Spectral Super-Resolution by Understanding Superposition Principle and Detecting Processes

Chandrasekhar Roychoudhuri* and Moncef Tayahi

*Photonics Lab., Physics Department, University of Connecticut, Storrs, CT 06269-5192, USA
bDepartment of Electrical Engineering, University of Nevada, Reno, NV 89557, USA
Tel: 1-860-486-2587; Fax: 1-860-1033; E-mail: chandra@phys.uconn.edu

Abstract- We present analytical and experimental results demonstrating a conceptual breakthrough on how to overcome the fundamental time-frequency bandwidth limit (or spectral super-resolution) for a light pulse in traditional and heterodyne spectrometry. This can be achieved either by a de-convolution process suggested by an analytical method or by heterodyne method, which we demonstrate.

Index Terms- Interference, superposition principle, time-frequency Fourier theorem, overcoming time-frequency bandwidth limit, super resolution in spectroscopy.

I. INTRODUCTION

The application oriented objective of this paper is to analytically and experimentally demonstrate that the traditional “time frequency bandwidth limit” $\delta \nu \delta t \geq 1$ is not a fundamental limit of nature [1]. In traditional spectrometric instruments $\delta \nu$ does represent actual spectral fringe broadening due to time-finite amplitude envelope of a light pulsed, but this broadening is due to spread of energy of the same source-generated carrier frequencies of the pulse and not due to generation of new optical carrier frequencies by the spectrometer or the amplitude envelope. Accordingly, one can obtain super resolution in all spectroscopic experiments. Separate, simultaneous measurement of the temporal envelope of the pulse is necessary while using traditional spectrometers. In heterodyne spectroscopy, the temporal envelope measurement is useful but not essential.

The second objective of this paper is to explain the conceptual foundation behind the derivation of the above mentioned results, which is an attempt to visualize the interaction process behind the measurements that register the effect of superposition of more than one light beams.

Our starting point is the universal observation that the EM fields (well defined light beams) really do not operate on (or, modify the energy of) each other [2] when they are superposed (occupy the same space and time), especially in the absence of materials (dipoles). Otherwise we could not have recognized each other on earth; we could not have discovered the systematic Doppler shifts of distant star lights (Expanding Universe!) and we could not have enjoyed the global internet revolution while diverse information channels are being carried by many different frequencies through the same, hair-thin fiber! In spite of non-interference of light, we continue to use the phrases, like “interference of light” (for centuries) and “single photon interference” (for many decades).

In the second (next) section we briefly show the similarity between the Maxwell and Fourier representations of linear superposition of simple harmonic oscillation function, which is supposed to vindicate the universal principle of superposition. Then we review our previously published experiments [3], which validates that the effect of superposition can be registered only in the presence of interacting detectors and the reported effect is “colored” by the different quantum properties of different detector dipoles.

IJMOT-2006-5-46 © 2006 ISRAMT
The third section gives the summary of the derivation of the classical spectrometric formulation based on direct time-domain propagation of an incident pulse in the real-space, instead of working in the Fourier transformed frequency-space. This gives the pulse-impulse response of a spectrometer when the pulse shape is known. The de-convolution of this pulse impulse-response gives the actual content of the carrier (E-vector) frequencies and allows us to obtain super resolution. The time integrated expression, by virtue of the Parseval’s energy conservation theorem, is equivalent to the traditional Fourier convolution expression for the pulse broadened “spectral” fringe. This establishes why we can, in most cases, ignore the difference between real-space and the frequency-space analyses.

The fourth section presents the result of heterodyne detection of amplitude modulated pulses, demonstrating that the carrier frequency content can be directly measured [4] irrespective of the pulse shape, provided the pulse is much longer than, first, the time constant of the photo detector, and second, the photo conductive electron emission process.

II. NON-INTERACTION OF LIGHT INSPITE OF FOURIER & MAXWELL

Maxwell’s free space wave equation [5] is given by:

$$\nabla^2 E = (1/c^2) \partial^2 E / \partial t^2$$

A simple CW solution to Eq.1, neglecting the arbitrary phase factor, is

$$b \exp[-i2\pi \nu t]$$.  

