FROG

In order to measure an event in time, you need a shorter one.
So how do you measure the shortest one?

If you've read the section on autocorrelation, you saw that measuring an ultrashort pulse required using the pulse to measure itself. But, in view of the above little dilemma, that wasn't good enough.

Frequency-Resolved Optical Gating (FROG) involves operating in a hybrid domain: the *time-frequency domain*. Measurements in the time-frequency domain involve both temporal *and* frequency resolution simultaneously. A well-known example of such a measurement is the *musical score*, which is a plot of a sound wave's short-time spectrum vs. time. Specifically, this involves breaking the sound wave up into short pieces and plotting each piece's spectrum (vertically) as a function of time (horizontally). So the musical score is a function of time as well as frequency. See Fig. 1. In addition, there's information on the top indicating intensity.

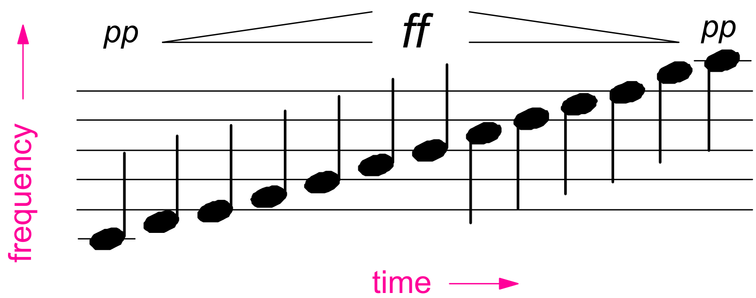

Fig. 1. The musical score is a plot of an acoustic waveform's frequency vs. time, with information on top regarding the intensity. Here the wave increases in frequency with time. It also begins at low intensity (pianissimo), increases to a high intensity (fortissimo), and then decreases again. Musicians call this waveform a "scale," but ultrafast laser scientists refer to it as a "linearly chirped pulse."

If you think about it, the musical score isn't a bad way to look at a waveform. For simple waveforms containing only one note at a time (we're not talking about symphonies here), it graphically shows the waveform's instantaneous frequency, ω, vs. time, and, even better, it has additional information on the top indicating the approximate intensity vs. time (e.g., fortissimo or pianissimo). Of course, the musical score can handle symphonies, too.

A mathematically rigorous version of the musical score is the spectrogram, $\Sigma_g(\omega,\tau)$:

$$\Sigma_g(\omega,\tau) \equiv \left| \int_{-\infty}^{\infty} E(t)\, g(t-\tau)\, \exp(-i\omega t)\, dt \right|^2$$

where $g(t-\tau)$ is a variable-delay gate function, and the subscript on the Σ indicates that the spectrogram uses the gate function, $g(t)$. Figure 2 is a graphical depiction of the spectrogram, showing a linearly chirped Gaussian pulse and a rectangular gate function, which gates out a piece of the pulse. For the case shown in Fig. 2, it gates a relatively weak, low-frequency region in the leading part of the pulse. The spectrogram is the set of spectra of all gated chunks of $E(t)$ as the delay, τ, is varied.

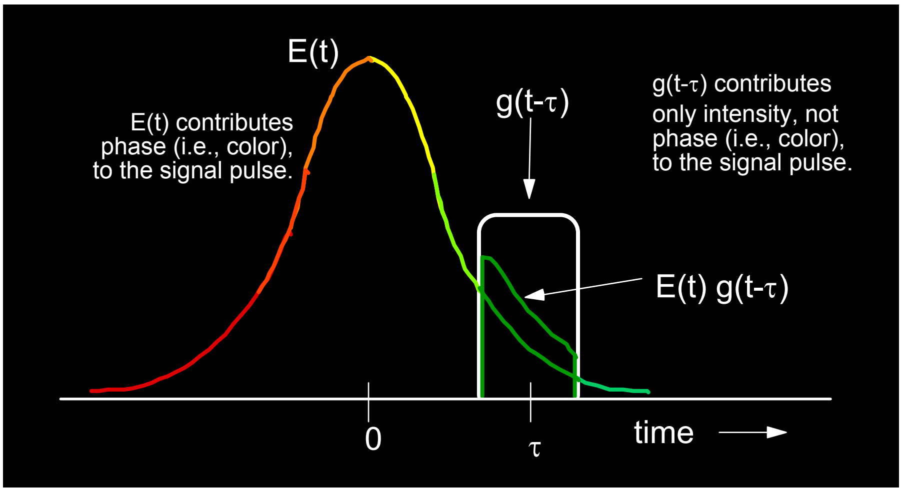

Fig. 2. Graphical depiction of the spectrogram. A gate function gates out a piece of the waveform (here a linearly chirped Gaussian pulse), and the spectrum of that piece is measured or computed. The gate is then scanned through the waveform and the process repeated for all values of the gate position (i.e., delay).

