



Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience

Transforms and Calculations: Behind the Mathematics of Psychophysiology

George H. Green^a & John C. LeMay^a

^a Cerebotix Institute, Reno, Nevada, USA

Published online: 25 Aug 2011.

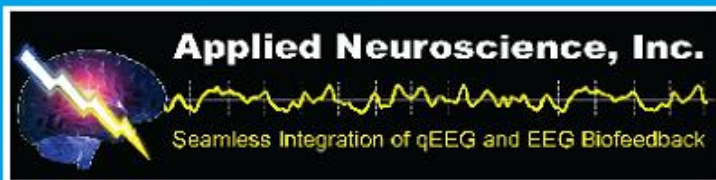
To cite this article: George H. Green & John C. LeMay (2011) Transforms and Calculations: Behind the Mathematics of Psychophysiology, *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and Applied Neuroscience*, 15:3, 214-231, DOI: [10.1080/10874208.2011.597255](https://doi.org/10.1080/10874208.2011.597255)

To link to this article: <http://dx.doi.org/10.1080/10874208.2011.597255>

PLEASE SCROLL DOWN FOR ARTICLE

© International Society for Neurofeedback and Research (ISNR), all rights reserved. This article (the “Article”) may be accessed online from ISNR at no charge. The Article may be viewed online, stored in electronic or physical form, or archived for research, teaching, and private study purposes. The Article may be archived in public libraries or university libraries at the direction of said public library or university library. Any other reproduction of the Article for redistribution, sale, resale, loan, sublicensing, systematic supply, or other distribution, including both physical and electronic reproduction for such purposes, is expressly forbidden. Preparing or reproducing derivative works of this article is expressly forbidden. ISNR makes no representation or warranty as to the accuracy or completeness of any content in the Article. From 1995 to 2013 the *Journal of Neurotherapy* was the official publication of ISNR (www.isnr.org); on April 27, 2016 ISNR acquired the journal from Taylor & Francis Group, LLC. In 2014, ISNR established its official open-access journal *NeuroRegulation* (ISSN: 2373-0587; www.neuroregulation.org).

THIS OPEN-ACCESS CONTENT MADE POSSIBLE BY THESE GENEROUS SPONSORS



TRANSFORMS AND CALCULATIONS: BEHIND THE MATHEMATICS OF PSYCHOPHYSIOLOGY

George H. Green, John C. LeMay

Cerebotix Institute, Reno, Nevada, USA

There are numerous scholarly documents with accurate and thorough explanations of the basis of the mathematical processes that have become essential to the field of psychophysiology. Review of many of these has revealed a pervasive emphasis on the technical and theoretical aspects of these formulae and theories with little or no emphasis on the primary and basic understanding of their development and application. This article specifically bridges the gap between the introduction of several cogent mathematical concepts and their ultimate applications within the field of applied psychophysiology, biofeedback, and neurofeedback. Special attention is given to the distinction between transforms and calculations and some of the statistical methods used to analyze them. Because the focus of this article is to enhance conceptual comprehension, integral, differential, and matrix mathematics are not referenced in any of the examples or explanations with the primary reliance on some algebra with verbal and pictorial descriptions of the processes. We suggest a comparison to an overuse of the black box model in which only the input and output are essential. Taking these processes out of the black box encourages the creative application of these mathematical principles as valuable tools for clinicians and researchers. Structured explanations emphasize the relevance of such important concepts as aliasing, autospectrum, coherence, common mode rejection, comodulation, cross spectral density, distribution, Fast Fourier Transform, phase synchrony, significance, standard deviation, statistical error, transform, *t* test, variance, and *Z* scores. The objective for providing these clarifications is to enhance the utility of these concepts.

In the years following World War II, behavioral psychology and the field of psychology in general received considerable attention. One of the principles that came to be embraced was the concept of the black box, which had been developed in the early 20th century. The black box was referred to in two contexts: complex equipment and the human brain. In both cases both the input and the output can be observed without requiring any knowledge of the inner workings of the box. From these observations, inferences can be drawn about the actual functioning of the black box and applications can be designed to best utilize the input–output relationship. B.F. Skinner embraced this notion by applying it to the

human organism, which he perceived as a black box in which internal processes are not significant behavioral predictors (Skinner, 1938, 1953).

As the field of psychophysiology becomes increasingly complex, the tendency to allow increasing numbers of tools and techniques to fall into the black box category may have finally brought us to what can be called the machinist's paradox.

As technology develops, fewer people maintain the basic knowledge and ability to build the tools that built the machines that subsequently built the current machines. In this paradox the users of the newest machines are forced to draw assumptions about the workings

Received 23 March 2011; accepted 27 May 2011.

Address correspondence to George H. Green, PhD, Cerebotix Institute, 3310 Smith Drive, Reno, NV 89509, USA. E-mail: drgreen@cerebotix.com

of their equipment simply because they are so far removed from the original process. Current computers are a good example of this paradox, because most users are not required to possess the knowledge necessary to build a computer. Even most expert programmers probably do not possess enough knowledge about the individual electronic components to construct a motherboard from scratch.

At this point the function of the black box can be inappropriately extended beyond its designed purpose to become a repository for complex or poorly understood processes. Such is the plight of the black box assumption.

In a controlled research environment, if the actual functioning of a device is generally accepted, the device may be given black box status for the purpose of the research. What started as a mechanism for allowing researchers to avoid unnecessary empirical and possibly tangentially related work, however, has evolved into a box into which we place the incomprehensible, the excessively complex, and the presumably unknowable. However, in the real world, the more you know about how your equipment processes data, the more effective you will be in your work. More important, the apparent and assumed output of the black box may not be the same as its actual output. The only way to understand this is by understanding the workings of the black box.

Complex mathematical principles not only underlie the modern biofeedback processes but are the vital blood that makes it possible. These principles have themselves become black boxes that are generally recognized by name and employed enthusiastically. Unfortunately, these processes may have become poorly understood by many clinicians and even researchers simply because the processes seem to do their jobs so well, there is no need to look inside them.

There are many important terms that describe and result from the mathematical processing of data. For example, disagreement and confusion frequently accompanies technical discussions involving the following terms: aliasing, autospectrum, coherence, common mode rejection, comodulation, cross spectral density,

distribution, Fast Fourier Transform (FFT), phase synchrony, significance, standard deviation, statistical error, transform, *t* test, variance, *Z* scores. It is possible to imagine that attempts at using or even defining such important terms can result in research errors, inaccurate clinical decisions, and invalid or potentially dangerous conclusions. These terms fall into the broad category of "Transforms and Calculations," which are elegantly practical blends of algebra, differential and integral calculus, trigonometry, and statistics.

The good news is that it may not be necessary to study all these branches of mathematics in order to develop an adequate appreciation for how these numeric processes work and what they accomplish. Instead of having to rebuild the black box, it is possible to open it and study the function of each piece.

DATA COLLECTION

The output of our biofeedback is only as good as the quality of the data we collect. When preparing to collect data, we are obliged to define

- *Constants*: the factors that are unchanging within the study
- *Statistical significance*: the measure of how likely a given outcome is to be the result of chance (Kaplan & Kaplan, 2010)
- *Independent variables*: the influencing factors to be studied
- *Dependent variables*: the responding factors to be studied
- *Accuracy*: the degree of statistical significance of the set of results
- *Precision*: the distribution of results within the data set
- *Errors*: influences outside the study that can alter the outcome

For example, in order to go to the store, some of the *constants* could be (a) the distance to be traveled, (b) the store's employees, (c) the store's business hours. If these values are fixed in relation to a study, then their influence is considered invariable.

The *independent variables* can include (a) the weather, (b) the method of transportation, (c) the time of day. These are the factors to be studied that will influence the outcome.

Dependent variables can be (a) how much you are going to buy, (b) how many of each item you actually purchase, (c) how much time spent chatting with a friend you bumped into. These are the factors to be measured.

With these definitions you could design studies such as these: "The Effect of Weather Conditions on Purchase Quantities," "Temporal Discourse Adjustments in a Shopping Environment Based on Methods of Transportation," or "Distribution of Purchased Items Measured Hourly in a 24-Hour Period."

Depending on the study, a constant can become a variable, and a variable can become a constant. If you study shopping at a variety of stores, then which store you select is a variable. On the other hand, if you study shopping only in a particular store, then the store becomes a constant.

