CALIBRATION OF TERAHERTZ SPECTROMETERS

Mira Naftaly and Richard A. Dudley

National Physical Laboratory, Teddington TW11 0LW, UK
Corresponding author: mira.naftaly@npl.co.uk

Abstract

Calibration methods for terahertz spectrometers are described. Frequency calibration is performed using an etalon and a gas absorption cell. Amplitude linearity is tested using a stack of Fresnel reflectors. The correct determination of the instrument dynamic range and signal-to-noise ratio is also discussed.

1 Introduction

With the growth of terahertz research and applications, a wide variety of systems and detectors are being employed to carry out studies in spectroscopy and imaging. In particular, the terahertz time-domain spectrometer (THz TDS) [1] has emerged as a key measurement device for spectroscopic investigations in the frequency range of 0.1-3 THz. To date, almost every type of material has been studied using THz TDS, including semiconductors, ceramics, polymers, metal films, liquid crystals, glasses, pharmaceuticals, DNA molecules, proteins, gases, composites, foams, oils, and many others.

Measurements with a TDS are made in the time domain; conversion from the time domain data to a frequency spectrum is achieved by applying the Fourier Transform, calculated numerically using the Fast Fourier Transform (FFT) algorithm. As in many other types of spectrometer, THz TDS requires that the sample data be referenced to similarly acquired data with no sample present. Unlike frequency-domain spectrometers which detect light intensity and measure absorption spectra, a TDS records both amplitude and phase information, and therefore yields both the absorption coefficient and the refractive index of the sample material.

The analysis of the data from a spectrometer relies on the twin assumptions that: a) the frequency scale is accurate; and b) the measurement of light amplitude or intensity is linear. The frequency scale of a THz TDS is derived from the displacement of the delay line; via FFT, positioning errors may give rise to frequency errors that are difficult to quantify. The measurement of the field amplitude in a THz TDS is required to be linear with a dynamic range of the order of 10 000. Therefore both frequency and linearity calibration are an important part of the design and maintenance of THz spectrometers.

2 THz TDS

Figure 1 is a schematic drawing of the NPL THz TDS system. It is a commonly used configuration incorporating a femtosecond laser, four off-axis parabolic mirrors, a biased GaAs emitter, and electro-optic detection with a ZnTe crystal and balanced photodiodes. Measurements are carried out in dry air in order to eliminate water absorption lines from the recorded spectra. Figure 2 shows a typical time-domain trace and the associated THz spectrum.

The examined samples are placed in the collimated part of the THz beam. The absorption coefficient (α) and refractive index (n) of the sample are calculated by comparing THz transmission through the sample with that through free-space, or preferably through two different thicknesses of the sample material, by using the equations [1]:

$$\alpha(\nu) = -\frac{2}{d_1 - d_2} \ln\left[\frac{E_2(\nu)}{E_1(\nu)}\right] \qquad (1)$$

$$n(\nu) = 1 + \frac{c\left[\phi_2(\nu) - \phi_1(\nu)\right]}{2\pi\nu(d_1 - d_2)} \qquad (2)$$

where $E_{1,2}(\nu)$ and $\phi_{1,2}(\nu)$ are the amplitude and phase of the THz field at the frequency ν, and $d_{1,2}$ are the sample thicknesses.

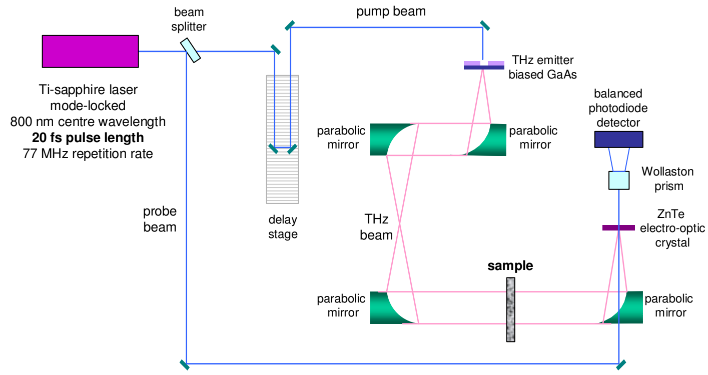

Figure 1. A schematic drawing of the THz TDS employed at NPL.

