treacherous technique because, while it does indeed improve the signal-to-noise of the spectrum, it does so only for features as narrow as the bands. For wide features, such as baseline fluctuations, the use of gain ranging does not provide any improvement. The fluctuations in the baseline behave as if one had used only a fixed gain A/D. The reason for this is that the centerburst of the interferogram from which the baseline is calculated, has indeed been measured at the lower resolution.

One way of solving this problem is a super clip apodization in which we take a small segment out of the interferogram (almost only the centerburst), transform it separately, and use it as a reference spectrum for ratioing. This normalizes the spectrum to a flat baseline (Fig. 6). In order for this technique to work, we require (1) a good quality beam splitter that lacks any perturbing bands, and (2) a weak sample absorption. Fortunately, unless the sample were a weak absorber, we would not be needing the high sensitivity. These techniques have, for example, allowed us to do on the fly GC-IR at subnanogram levels with reasonable

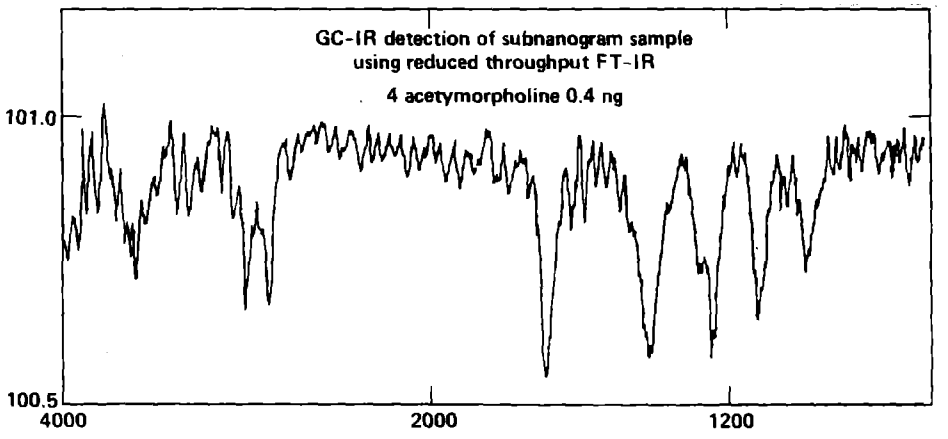


Fig. 7. GC-IR detection of subnanogram sample using reduced throughput FT-IR.

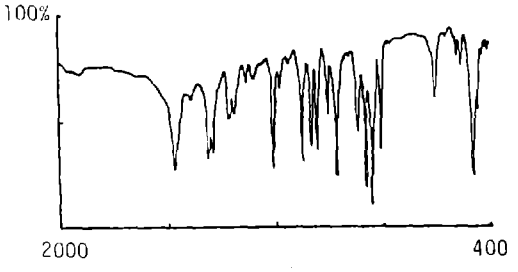


Fig. 8. Detection of small particle using light sensitivity microsampling in FT-IR.

signal-to-noise (Fig. 7), or measure microspots in an infrared microprobe with sample quantities in the low picogram range (Fig. 8).

However, signal-to-noise by itself is not everything. It has been our experience that, while the signal--to-noise ratio of a Fourier transform infrared spectrometer can be, and is very often, better than 2/1000 of a percent, its short-term reproducibility is ten times worse 0.02%, and in the long-term is at best some .1%. The accuracy of the photometric test methods we use to check out the instrument is barely .3%, and, if we do interlaboratory comparisons of infrared analyses, we get agreements of the order of 1%. If, beyond this, we want to do interlaboratory comparison of absolute infrared intensities, then the fluctuation is on the order of 2-5%.

After making all possible efforts to improve the hardware of the system, we can further improve the accuracy of the data by software correction of all known errors. But here we are limited by an insufficient understanding of where all these fluctuations come from.

One of the surprising things about FT-IR has always been how different two nominally identical instruments (same model, same type, different serial numbers) are. These differences are classic examples of how little we, at present, understand the reasons for the irreproducibility of infrared data. Worse yet, we have not been very careful about some problems we do know.

Here, I would like to mention the inexcusably bad quality of double-beam systems in most FT-IR instrumentation today. Double-beam systems in time, are limited by instrumentation stability, a serious problem. But double-beam systems in space suffer from plain sloppy design. It is as if manufacturers have become so confident in the performance and capabilities of FT-IR that they have thought that this exempts them from designing the mechanics of the instrument properly. We find that in most commercial FT-IR spectrophotometers that have a double-beam, the two beams are not matched in atmospheric path length not matched in energy level, not matched in the wavelength scale, not matched in angle, not matched in focal plane, not matched in polarization, and so it goes. The average constructional quality of double-beam systems in the old grating spectrometers was way ahead of current FT-IR instrument, and it is urgent that we recapture the elementary precautions of bygone times in our new instrumentation.

