INNOVATIONS IN AUDIO • AUDIO ELECTRONICS • THE BEST IN DIY AUDIO

www.audioxpress.com

audioxpress

ADVANCING THE EVOLUTION OF AUDIO TECHNOLOGY

Market Update
From Headphones to Hearables
Actively Enhancing Our Hearing
By J. Martins and Mike Klasco

Practical Test & Measurement
Objective Listening Effort Evaluation
By Hans W. Gierlich
and Jan Reimes (HEAD acoustics)

Market Update
Tackling Voice Capture and Processing Technologies
By J. Martins

Fresh From the Bench
Linear Audio Autoranger Mk II Interface
By Stuart Yaniger

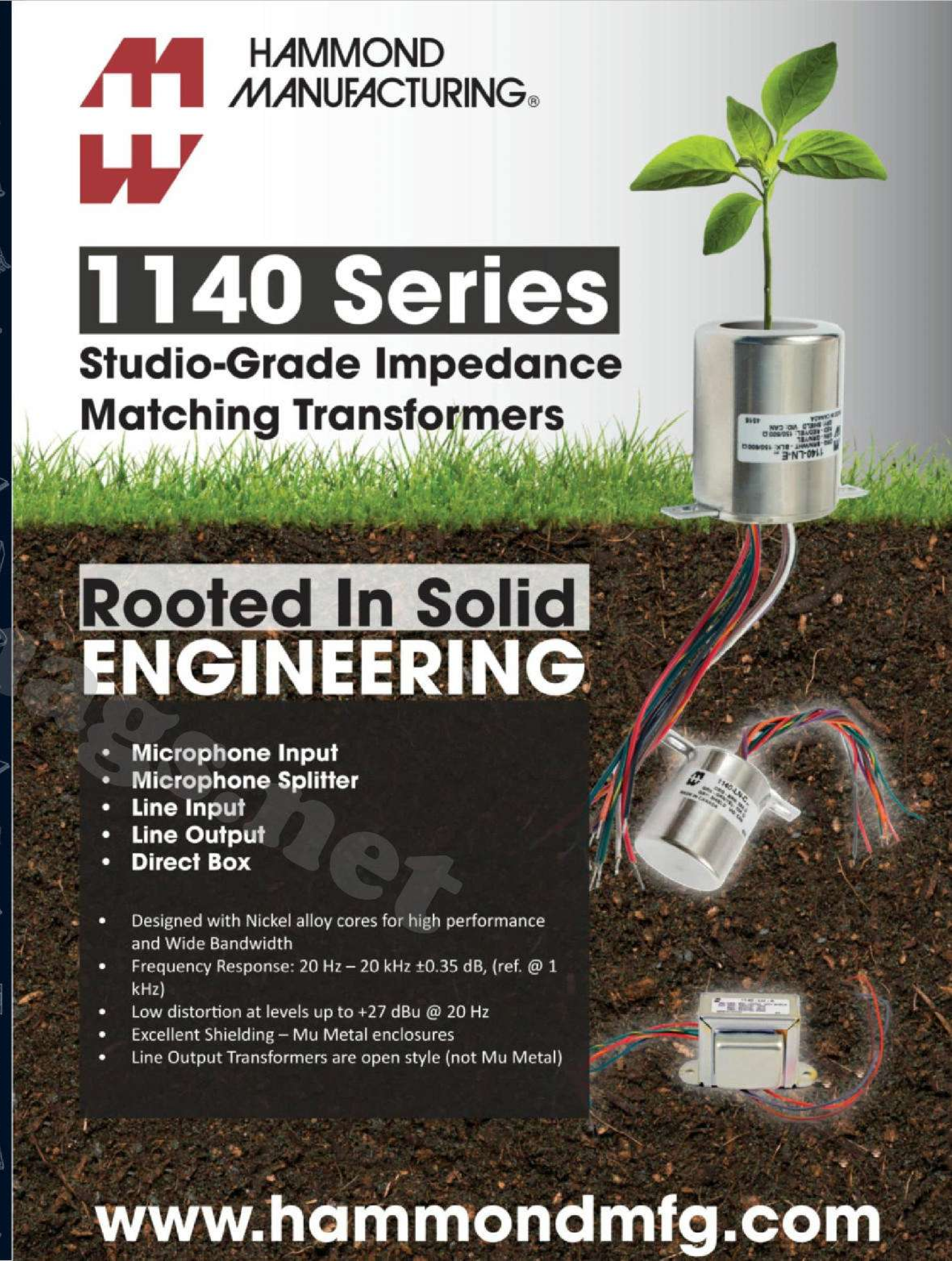
Audio Electronics
Powering Your Circuits: AC Power Rectification and Smoothing
By Ed Simon

Hollow-State Electronics
Triode Power Amplifiers
By Richard Honeycutt

Audio Scope
Loop Gain and Phase Margin Considerations in Applying Fully Differential Amplifiers to Active Filters
By Michael Steffes

$9.00US $10.00CAN
04
APRIL 2020


WE DO AUDIO
For over 30 years, we've been the design engineer's distributor of choice for innovative audio semiconductors and modules. Our knowledgeable technical experts are always on hand for reliable and honest advice.
It's as simple as that. We do audio. Talk to us.
Profusion
www.profusion.audio


HAMMOND MANUFACTURING®
1140 Series
Studio-Grade Impedance Matching Transformers
Rooted In Solid
ENGINEERING
• Microphone Input
• Microphone Splitter
• Line Input
• Line Output
• Direct Box
• Designed with Nickel alloy cores for high performance and Wide Bandwidth
• Frequency Response: 20 Hz – 20 kHz ±0.35 dB, (ref. @ 1 kHz)
• Low distortion at levels up to +27 dBu @ 20 Hz
• Excellent Shielding – Mu Metal enclosures
• Line Output Transformers are open style (not Mu Metal)
www.hammondmfg.com

April 2020 ISSN 1548-0628

www.audioxpress.com

audioXpress (US ISSN 1548-0628) is published monthly, at $50 per year for the US, at $65 per year for Canada, and at $75 per year Foreign/ROW, by KCK Media Corp., P.O. Box 417, Chase City, VA 23924, US Periodical Postage paid at Chase City, VA, and additional offices.

Head Office:
KCK Media Corp.
P.O. Box 417
Chase City, VA 23924, US
Phone: 434-533-0246
Fax: 888-980-1303

Subscription Management:
audioXpress
P.O. Box 417
Chase City, VA 23924, US
Phone: 434-533-0246
E-mail: custserv@audioxpress.com
Internet: www.audioxpress.com

Postmaster: Send address changes to:
audioXpress
P.O. Box 417
Chase City, VA 23924, US

Advertising:
Strategic Media Marketing, Inc.
2 Main Street
Gloucester, MA 01930, US
Phone: 978-281-7708
Fax: 978-281-7706
E-mail: audioxpress@smmarketing.us
Advertising rates and terms available on request.

Editorial Inquiries:
Send editorial correspondence and manuscripts to:
audioXpress, Editorial Department
P.O. Box 417
Chase City, VA 23924, US
E-mail: editor@audioxpress.com

State of the Headphones Industry

Headphones, earphones, and in particular, the fast-growing true wireless category, are currently the most sought-after consumer electronics devices. And naturally, the industry is researching improvements in design, materials, and user preferences to create differentiated new products. This is leading to technology innovations and raising the bar for consumer expectations, even for the most basic devices.

Incremental improvements in drivers, batteries, sensors, and powerful Bluetooth, amplification, and processing platforms are now available to manufacturers from many of the key technology vendors. This includes Qualcomm's QCC5100 and QCC302x ultra-low power Bluetooth audio System-on-Chips (SoCs), resulting in a significant number of new wireless products launched in 2019, all of which now offer improved specifications.

Still, for true wireless, Apple continues to lead the market, having sold an impressive number of its AirPods, and more recently having expanded its own market share with the launch of the new AirPods Pro with active noise cancellation (ANC) and wireless-charging. With its powerful combination of integrated silicon and well-tuned design, Cupertino captured the headphone market top position thanks to the growing popularity of its AirPods (and Beats headphones), leaving it in a strong position to maintain dominance in this market.

But of course, it's a continuous race, and all brands are making an effort to innovate and differentiate, while increasingly placing the focus on audio quality, which is perceived as the ideal angle to compete with Apple—even if "audio quality" is not exactly perceived in the same way by actual consumers. The jury is still out as to what features, designs, and form-factors consumers will prefer, given users' changing habits and the fact that wireless headphones and specifically true wireless and hearables are relatively new product categories.

In fact, no one seems to agree which key-features for headphones in general will foster adoption. From high-resolution audio support (the "Hi-Res Audio" stamp is big in Japan but apparently ignored elsewhere) to audio personalization or even auto-calibration, there isn't yet a single reference of a successful implementation in any of those possibilities.

A very positive development, the industry now seems to accept the conclusions of predicted listener frequency response preferences, according to the statistical models researched by Sean E. Olive, Todd Welti, and Omid Khonsaripour of Harman International. But while headphone manufacturers are tuning their products to approach the corresponding "Harman target curve," the reality is that the vast majority of consumers remain unaware of that discussion.

In 2019, product differentiation through the implementation of ANC became a new strong focus. From the popularity of using ANC headphones during flights, daily commuting, or working in open office spaces, which supported the early success stories from Bose and Sony, the industry is now evolving to new concepts where ANC has inspired innovative applications with transparency mode or enhancement modes, through selective-frequency filtering.

As a new generation of ANC products is now receiving mass-adoption, with Apple's AirPod Pros leading the segment, further research will be needed into users' preferences and adoption rate. Since advanced ANC products are relatively new, and most consumers have not yet had enough experiences to determine what they prefer—having the opportunity to compare different options—it is still too early to say what and how popular ANC features will be. And as manufacturers experiment with further active signal processing to combine captured environmental sounds with music reproduction, and communication and interaction with voice assistants, the more we will learn about how users will react to changes into the signal that is actual fed into their ears. In particular, more experience is needed into the combination of different form factors (e.g., loose-fit earbuds with ANC). Also, there's a lot to be learned about long-term exposure to the technology.

And when the discussion evolves toward hearables, augmented hearing and all the possibilities of interacting with voice assistants, using AI and edge-processing to offer additional advanced features—including voice translation—it suddenly becomes even clearer that we are in the early stages of an exciting new product category that will evolve on multiple fronts in the next few years. More importantly, it is a category that promises to continue to grow and surprise us.

J. Martins
Editor-in-Chief

The Team

President:	KC Prescott	Associate Editor:	Shannon Becker
Controller:	Chuck Fellows	Graphics:	Grace Chen
Editor-in-Chief:	João Martins	Advertising Coordinator:	Nathaniel Black

Technical Editor: Jan Didden

Regular Contributors: Bruce Brown, Bill Christie, Bob Cordell, Joseph D'Appolito, Vance Dickason, Jan Didden, Scott Dorsey, Gary Galo, Gerhard Haas, Chuck Hansen, Richard Honeycutt, Charlie Hughes, Mike Klasco/Nora Wong, Oliver Masciarotte, Nelson Pass, Christopher Paul, Bill Reeve, Fernando Rodrigues, Micheal Steffes, Stuart Yaniger

OUR NETWORK

audioxpress | circuit cellar | LIS Loudspeaker Industry Sourcebook | Voice Coil

SUPPORTING COMPANIES

NOT A SUPPORTING COMPANY YET?

Contact Peter Wostrel (audioxpress@smmarketing.us, Phone: 978-281-7708, Fax: 978-281-7706) to reserve your own ad space in the next edition of the magazine.

COLUMNISTS

Vance Dickason has been working as a professional in the loudspeaker industry since 1974. He is the author of *Loudspeaker Design Cookbook*—now in its seventh edition and published in English, French, German, Dutch, Italian, Spanish, and Portuguese—and *The Loudspeaker Recipes*. Vance is the editor of *Voice Coil: The Periodical for the Loudspeaker Industry*, a monthly publication. Although he has been involved with publishing throughout his career, he still works as an engineering consultant for a number of loudspeaker manufacturers.

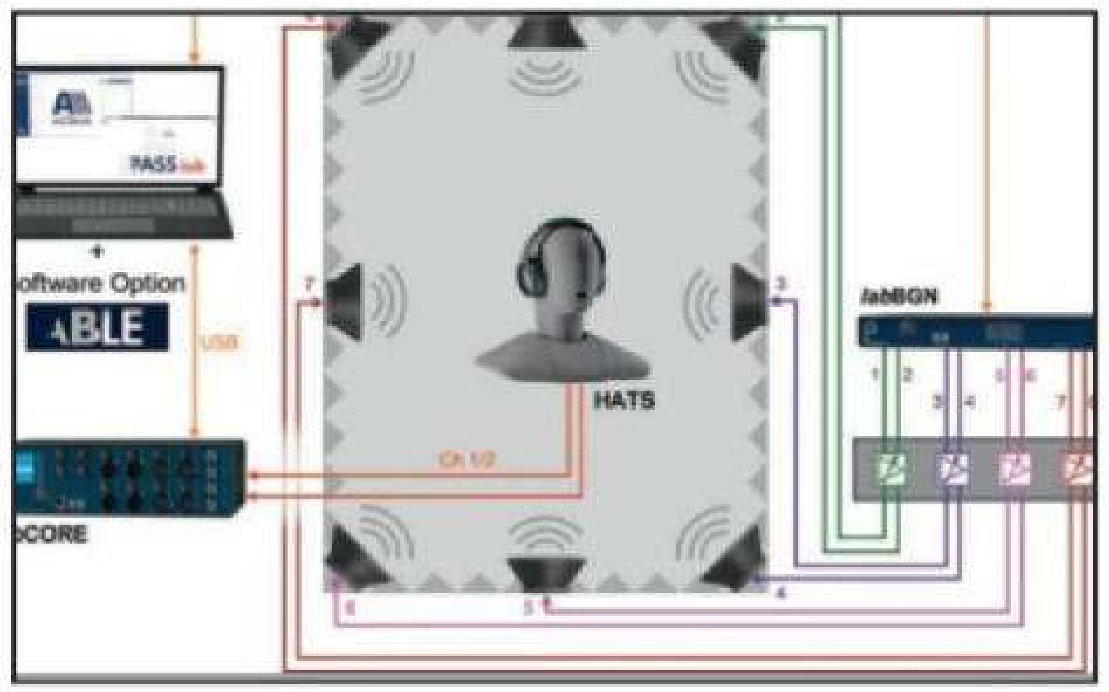
Dr. Richard Honeycutt fell in love with acoustics when his father brought home a copy of Leo Beranek's landmark text on the subject when Richard was in the ninth grade. Richard is a member of the North Carolina chapter of the Acoustical Society of America. Richard has his own business involving musical instruments and sound systems. He has been an active acoustics consultant since he received his PhD in electroacoustics from the Union Institute in 2004. Richard's work includes architectural acoustics, sound system design, and community noise analysis.

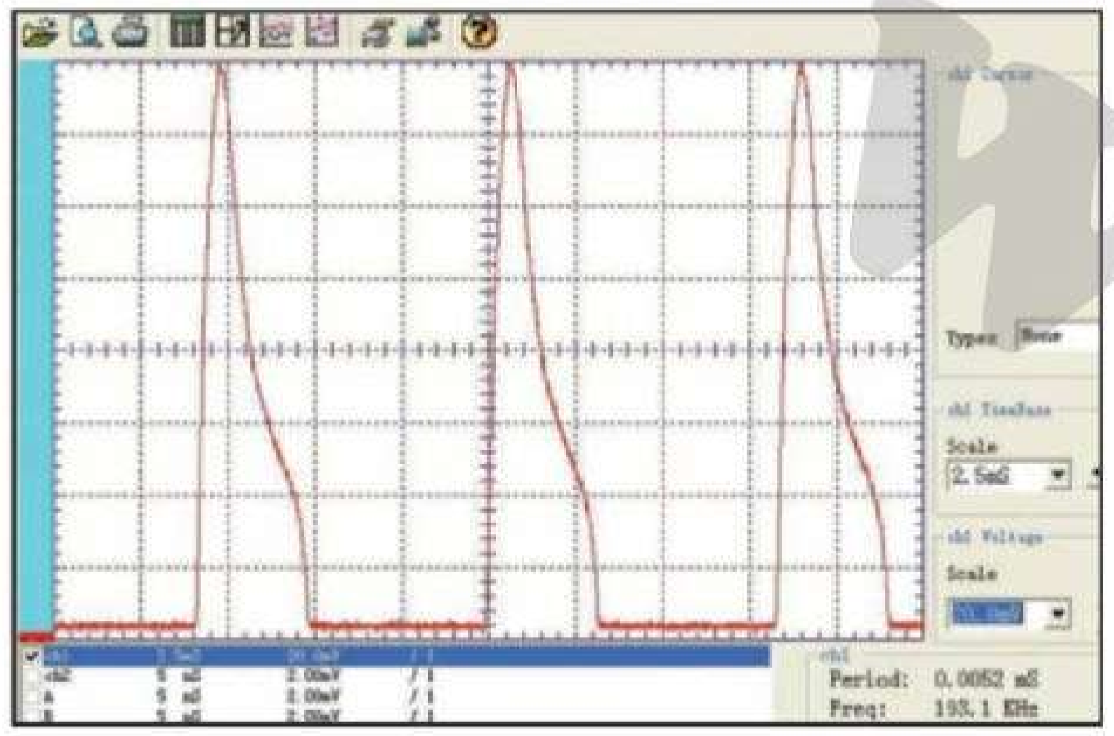
Mike Klasco is the president of Menlo Scientific, a consulting firm for the loudspeaker industry, located in Richmond, CA. He is a graduate from New York University, with post graduate work in signal processing, and he holds multiple patents licensed or assigned. For the past 35 years, Mike has worked on countless R&D projects for large and small companies. Mike specializes in materials and fabrication techniques to enhance audio performance.

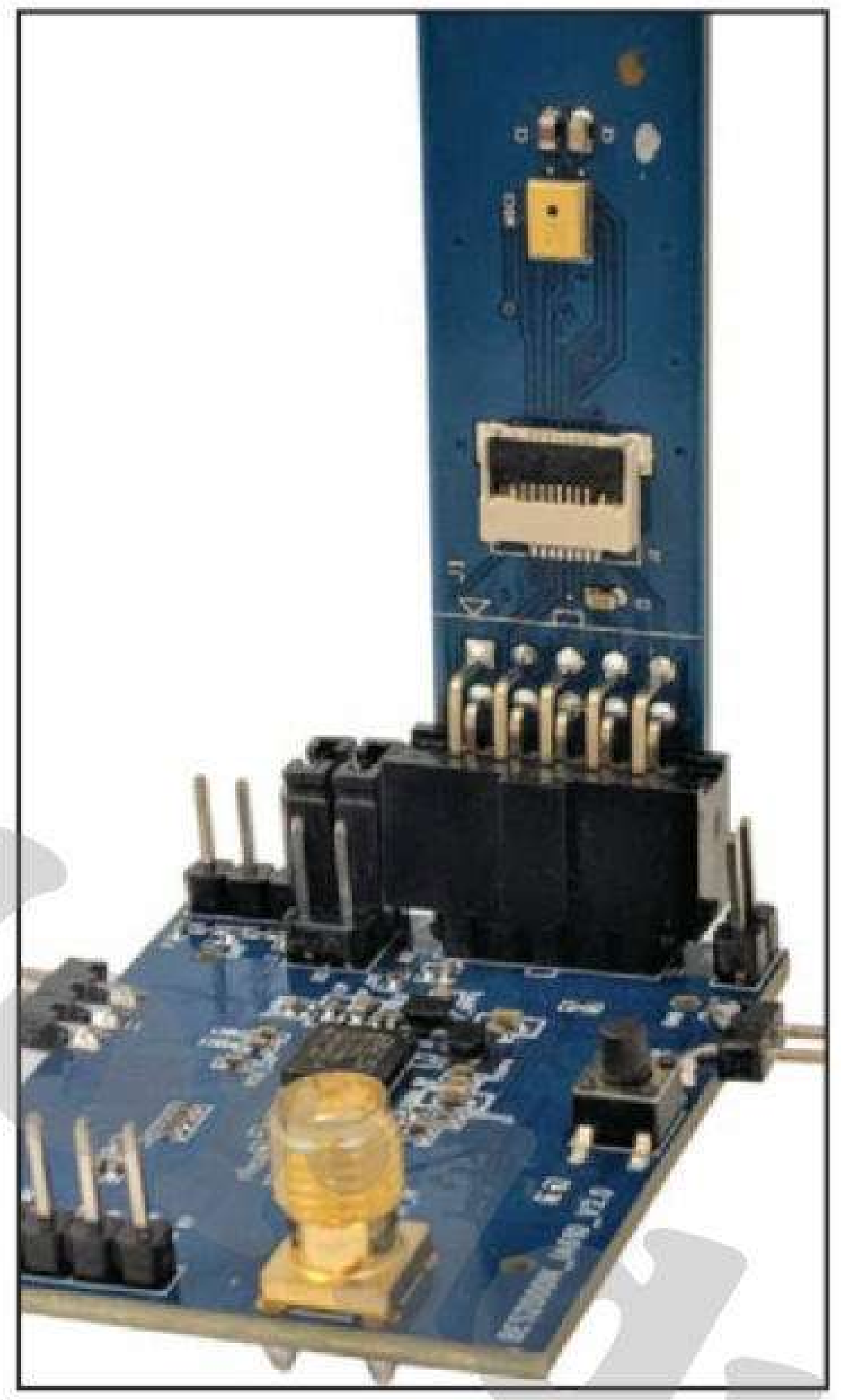
Michael Steffes has worked for more than three decades in high-speed signal path IC development, spanning IC design roles for the BurrBrown high-speed amplifier products, and product launch/application manager roles through Texas Instruments/Intersil. He continues to focus on high-performance ADC interface solutions, DSL/PLC line driver design and optimization, and signal path design optimization tools. In a career spanning five different major IC vendors, he has participated in the introduction of 150 industry leading amplifier products (setting many of the now standard practices in high speed amplifier characterization curves), publishing more than 100 contributed articles and application notes from 1989 onward.

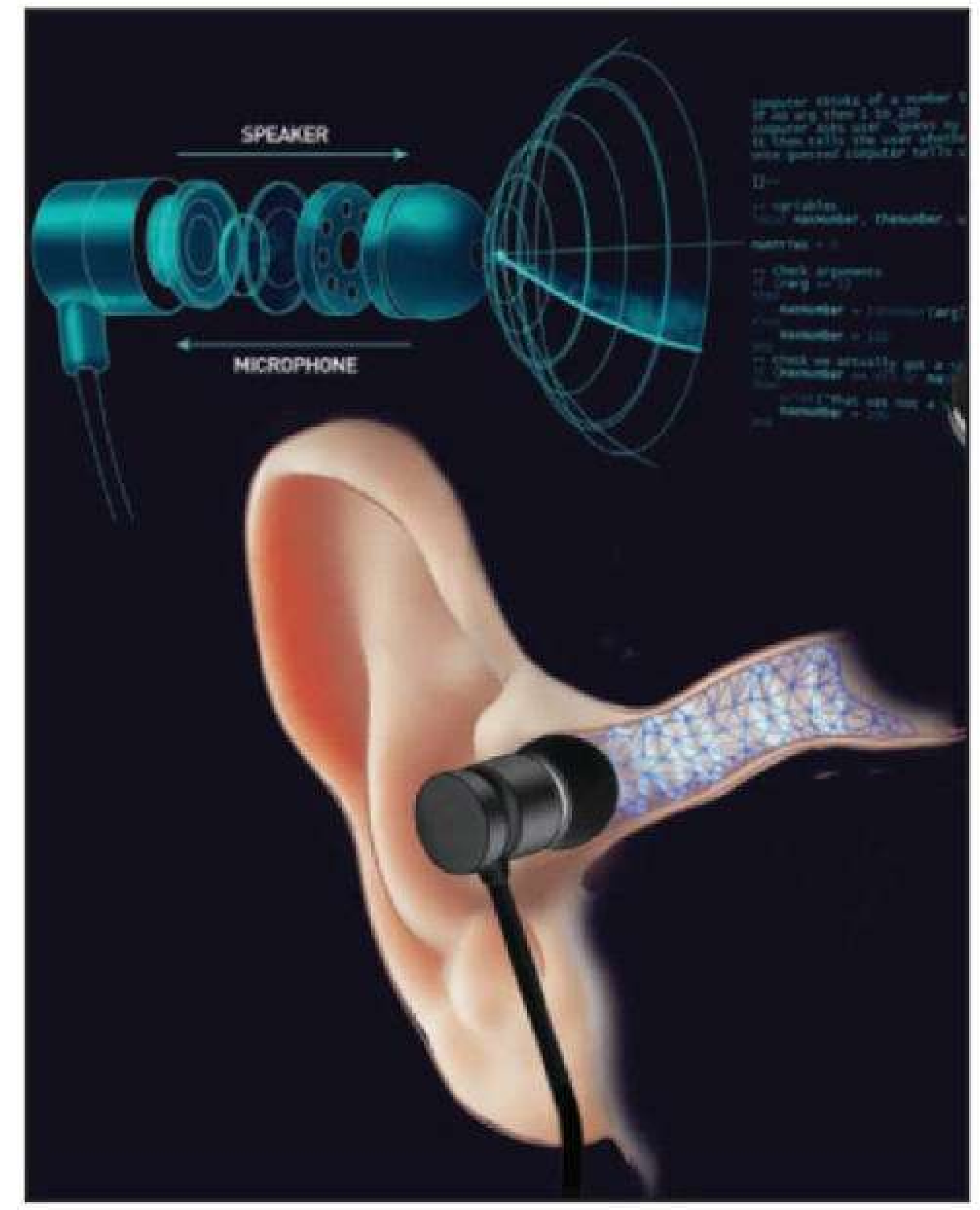
Contents

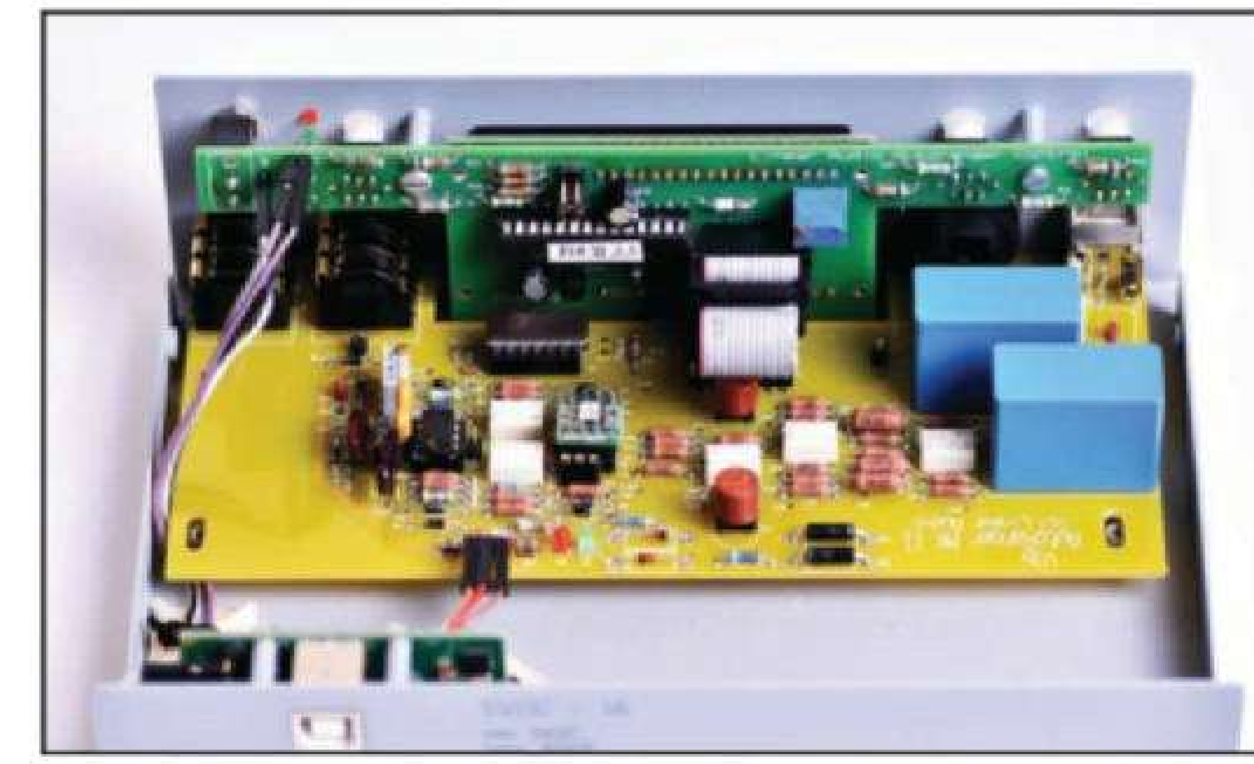
Volume 51 – No. 4 April 2020

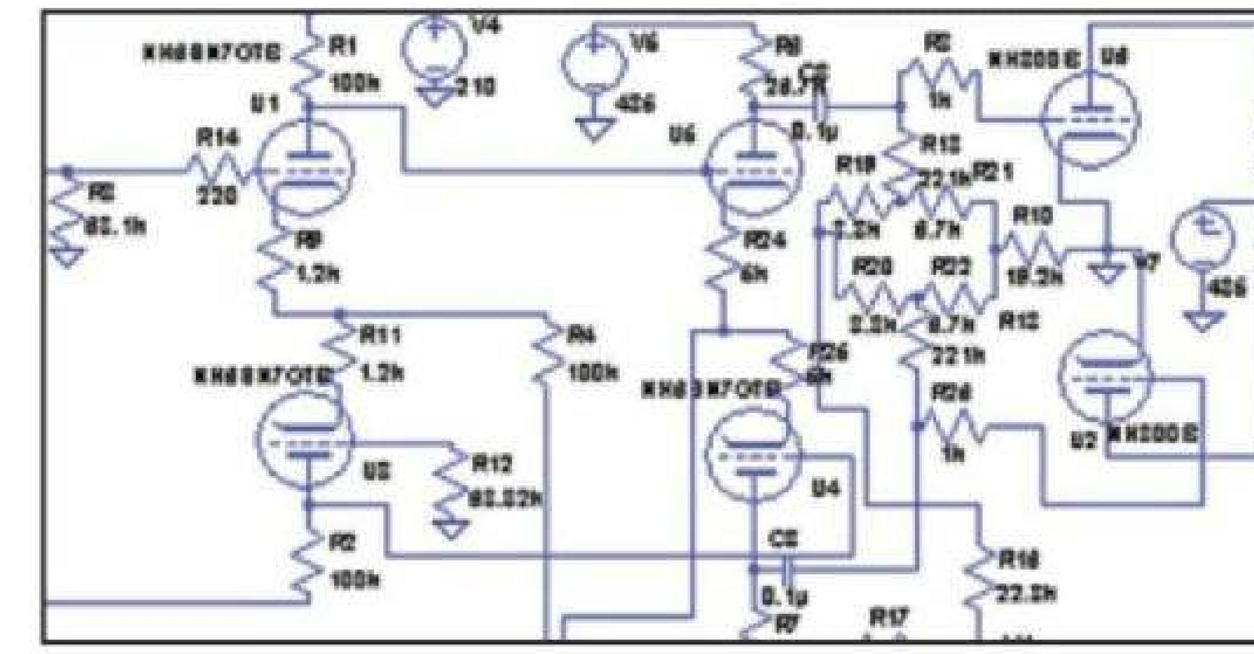












Websites

www.audioxpress.com
www.voicecoilmagazine.com
www.cc-webshop.com
www.loudspeakerindustrysourcebook.com
www.circuitcellar.com
www.linuxgizmos.com

@audioxp_editor audioxpresscommunity
linkedin.com/company/audioxpress

COVER ART
Photo Credit:
Cover photo courtesy of
AU-Stream ANC | True Wireless Earphone
https://ausounds.com
and
Microsoft Surface Earbuds
www.microsoft.com

Actively Enhancing Our Hearing

In this Market Update focused on headphones, we take a closer look at how the industry is transitioning from headphones and earphones to active designs, fostering the booming true wireless and hearables categories. Written from the developers' and manufacturers' perspective, *audioXpress* sums up the main market trends, current approaches, industry challenges, and new technologies available.

By
J. Martins and Mike Klasco

House of Marley Redemption True Wireless ANC Earbuds, a greener alternative to existing true wireless earbuds and the second-generation TWS design from the brand.

Electronics has traditionally not been within the comfort zone of headphone factories. Earphones and headphones have historically been "passive" meaning a wired connection to external amplification. The crystal radio sets preceding the 1920s were completely passive using just the signal strength to drive headphones with sensitive carbon granule drivers. Vacuum tube radios become common place by the early 1930s, enabling a wider range of headphone transducers as well as loudspeakers (using a tap to the power supply for juice for their electro-magnets).

In Asia, most headphone factories started as microspeaker, injection molders, and/or cable vendors. Times have changed and for a few years now active headphones and earphones account for more than 50% of the dollar value of the entire product earphone and headphone category. Headphone factories have addressed the active challenge either by bringing the electrical engineering and surface-mount printed circuit production in-house or by outsourcing design and production to a local Bluetooth partner.

Let's start with a definition of "active" earphones and headphones where the term is used as a catch-all phrase for devices containing electronics. The consumer market has seen active headphones for 30 years now, with the introduction of active noise cancelling (ANC) models from Bose, Telex, and Sennheiser. Yet, it was Bluetooth wireless operation that has enabled the market shift to active hearables. The earlier "cyborg look" of clip-on-the-ear Bluetooth hands-free devices is now considered tacky, but as Bluetooth technology graduated toward decent music quality the bulk of headphone and earphone designs embraced not just wireless but also the opportunity to add all sorts of signal processing.

More specifically, the Bluetooth wireless chipsets and SoCs (System-on-Chips) integrate DSP so it is now possible to add all sorts of signal processing to Bluetooth devices. The flexibility provided by flash programmable DSPs helps designers to easily differentiate products with new features without extended development cycles.

While headphone amplification is often integrated into the SoC, the RF aspects can be more rigorous. The Bluetooth team must select the antenna and design it into the product. Antennas do not work well if they are positioned too close to the battery or the headphone speakers, and you may remember an earlier iPhone a while back that needed to be properly held in your hand or the antenna performance would be impaired. Finally, there is the RF calibration process at the end of the line that the headphone vendor must be competent.

Active circuitry can confer many unique attributes, ranging from wireless, headphone response correction, ANC, personalization

In 1937, the company Elektrotechnische Fabrik Eugen Beyer (today's beyerdynamic) released the DT 48 dynamic headphones, the grandfather of hi-fi headphones, manufactured and improved for 56 years. The image shows a more modern version, as used by many audio engineers in pro audio applications.

(equalization compensation for a specific listener's hearing, immersive, and enhanced spatiality) and also biometrics for sport fitness and even an onboard artificial intelligence trainer coaching you.

Bluetooth Development

As the wireless technology standard for personal devices, Bluetooth gets credit as the main catalyst for forcing the earphone and the headphone factories to learn about battery power and circuitry. Bluetooth technology was first developed at Ericsson Sweden at the end of the 1980s as a wireless alternative to RS-232 data cables for exchanging data between fixed and mobile devices using short-wavelength UHF radio waves. While the range is short, the power drain is far lower than Wi-Fi, DECT, and other consumer wireless alternatives. Since 1998, Bluetooth has been managed by the Bluetooth Special Interest Group (SIG), which oversees development of the specification, manages the qualification program, and protects the trademarks.

Bluetooth 5 has arrived and is likely to be found in just about all products on the shelves and online. New developments in Bluetooth are ongoing (see the side text about Bluetooth LE Audio's announcement) but important enhancements range from adaptive performance trade-offs in sound quality vs. latency and also some fixes for marginal aspects of True Wireless Stereo (TWS) operation.

Aside from industry standard protocol enhancements there is the hardware. The last year has been a challenge for Bluetooth designs as Qualcomm's QCC 3 and 5 SoC series have both brought greatly enhanced audio, more robust wireless operation, and huge increases in battery operating time along with a stiff learning curve for designers. There are a dozen contenders for the chips especially for TWS products as the category is hot and the opportunity huge. Realtek, MediaTek, along with BEStechnics, Nordic Semiconductor, and Microsemi/Microchip to mention just a few of the chipset vendors are all fighting for the business.

Advanced and affordable, today's generation of active noise canceling (ANC) headphones includes the recent Sennheiser HD 450BT and HD 350BT around-the-ear Bluetooth 5.0 headphones, presented at CES 2020. The technology, simultaneously pioneered in the 1970s for aviation applications by Bose in the US and Sennheiser in Germany was made available to consumers in the 1980s.

Signal Processing for a Personalized User Experience

By the time you reach 30, your hearing is no longer what it was when you were 15. Aside from acquiring some dips in response, your left and right ears drift apart no longer matching too well, fuzzing up your stereo perception. Personalization assistive listening is about the user taking regular, self-administered hearing exams, which then compensate for the use's hearing deficiencies. Implemented in software/firmware algorithms that fit into the Bluetooth DSP or run on external memory that can be licensed from companies such as Even (EarPrint), Audiodo (former Hearezanz), Mimi Hearing Technologies, and others. There are quite a few benefits such as better intelligibility at lower listening levels and the closer match between left/right ears enhancing spatial perception.

The latest tweak to your hearing is to customize and fine-tune the Head-Related Transfer Function (HRTF). This term describes the acoustical properties of the head, upper torso, and external ear—elements that interact in complex ways to affect sounds reaching the eardrums. Multiple companies are currently exploring this promising field, such as Embody (currently working with Audeze on the Reveal+ personalized virtual studio plug-in for headphones), Mimi (which acquired French HRTF pioneer 3D Sound Labs), Dirac, and Genelec, which is promoting Aural ID, a development effort initiated in 2017, in joint-venture with Finnish HRTF modeling specialists IDA Audio. The generic approach is to compute personal ear and head interactions to create a personal data file characterizing the modification to sound arriving from any azimuth and elevation. This file consequently enables an audio engine to precisely render stereo or immersive content via headphones.

Typically, gathering personal HRTF information requires an anechoic room, placement of measurement microphones at the entry to the user's ear canals, and attention to setup and procedure details with multiple measurements. By contrast, the approach used by Aural ID software requires the user to provide only a 360° video of their head and shoulder region, for which a high-quality mobile phone camera is sufficient. Once the video is uploaded to the web-based calculation service, the calculation process first builds a detailed 3D model scaled to the dimensions of the head and upper torso, with special attention paid to modeling of the external ears.

For its 360 Reality Audio project, Sony is using a simpler approach to offer "a custom immersive musical field that is perfectly optimized for each individual user." This is done using an app that takes images of the ear, uploads it to the cloud, compares them to a database, and sends the closest HRTF profile, which is selected "leveraging Sony's unique algorithm." This generic solution works reasonably well for some people and not at all for others.

Augmented Hearing

One powerful vision for headphones and the consumer electronics industry is that of hearables and augmented hearing—from enhancing and protecting your hearing at loud concerts to blocking gun shots while hunting, in-ear recording of classroom lectures, binaural recording of music events, and even an all-knowing interactive computer consiglieri (advisor) that resides in your ears, and even translates other languages almost in real time.

Most of us spend too much of our life staring at our smartphones, but hearables may change that focus to listening to a voice in your head. Hearables, as with smartphones and smartwatches, are fast becoming a computing platform running apps and the race is clearly on the edge-processing and AI platforms for wearables and hearables.