Mathematically, any linear combination of this solution,

$$b_{total}(t) = \sum_n b_n \exp[-i2\pi \nu_n t]$$

will also satisfy Maxwell’s wave equation. Well before Maxwell, Fourier established a very useful theorem [6] for handling a time finite signal by its mathematical transform in the frequency space using the well-known integral:

$$a(t) = \int_0^\infty \tilde{a}(f) \exp[-i2\pi ft]df$$

The inverse transform is represented by:

$$\tilde{a}(f) = \int_0^\infty a(t) \exp[-i2\pi ft]dt$$

We have deliberately used different symbols \(\nu\) & \(f\) for the frequencies used by Maxwell’s wave equation and the Fourier’s time-frequency theorem to underscore the difference between the actual carrier frequencies for EM waves \(\nu\) and the generalized Fourier’s mathematical frequencies \(f\), which may or may not be identical for all cases of actual experiments. This point will be apparent later. Notice the mutually supporting mathematical equivalency between the summation of Eq.2 and the integral of Eq.3, which leads us to assume that Eq.2 is a physical phenomenon as if EM fields interact with each other all by themselves to generate a new time-finite resultant EM field with a new mean carrier frequency. The assumption is strengthened by the fact that Maxwell’s wave equation is derived from the so-called Maxwell’s four equations of electromagnetism, which are all essentially derived from experimental observations (Coulomb’s law, Ampere’s law, Faraday’s law of electromagnetic induction and Maxwell’s displacement current). It is the structure of the mathematics invented by us and the inherent properties of the sinusoidal functions that is behind this apparent mathematical congruency between the Eq.1, 2 and 3. Note that, as is the tradition, we have used time infinite sinusoidal solutions in the right hand sides of the Eq.2 and 3, which is non-causal in the real world because it attempts to override the inviolable law of conservation of energy. No real physical signal can have infinite spatial extension and time duration. Of course, approach has been developed to solve this issue by truncating the Fourier amplitudes [7] within a finite range of time T and frequency F and Parseval’s theorem of energy conservation makes the approach logically congruent:

$$\int_{-T}^{T} |a_n(t)|^2 dt = \int_{f}^{F} |\tilde{a}_n(f)|^2 df$$

So, it may appear that we can solve the problem by simply using a time finite solution to Maxwell’s wave equation and replace \(b_n\) by \(b_n(t)\):
\[ b_{\text{total}}(t) = \sum_{n} b_n(t) \exp[-i2\pi v_n t] \]  

(6)

Unfortunately, the real world’s observed interaction processes indicate that EM fields cannot be summed as in Eq.6 because they do not operate on each other by themselves even when they are physically superposed within the same space and time domain as we have underscored in the introduction. Something has to sum the fields or their effects. The Eq.6 is mathematically correct, but it does not represent the actual process of interaction between the field and the detector that is behind the observation of the superposition effects of light beams in the real world. We “see” light through the “eyes” of the detectors. It is the detector molecules that are linearly (in the low intensity regime) stimulated by each of the superposed EM fields with a strength given by the first order linear polarizability factor \( \chi^{(1)} \) [8]. If the quantum mechanical properties of the detector allow it to respond to all the superposed \( n \)-fields as dipole undulations \( d_n(t) \), then it is this detector molecule that sums the effects of all the superposed fields and makes the photo current or exposure (interference fringes) visible to us:

\[ I(t) = \left| d_{\text{total}}(t) \right|^2 = \left| \sum_{n} \chi_n^{(1)} b_n(t) e^{-i2\pi v_n t} \right|^2 \]

(7)

Eq.7 represents the real physical detection process.

To underscore the reality of Eq.7, we have carried out an experiment with two CW frequencies separated by 2 GHz, symmetrically centered on one of the Rb-resonance lines [3]. When the superposed beam is sent through a Rb-vapor tube, it did not show any resonance fluorescence, even though by simple trigonometry (according to two terms Fourier synthesis), we were supposed to get the matching resonance frequency (mean of the sum of the two superposed frequencies):

\[ b_{\text{total}}(t) = b_0 \cos 2\pi v_1 t + b_0 \cos 2\pi v_2 t = 2b_0 \cos 2\pi \frac{v_1 - v_2}{2} t \cos 2\pi \frac{v_1 + v_2}{2} t \]

(8)

This revalidates that light beams do not operate on each other by themselves. However, when we sent this same superposed beam on to a high-speed photo conductor, we found the traditional AC current undulating at the difference (beat) frequency. The valence and the conduction bands of the photo detector are broad.