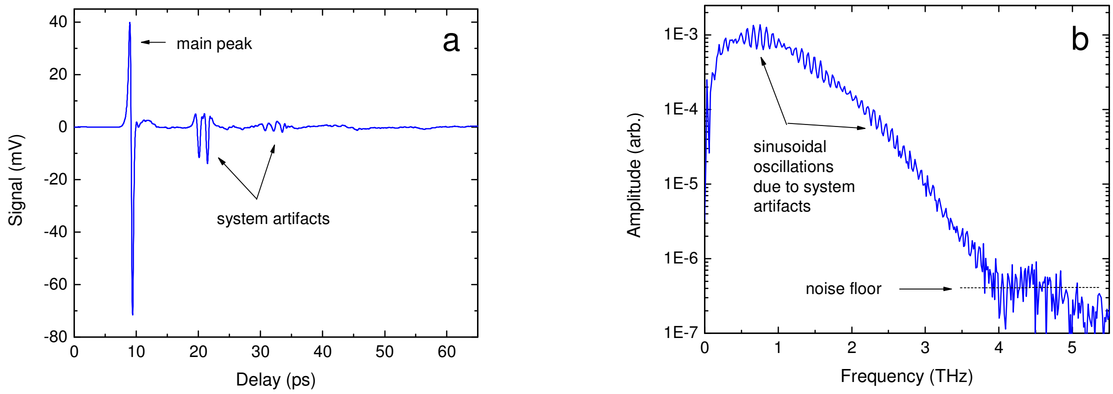

Figure 2. a) A typical time-domain trace. b) Associated THz spectrum calculated by FFT.

3 Frequency calibration using a CO gas cell

Gas absorption lines offer an obvious and readily available frequency standard in the THz band, due to their well-known frequencies and narrow linewidths. Indeed, the most widely used means of THz frequency verification is atmospheric water vapor which possesses many strong lines. However, many of the lines are doublets and triplets, and therefore require very high (sub-GHz) frequency resolution to define their peak maxima and profiles. The spacing of lines is particularly dense at higher frequencies above 2 THz, where the reduced signal-to-noise ratio and dynamic range of a TDS make accurate measurements far more difficult.

An examination of the HITRAN database of gas spectra [2] shows that carbon monoxide (CO) is particularly suitable as a THz frequency standard, having strong absorption lines in the range of 0.2-3 THz, spaced at equal intervals of 114 GHz with a distinctive amplitude envelope. The absorption spectrum of CO is well documented owing to its significance as a ubiquitous interstellar molecule and an important trace constituent of planetary atmospheres.

Figure 3 shows examples of gas cells used at NPL. The cylindrical aluminium cells have PTFE windows, which are transparent at THz, and are capable of containing gas at pressures of up to 7 bar. Figure 4 plots the absorption spectrum of CO measured at the pressure of 2 bar [3].

Frequency calibration of the THz TDS using CO absorption spectrum is depicted in Figure 5, which plots the frequencies of the absorption peaks and their deviation from the HITRAN database. As expected, peak frequencies are equally spaced with the peak separation of 114 GHz. The mean deviation of the data from database is 2 GHz, which is comparable with the system resolution in this experiment of 1.5 GHz, indicating that the frequency error of the TDS system is of the order of its resolution.

Figure 3. Gas cells for frequency standard.