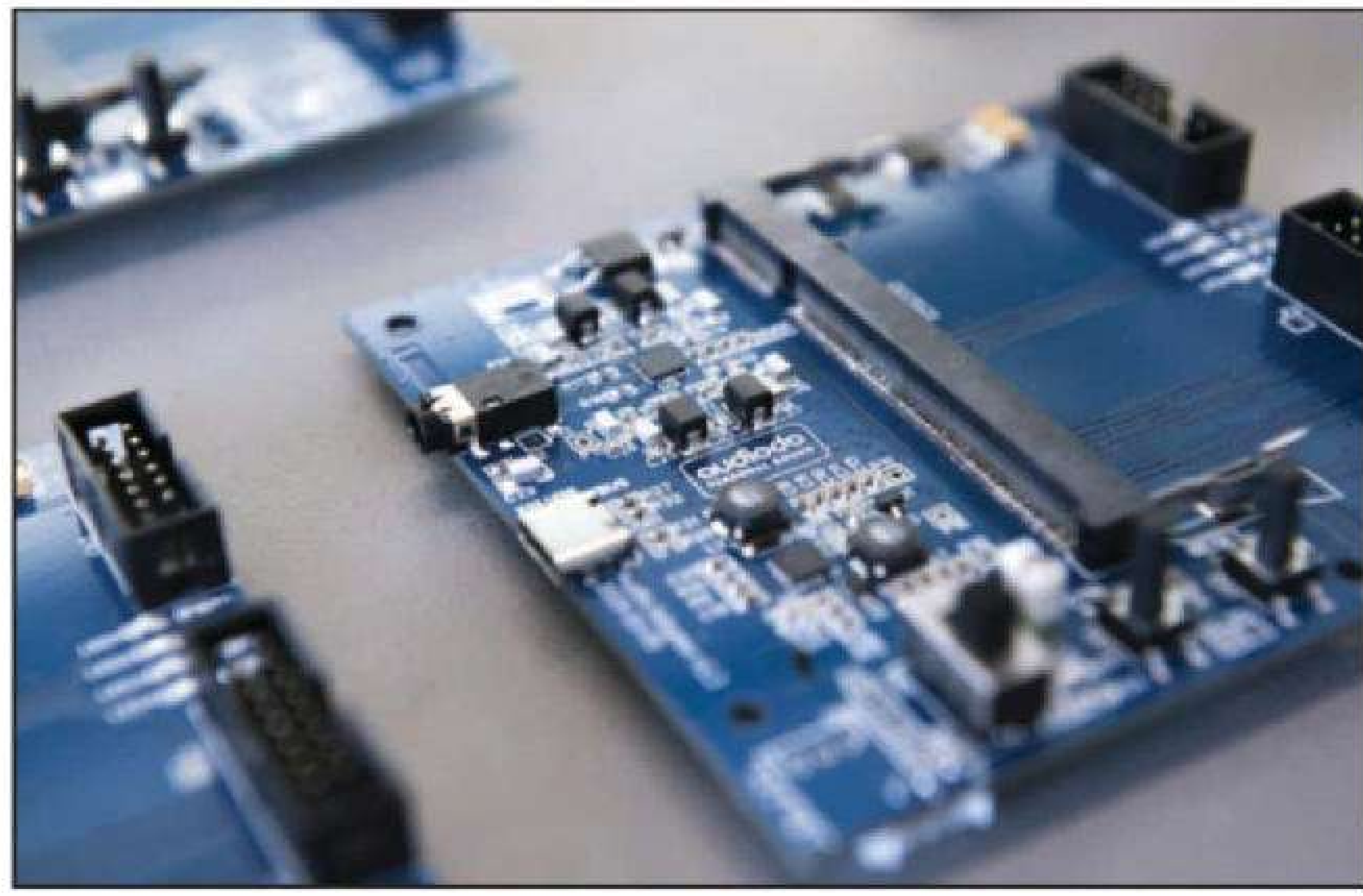
For now the concept is still dependent on being connected to the Internet through a smartphone to access Siri, Alexa, or other intelligent virtual assistants, conferring private guidance to the wearer.

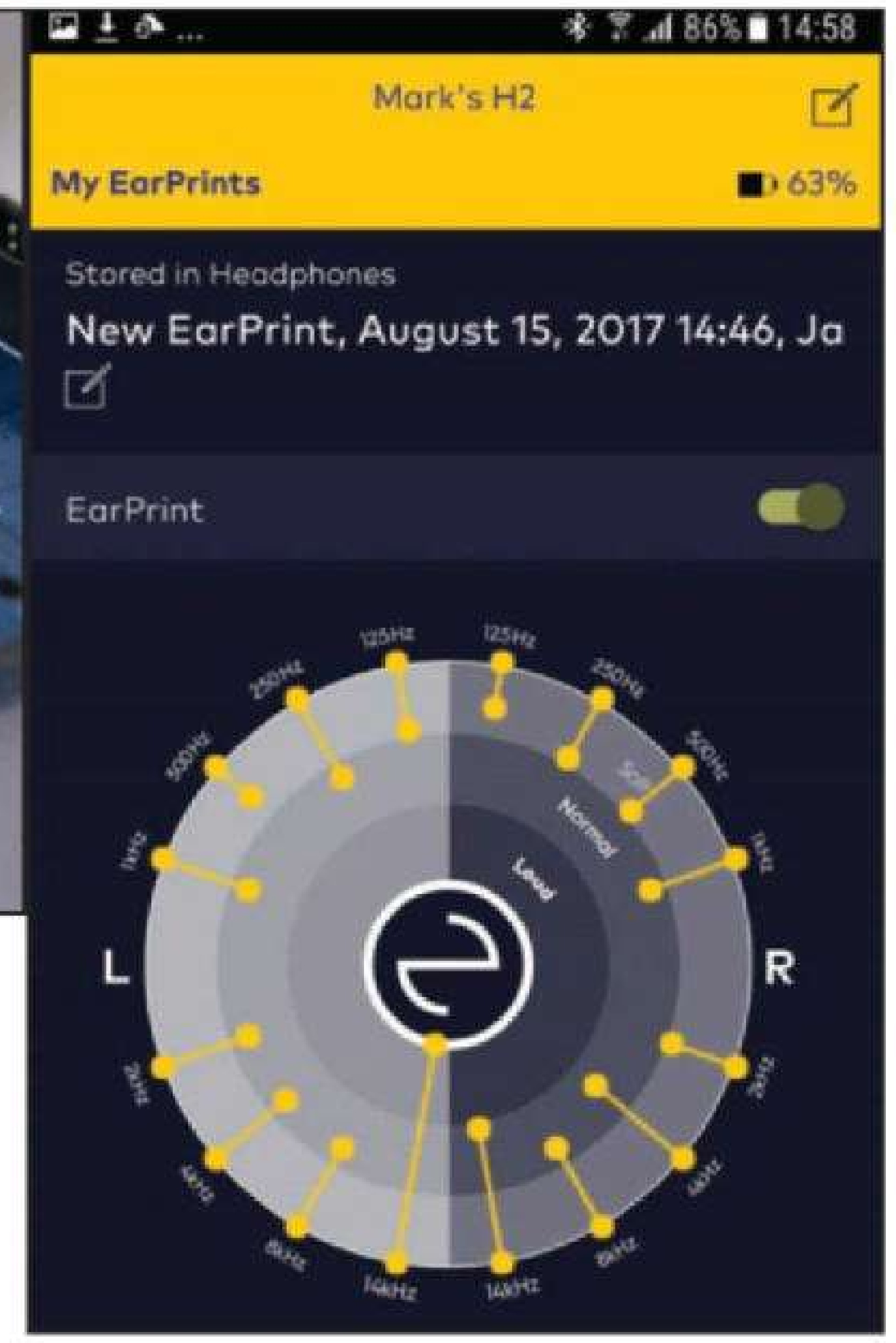
With consumers increasingly aware of active noise cancellation (ANC) technology, more robust Bluetooth wireless connectivity, and fast charging via USB-C, manufacturers are feeling the need to upgrade their existing models. Bowers & Wilkins adopted the latest available platform and technologies from Qualcomm to launch its 2019 models, which includes the flagship PX7 Wireless over-ear design with ANC and aptX Adaptive.

But soon, the concept of the autonomous multi-function hearable will challenge the smartphone just as the first iPhone extinguished the existence of both the iPod and the Blackberry personal digital assistant. Perhaps in the next few years, smartwatches and TWS will eliminate the smartphone.

The user interface for the most sophisticated hearables will migrate away from the user having to poke at the device's buttons on their ear and shift toward intuitive voice commands empowered with artificial intelligence to interpret what is being asked of it.

Based on the ultra-low power Qualcomm QCC5100 Series Bluetooth Audio SoCs with the new aptX Adaptive audio technology, all new Bowers & Wilkins headphones offer robust, low-latency, high-resolution wireless transmission at 24-bit/48 kHz streaming quality. The headphones also include the latest-generation ANC technology from Analog Devices that runs at 16 times the sampling rate to ensure best-in-class noise cancelling with no degradation to the audio quality.





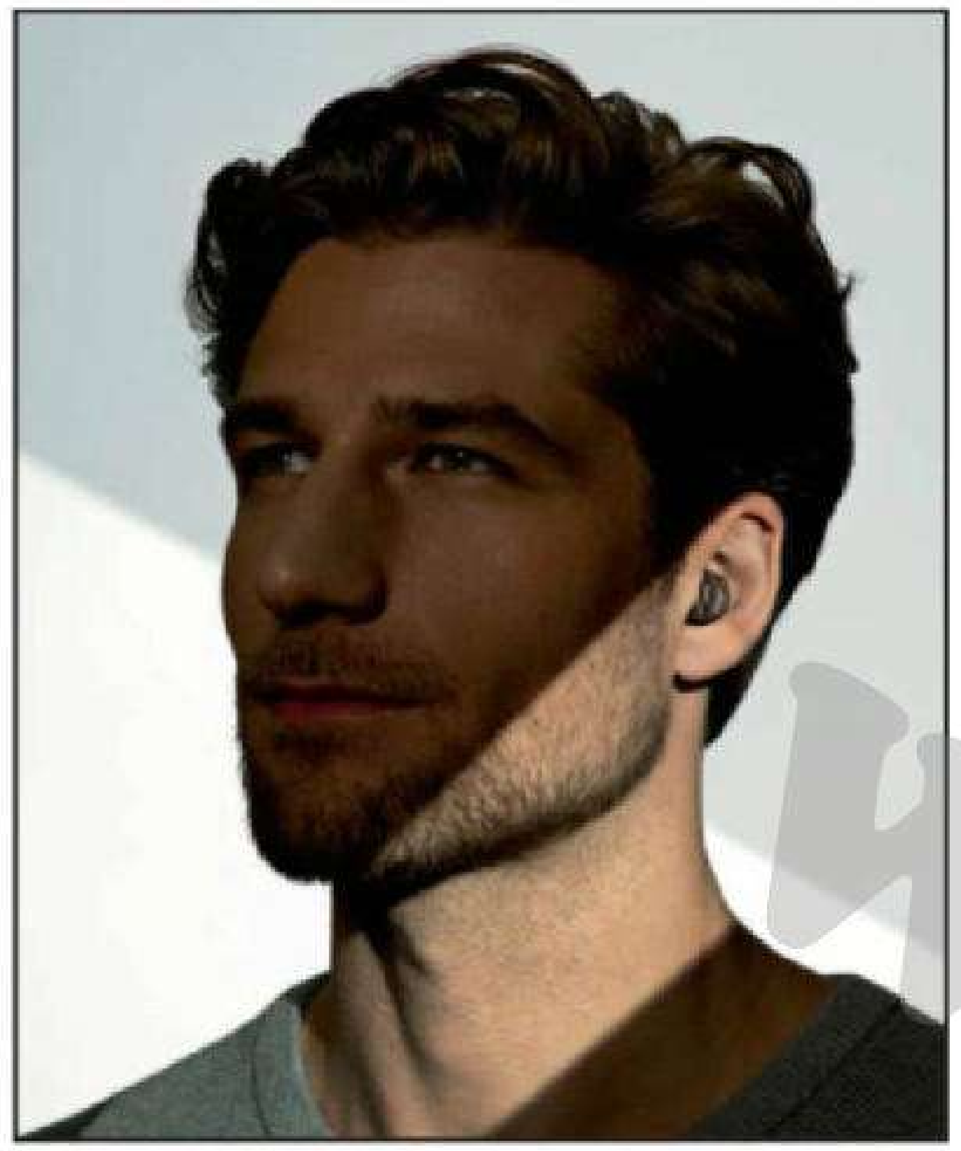
Personalization options for headphone manufacturers remain an attractive field to explore. Even with its EarPrint technology, and Audiodo Personal Sound (formerly Hearezanz) are among the companies targeting the complexity of hearing to offer personal audio profiles.

Today, the most viable and hottest hearable implementations are based on the evolution of TWS earphones, of which the most commercially successful example today are the Apple AirPods.

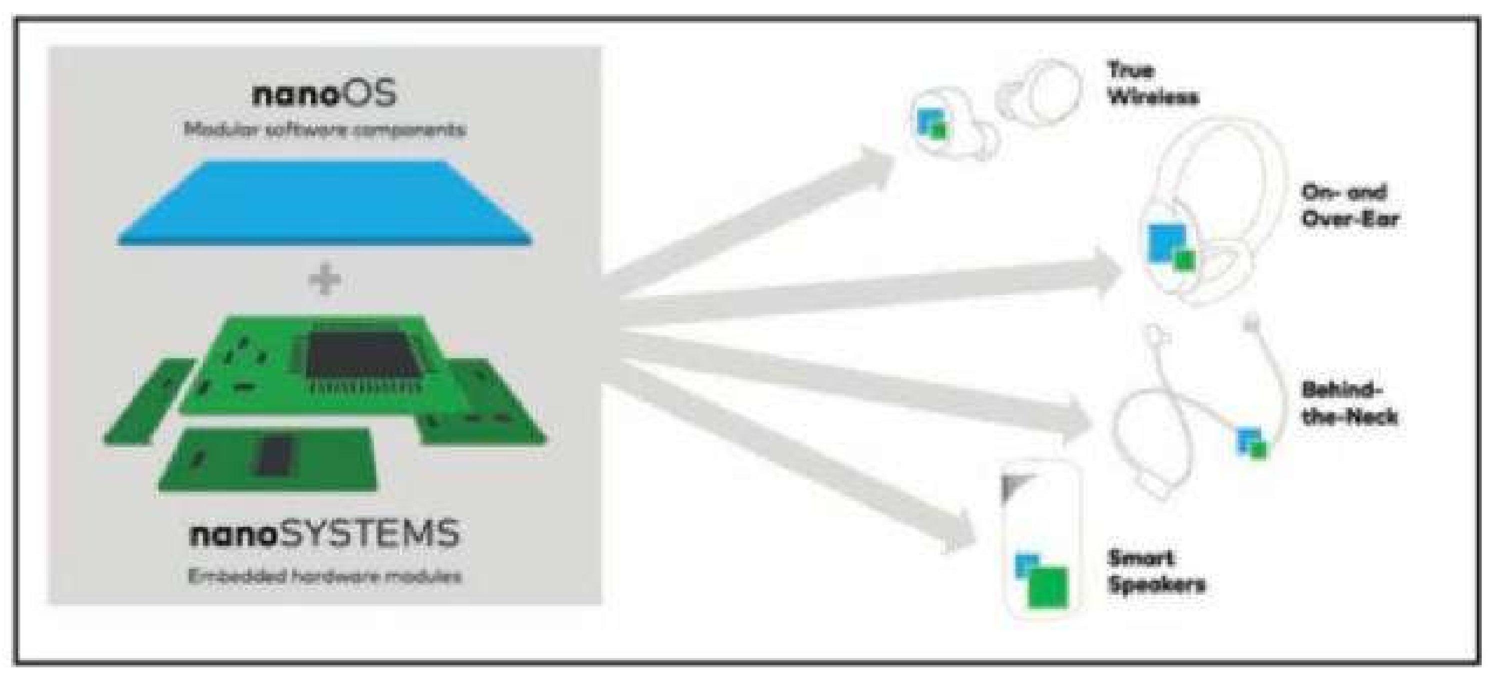
True Wireless Stereo functionality was first introduced by CSR (now part of Qualcomm) in one of its Bluetooth chips, about six years ago, allowing each ear to have an earphone with absolutely no wires, just as an in-ear hearing aid. Sending a stereo audio stream to two distinct earbuds is not possible using standard Bluetooth A2DP profile, since it is a point-to-point solution, and the industry has been busy working on different approaches.

But it appears the concept of a smart earpiece was first envisioned in 2005 by Personics Labs (aka Hearium Labs). Hearium was a bold visionary project in augmented reality, music listening, biometrics, and voice communications. It was extraordinarily

Headphones manufacturer Audeze has partnered with Embody to launch the Reveal+ personalized spatial plug-in for Audeze headphones and the corresponding Aural Map app, which enables users to create personal HRTF profiles generated from images of the ear, captured with a smartphone.



The new Phonak flagship product, the Virto Black was designed to blur the lines between a hearing aid and a hearable and encourage people to seek treatment sooner thanks to the stylish design and innovative technology.



True wireless pioneer, Bragi is now licensing its IP for product development, including software and hardware components for hearables and smart speakers.

Bugatone software technology uses the existing earphone driver to sense audio signatures and other functions.

ambitious and attempted to create the ubiquitous in-the-ear computer that Doppler and Bragi were later pursuing. The Hearium initiative yielded many patents but never made it to production.

Doppler Labs and Bragi were visionary leaders in TWS, both founded in 2013. Braggi, once the largest Kickstarter campaign in European history and the company behind the "World's First Hearable," followed with the Alexa-compatible Dash series. Anyone who has lived with TWS knows that trying to manipulate controls on the TWS is awkward and voice command—if reliable—is the way to go. And there are other considerable but surmountable challenges for hearables—such as the dense mechanical fabrication, ambitious signal processing, stable RF reception, and battery life (or lack of it!).

Doppler, after burning through a $50 million investment and unable to find further funding, closed at the end of 2017 (many of its engineers moved on to work on projects such as Amazon's Echo Buds.)

Bragi also dropped out of the hearables product market, focusing on licensing the company's IP to other manufacturers, and developing ultra-efficient artificial intelligence and edge-computing technologies for ultra-low-power devices. The company is now able to license technology for automatic activity recognition and tracking, head gestures, and other user context detection for interface interaction, and provides solutions for applications such as real-time translation, voice assistant integration, and personalized sound. Among other companies, Bragi is working with Syntiant—a semiconductor company from California and fellow campaigner—in the pursuit to move Artificial Intelligence (AI) and Machine Learning (ML) from cloud to edge devices.

While Hearium, Doppler, and Bragi may have been a bridge too far for their times, true wireless is now the fastest-growing category in consumer electronics, and gradually, many of the features envisaged for hearables by those companies are being introduced to the market.

As sales of Apple's AirPods nearly doubled to $6 billion this year and should take another leap forward next year, according to Toni Sacconaghi of Alliance Bernstein, Apple could sell 85 million AirPods in 2020, generating about $15 billion in revenue.

Apple's new AirPods Pro model basically reenergized sales for Apple by focusing on two main innovations: a simple but extremely effective implementation of ANC and adding wireless charging to the case (available as an option for the original AirPods).

And now the race is on, with Amazon, Google, and Microsoft all entering the true wireless category, followed by existing smartphone brands such as

Swedish audio brand Defunc says it conducted a consumer research survey of 200 people to test a number of different 3D-printed earbud tips, and determined that the form-factor adopted in its True Go true wireless earbuds provide an "improved true wireless experience." The Defunc True Go design combines the AirPods' shape with a silicone cover, which improves contact and fit.

Resources

Audiodo
https://www.audiodo.com

Aural ID
https://www.genelec.com/aural-id

Bluetooth Low Energy Audio
https://www.bluetooth.com/learn-about-bluetooth/bluetooth-technology/le-audio
https://audioxpress.com/newsbluetooth-sig-unveils-le-audio-the-next-generation-of-bluetooth-audio

Bugatone
http://www.bugatone.com

Embody
https://embody.co/audeze-reveal-plus

LifeBeam
https://vitrainer.com

Mediatek
https://www.mediatek.com

Qualcomm
https://www.qualcomm.com/products/audio

Samsung, Xiaomi, and Huawei—all mainly focusing on noise cancellation, sensors, and integrated AI voice assistants.

While racing toward being the first with more hearable "killer-features," companies are exploring compelling multi-faceted applications, including sport fitness earphones that monitor biometrics. A few years ago, biometrics did not reach the market penetration that was projected in other product categories, but the latest and next generation of TWS products have stronger appeal. While not everyone will want to know their temperature, heart rate, or blood oxygen levels, a smart coach that whispers in your ear how you are doing might be game changing, ("you have climbed 5000 steps today and you are still breathing").

LifeBeam's AI Personal Trainer is available in its Vi App. Initially released in 2016, LifeBeam released the prototype of Vi, a wearable AI earphone personal training device that combines voice instructions with bio-sensors. The voice-activated digital assistant combined features similar to Apple's Siri with the fitness tracking features of Fitbit. The earbuds have sensors that track heart rate, heart rate variability, and motion while running, and also track weather, elevation, and location. The current Vi app analyzes the data and provides real-time coaching based on fitness goals. From dedicated hardware to the Vi app, LifeBeam is now licensing its technology to Samsung (Samsung Simband platform) and other companies.

What's Next?

One challenge, especially for TWS, is cramming the mic(s), earphone driver, associated electronics, antenna and battery—all into the smallest possible package. While the Airpods "hanging out of your ears" look will not please everyone, it quickly became established as a new normal and even an iconic look. And most importantly, it served the purpose of creating space for the antenna and battery, and place the microphone closer to the mouth. So, no wonder everyone is copying the design. But other companies are experimenting with other designs, mostly with the goal of becoming invisible.

In sports-oriented products the mic and bio-sensor are fighting for the same straight to the ear channel real estate. Using the speaker (earphone driver) as a mic and other sensor functions has been a below the radar effort at a number of companies. NXP's smart-amps have a "speaker as mic" function for using the speakerphone driver in smartphones as the mic when there is high wind noise. And taking multi-function use of the speaker further is the purpose of Bugatone's "virtual sensor" technology.

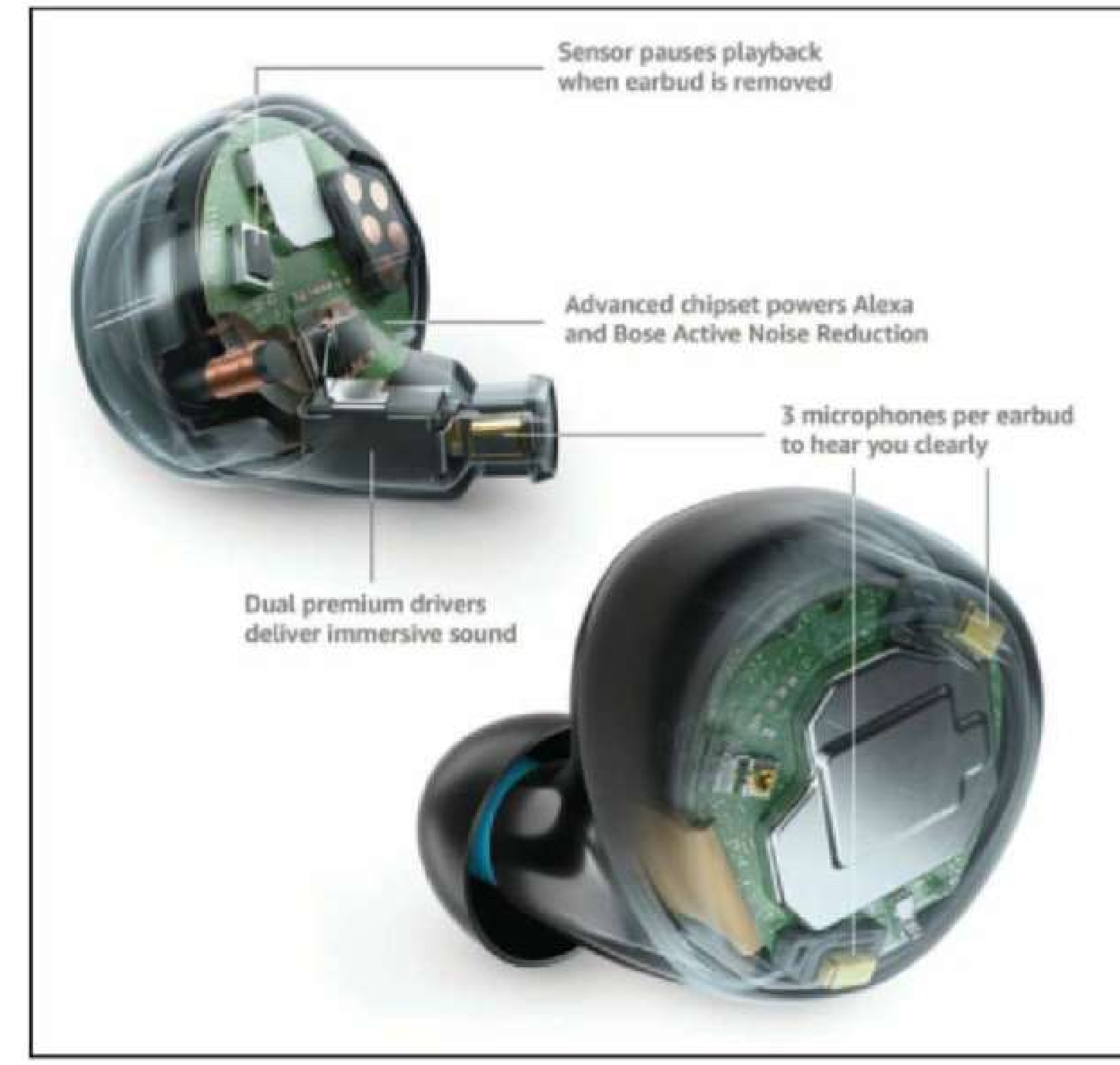
Bugatone uses the existing earphone driver to sense a half dozen functions without hardware changes including when the earphone is being worn (ear detection/auto off)—take the earphone out of your ear and the TWS shuts off—without adding a sensor. Another more sophisticated function is User ID—where the TWS can be trained to recognize authorized users—sensing the audio signature of the user's ear canal reflection. And other promising applications are using the earphone driver as the mic for ANC, with many engineering teams currently working on these and related techniques.

Active noise reduction, ANC, and selective frequency filtering (cancellation and amplification) are currently the most exciting fields of research for active earphone designs and toward the race of hearables and audio augmented hearing. As the market is quickly discovering, users are as excited to cancel environmental noises around them, as they are about the "transparency modes" that allow passing some sounds and cancelling others, while providing a more comfortable noise-level to listening to music. And gradually, they are discovering the ability to "spy on others," meaning amplify highly-directional selected frequencies (speech essentially).

When both Amazon and Microsoft recently promoted their own designs and vision for the true wireless earbuds category, both companies focused on different features. The Amazon Echo Buds extend the company's personal voice assistant Alexa outside the home and make it even more personal, and the Microsoft Surface Earbuds focus on ANC and augmented hearing, while adopting an interesting loose-fit design with a large round external surface, designed to be touched.

Both seem to be well-designed true wireless earbuds and Amazon even decided to create a technology partnership with Bose, licensing the ANC technology from the company. For the project, Amazon hired most of the former Doppler Labs engineering team, understanding they needed ANC not only as a selling point, but in order to make the voice assistant support more effective, combining audio processing abilities for optimal recognition outdoors.

The most interesting aspect of the Microsoft Surface Earbuds seems to be the decision to go for the loose-fit design, which successfully was used by Apple since the original EarPods and making the AirPods truly "universal." Apple decided to go for plugged in-ears for its AirPods Pro in order to create an effective seal for ANC, but we will soon see more loose-fit designs entering the market, including most probably also from Apple.



The new Amazon Echo Buds show that active noise cancelling (ANC) applications are starting to make a difference in the market. Market research firm Futuresource expects unit growth of 7% year-on-year, accounting for 6% of total headphone volumes.

Loose-fit are much more comfortable to wear for most people because they are non-occluding and do not form an airtight seal with the user's ear, allowing for longer usage without discomfort for most people. On the other hand, the absence of an airtight seal can affect the earbud's acoustic performance and makes implementing ANC much more difficult.

Yet, as Austrian company ams AG already demonstrated, loose-fit true wireless earbuds are compatible with ANC, even when more external noise can "leak" into the ear. For that scenario, ams designed Automatic Leakage Compensation (ALC) processing, which enables users to adjust for the air leakage between the earbud and the user's ear in real time, while still achieving a noise attenuation rating of 30 dB. Not dealing with the occlusion effect also helps to compensate the phase variations in selected frequencies, or what some people have called "the eardrum suck phenomenon." And more important, when combined with sensors, the form-factor could actually show some benefits in augmented hearing applications and hearables.

Bluetooth LE Audio

At CES 2020, the Bluetooth Special Interest Group (SIG) announced the upcoming release of LE Audio, the next generation of Bluetooth audio. Two years in the making, Bluetooth Low Energy (LE) Audio will enhance the performance of wireless audio, and introduce Audio Sharing, a new use case with the potential to once again change the way we experience audio and connect with the world around us.

"Twenty years ago, Bluetooth cut the audio cord and created the wireless audio market. With close to one billion Bluetooth audio products shipped last year, wireless audio is the largest Bluetooth market," said Mark Powell, CEO of the Bluetooth SIG. "The launch of LE Audio is a prime example of how the Bluetooth community is driving technology and product innovation and enabling delivery of even better and more capable Bluetooth audio products."

Bluetooth audio will support two operation modes. LE Audio will operate in the Bluetooth LE radio while Classic Audio—"classic" being the new designation for what we have today—operates on the Bluetooth Classic radio (BR/EDR). LE Audio will support development of the same audio products and use cases as Classic Audio, while introducing new features to improve their performance as well as enable new ones.

As one of its main building blocks, LE Audio will include a new high-quality, low-power audio codec, called Low Complexity Communication Codec (LC3) for Higher Quality and Lower Power. Providing high quality even at low data rates, LC3 will bring tremendous flexibility to developers, allowing them to make better design tradeoffs between key product attributes such as audio quality and power consumption.

"Extensive listening tests have shown that LC3 will provide improvements in audio quality over the SBC codec included with Classic Audio, even at a 50% lower bit rate," says Manfred Lutzky, Head of Audio for Communications at Fraunhofer IIS, the research institute responsible for LC3. "Developers will be able to leverage this power savings to create products that can provide longer

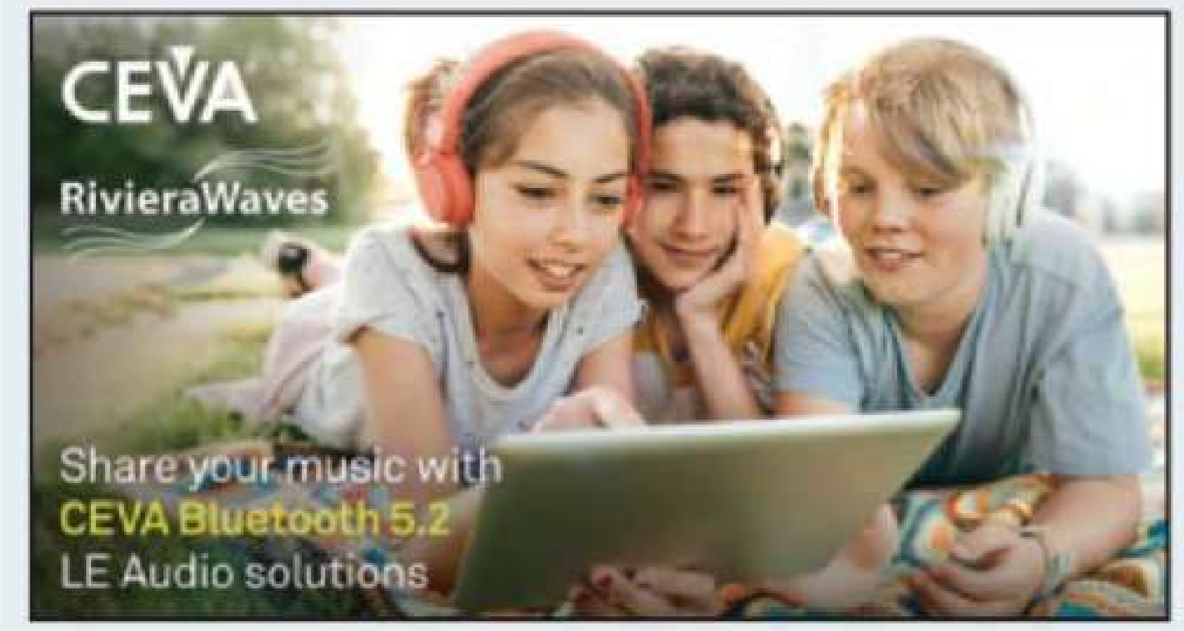
battery life or, in cases where current battery life is enough, reduce the form factor by using a smaller battery."

Bluetooth LE Audio also introduces Multi-Stream Audio, which enables the transmission of multiple, independent, synchronized audio streams between an audio source device, such as a smartphone, and one or more audio sink devices. Other exciting possibilities for LE Audio include Broadcast Audio, which opens significant new opportunities for innovation, including the enablement of a new Bluetooth use case, Audio Sharing. This enables an audio source device to broadcast one or more audio streams to an unlimited number of audio sink devices. This paves the way for assistive listening applications in classes and auditoriums, multi-language distribution during live shows, silent-discos, or simply being able to tune into a specific TV channel while running at the gym, or tune into the TV audio on the sports game at the bar.

Bluetooth Audio Sharing can be personal or location-based. With personal Audio Sharing, people will be able to share their Bluetooth audio experience with others around them (e.g., sharing music from a smartphone with family and friends). With location-based Audio Sharing, public venues (i.e., airports, bars, gyms, cinemas, and conference centers) can now share Bluetooth audio that augments the visitor experience.

One of the key application areas for Bluetooth LE Audio will be hearing aids and the specification received direct contributions from the European Hearing Instrument Manufacturers Association (EHIMA).

Bluetooth LE Audio quality improvements are courtesy of the new LC3 codec, or Low Complexity Communications Codec, which also enables lower power operation, extending the battery life of TWS earbuds and other audio streaming devices.



CEVA's RivieraWaves Bluetooth IP platforms provide comprehensive solutions for Bluetooth implementations and all the latest features of Bluetooth are supported.

The Bluetooth specifications that define LE Audio are expected to be released throughout the first half of 2020 and are part of the new Bluetooth Core Specification Version 5.2, which includes three primary updates. Products that support LE Audio will also support LE Isochronous Channels and brings an isochronous data transport option to the Bluetooth Low Energy radio. This will be an end-to-end solution that requires sinks and sources to use the same approach.

Industry Support

Following the Bluetooth LE Audio announcement, multiple vendors immediately announced products and solutions to support its implementation. Among those, CEVA, the wireless connectivity specialist, announced that it has enhanced its offerings for developers of true wireless stereo (TWS) earbuds and other wireless audio devices with the general release of the RivieraWaves Bluetooth Low Energy and Dual Mode IPs with support for LE Audio.

Nordic Semiconductor and Cirrus Logic Bluetooth LE Audio Evaluation Platform was designed to test source and sink devices, as required to complete an end-to-end LE Audio link.

CEVA's RivieraWaves Bluetooth IP platforms consist of a hardware baseband controller, plus a feature-rich software protocol stack. A flexible radio interface allows the platforms to be deployed with either RivieraWaves RF or various partners' RF IP, enabling optimal selection of foundry and process node. All the latest features of Bluetooth are supported, including Isochronous Channels for LE Audio, Direction Finding (AoA/AoD), Randomized Advertising Channel Indexing, Periodic Advertising Sync Transfer, GATT Caching, and other enhancements.

Nordic Semiconductor was another company that announced it has launched an LE Audio Evaluation Platform in partnership with San Diego, CA-based Bluetooth LE stack developer, Packetcraft. Nordic's LE Audio solution comprises a hardware reference design based on the company's nRF52832 Bluetooth LE System-on-Chip (SoC), Cirrus Logic's CS47L35 smart codec with an integrated low power Audio DSP, Packetcraft's Bluetooth LE host stack and link layer supporting LE Audio, and an LE Audio Software Development Kit (SDK). The platform allows developers to start evaluating the technology for wireless speakers, over-the-ear headphones, and true wireless earbuds.

Synopsys was one of the companies that collaborated with the Fraunhofer Institute for Integrated Circuits (IIS) to release an implementation of the Low Complexity Communication Codec (LC3), part of the next-generation Bluetooth LE Audio specification. Synopsys confirmed the availability of the LC3 Codec optimized for Synopsys' DesignWare ARC EM DSP and HS DSP processors.

Cadence Design Systems also announced that an LC3 implementation is available now for Cadence Tensilica HiFi DSPs. Cadence licensed reference code from Fraunhofer IIS to enable availability of LC3 in conjunction with the initial LE Audio launch, allowing audio manufacturers to immediately benefit from latest performance enhancements.

Finally, Tempow, a research and design lab focused on pushing the limits of Bluetooth technology, announced the release of a new Bluetooth Software Stack for hearables and speakers supporting the freshly revealed LE Audio. The company also announced a partnership program for chipset manufacturers to use Tempow's implementation. Telink Semiconductor was the first to join the program.

Resources

Bluetooth SIG, Inc. | www.bluetooth.com
Cadence Design Systems, Inc. | www.cadence.com
CEVA | www.ceva-dsp.com
Nordic Semiconductor | www.nordicsemi.com
Packetcraft, Inc. | www.packetcraft.com
Synopsys, Inc. | www.synopsys.com/designware
Tempow | www.tempow.com

ams AS3460 Digital Augmented Hearing Chip

ams showcased its new Earbud Innovation Demo Kit at CES 2020. The ready-to-use reference design shows how loose-fit true wireless earbuds can achieve selective noise cancellation for augmented hearing, while supporting features such as proximity sensing for automatic power saving and usability enhancements. Based on the new ams AS3460 Digital Augmented Hearing chip, the solution implements the company's Adaptive Leakage Compensation technology that dynamically adjusts the frequency coverage and depth of noise attenuation to cancel out noise which enters the ear directly, bypassing the earbud's silicone tip. This means that users who find loose-fit earbuds more comfortable to wear for long periods can enjoy the same noise cancellation performance as that of closed-fit earbuds, which plug the ear tightly to provide a physical barrier to ambient noise.