The spectrogram is a highly intuitive display of a waveform. Some examples of spectrograms are shown in Fig. 3, where you can see that the spectrogram intuitively displays the pulse instantaneous frequency vs. time. And pulse intensity vs. time is also evident in the spectrogram. Indeed, acoustics researchers can easily directly measure the intensity and phase of sound waves, which are many orders of magnitude slower than ultrashort laser pulses, but they often choose to display them using a time-frequency-domain quantity like the spectrogram. Importantly, knowledge of the spectrogram of $E(t)$ is sufficient to essentially completely determine $E(t)$ (except for a few unimportant ambiguities, such as the absolute phase, which are typically of little interest in optics problems).

Frequency-Resolved Optical Gating (FROG) measures a spectrogram of the pulse.

Okay, so a spectrogram is a good idea. But recall the dilemma of ultrashort pulse measurement: "In order to measure an event in time, you need a shorter one." In the spectrogram, then, isn't the gate function precisely that mythical shorter event, the one we *don't* have?

Indeed, that is the case.

So, as in autocorrelation, we'll have to use the pulse to measure itself. We must gate the pulse with itself. And to make a spectrogram of the pulse, we'll have to spectrally resolve the gated piece of the pulse.

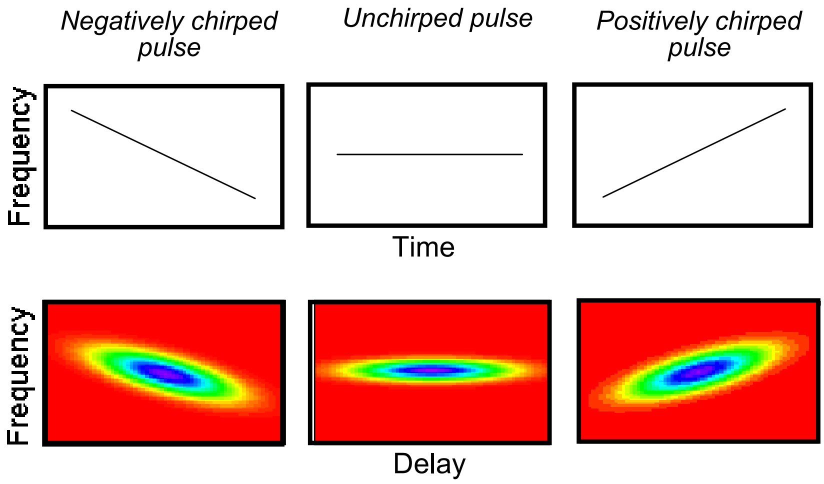

Fig. 3. Spectrograms (bottom row) for linearly chirped Gaussian pulses. The spectrogram, like the musical score, reflects the pulse frequency vs. time. It also yields the pulse intensity vs. time.

Will this work? It doesn't sound much better than autocorrelation, which also involves gating the pulse with itself (but without any spectral resolution). And autocorrelation isn't sufficient to determine even the intensity of the pulse, never mind its phase, too. So how do we resolve the dilemma?

And that's not the only problem. Even if this approach does somehow resolve the fundamental dilemma of ultrashort pulse measurement, spectrogram inversion algorithms assume that we know the gate function. After all, who would've imagined gating a sound wave *with itself* when it's so easy to do so electronically with detectors because acoustic time scales are so slow? So no one ever considered a spectrogram in which the unknown function gated itself—an idea, it would seem, that could occur to only a seriously disturbed individual. Unfortunately, we have no choice; we *must* gate the pulse with itself. But by gating the unknown pulse with itself—i.e., a gate that is also unknown—we can't use available spectrogram inversion algorithms. So all those nice things we said about the spectrogram don't necessarily apply to what we're planning to do. How will we avoid these problems?