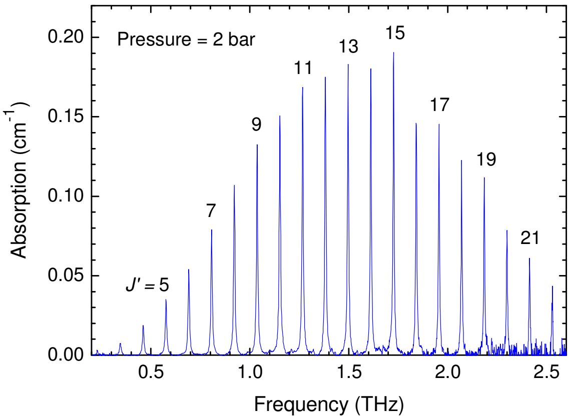

Figure 4. Absorption spectrum of CO (carbon monoxide).

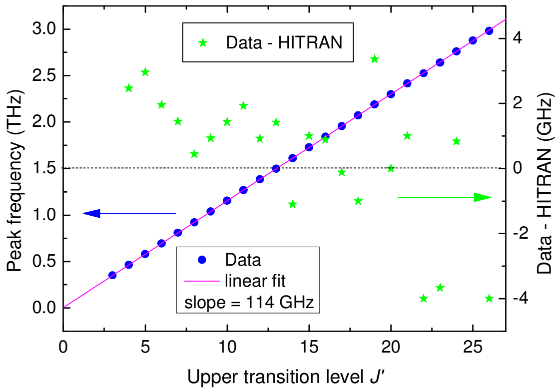

Figure 5. Left: Measured frequencies of CO absorption lines. Right: Deviation from the HITRAN database.

4 Frequency calibration using etalon

It is preferable if a frequency calibration standard should produce regularly spaced peaks across the entire THz band, of a uniform size and with a well-defined profile. This can be achieved by employing an etalon, which also has the advantage of being inexpensive, small, and convenient to use. The method utilizes the echoes produced by multiple reflections in thin plane-parallel samples inserted in the THz beam [3,4]. The etalon consists of a silicon wafer of a common type used in semiconductor industry, which for this purpose must be undoped (high resistivity) and optically polished, as seen in Figure 6. Undoped silicon has negligible absorption in the THz band; and due to its high refractive index of 3.42, the faces provide sufficient reflectivity to perform as an etalon. Figure 7 depicts the measured transmission spectrum of a Si wafer etalon together with the model.

The peaks and troughs in the etalon spectrum occur at frequencies given by [5]:

$$f_N = \frac{c}{2nl} N \qquad (peaks) \tag{3}$$
$$f_{N+1/2} = \frac{c}{2nl}(N + \frac{1}{2}) \qquad (troughs)$$

where n is the refractive index of the etalon material and l is its thickness. The integer N is the order of the peak. The amplitude transmission of an etalon as a function of frequency f can be calculated from [5]:

$$T = \left(1 + F\sin^2\frac{\delta}{2}\right)^{-1/2} \quad ; \quad F = \frac{4R}{(1-R)^2} \tag{4}$$
$$R = \left(\frac{n-1}{n+1}\right)^2 \quad ; \quad \delta = 4\pi nl\, f\,/\,c$$

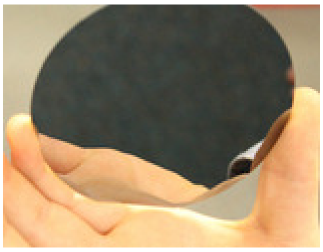

Figure 6. A silicon wafer etalon.

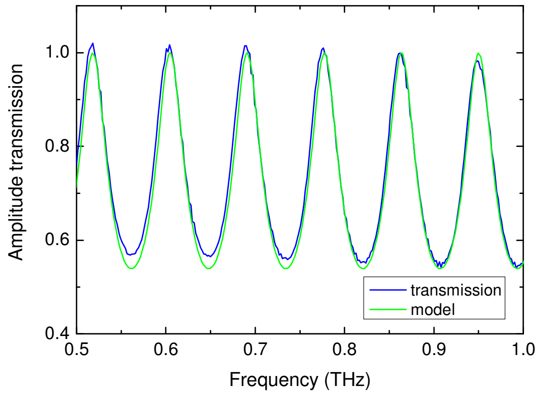

Figure 7. Measured transmission of a Si wafer etalon and model calculated from Eq. 4.

