Second Harmonic Imaging Microscopy

by Kraig Kumfer

While the study of optical phenomena that depend linearly upon the irradiance or associated electric field of light has generated a wealth of understanding of the physical world and powerful technologies to improve our lives, the study of nonlinear optical phenomena have provided new avenues towards understanding and instrumentation. This paper attempts to describe the physics of optical second harmonic generation (SHG) as it applies to the technique of second harmonic imaging microscopy (SHIM).

The interaction of light with matter is productively modeled as an interaction of an oscillating electric field (E-field) with a dielectric material. As a light wave passes through a dielectric, the E-field component will separate the positive and negative charges in the material. This charge separation establishes internal E-fields that oppose the driving field and generate forces that restore the separated charges back toward equilibrium. The E-fields that result from polarization of the dielectric are the source of scattering by a dielectric material. At optical frequencies, the oscillation of atomic electron clouds relative to their nuclei are the dominant source of dielectric polarization: more massive particles have moments of inertia that are too large to permit much physical displacement in the short period of the light's oscillation.

The electric polarization of a dielectric material can be well described by modeling the dielectrics as electrons connected to stationary nuclei by springs with spring constants that depend on the displacement of the electron. The induced dipole moment per unit volume is the electric polarization and is denoted by the vector **P**. A time-varying expression for **P** can be derived by modeling the forces exerted by the E-field and assuming a restorative force that is linear to the displacement (this is analogous to the simple harmonic motion of masses attached by an ideal spring); namely, for each electron:

$$m_e * d^2\mathbf{x}/dt^2 = q_e * \mathbf{E_0} \cos(\omega t) - k_E * \mathbf{x}(t), \text{ which can be integrated to}$$

$$\mathbf{x}(t) = q_e / m_e(\omega_0^2 - \omega^2) * \mathbf{E}(t).$$

Since the induced dipole moment of a unit volume with dipole density N is simply $N*q_e*\mathbf{x}$, substitution for **x** yields a useful expression for the electric polarization:

$$\mathbf{P}(t) = q_e*N*\mathbf{x}(t) = q_e^2*N / m_e(\omega_0^2 - \omega^2) * \mathbf{E}(t)$$

One consequence of this expression for **P** is that any E-field resulting from the electric polarization will oscillate with the same frequency as the driving E-field (i.e. $\mathbf{P}(t) \alpha \mathbf{E}(t)$). Therefore, this model predicts that the light scattered by a dielectric is of the same frequency as the incident light.

Expansion of the expression of the restorative force to include higher-order terms predicts a number of nonlinear interactions of materials with the strong E-fields of intense light. Just as real springs have spring constants that are not strictly linear with respect to displacement from equilibrium, the restorative forces that oppose the separation of charges are better described by a sum of linear and higher-order terms of the perturbing E-field. Using the same logical progression by which the relationship between **P**(t) and **E**(t) was described above (assuming only linear components of the restorative force), one can derive the relationship with respect to both linear and nonlinear terms of the restorative force. Schematically, the relationship between the magnitudes of **P**(t) and the various terms of **E**(t) is

$$P(t) = \varepsilon_0 * (\chi^{(1)}E + \chi^{(2)}E(t)E(t) + \chi^{(3)}E(t)E(t)E(t) + \ldots)$$

where ε_0 is the permittivity of free space and $\chi^{(n)}$ is the coefficient—called the dielectric susceptibility—that modifies the term that is the n-th order of E(t). Assuming that E-field varies such that $E(t)=E_0\sin(\omega t)$ and substituting into the above equation, the electric polarization can be rewritten as

$$P(t) = \varepsilon_0 * (\chi^{(1)}E_0\sin(\omega t) + \chi^{(2)}E_0^2(1 - \cos^2(\omega t)) + \chi^{(3)}E_0^3(\sin(\omega t))(1 - \cos^2(\omega t)) + \ldots)$$

and rearranged to

$$P(t) = \varepsilon_0 * (\chi^{(1)}E_0\sin(\omega t) + \chi^{(2)}E_0^2(1 - \cos(2\omega t))/2 + \chi^{(3)}E_0^3(3\sin(\omega t) - \sin(3\omega t))/4 + \ldots).$$

Notice that the 2nd-order component expands to include a term that varies at twice the frequency of the primary perturbing wave. This term predicts that a dielectric material could scatter light that is the 2nd harmonic of the incident light. The intensity of this scattered light should depend quadratically upon the intensity of the primary light. This effect, and the other effects described by nonlinear terms of E(t), are generally insignificant at the intensities of light usually encountered; however, laser techniques can generate sufficiently intense fields to make these effects observable. In fact, the first demonstration of 2nd harmonic generation (SHG) was reported in 1961,[1] only one year after the first report of an optical laser.[2] In addition to SHG, the above equation predicts a number of

other phenomena, including induction of a time-invariant E-field that depends upon $(E_0)^2$ and generation of a third harmonic with an intensity that depends upon $(E_0)^3$. Describing the extensive menagerie of nonlinear effects, however, is well beyond the scope of this paper: further description will be limited to what is immediately necessary to describe SHG.

