

Measuring Ultrashort Laser Pulses I: Autocorrelation and FROG

The dilemma

The goal: measuring the intensity and phase vs. time (or frequency)

Why?

The Spectrometer and Michelson Interferometer

1D Phase Retrieval



The Dilemma

In order to measure an event in time, you need a *shorter* one.

To study this event, you need a strobe light pulse that's shorter.



Photograph taken by Harold Edgerton, MIT

But then, to measure the strobe light pulse, you need a detector whose response time is even shorter.

And so on...

So, now, how do you measure the *shortest* event?

Ultrashort laser pulses are the shortest technological events ever created by humans.

It's routine to generate pulses shorter than 10⁻¹³ seconds in duration, and researchers have generated pulses tens of as (10⁻¹⁸ s) long.

Such a pulse is to one second as 5 cents is to the US national debt (or was, around the turn of the century).

Such pulses have many applications in physics, chemistry, biology, and engineering. You can measure any event—as long as you've got a pulse that's shorter.

So how do you measure the shortest pulse?



You must use the pulse to measure itself.

But that isn't good enough. It's only as short as the pulse. It's not shorter.

Techniques based on using the pulse to measure itself have not sufficed.

We must measure an ultrashort laser pulse's intensity and phase vs. time or frequency.

A laser pulse has the time-domain electric field:

$$E(t) = \operatorname{Re}\left\{\sqrt{I(t)} \exp\left[i(\omega_0 t - \phi(t))\right]\right\}$$

Intensity Phase

Equivalently, vs. frequency:

(neglecting the negative-frequency component)

$$\tilde{E}(\omega) = \sqrt{\frac{S(\omega)}{4}} \exp\left[-i\varphi(\omega)\right] \quad \longleftarrow \quad Spectrum \quad Spectral \\ Phase \quad \Theta = \frac{1}{2} \left[-i\varphi(\omega)\right] \quad \Theta = \frac{1}{2} \left[-i\varphi(\omega)\right]$$

Knowledge of the intensity and phase or the spectrum and spectral phase is sufficient to determine the pulse.

The phase determines the pulse's frequency (i.e., color) vs. time.

The instantaneous frequency:

$$\omega(t) = \omega_0 - d\phi/dt$$

Example: "Linear chirp"





We'd like to be able to measure, not only linearly chirped pulses, but also pulses with arbitrarily complex phases and frequencies vs. time.

Pulse Measurement in the Frequency Domain: *The Spectrometer*

The spectrometer measures the spectrum, of course. Wavelength varies across the camera, and the spectrum can be measured for a single pulse.



"Imaging spectrometers" allow many spectra to be measured simultaneously, one for each row of a 2D camera.

Pulse Measurement in the Time Domain: Detectors

Detectors are devices that emit electrons in response to photons.

Examples: Photo-diodes, Photo-multipliers



Detectors have very slow rise and fall times: ~ 1 nanosecond.

As far as we're concerned, detectors have infinitely slow responses. They measure the time integral of the pulse intensity from $-\infty$ to $+\infty$:

$$V_{detector} \propto \int_{-\infty}^{\infty} \left| E(t) \right|^2 dt$$

The detector output voltage is proportional to the pulse energy. By themselves, detectors tell us little about a pulse.

Pulse Measurement in the Time Domain: The Intensity Autocorrelator



Intensity Autocorrelation $A^{(2)}(\tau) \equiv \int_{0}^{\infty} I(t)I(t-\tau) dt$



Varying the delay yields varying overlap between the two replicas of the pulse. The intensity autocorrelation is only nonzero when the pulses overlap.

The Intensity Autocorrelation is always symmetrical with respect to delay.

This is easy to show:

$$A^{(2)}(\tau) = \int I(t) I(t-\tau) dt = \int I(t'+\tau) I(t') dt' = A^{(2)}(-\tau)$$

$$\uparrow^{t'} = t - \tau$$

$$\Rightarrow A^{(2)}(\tau) = A^{(2)}(-\tau)$$

This means that intensity autocorrelation cannot tell the "direction of time" of a pulse.

Of course, it also tells us nothing about the pulse phase either.

Square Pulse and Its Autocorrelation



Gaussian Pulse and Its Autocorrelation



Sech² Pulse and Its Autocorrelation



Since theoretical models of ultrafast lasers often predict sech² pulse shapes, people usually simply divide the autocorrelation width by 1.54 and call it the pulse width. Even when the autocorrelation is Gaussian...

Autocorrelations of more complex intensities

Autocorrelations nearly always have considerably less structure than the corresponding intensity.



An autocorrelation typically corresponds to more than one intensity. Thus the autocorrelation does not uniquely determine the intensity.