The peak/trough frequencies provide frequency calibration, while the transmission spectrum can be used to verify the measurement of the spectral amplitude profile.

The simplest and most straightforward method of analyzing the spectral data for the purposes of frequency calibration is as follows. Note the frequencies of the peaks and troughs in the transmission spectrum of the etalon obtained. Then calculate the expected peak/trough frequencies of the etalon from Eq. 3. A comparison of the two data sets will yield the difference between the measured and expected frequency value for each peak/trough. It is helpful to display the results by plotting these differences, i.e. the frequency errors, as a function of the etalon peak/trough position, as shown in Figure 8. Such a plot reveals any systematic frequency errors, as well as the digitizing errors and the noise in the data. It also helps to identify the band over which frequency measurements are valid within a defined uncertainty.

As seen in Figure 8, the frequency errors are distributed evenly around zero, confirming that there is no systematic frequency error in the system and that the errors arise from noise. The magnitude of the errors at frequencies below 2 THz is comparable with the system resolution of 1.5 GHz. The errors increase at higher frequencies, as the dynamic range of the TDS decreases and the system approaches its noise floor. The average magnitude of errors demonstrates the accuracy of frequency measurement over different frequency bands.

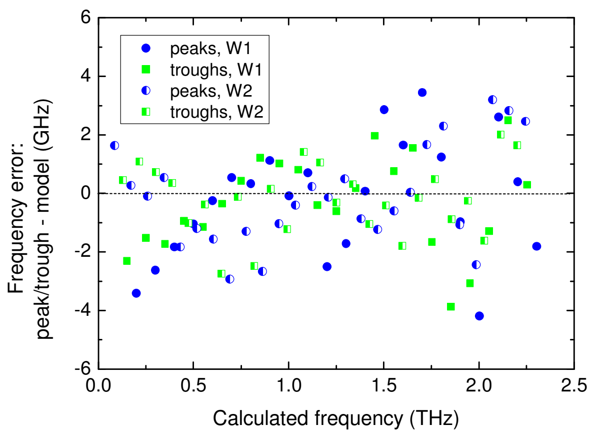

Figure 8. Differences between the measured and calculated peak/trough frequencies for two different Si wafer etalons.

5 Linearity calibration

In order to have confidence that the measured absorption spectrum of the sample material is correct, it is necessary that the amplitude scale of the spectrometer be linear; or if not, that its deviation from linearity be known and quantifiable. I.e. the recorded signal must be proportional to the THz field across the entire dynamic range of the system. However, in practice the linearity of THz data acquisition is seldom verified and no published literature exists.

Testing amplitude linearity of a THz TDS requires a calibration device whose loss is constant across the THz bandwidth and is capable of being varied accurately in equal steps spanning the dynamic range of the system being tested. The preferred solution is to employ as "loss elements" a stack of optically flat silicon plates [6]. As noted above, high-resistivity silicon has negligible absorption and dispersion in the THz band, and therefore transmission loss through a silicon plate is due solely to Fresnel reflection. The loss produced by a stack of plates separated by air gaps is multiplicative, to the power equal to the number of plates in the stack. The mount block is designed so as to allow easy alignment of the plates parallel to each other and normal to the incident THz beam, as seen in Figure 9. In the NPL test kit, the silicon plates were chosen to be 3 mm thick with 3 mm air gaps between them: the plates are robust and easy to handle; the thickness is sufficient to prevent the formation of standing waves; and the total length of the device is conveniently short.