It is possible to see how *error* terms or sources of error can be inadvertently assigned constant status or variable status. The four most common types of error are *random*, which are difficult to predict because their effect can vary differently with each element in the study population; *bias*, an error that is constant for the defined population; *Type I*, incorrectly identifying an outcome as significant (false positive); and *Type II*, finding no significance when in fact it is present (false negative; Graziano & Raulin, 2000).

Currently the amplifiers used in electroencephalograph (EEG) biofeedback have substantially greater input impedance values than were commonly found in EEG biofeedback equipment as recently as the 1990s. Although electrode impedance and skin preparation have become less of a factor, they remain as a potential source of error, which should not be ignored. Electrical impedance can be affected by a surprisingly large number of elements. Many of these can be sources of error, constants, independent variables, or even dependent variables. These can include electrode type, electrode condition, length and type

of wire, type of insulation, type and condition of the connectors, diameter of the wire, diameter of the electrode, age of paste, type of paste, degree of hydration in the paste, type of skin prep, method of skin attachment, skin type, method of skin prep, salinity of preparation products, chemicals that may be present in hair or scalp from normal grooming or from environmental influences, hair length, and type of impedance measuring device.

With this daunting list and the newer equipment characteristics, it is fairly easy to see how impedance is at risk of becoming (or may already be) a black box. Skin type, for example, varies among people and contributes to error, yet we assume it to be a constant. However, if you are testing for differences in skin type, it is an independent variable. If you think of skin type as a variable but determine that the impedance must be under a specified accepted value, then it becomes a constant for the experiment. If you ignore the impact of skin type, then it becomes part of the error term.

Constants can be defined as unchanging elements that influence the outcome that must be assumed and need to be acknowledged and reported. In a typical biofeedback experiment the room in which the experiment is conducted is most likely a constant because defining the numerous elements of a given room are unnecessary if the room is essentially the same and the elements within the room are not changing. Similarly, if the same equipment is used throughout the study, variables such as operating temperature of the equipment do not need to be reported unless those elements are being tested. Often constants are immeasurable values that can be qualified (identified) but not quantified (measured), so researchers have the option to report them within the context of error terms based on the objectives of the study.

A constant only maintains its status as a constant within a defined population. It may be variable between populations and must be redefined if the population itself is redefined. *T* tests allow us to test variables against constants in a population. By calculating the ratio of the difference between means and

the variability within the population, all constants and error terms will be included with the variability.

Variables, then, are elements that change. Independent variables are generally the elements we are studying. Dependent variables are the results of the study.

Once variables are identified, *accuracy* and *precision* in data collection must be addressed. Because error terms can have an additive or even multiplicative impact on data, they can introduce enough bias to invalidate an entire data series.

In the known universe all measurements are estimates. Accuracy itself is an estimate based on convention. The first decision about accuracy, therefore, is to determine the *degree of significance* that is acceptable for a given set of measurements. *Significance* and *significant figures* are statistical terms that are determined by the total number of digits used in measurement. As a rule, the last digit of a measurement is the approximation or estimate and is frequently either "5" or "0" because it necessarily is the result of rounding. The use of this digit

in calculations does not improve accuracy. A perfectly square table can be built with a substantially lower degree of accuracy (three or four significant figures) than a perfectly machined racing engine (six or more significant figures). In either case, the value of the last digit does not contribute to the accuracy of the result (Brown, LeMay, & Burstead, 2006).

Accuracy, therefore, is an absolute value that describes how close a given measurement approximates reality. Precision is a relative value that is reported as the standard error of the mean. It is possible to have measurements that are precise (small standard error) but not accurate as well as accurate but not precise (large standard error; Figure 1).

Most of the data collected for use in bio-feedback are *periodic* in the form of waves of electrical energy, although light and sound spectra are also used. Working with entire waves is ponderous and complex. By convention we reduce the wave data with *sampling* that transforms the data into segments of digital packages that are easier to manipulate and analyze. However, correctly selecting the type

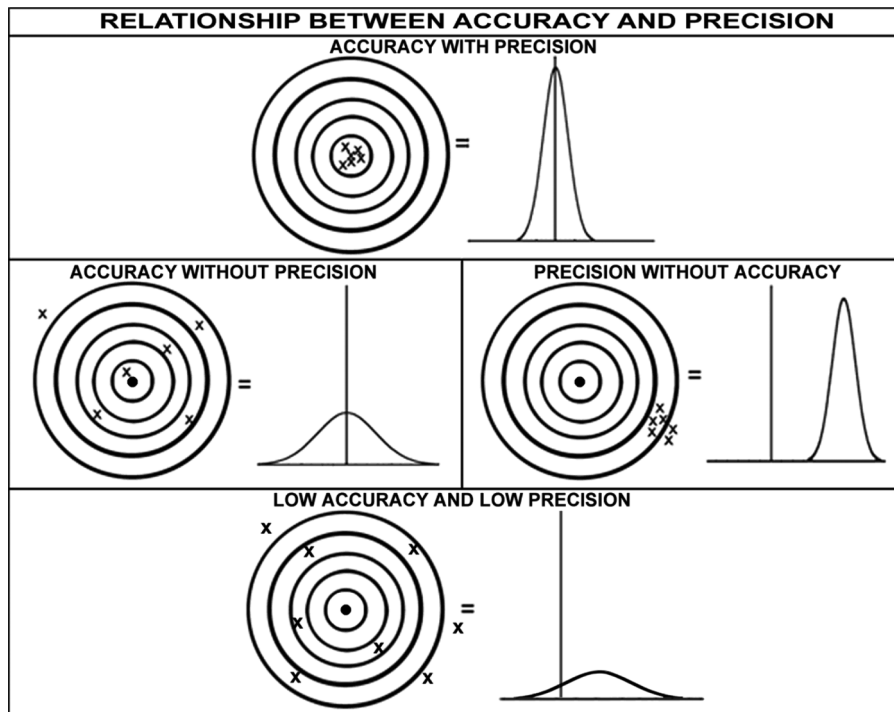


FIGURE 1. Relationship between accuracy and precision.

and rate of sampling is basic to collecting data that accurately represent the original observations (Marks, 1991).

When we sample, we are essentially agreeing that some of the original data can be discarded and the remaining information will accurately represent the entire data field (Figure 2). Reconstruction of original or missing data is accomplished with methods of *interpolation* or *extrapolation*. The four simplest and most popular forms of data reconstruction are

1. *Simple linear interpolation*, which determines missing data between samples by calculating simple means;
2. *Linear extrapolation*, which determines missing data outside the range of samples by projecting data from simple means;
3. *Cosine interpolation*, which creates a modest correction for the abrupt changes in values at existing data points by calculating new points along a cosine curve; and
4. *Aliasing*, which can occur automatically when the sampling rate is below the resolution or accuracy of the data and creates repetitions of existing sampled data in the empty spaces between samples (Harris, 2006).

Each of these techniques has strengths and weaknesses, but none of them are capable of bringing back discarded data (Figure 3).

The principle of *aliasing* provides a rational paradigm for appreciating the limitations inherent in working with sampling and accuracy. Leon Harmon (1973) created one of the first examples of pixelated oversampling of the image of Lincoln from a five-dollar bill. Figure 4 is a re-creation of Harmon's original

image and takes spatial aliasing a bit further to extreme oversampling, then shows the result of attempted reconstruction. Curiously, if you stand back far enough you can begin to see the Lincoln image appear to reemerge as your brain starts to use memory for interpolation. However, the ambiguity of these reconstructed data is evidenced by the fact that in the same reconstructed image it is also possible to see Hugh Jackman as Wolverine or Roddy McDowall as Cornelius from *Planet of the Apes*.

Spatial aliasing is the insertion of accurate data at the wrong coordinates. *Temporal aliasing* occurs when accurate data are inserted at the wrong time. The temporal aliasing effect is seen commonly when bicycle wheels are filmed. If the frame rate does not match both the number of spokes and the rate of revolution of the wheel, the wheel can appear to be moving in reverse or even standing still. This phenomenon is often referred to as the stroboscopic effect, which can be demonstrated by matching the flash rate of a stroboscope with the number of blades on an electric fan and its speed of revolution. The fan blades appear stationary while a pencil inserted in the fan is broken.

