An instrument for the acquisition and analysis of the nonlinear dielectric spectra of biological samples
A. Jones, J.J. Rowland, A.M. Woodward and D.B. Kell
Transactions of the Institute of Measurement and Control 1997; 19; 223
DOI: 10.1177/014233129701900502

The online version of this article can be found at:
http://tim.sagepub.com/cgi/content/abstract/19/5/223
An instrument for the acquisition and analysis of the nonlinear dielectric spectra of biological samples

by A. Jones*, BSc, PhD, J. J. Rowland*, MSc, CEng, MIEE, A. M. Woodward†, BSc, PhD, and Prof D. B. Kell‡, MA, DPhil

Nonlinear dielectric spectroscopy is, potentially, a powerful non-invasive method of monitoring biological cell activity, with possible applications in biotechnology and medicine. This paper describes the design and implementation of an integrated instrument for obtaining and analysing the nonlinear dielectric spectra of biological cell suspensions. The instrument is based on an IBM-compatible personal computer equipped with an AT&T digital signal processor and 16-bit analogue I/O. It allows collection of resultant power spectra in real time during optimal time sweeps over a wide range of excitation voltages and frequencies. Further analysis of the sets of spectra is undertaken by means of multivariate calibration. Examples are included of results obtained and it concludes with a brief discussion of the present limitations of the method.

Keywords: Nonlinear dielectrics; instrumentation; multivariate calibration; fermentation.

1. Background

Linear dielectric spectroscopy is a well established technique for characterising materials on the basis of the dependence of the permittivity of the material on the frequency of an applied sinusoidal electric field. In biotechnology, for example, this has been exploited to form the basis of the 'flugmeter' biomass monitor‡ (Beving et al, 1994; Davey et al, 1992, 1993; Davey and Kell, 1995; Harris et al, 1987; Kell et al, 1990; Markx et al, 1991) that measures the mass of live cells in a suspension and is used to monitor fermentation processes in the brewing and pharmaceutical industries. In linear dielectric spectroscopy, narrow-band filters are applied so as to exclude components of the response signal at frequencies other than the excitation frequency.

Nonlinear dielectric spectroscopy, with which this paper is concerned, is based on the detection and analysis of harmonics generated, upon the application of a sinusoidal electric field, as a result of nonlinearities in the medium being studied. There is established theory that relates such nonlinearities to changes in enzyme conformation within cell membranes, and consequently to changes in, for example, the metabolic state of the cells (Kell et al, 1988; Westerhoff et al, 1988; Astumian and Tsong, 1988). Previous experimental work (Woodward and Kell, 1990, 1991b; McShea et al, 1992) confirmed the existence of the effects that had been predicted from the theory. A full description of the technique and the extensive experimental work that has been undertaken to establish its capabilities is given in Woodward et al (1996).

This paper describes the design, development and validation of an integrated instrument that incorporates measurement and analysis functions to enable measurement of various aspects of cell activity based on the principles of nonlinear dielectric spectroscopy.

2. Measurement requirements

The requirement was for a versatile research tool that would permit automated data acquisition and analysis in the course of extensive experiments aimed at establishing the relationship between the nonlinear response and various parameters of cell activity. The instrument was to provide an automated means of data acquisition and analysis. It would apply a sinusoidal excitation voltage to a pair of electrodes, performing a frequency sweep at each of a range of different excitation voltage amplitudes, and would sample the resulting voltage waveform at the monitoring electrodes; at each combination of excitation amplitude and frequency it was to compute and store the power spectrum of the response. These sweeps were to be repeated for the duration of the experiment, often many hours.

