V. Jovanovic, "Tracing Distortion on Vinyl LPs"
J. Audio Eng. Soc., vol. 71, no. 10, pp. 616–637, (2023 October).
DOI: https://doi.org/10.17743/jaes.2022.0101

Tracing Distortion on Vinyl LPs

VLADAN JOVANOVIC, *AES Member*
(vladan.m.jovanovic@gmail.com)

Consultant, Bloomington, IN

Tracing errors arise because the reproducing stylus is of a different shape than the cutting chisel used to create the original acetate lacquer master for vinyl Long Play (LP) records. Tracing errors are typically the most significant source of distortion in vinyl reproduction and probably the main reason manufacturers of pickup cartridges seldom specified distortion figures for their products. In this paper, a historical overview of harmonic distortion results due to tracing errors is provided. In many cases, these results are 70–80 years old and, at least in some cases, seem largely forgotten by now. Some new simulation results are provided to verify various approximations proposed and used in the past.

0 INTRODUCTION

Tracing errors are the most significant source of distortion in vinyl reproduction (along with usually less detrimental vertical and horizontal tracking errors and elastic and nonelastic deformation of the vinyl). Under many reasonable sets of system parameters, tracing errors can create distortion at levels that might surprise most audiophiles and possibly even some audio experts.

Although tracking distortion has received some attention recently [1,2], tracing distortion has not been discussed in this Journal for more than 40 years [3].[1] Given the recent "vinyl revival" [4], the moment seems opportune to refresh the collective memory.

This paper presents a historical overview of the analytical and practical results on tracing distortion since 1937. As such, it does not contain much new material but provides verification of the previously derived approximate results via modern numerical methods on a computer, which were not available in the 1930s and 1940s when the problem was first tackled.

This paper also presents some old results in a way that might be easier to understand. Researchers in the past had to address the performance of discs that came in various sizes and with multiple rotation speeds, so that most of the results were presented in terms of the normalized variables that were often not straightforward to interpret. This task is much easier today because the present "vinyl revival" seems limited to Long Play (LP) discs with 33⅓ revolutions-per-minute (RPM) speed.

The rest of the paper is organized so that SEC. 1 explains the nature of the problem and provides its mathematical description. SEC. 2 offers descriptions and presents some numeric parameter values of the discs and styluses required to understand the problem quantitatively, while SEC. 3 presents a summary of the available theoretical results for tracing distortion. SEC. 4 deals with the peculiarities of these distortions with the vertical, horizontal, and 45°/45° stereo groove cuts, while SEC. 5 presents various numerical results for distortion in modern vinyl LPs. SEC. 6 describes several previously proposed methods to reduce problems due to tracing errors by predistorting the signal before the groove is cut. SEC. 7 shows how tracing distortion limits the maximum signal levels (and thus the dynamic range) of the recorded waveform, and SEC. 8 summarizes some considerations about the audibility of tracing distortion.

1 TRACING ERROR

The origin of tracing error can be explained with the aid of Fig. 1. The operation of the lacquer-cutting chisel, which must have sharp edges to do its job correctly, is illustrated on top. The reproduction stylus, which is always rounded to prevent recutting and is highly polished to minimize wear and tear of the vinyl, is shown at the bottom. Note that despite numerous different types of stylus tips (spherical, elliptical, hyperelliptical, Shibata, Van Den Hul, MicroLine, etc.), the point of the stylus contacting the groove practically always has a circular cross-section, no matter what shape it might take farther away from it. In Fig. 1, x is the linear distance along the groove, y the groove undulation, while $\Phi(\varepsilon)$ describes the curvature of the stylus tip vs. ε, the distance from its center.

From Fig. 2, it can be seen that at the instance marked by the vertical line, the stylus will not be touching the groove at the point where the chisel did (marked with a triangle)

[1] Tracing distortion was also mentioned in a historical summary paper on disc recordings from 1985 [5].

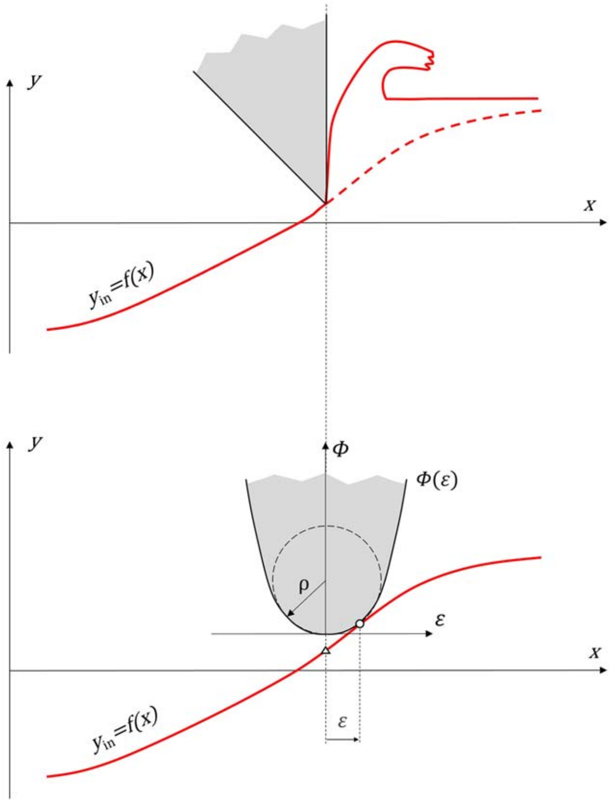

Fig. 1. Cutting chisel (top) and reproducing stylus tracing the same groove (bottom).[2]

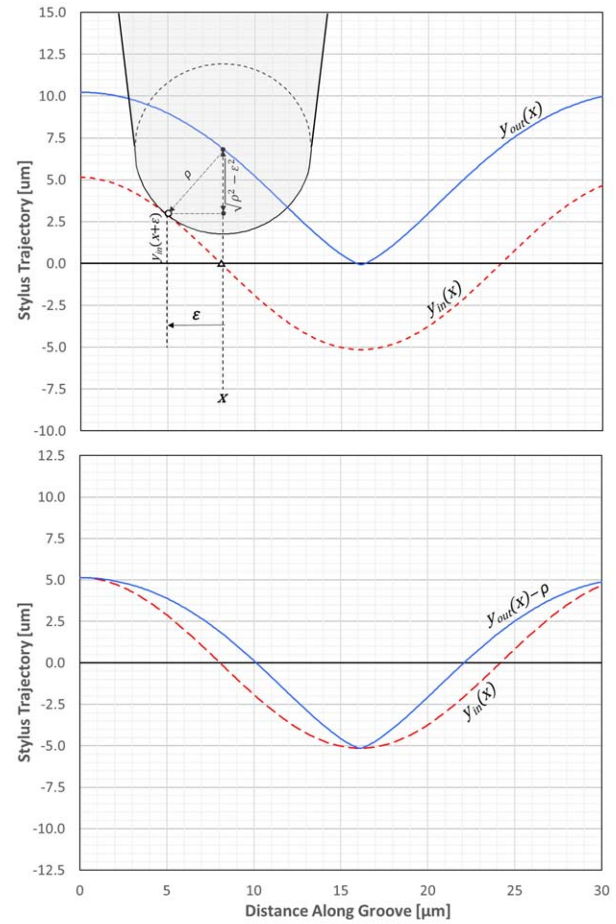

Fig. 2. Traced curve for a groove with the sinusoidal signal and spherical tip (v_{max} = 21 cm/s, R = 6 cm, ρ = 5 µm, f = 6.6 kHz).

but at some earlier point (marked with a circular dot, $\varepsilon < 0$ case). If the slope of the undulation is opposite, it could be touching it later ($\varepsilon > 0$, as in Fig. 1).

Mathematically, for a groove with a sinusoidal recorded signal, a stylus displacement vs. time is given by

$$y(t) = A\cos(2\pi f t),$$

where A is the amplitude and f is its frequency. Let Ω denote the angular speed of the record rotation

$$\Omega = 2\pi\, v_{RPM}/60,$$

where v_{RPM} is the disc rotation speed in RPM ($33\tfrac{1}{3}$ for LP records). The distance, x, that the stylus travels along the groove with radius R is

$$x = R\Omega t, \quad (1)$$

so the stylus displacement vs. distance along the track is

$$y_{in}(x) = A\cos\left(\frac{2\pi f}{\Omega R}x\right). \quad (2)$$

[2]Fig. 1 and Fig. 2 depict the so-called vertical cut, which is not representative of the engravings on either mono (horizontal) or stereo LPs (45°/45°cut). It is still a helpful starting point for explaining and analyzing tracing errors because the distortions for horizontal and 45°/45°cuts are much easier to explain once the simpler case of distortion mechanism in the vertical cut is understood.

A stylus's cross-section shape can be described by a function $\Phi(\varepsilon)$ in a coordinate system centered at the tip, with ε being the horizontal distance from the tip, as shown in Fig. 1. For a stylus with a spherically shaped endpoint

$$\Phi(\varepsilon) = \rho - \sqrt{\rho^2 - \varepsilon^2}, \quad (3)$$

where ρ is the stylus radius.

The point of contact at the location x will be at a value $x + \varepsilon_x$ for which the stylus shape curve $\Phi(\varepsilon_x)$ and the groove $y_{in}(x)$ have identical tangent lines. Slopes of the tangents being defined by the functions' first derivatives, the value ε_x at any point x along the groove needs to satisfy

$$\frac{\delta}{\delta\varepsilon}\Phi(\varepsilon_x) = \frac{\delta}{\delta x}y_{in}(x + \varepsilon_x), \quad (4)$$

It is then easy to show that for a sinusoidal recorded signal

$$\frac{\varepsilon_x}{\sqrt{\rho^2 - \varepsilon_x^2}} = -\frac{A\,2\pi f}{\Omega R}\sin\left(\frac{2\pi f}{\Omega R}(x + \varepsilon_x)\right), \quad (5)$$

needs to be solved for ε_x. The displacement seen by the stylus at some distance x will be determined by the position of its center. From Fig. 2, we then find

$$y_{out}(x) = y_{in}(x + \varepsilon_x) + \sqrt{\rho^2 - \varepsilon_x^2} = \\ = A\cos\left(\frac{2\pi f}{\Omega R}(x + \varepsilon_x)\right) + \sqrt{\rho^2 - \varepsilon_x^2}. \quad (6)$$

A major problem in this analysis is Eq. (5), which cannot be solved analytically (the same was true in the related case of tracking distortion, Eq. (3) in [1]).

With present-day computers, however, solving Eq. (5) numerically is relatively easy, with one such example shown in Fig. 2. Almost every paper on tracing distortion in the past included such a plot,[3] although the manual calculation of the resulting waveforms must have been incredibly tedious in the 1930s and 1940s.

The bottom part of Fig. 2 illustrates how the two waveforms compare when the stylus radius DC offset is removed from the output waveform $y_{out}(x)$. The offset stylus displacement goes through all the values of the input $y_{in}(x)$ but at the wrong times. These two waveforms are equal only at the two extreme values of the input sinusoid $y_{in}(x)$, because the tangents at the peaks are horizontal.

The name given to the solid (line) curve in Fig. 2 is a "poid," and it seems to have been invented by F. V. Hunt [6]. It is formally defined as the path traced by the center of the circle rolling over a sinusoid and appears to have not been previously described in mathematics. It is not included in most dictionaries but exists in the dictionary of IEEE standard terms [7].

2 SIZES AND SHAPES THAT MATTER

To better understand the nature of the problems involved in LP replay, it is illustrative to consider some of the dimensions involved. The playing time for one side of an LP varies but is usually supposed to be 22–23 min. Assuming 22.5 min, an LP would make about 750 revolutions while playing one side (22.5 × 33⅓). The outermost radius of the groove on a disc is standardized at 146.05 mm and the inner one at 60.325 mm [8],[4] so that the total length of the groove on a side is about half of a kilometer.

The valid range of radiuses for the grooves is thus about 86 mm, and that value divided by 750 gives about 115 μm (4.5 mils) of inter-groove[5] spacing on average. As the neighboring grooves cannot touch (see Fig. 3), this limits the maximum recorded groove modulation amplitude to about 50 μm (2 mils), a value most records seem to have stayed within [9].

[3] Such a figure first appeared in [6], but with a circle instead of a cone with a spherical tip, as depicted in our Fig. 2. Figures with a circle rolling over the groove were later often reproduced when the horizontal and 45°/45° cuts were discussed, confusing many readers.

[4] Original German DIN, Japanese JIS, and the first American RIAA standards prescribed slightly lower inner radius.

[5] One side of an LP has only one (although a spiral) groove, so this spacing relates to the neighboring sections of the same groove. Still, the "inter-groove" spacing and the "first" and "last" groove are the terms widely used within the recording industry.

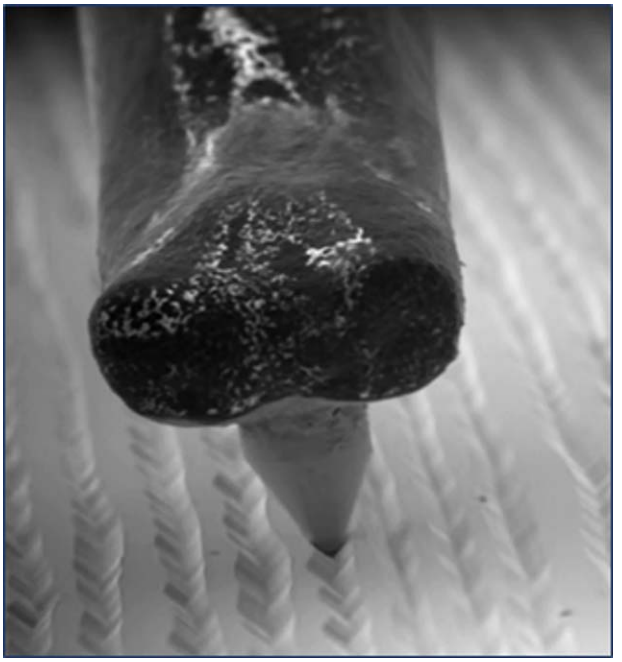

Fig. 3. The stylus in the groove (Courtesy of Ben Krasnow, www.youtube.com/c/AppliedScience)

Although 115 μm for inter-groove spacing appears to be about right on average, in the individual discs, the spacing might be considerably smaller, depending on how the recording was mastered. With soft music material, or the companders to limit the dynamic range (compress the loudest passages), the spacings could be smaller and the side playing time extended.[6] If the dynamic range was to be preserved, the grooves could also be spaced wider apart, with dynamic adjustments of the inter-groove spacings as required by the content. The most famous example of this is probably the iconic LP, "Overture 1812" by Telarc, shown in Fig. 4.[7]

In SEC. 3, it will be shown that the stylus speed of movement in the groove is often a parameter of more interest than the groove's amplitude itself. For reasons that are not clear now, within the industry, the peak velocity v_{max} was always expressed in cm/s, even in nonmetric countries. A simple calculation[8] shows that, with maximum amplitude $A_{max} = 50$ μm, the maximal recorded velocities are around $v_{max} = 3$ cm/s with a recorded tone frequency $f = 100$ Hz; it reaches 30 cm/s with 1 kHz and a whopping 300 cm/s with 10 kHz.

