Chapter 2
Developments from 1950 to 1980

Abstract This chapter presents some details of technical and experimental progress that was made during the period 1950–1980.

2.1 Developments of Techniques and Tools for Vibration Measurements

2.1.1 Electronics

The advancement of electronic equipment in the 1950–1980 period has been overwhelming. Initially the improvement focused primarily on analogue equipment. Signal generators as well as meters and analyzers were developed to more demanding specifications, aiming at better controllable and reproducible experiments. For instance, a high quality oscillator which tended to be somewhat sensitive to temperature variation—and therefore providing a slightly shifting frequency during warm-up—produced very little distortion or background noise. In contrast, the first digital signal generator produced extremely stable frequencies, but it took decades of additional developments to achieve the same precision in amplitude and signal purity as had been available in the analogue equipment.

Analogue set-ups could already be automated using electronic control techniques and analogue computing equipment. However, analogue control techniques began to lose impact in the late 1960s and 1970s, when digital control started to take over.

High quality equipment became available in laboratories for biophysical, electrophysiological, and psychophysical experiments, often with specialized fine-tuning for specific tasks. The benefit of these developments was not limited to vibration measurement. It was essential for all branches of hearing research. This also applies to laboratory computers, which appeared around 1970, and rapidly became indispensable, slightly after the step from slide rules to calculators.
Miniaturization of electronic equipment started gradually with the invention and application of transistors. These replaced vacuum tubes almost completely over a short period, and further miniaturization of chips—including electrode arrays—has not yet stopped.

### 2.1.2 The Mössbauer Technique

Innovative as the experiments of Von Békésy were, the excitation patterns he had obtained were far too wide to be related to the frequency tuning observed in psychophysics. In order to make the motion of the BM visible, Von Békésy had to apply a driving force to the stapes which would correspond to sound levels of up to 150 dB SPL, well above the pain threshold at approximately 120 dB. It was therefore conceivable that his measurements destroyed part of the mechanical processes which he attempted to measure. Moreover, the measurements were performed on cadaver cochleae. If the mechanics of the intact cochlea would depend on, e.g., blood supply this also would have been affected in his measurements. Extrapolating his excitation profiles to lower levels and ultimately to the threshold of hearing convinced Von Békésy and others that something was missing, because the results indicated membrane motions less than 1 pm at the threshold of hearing (at \( \approx 20 \mu Pa \)).

For the explanation of the problem of frequency selectivity, Von Békésy started investigating the neural mechanisms of lateral inhibition, known from neurons in the visual system. A different solution, which was proposed by Gold in 1948, might solve both the problem of the frequency selectivity and the higher motion amplitudes needed at the hearing threshold. With a background in electrical acoustics, Gold proposed an active amplification mechanism which would counteract the damping at low levels. He even suggested that such an amplification mechanism might suffer from slight mistuning which would lead to a feedback loop producing sound spontaneously [see, however, footnote 10 in Sect. 3.4 and comments in Sect. 7.3].

In 1967 another new technology was applied to hearing research by Johnstone and Boyle: the Mössbauer technique. With this technique, data could be obtained at significantly lower sound levels, viz. within the normal range of hearing—although, at its upper end.

In 1957 Mössbauer discovered resonant and recoil-free emission and absorption of gamma rays in solids. This causes very narrow line widths in the generated gamma spectrum and allows the measurement of a Doppler effect when the source and detector move relative to one another. A small source emitting narrow band gamma radiation is placed on the cochlear partition and a detector is tuned to the radiated energy when the source is at rest. With this configuration, the rate of detected photons becomes a function of the velocity
of the source relative to the detector. Unfortunately, this function shows strong compressive nonlinear behavior. This limits undistorted measurements of velocities to a limited dynamic range.

Applying the technique to the cochlea is not straightforward. The gamma ray sources have to be prepared and carefully placed in an intact, living cochlea, and the measurement system must be accurately calibrated. But it allows much smaller motions to be measured than can be achieved by light microscopy. Rhode perfected the technique and in 1971 he was able to show measurements from cochleae of 20 squirrel monkeys at stimulus levels from 70 dB to 100 dB SPL. These measurements were the first to show that the cochlea clearly behaved nonlinearly at these, still relatively high, levels.

