It Takes a Paradigm Shift to Reveal the Auditory Scene

John K. Bates, Jr., P.E., M.S. E.E.

It is astonishing how one simple incorrect idea can envelop the subject in a dense fog.

Francis Crick ["Lessons from biology," Nov. 1988

Chapter 1. A Clash of Paradigms

In the early 19th century, scientists Georg Ohm and Hermann von Helmholtz began a serious investigation of auditory perception. Working sequentially, they produced a hypothesis of sound perception that uses the mathematical Fourier transformation to convert sound waves into patterns of harmonic frequencies. Their hypothesis purports to explain how the ear could interpret sound as auditory patterns of groups of these frequencies. This method seemed correct for many simple auditory situations, although the theory also had some major known flaws. Despite this knowledge, the Fourier transformation became the established scientific paradigm that forms the basis of all subsequent theories and represents the current scientific paradigm.

A particular example of this situation is the Ohm-Helmholtz hypothesis, called spectrum analysis which was regarded as an explanation for human auditory perception but, for over 150 years, has failed to explain *any* of the critical auditory functions. For example, it cannot explain how the ear separates the eclectic sounds of the acoustic environment, the auditory scene. While spectral analysis has been helpful for single-channel auditory processing, scientists have long been bewildered by the paradigm's inability to solve major vital problems, such as responding to a multiplicity of other experimental anomalies. Nevertheless, all these unanswered anomalies signify the paradigm's failure to progress beyond the present state of the art. Logically it would be expected that a new paradigm is the only answer. It will be, and here I'm about to describe the paradigm. First of all, it must be a shift from the paradigm of the frequency domain to the paradigm of the dimension of time: Or more accurately, the new domain shall be the *time-space domain*, where "space" is composed of "neuron-logical" networks (not "neural" networks) of multidimensional logic memory arrays such as has recently been discovered in the human brain. And that is precisely what my auditory theory uses and predicts. These claims are not an exaggeration. I will show data to demonstrate the essential functions that do what I've just said. In any case, I think the story itself has been an interesting one.

How did this current flawed hypothesis remain for 150 years as the auditory paradigm? The answer to this question is given in the narrative of this book. A quick response is that it has failed because the auditory scientific community has never even *tried* to understand the basic principles of biological sensory systems. In this day of magnificent (and expensive) scientific achievements in the physical sciences, we don't know how the ear works or how any of the brain's sensory systems work. Of course, that boils down to understanding the brain itself: How could that be?

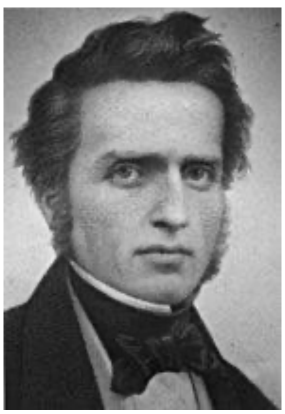
Here is the point: It's the scientist. They haven't asked why? They observe an object or process and immediately ask how? If they don't know *why* a thing does something, they can't judge *how* it can do what it does.

t all boils down to the fact that most science uses Aristotle's traditional scientific method of scientific discovery, whereas understanding animal sensory systems poses a complex system problem in system construction. So, is it not essential to know *why* the animal has ears? The quick answer is to say," It's to hear sounds." But it is more than that;

the answer is to help the animal *stay alive.* That answer means that staying alive requires a system, so the scientist must analyze the system's requirements and constraints to build and understand the how? That is what engineers do, not scientists. Therein lies the fundamental problem; these things have not been done, and the wrong decisions have been made, all from the beginning.

Let's examine this situation. It has been a portrayal of improper science. The story goes like this: Georg Ohm studied the newly-created mathematical Fourier transformation. He conceived that acoustic waves could be transformed into the "frequency domain." In this domain, the acoustic waves would become a spectrum, a pattern of sinusoidal harmonic frequencies. These frequencies could form graphic patterns that characterize specific meanings, such as the phonetic patterns of human speech sounds and other sound sources. This looked like a significant breakthrough, a theory of auditory perception that could make him even more famous. (He was already recognized for his discovery of electrical resistance.)