In SEC. 7, it will be demonstrated that other factors limit the recorded velocities to much smaller values than 300 cm/s. Most of the tests prescribed in the standards use a nominal peak velocity $v_{max} = 5$ cm/s.

[6] Todd Rundgren was famous for long LPs. E.g., side 2 of his album "Initiation" clocks at 36:00 minutes [10].

[7] "Telarc's" recording of "1812" lasts for about 15.5 minutes and takes a whole side – the last groove is at a radius of about 75mm. In the 1980s, it was often used for stress testing of cartridges and tonearms by many audiophiles.

[8] Eq.(16) from Section 3.

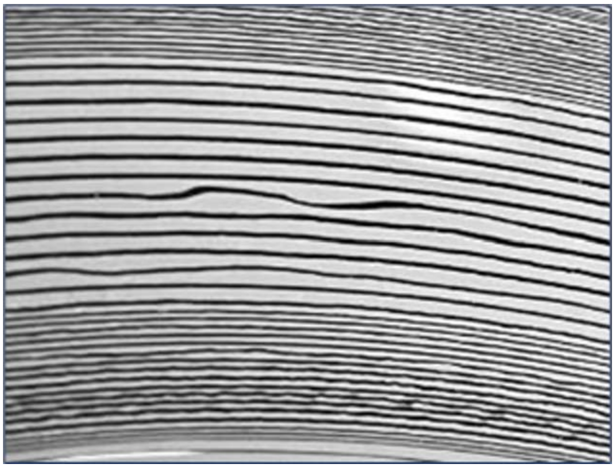

Fig. 4. Grooves with the canon shots on the author's "Overture 1812" LP by Telarc.

Another useful parameter to keep in mind is the recorded wavelength λ, i.e., the length of the groove corresponding to one period of the recorded sinusoidal tone. For a tone with frequency f, from Eq. (1)

$$\lambda = \frac{R\Omega}{f}. \tag{7}$$

If a tone with frequency $f = 10$ kHz is recorded, the wavelength would be $\lambda \approx 51$ μm (approximately 2 mils) at the maximal and about 21 μm (0.8 mils) at the minimal groove radius. The stylus is obviously dealing with the literally microscopic features engraved onto an LP disc.

This suggests that the stylus tips also have to be microscopic, and they actually are. Note that here we will be talking about the tip at the very end of the stylus, the part that actually fits into the groove in Fig. 3. The conical part to which the stylus tip (usually made out of diamond) is attached and the cantilever are much larger than the stylus.

When "microgroove" discs were first introduced (LPs in 1948 by "Columbia," and 7-in. singles in 1949 by "RCA"), all stylus tips were still spherical in shape. The initial design assumed the tip with a 25 μm (1 mil) radius [11], but by the time, the stereo disc arrived in 1957, the industry had settled on an unofficial standard of 18 μm (0.7 mils).

Because reducing the contact radius of the stylus (ρ in Fig. 1) reduces tracing errors, later in the 1960s, manufacturers developed the concept of an "elliptical" stylus. The name was somewhat misleading because most of these were the normal spherical styluses with two flat surfaces ground on the front and back of the tip, so that the horizontal cross-section in most cases looked like a rectangle with rounded sides rather than a true ellipse[9] (Fig. 5). The predominant size was 5 × 18 μm (0.2 × 0.7 mils approximately). Still, models with a smaller radius of 8 and 10 μm were also sold.

[9]"Shure," historically one of the most prominent makers of phonograph cartridges, often used the term "biradial" rather than "elliptical" for their styluses of this type.

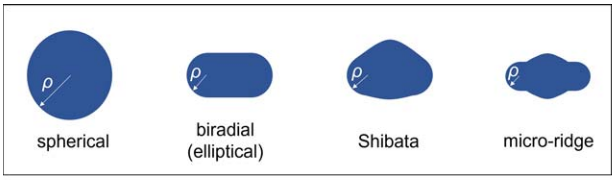

Fig. 5. Horizontal cross-sections of the styluses.

In the 1970s, when the ill-fated quadrophonic disc systems arrived (e.g., CD-4), phono cartridges needed to trace grooves with high-frequency content up to 45 kHz. The first such model was the famous "Shibata," but an avalanche of similar products followed,[10] as the laser cutting and improved polishing methods allowed for much more imaginative and complicated shapes. The contact area with the disc is still circular, but some later models reached the radiuses of 2–3 μm, about half the size of the best elliptical ones.

In what follows, we will only consider the kinematics of the stylus motion. When the mechanics of the problem is studied (e.g., in the analysis of the forces that cause the deformation of the vinyl), a parameter that often comes into play is stylus acceleration. Without going through the mathematics involved, it might be interesting to know that the stylus in the groove can experience accelerations of up to 1,000g [12] (g being the acceleration of gravity of 9.81 m/s^2). This is about 100 times larger than the maximum value experienced in military fighter jets [13].

3 TRACING DISTORTION

Nowadays, most engineers would probably solve Eq. (5) numerically on a computer and then perform a Fast Fourier Transform (FFT) of the waveform in Eq. (6) to obtain distortion results, but when the researchers first started looking into this problem, they did not have computers.

Unlike the tracking distortion case, in which the spectrum of the distorted waveform could be calculated without solving a transcendental equation (Eq. (3), [1]) similar to Eq. (5) for ε_x, in the case of tracing distortion, this proved impossible. Equipped with the mathematical skills and tenacity that would be considered enviable today, they still managed to obtain a series of successively more accurate approximative results, which is quite a fascinating story in its own right.

[10]Most came under company-specific, often trade-marked names, such as "Hyperelliptical," "Paroc," "Stereohedron," "Line Contact," "Fine Line," "MicroLine," "MicroRidge," "Van Den Hull," "SAS," etc. They created a lot of confusion among the audiophiles because, in most cases, they increased the size of the contact area with the disc in the vertical dimension, which reduced record wear and tear, but not necessarily in the horizontal, to reduce the tracing errors. For instance, the original "Shibata" and the first "Stereohedron" styluses by "Stanton" actually had the same horizontal contact radius of 5 μm as the contemporary best elliptical ones. In the later years, these radiuses became smaller.

The first to describe the tracing error mechanism was H. A. Frederick [14] in 1932,[11] but he did not offer any analysis of its effects. The first to analyze this problem was M. J. Di Toro in 1937 [15], who also seems to have coined the name "tracing distortion." By using an approximation to the original sinusoid that combined hyperbolic curves around the peaks with the linear segments between them, Di Toro managed to obtain reasonably accurate[12] but awkward and complicated formulas for the amplitudes of the harmonic distortion components, which provided little insight.

Quantitatively, he was also able to show that his results matched the experimentally measured values rather well, making him the first to report any measurements of tracing distortion. Qualitatively, Di Toro showed that the impact of tracing errors is twofold: he observed an increase in the power of the harmonics and a decrease in the power of the fundamental as the tracing errors got larger, but the relation of the results to the various parameters involved remained largely unclear.

One firm quantitative conclusion Di Toro made was that the harmonic distortion would become prohibitively high ($>10\%$) if the minimum radius of curvature in the recorded groove was less than 5 times the stylus tip radius. It will be shown here that this conclusion will be ignored in many analyses done several decades later. The other was that the harmonic distortion was a much more severe problem than the attenuation of the fundamental (a fraction of a dB when the total harmonic distortion (THD) reaches 10%).

Within less than a year, Di Toro's work was followed by a paper by J. Pierce and F. Hunt [6]. They were the first to explain the fundamental difference in distortion encountered with horizontal (lateral) and vertical ("hill-and-dale") groove recordings (to be discussed in SEC. 4) and provided a series of numerical solutions to Eq. (5) without any approximations. Combined with a series of 7-point Discrete Fourier Transform (DFT) results for thus calculated waveforms, this enabled them to create the nomogram graphs from which quite accurate distortion results could be read. They also devised an "empirical" formula for the n-th harmonic $\mathcal{A}_n$ of the traced waveform

$$\mathcal{A}_n = \frac{(kAk\rho)^{n-1}}{n^2}, \tag{8}$$

where A is the amplitude recorded in the groove, ρ the radius of the curvature of the stylus tip that is in contact with the groove, and k is a wavelength parameter given by

$$k = \frac{2\pi}{\lambda} = \frac{2\pi f}{R\Omega}, \tag{9}$$

as can be shown with the aid of Eq. (7).

F. Hunt subsequently joined forces with W. Lewis, and in 1941, they published a paper [19] in which they used the power series to derive the approximate formulas for the second and third harmonics of the waveform distorted by tracing errors. Normalized to the fundamental, they calculated their relative amplitudes as

$$D_2 = \frac{1}{4} kAk\rho, \tag{10}$$

$$D_3 = \frac{1}{8}(kAk\rho)^2, \tag{11}$$

showing that the Eq. (8) was right for the second harmonic but underestimated the amplitude of the third one by about 12%.[13] It also showed that, in many practical cases, the second harmonic would dominate the total distortion results.

By using the iterative numerical calculations to solve Eq. (5) and the 7- and 12-point DFT on Eq. (6), they further showed that the approximations from Eqs. (10) and (11) held quite well as long as the total harmonic distortion levels were reasonably small.

Lewis and Hunt also provided some very interesting results about distortions in the vertical vs. lateral cut records that will be addressed in SEC. 4, but they did not address the tracing loss on the fundamental frequency at all.

The final set of formulas for tracing distortion was published in 1949 by M. Corrington [20]. He used Lewis and Hunt's Taylor series approach but carried his calculations up to the seventh harmonic, evaluating between 5 and 7 first terms in the series for the amplitudes of the harmonics. Using the same notation $\mathcal{A}_n$ for the amplitude of the n-th harmonic as in Eq. (8), the first 3 components expressed by the first two terms of the series are as follows:

$$\mathcal{A}_1 = A\left[1 - \frac{1}{8}(k\rho kA)^2 + \cdots\right], \tag{12}$$

$$\mathcal{A}_2 = \mathrm{A}\left[\frac{1}{4}(k\rho kA) - \frac{1}{16}(k\rho)(kA)^3 + \cdots\right], \tag{13}$$

$$\mathcal{A}_3 = \mathrm{A}\left[\frac{1}{8}(k\rho kA)^2 - \frac{3}{32}(k\rho)^2(kA)^4 + \cdots\right]. \tag{14}$$

It is easy to see that $D_2 = \mathcal{A}_2/\mathcal{A}_1$ and $D_3 = \mathcal{A}_3/\mathcal{A}_1$ would match those in Eqs. (10) and (11) if only the first terms in Eqs. (12) to (14) are kept. More details about derivations from [19] and [20] can be found in APPENDIX A at the end of this paper.

The problem was probably considered solved at this point, and virtually nothing was published on tracing (and

[11] It might be interesting to mention that the famous conductor Leopold Stokowski participated in the discussion after Frederick's paper [14] was presented at the Society of Motion Picture Engineers meeting in Dec. 1931 in New York. Stokowski had a long history of cooperation with engineers from Bell Laboratories on improving systems for the reproduction of music from discs and over radio [16]. Among other things, on this occasion, he discussed the high-frequency limits necessary for the accurate reproduction of cymbal sounds.

[12] Di Toro's "exact" results were obtained by "harmonic analysis of the graphically determined traced curves." A similar graphical method was used by Olney [17], the first to calculate the tracking distortions, except that the Fourier coefficients were evaluated on an early analog computer – Di Toro did it manually, using the pre-FFT method of tabulated "schedules" [18].

[13] If the original papers are consulted, note that the results for our Eq. (8) to (11) were given for the velocity-sensitive cartridges, so there each result was multiplied by the order of the respective harmonic in them,

related tracking) distortion in the next almost 15 years. A pair of apparently independent papers in late 1962 and early 1963 [21,22] describing the previously unrecognized problem of vertical tracking error revived the interest in these questions. A series of new papers dealing with tracing and/or tracking distortion followed. About a dozen of these were authored by D. Cooper [23–33] in a span of about 3 years.

Cooper showed that tracing and tracking distortions can be viewed as "skew" transformations (see APPENDIX A for details). Despite many papers with often complex math, he did not derive any new formulas for tracing distortion. Quantitatively, he always used Eq. (10), neglecting all the harmonics after the second. Qualitatively, however, he showed that tracing distortion can be viewed as a phase (or frequency) modulation, with the modulating signal having the same frequency as the "carrier." This can be regarded as a case of "auto-modulation," where the signal modulates its own phase, with the modulation index

$$\beta = kAk\rho. \tag{15}$$

Cooper also showed that, under some simplifying assumptions discussed in APPENDIX A, the second harmonic of the tracing and tracking distortion would be in quadrature (orthogonal). This suggested that the powers of the distortion components would add up in the harmonic distortion, as well as in the eventual intermodulation distortion calculations.

It should be emphasized that all the formulas in this section so far were derived for groove displacements. All modern magnetic cartridges operate on Faraday's induction principle and are velocity-sensitive, i.e., their amplitude would be proportional to the rate of change of the engraved waveform, not the waveform itself. Because

$$\frac{\partial}{\partial t} A\cos(2\pi f t) = -A\,2\pi f \sin(2\pi f t),$$

this means that, in any distortion calculation, the amplitudes of the displacement components ($\mathcal{A}_n$ in our notation) should be multiplied by their respective order (n) first.

Since 1952, the signals from magnetic cartridges would also have been subject to the effects of the standardized RIAA equalization in preamplifiers.[14] This means that the fundamental and every distortion component would have to be adjusted according to the RIAA playback curve [8] per their actual frequencies to assess their magnitudes as they would be reproduced.

Another consequence of the velocity-sensitivity of cartridges is that in what follows, it will be more convenient to use the peak recorded velocity v_{max} as the parameter rather than the amplitude A of the groove displacement. Peak velocity is formally defined as

$$v_{max} = A\,2\pi f. \tag{16}$$

Again, for reasons that might be lost to history, v_{max} was always defined as a peak rather than a root-mean-square (RMS) value and practically always in cm/s units, even in the US.

[14]Before the RIAA standard, various manufacturers used many different equalization curves of their own design [34].