The Nyquist criterion (eg, Beauchamp and Yuen (1980) or Higgins (1990)) requires the sampling rate to be at least twice the frequency of the frequency component of interest. Hence, to detect, for example, the fifth harmonic, the sampling rate must be at least ten times the excitation frequency. The use of a fixed sampling frequency that satisfies the Nyquist criterion over the whole range of excitation frequencies and the harmonics of interest implies an undesirably large storage requirement at all but the highest excitation frequency, unless a computationally cumbersome decimating filter is used. We therefore chose to scan the sampling frequency so that it was a fixed multiple of the excitation frequency; a sufficiently large multiple (usually 32 times) was used that aliasing, even of undesirable
noise components, was not observed in the response. The instrument was to be capable of an excitation frequency range of at least 1–1500 Hz and an excitation voltage amplitude range of at least 0.1 to 1.5 V, although many experiments would use smaller ranges. A subsequent analysis phase would process and interpret the stored sets of spectral data.

3. Data acquisition and pre-processing

Spectrum analysis involves calculating a Fast Fourier Transform (Weaver, 1989) of an array of sampled points. Fourier methods assume that if the waveform is periodic then the available data characterise an integral number of periods, in other words that the waveform is commensurate with the buffer. If this is not the case, the transform will 'see' potentially large discontinuities at the extremities of the buffer and the spectrum will be distorted. The general solution to this problem is to multiply the samples in the buffer by corresponding values of a windowing function (Harris, 1978) that de-emphasises the samples at the extremities of the buffer and therefore removes the discontinuities.

In the present work, scanning the sampling frequency along with the excitation frequency ensures that the fundamental and its harmonics are always commensurate with the buffer. The only spectral data that are recorded are the amplitudes of the harmonics of interest (normally the first five), and windowing is not therefore necessary. Because this processing step needs to work on an integral number of waves, there is a constraint on the time taken to calculate each power spectrum. Ideally, this calculation should be completed in less than the time taken to acquire one wave, otherwise there is a need to re-synchronise acquisition with the excitation waveform, with corresponding loss of time resolution.

3.1 The hardware

A four-terminal electrode system, as described in Woodward and Kell (1990) and Woodward et al (1996) is normally used. The outer electrodes are driven by the excitation voltage waveform, with the response (voltage) signal sampled across the inner pair via a differential amplifier. This is in contrast to conventional four-point-probe resistivity measurements (Kell, 1987) on, for example, semiconductor materials, where a controlled current is applied to the drive electrodes to minimise effects due to the interface between those electrodes and the material being studied. With cell suspensions we have found that the nonlinear effects of interest are not observed if current drive, rather than voltage drive, is applied to the outer electrodes. Discussion of this point is outside the scope of this paper but the interested reader is referred to Woodward and Kell (1995).

Generation of the excitation signal along with sampling, processing and storage of the response waveform is undertaken by a suitably equipped IBM-compatible PC. To fulfil the signal processing requirements identified above, a digital signal processor (AT&T DSP32C) was used, carried on a Burr-Brown ZPB34 ISA bus expansion board. The DSP32C has built-in parallel and high-speed serial ports, which are directly controlled by registers on the processor. In addition, DMA (Direct Memory Access) to and from the DSP32C memory through the parallel port can be controlled by an external device. Using this facility, the PC can read and write DSP32C memory space. The serial port provides the connection to an analogue I/O board.

The analogue I/O board is a Burr-Brown ZPD1001-003. This carries two DACs and two ADCs, clockable at sampling rates up to 150 kHz. This maximum sampling rate gives the possibility of generating an excitation sine-wave approximation of 32 points up to a frequency in excess of 4.6 kHz, well in excess of the sampling frequency required for the highest frequency at which response harmonics are observed (Woodward and Kell, 1990). All four converters have sixteen bit precision over the range –3 V... +3 V, fulfilling the requirement identified by Woodward and Kell (1991a).

The clocking signal for the converters is generated by an 8254 programmable timer, also residing on a PC expansion board. This is configured to output a train of pulses to trigger the analogue converters at a sampling rate rate that is controlled by the PC writing a value to one of the timer registers in PC I/O address space. By this means a constant ratio between excitation frequency and sampling frequency can be maintained, ensuring that, throughout the swept frequency range, complete cycles at the fundamental frequency are generated and sampled at uniform time resolution, satisfying the Nyquist criterion and obviating the need for windowing of the sampled signal.

