Signal Processing Challenges for Active Noise Cancellation Headphones

Stefan Liebich, Johannes Fabry, Peter Jax, Peter Vary

Institute of Communication Systems, RWTH Aachen University
Email: {liebich,fabry,jax,vary}@iks.rwth-aachen.de

Abstract

The evolving field of headphonce, particularly headtubes, and current developments in low power audio processing chips raise the interest in active approaches to address the problem of ambient noise. Methods of Active Noise Cancellation (ANC) are well suited as a supplement to passive attenuation, as they work specifically well at low frequencies. They are offering more flexibility to account for changing acoustic conditions and requirements. This contribution gives an overview of the ANC headphone problem and addresses the specific algorithm design challenges.

1 Introduction

Humans in modern society are more than ever exposed to noise in their everyday life [1]. Acoustical passive solutions may already reduce the amount of noise at its source, e.g. directly at the motor of a car, or on the transmission path, e.g. by noise damping windows. However, in some situations the passive attenuation at the receiving end, the human ear, is more feasible and cost efficient. This is typically achieved by ear muffs or ear plugs. The attenuation at low frequencies is insufficient using this passive approach. It can be supplemented with active signal processing approaches which rely on the principle of destructive interference. A loudspeaker emits a cancellation signal to attenuate the present noise. This is typically known as Active Noise Cancellation (ANC). A prominent ANC application is the so-called Noise Cancelling headphone or in-ear headset. The performance of an ANC system is limited by the physical and technological constraints as well as the algorithmic design. The technological design of the headphone comprises the acoustic front-end — including headphone casing, technology and positioning of loudspeakers and microphones — and the electronic back-end — including the realization of the filter creating the cancellation signal. This filter can either be implemented with analog or digital electronics. The third part influencing the performance is the algorithmic realization of the filters. They can be implemented in a time-invariant way, based on offline optimization for an average situation, or in a time-variant way, with real-time adaptation to a current and individual situation.

In this contribution, we provide an overview of the challenges of building an ANC headphone, including acoustic front-end, electronic back-end and algorithmic realization, for the example of an in-ear headphone. We briefly introduce the two main topologies for ANC algorithms and describe the technological and application specific challenges that need to be considered for designing a system that works under real-life constraints.

2 Active Noise Cancellation Principles

To create as cancellation signal, information about the disturbance is necessary. This information can be provided by an internal microphone, recording the inner disturbance, and/or an external microphone, recording the ambient noise. Fig. 1 shows the structure of an ANC headphone, divided into acoustic front-end and electronic back-end. Using the external microphone signal $x(n)$ and a filter $W(z)$ to create the cancellation signal $y(n)$, the system is called a feedforward system. When using the internal microphone $e(n)$ and a filter $K(z)$ to create the cancellation signal $u(n)$

\[ u(n) = W(z) \cdot e(n) \]

\[ y(n) = K(z) \cdot x(n) \]

\[ \hat{y}(n) = \hat{G}(z) \cdot \hat{u}(n) \]

\[ \hat{G}(z) = \hat{G}_{f,DAC}(z) \cdot \hat{G}_{f,AA}(z) \cdot \hat{G}_{spk}(z) \cdot \hat{G}_{AI}(z) \cdot \hat{G}_{mic}(z) \cdot \hat{G}_{AA}(z) \cdot \hat{G}_{ADC}(z) \]

The acoustic feedback degrades the feedforward performance and can lead to instabilities.

2.1 Feedforward ANC

In a feedforward system, the cancellation signal $y(n)$ is created by filtering the outer disturbance signal $x(n)$ with the cancellation filter $W(z)$ to create $\hat{y}(n)$, neglecting the feedback filter for this introduction ($K(z) = 0$). It relies on the causality between

\[ G(z) = \hat{G}(z) \cdot \hat{G}_{f,AA}(z) \cdot \hat{G}_{spk}(z) \cdot \hat{G}_{AI}(z) \cdot \hat{G}_{mic}(z) \cdot \hat{G}_{AA}(z) \cdot \hat{G}_{ADC}(z) \]
The superposition of a sinusoidal disturber $\Delta$ with the relative amplitude deviation $\Delta A$ results in the

$$ P(z) = W(z) G(z), $$

(1)

As $G(z)$ includes an acoustic path, it is non-minimum phase for sufficiently high sampling rates and thus not invertible. One possibility is to calculate the optimal filter and only consider the causal part of it [5]. For the feedforward system design, the primary path $P(z)$ should have as much delay as possible and the secondary path $G(z)$ should incorporate as little delay as possible. In [6] we derived an analytical expression for the necessary accuracy of an anti-phase signal for a tonal disturbance. It can be interpreted as a qualitative measure for the achievable performance depending on the system latency in Sec. 3.2 resulting in Fig. 2. We regard the superposition of a sinusoidal disturber $d(t) = A \cos(\omega t)$ with a compensation signal $y(t) = B \cos(\omega t + \Delta \phi)$, with phase error $\Delta \phi$.

$$ e(t) = A \cos(\omega t) - B \cos(\omega t + \Delta \phi). $$

Relation (4) is applied to derive a qualitative measure for the achievable performance depending on the system latency in Sec. 3.2 resulting in Fig. 2.