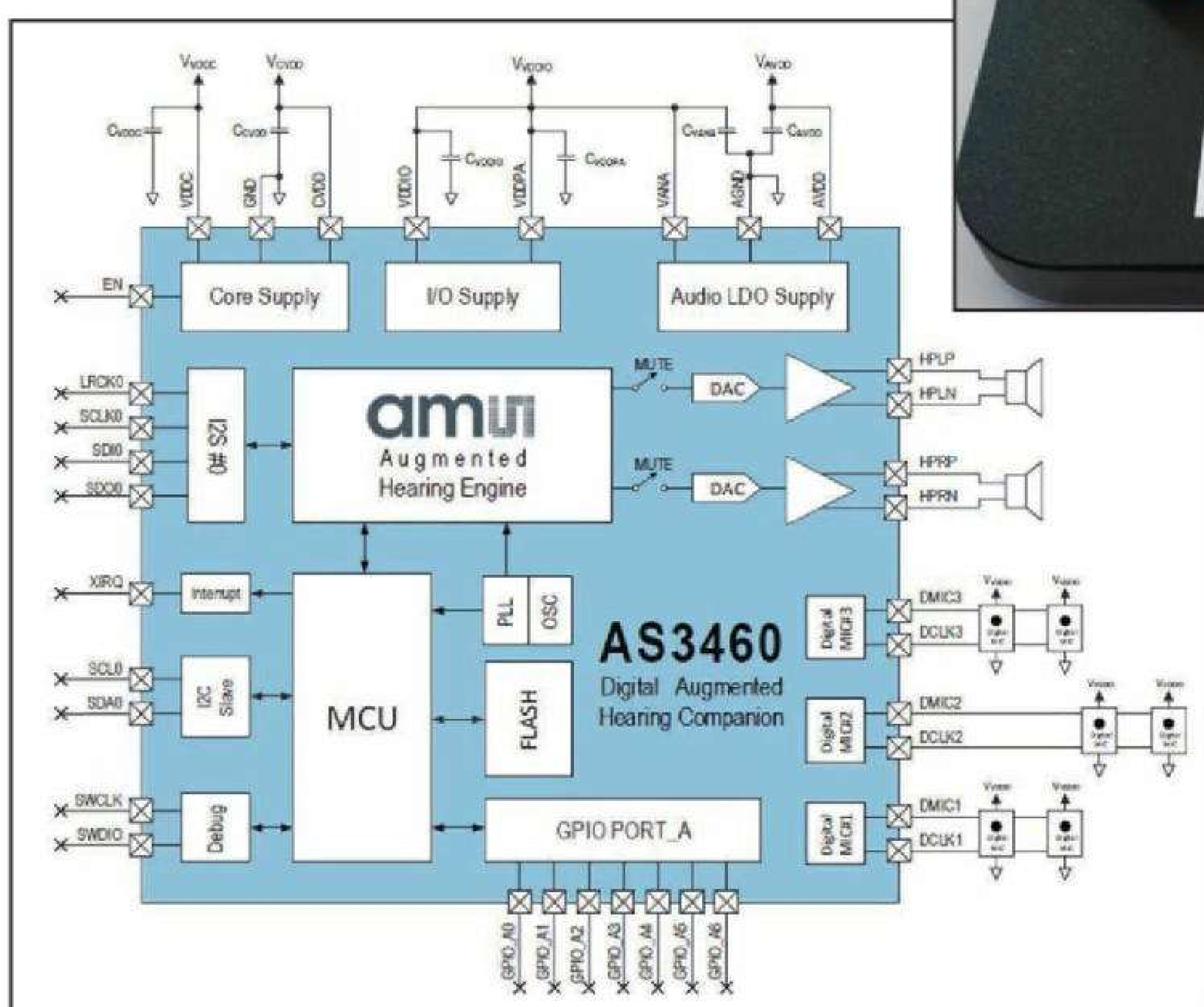
The multi-sensor innovation kit based on the ams AS3460 chip also supports new optical sensor features such as automatic in-ear detection. At CES 2020, ams also combined effective demonstrations on the role of wireless earbuds in personal health monitoring, sleep enhancement, selective sound awareness, and treatment of symptoms of hearing disorders (e.g., tinnitus).

The AS3460 is the first audio chip to provide up to 40 dB of noise attenuation in the loose-fit type of wireless earbuds over a broad range of audible frequencies, providing a quiet acoustic background in almost any environment. This means the user can take calls and listen to music at lower volume, which helps protect their long-term hearing health.

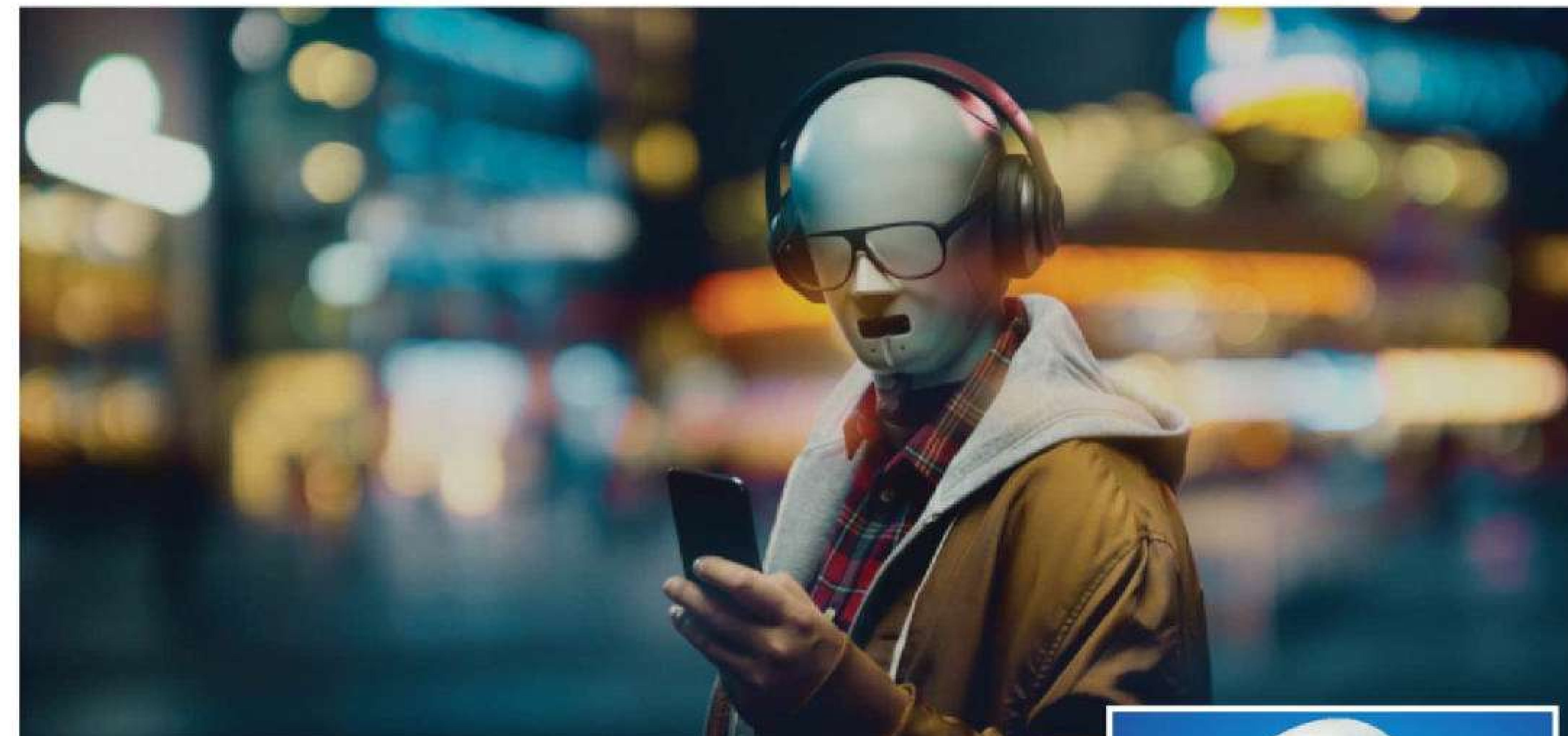
In the Earbud Innovation Demo Kit, ams also provides the AS3460's Automatic Preset Selection technology, which detects different acoustic environments such as the interior of a car, an airplane cabin, or a crowded room, and automatically applies an optimized noise filter without the user having to do anything—and without compromising safety or comfort as the user moves from one place to another.

The Earbud Innovation Demo Kit is supplied in a small, earbud-sized form factor, and is available to buy at the ams webshop.

https://ams.com/as3460.



ams, the Austrian sensor specialist, introduced the AS3460 Digital Augmented Hearing companion chip to bring high-performance noise cancellation for the first time to loose-fit wireless earbuds.



Tackling Voice Capture and Processing Technologies

It's not so much what you can do with Voice, but what voice can do for audio development. This Market Update discusses voice recognition and what voice capture and processing technologies can do for audio manufacturers and the audio industry in general.

The 2019 Amazon Echo smart speaker. The global smart speaker market reached a new high in 2019 with sales of 146.9 million units, an increase of 70% on 2018, according to Strategy Analytics. Amazon remained the leading brand with a share of 26.2%, down from 33.7% 2018.

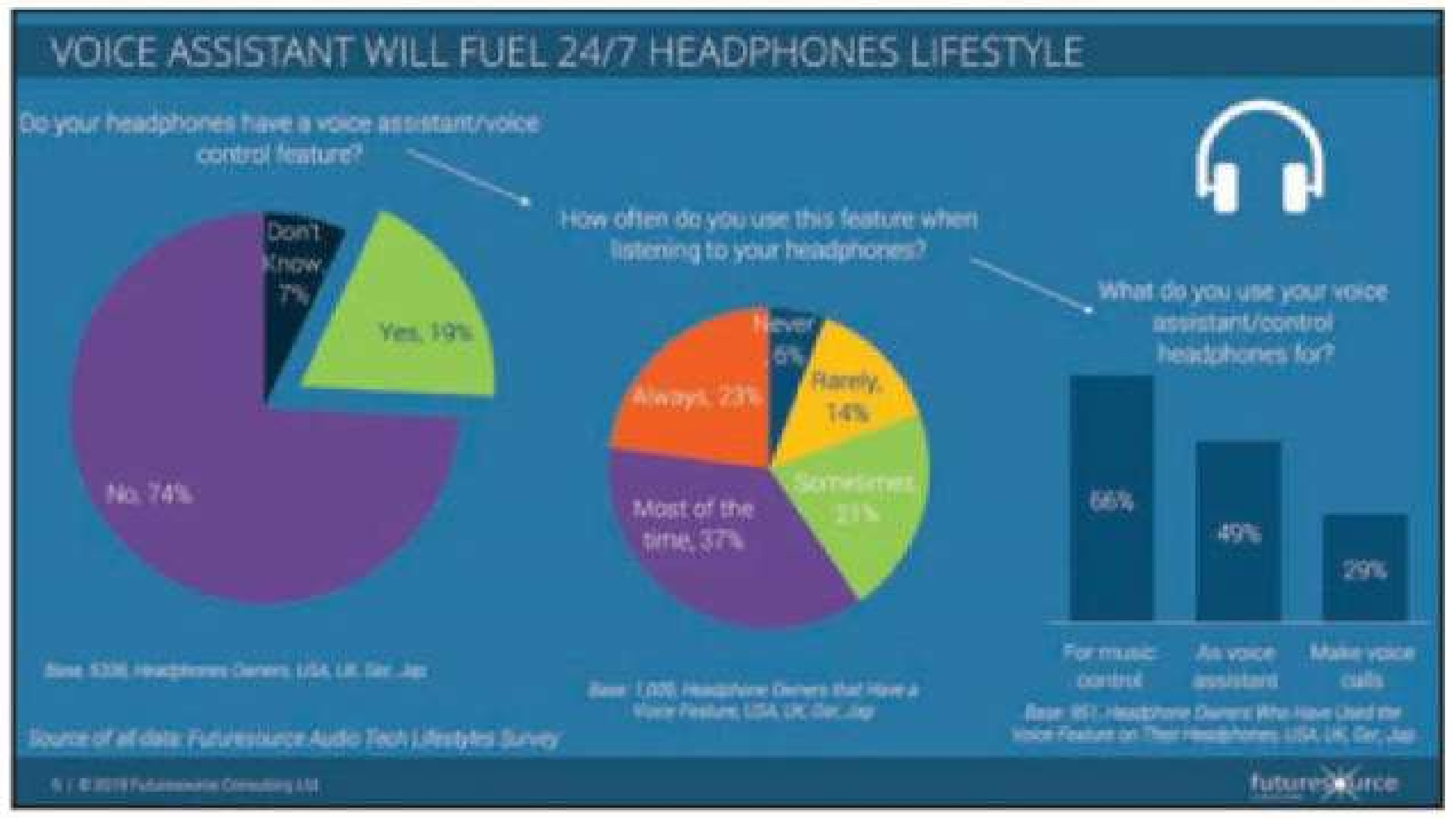
By
J. Martins
(Editor-in-Chief)

As Qualcomm's latest *State of Play Report 2019* reinforces, "Longer playback times, truly wireless earbud adoption, and Voice UI are all playing major roles in shaping the way people access and enjoy their audio." This annual report surveys 6,000 smartphone users from the US, U.K., China, Japan, India, and Germany for their thoughts on the future of audio. And even though those consumers continue to say that sound quality is the number one feature they desire in audio products in general (for 65% of those surveyed), more than 70% of all respondents rated extended battery life as an important purchase driver for wireless headphones, while requesting general information from voice assistants was named one of the main reasons for adopting smart speakers, right along with streaming music. As Qualcomm concludes, this suggests that both use cases are almost equally important to today's consumers and that they are growing more comfortable interacting with their voice assistants.

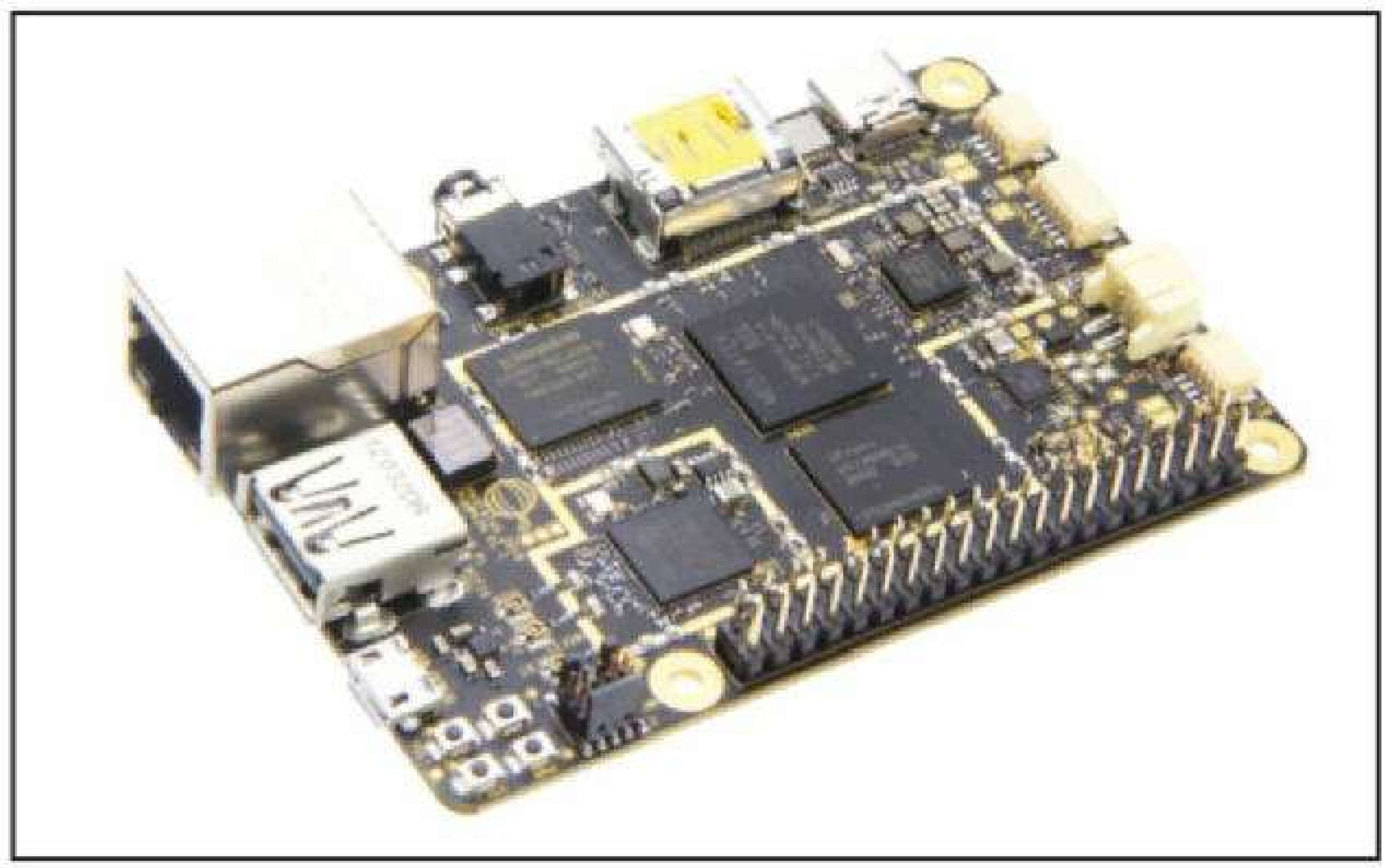
As consumers are becoming more familiar with personal voice assistants, voice interfaces, and voice recognition in general, it also seems that the excitement about the entire concept is quickly fading. This is noted in recent market surveys such as Qualcomm's *State of Play Report 2019* and *The Smart Audio Report* from NPR and Edison Research, which show that use of home control features remains stable year-on-year compared to 2018 and that use cases for smart speakers in the last three years did not change, with asking how's the weather remaining one of the dominant requests. People now understand what works reliably and they basically repeat.

It's already painfully obvious that voice assistants are not evolving to become smarter every day, neither is the technology expanding to more language markets as promised—certainly not at the same sophistication level, at least.

Essentially, voice recognition technologies as they have been implemented for now in smartphones, computers, smart speakers, domestic appliances, and smart-home solutions work reasonably well in English and Chinese, showing very basic capabilities in other languages, mostly depending on accurate



An interesting chart from Futuresource shows real market data about voice assistant usage on headphones. 2019 Audio Collaborative.



The Pumpkin Evaluation Kit Smart Audio Edition is a single-board computer (SBC) powered by a MediaTek MT8516 SoC for smart speaker and voice connected development.

DSP Concepts collaborated with StreamUnlimited, Pure, and renowned audio engineer Karl-Heinz Fink and his team, to bring the Braun Audio brand back to the premium audio market, including the industry's first multichannel product with hands-free, far-field voice control for Google Assistant.

voice-to-text conversion. The promise of artificial intelligence and machine learning turning voice engines exponentially "smarter" and even "learning" new languages and accents remains an exciting perspective for R&D departments and is the focus for academic research, but is not yet reflected in existing solutions.

The reality is setting in that our "voice engines" are not yet able to evolve as expected, and that progress is heavily dependent on basic human input. In fact, what is being advertised frequently as major progress in voice recognition technologies by Amazon, Google, and other voice engines, remains mostly in the realm of teaching the dog to do new tricks (or skills) with very little "intelligence" actually associated to the implementation itself. And yet, no one is willing to give up.

Smart Microphones

On the other hand, as the justifiable excitement about voice recognition applications continues—yes, even if basic, everyone recognizes that this is something that consumers want—so does the excitement around its enabling technologies and platforms. And that's the important thing. Far-field voice capture, microphones and sensors, edge-processing, and signal processing tools, including sophisticated noise reduction algorithms, sound signature recognition, and selective frequency processing are all evolving at an impressive pace thanks to the industry excitement generated around voice applications.

Everyone agrees that for "the machine" or the AI engine to be able to understand what is said and recognize who said it, we need to improve the voice signal that is captured and fed to the engine—under any circumstances, including in extremely noisy conditions or with multiples voices overlapping. Hence, the industry is focused on quickly improving all the enabling technologies, essentially adding the "smart" to microphones, amplifiers, and the entire audio chain.

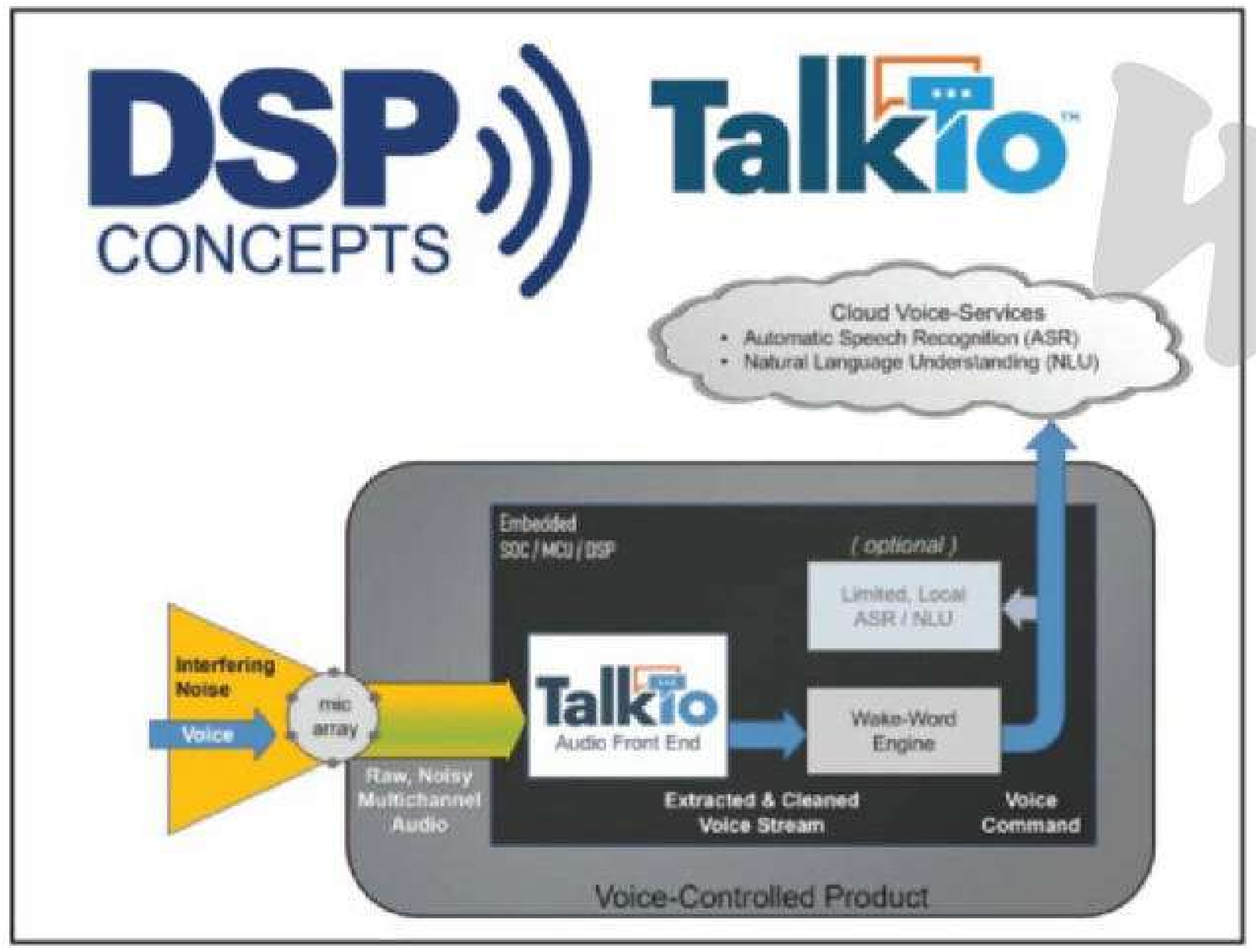
While a large chunk of the industry is focused on speakers and home applications, one of the pressing challenges for many manufacturers is also getting out of home and making voice recognition available in the car (one of the most obvious, and most promising application fields for voice control) and in personal, wearable devices.

And that's why all products, from smartwatches to true wireless earbuds are now receiving sophisticated edge-processing capabilities—to be able to cope with all the sensors and algorithms that are built in, even when they are not connected to the Internet. And even if they are not used for very sophisticated voice recognition applications yet, at least they will be able to power audio processing, such as active noise cancellation (ANC) and augmented hearing applications.

In the true wireless product category, where only now audio manufacturers are starting to be able to compete with Apple given the availability of dedicated low-power System-on-Chips (SoCs) such as the ones from Qualcomm, the topic of hearing augmentation implemented in true wireless stereo (TWS) products is something that everyone believes to be an opportunity, as the race toward hearables will continue. And voice recognition and voice engines are essential parts, mainly because everyone understands the threat of allowing companies such as Apple, Google, or Amazon to continue to dominate the virtual assistant space with their own products.

While the topic of virtual assistants being summoned directly through connected hearables is not consensual, most industry professionals see it as a major opportunity, and others feel it is one the greatest challenges and one of the reasons why the big players in the voice space could continue to "eat the audio brand's space," as a Futuresource analyst also describes it.

And that is why so much focus is being placed on low-power processing solutions and microelectromechanical systems (MEMS) microphones, preferably integrated, as the enabler for voice user interfaces.



DSP Concepts offers its complete suite of Audio Weaver audio processing tools with the TalkTo voice front-end to support advanced audio and voice capture applications.

The VocalFusion two-mic development kit from XMOS includes a XVF3510 processor base board with two Infineon IM69D130 PDM MEMS microphones and a Raspberry Pi HAT adapter board that provides additional I^2S audio and I^2C connectivity.

Improving Speech-Recognition

As very cleverly explained by Paul Beckmann, DSP Concepts CTO, one of the final obstacles to widespread voice interface adoption has been poor speech-recognition performance. "This inability for "the stupid thing" to hear you has now been solved by innovations "where the rubber meets the road" for voice-controlled devices: the Audio Front End (AFE)." This AFE, as Beckmann describes, is all the functional block that sits between a device's microphones and the rest of the voice-processing ecosystem, and something that DSP Concepts has been working to improve for the last few years. "Now, with the Audio Weaver processing platform available on nearly every IC, software-based AFEs such as TalkTo are enabling machines that can match and sometimes exceed a human's ability to understand speech in a noisy environment," Beckmann explains.

Basically, DSP Concepts developed Multichannel Acoustic Echo Canceller (AEC), which cancels the "known" sounds made by the device's own speakers and makes voice control possible even in Dolby Atmos soundbars; and a proprietary Adaptive Interference Canceller (AIC), which uses machine learning and advanced microphone-processing techniques to continually map and characterize the ambient sound-field. These two processing innovations available as the TalkTo advanced microphone-processing algorithms—part of the company's Audio Weaver platform—make it possible to feed a pristine, voice-only signal for use by the voice-ecosystem.

CEVA, one of the leading licensors of signal processing platforms and artificial intelligence

The rise of voice control has made audio processing critical. ON Semiconductor introduced its LC823455 Dual CPU processor with DSP, available in a complete SoC solution with high-resolution 32-bit/192 kHz audio processing and all the tools needed for wireless earbuds, wearables, and wireless speakers development with voice interfaces.

processors for connected devices, also partnered with DSP Concepts to use Audio Weaver tools and TalkTo on its CEVA-X2 and CEVA-BX family of DSPs. As CEVA confirms, Audio Weaver's graphic configuration tool enables rapid development of audio/voice DSP software, eliminating the need for native coding and kernel optimization, cutting product development time by up to 90%. CEVA's high-performance audio and voice DSPs are optimized for intensive sound processing applications, and the CEVA-X2 and CEVA-BX family of DSPs are ideal for any power-constrained audio/voice design, and are widely adopted in SoCs for smartphones, smart speakers, and any CE product with voice control.

DSP Concepts' Audio Weaver embedded audio design and processing tool is also the unifying link for an impressive number of development solutions that are available to manufacturers, starting with chipset maker MediaTek, which uses Audio Weaver and TalkTo as the audio platform for its i300 and i500 chipsets to power voice applications. MediaTek's chipsets also enable implementing all sorts of advanced multimedia features for connected devices.

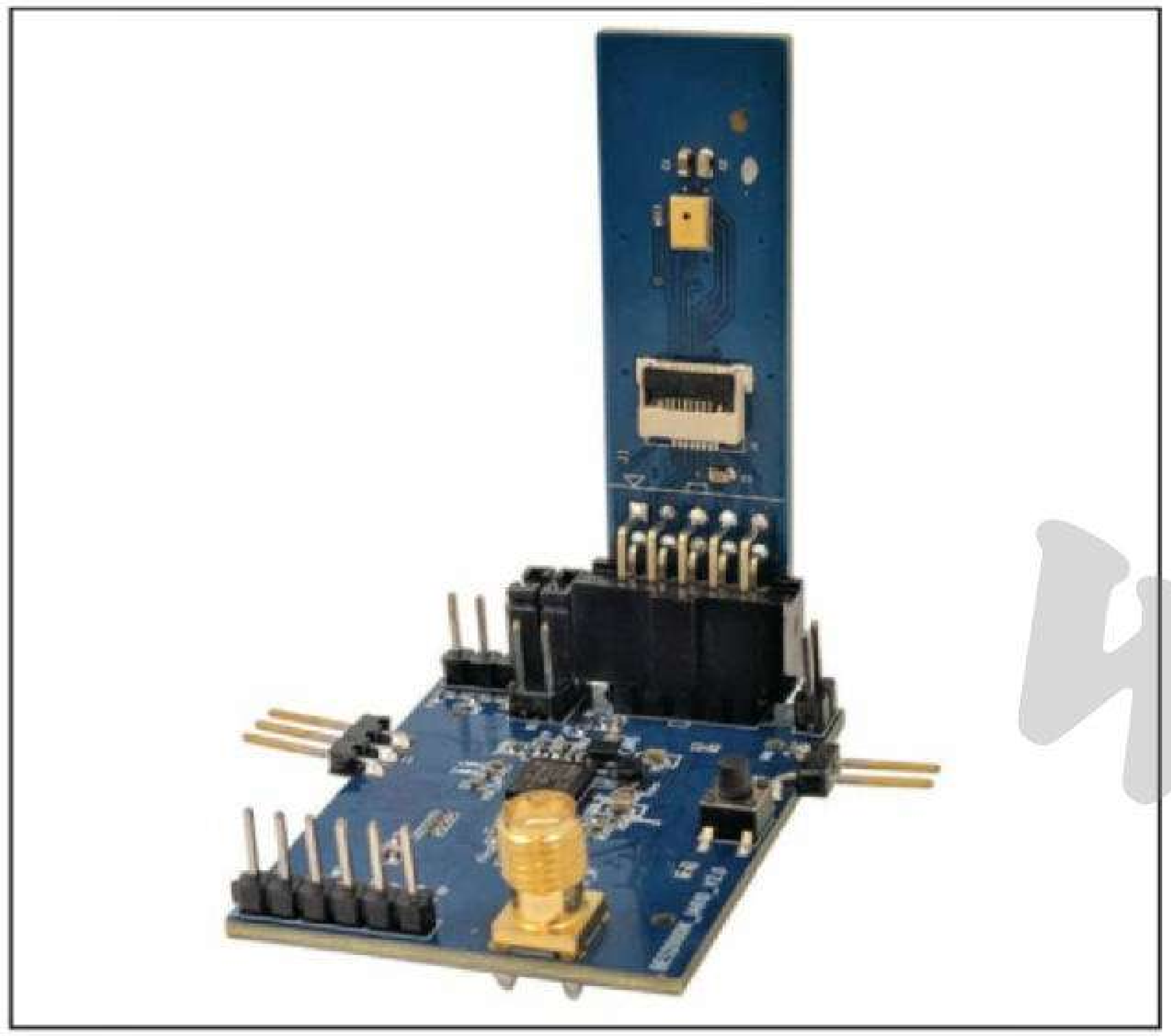
The other side of the equation is provided by the ability to trigger the use of the audio-processing power that is available on edge devices, only when a wake-word (e.g., "Alexa") or voice command (e.g., "Turn off the AC") is detected. This is powered by Sensory's TrulyHandsfree and TrulyNatural embedded technologies which run now on resource-constrained ICs and devices that are not connected.

A good example of a successful implementation for this technology is the recent Braun Audio line of "hi-fi smart speakers," resulting from a collaboration with DSP Concepts, Pure, StreamUnlimited, and audio consultant Karl-Heinz Fink. The Braun range, which resurrects a very popular concept in design from the 1950s, can work with voice recognition services such as Google Assistant—even from a distance (far-field) when the speakers are turned up to full volume. And, as Karl-Heinz Fink, who actively develops on Audio Weaver, states, all this power can be used to design also a better audio system. "Historically, audio engineers have wasted precious time jumping through hoops to get their ideas implemented. Now, thanks to Audio Weaver, we can focus on sound quality and ultimately develop a superior audio experience in significantly less time," says Fink.

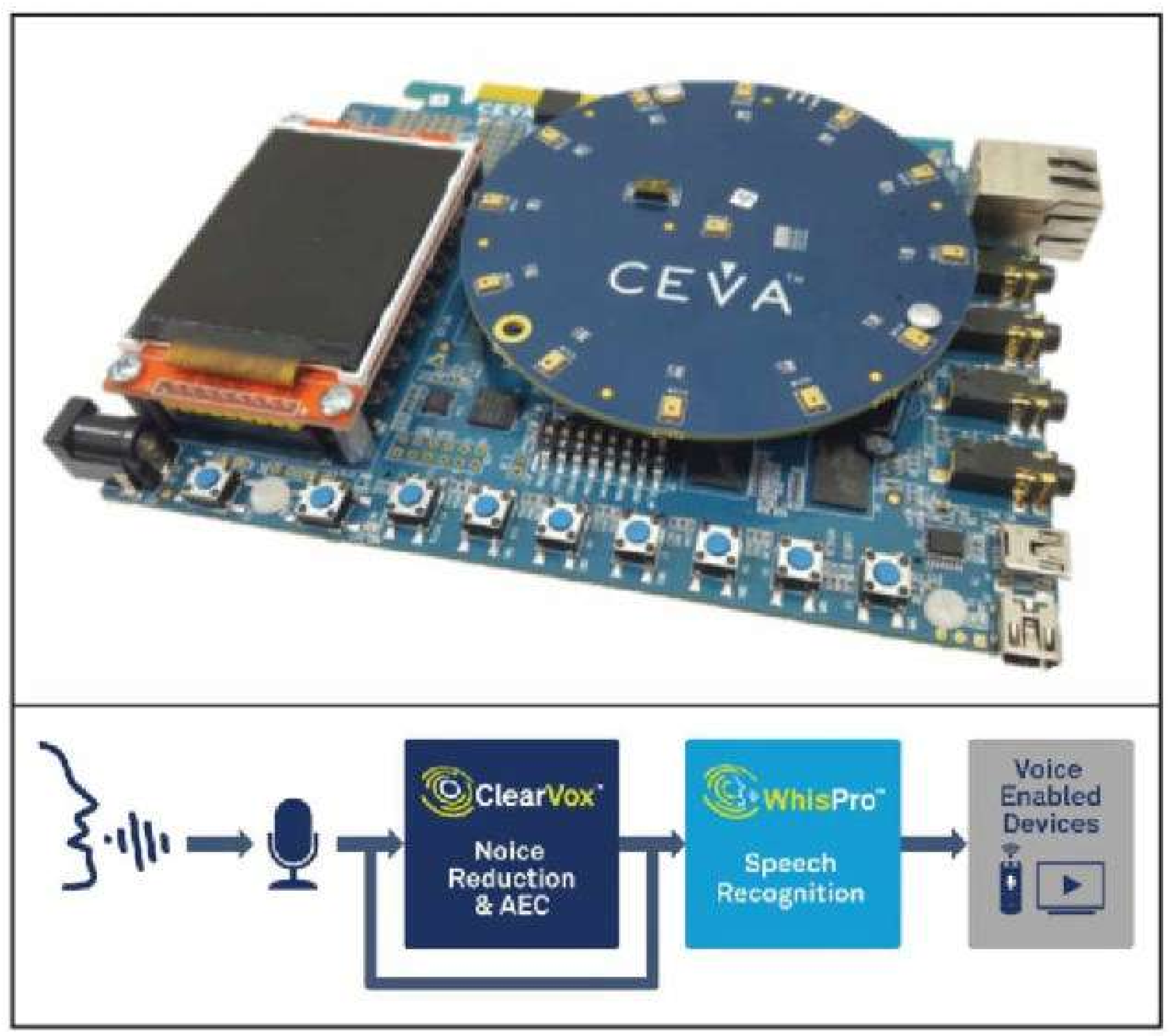
DSP Concepts is also working with Ambiq Micro, a provider of ultra-low power processors for hearables and wearables, and Sensory, to create new white-box solutions, development kits, and ODM solutions for voice-controlled ultra-low power consumer devices (e.g., smartwatches, earbuds, and smart home-end devices).

A similar solution that enables far-field voice capture in smart TVs and set-top boxes was also demonstrated by British company XMOS and has been now qualified by Amazon for Alexa Voice Services. The latest XMOS VocalFusion two-mic development kit enables manufacturers to build Amazon Alexa voice capabilities into their products more easily and more cost effectively, enabling all the performance of a four-mic solution. This new

Texas Instruments introduced the new TI Burr-Brown TLV320ADC5140 quad-channel audio ADC with support for far-field voice capture, achieving a 120 dB dynamic range for clear, high-fidelity audio in smart home applications.



The Knowles AISonic SmartMic Headset Development Kit for Amazon AVS



CEVA offers WhisPro speech recognition technology as an always-listening multi-trigger phrase technology for smartphones, smart speakers, Bluetooth earbuds, and other voice-enabled devices. WhisPro operates in tandem with CEVA's ClearVox front-end voice processing software technology and CEVA offers the WhisPro development board (pictured) and a smart speaker reference design with a complete set of drivers and add-on software for beamforming, AEC, direction of arrival, voice activity detection, etc.

XMOS development kit, which is based on the XMOS XVF3510 voice processor, is able to power next-generation acoustic algorithms, including full duplex stereo acoustic echo cancellation (AEC), barge-in, with Interference Canceller, Noise Suppression, Automatic Delay Estimator Control (ADEC), and Automatic Gain Control (AGC). The solution also features an Automatic Delay Estimator to adjust the audio reference signal latency dynamically, ensuring the AEC algorithms deliver a smooth, real-time barge-in experience. XMOS is now able to offer a complete range of solution for sophisticated voice-enabled audio systems, as illustrated by the company's recent collaboration project with Tymphany to create a reference design for soundbars with far-field voice capture.

Improving Mobile Audio

There are many possible examples of recent technologies introduced mainly to power voice recognition, but which can be effectively leveraged by manufacturers to implement many other powerful features. For example, there is Knowles latest AISonic audio edge processor, which is optimized for high-performance mobile, ear, and connected devices. The IA8201, the latest product in the Knowles AISonic family of audio edge processors, offers robust voice activation and multi-microphone audio processing optimized for power-sensitive applications, and is the first processor specifically designed for advanced audio features, with enough compute power to support advanced audio output, context awareness and gesture control functions. More importantly, this is a low-power solution that enables far field voice capture in noisy, real-world environments with "10 to 100 times the efficiency of generic processors," as the company states.

This audio processor packs 1.44 MB of memory, numerous interface options, plus a high compute 128-bit core (DMX) with Knowles' proprietary instruction set and a Cadence Tensilica HiFi3 core (HMD), both with Knowles' audio enhancements. The DMX is a four-way floating-point SIMD processor that powers beam-forming, barge-in AEC, while the HMD enables wake-on-voice applications. The solution is complemented with Knowles' open DSP platform that supports third-party algorithm providers. The Knowles AISonic family supports Sensory's latest TrulyHandfree wake word technology and even Waves Maxx's powerful processing algorithms.

Alango Technologies was one of the early adopters of the IA8201, which used it to implement its own advanced voice and audio processing algorithms, including sound personalization in hearables.

For more demanding applications, Knowles also offers the AISonic IA8508, a fully customizable, quad-core audio edge processor with four times the memory of the IA 8201, which allows for larger models and concurrent processing of multiple, high-performance algorithms for applications requiring intensive edge processing and privacy.

Also focused on development of headphones, headsets, and wireless earbuds with voice support, Knowles introduced the new AISonic SmartMic Headset Development Kit for Amazon AVS, which includes the Knowles AISonic IA611 SmartMic and a low-power audio-edge processor inside a tiny microphone. This OEM/ODM solution is fully compliant with the Alexa Mobile Accessory (AMA) protocol for Bluetooth accessories and is optimized for small battery-operated devices, interfacing to the Bestechnic BES2000i Bluetooth SoC, which controls overall headset functionality and provides Bluetooth and Bluetooth Low Engergy (BLE) connectivity. Retune DSP VoiceSpot technology provides the voice trigger and commands in the SmartMic Development Kit platform.

This solution, effectively tackles the challenges of speech recognition and voice control with car or street noise and was already adopted by Linkplay in the development of new true wireless designs for big consumer electronics brands.