Figure 1. Comparison of energy level diagrams for Rb sharp lines and detector broad bands with the quantum of energies as absorbed by the detectors from the two EM fields carrying two different optical frequencies.

This allows the detecting dipoles to simultaneously respond to all the allowed frequencies (here two), and the resultant current becomes:

\[ I(t) = \left| de^{-i2\pi v_1 t} + de^{-i2\pi v_2 t} \right|^2 = 2d^2 [1 + \cos 2\pi (v_1 - v_2)t] \]

(9)

The detailed “picture” in our view is that the undulating electric vector of the EM field induces the material dipoles to undulate with it. If the frequency matches with the quantum mechanically allowed transition frequency, then only there is transfer of energy from the field to the dipoles. For the superposition effects to be manifest, the detecting dipoles must be collectively allowed to respond to all the light beams simultaneously. This collective response to all the allowed fields of different frequencies makes the dipoles’ undulation strength vary with time. Accordingly, the rate of transfer of the number of electrons from the valance to the conduction band undulates with time and we get the “beat” current. The generation of the beat current also requires that the Poynting vectors of the two superposed light beams be perfectly collinear (completely match their wave fronts).
III. SUPER RESOLUTION IN CLASSICAL SPECTROMETRY BY TIME-DOMAIN ANALYSIS

It is important to recognize, as indicated by Eq.6, that all optical signals have to be time finite; even a CW laser has to be turned on and turned off. So, one should develop spectrometer behavior in terms of a time-finite optical signal with a given unique carrier frequency. This approach is also congruent with our understanding of light emitters like atoms and molecules. They have a finite life time and absorb and emit a finite amount of energy $\Delta E = h\nu$. We are making a rational assumption that when such a finite amount of energy is released by an atom or a molecule, it propagates out as a time-finite wave packet with a carrier frequency $\nu$.

Let us now propagate such a wave packet through a spectrometer. This propagation and evolution of a pulse in the real-space is depicted in the Fig. 2a, b. It is well recognized in classical optics that the finite width of the spectrometer fringes arise due to its limitation in producing a finite number N replicated beams out of the incident one [7, see Ch.7 & 8]. But, due to the periodic delay, $\tau$, the train of replicated short pulses are only partially superposed, causing a further broadening of the “spectral” fringe, which has nothing to do with the actual frequency content of the pulse. A fast detector array will register a time varying spectral fringe width variation given by the square modulus of the train of partially superposed pulses, where $d$ is the strength of the detector stimulation by the fields [1, 9, 10]:

$$\left|i_{out}(t)\right|^2 = \left|\sum_{n=0}^{N-1} Tr^n d(t - n\tau) e^{i2\pi \nu(t-n\tau)}\right|^2_{FP} (10a)$$

The Eq.10a and 10b correspond to the two classical spectrometers, Fabry-Perots (FP) and Gratings (Gt), respectively. The expression for the Michelson’s Fourier transform spectrometry can be obtained from Eq.10b by inserting $N = 2$ for two beam interferometry. The Eq.10 clearly indicates time varying fringe width that can be recorded by a streak camera. We should not interpret such time evolving fringe width as time evolving spectral content of the pulse.

![Figure 2. Schematic diagrams](image-url)
spectrometer with linear response does not produce any new optical frequencies. The time integrated fringe energy distributions or the pulse-impulse responses for an FP and a grating are given by:

\[ I_{\text{fps}}(v, \tau)_{\text{FP}} = \sum_{n=0}^{N-1} T^2 R^{2n} + 
\]

\[ 2 \sum_{n,m} N^{-1} T^2 R^{n+m} \gamma(n-m)\tau \cos[2\pi(n-m)v\tau] \]

\[ I_{\text{fps}}(v, \tau)_{\text{Grating}} = \frac{1}{N} + 
\]

\[ \frac{2}{N^2} \sum_{p=1}^{N-1} (N-p)\gamma(p\tau)\cos[2\pi pv\tau] \]

\[ I_{\text{fps}}(v, \tau)_{\text{Michelson}} = \frac{1}{2}[1 + \gamma(\tau)\cos 2\pi v\tau] \]