An important consequence of the equation above is that the 2nd-order effects must vanish in a material that is isotropic or that has an inversion center. For simplicity, only the magnitude of the vectors **P** and **E** are dealt with in the treatment above. Considering the direction associated with these vectors demonstrates that the 2nd-order term must be zero for any material that is inversionally invariant, meaning that there is at least one point about which the material is symmetric. This can be seen when considering the effects of negating the magnitude of the incident E-field. If an even-ordered term were finite, then the equation above would describe some contribution from the E^2 term to the electric polarization **P** in whatever direction the equation describes. The equation would predict the same contribution from E-field component with negative magnitude. Because there is no special direction defined by the medium (i.e. it has an inversion center), the only way this term would not create an arbitrarily preferred direction would be if its value were zero. Therefore, 2nd order effects, such as SHG, cannot occur in an isotropic or other inversionally invariant material.

The treatment of the susceptibility coefficients thus far ignores that they are actually tensors and has treated them as scalars. While this simplification is adequate for many analyses, it fails to adequately describe interesting phenomena. The electric polarization susceptibility is more fully characterized as a tensor whose components relate the magnitude of electric polarization of a given direction in response to a driving disturbance polarized in a given direction. For example, the first-order susceptibility tensor $\chi^{(1)}$ for a given material describes the linear part of the medium's electric polarization in response to the primary E-field of the incident light. $\chi^{(1)}$ can be regarded as a tensor that has nine components $\chi^{(1)}_{ij}$, where i and j denote spatial directions such that each component is a coefficient describing the magnitude of the electric polarization, **P,** in the $\hat{i}$-direction that results from the $\hat{j}$-component of the primary wave's E-field. Because i and j can independently take the unit vectors $\hat{x}$, $\hat{y}$, and $\hat{z}$, $\chi^{(1)}$ has nine components. For example, the $\chi^{(1)}_{xx}$ tensor component describes the $\hat{x}$-component of **P** that results from the $\hat{x}$-component of the E-field of the incident light. In this most general form, the electric susceptibility tensor $\chi^{(1)}$ contains six components for which $i \neq j$ (e.g. $\chi^{(1)}_{xz}$ and $\chi^{(1)}_{xy}$). For isotropic materials, these six components vanish and the remaining three components are equal: this is why $\chi^{(1)}$ can be effectively simplified to a scalar for isotropic

media. For birefringent materials, however, at least some of these components are nonzero,[3] and it is these nonzero components that give rise to the phenomenon of optical activity.[4]

The 2nd-order susceptibility tensor $\chi^{(2)}$ describes the magnitudes of SHG polarizations in response to the E-fields of two incident light waves. Thus, $\chi^{(2)}$ has twenty-seven components ($\chi^{(2)}_{ijk,}$ where i, j and k can independently take $\hat{x}$, $\hat{y}$, and $\hat{z}$) that describe the magnitude of the electric polarization that oscillates at twice the incident wave's frequency, $\mathbf{P^{(2)}}$, in the $\hat{i}$-direction and that results from the $\hat{j}$-component of the of incident E-field and the $\hat{k}$-component of another incident E-field. These two incident fields may be, but need not be, the E-fields of the same primary wave. Depending upon the polarization of the incident light wave(s) and symmetry of the material that permits SHG, the scattered 2nd harmonic light emanating from the material will be restricted to particular polarizations or will not be permitted at all (i.e. some or all of the tensor components may be zero). Because the equation describing SHG—namely, $P^{(2)}_i = \chi^{(2)}_{ijk} E^{(1)}_j E^{(1)}_k$—does not distinguish the order in which $E^{(1)}_j$ and $E^{(1)}_k$ calculated, nine of the twenty-seven tensor components are redundant (i.e. $\chi^{(2)}_{xxy} = \chi^{(2)}_{xyx}$, $\chi^{(2)}_{xxz} = \chi^{(2)}_{xzx}$, etc.) and only eighteen terms are needed to fully describe this tensor for any given medium.[5] To my intuition, the spatial symmetries of a material define the relationships among its tensor components.