Fig 1 shows the hardware used in the acquisition system. An alternative two-electrode arrangement has also been investigated, in which the excitation voltage is applied across two electrodes and the current in the circuit is sampled via a current sensing-amplifier to

---

<sup>2</sup>Burr-Brown Corporation, PO Box 11400, Tucson, AZ 85734, USA.
provide the response signal. Results obtained with this configuration have been found to be comparable with those for the four-point arrangement.

3.2 The software

Two sets of software are needed to generate a sweep of frequency and amplitude and to acquire and store the data from the resulting power spectra. A program running on the PC interacts with the user, the disks, the sampling rate generator and the DSP processor. The second program runs on the DSP board to provide high-speed data acquisition and pre-processing and passes the spectral data to the PC. Both sets of software were written using the C programming language supplemented, where needed, with assembly language.

Communication between the PC and the DSP board is supported by the MS-DOS device driver, "ZPB32DRV" from Burr Brown. However, this driver provides the PC with only low-level facilities, such as management of the DMA transfers between PC and DSP, DSP32C register access, and floating-point format conversion. Special provision had to be made for synchronisation between the two processors and was achieved by using one bit in a shared register as a semaphore, the operation of which is summarised in Table 1. An area of shared memory, protected by the semaphore, also had to be established in such a way that the compiler would not transparently allocate code or data to it. This was achieved by using the 'memory map' specification file facility of the AT&T compiler (AT&T, 1988a, 1988b), which allows an area of memory to be excluded from being allocated automatically by the compiler.

In deciding which process is to be 'in control' of the experiment the major considerations are memory usage and load sharing. In terms of code space, the master process requires more space; in terms of data space, the DSP program requires sample input buffers, a wave output buffer, plus space for the pre-processing. Since the ZPB34 only has 64 K of memory available (without expansion), the memory requirements suggest that the PC should be the master. In terms of load, the DSP processor has the stronger real-time constraints, as it has to complete preprocessing of one data block from the analogue board before the next arrives, otherwise it has to skip blocks, leading to longer overall acquisition times. These considerations led to the PC being deemed 'master'. Use of the 'xcc' package, which handles the shared memory requirement itself and makes the PC 'slave' to the DSP processor, was considered but was rejected because of its demands for DSP memory space. The master/slave program interface itself was implemented, like many communication systems, as a layered model.

The acquisition software takes advantage of the fact that the analogue I/O runs under DMA and interrupt control, so that preprocessing of the most recent data can be performed while the next set of samples is being acquired. In this arrangement the raw samples are copied to a temporary area, a very fast operation relative to the sampling rate, so the next acquisition cycle can start immediately, thus retaining synchronisation between the excitation waveform and the acquired waveform, which therefore remains commensurate with the buffer. In this way, acquisition and calculation are carried out concurrently and the data preprocessing has a full buffer acquisition period in which to complete.

The overall system is controlled from a configuration file; this is well suited to our present usage requirements but provides for the easy future addition of a graphical user interface. The configuration file allows control of all excitation parameters:

- The length of the buffer for the analogue output channel and hence the number of points per cycle of the excitation waveform,
- The number of samples per analogue acquisition buffer (the greater the ratio of acquisition buffer length to excitation buffer length, the greater is the frequency resolution of the resulting spectrum),
- The number of response harmonics that are to be recorded,
- The number of times the input is to be averaged into the acquisition buffer,
- The number of times the complete frequency/amplitude sweep is to be performed,
- The minimum allowed time, in seconds, between the start of frequency/amplitude sweeps (if a sweep completes in less than this time, a delay will be inserted before the next sweep starts),
- The list of excitation frequencies (for each amplitude value, all of these frequencies will be sequentially sampled),
- The list of excitation amplitudes,
- The ADC and DAC voltage scalings, that specify the number of ADC/DAC counts that correspond to given input and output voltages,
- The properties of any amplifiers on the ADC channels: eg, the gain of the differential amplifier or the input current to which an input voltage corresponds.

In addition, the file contains the name of the DSP32C executable which is to be downloaded and a piece of text that describes the experiment and is subsequently copied to the output data file.