We have written \( \gamma([m-n]\tau) \equiv \gamma(p\tau) \); this normalized autocorrelation function is defined as:

\[ \gamma(p\tau) = \frac{\int (d(t-n)\tau) \cdot (d(t-m)\tau) \, dt}{\int d^2(t) \, dt} \]

The limits of the integrations in the numerator and the denominator span the duration of the pulse train and the pulse pair, respectively, which is causal. Notice that unlike traditional definition of coherence [7] we have defined the autocorrelation function in terms of the detector stimulation since operationally EM fields cannot be summed as they do not operate on each other. In Fig.2 we have defined a spectrometer time constant \( \tau_0 \) as the duration of the partially overlapped \( N \)-pulses:

\[ \tau_0 = N\tau \]

It is not very difficult to demonstrate that in the limit as \( \delta\tau \to \tau_0 = N\tau \), \( \gamma(p\tau) \to 1 \) and the Eq.11a, b & c yield to classical text book formulas [7] for the Fabry-Perot, Grating and Michelson spectrometers, respectively:

\[ \frac{d}{dt} \int_{t_0 = N\tau}^{t_0 + N\tau} I_{\text{fps}}(v, \tau) \, dt = I_{\text{cw}}(v, \tau) \]

The operational implication is that whenever the duration of the incident light pulse is longer than the spectrometric time constant \( \tau_0 \), that is the pulse width \( \delta t > \tau_0 = N\tau \), the functional form of the spectral fringe is indistinguishable from the classical CW formulas we are accustomed to use. Accordingly, we claim that our time domain formulation represented by Eq.10 and 11 is a better and more generalized approach to spectrometry than the traditional approach since it starts with a general pulse and derives the CW formulas as a special case when the pulse is longer than \( \tau_0 \). This also gives us a new insight into the spectrometric fringe processing, a characteristic time constant, which has hitherto remained un-appreciated by us.

One can now appreciate that just as in the case of CW light, we de-convolves the CW-impulse response from the recorded fringes to recover the actual carrier frequency content, so should we de-convolve the pulse-impulse response function (Eq.11) when the light is pulsed. Thus when one knows the pulse envelope, one can achieve super resolution using the de-convolution of the pulse-impulse response function [9].

We will now show the equivalency of our time integrated fringe pattern with traditional assumption of the convolution of the CW-impulse response with the Fourier intensity spectrum of the pulse. The Fourier frequencies of the input amplitude envelope are given by:

\[ FT[a(t)e^{j2\pi v\tau}] = \tilde{a}(v-f) \]

The FT-Kernel is \( \exp(-j2\pi ft) \). The output amplitude, as per time-domain approach, is:

\[ i_{\text{out}}(t, v) = h(t) \otimes a(t)e^{j2\pi v\tau} \]

Then its Fourier transform and the corresponding energy are:

\[ \tilde{i}_{\text{out}}(f) = \tilde{h}(f) \cdot \tilde{a}(v-f) \]

\[ |\tilde{i}_{\text{out}}(f)|^2 = \tilde{H}(f) \cdot \tilde{A}(v-f) \]

Then we apply the Parseval’s theorem of conservation of energy to obtain:

\[ \int_{-\infty}^{\infty} |i_{\text{out}}(t)|^2 \, dt = \int_{-\infty}^{\infty} |\tilde{i}_{\text{out}}(f)|^2 \, df \]

\[ \int_{-\infty}^{\infty} \tilde{H}(f) \cdot \tilde{A}(v-f) \, df = \tilde{H}(v) \otimes \tilde{A}(v) \]
We now have the desired equivalency of the time-domain and the Fourier’s frequency-domain analysis, but only with the arbitrary assumption that actual carrier frequency of an EM field and the mathematical Fourier frequency are interchangeable, \( f \equiv \nu \):

\[
I_{\text{plb}}(\nu, \tau) \approx \int_{-\infty}^{\infty} i_{\text{out}}(t) e^{2\pi i \nu \tau} dt = I_{\text{cw}}(\nu) \otimes A(\nu) \quad (19)
\]