Hang on. You'll see.

In its simplest form, FROG is any autocorrelation-type measurement in which the autocorrelator signal beam is spectrally resolved. Instead of measuring the autocorrelator signal energy vs. delay, which yields an autocorrelation, FROG involves measuring the signal *spectrum* vs. delay.

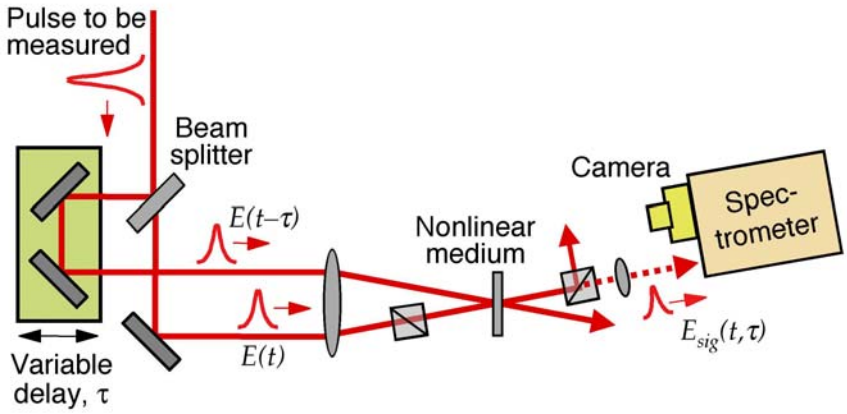

Fig. 4. FROG apparatus using the polarization-gate beam geometry.

As an example, let's consider, not an SHG autocorrelator, but a polarization-gate (PG) autocorrelation geometry. Ignoring constants, as usual, this autocorrelator's signal field is $E_{sig}(t,\tau) = E(t)\ |E(t-\tau)|^2$. Spectrally resolving yields the Fourier Transform of the signal field with respect to time, and we measure the squared magnitude, so the FROG signal trace is given by:

$$I_{FROG}^{PG}(\omega,\tau) \quad = \quad \left| \int_{-\infty}^{\infty} E(t) |E(t-\tau)|^2 \ \exp(-i\omega t)\ dt \right|^2$$

Note that the (PG) FROG trace is a spectrogram in which the pulse intensity gates the pulse field. In other words, the pulse gates itself. The traces obtained by such a technique look just like the spectrograms in Fig. 2. So making a FROG trace yields a very intuitive measure of the pulse. But how do we retrieve the pulse intensity and phase from its spectrogram?

It turns out that this inversion problem is well known. It is called the *two-dimensional phase-retrieval problem*.

Now, the two-dimensional phase-retrieval problem is a close relative of the *one-dimensional* phase-retrieval problem, which is well known to be unsolvable—many ambiguities exist, even in the presence of an additional constraint that might limit the number of spurious solutions. The one-dimensional phase-retrieval problem is bad news. It turns out that the retrieval of the pulse form its spectrum is equivalent to the one-dimensional phase-retrieval problem. And retrieving the intensity from the intensity autocorrelation is also the one-dimensional phase-retrieval problem. And those are unsolvable problems.

Almost certainly, the two-dimensional analog of a one-dimensional piece of mathematical bad news can only be worse news.

Quite unintuitively, however, the two-dimensional phase-retrieval problem has an essentially *unique solution* and is a *solved* problem when certain additional information regarding $E_{sig}(t,\tau)$ is available. This is in stark contrast to the one-dimensional problem, where many solutions can exist, despite additional information. Indeed, in the one-dimensional case, *infinitely* many additional solutions typically exist. On the other hand, the two-dimensional phase-retrieval problem, with a reasonable constraint, has only the usual "trivial" ambiguities, such as an absolute phase or a translation in time. In addition, there is an extremely small probability that another solution may exist, but this is generally not the case for a given trace. This is what is meant by *essentially unique*.

Okay, so the solution isn't totally unique, but it's good enough for practical measurements, where we don't care about the trivial ambiguities, and we probably won't be around long enough to do enough experiments to bump into one of the highly improbable ambiguities.

Now what type of constraint allows FROG retrieval to be essentially unique? It is that $E_{sig}(t,\tau) = E(t)\,|E(t-\tau)|^2$, which is a very strong constraint on the *mathematical form* that the signal field can have. There are other versions of FROG whose constraints are slightly different. For example, in second-harmonic-generation (SHG) FROG, $E_{sig}(t,\tau) = E(t)\,E(t-\tau)$. They're sufficient, too.