Editor's Note: For updates on solutions targeting voice capture and recognition, audioXpress *offers continuous coverage on its dedicated page available at: https://audioxpress.com/select/voice-recognition*

Resoures

Qualcomm, *The State of Play Report 2019*
https://assets.qualcomm.com/the-state-of-play-2019.html

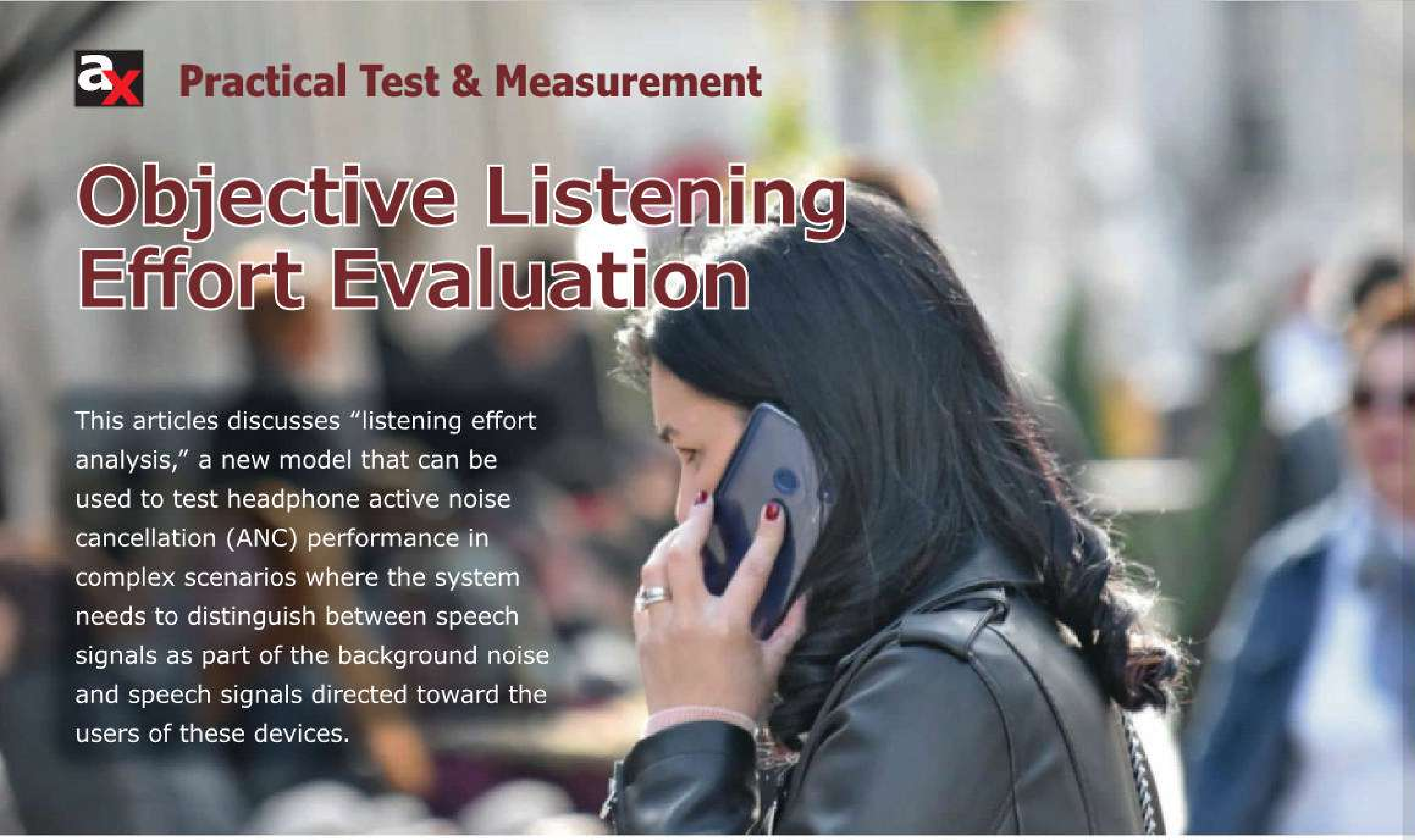
Strategy Analytics, *2019 Global Smart Home Forecast*
https://www.strategyanalytics.com/access-services/devices/connected-home/smart-home/market-data/report-detail/2019-global-smart-home-forecast---september-2019

Strategy Analytics, *2019 Global Smart Home Devices Forecast*
https://www.strategyanalytics.com/access-services/devices/connected-home/smart-home/market-data/report-detail/2019-global-smart-home-devices-forecast---september-2019

The Smart Audio Report from NPR and Edison Research
https://www.nationalpublicmedia.com/insights/reports/smart-audio-report

Sources

CEVA, Inc. | www.ceva-dsp.com
DSP Concepts | www.dspconcepts.com
Knowles Corp. | www.knowles.com
MediaTek, Inc. | www.mediatek.com
Sensory, Inc. | www.sensory.com
XMOS | www.xmos.com



Objective Listening Effort Evaluation

This articles discusses "listening effort analysis," a new model that can be used to test headphone active noise cancellation (ANC) performance in complex scenarios where the system needs to distinguish between speech signals as part of the background noise and speech signals directed toward the users of these devices.

By
Hans W. Gierlich and Jan Reimes
(HEAD acoustics)

Environmental noise in daily life is increasing, while users still expect to use their audio and communication equipment in all daily life situations. Smartphones providing speech enhancement of the receive-side signal and headphones providing noise cancellation functionality are a logical technical remedy for this issue and are becoming increasingly popular.

Besides reducing ambient noise, these devices include processing technologies that may amplify or attenuate speech signals but also may degrade the speech signal quality. When speech is part of the environmental noise attenuation may be intentional. However, speech often is the desired signal and in consequence should be enhanced rather than attenuated or degraded.

Although a lot of effort has been spent on the development of objective, perceptual-based methods [1], [2] until recently no method was available for determining the speech intelligibility or the Listening Effort at the receive side of telecommunication systems. So far, it has been impossible to predict the performance increase by any type of adaptive voice enhancement processing objectively and adequately.

To overcome this situation, some years ago HEAD acoustics started to develop a perceptual-based method targeting the objective prediction of Listening Effort for a variety of technical systems:

- Active Noise Cancellation (ANC)
- Near-end voice enhancement, including all types of signal processing such as companding, expanding, adaptive level control, adaptive frequency response shaping bandwidth-extension
- Voice enhancement used in real-time systems/ announcement systems

The application areas for such a method are:

- Smartphones and other smart devices (smart home)
- Headphones and headsets especially with ANC technologies
- Hands-free, conference and car hands-free
- In-Car Communication (ICC)

This work results in a new objective method for determining Listening Effort called "ABLE" (Assessment of Binaural Listening Effort), which was recently standardized in ETSI TS 103 558 [7] and which is available as a standardized option for the voice and audio quality measurement and analysis software ACQUA.

This article introduces the basics of the underlying auditory tests, a description of the prediction model and application examples.

Finding the Right Data Set

As a basis of any objective prediction model, listening examples are required to train and validate the model. These listening examples need to cover the entire quality range of devices for which the model is to be developed. Furthermore, these listening examples need to include all impairments which may arise from signal processing and which are relevant for the user's subjective impression.

When developing an objective test method, in consequence different applications, ambient noises, devices, implementations and possible speech processing algorithms need to be taken into account. In order to predict all these scenarios reliably, numerous noisy speech samples have to be available. To generate these types of listening examples, advanced signal processing techniques for simulation as well as a big variety of devices should be available. In addition, a variety of realistic noise scenarios where these devices are used in, need to be available.

A typical setup for generating such listening examples, which also is used for testing real devices, is shown in **Figure 1**. Here, the example of testing an ANC headset in conjunction with a Head-and-Torso Simulator (HATS) is illustrated. Pre-recorded background noise is played back in a laboratory using a standardized sound-field generation system (ETSI TS 103 224 [3]). This setup reproduces various background noises representative for the typical environmental noise situations with a high degree of accuracy around the device under test (DUT).

Auditory (Subjective) Test Databases

When evaluating Listening Effort subjects provide a self-assessment on a five-point categorical scale, similar to well-known speech quality testing methods. For evaluation of Listening Effort, the scale shown in **Table 1** is used.

The scale and the corresponding attributes are taken from Recommendation ITU-T P.800 [4]. Besides the aforementioned benefits of Listening Effort testing, recent studies (e.g., [5, 6]) indicate that speech enhancement benefits can be evaluated in a wider range of signal-to-noise ratios (SNRs) without reaching positive or negative saturation observed in intelligibility tests.

Auditory tests can be conducted in groups in parallel. A set of test conditions are combined in a test database. Each test database contains a set of speech samples processed with different devices and/ or algorithms ideally covering the complete range of Listening Effort and a set of reference conditions. Mapping of different test databases to each other is realized by mapping to the reference conditions.

The purpose of auditory tests is twofold:

- So-called training databases are used to develop and train the objective model for best performance.
- So-called validation databases are needed after model development is completed. The auditory test results of these databases were never seen before by the model. They provide information about the robustness and validity (generalization) of the objective model.

The Instrumental Model ABLE

ABLE, the instrumental model for Listening Effort was recently standardized in ETSI TS 103 558 [7]. A detailed description about the model can be found

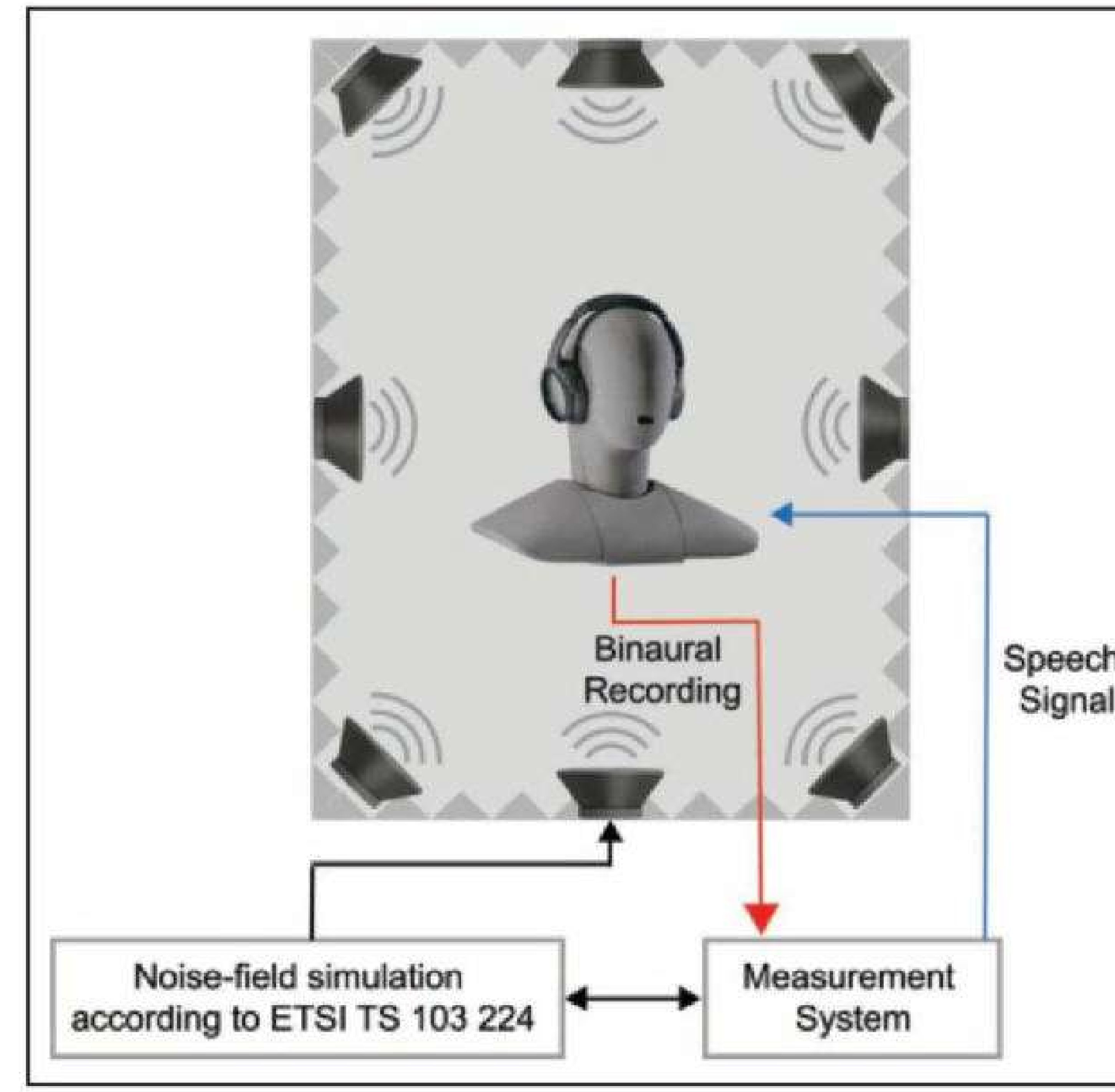


Figure 1: Test setup for Listening Effort evaluations with background noise

Score	Listening Effort
5	Complete relaxation possible, no effort required
4	No appreciable effort required
3	Attention necessary, moderate effort required
2	Considerable effort required
1	No meaning understood with any feasible effort

Table 1: For evaluation of Listening Effort, this scale is used.

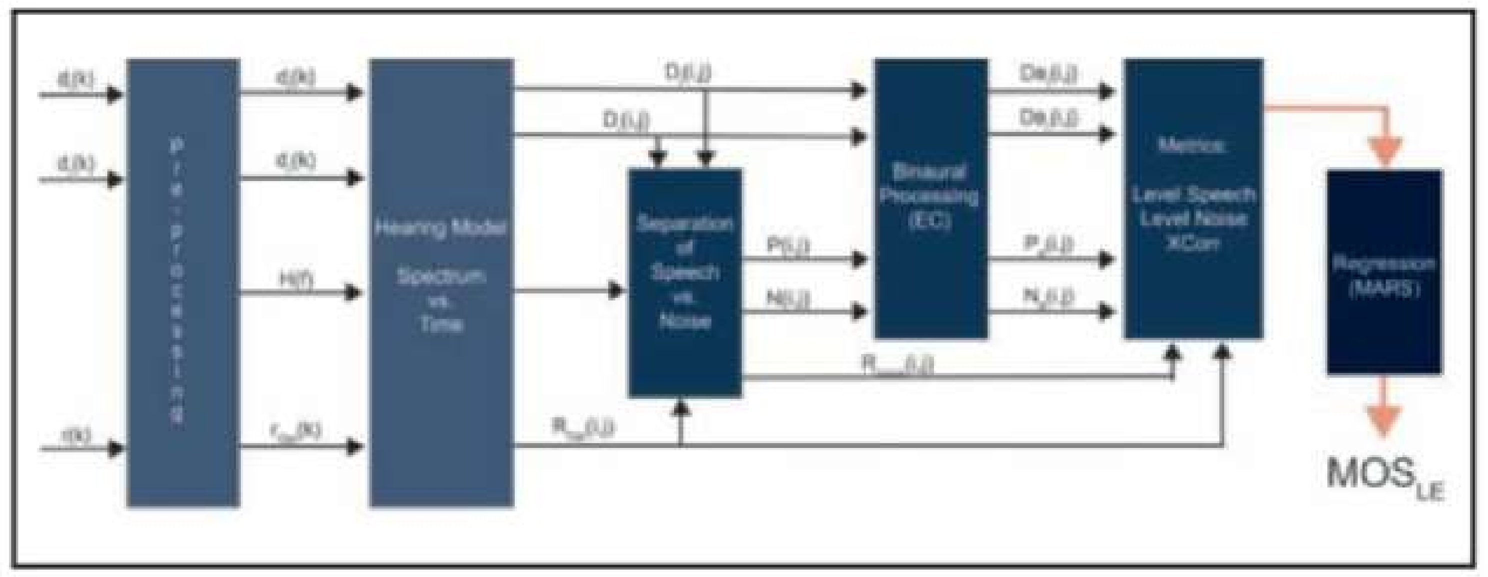


Figure 2: Flow chart of the prediction algorithm ABLE for Listening Effort (according to ETSI Standard 103 558)

in this standard. **Figure 2** provides an overview of the prediction algorithm, which consists of different stages. ABLE is a binaural model and in general the input to the model are binaurally recorded scenarios. The binaural input signals to the model are $d_L(k)$ and $d_R(k)$. The reference signal r(k) is the clean speech signal, which is inserted in the DUT.

After pre-processing, a transformation into the time-frequency domain based on a hearing model is applied (time index i, frequency bin j). These signal representations are denoted with capital letters D(egraded), N(oise), P(rocessed) and R(eference). The processing is performed for the left and right ear separately.

About the Authors

Hans W. Gierlich started his professional career in 1983 at the Institute for Communication Engineering at RWTH, Aachen. In February 1988, he received a Ph.D. in electrical engineering. In 1989, Gierlich joined HEAD acoustics GmbH in Aachen as vice president. Since 1999, he is head of the HEAD acoustics Telecom Division and in 2014, he was appointed to the board of directors. Gierlich is mainly involved in acoustics, speech signal processing and its perceptual effects, QOS and QOE topics, measurement technology and speech transmission quality. He is active in various standardization bodies such as ITU-T, 3GPP, GCF, IEEE, TIA, CTIA, DKE, and VDA and chairman of the ETSI Technical Committee for "Speech and Multimedia Transmission Quality".

Jan Reimes started his academic studies of electrical engineering at RWTH Aachen, Germany in 2000. In 2006, he received his diploma degree in information and communication technology with final thesis at the Institute of Communication Systems. From 2006 to 2015, he was employed at HEAD acoustics GmbH for the Telecom development department. In 2016, he joined the newly founded department "Telecom Research & Standardization". His fields of interest and research are in measurement technology and perceptual based models. He is active in a variety of standardization bodies such as ITU-T, ETSI and 3GPP.

About HEAD acoustics

HEAD acoustics GmbH is one of the world's leading companies for integrated acoustic solutions as well as sound and vibration analysis. In the telecom sector, the company enjoys global recognition due to the expertise and pioneering role in the development of hardware and software for the measurement, analysis and optimization of voice and audio quality as well as for customer-specific solutions and services. HEAD acoustics' range of services covers sound engineering for technical products, investigation of environmental noise, speech quality engineering as well as consulting, training and support. The medium-sized company from Herzogenrath near Aachen has subsidiaries in China, France, Italy, Japan, South Korea, the UK and the USA as well as numerous sales partners worldwide.

Pre-Processing

The purpose of the pre-processing stage is the temporal and level alignment of all input signals. The delay between degraded signals d(k) and the reference r(k) is calculated by a cross-correlation analysis for both ears. The reference signal r(k) is scaled to a fixed active speech level (ASL) of 79 dB_{SPL} ($r_{Opt}(k)$), which is assumed to be optimal regarding Listening Effort [8, 9]. The transfer functions H(f) between degraded and reference signal are calculated by the H1 methodology (cross-power spectral density) in order to exclude non-correlated noise components.

Hearing Model Calculation

An aurally adequate transformation is performed in the next step using the HEAD acoustics hearing model [10,11]. This transformation is applied to each of the binaurally recorded signals as well as to the clean speech reference signal. The transformation includes an auditory filter bank representation of the signal and a hearing-adequate envelope. This results in a frame resolution of about 8 ms. The initial frequency bandwidth Δf(f0) is set to 70 Hz. In total, 27 hearing-adequate bands up to 20 kHz are used. This time-frequency representations result in the hearing model spectra vs time DL(i, j), DR(I,j) and ROpt(i, j) (see Figure 2).

Separation of Speech and Noise

For the Listening Effort calculation, the degraded spectra D(i, j) are separated into sections with (processed) speech (P(i, j)) and sections representing the noise (N(i, j)) component. For this purpose, a (pseudo-)Wiener filter is used here for the determination of P(i, j). Further details can be found in [7].

For the determination of N˜ (i, j), first a soft mask M(i, j) of active/inactive time-frequency bins is determined with $R_{Comp}(i, j)$. The speech part classification is performed for each frequency band. For silent and paused frames, a mask value of 1.0 is set, for high activity of speech to 0.0. Weights for medium, low and uncertain activity frames are calculated by linear interpolation. An example of this threshold-based method for one single frequency band is provided in **Figure 3**.

The degraded spectra are multiplied by the masks and an estimate N˜ (i, j) of the noise signals per frequency band are calculated via deconvolution. N˜ (i, j) is then used to estimate P(i,j) with the aforementioned Wiener filter and the estimated noise spectra are used as the representation of the noise itself: N(i, j) ≈ N˜ (i, j).

Binaural Processing

Besides its ability to identify the sound source direction, the human ear is capable to improve the SNR when listening binaurally as compared to monaural listening. This is modeled in the binaural processing block of the prediction model.

The spectral components for left and right ears are combined by the short-term equalization/cancellation methodology as described in [12], which is an extension to the well-known model of Durlach [13]. This block requires the availability of the isolated speech and noise (masking) components.

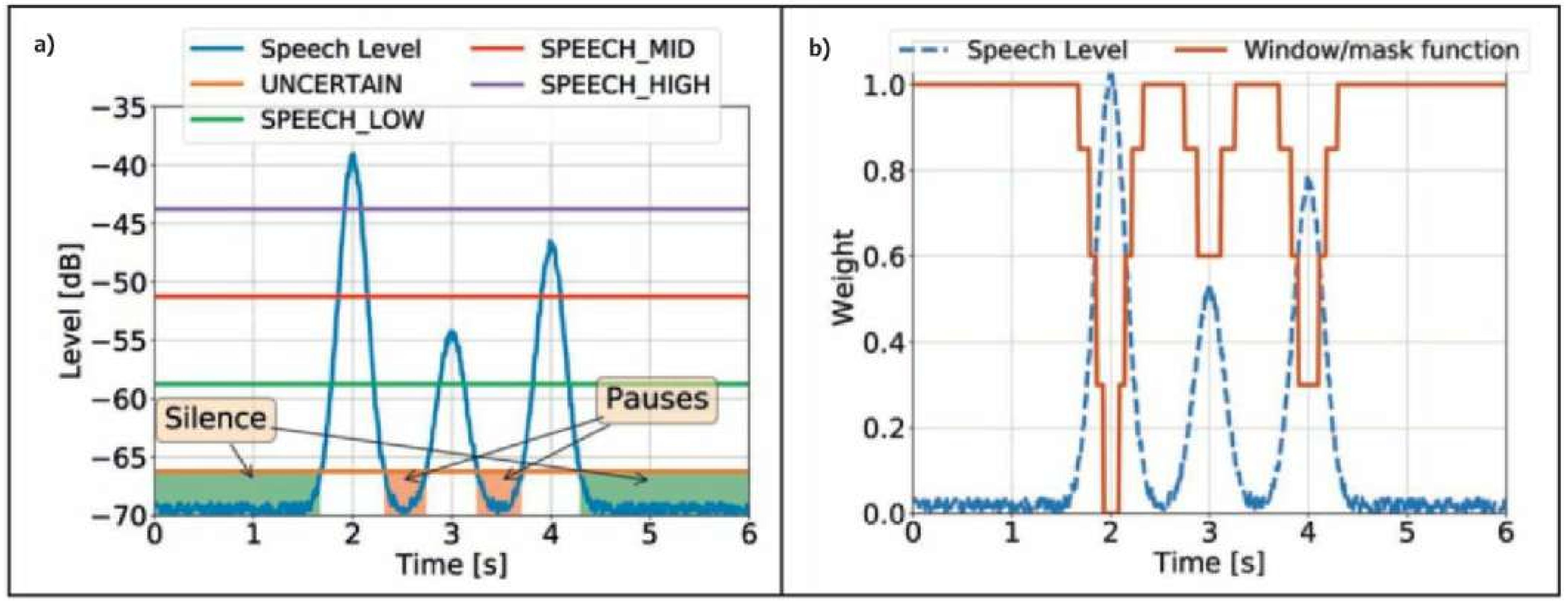


Figure 3: Frame classification (a) and mask generation (b)

At the output of this block combined and enhanced hearing model spectra vs time are provided.

Metric Calculation

A variety of metrics can be calculated based on the binaural spectra. For ABLE, level-based and correlation-based metrics were chosen. For the level-based metrics, the ASL is used for those segments, which were classified as active speech based on the frame classification of the pre-processing. In addition, an A-weighted noise level calculated from the noise spectrum.

The correlation-based metrics are derived from a spectral cross-correlation, which is carried out in four different ways. First, the threshold of hearing is applied [14] to the spectra. After this step, the nonlinear loudness transformation according to [10, 11] is applied to all spectra. Each band is sub-divided into sub-frames of 320 ms (N = 40 frames, only active indices) with 50 % overlap. These are aggregated to an overall metric. More information can be found in [7].

MOS Calculation

From the metric block, various single values are available and combined to an instrumental single Listening Effort score. Machine learning procedures are used here for the derivation of the instrumental Mean Opinion Score (MOS). Due to the limited amount of training databases available until now, multivariate adaptive regression splines (MARS) according to [15] are used in the current model in order to avoid overfitting. Parameters and coefficients of the regression are derived from data of similar studies related to the auditory assessment of Listening Effort [16, 17].

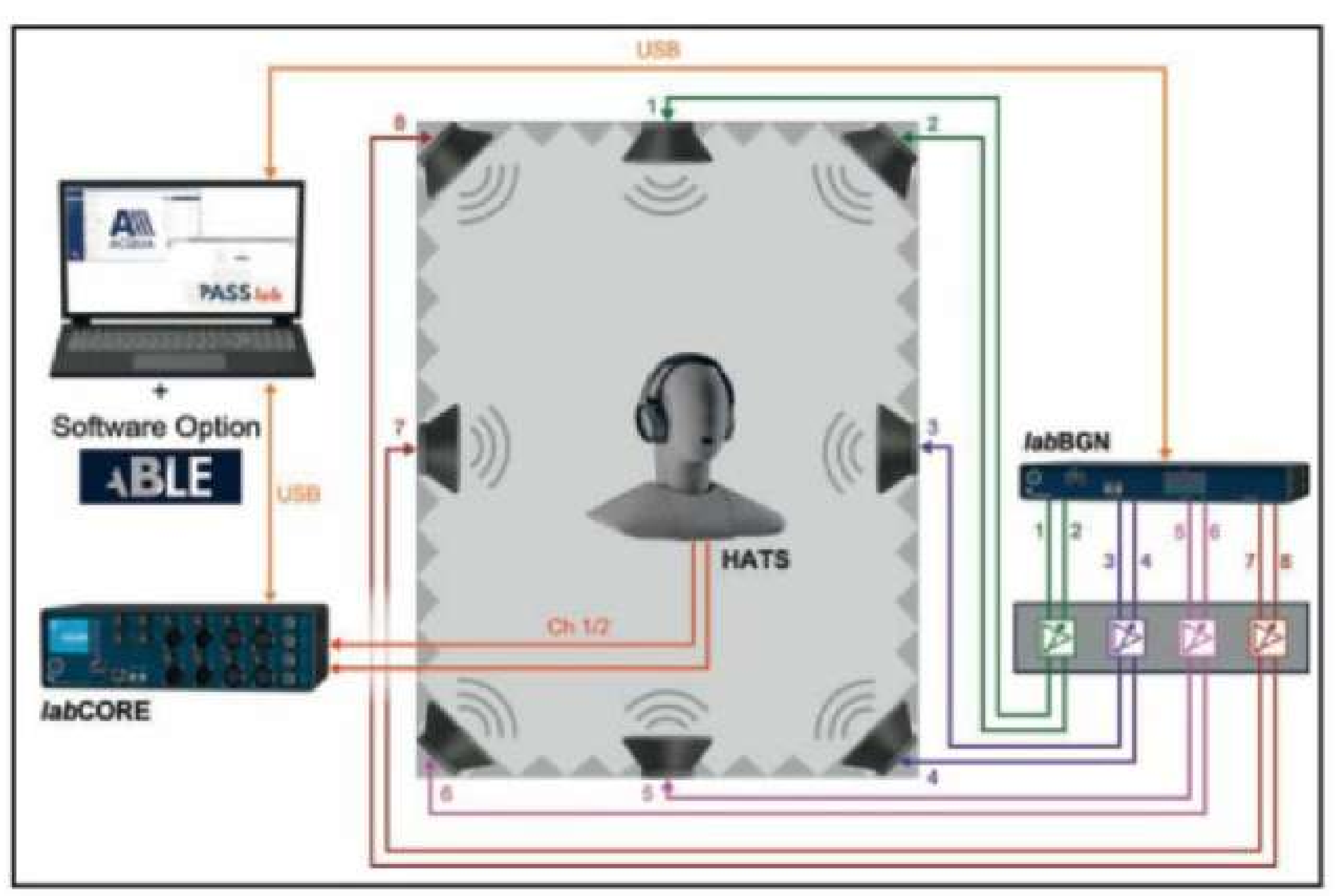


Figure 4: Test setup of ANC headphones in communication mode

Applications of ABLE

As outlined before, the objective model ABLE can be used for various devices and applications. Two rather different application examples are chosen to demonstrate the capabilities of ABLE – ANC headsets and In-Car Communication (ICC) systems.

ANC Headsets

Besides active noise cancellation, ANC headsets are often used for communication with ANC active; many devices also offer a so-called talk-through function, which allows passing the speech from a person or from an announcement system. A typical use case is the aircraft, where ANC functionality is in general highly desired, but communication with the service staff or listening to the announcements before/during takeoff and landing is required without removing the ANC headset.

To illustrate the capabilities of ABLE, four commercially available ANC headsets (here denoted as DUT A to D) were evaluated with the measurement using the setup described in the introduction (see Figure 1). DUT A, C, and D represent circumaural headsets, while DUT B is an in-ear device.

Two noise types were used for the noise field simulation according to [3]:

- Train compartment: This scenario was recorded in particular for this evaluation. On average, the noise level is ≈ 64 dB_{SPL}(A), measured close to the ears.
- Airplane: The aircraft noise is taken from the noise database of [3]. On average, the noise level is ≈ 80 dB_{SPL}(A), measured close to the ears.

When evaluating the Listening Effort for the different use cases, speech signals are inserted in various ways. The type and calibration of speech playback depends on the specific use case, two of them are described in the following clauses. The background noise playback is always presented synchronously with the speech signals, synchronized with the start of the measurement. So each device is exposed to exactly the same mixture of speech and noise.

The speech signal used for all evaluations is a speech sequence consisting of 16 American English sentences according to Annex C of ETSI TS 103 281 [2]. The binaural signals are captured at the drum reference point (DRP), a diffuse-field (DF) correction is applied (see [18]).

Communication Mode

A typical application of an ANC device is the usage as a headset (i.e., the device) is connected to a mobile phone or computer for telecommunication purposes. After connecting, the headset establishes a voice call and the speech signal is inserted into the downlink (receive direction). The measurement setup is shown in **Figure 4**. The headsets investigated were connected to a Bluetooth reference gateway. The connection was established with hands-free profile in wideband mode, which allows an audio bandwidth up to 8 kHz. The active speech level (according to ITU-T P.56 [19]), measured at the ear under silent conditions was adjusted to a range of 73 to 79 dB_{SPL} per ear, which is considered as an optimum level regarding Listening Effort and quality [8], [9]. This level was adjusted individually for each device. For all devices, ANC functionality was enabled and set to maximum cancellation operation mode where possible.

The results for the different headphones in communication mode for the different noises are shown in **Table 2**. As expected, the Listening Effort decreases significantly with increasing noise levels, which can be seen when comparing results between train compartment and airplane noise scenarios.

The speech level S_{act} values remain stable within a range of ~74-80 dB_{SPL}, while the noise cancellation performance and impact on MOS_{LE} differ, depending on scenario and device. The remaining noise level between the devices varies up to 17 dB. Even though DUT C provides quite low speech levels in both noise scenarios, this device outperforms the others with respect to residual noise level and MOS_{LE} = 2.8 is assessed for the most demanding noise type "airplane." The lowest MOS_{LE} is measured with device A. The rank order in performance also depends on the type of noise. While for airplane noise headset C outperforms the other devices, headset B performs best with train compartment noise.

In general, it would be expected that speech playback via headphone would result in even better values for MOS_{LE}. However, many devices include

Mode	Communication Mode							
Scenario	Airplane Scenario				Train Compartment Scenario			
Device	A	B	C	D	A	B	C	D
MOS_{LE}	1.8	2.2	2.8	2.3	2.8	3.7	3.5	2.9
SNR(A)[dB]	10.5	20.1	21.5	18.5	25	36.7	32	25.1
$S_{act}[dB_{SPL}]$	79.8	80.2	74.2	78.9	79.3	79.9	77.4	77
$N_A[dB_{SPL}(A)]$	69.4	60.1	52.7	60.4	54.3	43.2	45.4	51.9

Table 2: Results for in communication mode for the different background noises

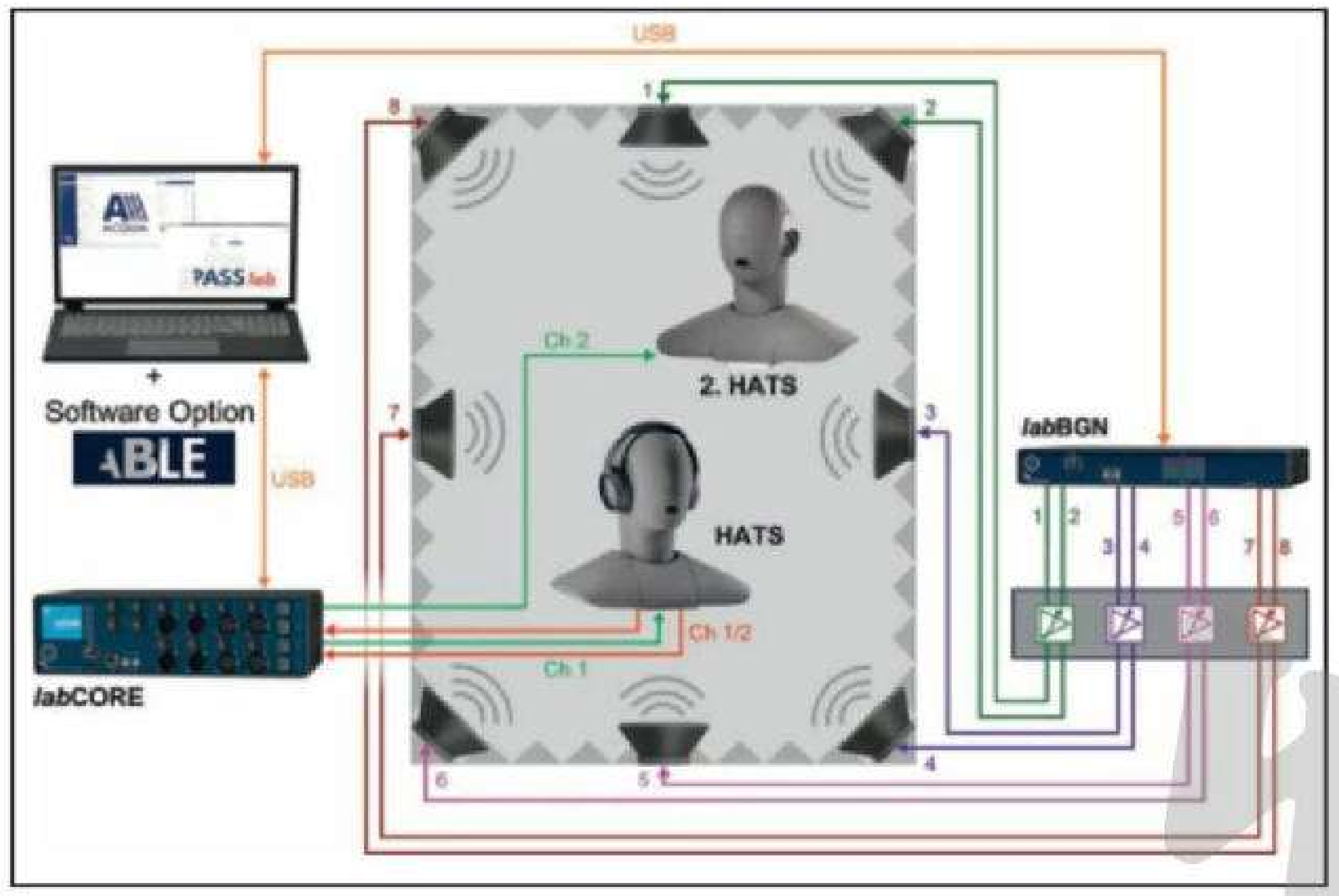


Figure 5: Test setup of ANC headphones with two HATS in talk-through mode

Scenario	Airplane	Train Compartment
Noise Level [dB_{SPL}(A)]	80	64
ASL Offset of Second Talker [dB]	9.1	4.8
ASL at MRP of Second Talker [dB_{PA}]	4.4	0.1

Table 3: Level adjustments for second talker setup

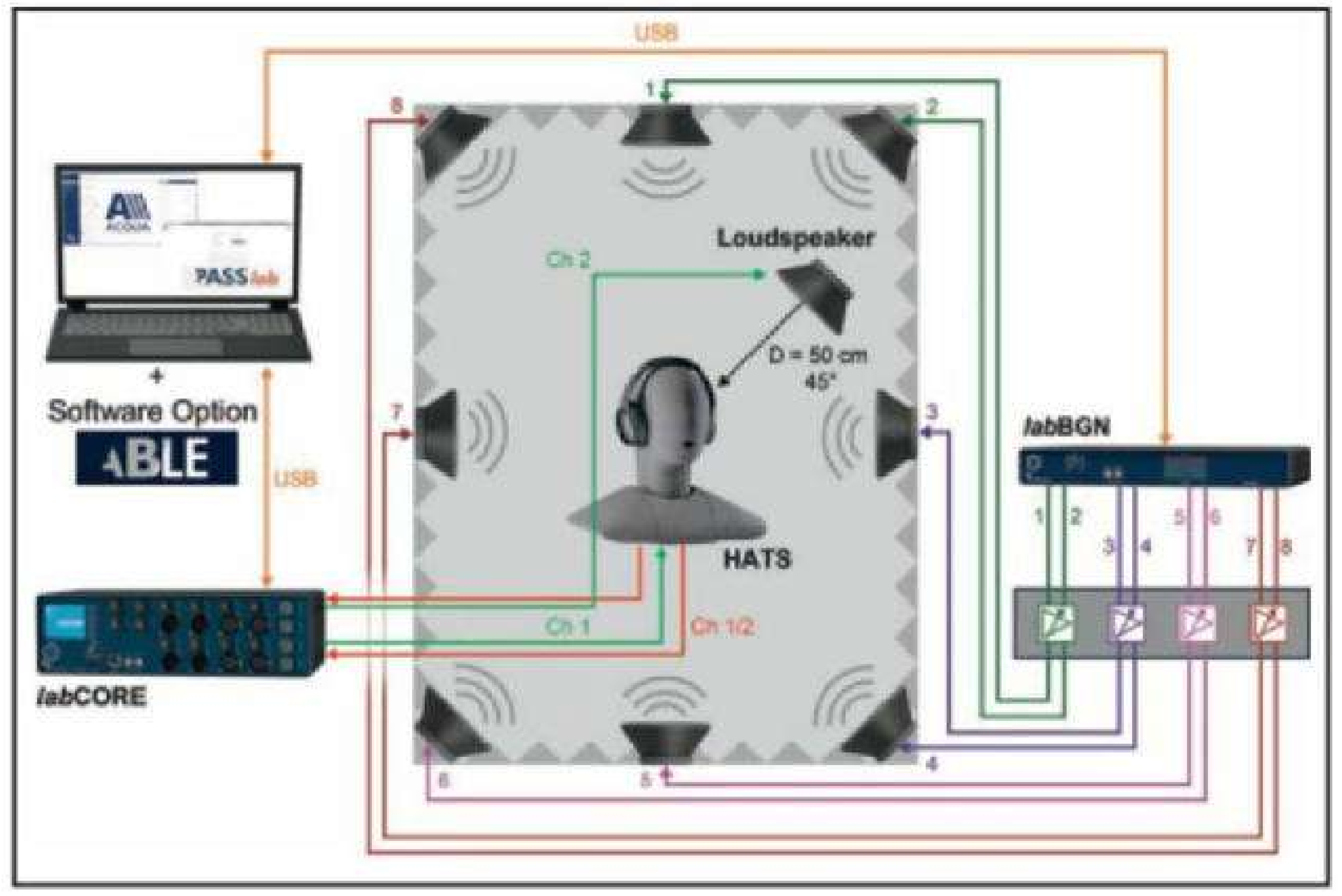


Figure 6: Test setup of ANC headphones with a loudspeaker simulating an announcement system

Scenario	Airplane	Train Compartment
Noise Level [dB_{SPL}(A)]	80	64
ASL at Listener [dB_{SPL}]	78.5	72
SNR at Listener [dB]	-1.5	8

Table 4: Level adjustments for loudspeaker setup

amplification of the audio signal in the low-frequency domain, which relatively reduces high-frequency components important for speech intelligibility. Further impairments may be due to non-optimum frequency response shaping, bandwidth limitations and speech coding which degrades the quality of the transmitted signal.