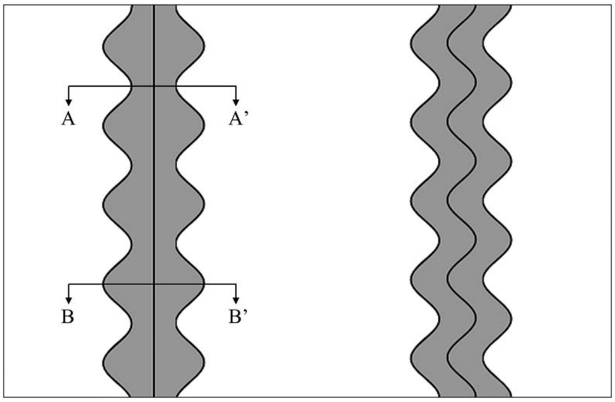

Fig. 6. Vertical (left) and horizontal groove cut (right) with the same recorded signal, view from the top.[15]

4 VERTICAL, HORIZONTAL, AND 45°/45° CUT

Thomas Edison's first phonograph cylinders, developed in 1877, used vertical cut (a.k.a., "hill-and-dale"), where the depth of the groove was proportional to the magnitude of the signal [35]. Less than a decade later, Emile Berliner, the father of the modern disc player ("gramophone"), patented a system with horizontal (lateral) indentation [36]. He argued that this, unlike the vertical method, would cause the stylus drag force to be constant, independent of the amplitude of the recorded signal, and introduced recordings in a disc format.

This started the first of many "format wars" that have plagued the audio industry ever since. It was fought on two fronts: vertical vs. horizontal cut and cylinder vs. disc. Berliner ultimately won on both; since about 1930, only discs with horizontal cut were produced.

Because vertical cut will be of interest when discussing stereo records, in Fig. 6, both cutting methods are illustrated. Looking at the cross-sections AA′ and BB′, as shown in Fig. 7, it can be seen that the stylus would always be at the distance of $\sqrt{2}\rho$ from the bottom of the groove. A cross-section along the center of the groove would look just like what was presented in Fig. 2, so one can conclude that the previously discussed distortion theory applies fully to the vertical cut case.

The situation with a horizontal cut is much more complicated, as depicted in Fig. 8. It is a redrawing of the original illustration made by Pierce and Hunt [6] and shows that the equality of the tangents at the point of contact would create "poid"-like curves on each side of the groove. At different

[15]The cutters' tip is usually rounded at the bottom (somewhat U-shaped rather than really V-shaped), so the edge at the bottom of the groove is not a visible line in either case. Showing the groves as they would appear if the cutter had a triangular vertical cross-section gives a better insight, however, and will be used in Fig. 6 to 8 and Fig. 10.

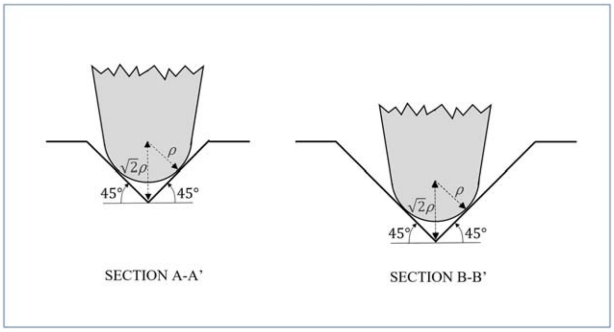

Fig. 7. Cross-sections as indicated in Fig. 6 for vertical cut.

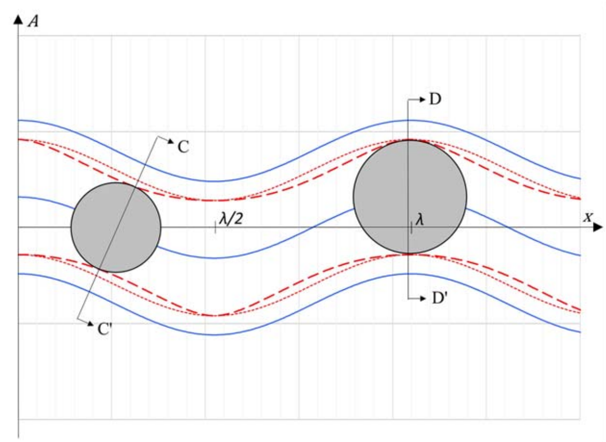

Fig. 8. Stylus contact point trajectories (dashed lines) and the hypothetical undistorted contact waveforms (dotted lines) on the sides of the laterally cut groove, top view. Gray circles are cross-sections of the stylus at the plane of contact.

points along the contact trajectory, however, the width of the area between contact points would be different, causing the stylus to rise in the parts with larger slope (CC′) and sink deeper into the groove when the slope is smaller (DD′). This was named the "pinch" effect, and the need to accommodate this vertical movement of the stylus when playing the laterally cut discs seems to have been a major surprise.

In 1938, when [6] appeared, most cartridges provided no degree of freedom for the stylus to move vertically. This caused excessive record wear and extra distortion. The problem was additionally complicated because, from Fig. 8, the stylus would go through two positive and two negative depth extremes per recorded wavelength. This meant that the vertical movements of the stylus would occur at twice the frequency of the waveform recorded horizontally. Because the piezoelectric and the massive magnetic cartridges available before World War 2 (WW2) had a very limited frequency range horizontally,[16] the requirement to track vertical movements over twice their horizontal band limit was very challenging at the time.

[16]Pre-WW2 disc lathe cutters had an upper limit of around 8 kHz, and the "sensational" FFRR (Full Frequency Range Recording" system introduced by "Decca" in 1945 had an upper limit of about 14 kHz [39].

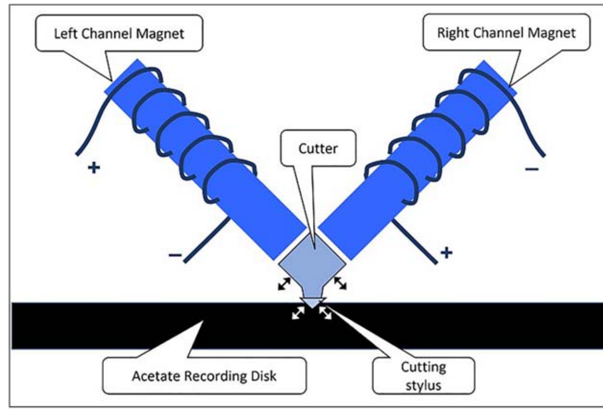

Fig. 9. Conceptual 45°/45° stereo cutter.

A possibly even more surprising conclusion that Lewis and Hunt reached in [6] was that with lateral cut, the *tracing error would produce no 2nd harmonic whatsoever, while the 3rd will be smaller than in vertical cut*. For small distortion values, Corrington [20] later found that the third harmonic in the horizontal cut would be one half of the value of the vertical cut, i.e., half of what Eq. (11) suggests.

The analysis in [6] for the horizontal cut is quite involved, an outline of which is given in APPENDIX B.1. But the conclusion about significantly lower distortions seems correct intuitively. From Fig. 8, the traced curves on two sides of the groove look quite symmetrical because the part seen as convex by the stylus on one side is seen as concave on the other side of the groove. The tracing errors are thus in counter phase and would cancel each other to a large extent, as the cartridge reacts to the movement of the center of the stylus in the horizontal plane. Fig. 8 is a very exaggerated example so that the variation of the vertical stylus positions can be seen to illustrate the point about "pinching."

One of the early ideas about introducing two channels on a stereo disc was to cut one channel vertically and the other horizontally. Another one, published on page 1 of the first issue of this Journal in 1953, was to cut two grooves for two channels separately in parallel [37]. Ultimately, the industry agreed on the solution where the left and right channels were cut at 45° angles, as depicted conceptually in Fig. 9. This system was standardized in 1958 but was based on the idea in UK engineer Alan Blumlein's patent filed way back in 1931 [38]..

This choice for the stereo cut was made to ensure backward compatibility. After a bit of math was sorted out, an arrangement could be made where the horizontal displacement of the groove would correspond to the sum of left and right signals (L+R); the vertical displacement would correspond to the difference (L-R or R-L),[17] depending on which of the two signals' polarity was inverted before cutting. (More details are in APPENDIX B.1). This way, the old

[17]Readers familiar with microphone techniques for stereo recordings might realize that a similar approach with sums and differences (often called "matrix") is used in the "M/S" method.

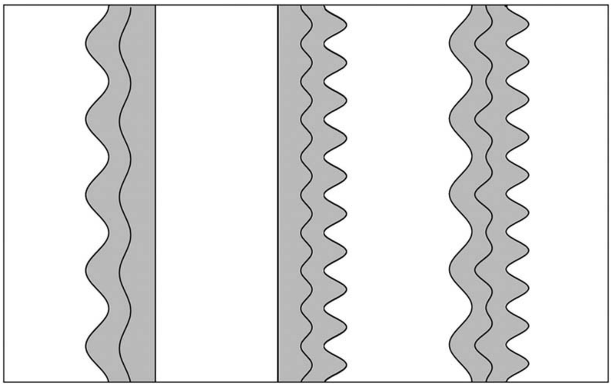

Fig. 10. From the left: stereo groove with the sinusoid in the left channel only, sinusoid of twice the frequency in the right channel only, and with both channels simultaneously (view from top).

monophonic cartridges that reacted to the groove's horizontal variations could play the new stereo discs correctly.[18]

The reproducing stylus can read the recorded pattern by ensuring that two sets of orthogonally positioned magnet-coil fixtures induce the voltages due to the movements in only one of the two 45° directions. Piezoelectric ("ceramic") stereo cartridges were also made and marketed in the 1960s. See [40] for an excellent summary of the early history of stereo cartridges.

According to eyewitness accounts [41,42], the technical standards committee of the Recording Industry Association of America (RIAA) received an in-depth analysis of the tracing distortion prepared by Corrington and Murakami at their December 1957 meeting when the stereo recordings were considered. This report was later published in [43] and showed that the *harmonic distortion in each of the 45° channels is the same as in the vertically cut discs*. Unlike the horizontal cut, stereo records would thus have the second and all other even harmonics.

This possibly surprising result could be explained with the aid of Fig. 10. If we imagine positioning our eye to align with either of the silent grooves in the figure, the other channel will be seen as a traditional vertical cut.

To deal with this distortion increase, the technical committee of the RIAA simultaneously recommended that the stylus tip radius for stereo should be reduced to 0.5 mils (12.7 μm), but in Europe, their engineering committee independently at about the same time, settled on 0.8 mils (20.3 μm), a value that some US cartridge manufacturers also preferred at that point. At the Electronic Industries Association (EIA) standards meeting held in January 1958, a compromise recommendation of 0.7 mils (17.8 μm) was adopted, a value that would appear for years in the dimension specifications for many styluses, although mainly for the dimension of a vertical cross-section.

5 DISTORTION RESULTS

In the horizontal tracking error case, the main distortion parameter (ε in [1]) is typically in the 0.01–0.02 range in all cases of practical interest. As a result, the distortion under any reasonable design is confined to the second harmonic and can be well approximated by the first term of a Taylor series for the Bessel functions [1].

The approximations in Eqs. (10) and (11) for tracing distortion work very well in many practical scenarios, but the equivalent parameter $\beta (= kAk\rho)$ from these expressions can assume much larger values under many combinations of realistic system parameters. For instance, with a velocity of 10 cm/s and stylus curvature radius ρ of 5 μm, in the last groove of an LP disc, the parameter β can be greater than 0.5 for frequencies $f > 7$ kHz.

At the recorded velocities of about 50 cm/s that some manufacturers claimed since the mid-1990s their cartridge could track,[19] such a condition would arise above 1.4 kHz, so the first-order approximations from Eqs. (10) and (11) can be somewhat inaccurate. Even the accuracy of the higher-order ones from Eqs. (12) to (14) can sometimes be problematic because the power series tend to converge very slowly in many cases of practical interest.

To illustrate this point, in Fig. 11, the distortion results calculated by solving Eq. (5) numerically and performing the FFT on the resulting waveform are shown, as already explained in Sec. 3. Curves were obtained for a fixed velocity of 10 cm/s, by calculating the frequencies which will give the resulting β on the abscissa. The results based on Eqs. (10) and (11) by Lewis and Hunt [19] are also shown, along with those obtained by Corrington, except that the fourth-order approximation from [20] was used, not the second-order versions from Eqs. (12) to (14).

Eqs. (10) and (11) work well for both harmonics at values, say, $\beta < 0.4$, overestimating the amplitude of the second harmonic by about 10% and of the third by about 20%. Corrington's fourth-order approximation performs better at the low end but diverges quickly at the high end. With larger velocities, both approximations diverge sooner, e.g., at 20 cm/s, discrepancies are already more than double those for 10 cm/s at $\beta \approx 0.1$. For velocities larger than that, one should probably rely exclusively on simulations. In some cases, even the fourth harmonic (also shown in Fig. 11) reaches values that cannot always be neglected. It probably would not be a surprise that in 1958 Corrington did not use his own formulas from [20] in his paper [43] on distortion in stereo reproduction but opted for numerical methods to obtain the required results on a computer.

[18] In practice, most old mono cartridges did not have enough vertical mobility to accommodate the variations encountered in stereo recordings. They were often found to damage the stereo discs, in some cases after just one playing. Stereo discs contained a warning on their sleeves about the dangers of playing them on old mono reproducers well into the 1970s.

[19] "Shure," one of the leading manufacturers of magnetic cartridges, introduced its iconic "V15" model in 1964. Throughout the years, many versions were produced, with ever-increasing "trackability" ratings. The last model, "V15 Type VxMR," made from 1996 to 2005, had a rating of 46cm/sec at 1 kHz [46].

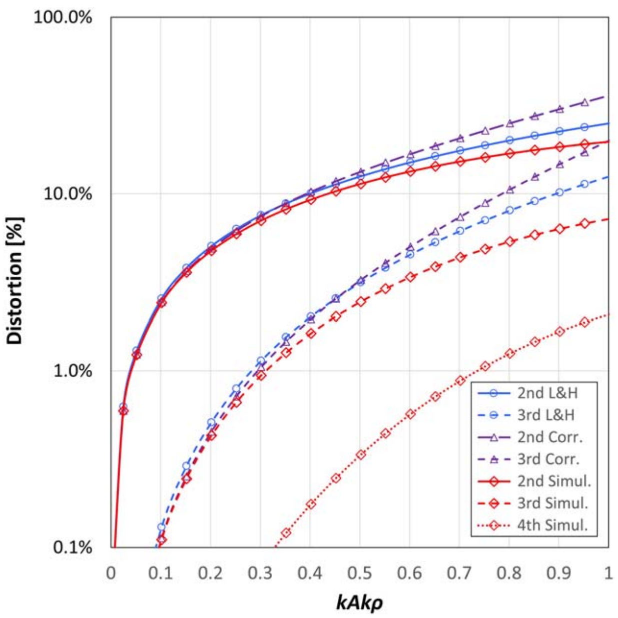

Fig. 11. Individual distortion levels for the second and third harmonic vs. $\beta = kAk\rho$ parameter at 10 cm/s recorded velocity (with 5-μm stylus, in the last LP groove), displacement-sensitive cartridge.