However, another physicist of lesser status, August Seebeck, was skeptical of Ohm's claim that the "first Fourier harmonic" was the pitch tone of human speech. In complex tonal sounds of many harmonics, such as human speech, he had observed that there was no fundamental "pitch" harmonic; he called it the "missing fundamental," which showed that Ohm's theory was incorrect. He had already devised an alternate approach called the "periodicity pitch." (That was incorrect too, but on the right track." So a serious rivalry grew up between Ohm and Seebeck. He



constructed an ingenious siren, a rotating disk with precisely placed holes that an air jet would blow through and generate tones of precise desired pitch and timbre to prove his point. He discovered that when he adjusted the siren so that its sound waves had *no* "*first harmonic*," listeners could still hear the pitch tone. This discovery destroyed Ohm's hypothesis. But Ohm wasn't convinced. He humiliated Seebeck by arbitrarily dismissing his theory. So Seebeck continued his argument among his scientific peers and demanded that Ohm admit his hypothesis was wrong. After debating this at length, the peers forced Ohm from the field. Then Ohm went to brilliant scientist Hermann von Helmholtz for support through experimentation on the Fourier hypothesis. These experiments convinced Helmholtz that the spectral theory was correct. One could question this decision since the experiments were subjective and observed *only* by himself.

Therefore, Seebeck did not give up. He invited both of them to witness his siren experiments, but they ignored him and claimed his experiment was somehow flawed, and after all," Wer das Teufel ist das guyisch Seebeck?"

Nevertheless, the spectral frequency hypothesis became the settled paradigm of auditory theory, disregarding the inconsistencies of Fourier-based anomalies that showed up over a century of experimentation. Apparently, what kept it alive is that, deep down among scientists, it has always been a significant source of intriguing scientific puzzles. It has never been regarded seriously as a perception system to be understood but as a seemingly endless source of profitable scientific fun.

This is where my own story comes into play. Looking back at poor Seebeck, I have much empathy for him," To quote Bill Clinton, "I feel his pain. (Notice Seebeck's angry expression in the photograph.) Seebeck and I have faced the same situation, showing that the prevailing auditory paradigm is wrong and having our work dismissed without cause. Seebeck's dismissal was understandable in that, while correctly indicating that the theory of spectrum transformation of sound waves is untenable, he could not propose an alternative basis for a time-domain theory of auditory perception. Seebeck could have foreseen that he would have to form an auditory theory based purely on the time domain because there is no alternative.

On the other hand, from what I have learned in my own experience, he would have struggled to build a plausible theory using early 19th-century technology. It's the technology, stupid! (Clinton, again) The best he could have done would have been an operational analysis of system requirements that a theory would have to explain. That *should have been the first step i*n any ventured scientific theory, not a frog leap to the next lily pad.

Only with the advent of integrated microcircuitry and digital computers in the 1960s could anyone conceive of signal processing systems that could implement the extensive logic arrays needed to implement my auditory modeling systems. The time domain is a much different animal than the frequency domain because you have to think in analog logic, where neurons live, not where math equations rule. Yes, the world of microtechnology made my time-to-space theory possible and plausible.

But Seebeck had won a moral victory over Ohm and Helmholtz. In contrast, I have been trumped by the full force of the Fourier paradigmistas, who have opposed and suppressed the general publication of my demonstrable facts. Some of those events make for exciting conspiracy theories.

I must let this book judge whether my theory wins or loses. I know that it should win in the long run. In a hopeful way, I would like this tale to be both interesting and convincing. My overriding problem is that explaining much of the technical stuff is outside the auditory box because my career thinking has primarily been in the time domain. [Reference: R.Steven Turner, etc., "The Ohm-Seebeck Dispute, etc."]

An unscientific situation

The scientists themselves have caused the fact that this erroneous situation has existed. Instead of studying the existential requirements and constraints of the sensory system, they seem to have been infected by physicist envy. Mathematical solutions are always complete and neat, although often, such as the Fourier transform, plentifully supplied with assumptions. For example, the Fourier series is a mathematical infinite series, but the frequency-to-spatial spectral transformation is based on the premise of having a limited time duration, not infinite. And therein lies the problem: It takes a period to get a specified proper spectral pattern, whereas the required instantaneous response has a time limit of zero, hence no pattern in the infinite spectral frequency bandwidth. What a dilemma that has been! When you're in the time domain mathematically, that is, the dilemma does not exist. Besides, time is easy to measure. The cochlea does an outstanding job at that.