Thus, the problem is solved. Indeed, it is solved in a particularly robust manner, with many other advantageous features, such as feedback regarding the validity of the data.

The two-dimensional phase-retrieval problem occurs frequently in imaging problems, where the squared magnitude of the Fourier transform of an image is often measured and where finite support is common. The two-dimensional phase-retrieval problem and its solution are the basis of an entire field, that of image recovery. If you're interested in reading more on it, please check out Henry Stark's excellent book on this subject, *Image Recovery*.

Another way to look at this issue is that phase retrieval is a type of de-convolution, which extracts information that's just beyond the resolution of the device and that initially doesn't seem to be there. For example, image de-convolution techniques can de-blur a photograph, thus retrieving details smaller in size than the apparent resolution of the camera that took the picture. After all, how else can CIA spy satellites read your license plate on the ground?

Indeed, recall Fig. 2, in which a shorter rectangular pulse gates the unknown longer pulse. This was the allegedly required shorter pulse. At the time you first looked at that figure, you were probably thinking, "Too bad we don't have an *infinitely short* gate pulse—a delta-function in time. That'd really do a nice job of measuring the pulse."

But you'd be wrong. If it really were the case that $g(t-\tau) = \delta(t-\tau)$, it's easy to do the integral and see that the resulting spectrogram would be completely independent of frequency. In fact, we would find that $\Sigma_g(\omega,\tau) = I(\tau)$. Thus, in this allegedly ideal case, the spectrogram reduces to precisely the pulse intensity vs. time! All phase-vs.-time information is lost! This is because the gated chunk of the pulse will be infinitely short and hence have infinitely broad spectrum, independent of the pulse color at the time.

So using too short a gate pulse is a bad idea. The time-frequency domain is subtle. Having time- and frequency-domain information simultaneously can be a bit unintuitive. Remember, you can't have perfect time and frequency resolution at the same time, or you'd violate the uncertainty principle. The better your time resolution the worse your frequency resolution. In any case, having both temporal and frequency resolution on the order of the pulse—which is what you have when you use the pulse to gate itself—is the way to go, and that's what happens in FROG. And this resolves the dilemma.

The pulse intensity and phase may be estimated simply by looking at the experimental FROG trace, or the iterative algorithm may be used to retrieve the precise intensity and phase vs. time or frequency. Figure 5 shows a couple of pulses measured using PG FROG.

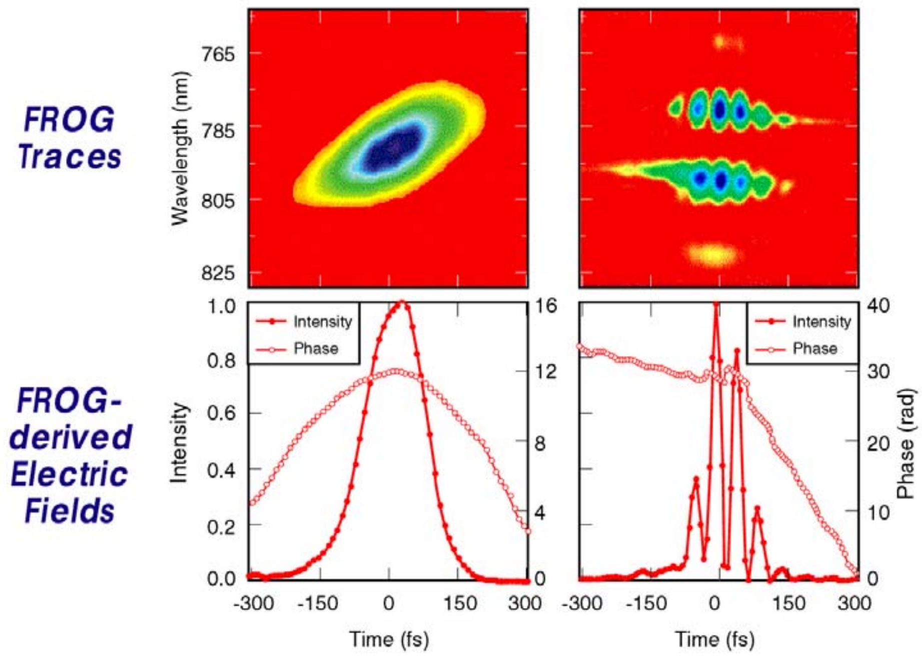

Fig. 5. Two pulses measured using PG FROG. Left: a linearly chirped pulse. Right: a complex pulse. Traces and figure provided by Prof. Bern Kohler, Ohio State University.