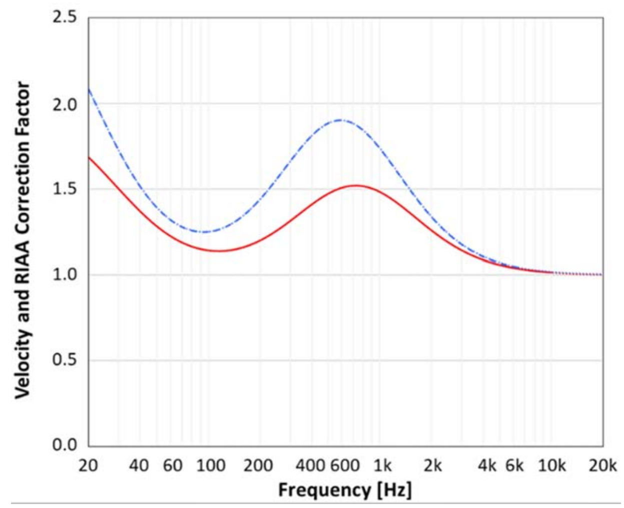

Fig. 12. Multiplicative factors due to velocity cartridges and RIAA equalization for the second (solid) and third (dashed line) harmonic vs. frequency of the fundamental.

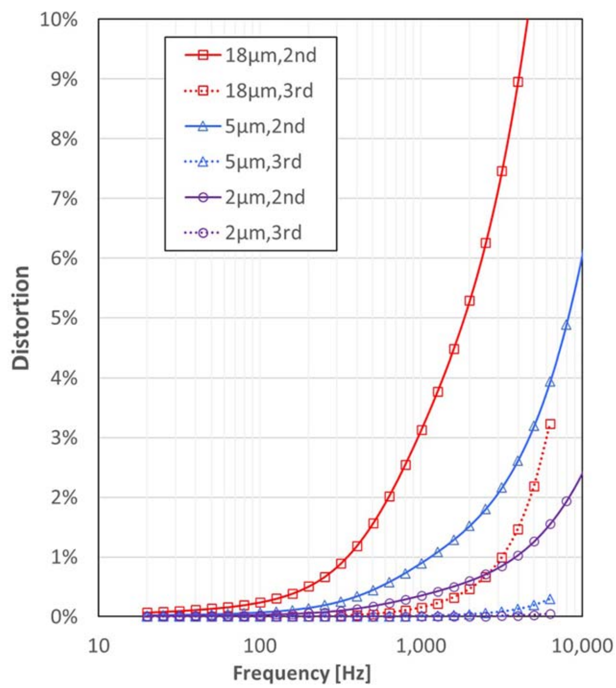

Fig. 13. Simulation results for second and third harmonic distortion in the middle groove (radius 10.3 cm) of an LP with stylus radius sizes of 18 μm, 5 μm, and 2 μm; 10 cm/s recorded velocity, velocity-sensitive cartridge, and RIAA equalization

While computers were available in 1958 when [43] appeared, the FFT algorithm was still not widely used in engineering. Although at least some elements of the FFT algorithm were developed by C. W. Gauss around 1805 [44], it was not well known and certainly not used on computers until Cooley and Tukey's paper [45] from 1965. This means that Corrington and Murakami in [43] likely still calculated the spectra using the DFT.

It should be remembered that the numeric results for the harmonics of the resulting signal $y_{out}(x)$ from Eq. (6) relate to the displacement of the stylus center. For magnetic cartridges, the output signal is proportional to its velocity, so that each component's amplitude must be multiplied by the order of the harmonic.

Furthermore, this output signal would be subjected to RIAA equalization, which boosts the high-frequency signals during the recording and attenuates them during playback. The RIAA equalization curve, however, is not linear, so the fundamentals of different frequencies would be attenuated differently.

In Fig. 12, we show how the second and third harmonics of the displacement signal would be amplified by the combined effects of the velocity cartridge and the RIAA equalization. Although the equalization curve prescribed by RIAA in [8] extends indefinitely, in practice, the frequency response of the cutters and audio amplifiers was often suppressed rather sharply above 20 kHz. To avoid further assumptions about that cutoff, in Fig. 13 and Fig. 14, we show the second harmonic results for the tones with fundamentals under 10 kHz and for the third under 6.7 kHz only.

Both figures indicate very high distortion values with a relatively moderate velocity of 10 cm/s,[20] and they will increase roughly proportionally at higher velocities. These numbers might be surprising for many audiophiles and possibly for some audio specialists too.

[20]To put these values into a bit of historical perspective, the German DIN 45500 standard from 1966 prescribed different requirements for various Hi-Fi components, with the maximum allowed distortion values usually around 1%.

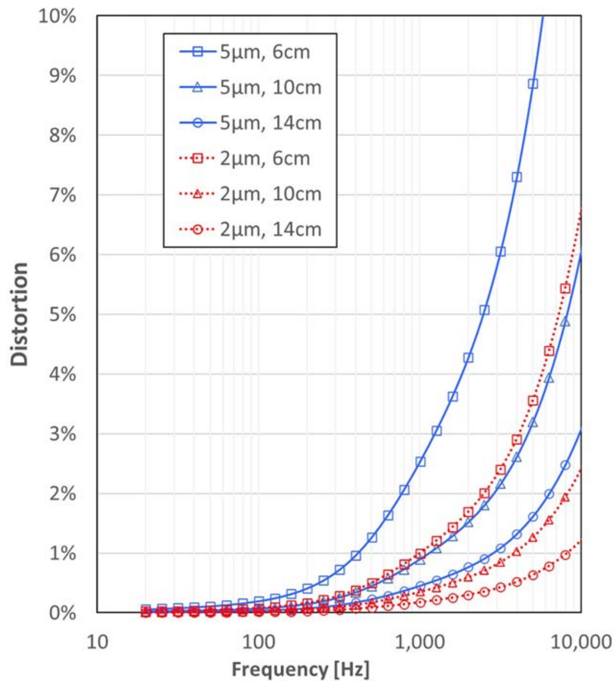

Fig. 14. Simulation results for second harmonic distortion in the first (14.6 cm), middle (10.3 cm), and last (6.0 cm) groove of an LP with stylus radiuses of 5 µm and 2 µm with 10 cm/s velocity, velocity-sensitive cartridge, and RIAA equalization.

The curves with very similar results, however, have been published previously (e.g., in [6] and [15]), but most later papers usually gave them as a function of the linear groove velocities (ΩR in our notation), not the groove radiuses. This was done to address the 33⅓, 45, and even 78 RPM speeds, and all different record sizes that existed at various times when these papers were prepared, so that the effects of tracing errors were a bit obscured.

Given the results in Fig. 13, it is not surprising that cartridge manufacturers strove to produce styluses with smaller curvatures than the initially recommended 18 µm radius as soon as stereo discs appeared. Looking at Fig. 14, however, it can be seen that even with the modern styluses of around 2 µm radius, the distortions at higher frequencies are at the levels most audiophiles consider unacceptable, especially as the groove radiuses decrease toward the end of the discs. The earliest researchers, Di Toro [15] and Pierce and Hunt [6], also performed experimental verification of their vertical and horizontal cut results and found that the measurements matched their theoretical results quite well in both cases.

This was somewhat surprising because most other sources of distortion that could have impacted their test results were not well understood at the time. They avoided this by selecting the system parameters that gave atypically large distortions, in the 10%–50% range, much higher than what the tracking and elastic groove deformation errors were likely to create. The agreement at the distortion levels below 10% was noticeably worse in [6], with measured results exceeding the theoretical values by a considerable margin, probably due to these other distortion sources.

Few other experimental results for harmonic distortion were reported in the technical audio literature or in popular magazines for audiophiles. In technical journals, when audio engineers needed to test tracing distortion (for instance, to compare results with and without the devices to reduce tracing distortions from Sec. 6), tests of intermodulation distortions (IMD) with two tones [47] were generally preferred because they could more easily create numerically large values, easier to assess correctly in the presence of the other distortion mechanisms.

To the best of this author's knowledge, the cartridge manufacturers never listed any distortion results (THD nor IMD) in any technical specifications for their products. The only example of reporting the THD in the popular literature he is aware of is in a 1964 paper [48], where the authors presented their first results for the "elliptical" styluses. They showed that with vertical cut, a 0.2 mils stylus would reduce the second harmonic by 20% and 40% with respect to the 0.5 and 0.7 mils styluses but still showed the THD results between 5% and 6% for 0.2 mils radius. They were obtained with a 21 cm/s test signal at 1 kHz frequency and 24 cm/s at 2 kHz, but the groove radiuses and the disc rotation speed were not specified.

6 CORRECTION OF TRACING ERRORS BY PREDISTORTION

The work on eliminating or at least reducing tracing distortion started almost as soon as stereo records were standardized [41]. The general idea was to introduce a predistortion in the signal driving the cutter for the original acetate from which the masters and stampers were made so that the distortions were reduced when the actual vinyl discs were played back by the reproducing stylus.

Interestingly, the first suggestion on how to achieve this was made almost as soon as the severity of the problem was first recognized. The idea was brought up in the spring of 1938 at the Society of Motion Picture Engineers (SMPE) meeting in Washington, DC, where Pierce and Hunt presented their paper [6].

Immediately after Pierce's presentation, W. MacNair, an SMPE Fellow, suggested that tracing distortion could be reduced if the original acetate disc was traced with a standard reproducing stylus and the signal from the cartridge, with its polarity inverted, used to cut a new acetate.[21] He argued that this "upside down" signal will be reproduced as the original sine signal, if played back with a stylus of the same size (except that its original polarity will be reversed). The explanation sounds complicated, but the concept is easy to understand with the aid of Fig. 15. The mathematical analysis becomes quite involved when the input signal is not sinusoidal, however.

[21] After the papers, the technical journals in those times regularly published the transcripts of the discussions held at the conferences after the paper was presented. Footnote 9 about conductor Stokowski's comments on Frederick's paper [14] was taken from such an appendix, and MacNair's suggestion was published as an appendix in [6].

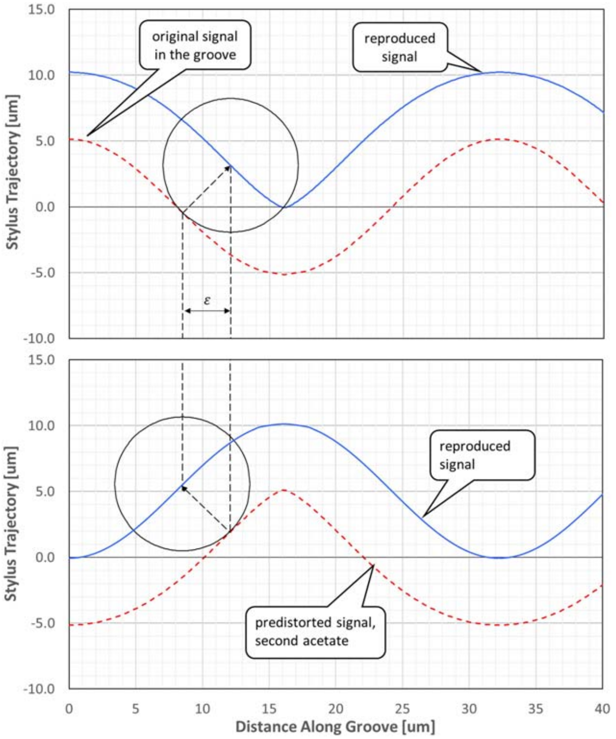

Fig. 15. McNair's predistortion method for tracing error cancelation. The dashed lines are engraved, and the solid ones are the traced waveforms (the reproduced waveform at the bottom is the traced one from the top, with inverted polarity).

E. Fox and J. Woodward in 1963 summarized the results of the work done in "RCA Labs" since the introduction of the stereo disc [49]. MacNair's idea with rerecording was shown experimentally to reduce the distortion components to about half of the original values with a combination of two sinusoidal signals as the input (IMD), but it doubled the acetate noise and was considered too cumbersome for use in commercial recordings.

As a practical alternative, the "RCA" team developed an electronic solution based on an analog delay line that could approximate the curve to be traced based on the best fit among 12 discrete possible contact points with the stylus. The device could not do its calculations for continuously varying groove radiuses either, so it provided options for 7 discrete radiuses to choose from [49]. It was an early 1960s technology that looks incredibly clunky by current standards. Still, it did its job quite well, reducing the most dominant IMD components by a factor of at least two in the worst-case scenarios, near the end of the disc.[22]

The "RCA" apparatus was named "Dynamic Recording Correlator" for reasons not completely clear. "Teldec," a joint venture of "Telefunken" and "Decca," also studied such an apparatus in 1965 [50], whereby the second harmonic distortion signal was synthesized and subtracted from the signal going into the cutter. Another solution, based on the ideas that D. Cooper covered in several of his papers, also went through some early stages of development [32], but it is not clear if it was ever commercialized.

A team from "Neumann," a famous manufacturer of lathe cutters from Germany, built a practical tracing error correction system based on the "Teldec" solution [50] by the late 1960s, using more advanced analog electronics [51]. Devices for tracing distortion elimination were also developed in Japan by teams from "Toshiba" in 1970 [52] and from "Nippon Columbia" in 1973 [53]. "EMI" from the UK reported one implementation using an analog shift register ("bucket-brigade" circuit) in 1977 [54].

Despite promising results from around the world, these methods appear to have hardly ever been used in commercial disc recordings. Part of the reason was an overwhelmingly negative reception that "RCA's" new "Dynagroove" system, which incorporated the "Dynamic Recording Correlator," received in the popular press when it was introduced in the USA in 1963.[23] For reasons that were not always clear, a significant portion of the scorn was directed at the "Dynamic Recording Correlator" itself.

Despite considerable fanfare and a launch that included an inaugural paper on the "Dynagroove" system published in this Journal [56] written by none other than H. Olson,[24] the system was retired within a few years only. "Neumann" offered the tracing distortion reduction module "TS 66" as an option in at least one of their amplifiers for driving the cutters, but it seems to have been used very sparingly in commercial recordings.

While subjective evaluation could have gone both ways, the objective problem for all these tracing predistortion systems was that they could adjust their algorithms for only one stylus curvature radius (in "Dynagroove" 18 μm [49] and 15 μm in "Neumann" [51]). By the time they arrived, styluses were already made in several different radius sizes (and shapes), with most higher-grade ones soon having radiuses smaller than either of these two values. The other problem was that the solutions possible with analog electronics from the 1960s and 1970s could not be very precise and—at least in "Neumann's" case—required a rather cumbersome periodic calibration [57].

[22]Complete removal of the distortions would not be possible even with a perfect implementation because elastic deformation of vinyl under stylus pressure cause the traced path to be different than calculated on a perfectly stiff groove surface [61].

[23]The campaign against "Dynagroove" was exceptionally feverish in the popular "Stereophile" magazine [55]. J. G. Holt, the magazine editor and a big advocate of subjective listening tests, seemed to have been on a personal crusade against anything that would alter the purity of the original recordings. He also strenuously objected when the IEC introduced an additional cutoff at 20 Hz into the RIAA's original playback equalization curve to deal with subsonic components due to record warps.

[24]Harry F. Olson was the head of the "RCA" laboratory for acoustic research, a celebrated author of many text and reference books in audio engineering, one of the founders and a president of AES, editor of this Journal from 1966 to 1969, a recipient of its John Potts Medal (now AES Gold Medal) award in 1949, member of the National Academy of Science, etc.