How can neat sets of mathematical equations explain life, whether on Earth or over there in the cosmos? To solve existential problems, you must dig deep into life's workings and understand what it takes to maintain it. Auditory scientists have not seen it that way. That is their greatest mistake. They've formed mathematical algorithms, but they are based on statistics. Statistics is another time-wasting process because it needs a time window, as the Fourier transform, to get useful data, which is then based on a prediction of future continuity of the present. We all know that predictions always get qualified as to their trustworthiness. Math does not apply to existential sensory systems; there is too much at stake in survival-threatening situations. That is why Nature has developed fundamental algorithms for building sensory systems that apply to all creatures. My objective has been to find them before making an auditory system.

As an engineer, the story of how auditory science has continued to make this mistake seems incredible, that mature scientists have never tried to understand even the kind of problem they've been trying to solve. Couldn't they see that auditory perception is fundamentally a complex system they must design and build, that it's not simply solving an amusing puzzle? [reference Thomas S. Kuhn] There can be no far-reaching theory to explain the operation of living systems: things that are alive. (OK … philosopher alert!) However, given that they have not understood the main problem, I will grant to scientists that finding the solution would have been such a daunting puzzle that even visualizing potential approaches would be hard to attempt. Besides, they ask, why risk a career on a wild goose chase? But that is because of their training and life's work, so it seems best for us to proceed with an engineering approach.

This is where the unique experiences of my career as an electrical engineer have given me the clues for finding the answer. I've been able to recall and use the analog time-domain technology of the pre-fast-Fourier era not available to smart young modern scientists. Therefore it has been relatively easy for me to understand and use the methods by which Nature extracts *meaning*, not information, from acoustic waveforms. So I intend this to be a lesson book on how I have arrived at and used these ancient concepts.

Against the classic scientific method, I will treat our search for an auditory theory as an engineering construction project. Let's consider that the purpose of an animal's auditory perception system is to extract the meaning in a sound waveform and report it to the brain. The brain sorts out the implications and decides how to use the information to maintain survival. The meanings the ear-brain needs are in the input signal waveform. Where else? Therefore, because the ear must operate synchronized with the waveform, it works in the time domain. Simple reasoning tells us that you will lose the meaningful waveform if you transform to an algorithm outside of the time domain. So why would you want to convert to the frequency domain? Helmholtz and his legions have unthinkingly done that, and we are now paying the price. Forgive Helmholtz: He never thought about waveforms, so he made an assumption. It was his followers that shared most of the blame. But nowadays, we can see the waveforms and extract meaning from them, and that is what the ear does, so that is what we shall do.

What I just said about the fallacy of converting to the frequency domain or any other domain implies that there can be no compromise. A few people have tried to get me to find some way to devise a hybrid. Why, I don't know, except to salve some hurt egos. Anyway, what will be needed is a change that can be accomplished only by a total shift in the central theory of auditory perception. There is no possibility that the present research scheme can ever qualify as a valid theory. Therefore almost everything I will propose and explain is opposite to the current auditory paradigm of thinking.

This oppositional viewpoint has been challenging for me. There is little or no "prior art" to help. The main outside help I've had is the literature on psychoacoustic and biophysical experiments by which I can compare results performed on simulations using computational models. The fact that I have been able to construct models is because, as I have realized lately, my models' biological-friendly time-domain technology is inherently the analog kind that is what was used in pre-digital vacuum tube/transistor signal processing. A good part of it was from my boyhood building ham-radio receivers, then to military radar systems of World War II. My work after the war continued in the radar field, designing, building, and flight-testing electronic intelligence ELINT airborne reconnaissance systems. It was the ELINT experience that pointed me inadvertently toward the field of auditory perception.