In the temporal aliasing model frame rate or flash rate is equivalent to sampling rate. The number of spokes or blades is analogous to the minimum resolution of the data, whereas the rate of revolution is analogous to the frequency of the waves we are sampling.

To work with the limitations of sampling and reconstruction, the *Cardinal Theorem of Interpolation Theory* was developed nearly simultaneously by Harry Nyquist and several others (Marks, 1991). Although the application of the theory can be complex, it has been

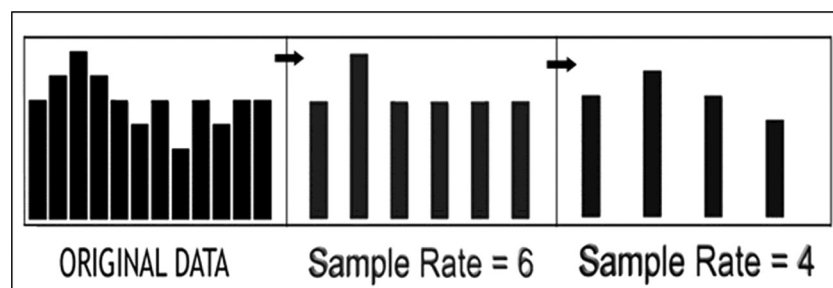


FIGURE 2. Sampling rate comparison.

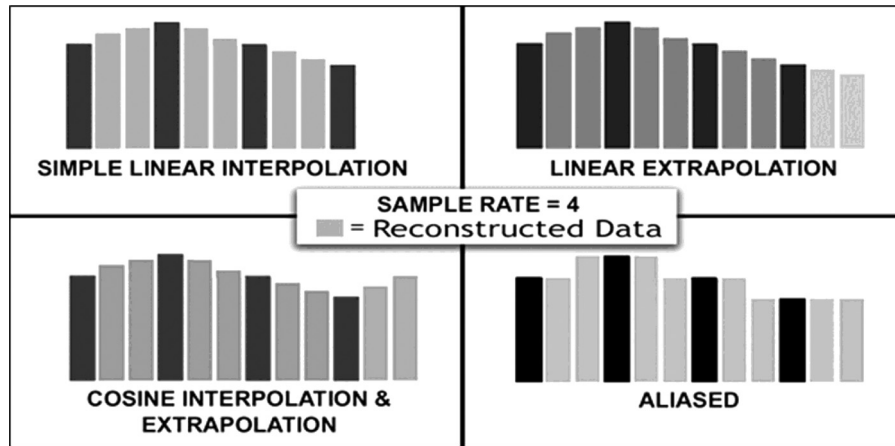


FIGURE 3. Some methods of data reconstruction.

reduced to a simple calculation that can be a useful guideline for determining sampling rates so that accurate reconstruction should always be possible. Basically, it states that *for the highest frequency in the original signal, the sampling rate should be greater than twice the frequency*. A useful rule of thumb developed by the CD industry states that the sampling rate should be at least 2.205 times the highest frequency (Hartman, 1997).

An audio CD has a sampling rate of 44,100 samples per second. When divided by the Nyquist approximation of 2.205, this yields

20,000 Hz, which is the highest frequency that is usually included in audio recordings. Professional recording studios generally sample performances at 48,000 samples per second so that after editing they are still above the 44,100 value. When they finally reduce it to 44,100 samples per second, the recording sounds like a perfect audio reproduction, which is essentially an accurate reconstruction of the original analog data.

For the purposes of EEG biofeedback based on Nyquist’s approximation, an 8 Hz wave should be sampled at a minimum of 18 samples per second. Similarly, a 30 Hz wave should be sampled at 67 samples per second or better. A 48 Hz wave should be sampled at 106 samples per second or better. If we consider that sample rates of 256 samples per second are considered the minimum recommended standards for EEG biofeedback, then application of Nyquist’s approximation indicates that data collection up to 116 Hz should be accurate (Sanei & Chambers, 2007).

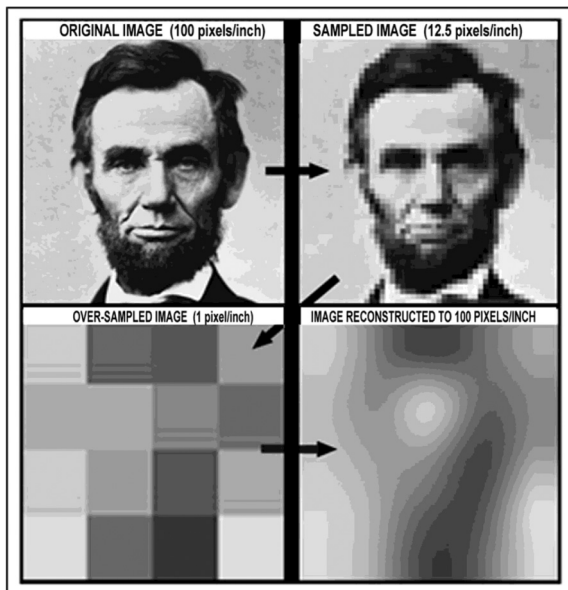


FIGURE 4. Example of oversampling.

TRANSFORMS AND CALCULATIONS

Transforms in mathematics are methods of modifying the basic structure of a data set to make it easier to analyze. Technically, these are *active transforms* because the data set is rearranged to new coordinates. *Passive transforms* are shifts in perspective or changes in vectors. These shifts can be made without

influencing the data coordinates directly. Within the context of digital sampling, the aliasing process involves active transform. Passive transform, which is accomplished by shifted vectors (altered perspective), has been assigned the term *alibi*. If aliased data are the placement of real data in a contiguous but incorrect location, then alibied data are real data that are perceived from a different and possibly incorrect perspective (Figure 5).

Once analyzed following active transform, the data set can be restored to its original domain by the inverse transform. The data do not change; only the coordinate system changes, which then provides new methods of analysis. Transforms in mathematics are an extensive and powerful collection of tools that allows us to analyze data sets in a variety of quantifiable ways. These tools permit observation and measurement of data relationships that might otherwise require far more complex methods of analysis or possibly escape discovery and measurement entirely.

In contrast to transforms, *calculations* use mathematical computations to determine a result. Interpolation and extrapolation are both examples of calculations. When working with EEG data, several types of transforms are employed in order to represent the data in such a manner that we can calculate treatment effects and other phenomena. Calculations provide new information, whereas transforms provide new perspectives on existing data.

SOME BASIC TRANSFORMS

Most aspects of biofeedback involve the use of transforms. To analyze raw EEG data, transforms are used to allow us to identify characteristics

that would be difficult or impossible to work with in their original form. Even the graphical user interface, which provides the visual link essential to most types of biofeedback, depends heavily on transforms. Information that is rendered as images in three-dimensional space is entirely a product of mathematical transforms. Even simple bar graphs or two-dimensional images are transformative of the calculated data.

Examining four of the basic transforms will provide a valuable starting point for a more detailed understanding of EEG data modification. These basic transforms are *translation*, *reflection*, *scaling*, and *rotation*. These four can be supplemented by numerous more complex transforms that allow a variety of data perspectives in wave analysis. The Log Base 10 transform is demonstrated in Figure 12. Additional transforms that are frequently referenced in the literature include Euler, Laplace, Bessel, Radon, Gauss, Gabor, Zak, Box-Cox, square root, cube root, and hyperbolic arcsine, each providing a unique method of exploring data (Sanei & Chambers, 2007).

The *translation transform* can be expressed mathematically as

$$(x', y') = (x + X, y + Y)$$

To use the translation transform, simply slide the data intact along a single Cartesian axis in one direction (Figure 6). You are adding or subtracting a single value from one axis.

Imagine driving in a car. Your movement down the street is a translation of data in the set that includes your car and its contents. You are adding or subtracting a single value from one axis.

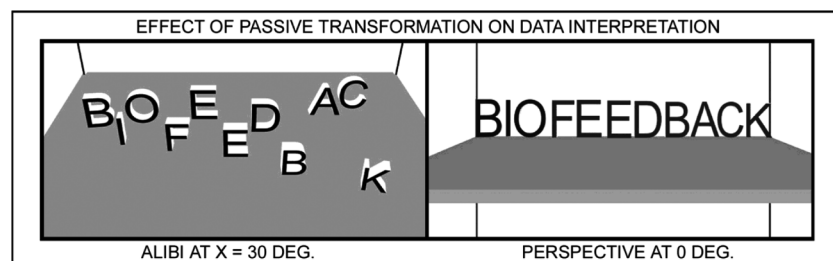


FIGURE 5. Example of passive transform.