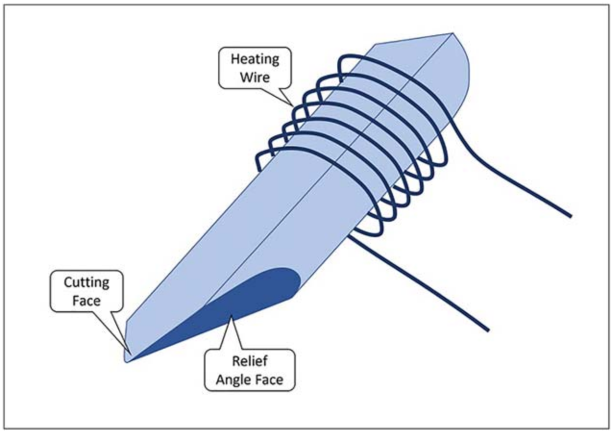

Fig. 16. Shape of the cutting stylus.

7 MAXIMUM RECORDED VELOCITY

The dynamic range of LP discs is frequently discussed on internet forums in often heated debates on the advantages and disadvantages of analog vs. digital recordings. The range obviously depends on the noise levels LPs have and the maximum signal level they can support.

What is usually called “vinyl noise” in popular audio literature covers several noise sources in addition to those coming from the microvibrations of the stylus due to the nonhomogenous, granular molecular structure of the vinyl itself. It is an interesting subject in its own right, but it is beyond the scope of this paper.[25]

The maximum possible recorded velocity, however, is limited by three factors recognized as far back as 1960 by J. Stafford [60]. All three are related to the stylus and groove geometry; one of them is tracing distortion.

The first and easy-to-understand factor is the maximum amplitude that can be recorded in the groove. It has to be small enough to ensure that the neighboring grooves do not touch each other. In SEC. 2, a quick calculation was provided that suggested that the maximum amplitude in typical recordings $A_{max} \approx 50$ μm, so from Eq. (16)

$$v_{max} \leq 2\pi f A_{max}. \tag{18}$$

Maximum possible velocity would thus be increasing with the frequency, and we can easily adjust the curve for any other value A_{max} that shorter-lasting LP sides with larger inter-groove spacings would allow.

The second constraint is a bit more challenging to explain, but it is related to the shape of the cutting stylus, whose cross-sections are always triangular, as depicted in Fig. 16. For some reason, these sides in practice were at 45° (a.k.a. “relief angle”) relative to the front, cutting side.

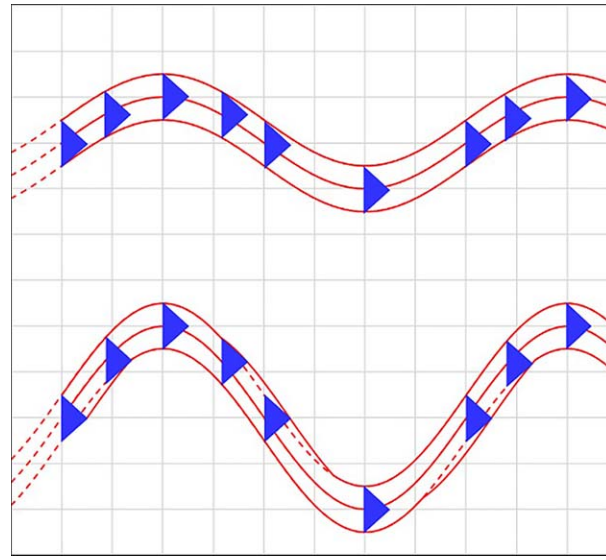

Fig. 17. Cutting stylus trajectory without (top) and with the “slope overload” distortion (bottom). Horizontal cut, view from above.

(A smaller angle of 35° appeared to have been used for some quadrophonic recordings only, like CD-4 [62].)

If we keep the amplitude of the recorded signal constant and increase its frequency, at some point, the cutter’s back sides will start interfering with the recorded groove sides. Fig. 17 depicts a horizontal cut, but it is easy to see that the same mechanism would exist in vertical and stereo recordings. This is called “slope overload” and will occur when the slope of the signal becomes larger than the relief angle. If we let α denote the relief angle, and because the slope of the recorded signal is its derivative, from Eqs. (2) and (16), we will thus have a requirement

$$\left| -\frac{v_{max}}{\Omega R} \sin\left(\frac{2\pi f}{\Omega R}\right) \right| \leq \frac{v_{max}}{\Omega R} \leq \tan(\alpha),$$

or

$$v_{max} \leq \Omega R \tan(\alpha). \tag{19}$$

This constraint is independent of frequency.

A third constraint introduced in [60] was based on the tracing requirement. Stafford noticed that the stylus would not be able to trace the engraved signal if the maximum curvature radius of the signal became smaller than the stylus radius ρ. Fig. 18 depicts a situation when the stylus cannot reach the bottom part of the curve around $x = 10.5$ μm because of this condition.

The radius of curvature $r(x)$ of any twice differentiable function $y(x)$ is given by

$$r(x) = \frac{\left|1 + y'(x)^2\right|^{\frac{3}{2}}}{|y''(x)|}. \tag{20}$$

[25] The latest technical papers this author could find on vinyl noise are more than 50 years old [58, 59]. Various measurements can be found on the internet nowadays, but they are of somewhat limited usefulness. Generally, they do not indicate the resolution bandwidth or FFT size, disc speed, or groove radius. More often than not, they do not offer any reference point for their measurements (e.g., level for 5 cm/sec velocity) either.

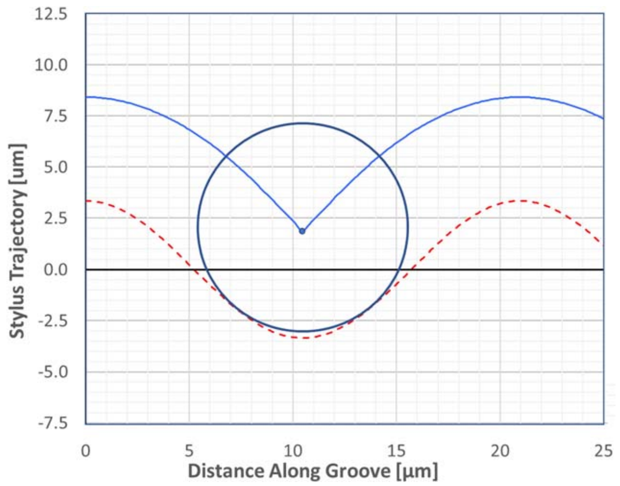

Fig. 18. An example where the curvature of the groove is smaller than the stylus radius. Parameters are the same as in Fig. 2, except that the recorded frequency is $f = 10$ kHz.

Using Eq. (2) in (20) and carrying out the necessary calculations gives

$$r(x) = \frac{\left|1 + \left[\frac{v_{max}}{\Omega R} \sin\left(\frac{2\pi f}{\Omega R} x\right)\right]^2\right|}{\left|\frac{2\pi f v_{max}}{\Omega^2 R^2} \cos\left(\frac{2\pi f}{\Omega R} x\right)\right|}. \quad (20)$$

Minimum radius values r_{min} will appear when the argument of both trigonometric functions is an integer multiple of π, because then the cosine term in the denominator is the largest and the sine term in the numerator the smallest. This value r_{min} has to be larger than the stylus radius, i.e.,

$$r_{min} = r(n\pi) = \frac{\Omega^2 R^2}{2\pi f v_{max}} \geq \rho, \quad (21)$$

and from Eq. (21), the third constraint on maximum velocity is obtained as

$$v_{max} \leq \frac{\Omega^2 R^2}{2\pi f \rho}, \quad (22)$$

which obviously decreases with frequency.

In [60], Stafford combined Eqs. (18), (19), and (22) in a graph like the one in Fig. 19. The sloped line on the left is due to the amplitude, horizontal lines in the middle to the slope overload, and the downward sloped ones on the right to the tracing curvature constraint.

Assuming the same or slightly different values of the pertinent parameters, graphs like this would be reproduced in many technical and popular papers for years [9, 63]. It was usually done when discussing "trackability," a parameter that authors from "Shure" evangelized as an answer to a simple question—which velocities are cartridge styluses supposed to be able to track?

A serious problem with the graphs like the one in Fig. 19 is that the limit on the right-hand side assumes that the groove curvature radius has to be no less than the stylus curvature radius. This would ensure that the stylus would not lose contact with the bottom of the groove indeed, but Di Toro's analysis way back in 1937 showed that, for reasonable distortions, the minimal groove curvature has to be several times larger than the stylus radius.

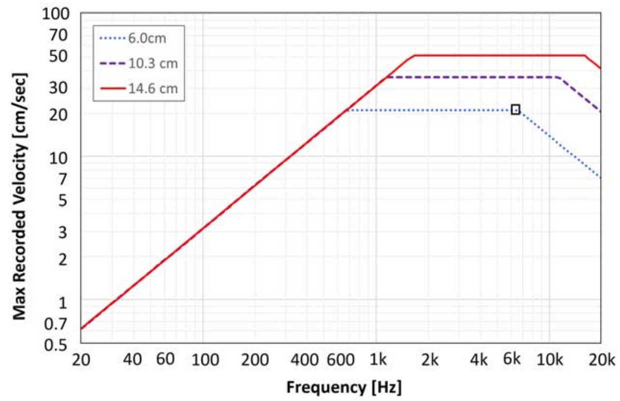

Fig. 19. Maximum possible velocities on an LP based on the analysis by Stafford [60] (relief angle α = 45°, stylus radius ρ = 5 µm, and maximum recorded amplitude A_{max} = 50 µm).

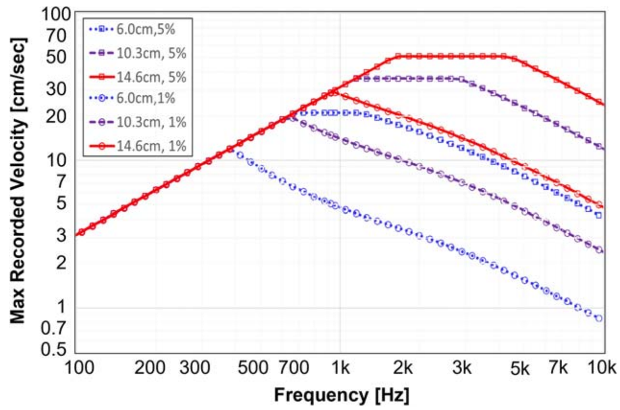

Fig. 20. Maximum possible velocities on an LP with the same parameters as in Fig. 18 for 1% and 5% harmonic distortion.

Consider the point marked with a square in Fig. 19, at the corner where slope overload and curvature lines intersect. It corresponds to the velocity of 21 cm/s at a frequency of 6.6 kHz in the inner groove of an LP with a radius of 6 cm, which turns out to be precisely the parameters from our example in Fig. 2. Very few audio experts would be likely to consider the traced curve from that figure to be an acceptable proxy for the sinusoidal input signal. THD in that example is actually a whopping 19.60% (the second harmonic is at 18.46% and the third at 5.61%; even the sixth harmonic is above 1%).

If more realistic limits on the allowed distortion are imposed, possible velocities will be much smaller. An example of such a calculation, assuming that the worst second harmonic distortion should be at 1% or 5%, is presented in Fig. 20. The first to create such plots was Cooper in 1963 [25]. Still, the industry by and large continued to use the plots as in Fig. 19.

We can see that in the 1% case, the recorded velocities can never reach the horizontal limit due to slope overload constraints. The actual maximums are at about 12, 19, and 29 cm/s in the innermost, middle, and outermost groove of an LP, respectively. With a 5% limit, the results would be different, but ultimately the question of what maximum velocity can be reproduced reliably on an LP depends on how much distortion we are willing to accept at which frequencies and at which groove radiuses.

Obviously, an important question to assess the extent of the potential distortion would be what kind of recorded velocities can be found on LPs. A study "on a large collection of 'difficult to track' records" reported in 1973 [9] showed that out of a few hundred samples of the most "difficult-to-track passages," the recorded velocities were larger than 50 cm/s in 5 cases, all having the "dominant frequency" in the 4 to 10 kHz range.

Another often referenced data point is from the 1977 paper by T. Holman [64]. Citing another paper, which turns out to be the minutes from a meeting of an audio club in Boston [65], this author reported that a maximum recorded velocity of 105 cm/s had been measured at a frequency of about 7 kHz on a jazz record issued in 1963.[26]

Neither of these publications explained anything about the measurement methodology, but the number 105 cm/s can often be found on the internet when the dynamic range of an LP is discussed. It even appeared in at least one published paper [66], purported to be the maximum level that can appear on an LP.[27]

Even if the number 105 was measured correctly,[28] it is probably clear by now that this signal could not have been tracked with the cartridges available in the 1970s. And even if the stylus could stay in the groove, the distortion would have reached atrocious values when replaying this signal. The question of the maximum velocity that can be reproduced reasonably well from an LP cannot be answered simply, but the 105 cm/s value does not appear to be the one to use in any further analyses.

8 AUDIBILITY OF TRACING DISTORTION

Some distortion results shown here are very high and might look outright unbelievable to many readers with any experience in vinyl. Still, the theory, simulations, and measurement results all suggest the results are correct.

If that is the case, however, one cannot but wonder why the reduced distortion was never mentioned as a reason for a switch from vertical to horizontal cut in the early 20th century. Or why there was no backlash when stereo discs were introduced in the middle of that century, with their considerable increase in the second harmonic distortion relative to their mono predecessors.

Some of the early authors appeared to have been surprised themselves by the sizeable distortion levels their calculations gave. For instance, Lewis and Hunt [19] noted that the discs, even in the 1940s, sounded better than what their results for THD and IMD would suggest.

One argument made early, and occasionally mentioned in the popular audio press, especially during the golden age of Hi-Fi in the 1970s and 1980s, was that we do not perceive these distortions badly because they are almost exclusively in the second harmonic. Although it seems correct that the second harmonic is the least bothersome to many listeners (and some audiophiles like it enough to go the extra mile to acquire amplifiers with nonlinearities that would add it in abundance), it can definitely be audible.

Regarding the later stereo recordings, another argument sometimes mentioned was that tracing error introduces the predominant second harmonic distortion in the vertical plane cut, i.e., in the (L-R) signal, and not in the horizontal (L+R) one. Because we are hearing both channels simultaneously in stereo, it suggested to some that the distortion would not be impacting the sum we hear, but mainly the difference signal, which would in turn distort the stereo picture only.

Another way to think of this argument is that most of this distortion would disappear if we switched the amplifier to the mono mode to reproduce (L+R) signal from both channels, an option that existed on many amplifiers in the 1970s and 1980s. Cooper argued against this assertion in the closely related case of the vertical tracking error in 1963 [70], but no subjective tests appear to have ever been performed to verify either viewpoint.