However, while discussing the accuracy of FT-IR, let us remind ourselves that not all the blame should lie with the instrument. Once we have instruments with the accuracy of today's FT-IR instrumentation, it is ridiculous to behave as if classical sample handling procedures for the infrared could possibly be enough to give reproducibilities approaching our instrumental SNR ratios.

We should point out here that the analytical balance is capable of signal-to-noise ratios of $2\text{-}10\cdot10^6$. And yet, when, in the early part of the century, T. W. Richards wanted to measure by gravimetry the atomic weight of silver to four decimals, it took him 26 years of effort to achieve it. Even in a technique as old as the analytical balance, we still have not gotten within three orders of magnitude of its ultimate capabilities. There is no reason then to expect ultimate performance FT-IR very soon either.

Interestingly, one of the problems of the state-of-the-art of FT-IR in that suddenly, and without much warning, the reproducibility of FT-IR instrumentation has largely surpassed the reproducibility of laboratory glassware. The observed convergence of reproducibility data from many manufacturers and researchers may be attributable to the common bottleneck of laboratory glassware. We are getting to the point where these instruments are accurate enough that they would justify gravimetry as an across-the-board method of preparing accurate samples.

In discussing the potential and future of FT-IR and its ultimate limits, it is also necessary to discuss possible new spectroscopies on the horizon, that will one day replace FT-IR as FT-IR has now replaced grating spectroscopy.

Other transform spectroscopies, such as Hadamard transform, correlation transform, sampling transforms, pattern recognitions, Walsh transforms, equiabsorbance contours, and Z transforms have various specialized roles in infrared spectroscopy. Of these, only Hadamard transform spectroscopy has been extensively discussed as a general alternative or successor to FT-IR spectroscopy. Of the suggested applications, the one for which Hadamard transform does not seem promising is the measurement of full-range reasonable resolution spectra. Here, problems in the Hadamard transform, due to diffraction and its effect on the decoding process, limit the number of resolution elements that can be simultaneously viewed. When several thousand resolution elements are being measured, Hadamard transform spectroscopy would not achieve the same performance of FT-IR.

However, in two other applications, Hadamard transform has undoubted advantages over FT-IR and will, no doubt, make a significant impact. These are the very high resolution measurement of fractions of a spectrum, because Hadamard transform spectroscopy need not contain all frequencies in the spectrum down to zero. It can measure short segments of spectrum at very high accuracies with moderate instrumental and computer demands. It is, for this application, intrinsically superior to FT-IR spectroscopy.

Hadamard transforms also enjoy greater simplicity of computation than do Fourier transforms. For current computer technology, this does not matter for normal Fourier transform requirements. But new requirements for imaging infrared spectroscopy, microprobes, and mapping systems may require multidetector systems capable of measuring simultaneously many points of an image. In such systems, the computation load is orders of magnitude larger than that of normal Fourier transform spectroscopy, and the computational advantage of Hadamard becomes crucial.

In discussing possible replacements for Fourier transform spectroscopy, the tunable laser has been mentioned quite often. For the last ten years, we have been hearing "The lasers are coming, the lasers are coming." Well, they are still coming. And, as things look, they will still be coming for a good many years. Tunable lasers have only limited tuning capability, over extremely short ranges, with a lack of continuity even over those, with gross problems in tracking accurately these wavelengths, and with substantial wavelength stability problems. Furthermore, the high resolution, which is the lone claim to fame of the tunable laser, is not one of the major applications of FT-IR today. In fact, this extreme resolution is useful only for low-pressure gas studies which are in no way the main line effort going on in infrared spectroscopy. At the same time, tunable lasers now have quite high noise levels, and even with the best of electronic corrections, are far from achieving the signal-to-noise ratios that are customary in state-of-the-art FT-IRs. If one combines this with the considerable slowdown of research spending in tunable lasers, one can expect these problems will be with us for quite a while yet.

Another alternative to FT-IR seems a good deal closer to date. This is the scanning Fabry-Perot spectrometer. A scanning Fabry-Perot spectrometer, with a piezoelectrically-driven etalon plate, preceeded by an order sorting filter wheel or by an order sorting second interferometer, has the full Jacquinot advantage of Fourier transform spectroscopy. It furthermore can, with simple design modifications, have the field-enhanced throughput advantage as well. While it lacks the Fellgett advantage, it has the ability of scanning quickly to the wavelengths of interest and stopping there, which gives one the slew scan advantage in SNR whenever one does not need the entire spectrum. It has a higher scan speed capability, and, of course, it is a vastly cheaper and simpler system. For applications where quantitative analysis is important and where, therefore, full-scan spectra will not be taken all the time, Fabry-Perot spectroscopy is a quite reasonable alternative to FT-IR.

When we set out to introduce FT-IR to the infrared spectroscopy domain, we did so as infrared spectroscopists, not as FT-IR spectroscopists. We went into FT-IR because it was better infrared spectroscopy. In the years to come, we will be exploring possible successors to FT-IR in the same spirit. The goal of the exercise is better infrared spectroscopy with whatever means we are able to generate.