Single-pulse transmission loss through a stack of N plates, as measured by a THz TDS, for field amplitude E is given by [5]:

$$E_N / E_0 = (1-R)^N \quad ; \quad R = [(n-1)/(n+1)]^2 \qquad (5)$$

where R is the Fresnel reflectivity, and n is the refractive index. For Si, n = 3.42 and R = 0.3. Therefore for a THz TDS, the transmission factor per each Si plate is E_1/E_0 = 0.7.

The simplest method of linearity calibration is to test the linearity of the time-domain signal. This involves plotting the amplitude of the time-domain peak maximum against the number of Si plates in the beam path, and gives a single frequency-averaged result. For a linear system, a semi-log plot of the data will be linear with a slope of 0.7, as shown in Figure 10. It is seen that at low signal levels where the system approaches its noise floor there is a slight deviation from linearity. The contribution of noise causes a positive deviation because the peak amplitude must always be positive, and therefore the average of the noise has a positive value, so that the error increases with the relative noise.

A more detailed, frequency-resolved method of testing the linearity of a TDS involves calculating THz spectra. The amplitude at chosen frequencies is plotted against the number of silicon plates in the beam path, as seen in Figure 11a. As previously, the semi-log plots are expected to be linear with a slope of 0.7. Increased positive errors at higher frequencies indicate the limits of the dynamic range of the system. The data at 3 THz show the strongest deviation, because at that frequency the system is close to its noise floor, with the dynamic range of <10.

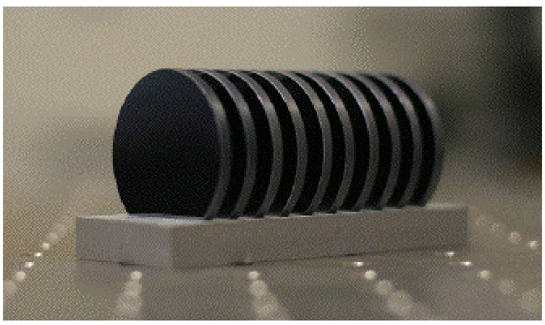

Figure 9. A stack of silicon plates for calibrating the linearity of amplitude measurement.

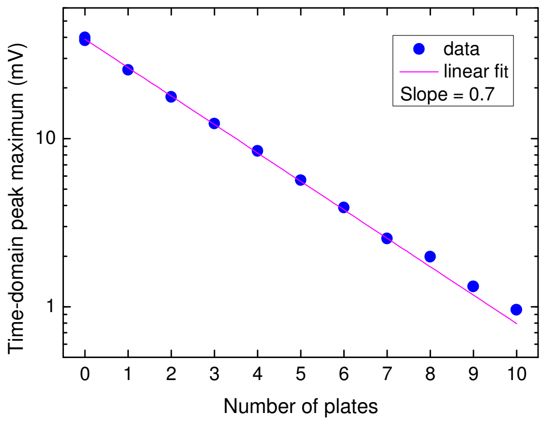

Figure 10. Amplitude linearity test of the time-domain peak maximum (frequency-averaged).