It is probably not surprising that Jacobs and Wittman wrote a full-length paper [71] in 1964 devoted exclusively to this subject, titled "*Psychoacoustics, the Determining Factor in Stereo Disc Distortion*." Trying to explain why discs do not sound nearly as bad as the distortion values indicate, they first argued about several—always at least to some extent controversial—points, like the listener's ear training, masking of the harmonics, and the importance of distortion figures for stationary vs. transient waveforms.

More aptly, they reported that their critical listening tests showed that it was impossible to distinguish between the original master tapes and the signals picked up from the

[26] The LP is "Hey! Heard the Herd" by Woody Herman, issued by *"Verve Records"* (V-8558). It seems to be an acoustic jazz recording with an electric organ [67]. It is somewhat puzzling how such a high velocity could be present on a music piece with such a lineup, and how it could have been recorded with the 1963 technology.

[27] Bauman in [66] then went on to claim that with this maximum velocity, the dynamic range of LPs can exceed that of digital CDs. In this analysis, however, he considered the thermal noise of the preamplifier as the only source of noise in LP reproduction, disregarding all the components of the vinyl noise at its input, which will be amplified just like the desired signals.

[28] Neither of the papers gives any details on how these velocities have been measured. They were above the tracking ability of the cartridges in the 1970s, so direct cartridge output signal level measurements would probably not have worked. Other methods to measure the recorded velocity at the time [68] included measurements of the optical light patterns after the disc was illuminated with a specially positioned light source, microscopic photography, and "variable speed" measurements, which in this case probably would have involved running a turntable at lower rotational speed so that the calibrated test cartridge would stay within its tracking ability [69]. Neither method worked well at higher audio frequencies [68], even with sinusoidal signals, and would have had much more problems with more complex musical ones. It is even harder to imagine how the "dominant frequency" on a very short musical passage could be reliably evaluated with the technologies available in the 1970s.

records made using the same tapes. The THD added in the reproduction of the discs they used in these tests was measured at various levels between 2% and about 8%. However, their version of this subjective A/B testing, without any reported statistical results and no detailed description of the test conditions, would probably not be accepted as reliable by today's standards.

Cooper, on the other hand, after showing that tracking distortion can be viewed as phase modulation (PM) [23], suggested this as a possible explanation for the better than expected listening experience. Tracing distortion, a form of PM, could thus be less bothersome because of our limited sensitivity to phase inaccuracies.

The original hypothesis of G. Ohm (of the Ohm's law fame) that our hearing is based only on the power spectrum, not on the relative phases of the spectral components, was disproved long ago. Still, it turns out that the phase distortions are, in the words of Lipshitz, Pocock, and Vanderkooy [72], "generally inaudible." (These authors also cite numerous references on the subject, starting with Ohm's original paper.) This conclusion, however, was reached for the cases when the overall phase characteristic is nonlinear, not when the test signal is phase modulated.

As Cooper noted in 1963 [23], the papers that reported the subjective testing results on the audibility of PM were few and far between, and this seems to be the case to this day. The situation is somewhat better in the case of frequency modulation (FM), whose audibility has been studied more, partly because of the wow and flutter. This refers to the fluctuations of the turntable's speed that create an unintentional FM of the reproduced signals, as the vibrato on strings and vocals does intentionally in music.

With sinusoidally modulated waveforms, the PM and FM will be of the same form (hence the joint name angular modulations), with the equivalent FM carrier peak frequency deviation

$$\Delta f_c = \beta\, f_m, \tag{23}$$

where β is the modulation index for PM from Eq. (15), and f_m is the frequency of the modulating signal.

Again, at least for low indexes β, tracing distortion can be viewed as a case of phase automodulation, where the carrier frequency f_c is equal to f_m, the modulating frequency (i.e., when the reproduced tone is at f_c= 1 kHz, we would have the modulating frequency f_m= 1 kHz).

Most of the published research on the audibility of FM modulation so far dealt with modulating signals whose frequencies f_m were in a much lower range, e.g., $f_m \leq 64$ Hz [73], with most of the other results limited to $f_m \leq 10$ Hz (see [74] and references therein). These results showed that listeners are most sensitive to the modulations with f_m in the 2–4-Hz range. The sensitivity falls quite quickly on both sides, but there appears to be no curve proposed to extend the results beyond these ranges.

It is tempting to use the filtering characteristics from the standards for the weighted wow and flutter measurements to estimate how much the higher modulating frequencies would be attenuated, because the weighting filters are supposed to correlate with our perceptual experience. In the most widely used standard for wow and flutter measurements [75], the filter is centered at 4 Hz and then attenuates at the 6 dB/octave rate. Hence, the weighted wow and flutter filters attenuate components at 125 Hz by 20 dB, but it is undefined above 200 Hz.

Extending this curve to, e.g., 1 kHz, however, seems unjustified. The standard states explicitly that the method is devised to measure frequency deviations up to 100 Hz only and states that components above 100 Hz (dubbed "scrape flutter") can be heard as "a noise added to the signal." Whatever that is supposed to mean.

Another significant result from the tests related to wow and flutter is that the "just noticeable" peak frequency deviation, as judged by the listeners in controlled tests, tends to vary with the frequency and the sound pressure value, but is generally around 0.1% of the carrier frequency f_c [76]. The inaudibility requirement would be

$$\frac{\Delta f_c}{f_c} < 0.001. \tag{24}$$

This is the value that became the basis for the wow and flutter requirements for professional and high-quality audio equipment like turntables and tape recorders.

If similar results would extend to the larger modulating frequencies f_m in the audio range, then based on Eq. (24), any tracing distortion with $\beta > 0.001$ would be audible since in our case $f_c = f_m$. Unfortunately, we saw that values of β several orders of magnitude larger than this can occur under normal LP reproduction. Based on the available test results with low f_m frequencies and wow and flutter requirements, it seems impossible to tell if and by how much this $\beta < 0.001$ requirement might be relaxed by our higher sensitivity thresholds at higher f_m values.

Interestingly, the only results on phase distortions with the $f_c = f_m$ condition that this author is aware of were reported by Tollerton in 2009 [77] and included a digital simulation of the distortions due to horizontal or vertical tracking errors. Even though Tollerton created portable digital files with various music and test tones inputs under varying levels of distortion, the listening tests in [77] were performed with two subjects only. Unfortunately, the plans to repeat these tests with a larger number of test subjects and to extend them to the tracing distortion apparently never materialized.

9 SUMMARY AND CONCLUSIONS

In this paper, we explained the tracing error mechanism and provided a historical survey of the theoretical and experimental results for the distortion it produces. Differences between vertical, horizontal/lateral, and stereo 45°/45° were described, as well as the methods to alleviate tracing errors that were considered throughout the years.

Unlike some older work, the paper also presents distortion results for LP records that are straightforward to interpret. Many of these values will likely seem unbelievable to most audiophiles and possibly to some audio professionals—some of these analyses and results are so old that it often seems almost everybody has forgotten them by now.

Based on the simulation results, one of the conclusions of this paper is that the approximations from Eqs. (10) and (11), widely used in past analyses of this problem, can seriously overestimate the actual distortion. For any recorded velocity larger than perhaps 10 cm/s, great caution should be exercised in using them. Relatively straightforward numerical methods described in SEC. 3 seem a safer bet and might be a preferable approach to use nowadays.

Another conclusion of this paper is that we do not appear to have a clear understanding of the audibility of the tracing (and tracking) distortion. Several theories about why the LPs sound much better than the tracing distortion results suggest are touched upon, but it seems that none provide a satisfactory explanation.

Papers providing historical surveys of the problems first recognized 90 years ago are probably not expected to come up with any significant recommendations for future work. Still, it seems obvious to this author that much more work is needed in subjective testing of the phase/frequency-modulated signals with the modulating frequencies over the whole audio range, not just in 0–10 Hz or 0–60 Hz.

Very precise digital audio files with various degrees of distortion due to tracing and tracking errors can be easily generated today. Proper tests can then be executed on large enough sets of prequalified test subjects in labs equipped for such work. The files to be evaluated can include music and samples with test tones having different values of tracing and tracking distortion and can come in mono and stereo versions so that many audibility questions might be answered.

If this and several recent papers on tracking distortion generate enough interest among readers, another conclusion would probably be that similar survey papers should be prepared on the other sources of imperfections in vinyl LP replay. Prime candidates are distortions due to vinyl deformation, vertical tracking error, and anti-skating problems.

Another survey article that many readers might find interesting could deal with IMD in general, and IMD due to tracking and tracing errors in particular—many audio specialists seem to have forgotten the intricacies of these tests. In the case of tracing and tracking errors, the IMD might be of particular interest because the distortion levels involved are much higher than the IMD levels produced by amplifiers with similar levels of harmonic distortion.

Yet another question worth looking at would probably be vinyl noise. The energy (and sometimes the vitriol) involved in addressing the dynamic range of LPs vs. digital on the internet seems inversely proportional to the availability of reliable measurements, the last of which appeared more than 50 years ago.

An exciting option to consider now would be to try to figure out how to remove tracing distortion not on the recording but on the *replay* side. Present technologies might allow us to achieve this by performing an inverse of the "skew" transformation. If this can be accomplished, there should be no problem using the radius of the stylus as a parameter in the calculations that such systems are to perform, so that the cancelation algorithm can be tailored to the owner's actual cartridge used in the replay.

One of the reviewers pointed out that a solution along these lines has already been implemented in software by "Pspatial Audio," but the available documentation ([82], pp. 357–358) does not describe the details of their "deskew" algorithm.

10. REFERENCES

[1] V. M. Jovanovic, "New Analytical Results for Löfgren C Tonearm Alignment," *J. Audio Eng. Soc.*, vol. 70, no. 3, pp. 199–207 (2022 Mar.). https://doi.org/10.17743/jaes.2021.0061

[2] A. P. Hickman, "Analysis of Löfgren's Tonearm Optimization," *J. Audio Eng. Soc.*, vol. 71, nos. 1/2, pp. 53–61 (2023 Jan.). https://doi.org/10.17743/jaes.2022.0062/

[3] J. A. V. White and A. R. J. Gust, "Measurement of FM Distortion in Phonographs," *J. Audio Eng. Soc.*, vol. 27, no. 3, pp. 121–133 (1979 Mar.).

[4] F. Rumsey, "The Vinyl Frontier," *J. Audio Eng. Soc.*, vol. 62 nos. 7/8, pp. 559–562 (2014 Jul.).

[5] J. G. Frayne, "History of Disk Recording," *J. Audio Eng. Soc.*, vol. 33, no. 4, pp. 263–270 (1985 Apr.).

[6] J. A. Pierce and F. V. Hunt, "Distortion in Sound Reproduction from Phonograph Records," *J. Soc. Motion Picture Eng.*, vol. 31, no. 2, pp. 157–182 (1938 Aug.).

[7] "The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition," in *IEEE Std 100-2000* (2000 Dec.).

[8] RIAA, "*Dimensional Standards, Disc Phonograph Records for Home Use*," Bulletin No. E 4 (1978 Nov.).

[9] J. Kogen, B. Jakobs, and F. Karlov, "Trackability – 1973," *Audio*, vol. 57, no. 8, pp. 16–22 (1973 Aug.).

[10] Wikipedia, "Initiation\ (Todd Rundgren Album)," https://en.wikipedia.org/wiki/Initiation_(Todd_Rundgren_album)#Track_listing (accessed May 7, 2023).

[11] P. C. Goldmark, R. Snepvangers, and W. S. Bachman, "The Columbia Long-Playing Microgroove Recording System," *Proc. IRE*, vol. 37, no. 8, pp. 923–927 (1949 Aug.).

[12] F. V. Hunt, "The Rational Design of Phonograph Pickups," *J. Audio Eng. Soc.*, vol. 10, no. 4, pp. 274–289 (1962 Oct.).

[13] FAA, "Acceleration in Aviation: G-Force," https://www.faa.gov/sites/faa.gov/files/pilots/safety/pilotsafetybrochures/Acceleration.pdf (accessed May 7, 2023).

[14] H. A. Frederick, "Vertical Sound Records," *J. Soc. Motion Picture Eng.*, vol. 18, no. 2, pp. 141–163 (1932 Feb.)

[15] M. J. Di Toro, "Distortion in the Reproduction of Hill-and-Dale Recording," *J. Soc. Motion Picture Eng.*, vol. 29, no. 5, pp. 493–509 (1937 Nov.).

[16] G. E. O'Brien, "The New Age of Sound: How Bell Telephone Laboratories and Leopold Stokowski Modernized Music," *Acoust. Today*, vol. 14, no. 2, pp. 17–24 (2018 Summer).

[17] B. Olney, “Phonograph Pickup Tracking Error vs. Distortion and Record Wear,” *Electronics*, vol. 10, no. 11, pp. 19–23 (1937 Nov.).

[18] F. W. Grover, “Analysis of Alternating-Current Waves by the Method of Fourier, with Special Reference to Methods of Facilitating the Computation,” *Bull. Bur. Stand.*, vol. 9, no. 4, pp. 567–646. (1913 Oct.).

[19] W. D. Lewis and F. V. Hunt, “A Theory of Tracing Distortion in Sound Reproduction from Phonograph Records,” *J. Acoust. Soc. Am.*, vol. 12, no. 3, pp. 348–365 (1941 Jan.).

[20] M. S. Corrington, “Tracing Distortion in Phonograph Records,” *RCA Rev.*, vol. 10, no. 2, pp. 241–253 (1949 Jun.).

[21] E. R. Madsen, “Vertical Tracking Angle – A Source of IM Distortion,” *Audio*, vol. 46, no. 11, pp. 21–24 (1962 Nov.).

[22] C. R. Bastiaans, “Further Thoughts on Geometric Conditions in the Cutting and Playing of Stereo Discs,” *J. Audio Eng. Soc.*, vol. 11, no. 1, pp. 6–15 (1963 Jan.).

[23] D. H. Cooper, “Tracking Distortion as Phase Modulation,” *IEEE Trans. Audio*, vol. AU-11, no. 2, pp. 41–46 (1963 Mar.-Apr.).

[24] D. H. Cooper, “Further Comments on Geometric Conditions in the Cutting of Stereo Records,” *J. Audio Eng. Soc.*, vol. 11, no. 3, pp. 280, 282 (1963 Jul.).

[25] D.H. Cooper, “Unified Analysis of Tracing and Tracking Errors,” *IEEE Trans. Audio*, vol. AU-11, no. 4, pp. 141–148 (1963 July-Aug.).

[26] D. H. Cooper, “Compensation for Tracing and Tracking Error,” *J. Audio Eng. Soc.*, vol. 11, no. 4, pp. 318–322 (1963 Oct.).

[27] D. H. Cooper, “Integrated Treatment of Tracing and Tracking Error,” *J. Audio Eng. Soc.*, vol. 12, no. 1, pp. 2–7 (1964 Jan.).

[28] D. H. Cooper, “Interaction of Tracing and Tracking Error,” *J. Audio Eng. Soc.*, vol. 13, no. 2, pp. 138–141 (1965 Apr.).