From 1971 on, several researchers started measuring the vibrations of the cochlear partition using the Mössbauer technique, showing the motion of the cochlear partition to be a nonlinear function of stimulus level. At first, attempts to reproduce the measured nonlinearity, even by Rhode himself, failed to reproduce the data. In hindsight, this was probably due to the vulnerability of the nonlinearity, which, as we know now, rapidly decreases after death of the animal or any damage to the structure of the cochlea.

2.1.3 Optical Techniques: Application of Lasers

The possibility to use (phase-) coherent light sources allowed a big step in the improvement of optical resolution, both for still and moving pictures. The spatial resolution was no longer limited to the order of magnitude of the wavelength of the source, but became dependent on the accuracy with which the phase of the light wave could be controlled and measured.

The principle of interference methods was recognized at the end of the 1950s, and patented by Minsky in 1961, but application of laser interferometry and confocal laser scanning techniques in biophysical experiments had to wait until the 1990s.

In laser interferometry, a laser bundle is split into two coherent beams. One beam—the reference beam—is sent to a detector as directly as possible. The other is focused on a moving reflector, the target, and the reflected light is sent to the detector. By studying the pattern created by the interference of the reference beam and the reflected beam, the velocity of the target can be deduced. Applications to cochlear and lateral line hair cell research were developed adjacently by Khanna (1986); Khanna et al. (1986) and van Netten (1988).
2.1.4 Pressure Measurement

Two lines of development of pressure measurements are important for cochlear mechanics (CMs). The first involves the outer and middle ear, the second the cochlea.

Middle ear pressure measurements are measurements of acoustic pressure (in air) within the middle ear cavity. The employed microphones require proper coupling to the middle ear structure. Basically, they are specializations of normal audio-equipment (e.g., Puria et al. 1997). The most accurate microphones tend to be of the condensor type. They have an impressive dynamical ranges, and a rather wide spectral sensitivity. Both the low-level boundary and the frequency characteristic improve with increasing microphone size. However, over the period 1950–1980 the common diameter of high-quality microphones reduced from $1$ to $\frac{1}{4}$” without significant reduction of quality.

The intracochlear pressure is a basic experimental variable in cochlear mechanics, but experimental accessibility is demanding. Results of reliable pressure measurements within the (cat) cochlear were first presented by Nedzelntsky (1974, 1980). The measurements were difficult. Within the scala vestibuli (SV), data could be obtained over a 40–105 dB SPL range (measured at the tympanic membrane); for the scala tympani (ST), this range reduced to 75–105 dB SPL. The experimental difficulty is probably one of the reasons that the method did not become very popular, although it reappears from time to time, with slightly improved techniques. Dancer and Franke presented guinea pig results in 1980. Like Nedzelntsky, they used a fluid filled thin probe connected to a sensitive sensor, in this case a piezoresistive transducer. The outer tip diameter was 0.25–0.35 mm.

Pressure measurements are important components of reliable power flow measurements, which require the independent measurement of, e.g., local pressure and local volume velocity. The majority of techniques in current use focuses successfully on velocity measurements, and associated pressure data are usually lacking. They might be estimated on the basis of a local impedance, but that assumes a sufficiently valid linear analysis and associated data. The independent measurement is more reliable. However, it also runs into the problem of measurements at the edge of what is fundamentally possible, viz. the fundamental uncertainty principle. In the present case the product $\Delta p \Delta U$ is bounded ($p =$ local pressure, $U =$ local volume velocity). But there is also uncertainty along the time dimension: energy and power are not instantaneous quantities, but they imply averaging across some time window. The points are addressed in more detail in Sect. 5.1.3.1.

2.1.5 Development of Computing Power

Around 1950 the slide rule and computation tables were still very much in use for the analysis of numerical problems. These provided the only computation option
for cases where the analytic approach or approximation was not feasible. Computer centers started to develop, both in industrial and academic environments. Initially, however, these were primarily tied to (applied) mathematics departments. It took more than a decade before lab computers became available and another before they became powerful enough to gradually become indispensable. Among the first minicomputers used for data processing are the TX-0 and the LINC at MIT (e.g., Kiang et al. (1965), Sachs and Kiang (1968), and Goblick and Pfeiffer (1969)), predecessors of the Digital Equipment Corporation (DEC) PDP-line. The PDP-1 was produced in 1963, and in 1965 the PDP-8 systems became one of the first successful minicomputers. Around 1970 many PDP-4 and PDP-8 systems were used in many auditory research laboratories. There was some competition from systems from Honeywell (DDP), and Data General (Nova, Eclipse). Personally I have been interested in this development using minicomputers since the early 1970s for simple modeling (DDP, DEC, and Data General).