### 2.2 Feedback ANC

When the internal microphone signal $e(n)$ is used to create the cancellation signal $u(n)$, the system is called a feedback system. To separate the feedback system, the feedforward filter is neglected here ($W(z) = 0$). The error $e(n)$ is filtered by $K(z)$ and fed back into the ear by the loudspeaker. A feedback system requires a low latency from disturbance recording to interference with cancellation signal. For the feedback system design, the secondary path $G(z)$ should have as little delay as possible. The complete latency is encapsulated in the so-called open-loop transfer function $L(z) = K(z) \cdot G(z)$. The relation of $e(n)$ to $d(n)$ in the z-domain describes the overall active attenuation of the feedback system, also known as the feedback sensitivity function

$$ S_{FB}(z) = \left. \frac{E(z)}{D(z)} \right|_{FB \; on} = \frac{1}{1 + K(z) G(z)}, $$

(5)

To minimize the error, (5) has to become small, and thus $K(z)$ has to become large without creating instability. Furthermore, the variation of $G(z)$, typically referred to as its uncertainty, should be as small as possible, especially in the frequency range of desired attenuation [7].

### 3 Technological challenges

The technological challenges involve the design of the acoustic front-end, including loudspeakers, microphones and cabling, as well as the design of the electronic back-end, which processes the microphone signals and drives the loudspeaker. The requirements of the feedforward and feedback ANC algorithms are as much delay as possible in the primary path $P(z)$ and as little delay as possible in the secondary path $G(z)$. These two paths are defined by the design of the acoustic front-end and the electronic back-end, which will be broken down in more detail in the following.

### 3.1 Acoustic Front-End

The choice of the loudspeaker technology induces a few restrictions based on their acoustic properties. Generally, in consumer headphones dynamic drivers are dominating. They are able to provide high sound pressure levels (SPL) at low frequencies and are well suited to provide sound for the whole audible frequency range. However, they have the drawback of large size and high energy consumption. For hearing aids usually magnetic or balanced armature (BA) drivers are employed. Their main advantage is a lower energy consumption. Furthermore, they are typically optimized for a specific frequency range and thus often combined to multi-loudspeaker systems with crossover circuitry to cover the whole audible frequency range. However, BA drivers are only able to provide sufficient sound pressure levels at low frequencies for a closed volume. This can be a downside when considering semi-closed or open applications.

Another important aspect is the maximum SPL and the total harmonic distortion for large SPL. Here, dynamic drivers typically provide better properties. When non-linear distortion is audible to the user this is not only unpleasant, but the non-linearities also degrade the performance of algorithms, which usually anticipate a linear system.

The loudspeaker and the internal microphone should be kept in close distance to each other to minimize the delay within the secondary path $G_A(s)$. The reference and the error microphones should be positioned such that the group delay of $P_A(s)$ is maximal. A good sealing of the ear canal is beneficial as it introduces additional group delay to $P_A(s)$ and keeps $G_A(z)$ small. The acoustic feedback $G(z)$ needs to be considered in the design of the acoustic front-end and the algorithms, as the feedforward ANC system could get instable if $G_A(z)$ is too large.