Talk-Through Mode

Speech signals originated from the surroundings (e.g., another talker or public address systems) are not easy to identify as wanted signals and may be attenuated in the same way as the environmental noise. To overcome this challenge, many devices employ a so-called talk-through (TT) functionality, which is evaluated more in detail.

To evaluate the talk-through performance of the devices, the talker can be simulated by a second HATS positioned at a distance of d = 50 cm to the left of the listener as shown in **Figure 5**. The artificial mouth of the second artificial head is equalized up to 20 kHz, providing most natural presentation of the second talker's voice. The speech signal level is -4.7 dB_{Pa} in silence, in noisy scenarios the speech signal level is increased considering the Lombard Effect (humans talk louder in noise scenarios) as described in [20]. The ANC functionality is enabled and set to maximum cancellation operation mode for all devices. The corresponding level settings are shown in **Table 3**.

Besides the communication to a human, the talk-through mode is useful for listening to announcement systems (e.g., in an airplane). For this simulation, an equalized loudspeaker is positioned at a distance of 50 cm a height h = 50 cm above the horizontal plane of the artificial head and moved 45° in front of the listener (see **Figure 6**). The loudspeaker output is calibrated to obtain a typical and reasonable level and SNR at the listener's ears (without wearing a device) for such scenarios. The corresponding level settings are shown in **Table 4**.

Table 5 and **Table 6** show the prediction results for the talk-through mode with a second talker and a PA-simulation (the loudspeaker scenario) for both noise scenarios across the four headsets.

In the talk-through scenario with a second talker (Table 5), the speech level S_{act} decreases compared to the communication mode (in some cases more than 10 dB), due to active and passive attenuation of any sound field outside the headset. For the second talker scenario, the noise cancellation performance remains comparable across devices and noise types. The speech attenuation is higher than in communication mode, in some cases up to 9 dB (DUT D for train compartment).

Obviously, the noise cancellation performance is reduced compared to communication mode, where speech is simultaneously played back. Nevertheless, the MOS_{LE} is comparable. Listening Effort is less sensitive against the speech level itself. Especially for low SNR ranges, the noise level reduction seems more important.

Mode	Talk-Through Mode: Second Talker							
Scenario	Airplane Scenario				Train Compartment Scenario			
Device	A	B	C	D	A	B	C	D
MOS_{LE}	1.9	2.3	2.7	2.1	2.6	3.7	3.6	3.4
SNR(A)[dB]	5.1	8.5	10.1	9	12.3	23.2	16.8	18.2
$S_{act}[dB_{SPL}]$	72.6	64.8	61.3	66.7	67.9	60.5	55.3	61.3
$N_A[dB_{SPL}(A)]$	67.5	56.3	51.3	57.6	55.6	37.3	38.5	43.2

Table 5: Talk-through performance with a second talker

Mode	Talk-Through Mode: Loudspeaker							
Scenario	Airplane Scenario				Train Compartment Scenario			
Device	A	B	C	D	A	B	C	D
MOS_{LE}	2	2.2	1.9	2.2	1.9	2.2	2	2.1
SNR(A)[dB]	10.3	10.9	4.5	9.9	5.5	8.6	-1.2	6
$S_{act}[dB_{SPL}]$	74.2	68.6	69.3	66.7	73.1	65.5	54.2	63.7
$N_A[dB_{SPL}(A)]$	63.8	57.6	64.8	56.7	67.6	56.9	55.4	57.7

Table 6: Talk-through performance with an external loudspeaker (PA-system simulation)

Table 6 shows the prediction results for the talk-through mode with an additional loudspeaker for the two noise scenarios. In this setup, a dedicated talk-through mode was activated. The noise cancellation performance decreases although the speech level increases compared to the previous talk-through mode. The Listening Effort does not seem to benefit from this increase; results for MOS_{LE} are consistently lower. The concurrent gain of noise level is predominating here.

It becomes clear that the dominating factors for the assessment of Listening Effort are the (residual) noise and the (mere) speech level. However, it is not sufficient considering just the SNR for the Listening Effort. It seems more important to investigate at which speech/noise level a certain SNR is obtained. This can be seen for the talk-through scenarios, where wide ranges of speech and noise levels are covered.

In-Car Communication (ICC)

Another application area for ABLE is evaluation of ICC systems. ICC systems are used to enhance the communication between passengers in a car. Conversation between passengers may be impaired due to the driving noise present and the high speech attenuation, which is caused by e.g., an acoustically high-quality treated car or higher distances between talker and listener. The bigger the car, the bigger the challenge.

The performance of ICC systems is measured by a variety of parameters, which can be found in the newly released Recommendation ITU-T P.1150 [21]. One very important question is how ICC systems can improve the Listening Effort in such situations. ABLE is ideally suited for this application. Based on a variety of subjective studies, a requirement for the enhancement provided by ICC systems was developed (see Figure 4).

From **Figure 7**, it can be seen that the required enhancement provided by ICC systems depends on the measured performance without ICC. The explanation for this requirement is rather easy. For lower Listening Effort values, the noise level in the car is generally high. However, the speech signal amplification introduced by the ICC system is limited. Even in noise, too high speech levels decrease the Listening Effort and even may damage the user's ear.

On the other hand, with low noises in a car there is often not much enhancement needed. Under quiet conditions, Listening Effort is already sufficiently good in smaller cars. In bigger cars and cabins with three rows, Listening Effort may already be quite poor, just due to the insufficient level of the direct speech. Therefore, any reinforcement introduced by the ICC

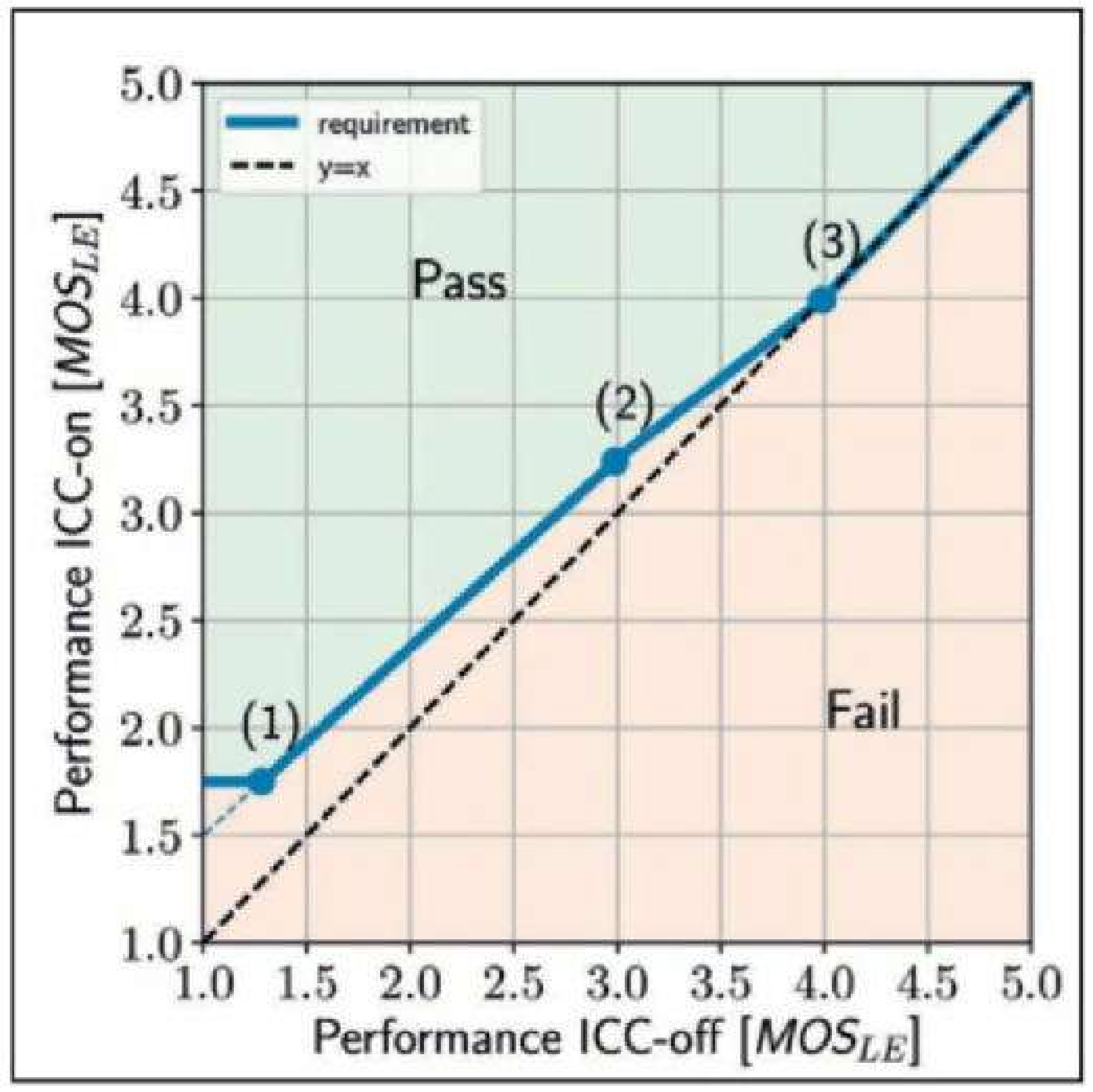


Figure 7: Minimum enhancement of Listening Effort by ICC systems

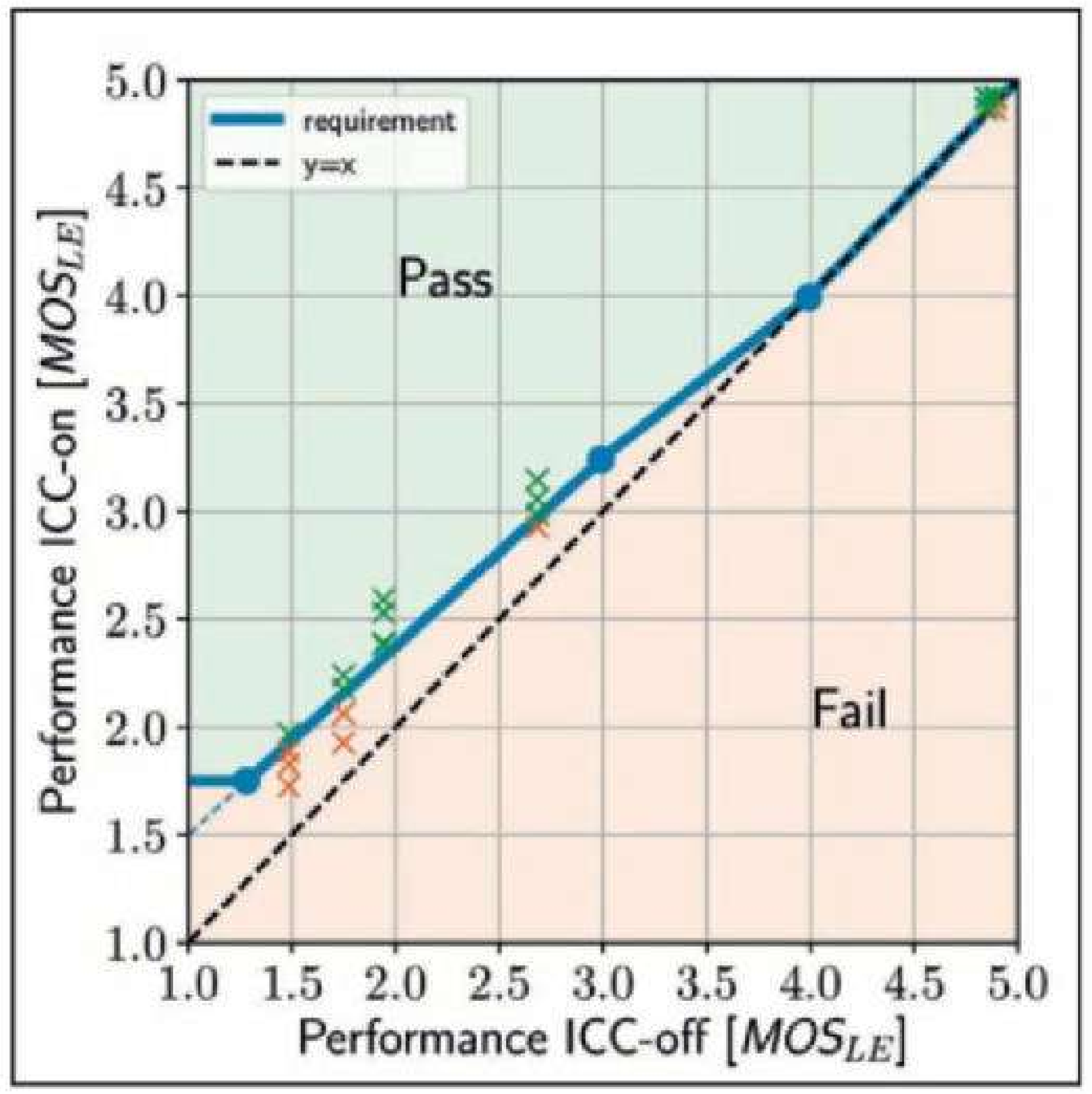


Figure 8: Examples for enhancement of Listening Effort by ICC systems

system would well improve the situation. This is the best operation range of ICC systems.

Figure 8 provides some examples how these requirements are applied on several ICC systems under different driving/noise conditions. It can be seen that it is technically possible to achieve such a degree of improvements. ABLE is a suitable methodology for predicting the Listening Effort improvement according to ITU-T P.1150 [21], but can also be used for optimization of ICC systems as well.

Conclusions and Outlook

With ABLE, a new perceptual-based method is available allowing the objective evaluation of Listening Effort for a variety of devices, for a variety of applications in various situations. Besides the qualification of devices, ABLE is perfectly suited for the performance optimization of (tele-) communication devices. Work for further enhancement and validation of this methodology is ongoing in ETSI Technical Committee STQ.

ABLE is commercially available. In conjunction with background noise and room simulation techniques [3], a laboratory setup including an artificial head, turntable (depending on application) or car environment as well as a fully automated test system, reliable measurements can be conducted in a laboratory environment. This setup enables developers and test labs to efficiently and reproducibly optimize and qualify numerous types of devices.

Resources

[1] Recommendation ITU-T P.863: Perceptual objective listening quality assessment, 03/2018.

[2] ETSI TS 103 281: Speech quality in the presence of background noise: Objective test methods for super-wideband and fullband terminals, V1.3.1, 05-2019.

[3] ETSI TS 103 224: A sound field reproduction method for terminal testing including a background noise database V1.4.1, 08-2019.

[4] ITU-T Recommendation P.800: Methods for subjective determination of transmission quality, 08/1996.

[5] Pusch, A. et al., "Binaural listening effort in noise and reverberation," in Fortschritte der Akustik DAGA 2018, volume 44, pp. 543–546, Berlin, Germany, 2018.

[6] Rennies, J. and Kidd, G., "Binaural listening effort in noise and reverberation," in Fortschritte der Akustik - DAGA 2018, volume 44, pp. 615– 616, Berlin, Germany, 2018.

[7] ETSI TS 103 558: Methods for objective assessment of listening effort, 11-2019.

[8] ITU-T: Handbook on Telephonometry, ITU, 1992, ISBN 92-61-04911-7.

[9] ITU-T: Practical Procedures for Subjective Testing, ITU, 2011.

[10] Sottek, R.: Modelle zur Signalverarbeitung im menschlichen Gehör, Ph.D. thesis, RWTH Aachen, 1993, Techn. Hochsch., Diss., 1993.

[11] Sottek, R.: "A Hearing Model Approach to Time-Varying Loudness," Acta Acustica united with Acustica, 102(4), p. 725–744, 2016, ISSN 16101928.

[12] Wan, R., Durlach, N. I., and Colburn, H. S.: "Application of a short-time version of the Equalization-Cancellation model to speech intelligibility experiments with speech maskers," The Journal of the Acoustical Society of America, 136(2), pp. 768–776, 2014, doi:10.1121/1. 4884767.

[13] Durlach, N. I.: "Binaural signal detection: Equalization and cancellation theory," Foundations of Modern Auditory Theory, Vol. 2, 1972.

[14] ISO/DIS 532-1, Acoustics – Reference zero for the calibration of audiometric equipment – Part 7: Reference threshold of hearing under freefield and diffuse-field listening conditions, International Organization for Standardization, 2005.

[15] Friedman, J. H.: "Multivariate adaptive regression splines," Ann. Statist, 1991.

[16] Reimes, J.: "Auditory Evaluation of Receive-Side Speech Enhancement Algorithms," in Fortschritte der Akustik - DAGA 2016, DEGA e.V., Berlin, 2016.

[17] Reimes, J. and Gierlich, H.W.: "Auditory Listening Effort Assessment for ICC systems," Contribution to ITU-T SG12/Q4, T17-SG12-C-0157, 2018, Geneva/CH.

[18] Recommendation ITU-T P.581: Use of head and torso simulator for hands-free and handset terminal testing, 02/2014

[19] Recommendation ITU-T P.56: Objective measurement of active speech level, 12/2011.

[20] Recommendation ITU-T P.1100: Narrowband hands-free communication in motor vehicles, 01/2019.

[21] Recommendation ITU-T P.1150: In-Car communication audio specification, 01/2020.

Loop Gain and Phase Margin Considerations in Applying Fully Differential Amplifiers to Active Filters

In this article, Michael Steffes discusses some of the design considerations when applying newer, low power, fully differential amplifiers (FDAs) to high dynamic range measurement mic interfaces to the delta-sigma analog-to-digital converters (ADCs).

By
Michael Steffes
(United States)

Modern high-speed fully differential amplifiers (FDAs) provide an attractive option interfacing from single-ended output measurement mics to differential input audio ADCs implementing an active filter. These recent FDAs have been very aggressively designed for low supply current with significant gain bandwidth product (GBP) to deliver extremely low harmonic distortion on a low power budget. The multiple feedback (MFB) active filter implementation is preferred using the FDA but loop gain (LG) and phase margin (PM) issues need a careful review for best end system performance.

The second-order low-pass MFB stage often used in an FDA interface will always have some internal noise gain (NG) profile over frequency. Rarely considered, the NG peaking can incrementally impair the output distortion if allowed to peak excessively. Subtracting that NG (in decibels) from the amplifier's open-loop (AOL) response curve will give the LG across frequency.

The LG acts to correct the amplifier's open-loop nonlinearity to deliver a much lower closed-loop harmonic distortion performance. This is one reason the 135 MHz GBP Texas Instruments THS4551 (see Reference 1) is being used for this apparently modest 50 kHz signal band requirement. Higher GBP normally means higher LG in those lower audio frequencies contributing significantly to those measured less than -130 dBc harmonic distortion terms below 50 kHz (refer to Figure 13 in Reference 1).

There are literally infinite combinations of RC values that will implement the same closed-loop response shape. The specific RC values used in the second-order part of a 65 kHz third-order low-pass filter in **Figure 1** are targeting:

- Nominal audio frequency gain of 1.7 V/V (or 4.61 dB)
- F_0 = 65 kHz (Butterworth filter poles are on a circle of equal F_0 in the s-plane)
- Q = 1

while also attempting to reduce the NG zero induced peaking effects. The NG will always have poles equal to the desired filter poles but will also have zeroes that depend partially (but not completely) on the desired signal response shape.

Regardless of where the RC values have come from, the MFB circuit will have a NG response, which is given by Equation 1 (or refer to Equation 19 in Reference 2). This has been written in terms of the desired filter response shape (where A_V is the inverting gain magnitude) showing the remaining $1/(R_2C_2)$ degree of freedom in the numerator (zero) expression. The NG magnitude over frequency will start out at the DC noise gain ($1+A_V$), show a midrange peaking, then approach 1 V/V (0 dB) at higher frequencies. The RC solutions shown in Figure 1 have used that extra degree of freedom to slightly reduce the NG peaking by pulling the lower NG zero up in frequency.

$$NG = \frac{s^2 + s\left(\frac{\omega_0}{Q} + \frac{1}{R_2C_2}\right) + (1+A_V)\omega_0^2}{s^2 + s\left(\frac{\omega_0}{Q}\right) + \omega_0^2} \qquad \text{Eq. 1}$$

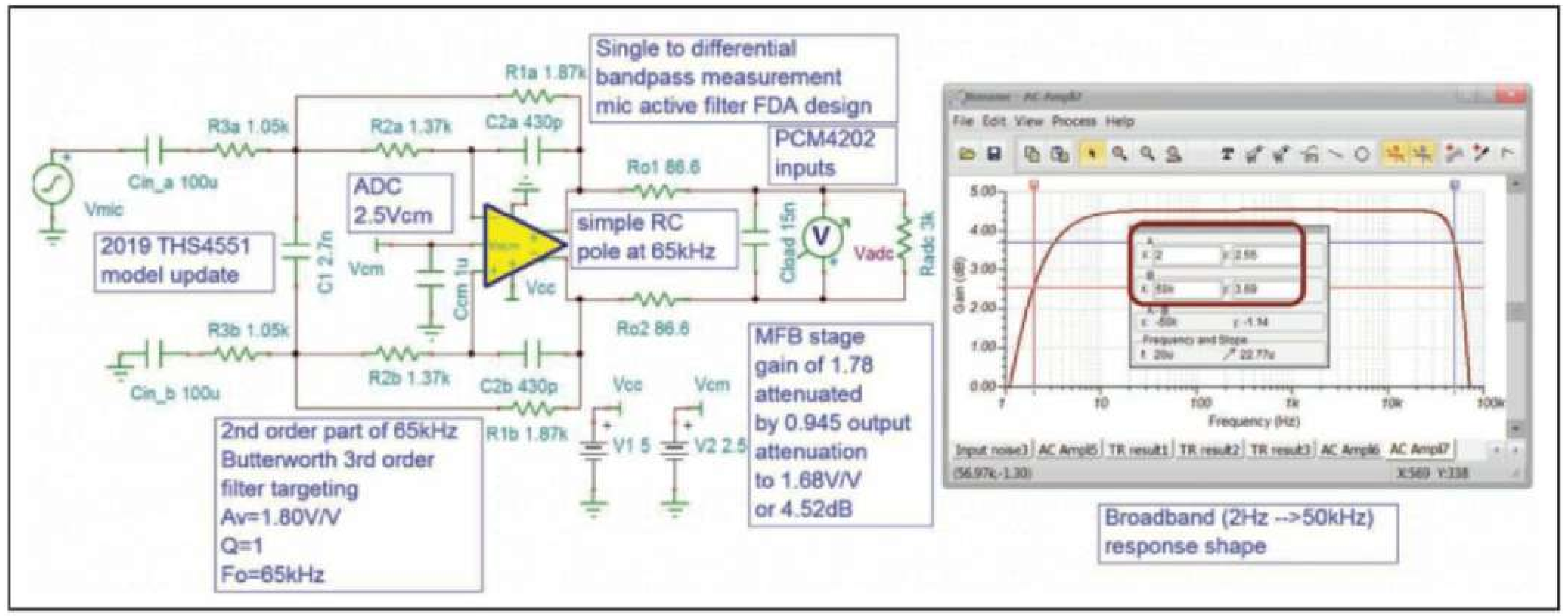


Figure 1: An AC coupled third-order low-pass filter to the PCM4202 dual audio ADC is shown here.

Generating the magnitude plot over frequency for Equation 1, along with the single-pole AOL response for the THS4551, will give the plot shown in **Figure 2**, where the difference between those curves is the LG. The minimum LG occurs in the region of the NG peak frequency reported here along with the LG at F_0.

The Final Design Element

The final design element considered here will be the loop gain phase margin. Every op-amp or FDA implementation will have some nominal phase margin. Significant modeling efforts have improved these recent models to more accurately predict the phase margin for any design. Starting from a reasonable estimate of the nominal PM, paths to increasing the phase margin (with minimal impact on the desired amplifiers operation) can be very useful.

While there are several valid approaches to simulating phase margin, the approach shown in **Figure 3** captures all the terms with a reasonably simple approach (refer to Reference 3). This is the closed-loop second-order MFB filter of Figure 1, reconfigured to produce a broadband LG simulation where the loop phase is also reported. The differential feedback voltage is being sensed into a dependent source and converted to single ended with an inversion to report the phase margin (instead of phase shift) directly.

The various elements to this PM simulation include:

- The FDA loop is broken using the two very large inductors. These provide a simulation trick to close the loop at DC to find a mid-scale operating point. Those open up on the first small frequency AC step while the two large capacitors short out passing the differential stimulus to the FDA inputs. That input signal generates the open-loop gain and phase response to the FDA output pins—including any open loop output impedance effects to the loading.
- It is important to keep the load in place as that sometimes can have a strong influence on PM.
- The rest of the MFB feedback RC elements are passing a feedback signal back to the differential feedback point that is sensed using a dependent source.
- The only added requirement is to duplicate at those differential inputs the specified input impedance. That 100 kΩ||1.2 pF differential input loading has a small effect here, but are sometimes critical to simulation accuracy.

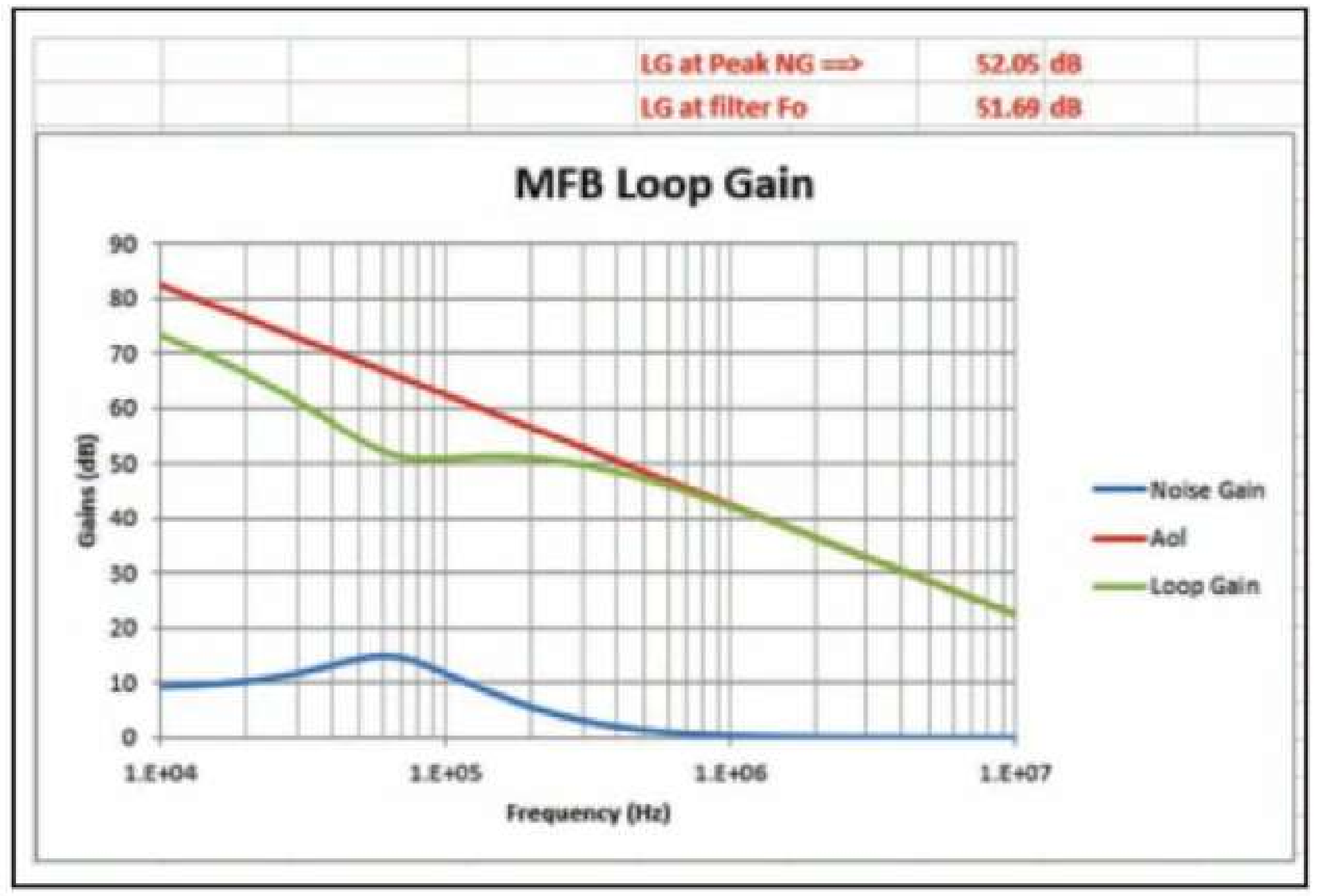


Figure 2: These are the amplifier's open-loop and noise gain plots for Figure 1, showing loop gain over frequency.

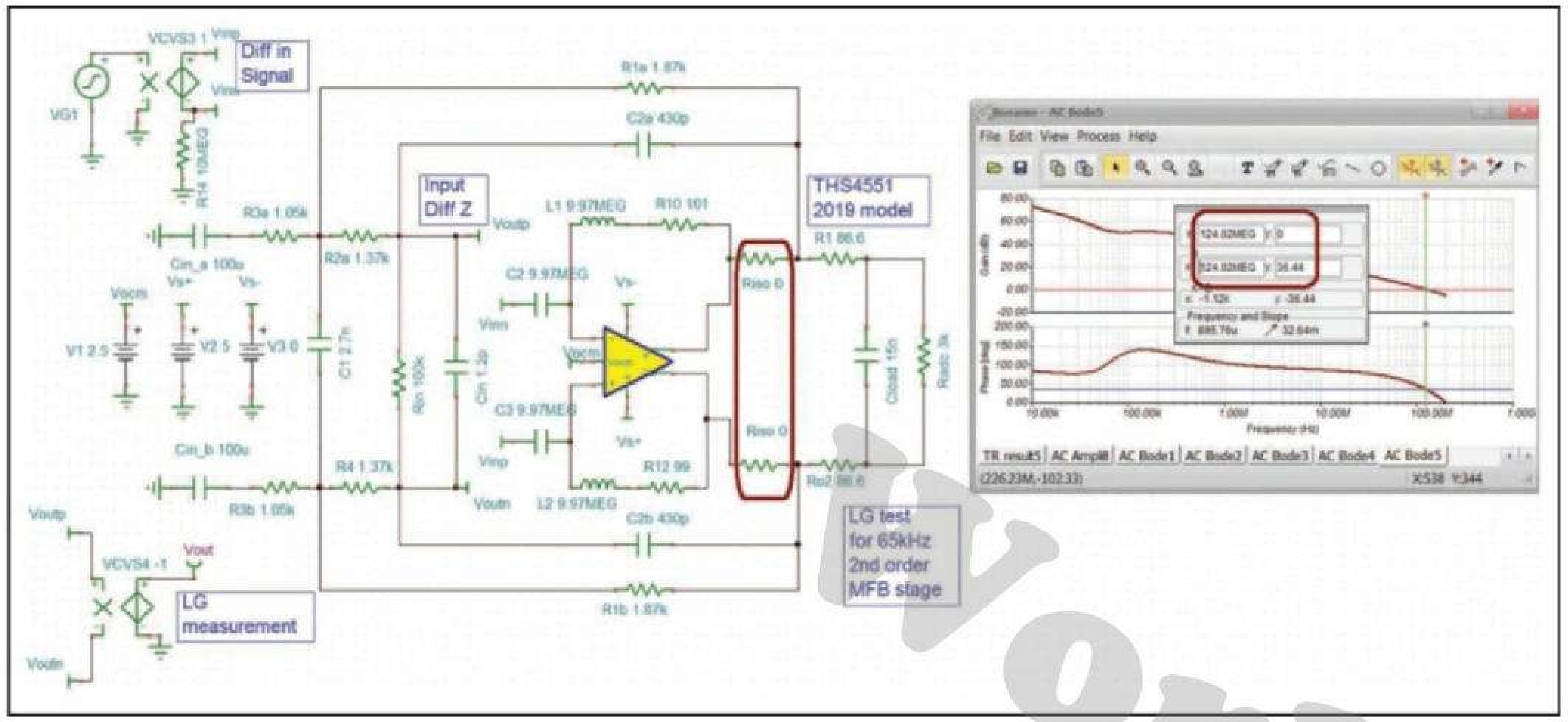


Figure 3: This is the loop gain phase margin simulation for the active filter of figure 1.

R_{ISO} Elements

Two new elements are added to this simulation over the original MFB circuit. Those inside the loop R_{ISO} resistors are set to zero here where the loop gain phase margin simulates to a comfortable 36°. The gain portion of the LG response shown in Figure 3 is simply a broadband version of the LG calculation shown in Figure 2, where the flat region around F_0 is apparent. Here, we are looking for the 0 dB crossover frequency and then checking the phase shift around the loop—inverted here to add 180° reporting phase margin directly.

That 124 MHz LG = 0 dB crossover frequency is much higher than many audio designers would consider—but critical to assess PM using this relatively high GBP solution to get the very low harmonic distortion using only 7 mW.

Those R_{ISO} elements have proven very useful in the wider range of MFB designs where the reactive open-loop output impedance (refer to Figure 68, Reference 1) interact with the MFB feedback capacitor (and/or a capacitive load) to yield low phase margin in some cases. This effect is very common for rail-to-rail output amplifiers but only recently modelled accurately.

For low PM designs, increasing R_{ISO} from 0 Ω has proven very effective with minimal performance interaction. For instance, re-running the LG simulation shown in Figure 3 with R_{ISO} = 20 Ω on each side increases the PM to 49°. This begs the question that perhaps 20 Ω on chip resistors at the output should have been included on THS4551 die. Aiming at a wide range of applications, it was decided this rail-to-rail output device could not afford the heavier load internal IR drop those resistors would introduce.

This concludes a review of some of the designer considerations applying these newer, low power, FDAs to high dynamic range measurement mic interfaces to the delta-sigma ADCs. While the THS4551 has been used as an example, ongoing vendor developments continue to improve your design options.

Reference

[1] "Low Noise, Precision, 150MHz, Fully Differential Amplifier," Texas Instruments THS4551 Datasheet, revised July 2017, http://www.ti.com/lit/ds/symlink/ths4551.pdf

[2] "Design Methodology for MFB Filters in ADC Interface Applications," Texas Instruments application note, Michael Steffes, February 2006, http://www.ti.com/lit/an/sboa114/sboa114.pdf

[3] "Stability Issues and Resolutions for High Speed Fully Differential Amplifiers (FDAs), Insight #8," Planet Analog, Michael Steffes, May 1, 2019, https://www.planetanalog.com/stability-issues-and-resolutions-for-high-speed-fully-differential-amplifiers-fdas-insight-8

Sound Cards for Data Acquisition in Audio

The Linear Audio Autoranger Mk II Interface

Stuart Yaniger explores new possibilities in audio measurement systems, expanding on the basics of using sound cards combined with new accessible-to-all software tools, and a recently improved measurement interface option.

Photo 1: The Autoranger Mk II is a flexible measurement interface with a variety of I/O options.

By
Stuart Yaniger
(United States)

Photos courtesy of Cynthia Wenslow

In the years since my series of articles on the basics of using sound cards (in the broadest sense—actual cards, external ADC/DACs, and pro-sumer interfaces) as the core to an audio measurement system, the use of these tools has mushroomed among audio hobbyists and small professional firms that can't justify the expense of a full-bore audio measurement system from leading test and measurement manufacturers. Several things have, in my view, driven this expansion, including:

- the increasing display of sophisticated measurements in magazine and blog equipment reviews
- the increasing familiarity of audiophiles with computer processing of audio signals
- the increasing use of room measurement and correction software built into audio components
- the increasing availability of free and easy-to-use measurement software for free-standing measurement

Regarding the last point, the most popular choice is typically Room Equalization Wizard (REW), a freeware package that has taken the audio enthusiast world by storm. As its name implies, REW is mostly targeted toward room measurement and correction, but under that exterior is a very powerful and versatile measurement package that audio enthusiasts have been using for frequency response and distortion measurements. It also has the nice feature of being able to automatically generate filter parameters for room correction after doing acoustic measurements. The progress in this software and its applications since my first article on software options has been impressive, and the support that REW gives users through several Internet audio forums is noteworthy.

For perhaps the majority of users and measurements needs, a simple sound card and REW will do the job nicely. This is especially true with the latest generation of USB microphones that do not require phantom power. But for more sophisticated measurements, more sophisticated tools are needed, and there's been a lot of progress in this area. In this article, I'm exploring the Linear Audio Autoranger Mk II, a hardware product that could improve the measurement capabilities in small labs. In a forthcoming article, I will discuss a software product that will also improve measurement capabilities, the latest version of Virtins Multi-Instrument.

The Autoranger Mk II

Sound cards have several limitations when you want to measure voltages. Most notably, they have a limited range of input voltages so that if, for example, you want to measure the output of a power amp, you will quickly exceed the maximum input voltage limitation. This requires rigging up an external attenuator, and for best performance, you'll need several of them to translate the amplifier's output through a sweep to full power to the roughly 1 V to 2 V that is in the optimum performance range for most sound cards. And of course, sound card systems need to be calibrated for each attenuator used, as well as having to be redone any time a volume control is changed.

Photo 2: The interface comprises three circuit boards, which separately accommodate display/logic, analog signal circuits, and the SilentSwitcher power supply.