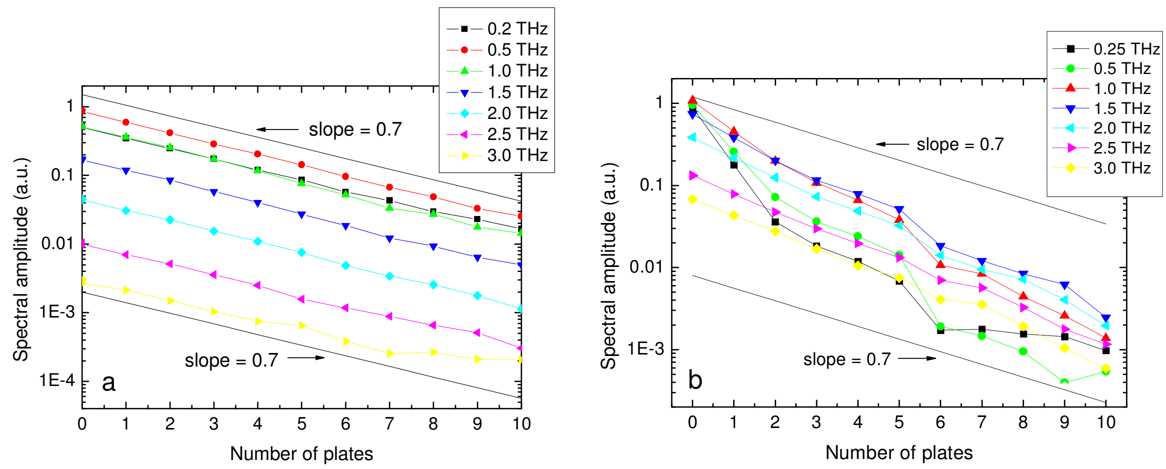

Figure 11. Frequency-resolved linearity tests on a THz TDS: a) after re-alignment, showing satisfactory linearity; b) before re-alignment, showing severe non-linearities.

It is important to note that the linearity of a THz TDS should not be assumed, but ought to be experimentally verified. Figure 11b presents an example of a TDS which is severely non-linear, especially at low frequencies, owing to issues of alignment. A possible cause of the nonlinearity may be variations in the THz-probe overlap on the detector crystal (ZnTe): the effect is strongest at low frequencies since the diffraction-limited waist of the THz beam decreases with frequency. Figure 11a shows results from the same TDS after re-alignment where the probe beam was defocused. This resulted in nonlinearities being drastically reduced, demonstrating that simple measures, such as attention to correct alignment, can significantly improve the linearity of amplitude measurement of a THz TDS system.

Average-power detectors, such as pyroelectric sensors and Golay cells, can also have their linearity tested using a stack of silicon plates. For such detectors the Fresnel loss per plate must take into account multiple reflections, calculated from [5]:

$$\frac{I_1}{I_0} = (1-R)^2 \sum_{n=0}^{\infty} R^{2n} = \frac{1-R}{1+R} \tag{6}$$

The transmission factor per Si plate for an average-power detector is therefore $I_1/I_0 = 0.54$. Figure 12 plots the results of linearity tests for a pyroelectric sensor, which is seen to be linear within its dynamic range; and for a Golay cell, which is not.

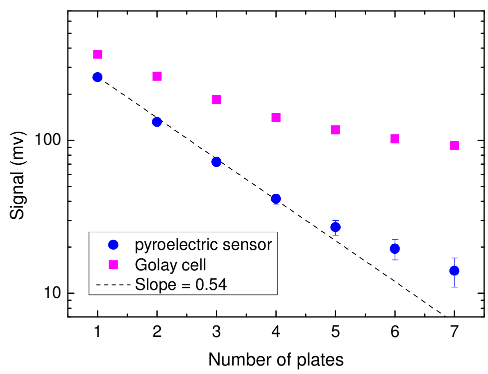

Figure 12. Power linearity tests on a pyroelectric sensor and a Golay cell.

6 Dynamic range and signal-to-noise ratio of a THz TDS

When describing the performance of a THz TDS system, it is common practice to quote its dynamic range (DR) and/or signal-to-noise ratio (SNR). However, there are many disparate and often contradictory methods of defining these quantities in relation to a THz TDS, and no commonly agreed standard exists. The situation is made more confusing by the fact that the value of DR or SNR is usually quoted without mentioning the method by which it is calculated.

The SNR and DR of a system that measures amplitude, such as a THz TDS, are defined as:

$$SNR = \frac{mean\ \ magnitude\ \ of\ \ amplitude}{standard\ \ deviation\ \ of\ \ amplitude} \quad (7)$$

$$DR = \frac{maximum\ \ magnitude\ \ of\ \ amplitude}{RMS\ \ of\ \ noise\ \ floor} \quad (8)$$

SNR and DR have different, although complementary, implications on system performance. SNR indicates the minimum detectable signal change; while DR describes the maximum quantifiable signal change.