Much of what I have learned has come from the "trial-and-error school of hard knocks." So, building an auditory simulating substructure has taken a long time in different places and unusual assignments that generally had little in common. In reflecting on why my ideas had not been discovered before I did and why explanations of my methods have led to blank looks, I decided that it was because no one else in this field has had a similar experience. Therefore, it has been nearly impossible to communicate. Consequently, I feel that including my mixed background of experience could help pave the way for a better understanding of how I evolved an auditory model.

Understanding the problem: an ELINT analogy

As I've been saying, my thinking comes from the mind of an engineer because that is what I am. Since 1949 when I began my career in electronic intelligence, ELINT, systems, and operation, I've been designing, building, or thinking about complex airborne microwave radio signal reconnaissance systems that collected information over a wide frequency range. These systems operated in real-time, i.e., in sync with the signal. Accordingly, the process had to be using *time-domain methods*. Systems of auditory perception, ears, operate in the same way too. In radio and audio, the main problem of signal analysis was separating signals from multiple sources. On top of the separation, the analysis objective was to select the priority signal for attention in a given situation: Which meant that the situation had to be *known* at any instant.

With all those "extra" requirements, it's evident that simply separating sounds is not enough. You don't deal out a deck of cards and walk away thinking you've played a game. No --- you must sort the cards in your hand, evaluate the cards' meanings, and decide which one to play. That is the kind of thing auditory perception does; select, reject, plan, and choose, except for one more requirement: All of this analysis, thinking, and action must be done simultaneously or as close to that as possible. So how well could that be done in a system (imaginary, of course) using the current frequency-based paradigm? First, it would be impossible because we have the demand for the near-instantaneous acquisition of information from within the waveform. As we now know, all that is lost. So what is left in a spectral analysis that might be used? Specifically, what could the signal's spectrum tell us? Getting a spectrum takes much time if you intend to accurately distinguish the sources by their spectral patterns, which would be overlapping and obscure each other most of the time. In any case, the time needed to form a spectrum would be unacceptable. This is such a feckless concept that, oddly, the auditory community hasn't folded their hands and gone to the time domain long ago. (It's very odd.) Still, money is being poured into a game they can't win.

On the other hand, I can see their dilemma. They speculate: Should we move on to something else if the frequency domain is wrong and the time domain is correct, but none of us scientists know what to do? Why, yes! "Let's make music with many frequencies, tones, pitches, and perception issues." So that's where the technology has been going. They ignore that science that lurks in the music of multiple sound sources in choral and orchestral music. But this is not facing the problem. Without a viable auditory theory, soon there be no more auditory scientists, just composers of incomprehensible "music."

Now that I have commented on the fundamental problem, you scientists, it is time for us to examine the engineering problem. To begin a project, engineers must understand the problem to be solved. So, the first question is: What is there about the sense of hearing that is so hard for scientists to understand? By and large, you need to be very curious about things. You should focus on an interesting problem and ask why it's a problem. For example, I looked back through the history of auditory research. I became convinced that researchers have only paid attention to the "how" part of the problem and have never asked the more profound question: Why do the ear and brain exist? Evidently, they live to hear and understand sound, but not very interesting. So what happens is that the why-question goes unanswered. I decided to study that question, and I found the answer. Unsurprisingly, the answer is that sensory systems exist to enable living things to survive. And then, after knowing why they exist, it became possible to know *what had to be done* to maintain survival. So, I figured out how to use sound waves to keep an animal's survival.

It is hard to get a handle on something you cannot see. The intangibility of sound might explain why understanding auditory perception is difficult. In the other sensory systems, seeing, smelling, tasting, and touching objects are tangible and lasting. It's easy for people to mutually observe or sense them, discuss them, and think about them over extended periods. But with sounds, there are no lingering vestiges of a sound to think about. Sounds are ephemeral objects that exist only in memories stored in brains, and they aren't usually the same among other listeners. In a prime example, visual objects can be studied indefinitely and in great detail. That is why psychologists and philosophers have based most of their studies and hypotheses on visual experiments.