There are many different beam geometries for FROG. Essentially any spectrally resolved autocorrelation works, and other geometries do also. The most common and most sensitive FROG beam geometry is second-harmonic-generation (SHG) FROG. (GRENOUILLE is a type of SHG FROG.) The SHG FROG beam geometry is shown in Fig. 6. SHG FROG traces are shown in Fig. 7, which shows that SHG FROG has symmetrical traces and hence has an ambiguity in the direction of time. And Fig. 8 shows an SHG FROG measurement of one of the shortest pulses ever created.

There are many nice features of FROG. FROG is very accurate. Any known systematic error in the measurement can be modeled in the algorithm, although this is not usually necessary, except for extremely short pulses (< 10 fs) or for exotic wavelengths. Also, unlike other ultrashort pulse measurement methods, FROG completely determines the pulse with essentially infinite temporal resolution. It does this by using the time domain to obtain long-time resolution and the frequency domain for short-time resolution. As a result, if the pulse spectrogram is entirely contained within the measured

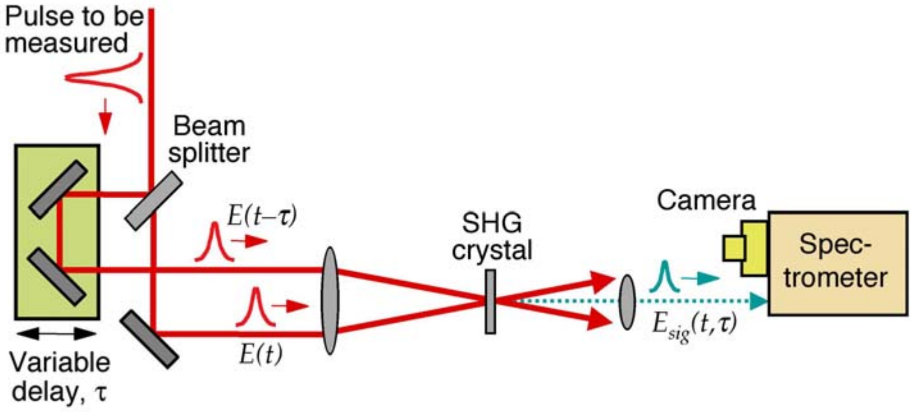

Fig. 6. SHG FROG, the most common and most sensitive version of FROG.

trace, then there can be no additional long-time pulse structure (since the spectrogram is effectively zero for off-scale delays), and there can be no additional short-time pulse structure (since the spectrogram is essentially zero for off-scale frequency offsets).