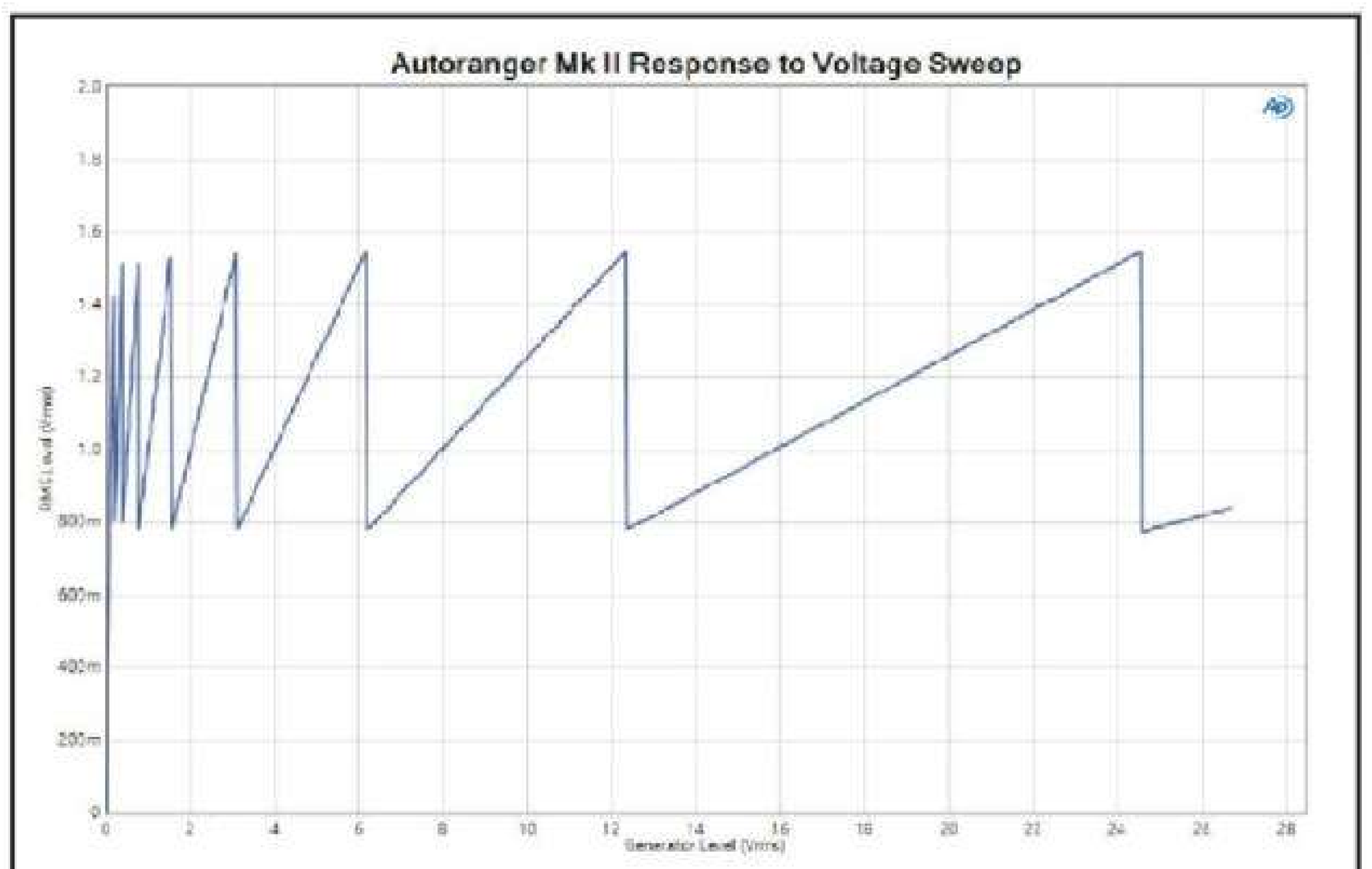


Figure 1: Sweeping the input voltage while monitoring the output shows the autoranging performance, holding the output to a level suitable for most soundcards.

About the Author

Stuart Yaniger has been designing and building audio equipment for nearly half a century, and currently works as an R&D director for a polymer manufacturing company. His professional research interests have spanned theoretical physics, electronics, chemistry, spectroscopy, aerospace, biology, and sensory science. One day, he will figure out what he would like to be when he grows up.

Another potential issue is the need for both balanced and single-ended inputs and outputs for the measurement and the stimulus, respectively. While this is generally available on pro-sumer recording interfaces, it's not universal, and most USB cards intended for gaming or home theater and PCIe solutions are single ended only. Even rarer is any sound card or recording interface that allows inputs to be floated, either to break up ground loops or to measure signals in the presence of DC.

This is something I discussed in Part 4 of my earlier series on sound card measurement (*audioXpress*, September 2015). To overcome these limitations, I ended up using the excellent measurement interface box from my buddy Pete Millett (see Resources); this product is a measurement interface box that goes between the device under test (DUT) and the sound card.

As described in my article and the write-up on Millett's website, the Millett box provides manually switched attenuation of 1, 10, 100, and 1000 (or 0 dB, -20 dB, -40 dB, and -60 dB). It also has balanced and unbalanced inputs and outputs that can be floated, with as much as a 200 VDC offset. Additionally, its 100 kΩ input impedance is about 5 to 10 times higher than that of typical sound card inputs, so it will tend to have less effect on a high impedance circuit measurement. Conveniently, the Millett box also has a built in AC voltmeter, which can be useful for both calibration and basic measurement.

A new sound card measurement interface option for audio measurement first appeared about two years ago, the Autoranger, designed and sold by *audioXpress'* Technical Editor Jan Didden through his Linear Audio website. This interface performs a similar task as the Millett box, but is significantly more versatile. Like the Millett box, the Autoranger was available as a kit or fully assembled. I got an early version of the Autoranger and gave Didden feedback on usability and performance.

In response to the feedback from many other early adopters, the design was recently updated to the Autoranger Mk II, with lower distortion and noise than the original unit. I installed the new Mk II boards for evaluation.

As its name implies, the Autoranger Mk II will automatically adjust input gain to scale the input signal to provide a nominally 0.4 V or 1 V signal (user selectable) to the sound card. In the case of the Autoranger Mk II, the voltage gain varies from +18 dB to -42 dB in 6 dB steps, which provides great flexibility regarding getting the measured signal close to the optimum range for the sound card. The user also has the option of switching the gain range manually.

The Autoranger Mk II is housed in a rather generic powder-coated steel case, with silk-screened legends. There is a CE designation on the back. Also on its back is a USB plug for power, which can either be provided by a computer, a plug-in USB charger, or a 5 V portable battery power bank. The allowable power input voltage range is 3 V to 10 V. No matter which power option is used, the signal circuitry is isolated by the internal switching power supply, Didden's SilentSwitcher, which is also available as a standalone module for other projects (see Resources). The very low noise and isolation provided by the SilentSwitcher is integral to the Autoranger Mk II's noise and distortion performance.

The front panel has two input jacks, a BNC for single-ended sources, and a combination XLR/TRS for balanced sources (see **Photo 1**). The user can select between them with a button push. Helpfully, the button for the chosen input illuminates to indicate that it's active. Outputs are separate for single-ended or balanced sources, and are each a TRS jack.

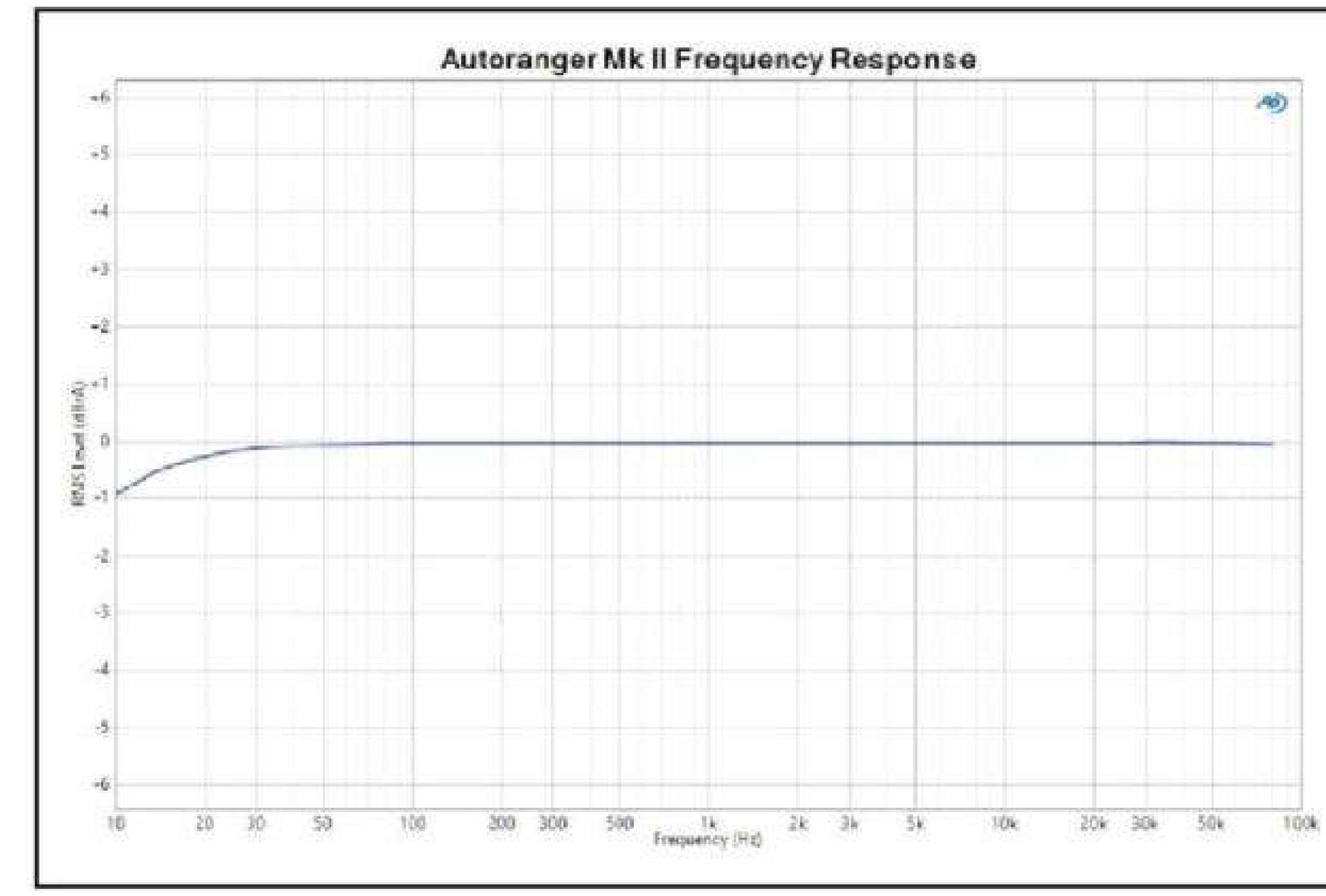


Figure 2: The Autoranger Mk II's frequency response is absolutely flat to frequencies beyond the capability of the AP's signal generator.

The large LED display indicates the measured input RMS voltage, the output voltage to the sound card, the gain, and input configuration (balanced or single-ended). By moving internal jumpers, the user can select the display color (red, green, or blue) and adjust brightness. Contrast is adjusted via an internal trimmer. The voltage measurement capability is particularly useful for calibrating the sound card in whatever measurement software is being used.

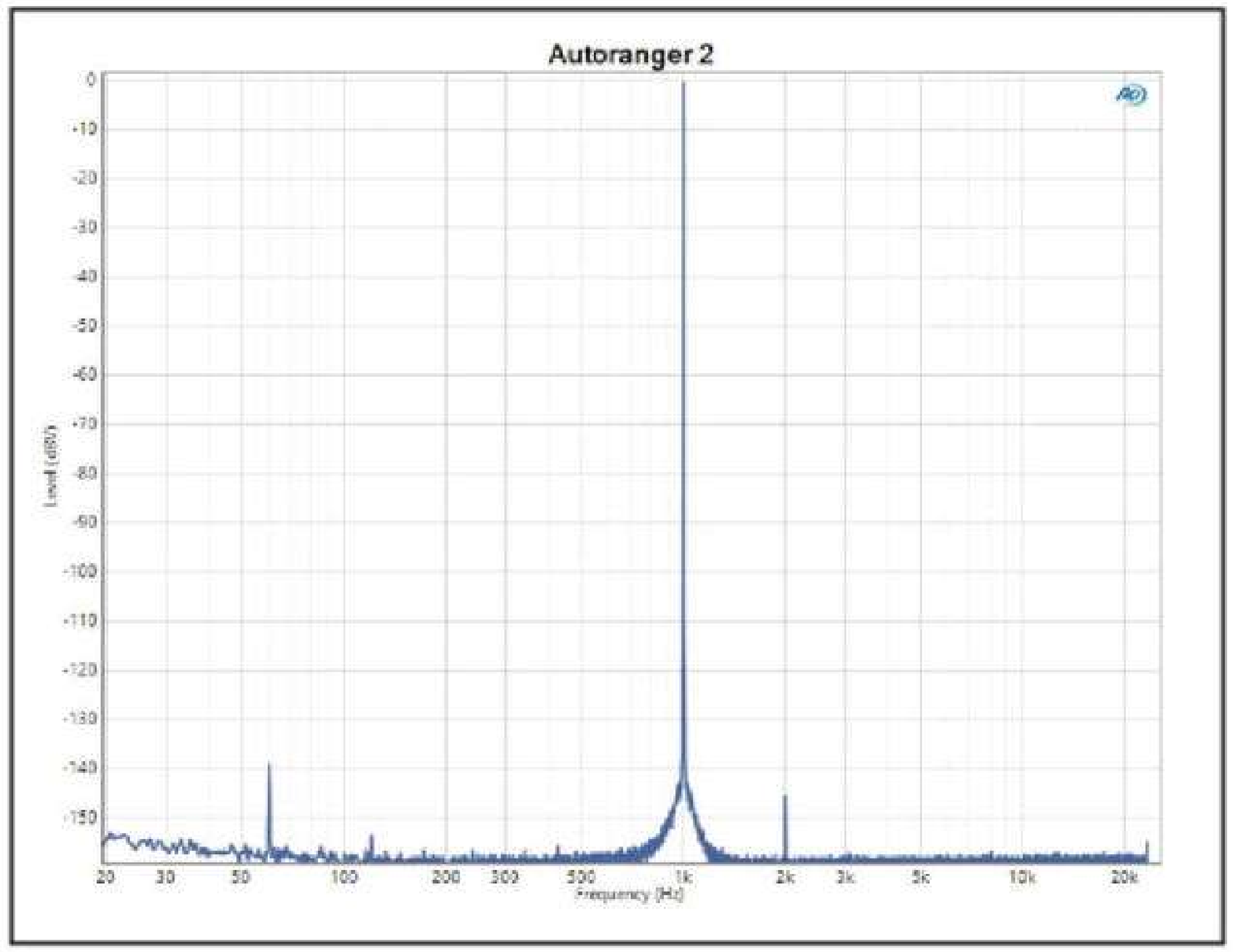


Figure 3: The Autoranger Mk II shows astonishingly low distortion and noise, requiring a state-of-the-art audio analyzer to characterize.

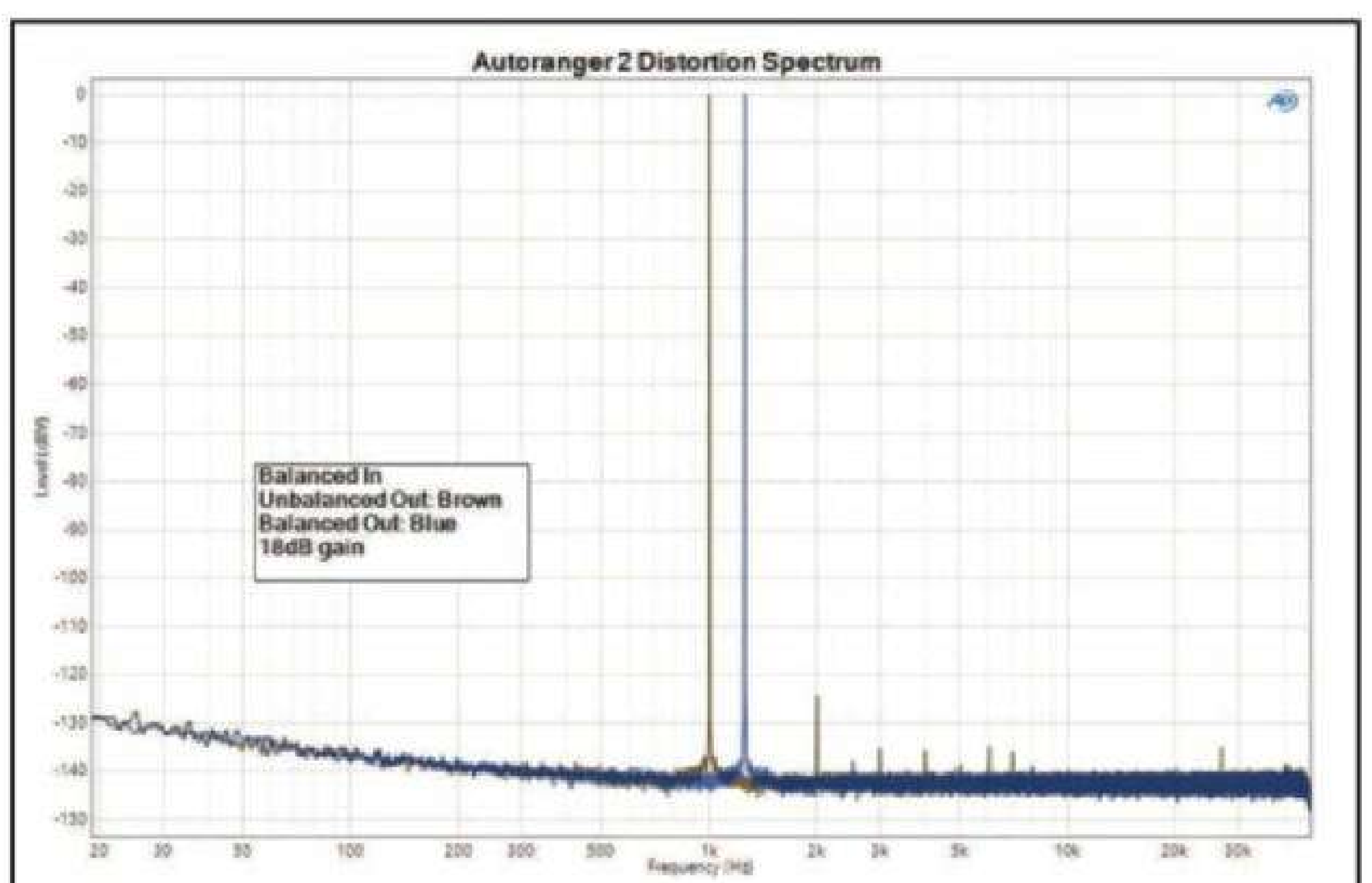


Figure 4: Switching between balanced and single-ended modes, the only significant difference in distortion is a small amount of added second harmonic in the latter mode.

Internally, three PCBs contain the SilentSwitcher power supply, the display electronics and logic, and the analog signal circuits (see **Photo 2**). No doubt someone could put all of this in a surface-mount package to reduce the interface size and make a slicker package, but for its intended bench use and for the convenience of people who want to build this from a kit, the use of larger through-hole components and a standard metal case doesn't present a huge disadvantage.

To switch between Automatic and Manual gain adjustment, the user goes through a multiple button press sequence—first, the Hold button freezes the gain range at whatever value it is at the time of that button push, then the gain range can be manually incremented up and down in 6 dB steps by pressing the Bal or SE buttons, respectively. Helpfully, there are two LEDs on the front panel to indicate under-range and over-range, so you can get a visual indication of whether the manually chosen range is appropriate for your measurement.

One important limitation: the Autoranger Mk II's input impedance is either 10 kΩ or 20 kΩ (which is specified when the unit is ordered). That will present no difficulty for modern sources (e.g., DACs, preamps, power amps) but precludes the use of the Autoranger Mk II in measuring internal voltages at high impedance points or measuring older tube equipment, which would be excessively loaded down by the Autoranger Mk II. As far as I can see, there's no disadvantage to using the 20 kΩ impedance version, so that's what I'd recommend.

According to Didden, raising the input impedance makes it much more difficult to design the attenuator for flat response and low noise. The world is still awaiting my dream, an interface with a 1M input impedance that can be used with oscilloscope probes. Besides the lower loading from the higher impedance, the ability to use scope probes allows a 10x option, increasing the input voltage range and drastically reducing the effect of the probe's cable capacitance on the circuit being measured. This would especially be a blessing for those of us who work with vacuum tube and field-effect transistor (FET) circuits.

Measurements

The test setup here varied between an Audio Precision APx525 and an APx555 B Series, with the latter used for distortion measurements. I first connected the Autoranger Mk II via balanced in and out, with the nominal output setting at 1 V. The AP was then set to step a 1 kHz sine wave from 1 mV to its maximum output voltage (26.6 V) while monitoring the Autoranger Mk II's output.

This measurement is shown in **Figure 1**, and it can be seen that the Autoranger Mk II switches ranges when its output voltage reaches about 1.55 V, with the next range being 6 dB (or 2x) lower at 0.775 V. At the lower voltage ranges, the graph shows a lower transition voltage, but this is a measurement artifact resulting from the finite voltage step size in the applied voltage sweep. Other than this measurement anomaly, the transitions are all quite consistent. I note, however, that these threshold voltages for switching are significantly higher than the numbers given in the User Manual, though for most sound cards, this should not be a concern. For the nominal output setting of 0.4 V, the outputs ranged from 350 mV to 660 mV, again higher than the numbers given in the User Manual.

It would be nice to see the documentation updated for the Autoranger Mk II version. One thing I noticed was that if I made the sweep slow enough, I could get the Autoranger Mk II to act a bit chaotically near the switching transition. Admittedly, this took some effort and is not indicative of normal use, but perhaps a bit more hysteresis in the transition could eliminate this behavior entirely.

The frequency response of the Autoranger Mk II in-balanced out mode is shown in **Figure 2**. On the low end, it starts rolling off at 20 Hz, where it's about -0.3 dB. On the high end, it is absolutely flat to the 80 kHz limit of the AP's sine wave generator. The claimed high-frequency limit is 600 kHz, and I see no reason to doubt that.

Figure 3 shows the 1 kHz distortion at unity gain. The only distortion component visible is the second harmonic at -145 dB, which is less than 0.000006%! There's a very small amount of 60 Hz pickup, which may have been from my cabling, but at -140 dBV (125 nV), I didn't spend much effort in tracing it down. This is superb performance which challenges even the APx555 B Series, and the fact that any of these distortion products can be seen suggests that the noise of this unit is not the measurement limitation.

Switching to 18 dB gain, I compared the spectra from both balanced and single-ended outputs (see **Figure 4**). For these measurements, I displaced the frequencies slightly to make differences on the graph more visible but not otherwise affect the measurement.

The only significant differences are the visible even-order harmonics in the single-ended outputs, most notably the second harmonic at -124 dB (0.00006%). This is unsurprising since symmetric circuits tend to cancel even-order harmonics, and the single-ended output is not symmetric by definition. Nonetheless, this is a superbly low distortion result, and is better than most signal sources and analyzers. There's no hum visible in either measurement, so I inadvertently did a better job of cabling than during the test run of Figure 3.

Increasing the frequency to 5 kHz gives the results shown in **Figure 5**. The second harmonic for the balanced out has risen a bit, but the other distortion components seem to be basically unchanged. There's a spurious 1 kHz pickup in the single-ended results, which I see from time to time in low-level measurements—I believe that it's a pickup from a

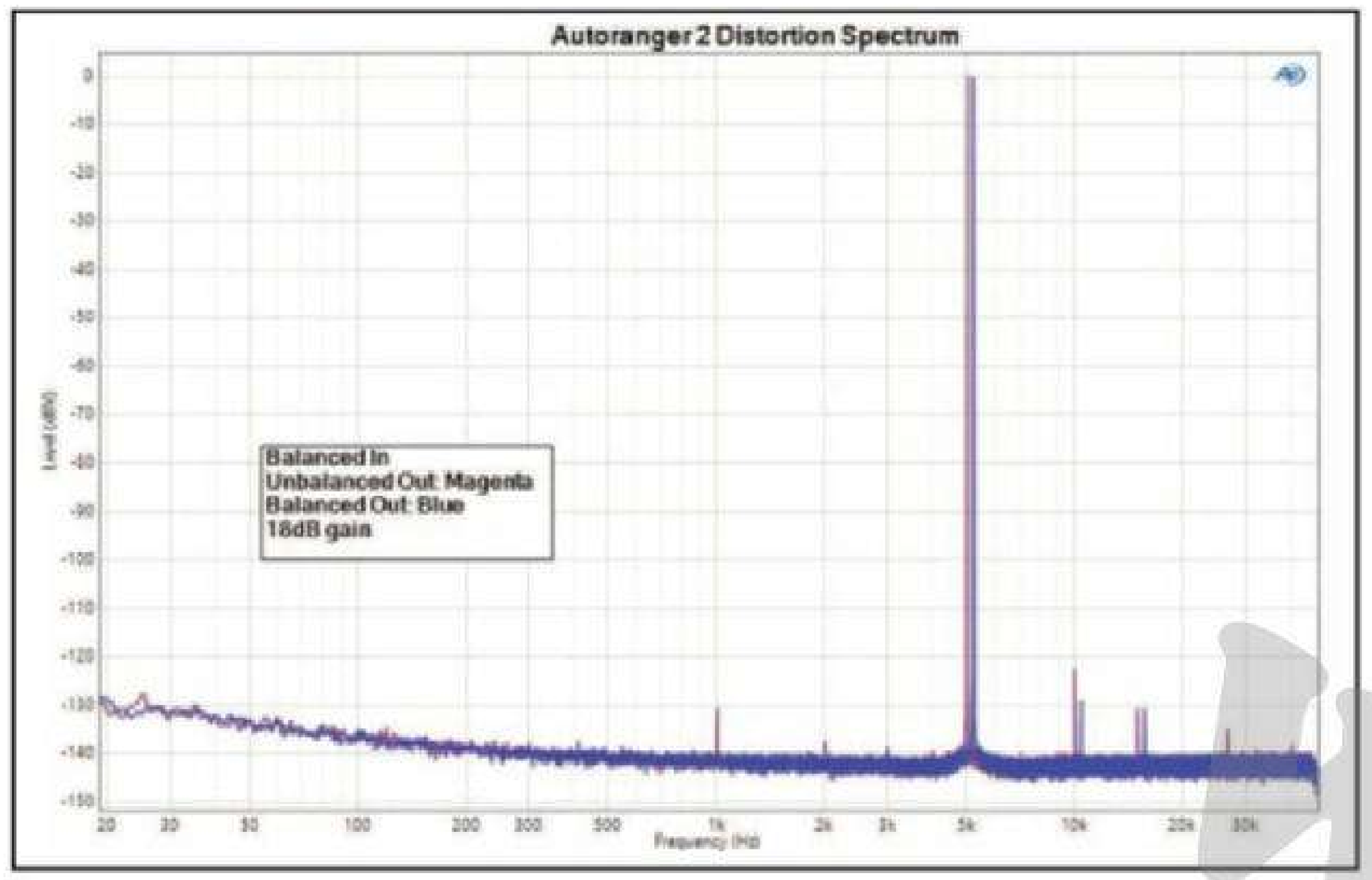


Figure 5: Distortion at higher frequencies (in this case, 5 kHz) remains impressively low.

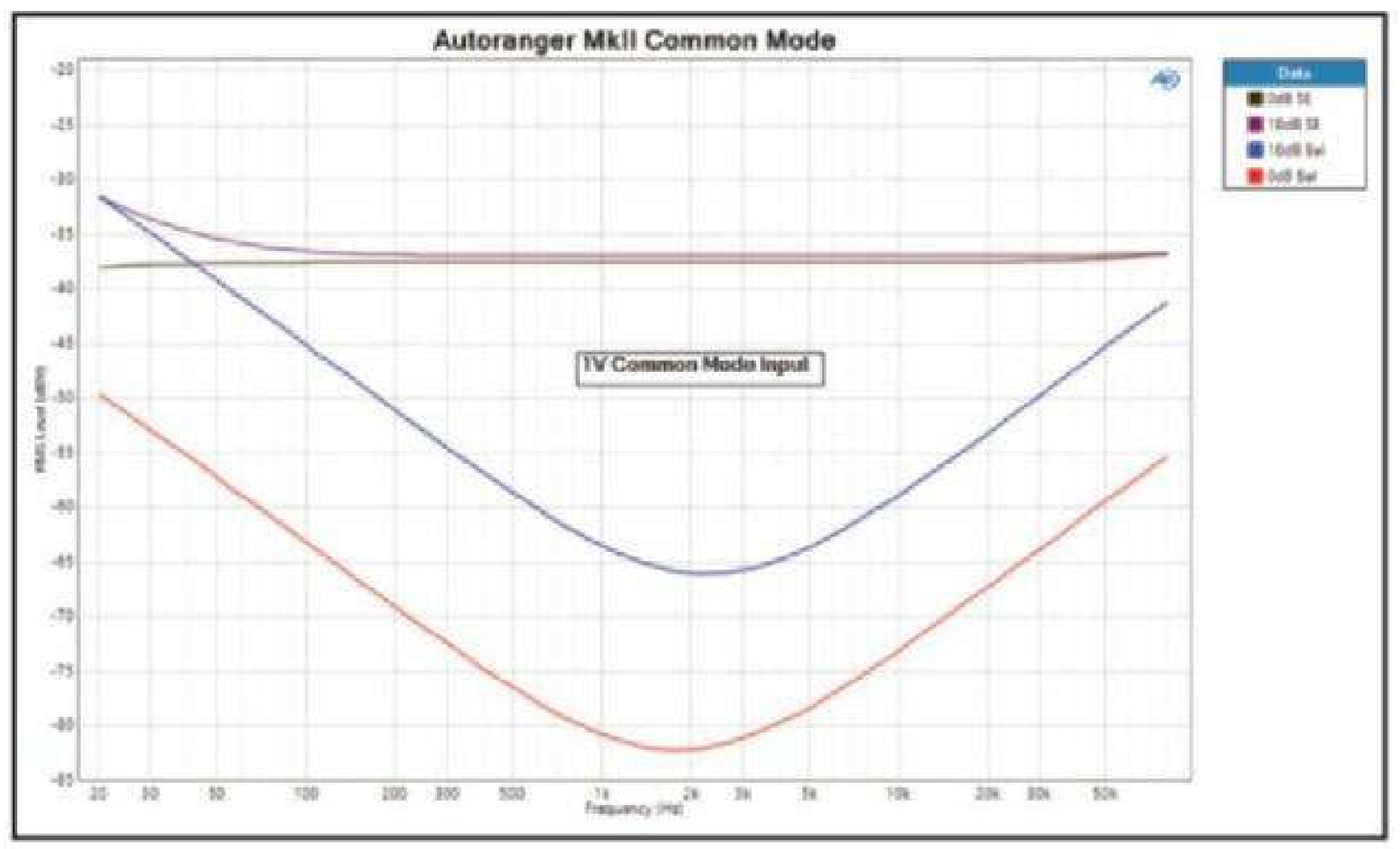


Figure 6: The response of the Autoranger Mk II's balanced input to a 1 V common mode signal shows moderate (and frequency dependent) common mode rejection.

Resources

J. Didden, "The L|A Autoranger," https://linearaudio.nl/la-autoranger and https://www.youtube.com/watch?v=K4TRhEjaIcs

J. Didden, "The SilentSwitcher," https://linearaudio.nl/silentswitcher

P. Millett, "Sound Card Interface" http://www.pmillett.com/ATEST.htm

S. Yaniger, "Sound Cards for Data Acquisition in Audio, Part 4" *audioXpress*, September 2015, https://audioxpress.com/article/practical-test-measurement-sound-cards-for-data-acquisition-in-audio-measurements-part-4

Sources

REW software
Room EQ Wizard | https://www.roomeqwizard.com

Virtins Multi Instrument software
Virtins Technology | https://virtins.com/multi-instrument.shtml

nearby USB cable, but it is a nice illustration of the inherent common-mode noise rejection of balanced circuits.

Speaking of which, **Figure 6** shows the response to a 1 V common mode input signal with the Autoranger Mk II set to 0 dB and 18 dB of gain. This was measured both at the single-ended and balanced outputs. The balanced output curves show a minimum of common mode response (maximum of common mode rejection) at about 1.8 kHz, with common mode rejection (CMR) deteriorating above and below that frequency.

Not surprisingly, they differ by about 18 dB through most of their range, indicating that the limitation on CMR is in the input stage. Minima for the output at 0 and 18 dB gain are at -82 dB and -67 dB, respectively. The common mode responses measured at the single ended outputs are higher (lower CMR) at about -38 dB, and flatter with frequency, which may be a clue about the source of the CMR limits for that topology.

Concluding Thoughts

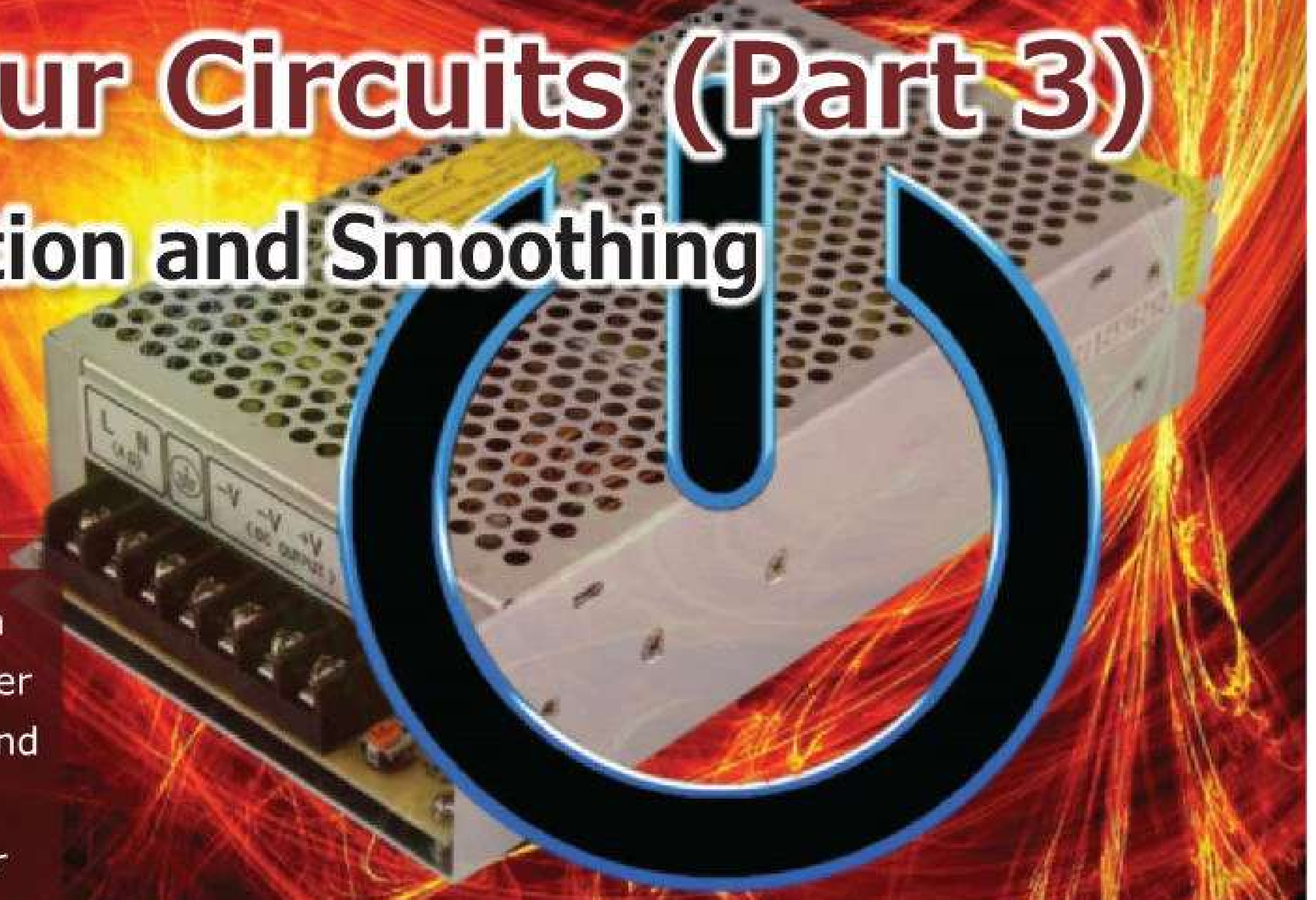
It's been interesting to see the evolution of this product over the past few years. The distortion and noise results show that a lot of attention was paid to the details of the circuit and implementation, and it is unlikely that the Autoranger Mk II would be the bottleneck in any sound card-based measurement system.

On the contrary, this product enables users with high-quality sound cards or recording interfaces to perform measurements to the level that is usually reserved for extremely expensive professional-grade audio analyzers. Indeed, its performance required the use of the newest and most elaborate Audio Precision system to characterize, and it greatly exceeded the performance of the APx525 I use day-to-day in my lab.

The Autoranger Mk II is also easy and intuitive to use, and reconfiguring it is rapid and convenient. The only remaining limitation is still input impedance—the Autoranger Mk II is more than suitable for measuring equipment, but getting measurements at the circuit level remains an unconquered challenge. Although the Autoranger Mk II gets closer to the ideal measurement interface than any previous products, the world still awaits an interface with the same input characteristics as oscilloscopes and the ability to use 10x probes.

In an upcoming article, I'll combine the Autoranger Mk II with a high-performance recording interface and the latest version of Virtins Multi Instrument to demonstrate near state-of-the-art measurement capability and flexibility.

Audio Electronics



Powering Your Circuits (Part 3)

AC Power Rectification and Smoothing

By
Ed Simon
(United States)

Ed Simon has provided *audioXpress* with an article series that details ways to power circuits. After discussing battery power and AC power supplies, this article expands the discussion to transformers and power rectification.

There are two common solutions to the unbalanced load problem. The first solution is to use a center-tapped winding on the secondary of the transformer. This enables the use of a dual-diode vacuum tube to reduce ripple, improve the current capability, and make better use of the transformer. Today, we mostly use semiconductor diodes as they handle reasonably high voltages and currents and cost much less.

What happens in the split winding is that when the AC is positive at the top diode with reference to the center tap, it is negative to the bottom diode (see **Figure 1**). As a result, the top diode will charge the capacitor. On the other half of the cycle, when the top diode is now not passing current due to the reversal of the AC cycle, the bottom diode is now conducting. That enables the transformer to work on both halves of the cycle. It can now pass more power before saturating.

The second advantage is that the capacitor now gets current twice each cycle. So the ripple would be halved with the same size capacitor. Also, since each diode now provides half of the final current, they can be sized smaller or it will have slightly less voltage drop.

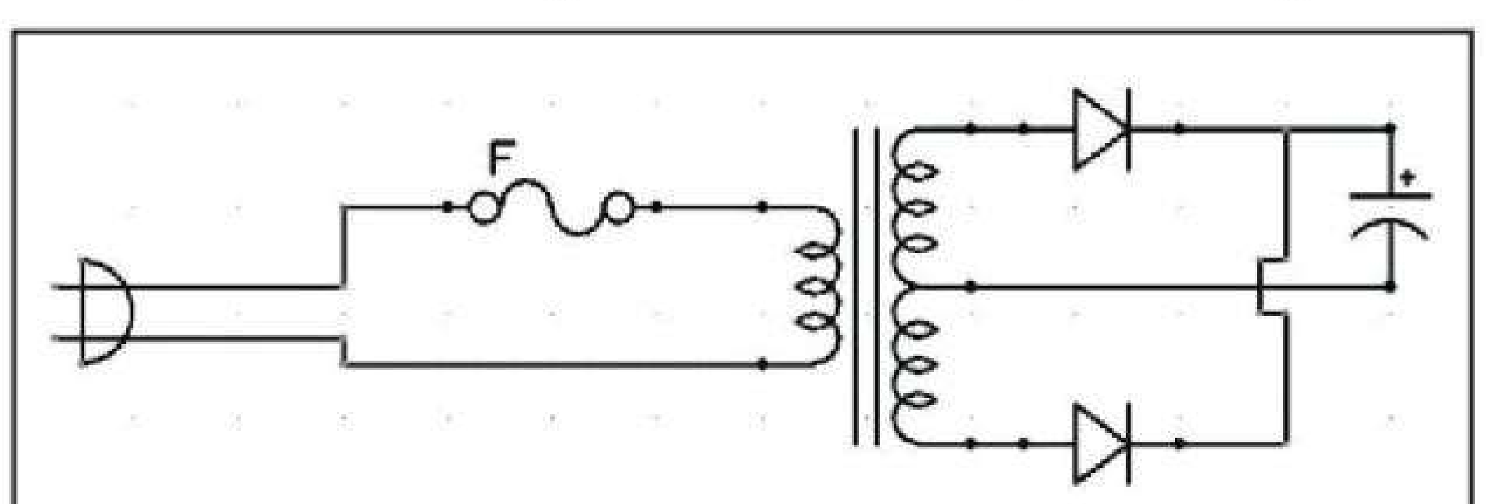


Figure 1: When the AC is positive at the top diode with reference to the center tap, it is negative to the bottom diode.