Even nice autocorrelations have ambiguities.

These complex intensities have nearly Gaussian autocorrelations.



Conclusions drawn from an autocorrelation are unreliable.

Autocorrelations of more complex pulses: a double pulse



Autocorrelation of Very Complex Pulses

As the intensity increases in complexity, its autocorrelation approaches a broad diffuse background with a coherence spike.



A practical autocorrelator



A more practical autocorrelator



0,1 USD/pc



A more practical autocorrelator



Early prototype, ca. 2004



- finite beam size,
- finite optics size,
- Fresnel reflections,
- plate wedge,
- dispersion of the elements,
- AR coatings









ACrot ver 1.3 (DC motor)

for MINIMODS project Work package No. 1 Task 1.1 Development of miniaturised autocorrelator











Single-shot autocorrelation

Crossing beams at an angle also maps delay onto transverse position.



 $\tau(x) = 2(x/c)\sin(\theta/2) \approx x\theta/c$

Imaging the nonlinear medium onto an array detector allows us to measure a pulse on a single laser shot if we use a large beam and a large beam angle to achieve the desired range of delays.

So single-shot SHG AC has no geometrical smearing! (SHG FROG, too)

Single-Shot Autocorrelation

Crossing beams at a large angle, focusing with a cylindrical lens, and detecting vs. transverse position (*x*) yields $A^{(2)}(\tau)$ for a single pulse.



The beam must have constant intensity vs. *x* to avoid biases.

Other Practical Issues in Autocorrelation

Minimal amounts of glass must be used in the beam before the crystal to minimize the GVD introduced into the pulse by the autocorrelator.

Conversion efficiency must be kept low, or distortions due to "depletion" of input light fields will occur.

In single-shot measurements, the beam must have a **constant intensity vs. position**. In multi-shot measurements, the **beam overlap** in space must be maintained as the delay is scanned.

It's easy to introduce **systematic error**. The only feedback on the measurement quality is that it should be maximal at $\tau = 0$ and symmetrical.

Third-order Autocorrelation

Some ambiguity problems in autocorrelation can be overcome by using a third-order nonlinearity, such as the Optical Kerr effect.



The third-order autocorrelation is **not** symmetrical, so it yields slightly more information, but still not the full pulse.

$$A^{(3)}(\tau) \equiv \int_{-\infty}^{\infty} I(t)I(t-\tau)^2 dt$$

When a shorter reference pulse is available: The Intensity Cross-Correlation

If a reference, very short pulse is available, then it can be used to measure the unknown pulse. In this case, we perform sum-frequency generation, and measure the energy vs. delay.



If the reference (unknown) pulse is much shorter than the unknown pulse, then the intensity cross-correlation fully determines the unknown pulse intensity.

Interferometric Autocorrelation

What if we use a **collinear beam geometry**, and allow the autocorrelator signal light to interfere with the SHG from each individual beam?



Also called the "Fringe-Resolved Autocorrelation"

Interferometric Autocorrelation Math

The measured intensity vs. delay is:

$$IA^{(2)}(\tau) = \int_{-\infty}^{\infty} \left[E^{2}(t) + E^{2}(t-\tau) - 2E(t)E(t-\tau) \right] \left[E^{*2}(t) + E^{*2}(t-\tau) - 2E^{*}(t)E^{*}(t-\tau) \right] dt$$

Multiplying this out:

$$IA^{(2)}(\tau) = \int_{-\infty}^{\infty} \left\{ \left| E^{2}(t) \right|^{2} + E^{2}(t) E^{*2}(t-\tau) - 2E^{2}(t) E^{*}(t) E^{*}(t-\tau) + E^{2}(t-\tau) E^{*2}(t) + \left| E^{2}(t-\tau) \right|^{2} - 2E^{2}(t-\tau) E^{*}(t) E^{*}(t-\tau) + -2E(t)E(t-\tau) E^{*2}(t) - 2E(t)E(t-\tau) E^{*2}(t-\tau) + 4 \left| E(t) \right|^{2} \left| E(t-\tau) \right|^{2} \right\} dt$$

$$= \int_{-\infty}^{\infty} \left\{ I^{2}(t) + E^{2}(t)E^{*2}(t-\tau) - 2I(t)E(t)E^{*}(t-\tau) + E^{2}(t-\tau)E^{*2}(t) + I^{2}(t-\tau) - 2I(t-\tau)E^{*}(t)E(t-\tau) + -2I(t)E(t-\tau)E^{*}(t) - 2I(t-\tau)E(t)E^{*}(t-\tau) + 4I(t)I(t-\tau) \right\} dt$$

where $I(t) \equiv |E(t)|^2$

The Interferometric Autocorrelation is the sum of four different quantities.