In the case of TDS, SNR and DR may be evaluated either with respect to the time-domain trace or to the spectrum calculated via FFT [7]. The noise in the time-domain trace is evaluated directly from experimental data. However, there is no simple analytical relationship between the values of SNR and DR in the time-domain data and those in the FFT spectrum. Therefore the method of calculating SNR and DR must be directly related to the type of measurement being carried out, i.e. whether it uses time-domain data, or FFT-derived spectra.

Figure 13 shows a typical time-domain trace, which in this case is a mean of 9 runs, together with the standard deviation (SD) of the data. It is seen that the SD varies with the signal, and is largest where the signal is strongest. Also in Figure 13 is plotted the SNR calculated in the accepted way as the ratio of mean signal to its SD. The SNR fluctuates strongly from point to point, and is therefore a poor indicator of the measurement accuracy or amplitude resolution.

A better estimate of the SNR and the DR of the time-domain data can be obtained by evaluating the standard deviation of the peak signal and that of the noise in the absence of THz signal. The SNR is then given by the ratio of the mean peak maximum signal to its SD; while the DR is the ratio of the mean maximum peak to the SD of noise. In the presented data set the SNR = 120, while the DR = 9900. Note that here the DR is a factor of 80 greater than the SNR.

This method of evaluating the DR is supported by the fact that it is possible to measure THz transmission through strongly attenuating samples which reduce the measured amplitude by a factor of 1000 or more.

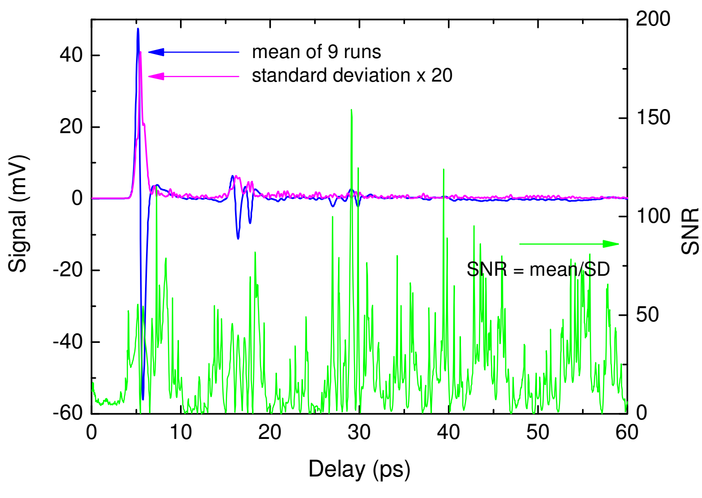

Figure 13. Left: A typical time-domain trace (mean of 9 runs), and the SD of the data.
Right: SNR, calculated as the ratio of the mean value and its SD.

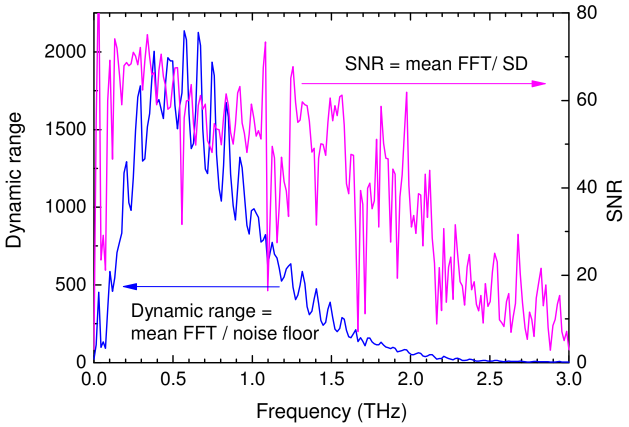

Figure 14. Left: The DR of the amplitude spectrum, calculated as the ratio of the amplitude and the noise floor. Right: The SNR of the amplitude spectrum, calculated as the ratio of the amplitude and its standard deviation.