### 3.2 Electronic Back-End

In the field of ANC ultra low latency processing is necessary to achieve good performance. Additional latency in creating the cancellation signal directly imposes boundaries on the achievable performance. Figure 2 gives a novel view on the upper frequency to achieve a given attenuation depending on the latency of a system. It is deduced for a sinusoidal disturbance signal $d(n)$ with perfect amplitude estimation $\Delta A_{rel} = 0 \; \text{dB}$ following (4). The attenuations $0, 5, 10, 15, 20 \; \text{dB}$ correspond to phase deviations $\Delta \phi = \{60^\circ, 32.7^\circ, 18.2^\circ, 10.2^\circ, 5.8^\circ\}$. The minimum achievable latency with different soundcards and systems is indicated on top of the figure (Analog Devices ADAU 1777 [8], dSPACE DS1005 with DS2004 and DS2102 extension boards (48 kHz), RME Babyface (96 kHz), RME Madi FX with real-time Linux (48 kHz), Bele IO card for BeagleBone (88.2 kHz) [9]). For analog circuits the latency in Fig. 2 corresponds to the group delay of the implemented filter. Especially for feedforward systems, this gives a good idea of the influence of additional latency. For feedback control algorithms the boundaries imposed by time-delays in the secondary path are, e.g., addressed in [10]. To keep latency low, ANC implementations in the past have mostly been realized in analog circuitry. However, advances in integrated circuits and low delay AD and DA conversion make digital solutions feasible. A comparison of analog and digital realizations is given in [12]. The largest disadvantage of digital solutions is the additional delay imposed by the AD and DA conversion. This problem can be tackled with successive approximation register (SAR) converters, which incorporate very fast conversion times but only a limited resolution of usually 16 to 20 bit. It can be approached even with a dedicated sigma-delta-converter used at high sampling rates ($f_s = 192 \; \text{kHz}$ or above). When choosing an electronic back-end,
one also needs to consider constraints on energy consumption, size and computational capabilities.

4 Application Specific Challenges

4.1 Variation of the Secondary path

The secondary path $G(z)$ varies depending on the acoustic load coupled to the headphone, specifically different ear canal shapes, different fittings of the headphone and during jaw movements. Extreme cases such as covered (closed) or open nozzle of the headphones should be considered and tested regarding their stability. Measurements of these situations of the Bose QC20 headphones [13] (without Bose ANC Electronics) are illustrated in Fig. 3, including a percentile plot for the variation across different persons.

We can see that most of the persons have a very similar secondary path with a few outliers in the 95% interquartile range around the median, where the headphone did not completely occlude the ear canal. The closed characteristic (solid) has a similar characteristic at low frequencies, however shows different behavior above 300 Hz. The open case has a significantly lower response at low frequencies. Also, the phase is showing deviations of roughly up to 70° between different persons, which severely influences adaptive algorithms.

4.2 Direction-of-Arrival Dependency of the Primary Path

We focus on ear-mounted audio devices, which are influenced by head and ears of the user. Thus, we investigated the DOA-dependency of the Bose QC20 in [6] by measurements in a semi-anechoic chamber with a fast acquisition HRTF measurement system [14] and found a significant dependency, as illustrated in Fig. 4. The performance of the ANC system depends on the DOA and is not constant for directional disturbances. The DOA dependency needs to be considered for ANC algorithms to attain an optimal performance. Note the different frequency scale in Fig. 3 and 4, due to different measurement methods.

4.3 Perceptual Sound Quality

The design goal for ANC solutions is typically the minimization of the SPL, or an overall reduction of noise. However, this does not take into account the human perception of noise. Apart from considering the perception in the design process [15–17], it can be interesting to evaluate ANC solutions with perceptually motivated rating metrics [18, 19]. In addition to the remaining ambient noise audible system noise is induced by ANC systems, especially in situations with low ambient noise. System noise stems from quantization of the AD- and DA-converters, microphone, amplifier noise or even intentionally emitted as an excitation signal for online secondary path estimation (OSPE). Headphones or headsets are usually worn to playback a desired signal $a(n)$, such as music or speech. Thus, the effect of the ANC on the desired signal is an important, however, largely disregarded subject. Especially in the presence of a feedback ANC system, an adequate preprocessing of the desired signal $a(n)$, considering the influence of the feedback sensitivity $S(n)$ and the loudspeaker characteristic, is necessary [7]. It may also be desirable to have a better perception of the surrounding such as in the case of hearing aids. Principles of correctly processing the ambient signals to create a transparent perception of the occluding device are, e.g., presented in [11] and in more detail in [20].

4.4 Further deteriorating acoustic effects

Besides variability in $P(z)$ and $G(z)$, further acoustic effects deteriorate the performance of the ANC algorithms if applied for speech communication. They include audible disturbance sounds and additional disturbances in the microphone signals. One major problems with in-ear headsets is mechanically induced noise through the cables, also known as microphonics. Microphonics are the transmission of mechanical vibrations through electrical components, tackled e.g. by coated cables. They are typically tackled with coated cables, clips to fix the cables to clothing, by

Figure 2: Upper frequency to achieve a given attenuation of $Ad = \{0, 5, 10, 15, 20\}$ dB depending on the latency of the system (examples indicated).

Figure 3: Percentile plot for secondary path measurements for 23 people from listening test in [11] and additional lines for open and closed case.