To conclude the period covered in this chapter: around 1980, the use of minicomputers in auditory research labs had become common. They were employed for experiment control, data processing, and simple modeling. It took another 15 years before these were replaced by still rather expensive workstations, and then by the gradually more powerful, and economical, modern laptops and workstations.

At the same time, another line of computing power virtually disappeared. Based on the properties of op-amps (operational amplifiers), economical analog computers were developed and applied in signal analysis and control theory. The advantage of these setups is that they operate using continuous time (no time sampling). However, the amplitude range has natural upper and lower bounds, determined by power supply and internal noise, respectively. Second order differential equations, as used throughout this book, can be represented by a three op-amp circuit. (The interested student should be able to find the relevant literature in electrical engineering textbooks as referenced in the 2005 review paper by Lundberg).

2.2 Anatomical and Physiological Progress

2.2.1 Cochlea: Application of Electron Microscopy

In 1955 the TEM (see Sect. 1.3.1.2) was first introduced into the field of otology (the study of the anatomy and physiology of the ear) by Engström. It meant a marked improvement of the amount of detail that could be observed. Studies by, e.g., Spoendlin and Flock now showed the structure of the hair bundles on top of the hair cells and showed these to contain actin, leading to their official name of stereovilli (from the Greek stereos = stiff and villi as the official term for actin-filaments attached to a cell). Kimura showed in 1965 that the longest outer hair cell stereovilli are firmly connected to the tectorial membrane by showing the imprints left on the underside of the TM. The TEM technique operates on thin slices of tissue,
and the high optical sensitivity of this microscope implies that it is very sensitive to the slice preparation techniques. These have been improved and (semi-) automated over time.

In 1969 the scanning electron microscope was used by Lim for a detailed study of the structure of the organ of Corti (e.g., Lim and Melnick 1971). The greater depth of view of the SEM pictures allowed a good representation of its complicated 3-dimensional structure. Although SEM works on surfaces instead of on slices, the surfaces needed are too small to allow access in intact preparations, but structure within a preparation can be conserved much better than for TEM. On the other hand, in order to receive significant reflection from the surface, this has to reflect electron beams, which requires proper coating of the surface. In other words, the natural surface is potentially distorted both by properties of the coating, and by the electron beam.

Nevertheless, both TEM and SEM have provided marvelous images of the static structure.

### 2.2.2 Outer Hair Cells

One of the benefits of the introduction of EM into the lab was that they provided convincing evidence that both inner and outer hair cells are directionally sensitive (Flock et al. 1962; Engström et al. 1962). They confirmed that the structure of the hair bundle, and its orientation are relevant, and Engström et al. reported that mammalian cochlear hair cells no longer have a kinocilium but only a remaining basal body.

Flock et al. confirmed that the hair bundle structure and the location of the basal body of the kinocilium indicate the directional sensitivity. This was found for guinea pig OHCs, and for vestibular hair cells. Up to 1956 microphonic potential data from lateral line had been interpreted as indicators of bi-directional sensitivity of cupular hair cells (e.g., Kuiper 1956), but now that hypothesis was replaced by the notion that two sub-populations with opposite orientation can account for the microphonic data—as long as there is a nonlinear stimulus–response relation, because otherwise the microphonic response would disappear. These ideas were confirmed and accepted almost immediately.

Up to 1970 the difference in function between IHCs and OHCs remained a matter of speculation. The morphological differences had been established convincingly, and also the results of the first extensive studies of the afferent (and also the efferent) innervation had been presented (Spoendlin 1970). At the time there seemed to be no relevant role for the OHCs, in particular because of the distribution of afferent nerve fibers: Spoendlin estimated that at least 90% of the approximately 50,000 fibers innervate IHC and that 5–10% might innervate OHCs. At first, this was a matter of serious dispute (e.g., Eldredge 1967), but later studies confirmed the basic distinction and disproved the objections.
During the 1970s and 1980s, it became clear that there was a correlation between HC-damage caused by noise exposure and auditory nerve response. More in particular: the amount of damage depended on exposure level and duration, and damage started at the OHCs, beginning at the outer row, and finally reaches IHCs. As long as IHCs are (almost) intact, the auditory nerve fibers do give a response, but the tip of the tuning curve deteriorates with increasing damage level. IHC damage, finally, leads to deterioration of the cell and of the nerve fibers. Active nerve fiber labeling techniques have been used to confirm the nerve fiber—IHC connections (e.g., Liberman 1984; Liberman and Dodds 1984a,b; Liberman and Kiang 1984).