Similar considerations apply to calculations of SNR and DR of spectra obtained via FFT. Figure 14 (left ordinate) plots the dynamic range of the spectral measurement, calculated as the ratio of mean amplitude and the noise floor (see Figure 2b). The right ordinate of Figure 14 plots the SNR, given as the ratio of mean amplitude and its standard deviation. It is seen that both the dynamic range and the SNR vary strongly with frequency.

It is a widely accepted custom to quote the maximum value (in this case ~2000) as the DR of a THz TDS. Provided that the frequency dependence of the DR is borne in mind, this approach is justifiable to a degree, because the great majority of THz TDS produce similar spectral profiles. Nevertheless, it would be preferable to quote the DR values at a range of frequencies, or to provide a DR frequency profile. In performing spectral measurements, great care must be exercised because the DR drops steeply with frequency, which limits the type of samples that can be examined.

As in the case of time-domain data, the SNR is much smaller than the DR over most of the frequency range. In the presented data, the maximum SNR is ~70, i.e. a factor of 30 lower than the DR. Note that neither the DR nor the SNR of spectral amplitude is directly related to those of the time-domain data (see Table 1).

Table 1. SNR and DR values of the sample data

	SNR	DR
Time-domain data	120	9900
FT amplitude spectrum (frequency dependent)	Maximum: 70	Maximum: 2000

The DR and SNR of a THz TDS system must be evaluated in relation to the type of measurement being carried out. If the measurement directly utilises the time-domain data, then the DR and SNR must be derived from that data. If the measurement concerns spectroscopic data, then the DR and SNR must be calculated from the FFT amplitude spectra. It is important to bear in mind that there is no simple analytical relationship between the DR and SNR values of the time-domain data and those of the amplitude spectrum. It is therefore important to avoid confusion and always to define clearly the parameters referred to and the method of calculating them.

An important consequence of the typical values of the SNR and DR is that the large DR allows the examination of strongly attenuating samples, while the much lower SNR limits the accuracy and amplitude resolution of these measurements. Moreover, in the time domain the DR of a THz TDS is much larger than in the frequency domain and can be fully utilised for the measurement of low-transmission samples. In contrast, in the frequency domain the variable DR limits the usable spectral bandwidth for lossy samples.

Acknowledgements

Financial support for work at NPL was provided by the National Measurement Office, an Executive Agency of the Department for Business, Innovation and skills.

References

1. M.C. Beard, G.M. Turner, C.A. Schmuttenmaer, "Terahertz spectroscopy", *j. Phys. Chem.,* **106** (2002) 1746-7159.
2. The HITRAN database of gas absorption spectra: http://www.cfa.harvard.edu/hitran/docs.html.
3. M Naftaly, RA Dudley, JR Fletcher, F Bernard, C Thomson, Z.Tian, "Frequency Calibration of Terahertz Time-Domain Spectrometers", *J. Opt. Soc. Am. B*, **26/**7 (2009) 1357-1362.
4. M Naftaly, RA Dudley, JR Fletcher, "An etalon-based method for frequency calibration of terahertz time-domain spectrometers (THz TDS)", *Opt. Commun.*, **283** (2010) 1849-1853.
5. Guenther R., *Modern Optics*, John Wiley & Sons, Inc., 1990.
6. M Naftaly, R Dudley, "Linearity calibration of amplitude and power measurements in terahertz systems and detectors", *Opt. Lett.*, **34/**5 (2009) 674-676.
7. M Naftaly, RA Dudley, "Methodologies for determining the dynamic ranges and signal-to-noise ratios of terahertz time-domain spectrometers", *Opt. Lett.*, **34/**8 (2009) 1213-1215.