$$= \int_{-\infty}^{\infty} I^{2}(t) + I^{2}(t-\tau) dt$$
 Constant (uninteresting)
+ $4 \int_{-\infty}^{\infty} I(t)I(t-\tau) dt$ Intensity autocorrelation

$$- 2 \int_{-\infty}^{\infty} [I(t) + I(t-\tau)] E(t) E^{*}(t-\tau) dt + c.dt$$

Sum-of-intensities-weighted "interferogram" of E(t)(oscillates at ω in delay)

+
$$\int_{-\infty}^{\infty} E^2(t) E^{2*}(t-\tau) dt + c.c.$$

Interferogram of the second harmonic; equivalent to the spectrum of the SH (oscillates at 2ω in delay)

The interferometric autocorrelation simply combines several measures of the pulse into one (admittedly complex) trace. Conveniently, however, they occur with different oscillation frequencies: 0, ω , and 2ω .

Interferometric Autocorrelation: Examples


Does the interferometric autocorrelation yield the pulse intensity and phase?

No. The claim has been made that the Interferometric Autocorrelation, combined with the pulse interferogram (i.e., the spectrum), could do so (except for the direction of time).

Naganuma, IEEE J. Quant. Electron. 25, 1225-1233 (1989).

But the required iterative algorithm rarely converges.

The fact is that the interferometric autocorrelation yields little more information than the autocorrelation and spectrum.

We shouldn't expect it to yield the full pulse intensity and phase. Indeed, very different pulses have very similar interferometric autocorrelations.

Example of Interferometric Autocorrelation ambiguity



Despite very different pulse lengths, these pulses have nearly identical IAs!

Interferometric Autocorrelation: Practical Details and Conclusions

A good check on the interferometric autocorrelation is that it should be symmetrical, and the peak-to-background ratio should be 8.

This device is difficult to align; there are five very sensitive degrees of freedom in aligning two collinear pulses.

Dispersion in each arm must be the same, so it may be necessary to insert a compensator plate in one arm.

The typical ultrashort pulse is still many wavelengths long. So many fringes must typically be measured.

It is difficult to distinguish between different pulse shapes and, especially, different phases from interferometric autocorrelations.

Like the intensity autocorrelation, it must be curve-fit to an assumed pulse shape and so should only be used for rough estimates.

Nonlinear fluorescence and absorption are also used for autocorrelation, interferometric or not.



Resolving the sub- λ fringes yields *interferometric* autocorrelation; otherwise not.

Autocorrelation and related techniques yield little information about the pulse.

Perhaps it's time to ask how researchers in other fields deal with their waveforms...



Consider, for example, acoustic waveforms.

Most people think of acoustic waves in terms of a musical score.



It's a plot of frequency vs. time, with information on top about the intensity.

The musical score lives in the time-frequency domain.

A mathematically rigorous form of a musical score is the spectrogram.

If E(t) is the waveform of interest, its spectrogram is:

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

where $g(t-\tau)$ is a variable-delay gate function and τ is the delay.

Without $g(t-\tau)$, $\Sigma_E(\omega, \tau)$ would simply be the spectrum.

The spectrogram is a function of ω and τ . It is the set of spectra of all temporal slices of E(t).

The Spectrogram of a waveform E(t)

We must compute the spectrum of the product: $E(t) g(t-\tau)$



The spectrogram tells the color and intensity of E(t) at the time, τ .

Spectrograms for Linearly Chirped Pulses



Like a musical score, the spectrogram visually displays the frequency vs. time (and the intensity, too).

Properties of the Spectrogram

Algorithms exist to retrieve E(t) from its spectrogram.

The spectrogram essentially uniquely determines the waveform intensity, I(t), and phase, $\phi(t)$.

There are a few ambiguities, but they're "trivial."

The gate need not be—and should not be—much shorter than E(t).

Suppose we use a delta-function gate pulse:

$$\int_{-\infty}^{\infty} E(t) \,\delta(t-\tau) \exp(-i\omega t) \,dt \Big|^2 = |E(\tau) \exp(-i\omega \tau)|^2$$
$$= |E(\tau)|^2 = \text{The Intensity.}$$
No phase information

The spectrogram resolves the dilemma! It doesn't need the shorter event! It temporally resolves the slow components and spectrally resolves the fast components.

Frequency-Resolved Optical Gating (FROG)

FROG involves gating the pulse with a variably delayed replica of itself in an instantaneous nonlinear-optical medium and then spectrally resolving the gated pulse vs. delay.