In this circuit, if we use a 6.3 V rated transformer, we may have an unloaded circuit so that the 6.3 V would be 20% higher. If we look, we would see that the primary voltage may be rated for as little as 110 V. A typical AC power line could be 132 V and still follow power company guidelines. So, we should design for an AC output voltage of:

$$6.3 \times 1.2 \times 132 / 110 = 9.072 \text{ VAC}$$

Allowing for no diode drop would give us $9.072 \times 1.414 = 12.83$ VDC. That would be the most we would expect. Now if we used a single diode to rectify the AC voltage, the capacitor would try to charge all the way to 12.83 DCVs (ignoring losses.) When the AC voltage swings to the minimum on the other half cycle, we would have -12.83 V. So the voltage across the diode would be 25.66 V. The next highest standard Peak Inverse Voltage (PIV) rating for diodes is 50 V. Most common values are 25 V, 50 V, 100 V, 200 V, 400 V, 600 V, 800 V, and 1,000 V. There are special low-loss Schotky diodes that have lower forward voltage losses, but they also have lower available P.I.V. ratings.

If you were building a large power amplifier, the difference in secondary voltage would be more of an issue as you normally want as much DC voltage as the output transistors can safely handle, but not enough to blow them out. Some people like to oversize the power transformer in terms of current capacity. That often results in tighter regulation and always permits faster capacitor charging times.

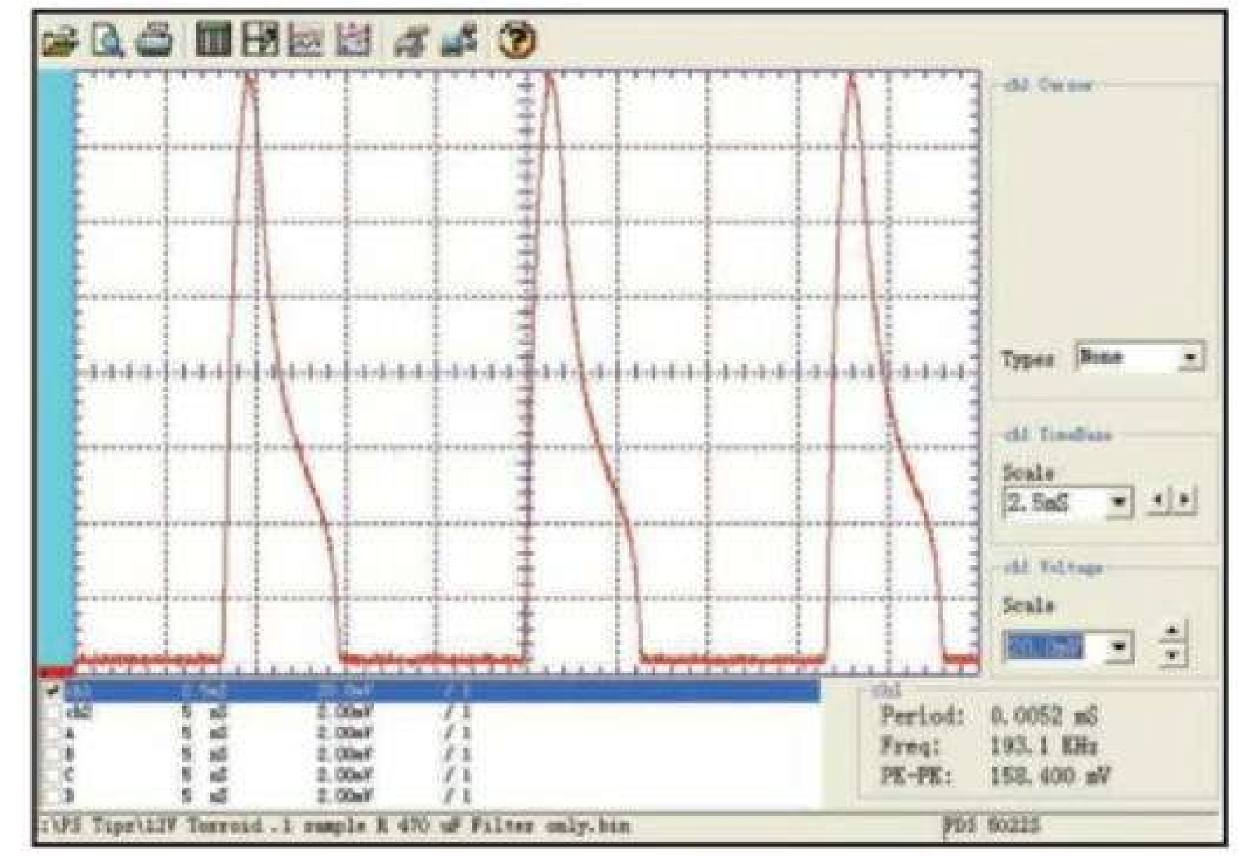


Figure 2: A toroidal transformer is shown with two primary windings and two secondaries.

The Charging Current at Extremes

Let us look at what happens to the charging current when we go to extremes. **Figure 2** shows a toroidal transformer with two primary windings and two secondaries. If you are using it on 120 V, you place the primary windings in parallel. For 220 V to 240 V, you can place them in series. The two secondaries can be used in parallel for higher current or in series as used here to provide a center tap, which can be used with just two diodes to provide DC or with four diodes to provide positive and negative rails.

The transformer is rated for 24 VAC RMS at 12.5 A. It is filtered by 470 µF and loaded by 50 Ω to provide around 1 V RMS of ripple. The current monitor is a 0.1 Ω non-inductive 10 W resistor in the center tap circuit leg.

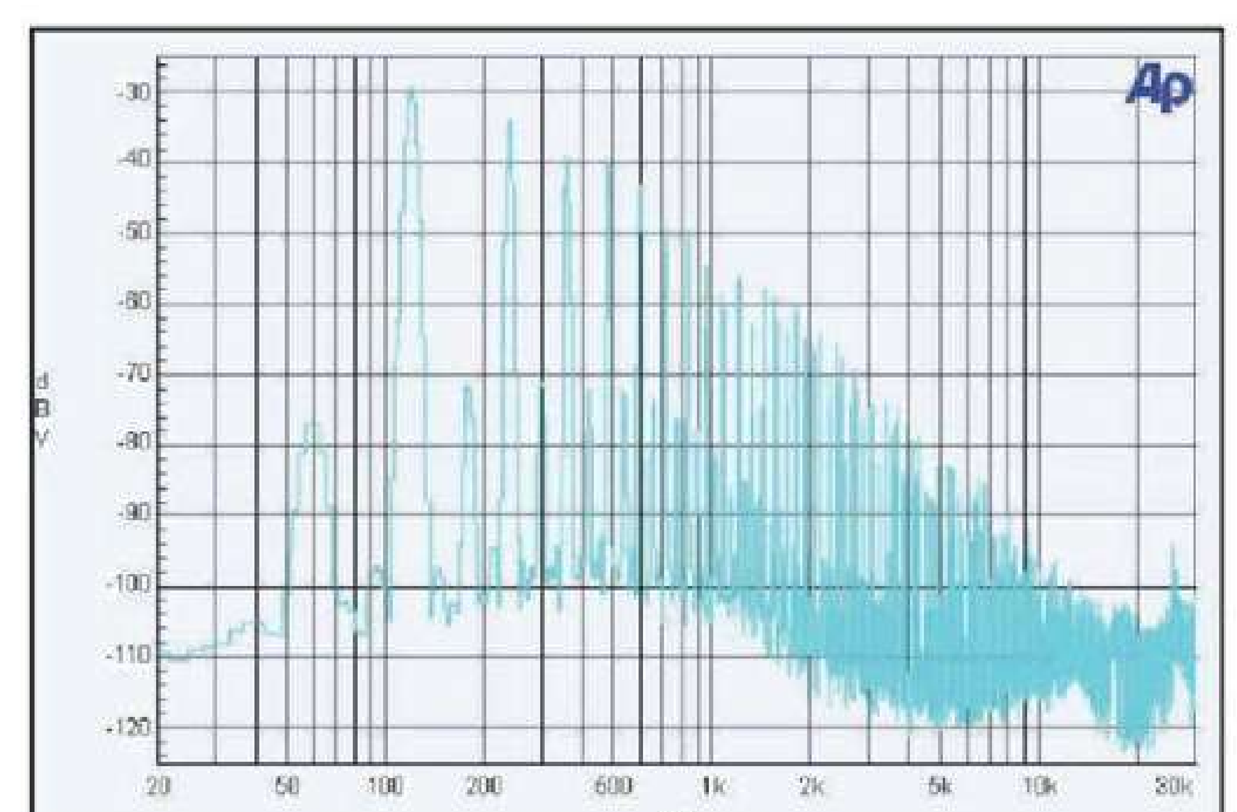


Figure 3: We could actually write the equation that would give us this waveform, but it is easier to cheat and measure it.

There is no current drawn for most of the AC power cycle time. Only when the transformer's unloaded output voltage rises above the filter capacitor's stored voltage does current flow. As this is a larger transformer, the regulation is fairly tight so the flow starts and quickly charges the small filter capacitor very close to the peak of the sine wave.

After the voltage peak, the current begins to drop—enabling the "float" to keep providing the filter current until it the input voltage drops even lower, however there is some inductance in every transformer and the resulting stored energy has to go somewhere so it extends the charging time until the current finally shuts off.

Now Jean-Baptiste Joseph Fourier postulated that all periodic waveforms can be made up from a sum of sine waves. This work has been greatly expanded, but is still all classed under the umbrella of Fourier Theory. We could actually write the equation that would give us this waveform, but it is easier to cheat and measure it (see **Figure 3**).

Before, we were looking at the AC voltage distortion in the transformer's output voltage and the current induced into a chassis. Here we are looking at current flow in the output of an oversized transformer. As you can see from the Fast Fourier Transform (FFT) measurement of the current, there are lots of harmonics. Since the reference RMS value is 1 V across a 0.1 Ω sense resistor, we can see there is still about 3 µV or 30 µA of noise at 20,000 Hz! All of this from what started as 60 Hz AC power. We now see noise from both the transformer and the result of rectification. If we look at the output voltage of the power supply with our FFT analyzer, we see what at first looks just like the current plot (see **Figure 4**).

If you look closely, you can see the 120 Hz ripple is about 0 dBV or about 1 V RMS. The ripple around 3,000 Hz (an area of particular hearing acuity) is still around 100 µV. With an oscilloscope, the noise or ripple on the output of the power supply looks like the image shown in **Figure 5**.

Many loads have a rated power supply rejection ratio (PSRR) that will reduce the effect of this noise on the final signal. It is important to note that an IC op-amp may have a PSSR rating of

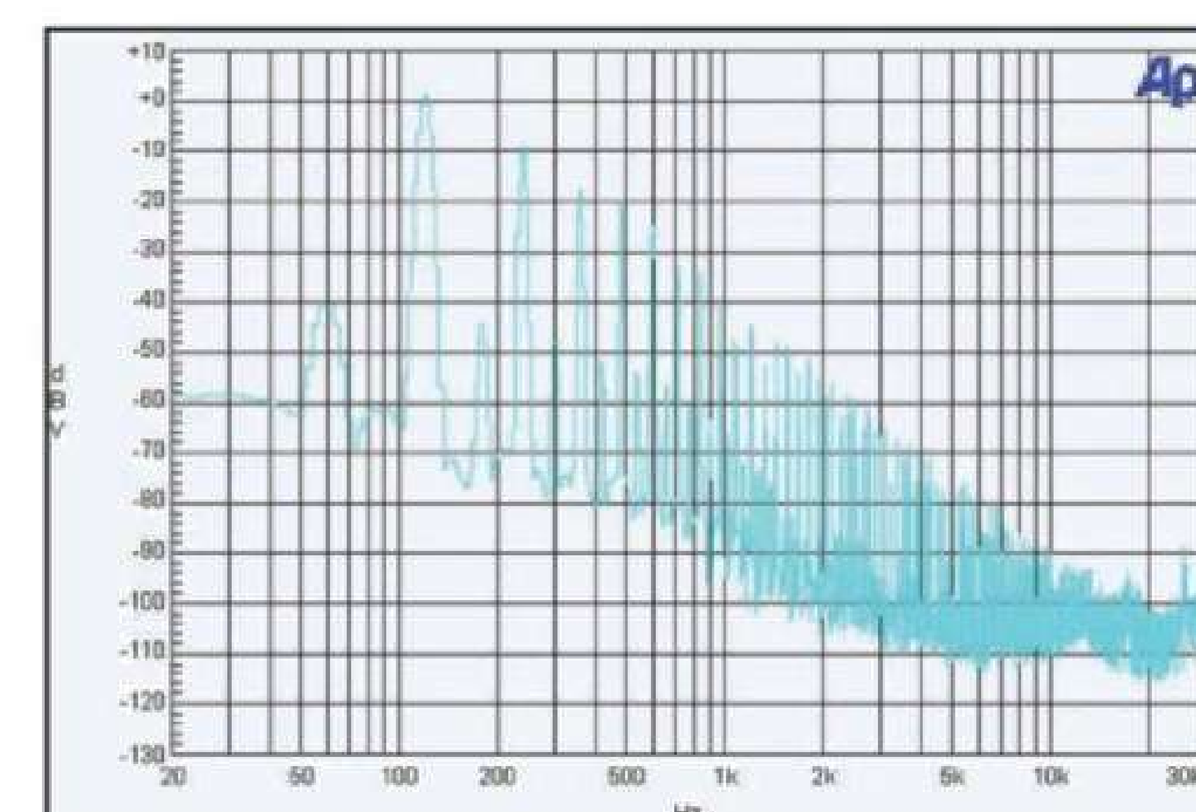


Figure 4: If we look at the output voltage of the power supply with our FFT analyzer, we see what at first looks just like the current plot.

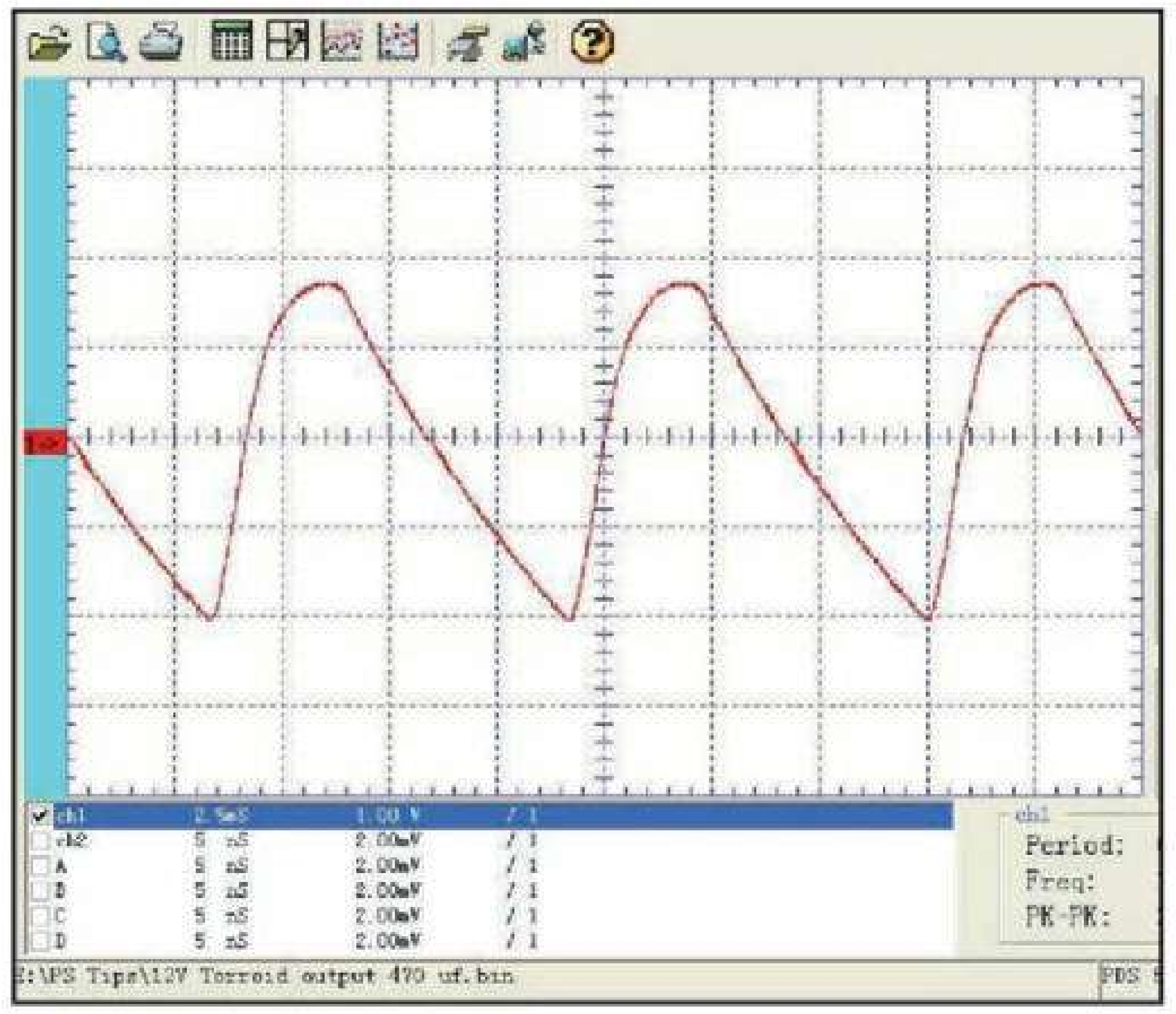


Figure 5: The ripple around 3,000 Hz (an area of particular hearing acuity) is still around 100 µV. With an oscilloscope, the noise or ripple on the output of the power supply looks like this.

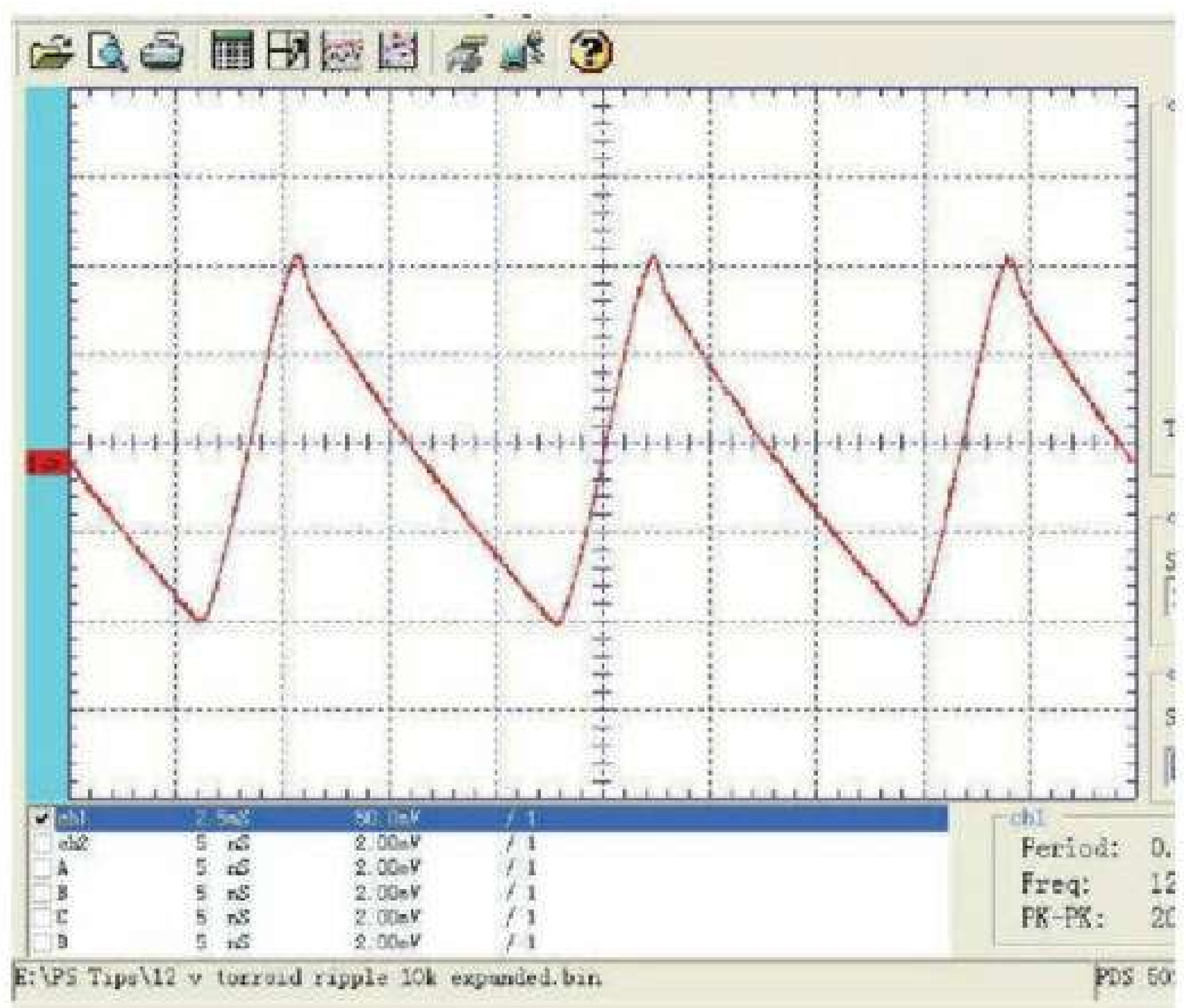


Figure 6: This is the same supply ripple as shown in Figure 5, but with a 10,000 µF axial filter capacitor.

About the Author

Ed Simon has a BSEE from Carnegie-Mellon University. He was employed by Cameradio, a regional electronics parts distributor from 1973-1975. In 1975, he started Edward Simon & Co. He has published more than 40 articles and taught courses for the NSCA, Carnegie-Mellon University and the University of Pittsburgh among others. His company specializes in sound and communication systems, specializing in systems for public facilities, houses of worship, and allied engineering. The business has completed more than 400 church installations, and 150 other installations including shopping malls, teleconferencing facilities, performing arts centers, television studios, and radio stations.

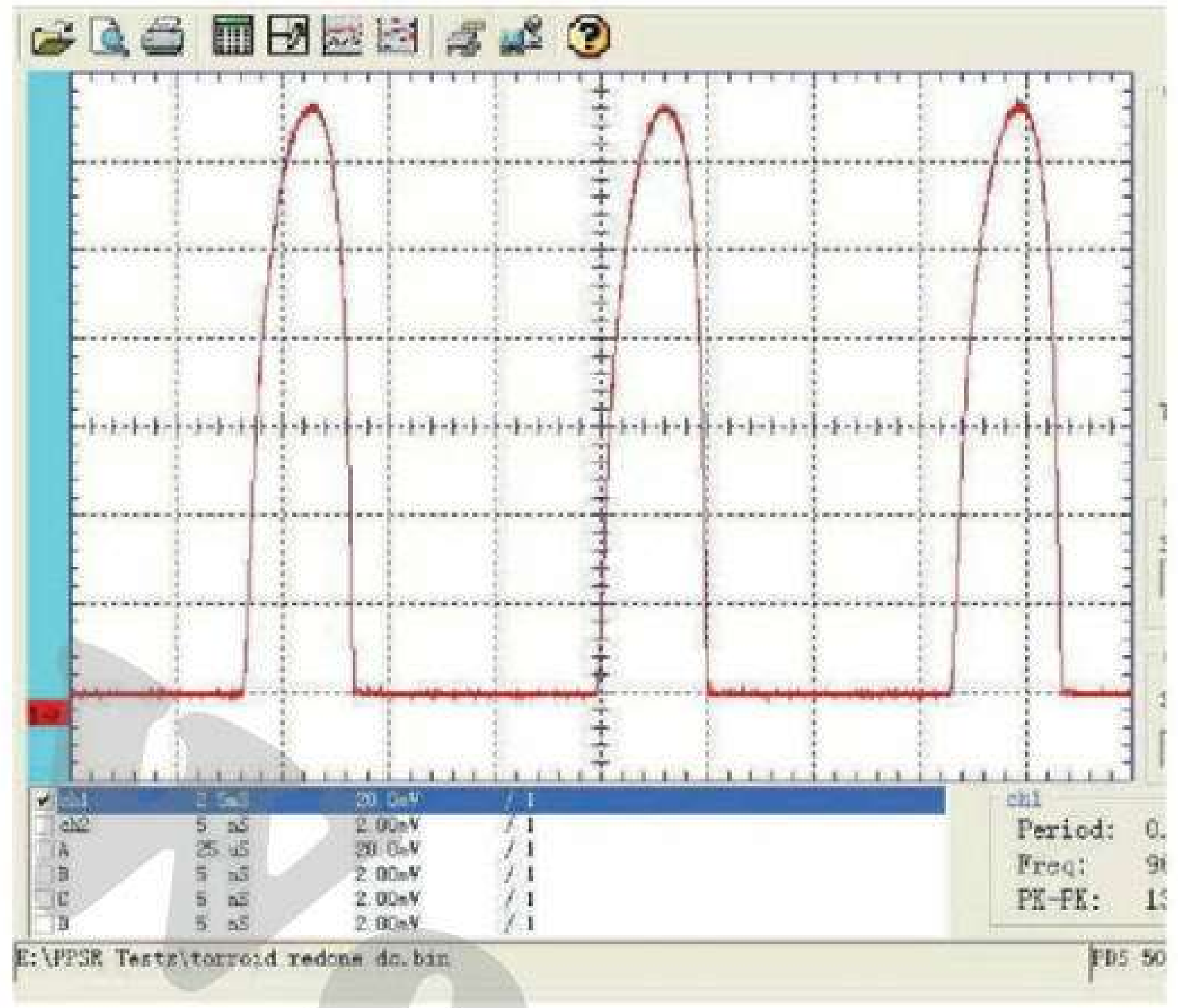


Figure 7: When we use a larger filter capacitor, it takes longer to charge.

150 dB. This is often at DC and decreases as frequency increases. It also may go down as the gain of the op-amp stage goes up. When there is a differential input, this will have a common mode rejection ratio (CMRR) that also reduces the complete circuits' PSSR, particularly when the input impedance is not the same for both inverting and non-inverting inputs. Of course, additional gain stages will increase the noise that does come through.

Lessening the Power Supply Noise

As shown, most audio enthusiasts would find this level of power supply noise excessive. So the first step to improve it would be to increase the filter capacitor size (see **Figure 6**).

The scale has been changed from 1 V per division to 50 mV per division. As expected the ripple has been reduced from a peak value of 3.76 V to 208 mV or from 23% to 1.3%. Close to the 20 times improvement that would be expected from the increase in capacitor size, but it really is just enough off of that to hint that there may be limits to what can be accomplished by only increasing capacitor size.

When we use a larger filter capacitor, as shown in **Figure 7**, it takes longer to charge. A close look will show that the charging current does not rises as abruptly as before. What is happening is that while the transformer is not delivering current the internal losses that we modeled as a resistor are very small so that the voltage on the secondaries is about 5% higher than it would be under load.

As the unloaded voltage rises, it becomes larger than the stored charge on the filter capacitor. The diode then begins to conduct. As the voltage rises into the capacitor the load and loss increase so the capacitor voltage rise slows as it begins to better match the loaded change in the AC line. This increases charging time on the front slope.

In the next article, we will detail what happens at the top of the AC waveform in more detail, and we will look at Full-Wave Bridge (FWB) circuits before we start discussing DC power regulation.

Triode Power Amplifiers

The first hollow-state amplifiers used triodes (since tetrodes and pentodes had not yet been invented), and many hollow-state audio enthusiasts today prefer the sound of triode amplifiers. Let's look at the performance implications of some design features of a couple of generic triode power amplifiers.

By
Richard Honeycutt
(United States)

The oldest architecture for a triode power amplifier is a Class A single-ended (SE) grounded-cathode stage as shown in **Figure 1**. Since vacuum tubes require a relatively high load impedance to produce much power, especially with significant efficiency, an output transformer is usually used to couple the 4 Ω, 8 Ω, or 16 Ω speaker to the amplifier. In order to prevent the DC component of the plate current from causing transformer saturation (and hence, distortion), transformers for this type of stage are usually built with a small gap in the magnetic circuit, requiring extra turns in the primary winding to provide sufficient inductance. Naturally, the resulting transformers weigh more and cost more than they otherwise would.

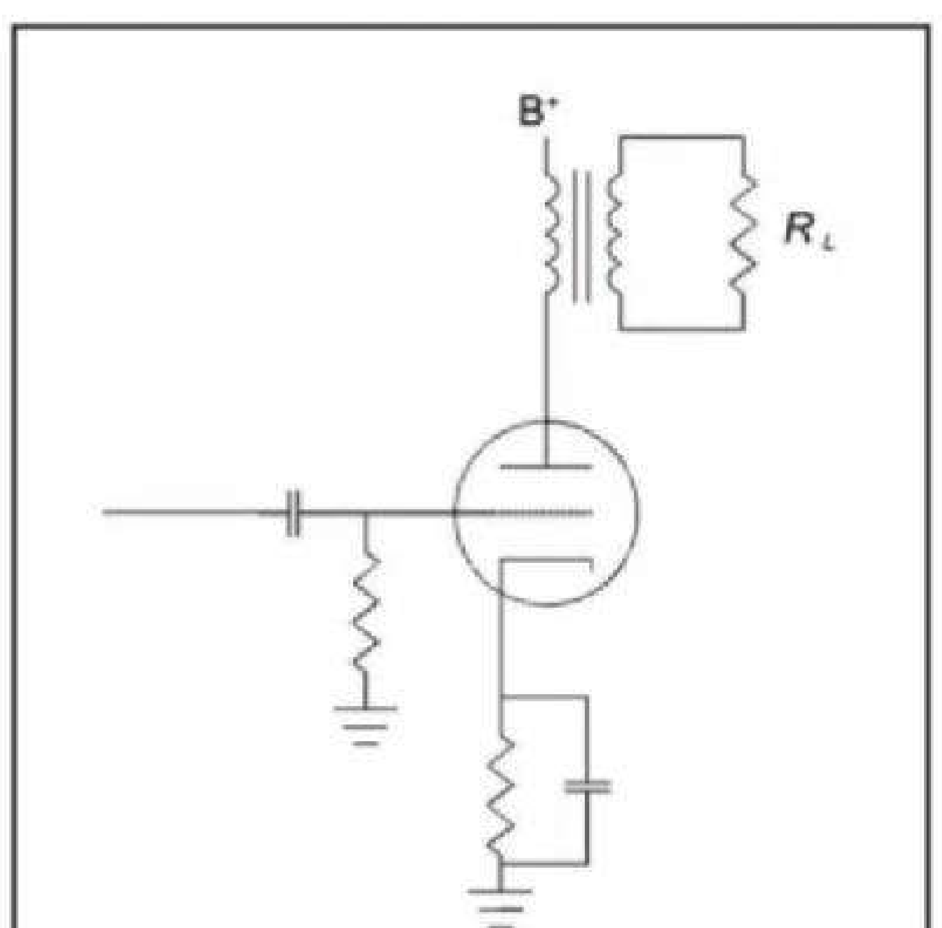


Figure 1: The single-ended, transformer-coupled stage comprised the first architecture for a triode hollow-state audio power amplifier.

The output transformer design affects the amplifier's distortion, maximum power output, and frequency response. Thus through the years, designers have developed output transformerless (OTL) hollow-state amplifiers. The most obvious way to do this with a triode is to capacitor-couple the speaker to a cathode-follower output stage. However, to achieve a 20-Hz low-frequency cutoff (-3 dB) for an 8 Ω speaker, you'd need a 995 µF capacitor.

Also, even using a triode-connected KT88 or 6146B output tube, you only get 175 mA maximum current to the speaker (actually, somewhat less, since a resistor is required to pass DC from cathode to ground). This amounts to about a quarter of a watt into an 8 Ω speaker. So even if you use very efficient speakers and can get by with only a 10 W power amplifier, you'd need 40 power triodes in parallel (with 40 995 µF output capacitors). Or you could have high-impedance speakers custom-made, but that opens another can of worms!

Clearly, the capacitor-coupled SE Class A triode power amplifier is not practical for an amplifier that will drive speakers, although it is certainly feasible for a headphone amplifier, which would drive a higher-impedance load (typically 32 Ω for modern headphones) with much lower output power requirements.

Another approach to output transformerless (OTL) amplifier design is explored on the Tube CAD Journal (see Resources).This is a bridged or Circlotron design whose concept drawing is shown in **Figure 2**. By using paired triodes with each unit of each pair operating under identical conditions, it provides an AC signal output that is free of DC, thus eliminating the need for the capacitor. While it eliminates output capacitors, the bridged approach still suffers from the difficulty of providing enough output power to drive most speakers. That being said, a triode-based Circlotron amplifier using four power triodes could drive a Klipschorn AK6 to 105 dB peaks (at 1 m distance).

A problem with Class A power amplifiers—even those using output transformers—is low efficiency. The maximum efficiency you can achieve with a Class A stage is 25%. This means that the tube must dissipate three times as much power as is fed to the load. Thus, a KT88, with its 35 W maximum plate dissipation, can only output a bit under 12 W as a Class A amplifier; a 300B, with its 40 W maximum, can squeak out a little more. Paralleling two output tubes doubles this power. But by using Class B (push-pull or peak-to-peak) operation, we can increase the efficiency to somewhat over 70%, enabling a KT88 pair to put out 100 W of audio. The 300B, with its lower maximum plate current (70 mA vs 175 mA for the KT88) was able to output 20 W from a peak-to-peak (PP) pair in the classic Western Electric model 92B power amplifier, whose schematic is shown in **Figure 3**.

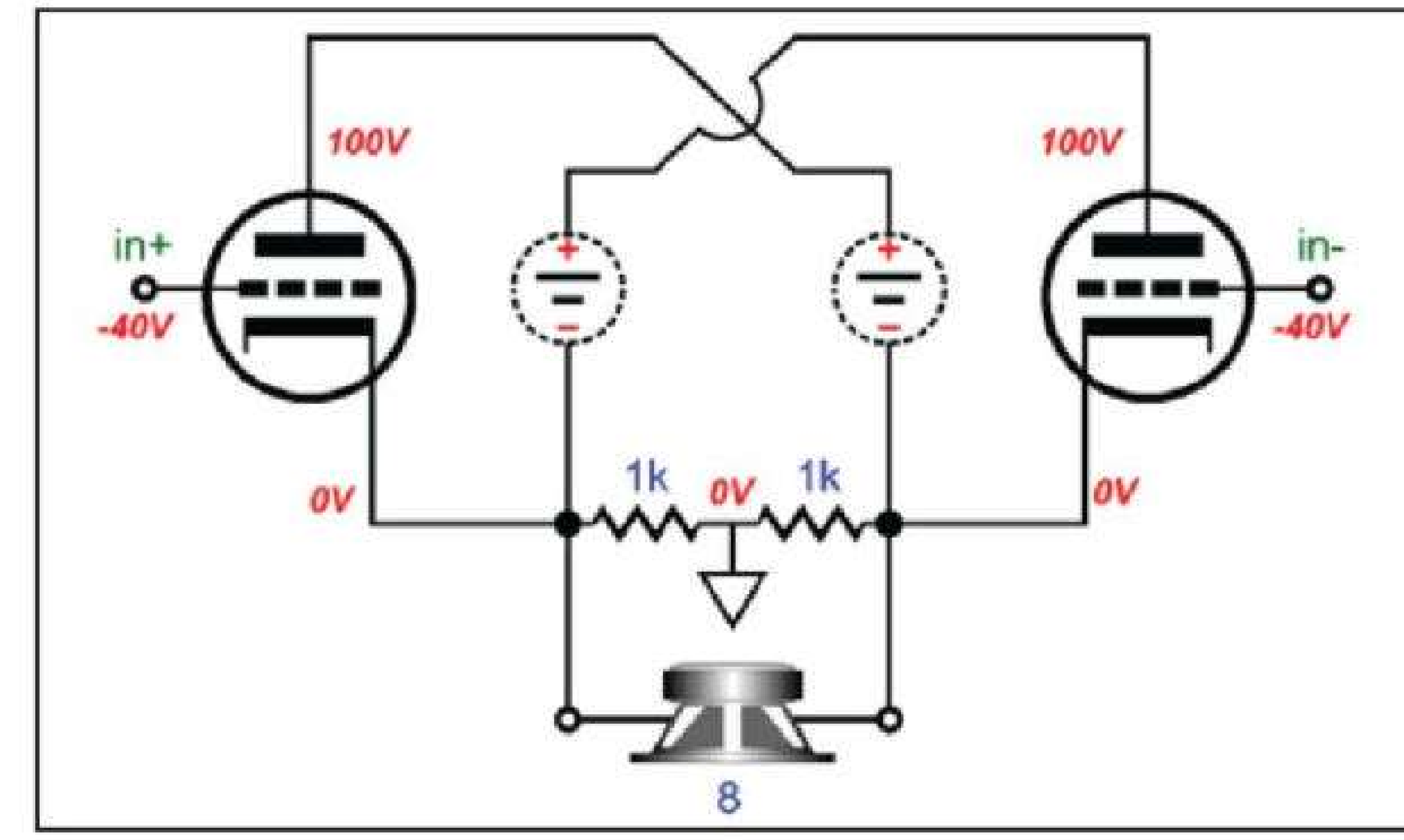


Figure 2: This Circlotron OTL triode power amplifier concept is explored on the Tube CAD Journal website.

Actually, Class B operation is defined as having each tube biased so that it conducts only on either the positive or the negative half-cycle. In order to decrease crossover distortion that occurs when the signal switches from one tube to the other, PP pairs in audio amplifiers are usually biased as Class AB1, which keeps each tube conducting for a bit more than a half-cycle, but has little effect on efficiency. (These considerations were discussed in my May and June 2012 Hollow-State Electronics articles.)

Distortion Signature

In addition to differences in efficiency, Class A and Class B amplifiers have different distortion

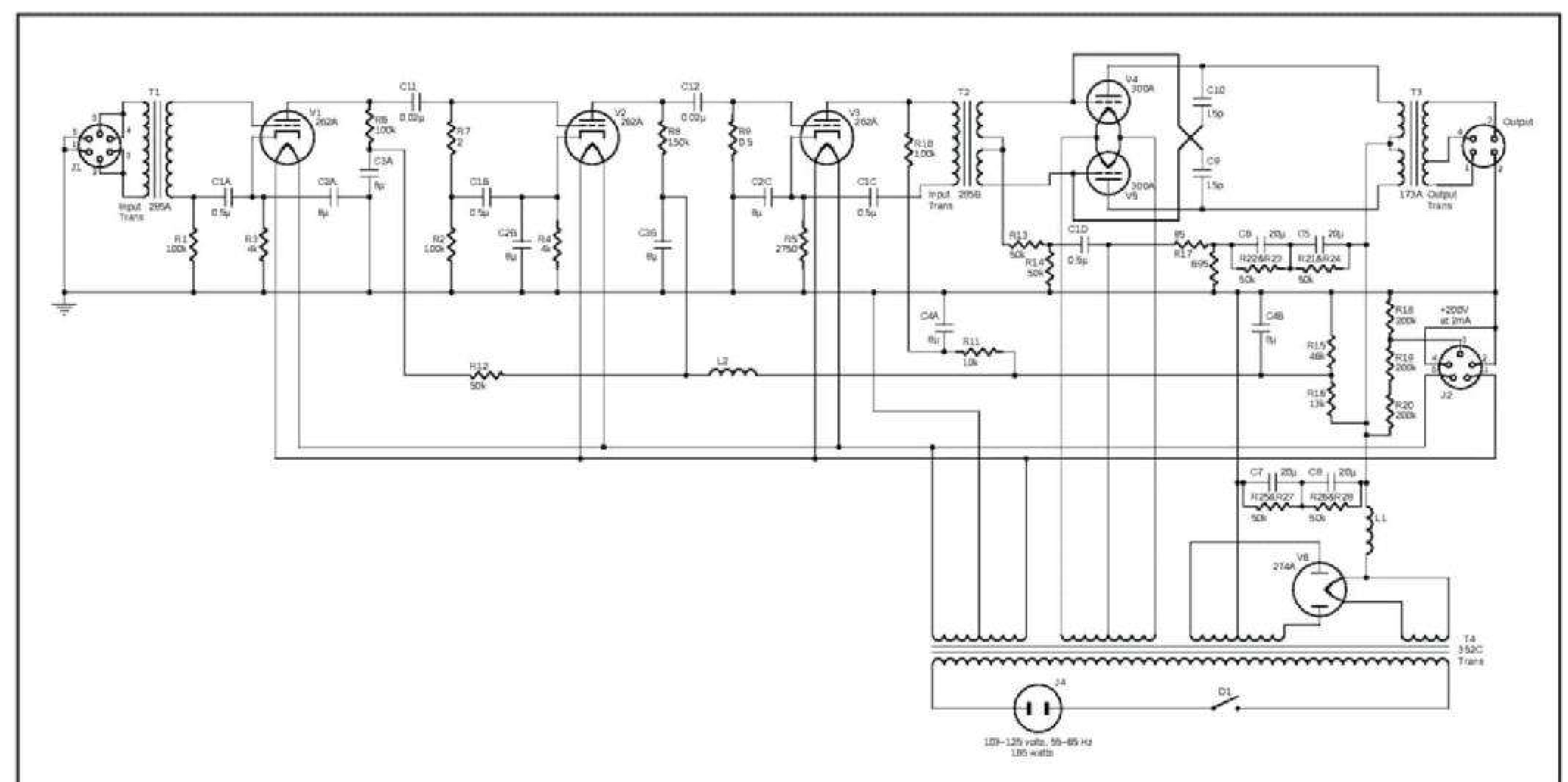


Figure 3: The Western Electric Model 92B PP triode power amplifier provided 20 W maximum output power.