Figure 4: Direction-of-arrival dependency of the primary path ($M = 4608$ directions), including one dedicated direction (lateral left for left headphone) [6].
laying the cable from the headphone upwards around the ear or with wireless connection, e.g. bluetooth. One other problem of headphones with closed fitting, which completely occludes the ear canal, is the *occlusion effect*. It manifests as a low frequency amplification of the own voice [11, 21]. It results from a change in the acoustic impedance of the ear canal due to the closure, which affects bone conducted sound $z_B(t)$ transmitted into the ear canal [22, 23]. Similar problems occur with stomping or other vibrations induced into the body. Both of these influences reduce the signal-to-noise ratio (SNR) inside the ear canal. It specifically is a problem for algorithms which rely on adaptive algorithms based on $e(n)$. Also, wind-noise can be a problem, which occurs at the external microphones due to turbulences [24].

5 Algorithmic realization

As described in Sec. 2, ANC can be realized as a feedforward or a feedback system. It is also possible to combine both topologies in a hybrid system. The challenge for ANC is to design filters $W(z)$ and $K(z)$ in a way to achieve performance and guarantee stability during varying conditions, especially within the primary and secondary paths, $P(z)$ and $G(z)$. The filters can either be pre-optimized and implemented as *time-invariant* filters or be adapted online as *time-variant* filters [25], [26], [27], [5]. There are also publications on combinations of time-variant filters with time-variant components, e.g. in [28] or [6]. It is essential to fundamentally understand the specific problem, as ANC systems always need to be customized. An overview with focus on algorithmic challenges is given in [29].

5.1 Time-Invariant Design

Time-invariant methods use offline optimizations for filters $W(z)$ and $K(z)$. They require knowledge about the *nominal* paths, which describe the actual use case, and the *perturbation*, which are all other use cases that may occur. Specifically, the variations of the paths $P(z)$ and $G(z)$ need to be considered. Feedforward filters can be designed with methods of optimal filtering [5] or robust control [30]. One way to design a time-invariant feedforward filter in the minimum-mean-square error sense is demonstrated in [6]. However, it should be noted that the performance of time-invariant feedforward filters degrades for varying paths $P(z)$ and $G(z)$. For feedback filters, considering variations of the paths is even more important, as they are prone to instabilities. They can be designed heuristically, however, robust control offers a structured approach, which directly includes uncertainties in the design process, e.g. presented in [7, 30]. The goal is to achieve robust stability, i.e. stability for variations of different persons as well as extreme cases as indicated in Fig. 3, and nominal performance, referring to the performance for the nominal case. Other approaches emphasizing on the performance are, e.g. based on $H_2$-metric [31] or the cepstral domain [32].

5.2 Time-Variant Design

A timely topic in ANC research is the design of time-variant, adaptive filters, which adjust to changing acoustic conditions. Applying adaptive filters to the ANC problem can be viewed as an online system identification process, like, e.g., echo cancellation, with the secondary path $G(z)$ as an additional unknown component. The system that shall be identified, the primary path $P(z)$, is estimated by the combination of adapted filter $W(z)$ and the non-controllable secondary path $G(z)$. This situation leads to the optimal filter given in (2) [5]. The unknown component $G(z)$, limiting the performance, also needs to be considered for the adaptation, since the reference signal $x(n)$ is prefifled with an estimation $\hat{G}(z)$ of the secondary path [33, 34]. This estimation $\hat{G}(z)$ either needs to be determined beforehand or estimated online. In [35] the authors specifically address the problem of secondary path models for over-the-ear ANC headphones. Ways to detect irregularities in $G(z)$ are presented in [36]. Methods for online secondary path estimation (OSPE) are, e.g. given in [37–41]. However, they typically require an additional excitation signal, the *added noise*, audible to the headphone user. Tracking secondary path changes is still a challenging task, considering the requirement for minimal disturbance by the added noise. Overall, the time-variant filters inherit the challenges of adaptive systems, which are residual noise, convergence times, divergence, which leads to instabilities, and computational complexity.

6 Conclusion

This contribution gives an overview of the challenges for implementing Active Noise Cancellation (ANC) in headphones. We address technological challenges, including the acoustic front-end design, which concerns building a headphone, and the electronic back-end, which comprises analog or digital filter implementation. Furthermore, we give an overview of challenges for the application specific realization, such as variations of the secondary path, DOA-dependency of the primary path, perceptual noise and sound quality as well as further deteriorating acoustic effects. We conclude with a brief overview of the different approaches to achieve performance and stability.

References