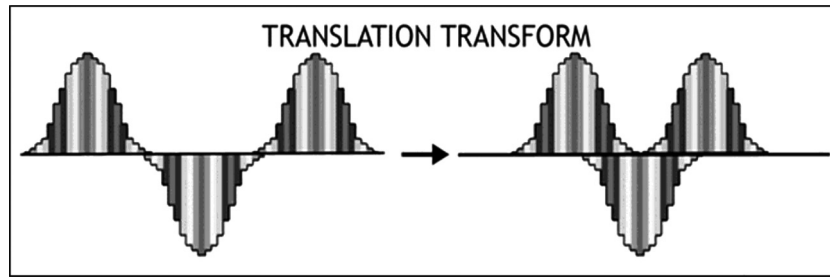


FIGURE 6. Example of the translation transform for data subset $y > 0$, $x' = x - |x|/2$.

The reflection transform can be expressed mathematically as

$$(x', y') = (-x, y) = (x, -y)$$

The reflection transform changes your orientation on one axis by changing your sign. Positive becomes negative, and negative would become positive (Figure 7). If you were heading east, then after the reflection transform you would be heading west.

The scaling transform can be expressed mathematically as

$$(x', y') = (mx, my) \text{ where } m > 1 \\ = \text{dilation and } m < 1 = \text{reduction}$$

The scaling transform includes the characteristics of reduction and dilation and is accomplished by multiplying your Cartesian coordinates by a single value (Figure 8). Using the car analogy, you and your car would be

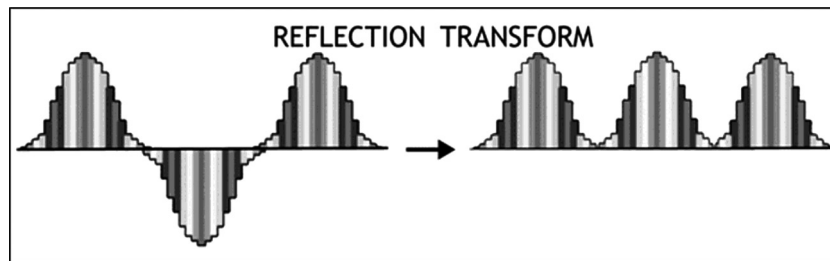


FIGURE 7. Example of the reflection transform for data subset $y < 0$, $y' = -1(y)$.

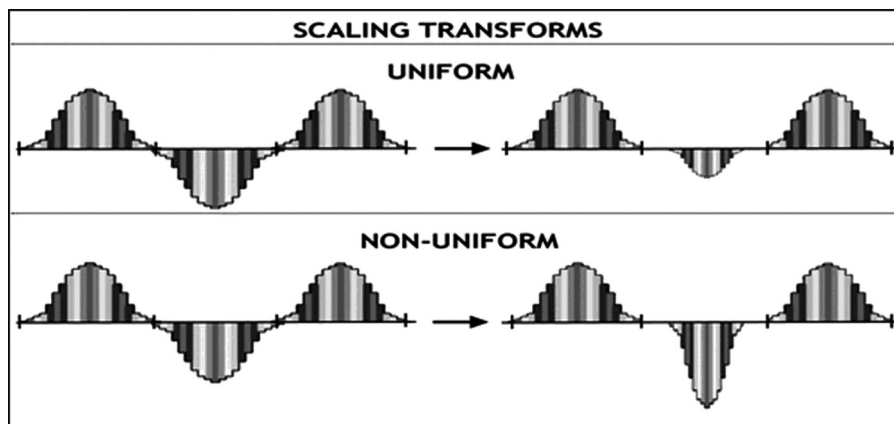


FIGURE 8. Examples of the scaling transform for data subset $y < 0$: for uniform scale, $x', y' < x, y$; for nonuniform scale, $x' < x$, $y' > y$.

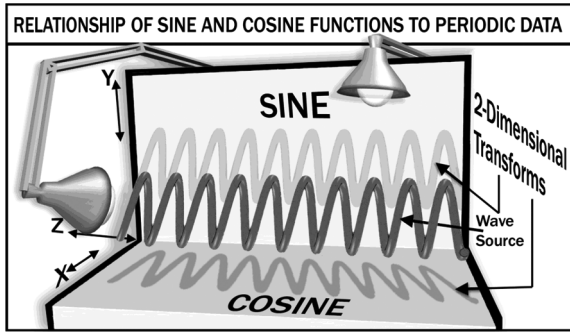


FIGURE 9. Application of sine and cosine functions in the analysis of periodic data.

expanded or miniaturized. If the scaling value is different for each coordinate, the scaling transform is no longer uniform because the relationship between all elements in the data set is not maintained.

The *rotational transform* is another matter. Because the data set is being modified in spatial orientation, the transform involves three-dimensional coordinates. Furthermore, because the reorientation is rotational, this

transformation requires adjustments made based on angular changes.

$$\begin{aligned}
 x' &= x(\cos \theta) - y(\sin \theta) \\
 y' &= x(\sin \theta) + y(\cos \theta) \\
 z' &= \{1, 0, 0\} = z
 \end{aligned}$$

In this case *polar movement* (rotation around a specified axis) occurs as the data set is fixed to a single axis ("z" in the prior example), whereas the values of the other two axes are allowed to change. Cosine values of the angle of change provide the two-dimensional representation of the circular spatial movement around the stationary axis, which acts as the origin for this portion of the transform. Hence, multiplying a Cartesian value by the cosine of the angle of change generates one two-dimensional image of the data. When you have generated this value for two coordinates, piecing them together provides the three-dimensional spatial data, which can be demonstrated by observing the shadows cast at 90° by a helix (Figure 9).

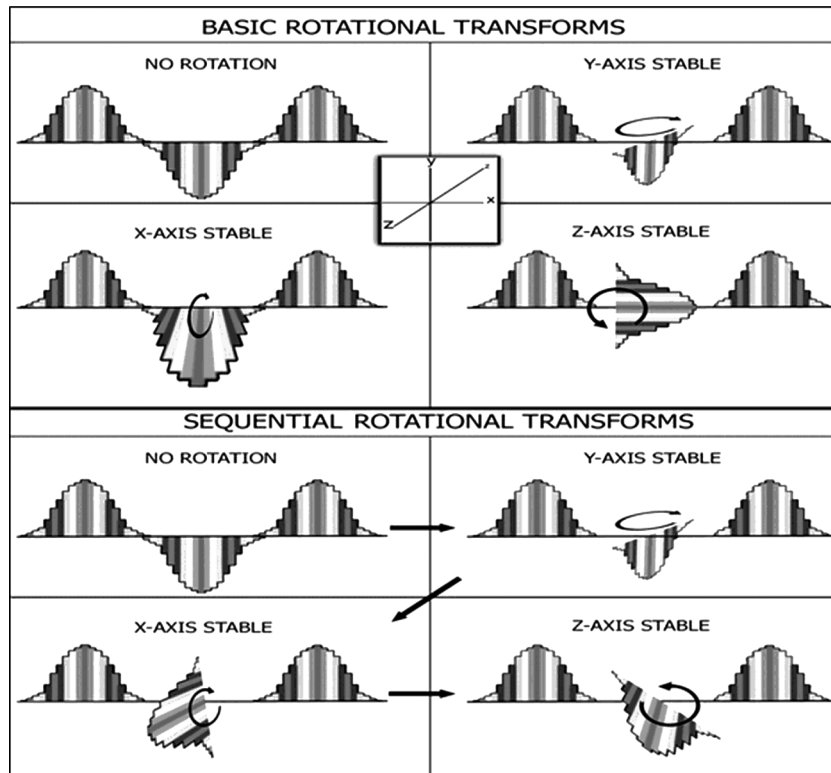


FIGURE 10. Examples of rotational transforms for data subset $y < 0$.