<table>
<thead>
<tr>
<th>PC action</th>
<th>Semaphore value</th>
<th>DSP32C action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wait for semaphore = 1</td>
<td>0</td>
<td>Busy</td>
</tr>
<tr>
<td>Waiting</td>
<td>0</td>
<td>Write semaphore = 1</td>
</tr>
<tr>
<td>Transfer to or from DSP32C memory</td>
<td>1</td>
<td>Wait for semaphore = 0</td>
</tr>
<tr>
<td>Write semaphore = 0</td>
<td>1</td>
<td>Waiting</td>
</tr>
<tr>
<td>Continue</td>
<td>0</td>
<td>Continue</td>
</tr>
</tbody>
</table>

3AT&T, 555 Union Boulevard, Allentown, PA 18103, USA.
4Supplied by Bores Signal Processing, Fordwater, Pond Road, Woking, Surrey, UK.
The hardware and software implementation was tested extensively against known signal sources to help ensure its validity.

4. The analysis system

During development of the instrument and investigation of suitable analysis techniques, the biological system chosen for study was a suspension of baker’s yeast (Saccharomyces cerevisiae) to which glucose was added, following which the glucose was gradually metabolised by the yeast, causing effects within the cells that had been shown qualitatively (Woodward and Kell, 1990) to be detectable by nonlinear dielectric spectroscopy. Such an experiment would result in a multi-dimensional data set containing up to around 5000 floating point values. Visual inspection of these data sets, with each represented as multiple three-dimensional plots, indicated that data points relevant to the effects of interest were distributed across various regions of the multi-dimensional data space and that no simple approach to data analysis was likely to be effective in our quest for a (quantitative) measurement technique.

To explore the suitability of multivariate statistical methods for quantitative interpretation of the data, Principal Components Analysis (PCA) (Hotelling, 1933) was applied. PCA identifies linear combinations (the ‘components’) of data points, where each combination represents a hidden ‘factor’ within the data set. If these factors are arranged in decreasing order of variance then the first few, the ‘Principal Components’, should characterise hidden factors that relate to significant effects within the data, while components of lower variance reflect less important effects along with measurement errors and noise. An implementation of PCA using the NIPALS algorithm (Wold, 1966) was written in ‘C’.

An example of the resulting output data is given in Fig 2. This is a graph of the fifth principal component value for a sequence of frequency/amplitude sweeps for a yeast culture. Glucose was added to the culture at sample number 20 (1 hour from the start of the sequence). The strong peak and decay corresponds (qualitatively) to the response one would expect when cells detect the presence of glucose, metabolise it and return to inactivity. Examination of several data sets in this way showed that the first few components did indeed contain the most apparently relevant variations.

The next stage in identifying appropriate analysis methods was to turn to the relatively new discipline of ‘chemometrics’ (Massart et al, 1988; Brereton, 1992), which uses multivariate statistical methods to form calibration models; formation of such a model to represent the relationship between glucose metabolism in yeast and the corresponding nonlinear dielectric spectra would provide, at least in principle, a method of measuring metabolic activity in which the model would predict, from the nonlinear dielectric spectra, the parameters of interest. The chemometric methods used are based on Partial Least Squares regression (PLS) (eg, Martens and Næs (1989)) and involve a process of supervised learning. Known values of a single parameter of interest are determined by independent means and are then presented to the PLS algorithm along with the measured data (the ‘X’ data) for which a calibration model is required. Where prediction of a single (‘Y’) parameter is sought, the method is often referred to as PLS1; PLS2 extends the PLS method so as to form a model that incorporates, and permits prediction of, multiple ‘Y’ parameters.

Software was written to incorporate these techniques into the instrumentation. This was based around three libraries that were specially written to provide a basis for a compact, efficient and extensible implementation of the analysis software. The first of these was a matrix library; use of “off the shelf” matrix software was considered but in view of the relatively modest set of matrix operations needed for this application a compact, tailored, implementation was preferred. The completed library was tested extensively and the results compared with those obtained from using the Paris maths library. The data library defines 13 major data structures for storing the results of all stages of the analysis cycle and contains functions for manipulating those structures. Facilities for overall coordination of the system were provided by the menu library, which supports the construction of menus and the association of ‘C’ functions with the various menu options. The software structure is such as to allow relatively easy addition of a graphical user interface simply by replacing the menu library. Portability was designed into the software, so that it could be recompiled to allow the analysis to be performed either on the PC itself or on a separate MSDOS or Unix workstation.