Consequently, their theories of human "consciousness" are founded mostly on their ability to see and understand visual things. But the necessary tangibility for understanding sound is sealed within the ear's cochlea, its tiny, bony box of secrets. Unlike the eyeball with its camera-like lens and retina, one cannot easily conceive how the cochlea converts sound waves into neural "spikes" that are meaningful to the brain. For over a century, scientists have been trying to discover or hypothesize ways to find observable evidence of the ear's performance to formulate a valid theory of audition.

My first clue about why and how the cochlea does what it does comes from my solution to a decades-old ELINT problem: separating radar pulse trains to measure their repetition rates and pitch frequencies. I called it a pulse repetition rate sorting matrix. It could operate over a wide range of frequencies. It had a set of logic matrixes connected in parallel with feedback to prevent recognizing false submultiples. It turned out that this used a logarithmic system that responded almost precisely like the ear's cochlea. Of course, I immediately knew I had the answer to what goes on in that little box. So from there, I was beginning to see the auditory system. That event was a bit of good luck. I got by successfully with an algorithm that produced the cochlear function's output but did not simulate its function, which uses a traveling wave to "scan" the cochlear hair cells. But it wasn't enough to convince the hard-nosed skeptics. Finally, I had to design a true-functioning cochlear model with a traveling wave. (This was in 2013.) It is successful and should be convincing. But I haven't had the facilities to do any more on the system. It makes me sad.

This answers the first question. Few scientists can mentally visualize the puzzle to be solved, and so do thought experiments. (Note that Einstein's uncle was an engineer, so maybe he had inherited the ability to do his famous "relativity" thought experiments from his father.)

Reprise of Ohm, Helmholtz and Seebeck

Here is another look at how the existing auditory paradigm became established. Remember that the auditory community has always had a frustrated, desperate need to visualize how the cochlea operates mentally. In the early 19th century, Georg Ohm discovered that the mathematical Fourier Series could convert sounds into harmonic patterns representing known sounds, such as spoken words. Then, Herman von Helmholtz came up with the idea that the ear's cochlea might have a neural structure of a row of resonating cavities like a piano keyboard that might convert musical tones into a spectrum of harmonic frequencies. (This is called a *tonotopic* arrangement.) He performed experiments by listening to tones and their "partials" (harmonics) that confirmed this assumption. This discovery was his "breakthrough." In the early 20th century, Georg von Beckesy experimentally backed up Helmholtz's spectral theory, which helped the spectrum to become the "settled theory" of auditory perception. This led to the discovery that a graphic display of a sound's spectrum could be evidence that characterizes its source. In this graphic form, the Fourier Spectrum converts sounds into tangible images that could be visible and permanent images. Now, scientists can present experimental evidence and write lots of scientific papers. From that beginning, the Fourier transformation became the basis of all acoustic and auditory theory. And thus, it conformed to Thomas Kuhn's requirements that define a scientific paradigm.

One powerful reason for its persistence is that, in areas other than the auditory, it became helpful in improving communication channels where acoustic information is processed one source at a time. Along with communications, the theory spread into other applications where complex patterns were analyzed, even in vision. In the 1940s, especially after World War II, communications theory brought more advancement, and with Bell Laboratories' Claude Shannon theory of information flow in communication channels. Then, in 1946 Dennis Gabor introduced some negative reality into the prevailing exuberance of Fourier analysis by showing mathematically that it had a fundamental problem with processing sound; a time-frequency limitation on the ability to separate sound sources because the accuracy of spectrum patterns depends on the amount of time allowed for the pattern to exist. Short intervals give poor shape resolution, creating poor pattern recognition among spoken words. It is impossible to create an auditory theory without the ability to separate individual sounds from different sources.

Hence, there would be poor automatic speech recognition and separation from a "noisy" background. In communications work, this noise was encountered mainly as the sound or internal system noise that interfered with detecting weak radio signals or communications systems. But in the late 1940s, the interfering background began to be regarded as another kind of noise created by many sounds that may or may not be important. Auditory psychologist Colin Cherry summed up the problem of separating sounds with his jocular "cocktail party problem": How do people in a cocktail party's acoustic melee and confusion manage to understand each other? How does the ear do that? But overall, there was not much interest in sound separation at that time, except, maybe, in the area of automatic speech recognition. But technical interest was mostly in generating "workarounds" rather than confronting the problem. Thus, it wasn't until Albert S. Bregman published his "Auditory Scene Analysis" in 1990 that sound separation came to the front of the scene.