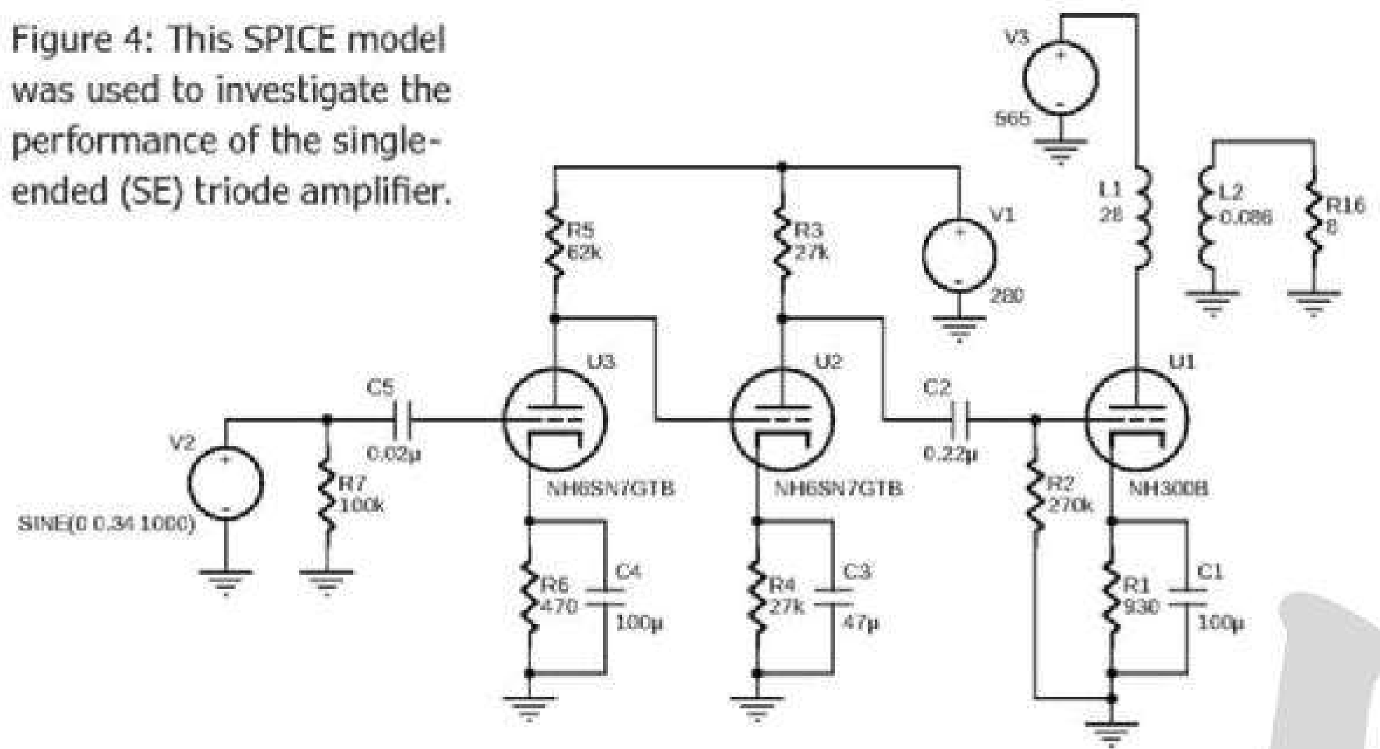


Figure 4: This SPICE model was used to investigate the performance of the single-ended (SE) triode amplifier.

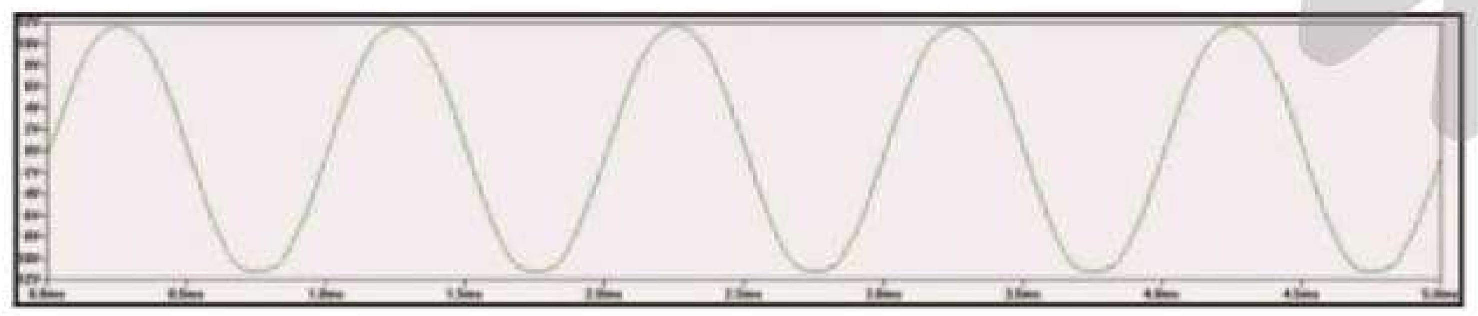


Figure 5: The SE stage produced this output wave at the onset of soft clipping.

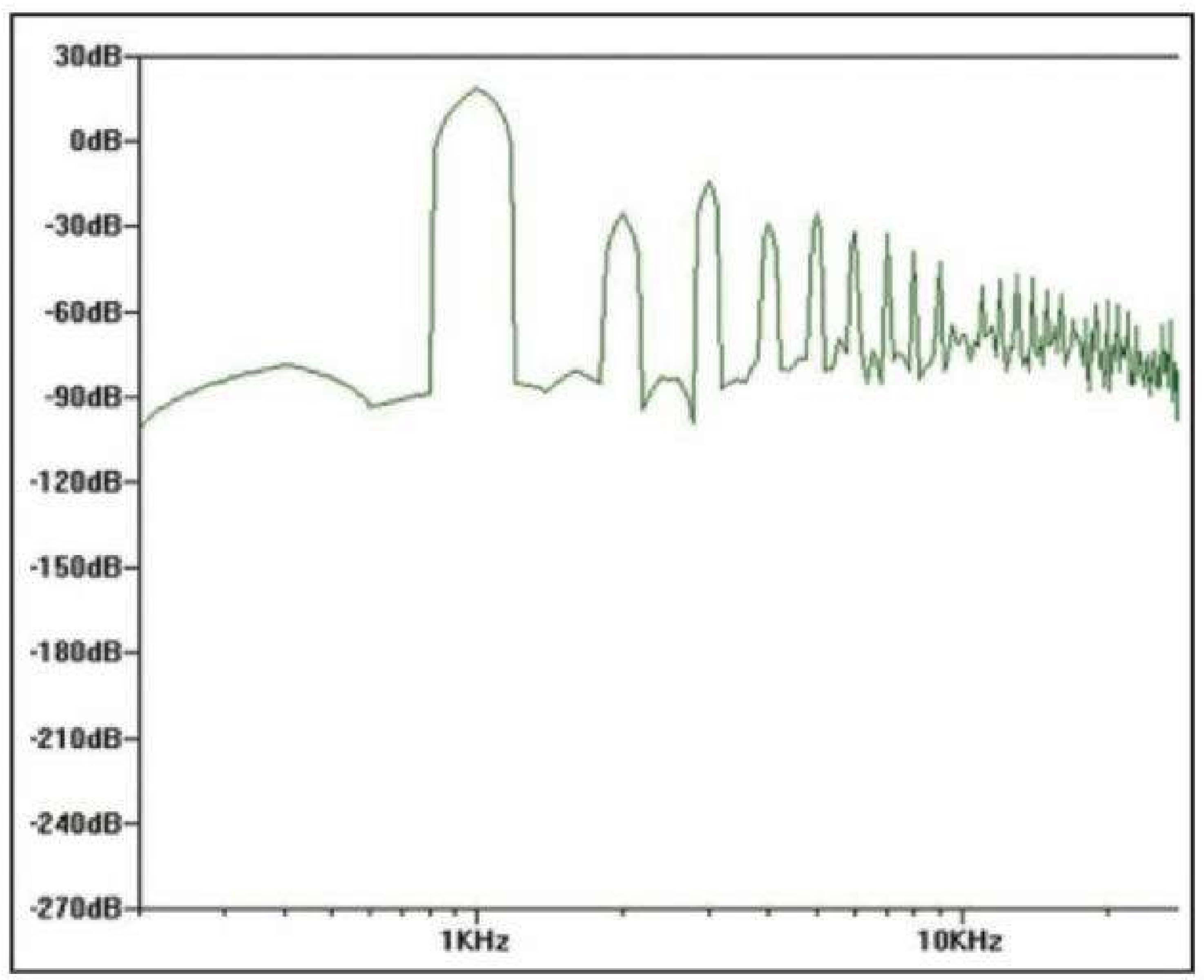


Figure 6: Although this Class A amplifier produces more odd- than even-order distortion, you can still see significant second, fourth, and other even-order harmonics.

Resource

"John Broskie's Guide to Tube Circuit Analysis & Design," Tube CAD Journal, March 2018, https://www.tubecad.com/2018/03/blog0416.htm.

R. Honeycutt, "Hollow-State Power Amplifiers," *audioXpress*, May 2012.

R. Honeycutt, "Hollow-State Class AB and Class B Amplifiers," *audioXpress*, June 2012.

signatures: Class A triode amps produce more even-order distortion than Class B designs, which cancel out most even-order distortion. While everyone agrees that Class B amps have less total distortion, some listeners prefer the sound of even-order distortion over odd-order distortion enough that they are willing to put up with more total harmonic distortion (THD) in order to reduce the relative amount of odd-order distortion products.

I decided to use LTspice to model a generic single-ended (SE) triode power amplifier and a generic PP one, both using 300B output tubes, to compare the performance. In both cases, the effects of the transformers on distortion and frequency response do not appear, since I did not have accurate SPICE models of either SE or PP output transformers, and thus, I modeled the transformers as ideal units. This enables us to focus our attention on the tube performance and the effect of the circuit architecture.

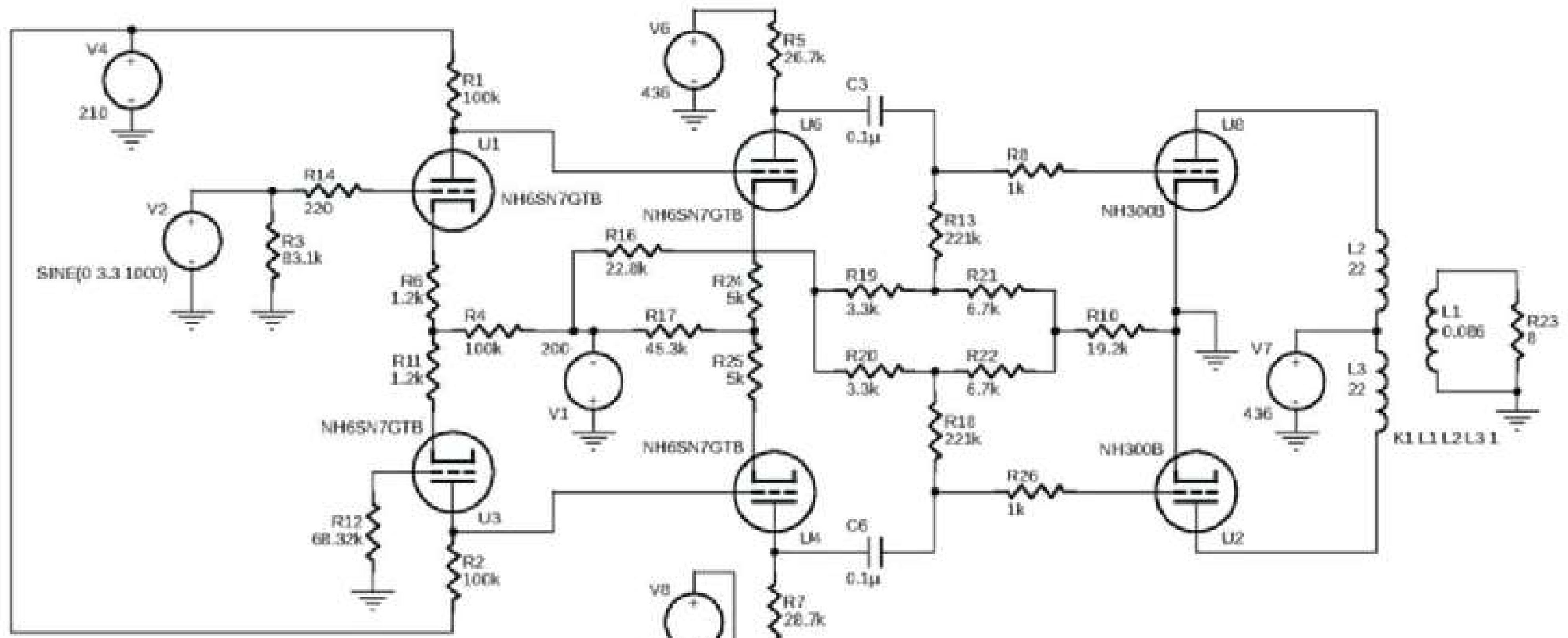
Figure 4 shows the model of the SE amplifier. Care was taken to ensure that the two voltage-amplifier stages did not clip before the output stage. **Figure 5** shows the output wave produced when the stage was driven very slightly into soft clipping. Notice the rounding of the negative half-cycle, which is a result of second-harmonic distortion, as you can see from the Fast Fourier Transform (FFT) shown in **Figure 6**. (Odd-order harmonics distort the waveform symmetrically, while even-order ones give asymmetrical distortion.)

Figure 7 shows the model of the PP amplifier. Comparing this circuit with that of the Western Electric Model 92B shown earlier, you will notice that the interstage transformers have been replaced by tube stages. Its output wave at the onset of clipping is shown in **Figure 8**. Even though we do not see the rounding of the negative peaks that was characteristic of the SE stage, with careful examination, you can see very slight flattening of the extreme negative portion (and to a lesser extent, the extreme positive portion) of the wave. The FFT shown in **Figure 9** illustrates the decrease in even-order harmonic components compared to those of the SE circuit.

Frequency Response

Another interesting difference appears when we compare the frequency responses of the SE and PP amplifiers, as you can see in **Figure 10** and **Figure 11**. The gentler roll-off of the SE stage, beginning at a lower frequency, has nothing to do with the circuit architecture, but results from the slightly higher grid-to-plate capacitance of the

Figure 7: This model was used for a generic peak-to-peak (PP) triode power amplifier.



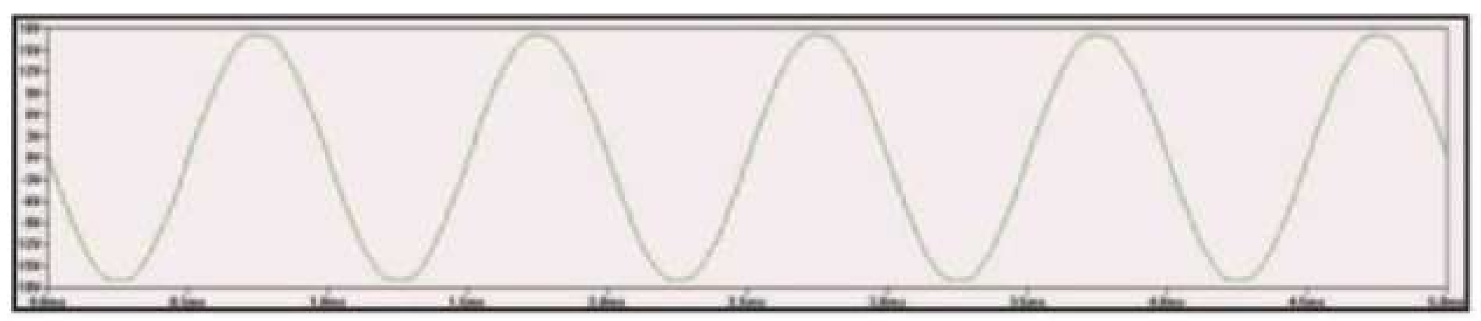


Figure 8: The output wave of the PP circuit shows less total distortion than that of the SE circuit.

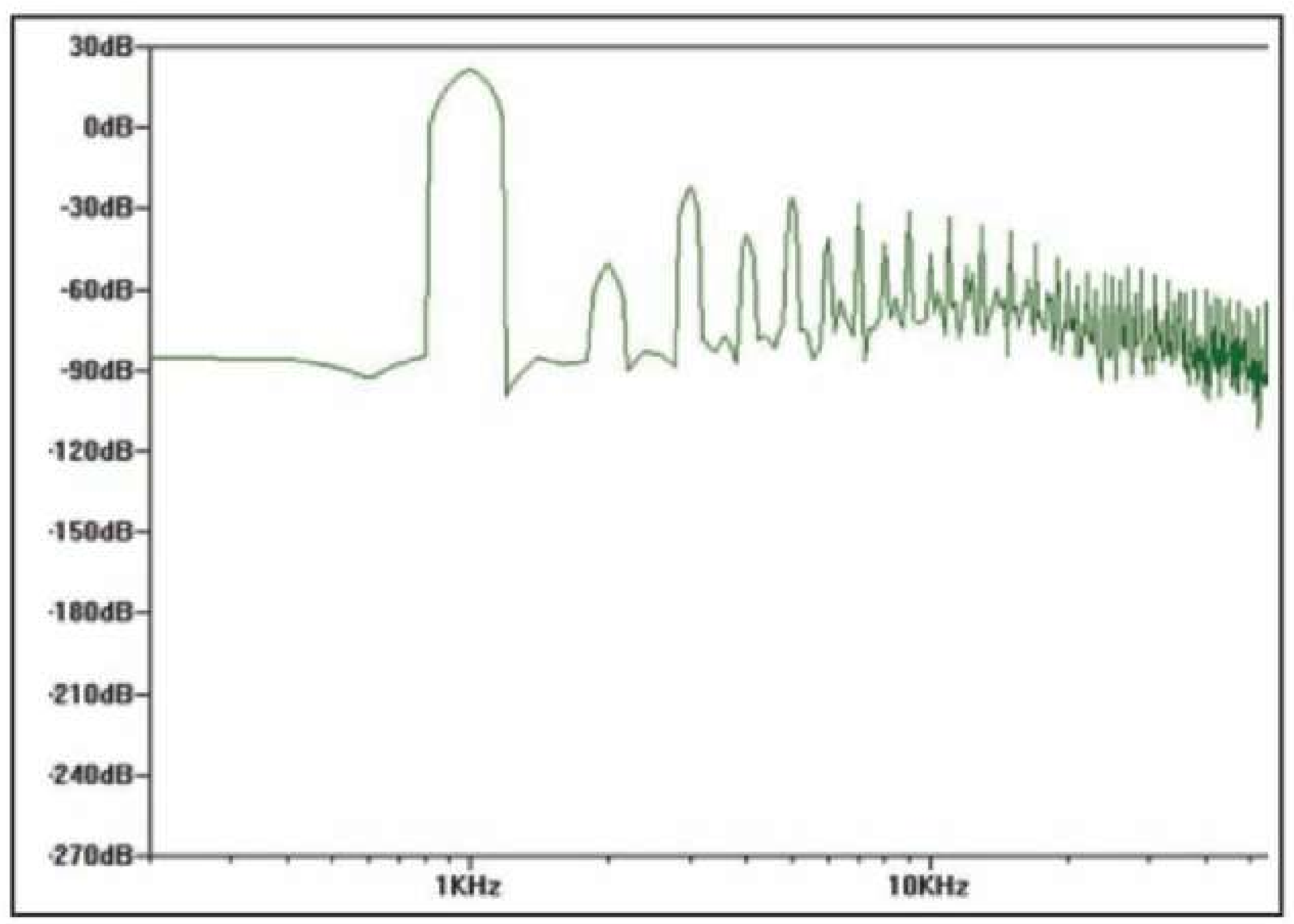


Figure 9: The Fast Fourier Transform (FFT) of the output wave of the PP circuit shows lower even-order harmonic levels than that of the SE stage.

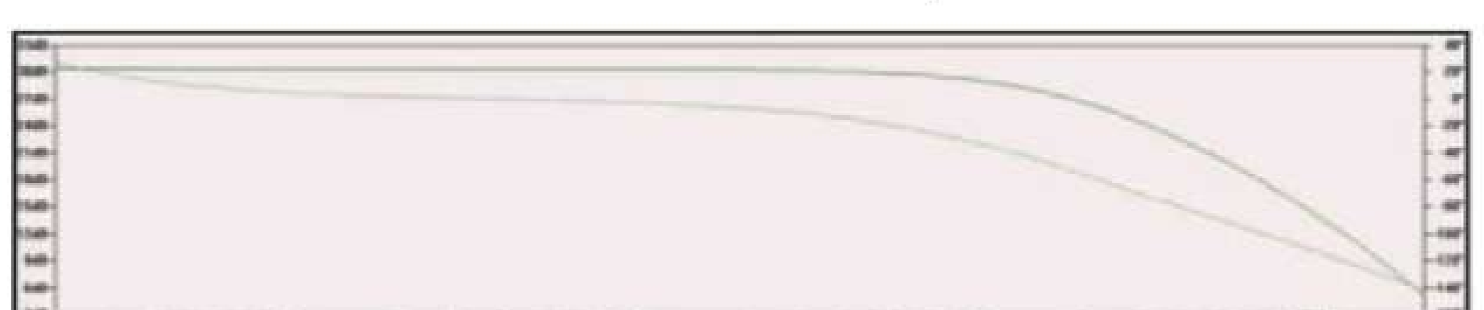


Figure 10: The frequency response of the SE stage rolls off gently, starting just below 10 kHz. (The solid curve shows the magnitude response; the dotted curve, the phase response.)

300B, compared to the 6SN7s used for voltage amplifiers in the PP circuit.

Tube Choices

The best-known vintage power triodes are the Western Electric 300 series (300, 300A, and 300B), now manufactured by several companies. But today's designer has other choices of power triodes. The 2A3 has a maximum plate voltage of 300 V and a maximum plate dissipation of 15 W. The 811 has a maximum plate voltage of 1250 V and a maximum plate dissipation of 45 W. The 6AS7 has a maximum plate voltage of 250 V and a maximum plate dissipation of 13 W. In addition, power pentodes and beam power tubes can be connected as triodes, as illustrated in **Figure 12**. **Table 1** was taken from a section of the Genelex datasheet for the KT88. It shows the performance of that tube when used in a PP triode connection, in which application it has a µ of 8.

Interstage Transformers

You may wonder how much difference in sound the transformer interstage coupling of the old Western Electric and other vintage amplifiers makes, compared to the extra tube stages used in modern hollow-state amplifiers. First of all, transformers provide noise-free gain, as long as they are well balanced during the manufacturing process, and also well shielded against induced hum and other electrostatic and electromagnetic noise. But they do somewhat limit high-frequency and low-frequency response, although with careful transformer and circuit design, these effects can be minimized. Transformers can also add distortion if they are overdriven. This problem, too, can be minimized if the proper transformer is chosen for the application. Making such choices is an art and science itself!

But the primary reasons that interstage transformers are not much used any more are their cost, weight, and difficulty in sourcing. For years, the Gibson Guitar company made amps that used transformers as phase splitters, and my experience has been that these interstage transformers were among the most common failure items for those amps. While these failures may have been caused by corrosion of the windings (I really don't know), such experiences have not encouraged me to design interstage transformers into circuits, if I can avoid doing so. While tubes do need replacing from time to time, that job is much easier than replacing a transformer.

Whether you design hollow-state amplifiers as a hobby or as a profession, or just want to better understand how Class A and Class AB1 amps compare, hopefully this article has been helpful to you.

Figure 11: The frequency response of the PP stage stays flat to above 20 kHz, then rolls off more rapidly.

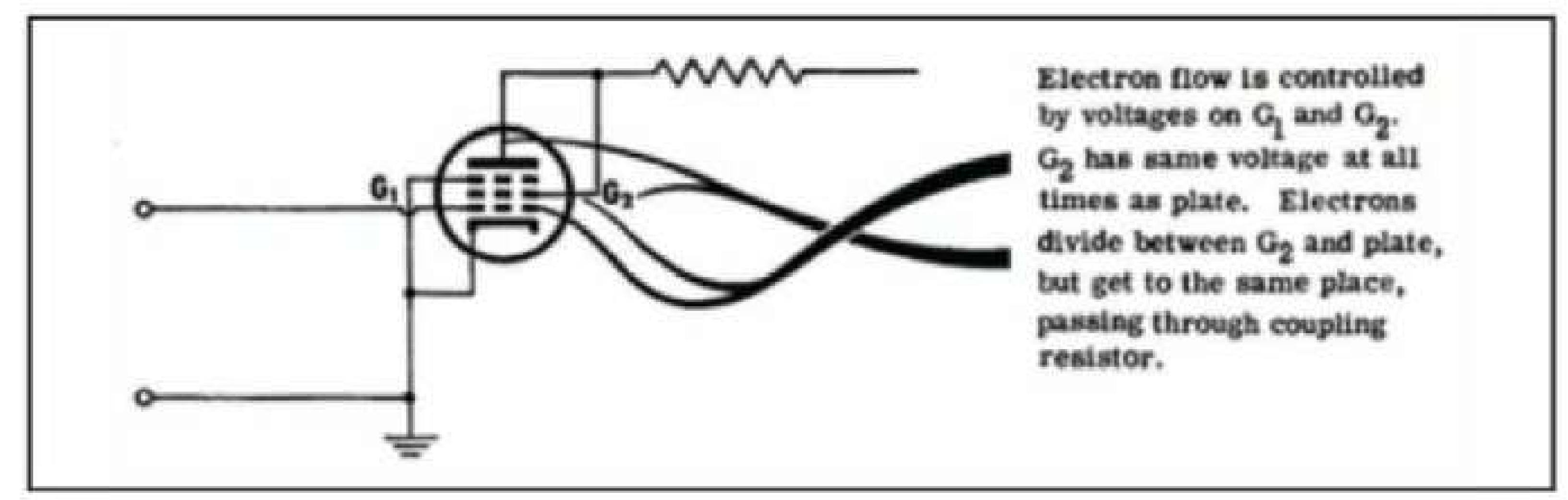


Figure 12: Pentodes can be connected for use as triodes (from N.H. Crowhurst's *Basic Audio*: www.vias.org/crowhurstba/crowhurst_basic_audio_vol2_053.html).

Push-Pull Triode Connection. Cathiode Bias.			
$V_{a(b)}$	400	485	V
V_a,g2	350	425	V
I_a+g2(0)	2×67	2×85	mA
I_a+g2(max signal)	2×72	2×90	mA
P_a+g2(0)	2×24	2×40	W
*R_k	2×525±5%	2×525±5%	Ω
V_{g1} (approximate)	-38	-48	V
$V_{in(g-g)}$	60	70	V
$R_{L(a-a)}$	4	4	kΩ
Z_{out}	2.5	2.5	kΩ
P_{out}	15	27	W
**D	1–3	1–3	%
Intermodulation	6	6	%

*Separate bias resistors are essential
**The distortion will vary according to the degree of matching

Table 1: Taken from a Genelex datasheet for the KT88, the data shows the performance of the tube when used in a PP triode connection.

Here are a few places where you might find a copy of *audioXpress* and possibly meet one of our authors and staff members.

May 11–15, 2020
179th Meeting of the Acoustical Society of America (ASA)
Chicago, Illinois
acousticalsociety.org

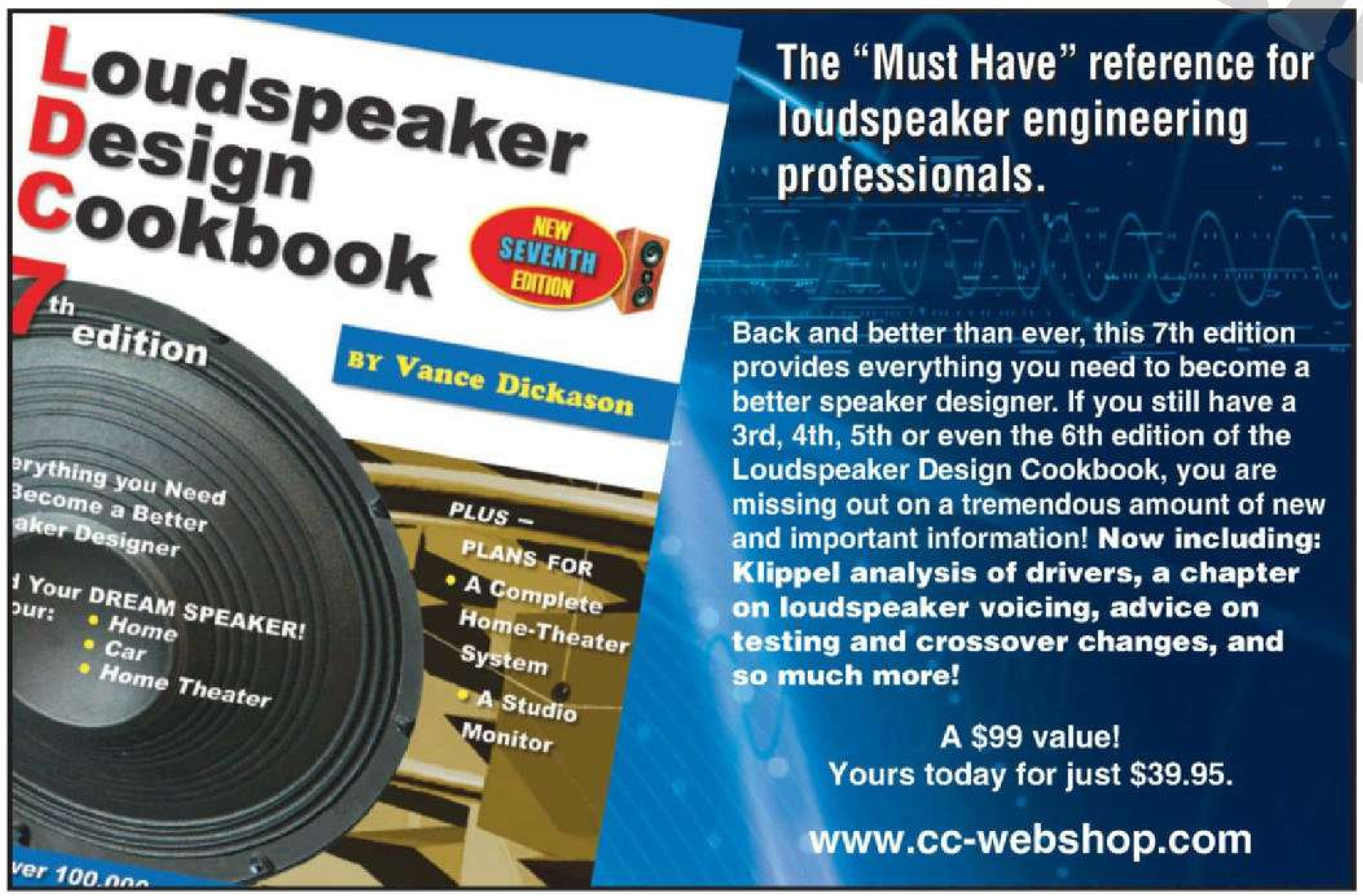
The Acoustical Society of America (ASA) holds two meetings each year, one of which is held at various locations in the US and the second one abroad (this year in Cancun, Mexico). At each meeting, invited and contributed papers and poster presentations, usually number between 850 and 1,100, and reflect the diversity and interests of the membership.

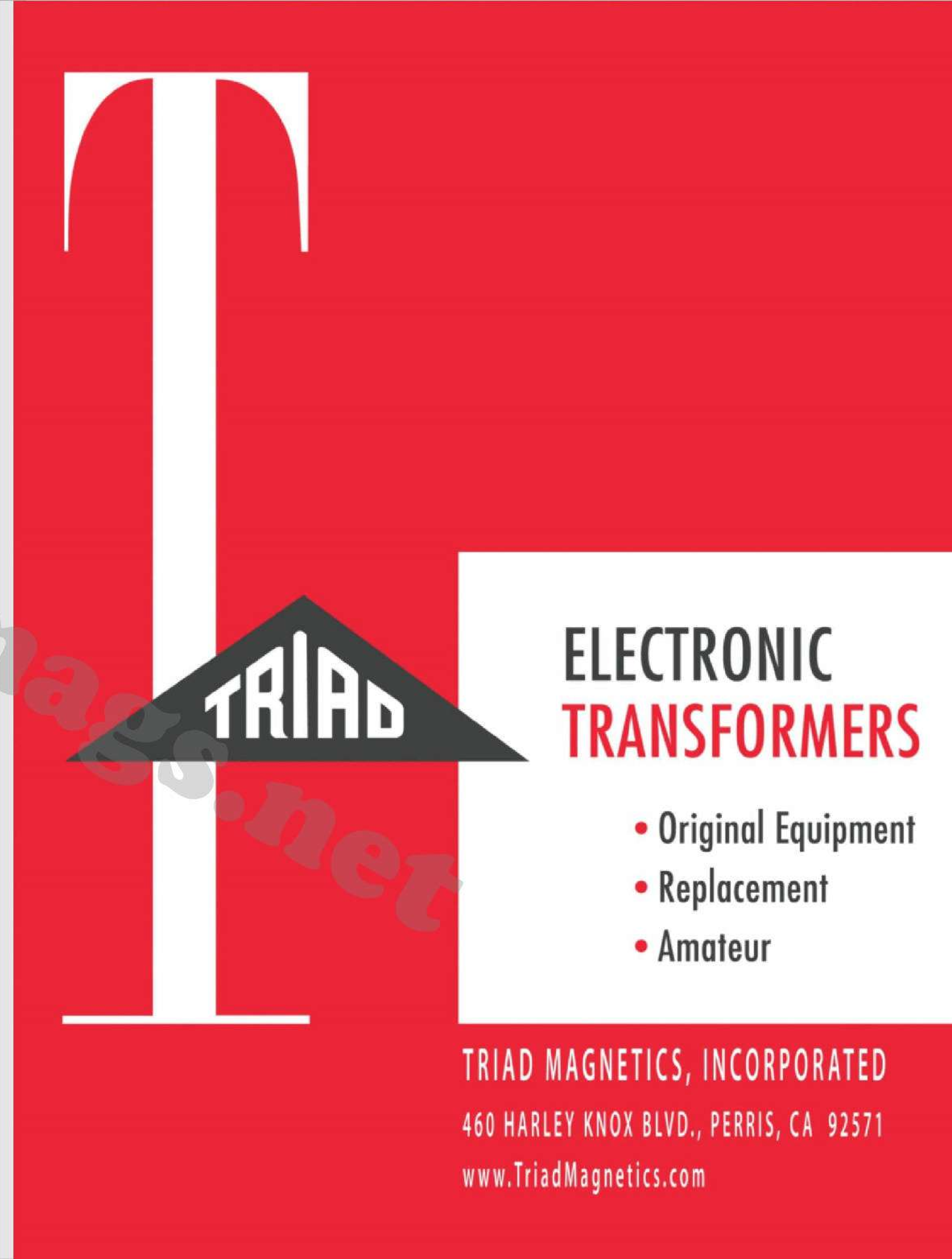
Tutorials, lectures, technical and standards committee meetings, and visits to university, government, and industrial laboratories are regular features, along with receptions and a variety of social programs. These meetings offer opportunities for students and young researchers as well as experienced acousticians to share information. Members who are able to regularly attend meetings find the personal contacts a rewarding experience.

2020 will be the International Year of Sound! IYS2020 will be a global initiative to highlight the importance of sound in all aspects of life on earth and will work toward an understanding of sound-related issues at the international level. For more information, visit https://sound2020.org.

May 25-28, 2020
Audio Engineering Society (AES) Vienna 2020
148th Audio Engineering Society International Pro Audio Convention
Austria Center, Vienna
http://www.aes.org/events/148

The Audio Engineering Society (AES) has announced that AES Vienna 2020 convention co-chairs are Piotr Majdak (Senior Researcher in Acoustics at the Austrian Academy of Sciences' Acoustics Research Institute) and Nadja Wallaszkovits (Chief Audio Engineer for the Phonogrammarchiv sound archives of the Austrian Academy of Sciences and 2019 AES President). The co-chairs will oversee the convention's Technical Program, Exhibition, and Society functions, in conjunction with an assembled team of industry professionals, students, event partners and sponsors. For end users, researchers, educators, and students, AES Vienna will be the place to commemorate the International Year of Sound and explore "A World of Audio" with audio professionals and peers from around the globe. Convention program topics will include Studio Recording, Live Sound, Networked Audio, Broadcast & Streaming, Music Production, Post-Production, Game Audio, Spatial Sound, Audio for AR/VR/XR, Product Development and more.





Full-Features and Versatility
Put You in Control!

Dayton Audio PPA800DSP
2-Way Plate Amplifier

Adjustable Filters

Parametric EQ

The perfect amp for your next project!

Tailor to fit your needs

With customizable DSP and 800 watts of clean power **(600W RMS for low frequencies; 200W RMS for highs)**, Dayton Audio's PPA800DSP is perfect for building your own powered PA speakers or even retrofitting existing passive speakers. With the ability to set crossover frequency, parametric EQ, and delay for each channel the possibilities are endless. In addition, the 4 additional programmable EQ presets and Bluetooth input gives this amplifier all of the features you would expect from todays powered PA speakers.

- ***Bluetooth input (summed to mono) for use with phones, tablets, or laptops***
- ***2-channel mixer built right in, plus line out for subs or additional speakers***
- ***A total of 5 programmable EQ presets per output for various applications***
- ***Balanced XLR and unbalanced 1/4" inputs for flexibility***
- ***Easy OEM customization using the 1" x 2-1/2" blank plate at the top of the amp***

DAYTONAUDIO

Custom OEM inquiries welcome.

SAVE $10 Code: AXMQ1Q4 Expires 4/30/2020
off $100 order + FREE SHIPPING!

Not valid with any other offer, on prior purchases or backorders.
Not valid on Commercial Sales accounts. No cash value.
Valid for one time use.

Distributed by:

PARTS EXPRESS
AUDIO COMPONENTS & SOLUTIONS

parts-express.com
1-800-338-0531