Use any ultrafast nonlinearity: Second-harmonic generation, etc.



 $E_{sig}(t,\tau) \propto E(t) \left| E(t-\tau) \right|^2$



The gating is more complex for complex pulses, but it still works. And it also works for other nonlinear-optical processes.

FROG Traces for Linearly Chirped Pulses



Like a musical score, the FROG trace visually reveals the pulse frequency vs. time—for simple and complex pulses.

FROG Traces for More Complex Pulses



The FROG trace is a spectrogram of E(t).

Substituting for $E_{sig}(t, \tau)$ in the expression for the FROG trace:

$$E_{sig}(t,\tau) \propto E(t) |E(t-\tau)|^2$$

$$I_{FROG}(\omega,\tau) \propto \left| \int E_{sig}(t,\tau) \exp(-i\omega t) dt \right|^2$$

yields:

$$I_{FROG}(\omega,\tau) \propto \left| \int E(t) g(t-\tau) \exp(-i\omega t) dt \right|^{2}$$
$$g(t-\tau) = |E(t-\tau)|^{2}$$

where:

Unfortunately, spectrogram inversion algorithms require that we know the gate function.

Generalized Projections

A projection maps the current guess for the waveform to the closest point in the constraint set.



Convergence is guaranteed for convex sets, but generally occurs even with non-convex sets and in particular in FROG.

A few more slides on phase retrievalat the end of the lecture notes

Ultrashort pulses measured using FROG



Data courtesy of Profs. Bern Kohler and Kent Wilson, UCSD.

Second-harmonic-generation FROG

SHG FROG is simply a spectrally resolved autocorrelation.



SHG FROG is the most sensitive version of FROG.

SHG FROG

 $E_{sig}(t,\tau) \propto E(t)E(t-\tau)$



SHG FROG is also a spectrogram, but its interpretation is more complex.

SHG FROG traces are symmetrical with respect to delay.



SHG FROG has an ambiguity in the direction of time, but it can be removed.

SHG FROG traces for complex pulses



SHG FROG traces are symmetrized PG FROG traces.

SHG FROG Measurements of a 4.5-fs Pulse



Baltuska, Pshenichnikov, and Weirsma, J. Quant. Electron., 35, 459 (1999).

Generalizing FROG to arbitrary nonlinear-optical interactions

FROG is simply a frequency-resolved nonlinear-optical signal that is a function of time and delay.



Use any fast nonlinear-optical process to create the signal field.

Pulse retrieval remains equivalent to the 2D phase-retrieval problem.

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

Many interactions have been used, e.g., polarization rotation in a fiber.

When a known reference pulse is available: Cross-correlation FROG (XFROG)

If a known pulse is available (it need not be shorter), then it can be used to fully measure the unknown pulse. In this case, we perform sum-frequency generation, and measure the spectrum vs. delay.



XFROG completely determines the intensity and phase of the unknown pulse, provided that the gate pulse is not too long or too short. If a reasonable known pulse exists, use XFROG, not FROG.

Linden, et al., Opt. Lett., 24, 569 (1999).

Example of XFROG measurement: microstructure-fiber ultrabroadband continuum.





The continuum has many applications, from medical imaging to metrology.

It's important to measure it.

Measuring the continuum with XFROG

It's better to gate a complicated pulse with a simple (known) one.



We angle-dither the crystal to increase the phase-matching bandwidth.

XFROG measurement of the continuum



While the large-scale structure of each trace is identical, the measured trace lacks the fine-scale structure of the retrieved trace.

Wave mixing for CW and pulsed fields



type I and type II in KDP



Poor Man's FROG – the first prototype



PM FROG – does it actually do the job?



C. Radzewicz, P. Wasylczyk, J. S. Krasiński, A Poor Man's FROG, Optics Comm. 186, 329 (2000)

GRating-Eliminated No-nonsense Observation of Ultrafast Incident Laser Light E-fields (GRENOUILLE)



C. Radzewicz, P. Wasylczyk, and J. S. Krasinski, Opt. Comm. 2000. P. O'Shea, M. Kimmel, X. Gu and R. Trebino, Optics Letters, 2001.

The thick crystal

Suppose white light with a large divergence angle impinges on an SHG crystal. The SH generated depends on the angle. And the angular width of the SH beam created varies inversely with the crystal thickness.



GRENOUILLE Beam Geometry



Yields a complete single-shot FROG. Uses the standard FROG algorithm. Never misaligns. Is more sensitive. Measures spatio-temporal distortions!