This conclusion was reached at the time that the source(s) of cochlear nonlinearity and a cochlear amplifier were still unknown. Now the pieces of the puzzle fell into place, and the role of intact OHCs appeared to become clear, or at least to possibly fill the gap. Detailed study of the OHC cell body indicates that its cytoskeleton is involved in electromotile responses of the cell (e.g., Santos-Sacchi 1992; Kalinec et al. 1992; Frank et al. 1999). Although this tells us in what direction progress can be made, it has not yet provided answers to all open questions.

### 2.2.3 Auditory Nerve and Beyond

Progress in auditory neurophysiology over the 1950–1980 period was also impressive. The basic techniques had just started (see Chap. 1), but biophysical understanding of the basics of the neural action potentials is attributed to Hodgkin et al. (1952); Hodgkin and Huxley (1952b,a,c) and did not develop before the early 1950s.

The primary auditory nerve fibers can be considered the output channels of the cochlea, and in this view the nerve response is the relevant output response of the cochlear system. However, the nerve response is not a simple analogue response, but a spike (=neural action potential) train, and the relation between stimulus parameters and response characteristics is not trivial. Although the response remains a deterministic, causal process, it becomes complicated enough to be characterized more efficiently in stochastic terms. Details of the transform of the continuous stimulus to discrete neural response can be avoided by using threshold detection techniques and constant response techniques. With the threshold technique, which requires the detection of a signal response in comparison to the background noise (spontaneous activity) it was possible to measure tuning curves–band-filter transfer functions—of single auditory nerve fibers (e.g., Kiang et al. 1965). It was immediately clear that the results were much more in accordance with psychophysical tuning than with the BM tuning observed by von Békésy. The same group also measured a large set of click responses, using positive and negative clicks, termed condensations click with inward stapes motion and rarefaction for the opposite. The Post Stimulus Time Histograms (PSTH) in response to clicks show a response profile that is equivalent to a half-wave rectified impulse response; click potential reversal evokes the other half-wave rectified part.
Response strength increases with stimulus strength, but saturates at some spike rate. The response area, used as a term for the CP range innervated by responding nerve fibers, also increases with level. In other words, as the level increases, the number of responding nerve fibers increases. This might seem a solution to the limitation caused by the saturation of the responses in single fibers, but, unfortunately, this only works for narrow band signals. The saturation properties are not analyzed in detail; they are addressed briefly in Sect. 7.1.1.

At the same time, other groups were studying time domain properties of the neural responses: in particular, the group of Rose and Hind focused on questions of waveform representation in the precise timing of the spikes. They found—for sufficiently low frequencies—a good synchronization to periodic stimulus properties. For a single tone, the response follows the half-wave rectified sine wave, both in synchronized period histograms (PSTH) and in interspike interval histograms (ISI) (e.g., Rose et al. 1967).

But besides the interest in these linear properties, there was also increasing interest in aspects that appeared to be in conflict with a linear interpretation. Without aiming to be complete, we refer the broad interests in:

- Studies of 2-tone suppression, apparently a nonlinear 2-tone interaction rather than an adapting feedback sharpening process (e.g., Nomoto et al. 1964; Sachs and Kiang 1968; Hind et al. 1970; Arthur et al. 1971; Abbas and Sachs 1976);
- CT phenomena, where auditory nerve data gave the first strong indication that the origin of the phenomenon had mechanical origin (i.e., prior to the nerve) in the cochlea (e.g., Goldstein and Kiang 1968; Dallos 1969; Smoorenburg et al. 1976);
- Changes in shape of tuning curves and excitation patterns with level (e.g., Goblick and Pfeiffer 1969; Anderson et al. 1971; Kiang and Moxon 1974).

More limitations of the linear cochlea are listed in the next chapter in Sect. 3.8. They are discussed in Chap. 4.