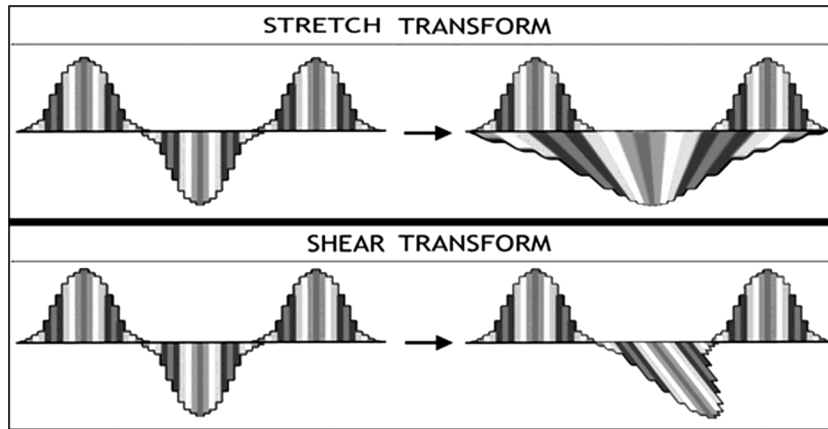


FIGURE 11. Examples of the stretch and shear transforms for data subset $y < 0$.

If you perform this transform for each of the three coordinates, the result will be a complex spatial transformation (Figure 10). Because all motion is relative to some object, all objects are in a state of dynamic translation. In addition, all movement is influenced by a variety of gravity fields and other forces, which make the concept of a perfectly straight line limited in application. Consequently, all objects are subject to continuous rotational transformation and spatial curvilinear translation. Understanding and accurately analyzing these complex movements requires mathematics beyond geometry and trigonometry (Shoemake, 1994).

Three other basic transforms are commonly employed: *the stretch/squish transform*, *the shear transform*, and *the logarithmic transform*. The stretch transform increases dimension

along a single axis while having no influence on the other perpendicular dimensions. A square or rectangle would transform into a trapezoid. The shear transform describes the linear translation of a single side of an object along its axis such that the resultant figure retains the initial relationship of opposing sides while transforming the relationship of adjacent sides. In this case, a square or rectangle would be transformed into a parallelogram of the same area (Figure 11).

The logarithmic transforms allow data sets that contain one or more expanding variables to be plotted on a linear map. Values that are multiplicative, exponential, and percentage distribute in nonlinear fashion. Log transforms convert these to values that are additive and more readily subject to analysis. Figure 12 shows both

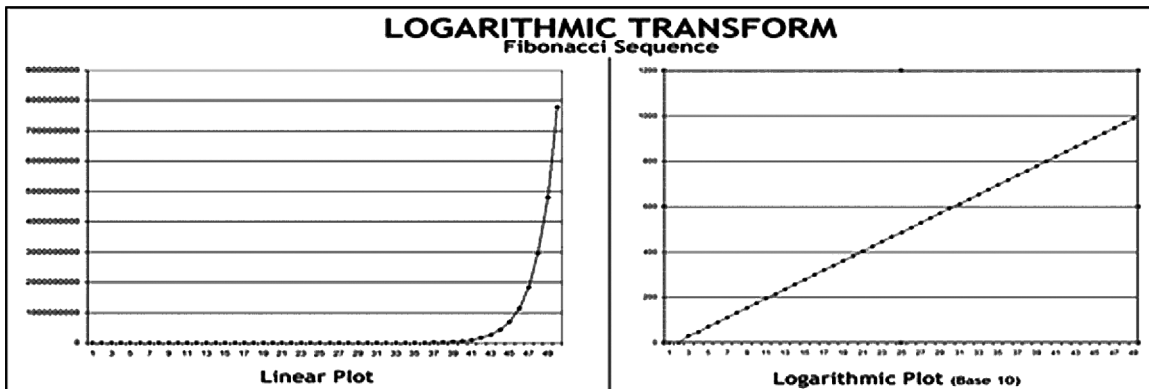


FIGURE 12. Example of the logarithmic transform of the first 50 values in the Fibonacci sequence.

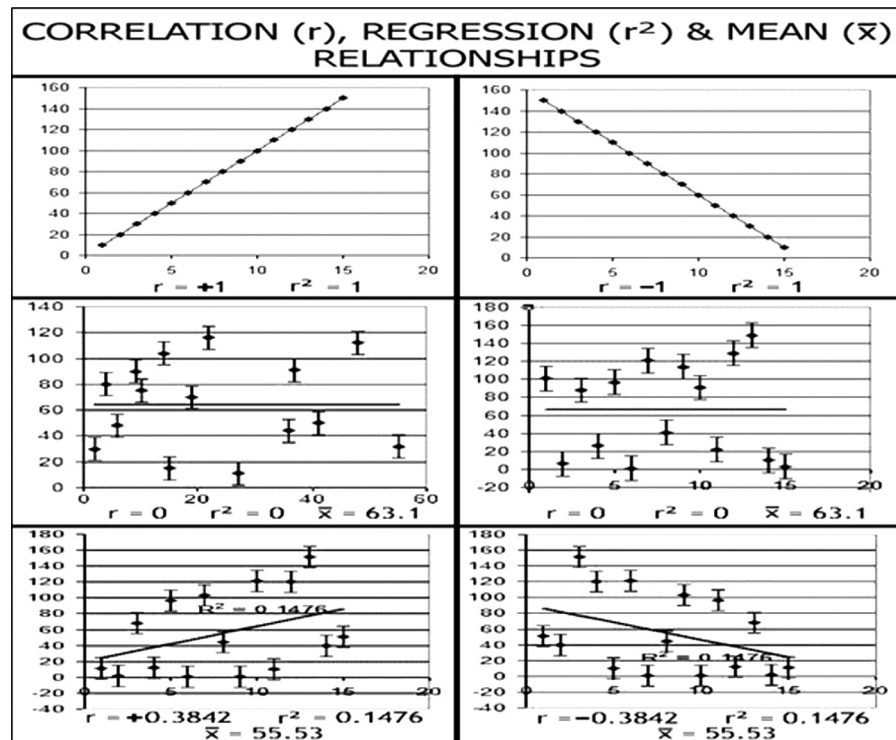


FIGURE 13. The correlation, regression, and mean relationship.

the linear plotting and the logarithmic plotting of the first 50 numbers of the Fibonacci Sequence as defined by the linear recurrence equation:

$$F(n) = F(n - 1) + F(n - 2).$$

Once data have been transformed onto a linear map, they can be analyzed by *linear regression* and *correlation*. Even a simple regression—the best straight line you can make through your data—can provide a valuable new perspective as in Figure 13.

Understanding the distinction between regression and correlation is essential to the data interpretation process because these are powerful tools. Correlation (r) is predictive because you are calculating the relationship of the data to the regression line by determining the likelihood that the data will fall along that line. *Correlation coefficient* values range from -1 for an inverse proportional relationship to $+1$ for a direct proportional relationship. In Figure 13, note the differences and similarities between the paired graphs.

If the r value is squared, r^2 , the result mathematically is referred to as the *Coefficient of Determination* and is restricted to a range of 0 to $+1$. An r^2 value of 0 indicates no relationship between the data and the calculated regression line. An r^2 value of $+1$ indicates that the data and the regression line are identical. The equation for r (Figure 14) becomes logical when it is reduced to its elements.

In other words, the correlation coefficient (also referred to as Pearson's correlation) represents the amount that the two variables change together (covariance) when compared against their dispersion around the mean. Another way to phrase that could be the amount of the dispersion coming from the relationship between the two variables.

$$r = \frac{\text{Covariance}(XY)}{\text{Standard Dev}(X) \cdot \text{Standard Dev}(Y)}$$

FIGURE 14. Formula for correlation coefficient.

USING TRANSFORMS AND CALCULATIONS TO IMPROVE THE NEUROFEEDBACK PROCESS

The specific method of completing the information loop that creates the feedback is based on reward systems, which themselves are determined by calculations. The presentation of psychophysiological information in the form of biofeedback uses the analyzed data and calculates a data stream that will be transformed into usable information by means of some sort of display apparatus. The nature of the display signal can be visual, auditory, and/or tactile and will drive the display in either an analog or binary manner. Analog or proportional signals are presented as a continuum of data, whereas binary signals are restricted to either the “on” state or the “off” state. This section addresses the calculations that produce the signals that drive the biofeedback.