Testing GRENOUILLE

Compare a GRENOUILLE measurement of a pulse with a triedand-true FROG measurement of the same pulse:





Read more about GRENOUILLE in the cover story of OPN, June 2001

Retrieved pulse in the time and frequency domains


Disadvantages of GRENOUILLE

Its low spectral resolution limits its use to pulse lengths between ~ 20 fs and ~ 1 ps.

Like other single-shot techniques, it requires good spatial beam quality.



Improvements on the horizon:

Inclusion of GVD and GVM in FROG code to extend the range of operation to shorter and longer pulses.

PM FROG on the market (aka GRENOUILLE, Swamp Optics)





Consider FROG as a two-dimensional phase-retrieval problem.

If $E_{sig}(t, \tau)$ is the 1D Fourier transform with respect to Ω of some new signal field, $\hat{E}_{sig}(t, \Omega)$, then:

The input pulse, E(t), is easily obtained from $\hat{E}_{sig}(t,\Omega)$: $E(t) \propto \hat{E}_{sig}(t,0)$ $\hat{E}_{sig}(t,0) = \int E(t) \left| E(t-\tau) \right|^2 \exp(i(0)\tau) d\tau$ $I_{FROG}(\omega, \tau)$ $= E(t) \int \left| E(t-\tau) \right|^2 d\tau$ $= E(t) \int \left| E(\tau') \right|^2 d\tau' \quad \leftarrow \quad \tau' = t - \tau$ So we must invert this in $\propto E(t)$ This integral-inversion p for which the solution exists and is (essentially) unique. And simple algorithms exist for finding it.

1D vs. 2D Phase Retrieval

1D Phase Retrieval: Suppose we measure $S(\omega)$ and desire E(t), where:

$$S(\omega) = \left| \int_{-\infty}^{\infty} E(t) \exp(-i\omega t) dt \right|$$

Given $S(\omega)$, there are infinitely many solutions for E(t). We lack the spectral phase.



2D Phase Retrieval: Suppose we measure $S(k_x, k_y)$ and desire E(x, y):

$$S(k_x, k_y) = \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x, y) \exp(-ik_x x - ik_y y) \, dx \, dy \right|$$

Stark, Image Recovery, Academic Press, 1987.

Given $S(k_x, k_y)$, there is essentially one solution for E(x, y)!!! It turns out that it's possible to retrieve the 2D spectral phase!

These results are related to the Fundamental Theorem of Algebra.

Phase Retrieval and the Fundamental Theorem of Algebra

The Fundamental Theorem of Algebra states that all polynomials can be factored:

$$f_{N-1}z^{N-1} + f_{N-2}z^{N-2} + \dots + f_1z + f_0 = f_{N-1}(z-z_1)(z-z_2)\dots(z-z_{N-1})$$

The Fundamental Theorem of Algebra fails for polynomials of two variables. Only a set of measure zero can be factored.

$$f_{N-1,M-1}y^{N-1}z^{M-1} + f_{N-1,M-2}y^{N-1}z^{M-2} + \dots + f_{0,0} = ?$$

Why does this matter?

The existence of the 1D Fundamental Theorem of Algebra implies that 1D phase retrieval is impossible.

The non-existence of the 2D Fundamental Theorem of Algebra implies that 2D phase retrieval is possible.

1D Phase Retrieval & the Fundamental Theorem of Algebra

The Fourier transform $\{F_0, ..., F_{N-I}\}$ of a discrete 1D data set, $\{f_0, ..., f_{N-I}\}$, is:

$$F_{k} \equiv \sum_{m=0}^{N-1} f_{m} e^{-imk} = \sum_{m=0}^{N-1} f_{m} z^{m} \quad \text{where } z = e^{-ik}$$

The Fundamental Theorem of Algebra states that any polynomial, $f_{N-1}z^{N-1} + ... + f_0$, can be factored to yield: $f_{N-1}(z-z_1)(z-z_2) ... (z-z_{N-1})$

So the magnitude of the Fourier transform of our data can be written:

$$|F_k| = |f_{N-1}(z-z_1)(z-z_2)...(z-z_{N-1})|$$
 where $z = e^{-ik}$

Complex conjugation of any factor(s) leaves the magnitude unchanged, but changes the phase, yielding an ambiguity! So 1D phase retrieval is impossible!

2D Phase Retrieval and the Fundamental Theorem of Algebra

The Fourier transform $\{F_{0,0}, \dots, F_{N-1,N-1}\}$ of a discrete 2D data set, $\{f_{0,0}, \dots, f_{N-1,N-1}\}$, is:



But we cannot factor polynomials of two variables. So we can only complex-conjugate the entire expression (yielding a trivial ambiguity).

Only a set of polynomials of measure zero can be factored. So 2D phase retrieval is possible! And the ambiguities are very sparse.