An important consideration of SHG that is important for SHIM, though not a directly related to the susceptibility coefficient concerns the interference among the generated 2nd harmonics. As a coherent primary wave propagates through a harmonophore (a medium that permits SHG) with real thickness, it will generate 2nd harmonics along the entire distance. A complication arises from the fact that all real dieletrics are dispersive: the generated 2nd harmonic wave will experience index of refraction $n_{2\omega}$, the fundamental will experience n_ω, and $n_\omega \neq n_{2\omega}$. Generated 2nd harmonic waves will be in phase with the fundamental but generally not with the previously or subsequently generated 2nd harmonic waves. If the harmonophore is sufficiently thin, the generated 2nd harmonics will be close enough in phase to mostly interfere constructively, but destructive interference will occur as the thickness is increased. The thickness at which the interference of the 2nd harmonics first reaches a maximum is called the coherence length (Z_c) and is related to the fundamental wavelength (λ_ω) and the two indices of refraction as follows:[6]

$$Z_c = \lambda_\omega / (4*|n_\omega - n_{2\omega}|).$$

The intensity of the generated 2nd harmonic reaches a local maximum when it passes through whole number multiples of Z_c. A practice called index matching permits coherent SHG through the length of harmonophore of any

thickness. This technique involves using a birefringement harmonophore such that the 2nd harmonic waves are scattered with polarizations that are orthogonal to both the propagation and polarization of the fundamental wave. If the material also has indices of refraction for perpendicular orientations such that they exactly compensate for the dispersion, then SHG can occur in phase through the thickness of the material. While phase matching is unlikely to be found in any given sample in a random orientation, it is useful for designing a system for optimal SHG. The concept of a coherence length, however, is generally applicable to any material and any sample to be imaged.

Second harmonic generation imaging microscopy (SHIM) takes advantage of several useful properties of 2nd harmonic generation to construct an image of the sample. The relatively long excitation wavelength used for SHIM results in less scattering of the illuminated light than most other microscopy techniques. Although liquids and solids are composed of scatterers at high density, the samples that are generally of interest for microscopists are not sufficiently homogenous to restrict constructive interference to the forward propagation of the beam. The inhomogeneities of the samples give rise to Reyleigh scattering, which redistributes energy from the intended target to the surrounding medium. Because this scattering depends upon $1/\lambda^4$, the longer wavelengths used for illumination in SHIM lead to greatly reduced scattering. This, in turn, permits a deeper penetration of illumination into the sample.

A 2nd useful property of SHIM is the small volume in which SHG is appreciable. Because the intensity of the generated 2nd harmonic depends upon the *square* of the intensity of the incident illumination, the intensity of the generated 2nd harmonic depends upon $1/r^4$, where r is the distance from the focal point of the illumination. This contrasts sharply with linear microscopy techniques, such as one-photon fluorescence microscopy, in which the generated signal depends upon $1/r^2$. Constraining significant SHG to a smaller volume about the focal point decreases the size of the resulting voxel. Thus, the nonlinear relation between illumination intensity and resulting SHG intensity increases the spatial resolution of the collected image.

The fact that SHG is only appreciable in relatively rare materials provides another useful property of SHIM: the rarity of harmonophores in a sample provides an endogenous source of contrast. This rarity is due to the requirements that a harmonophore must be noncentrosymmetric and must either be index matched or have a thickness that is near a multiple of its coherence length. The presence of a structure that permits SHG in a medium that generally does not allows that structure to result in a feature of the image that can easily be identified. The lack of endogenous sources of contrast appreciable by other microscopy techniques has required applying stains, dyes or

fluorophores to the sample to more easily identify features of the resulting image. While useful for identifying interesting structures, each of these approaches alters the sample in order to image it. In fact, many of the common approaches to increasing the contrast of biological samples require a fixation step that stops any biological processes. The contrast provided by the endogenous harmonophores allows their identification without first altering the sample. By imaging without applying exogenous components, the image is more likely to represent the structure of the unperturbed sample. This is particularly useful for imaging biological materials because it is possible to image a sample while it is still alive and responding to experimental conditions.