The system supports PCA, PLS1, PLS2 and other methods. In each case, the user is prompted for the number of factors, components, etc. The resulting model is stored in the appropriate data set, along with the residuals, and may be applied to a raw data set to give a prediction of Y values for the set. A script file facility allows a set of keypresses to be recorded and stored for later execution so as to reproduce the stored sequence of operations. The package generates plot files for use by

---

% Available by ftp from ftp://megrez.ceremab.u-bordeaux.fr/pub/pari/
gnuplot, from which 2-D and 3-D graphs may be generated for raw data, residuals and weightings sets.

4.1 Model validation

Once created, a PLS model needs to be validated carefully before any attempt is made to generate predictions from it. Validation is, in ideal circumstances, achieved by testing the model against a very large number of previously unseen pairs of X and Y data and ascertaining that the error (usually expressed as the mean, or root mean, squared error of prediction) between the expected and predicted values is acceptable for the intended application. This ascribes a confidence level to subsequent predictions and guards against 'overfitting', where the model is able to reproduce the calibration set but is not able to generalise.

In practice, data sets are often not easy to obtain and the number of data points available for inclusion in the validation set is rather limited. In these circumstances, other techniques can be used to assess the validity of a model. Partitioning the available data into a calibration set and a validation set, so that each contains a representative spread of 'X' and 'Y' values, can be used with care when the complete data set is large. In another approach, known as 'full cross validation', the data set is split into a number of parts and in each case one of these parts is used to determine the prediction error of a model formed on the remainder of the data; finally a model is formed on the whole data set and the mean of the errors of the individual parts is taken to be the error for the whole model. In the limit, this technique is used to predict each 'Y' data point in turn from successive models formed with the remaining 'X,Y' data pairs. Support for these validation techniques is incorporated into the analysis software.

5. Example results

It is emphasised that the results presented here are examples intended only to illustrate the capabilities of the instrument and not in any way to constitute a detailed validation of nonlinear dielectric spectroscopy as a technique for biological monitoring, in respect of which the reader is referred to Woodward et al (1996).

In our simple experiments to explore the relationship between the spectral data and the metabolism of glucose by yeast, the 'Y' parameter was the concentration of glucose in the suspension, measured by means of a 'Reflolux' blood glucose monitor, and the 'X' data were the collected harmonic amplitudes from the repeated frequency/amplitude sweeps. These values were collected throughout the period of the experiment, which was typically several hours.

The first example illustrates the use of the instrument to monitor glucose concentration in a yeast suspension over a period of six hours: an hour after data acquisition commenced, glucose was added to the suspension and monitoring continued until well after all the glucose had been metabolised. Over the six hour period, measurements of glucose concentration, with the 'Reflolux', and complete spectral sweeps, were taken every six minutes. Thus there were 60 known (X,Y) data pairs. A four factor PLS model was formed on the odd-numbered pairs and was used to predict the even-numbered pairs. The validity of the model is indicated by the similarity between the predicted and actual curves in Fig 3 and is represented more appropriately (Kell and Sonnleitner, 1995) by the plot of predicted glucose values versus measured values shown in Fig 4.