[29] D. Cooper, “Power-Series corrections of Tracing Distortion,” *IEEE Trans. Audio*, vol. AU-13, no. 2, pp. 27–30(1965 Mar.-Apr.).

[30] D. H. Cooper, “Construction of Tracing Correlator Waveforms,” *J. Audio Eng. Soc.*, vol. 13, no. 3, pp. 236–240 (1965 Jul.).

[31] D.H. Cooper, “Continuous Delay Regulator for Controlling Recording Errors,” J. Audio Eng. Soc., vol. 14, no. 1, pp. 7-16, (1966 Jan.).

[32] D. H. Cooper and J. E. Stifle, “Skew Sampler Construction: Interim Report,” *J. Audio Eng. Soc.*, vol. 14, no. 2, pp. 164, 166 (1966 Apr.).

[33] D. H. Cooper, “Waveform Stretching Using Time-Varying Memories: A Sampling Approach,” *Proc. IEEE*, vol. 55, no. 5, pp. 712–712 (1967 May).

[34] P. Copeland, *Manual of Analogue Sound Restoration Techniques* (The British Library, London, UK, 2008).

[35] T. Edison, “Improvement in Phonograph or Speaking Machines,” *US Patent* 200,521 (1878 Feb.).

[36] E. Berliner, “Gramophone,” *US Patent* 372,786 (1887 Nov.).

[37] E. Cook, “Binaural Disc Recording,” *J. Audio Eng. Soc.*, vol. 1, no. 1, pp. 1–3 (1953 Jan.).

[38] A.D. Blumlein, “Improvements in and Relating to Sounds Transmission, Sound Recording and Sound Reproducing Systems,” *GB Patent* 394,325 (1931 Dec.).

[39] A. Millard, “*America on Record*,” 2nd ed. (Cambridge University Press, New York, NY, 2005).

[40] C. R. Anderson, “Fifty Years of Stereo Pickups: A Capsule History," *J. Acoust. Soc. Am.*, vol. 77, no. 4, pp. 1320–1326 (1985 Apr.).

[41] H. E. Roys, “Reminiscing — The Stereophonic Record*,”* *J. Acoust. Soc. Am.*, vol. 77, no. 4, pp. 1320–1326 (1985 Apr.).

[42] H. E. Roys, “The Coming of Stereo,” *J. Audio Eng. Soc.*, vol. 25, nos. 10/11, pp. 824–827 (1977 Nov.).

[43] M. S. Corrington and T. Murakami, “Tracing Distortion in Stereophonic Disk Recording,” *RCA Rev*., vol. 19, no. 2, pp. 216–231 (1958 Jun.).

[44] M. Heideman, D. Johnson, and C. Burrus, “Gauss and the History of the Fast Fourier Transform,” *IEEE ASSP Mag.*, vol. 1, no. 4, pp. 14–21 (1984 Oct.).

[45] J. W. Cooley and J. W. Tukey, “An Algorithm for the Machine Calculation of Complex Fourier Series,” *Math. Comput.*, vol. 19, no. 90, pp. 297–301 (1965 Apr.).

[46] Shure, "Data for Various V15 Models Can Be Found in Their Respective User Guides," https://www.shure.com/en-US/support/downloads/documents?discontinued=discontinued&current=current&search=V15 (accessed May 7, 2023).

[47] H. E. Roys, “Analysis by the Two-Frequency Intermodulation Method of Tracing Distortion Encountered in Phonograph Reproduction,” *RCA Rev*., vol. 10, no. 2, pp. 254–269 (1949 Jun.).

[48] J. Kogen and R. Samson, “The Elliptical Stylus,” *Audio*, vol. 48, no. 5, pp. 33–36, 60–61 (1964 May).

[49] E. C. Fox and J. G. Woodward, “Tracing Distortion – Its Causes and Correction in Stereodisk Recording Systems,” *J. Audio Eng. Soc.*, vol. 11, no. 4, pp. 294–301 (1963 Oct.).

[50] H. Redlich and H. J. Klemp, “A New Method of Disc Recording for Reproduction with Reduced Distortion: The Tracing Simulator,” *J. Audio Eng. Soc.*, vol. 13, no. 2, pp. 111–118 (1965 Apr.).

[51] D. Braschoss, “Development and Application of a New Tracing Simulator,” *J. Audio Eng. Soc.*, vol. 19, no. 2, pp. 108–114 (1971 Feb.).

[52] M. Ohkawa, M. Kuriyagawa, and S. Makino, “Clear Sound Records Using the Pinchless Tracing Simulator (PTS) System,” presented at the *39th Convention of the Audio Engineering Society*, (1970 Oct.), paper 747.

[53] S. Washizawa, T. Nakatani, and T. Shiga, “Development of Skew Sampling Compensator for Tracing Error,” *J. Audio Eng. Soc.*, vol. 21, no. 8, pp. 630–634 (1973 Oct.).

[54] E. G. Trendell, “Tracing Distortion Correction,” *J. Audio Eng. Soc.*, vol. 25, no. 5, pp. 273–277 (1977 May)

[55] J. Gordon Holt, “Down with the Dynagroove!” https://www.stereophile.com//asweseeit/95/ (accessed May 7, 2023)

[56] H. F. Olson, “The RCA Victor DYNAGROOVE System,” *J. Audio Eng. Soc.*, vol. 12, no. 2, pp. 98–114 (1964 Apr.).

[57] Neumann, “Electrical Calibration of the TS 66 Tracing Simulator,” 1973(?), https://www.lathetrolls.com/viewtopic.php?f=1&t=4949 (accessed May 7, 2023).

[58] D. W. Gravereaux, A. J. Gust, and B. B. Bauer, “The Dynamic Range of Disc and Tape Records,” *J. Audio Eng. Soc.*, vol. 18, no. 5, pp. 530–535 (1970 Oct.).

[59] J. M. Eargle, “Performance Characteristics of the Commercial Stereo Disc,” *J. Audio Eng. Soc.*, vol. 17, no. 4, pp. 416–422 (1969 Aug.).

[60] J. W. Stafford, “Maximum Peak Velocity Capabilities of the Disk Record,” *J. Audio Eng. Soc.*, vol. 8, no. 3, pp. 162–165 (1960 Jul.).

[61] D. A. Barlow and G. R. Garside, “Groove Deformation and Distortion in Recordings,” *J. Audio Eng. Soc.*, vol. 26, nos. 7/8, pp. 498–510 (1978 Aug.).

[62] T. Inoue, I. Owaki, K. Totsuka, S. Nakamura, and J. Eargle, “Improvements in Cutting Styli for CD-4 Discs,” *J. Audio Eng. Soc.*, vol. 23, no. 8, pp. 641–643 (1975 Oct.)

[63] C. R. Anderson, J. H.. Kogen, and R. S. Samson, “Optimizing the Dynamic Characteristics of a Phonograph Pickup,” *J. Audio Eng. Soc.*, vol. 14, no. 2, pp. 145–152 (1966 Apr.).

[64] T. Holman, “Dynamic Range Requirements of Phonographic Preamplifiers,” *Audio*, vol. 61, no. 7, pp. 72–79 (1977 Jul.).

[65] C. Lu, “November BAS Meeting,” *B.A.S. Speaker*, vol. 5, no. 3, pp. 17–21 (1976 Dec.).

[66] R. M. Bauman, “The Dynamic Range Potential of the Phonograph,” *Audio Electronics*, vol. 27, no. 4 (1996 Dec.).

[67] Discogs, “The Woody Herman Big Band – Hey! Heard the Herd?,” https://www.discogs.com/release/7925075 (accessed May 7, 2023)

[68] “IRE Standards on Recording and Reproducing: Methods of Calibration of Mechanically-Recorded Lateral Frequency Records, 1958,” *Proc. IRE*, vol. 46, no. 12, pp. 1940–1946 (1958 Dec.).

[69] H.E. Hayne and H.E. Roys, “A Variable Speed Turntable and Its Use in the Calibration of Disk Reproducing Pickups,” *Proc. IRE*, vol. 38, no. 3, pp. 239–243 (1950 Mar.).

[70] D. H. Cooper, “Vertical Tracking Distortion in Practice,” *Audio*, vol. 47, no. 9, pp. 40, 71 (1963 Sep.).

[71] J. E. Jacobs and P. Wittman, “Psychoacoustics, the Determining Factor in Stereo Disc Distortion,” *J. Audio Eng. Soc.*, vol. 12, no. 2, pp. 115–123 (1964 Apr.).

[72] S. P. Lipshitz, M. Pocock, and J. Vanderkooy, “On the Audibility of Midrange Phase Distortion in Audio Systems,” *J. Audio Eng. Soc.*, vol. 30, no. 9, pp. 580–595 (1982 Sep.).

[73] L. Demany and C. Semal, “Detection Thresholds for Sinusoidal Frequency Modulation,” *J. Acoust. Soc. Am.*, vol. 85, no. 3, pp. 1295–1301 (1989 Mar.).

[74] S. M. A. Ernst and B. C. J. Moore, “Mechanisms Underlying the Detection of Frequency Modulation,” *J. Acoust. Soc. Am.*, vol. 128, no. 6, pp. 3642–3648 (2010 Dec.).

[75] IEEE, “ Method of Measurement of Weighted Peak Flutter of Sound Recording and Reproducing Equipment," *Standard 193-1971* (1972 Mar.).

[76] H. Sakai, “Perceptibility of Wow and Flutter,” *J. Audio Eng. Soc.*, vol. 181, no. 3, pp. 290–298 (1970 Jun.).

[77] R. Tollerton, “Digital Simulation of Phonograph Tracking Distortion,” presented at the *127th Convention of Audio Engineering Society* (2009 Oct.), paper 7924.

[78] M. Abramowitz and I. A Stegun (Ed.), “Handbook of Mathematical Functions, Nat. Bureau Standards, 10th printing (1972).

[79] Mathworld Wolfram, “Trigonometric Formulas,” https://mathworld.wolfram.com/TrigonometricPowerFormulas.html (accessed May 7, 2023)

[80] H. Taub and D. L. Schilling, *Principles of Communications Systems,* (McGraw-Hill, New York, NY, 1987).

[81] J. P. Ballantine and A. R. Jerbert, “Distance from a Line, or Plane, to a Point,” *Amer. Math. Mon.*, vol. 59, no. 4, pp. 242–243 (1952 Apr.).

[82] R. Brice, “Needle-Drop Handbook,” 3rd ed., (*Transform Media Limited*, Liphook, UK, 2022).

APPENDIX A.1 ANALYSES OF TRACING DISTORTION

Most tracing distortion analyses in the past did not follow the method described in SEC. 1. Instead, one approach was to use parametric equations for the movement of the center of the spherical stylus, which worked well for numerical evaluations [6, 43].

In [19] and [20], a more general approach was adopted, wherein the stylus did not have to be of circular cross-section. If we let $\phi(\varepsilon)$ denote the shape of the cross-section of the stylus, as in SEC. 3, only $\phi(0) = 0$ and $\phi'(0) = 0$ are required. Because $\phi(\varepsilon)$ does not necessarily have a center as a spherical cross-section, the stylus tip bottom can be used as a reference point for calculations.

Instead of parametric equations, Lewis and Hunt in [19] envisioned a coordinate system (X, Y) traveling along the groove at the groove speed V (from Eq. (1), $V = R\Omega$), as depicted in Fig. A1. Keeping largely their original notation, one can write the recorded waveform as

$$Y = A\,\psi(X + Vt), \qquad \text{(A.1)}$$

where $\psi(\cdot)$ is the waveform normalized to unity. Based on Fig. A1, at some instant t when $X = 0$, the displacement of the tip's bottom point, denoted by $S(t)$, will be given by

$$S(t) = A\psi(\varepsilon + Vt) - \phi(\varepsilon). \qquad \text{(A.2)}$$

The tangentiality requirement at the point of contact is

$$A\,\psi'(\varepsilon + Vt) = \phi'(\varepsilon). \qquad \text{(A.3)}$$

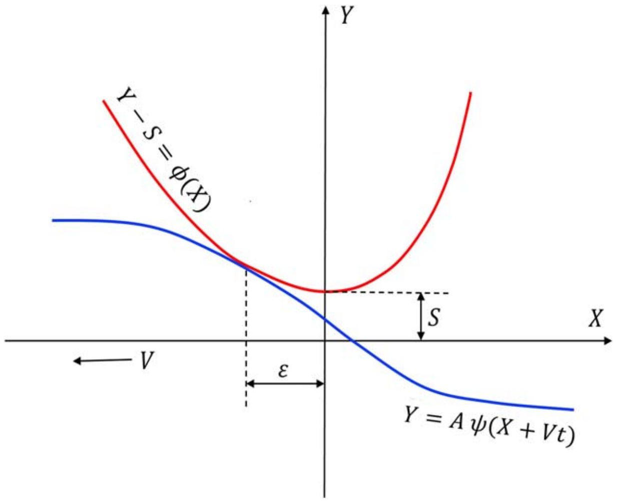

Fig. A1 Stylus-groove contact geometry [19].

Expanding both sides of Eq. (A.3) into a Taylor series gives

$$\sum_{i=0}^{\infty} \frac{1}{i!}\, \phi^{(i+1)}(0)\ \varepsilon^i = A \sum_{j=0}^{\infty} \frac{1}{j!}\, \psi^{(j+1)}(Vt)\ \varepsilon^i. \quad (A.4)$$

Because the contact point distance ε will depend on the amplitude A ($\varepsilon = 0$ if $A = 0$, and would increase when A increases), it makes sense to expand it in the power series in terms of the amplitude A as

$$\varepsilon = \sum_{k=0}^{\infty} a_k\ A^k. \quad (A.5)$$

In [20], Corrington used Eq. (A.5) to calculate the first terms of the expressions for ε^i for $i = 1 \ldots,7$, substituted these values into Eq. (A.4), and evaluated the coefficients associated with A, A^2, ... A^7 on both sides. By equating these coefficients, he was able to get 7 equations for the coefficients a_1 to a_7 in terms of the amplitude A and derivatives $\phi^{(i)}(0)$ and $\psi^{(j)}(Vt)$, which in turn gave him the expressions for the ε^2 to ε^7 in terms of these same variables.

Expanding (A.2) into the series similarly results in

$$S(t) = \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} \left(A\ \psi^{(n+1)}(Vt) - \phi^{(n)}(0)\right), \quad (A.6)$$

which is a polynomial in A that can be written as

$$S(t) = \sum_{m=1}^{\infty} b_m\ A^m. \quad (A.7)$$

Corrington then derived formulas for the coefficients b_1 to b_7 in terms of the derivatives of $\phi(0)$ and $\psi(Vt)$, which will not be reproduced here because the first four required almost a full page (but can be found in [20, p. 244]).