Earlier in this article several terms that either use or embody transforms and calculations were identified as potential candidates for inadvertent black box status. In essence the following terms are the mathematical building blocks on which the field of biofeedback rests.

Distribution

For our purposes, *distribution* is a shortened form of the term “probability distribution” and describes the likelihood that a variable will fall within a specified range around a mean value. The two most commonly referenced distributions are *Gaussian* and *categorical*. Gaussian distribution is also known as *normal distribution* and is characterized by its familiar bell curve graph. Approximation to a Gaussian distribution has been established as one of the standards for QEEG (Thatcher & Lubar, 2009). Categorical distribution describes the likelihood that a single event will have a specific result. A good example of this is coin tossing. A categorical distribution of one to two means that a given side of the coin is expected to show once for every two tosses.

Consider the following equation (Figure 15):

Calculating the *standard error of the mean*—which itself is arguably the most routinely

STANDARD ERROR OF THE MEAN	
$SE_{\bar{x}} =$	$\sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$

FIGURE 15. Standard error formula.

employed statistical tool—is traditionally taught as a reasonably simple recipe. However, it has also fallen into the black box along with its component elements. In fact, these elements are often employed individually without necessarily acknowledging their interrelationship or the value of that interrelationship. The following four steps construct the calculation of the standard error of the mean.

1. *Sum of squares (SS) = $\{\sum x^2 - [(\sum x)^2/n]\}$* : This value is describing the dispersion around the mean by calculating how far each sample is from the mean. By squaring this value two qualities are added: (a) negative values are removed, and (b) larger differences are amplified (transformed).
2. *Variance, $\sigma^2 = SS/n$ (or $s^2 = SS/n-1$)*: This value is the SS calculated down to the level of the mean, which allows it to be a useful comparison tool by describing population variation away from the mean. The distinction between σ^2 (sigma squared) and s^2 is that σ^2 is calculated when working with the values for an entire data set. By calculating s^2 with $(n - 1)$ as the denominator, the variance will be larger and will compensate for calculations made using an incomplete data set.
3. *Standard deviation, $\sigma = \sqrt{\sigma^2}$ (or $s = \sqrt{s^2}$)*: The square root of the variance provides the basis for the distribution curve around the mean by returning the value to the magnitude of the population of samples found in the original data. When this value is added and subtracted from the population mean, the values contained within that range will account for 66.7% of the total population assuming that the data have a normal distribution.

4. *Standard error* = (s / \sqrt{n}) : Reported generally after a mean value and preceded by $+/-$ the standard error provides the standard deviation for a particular sample. Dividing by the square root of the sample size removes it from the population range and places it in the sample range. A quick and simple statistical test of significance between two samples is to subtract the standard error from the larger one and add the standard error to the smaller one. If the resulting values overlap, the general rule is that the difference is probably not statistically significant.

Many statistical tools are based mathematically on ratio relationships that, when phrased in words, can clarify both function and applicability. A ratio is often written as a fraction, such as

$$A/B$$

Beyond the mathematical phraseology for this term—*A over B*, *A divided by B*, *ratio of A to B*—are more descriptive phrasings that describe the nature of the relationship between the two values. The three most common phrasings are

1. the number of times *B* fits into *A*,
2. the amount of *B* represented by *A*, and
3. how *A* changes or varies with respect to *B*.

As some of the more complex terms are defined, because they are based on ratio calculations, this more descriptive language should help define their identities and applications. Figure 14 is a good example of a useful ratio relationship in the calculation of the correlation coefficient, *r*.

Z Scores and the *t* Test

Z scores and the *t* test (Figure 16) are two of the most often applied of the *dimensionless* ratios. Dimensionless ratios are those in which the

$$z = \frac{x - \mu}{\sigma} \quad t = \frac{x - \bar{x}}{s}$$

FIGURE 16. Formulae for *Z* scores and *t* test.

units of measurement are the same for the numerator and denominator such as the correlation coefficient. Mathematically, the units cancel out, which leaves a numeric value without a unit of measurement. These two tests are essentially identical in terms of their algebra.

For *Z*, *x* is a raw score to be standardized, μ is the mean of the population, and σ is the standard deviation of the population. For *t*, \bar{x} is the mean of the subset, and *s* is the standard deviation of the subset which sometimes is substituted with the standard error. Both statistics calculate how a sample value varies from the mean when compared to the relative dispersion of the data set.

The *Z* score assumes the data set includes the entire population or that the true population parameters are known. Because it deals with entire populations of data, the *Z* distribution is considered the same as the normal distribution, and both extremes—called *tails*—extend to infinity. The one-tailed *Z* test defines significance in *direction* such as determining whether the sample is larger than the population mean. The two-tailed *Z* test defines significance in *difference* only without specifying direction. The *t* test compares the sample with a known subset of the population that may not represent the entire population. Because the *t* test uses incomplete data, the *t*-distribution is frequently not normal.

The values generated by both the *Z* and the *t* provide an estimate of statistical significance when compared to their respective distributions and adjusted for sample size.

PREPARATION OF PERIODIC DATA FOR BIOFEEDBACK

Because EEG signals are complex periodic data (i.e., data in the form of complex waves), a variety of transforms and calculations reduce the signal complexity to the final analog or binary output needed for conversion to biofeedback. The simplest waveform is referred to as a sine wave, which is recognized by its characteristic smooth uniformity. Determining the sine of an angle is as appropriate for calculating degrees of arc as it is for calculating the characteristics of a right triangle.

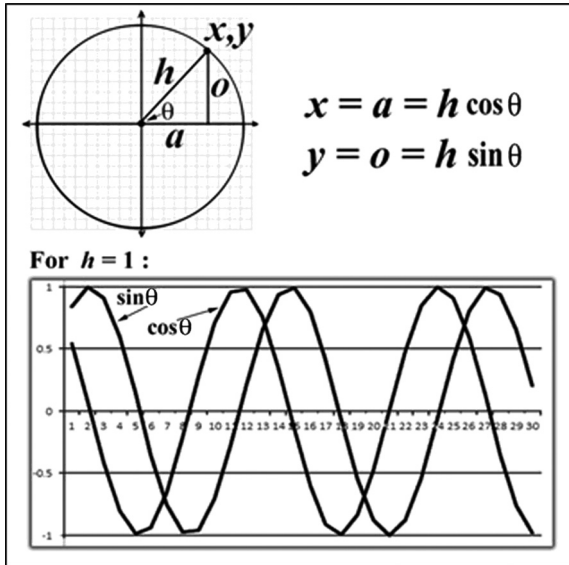


FIGURE 17. The sine-cosine relationship.

By applying trigonometric functions to the analysis of periodic data, mathematical descriptions of waves and their components can be readily derived. Because a wave is periodic in nature, such repetition can be analyzed with the same tool used to analyze circular information.

Figure 17 illustrates the relationship between a circle and both sine and cosine waves. In the example the radius of the circle is also the hypotenuse (h) of the angle, θ , of a right triangle. Notice that the adjacent arm (a) of the triangle corresponds to the x -coordinate of the hypotenuse, and the opposite arm (o) corresponds to the y -coordinate. Because the sine of θ equals the ratio of the opposite to the hypotenuse, the sine

is a calculation of the y -coordinate. The cosine of θ equals the ratio of the adjacent to the hypotenuse, which means that it is a calculation of the x -coordinate. When the sine and cosine of the angle are plotted against the size of the angle, the familiar waveforms are the result. These relationships are the basic tools that allow detailed analysis of EEG data.

THE FOURIER TRANSFORM

Of all the mathematical discoveries that have benefited EEG research, perhaps the contributions of Joseph Fourier (1768–1830) have had the greatest impact. Indeed, the Fourier Transforms may have made the field of neurofeedback feasible. Fourier conceptualized reduction of waves into component data represented by series of sines and cosines. His math made it possible to analyze waveforms with both accuracy and precision. The Fourier Transform reorganizes a wave signal into a format in which the individual frequencies can be analyzed.

Periodic data are comprise three *domains*: time, amplitude, and frequency. Normally, data are collected as a function of time, which means that the data set is in the time domain. By transforming the data to the amplitude domain, we can quantify and display amplitude data (Figure 18).