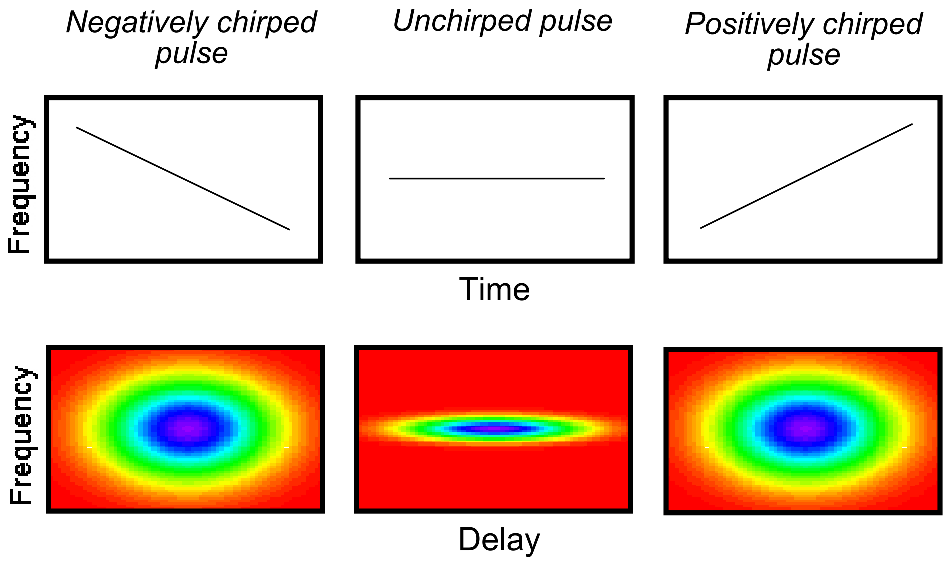

Fig. 7. SHG FROG traces for linearly chirped pulses. Note that the traces are necessarily symmetrical, so the direction of time is not determined. This and a few "trivial" ambiguities are the only known undetermined parameters in SHG FROG.

Interestingly, this extremely high temporal resolution can be obtained by using delay increments that are as large as the time scale of the structure. Again, this is because the short-time information is obtained from large frequency-offset measurements. Thus, as long as the measured FROG trace contains all the nonzero values of the pulse FROG trace, the result is rigorous.

Another useful and important feature that's unique to FROG is the presence of feedback regarding the validity of the measurement data. FROG actually contains two different types of feedback. The first is probabilistic, rather than deterministic, but it is still very helpful. It results from the fact that the FROG trace is a time-frequency plot, that is, an NxN array of points, which are then used to determine N intensity points and N phase points, that is, 2N points. There is thus significant over-determination of the pulse intensity and phase—there are many more degrees of freedom in the trace than in the pulse. As a result, the likelihood of a trace composed of randomly generated points corresponding to an actual pulse is very small. Similarly, a measured trace that has been contaminated by systematic error is unlikely to correspond to an actual pulse. Thus, convergence of the FROG algorithm to a pulse whose trace agrees well with the measured trace virtually assures that the measured trace is free of systematic error. Conversely, non-convergence of the FROG algorithm (which rarely occurs for valid traces) indicates the presence of systematic error. To appreciate the utility of this feature, recall that intensity autocorrelations have only three constraints: a maximum at zero delay, zero for large delays, and even symmetry with respect to delay. These constraints do not limit the autocorrelation trace significantly, and one commonly finds that the autocorrelation trace can vary quite a bit in width during alignment while still satisfying these constraints.

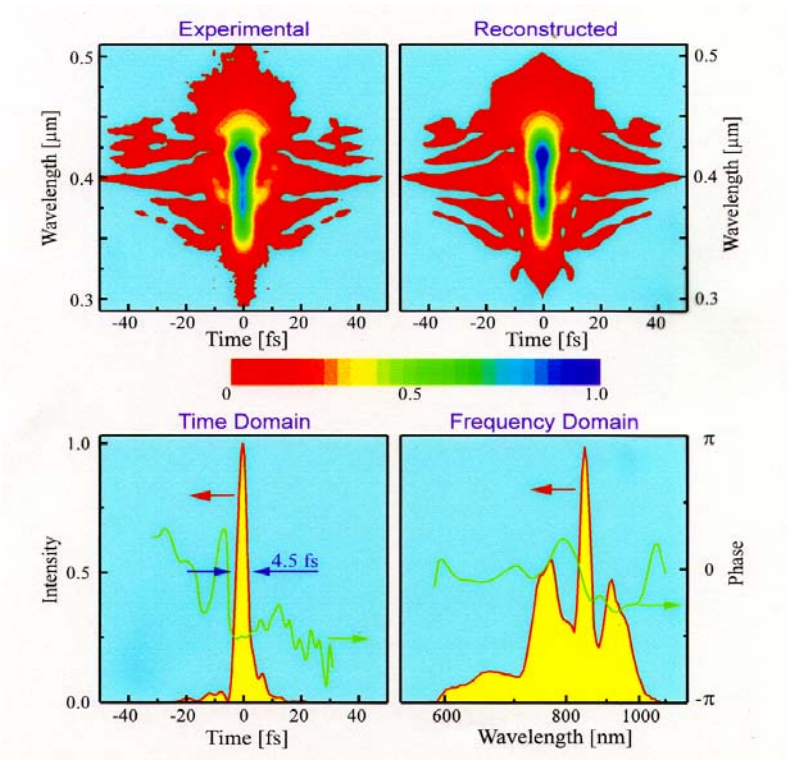

Fig. 8. One of the shortest events ever measured, a 4.5-fs pulse, measured using SHG FROG. Baltuska, Pshenichnikov, and Weirsma, J. Quant. Electron., 35, 459 (1999).