Assume the case of a cosine modulation

$$\psi(Vt) = \cos(kx). \quad (A.8)$$

The derivatives of Eq. (A.8) can be easily calculated, and so can the derivatives of $\phi(\varepsilon)$ for a spherical tip

$$\phi(\varepsilon) = \rho - \sqrt{\rho^2 - \varepsilon^2}, \quad (A.9)$$

at $\varepsilon = 0$ (e.g., $\phi^{(0)}(0) = 0$, $\phi^{(1)}(0) = 0$, $\phi^{(2)}(0) = 1/\rho$, etc.).

As the final step, the powers of the cosine and sine functions from the $\phi^{(n)}(x)$ terms had to be expressed in terms of multiple angles [78], and the coefficients $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \ldots$ associated with each sine or cosine component at kx, $2kx$, $3kx \ldots$ identified.

All odd-order harmonics will be at phase zero, while the even-order ones will be at $\pm\pi/2$. This is because the odd derivatives of $\cos(kx)$ take the form $\pm k^n \sin(kx)$, and the powers of sines can be expressed only in the sines of multiple angles (and analogously for cosines [79]). Thus

$$y_{out}(x) = \mathcal{A}_0 + \mathcal{A}_1 \cos(kx) + \mathcal{A}_2 \sin(2kx) + \\ + \mathcal{A}_3 \cos(3kx) + \mathcal{A}_4 \sin(4kx) \ldots \quad (A.10)$$

The truncated versions of the formulas for the first three harmonics $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ are given in Eqs. (12) to (14) in SEC. 3.

The complete set of Corrington's results for the first seven harmonics required 4 pages [20, pp. 245–247] and will not be reproduced here. But even more impressive were his formulas for the intermodulation distortion (IMD) with two tones present for the lateral recording, which required 3 pages for the first four intermodulation products. IMD is beyond the scope of the present paper, but one cannot but admire the tenacity and mathematical prowess of these early researchers. Evaluating these formulas would be very time-consuming even nowadays, using modern software tools for symbolic algebra computations.

Besides a paper [43] that used the numerical evaluation of distortion on a computer for the stereo record case, nothing was published on tracing distortion until 1963. As mentioned earlier, the most prolific author on this subject was D. Cooper. He first tried to show that both tracing and tracking distortion could be viewed as a so-called skew transformation[29] [25]. Instead of focusing on the spherical tip, Cooper assumed a general form of the transformation in the tracing case to be

$$\xi = x + G(y') \quad (A.11)$$

$$\eta' = y', \quad (A.12)$$

where $G(\cdot)$ is the "inverse of $\phi'(\cdot)$," a derivative of the stylus shape cross-section shape function $\phi(\cdot)$. In the case of a spherical tip, he expressed it as

$$G = \frac{y'}{\sqrt{1 + (y')^2}}, \quad (A.13)$$

[29] A skew transformation relates the coordinates in a standard coordinate system with the coordinates in a skewed one, where the vertical axis is at an angle other than 90° relative to the horizontal one.

but then chose to use a parabolic stylus tip[30] with a curvature at the nominal contact point of ρ, i.e.

$$\phi(\varepsilon) = \frac{\varepsilon^2}{2\rho}, \tag{A.14}$$

which gave him a much simpler expression for $\xi - x$ ($= \varepsilon$ in our notation) in Eq. (A.11). In the case of the recorded tone

$$y(x) = A\cos(kx) \tag{A.15}$$

it reduces to

$$\varepsilon = -\rho k A \sin[k(x+\varepsilon)]. \tag{A.16}$$

Cooper then proceeded to insert Eq. (A.16) repeatedly into a $\eta = y$ version of Eq. (A.12) to obtain what he called the "continued function" (inspired by the "continued fraction"). In our notation from Sec. 1, it would be

$$y_{out}(x) = A\cos\{kx - k\rho k A \sin[kx - k\rho k A \sin(\ldots)]\}. \tag{A.17}$$

If $k\rho k A$ was small, he then argued that Eq. (A.17) can be approximated by

$$y_{out}(x) \approx A\cos\{kx - k\rho k A \sin(kx)\}. \tag{A.18}$$

Eq. (A.18) is a canonical expression for the phase-modulated waveform, with a modulation index $k\rho k A = \beta$.

At this point, Cooper seems to have been the first to note that tracing distortion is essentially a phase modulation process. One can argue that this conclusion was obvious: from Fig. 2, we see that the waveform traced by the stylus goes through all the same amplitudes as the original, just at the slightly wrong instances in time. Still, it was never stated explicitly before.

A complete spectral analysis of Eq. (A.18) could have been done here in terms of Bessel's functions, but Cooper—usually not shy to introduce very complex math in his papers—opted to consider only the second harmonic. It corresponds to the first principal component on the sum of the carrier and modulating frequencies (equal here).

For small modulation indexes ($k\rho k A$ in this case), from the theory of FM [80], it is well known that the relative amplitude of that component would equal half of the modulation index itself. In the notation from Sec. 3, the second harmonic distortion factor would be

$$D_2 = \frac{1}{2} k A k \rho, \tag{A.19}$$

which is twice what the previous analyses [19, 20] gave.[31]

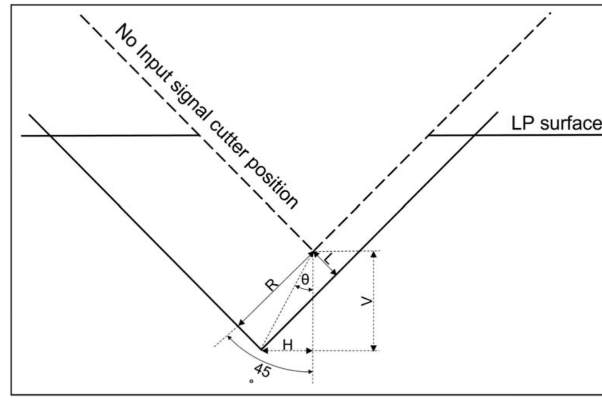

Fig. B.1 Cutter positions in a neutral (no input signal) state and with intended displacements of L and R in the left and right channel.

APPENDIX B.1 HORIZONTAL AND 45°/45° CUT

In the 45°/45° cut, each side of the groove is indented in proportion to the left and right channel signals. If the intended displacements are denoted by L and R, the cross-cut of the disc will look as shown in Fig. B1.

To calculate the resultant displacements H and V in the horizontal and vertical directions, from Fig. B1, we see that

$$\frac{H}{V} = \tan(\theta) \tag{B.1}$$

$$\frac{L}{R} = \tan(\pi/4 - \theta) \tag{B.2}$$

From Fig. B1, it is also evident that

$$L^2 + R^2 = H^2 + V^2 \tag{B.3}$$

By using the addition formula for tangents [78, p. 72]

$$\tan\left(\frac{\pi}{4} - \theta\right) = \frac{1 - \tan(\theta)}{1 + \tan(\theta)}, \tag{B.4}$$

from Eqs. (B.3) and (B.1)

$$\frac{L}{R} = \frac{1 - \tan(\theta)}{1 + \tan(\theta)} = \frac{V - H}{V + H}. \tag{B.5}$$

Solving Eq. (B.5) for V then gives

$$V = H\frac{L + R}{R - L}, \tag{B.6}$$

[30] Cooper stated that the assumption about the spherical tip "invokes complications of no theoretical interest" and is "unwarranted," without providing any practical justification for the parabolic one. As long as tracing errors and distortion are small, however, all shapes with the same radius of curvature at the end of the tip will give very similar results.

[31] In [25, p.144], Cooper argued that "the extra factor 1/2 appears for tracing distortion, upon converting to the amplitude mode, because of the reduced sensitivity to the second harmonic that such a mode conversion entails." However, this whole analysis was done for the displacement mode, not the velocity mode, and he did not include this factor in any of the tracking distortion calculations performed in parallel. This author believes that the problem was in Cooper's equation given as Eq. (A.18) here, which approximates the displacement at the movable point of contact, not at the fixed point on the stylus (like the center of the spherical tip, or the bottom of the more complex shapes). This can be verified by comparing Eq. (6) with Eq. (A.18), the latter of which misses the second term. In his later work, Cooper used the fixed point on the stylus as the reference [27].

By inserting this into Eq. (B.3), the horizontal/lateral component of the resultant displacement is

$$H = \frac{R - L}{\sqrt{2}}, \tag{B.7}$$

and from Eq. (B.6) we find the vertical one as

$$V = \frac{R + L}{\sqrt{2}}, \tag{B.8}$$

This seems to contradict the statement from SEC. 4 that, to preserve the backward compatibility on mono reproducers, the horizontal displacement has to be proportional to the sum of the left and right channel signals. However, the intended displacement can be made if one channel operates in the opposite direction (e.g., decrease it from the nominal position when the signal is positive and increase it when it is negative, while doing the opposite in the other channel). In practice, this can be accomplished simply by reversing the polarity of the two stereo signals at the cutter's coils. This is the arrangement that was actually already shown in Fig. 9. Assuming that the right channel was inverted, we would then have

$$R = -K \cdot s_R(t) \tag{B.9}$$

$$L = K \cdot s_L(t). \tag{B.10}$$

In Eqs. (B.9) and (B.10), K is the proportionality constant dependent on the cutter, while $s_R(t)$ and $s_L(t)$ stand for the right and left stereo input signals, respectively. In such a case, the lateral displacement will be

$$H = \frac{K\left[s_R(t) + s_L(t)\right]}{\sqrt{2}}, \tag{B.11}$$

and the vertical displacement will be proportional to the difference between the right and left channel signals.

With this in mind, we can now explain the lack of even harmonics in the lateral cut mentioned in SEC. 4. Following the classic analysis by Lewis and Hunt [19], assume a cosine recorded signal

$$s_L(kx) = A\cos(kx). \tag{B.11}$$

Looking at the left groove perpendicularly from the right at the 45° elevation angle, we see that the stylus will behave as if the cut was vertical. If the point of contact with the left groove is marked with L, the stylus would thus follow a poid curve in the LL' direction. Based on the analysis in APPENDIX A, that trajectory can be described as

$$\mathcal{S}_L(kx) = \mathcal{A}_0 + \mathcal{A}_1\cos(kx) + \mathcal{A}_2\sin(2kx) + \\ + \mathcal{A}_3\cos(3kx) + \mathcal{A}_4\sin(4kx) + \ldots, \tag{B.12}$$

where the coefficients $\mathcal{A}_i$ are obtained from Eqs. (12) to (14) for the amplitudes of the harmonics.

The center of the stylus C in Fig. B2 will be at a distance $\rho + \mathcal{S}_L(kx)$ from the left groove plane defined by

$$y + z = 0. \tag{B.13}$$

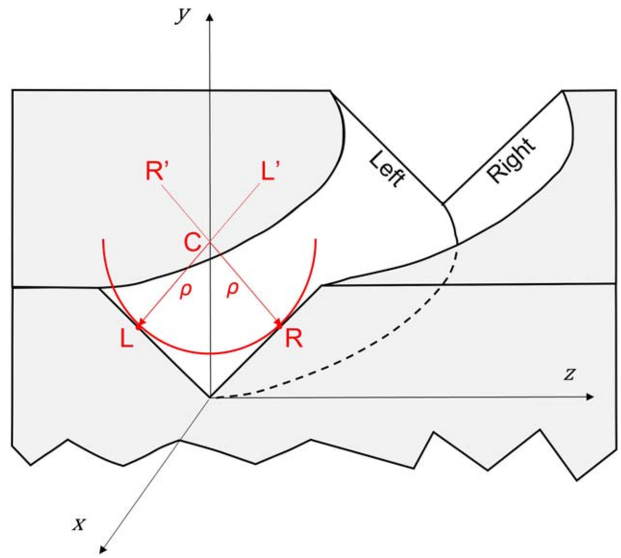

Fig. B.2 Geometry of the laterally cut groove.

The distance D between this plane and any point with the coordinates (y_0, z_0) is [81]

$$\frac{y_0 + z_0}{\sqrt{2}} = D. \tag{B.13}$$

Letting the coordinates of point C be (Y, Z), we obtain

$$Y + Z = \sqrt{2}\left[\rho + \mathcal{S}_L(kx)\right]. \tag{B.14}$$

Analogously, on the other wall of the groove with the mono lateral recording, from (B.9)

$$s_R(kx) = -A\cos(kx), \tag{B.15}$$

but to calculate the harmonic amplitudes, we will now have to set the amplitude to $-A$ in Eqs. (12) to (14). Because the odd harmonics in them contain only odd, and the even harmonics only even powers of A, the expression will look like the one in Eq. (B.12) but with the odd harmonics having negative signs, i.e.,

$$\mathcal{S}_R(kx) = \mathcal{A}_0 - \mathcal{A}_1\cos(kx) + \mathcal{A}_2\sin(2kx) + \\ - \mathcal{A}_3\cos(3kx) + \mathcal{A}_4\sin(4kx) + \ldots \tag{B.16}$$

The right groove plane can be defined analogously by

$$y - z = 0, \tag{B.17}$$

and following the same logic

$$Y - Z = \sqrt{2}\left[\rho + \mathcal{S}_R(kx)\right]. \tag{B.18}$$

Solving for Z from Eqs. (B.14) and (B.18)

$$Z = \frac{\sqrt{2}}{2}\left[\mathcal{S}_L(kx) - \mathcal{S}_R(kx)\right],$$

and from Eqs. (B.12) and (B.16)

$$Z = \frac{\sqrt{2}}{2}\left\{\mathcal{A}_1\cos(kx) + \mathcal{A}_3\cos(3kx) + \ldots\right\}. \tag{B.19}$$

The vertical position of center C will similarly be

$$Y = \sqrt{2}\rho + \frac{\sqrt{2}}{2}\left[\mathcal{S}_L(kx) + \mathcal{S}_R(kx)\right],$$

and finally

$$Y = \sqrt{2}\rho + \frac{\sqrt{2}}{2}\{\mathcal{A}_0 + \mathcal{A}_2 \sin(2kx) + \\ + \mathcal{A}_4 \sin(4kx) + \ldots\}. \qquad \text{(B.20)}$$

From Eq. (B.19), the laterally cut records will have no even-order harmonic distortion components in the horizontal plane. In contrast, in the vertical plane, the stylus motion with the lateral cut would only have even harmonics, with the first being at twice the fundamental frequency.

THE AUTHORS

Vladan Jovanovic

As a student, Vladan Jovanovic worked as a rock journalist and editor of a Hi-Fi column in the Yugoslav music magazine "*Džuboks*." In 1982, he won the Belgrade Chamber of Commerce Award for his B.Sc. thesis on the record player's tonearm geometry. After getting his Ph.D. (EE) from the University of Belgrade, Serbia, he worked in R&D for digital telecommunications. From 1991 to 1993, he was a Research Associate at the University of Toronto, Canada. From 1993 to 2006, he worked with mobile telephone operators in Canada and the US, and "Lucent Technologies. From 2006 to 2008 he was a CTO of "Newfield Wireless," and has worked as a private Consultant since. Dr. Jovanovic authored over 30 technical papers and holds seven patents. Besides working as an expert witness for mobile telephony in the US courts, in his retirement, he has also written the book *Electronics of Rock and Roll*.