One can now appreciate why in most of the situation the classical convolution concept has given us the correct experimental spectral fringe broadening. However, this broadening does not represent any real carrier frequency distribution. The energy due to the same carrier frequency is just spread over a broader area because of incomplete super position of the replicated pulse by the spectrometer. Thus, while \( \delta \nu \delta t \geq 1 \) correctly represents the fringe broadening; \( \delta \nu \) represents energy spread, and not new frequencies. Thus, we do not have a fundamental limit on the measurement of the time-frequency-bandwidth as has been traditionally assumed. Only time-domain (real-space) analysis reveals this deeper physical understanding of the actual process behind the spectral fringe evolution in time and its broadening due to time integrated recording. We can now apply this understanding to obtain super-resolution for a pulse that contains a normalized carrier frequency distribution \( S(\nu) \) instead of a single carrier frequency assumed while deriving Eq.11a & b. The time integrated fringe broadening \( I_{\text{plb}}^{mf}(\nu) \) due to multi frequency (denoted as superscript \( mf \)) will now be given by the convolution of the single carrier frequency \( I_{\text{plb}}(\nu) \) and \( S(\nu) \):

\[
I_{\text{plb}}^{mf}(\nu) = S(\nu) \otimes I_{\text{plb}}(\nu) \quad (20)
\]

If \( a(t) \) is separately determined by non-collinear autocorrelation method [11] for very short pulses, or by direct intensity detection for longer pulses, \( I_{\text{plb}}(\nu) \) is known from Eq.11. Then \( S(\nu) \) can be recovered with super-resolution by de-convolution of \( I_{\text{plb}}(\nu) \) from \( I_{\text{plb}}^{mf}(\nu) \).

IV. SUPER-RESOLUTION BY HETERODYNE SPECTROSCOPY

As underscored in section two, a broad band detector designed with fast electronic response, gives an undulating current (beat frequency) that exactly equals to the difference between the two carrier frequencies. When the reference light beam is a single, CW (very narrow) frequency and the signal under study (pulsed or CW) contains more than one carrier frequencies, the detector current will consist of many undulating beat frequencies. We demonstrate by this experiment that pure amplitude modulation does not generate any new optical frequencies, inspite of the popularly held belief supported by spectral measurements whose coincidental correctness can be appreciated from Eq.18 and 19.

We take two continuously running semiconductor lasers around 1550 nm, an external cavity laser (EC-L) and a distributed feedback laser (DFB-L). The two laser beams are combined on a 30 GHz photo detector whose output signal is analyzed by a 25 GHz electronic spectrum analyzer (ESA). The EC-L is conveniently tuned to within 20 GHz of the line center of the DFB-L so the ESA can display the beat signal and the intrinsic combined line width of the two lasers. For our case, the intrinsic line width of the DFB-L is below 20 MHz. The ECL line width is below 100 KHz. The top picture of Fig.3 shows a sharp ~15 GHz beat line when both the lasers are running CW:

\[
I(t) = |d_1 e^{-j2\pi \nu_1 t} + d_2 e^{-j2\pi \nu_2 t}|^2 \quad (21)
\]

\[= d_1^2 + d_2^2 + 2d_1 d_2 \cos 2\pi (\nu_1 - \nu_2) t \]

The detector and the ESA are DC-coupled that drops \( d_1^2 \) and \( d_2^2 \). The ESA then recognizes the third term as a harmonic frequency of 15GHz and displays it as a sharp line.

Next, we amplitude modulated the DFB-L with the help of an external Lithium-Niobate Mach-Zehnder (LN-MZ) modulator. We have applied pseudo random communication data (fairly good square pulses) on the LN-MZ at various frequencies from 1 to 10 GHz. The bottom
picture of Fig.3 shows the case for 2.5 GHz pseudo random data. Notice that the FWHM of the beat line has remained essentially unchanged. This indicates that the carrier frequency content of the modulated AM pulses from the DFB-L is essentially the same as before. If time-frequency Fourier theorem were representing a real physical property of an AM signal, we should have recorded a 2.5 GHz FWHM spectral broadening of this beat line (note that the vertical scale is logarithmic). However, notice that there is a separate sinc² curve whose first zero is at 2.5 GHz (inverse of 0.4 ns). Since the function of an ESA (with memory and built-in algorithm) is designed to present all time varying currents received by it in terms of harmonic contents, the bottom picture separately shows that the photo current was turning on and off at 0.4 ns intervals (inverse of 2.5 GHz).