A second example is included to indicate the potential applicability of the method to the measurement of other parameters of interest. The instrument was used to monitor a yeast fermentation in batch culture under controlled conditions. Multiple parameters were measured by largely wet-chemical 'conventional' methods at intervals of approximately one hour. For each complete spectral sweep, representing a single data sample, 100 power spectra were averaged at each point in a logarithmic frequency range from 10–100 Hz in 30 steps and a linear voltage range from 0.5–1.5 V in six steps. A complete sweep was acquired at regular intervals of approximately three minutes, so that the resulting predictions of the 'Y' parameters exhibit a time resolution about 20 times greater than the manually
Fig 5a, 5b, 5c, 5d. PLS2 predictions and measured data from a yeast fermentation in batch culture under controlled conditions. Multiple parameters were measured by largely wet-chemical methods at intervals of approximately one hour. A complete excitation sweep was executed by the instrument at regular intervals of approximately three minutes, thus exhibiting a time resolution about 20 times greater than the manually collected data. A two-factor PLS2 model was formed on the manually measured 'Y' values and their corresponding 'X' data samples. This model was then used to predict the 'Y' parameters corresponding to the remaining 'X' samples. The figure shows plots of four of the 'Y' parameters, in which correlation between the measured points and the predicted variation can be observed.

6. Discussion and conclusions

Results such as those given above indicate the suitability of this instrument for exploiting nonlinear dielectric spectroscopy in monitoring fermentations. Application of the method has up to now been limited in that a calibration model formed from fermentation data, such as that described above, does not always with sufficient accuracy predict the parameters of interest in another, similar fermentation. This limitation has been investigated and the cause identified (Woodward et al., 1996) as large non-reproducible variations in the characteristics of the interface between the excitation electrodes and the medium supporting the cells. As noted earlier, driving the excitation electrodes from a high-impedance source, so that the current through the cell suspension is effectively independent of variation in the electrode characteristics, removes the harmonics of interest from the response signal and does not therefore offer a solution (Woodward and Kell, 1995). Work is continuing on various approaches to the improvement of electrode reproducibility; Zhang (1996) has undertaken an extensive investigation of the performance of different electrode types and Woodward et al. (1996) include a discussion of the topic.

Application of the technique to monitoring photosynthetic cells is reported by McShea et al. (1992). Another approach to the measurement of nonlinear dielectric spectra, discussion of the biological aspects, and a description of the extensive experimental
investigations that have been undertaken to establish the capabilities of the technique, are given in Woodward et al. (1996).

The work described in this paper has resulted in a powerful and convenient means of acquiring and analysing nonlinear dielectric spectra. Further details of the integrated instrument and its capabilities are given in Jones (1995).

7. Acknowledgements

This work was supported by an EPSRC studentship awarded to Alun Jones and by BBSRC grant No GR/H87858. We would like to thank Drs Hazel and Chris Davey for their help with the fermenter experiments.

8. References


AT&T. 1988a. WE DSP32 and DSP32C C language compiler library reference manual. AT&T, 555 Union Boulevard, Allentown, PA 18103, USA.

AT&T. 1988b. WE DSP32 and DSP32C support software library user manual. AT&T, 555 Union Boulevard, Allentown, PA 18103, USA.


Research, Development and Application Notes
A shorter form of Transaction Paper

For various reasons, but particularly due to pressure of work in industry, publication of an important advance in the application of measurement and control techniques is delayed or even prevented altogether. This means that wider application may be delayed; it also denies to the author recognition of his work outside his immediate environment. Although, in industry, there is not the same pressure to publish – in fact one might say the pressure is often in the opposite direction – it must be for the general good that as much as possible is published widely.

Bearing this in mind, the Institute is prepared to publish in its Transactions what are described as ‘Research, Development and Application Notes’. These would be of, say, 2000–3000 words, and be intermediate between an abstract and a full paper. They would, of course, be refereed and the work described would have to be normally publishable in the Transactions rather than in the journal Measurement and Control, but the degree of detail would necessarily be less. The objective, the methods used and the degree of success would have to be described, and perhaps wider implications, the need for the work, etc., outlined.

In addition to completed work, these notes could be used to describe major landmarks in a long programme of research or development. They could also be used to describe minor developments of research already reported in the Transactions or elsewhere.

It is bound to be beneficial to the author to have his work brought quickly to the attention of his profession. It is equally important for the technical excellence of any industrial R&D organisation to be recognised outside the company within which it operates. Cross fertilisation between R&D Groups and individuals is a vital requirement for rapid advances in technology. We hope readers will take advantage of this facility.