Another feedback mechanism in FROG has proven extremely effective in revealing systematic error in SHG FROG measurements of ~10-fs pulses, where crystal phase-matching bandwidths are insufficient for the massive bandwidths of the pulses to be measured. It involves computing the *marginals* of the FROG trace, that is, integrals of the trace with respect to delay or frequency. The marginals can be compared to the independently measured spectrum or autocorrelation, and expressions have been derived relating these quantities. Comparison with the spectrum is especially useful. Marginals can even be used to correct an erroneous trace.

In practice, FROG has been shown to work very well in the IR, visible, and UV. Work is underway to extend FROG to other wavelength ranges, such as the x-ray. It has been used to measure pulses from a few fs to many ps in length. It has measured pulses from fJ to mJ in energy. And it can measure simple near-transform-limited pulses to extremely complex pulses with time-bandwidth products in excess of 1000. It can use nearly any fast nonlinear-optical process that might be available. FROG has proven to be a marvelously general technique that works. If an autocorrelator can be constructed to measure a given pulse, then making a FROG is straightforward since measuring the spectrum of it is usually easy.

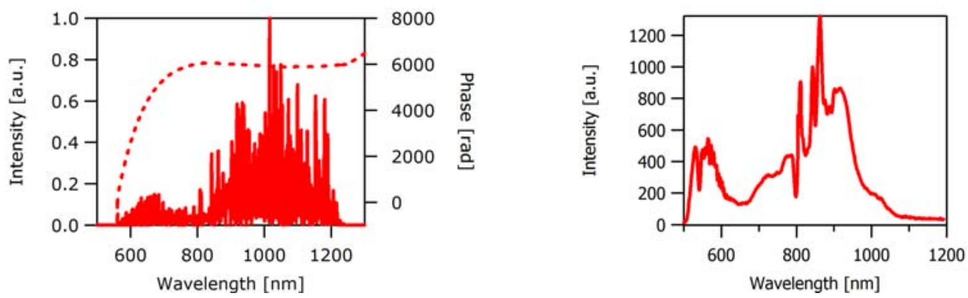

Fig. 9. Measurements of the spectrum of a broadband continuum pulse. The FROG measurement (left) reveals the spectral structure, which washes out in the spectrometer measurement (right).

FROG has other advantages. Figure 9 shows two different measurements of the spectrum of a very broadband light pulse ("continuum"). On the left is a FROG measurement (accumulated over ~10^9 laser shots), and on the right is a simple spectrometer measurement (accumulated over 10^6 laser shots). The continuum spectrum contained much fine-scale structure that fluctuated greatly form pulse to pulse, and which averaged out in the spectrometer spectrum. FROG, on the other hand, because it has both time and frequency resolution, sees the structure. This structure was confirmed by single-shot spectral measurements.

What FROG doesn't measure

We've been saying that FROG measures the complete intensity and phase vs. time or frequency. Actually, there are a few aspects of the intensity and phase that FROG does not measure (the "trivial" ambiguities). First, since FROG is a magnitude-squared quantity, it doesn't measure the absolute phase, φ_0, in the Taylor expansion of the spectral phase. Also, because FROG involves the pulse gating itself, there is no absolute time reference, so FROG doesn't measure the pulse arrival time, which corresponds in the frequency domain to φ_1, the first-order term coefficient in the spectral-phase Taylor series. In other words, the linear component of the slope of the spectral phase will vary randomly, but this is reasonable. So φ_0 and φ_1 are the only two parameters not measured in FROG, although a few versions of FROG have their own unmeasured parameters in specific situations, and these are discussed in *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*. There is, however, a direction-of-time ambiguity in SHG FROG, which means that a pulse and its time-reversed replica are both possible, but this ambiguity can be removed by having some (almost any) additional information available.

In any case, it is common to see the phase jump around apparently randomly due to these undetermined, but not very important, quantities. Please don't interpret this to mean that the FROG algorithm isn't operating properly. Also, by definition, the phase becomes undetermined when the intensity goes to zero. So you'll see the phase jumping around in the pulse wings, where the intensity is nearly zero, too. This is also as it should be.