Figure 3. Output from an electronic spectrum analyzer (ESA) fed by the photo current from a high speed detector illuminated by the superposed light beams of two different frequencies. The top photo corresponds to two CW light beams separated by ~15 GHz, the beat frequency. The bottom photo corresponds to the external amplitude modulation of one of the lasers by 0.4 ns square pulses (2.5 GHz pseudo random data). The carrier frequency (beat) signal remains essentially unchanged, while the presence of AM is separately displayed as the Fourier transform of the square pulses (first zero at the 2.5 GHz location [4, 10]).

Thus, heterodyne spectroscopy of pulsed light can give us super resolution (actual carrier frequency content) even when the light is pulsed. If the temporal response time of the combination of ESA and detector is faster than the rise time of the pulse under study, one can also obtain a quantitative value of the pulse envelope given as the inverse transform of the square root of the ESA pulse spectrum. This can be appreciated from the Eq.22 where $d_2(t)$ represents one single square pulse

$$I(t) = |d_1 e^{-i 2\pi \nu t} + d_2(t) e^{-i 2\pi \nu t}|^2$$

$$= d_1^2 + d_2^2 + 2d_1 d_2(t) \cos 2\pi (\nu_1 - \nu_2) t$$

out of the pseudo random pulse train. As before $d_1^2$ is filtered out. The ESA memory and software is designed to recognize the harmonic signal 15GHz of the third term; but it also separately recognizes the presence of $d_2(t)$, which repeatedly interrupts the 15GHz harmonic signal. The ESA is designed to present this interrupting signal $d_2(t)$ as the square modulus of its Fourier transform. In our case $d_2(t)$ is a square envelope of width 0.4ns; the square modulus of its FT is a sinc² whose first zero is at 2.5GHz as is evident in the lower photo of Fig.3.

V. CONCLUSION

We conclude by underscoring that we need to give careful attention to the actual detection (interaction) processes in nature for proper interpretation of mathematics, even when it predicts the measured value correctly [1, 10]. We do not “see” light in the absence of interacting materials. The superposition effect for light becomes manifest only when detecting materials can respond to all the superposed light beams simultaneously, which is dictated by their
quantum mechanical behavior. In other words, the detecting dipoles do the summing of all the undulations induced on them by the superposed fields. The effects of those frequencies to which the detectors are not allowed to respond to by quantum conditions, will not be manifested even when the fields are superposed on the detector. These understandings when applied appropriately to pulsed light, one can achieve spectral super-resolution, which we interpret as determination of the actual carrier frequency content of a light pulse with arbitrary precision without being limited by the traditional ‘time frequency bandwidth’, $\delta v \delta t \geq 1$. This relation is only a corollary of the time-frequency Fourier theorem, a product of the half-widths of the two functions that form the Fourier transform pair; a rigorous proof uses Schwartz’s mathematical inequality theorem. Neither Fourier theorem nor the Schwartz’s inequality theorem represent any fundamental functional principle of nature, unlike say, Huygens-Fresnel principle, which accurately models diffraction and propagation behavior of all wave phenomena.

We have validated this assertion of super-resolution for classical spectroscopy analytically using a process of de-convolution of a function that we call pulse impulse response function (Eq.11), which is derived by using the knowledge of the pulse envelope. For heterodyne spectroscopy, we have actually demonstrated by experiment that quantitative information for both the carrier frequency content and the pulse envelope function can be determined separately with many orders of magnitude better precision than the traditional limit set by Fourier’s corollary, $\delta v \delta t \geq 1$.

VI. ACKNOWLEDGEMENT

A part of this research has been supported by the Nippon Sheet Glass Corporation.

VII. REFERENCES