The concept of the auditory scene became an inspiration to the entire field of auditory science. It was simply the scientific description of the cocktail party, a "to-do list" of things to accomplish to create an auditory theory. But it offered no solutions for any of the things to be done. In spite of all the theoretical obstacles, though, it brought hundreds of new scientists into the research effort.

Thus, in 1992 Bregman created an email Auditory List to discuss auditory perception centered at the McGill University Department of Psychology. It is currently maintained by Dr. Dan Ellis of Columbia University. To understand the depth of the problem, it is instructive to observe in the archives of the List the rise and fall of research in the auditory scene over the decade of the 1990s. Unsurprisingly, following the 90s decade, the emphasis has been on sound separation. It, alone, became the central problem to be solved. It was so serious that research centers and competitions were established, hoping someone could solve the problem. But it withstood all attacks, and how the ear separates sounds remains a mystery. That is, it's a mystery to the occupants of the reigning paradigm. But, to me, *why* it reigns a mystery is a mystery. It would not be a mystery if the auditory community saw that the answer was in plain sight. Even more mysterious is that I have pointed out exactly where it is in my many publications and presentations, but they refuse to acknowledge it. It must be a case of psychological mass denial, a *fear* that it would be necessary to shift one's life's work from the auditory paradigm to an unexplored one *founded in the time domain, an essentially unexplored domain.* Therefore, there is no "prior art" to judge my theory.

I'm asserting that the original choice of a Fourier power spectrum as the fundamental principle of the ear (or all animal ears) was and has been a drastic mistake. I do not blame Ohm and Helmholtz. All they did was show a possible solution. In 1844, their hypothesis was the only one with any experimental justification. At the same time, though, August Seebeck had shown that the frequency spectrum was a flawed model.

Nevertheless, science philosophers ignored the warning. And so, at some point along the way, none were curious enough to wonder how animals could do such amazing things in their abilities to process sound. If they had done their homework, they would have asked why animals have ears, sensory *systems,* and brains, and they might have found the answers. But they did not, so a *huge* body of research has become dependent on a flawed paradigm. This is because scientists have yet to understand, or even *tried* to understand, the problem they wanted to solve.

Now, the engineering approach:

And so here is where I found myself entering the auditory scene. I never intended to be here. it has been a slow process, a long and eventful story, and an evolutionary process in more ways than one. Also, I was able to exploit some lucky events that aimed me surreptitiously in an auditory direction.

In electrical engineering school, I learned the principles of Fourier transformation and the essential functions of frequency domain systems operating in the frequency domain. Still, I never became deeply involved with spectral analysis used in signal processing. It was simply not relevant to the requirements of my work. I designed systems for extracting reconnaissance information of electronic intelligence ELINT directly in the time domain that required a kind of “analog” signal processing. At that time, I never thought of that analog analogy; in those days, it was the only way to do real-time “computing.” But the memory of all those techniques came back. I saw how to use those methods to solve the problems of current system designs brought about by using the Fourier transformation. It was straightforward: *get rid of the transform and replace it using the early time-domain concepts* I had developed in my theory of auditory signal processing. I knew it would be possible to use those time-domain analog ideas because they led me to a couple of inventions that could solve the basic problems of the cocktail party effect.

Later I will show algorithms derived from fundamental principles that apply to animals’ survival. Algorithms are analogous to mathematical equations. If the algorithm is taken from fundamental principles, it should be as “general purpose” as a mathematical equation. To design a system, one plug in the parameter numbers and builds the system. Loosely speaking, that is the way I arrived at the auditory model.

I built a simple low-level model to demonstrate the fundamental principles of survival using the auditory sensor. It showed crude, primitive cognitive performance to test the concept, which was successful. Because a computer program can also be a general-purpose algorithm, a successfully-performing one should represent a correct auditory model or theory. Suddenly I realized that now I was aiming directly at a complete theory of auditory perception! That was an unexpected, exciting feeling*. It was an objective far out of my league.* Yet it seemed attainable. So I’m taking it seriously.