SHIM requires instrumentation similar to that used for multiphoton excitation fluorescence microscopy, including similar lasers, scanning mirrors, lenses, filters, detectors, computers and software. Generally, the laser is raster-scanned across the sample my means of two mirrors mounted on galvonometers in order to scan the focal point along the x- and y-axes. The laser then passes through the objective to focus the light on the sample, which may scatter a 2^{nd} harmonic. The scattered signal is collected by the objective lenses (to collect back-scattered light), a condenser lens (to collect forward-scattered light) or both. Placing a filter, such as a substrate coated with appropriately designed layers of thin films, permits transmission of the 2^{nd} harmonic signal while reflecting the undesired fundamental. The filtered signal can be detected by a high-speed sensor, such as a photomultiplier tube, and the resulting electronic signal interpreted and saved by a computer. By synchronizing collecting signals from the detector and the position of the galvonometers, the computer reassembles the saved pixels into a picture.

SHG creates some constraints on the instrument that are not significant for linear microscopy techniques. For instance, while SHG's dependence upon the square of the illumination intensity requires high-powered illumination, many samples will become ionized or otherwise damaged by the sustained application of intense illumination. By using a laser that emits light in short pulses, it is possible to obtain the high instantaneous power that is needed for SHG without delivering a time-averaged power high enough to destroy the sample. Commercially available lasers that generate pulses shorter than 100ps permit both the high instantaneous power and low time-averaged power. Also, because the scattered signal oscillates at twice the frequency of the illumination, collecting backscattered signal requires that the objective must be transparent to both the fundamental and 2^{nd} harmonics.

Controlling the laser and signal polarizations permits characterization of the symmetries of the harmonophores. By illuminating the sample with light that is polarized in a single direction and measuring the emitted SHG signal that is polarized in each direction, it is possible to empirically determine the values of the

relevant tensor elements of the susceptibility tensor. For example the element $\chi^{(2)}_{xyy}$ can be determined by illuminating the sample with light polarized in the ŷ-direction and detecting the SHG signal that is polarized in the $\hat{x}$ -direction. By extending this experimental approach, and fully using the instrument, it is possible to determine the overall symmetries of the unit cell of the crystalline harmonophore.[7] This information can be combined with independent models of how the harmonophore could be arranged in the sample.

SHIM is a particularly useful, emerging technique for imaging biological materials.[8] This technique has been successfully used to image and investigate the crystalline, supramolecular structures formed by collagen,[9] myosin[10] and microtubules[11] without the use of exogenous dyes in living specimens.

References / Endnotes

Unless otherwise indicated, all of the physics of SHG presented here was derived from the treatments of the topic presented in two books: Eugene Hecht's *Optics, 4th edition* (2002; Addison Wesley; ISBN 0805385665) and George C. Baldwin's An Introduction to Nonlinear Optics, 1st edition (1969; Plenum Press; ISBN 030620004X).
This treatment is not intended to exhaust any discussed aspect and practically ignores all physical, chemical and biological discoveries made by means of SHIM.

[1] Franken, PA et al. (1961). Generation of Optical Harmonics. *Physical Review Letters* **7**, 118-119.

[2] Maiman, TH. (1960). Stimulated Optical Radiation in Ruby. *Nature* **187**, 493-494.

[3] See Hecht, p 341, for a more authoritative statement to this effect.

[4] See Hecht, section 8.10, for a fuller explanation and references for further reading.

[5] See Baldwin, pp. 81-83.

[6] This treatment of coherence length is a conflation derived from Hecht (p641) and Baldwin (p76).

[7] DA Kleinman explicitly suggests a very similar, and more directed, experimental setups. Kleinman DA. (1962). Nonlinear Dielectric Polarization in Optical Media. *Physical Review* **126**, 1977-1979.

[8] For a useful review, see Campagnola PJ and Loew LM. (2003). Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. Nature Biotechnology 21:1356-1360.

[9] Freund I, Deutsch M and Sprecher A. (1986). Optical Second-harmonic Microscopy. Crossed-beam Summation, and Small-angle Scattering in Rat-tail Tendon. *Biophysical Journal* **50**, 693-712.

[10] The following report contains images that I find particularly spectacular. Plotnikov SV et al. (2005). Characterization of the Myosin-Based Source for Second-Harmonic Generation from Muscle Sarcomeres. *Biophysical Journal* **90**, 693-703.

[11] Campagnola PJ et al. (2002). Three-Dimensional High-Resolution Second-Harmonic Generation Imaging of Endogenous Structural Proteins in Biological Tissues. *Biophysical Journal* **81**, 493-508.