By transforming the data to the frequency domain (Figure 19) it is possible to calculate the characteristics for given frequencies. *Energy* can be described as the *variance of the frequency*, and power can be defined as *energy*

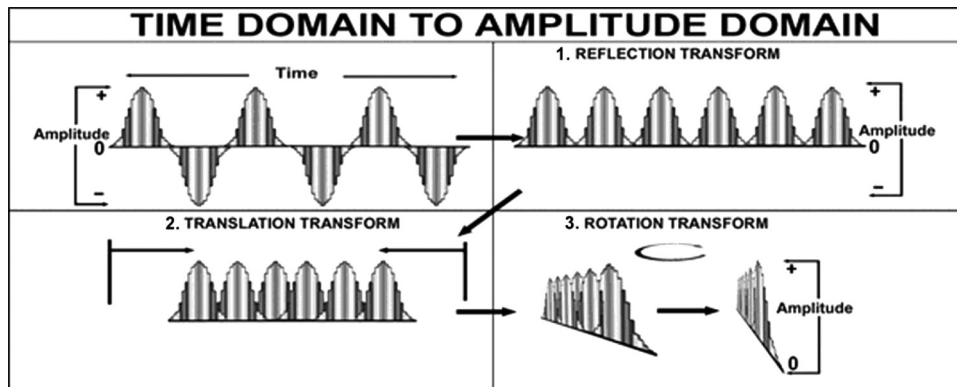


FIGURE 18. Transform from time to amplitude domain: 1) $y' = -1(y)$; 2) $x' = x/m$; 3) $z'(\sin\theta) = z(\cos\theta)$.

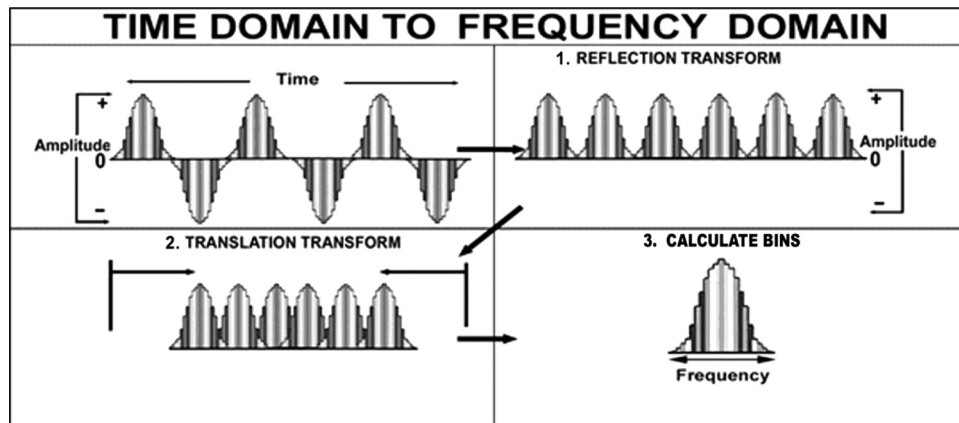


FIGURE 19. Transform from time to frequency domain: 1) $y' = -1(y)$; 2) $x' = x/m$; 3) calculate frequency bins.

per unit of time. When the wave data are transformed from the time domain into the frequency domain, this representation is commonly referred to as *spectral density* and represents the mathematical description of the power distribution in relation to frequency.

From the frequency transform both the *autospectral density* and the *cross spectral density* can be determined. The autospectral density value is calculated as variations in the power spectrum across time. The cross spectral density is calculated as variations in the power spectrum between waves.

Figure 20 further illustrates the basic Fourier Transform process. Technically, Figure 20 is calculating the *Discrete Time Short Time Fourier Transform* because each Fourier Transform is extended into a discrete time packet.

The Fourier Transform as it is applied to EEG analysis has two predominant formats: *discrete* and *fast*. The discrete transform works with a defined sample of a larger signal. The amount of data in a given signal is substantial and can involve literally millions of mathematical operations that can become ponderous. The FFT recognizes that for a finite signal sample, the actual number of operations required to be able to rebuild the original data set can be significantly less. The original number (N) of operations (O) required is $O(N^2)$, whereas the reduced number is $O(N/\log N)$. Numerically, for a set of 1,000 samples, this would reduce the operations from 1,000,000 to 333 (Reddy, 2005).

The familiar output of the Fourier Transform is the *Continuous Time Short Time Fourier Transform* spectral display, which is often referred to

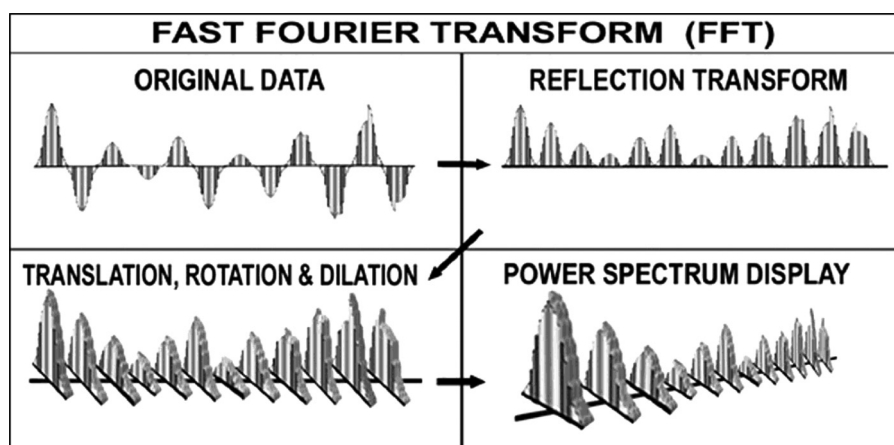


FIGURE 20. Graphic representation of an interpretation of a Fast Fourier Transform.

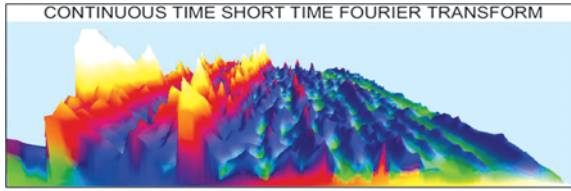


FIGURE 21. Power spectrum as Continuous Time Short Time Fourier Transform. (Color figure available online.)

as a *spectrogram*. This output device allows inspection and analysis of EEG data with the emphasis on the frequency–amplitude relationship as it changes across time (Figure 21).

COMPARING WAVES

Once we have analyzed the EEG waves within the parameters of study, it is generally desirable to compare EEG activity between various sites. To accomplish this comparison, another set of mathematical tools is employed that allows us to create new data sets by combining information from more than one location.

Figures 9 and 17 illustrate the principle of measuring periodic data as degrees of a circle. As periodic information passes through positive values to negative then back to $y=0$, it will have passed through 360° of arc. *Phase synchrony* as it applies to EEG describes the difference along the time axis between two waves. The relative position of one wave is calculated as degrees of arc around a circle ahead of (+) or behind (–) the other wave (Figure 22).

One of the most useful mathematical tools for comparing wave data and indeed for comparing any data sets is the use of the mathematical ratio. The basic ratio model that most ratios used in neurofeedback follow is the *ratio of differences*. All ratios that depend on this characteristic use some variation of the relationship that defines the differences between two data sets relative to the combination of the two data sets. In fact, this relationship has already been introduced in this article. The Correlation Coefficient (r ; Figure 13) and both the Z scores and t test (Figure 16) are applications of this principle.