### 2.3 Auditory Perception

Auditory perception involves the entire auditory system. Here we focus on some specific topics, which traditionally have been assumed to depend largely on primary auditory system properties.

Firstly, this concerns studies of temporal and spectral sensitivity and selectivity, such as the relation between critical bands and effective bandwidths of tuning curves, between auditory masking patterns and cochlear excitation patterns, and the time–frequency uncertainty relation $\Delta f \Delta t \geq 1$ in relation to sharpness of tuning ($Q_{3dB}$) and time resolution.

Although the tools and theories largely originate from linear systems analysis, it had been clear from the beginning of masking studies that the data showed unmistakable nonlinear effects. Wegel and Lane’s classical study clearly showed
that the growth of masking with masker level is nonlinear, especially if probe frequency and masker frequency are different (Wegel and Lane 1924). Moreover, there was a difference in the masking behavior above and below the masker frequency. In short, the masking pattern broadens significantly with increasing masker level, in particular on the hf-side of the masker.

The shape of the masking pattern of a fixed, narrow band masker has the characteristics of a band-filter. Its bandwidth became subject to several discussions. Several definitions are used, and sometimes not clearly separated: the $Q_{3\text{dB}}$ and $Q_{10\text{dB}}$ are defined as the ratio of the center frequency and the bandwidth determined at, respectively, 3 dB and 10 dB below the peak; the $\text{ERB}$—the equivalent rectangular bandwidth—is the width of the rectangular filter covering the same spectral area $\int H(f) \, df$, using the same peak level.

The critical band, CB, was introduced as by Fletcher (1940) a subjective effective bandwidth. It was introduced as the maximum bandwidth of a narrow band of noise over which all noise power integrates when masking a tone. Later, it played an important role in loudness computations (Zwicker). Some controversy remained between European (+Canada) and US laboratories about the width of the CB, where a difference by a factor of about 2.5 remains (Swets et al. 1962). The greater CB made it to a standard for loudness computation (ISO R 532, 1975 [to be revised status], in the US recently replaced by ANSI S3.4-2005).

Another line with a historical basis was picked up again in the 1960s, and concerns the study of aural CTs (e.g., Plomp 1965; Goldstein 1967). Since the description of the effect by the Italian violinist Tartini (1692–1770) in 1754—although von Helmholtz (1863, Chap. VII) mentions that the earlier inventor (1745) was the German organist Sorge—the phenomenon of the perception of combination tones had been known in the musical environment. Both Sorge and Tartini reported CTs at frequencies below the primaries or differential tones, Helmholtz also claims to have perceived summational tones.

Meanwhile the perception of the lf intermodulation frequencies (i.e., $f_{\text{CT}} = m f_1 \pm n f_2$, where $f_1$ and $f_2$ are the primary components, $f_{\text{CT}} \ll f_1 \ll f_2$) has been well documented and quantified. A difference has sometimes been reported and claimed between even-order and odd order ($m + n = \text{even or odd}$) CTs. The difference tone $f_2 - f_1$ is of the even-order type, the most common $2f_1 - f_2$ is odd. Unfortunately, the latter is also often called a cubic difference tone, because a cubic nonlinearity can generate this type of distortion component. Hence the suggestion that there is a cubic nonlinearity in the system. However, there is ample evidence that the nonlinearity cannot be cubic; in fact it appears to be much closer to a cubic root, or a power of 1/3 rather than 3 (e.g., Smoorenburg 1972). The perception of hf intermodulation products or overtones has not been confirmed. Aural CTs are generated only if the primary components are within a certain proximity. The effect is optimum at a primary frequency ratio of about 6/5.

A different line of research addressed psychophysical 2-tone suppression, which appears to be related to neurophysiological 2-tone suppression. This phenomenon was studied extensively during the 1970s. One of the breakthroughs was the introduction of the “pulsation threshold” technique by Houtgast (1972). The basic
result, similar in psychoacoustics and in neurophysiology, is that apparently a strong enough tone can not only mask a tone, or generate a CT, but can also suppress the response to an adjacent (in frequency) simultaneously presented tone.

References


References

Cochlear Mechanics
Introduction to a Time Domain Analysis of the Nonlinear Cochlea
Duifhuis, H.
2012, XIV, 262 p., Hardcover
ISBN: 978-1-4419-6116-7