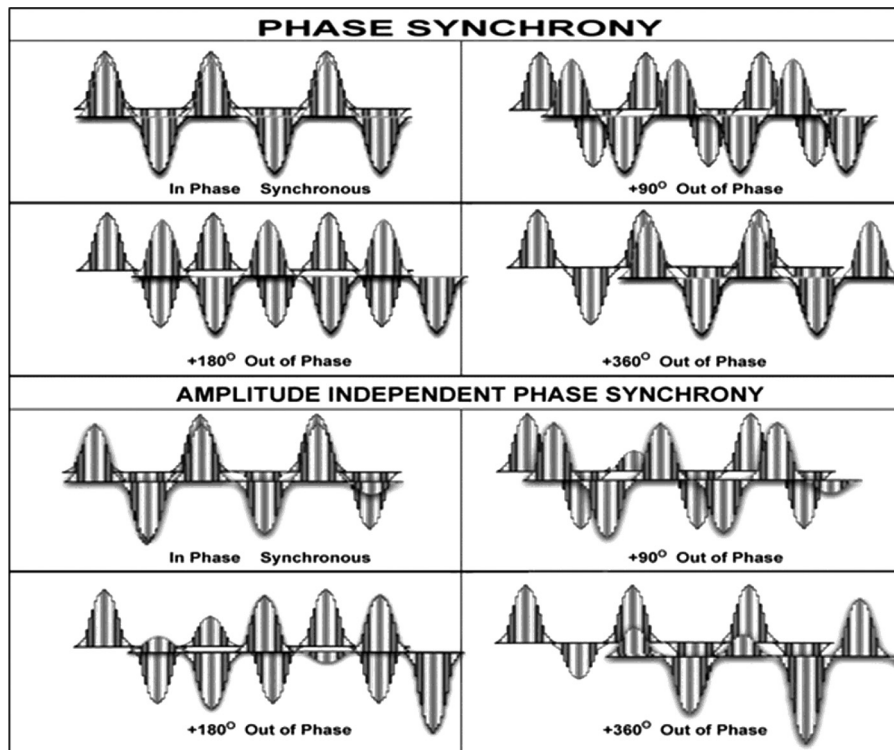


FIGURE 22. Examples of phase synchrony.

Common Mode Rejection is one of the ratios on which we depend in EEG biofeedback. In fact, the complete term is Common Mode Rejection Ratio (CMRR). The mathematics can be represented as follows:

$$CMRR = 10 \log_{10}(A_v/A_{cm})^2$$

A_v is the differential voltage gain, and A_{cm} is the common mode voltage gain. As this ratio of squares gets larger, the signals common to the two electrical outputs will be attenuated while the signals that are different will be amplified. The CMRR for biological data is typically greater than 1000:1. By calculating the CMRR as a log value, the large variations found in EEG data can be represented by smaller numbers. Because the CMRR is a ratio of exponents (squares in this case), the unit of measurement is the *decibel*. A CMRR of 1000:1 would be +30 dB. With modern operational amplifiers CMRR values of +80 dB are possible and desirable (Sanei & Chambers, 2007).

Inherent in EEG biofeedback applications is the placement of *active electrodes*, *reference electrodes*, and *ground electrodes*. The function of the reference electrode is to provide a passive measurement that is subtracted from the activity moving from the active electrode to the ground. The common mode rejection ratio is used to eliminate unrelated data by subtracting out information that we relegate to the category of "noise."

By comparing the function of electrodes to a soccer game, we can let the two teams represent active electrodes and the referee represent the reference. Because the referee is in the middle of the game, his observations of the actions of the players are assumed to be accurate. If we consider the boundaries of the playing field to be the equivalent of the ground electrode, then it is possible to identify the function of the ground. The ground is essential for the actions to be completed (a goal or out of bounds), and the reference is the standard against which the actions are measured. Were the referee placed anywhere other than on the field, his measurements would be suspect at best if not inaccurate.

If the referee has good stamina and a good understanding of the game, then his similarities to the players make his observations more reliable. However, if the referee had no prior soccer experience and was physically out of shape, these differences from the players would introduce additional factors into his observations reducing his accuracy. In this case the CMRR would be low. For CMRR a higher value is desirable and a lower value indicates a greater amount of "noise" in the signal.

Comodulation is another valuable application of the ratio principle for analyzing EEG data. An exceptional definition of comodulation was offered by Jacobson (2008): "Comodulation refers to the property that for a given source, there are likely to be relationships among its spectral components, such that they will start/stop at the same time and will rise/fall in amplitude and increase/decrease in frequency at the same rate" (p. 19). Essentially, the principle of comodulation is that one of three conditions will be analyzed: (a) Spectral components of two waves will start and stop, (b) amplitudes will rise and fall, or (c) frequencies will increase or decrease together. Consequently, there are several types of comodulation possible. Researchers have employed amplitude-phase comodulation as well as phase-phase comodulation. The form commonly referenced in EEG literature is amplitude-amplitude power comodulation and is generally reported as a correlation coefficient (Shirvalkar, Rapp, & Shapiro, 2010).

Collura (2008) noted the similarity between the equation for comodulation and Pearson's correlation while pointing out that the comodulation measurements are "amplitudes across time."

Coherence is somewhat more complex in theory. Collura presents an excellent working definition of coherence. Basically, it is the ratio relationship between cross spectral density and the autospectral density as previously described. The calculation of spectral density utilizes FFT values to provide a numeric description of the similarity between waves of individual frequency characteristics compared to the behavior of waves with those same

characteristics across a defined period. It is important to note that two waves with morphological similarities that make them highly coherent may also have low phase synchrony as well as low comodulation. These three factors are separate calculations that define unique properties.

The final analytic tool in this article is a remarkably valuable application of the ratio relationship called the *Bray Curtis Dissimilarity* (Bray & Curtis, 1957). This simple ratio provides two useful functions when comparing two sets of data: It creates a set of normalized values, and it establishes a unique index of dissimilarity for each pair of data sets. In its simplest form the equation describes the amount of difference found in a unified data set:

$$BCD = (A - B)/(A + B)$$

The normalized values fall between 0 and 1, with 0 indicating no dissimilarity and 1 indicating complete dissimilarity. If the values range between 0 and -1 , this indicates that the second value (B) is weighted more heavily than the first. If weighting is not a consideration, then use absolute values in the numerator. If percentage of difference is required, multiply the BCD by 100.

Determining the characteristics of a defined difference as it relates to the combined field of two data sets is in essence the objective of a substantial portion of mathematical analytic techniques. Although the complete understanding of all the mathematical formulae involved is outside the scope of many researchers and clinicians, grasping the general principles allows the knowledgeable implementation of many wonderful tools in potentially creative applications that removes the shadowy darkness from the realm of many black boxes.

REFERENCES

- Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs*, 27, 325–349.
- Brown, T., LeMay, H. E., & Burstead, B. E. (2006). *Chemistry: The central science*. Upper Saddle River, NJ: Prentice Hall.
- Collura, T. F. (2008). Toward a coherent view of brain connectivity. *Journal of Neurotherapy*, 12(2–3), 99–111.
- Graziano, A. M., & Raulin, M. L. (2000). *Research methods: A process of inquiry*. Boston, MA: Allyn & Bacon.
- Harmon, L. D. (1973). The recognition of faces. *Scientific American*, 229(5), 71–82.
- Harris, F. J. (2006). *Multirate signal processing for communication systems*. Upper Saddle River, NJ: Prentice Hall.
- Hartman, W. M. (1997). *Signals, sound, and sensation*. New York, NY: Springer-Verlag.
- Jacobson, D. B. (2008, September). *Combined channel instantaneous frequency analysis for audio source separation based on comodulation*. Manuscript submitted for publication. Retrieved from <http://stuff.mit.edu/people/bdj/Thesis20j.pdf>
- Kaplan, E., & Kaplan, M. (2010). *Bozo sapiens: Why to err is human*. New York, NY: Bloomsbury.
- Marks, R. J., II. (1991). *Introduction to Shannon sampling and interpolation theory*. New York, NY: Springer-Verlag.
- Reddy, D. C. (2005). *Biomedical signal processing: principles and techniques*. New Dehli, India: Tata McGraw-Hill.
- Sanei, S., & Chambers, J. A. (2007). *EEG signal processing*. West Sussex, UK: Wiley & Sons.
- Shirvalkar, P. R., Rapp, P. R., & Shapiro, M. L. (2010). Bidirectional changes to hippocampal theta-gamma comodulation predict memory for recent spatial episodes. *PNAS*, 107, 7054–7059.
- Shoemake, K. (1994). Euler angle conversion. In P. Heckbert (Ed.), *Graphics gems IV*, (pp. 220–229). San Diego, CA: Academic Press.
- Skinner, B. F. (1938). *The behavior of organisms*. New York, NY: Appleton-Century-Crofts.
- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Appleton-Century-Crofts.
- Thatcher, R. W., & Lubar, J. F. (2009). History of scientific standards of QEEG normative databases. In T. Budzinski, H. Budzynski, J. Evans, & A. Abarbanel (Eds.), *Introduction quantitative EEG and neurofeedback*, (pp. 29–59). New York, NY: Academic Press.